Gerardo F. Torres del Castillo

Differentiable Manifolds

A Theoretical Physics Approach

Birkhäuser

Birkhäuser

Gerardo F. Torres del Castillo

Differentiable Manifolds

A Theoretical Physics Approach

Birkhäuser

Gerardo F. Torres del Castillo
Instituto de Ciencias
Universidad Autónoma de Puebla
Ciudad Universitaria
72570 Puebla, Puebla, Mexico
gtorres@fcfm.buap.mx

ISBN 978-0-8176-8270-5 e-ISBN 978-0-8176-8271-2
DOI 10.1007/978-0-8176-8271-2
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011939950

Mathematics Subject Classification (2010): 22E70, 34C14, 53B20, 58A15, 70H05

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.birkhauser-science.com)

Preface

The aim of this book is to present in an elementary manner the basic notions related with differentiable manifolds and some of their applications, especially in physics. The book is aimed at advanced undergraduate and graduate students in physics and mathematics, assuming a working knowledge of calculus in several variables, linear algebra, and differential equations. For the last chapter, which deals with Hamiltonian mechanics, it is useful to have some previous knowledge of analytical mechanics. Most of the applications of the formalism considered here are related to differential equations, differential geometry, and Hamiltonian mechanics, which may serve as an introduction to specialized treatises on these subjects.

One of the aims of this book is to emphasize the connections among the areas of mathematics and physics where the formalism of differentiable manifolds is applied. The themes treated in the book are somewhat standard, but the examples developed here go beyond the elementary ones, trying to show how the formalism works in actual calculations. Some results not previously presented in book form are also included, most of them related to the Hamiltonian formalism of classical mechanics. Whenever possible, coordinate-free definitions or calculations are presented; however, when it is convenient or necessary, computations using bases or coordinates are given, not underestimating their importance.

Throughout the work there is a collection of exercises, of various degrees of difficulty, which form an essential part of the book. It is advisable that the reader attempt to solve them and to fill in the details of the computations presented in the book.

The basic formalism is presented in Chaps. 1 and 3 (differentiable manifolds, differentiable mappings, tangent vectors, vector fields, and differential forms), after which the reader, if interested in applications to differential geometry and general relativity, can continue with Chaps. 5 and 6 (even though in the definitions of a Killing vector field and of the divergence of a vector field given in Chap. 6, the definition of the Lie derivative, presented in Chap. 2, is required). Chapter 7 deals with Lie groups and makes use of concepts and results presented in Chap. 2 (one-parameter groups and Lie derivatives). Chapters 2 and 4 are related with differential equations and can be read in an independent form, after Chaps. 1 and 3. Finally, for Chap. 8, which deals with Hamiltonian mechanics, the material of Chaps. 1, 2, and 3, is necessary and, for some sections, Chaps. 6 and 7 are also required.

Some of the subjects not treated here are the integration of differential forms, cohomology theory, fiber bundles, complex manifolds, manifolds with boundary, and infinite-dimensional manifolds.

This book has been gradually developed starting from a first version in Spanish (with the title *Notas sobre variedades diferenciables*) written around 1981, at the Centro de Investigación y de Estudios Avanzados del IPN, in Mexico, D.F. The previous versions of the book have been used by the author and some colleagues in courses addressed to advanced undergraduate and graduate students in physics and mathematics.

I would like to thank Gilberto Silva Ortigoza, Merced Montesinos, and the reviewers for helpful comments, and Bogar Díaz Jiménez for his valuable help with the figures. I also thank Jessica Belanger, Tom Grasso, and Katherine Ghezzi at Birkhäuser for their valuable support.

Puebla, Puebla, Mexico Gerardo F. Torres del Castillo

Contents

Chapter 1
Manifolds

The basic objective of the theory of differentiable manifolds is to extend the application of the concepts and results of the calculus of the $\mathbb{R}^n$ spaces to sets that do not possess the structure of a normed vector space. The differentiability of a function of $\mathbb{R}^n$ to $\mathbb{R}^m$ means that around each interior point of its domain the function can be approximated by a linear transformation, but this requires the notions of linearity and distance, which are not present in an arbitrary set.

The essential idea in the definition of a manifold should already be familiar from analytic geometry, where one represents the points of the Euclidean plane by a pair of real numbers (e.g., Cartesian or polar coordinates). Roughly speaking, a manifold is a set whose points can be labeled by coordinates.

In this chapter and the following three, the basic formalism applicable to any finite-dimensional manifold is presented, without imposing any additional structure. In Chaps. 5 and 6 we consider manifolds with a connection and a metric tensor, respectively, which are essential in differential geometry.

1.1 Differentiable Manifolds

Let M be a set. A *chart* (or *local chart*) on M is a pair (U, ϕ) such that U is a subset of M and ϕ is a one-to-one map from U onto some *open* subset of $\mathbb{R}^n$ (see Fig. 1.1). A chart on M is also called a *coordinate system* on M. Defining a chart (U, ϕ) on a set M amounts to labeling each point $p \in U$ by means of n real numbers, since $\phi(p)$ belongs to $\mathbb{R}^n$, and therefore consists of n real numbers that depend on p; that is, $\phi(p)$ is of the form

$$\phi(p) = \big(x^1(p), x^2(p), \ldots, x^n(p)\big). \tag{1.1}$$

This relation defines the n functions $x^1, x^2, \ldots, x^n$, which will be called the coordinate functions or, simply coordinates, associated with the chart (U, ϕ). The fact that ϕ is a one-to-one mapping ensures that two different points of U differ, at least, in the value of one of the coordinates.

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_1,

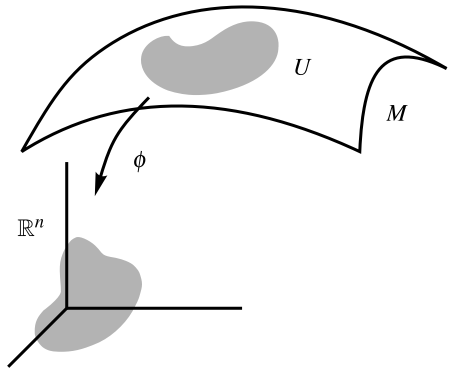

Fig. 1.1 A coordinate system in a set M; with the aid of ϕ, each point of U corresponds to some point of $\mathbb{R}^n$. The image of U under ϕ must be an open subset of some $\mathbb{R}^n$

We would also like close points to have close coordinates, but that requires some notion of nearness in M, which can be given by the definition of a distance between points of M or, more generally, by assigning a topology to M. We are not assuming that the reader is acquainted with the basic concepts of topological spaces and in most applications we will be dealing with sets possessing a natural notion of nearness (see, however, the comment after Exercise 1.2). Hence, we shall not make use of the concepts required for an adequate general discussion. For a more rigorous treatment see, e.g., Crampin and Pirani (1986), Conlon (2001), Boothby (2002), and Lee (2002).

These concepts have many applications in physics. For instance, if M is the configuration space of a mechanical system with n degrees of freedom, a choice of the so-called generalized coordinates is equivalent to the definition of a chart on M; when M is the set of equilibrium states of a thermodynamical system, the coordinates associated with a chart on M are, typically, the pressure, the temperature, and the volume of the system.

The coordinates associated with any chart (U, ϕ) must be functionally independent among themselves, since the definition of a chart requires that $\phi(U)$ $(\equiv \{\phi(p) \,|\, p \in U\})$ be an open subset of $\mathbb{R}^n$. If, for instance, the coordinate x^n could be expressed as a function of $x^1, x^2, \ldots, x^{n-1}$, then the points $\phi(p)$ $(p \in U)$ would lie in a hypersurface of $\mathbb{R}^n$, which is not an open subset of $\mathbb{R}^n$.

Frequently, a chart (U, ϕ) on M will not cover all of M, that is, U will be a proper subset of M; moreover, it is possible that a given set M cannot be covered by a single chart, as in the case of the circle or the sphere, where at least two charts are necessary to cover all the points of M (see the examples below). Hence, in order to cover all of M, it may be necessary to define two or more charts and, possibly, some points of M will lie in the domain of more than one chart.

A function $F : \mathbb{R}^n \to \mathbb{R}^m$ given by $F(q) = (f_1(q), f_2(q), \ldots, f_m(q))$ is *differentiable* of class C^k if the real-valued functions $f_1, f_2, \ldots, f_m$ have kth continuous partial derivatives; two charts on M, (U, ϕ) and (V, χ), are said to be C^k*-related* (or C^k*-compatible*) if $U \cap V = \emptyset$ (the empty set), or if $\phi \circ \chi^{-1} : \chi(U \cap V) \to \phi(U \cap V)$ and $\chi \circ \phi^{-1} : \phi(U \cap V) \to \chi(U \cap V)$, whose domains are open in $\mathbb{R}^n$, are dif-

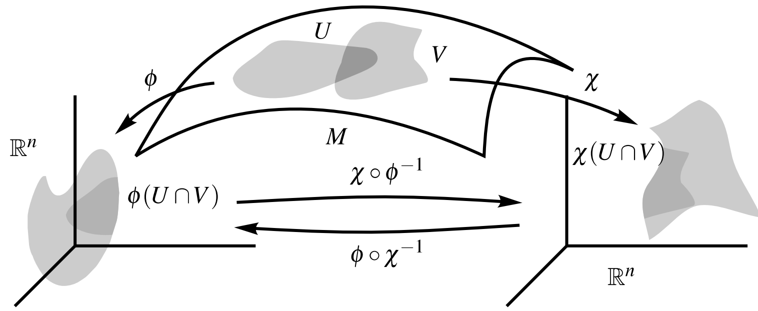

Fig. 1.2 Two coordinate systems whose domains have a nonempty intersection. A point p belonging to $U \cap V$ corresponds to two points of $\mathbb{R}^n$, $\phi(p)$ and $\chi(p)$; the charts (U, ϕ) and (V, χ) are C^k-related if the maps $\phi(p) \mapsto \chi(p)$ and $\chi(p) \mapsto \phi(p)$ are differentiable functions of class C^k

ferentiable of class C^k (see Fig. 1.2). If $x^1, x^2, \ldots, x^n$ are the coordinates associated to (U, ϕ) and $y^1, y^2, \ldots, y^n$ are the coordinates associated to (V, χ), the fact that (U, ϕ) and (V, χ) be C^k-related amounts to the fact that, for all $p \in U \cap V$, $y^1(p), y^2(p), \ldots, y^n(p)$ be differentiable functions of class C^k of $x^1(p), x^2(p), \ldots, x^n(p)$, and conversely.

A C^k *subatlas* on M is a collection of charts on M, $\{(U_i, \phi_i)\}$, such that for any pair of indices i, j, (U_i, ϕ_i) and (U_j, ϕ_j) are C^k-related and $M = U_1 \cup U_2 \cup \cdots$ (so that each point of M is in the domain of at least one chart). The collection of all the charts C^k-related with the charts of a C^k subatlas, on M, form a C^k *atlas* on M.

Definition 1.1 A C^k *manifold* of dimension n is a set M with a C^k atlas; if $k \geq 1$, it is said that M is a *differentiable manifold*. If $k = 0$, it is said that M is a *topological manifold*.

In the space $\mathbb{R}^n$, the pair $(\mathbb{R}^n, \mathrm{id})$ (where id denotes the identity map) is a chart that, by itself, forms a C^∞ subatlas. The infinite collection of all the coordinate systems C^∞-related with this chart form a C^∞ atlas with which $\mathbb{R}^n$ is a C^∞ manifold of dimension n. When we consider $\mathbb{R}^n$ as a differentiable manifold, it is understood that this is its atlas.

For instance, the usual polar coordinates of the Cartesian plane belong to the atlas of $\mathbb{R}^2$; one can readily verify that the pair (V, χ), with $V = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ and

$$\chi(x, y) = \left(\sqrt{x^2 + y^2}, \arctan y/x\right)$$

is a chart on $\mathbb{R}^2$ with $\chi(V) = (0, \infty) \times (-\pi/2, \pi/2)$, which is an open set in $\mathbb{R}^2$. Taking $(U, \phi) = (\mathbb{R}^2, \mathrm{id})$, one readily verifies that $(\chi \circ \phi^{-1})(x, y) = \chi(x, y) = (\sqrt{x^2 + y^2}, \arctan y/x)$ and $(\phi \circ \chi^{-1})(r, \theta) = \chi^{-1}(r, \theta) = (r \cos\theta, r \sin\theta)$ are differentiable of class C^∞ in $\phi(U \cap V) = V$ and $\chi(U \cap V) = \chi(V) = (0, \infty) \times (-\pi/2, \pi/2)$, respectively.

Let M be a manifold. A subset A of M is said to be *open* if for any chart (U, ϕ) belonging to the atlas of M, the set $\phi(A \cap U)$ is open in $\mathbb{R}^n$.

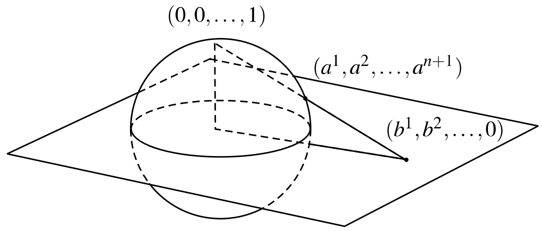

Fig. 1.3 The stereographic projection establishes a one-to-one correspondence between the points of the n-sphere, excluding the "north pole" $(0, 0, \ldots, 1)$, and the points of the plane $x^{n+1} = 0$. The point $(a^1, a^2, \ldots, a^{n+1})$ is a point of the n-sphere different from $(0, 0, \ldots, 1)$

Exercise 1.2 Show that the collection τ of open subsets of a manifold M is a *topology* of M, that is, show that M and the empty set belong to τ, that the union of any family of elements of τ belongs to τ, and that the intersection of any finite family of elements of τ belongs to τ. We say that this topology is induced by the manifold structure given in M.

When a given set, M, already possesses a topology and one wants to give it the structure of a manifold in such a way that the topology induced by the manifold structure coincides with the topology originally given, one demands that for each chart (U, ϕ), in the atlas of M, the map ϕ be continuous and have a continuous inverse; as a consequence, U must be an open set of M. (A map is continuous if and only if the preimage of any open set is open.)

Example 1.3 Almost all the points of the n-sphere

$$S^n \equiv \left\{ \left(a^1, \ldots, a^{n+1}\right) \in \mathbb{R}^{n+1} \mid \left(a^1\right)^2 + \cdots + \left(a^{n+1}\right)^2 = 1 \right\}$$

$(n \geq 1)$ can be put into a one-to-one correspondence with the points of $\mathbb{R}^n$ by means of the *stereographic projection* defined in the following way. Any point $(a^1, \ldots, a^{n+1}) \in S^n$, different from $(0, 0, \ldots, 1)$, can be joined with $(0, 0, \ldots, 1)$ by means of a straight line that intersects the hyperplane $x^{n+1} = 0$ at some point $(b^1, \ldots, b^n, 0)$ (see Fig. 1.3). The condition that the three points $(a^1, \ldots, a^{n+1})$, $(0, 0, \ldots, 1)$, and $(b^1, \ldots, b^n, 0)$ lie on a straight line amounts to

$$\left(b^1, \ldots, b^n, 0\right) - (0, 0, \ldots, 1) = \lambda\left[\left(a^1, \ldots, a^{n+1}\right) - (0, 0, \ldots, 1)\right], \tag{1.2}$$

for some $\lambda \in \mathbb{R}$. By considering the last component in the vector equation (1.2) we have $0 - 1 = \lambda(a^{n+1} - 1)$; hence, $\lambda = 1/(1 - a^{n+1})$. Substituting this value of λ into (1.2) we find that the mapping $\phi : S^n \setminus \{(0, 0, \ldots, 1)\} \to \mathbb{R}^n$ defined by $\phi(a^1, \ldots, a^{n+1}) \equiv (b^1, \ldots, b^n)$ is given by

$$\phi\left(a^1, \ldots, a^{n+1}\right) = \frac{1}{1 - a^{n+1}}\left(a^1, \ldots, a^n\right). \tag{1.3}$$

The pair (U, ϕ), with $U \equiv S^n \setminus \{(0, 0, \ldots, 1)\}$, is a chart of coordinates, since ϕ is injective and $\phi(U) = \mathbb{R}^n$ (which is an open set in $\mathbb{R}^n$).

In a similar manner, joining the points of S^n with $(0, 0, \ldots, -1)$ by means of straight lines, another projection is obtained, $\chi : S^n \setminus \{(0, 0, \ldots, -1)\} \to \mathbb{R}^n$, given by

$$\chi\big(a^1, \ldots, a^{n+1}\big) = \frac{1}{1 + a^{n+1}}\big(a^1, \ldots, a^n\big) \tag{1.4}$$

so that (V, χ), with $V \equiv S^n \setminus \{(0, 0, \ldots, -1)\}$, is a second chart of coordinates which is C^∞-related with (ϕ, U). In effect, from (1.3) and (1.4) we find that

$$\begin{aligned}
\phi^{-1}\big(b^1, \ldots, b^n\big) &= \frac{1}{1 + \sum_{i=1}^n (b^i)^2}\left(2b^1, \ldots, 2b^n, -1 + \sum_{i=1}^n \big(b^i\big)^2\right), \\
\chi^{-1}\big(b^1, \ldots, b^n\big) &= \frac{1}{1 + \sum_{i=1}^n (b^i)^2}\left(2b^1, \ldots, 2b^n, 1 - \sum_{i=1}^n \big(b^i\big)^2\right),
\end{aligned} \tag{1.5}$$

and therefore

$$\big(\chi \circ \phi^{-1}\big)\big(b^1, \ldots, b^n\big) = \big(\phi \circ \chi^{-1}\big)\big(b^1, \ldots, b^n\big) = \frac{(b^1, \ldots, b^n)}{\sum_{i=1}^n (b^i)^2}.$$

We have $U \cap V = S^n \setminus \{(0, 0, \ldots, 1), (0, 0, \ldots, -1)\}$; hence $\phi(U \cap V) = \chi(U \cap V) = \mathbb{R}^n \setminus \{(0, 0, \ldots, 0)\}$, where the compositions $\chi \circ \phi^{-1}$ and $\phi \circ \chi^{-1}$ are differentiable of class C^∞. Since $S^n = U \cup V$, the charts (U, ϕ) and (V, χ) form a C^∞ subatlas for S^n.

The Cartesian product of two differentiable manifolds, M and N, acquires the structure of a differentiable manifold in a natural way. If $\{(U_i, \phi_i)\}$ and $\{(V_j, \psi_j)\}$ are subatlases of M and N, respectively, one can verify that $\{(U_i \times V_j, \rho_{ij})\}$ is a subatlas for $M \times N$, with $\rho_{ij}(p, q) \equiv (x^1(p), \ldots, x^n(p), y^1(q), \ldots, y^m(q))$, where $(x^1(p), \ldots, x^n(p)) = \phi(p)$ and $(y^1(q), \ldots, y^m(q)) = \psi_j(q)$.

Differentiability of Maps If f is a real-valued function defined on a differentiable manifold M, $f : M \to \mathbb{R}$, and (U, ϕ) is a chart belonging to the atlas of M, the composition $f \circ \phi^{-1}$ is a real-valued function defined on an open subset of $\mathbb{R}^n$, which may be differentiable or not (see Fig. 1.4). The differentiability of the composition $f \circ \phi^{-1}$ does not depend on the chart chosen, since the charts of the atlas of M are C^k-related (for some $k \geq 1$). From the identities

$$f \circ \phi^{-1} = \big(f \circ \chi^{-1}\big) \circ \big(\chi \circ \phi^{-1}\big), \qquad f \circ \chi^{-1} = \big(f \circ \phi^{-1}\big) \circ \big(\phi \circ \chi^{-1}\big)$$

it follows that $f \circ \phi^{-1}$ is differentiable if and only if $f \circ \chi^{-1}$ is. Hence, it makes sense to state the following definition. Let M be a differentiable C^k manifold. A function $f : M \to \mathbb{R}$ is *differentiable* of class C^r $(r \leq k)$ if $f \circ \phi^{-1}$ is differentiable of class C^r for every chart (U, ϕ) in the atlas of M.

For a fixed coordinate system (U, ϕ) belonging to the atlas of M, and a real-valued function $f : M \to \mathbb{R}$, letting $F \equiv f \circ \phi^{-1}$, we have [see (1.1)]

$$f(p) = \big(f \circ \phi^{-1}\big)\big(\phi(p)\big) = F\big(\phi(p)\big) = F\big(x^1(p), x^2(p), \ldots, x^n(p)\big),$$

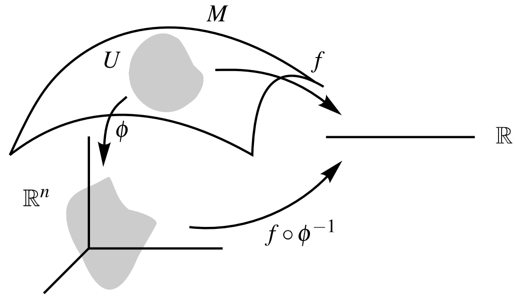

Fig. 1.4 With the aid of a coordinate system on M, a real-valued function f defined on M is represented by the function $f \circ \phi^{-1} : \mathbb{R}^n \to \mathbb{R}$

for $p \in U$. Thus, we write $f = F(x^1, x^2, \ldots, x^n)$; in this manner, the function f is expressed in terms of a real-valued function defined in (a subset of) $\mathbb{R}^n$.

Exercise 1.4 Let M be a C^k manifold. Show that the coordinates associated with any chart in the atlas of M are differentiable functions of class C^k. (*Hint*: if $\phi(p) = (x^1(p), x^2(p), \ldots, x^n(p))$, then $x^i = \pi^i \circ \phi$ where $\pi^i : \mathbb{R}^n \to \mathbb{R}$ is defined by $\pi^i(a^1, a^2, \ldots, a^n) = a^i$.)

If M is a C^k manifold and N is a C^l manifold, a map ψ from M into N is *differentiable* of class C^r (with $r \leq \min\{k, l\}$) if for any pair of charts (U, ϕ) on M and (V, χ) on N, the map $\chi \circ \psi \circ \phi^{-1}$ is differentiable of class C^r; that is, $\psi : M \to N$ is differentiable if, for $p \in M$, the coordinates of $\psi(p)$ depend differentiably on the coordinates of p (see Fig. 1.5). In fact, if $x^1, x^2, \ldots, x^n$ are the coordinates associated with the chart (U, ϕ) on M and $y^1, y^2, \ldots, y^m$ are the coordinates associated with the chart (V, χ) on N, we have

$$\begin{aligned}\big(y^1\big(\psi(p)\big), \ldots, y^m\big(\psi(p)\big)\big) &= \chi\big(\psi(p)\big) = \big(\chi \circ \psi \circ \phi^{-1}\big)\big(\phi(p)\big) \\ &= \big(\chi \circ \psi \circ \phi^{-1}\big)\big(x^1(p), \ldots, x^n(p)\big).\end{aligned}$$

A *diffeomorphism* ψ is a one-to-one map from a differentiable manifold M to a differentiable manifold N such that ψ and ψ^{-1} are differentiable; two differentiable manifolds M and N are *diffeomorphic* if there exists a diffeomorphism ψ from M onto N.

Exercise 1.5 Show that the set of diffeomorphisms of a manifold onto itself forms a group with the operation of composition.

Let M be a C^k manifold of dimension n. A subset N of M is a *submanifold* of M, of dimension m ($m \leq n$), if there exists a C^k subatlas of M, $\{(U_i, \phi_i)\}$, such that

$$\phi_i(N \cap U_i) = \big\{\big(a^1, a^2, \ldots, a^n\big) \in \mathbb{R}^n \mid a^{m+1} = a^{m+2} = \cdots = a^n = 0\big\}.$$

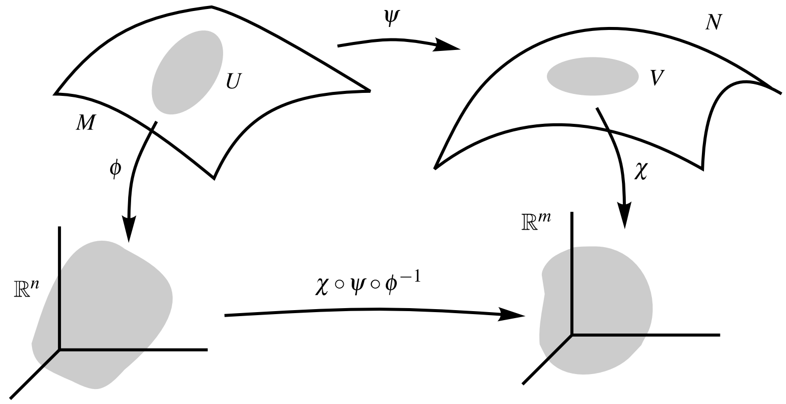

Fig. 1.5 The map $\psi : M \to N$ is locally represented by the map $\chi \circ \psi \circ \phi^{-1}$. ψ is differentiable if the compositions $\chi \circ \psi \circ \phi^{-1}$ are differentiable for any pair of charts (U, ϕ) on M and (V, χ) on N

Let π be the canonical projection from $\mathbb{R}^n$ onto $\mathbb{R}^m$ given by $\pi(a^1, a^2, \ldots, a^n) = (a^1, a^2, \ldots, a^m)$. The collection $\{(N \cap U_i, \pi \circ \phi_i)\}$ is a C^k subatlas on N, and N becomes a C^k manifold of dimension m with the atlas generated by this subatlas; in other words, N is a submanifold of dimension m if there exist coordinate systems (U, ϕ) on M such that if U intersects N, then $N \cap U = \{p \in U \mid x^{m+1}(p) = x^{m+2}(p) = \cdots = x^n(p) = 0\}$, where $x^1, x^2, \ldots, x^n$ are the coordinates associated to (U, ϕ).

With the aid of the following theorem we can construct or identify many examples of submanifolds.

Theorem 1.6 *Let $f^1, f^2, \ldots, f^m$ be real-valued differentiable functions defined on M. The set $N \equiv \{p \in M \mid f^1(p) = f^2(p) = \cdots = f^m(p) = 0\}$ is a submanifold of dimension $n - m$ of M if for any chart (U, ϕ) of the atlas of M such that U intersects N, the matrix with entries $D_i(f^j \circ \phi^{-1})|_{\phi(p)}$ $(1 \leq i \leq n, 1 \leq j \leq m)$ is of rank m for $p \in N$. (D_i stands for the ith partial derivative.)*

Proof Let $p \in N$ and let (U, ϕ) be a chart on M with $p \in U$. Assuming that the determinant of the square matrix $D_i(f^j \circ \phi^{-1})|_{\phi(p)}$ $(1 \leq i, j \leq m)$ is different from zero (which can be achieved by appropriately labeling the coordinates if necessary) and denoting by $x^1, x^2, \ldots, x^n$ the coordinates associated with (U, ϕ), the relations

$$\begin{aligned} &y^1 \equiv f^1, \qquad y^2 \equiv f^2, \quad \ldots, \quad y^m \equiv f^m, \\ &y^{m+1} \equiv x^{m+1}, \quad \ldots, \quad y^n \equiv x^n \end{aligned} \tag{1.6}$$

define a coordinate system in some subset V of U, that is, the x^i can be written as differentiable functions of the y^i. In the coordinates y^i the points p of N satisfy $y^1(p) = y^2(p) = \cdots = y^m(p) = 0$. Therefore, N is a submanifold of M of dimension $n - m$. □

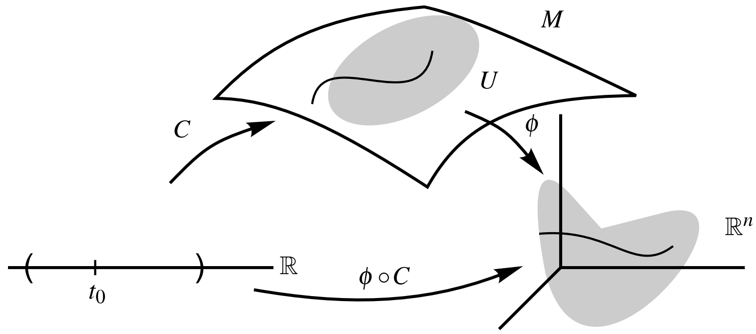

Fig. 1.6 A curve in M and its image in a coordinate system. C is differentiable if $\phi \circ C$ is differentiable for any chart (U, ϕ) on M

Example 1.7 Let $M = \mathbb{R}^3$ and $N \equiv \{p \in \mathbb{R}^3 \mid f(p) = 0\}$ with $f = x^2 + y^2 - z$, where (x, y, z) are the natural coordinates of $\mathbb{R}^3$. The matrix $(D_i(f \circ \phi^{-1})|_{\phi(p)})$, mentioned in Theorem 1.6, is the row matrix $(2x(p)\ \ 2y(p)\ \ -1)$, whose rank is equal to 1 at all the points of N (actually, it is equal to 1 everywhere). Thus, we conclude that N is a submanifold of $\mathbb{R}^3$ of dimension two. However, in order to see in detail how the proof of the theorem works, we shall explicitly show that N satisfies the definition of a submanifold given above.

It is convenient to relabel the coordinates, so that f takes the form $f = y^2 + z^2 - x$, because in that way the first entry of the matrix $(D_i(f \circ \phi^{-1})|_{\phi(p)})$ is always different from zero. Then, following the steps of the proof of the theorem, we introduce the coordinate system (u, v, w) [see (1.6)],

$$u = f = y^2 + z^2 - x, \qquad v = y, \qquad w = z,$$

on all of $\mathbb{R}^3$. From these expressions and their inverses, $x = v^2 + w^2 - u$, $y = v$, $z = w$, we see that the two coordinate systems are C^∞-related, and in terms of the coordinate system (u, v, w), each point $p \in N$ satisfies $u(p) = 0$.

Exercise 1.8 Show that if $x^1, x^2, \ldots, x^n$ are the natural coordinates of $\mathbb{R}^n$ (that is, the coordinates associated with the chart $(\mathbb{R}^n, \mathrm{id})$ of $\mathbb{R}^n$), then $\mathrm{S}^{n-1} \equiv \{p \in \mathbb{R}^n \mid (x^1(p))^2 + (x^2(p))^2 + \cdots + (x^n(p))^2 = 1\}$, is a submanifold of $\mathbb{R}^n$ of dimension $n - 1$.

Definition 1.9 Let M be a C^k manifold. A *differentiable curve*, C, of class C^r, in M, is a differentiable mapping of class C^r from an open subset of $\mathbb{R}$ into M; that is, $C : I \to M$ is a differentiable curve of class C^r in M if I is an open subset of $\mathbb{R}$ and $\phi \circ C$ is a differentiable map of class C^r for every chart (U, ϕ) of the atlas of M (see Fig. 1.6).

In what follows it will be assumed that all the objects dealt with (manifolds, maps, curves, etc.) are of class C^∞.

The set of all differentiable functions from M to $\mathbb{R}$ will be denoted by $C^\infty(M)$. This set is a ring with the operations given by

$$\begin{aligned}(f+g)(p) &\equiv f(p)+g(p)\\ (af)(p) &\equiv af(p)\\ (fg)(p) &\equiv f(p)g(p) \quad \text{for } f,g\in C^\infty(M),\ a\in\mathbb{R},\ \text{and } p\in M.\end{aligned}\tag{1.7}$$

If ψ is a differentiable map from M to a differentiable manifold N and $f\in C^\infty(N)$, the *pullback* of f under ψ, $\psi^* f$, is defined by

$$\psi^* f \equiv f\circ\psi. \tag{1.8}$$

From the relation $(\psi^* f)\circ\phi^{-1} = (f\circ\chi^{-1})\circ(\chi\circ\psi\circ\phi^{-1})$ it follows that $\psi^* f\in C^\infty(M)$. That is, $\psi^*: C^\infty(N)\to C^\infty(M)$ (ψ^* is applied to functions defined on N to produce functions defined on M; hence the name *pullback* for ψ^*).

Exercise 1.10 Show that $\psi^*(af+bg) = a\psi^* f + b\psi^* g$ and $\psi^*(fg) = (\psi^* f)(\psi^* g)$ for $f,g\in C^\infty(N)$ and $a,b\in\mathbb{R}$.

Exercise 1.11 Show that a map $\psi: M\to N$ is differentiable if and only if $\psi^* f\in C^\infty(M)$ for $f\in C^\infty(N)$.

Exercise 1.12 Show that if $\psi_1: M_1\to M_2$ and $\psi_2: M_2\to M_3$ are differentiable maps, then $(\psi_2\circ\psi_1)^* = \psi_1{}^*\circ\psi_2{}^*$.

1.2 The Tangent Space

If C is a differentiable curve in M and $f\in C^\infty(M)$, then $C^* f = f\circ C$ is a differentiable function from an open subset $I\subset\mathbb{R}$ into $\mathbb{R}$ (see Fig. 1.7). If $t_0\in I$, the *tangent vector to C* at the point $C(t_0)$, denoted by C'_{t_0}, is defined by

$$C'_{t_0}[f] \equiv \frac{\mathrm{d}}{\mathrm{d}t}(C^* f)\Big|_{t_0} = \lim_{t\to t_0}\frac{f(C(t)) - f(C(t_0))}{t-t_0}. \tag{1.9}$$

Hence, C'_{t_0} is a map from $C^\infty(M)$ into $\mathbb{R}$ with the properties (see Exercise 1.10)

$$\begin{aligned}C'_{t_0}[af+bg] &= \frac{\mathrm{d}}{\mathrm{d}t}\big(C^*(af+bg)\big)\Big|_{t_0}\\ &= \frac{\mathrm{d}}{\mathrm{d}t}(a\,C^* f + b\,C^* g)\Big|_{t_0}\\ &= a\,C'_{t_0}[f] + b\,C'_{t_0}[g], \quad \text{for } f,g\in C^\infty(M),\ a,b\in\mathbb{R},\end{aligned}$$

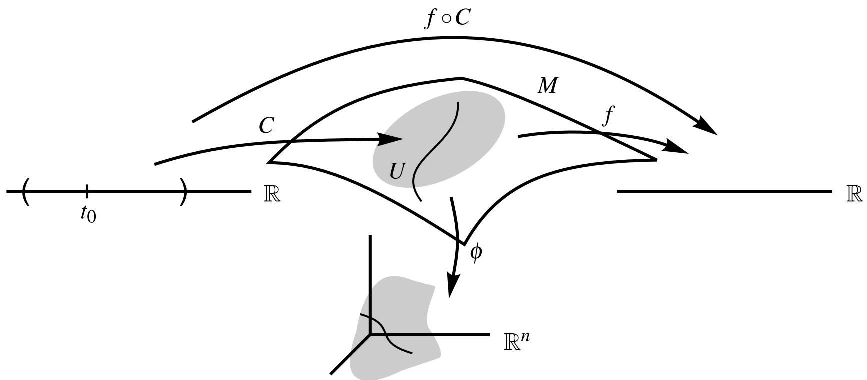

Fig. 1.7 Composition of a curve in M with a real-valued function f. The derivative of $f \circ C$, which is a function from $\mathbb{R}$ into $\mathbb{R}$, represents the directional derivative of f along C

and

$$\begin{aligned} C'_{t_0}[fg] &= \frac{\mathrm{d}}{\mathrm{d}t}\big(C^*(fg)\big)\Big|_{t_0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\big((C^*f)(C^*g)\big)\Big|_{t_0} \\ &= f\big(C(t_0)\big)\, C'_{t_0}[g] + g\big(C(t_0)\big)\, C'_{t_0}[f], \quad \text{for } f, g \in C^\infty(M). \end{aligned}$$

The real number $C'_{t_0}[f]$ is the rate of change of f along C around the point $C(t_0)$. The properties of the tangent vector to a curve lead to the following definition.

Definition 1.13 Let $p \in M$. A *tangent vector* to M at p is a map, v_p, of $C^\infty(M)$ in $\mathbb{R}$ such that

$$\begin{aligned} v_p[af + bg] &= a v_p[f] + b v_p[g] \\ v_p[fg] &= f(p)\, v_p[g] + g(p)\, v_p[f], \end{aligned} \tag{1.10}$$

for $f, g \in C^\infty(M)$, $a, b \in \mathbb{R}$.

For a constant function, c (denoting by c both the function and its value, i.e., $c(p) = c$ for all $p \in M$), we have

$$\begin{aligned} v_p[c] &= v_p[c \cdot 1] = c\, v_p[1] = c\, v_p[1 \cdot 1] \\ &= c\big(1 \cdot v_p[1] + 1 \cdot v_p[1]\big) = 2c\, v_p[1] = 2v_p[c]; \end{aligned}$$

therefore,

$$v_p[c] = 0. \tag{1.11}$$

The *tangent space* to M at p, denoted by T_pM (or by the symbols $T_p(M)$ and M_p), is the set of all the tangent vectors to M at p. The set T_pM is a real vector space with the operations defined by

$$\begin{aligned}(v_p + w_p)[f] &\equiv v_p[f] + w_p[f],\\ (a v_p)[f] &\equiv a\big(v_p[f]\big),\end{aligned} \tag{1.12}$$

for $v_p, w_p \in T_pM$, $f \in C^\infty(M)$, and $a, b \in \mathbb{R}$. Hence, 0_p, the zero vector of T_pM, satisfies $0_p[f] = 0$ for $f \in C^\infty(M)$.

If (U, ϕ) is a chart on M, with coordinates $x^1, x^2, \dots, x^n$ and $p \in U$, the tangent vectors, $(\partial/\partial x^1)_p, (\partial/\partial x^2)_p, \dots, (\partial/\partial x^n)_p$, are defined by

$$\left(\frac{\partial}{\partial x^i}\right)_p [f] \equiv D_i\big(f \circ \phi^{-1}\big)\big|_{\phi(p)}, \quad \text{for } f \in C^\infty(M), \tag{1.13}$$

where D_i denotes the partial derivative with respect to the ith argument; that is,

$$\begin{aligned}\left(\frac{\partial}{\partial x^i}\right)_p [f] = \lim_{t\to 0} \frac{1}{t}\Big[&\big(f \circ \phi^{-1}\big)\big(x^1(p), \dots, x^i(p) + t, \dots, x^n(p)\big)\\ &- \big(f \circ \phi^{-1}\big)\big(x^1(p), \dots, x^i(p), \dots, x^n(p)\big)\Big].\end{aligned} \tag{1.14}$$

Using the definition (1.13) one readily verifies that, in effect, $(\partial/\partial x^i)_p$ satisfies the conditions (1.10) and therefore $(\partial/\partial x^i)_p \in T_pM$.

Taking $f = x^j$ in (1.14) and noting that

$$\big(x^j \circ \phi^{-1}\big)\big(x^1(p), x^2(p), \dots, x^n(p)\big) = \big(x^j \circ \phi^{-1}\big)\big(\phi(p)\big) = x^j(p)$$

and, similarly,

$$\big(x^j \circ \phi^{-1}\big)\big(x^1(p), x^2(p), \dots, x^i(p) + t, \dots, x^n(p)\big) = \begin{cases} x^j(p) & \text{if } i \neq j,\\ x^j(p) + t & \text{if } i = j\end{cases}$$

(for t sufficiently small, so that all the points belong to U), we find that

$$\left(\frac{\partial}{\partial x^i}\right)_p \big[x^j\big] = \delta_i^j \equiv \begin{cases} 0 & \text{if } i \neq j,\\ 1 & \text{if } i = j.\end{cases} \tag{1.15}$$

The set $\{(\partial/\partial x^i)_p\}_{i=1}^n$ is linearly independent since if $a^i(\partial/\partial x^i)_p = 0_p$ (here and in what follows, any index that appears twice, once as a subscript and once as a superscript, implies a sum over all the values of the index, for instance, $a^i(\partial/\partial x^i)_p = \sum_{i=1}^n a^i(\partial/\partial x^i)_p$), then using (1.15) we have

$$0 = 0_p\big[x^j\big] = a^i \left(\frac{\partial}{\partial x^i}\right)_p \big[x^j\big] = a^i \delta_i^j = a^j.$$

Theorem 1.14 *If (U, ϕ) is a chart on M and $p \in U$, the set $\{(\partial/\partial x^i)_p\}_{i=1}^n$ is a basis of $T_p M$ and*

$$v_p = v_p\big[x^i\big]\left(\frac{\partial}{\partial x^i}\right)_p \tag{1.16}$$

for $v_p \in T_p M$.

Proof We only have to prove that any tangent vector to M at p can be expressed as a linear combination of the vectors $(\partial/\partial x^i)_p$. Let $f \in C^\infty(M)$. The composition $F \equiv f \circ \phi^{-1}$ is a real-valued function defined on $\phi(U)$, which is an open set of $\mathbb{R}^n$. For an arbitrary point $q \in U$, we have $f(q) = (f \circ \phi^{-1}) \circ (\phi(q)) = F(\phi(q))$ and, similarly, $f(p) = F(\phi(p))$. According to the mean value theorem for functions from $\mathbb{R}^n$ in $\mathbb{R}$, for a real-valued differentiable function, F, defined in some open subset of $\mathbb{R}^n$, given two points $(a^1, \dots, a^n)$ and $(b^1, \dots, b^n)$ such that the straight line segment joining them is contained in the domain of F, we have

$$F\big(b^1, \dots, b^n\big) - F\big(a^1, \dots, a^n\big) = \big(b^i - a^i\big) D_i F|_{(c^1, \dots, c^n)}, \tag{1.17}$$

where $(c^1, \dots, c^n)$ is some point on the straight line segment joining the points $(a^1, \dots, a^n)$ and $(b^1, \dots, b^n)$ [i.e., $(c^1, \dots, c^n) = (1 - t_0)(a^1, \dots, a^n) + t_0(b^1, \dots, b^n)$, for some $t_0 \in (0, 1)$]. Applying the formula (1.17) with $(a^1, \dots, a^n) = (x^1(p), \dots, x^n(p)) = \phi(p)$ and $(b^1, \dots, b^n) = (x^1(q), \dots, x^n(q)) = \phi(q)$ we obtain

$$F\big(\phi(q)\big) = F\big(\phi(p)\big) + \big[x^i(q) - x^i(p)\big] D_i F|_{(c^1, \dots, c^n)}. \tag{1.18}$$

Taking p fixed, the real numbers $D_i F|_{(c^1, \dots, c^n)}$ depend on q and will be denoted by $g_i(q)$; then (1.18) amounts to

$$f(q) = f(p) + \big[x^i(q) - x^i(p)\big]\, g_i(q)$$

or, since q is an arbitrary point in a neighborhood of p,

$$f = f(p) + \big[x^i - x^i(p)\big]\, g_i. \tag{1.19}$$

Using (1.10), (1.11), and the expression (1.19), taking into account that $f(p)$ as well as $x^i(p)$ are real numbers, while f, x^i, and g_i are real-valued functions defined in a neighborhood of p, for any $v_p \in T_p M$ we have

$$\begin{aligned} v_p[f] &= v_p\big[f(p)\big] + \big[x^i(p) - x^i(p)\big]\, v_p[g_i] + g_i(p)\, v_p\big[x^i - x^i(p)\big] \\ &= g_i(p)\, v_p\big[x^i\big], \end{aligned}$$

but $g_i(p) = D_i F|_{\phi(p)} = (\partial/\partial x^i)_p[f]$ [see (1.13)]. Therefore

$$v_p[f] = v_p\big[x^i\big]\left(\frac{\partial}{\partial x^i}\right)_p [f]$$

and, since f is arbitrary, we obtain the expression (1.16). As a corollary of this result we find that the dimension of $T_p M$ coincides with the dimension of M. □

According to (1.16), the tangent vector to a differentiable curve C in M ($C : I \to M$), at the point $C(t_0)$ is given by

$$C'_{t_0} = C'_{t_0}[x^i]\left(\frac{\partial}{\partial x^i}\right)_{C(t_0)}.$$

But, from (1.9), $C'_{t_0}[x^i] = \mathrm{d}(x^i \circ C)/\mathrm{d}t|_{t_0}$; therefore

$$C'_{t_0} = \left.\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\right|_{t_0}\left(\frac{\partial}{\partial x^i}\right)_{C(t_0)}. \tag{1.20}$$

Exercise 1.15 Let $v_p \in T_p M$. Show that there exists a curve C such that $v_p = C'_{t_0}$.

If (V, χ) is a second chart on M with coordinate functions $y^1, y^2, \ldots, y^n$, and $p \in U \cap V$, then we have another basis for $T_p M$ given by $\{(\partial/\partial y^i)_p\}_{i=1}^n$. From (1.16) we see that

$$\left(\frac{\partial}{\partial y^i}\right)_p = \left(\frac{\partial}{\partial y^i}\right)_p [x^j]\left(\frac{\partial}{\partial x^j}\right)_p.$$

It is convenient to write $(\partial f/\partial x^i)_p$ instead of $(\partial/\partial x^i)_p[f]$, keeping in mind the definition (1.13), so that (1.15) becomes $(\partial x^j/\partial x^i)_p = \delta_i^j$ and the foregoing relation can be expressed in the simpler form

$$\left(\frac{\partial}{\partial y^i}\right)_p = \left(\frac{\partial x^j}{\partial y^i}\right)_p\left(\frac{\partial}{\partial x^j}\right)_p \tag{1.21}$$

and, similarly,

$$\left(\frac{\partial}{\partial x^j}\right)_p = \left(\frac{\partial y^i}{\partial x^j}\right)_p\left(\frac{\partial}{\partial y^i}\right)_p, \tag{1.22}$$

which means that the two bases of $T_p M$, $\{(\partial/\partial x^i)_p\}_{i=1}^n$ and $\{(\partial/\partial y^i)_p\}_{i=1}^n$, are related by means of the matrix $c_i^j(p) = (\partial x^j/\partial y^i)_p$, whose inverse is the matrix $\tilde{c}_j^k(p) = (\partial y^k/\partial x^j)_p$.

Let M and N be two differentiable manifolds, and let $\psi : M \to N$ be a differentiable map (see Fig. 1.8). The map ψ induces a linear transformation between the tangent spaces $T_p M$ and $T_{\psi(p)} N$ called the *Jacobian* (or *differential*) of ψ at p, denoted by ψ_{*p} (or by $\mathrm{d}\psi_p$). If $v_p \in T_p M$, $\psi_{*p}(v_p)$ is defined as the tangent vector to N at $\psi(p)$ such that for $f \in C^\infty(N)$

$$\psi_{*p}(v_p)[f] \equiv v_p[\psi^* f]. \tag{1.23}$$

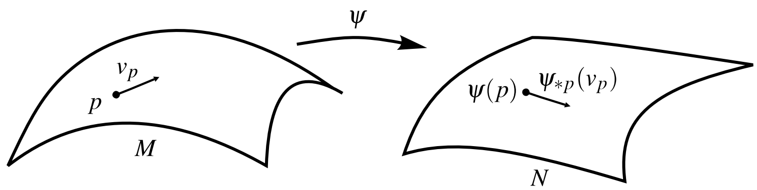

Fig. 1.8 If $\psi : M \to N$ is a differentiable mapping from M into N, its Jacobian, or differential, maps tangent vectors to M into tangent vectors to N

Exercise 1.16 Show that if $v_p \in T_p M$, then $\psi_{*p}(v_p) \in T_{\psi(p)}N$ and that ψ_{*p} is linear.

If $(x^1, x^2, \ldots, x^n)$ is a coordinate system on M about the point p and $(y^1, y^2, \ldots, y^m)$ is a coordinate system on N about $\psi(p)$, since $\psi_{*p}(\partial/\partial x^i)_p \in T_{\psi(p)}N$, using (1.16) we obtain the relation

$$\psi_{*p}\left(\frac{\partial}{\partial x^i}\right)_p = \psi_{*p}\left(\frac{\partial}{\partial x^i}\right)_p \left[y^j\right]\left(\frac{\partial}{\partial y^j}\right)_{\psi(p)}.$$

But from the definitions (1.23) and (1.8), $\psi_{*p}(\partial/\partial x^i)_p[y^j] = (\partial/\partial x^i)_p[\psi^* y^j] = (\partial/\partial x^i)_p[y^j \circ \psi]$; therefore

$$\psi_{*p}\left(\frac{\partial}{\partial x^i}\right)_p = \left(\frac{\partial(y^j \circ \psi)}{\partial x^i}\right)_p \left(\frac{\partial}{\partial y^j}\right)_{\psi(p)}. \tag{1.24}$$

In other words, the matrix with entries $(\partial(y^j \circ \psi)/\partial x^i)_p$ represents the linear transformation ψ_{*p} with respect to the bases $\{(\partial/\partial x^i)_p\}_{i=1}^n$ and $\{(\partial/\partial y^j)_{\psi(p)}\}_{j=1}^m$ (compare with the usual definition of the Jacobian matrix in the calculus of several variables).

If $\psi_1 : M_1 \to M_2$ and $\psi_2 : M_2 \to M_3$ are differentiable maps between differentiable manifolds, then, for $v_p \in T_p M_1$ and $f \in C^\infty(M_3)$, using (1.23) and Exercise 1.12, we have

$$\begin{aligned}(\psi_2 \circ \psi_1)_{*p}(v_p)[f] &= v_p\big[(\psi_2 \circ \psi_1)^* f\big] = v_p\big[({\psi_1}^* \circ {\psi_2}^*) f\big] \\ &= v_p\big[{\psi_1}^*({\psi_2}^* f)\big] = \psi_{1*p}(v_p)[{\psi_2}^* f] \\ &= \psi_{2*\psi_1(p)}\big(\psi_{1*p}(v_p)\big)[f],\end{aligned}$$

i.e.,

$$(\psi_2 \circ \psi_1)_{*p} = \psi_{2*\psi_1(p)} \circ \psi_{1*p}. \tag{1.25}$$

This relation is called the *chain rule*.

If $\psi : M \to N$ is a differentiable map between differentiable manifolds and $C : I \to M$ is a curve in M, the composition $\psi \circ C$ is a curve in N. According

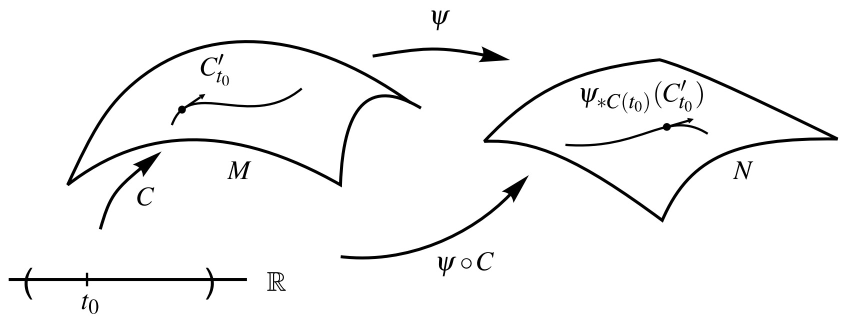

Fig. 1.9 The tangent vectors of the curve $\psi \circ C$ are obtained applying the Jacobian of ψ to the tangent vectors of C

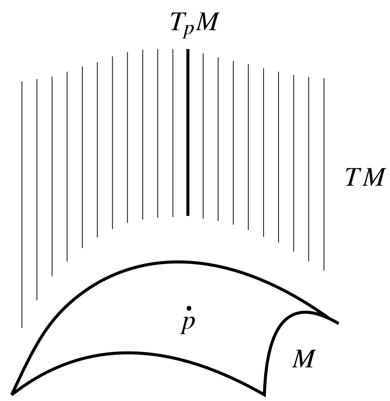

Fig. 1.10 The tangent bundle of M is formed by the union of the tangent spaces to M at all the points of M. Each tangent space to M is represented here by a *vertical line*

to (1.9) and (1.23), the tangent vector to $\psi \circ C$ at the point $(\psi \circ C)(t_0) = \psi(C(t_0))$ satisfies

$$\begin{aligned}(\psi \circ C)'_{t_0}[f] &= \frac{\mathrm{d}}{\mathrm{d}t}(f \circ \psi \circ C)\Big|_{t_0} = C'_{t_0}[f \circ \psi] \\ &= C'_{t_0}[\psi^* f] = \psi_{*C(t_0)}(C'_{t_0})[f], \quad \text{for } f \in C^\infty(N).\end{aligned}$$

Hence

$$(\psi \circ C)'_{t_0} = \psi_{*C(t_0)}(C'_{t_0}), \tag{1.26}$$

which means that the tangent vectors to the image of a curve C under the map ψ are the images of the tangent vectors to C under the Jacobian of ψ (see Fig. 1.9).

The Tangent Bundle of a Manifold The *tangent bundle* of a differentiable manifold M, denoted by TM, is the set of all tangent vectors at all points of M; that is, $TM = \bigcup_{p \in M} T_p M$. The *canonical projection*, π, from TM on M is the mapping that associates to each element of TM the point of M at which it is attached; that is, if $v_p \in T_p M$, then $\pi(v_p) = p$. Therefore, $\pi^{-1}(p) = T_p M$ (see Fig. 1.10).

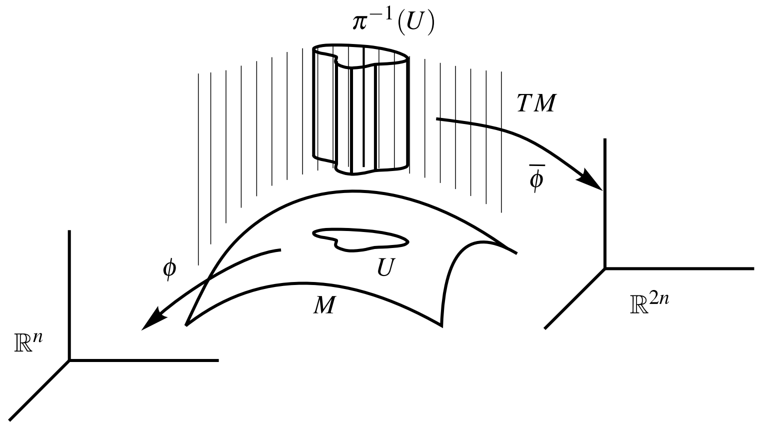

Fig. 1.11 Each coordinate system on M, (U,ϕ), induces a coordinate system on TM, $(\pi^{-1}(U),\overline{\phi})$

The tangent bundle has the structure of a differentiable manifold induced by the structure of M in a natural way. If (U,ϕ) is a coordinate system on M, each $v_p \in \pi^{-1}(U)$ is a linear combination of the vectors $(\partial/\partial x^i)_p$, with real coefficients that depend on v_p. Hence, we can write

$$v_p = \dot{q}^i(v_p)\left(\frac{\partial}{\partial x^i}\right)_p, \tag{1.27}$$

which defines n functions $\dot{q}^i : \pi^{-1}(U) \to \mathbb{R}$. (This notation comes from that commonly employed in Lagrangian mechanics, when M is the configuration space of a mechanical system.) From (1.15) we also have

$$\dot{q}^i(v_p) = v_p\big[x^i\big]. \tag{1.28}$$

Defining the n functions $q^i : \pi^{-1}(U) \to \mathbb{R}$, by $q^i \equiv x^i \circ \pi = \pi^* x^i$, the pair $(\pi^{-1}(U),\overline{\phi})$, with

$$\overline{\phi}(v_p) \equiv \big(q^1(v_p),\ldots,q^n(v_p),\dot{q}^1(v_p),\ldots,\dot{q}^n(v_p)\big),$$

is a chart on TM (see Fig. 1.11). (The image of $\pi^{-1}(U)$ under $\overline{\phi}$ is $\phi(U)\times\mathbb{R}^n$, which is an open subset of $\mathbb{R}^{2n}$, and the injectivity of $\overline{\phi}$ follows from the injectivity of ϕ and the fact that if two vectors have the same components with respect to a basis, they must be the same vector.)

Two coordinate systems (U,ϕ) and (U',ϕ') on M, C^k-related, induce the coordinate systems $(\pi^{-1}(U),\overline{\phi})$ and $(\pi^{-1}(U'),\overline{\phi'})$ on TM, which are C^{k-1}-related, since from (1.27) [or from (1.21) and (1.22)] it follows that the coordinates $\dot{q}^i$ and $\dot{q}'^i$ are related by

$$\dot{q}^i = \dot{q}'^j\,\pi^*\left(\frac{\partial x^i}{\partial x'^j}\right), \qquad \dot{q}'^i = \dot{q}^j\,\pi^*\left(\frac{\partial x'^i}{\partial x^j}\right),$$

where x^i denotes the coordinates associated with (U, ϕ), x'^i those associated with (U', ϕ'), while $\dot{q}^i$ and $\dot{q}'^i$ are the coordinates induced on *TM* by (U, ϕ) and (U', ϕ'), respectively. Thus, if $\{(U_i, \phi_i)\}$ is a subatlas on M, $\{(\pi^{-1}(U_i), \overline{\phi}_i)\}$ is a subatlas on *TM* that defines a differentiable manifold structure.

Since, by definition, $\pi^* x^i = q^i$, the projection π is differentiable. Moreover, from (1.24) we obtain

$$\pi_{*v}\left(\frac{\partial}{\partial q^i}\right)_v = \left(\frac{\partial}{\partial x^i}\right)_{\pi(v)}, \qquad \pi_{*v}\left(\frac{\partial}{\partial \dot{q}^i}\right)_v = 0, \tag{1.29}$$

for $v \in \pi^{-1}(U)$.

Exercise 1.17 With the notation employed above, show that

$$\pi^*\left(\frac{\partial x'^i}{\partial x^j}\right) = \frac{\partial q'^i}{\partial q^j} \quad \text{and} \quad \pi^*\left(\frac{\partial f}{\partial x^j}\right) = \frac{\partial(\pi^* f)}{\partial q^j},$$

for $f \in C^\infty(M)$.

The tangent bundle and the cotangent bundle (defined in Sect. 8.1) of a manifold are two examples of vector bundles and fiber bundles. We are not giving here the definitions of these more ample concepts, since we will not make use of them. However, the vector bundles and the fiber bundles are two very useful concepts in manifold theory and topology. Some introductory presentations can be found in Crampin and Pirani (1986), Lee (1997), Isham (1999), and Conlon (2001).

1.3 Vector Fields

A *vector field* $\mathbf{X}$, on M, is a function that to each point p of M assigns a tangent vector $\mathbf{X}(p) \in T_p M$. The tangent vector $\mathbf{X}(p)$ is also denoted by $\mathbf{X}_p$. A vector field may not be defined in all of M (for instance, its domain may be the image of a curve); but when a vector field is defined in all of M we say that it is defined globally, otherwise we say that it is defined only locally.

Since a vector field gives us a tangent vector at each point of its domain and a tangent vector can be applied to real-valued differentiable functions to yield real numbers, given a vector field $\mathbf{X}$ and $f \in C^\infty(M)$, we can form a real-valued function $\mathbf{X}f$, defined by

$$(\mathbf{X}f)(p) \equiv \mathbf{X}_p[f]. \tag{1.30}$$

Since $\mathbf{X}_p \in T_p M$, from (1.10) it follows that

$$\mathbf{X}(af + bg) = a\mathbf{X}f + b\mathbf{X}g \quad \text{and} \quad \mathbf{X}(fg) = f\mathbf{X}g + g\mathbf{X}f, \tag{1.31}$$

for $f, g \in C^\infty(M)$ and $a, b \in \mathbb{R}$.

A vector field $\mathbf{X}$ is differentiable (of class C^∞) if for all $f \in C^\infty(M)$, the function $\mathbf{X}f$ also belongs to $C^\infty(M)$. The set of all differentiable vector fields on M will be denoted by $\mathfrak{X}(M)$. Vector fields can be combined by means of the operations

$$\begin{aligned}(a\mathbf{X}+b\mathbf{Y})_p &\equiv a\mathbf{X}_p + b\mathbf{Y}_p,\\ (f\mathbf{X})_p &\equiv f(p)\mathbf{X}_p\end{aligned} \tag{1.32}$$

for $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$, $a, b \in \mathbb{R}$, and $f \in C^\infty(M)$. Making use of the definitions above, one verifies directly that $a\mathbf{X} + b\mathbf{Y}$ and $f\mathbf{X}$ are vector fields.

Exercise 1.18 Let $\mathbf{X}$ and $\mathbf{Y}$ be two vector fields on M. Show that

$$(a\mathbf{X}+b\mathbf{Y})f = a\mathbf{X}f + b\mathbf{Y}f, \tag{1.33}$$

$$(g\mathbf{X})f = g(\mathbf{X}f), \tag{1.34}$$

for $a, b \in \mathbb{R}$ and $f, g \in C^\infty(M)$.

If (U, ϕ) is a chart on M, we have n vector fields, $(\partial/\partial x^i)$, $i = 1, 2, \ldots, n$, on U defined by $(\partial/\partial x^i)(p) \equiv (\partial/\partial x^i)_p$. These vector fields are differentiable since from (1.30) and (1.13) we see that, for any $f \in C^\infty(M)$,

$$\left(\frac{\partial}{\partial x^i}\right) f = \left[D_i\left(f \circ \phi^{-1}\right)\right] \circ \phi. \tag{1.35}$$

Below $(\partial/\partial x^i) f$ will be also written as $\partial f/\partial x^i$, keeping in mind that these functions are defined by (1.35).

Since the tangent vectors $(\partial/\partial x^i)_p$ form a basis of T_pM, any vector field $\mathbf{X}$ evaluated at the point p must be a linear combination of the vectors $(\partial/\partial x^i)_p$ with real coefficients, which may depend on p. Therefore

$$\mathbf{X}_p = X^i(p)\left(\frac{\partial}{\partial x^i}\right)_p.$$

This relation defines n real-valued functions $X^1, X^2, \ldots, X^n$ in the intersection of U and the domain of $\mathbf{X}$. Making use of the operations (1.32) we have

$$\mathbf{X}_p = X^i(p)\left(\frac{\partial}{\partial x^i}\right)(p) = \left[X^i\left(\frac{\partial}{\partial x^i}\right)\right](p),$$

hence

$$\mathbf{X} = X^i\left(\frac{\partial}{\partial x^i}\right). \tag{1.36}$$

(Strictly speaking, the left-hand side of this last equation is the restriction of $\mathbf{X}$ to the intersection of U and the domain of $\mathbf{X}$, denoted by $\mathbf{X}|_V$, where V is the intersection of U and the domain of $\mathbf{X}$.)

Exercise 1.19 Let $\mathbf{X} = X^i(\partial/\partial x^i)$. Show that the functions X^i are given by $X^i = \mathbf{X}x^i$ and that $\mathbf{X}$ is differentiable if and only if the functions X^i are.

Exercise 1.20 Let $(x^1, x^2, \ldots, x^n)$ and $(x'^1, x'^2, \ldots, x'^n)$ be two coordinate systems. Show that if $\mathbf{X} = X^i(\partial/\partial x^i)$ and $\mathbf{X} = X'^j(\partial/\partial x'^j)$, then

$$X'^j = X^i \frac{\partial x'^j}{\partial x^i},$$

in the intersection of the domains of $\mathbf{X}$ and those of the two coordinate systems. (This last expression is the definition of a *contravariant* vector field in the tensor formalism.)

There is another operation between vector fields, called the *Lie bracket*, with which $\mathfrak{X}(M)$ becomes a Lie algebra over $\mathbb{R}$ (see Appendix A). If $\mathbf{X}$ and $\mathbf{Y}$ are vector fields on M, their Lie bracket is defined by

$$[\mathbf{X}, \mathbf{Y}]f \equiv \mathbf{X}(\mathbf{Y}f) - \mathbf{Y}(\mathbf{X}f) \quad \text{for } f \in C^\infty(M). \tag{1.37}$$

Then $[\mathbf{X}, \mathbf{Y}] = -[\mathbf{Y}, \mathbf{X}]$.

Exercise 1.21 Show that if $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(M)$ then $[\mathbf{X}, \mathbf{Y}] \in \mathfrak{X}(M)$ and $[\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]] + [\mathbf{Y}, [\mathbf{Z}, \mathbf{X}]] + [\mathbf{Z}, [\mathbf{X}, \mathbf{Y}]] = 0$.

Exercise 1.22 Show that $[f\mathbf{X}, g\mathbf{Y}] = fg[\mathbf{X}, \mathbf{Y}] + f(\mathbf{X}g)\mathbf{Y} - g(\mathbf{Y}f)\mathbf{X}$, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$ and $f, g \in C^\infty(M)$.

If (U, ϕ) is a chart on M with coordinates $x^1, x^2, \ldots, x^n$, from (1.35) we have

$$\begin{aligned}
&\big[(\partial/\partial x^i), (\partial/\partial x^j)\big]f \\
&= \left(\frac{\partial}{\partial x^i}\right)\big\{\big[D_j\big(f \circ \phi^{-1}\big)\big] \circ \phi\big\} - \left(\frac{\partial}{\partial x^j}\right)\big\{\big[D_i\big(f \circ \phi^{-1}\big)\big] \circ \phi\big\} \\
&= \big\{D_i D_j\big(f \circ \phi^{-1}\big) - D_j D_i\big(f \circ \phi^{-1}\big)\big\} \circ \phi \\
&= 0,
\end{aligned}$$

for $f \in C^\infty(M)$; hence

$$\big[(\partial/\partial x^i), (\partial/\partial x^j)\big] = 0. \tag{1.38}$$

Exercise 1.23 Show that if $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$ are given by $\mathbf{X} = X^i(\partial/\partial x^i)$ and $\mathbf{Y} = Y^j(\partial/\partial x^j)$, then $[\mathbf{X}, \mathbf{Y}] = (\mathbf{X}Y^i - \mathbf{Y}X^i)(\partial/\partial x^i)$. (*Hint*: use the result of the first part of Exercise 1.19.)

Exercise 1.24 Compute the Lie brackets of the vector fields

$$\mathbf{X} = \big(1 + r^2\big) \sin\theta \frac{\partial}{\partial r} + \frac{1 - r^2}{r} \cos\theta \frac{\partial}{\partial \theta},$$

$$\mathbf{Y} = -\left(1+r^2\right)\cos\theta\frac{\partial}{\partial r} + \frac{1-r^2}{r}\sin\theta\frac{\partial}{\partial\theta},$$
$$\mathbf{Z} = \frac{\partial}{\partial\theta}.$$

As shown above, each coordinate system gives rise to a set of n vector fields that can be used to express an arbitrary vector field in the local form (1.36), and that satisfy the relations (1.38). Nevertheless, an arbitrary vector field can also be expressed in a form analogous to (1.36) in terms of any set of n vector fields such that at each point of their common domain forms a basis for the tangent space. The use of this kind of set of vector fields, not necessarily associated with coordinate systems, may be convenient when there exists some additional structure on the manifold (e.g., a connection, a metric tensor or a Lie group structure), as shown, e.g., in Sects. 5.3, 6.2, 6.3, 6.4, 7.2, and 7.5.

While any differentiable mapping from a manifold into another manifold allows us to map tangent vectors to the first manifold into tangent vectors to the second one (by means of the Jacobian of the map), not any differentiable map between manifolds allows us to map a vector field on the first manifold into a vector field on the second one. For instance, if a differentiable map $\psi : M \to N$ is not injective, there exist two different points p and q, belonging to M, which have the same image under ψ; however, for a vector field $\mathbf{X}$ on M, the tangent vectors $\psi_{*p}\mathbf{X}_p$ and $\psi_{*q}\mathbf{X}_q$ need not coincide.

Let $\psi : M \to N$ be a differentiable map between differentiable manifolds. If $\mathbf{X} \in \mathfrak{X}(M)$ and $\mathbf{Y} \in \mathfrak{X}(N)$, we say that $\mathbf{X}$ and $\mathbf{Y}$ are ψ*-related* if

$$\mathbf{Y}_{\psi(p)} = \psi_{*p}\mathbf{X}_p, \quad \text{for } p \in M. \tag{1.39}$$

From (1.30) and (1.23) it follows that if $f \in C^\infty(N)$, then

$$\begin{aligned}(\mathbf{Y}f)\big(\psi(p)\big) &= \mathbf{Y}_{\psi(p)}[f] = \psi_{*p}\mathbf{X}_p[f] = \mathbf{X}_p[f\circ\psi] \\ &= \big(\mathbf{X}(f\circ\psi)\big)(p), \quad \text{for } p \in M,\end{aligned}$$

that is

$$(\mathbf{Y}f)\circ\psi = \mathbf{X}(f\circ\psi), \quad \text{for } f \in C^\infty(N). \tag{1.40}$$

For example, according to Exercise (1.17), the vector fields $\partial/\partial q^j$ and $\partial/\partial x^j$ are π-related.

If $\mathbf{X}_1, \mathbf{X}_2 \in \mathfrak{X}(M)$ are ψ-related with $\mathbf{Y}_1, \mathbf{Y}_2 \in \mathfrak{X}(N)$, respectively, then $[\mathbf{X}_1, \mathbf{X}_2]$ is ψ-related with $[\mathbf{Y}_1, \mathbf{Y}_2]$, since, by hypothesis, $(\mathbf{Y}_1 f)\circ\psi = \mathbf{X}_1(f\circ\psi)$ and $(\mathbf{Y}_2 g)\circ\psi = \mathbf{X}_2(g\circ\psi)$, for $f, g \in C^\infty(N)$ [see (1.40)]. Taking $g = \mathbf{Y}_1 f$, we have

$$\big[\mathbf{Y}_2(\mathbf{Y}_1 f)\big]\circ\psi = \mathbf{X}_2\big((\mathbf{Y}_1 f)\circ\psi\big) = \mathbf{X}_2\big(\mathbf{X}_1(f\circ\psi)\big),$$

and a similar relation is obtained by interchanging the indices 1 and 2. Then

$$\big([\mathbf{Y}_1, \mathbf{Y}_2]f\big)\circ\psi = \big(\mathbf{Y}_1(\mathbf{Y}_2 f) - \mathbf{Y}_2(\mathbf{Y}_1 f)\big)\circ\psi$$

$$= \mathbf{X}_1\big(\mathbf{X}_2(f \circ \psi)\big) - \mathbf{X}_2\big(\mathbf{X}_1(f \circ \psi)\big)$$
$$= [\mathbf{X}_1, \mathbf{X}_2](f \circ \psi).$$

Since a vector field on M, $\mathbf{X}$, is a function that maps each point $p \in M$ into an element $\mathbf{X}_p \in T_pM \subset TM$, $\mathbf{X}$ is a function from M into TM such that $\pi \circ \mathbf{X} = \mathrm{id}_M$, where π is the projection of the tangent bundle TM onto M.

Exercise 1.25 Show that a vector field $\mathbf{X}$ on M is differentiable if and only if the function $p \mapsto \mathbf{X}_p$, from M into TM, is differentiable. (*Hint*: prove that if $\mathbf{X}$ is given locally by (1.36), then $\mathbf{X}^* q^i = x^i$ and $\mathbf{X}^* \dot{q}^i = X^i$, where the q^i and $\dot{q}^i$ are the coordinates induced on TM by a system of coordinates x^i on M.)

1.4 1-Forms and Tensor Fields

Let $f \in C^\infty(M)$; the *differential* of f at the point p ($p \in M$), denoted by $\mathrm{d}f_p$, is defined by

$$\mathrm{d}f_p(v_p) \equiv v_p[f], \quad \text{for } v_p \in T_pM. \tag{1.41}$$

The map $\mathrm{d}f_p$ is a linear transformation from T_pM in $\mathbb{R}$, since if $v_p, w_p \in T_pM$ and $a, b \in \mathbb{R}$, from (1.41) and (1.12) we have

$$\begin{aligned} \mathrm{d}f_p(av_p + bw_p) &= (av_p + bw_p)[f] \\ &= av_p[f] + bw_p[f] \\ &= a\,\mathrm{d}f_p(v_p) + b\,\mathrm{d}f_p(w_p). \end{aligned}$$

This means that $\mathrm{d}f_p$ belongs to the dual space of T_pM, denoted by T_p^*M. By definition, the elements of T_p^*M are the linear transformations from T_pM in $\mathbb{R}$, which are called *covectors* or *covariant vectors*, while T_p^*M is called the *cotangent space* to M at p. The space T_p^*M is a vector space over $\mathbb{R}$ with the operations

$$(\alpha_p + \beta_p)(v_p) \equiv \alpha_p(v_p) + \beta_p(v_p), \qquad (a\alpha_p)(v_p) \equiv a\big(\alpha_p(v_p)\big), \tag{1.42}$$

for $\alpha_p, \beta_p \in T_p^*M$, $v_p \in T_pM$, and $a \in \mathbb{R}$.

A *covector field* α on M is a map that assigns to each $p \in M$ an element $\alpha(p) \in T_p^*M$. The covector $\alpha(p)$ will also be denoted by α_p. A covector field α is differentiable (of class C^∞) if for all $\mathbf{X} \in \mathfrak{X}(M)$ the function $\alpha(\mathbf{X})$ defined by

$$\big(\alpha(\mathbf{X})\big)(p) \equiv \alpha_p(\mathbf{X}_p) \tag{1.43}$$

is differentiable (of class C^∞).

The function $\alpha(\mathbf{X})$ is also denoted by $\mathbf{X} \lrcorner \alpha$ (which allows us to reduce the repeated use of parentheses with various purposes) and by $i(\mathbf{X})\alpha$, $i_{\mathbf{X}}\alpha$, or $\langle \mathbf{X}, \alpha \rangle$. This operation is called *contraction* or *interior product*.

The set of all differentiable covector fields on M will be denoted by $\Lambda^1(M)$. The set $\Lambda^1(M)$ is a module over $C^\infty(M)$ with the operations given by

$$\begin{aligned}(\alpha+\beta)_p &\equiv \alpha_p + \beta_p,\\ (f\alpha)_p &\equiv f(p)\,\alpha_p,\end{aligned} \tag{1.44}$$

for $\alpha, \beta \in \Lambda^1(M)$ and $f \in C^\infty(M)$. The elements of $\Lambda^1(M)$ are called *linear differential forms* or 1-*forms*.

If $f \in C^\infty(M)$, the differential of f, denoted by $\mathrm{d}f$ and given by $\mathrm{d}f(p) \equiv \mathrm{d}f_p$, is a differentiable covector field or 1-form [i.e., $\mathrm{d}f \in \Lambda^1(M)$], since if $\mathbf{X} \in \mathfrak{X}(M)$, then from (1.41) and (1.30) it follows that

$$\big(\mathrm{d}f(\mathbf{X})\big)(p) = \mathrm{d}f_p(\mathbf{X}_p) = \mathbf{X}_p[f] = (\mathbf{X}f)(p),$$

for $p \in M$; that is,

$$\mathrm{d}f(\mathbf{X}) = \mathbf{X}f \tag{1.45}$$

(or, equivalently, $\mathbf{X} \lrcorner \mathrm{d}f = \mathbf{X}f$), which is a differentiable function for all $\mathbf{X} \in \mathfrak{X}(M)$, thus verifying that $\mathrm{d}f$ is, indeed, a differentiable covector field.

From (1.45), (1.31), (1.44), and (1.42) it follows that the map $\mathrm{d} : C^\infty(M) \to \Lambda^1(M)$, which sends f into $\mathrm{d}f$, satisfies

$$\begin{aligned}\mathrm{d}(af+bg)(\mathbf{X}) &= \mathbf{X}(af+bg) = a\mathbf{X}f + b\mathbf{X}g\\ &= a\,\mathrm{d}f(\mathbf{X}) + b\,\mathrm{d}g(\mathbf{X}) = (a\,\mathrm{d}f + b\,\mathrm{d}g)(\mathbf{X}),\end{aligned}$$

for $\mathbf{X} \in \mathfrak{X}(M)$; therefore

$$\mathrm{d}(af+bg) = a\,\mathrm{d}f + b\,\mathrm{d}g, \quad \text{for } f, g \in C^\infty(M) \text{ and } a, b \in \mathbb{R}. \tag{1.46}$$

Similarly, from (1.45), (1.31), and (1.42),

$$\begin{aligned}\mathrm{d}(fg)(\mathbf{X}) &= \mathbf{X}(fg) = f\mathbf{X}g + g\mathbf{X}f\\ &= f\,\mathrm{d}g(\mathbf{X}) + g\,\mathrm{d}f(\mathbf{X}) = (f\,\mathrm{d}g + g\,\mathrm{d}f)(\mathbf{X}), \quad \text{for } \mathbf{X} \in \mathfrak{X}(M),\end{aligned}$$

hence,

$$\mathrm{d}(fg) = f\,\mathrm{d}g + g\,\mathrm{d}f, \quad \text{for } f, g \in C^\infty(M). \tag{1.47}$$

If (U, ϕ) is a chart on M, then (1.41) and (1.15) imply that the differential of the coordinate functions $x^1, x^2, \ldots, x^n$ satisfies

$$\mathrm{d}x^i_p\left(\left(\frac{\partial}{\partial x^j}\right)_p\right) = \left(\frac{\partial}{\partial x^j}\right)_p \left[x^i\right] = \delta^i_j. \tag{1.48}$$

This relation implies that $\{\mathrm{d}x^i_p\}_{i=1}^n$ is a basis of T^*_pM, since if a linear combination, with real coefficients, $a_i\,\mathrm{d}x^i_p$, is equal to the zero covector, we have

$0 = (a_i \, \mathrm{d}x^i_p)((\partial/\partial x^j)_p) = a_i \delta^i_j = a_j$. Furthermore, if $\alpha_p \in T^*_p M$, for any tangent vector $v_p \in T_p M$ expressed in the form $v_p = v_p[x^i](\partial/\partial x^i)_p$ [see (1.16)], then we find

$$\alpha_p(v_p) = \alpha_p\left(v_p\left[x^i\right]\left(\frac{\partial}{\partial x^i}\right)_p\right) = v_p\left[x^i\right]\alpha_p\left(\left(\frac{\partial}{\partial x^i}\right)_p\right),$$

but, according to (1.41), $v_p[x^i] = \mathrm{d}x^i_p(v_p)$. Therefore

$$\alpha_p(v_p) = \alpha_p\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\mathrm{d}x^i_p(v_p) = \left[\alpha_p\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\mathrm{d}x^i_p\right](v_p),$$

and since v_p is arbitrary, we have

$$\alpha_p = \alpha_p\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\mathrm{d}x^i_p \tag{1.49}$$

[cf. (1.16)].

If α is a covector field on M, using (1.49), (1.43), and (1.44) it follows that the covector $\alpha(p) \in T^*_p M$ is expressed as

$$\begin{aligned}\alpha(p) &= \alpha(p)\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\mathrm{d}x^i_p = \left[\alpha\left(\left(\frac{\partial}{\partial x^i}\right)\right)\right](p)\,\mathrm{d}x^i(p)\\ &= \left[\alpha\left(\left(\frac{\partial}{\partial x^i}\right)\right)\mathrm{d}x^i\right](p);\end{aligned}$$

that is,

$$\alpha = \alpha\left(\left(\frac{\partial}{\partial x^i}\right)\right)\mathrm{d}x^i. \tag{1.50}$$

Denoting the real-valued functions $\alpha((\partial/\partial x^i))$ by α_i we conclude that any covector field is locally expressed (i.e., in the domain of a local chart of coordinates) in the form

$$\alpha = \alpha_i \, \mathrm{d}x^i. \tag{1.51}$$

Exercise 1.26 Show that α is a differentiable covector field if and only if the functions α_i are differentiable.

Exercise 1.27 Let $(x^1, x^2, \ldots, x^n)$ and $(x'^1, x'^2, \ldots, x'^n)$ be two coordinate systems. Show that if $\alpha = \alpha_i \, \mathrm{d}x^i$ and $\alpha = \alpha'_j \, \mathrm{d}x'^j$, then

$$\alpha'_j = \alpha_i \frac{\partial x^i}{\partial x'^j}$$

in the common domain of α and the two systems of coordinates (cf. Exercise 1.20). (This relation is taken as the definition of a covariant vector field in the tensor formalism.)

The local expression of the differential of a function $f \in C^\infty(M)$ is, according to (1.50), $\mathrm{d}f = [\mathrm{d}f((\partial/\partial x^i))]\,\mathrm{d}x^i$; but, by virtue of (1.45), $\mathrm{d}f((\partial/\partial x^i)) = (\partial/\partial x^i)f$, so that

$$\mathrm{d}f = \frac{\partial f}{\partial x^i}\,\mathrm{d}x^i, \tag{1.52}$$

which agrees with the expression for the total differential of a function of several variables, as defined in textbooks on the calculus of several variables.

Example 1.28 The linear differential forms and the differential forms of degree greater than 1, defined in Chap. 3, correspond to the integrands of the line integrals, surface integrals, and so on, encountered in various areas of mathematics and physics [see, e.g., Guillemin and Pollack (1974), do Carmo (1994), Lee (2002)]. If $C : [a, b] \mapsto M$ is a differentiable curve in M (that is, C is the restriction to $[a, b]$ of a differentiable map of an open subset of $\mathbb{R}$ containing $[a, b]$ to M) and α is a linear differential form on M, then the line integral of α on C is defined by

$$\int_C \alpha \equiv \int_a^b \alpha_{C(t)}\big(C'(t)\big)\,\mathrm{d}t, \tag{1.53}$$

where the integral on the right-hand side is the Riemann integral of the real-valued function $t \mapsto \alpha_{C(t)}(C'(t))$. As is well known, the value of $\int_C \alpha$ depends on C only through its image and the direction in which these points are traversed.

If α is the differential of a function f, according to the definitions (1.53), (1.41), and (1.9) we have

$$\int_C \mathrm{d}f = \int_a^b \mathrm{d}f_{C(t)}\big(C'(t)\big)\,\mathrm{d}t = \int_a^b C'_t[f]\,\mathrm{d}t = \int_a^b \frac{\mathrm{d}}{\mathrm{d}t}(C^* f)\,\mathrm{d}t$$
$$= f\big(C(b)\big) - f\big(C(a)\big).$$

Hence, if C is a closed curve [that is, $C(a) = C(b)$],

$$\int_C \mathrm{d}f = 0.$$

For instance, if $M \equiv \mathbb{R}^2 \setminus \{(0, 0)\}$, recalling that

$$\big(\mathrm{d}x^i\big)_{C(t)}(C'_t) = C'_t\big[x^i\big] = \frac{\mathrm{d}}{\mathrm{d}t}\big(x^i \circ C\big)$$

[see (1.41) and (1.9)], the line integral of the 1-form

$$\alpha = \frac{x\,\mathrm{d}y - y\,\mathrm{d}x}{x^2 + y^2}$$

on the closed curve $C : [0, 2\pi] \to M$, defined by $C(t) = (\cos t, \sin t)$, has the value

$$\int_C \alpha = \int_0^{2\pi} \frac{\cos t \cos t + \sin t \sin t}{\cos^2 t + \sin^2 t}\,\mathrm{d}t = 2\pi,$$

which is different from zero and, therefore, α is not the differential of some function defined on M [see also Guillemin and Pollack (1974), do Carmo (1994)].

Tensor Fields A *tensor of type* $\binom{0}{k}$ (or *covariant tensor of rank* k) at p is a multilinear map $t_p : T_pM \times \cdots \times T_pM$ (k times) $\to \mathbb{R}$. A tensor of type $\binom{0}{1}$ is a covector. The set of tensors of type $\binom{0}{k}$ at p is a real vector space if for any pair of tensors of type $\binom{0}{k}$ at p, t_p and s_p, we define

$$(at_p + bs_p)(v_1, \ldots, v_k) \equiv at_p(v_1, \ldots, v_k) + bs_p(v_1, \ldots, v_k), \qquad (1.54)$$

for $v_1, \ldots, v_k \in T_pM$ and $a, b \in \mathbb{R}$.

If t_p is a tensor of type $\binom{0}{k}$ at p and s_p is a tensor of type $\binom{0}{l}$ at p, the *tensor product* $t_p \otimes s_p$ is defined by

$$(t_p \otimes s_p)(v_1, \ldots, v_{k+l}) \equiv t_p(v_1, \ldots, v_k)\, s_p(v_{k+1}, \ldots, v_{k+l}), \qquad (1.55)$$

for $v_1, \ldots, v_{k+l} \in T_pM$. Then $t_p \otimes s_p$ is a tensor of type $\binom{0}{k+l}$ at p.

Exercise 1.29 Show that

$$\begin{aligned} (at_{1p} + bt_{2p}) \otimes s_p &= at_{1p} \otimes s_p + bt_{2p} \otimes s_p, \\ t_p \otimes (as_{1p} + bs_{2p}) &= at_p \otimes s_{1p} + bt_p \otimes s_{2p}, \\ (r_p \otimes s_p) \otimes t_p &= r_p \otimes (s_p \otimes t_p). \end{aligned}$$

If t_p is a tensor of type $\binom{0}{k}$ at p and $v_1, \ldots, v_k \in T_pM$, making use of the multilinearity of t_p, of the definition of the tensor product, and expressing the vectors v_i in the form $v_i = v_i[x^j](\partial/\partial x^j)_p = \mathrm{d}x^j_p(v_i)(\partial/\partial x^j)_p$ $(i = 1, 2, \ldots, k)$, according to the definition (1.55) we have

$$\begin{aligned} t_p(v_1, \ldots, v_k) &= t_p\left(\mathrm{d}x^i_p(v_1)\left(\frac{\partial}{\partial x^i}\right)_p, \ldots, \mathrm{d}x^m_p(v_k)\left(\frac{\partial}{\partial x^m}\right)_p\right) \\ &= \mathrm{d}x^i_p(v_1)\cdots \mathrm{d}x^m_p(v_k)\, t_p\left(\left(\frac{\partial}{\partial x^i}\right)_p, \ldots, \left(\frac{\partial}{\partial x^m}\right)_p\right) \\ &= \left[t_p\left(\left(\frac{\partial}{\partial x^i}\right)_p, \ldots, \left(\frac{\partial}{\partial x^m}\right)_p\right)\mathrm{d}x^i_p \otimes \cdots \otimes \mathrm{d}x^m_p\right](v_1, \ldots, v_k); \end{aligned}$$

therefore,

$$t_p = t_p\left(\left(\frac{\partial}{\partial x^i}\right)_p, \left(\frac{\partial}{\partial x^j}\right)_p, \ldots, \left(\frac{\partial}{\partial x^m}\right)_p\right)\mathrm{d}x^i_p \otimes \mathrm{d}x^j_p \otimes \cdots \otimes \mathrm{d}x^m_p. \qquad (1.56)$$

A *tensor field of type* $\binom{0}{k}$ (or *covariant tensor field of rank* k), t, on M is a map that associates with each point $p \in M$ a tensor of type $\binom{0}{k}$, $t(p)$ or t_p, at p. If t

is tensor field of type $\binom{0}{k}$ and $\mathbf{X}_1, \ldots, \mathbf{X}_k$ are k vector fields on M, $t(\mathbf{X}_1, \ldots, \mathbf{X}_k)$ is the real-valued function given by $[t(\mathbf{X}_1, \ldots, \mathbf{X}_k)](p) \equiv t_p(\mathbf{X}_1(p), \ldots, \mathbf{X}_k(p))$. We say that t is differentiable if $t(\mathbf{X}_1, \ldots, \mathbf{X}_k)$ is a differentiable function for all $\mathbf{X}_1, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$.

Exercise 1.30 Show that t is differentiable if and only if the functions $t_{ij\ldots m} \equiv t(\partial/\partial x^i, \partial/\partial x^j, \ldots, \partial/\partial x^m)$ (the components of t with respect to the basis induced by the coordinates x^i) are differentiable.

The sum, the product by scalars, the product by real-valued functions, and the tensor product of tensor fields are defined pointwise:

$$(at + bs)_p \equiv at_p + bs_p,$$
$$(ft)_p \equiv f(p)t_p,$$
$$(t \otimes s)_p \equiv t_p \otimes s_p,$$

for $a, b \in \mathbb{R}$, s, t tensor fields on M, and $f : M \to \mathbb{R}$. Using these operations, any tensor field of type $\binom{0}{k}$ has the local expression

$$t = t\left(\left(\frac{\partial}{\partial x^i}\right), \left(\frac{\partial}{\partial x^j}\right), \ldots, \left(\frac{\partial}{\partial x^m}\right)\right) \mathrm{d}x^i \otimes \mathrm{d}x^j \otimes \cdots \otimes \mathrm{d}x^m. \tag{1.57}$$

If t is a tensor field of type $\binom{0}{k}$ on M and $\mathbf{X}_1, \ldots, \mathbf{X}_k$ are k vector fields on M, owing to the linearity of t_p in each of its arguments, for any function $f : M \to \mathbb{R}$,

$$\begin{aligned}
\big[t(\mathbf{X}_1, \ldots, f\mathbf{X}_i, \ldots, \mathbf{X}_k)\big](p) &= t_p\big(\mathbf{X}_1(p), \ldots, (f\mathbf{X}_i)(p), \ldots, \mathbf{X}_k(p)\big) \\
&= t_p\big(\mathbf{X}_1(p), \ldots, f(p)\mathbf{X}_i(p), \ldots, \mathbf{X}_k(p)\big) \\
&= f(p)t_p\big(\mathbf{X}_1(p), \ldots, \mathbf{X}_i(p), \ldots, \mathbf{X}_k(p)\big) \\
&= f(p)\big[t(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k)\big](p),
\end{aligned}$$

for $p \in M$, that is,

$$t(\mathbf{X}_1, \ldots, f\mathbf{X}_i, \ldots, \mathbf{X}_k) = f\, t(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k), \quad 1 \leq i \leq k.$$

Similarly, we conclude that

$$t(\mathbf{X}_1, \ldots, \mathbf{X}_1 + \mathbf{X}'_i, \ldots, \mathbf{X}_k) = t(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k) + t(\mathbf{X}_1, \ldots, \mathbf{X}'_i, \ldots, \mathbf{X}_k).$$

Note that, for instance, the Lie bracket is not a tensor since $[\mathbf{X}, f\mathbf{Y}] = f[\mathbf{X}, \mathbf{Y}] + (\mathbf{X}f)\mathbf{Y}$ (see Exercise 1.22).

Conversely, if t is a map that to each set of k vector fields on M associates a function of M in $\mathbb{R}$ with the property that for any pair of functions $f, g : M \to \mathbb{R}$,

$$\begin{aligned}
&t(\mathbf{X}_1, \ldots, f\mathbf{X}_i + g\mathbf{X}'_i, \ldots, \mathbf{X}_k) \\
&\quad = f\,t(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k) + g\,t(\mathbf{X}_1, \ldots, \mathbf{X}'_i, \ldots, \mathbf{X}_k),
\end{aligned}$$

$1 \leq i \leq k$, then t is a tensor field of type $\binom{0}{k}$. In effect, the property for t assumed ensures that, locally, t is of the form $t = t((\partial/\partial x^i), \ldots, (\partial/\partial x^m))\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^m$, since if $\mathbf{X}_1, \ldots, \mathbf{X}_k$ are vector fields on M, writing them in the form $\mathbf{X}_i = \mathrm{d}x^j(\mathbf{X}_i)(\partial/\partial x^j)$, $i = 1, \ldots, k$, we have

$$\begin{aligned} t(\mathbf{X}_1, \ldots, \mathbf{X}_k) &= t\left(\mathrm{d}x^i(\mathbf{X}_1)\left(\frac{\partial}{\partial x^i}\right), \ldots, \mathrm{d}x^m(\mathbf{X}_k)\left(\frac{\partial}{\partial x^m}\right)\right) \\ &= \mathrm{d}x^i(\mathbf{X}_1)\cdots \mathrm{d}x^m(\mathbf{X}_k)\, t\left(\left(\frac{\partial}{\partial x^i}\right), \ldots, \left(\frac{\partial}{\partial x^m}\right)\right) \\ &= \left[t\left(\left(\frac{\partial}{\partial x^i}\right), \ldots, \left(\frac{\partial}{\partial x^m}\right)\right)\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^m\right](\mathbf{X}_1, \ldots, \mathbf{X}_k). \end{aligned}$$

A *tensor of type* $\binom{k}{0}$ (or *contravariant tensor of rank* k) at p is a multilinear mapping $t_p : T_p^*M \times \cdots \times T_p^*M$ (k times) $\to \mathbb{R}$. The set of tensors of type $\binom{k}{0}$ at p forms a vector space defining the sum and the multiplication by real scalars in an analogous manner to the operations for tensors of type $\binom{0}{k}$. Similarly, if t_p is a tensor of type $\binom{k}{0}$ at p and s_p is a tensor of type $\binom{l}{0}$ at p, the tensor product $t_p \otimes s_p$, given by

$$(t_p \otimes s_p)(\alpha_1, \ldots, \alpha_{k+l}) \equiv t_p(\alpha_1, \ldots, \alpha_k)\, s_p(\alpha_{k+1}, \ldots, \alpha_{k+l}),$$

for $\alpha_1, \ldots, \alpha_{k+l} \in T_p^*M$, is a tensor of type $\binom{k+l}{0}$ at p.

If t_p is a tensor of type $\binom{k}{0}$ at p and $\alpha_1, \ldots, \alpha_k \in T_p^*M$, expressing each covector α_i in the form $\alpha_i = \alpha_i((\partial/\partial x^j)_p)\,\mathrm{d}x_p^j$ [see (1.49)], we have

$$\begin{aligned} t_p(\alpha_1, \ldots, \alpha_k) &= t_p\left(\alpha_1\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\mathrm{d}x_p^i, \ldots, \alpha_k\left(\left(\frac{\partial}{\partial x^m}\right)_p\right)\mathrm{d}x_p^m\right) \\ &= \alpha_1\left(\left(\frac{\partial}{\partial x^i}\right)_p\right)\cdots\alpha_k\left(\left(\frac{\partial}{\partial x^m}\right)_p\right)t_p(\mathrm{d}x_p^i, \ldots, \mathrm{d}x_p^m). \end{aligned}$$

Defining $v_p(\alpha_p) \equiv \alpha_p(v_p)$ for $v_p \in T_pM$ and $\alpha_p \in T_p^*M$ (which amounts to the identification of T_pM with the dual space of T_p^*M), we have

$$t_p(\alpha_1, \ldots, \alpha_k) = \left[t_p\big(\mathrm{d}x_p^i, \ldots, \mathrm{d}x_p^m\big)\left(\frac{\partial}{\partial x^i}\right)_p \otimes \cdots \otimes \left(\frac{\partial}{\partial x^m}\right)_p\right](\alpha_1, \ldots, \alpha_k),$$

and therefore

$$t_p = t_p\big(\mathrm{d}x_p^i, \ldots, \mathrm{d}x_p^m\big)\left(\frac{\partial}{\partial x^i}\right)_p \otimes \cdots \otimes \left(\frac{\partial}{\partial x^m}\right)_p.$$

A *tensor field of type* $\binom{k}{0}$ (or *contravariant tensor field of rank* k), t, on M is a map that associates to each point $p \in M$ a tensor of type $\binom{k}{0}$, $t(p)$ or t_p, at p.

The tensor field t is differentiable if for any k 1-forms, $\alpha_1, \ldots, \alpha_k$, the function $t(\alpha_1, \ldots, \alpha_k)$, defined by $[t(\alpha_1, \ldots, \alpha_k)](p) \equiv t_p(\alpha_1(p), \ldots, \alpha_k(p))$, is differentiable. Any tensor field of type $\binom{k}{0}$ on M is expressed locally as

$$t = t\big(\mathrm{d}x^i, \ldots, \mathrm{d}x^m\big)\left(\frac{\partial}{\partial x^i}\right) \otimes \cdots \otimes \left(\frac{\partial}{\partial x^m}\right).$$

Again, it turns out that t is differentiable if and only if the functions $t^{i \ldots m} \equiv t(\mathrm{d}x^i, \ldots, \mathrm{d}x^m)$ are differentiable. Furthermore, any map t that to each set of k covector fields associates a function from M into $\mathbb{R}$ is a tensor field of type $\binom{k}{0}$ if and only if for $\alpha_1, \ldots, \alpha_i, \alpha'_i, \ldots, \alpha_k$, covector fields on M,

$$t(\alpha_1, \ldots, f\alpha_i + g\alpha'_i, \ldots, \alpha_k) = f\,t(\alpha_1, \ldots, \alpha_i, \ldots, \alpha_k) + g\,t(\alpha_1, \ldots, \alpha'_i, \ldots, \alpha_k),$$

for $f, g : M \to \mathbb{R}$.

A *mixed tensor of type* $\binom{k}{l}$ at p, is a multilinear map from the Cartesian product of k copies of $T^*_p M$ and l copies of $T_p M$ in $\mathbb{R}$. The tensors of type $\binom{k}{l}$ at p form a real vector space where the sum and the product by real scalars are defined in the natural way. The tensor product of a tensor of type $\binom{k}{l}$ by a tensor of type $\binom{k'}{l'}$ is a tensor of type $\binom{k+k'}{l+l'}$. A basis for the vector space of the tensors of type $\binom{k}{l}$ at p is formed by the tensor products of k vectors $(\partial/\partial x^i)_p$ and l covectors $\mathrm{d}x^i_p$; therefore, this space has dimension n^{k+l}.

A *tensor field of type* $\binom{k}{l}$ on M is a map that to each point $p \in M$ associates a tensor of type $\binom{k}{l}$ at p; a tensor field of type $\binom{0}{0}$ on M is a function of M in $\mathbb{R}$. A tensor field, t, of type $\binom{k}{l}$ is differentiable if for $\mathbf{X}_1, \ldots, \mathbf{X}_l \in \mathfrak{X}(M)$ and $\alpha_1, \ldots, \alpha_k \in \Lambda^1(M)$, the function of M into $\mathbb{R}$ that to each point $p \in M$ associates the value of t_p on $\mathbf{X}_1(p), \ldots, \mathbf{X}_l(p), \alpha_1(p), \ldots, \alpha_k(p)$ (taken in an appropriate order) is differentiable.

The sum, the product by scalars, the product by real-valued functions, and the tensor product of mixed tensor fields are defined pointwise:

$$\begin{aligned}(at + bs)_p &\equiv at_p + bs_p \quad \text{(when } t \text{ and } s \text{ are of the same type)}\\ (ft)_p &\equiv f(p)t_p,\\ (t \otimes s)_p &\equiv t_p \otimes s_p,\end{aligned}$$

for $a, b \in \mathbb{R}$, $f : M \to \mathbb{R}$, and t, s mixed tensor fields on M. The set of differentiable tensor fields of type $\binom{k}{l}$ on M, denoted by $T^k_l(M)$, is a module over the ring $C^\infty(M)$.

Chapter 2
Lie Derivatives

In this chapter several additional useful concepts are introduced, which will be extensively employed in the second half of this book. It is shown that there is a one-to-one relation between vector fields on a manifold and families of transformations of the manifold onto itself. This relation is essential in the study of various symmetries, as shown in Chaps. 4, 6, and 8, and in the relationship of a Lie group with its Lie algebra, treated in Chap. 7.

2.1 One-Parameter Groups of Transformations and Flows

Definition 2.1 Let M be a differentiable manifold. A *one-parameter group of transformations*, φ, on M, is a differentiable map from $M \times \mathbb{R}$ onto M such that $\varphi(x,0) = x$ and $\varphi(\varphi(x,t),s) = \varphi(x,t+s)$ for all $x \in M$, $t, s \in \mathbb{R}$.

If we define $\varphi_t(x) \equiv \varphi(x,t)$, then, for each $t \in \mathbb{R}$, φ_t is a differentiable map from M onto M and $\varphi_{t+s}(x) = \varphi(x,t+s) = \varphi(\varphi(x,t),s) = \varphi(\varphi_t(x),s) = \varphi_s(\varphi_t(x)) = (\varphi_s \circ \varphi_t)(x)$, that is,

$$\varphi_{t+s} = \varphi_s \circ \varphi_t = \varphi_t \circ \varphi_s$$

(since $t+s = s+t$). φ_0 is the identity map of M since $\varphi_0(x) = \varphi(x,0) = x$ for all $x \in M$. We have then $\varphi_t \circ \varphi_{-t} = \varphi_{-t} \circ \varphi_t = \varphi_0$, which means that each map φ_t has an inverse, φ_{-t}, which is also differentiable. Therefore, each φ_t is a diffeomorphism of M onto itself. Thus, the set of transformations $\{\varphi_t \mid t \in \mathbb{R}\}$ is an Abelian group of diffeomorphisms of M onto M, and the map $t \mapsto \varphi_t$ is a homomorphism from the additive group of the real numbers into the group of diffeomorphisms of M.

Each one-parameter group of transformations φ on M determines a family of curves in M (the *orbits* of the group). The map $\varphi_x : \mathbb{R} \to M$ given by $\varphi_x(t) = \varphi(x,t)$ is a differentiable curve in M for each $x \in M$. Since $\varphi_x(0) = \varphi(x,0) = x$, the tangent vector to the curve φ_x at $t = 0$ belongs to $T_x M$. The *infinitesimal generator* of φ is the vector field $\mathbf{X}$ such that $\mathbf{X}_x = (\varphi_x)'_0$. In other words, the infinitesimal

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_2,

generator of φ is a vector field tangent to the curves generated by the one-parameter group of transformations.

Example 2.2 Let $M = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ and let $\varphi : M \times \mathbb{R} \to M$ be given by

$$\varphi\big((x_0, y_0), t\big) = \frac{(2x_0, 2y_0 \cos t + (1 - x_0{}^2 - y_0{}^2)\sin t)}{1 + x_0{}^2 + y_0{}^2 + (1 - x_0{}^2 - y_0{}^2)\cos t - 2y_0 \sin t}. \tag{2.1}$$

The map (2.1) is differentiable because it is the composition of differentiable functions and the denominator does not vanish for $x_0 \neq 0$ (it can be verified that the denominator in (2.1) is equal to $2[(x_0 \sin(t/2))^2 + (y_0 \sin(t/2) - \cos(t/2))^2]$). Furthermore, $\varphi((x_0, y_0), t) \in M$ for any $(x_0, y_0) \in M$, $t \in \mathbb{R}$, and $\varphi((x_0, y_0), 0) = (x_0, y_0)$. Finally, a direct but lengthy computation shows that (2.1) satisfies the relation $\varphi(\varphi((x_0, y_0), t), s) = \varphi((x_0, y_0), t + s)$, and therefore we have a one-parameter group of transformations on M.

For $(x_0, y_0) \in M$ fixed, $\varphi_{(x_0,y_0)}(t) \equiv \varphi((x_0, y_0), t)$ is a differentiable curve in M whose tangent vector at $t = 0$ can be obtained using (1.20), that is,

$$(\varphi_{(x_0,y_0)})'_0 = \frac{\mathrm{d}}{\mathrm{d}t}(x \circ \varphi_{(x_0,y_0)})\bigg|_{t=0} \left(\frac{\partial}{\partial x}\right)_{(x_0,y_0)} + \frac{\mathrm{d}}{\mathrm{d}t}(y \circ \varphi_{(x_0,y_0)})\bigg|_{t=0} \left(\frac{\partial}{\partial y}\right)_{(x_0,y_0)}$$

with

$$\begin{aligned} (x \circ \varphi_{(x_0,y_0)})(t) &= \frac{2x_0}{1 + x_0{}^2 + y_0{}^2 + (1 - x_0{}^2 - y_0{}^2)\cos t - 2y_0 \sin t}, \\ (y \circ \varphi_{(x_0,y_0)})(t) &= \frac{2y_0 \cos t + (1 - x_0{}^2 - y_0{}^2)\sin t}{1 + x_0{}^2 + y_0{}^2 + (1 - x_0{}^2 - y_0{}^2)\cos t - 2y_0 \sin t} \end{aligned} \tag{2.2}$$

[see (2.1)]. Calculating the derivatives of the expressions (2.2) with respect to t at $t = 0$, one finds that the infinitesimal generator of the one-parameter group (2.1), $\mathbf{X}$, is given by

$$\begin{aligned} \mathbf{X}_{(x_0,y_0)} \equiv (\varphi_{(x_0,y_0)})'_0 &= x_0 y_0 \left(\frac{\partial}{\partial x}\right)_{(x_0,y_0)} + \frac{1 - x_0{}^2 + y_0{}^2}{2} \left(\frac{\partial}{\partial y}\right)_{(x_0,y_0)} \\ &= \left(xy\frac{\partial}{\partial x} + \frac{1 - x^2 + y^2}{2}\frac{\partial}{\partial y}\right)_{(x_0,y_0)} \end{aligned}$$

[see (1.32)]; thus,

$$\mathbf{X} = xy\frac{\partial}{\partial x} + \frac{1 - x^2 + y^2}{2}\frac{\partial}{\partial y}. \tag{2.3}$$

The (images of the) curves defined by the one-parameter group (2.1), to which $\mathbf{X}$ is tangent, are circle arcs. In order to simplify the notation, we shall write x and y in place of $x \circ \varphi_{(x_0,y_0)}$ and $y \circ \varphi_{(x_0,y_0)}$, respectively; then, from (2.2), eliminating the parameter t, we see that

$$\left(x - \frac{1 + {x_0}^2 + {y_0}^2}{2x_0}\right)^2 + y^2 = \left(\frac{1 + {x_0}^2 + {y_0}^2}{2x_0}\right)^2 - 1,$$

which is the equation of a circle centered at a point of the x axis.

Exercise 2.3 Show that the following families of maps $\varphi_t : \mathbb{R}^2 \to \mathbb{R}^2$ form one-parameter groups of transformations and find their infinitesimal generators:

(a) $\varphi_t(x, y) = (x \cos t - y \sin t, x \sin t + y \cos t)$.
(b) $\varphi_t(x, y) = (x + at, y + bt)$, with $a, b \in \mathbb{R}$.
(c) $\varphi_t(x, y) = (\mathrm{e}^{at} x, \mathrm{e}^{bt} y)$, with $a, b \in \mathbb{R}$.

Exercise 2.4 Let φ be a one-parameter group of transformations on M and let $\mathbf{X}$ be its infinitesimal generator. Show that if $y = \varphi_x(t_0)$, for some $t_0 \in \mathbb{R}$, then $(\varphi_x)'_{t_0} = (\varphi_y)'_0$ and, therefore, $(\varphi_x)'_{t_0} = \mathbf{X}_{\varphi_x(t_0)}$.

Given a differentiable vector field, $\mathbf{X}$, on M, there does not always exist a one-parameter group of transformations whose infinitesimal generator is $\mathbf{X}$; it is said that $\mathbf{X}$ is *complete* if such a one-parameter group of transformations exists.

Integral Curves of a Vector Field

Definition 2.5 Let $\mathbf{X}$ be a vector field on M. A curve $C : I \to M$ is an *integral curve* of $\mathbf{X}$ if $C'_t = \mathbf{X}_{C(t)}$, for $t \in I$. If $C(0) = x$ we say that C starts at x. (According to Exercise 2.4, if φ is a one-parameter group of transformations and $\mathbf{X}$ is its infinitesimal generator, then the curve φ_x is an integral curve of $\mathbf{X}$ that starts at x.)

If $(x^1, x^2, \ldots, x^n)$ is a local coordinate system on M and $\mathbf{X}$ is expressed in the form $\mathbf{X} = X^i(\partial/\partial x^i)$, the condition that C be an integral curve of $\mathbf{X}$ amounts to the system of ordinary differential equations (ODEs) [see (1.20)]

$$\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t} = X^i \circ C. \tag{2.4}$$

More explicitly, writing the right-hand side of the previous equation in the form

$$\begin{aligned}
\big(X^i \circ C\big)(t) &= \big(X^i \circ \phi^{-1}\big)\big(\phi\big(C(t)\big)\big) \\
&= \big(X^i \circ \phi^{-1}\big)\big(x^1\big(C(t)\big), x^2\big(C(t)\big), \ldots, x^n\big(C(t)\big)\big) \\
&= \big(X^i \circ \phi^{-1}\big)\big(\big(x^1 \circ C\big)(t), \big(x^2 \circ C\big)(t), \ldots, \big(x^n \circ C\big)(t)\big),
\end{aligned}$$

one finds that equations (2.4) correspond to the (autonomous) system of equations

$$\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t} = \big(X^i \circ \phi^{-1}\big)\big(x^1 \circ C, x^2 \circ C, \ldots, x^n \circ C\big) \tag{2.5}$$

for the n functions $x^i \circ C$ of $\mathbb{R}$ to $\mathbb{R}$. (Note that each composition $X^i \circ \phi^{-1}$ is a real-valued function defined in some subset of $\mathbb{R}^n$.) According to the fundamental theorem for systems of ODEs, given $x \in M$, there exists a unique integral curve of $\mathbf{X}$, C, starting at x. (That is, if D is another integral curve of $\mathbf{X}$ starting at x, then $D = C$ in the intersection of their domains.)

Let C be an integral curve of $\mathbf{X}$ starting at x, and let $\varphi(x,t) \equiv C(t)$. The curve D defined by $D(t) \equiv C(t+s)$, with s fixed, is an integral curve of $\mathbf{X}$, since for an arbitrary function $f \in C^\infty(M)$

$$\begin{aligned} D'_t[f] &= \lim_{h\to 0} \frac{f(D(t+h)) - f(D(t))}{h} \\ &= \lim_{h\to 0} \frac{f(C(t+h+s)) - f(C(t+s))}{h} \\ &= C'_{t+s}[f] = \mathbf{X}_{C(t+s)}[f] = \mathbf{X}_{D(t)}[f]. \end{aligned}$$

The curve D starts at $D(0) = C(s)$ and by virtue of the uniqueness of the integral curves, we have

$$D(t) = \varphi\big(C(s), t\big) = \varphi\big(\varphi(x,s), t\big).$$

On the other hand, from the definition of D,

$$D(t) = C(t+s) = \varphi(x, t+s);$$

therefore,

$$\varphi\big(\varphi(x,s), t\big) = \varphi(x, t+s) \tag{2.6}$$

(cf. Definition 2.1).

In some cases φ is not defined for all $t \in \mathbb{R}$, and for that reason it is not a one-parameter group of transformations. However, for each $x \in M$ there exist a neighborhood, U of x and an $\varepsilon > 0$ such that φ is defined on $U \times (-\varepsilon, \varepsilon)$ and is differentiable. The map φ is called a *flow* or *local one-parameter group of transformations* and $\mathbf{X}$ is its infinitesimal generator.

If $\mathbf{X}$ is the infinitesimal generator of a one-parameter group of transformations or a flow, the transformations φ_t are also denoted by $\exp t\mathbf{X}$. Then, the relation (2.6) is expressed as $\exp t\mathbf{X} \circ \exp s\mathbf{X} = \exp(t+s)\mathbf{X}$.

Example 2.6 Let $M = \mathbb{R}$ with the usual coordinate system, $x = \mathrm{id}$. The integral curves of the vector field $\mathbf{X} = x^2\, \partial/\partial x$ are determined by the single differential equation [see (2.4)]

$$\frac{\mathrm{d}(x \circ C)}{\mathrm{d}t} = x^2 \circ C = (x \circ C)^2 \tag{2.7}$$

[the previous equality follows from (1.7), which gives $x^2(p) = (x(p))^2$; hence, $(x^2 \circ C)(t) = x^2(C(t)) = [x(C(t))]^2 = ((x \circ C)(t))^2 = (x \circ C)^2(t)$]. The solution of (2.7) is $(x \circ C)(t) = -1/(t+a)$, where a is a constant or, simply, since

$x = \mathrm{id}$, $C(t) = -1/(t+a)$. If the integral curve of $\mathbf{X}$ starts at x_0, then $C(0) = -1/a = x_0$, i.e., $a = -1/x_0$. Since φ_{x_0} is the integral curve of $\mathbf{X}$ starting at x_0 (see Definition 2.5), we have

$$\varphi_{x_0}(t) = -\frac{1}{t - 1/x_0} = \frac{x_0}{1 - x_0 t},$$

and therefore

$$\varphi(x_0, t) = \frac{x_0}{1 - x_0 t} \tag{2.8}$$

is the local one-parameter group generated by $x^2\,\partial/\partial x$.

The expression (2.8) is not defined for $t = 1/x_0$, and therefore we are not dealing with a one-parameter group of transformations, despite the fact that $\mathbf{X}$ is differentiable. However, the flow (2.8) satisfies the relation (2.6), since, according to (2.8),

$$\begin{aligned}\varphi\big(\varphi(x_0, s), t\big) &= \frac{\varphi(x_0, s)}{1 - \varphi(x_0, s)t} = \frac{x_0/(1 - x_0 s)}{1 - t x_0/(1 - x_0 s)} = \frac{x_0}{1 - x_0(t+s)} \\ &= \varphi(x_0, t+s),\end{aligned}$$

whenever all the expressions involved are defined.

Example 2.7 Let $M = \mathbb{R}^2$ and let $\mathbf{X} = y\,\partial/\partial x + x\,\partial/\partial y$, where (x, y) are the usual coordinates of $\mathbb{R}^2$. Equations (2.4) are in this case

$$\frac{\mathrm{d}(x \circ C)}{\mathrm{d}t} = y \circ C, \qquad \frac{\mathrm{d}(y \circ C)}{\mathrm{d}t} = x \circ C.$$

By adding and subtracting these equations we obtain

$$\frac{\mathrm{d}(x \circ C + y \circ C)}{\mathrm{d}t} = x \circ C + y \circ C, \qquad \frac{\mathrm{d}(x \circ C - y \circ C)}{\mathrm{d}t} = -(x \circ C - y \circ C),$$

whose solutions are $(x \circ C + y \circ C)(t) = (x_0 + y_0)\,\mathrm{e}^t$ and $(x \circ C - y \circ C)(t) = (x_0 - y_0)\,\mathrm{e}^{-t}$, where x_0 and y_0 are the initial values of $x \circ C$ and $y \circ C$, respectively. Hence, $(x \circ C)(t) = x_0 \cosh t + y_0 \sinh t$, $(y \circ C)(t) = x_0 \sinh t + y_0 \cosh t$, and

$$\varphi_{(x_0, y_0)}(t) = (x_0 \cosh t + y_0 \sinh t, x_0 \sinh t + y_0 \cosh t). \tag{2.9}$$

Since $(x \circ C)^2 - (y \circ C)^2 = {x_0}^2 - {y_0}^2$, the (images of the) integral curves of $\mathbf{X}$ are hyperbolas or straight lines. The expression (2.9) is defined for all $t \in \mathbb{R}$, and therefore it corresponds to a one-parameter group of transformations. Substituting (2.9) into (2.6) one finds the well-known addition formulas

$$\begin{aligned}\cosh(t+s) &= \cosh t \cosh s + \sinh t \sinh s, \\ \sinh(t+s) &= \sinh t \cosh s + \cosh t \sinh s.\end{aligned}$$

Exercise 2.8 Let $\psi : M_1 \to M_2$ be a differentiable map and let φ_1 and φ_2 be one-parameter groups of transformations or flows on M_1 and M_2, respectively. Show that if $\varphi_{2t} \circ \psi = \psi \circ \varphi_{1t}$, then the infinitesimal generators of φ_1 and φ_2 are ψ-related, i.e., show that $\psi_{*x}\mathbf{X}_x = \mathbf{Y}_{\psi(x)}$, where $\mathbf{X}$ and $\mathbf{Y}$ are the infinitesimal generators of φ_1 and φ_2, respectively.

Example 2.9 An integration procedure distinct from that employed in the preceding examples is illustrated by considering the vector field $\mathbf{X} = \frac{1}{2}(x^2 - y^2)\,\partial/\partial x + xy\,\partial/\partial y$ on $M \equiv \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$. (The one-parameter group generated by this vector field is also found, by another method, in Example 6.12.) The system of equations (2.4) is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{2}\big(x^2 - y^2\big), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = xy, \tag{2.10}$$

where, in order to simplify the notation, we have written x and y in place of $x \circ C$ and $y \circ C$, respectively. Eliminating the variable t from these equations (with the aid of the chain rule) we obtain the ODE $\mathrm{d}y/\mathrm{d}x = 2xy/(x^2 - y^2)$. Noting that the right-hand side of the last equation is the quotient of two homogeneous functions of the same degree, it is convenient to introduce $u \equiv y/x$, so that $\mathrm{d}u/\mathrm{d}x = u(1 + u^2)/[x(1 - u^2)]$, which by the standard procedures leads to

$$\frac{\mathrm{d}x}{x} = \frac{(1 - u^2)\,\mathrm{d}u}{u(1 + u^2)} = \left(\frac{1}{u} - \frac{2u}{1 + u^2}\right)\mathrm{d}u,$$

whose solution is given by $x = cu/(1 + u^2) = cy/[x(1 + y^2/x^2)]$, where c is some constant. Hence $x^2 + y^2 = cy$, which corresponds to the circle centered at $(0, c/2)$ and radius $c/2$.

In order to obtain the parametrization of these curves, one can substitute $x = \pm\sqrt{cy - y^2}$ into the second of equations (2.10), which yields $\mathrm{d}y/\mathrm{d}t = \pm y\sqrt{cy - y^2}$, or, putting $v = 1/y$, $\mathrm{d}v/\mathrm{d}t = \mp\sqrt{cv - 1}$; hence $2\sqrt{cv - 1} = \mp c(t - t_0)$, where t_0 is a constant. Thus, from the foregoing relations we find that

$$y = \frac{4c}{4 + c^2(t - t_0)^2}, \qquad x = -\frac{2c^2(t - t_0)}{4 + c^2(t - t_0)^2}. \tag{2.11}$$

For the integral curve of $\mathbf{X}$ starting at (x_0, y_0), from (2.11) we have $y_0 = 4c/(4 + c^2{t_0}^2)$ and $x_0 = 2c^2 t_0/(4 + c^2{t_0}^2)$, which imply that

$$c = \frac{{x_0}^2 + {y_0}^2}{y_0}, \qquad t_0 = \frac{2x_0}{{x_0}^2 + {y_0}^2}$$

and, substituting these expressions into (2.11), we obtain

$$\varphi\big((x_0, y_0), t\big) = \frac{2({x_0}^2 + {y_0}^2)(2x_0 - ({x_0}^2 + {y_0}^2)t, 2y_0)}{[({x_0}^2 + {y_0}^2)t - 2x_0]^2 + 4{y_0}^2}$$

$$= \frac{(x_0 - ({x_0}^2 + {y_0}^2)t/2, y_0)}{(1 - x_0 t/2)^2 + {y_0}^2 (t/2)^2}. \tag{2.12}$$

From (2.12) we see that the integral curves of $\mathbf{X}$ are defined for all $t \in \mathbb{R}$, and therefore $\mathbf{X}$ is complete and (2.12) corresponds to a one-parameter group of transformations. Further examples are given in Examples 4.1, 6.11, 6.12, 6.20, 7.40, and 7.41.

Exercise 2.10 Find the integral curves of the vector field $\mathbf{X} = \frac{1}{x^2+y^2}(x\frac{\partial}{\partial x} - y\frac{\partial}{\partial y})$ on $\mathbb{R}^2 \setminus \{(0,0)\}$ and the one-parameter group of diffeomorphisms generated by $\mathbf{X}$.

From equations (2.10) one notices that if one looks for the integral curves of $f\mathbf{X}$, where f is some real-valued differentiable function, on eliminating the variable t the function f disappears and one obtains the same equation for $\mathrm{d}y/\mathrm{d}x$ as obtained in the preceding example. Therefore, the same circles are obtained. For any vector field $\mathbf{X}$, the integral curves of $\mathbf{X}$ and $f\mathbf{X}$, with $f \in C^\infty(M)$, only differ in the parametrization. If φ_t denotes the flow or one-parameter group generated by $\mathbf{X}$ and σ is a function of some open subset of $\mathbb{R}$ in the domain of the curve φ_x, then the tangent vector to the curve $\psi_x \equiv \varphi_x \circ \sigma$ satisfies, for $g \in C^\infty(M)$,

$$\begin{aligned}
(\psi_x)'_{t_0}[g] &= \frac{\mathrm{d}}{\mathrm{d}t} g\big(\psi_x(t)\big)\Big|_{t=t_0} = \frac{\mathrm{d}}{\mathrm{d}t}\big((g \circ \varphi_x) \circ \sigma\big)(t)\Big|_{t=t_0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}(g \circ \varphi_x)\Big|_{\sigma(t_0)} \frac{\mathrm{d}\sigma}{\mathrm{d}t}\Big|_{t_0} = (\varphi_x)'_{\sigma(t_0)}[g] \frac{\mathrm{d}\sigma}{\mathrm{d}t}\Big|_{t_0} \\
&= \frac{\mathrm{d}\sigma}{\mathrm{d}t}\Big|_{t_0} \mathbf{X}_{\varphi_x(\sigma(t_0))}[g],
\end{aligned} \tag{2.13}$$

where we have made use of the chain rule for functions from $\mathbb{R}$ into $\mathbb{R}$ and of the result of Exercise 2.4. The expression (2.13) coincides with $(f\mathbf{X})_{\varphi(\sigma(t_0))}[g]$ if we choose σ in such a way that

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = f\big(\varphi_x\big(\sigma(t)\big)\big). \tag{2.14}$$

Hence, if additionally we impose the condition $\sigma(0) = 0$, the curve $\psi_x = \varphi_x \circ \sigma$ is an integral curve of $f\mathbf{X}$ starting at x.

Example 2.11 The integral curves of $f\mathbf{X}$, where $\mathbf{X}$ is the vector field considered in Example 2.9 and f is any function belonging to $C^\infty(M)$, can be obtained by solving equation (2.14) with φ_x given by (2.12), i.e.,

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}t} = f\Bigg(&\frac{({x_0}^2 + {y_0}^2)[4x_0 - 2({x_0}^2 + {y_0}^2)\sigma(t)]}{[({x_0}^2 + {y_0}^2)\sigma(t) - 2x_0]^2 + 4{y_0}^2}, \\
&\frac{4y_0({x_0}^2 + {y_0}^2)}{[({x_0}^2 + {y_0}^2)\sigma(t) - 2x_0]^2 + 4{y_0}^2}\Bigg).
\end{aligned} \tag{2.15}$$

If we take, for example, $f(x, y) = y^{-1}$, equation (2.15) becomes

$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{[({x_0}^2 + {y_0}^2)\sigma(t) - 2x_0]^2 + 4{y_0}^2}{4y_0({x_0}^2 + {y_0}^2)}$$

and with the change of variable $({x_0}^2 + {y_0}^2)\sigma(t) - 2x_0 = 2y_0 \tan u$ we have $\mathrm{d}u/\mathrm{d}t = 1/2$. Hence $u = (t - t_0)/2$, where t_0 is some constant and

$$\sigma(t) = \frac{2x_0 + 2y_0 \tan \frac{1}{2}(t - t_0)}{{x_0}^2 + {y_0}^2}. \tag{2.16}$$

The condition $\sigma(0) = 0$ amounts to $0 = x_0 - y_0 \tan \frac{1}{2} t_0$, which, substituted into (2.16), yields

$$\sigma(t) = \frac{2 \tan \frac{1}{2}t}{y_0 + x_0 \tan \frac{1}{2}t} = \frac{2 \sin \frac{1}{2}t}{x_0 \sin \frac{1}{2}t + y_0 \cos \frac{1}{2}t}. \tag{2.17}$$

Thus, the flow generated by $f\mathbf{X} = y^{-1}[\frac{1}{2}(x^2 - y^2)\,\partial/\partial x + xy\,\partial/\partial y]$ is given by $\psi((x_0, y_0), t) = \varphi((x_0, y_0), \sigma(t))$, where φ is the one-parameter group generated by $\mathbf{X}$, given by (2.12), and σ is the function (2.17), i.e.,

$$\begin{aligned} &\psi\big((x_0, y_0), t\big) \\ &\quad= \frac{1}{2y_0}\big(0, {x_0}^2 + {y_0}^2\big) \\ &\qquad+ \frac{1}{2y_0}\Big(2x_0 y_0 \cos t - \big({y_0}^2 - {x_0}^2\big) \sin t, \big({y_0}^2 - {x_0}^2\big) \cos t + 2x_0 y_0 \sin t\Big) \end{aligned} \tag{2.18}$$

[cf. (2.12)]. Even though the expression (2.18) is defined for all $t \in \mathbb{R}$, the variable t has to be restricted to some open interval of length 2π where $\psi((x_0, y_0), t) \neq (0, 0)$, taking into account that the manifold being considered is $M = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$. It may be noticed that $f\mathbf{X}$ is differentiable on M because y does not vanish there. Whereas $\mathbf{X}$ is complete, $f\mathbf{X}$ is not. The expression (2.18) shows that the images of the integral curves of $f\mathbf{X}$ (and of $\mathbf{X}$) are arcs of circles.

Second-Order ODEs A vector field $\mathbf{X}$ on the tangent bundle TM such that, for $v \in TM$,

$$\pi_{*v}\mathbf{X}_v = v, \tag{2.19}$$

where π is the canonical projection of TM on M, corresponds to a system of second-order ODEs. (Equation (2.19) makes sense because v is a tangent vector to M at $\pi(v)$, that is, $v \in T_{\pi(v)}M$, and π_{*v} applies $T_v(TM)$ into $T_{\pi(v)}M$.) In effect, using the local expression $\mathbf{X} = A^i\,\partial/\partial q^i + B^i\,\partial/\partial \dot{q}^i$ as well as (1.29) and (1.27), the relation

(2.19) amounts to

$$A^i(v)\left(\frac{\partial}{\partial x^i}\right)_{\pi(v)} = \dot{q}^i(v)\left(\frac{\partial}{\partial x^i}\right)_{\pi(v)},$$

that is, $A^i = \dot{q}^i$. Hence, any vector field on *TM* satisfying (2.19) locally is of the form

$$\mathbf{X} = \dot{q}^i \frac{\partial}{\partial q^i} + B^i \frac{\partial}{\partial \dot{q}^i}$$

in a coordinate system induced by a coordinate system on M (see Sect. 1.2), where the B^i are n arbitrary real-valued functions defined on *TM*. The integral curves of **X** are determined by the equations

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \dot{q}^i, \qquad \frac{\mathrm{d}\dot{q}^i}{\mathrm{d}t} = B^i,$$

which are equivalent to the system of n second-order ODEs

$$\frac{\mathrm{d}^2 q^i}{\mathrm{d}t^2} = \left(B^i \circ \overline{\phi}^{-1}\right)\left(q^1, \ldots, q^n, \frac{\mathrm{d}q^1}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}q^n}{\mathrm{d}t}\right).$$

Exercise 2.12 Let $\varphi(x, y, t) = (F_1(x, y, t), F_2(x, y, t))$ be a one-parameter group of transformations on $\mathbb{R}^2$ [which, among other things, implies that F_1 and F_2 are differentiable functions from $\mathbb{R}^3$ into $\mathbb{R}$ such that $F_1(x, y, 0) = x$ and $F_2(x, y, 0) = y$], and let

$$F_3(x, y, z, t) \equiv \frac{D_1 F_2 + z D_2 F_2}{D_1 F_1 + z D_2 F_1}, \tag{2.20}$$

where D_i represents partial differentiation with respect to the ith argument. Show that $\varphi^{(1)}(x, y, z, t) \equiv (F_1(x, y, t), F_2(x, y, t), F_3(x, y, z, t))$ is a (possibly local) one-parameter group of transformations on $\mathbb{R}^3$ (known as the extension or *first prolongation* of φ). Show that if $\xi\,(\partial/\partial x) + \eta\,(\partial/\partial y)$ is the infinitesimal generator of φ, then the infinitesimal generator of $\varphi^{(1)}$ is

$$\xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y} + \left[\eta_x + z(\eta_y - \xi_x) - z^2 \xi_y\right] \frac{\partial}{\partial z}, \tag{2.21}$$

where the subscripts denote partial differentiation (e.g., $\eta_x \equiv \partial\eta/\partial x$). [Strictly speaking, in (2.21), in place of x, y, ξ, η, their pullbacks under the projection of $\mathbb{R}^3$ onto $\mathbb{R}^2$ should appear.] The prolongation of a one-parameter group of diffeomorphisms is employed in the study of the symmetries of an ODE; see, e.g., Hydon (2000).

Canonical Lift of a Vector Field A differentiable mapping $\psi : M_1 \to M_2$ gives rise to a differentiable mapping $\overline{\psi} : TM_1 \to TM_2$, defined by

$$\overline{\psi}(v_p) \equiv \psi_{*p}(v_p), \quad \text{for } v_p \in T_p(M_1).$$

Since $\psi_{*p}(v_p) \in T_{\psi(p)}M_2$, we see that $\pi_2 \circ \overline{\psi} = \psi \circ \pi_1$, where π_1 is the canonical projection of TM_1 on M_1 and, similarly, π_2 is the canonical projection of TM_2 on M_2. Making use of the chain rule (1.25), one can readily verify that if $\psi_1 : M_1 \to M_2$ and $\psi_2 : M_2 \to M_3$ are two differentiable mappings, then $\overline{(\psi_2 \circ \psi_1)} = \overline{\psi}_2 \circ \overline{\psi}_1$. Hence, if $\{\varphi_t\}$ is a one-parameter group of diffeomorphisms on a manifold M, the mappings $\overline{\varphi_t}$ form a one-parameter group of diffeomorphisms on TM.

The local expression of the transformations φ_t is given by the functions $\varphi_t{}^* x^i$, where the x^i form some coordinate system on M. Then, in terms of the coordinates q^i, $\dot{q}^i$ induced on TM by the x^i, the transformations $\overline{\varphi_t}$ are locally given by the functions $\overline{\varphi_t}{}^* q^i$ and $\overline{\varphi_t}{}^* \dot{q}^i$. Since $\pi \circ \overline{\varphi_t} = \varphi_t \circ \pi$ and, by definition, $q^i = \pi^* x^i$, we obtain

$$\overline{\varphi_t}{}^* q^i = (\overline{\varphi_t}{}^* \circ \pi^*) x^i = (\pi \circ \overline{\varphi_t})^* x^i = (\varphi_t \circ \pi)^* x^i = \pi^* \big(\varphi_t{}^* x^i\big)$$

and, making use of the definitions of $\overline{\varphi_t}$ and of the coordinates $\dot{q}^i$ [see (1.27)], we find that

$$\begin{aligned}\big(\overline{\varphi_t}{}^* \dot{q}^i\big)(v_p) &= \dot{q}^i\big(\overline{\varphi_t}(v_p)\big) = \dot{q}^i\big(\varphi_{t*p}(v_p)\big) = \big(\varphi_{t*p}(v_p)\big)\big[x^i\big] = v_p\big[\varphi_t{}^* x^i\big] \\ &= \dot{q}^j(v_p)\left(\frac{\partial}{\partial x^j}\right)_p \big[\varphi_t{}^* x^i\big] = \left[\dot{q}^j\, \pi^*\left(\frac{\partial(\varphi_t{}^* x^i)}{\partial x^j}\right)\right](v_p),\end{aligned}$$

i.e.,

$$\overline{\varphi_t}{}^* \dot{q}^i = \dot{q}^j\, \pi^*\left(\frac{\partial(\varphi_t{}^* x^i)}{\partial x^j}\right).$$

Recalling that the infinitesimal generator, $\mathbf{X}$, of φ_t, is given by $\mathbf{X} = X^i\, \partial/\partial x^i$ with $X^i = (\mathrm{d}/\mathrm{d}t)(\varphi_t{}^* x^i)|_{t=0}$, from the expressions obtained above we find that the infinitesimal generator, $\overline{\mathbf{X}}$, of $\overline{\varphi_t}$ is locally given by

$$\overline{\mathbf{X}} = \big(\pi^* X^i\big)\frac{\partial}{\partial q^i} + \dot{q}^j\, \pi^*\left(\frac{\partial X^i}{\partial x^j}\right)\frac{\partial}{\partial \dot{q}^i}. \tag{2.22}$$

The vector field $\overline{\mathbf{X}}$ is called the *canonical lift* of $\mathbf{X}$ to TM.

Exercise 2.13 Find the one-parameter group of diffeomorphisms on the tangent bundle $T\mathbb{R}^2$ induced by the one-parameter group of diffeomorphisms on $\mathbb{R}^2$ defined by $\varphi_t(x, y) = (\mathrm{e}^{at}x, \mathrm{e}^{bt}y)$, with $a, b \in \mathbb{R}$. Show that its infinitesimal generator is

$$aq^1\frac{\partial}{\partial q^1} + bq^2\frac{\partial}{\partial q^2} + a\frac{\partial}{\partial \dot{q}^1} + b\frac{\partial}{\partial \dot{q}^2},$$

where the q^i and $\dot{q}^i$ are the coordinates on $T\mathbb{R}^2$ induced by the Cartesian coordinates x, y.

Exercise 2.14 Show that $\overline{[\mathbf{X}, \mathbf{Y}]} = [\overline{\mathbf{X}}, \overline{\mathbf{Y}}]$, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$.

Exercise 2.15 A (time-independent) Lagrangian is a real-valued function defined in *TM*. A differentiable curve C in M is a solution of the Euler–Lagrange equations corresponding to the Lagrangian L if, locally,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{q}^i}\big(\overline{C}(t)\big)\right] - \frac{\partial L}{\partial q^i}\big(\overline{C}(t)\big) = 0, \quad i = 1, 2, \ldots, n,$$

where $\overline{C}$ is the curve in *TM* defined by $\overline{C}(t) = C'_t$. The vector field $\mathbf{X}$ on M represents a *symmetry* of the Lagrangian L if $\overline{\mathbf{X}}L = 0$. Show that if $\mathbf{X}$ represents a symmetry of L, then

$$X^i\big(C(t)\big)\frac{\partial L}{\partial \dot{q}^i}\big(\overline{C}(t)\big)$$

is a *constant of motion*, i.e., it does not depend on t. (Note that $\pi(\overline{C}(t)) = C(t)$, hence $q^i(\overline{C}(t)) = x^i(C(t))$, and that, according to (1.28) and (1.20), $\dot{q}^i(\overline{C}(t)) = C'_t[x^i] = \mathrm{d}(x^i \circ C)/\mathrm{d}t = \mathrm{d}(q^i(\overline{C}(t)))/\mathrm{d}t$.)

2.2 Lie Derivative of Functions and Vector Fields

Let φ be a one-parameter group of transformations or a flow on M. As pointed out above, the map $\varphi_t : M \to M$, defined by $\varphi_t(x) = \varphi(x, t)$, is a differentiable mapping. For $f \in C^\infty(M)$, $\varphi_t^* f = f \circ \varphi_t$ also belongs to $C^\infty(M)$; the limit $\lim_{t\to 0} \frac{\varphi_t^* f - f}{t}$ represents the rate of change of the function f under the family of transformations φ_t.

If $\mathbf{X}$ is the infinitesimal generator of φ, the curve φ_x given by $\varphi_x(t) = \varphi(x, t)$ is the integral curve of $\mathbf{X}$ that starts at x; therefore

$$\begin{aligned}
\left(\lim_{t\to 0}\frac{\varphi_t^* f - f}{t}\right)(x) &= \lim_{t\to 0}\frac{f(\varphi_t(x)) - f(x)}{t}\\
&= \lim_{t\to 0}\frac{f(\varphi(x,t)) - f(x)}{t}\\
&= \lim_{t\to 0}\frac{f(\varphi_x(t)) - f(\varphi_x(0))}{t}\\
&= (\varphi_x)'_0[f] = \mathbf{X}_x[f]\\
&= (\mathbf{X}f)(x),
\end{aligned}$$

which shows that, for any differentiable function, the limit $\lim_{t\to 0}\frac{\varphi_t^* f - f}{t}$ exists and depends on φ only through its infinitesimal generator. This limit is called the *Lie derivative* of f with respect to $\mathbf{X}$ and is denoted by $£_{\mathbf{X}} f$. From the expression

$$£_{\mathbf{X}} f = \mathbf{X} f \tag{2.23}$$

one can derive the properties of the Lie derivative of functions.

Exercise 2.16 Show that if $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, then $\pounds_\mathbf{X}(\pounds_\mathbf{Y} f) - \pounds_\mathbf{Y}(\pounds_\mathbf{X} f) = \pounds_{[\mathbf{X},\mathbf{Y}]} f$.

Let M and N be differentiable manifolds and let $\psi : M \to N$ be a diffeomorphism. If $\mathbf{X}$ is a vector field on N, then there exists a unique vector field $\mathbf{Y}$ on M such that $\mathbf{Y}$ and $\mathbf{X}$ are ψ-related. Indeed, since $\psi^{-1} \circ \psi$ is the identity map of M, using the chain rule (1.25) we find that $(\psi^{-1})_{*\psi(x)}$ is the inverse of ψ_{*x} and, therefore, the condition that $\mathbf{Y}$ and $\mathbf{X}$ be ψ-related (i.e., $\psi_{*x}\mathbf{Y}_x = \mathbf{X}_{\psi(x)}$) has a unique solution, given by

$$\mathbf{Y}_x = \left(\psi^{-1}\right)_{*\psi(x)} \mathbf{X}_{\psi(x)}.$$

The vector field $\mathbf{Y}$ is, by definition, the *pullback* of $\mathbf{X}$ under ψ and will be denoted by $\psi^*\mathbf{X}$, that is,

$$(\psi^*\mathbf{X})_x \equiv \left(\psi^{-1}\right)_{*\psi(x)} \mathbf{X}_{\psi(x)}. \tag{2.24}$$

Note that since $\psi^*\mathbf{X}$ and $\mathbf{X}$ are ψ-related,

$$(\psi^*\mathbf{X})(\psi^* f) = \psi^*(\mathbf{X} f), \tag{2.25}$$

for $f \in C^\infty(N)$ [see (1.40)].

Exercise 2.17 Show that $\psi^*(f\mathbf{X}) = (\psi^* f)(\psi^*\mathbf{X})$ and that $\psi^*(a\mathbf{X} + b\mathbf{Y}) = a\psi^*\mathbf{X} + b\psi^*\mathbf{Y}$ for $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(N)$, $f \in C^\infty(N)$, and $a, b \in \mathbb{R}$.

Exercise 2.18 Show that if $\psi : M \to N$ is a diffeomorphism and φ is a one-parameter group of transformations on N whose infinitesimal generator is $\mathbf{X}$, then $\chi_t \equiv \psi^{-1} \circ \varphi_t \circ \psi$ is a one-parameter group of transformations on M whose infinitesimal generator is $\psi^*\mathbf{X}$ (cf. Exercise 2.8).

Exercise 2.19 Show that if $\psi_1 : M_1 \to M_2$ and $\psi_2 : M_2 \to M_3$ are diffeomorphisms, then $(\psi_2 \circ \psi_1)^*\mathbf{X} = (\psi_1{}^* \circ \psi_2{}^*)\mathbf{X}$, for $\mathbf{X} \in \mathfrak{X}(M_3)$.

Let φ be a one-parameter group of transformations or a flow on M and let $\mathbf{X}$ be its infinitesimal generator. For any vector field $\mathbf{Y}$ on M, the limit $\lim_{t\to 0} \frac{\varphi_t^*\mathbf{Y} - \mathbf{Y}}{t}$, if it exists, is called the *Lie derivative* of $\mathbf{Y}$ with respect to $\mathbf{X}$ and is denoted by $\pounds_\mathbf{X}\mathbf{Y}$.

Proposition 2.20 *Let* $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$; *then the Lie derivative of* $\mathbf{Y}$ *with respect to* $\mathbf{X}$ *exists and is equal to the Lie bracket of* $\mathbf{X}$ *and* $\mathbf{Y}$.

Proof Let f be an arbitrary differentiable function, then, using (2.25),

$$\begin{aligned}
\pounds_\mathbf{X}(\mathbf{Y} f) &= \lim_{t\to 0} \frac{\varphi_t^*(\mathbf{Y} f) - \mathbf{Y} f}{t} \\
&= \lim_{t\to 0} \frac{(\varphi_t^*\mathbf{Y})(\varphi_t^* f) - \mathbf{Y} f}{t}
\end{aligned}$$

$$= \lim_{t\to 0}\left[(\varphi_t^*\mathbf{Y})\frac{\varphi_t^* f - f}{t} + \frac{\varphi_t^*\mathbf{Y} - \mathbf{Y}}{t} f\right]$$

$$= \mathbf{Y}(£_{\mathbf{X}} f) + (£_{\mathbf{X}}\mathbf{Y}) f, \tag{2.26}$$

but $£_{\mathbf{X}} f = \mathbf{X} f$; therefore

$$\mathbf{X}(\mathbf{Y} f) = £_{\mathbf{X}}(\mathbf{Y} f) = \mathbf{Y}(\mathbf{X} f) + (£_{\mathbf{X}}\mathbf{Y}) f,$$

hence

$$(£_{\mathbf{X}}\mathbf{Y}) f = \mathbf{X}(\mathbf{Y} f) - \mathbf{Y}(\mathbf{X} f) = [\mathbf{X}, \mathbf{Y}] f,$$

which means that

$$£_{\mathbf{X}}\mathbf{Y} = [\mathbf{X}, \mathbf{Y}]. \tag{2.27}$$

□

As in the case of the relation (2.23), the formula (2.27) allows us to readily obtain the properties of the Lie derivative of vector fields. Furthermore, the relation (2.27) allows us to give a geometrical meaning to the Lie bracket.

Exercise 2.21 Show that if $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$ and $f \in C^\infty(M)$, then $£_{\mathbf{X}}(f\mathbf{Y}) = f\, £_{\mathbf{X}}\mathbf{Y} + (£_{\mathbf{X}} f)\mathbf{Y}$ [cf. (2.26)]. Also show that $£_{\mathbf{X}}(\mathbf{Y} + \mathbf{Z}) = £_{\mathbf{X}}\mathbf{Y} + £_{\mathbf{X}}\mathbf{Z}$. (*Hint*: use (2.23), (2.27), and Exercise 1.22.)

Exercise 2.22 Show that if $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(M)$, then $£_{\mathbf{X}}(£_{\mathbf{Y}}\mathbf{Z}) - £_{\mathbf{Y}}(£_{\mathbf{X}}\mathbf{Z}) = £_{[\mathbf{X},\mathbf{Y}]}\mathbf{Z}$ (cf. Exercise 2.16).

Example 2.23 The Lie derivative frequently appears in connection with symmetries. The vector field $\mathbf{Y} \in \mathfrak{X}(M)$ is invariant under the one-parameter group of diffeomorphisms φ_t if $£_{\mathbf{X}}\mathbf{Y} = 0$, where $\mathbf{X}$ is the infinitesimal generator of φ_t. For instance, in order to find all the vector fields on $\mathbb{R}^2$ invariant under rotations about the origin, it is convenient to employ polar coordinates (r, θ), so that, locally, $\mathbf{X} = \partial/\partial\theta$. The condition $£_{\mathbf{X}}\mathbf{Y} = 0$ amounts to

$$0 = \big[(\partial/\partial\theta), Y^1(\partial/\partial r) + Y^2(\partial/\partial\theta)\big] = \frac{\partial Y^1}{\partial\theta}\frac{\partial}{\partial r} + \frac{\partial Y^2}{\partial\theta}\frac{\partial}{\partial\theta},$$

where Y^1, Y^2 are the components of $\mathbf{Y}$ with respect to the natural basis induced by the coordinates (r, θ). Hence, $\mathbf{Y}$ is invariant under rotations about the origin if and only if Y^1, Y^2 are functions of r only.

Exercise 2.24 Show that if φ_t and ψ_t are two one-parameter groups of diffeomorphisms on M that commute with each other, i.e., $\varphi_t\psi_s = \psi_s\varphi_t$ for all $t, s \in \mathbb{R}$, then the Lie bracket of their infinitesimal generators is equal to zero (cf. Exercise 2.18). (The converse is also true: two vector fields $\mathbf{X}, \mathbf{Y}$ on M such that $[\mathbf{X}, \mathbf{Y}] = 0$ generate (local) one-parameter groups that commute.)

2.3 Lie Derivative of 1-Forms and Tensor Fields

Let $\psi : M \to N$ be a differentiable map. If t is a tensor field of type $\binom{0}{k}$ on N, the pullback of t under ψ, $\psi^* t$, is the tensor field on M such that

$$(\psi^* t)_p(u_p, \ldots, w_p) \equiv t_{\psi(p)}(\psi_{*p} u_p, \ldots, \psi_{*p} w_p), \tag{2.28}$$

for $u_p, \ldots, w_p \in T_p M,\ p \in M$. Given that ψ_{*p} is a linear transformation, it can readily be verified that effectively $\psi^* t$ is a tensor field of type $\binom{0}{k}$ on M.

Exercise 2.25 Let $\psi : M \to N$ be a differentiable map and let α be a linear differential form on N. Show that

$$\int_C \psi^* \alpha = \int_{\psi \circ C} \alpha,$$

for any differentiable curve C in M (see Example 1.28).

If $f \in C^\infty(N)$, the differential of f, $\mathrm{d}f$, is a tensor field of type $\binom{0}{1}$. Therefore, from (2.28)

$$(\psi^* \mathrm{d}f)_p(v_p) = \mathrm{d}f_{\psi(p)}(\psi_{*p} v_p),$$

for $v_p \in T_p M$. But from the definitions of $\mathrm{d}f$ and of the Jacobian [see (1.41) and (1.23)], we have

$$\mathrm{d}f_{\psi(p)}(\psi_{*p} v_p) = \psi_{*p} v_p[f] = v_p[\psi^* f] = \mathrm{d}(\psi^* f)_p(v_p).$$

Thus

$$\psi^* \mathrm{d}f = \mathrm{d}(\psi^* f). \tag{2.29}$$

If t and s are tensor fields of type $\binom{0}{k}$ on N and $a, b \in \mathbb{R}$, we have

$$\begin{aligned}
&\big(\psi^*(at + bs)\big)_p(u_p, \ldots, w_p) \\
&\quad = (at + bs)_{\psi(p)}(\psi_{*p} u_p, \ldots, \psi_{*p} w_p) \\
&\quad = (at_{\psi(p)} + bs_{\psi(p)})(\psi_{*p} u_p, \ldots, \psi_{*p} w_p) \\
&\quad = at_{\psi(p)}(\psi_{*p} u_p, \ldots, \psi_{*p} w_p) + bs_{\psi(p)}(\psi_{*p} u_p, \ldots, \psi_{*p} w_p) \\
&\quad = a(\psi^* t)_p(u_p, \ldots, w_p) + b(\psi^* s)_p(u_p, \ldots, w_p) \\
&\quad = (a\psi^* t + b\psi^* s)_p(u_p, \ldots, w_p),
\end{aligned}$$

for $u_p, \ldots, w_p \in T_p M$, that is,

$$\psi^*(at + bs) = a\psi^* t + b\psi^* s. \tag{2.30}$$

Similarly, if $f : N \to \mathbb{R}$

$$
\begin{aligned}
\big(\psi^*(ft)\big)_p(u_p, \ldots, w_p) &= (ft)_{\psi(p)}(\psi_{*p}u_p, \ldots, \psi_{*p}w_p) \\
&= f\big(\psi(p)\big)t_{\psi(p)}(\psi_{*p}u_p, \ldots, \psi_{*p}w_p) \\
&= (\psi^* f)(p)(\psi^* t)_p(u_p, \ldots, w_p) \\
&= \big((\psi^* f)(\psi^* t)\big)_p(u_p, \ldots, w_p);
\end{aligned}
$$

hence

$$
\psi^*(ft) = (\psi^* f)(\psi^* t). \tag{2.31}
$$

Finally, if t and s are tensor fields of type $\binom{0}{k}$ and $\binom{0}{l}$ on N, respectively, we have

$$
\begin{aligned}
\big(\psi^*(t \otimes s)\big)_p &= (t \otimes s)_{\psi(p)}(\psi_{*p}u_p, \ldots, \psi_{*p}w_p) \\
&= t_{\psi(p)}(\psi_{*p}u_p, \ldots)\, s_{\psi(p)}(\ldots, \psi_{*p}w_p) \\
&= (\psi^* t)_p(u_p, \ldots)(\psi^* s)_p(\ldots, w_p) \\
&= \big((\psi^* t) \otimes (\psi^* s)\big)_p(u_p, \ldots, w_p),
\end{aligned}
$$

for $u_p, \ldots, w_p \in T_pM$, and therefore

$$
\psi^*(t \otimes s) = (\psi^* t) \otimes (\psi^* s). \tag{2.32}
$$

Exercise 2.26 Let $\psi_1 : M_1 \to M_2$ and $\psi_2 : M_2 \to M_3$ be differentiable maps. Show that $(\psi_2 \circ \psi_1)^* t = ({\psi_1}^* \circ {\psi_2}^*)t$, for $t \in T^0_k(M_3)$.

Thus, if t is a tensor field of type $\binom{0}{k}$ on N, given locally by $t = t_{i\ldots j}\, \mathrm{d}y^i \otimes \cdots \otimes \mathrm{d}y^j$, the pullback of t under ψ is given by

$$
\begin{aligned}
\psi^* t &= \psi^*\big(t_{i\ldots j}\, \mathrm{d}y^i \otimes \cdots \otimes \mathrm{d}y^j\big) \\
&= (\psi^* t_{i\ldots j})\big(\psi^*\, \mathrm{d}y^i\big) \otimes \cdots \otimes \big(\psi^* \mathrm{d}y^j\big) \\
&= (\psi^* t_{i\ldots j})\, \mathrm{d}\big(\psi^* y^i\big) \otimes \cdots \otimes \mathrm{d}\big(\psi^* y^j\big).
\end{aligned}
$$

But $\mathrm{d}(\psi^* y^i) = (\partial(\psi^* y^i)/\partial x^l)\, \mathrm{d}x^l$, where $(x^1, \ldots, x^n)$ is a coordinate system on M; hence

$$
\psi^* t = (\psi^* t_{i\ldots j}) \frac{\partial(\psi^* y^i)}{\partial x^l} \cdots \frac{\partial(\psi^* y^j)}{\partial x^m}\, \mathrm{d}x^l \otimes \cdots \otimes \mathrm{d}x^m. \tag{2.33}
$$

This expression shows that $\psi^* t$ is differentiable if t is.

Example 2.27 In the standard treatment of ODEs one encounters expressions of the form $P\, \mathrm{d}x + Q\, \mathrm{d}y = 0$. The left-hand side of this equation can be regarded as a

1-form on some manifold, M, with local coordinates (x, y) (assuming that the functions P and Q are differentiable) and the equality to zero is to be understood considering curves, $C : I \to M$, such that $C^*(P\,\mathrm{d}x + Q\,\mathrm{d}y) = 0$. (That is, $P\,\mathrm{d}x + Q\,\mathrm{d}y$ is not equal to zero as a covector field on M; it is only its pullback under C that vanishes.) Then, for one of these curves, using the properties (2.29), (2.30), and (2.31), we have

$$(P \circ C)\,\mathrm{d}(x \circ C) + (Q \circ C)\,\mathrm{d}(y \circ C) = 0. \tag{2.34}$$

Since $x \circ C$ and $y \circ C$ (as well as $P \circ C$ and $Q \circ C$) are functions from I to $\mathbb{R}$, we can write [see (1.52)]

$$\mathrm{d}(x \circ C) = \frac{\mathrm{d}(x \circ C)}{\mathrm{d}t}\,\mathrm{d}t \quad \text{and} \quad \mathrm{d}(y \circ C) = \frac{\mathrm{d}(y \circ C)}{\mathrm{d}t}\,\mathrm{d}t,$$

where t is the usual coordinate of $\mathbb{R}$. Hence, from (2.34), we get the equivalent expression

$$(P \circ C)\frac{\mathrm{d}(x \circ C)}{\mathrm{d}t} + (Q \circ C)\frac{\mathrm{d}(y \circ C)}{\mathrm{d}t} = 0.$$

This equation alone does not determine the two functions $x \circ C$ and $y \circ C$. If, for instance, $\mathrm{d}(x \circ C)/\mathrm{d}t \neq 0$ in I (which holds if Q does not vanish), using the chain rule (regarding $x \circ C$ as the independent variable instead of t), one finds that

$$\frac{\mathrm{d}(y \circ C)}{\mathrm{d}(x \circ C)} = -\frac{P \circ C}{Q \circ C}.$$

In this manner, writing x in place of $x \circ C$ and similarly for the other functions, one obtains the first-order ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{P(x, y)}{Q(x, y)}, \tag{2.35}$$

where it is assumed that y is a function of x. According to the existence and uniqueness theorem for the solutions of the differential equations, through each point of M there passes one of these curves. In this way, equation (2.35) corresponds to the expression $P\,\mathrm{d}x + Q\,\mathrm{d}y = 0$.

Now we want to find one-parameter groups of diffeomorphisms, φ_s, on M such that, when applied to a solution curve of the differential equation expressed in the usual form, $P\,\mathrm{d}x + Q\,\mathrm{d}y = 0$, they yield another solution curve. More precisely, this corresponds to finding the one-parameter groups of diffeomorphisms such that if $C^*\alpha = 0$, where $\alpha \equiv P\,\mathrm{d}x + Q\,\mathrm{d}y$, then $(\varphi_s \circ C)^*\alpha = 0$, for all $s \in \mathbb{R}$. The previous equality amounts to $C^*({\varphi_s}^*\alpha) = 0$ (see Exercise 2.26), which is equivalent to the existence of a function $\chi_s \in C^\infty(M)$ (which may depend on s) such that ${\varphi_s}^*\alpha = \chi_s\alpha$. A one-parameter group of diffeomorphisms, φ_s, such that ${\varphi_s}^*\alpha = \chi_s\alpha$ is a *symmetry of the equation* $\alpha = 0$. (As shown in Sect. 4.3, knowing a symmetry of the equation $\alpha = 0$, or its infinitesimal generator, allows us to find the solution of the differential equation.)

Let φ be a one-parameter group of transformations or a flow on M with infinitesimal generator $\mathbf{X}$, and let t be a tensor field of type $\binom{0}{k}$ on M. If the limit $\lim_{h\to 0}\frac{\varphi_h^* t - t}{h}$ exists, it is called the Lie derivative of t with respect to $\mathbf{X}$ and is denoted by $\pounds_{\mathbf{X}} t$. The properties of the Lie derivative of tensor fields of type $\binom{0}{k}$ follow from the properties of the pullback of tensor fields. That is, given two tensor fields of type $\binom{0}{k}$ on M, s, and t, it follows from (2.32) that

$$\begin{aligned}
\pounds_{\mathbf{X}}(t\otimes s) &= \lim_{h\to 0}\frac{\varphi_h^*(t\otimes s) - t\otimes s}{h}\\
&= \lim_{h\to 0}\frac{(\varphi_h^* t)\otimes(\varphi_h^* s) - t\otimes s}{h}\\
&= \lim_{h\to 0}\left[(\varphi_h^* t)\otimes\frac{\varphi_h^* s - s}{h} + \frac{\varphi_h^* t - t}{h}\otimes s\right]\\
&= t\otimes(\pounds_{\mathbf{X}} s) + (\pounds_{\mathbf{X}} t)\otimes s.
\end{aligned}\tag{2.36}$$

If t and s are of type $\binom{0}{k}$ and $a, b \in \mathbb{R}$, by (2.30) we have

$$\begin{aligned}
\pounds_{\mathbf{X}}(at + bs) &= \lim_{h\to 0}\frac{\varphi_h^*(at+bs) - (at+bs)}{h}\\
&= \lim_{h\to 0}\frac{a\varphi_h^* t + b\varphi_h^* s - at - bs}{h}\\
&= a\pounds_{\mathbf{X}} t + b\pounds_{\mathbf{X}} s.
\end{aligned}\tag{2.37}$$

For $f \in C^\infty(M)$, using (2.31) we have

$$\begin{aligned}
\pounds_{\mathbf{X}}(ft) &= \lim_{h\to 0}\frac{\varphi_h^*(ft) - ft}{h}\\
&= \lim_{h\to 0}\frac{(\varphi_h^* f)(\varphi_h^* t) - ft}{h}\\
&= \lim_{h\to 0}\left[\varphi_h^* f\,\frac{\varphi_h^* t - t}{h} + \frac{\varphi_h^* f - f}{h}t\right]\\
&= f(\pounds_{\mathbf{X}} t) + (\pounds_{\mathbf{X}} f)t.
\end{aligned}\tag{2.38}$$

Furthermore, by (2.29), the Lie derivative of $\mathrm{d}f$ with respect to $\mathbf{X}$ is

$$\begin{aligned}
\pounds_{\mathbf{X}}\,\mathrm{d}f &= \lim_{h\to 0}\frac{\varphi_h^*\,\mathrm{d}f - \mathrm{d}f}{h}\\
&= \lim_{h\to 0}\frac{\mathrm{d}(\varphi_h^* f) - \mathrm{d}f}{h}\\
&= \mathrm{d}(\pounds_{\mathbf{X}} f).
\end{aligned}\tag{2.39}$$

Using these properties of the Lie derivative we can find the components of the Lie derivative of any tensor field of type $\binom{0}{k}$. If t is given locally by $t = t_{i\ldots j}\,\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j$, we have

$$\begin{aligned}
£_{\mathbf{X}} t &= £_{\mathbf{X}}\big(t_{i\ldots j}\,\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j\big)\\
&= (£_{\mathbf{X}} t_{i\ldots j})\,\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j\\
&\quad + t_{i\ldots j}\big(£_{\mathbf{X}}\,\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j + \cdots + \mathrm{d}x^i \otimes \cdots \otimes £_{\mathbf{X}}\,\mathrm{d}x^j\big)\\
&= (\mathbf{X} t_{i\ldots j})\,\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j\\
&\quad + t_{i\ldots j}\big[\mathrm{d}\big(£_{\mathbf{X}} x^i\big) \otimes \cdots \otimes \mathrm{d}x^j + \cdots + \mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}\big(£_{\mathbf{X}} x^j\big)\big].
\end{aligned}$$

Expressing $\mathbf{X}$ in the form $\mathbf{X} = X^l(\partial/\partial x^l)$ and using (2.23) we find that

$$£_{\mathbf{X}} x^i = \mathbf{X} x^i = X^l\left(\frac{\partial}{\partial x^l}\right)x^i = X^i;$$

hence, $\mathrm{d}(£_{\mathbf{X}} x^i) = \mathrm{d}X^i = (\partial X^i/\partial x^l)\,\mathrm{d}x^l$, and

$$\begin{aligned}
£_{\mathbf{X}} t &= (\mathbf{X} t_{i\ldots j})\,\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j\\
&\quad + t_{i\ldots j}\left(\frac{\partial X^i}{\partial x^l}\,\mathrm{d}x^l \otimes \cdots \otimes \mathrm{d}x^j + \cdots + \mathrm{d}x^i \otimes \cdots \otimes \frac{\partial X^j}{\partial x^l}\,\mathrm{d}x^l\right)\\
&= \left(X^l \frac{\partial t_{i\ldots j}}{\partial x^l} + t_{l\ldots j}\frac{\partial X^l}{\partial x^i} + \cdots + t_{i\ldots l}\frac{\partial X^l}{\partial x^j}\right)\mathrm{d}x^i \otimes \cdots \otimes \mathrm{d}x^j. \qquad (2.40)
\end{aligned}$$

Example 2.28 According to the results of Example 2.27, if $\mathbf{X}$ is the infinitesimal generator of a one-parameter group of diffeomorphisms that maps solutions of the differential equation $P\,\mathrm{d}x + Q\,\mathrm{d}y = 0$ into solutions of the same equation, then $£_{\mathbf{X}}(P\,\mathrm{d}x + Q\,\mathrm{d}y) = \nu(P\,\mathrm{d}x + Q\,\mathrm{d}y)$, where ν is some real-valued function. Writing

$$\mathbf{X} = \xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y},$$

by means of the relation (2.40) we find that

$$\begin{aligned}
\xi\frac{\partial P}{\partial x} + \eta\frac{\partial P}{\partial y} + P\frac{\partial \xi}{\partial x} + Q\frac{\partial \eta}{\partial x} &= \nu P,\\
\xi\frac{\partial Q}{\partial x} + \eta\frac{\partial Q}{\partial y} + Q\frac{\partial \eta}{\partial y} + P\frac{\partial \xi}{\partial y} &= \nu Q,
\end{aligned}$$

which can be conveniently expressed in the form (eliminating the unknown function ν)

$$\xi\frac{\partial f}{\partial x} + \eta\frac{\partial f}{\partial y} = \frac{\partial \eta}{\partial x} + \left(\frac{\partial \eta}{\partial y} - \frac{\partial \xi}{\partial x}\right)f - \frac{\partial \xi}{\partial y}f^2, \qquad (2.41)$$

where $f \equiv -P/Q$ [cf. (2.35)]. This equation, for the two functions ξ and η, has infinitely many solutions and turns out to be more convenient for finding the symmetries of the differential equation $P\,dx + Q\,dy = 0$ than the condition $\varphi_s{}^*\alpha = \chi_s \alpha$. This is so since, whereas φ must satisfy the conditions defining a one-parameter group of diffeomorphisms, the functions ξ and η only have to be differentiable. A practical way of finding some solution of (2.41) consists in proposing expressions for ξ and η containing some constants to be determined (see, e.g., Hydon 2000).

If t is a tensor field of type $\binom{0}{k}$ on M and $\mathbf{X}$ is a vector field on M, the *contraction* of t with $\mathbf{X}$, denoted by $\mathbf{X} \lrcorner t$, is the tensor field of type $\binom{0}{k-1}$ on M given by

$$(\mathbf{X} \lrcorner t)_p(v_p, \ldots, w_p) \equiv k\, t_p(\mathbf{X}_p, v_p, \ldots, w_p), \tag{2.42}$$

for $v_p, \ldots, w_p \in T_p M$ (the constant factor k appearing on the right-hand side is introduced for later convenience). If t is a tensor field of type $\binom{0}{0}$ on M, that is, t is a function from M into $\mathbb{R}$, we define $\mathbf{X} \lrcorner t \equiv 0$. Note that if α is a 1-form on M, $\mathbf{X} \lrcorner \alpha$ is the function $\alpha(\mathbf{X})$ [see (1.43)].

The contraction commutes with the pullback under diffeomorphisms; for if $\psi : M \to N$ is a diffeomorphism, t a tensor field of type $\binom{0}{k}$ on N, and $\mathbf{X}$ a vector field on N, then, since $\psi^*\mathbf{X}$ and $\mathbf{X}$ are ψ-related, we have

$$\begin{aligned}
\big[(\psi^*\mathbf{X}) \lrcorner (\psi^* t)\big]_p(v_p, \ldots, w_p) &= k(\psi^* t)_p\big((\psi^*\mathbf{X})_p, v_p, \ldots, w_p\big) \\
&= k\, t_{\psi(p)}\big(\psi_{*p}(\psi^*\mathbf{X})_p, \psi_{*p} v_p, \ldots, \psi_{*p} w_p\big) \\
&= k\, t_{\psi(p)}(\mathbf{X}_{\psi(p)}, \psi_{*p} v_p, \ldots, \psi_{*p} w_p) \\
&= (\mathbf{X} \lrcorner t)_{\psi(p)}(\psi_{*p} v_p, \ldots, \psi_{*p} w_p) \\
&= \big[\psi^*(\mathbf{X} \lrcorner t)\big]_p(v_p, \ldots, w_p),
\end{aligned}$$

for $v_p, \ldots, w_p \in T_p M$, that is,

$$\psi^*(\mathbf{X} \lrcorner t) = (\psi^*\mathbf{X}) \lrcorner (\psi^* t). \tag{2.43}$$

Hence, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$ and $t \in T^0_k(M)$, we have

$$\pounds_{\mathbf{X}}(\mathbf{Y} \lrcorner t) = (\pounds_{\mathbf{X}}\mathbf{Y}) \lrcorner t + \mathbf{Y} \lrcorner (\pounds_{\mathbf{X}} t). \tag{2.44}$$

Thus, if $t \in T^0_k(M)$ and $\mathbf{X}, \mathbf{Y}_1, \ldots, \mathbf{Y}_k \in \mathfrak{X}(M)$, repeatedly applying this relation, we obtain

$$\begin{aligned}
&\mathbf{X}\big(t(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)\big) \\
&\quad = \pounds_{\mathbf{X}}\big(t(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)\big) \\
&\quad = \frac{1}{k!}\pounds_{\mathbf{X}}(\mathbf{Y}_k \lrcorner \mathbf{Y}_{k-1} \lrcorner \cdots \lrcorner \mathbf{Y}_1 \lrcorner t)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{k!}\big[(\pounds_{\mathbf{X}}\mathbf{Y}_k)\lrcorner\mathbf{Y}_{k-1}\lrcorner\cdots\lrcorner\mathbf{Y}_1\lrcorner t + \mathbf{Y}_k\lrcorner(\pounds_{\mathbf{X}}\mathbf{Y}_{k-1})\lrcorner\mathbf{Y}_{k-2}\lrcorner\cdots\lrcorner\mathbf{Y}_1\lrcorner t \\
&\quad + \cdots + \mathbf{Y}_k\lrcorner\mathbf{Y}_{k-1}\lrcorner\cdots\lrcorner(\pounds_{\mathbf{X}}\mathbf{Y}_1)\lrcorner t + \mathbf{Y}_k\lrcorner\mathbf{Y}_{k-1}\lrcorner\cdots\lrcorner\mathbf{Y}_1\lrcorner(\pounds_{\mathbf{X}}t)\big] \\
&= t(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \pounds_{\mathbf{X}}\mathbf{Y}_k) + t(\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \pounds_{\mathbf{X}}\mathbf{Y}_{k-1}, \mathbf{Y}_k) + \cdots \\
&\quad + t(\pounds_{\mathbf{X}}\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k) + (\pounds_{\mathbf{X}}t)(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) \\
&= (\pounds_{\mathbf{X}}t)(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) + \sum_{i=1}^{k} t(\mathbf{Y}_1, \ldots, \pounds_{\mathbf{X}}\mathbf{Y}_i, \ldots, \mathbf{Y}_k),
\end{aligned}
$$

that is,

$$
(\pounds_{\mathbf{X}}t)(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) = \mathbf{X}\big(t(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)\big) - \sum_{i=1}^{k} t\big(\mathbf{Y}_1, \ldots, [\mathbf{X}, \mathbf{Y}_i], \ldots, \mathbf{Y}_k\big). \quad (2.45)
$$

Exercise 2.29 Show that all the properties of the Lie derivative of tensor fields of type $\binom{0}{k}$ follow from (2.45).

Exercise 2.30 Show that if $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$ and $t \in T_k^0(M)$, then $\pounds_{\mathbf{X}}(\pounds_{\mathbf{Y}}t) - \pounds_{\mathbf{Y}}(\pounds_{\mathbf{X}}t) = \pounds_{[\mathbf{X},\mathbf{Y}]}t$.

Exercise 2.31 Show that if $\mathbf{X} \in \mathfrak{X}(M)$ and $t \in T_k^0(M)$, then $\pounds_{\mathbf{X}}(\mathbf{X}\lrcorner t) = \mathbf{X}\lrcorner(\pounds_{\mathbf{X}}t)$.

Exercise 2.32 Let t be a differentiable tensor field of type $\binom{k}{l}$ on M. Assuming that the first k arguments of t are covectors and defining $\pounds_{\mathbf{X}}t$ by

$$
\begin{aligned}
&(\pounds_{\mathbf{X}}t)(\alpha_1, \ldots, \alpha_k, \mathbf{Y}_1, \ldots, \mathbf{Y}_l) \\
&\quad \equiv \mathbf{X}\big(t(\alpha_1, \ldots, \alpha_k, \mathbf{Y}_1, \ldots, \mathbf{Y}_l)\big) \\
&\qquad - \sum_{i=1}^{k} t(\alpha_1, \ldots, \pounds_{\mathbf{X}}\alpha_i, \ldots, \alpha_k, \mathbf{Y}_1, \ldots, \mathbf{Y}_l) \\
&\qquad - \sum_{i=1}^{l} t(\alpha_1, \ldots, \alpha_k, \mathbf{Y}_1, \ldots, \pounds_{\mathbf{X}}\mathbf{Y}_i, \ldots, \mathbf{Y}_l),
\end{aligned}
$$

for $\alpha_1, \ldots, \alpha_k \in \Lambda^1(M)$, $\mathbf{Y}_1, \ldots, \mathbf{Y}_l \in \mathfrak{X}(M)$, show that $\pounds_{\mathbf{X}}t$ is a differentiable tensor field of type $\binom{k}{l}$ and that $\pounds_{\mathbf{X}}(t \otimes s) = (\pounds_{\mathbf{X}}t) \otimes s + t \otimes (\pounds_{\mathbf{X}}s)$ for any pair of mixed tensor fields.

Chapter 3
Differential Forms

Differential forms are completely skew-symmetric tensor fields. They are applied in some areas of physics, mainly in thermodynamics and classical mechanics, and of mathematics, such as differential equations, differential geometry, Lie groups, and differential topology. Many of the applications of differential forms are presented in subsequent chapters.

3.1 The Algebra of Forms

Definition 3.1 Let M be a differentiable manifold. A *differential form* of degree k, or *k-form*, ω, on M, is a completely skew-symmetric differentiable tensor field of type $\binom{0}{k}$ on M, that is,

$$\omega(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_k) = -\omega(\mathbf{X}_1, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k), \qquad (3.1)$$

$1 \leq i < j \leq k$, for $\mathbf{X}_1, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$; a 0-form is a differentiable real-valued function on M.

Starting from an arbitrary tensor field, t, of type $\binom{0}{k}$, one can construct a completely skew-symmetric tensor field of the same type. Let S_k be the group of all permutations of the numbers $(1, 2, \ldots, k)$ and let $\operatorname{sgn}\sigma$ be the sign of the permutation $\sigma \in \mathrm{S}_k$ ($\operatorname{sgn}\sigma = 1$ if σ is even, $\operatorname{sgn}\sigma = -1$ if σ is odd). We define $\mathscr{A}t$ by

$$\mathscr{A}t(\mathbf{X}_1, \ldots, \mathbf{X}_k) \equiv \frac{1}{k!} \sum_{\sigma \in \mathrm{S}_k} (\operatorname{sgn}\sigma)\, t(\mathbf{X}_{\sigma(1)}, \ldots, \mathbf{X}_{\sigma(k)}), \qquad (3.2)$$

for $\mathbf{X}_1, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$. It can readily be seen that $\mathscr{A}t$ is completely skew-symmetric, and that if t and s are tensor fields of type $\binom{0}{k}$, then $\mathscr{A}(t+s) = \mathscr{A}t + \mathscr{A}s$ and $\mathscr{A}(ft) = f\mathscr{A}t$, for $f : M \to \mathbb{R}$; furthermore if t is skew-symmetric, then $\mathscr{A}t = t$, so that $\mathscr{A}^2 = \mathscr{A}$.

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_3,

The set of the k-forms on M, which will be denoted here by $\Lambda^k(M)$, is a submodule of $T^0_k(M)$, since the sum of two k-forms, the product of a k-form by a scalar and the product of a k-form by a function $f \in C^\infty(M) = \Lambda^0(M)$ are also k-forms, as can be verified directly from the definition of the operations in $T^0_k(M)$. By contrast, the tensor product of a k-form by an l-form is skew-symmetric, separately, in its first k arguments and in its last l arguments, but it is not necessarily completely skew-symmetric in its $k+l$ arguments (except in the case where k or l is zero); nevertheless, from the tensor product of two differential forms one can obtain a completely skew-symmetric tensor field with the aid of the map $\mathscr{A}$.

Definition 3.2 If ω is a k-form and η is an l-form on M, the *exterior*, or *wedge, product* of ω by η, $\omega \wedge \eta$, is defined by

$$\omega \wedge \eta = \mathscr{A}(\omega \otimes \eta). \tag{3.3}$$

(Some authors employ the definition

$$\omega \wedge \eta = \frac{(k+l)!}{k!\,l!}\mathscr{A}(\omega \otimes \eta),$$

with which some numerical factors that appear in several expressions [e.g., (2.42), (3.7), and (3.28)] are avoided, but it makes it necessary to introduce some factors in other expressions. However, some important formulas, such as (3.27), (3.35), and (3.39), are equally valid whether one makes use of the conventions followed here in the definitions of the exterior product, the contraction, and the exterior derivative, or in the alternative conventions.)

The exterior product of ω by η is then a $(k+l)$-form. (Note that if ω is a k-form and f is a 0-form, we have $f \wedge \omega = \mathscr{A}(f \otimes \omega) = \mathscr{A}(f\omega) = f\mathscr{A}\omega = f\omega = \omega \wedge f$.)

From the properties of $\mathscr{A}$ it follows that if $\omega, \omega_1, \omega_2 \in \Lambda^k(M)$ and $\eta \in \Lambda^l(M)$, then

$$(a\omega_1 + b\omega_2) \wedge \eta = a(\omega_1 \wedge \eta) + b(\omega_2 \wedge \eta) \tag{3.4}$$

and

$$(f\omega) \wedge \eta = \omega \wedge (f\eta) = f(\omega \wedge \eta), \tag{3.5}$$

for $a, b \in \mathbb{R}$, $f \in \Lambda^0(M)$. The exterior product is associative but not always commutative [see (3.23)]. If α, β, and γ are differential forms on M, it can be shown that

$$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) = \mathscr{A}(\alpha \otimes \beta \otimes \gamma). \tag{3.6}$$

If α and β are 1-forms, applying the definition of the exterior product we have

$$\begin{aligned}(\alpha \wedge \beta)(\mathbf{X}_1, \mathbf{X}_2) &= \mathscr{A}(\alpha \otimes \beta)(\mathbf{X}_1, \mathbf{X}_2) \\ &= \frac{1}{2!}\big[(\alpha \otimes \beta)(\mathbf{X}_1, \mathbf{X}_2) - (\alpha \otimes \beta)(\mathbf{X}_2, \mathbf{X}_1)\big]\end{aligned}$$

$$= \frac{1}{2}\big[\alpha(\mathbf{X}_1)\beta(\mathbf{X}_2) - \alpha(\mathbf{X}_2)\beta(\mathbf{X}_1)\big]$$

$$= \frac{1}{2}(\alpha \otimes \beta - \beta \otimes \alpha)(\mathbf{X}_1, \mathbf{X}_2), \quad \text{for } \mathbf{X}_1, \mathbf{X}_2 \in \mathfrak{X}(M),$$

that is,

$$\alpha \wedge \beta = \frac{1}{2}(\alpha \otimes \beta - \beta \otimes \alpha) = -\beta \wedge \alpha, \quad \text{for } \alpha, \beta \in \Lambda^1(M). \tag{3.7}$$

Combining the definition of the contraction (2.42) with (3.7) one finds that if $\mathbf{X}$ is a vector field on M, then for any $v_p \in T_pM$,

$$\begin{aligned}
\big[\mathbf{X} \lrcorner (\alpha \wedge \beta)\big]_p(v_p) &= 2(\alpha \wedge \beta)_p(\mathbf{X}_p, v_p) \\
&= (\alpha \otimes \beta - \beta \otimes \alpha)_p(\mathbf{X}_p, v_p) \\
&= \alpha_p(\mathbf{X}_p)\beta_p(v_p) - \beta_p(\mathbf{X}_p)\alpha_p(v_p) \\
&= \big[(\mathbf{X} \lrcorner \alpha)\beta - (\mathbf{X} \lrcorner \beta)\alpha\big]_p(v_p),
\end{aligned}$$

which means that

$$\mathbf{X} \lrcorner (\alpha \wedge \beta) = (\mathbf{X} \lrcorner \alpha)\beta - (\mathbf{X} \lrcorner \beta)\alpha, \quad \text{for } \alpha, \beta \in \Lambda^1(M). \tag{3.8}$$

Let $(x^1, \ldots, x^n)$ be a local coordinate system on M. A k-form possesses the local expression [see (1.57)]

$$\omega = \omega_{i_1 \ldots i_k}\, \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}, \tag{3.9}$$

with

$$\omega_{i_1 \ldots i_k} = \omega\left(\left(\frac{\partial}{\partial x^{i_1}}\right), \ldots, \left(\frac{\partial}{\partial x^{i_k}}\right)\right). \tag{3.10}$$

As a consequence of the skew-symmetry of ω, its components $\omega_{i_1 \ldots i_k}$ are completely skew-symmetric in all their indices and $\omega = \mathscr{A}(\omega)$. Therefore, making use of the properties of $\mathscr{A}$ we have

$$\begin{aligned}
\omega &= \mathscr{A}(\omega) \\
&= \omega_{i_1 \ldots i_k} \mathscr{A}\big(\mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}\big) \\
&= \omega_{i_1 \ldots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.
\end{aligned} \tag{3.11}$$

Since the differentials of the coordinates are 1-forms, from (3.7) it follows that

$$\begin{aligned}
&\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_j} \wedge \cdots \wedge \mathrm{d}x^{i_l} \wedge \cdots \wedge \mathrm{d}x^{i_k} \\
&\quad = -\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \wedge \cdots \wedge \mathrm{d}x^{i_j} \wedge \cdots \wedge \mathrm{d}x^{i_k},
\end{aligned} \tag{3.12}$$

and therefore $\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} = 0$ if one of the values of the indices $i_1, \ldots, i_k$ appears more than once. Hence, if ω is a k-form with $k > n$ then $\omega = 0$, since

$\omega = \omega_{i_1 \dots i_k} \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$, and for $k > n$ necessarily some value of the indices $i_1, \dots, i_k$ will appear more than once.

Example 3.3 Let M be a manifold of dimension two, with local coordinates (q, p). Given a function $H \in C^\infty(M \times \mathbb{R})$, there exists only one vector field, $\mathbf{X}$, on the manifold $M \times \mathbb{R}$, such that $\mathbf{X}t = 1$, where t is the usual coordinate of $\mathbb{R}$, and

$$\mathbf{X} \lrcorner (\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) = 0. \tag{3.13}$$

Indeed, the condition $\mathbf{X}t = 1$ is equivalent to $\mathbf{X}$ having the local expression

$$\mathbf{X} = A \frac{\partial}{\partial q} + B \frac{\partial}{\partial p} + \frac{\partial}{\partial t}, \tag{3.14}$$

where A and B are functions of $M \times \mathbb{R}$ in $\mathbb{R}$. From (1.52) and (3.12) one finds that (3.13) amounts to

$$\mathbf{X} \lrcorner \left(\mathrm{d}p \wedge \mathrm{d}q - \frac{\partial H}{\partial q} \mathrm{d}q \wedge \mathrm{d}t - \frac{\partial H}{\partial p} \mathrm{d}p \wedge \mathrm{d}t \right) = 0, \tag{3.15}$$

and making use of (3.8) one has

$$\begin{aligned}
0 &= \mathrm{d}p(\mathbf{X})\, \mathrm{d}q - \mathrm{d}q(\mathbf{X})\, \mathrm{d}p - \frac{\partial H}{\partial q} \mathrm{d}q(\mathbf{X})\, \mathrm{d}t + \frac{\partial H}{\partial q} \mathrm{d}t(\mathbf{X})\, \mathrm{d}q \\
&\quad - \frac{\partial H}{\partial p} \mathrm{d}p(\mathbf{X})\, \mathrm{d}t + \frac{\partial H}{\partial p} \mathrm{d}t(\mathbf{X})\, \mathrm{d}p \\
&= \left[\mathrm{d}p(\mathbf{X}) + \frac{\partial H}{\partial q} \mathrm{d}t(\mathbf{X}) \right] \mathrm{d}q - \left[\mathrm{d}q(\mathbf{X}) - \frac{\partial H}{\partial p} \mathrm{d}t(\mathbf{X}) \right] \mathrm{d}p \\
&\quad - \left[\frac{\partial H}{\partial q} \mathrm{d}q(\mathbf{X}) + \frac{\partial H}{\partial p} \mathrm{d}p(\mathbf{X}) \right] \mathrm{d}t,
\end{aligned}$$

which means that the expressions inside the brackets must be separately equal to zero. Then, making use of (3.14) and (1.45), one obtains $A = \partial H/\partial p$, $B = -\partial H/\partial q$, that is,

$$\mathbf{X} = \frac{\partial H}{\partial p} \frac{\partial}{\partial q} - \frac{\partial H}{\partial q} \frac{\partial}{\partial p} + \frac{\partial}{\partial t}, \tag{3.16}$$

thus proving the assertion above. The integral curves of $\mathbf{X}$ are determined by the equations

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}, \qquad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}, \tag{3.17}$$

which, in the context of classical mechanics, are known as the *Hamilton equations*.

According to this, if (Q, P, t) is a second coordinate system on $M \times \mathbb{R}$ such that

$$\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t = \mathrm{d}P \wedge \mathrm{d}Q - \mathrm{d}K \wedge \mathrm{d}t, \tag{3.18}$$

where K is some function of $M \times \mathbb{R}$ in $\mathbb{R}$ [see (3.13)], then equations (3.17) are equivalent to

$$\frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\partial K}{\partial P}, \qquad \frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial K}{\partial Q},$$

that is, the form of the Hamilton equations is maintained if (3.18) holds. (Note that Q, P need not be coordinates on M, in the same manner as the polar coordinates (r, θ) on $\mathbb{R}^2$, which can be identified with $\mathbb{R} \times \mathbb{R}$, are not formed by one coordinate function on the first copy of $\mathbb{R}$ and one coordinate on the second copy of $\mathbb{R}$.) The relationship between the coordinate systems (q, p, t) and (Q, P, t) is called a *canonical transformation*. (See also Sect. 8.7.)

Exercise 3.4 Show that the relationship between two coordinate systems on $P \times \mathbb{R}$, (q, p, t) and (Q, P, t), is a canonical transformation if and only if

$$\frac{\partial P}{\partial p}\frac{\partial Q}{\partial q} - \frac{\partial P}{\partial q}\frac{\partial Q}{\partial p} = 1. \tag{3.19}$$

(Among other things, this means that if the condition (3.19) holds, then there exists a function $K \in C^\infty(P \times \mathbb{R})$ such that equation (3.18) is satisfied.) Usually, a canonical transformation is defined as a transformation satisfying (3.19).

From (3.12) it follows that if $n = \dim M$, then the exterior product of n differentials of the coordinates satisfies

$$\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_n} = \varepsilon^{i_1 \dots i_n}\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n, \tag{3.20}$$

where

$$\varepsilon^{i_1 \dots i_n} \equiv \begin{cases} 1 & \text{if } (i_1, \dots, i_n) \text{ is an even permutation of } (1, 2, \dots, n), \\ -1 & \text{if } (i_1, \dots, i_n) \text{ is an odd permutation of } (1, 2, \dots, n), \\ 0 & \text{if one of the values of the indices appears repeated.} \end{cases} \tag{3.21}$$

Hence, if $\omega \in \Lambda^n(M)$, using the fact that the components of ω are totally skew-symmetric and that there exist $n!$ permutations for a set of n objects, we have

$$\omega = \omega_{i_1 \dots i_n}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_n} = n!\, \omega_{12 \dots n}\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n. \tag{3.22}$$

Let $\omega \in \Lambda^k(M)$ and $\eta \in \Lambda^l(M)$ be given locally by $\omega = \omega_{i_1 \dots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ and $\eta = \eta_{j_1 \dots j_l}\, \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l}$, using the associativity of the exterior product and its skew-symmetry for the 1-forms [see (3.7)], we have

$$\begin{aligned} \omega \wedge \eta &= \omega_{i_1 \dots i_k} \eta_{j_1 \dots j_l}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l} \\ &= (-1)^{kl} \omega_{i_1 \dots i_k} \eta_{j_1 \dots j_l}\, \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \\ &= (-1)^{kl} \eta \wedge \omega. \end{aligned} \tag{3.23}$$

This implies that a form of even degree commutes under the exterior product with any form and that the exterior product of two differential forms of odd degrees is anticommutative. The set of all the differential forms on M forms an associative algebra with the exterior product.

Expression (3.11) shows that any k-form, with $k > 1$, can be expressed locally in terms of the exterior products of the differentials of the coordinates of some chart; however, from (3.12) it follows that such products are not independent among themselves, so that the equation $c_{i_1\ldots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} = 0$ *does not* imply that the coefficients $c_{i_1\ldots i_k}$ are equal to zero, but only that the *totally skew-symmetric part* of $c_{i_1\ldots i_k}$, given by

$$c_{[i_1\ldots i_k]} \equiv \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, c_{i_{\sigma(1)}\ldots i_{\sigma(k)}}, \tag{3.24}$$

is zero. This fact follows from the definitions (3.2) and (3.3); for if $c_{i_1\ldots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} = 0$, then

$$\begin{aligned} 0 &= \left(c_{i_1\ldots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}\right)\left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_k}}\right) \\ &= \left[\mathscr{A}\left(c_{i_1\ldots i_k}\, \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}\right)\right]\left(\frac{\partial}{\partial x^{j_1}}, \ldots, \frac{\partial}{\partial x^{j_k}}\right) \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, c_{i_1\ldots i_k} \delta^{i_1}_{j_{\sigma(1)}} \cdots \delta^{i_k}_{j_{\sigma(k)}} \\ &= \frac{1}{k!} \sum_{\sigma \in S_k} (\operatorname{sgn} \sigma)\, c_{j_{\sigma(1)}\ldots j_{\sigma(k)}}. \end{aligned}$$

Since $\psi^*(\omega \otimes \eta) = (\psi^*\omega) \otimes (\psi^*\eta)$ for any differentiable map $\psi : M \to N$ and tensor fields ω, η on N [see (2.32)], from (2.30), (3.2), and (3.3) it follows that

$$\psi^*(\omega \wedge \eta) = (\psi^*\omega) \wedge (\psi^*\eta), \tag{3.25}$$

for $\omega \in \Lambda^k(N)$, $\eta \in \Lambda^l(N)$ and, therefore,

$$\pounds_{\mathbf{X}}(\omega \wedge \eta) = (\pounds_{\mathbf{X}}\omega) \wedge \eta + \omega \wedge (\pounds_{\mathbf{X}}\eta), \tag{3.26}$$

for $\omega \in \Lambda^k(M)$, $\eta \in \Lambda^l(M)$, $\mathbf{X} \in \mathfrak{X}(M)$.

If ω is a k-form on M and $\mathbf{X} \in \mathfrak{X}(M)$, the contraction $\mathbf{X} \lrcorner\, \omega$ is a $(k-1)$-form; in other words, $\mathbf{X} \lrcorner$ is a map of $\Lambda^k(M)$ into $\Lambda^{k-1}(M)$. The operation of contraction is also called *interior product* and $\mathbf{X} \lrcorner\, \omega$ is also denoted by $i(\mathbf{X})\omega$ or by $i_{\mathbf{X}}\omega$. If $\mathbf{Y}$ is another vector field on M, then we have $\mathbf{Y} \lrcorner (\mathbf{X} \lrcorner\, \omega) = -\mathbf{X} \lrcorner (\mathbf{Y} \lrcorner\, \omega)$, by virtue of the skew-symmetry of ω; therefore $\mathbf{X} \lrcorner (\mathbf{X} \lrcorner\, \omega) = 0$, for $\omega \in \Lambda^k(M)$, $\mathbf{X} \in \mathfrak{X}(M)$.

By means of a lengthy computation it can be shown that if ω is a k-form and η is an l-form, then

$$\mathbf{X} \lrcorner (\omega \wedge \eta) = (\mathbf{X} \lrcorner\, \omega) \wedge \eta + (-1)^k \omega \wedge (\mathbf{X} \lrcorner\, \eta), \tag{3.27}$$

for $\mathbf{X} \in \mathfrak{X}(M)$ [cf. (3.8)]. Owing to this relation it is said that the map $\mathbf{X}\lrcorner : \Lambda^k(M) \to \Lambda^{k-1}(M)$ is an *antiderivation*.

Exercise 3.5 Show that if $\omega = \omega_{i_1 \dots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ and $\mathbf{X} = X^j(\partial/\partial x^j)$, then $\mathbf{X}\lrcorner\omega = k\, X^j \omega_{j i_1 \dots i_{k-1}}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_{k-1}}$.

3.2 The Exterior Derivative

Definition 3.6 Let ω be a k-form on M; its *exterior derivative*, $\mathrm{d}\omega$, is given by

$$\begin{aligned}(k+1)\,\mathrm{d}\omega(\mathbf{X}_1, \dots, \mathbf{X}_{k+1}) \\ \equiv \sum_{i=1}^{k+1} (-1)^{i+1} \mathbf{X}_i \big(\omega(\mathbf{X}_1, \dots, \widehat{\mathbf{X}_i}, \dots, \mathbf{X}_{k+1})\big) \\ + \sum_{i<j} (-1)^{i+j} \omega\big([\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_1, \dots, \widehat{\mathbf{X}_i}, \dots, \widehat{\mathbf{X}_j}, \dots, \mathbf{X}_{k+1}\big), \end{aligned} \tag{3.28}$$

for $\mathbf{X}_1, \dots, \mathbf{X}_{k+1} \in \mathfrak{X}(M)$, where the symbol $\widehat{\ }$ on $\mathbf{X}_i$ indicates that $\mathbf{X}_i$ is omitted. The coefficient $(k+1)$ on the left-hand side of the definition has the same origin as the coefficient appearing in the definition of the contraction; both are included in order for the contraction and the exterior differentiation to be antiderivations of the algebra of forms of M.

It is convenient to present in a more explicit way the definition (3.28) for the degrees that will be encountered more frequently in what follows. When $k = 0$, the definition (3.28) gives, for $f \in \Lambda^0(M)$,

$$\mathrm{d}f(\mathbf{X}) = \mathbf{X}f. \tag{3.29}$$

Comparing with (1.45), we see that the exterior derivative of a function f is just the differential of f. In the case of a 1-form α we have

$$2\,\mathrm{d}\alpha(\mathbf{X}, \mathbf{Y}) = \mathbf{X}\big(\alpha(\mathbf{Y})\big) - \mathbf{Y}\big(\alpha(\mathbf{X})\big) - \alpha\big([\mathbf{X}, \mathbf{Y}]\big), \tag{3.30}$$

and for a differential form of degree 2, ω,

$$\begin{aligned}3\,\mathrm{d}\omega(\mathbf{X}, \mathbf{Y}, \mathbf{Z}) = \mathbf{X}\big(\omega(\mathbf{Y}, \mathbf{Z})\big) + \mathbf{Y}\big(\omega(\mathbf{Z}, \mathbf{X})\big) + \mathbf{Z}\big(\omega(\mathbf{X}, \mathbf{Y})\big) \\ - \omega\big([\mathbf{X}, \mathbf{Y}], \mathbf{Z}\big) - \omega\big([\mathbf{Y}, \mathbf{Z}], \mathbf{X}\big) - \omega\big([\mathbf{Z}, \mathbf{X}], \mathbf{Y}\big), \end{aligned} \tag{3.31}$$

where we have made use of the skew-symmetry of ω.

Exercise 3.7 Show that the expressions (3.30) and (3.31) effectively define differential forms. Using (3.29)–(3.31), show that for $f \in C^\infty(M)$, $\mathrm{dd}f = 0$ and that if α is a 1-form, then $\mathrm{dd}\alpha = 0$.

Exercise 3.8 Show that the definitions (3.29)–(3.31) imply that $\pounds_{\mathbf{X}}\omega = \mathbf{X} \lrcorner\, \mathrm{d}\omega + \mathrm{d}(\mathbf{X} \lrcorner\, \omega)$, if ω is a differential form of degree 0, 1, or 2, and $\mathbf{X}$ is a differentiable vector field.

From the definition (3.28) it follows that $\mathrm{d}\omega$ is completely skew-symmetric and $\mathbb{R}$-linear in each of its $k+1$ arguments [this is more easily seen in the specific cases (3.29)–(3.31)]. In order to show that it is also a tensor field, it is sufficient to show that $\mathrm{d}\omega$ is $C^\infty(M)$-linear in its first argument, $\mathrm{d}\omega(f\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{k+1}) = f\,\mathrm{d}\omega(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{k+1})$, since

$$\mathrm{d}\omega(\mathbf{X}_1, \ldots, f\mathbf{X}_i, \ldots, \mathbf{X}_{k+1}) = (-1)^{i-1}\,\mathrm{d}\omega(f\mathbf{X}_i, \mathbf{X}_1, \ldots, \widehat{f\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1}),$$

by virtue of the skew-symmetry of $\mathrm{d}\omega$. Making use of the definition (3.28) we find that, for $f \in \Lambda^0(M)$ $(= C^\infty(M))$,

$$\begin{aligned}
&(k+1)\,\mathrm{d}\omega(f\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{k+1}) \\
&\quad = (f\mathbf{X}_1)\big(\omega(\mathbf{X}_2, \ldots, \mathbf{X}_{k+1})\big) \\
&\qquad + \sum_{i=2}^{k+1}(-1)^{i+1}\mathbf{X}_i\big(\omega(f\mathbf{X}_1, \ldots, \widehat{\mathbf{X}}_i, \ldots, \mathbf{X}_{k+1})\big) \\
&\qquad - \sum_{j>i}(-1)^j\omega\big([f\mathbf{X}_1, \mathbf{X}_j], \mathbf{X}_2, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big) \\
&\qquad + \sum_{1<i<j}(-1)^{i+j}\omega\big([\mathbf{X}_i, \mathbf{X}_j], f\mathbf{X}_1, \ldots, \widehat{\mathbf{X}}_i, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big).
\end{aligned}$$

Using now (1.34), the fact that ω is a tensor field, and that $[f\mathbf{X}_1, \mathbf{X}_j] = f[\mathbf{X}_1, \mathbf{X}_j] - (\mathbf{X}_j f)\mathbf{X}_1$ (see Exercise 1.32), this expression becomes

$$\begin{aligned}
&(k+1)\,\mathrm{d}\omega(f\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{k+1}) \\
&\quad = f\left[\mathbf{X}_1\big(\omega(\mathbf{X}_2, \ldots, \mathbf{X}_{k+1})\big)\right] \\
&\qquad + \sum_{i=2}^{k+1}(-1)^{i+1}\mathbf{X}_i\big(f\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}}_i, \ldots, \mathbf{X}_{k+1})\big) \\
&\qquad - \sum_{j>1}(-1)^j\omega\big(f[\mathbf{X}_1, \mathbf{X}_j] - (\mathbf{X}_j f)\mathbf{X}_1, \mathbf{X}_2, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big) \\
&\qquad + \sum_{1<i<j}(-1)^{i+j}f\omega\big([\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_1, \ldots, \widehat{\mathbf{X}}_i, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big)
\end{aligned}$$

$$
\begin{aligned}
&= f\left[\mathbf{X}_1\big(\omega(\mathbf{X}_2, \ldots, \mathbf{X}_{k+1})\big)\right] \\
&\quad + \sum_{i=2}^{k+1}(-1)^{i+1}\big[f\mathbf{X}_i\big(\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1})\big) \\
&\quad + (\mathbf{X}_i f)\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1})\big] \\
&\quad - \sum_{j>1}(-1)^j\big\{f\omega\big([\mathbf{X}_1, \mathbf{X}_j], \mathbf{X}_2, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big) \\
&\quad - (\mathbf{X}_j f)\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1})\big\} \\
&\quad + \sum_{1<i<j}(-1)^{i+j} f\omega\big([\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big) \\
&= f(k+1)\,\mathrm{d}\omega(\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}).
\end{aligned}
$$

From the definition (3.28) we also see that $\mathrm{d}\omega$ is differentiable, and we conclude that $\mathrm{d}\omega$ is a $(k+1)$-form or, equivalently, that d is a map from $\Lambda^k(M)$ into $\Lambda^{k+1}(M)$.

If ω_1 and ω_2 are k-forms and $a, b \in \mathbb{R}$, from the definition of d we directly see that

$$
\mathrm{d}(a\omega_1 + b\omega_2) = a\,\mathrm{d}\omega_1 + b\,\mathrm{d}\omega_2. \tag{3.32}
$$

On the other hand, for $\omega \in \Lambda^k(M)$ and $f \in \Lambda^0(M)$ we have

$$
\begin{aligned}
&(k+1)\,\mathrm{d}(f\omega)(\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}) \\
&= \sum_{i=1}^{k+1}(-1)^{i+1}\mathbf{X}_i\big((f\omega)(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1})\big) \\
&\quad + \sum_{i<j}(-1)^{i+j}(f\omega)\big([\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big) \\
&= \sum_{i=1}^{k+1}(-1)^{i+1}\big[f\mathbf{X}_i\big(\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1})\big) \\
&\quad + (\mathbf{X}_i f)\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1})\big] \\
&\quad + \sum_{i<j}(-1)^{i+j} f\omega\big([\mathbf{X}_i, \mathbf{X}_j], \mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \widehat{\mathbf{X}_j}, \ldots, \mathbf{X}_{k+1}\big) \\
&= (k+1) f\,\mathrm{d}\omega(\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}) \\
&\quad + \sum_{i=1}^{k+1}(-1)^{i+1}(\mathbf{X}_i f)\,\omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1})
\end{aligned}
$$

$$
\begin{aligned}
&= (k+1) f \,\mathrm{d}\omega(\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}) \\
&\quad + \sum_{i=1}^{k+1} (-1)^{i+1} \,\mathrm{d}f(\mathbf{X}_i)\, \omega(\mathbf{X}_1, \ldots, \widehat{\mathbf{X}_i}, \ldots, \mathbf{X}_{k+1}) \\
&= (k+1) f \,\mathrm{d}\omega(\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}) + (k+1)(\mathrm{d}f \wedge \omega)(\mathbf{X}_1, \ldots, \mathbf{X}_{k+1}),
\end{aligned}
$$

that is,

$$
\mathrm{d}(f\omega) = f\,\mathrm{d}\omega + \mathrm{d}f \wedge \omega. \tag{3.33}
$$

Expressing $\omega \in \Lambda^k(M)$ in local coordinates as $\omega = \omega_{i_1\ldots i_k}\,\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$, with $\omega_{i_1\ldots i_k} \in \Lambda^0(M)$, from the properties (3.32) and (3.33) we find that

$$
\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}\big(\omega_{i_1\ldots i_k} \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}\big) \\
&= \omega_{i_1\ldots i_k}\,\mathrm{d}\big(\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}\big) + \mathrm{d}\omega_{i_1\ldots i_k} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.
\end{aligned}
$$

The exterior derivative of $\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ is equal to zero, as can be seen by applying the definition (3.28) to calculate $\mathrm{d}(\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k})((\partial/\partial x^{j_1}), \ldots, (\partial/\partial x^{j_{k+1}}))$, using that $[(\partial/\partial x^i), (\partial/\partial x^j)] = 0$. Thus, $\mathrm{d}\omega$ is given locally by

$$
\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}\omega_{i_1\ldots i_k} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \\
&= \left(\frac{\partial}{\partial x^l}\right)\omega_{i_1\ldots i_k}\,\mathrm{d}x^l \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.
\end{aligned} \tag{3.34}
$$

Exercise 3.9 Derive the expression (3.34) for the components of $\mathrm{d}\omega$ without employing (3.33), directly from (3.10) and (3.28). Making use of (3.34), demonstrate the validity of (3.33).

With the aid of the local expression (3.34) for d one can show that the exterior differentiation is an *antiderivation* of the algebra of forms, that is, if $\omega \in \Lambda^k(M)$ and $\eta \in \Lambda^l(M)$, then

$$
\mathrm{d}(\omega \wedge \eta) = \mathrm{d}\omega \wedge \eta + (-1)^k \omega \wedge \mathrm{d}\eta. \tag{3.35}
$$

Indeed, expressing ω and η as $\omega = \omega_{i_1\ldots i_k}\,\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ and $\eta = \eta_{j_1\ldots j_l}\,\mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l}$, respectively, and using the expression for the differential of a product of functions (1.47), from (3.34) we have

$$
\begin{aligned}
\mathrm{d}(\omega \wedge \eta) &= \mathrm{d}\big(\omega_{i_1\ldots i_k}\eta_{j_1\ldots j_l}\,\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l}\big) \\
&= \mathrm{d}(\omega_{i_1\ldots i_k}\eta_{j_1\ldots j_l}) \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l} \\
&= \big[(\mathrm{d}\omega_{i_1\ldots i_k})\eta_{j_1\ldots j_l} + \omega_{i_1\ldots i_k}\,\mathrm{d}\eta_{j_1\ldots j_l}\big] \\
&\qquad \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l}
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\mathrm{d}\omega_{i_1\ldots i_k} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}\right) \wedge \left(\eta_{j_1\ldots j_l}\mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l}\right) \\
&\quad + (-1)^k \left(\omega_{i_1\ldots i_k}\mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}\right) \wedge \left(\mathrm{d}\eta_{j_1\ldots j_l} \wedge \mathrm{d}x^{j_1} \wedge \cdots \wedge \mathrm{d}x^{j_l}\right) \\
&= \mathrm{d}\omega \wedge \eta + (-1)^k \omega \wedge \mathrm{d}\eta.
\end{aligned}
$$

(Note that (3.33) is a particular case of (3.35).)

A k-form whose exterior derivative is zero is called *closed*; a k-form is *exact* if it is the exterior derivative of some $(k-1)$-form. Any exact differential form is closed since if $\omega = \mathrm{d}\eta$ with $\eta \in \Lambda^l(M)$, locally we have

$$
\omega = \mathrm{d}\eta_{i_1\ldots i_l} \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} = \left(\frac{\partial}{\partial x^j}\right)\eta_{i_1\ldots i_l}\, \mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l},
$$

with the functions $\eta_{i_1\ldots i_l}$ being the components of η; then

$$
\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}\left[\left(\frac{\partial}{\partial x^j}\right)\eta_{i_1\ldots i_l}\right] \wedge \mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l} \\
&= \left(\frac{\partial}{\partial x^m}\right)\left(\frac{\partial}{\partial x^j}\right)\eta_{i_1\ldots i_l}\, \mathrm{d}x^m \wedge \mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_l},
\end{aligned}
$$

which is equal to zero since $(\partial/\partial x^m)$ and $(\partial/\partial x^j)$ commute, whereas $\mathrm{d}x^m \wedge \mathrm{d}x^j = -\mathrm{d}x^j \wedge \mathrm{d}x^m$. In other words, the exterior differentiation has the property that

$$
\mathrm{d}^2 = 0 \tag{3.36}
$$

(see also Exercise 3.7). Note that there do not exist exact 0-forms and that any n-form is closed because any $(n+1)$-form is zero.

Example 3.10 According to the first and the second law of thermodynamics, for a given thermodynamical system there exist two real-valued functions, U (the internal energy) and S (the entropy) defined on the set of equilibrium states of the system, such that

$$
T\,\mathrm{d}S = \mathrm{d}U + P\,\mathrm{d}V.
$$

This is the case if the only way in which one can do mechanical work on the system is by compression, where T, P, and V are the absolute temperature, pressure, and volume, respectively. Using the properties of the exterior derivative we obtain $\mathrm{d}T \wedge \mathrm{d}S = \mathrm{d}P \wedge \mathrm{d}V$ and therefore, $\mathrm{d}P \wedge \mathrm{d}V \wedge \mathrm{d}T = 0$, which implies that P, V, and T cannot be functionally independent; that is, there must exist an "equation of state" expressing, e.g., P as some function of V and T. In a similar manner, combining the expressions above, one finds that any set formed by three of the functions T, S, U, P, and V is functionally dependent. (For instance, $\mathrm{d}U \wedge \mathrm{d}S \wedge \mathrm{d}V = (T\,\mathrm{d}S - P\,\mathrm{d}V) \wedge \mathrm{d}S \wedge \mathrm{d}V = 0$.) Therefore, the manifold of the equilibrium states is two-dimensional.

Exercise 3.11 Compute the exterior derivative of the 2-form

$$\omega = \frac{1}{(x^2+y^2+z^2)^{3/2}}(x\,\mathrm{d}y \wedge \mathrm{d}z + y\,\mathrm{d}z \wedge \mathrm{d}x + z\,\mathrm{d}x \wedge \mathrm{d}y) \tag{3.37}$$

defined on $\mathbb{R}^3 \setminus \{(0,0,0)\}$, where (x, y, z) are the natural coordinates of $\mathbb{R}^3$. Find the *local* expression of this form in terms of the spherical coordinates (r, θ, ϕ) (with $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$).

Exercise 3.12 Consider the 1-forms $\omega^1, \omega^2, \omega^3$ defined by

$$\begin{aligned}
\omega^1 &= h\,\mathrm{d}x^1 - x^1\,\mathrm{d}h - x^2\,\mathrm{d}x^3 + x^3\,\mathrm{d}x^2,\\
\omega^2 &= h\,\mathrm{d}x^2 - x^2\,\mathrm{d}h - x^3\,\mathrm{d}x^1 + x^1\,\mathrm{d}x^3,\\
\omega^3 &= h\,\mathrm{d}x^3 - x^3\,\mathrm{d}h - x^1\,\mathrm{d}x^2 + x^2\,\mathrm{d}x^1,
\end{aligned}$$

in terms of a local coordinate system (x^1, x^2, x^3), with $h \equiv \sqrt{1 - \sum_{i=1}^3 (x^i)^2}$. Show that

$$\mathrm{d}\omega^1 = -2\omega^2 \wedge \omega^3, \qquad \mathrm{d}\omega^2 = -2\omega^3 \wedge \omega^1, \qquad \mathrm{d}\omega^3 = -2\omega^1 \wedge \omega^2.$$

(The forms ω^i arise in connection with the group SU(2); see Sect. 7.3.)

If M and N are two differentiable manifolds and $\psi : M \to N$ is a differentiable map, then we have

$$\psi^*(\mathrm{d}\omega) = \mathrm{d}(\psi^*\omega), \quad \text{for } \omega \in \Lambda^k(N). \tag{3.38}$$

In effect, expressing ω as $\omega = \omega_{i_1\ldots i_k}\,\mathrm{d}y^{i_1} \wedge \cdots \wedge \mathrm{d}y^{i_k}$, where $(y^1, \ldots, y^m)$ is a coordinate system on N, we have [see (3.25)]

$$\begin{aligned}
\psi^*(\mathrm{d}\omega) &= \psi^*\big(\mathrm{d}\omega_{i_1\ldots i_k} \wedge \mathrm{d}y^{i_1} \wedge \cdots \wedge \mathrm{d}y^{i_k}\big)\\
&= \psi^*(\mathrm{d}\omega_{i_1\ldots i_k}) \wedge \psi^*\big(\mathrm{d}y^{i_1}\big) \wedge \cdots \wedge \psi^*\big(\mathrm{d}y^{i_k}\big),
\end{aligned}$$

but $\psi^*(\mathrm{d}f) = \mathrm{d}(\psi^* f)$ for $f \in \Lambda^0(N)$ [see (2.29)]. Using the fact that d is an antiderivation and that $\mathrm{d}^2 = 0$ on functions [see Exercise 3.7 or (3.36)], it follows that

$$\begin{aligned}
\psi^*(\mathrm{d}\omega) &= \mathrm{d}(\psi^*\omega_{i_1\ldots i_k}) \wedge \mathrm{d}\big(\psi^* y^{i_1}\big) \wedge \cdots \wedge \mathrm{d}\big(\psi^* y^{i_k}\big)\\
&= \mathrm{d}\big[(\psi^*\omega_{i_1\ldots i_k})\,\mathrm{d}\big(\psi^* y^{i_1}\big) \wedge \cdots \wedge \mathrm{d}\big(\psi^* y^{i_k}\big)\big]\\
&= \mathrm{d}\big[(\psi^*\omega_{i_1\ldots i_k})\psi^*\big(\mathrm{d}y^{i_1}\big) \wedge \cdots \wedge \psi^*\big(\mathrm{d}y^{i_k}\big)\big]\\
&= \mathrm{d}\big[\psi^*\big(\omega_{i_1\ldots i_k}\,\mathrm{d}y^{i_1} \wedge \cdots \wedge \mathrm{d}y^{i_k}\big)\big]\\
&= \mathrm{d}(\psi^*\omega).
\end{aligned}$$

Exercise 3.13 Show that $\pounds_{\mathbf{X}}(\mathrm{d}\omega) = \mathrm{d}(\pounds_{\mathbf{X}}\omega)$, for $\omega \in \Lambda^k(M)$ and $\mathbf{X} \in \mathfrak{X}(M)$.

The Lie derivative is related with the exterior derivative in the following manner. (Cf. Exercise 3.8.)

Proposition 3.14 *For* $\mathbf{X} \in \mathfrak{X}(M)$ *and* $\omega \in \Lambda^k(M)$ *we have*

$$\pounds_{\mathbf{X}}\omega = \mathbf{X} \lrcorner \mathrm{d}\omega + \mathrm{d}(\mathbf{X} \lrcorner \omega). \tag{3.39}$$

Proof Making use of the definitions of the contraction and of the exterior derivative, for $\mathbf{X}, \mathbf{Y}_1, \ldots, \mathbf{Y}_k \in \mathfrak{X}(M)$ we obtain

$$\begin{aligned}
(\mathbf{X} \lrcorner \mathrm{d}\omega)(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) &= (k+1)\,\mathrm{d}\omega(\mathbf{X}, \mathbf{Y}_1, \ldots, \mathbf{Y}_k) \\
&= \mathbf{X}\big(\omega(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)\big) \\
&\quad + \sum_{i=1}^{k} (-1)^i \mathbf{Y}_i \big(\omega(\mathbf{X}, \mathbf{Y}_1, \ldots, \widehat{\mathbf{Y}_i}, \ldots, \mathbf{Y}_k)\big) \\
&\quad + \sum_{j=1}^{k} (-1)^j \omega\big([\mathbf{X}, \mathbf{Y}_j], \mathbf{Y}_1, \ldots, \widehat{\mathbf{Y}_j}, \ldots, \mathbf{Y}_k\big) \\
&\quad + \sum_{i<j} (-1)^{i+j} \omega\big([\mathbf{Y}_i, \mathbf{Y}_j], \mathbf{X}, \ldots, \widehat{\mathbf{Y}_i}, \ldots, \widehat{\mathbf{Y}_j}, \ldots, \mathbf{Y}_k\big);
\end{aligned}$$

on the other hand,

$$\begin{aligned}
&\big(\mathrm{d}(\mathbf{X} \lrcorner \omega)\big)(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) \\
&= \frac{1}{k}\left[\sum_{i=1}^{k} (-1)^{i+1} \mathbf{Y}_i \big((\mathbf{X} \lrcorner \omega)(\mathbf{Y}_1, \ldots, \widehat{\mathbf{Y}_i}, \ldots, \mathbf{Y}_k)\big)\right. \\
&\quad \left. + \sum_{i<j} (-1)^{i+j} (\mathbf{X} \lrcorner \omega)\big([\mathbf{Y}_i, \mathbf{Y}_j], \mathbf{Y}_1, \ldots, \widehat{\mathbf{Y}_i}, \ldots, \widehat{\mathbf{Y}_j}, \ldots, \mathbf{Y}_k\big)\right] \\
&= \sum_{i=1}^{k} (-1)^{i+1} \mathbf{Y}_i \big(\omega(\mathbf{X}, \mathbf{Y}_1, \ldots, \widehat{\mathbf{Y}_i}, \ldots, \mathbf{Y}_k)\big) \\
&\quad + \sum_{i<j} (-1)^{i+j} \omega\big(\mathbf{X}, [\mathbf{Y}_i, \mathbf{Y}_j], \mathbf{Y}_1, \ldots, \widehat{\mathbf{Y}_i}, \ldots, \widehat{\mathbf{Y}_j}, \ldots, \mathbf{Y}_k\big).
\end{aligned}$$

By adding these two relations and using the expression for $\pounds_{\mathbf{X}}\omega$ given at the end of Chap. 2 [equation (2.45)] we obtain the proposed relation. (Another proof is given in Sect. 4.1.) □

Exercise 3.15 Compute the Lie derivatives of the 2-form (3.37) with respect to the vector fields

$$y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}, \qquad z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}, \qquad x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}.$$

Poincaré's Lemma Given a closed k-form, ω, there does not always exist a $(k-1)$-form η such that $\omega = \mathrm{d}\eta$. That is, not every closed form is exact. However, as shown below, such an η always exists locally, and this result is known as the Poincaré Lemma. The global existence of η (that is, on all of M) depends on the properties of M (see also do Carmo 1994).

A well-known, illustrative example is given by the 1-form

$$\alpha = \frac{x\,\mathrm{d}y - y\,\mathrm{d}x}{x^2 + y^2} \tag{3.40}$$

on $M = \mathbb{R}^2 \setminus \{(0,0)\}$ (which is the analog of the 2-form (3.37), considered in Exercise 3.11). One readily verifies that α is closed:

$$\mathrm{d}\alpha = \left[\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) + \frac{\partial}{\partial y}\left(\frac{y}{x^2+y^2}\right)\right]\mathrm{d}x \wedge \mathrm{d}y = 0,$$

but there does not exist a function defined on all of M whose differential coincides with α (see Example 1.28).

However, on the simply connected set $\mathbb{R}^2 \setminus \{(x,0)\,|\,x \geq 0\}$ (the plane with the positive x axis removed), $\alpha = \mathrm{d}\theta$, where θ is the standard coordinate function used in the polar coordinates, with its values restricted to the interval $(0, 2\pi)$. (Substituting $x = r\cos\theta$, $y = r\sin\theta$ into the expression for α one finds $\alpha = r^{-2}[r\cos\theta(r\cos\theta\,\mathrm{d}\theta + \sin\theta\,\mathrm{d}r) - r\sin\theta(-r\sin\theta\,\mathrm{d}\theta + \cos\theta\,\mathrm{d}r)] = \mathrm{d}\theta$.) On M, the angle θ is not a well-defined (single-valued) differentiable function.

In order to prove that a closed k-form, with $k \geq 1$, is locally exact, we consider a one-parameter group of diffeomorphisms φ_t on M; then, for any k-form ω we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}{\varphi_t}^*\omega &= \lim_{h\to 0}\frac{{\varphi_{t+h}}^*\omega - {\varphi_t}^*\omega}{h} \\
&= {\varphi_t}^*\lim_{h\to 0}\frac{{\varphi_h}^*\omega - \omega}{h} \\
&= {\varphi_t}^*\pounds_{\mathbf{X}}\omega,
\end{aligned}$$

where $\mathbf{X}$ is the infinitesimal generator of φ_t. Making use of the relation $\pounds_{\mathbf{X}}\omega = \mathbf{X}\lrcorner\,\mathrm{d}\omega + \mathrm{d}(\mathbf{X}\lrcorner\,\omega)$, if ω is closed

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}{\varphi_t}^*\omega &= {\varphi_t}^*\big[\mathbf{X}\lrcorner\,\mathrm{d}\omega + \mathrm{d}(\mathbf{X}\lrcorner\,\omega)\big] \\
&= {\varphi_t}^*\mathrm{d}(\mathbf{X}\lrcorner\,\omega) \\
&= \mathrm{d}\big[{\varphi_t}^*(\mathbf{X}\lrcorner\,\omega)\big]
\end{aligned}$$

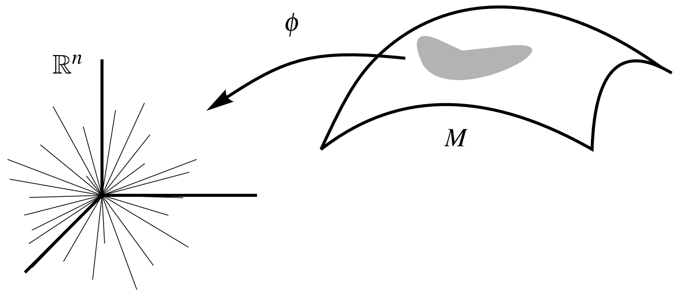

Fig. 3.1 The images in $\mathbb{R}^n$ of the orbits of the group (3.42) under ϕ are radial segments

and integrating on the parameter t, from t_0 to 0, we find

$$\mathrm{d}\int_{t_0}^{0} {\varphi_t}^*(\mathbf{X}\lrcorner\,\omega)\,\mathrm{d}t = \int_{t_0}^{0}\left(\frac{\mathrm{d}}{\mathrm{d}t}{\varphi_t}^*\omega\right)\mathrm{d}t = \omega - {\varphi_{t_0}}^*\omega. \tag{3.41}$$

(Note that in this last integral, t is only an integration variable; the product of the differential form appearing in the integrand by $\mathrm{d}t$ is not an exterior product. See the examples given below.) With the aid of the group φ_t given by

$${\varphi_t}^* x^i = \mathrm{e}^t x^i, \tag{3.42}$$

where $(x^1, \ldots, x^n)$ is some local coordinate system on M (see Fig. 3.1) [then $\mathbf{X} = x^i(\partial/\partial x^i)$], we have ${\varphi_t}^*\,\mathrm{d}x^i = \mathrm{e}^t\,\mathrm{d}x^i$; therefore, if

$$\lim_{t_0\to-\infty} {\varphi_{t_0}}^*\omega = 0, \tag{3.43}$$

then from (3.41) it follows that

$$\omega = \mathrm{d}\int_{-\infty}^{0} {\varphi_t}^*(\mathbf{X}\lrcorner\,\omega)\,\mathrm{d}t, \tag{3.44}$$

assuming that the integrand is well behaved for $t \in (-\infty, 0]$ (see the discussion below). Thus we express ω as the exterior derivative of a $(k-1)$-form.

For instance, one can verify that the 2-form $\omega = 5x\,\mathrm{d}y\wedge\mathrm{d}z - 3y\,\mathrm{d}z\wedge\mathrm{d}x - (x^2+2z)\,\mathrm{d}x\wedge\mathrm{d}y$ is closed. In order to apply (3.44) we start by computing the contraction $\mathbf{X}\lrcorner\,\omega$

$$\begin{aligned}\mathbf{X}\lrcorner\,\omega &= \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right)\lrcorner\left[5x\,\mathrm{d}y\wedge\mathrm{d}z - 3y\,\mathrm{d}z\wedge\mathrm{d}x - \left(x^2+2z\right)\mathrm{d}x\wedge\mathrm{d}y\right]\\ &= 5x(y\,\mathrm{d}z - z\,\mathrm{d}y) - 3y(z\,\mathrm{d}x - x\,\mathrm{d}z) - \left(x^2+2z\right)(x\,\mathrm{d}y - y\,\mathrm{d}x)\\ &= \left(x^2y - yz\right)\mathrm{d}x - \left(x^3 + 7xz\right)\mathrm{d}y + 8xy\,\mathrm{d}z.\end{aligned}$$

Then,

$$\int_{-\infty}^{0} \varphi_t{}^*(\mathbf{X}\lrcorner\omega)\,\mathrm{d}t = \int_{-\infty}^{0} \mathrm{d}t\,\Big[\big(\mathrm{e}^{4t}x^2y - \mathrm{e}^{3t}yz\big)\,\mathrm{d}x - \big(\mathrm{e}^{4t}x^3 + \mathrm{e}^{3t}7xz\big)\,\mathrm{d}y + \mathrm{e}^{3t}8xy\,\mathrm{d}z\Big]$$

[note that the integrand satisfies the condition (3.43)] and, with the change of variable $s = \mathrm{e}^t$, we have

$$\int_{-\infty}^{0} \varphi_t{}^*(\mathbf{X}\lrcorner\omega)\,\mathrm{d}t = \int_{0}^{1} \mathrm{d}s\,\Big[\big(s^3x^2y - s^2yz\big)\,\mathrm{d}x - \big(s^3x^3 + s^27xz\big)\,\mathrm{d}y + s^28xy\,\mathrm{d}z\Big]$$
$$= \left(\frac{1}{4}x^2y - \frac{1}{3}yz\right)\mathrm{d}x - \left(\frac{1}{4}x^3 + \frac{1}{3}7xz\right)\mathrm{d}y + \frac{1}{3}8xy\,\mathrm{d}z.$$

A direct computation shows that, indeed, the exterior derivative of this 1-form coincides with the 2-form ω originally given.

Note that the 2-form ω, defined in (3.37), and the 1-form α, defined in (3.40), are both closed and satisfy $\mathbf{X}\lrcorner\,\omega = 0$, $\mathbf{X}\lrcorner\,\alpha = 0$, but they do not satisfy the condition (3.43) (in fact, both forms are invariant under the group (3.42), $\varphi_t{}^*\omega = \omega$ and $\varphi_t{}^*\alpha = \alpha$). Therefore (3.44) cannot be applied to them.

However, by simply making use of another coordinate system, we can locally express these forms as exterior derivatives of some appropriate forms. For instance, the 1-form α defined in (3.40) can also be expressed as

$$\alpha = \frac{(x'+1)\,\mathrm{d}y' - y'\,\mathrm{d}x'}{(x'+1)^2 + y'^2},$$

in terms of the coordinate system (x', y') related to (x, y) by $x' = x - 1$, $y' = y$. Dropping the primes we have

$$\mathbf{X}\lrcorner\alpha = \frac{y}{(x+1)^2 + y^2}, \qquad \varphi_t{}^*(\mathbf{X}\lrcorner\alpha) = \frac{\mathrm{e}^t y}{(\mathrm{e}^t x + 1)^2 + (\mathrm{e}^t y)^2}.$$

Now, condition (3.43) is satisfied and, putting $s = \mathrm{e}^t$, we find

$$\int_{-\infty}^{0} \varphi_t{}^*(\mathbf{X}\lrcorner\alpha)\,\mathrm{d}t = \int_{0}^{1} \mathrm{d}s\,\frac{y}{(sx+1)^2 + (sy)^2} = \arctan\frac{(x^2+y^2)s + x}{y}\bigg|_{s=0}^{1}$$
$$= \arctan\frac{y}{x+1},$$

provided that $y \neq 0$, or that $x > -1$ if $y = 0$, so that the integrand does not become singular for $s \in [0, 1]$ (in other words, $(x, y) \in \mathbb{R}^2 \setminus \{(x, y)\,|\,x \leq -1, y = 0\}$).

Example 3.16 Using the properties (3.32), (3.35), and (3.36), the condition (3.18) defining a canonical transformation can be expressed in the form

$$\mathrm{d}(p\,\mathrm{d}q - H\,\mathrm{d}t - P\,\mathrm{d}Q + K\,\mathrm{d}t) = 0,$$

which is locally equivalent to the existence of a function F such that

$$p\,\mathrm{d}q - P\,\mathrm{d}Q + (K - H)\,\mathrm{d}t = \mathrm{d}F.$$

If q and Q are functionally independent, then (q, Q, t) can be used as local coordinates on $P \times \mathbb{R}$ and from the last equation we have

$$p = \frac{\partial F}{\partial q}, \qquad P = -\frac{\partial F}{\partial Q}, \qquad K = H + \frac{\partial F}{\partial t}.$$

Assuming that $\partial^2 F/\partial Q \partial q \neq 0$, these expressions allow us to express P and Q in terms of p, q, and t. For this reason, F is a *generating function* of the canonical transformation.

Example 3.17 Using the notation and results of Example 3.3, if $\mathbf{Y}$ is the infinitesimal generator of a one-parameter group of diffeomorphisms on $P \times \mathbb{R}$ which maps any solution of the Hamilton equations (3.17) into another solution, then there exists some function $\nu \in C^\infty(P \times \mathbb{R})$ such that $£_{\mathbf{Y}}(\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) = \nu(\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t)$ (cf. Example 2.28). In particular, if $£_{\mathbf{Y}}(\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) = 0$, making use of the identity (3.39) and the fact that the form $\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t$ is closed, we have

$$\mathrm{d}\big[\mathbf{Y} \lrcorner (\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t)\big] = 0,$$

which is equivalent to the local existence of a function, χ, such that

$$\mathbf{Y} \lrcorner (\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) = \mathrm{d}\chi. \tag{3.45}$$

The function χ is a *constant of motion*, that is, its value is constant along the curves in $P \times \mathbb{R}$ which are a solution of the Hamilton equations (3.17); this amounts to $\mathbf{X}\chi = 0$, where $\mathbf{X}$ is the vector field (3.16). In fact, making use of (3.13) and (3.45) we have

$$\begin{aligned}
\mathbf{X}\chi &= \mathbf{X} \lrcorner \mathrm{d}\chi \\
&= \mathbf{X} \lrcorner \mathbf{Y} \lrcorner (\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) \\
&= -\mathbf{Y} \lrcorner \mathbf{X} \lrcorner (\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) \\
&= 0.
\end{aligned}$$

Exercise 3.18 Using the definitions given in Example 3.17, show that the function ν appearing in $£_{\mathbf{Y}}(\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t) = \nu(\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t)$ is a constant of motion. (See also Sect. 8.7.)

Chapter 4
Integral Manifolds

We have met the concept of integral curve of a vector field in Sect. 2.1 and we have seen that finding such curves is equivalent to solving a system of ODEs. In this chapter we consider a generalization of this relationship defining the integral manifolds of a set of vector fields or of differential forms. We shall show that the problem of finding these manifolds is equivalent to that of solving certain systems of differential equations.

4.1 The Rectification Lemma

As shown in Sect. 1.3, any vector field, $\mathbf{X}$, on a differentiable manifold M can be expressed locally in the form $\mathbf{X} = X^i(\partial/\partial x^i)$, where $(x^1, x^2, \ldots, x^n)$ is a coordinate system on M [see (1.36)]. As we shall see now, for each point of M where $\mathbf{X}$ does not vanish, there exists a local coordinate system, $(x'^1, x'^2, \ldots, x'^n)$, such that

$$\mathbf{X} = \frac{\partial}{\partial x'^1}. \tag{4.1}$$

This result, known as the *rectification*, or *straightening-out, lemma*, ensures that, in some neighborhood of each point where $\mathbf{X}$ is different from zero, there exists a coordinate system adapted to $\mathbf{X}$, in which $\mathbf{X}$ has a very simple form.

Assuming that there exists a coordinate system $(x'^1, x'^2, \ldots, x'^n)$ such that $\mathbf{X} = \partial/\partial x'^1$, the functions $x'^2, \ldots, x'^n$ must satisfy $\mathbf{X}x'^2 = 0$, $\mathbf{X}x'^3 = 0, \ldots,$ $\mathbf{X}x'^n = 0$; that is, the coordinates $x'^2, \ldots, x'^n$ must be $n-1$ functionally independent solutions of the linear partial differential equation (PDE)

$$X^1\frac{\partial f}{\partial x^1} + X^2\frac{\partial f}{\partial x^2} + \cdots + X^n\frac{\partial f}{\partial x^n} = 0. \tag{4.2}$$

Given $n-1$ functionally independent solutions of (4.2), the coordinates $x'^2, \ldots, x'^n$ can be chosen as any set of $n-1$ functionally independent functions of them. In

G.F. Torres del Castillo, *Differentiable Manifolds*,

DOI 10.1007/978-0-8176-8271-2_4,

contrast, x'^1 must satisfy the (inhomogeneous) equation $\mathbf{X}x'^1 = 1$, that is,

$$X^1 \frac{\partial x'^1}{\partial x^1} + X^2 \frac{\partial x'^1}{\partial x^2} + \cdots + X^n \frac{\partial x'^1}{\partial x^n} = 1. \tag{4.3}$$

Given a solution of this equation, one can obtain another solution by adding to it any solution of (4.2).

The problem of finding the solutions of (4.2) is related to that of finding the integral curves of $\mathbf{X}$ (see Sect. 2.1), because if $\mathbf{X}$ has the form (4.1), then its integral curves are given by $(x'^1 \circ C)(t) = t + \text{const}$, $x'^2 \circ C = \text{const}$, $x'^3 \circ C = \text{const}, \dots, x'^n \circ C = \text{const}$. Therefore, if we have the integral curve of $\mathbf{X} = X^i(\partial/\partial x^i)$ that starts at an arbitrary point of some neighborhood, in the original coordinate system $(x^1, x^2, \dots, x^n)$, then we have n functions $x^i \circ C$ satisfying the system (2.5), which must contain n arbitrary constants (which determine the starting point of the curve C). Combining these n expressions to eliminate the parameter of the curve [the variable t in equations (2.5)], one obtains $n-1$ equations that are equivalent to the $n-1$ equations $x'^2 \circ C = \text{const}$, $x'^3 \circ C = \text{const}, \dots, x'^n \circ C = \text{const}$.

The coordinate x'^1 (which is defined up to an additive function of x'^2, x'^3, $\dots, x'^n$) can be found noting that the contraction of $\mathbf{X}$ with any of the 1-forms $\mathrm{d}x^1/X^1$, $\mathrm{d}x^2/X^2, \dots,$ and $\mathrm{d}x^n/X^n$, among many others, is equal to 1. Since any 1-form on a manifold of dimension one is locally exact and since the integral curves of $\mathbf{X}$ are manifolds of dimension one, on these curves each of the 1-forms $\mathrm{d}x^1/X^1$, $\mathrm{d}x^2/X^2, \dots, \mathrm{d}x^n/X^n$, is, locally, the differential of a function that can be chosen as x'^1.

Example 4.1 Let us consider the vector field $\mathbf{X}$ given locally by

$$\mathbf{X} = -x^2 \frac{\partial}{\partial x} - xy \frac{\partial}{\partial y} + (xz - y) \frac{\partial}{\partial z}, \tag{4.4}$$

where (x, y, z) is a coordinate system on some manifold of dimension three. Its integral curves are determined by the system of ODEs

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -x^2, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = -xy, \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = xz - y, \tag{4.5}$$

where, as in the previous examples, we have written x, y, and z in place of $x \circ C$, $y \circ C$, and $z \circ C$, respectively. The first of equations (4.5) is readily integrated, giving

$$\frac{1}{x(t)} = t + \frac{1}{x_0},$$

where x_0 is the value of the coordinate x at the starting point of the curve, or, equivalently,

$$x(t) = \frac{x_0}{x_0 t + 1}. \tag{4.6}$$

Substituting this expression into the second equation (4.5) we find that

$$y(t) = \frac{y_0}{x_0 t + 1}, \tag{4.7}$$

where y_0 is the value of the coordinate y at the starting point of the curve. Substituting now (4.6) and (4.7) into the third equation (4.5) we obtain a linear equation whose solution is

$$z(t) = z_0 + (x_0 z_0 - y_0)t, \tag{4.8}$$

where z_0 is the value of the coordinate z at the initial point of the curve. Eliminating the parameter t from (4.6)–(4.8) one finds that

$$\frac{x(t)}{y(t)} = \frac{x_0}{y_0}, \qquad x(t)z(t) - y(t) = x_0 z_0 - y_0,$$

which means that the (images of the) integral curves of $\mathbf{X}$ are the intersections of the surfaces $x/y = \text{const}$, $xz - y = \text{const}$, and that the coordinates x'^2 and x'^3 can be chosen as x/y and $xz - y$. (One can readily verify that $\mathbf{X}(x/y) = 0$ and $\mathbf{X}(xz - y) = 0$.)

According to the discussion above, on the curves $x/y = \text{const}$, $xz - y = \text{const}$, the 1-forms $-\mathrm{d}x/x^2$, $-\mathrm{d}y/xy$, and $\mathrm{d}z/(xz - y)$ (whose contractions with $\mathbf{X}$ are equal to 1) are the differentials of possible choices for x'^1. In fact, $-\mathrm{d}x/x^2 = \mathrm{d}(1/x)$, so that we can choose $x'^1 = 1/x$. Alternatively, by imposing the conditions $x/y = \text{const}$, $xz - y = \text{const}$, we have, for instance,

$$\frac{\mathrm{d}z}{xz - y} = \mathrm{d}\left(\frac{z}{xz - y}\right),$$

which gives another acceptable choice for x'^1, namely, $x'^1 = z/(xz - y)$. The functions $1/x$ and $z/(xz - y)$ differ by $-(x/y)^{-1}(xz - y)^{-1}$, which is effectively a function of x/y and $xz - y$ only, as stated above.

In order to find the images of the integral curves of a vector field $\mathbf{X}$ (and to identify a set of coordinates $x'^2, \ldots, x'^n$) it is not necessary to integrate equations (2.5), with the subsequent elimination of the parameter t; the parameter can be eliminated from the beginning (see, e.g., Example 2.9), which leads to a set of equations that is usually expressed in the form

$$\frac{\mathrm{d}x^1}{X^1} = \frac{\mathrm{d}x^2}{X^2} = \cdots = \frac{\mathrm{d}x^n}{X^n} \tag{4.9}$$

[see, e.g., Sneddon (2006, Chap. 2)].

The fact that a vector field can be expressed in the form (4.1) has several applications. For instance, it allows us to simplify the demonstration of some propositions involving vector fields, as can be seen in connection with Proposition 3.14. If ω is a k-form and $\mathbf{X}$ is a vector field, on some neighborhood of each point where $\mathbf{X}$ does not vanish, one can find a coordinate system $(x^1, x^2, \ldots, x^n)$ such that $\mathbf{X} = \partial/\partial x^1$, using then the expression $\omega = \omega_{i_1 \ldots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ and the properties of the Lie derivative of differential forms, we have $£_{\mathbf{X}}\, \mathrm{d}x^i = \mathrm{d}(\mathbf{X}x^i) = 0$ [see (2.39) and (2.23)], hence, according to (3.26)

$$£_{\mathbf{X}}\omega = \frac{\partial \omega_{i_1 \ldots i_k}}{\partial x^1}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}.$$

On the other hand, using (3.34) and (3.27)

$$\begin{aligned}
\mathbf{X} \lrcorner \mathrm{d}\omega &= \frac{\partial}{\partial x^1} \lrcorner \left(\frac{\partial \omega_{i_1 \ldots i_k}}{\partial x^j}\, \mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \right) \\
&= \frac{\partial \omega_{i_1 \ldots i_k}}{\partial x^j} \frac{\partial}{\partial x^1} \lrcorner \left(\mathrm{d}x^j \wedge \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \right) \\
&= \frac{\partial \omega_{i_1 \ldots i_k}}{\partial x^1}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} - k \frac{\partial \omega_{1 i_2 \ldots i_k}}{\partial x^j}\, \mathrm{d}x^j \wedge \mathrm{d}x^{i_2} \wedge \cdots \wedge \mathrm{d}x^{i_k}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{d}(\mathbf{X} \lrcorner \omega) &= \mathrm{d}\left(k\, \omega_{1 i_2 \ldots i_k} \mathrm{d}x^{i_2} \wedge \cdots \wedge \mathrm{d}x^{i_k} \right) \\
&= k \frac{\partial \omega_{1 i_2 \ldots i_k}}{\partial x^j}\, \mathrm{d}x^j \wedge \mathrm{d}x^{i_2} \wedge \cdots \wedge \mathrm{d}x^{i_k}.
\end{aligned}$$

Thus, $£_{\mathbf{X}}\omega = \mathbf{X} \lrcorner \mathrm{d}\omega + \mathrm{d}(\mathbf{X} \lrcorner \omega)$.

The form (4.1) is also useful in the solution of ODEs. In many cases it is possible to find explicitly a vector field, $\mathbf{X}$, that generates some symmetry of a given ODE, which means that the image of any solution of the equation under the flow generated by $\mathbf{X}$ is a solution of the same equation [see, e.g., Stephani (1989), Hydon (2000), and Sect. 4.3]. The form of the ODE is simplified making use of a coordinate system in which $\mathbf{X}$ has the expression (4.1).

It should be clear that the expression (4.1) is not valid at the points where $\mathbf{X}$ vanishes. Whereas the rectification lemma establishes that all vector fields look the same wherever they do not vanish, there exist several different behaviors for a vector field in a neighborhood of a point where it vanishes [see, e.g., Guillemin and Pollack (1974, Chap. 3)].

If $\mathbf{Y}$ is a second vector field, it is not always possible to find a coordinate system $(x'^1, x'^2, \ldots, x'^n)$, such that $\mathbf{Y} = \partial/\partial x'^2$, simultaneously with $\mathbf{X} = \partial/\partial x'^1$. A necessary condition for this to happen is that $[\mathbf{X}, \mathbf{Y}] = 0$, since $[\partial/\partial x'^1, \partial/\partial x'^2] = 0$ [see (1.38)]. This condition is also sufficient; in general, if $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$ satisfy $[\mathbf{X}_i, \mathbf{X}_j] = 0$, for $1 \leq i, j \leq k$, then, in a neighborhood of each point where $\{\mathbf{X}_1, \ldots, \mathbf{X}_k\}$ is linearly independent, there exists a coordinate system $x'^1, \ldots, x'^n$ such that $\mathbf{X}_1 = \partial/\partial x'^1, \ldots, \mathbf{X}_k = \partial/\partial x'^k$. The proof is similar to that given in the case with $k = 1$.

4.2 Distributions and the Frobenius Theorem

As shown in Example 2.27, the solutions of the first-order ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{P(x, y)}{Q(x, y)}$$

are the images of the curves C such that $C^*(P\,\mathrm{d}x + Q\,\mathrm{d}y) = 0$. A similar formulation can be given in the case of an ODE of order m, making use of a set of m 1-forms on a manifold of dimension $m+1$. For instance, given a second-order ODE of the form

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = F(x, y, \mathrm{d}y/\mathrm{d}x), \tag{4.10}$$

where F is a differentiable real-valued function of three variables, we introduce an auxiliary variable z and define the two 1-forms

$$\alpha^1 \equiv \mathrm{d}y - z\,\mathrm{d}x \quad \text{and} \quad \alpha^2 \equiv \mathrm{d}z - F(x, y, z)\,\mathrm{d}x. \tag{4.11}$$

Then, considering (x, y, z) as local coordinates of some manifold M, the solutions of (4.10) are given by the images of the curves C in M such that $C^*\alpha^1 = 0$, $C^*\alpha^2 = 0$ (see Sect. 4.3, below).

The 1-forms are also employed in classical mechanics to express *constraints*. When a mechanical system is subject to a constraint represented by a 1-form α, the possible curves in the configuration space must satisfy the condition $C^*\alpha = 0$, and a mechanical systems may have more than one of such constraints.

For instance, for a block sliding under the influence of gravity on a wedge of angle θ, which lies on a horizontal table, there are two constraints, given by

$$\begin{aligned} \alpha^1 &\equiv \mathrm{d}y - \tan\theta(\mathrm{d}x - \mathrm{d}\tilde{x}), \\ \alpha^2 &\equiv \mathrm{d}\tilde{y}, \end{aligned} \tag{4.12}$$

where θ is the angle of the wedge, (x, y) and $(\tilde{x}, \tilde{y})$ are Cartesian coordinates of the block and the wedge, respectively (see Fig. 4.1). The condition $C^*\alpha^2 = 0$, that is, $C^*\mathrm{d}\tilde{y} = 0$, means that $\tilde{y}$ has to remain constant along the admissible curves in the configuration space.

Another well-known example of a mechanical system with constraints corresponds to a vertical disk, of radius a, say, that rolls without slipping on a horizontal plane (see Fig. 4.2). The constraints can be expressed by means of the 1-forms $\alpha^1 = \mathrm{d}x - a\cos\theta\,\mathrm{d}\phi$ and $\alpha^2 = \mathrm{d}y - a\sin\theta\,\mathrm{d}\phi$, where (x, y) are Cartesian coordinates of the contact point of the disk with the plane, θ is the angle between the x axis and the plane of the disk, and ϕ is the angle between a given radius of the disk and the line joining the center of the disk with the point of contact with the plane.

A final example is given by a sphere of radius a that rolls without slipping on a plane surface; there are two constraints represented by the 1-forms

$$\begin{aligned} \beta^1 &\equiv \mathrm{d}x + a(\sin\theta\cos\phi\,\mathrm{d}\psi - \sin\phi\,\mathrm{d}\theta), \\ \beta^2 &\equiv \mathrm{d}y + a(\sin\theta\sin\phi\,\mathrm{d}\psi + \cos\phi\,\mathrm{d}\theta), \end{aligned} \tag{4.13}$$

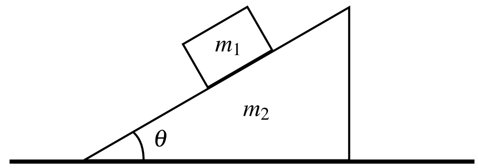

Fig. 4.1 The block remains in contact with the wedge, which lies on a horizontal table

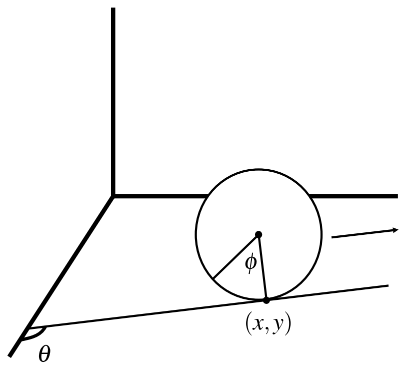

Fig. 4.2 The disk rolls without slipping on a horizontal plane. θ is the *angle* between the velocity of the disk and the positive x axis

where (ϕ, θ, ψ) are Euler angles (see, e.g., Sect. 8.6) and (x, y) are Cartesian coordinates of the point of contact between the sphere and the plane.

An important difference between the sets of 1-forms (4.12) and (4.13) is that in the first case the curves C in the configuration space satisfying $C^*\alpha^1 = 0 = C^*\alpha^2$ lie on a submanifold defined by $y - \tan\theta(x - \tilde{x}) = \text{const}$, $\tilde{y} = \text{const}$ (which is related to the facts that $\alpha^1 = \mathrm{d}[y - \tan\theta(x - \tilde{x})]$ and $\alpha^2 = \mathrm{d}\tilde{y}$), whereas, as we shall be able to show below, in the case of the 1-forms (4.13) the curves C satisfying the conditions $C^*\beta^1 = 0 = C^*\beta^2$ *are not* contained in submanifolds of the form $y^1 = \text{const}$, $y^2 = \text{const}$ (see Exercise 4.8, below). Owing to this difference, the 1-forms (4.12) constitute *holonomic constraints* and the 1-forms (4.13) represent *non-holonomic constraints* (a precise definition is given below).

Now we shall introduce some definitions. Let $\alpha^1, \ldots, \alpha^k$ be 1-forms on M and let $p \in M$ such that $\{\alpha^1_p, \ldots, \alpha^k_p\}$ is linearly independent. Then, the set of vectors $v_p \in T_pM$ such that $\alpha^1_p(v_p) = \alpha^2_p(v_p) = \cdots = \alpha^k_p(v_p) = 0$ forms a vector subspace of T_pM of dimension $(n - k)$, with n being the dimension of M. A *distribution* of dimension l on M is a map, $\mathscr{D}$, that assigns to each point $p \in M$ a vector subspace, $\mathscr{D}_p$, of T_pM of dimension l. Thus, a set of k independent 1-forms $\{\alpha^1, \ldots, \alpha^k\}$ defines a distribution of dimension $(n - k)$ given by $\mathscr{D}_p = \{v_p \in T_pM \mid \alpha^i_p(v_p) = 0,\ i = 1, \ldots, k\}$. The sets of 1-forms $\{\alpha^1, \ldots, \alpha^k\}$ and $\{\beta^1, \ldots, \beta^k\}$ define the same distribution if and only if there exist k^2 real-valued functions b^i_j, $i, j = 1, 2, \ldots, k$, such that $\beta^i = b^i_j\alpha^j$, with $\det(b^i_j)$ nowhere zero.

An *integral manifold* of the distribution $\mathscr{D}$ is a submanifold N of M such that the tangent space to N at p is contained in $\mathscr{D}_p$. If the distribution $\mathscr{D}$ is defined by the k

1-forms $\alpha^1, \ldots, \alpha^k$, as explained above, the submanifold N is an integral manifold of $\mathscr{D}$ if and only if $i^*\alpha^1 = 0, \ldots, i^*\alpha^k = 0$, where $i : N \to M$ is the inclusion map (cf. Example 2.27). (If N is a subset of M, the *inclusion* mapping, $i : N \to M$, sends each point of N into the same point, considered as an element of M.)

In classical thermodynamics, a quasistatic adiabatic process is (the image of) a curve C in the space of equilibrium states such that $C^*(\mathrm{d}U + P\,\mathrm{d}V - \mu\,\mathrm{d}\nu) = 0$, where U, P, V, μ, and ν are the internal energy, pressure, volume, chemical potential, and mole number, respectively. In this case, the space of equilibrium states is a manifold of dimension three and the distribution defined by the 1-form $\mathrm{d}U + P\,\mathrm{d}V - \mu\,\mathrm{d}\nu$ (the heat 1-form) is of dimension two.

A distribution $\mathscr{D}$ of dimension l is *completely integrable* in $U \subset M$ if each point $p \in U$ is contained in an integral manifold of $\mathscr{D}$ of dimension l. An integral manifold N of the distribution $\mathscr{D}$ is maximal if any integral manifold N' of $\mathscr{D}$, such that $N \subset N'$, coincides with N.

The second law of thermodynamics states that the distribution defined by the heat 1-form is completely integrable; its integral manifolds are given by $S = \text{const}$, where S is the entropy [see, e.g., Sneddon (2006, Chap. 1)].

Lemma 4.2 *Let $\alpha^1, \ldots, \alpha^k$ be independent 1-forms on M. The distribution $\mathscr{D}$ defined by $\{\alpha^1, \ldots, \alpha^k\}$ is completely integrable in $U \subset M$ if there exist k functionally independent differentiable functions y^i defined in U such that*

$$\alpha^i = c^i_j\,\mathrm{d}y^j, \tag{4.14}$$

where the c^i_j are differentiable functions. (Note that the condition that the 1-forms α^i be independent implies that the matrix (c^i_j) be non-singular, that is, invertible.)

Proof The fact that the functions y^i be functionally independent implies that the set N, formed by the points $p \in U$ such that $y^i(p) = a^i$, where $a^1, \ldots, a^k$ are fixed real numbers, is a submanifold of M of dimension $n - k$ (see Theorem 1.6). Let v_p be a tangent vector to N at p; then $v_p[y^i] = 0$, since at the points of N the y^i are constant. Therefore, using (4.14) and (1.41)

$$\alpha^i_p(v_p) = c^i_j(p)\,\mathrm{d}y^j_p(v_p) = c^i_j(p)v_p\big[y^j\big] = 0, \tag{4.15}$$

that is, $v_p \in \mathscr{D}_p$. Thus, N is an integral manifold of $\mathscr{D}$. □

In particular, a single 1-form, α, defines a distribution of dimension $n - 1$ at the points of M where α does not vanish. According to the preceding lemma, the distribution given by α is completely integrable if there exists a function, y, such that

$$\alpha = \nu\,\mathrm{d}y, \tag{4.16}$$

where ν is some real-valued differentiable function. In the terminology employed in the textbooks on differential equations, when a linear differential form α is of the form (4.16), it is said to be *integrable* (and one says that $1/\nu$ is an *integrating factor* for α).

Theorem 4.3 *A* 1*-form* α *is locally integrable if and only if*

$$\alpha \wedge \mathrm{d}\alpha = 0. \tag{4.17}$$

(*It should be noticed that* $\alpha \wedge \mathrm{d}\alpha$ *is a* 3*-form and therefore is identically equal to zero if the dimension of* M *is one or two*.)

Proof If α has the local expression (4.16) then

$$\alpha \wedge \mathrm{d}\alpha = \nu\, \mathrm{d}y \wedge (\mathrm{d}\nu \wedge \mathrm{d}y) = 0.$$

Let us assume now that $\alpha \wedge \mathrm{d}\alpha = 0$; and let $n = \dim M$. If $n = 1$, the assertion is trivially true since any 1-form has the local expression $\alpha = \alpha_1\, \mathrm{d}x^1$. For $n = 2$, we have $\alpha = \alpha_1\, \mathrm{d}x^1 + \alpha_2\, \mathrm{d}x^2$. Making use of the functions α_1 and α_2 we construct the first-order ODE

$$\frac{\mathrm{d}x^2}{\mathrm{d}x^1} = -\frac{\alpha_1}{\alpha_2} \tag{4.18}$$

whose general solution must contain an arbitrary constant. Assuming that $F(x^1, x^2) = \text{const}$ represents the general solution of (4.18), differentiating implicitly with respect to x^1 and using (4.18), we have

$$0 = \frac{\partial F}{\partial x^1} + \frac{\partial F}{\partial x^2}\frac{\mathrm{d}x^2}{\mathrm{d}x^1} = \frac{\partial F}{\partial x^1} - \frac{\alpha_1}{\alpha_2}\frac{\partial F}{\partial x^2},$$

thus

$$\begin{aligned} \mathrm{d}F &= \frac{\partial F}{\partial x^1}\, \mathrm{d}x^1 + \frac{\partial F}{\partial x^2}\, \mathrm{d}x^2 \\ &= \frac{\alpha_1}{\alpha_2}\frac{\partial F}{\partial x^2}\, \mathrm{d}x^1 + \frac{\partial F}{\partial x^2}\, \mathrm{d}x^2 \\ &= \frac{1}{\alpha_2}\frac{\partial F}{\partial x^2}\big(\alpha_1\, \mathrm{d}x^1 + \alpha_2\, \mathrm{d}x^2\big), \end{aligned}$$

that is,

$$\alpha = \left(\frac{\alpha_2}{\partial F/\partial x^2}\right) \mathrm{d}F.$$

Therefore, any 1-form in two variables is locally integrable.

Considering now an arbitrary value of n, the 1-form α has the local expression $\alpha = \alpha_i\, \mathrm{d}x^i$. On the submanifold N given by $x^n = \text{const} \equiv (x^n)_0$, α becomes

$$\tilde{\alpha} = \sum_{i=1}^{n-1} \tilde{\alpha}_i\, \mathrm{d}x^i,$$

where the real-valued functions $\tilde{\alpha}_i$ depend only on $x^1, \dots, x^{n-1}$, with $(x^n)_0$ as a parameter (see the example given below). More precisely, $\tilde{\alpha}$ is the pullback of α under the inclusion mapping $i : N \to M$ and, by abuse of notation, we denote by $x^1, \dots, x^{n-1}$ the restrictions to N (or pullbacks under i) of the coordinates of M.

Since $i^*(\alpha \wedge \mathrm{d}\alpha) = (i^*\alpha) \wedge \mathrm{d}(i^*\alpha)$ [see (3.25) and (3.38)], the 1-form $\tilde{\alpha}$ satisfies the condition $\tilde{\alpha} \wedge \mathrm{d}\tilde{\alpha} = 0$; therefore, assuming, by induction, that the proposition holds for manifolds of dimension $n-1$, there exist real-valued functions μ and f which depend parametrically on $(x^n)_0$, with

$$\tilde{\alpha} = \mu\big(x^1, \dots, x^{n-1}, (x^n)_0\big)\, \mathrm{d}f\big(x^1, \dots, x^{n-1}, (x^n)_0\big),$$

therefore, eliminating the restriction on x^n,

$$\begin{aligned} \alpha &= \mu\, \mathrm{d}f + \left(\alpha_n - \mu \frac{\partial f}{\partial x^n}\right) \mathrm{d}x^n \\ &\equiv \mu\, \mathrm{d}f + b\, \mathrm{d}x^n. \end{aligned}$$

Substituting this expression into the equation $\alpha \wedge \mathrm{d}\alpha = 0$ we find that

$$\begin{aligned} 0 &= \big(\mu\, \mathrm{d}f + b\, \mathrm{d}x^n\big) \wedge \big(\mathrm{d}\mu \wedge \mathrm{d}f + \mathrm{d}b \wedge \mathrm{d}x^n\big) \\ &= \mathrm{d}x^n \wedge \mathrm{d}f \wedge (\mu\, \mathrm{d}b - b\, \mathrm{d}\mu) \\ &= \mu^2\, \mathrm{d}x^n \wedge \mathrm{d}f \wedge \mathrm{d}\left(\frac{b}{\mu}\right), \end{aligned}$$

which is equivalent to the fact that b/μ is a function of x^n and f only, $b/\mu = g(x^n, f)$. Then

$$\begin{aligned} \alpha &= \mu\left(\mathrm{d}f + \frac{b}{\mu}\, \mathrm{d}x^n\right) \\ &= \mu\big(\mathrm{d}f + g\big(x^n, f\big)\, \mathrm{d}x^n\big). \end{aligned} \tag{4.19}$$

Since $\mathrm{d}f + g(x^n, f)\, \mathrm{d}x^n$ is a 1-form in two variables, it is integrable and therefore α is integrable. □

Example 4.4 Let

$$\alpha = 2z(y+z)\, \mathrm{d}x - 2xz\, \mathrm{d}y + \big[(y+z)^2 - x^2 - 2xz\big]\, \mathrm{d}z,$$

where x, y, z are local coordinates of a manifold of dimension three. One can verify that

$$\mathrm{d}\alpha = 2(x+y+z)\, \mathrm{d}y \wedge \mathrm{d}z + 2(x+y+3z)\, \mathrm{d}z \wedge \mathrm{d}x - 4z\, \mathrm{d}x \wedge \mathrm{d}y$$

and that $\alpha \wedge \mathrm{d}\alpha = 0$; therefore, α is integrable, at least locally.

Following the procedure shown in the proof of the foregoing theorem, making $z = \text{const} \equiv z_0$, we obtain the 1-form in two variables

$$\tilde{\alpha} = 2z_0(y+z_0)\, \mathrm{d}x - 2xz_0\, \mathrm{d}y.$$

Then, equation (4.18) is, in this case, the first-order linear (and also separable) ODE

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{y + z_0}{x},$$

whose general solution is given by $F(x, y) = (y + z_0)/x = \text{const}$ and, as can be verified directly, $\tilde{\alpha} = -2x^2 z_0\, \mathrm{d}F$. Taking now $F(x, y, z) = (y + z)/x$ we find that

$$\begin{aligned}
\alpha &= -2x^2 z\, \mathrm{d}F + \left[(y+z)^2 - x^2\right] \mathrm{d}z \\
&= -2x^2 z \left\{ \mathrm{d}F + \left[\frac{1}{2z} - \frac{(y+z)^2}{2x^2 z} \right] \mathrm{d}z \right\} \\
&= -2x^2 z \left(\mathrm{d}F + \frac{1 - F^2}{2z}\, \mathrm{d}z \right)
\end{aligned}$$

[cf. (4.19)]. The 1-form inside the parentheses in the last equality is, in effect, a 1-form in two variables, which must be integrable. In this case, we can see directly that

$$\begin{aligned}
\mathrm{d}F + \frac{1 - F^2}{2z}\, \mathrm{d}z &= \left(1 - F^2\right) \left(\frac{\mathrm{d}F}{1 - F^2} + \frac{\mathrm{d}z}{2z} \right) \\
&= \frac{(1 - F)^2}{2z}\, \mathrm{d}\left[\frac{z(1 + F)}{1 - F} \right].
\end{aligned}$$

Hence,

$$\alpha = (y + z - x)^2\, \mathrm{d}\left[\frac{(y + z + x)z}{y + z - x} \right]. \tag{4.20}$$

Going back to the more general case of Lemma 4.2, we can see that a necessary condition for the existence of the functions y^i, appearing in (4.14), is obtained by applying the exterior derivative d to the relations $\alpha^i = c^i_j\, \mathrm{d}y^j$. Expressing $\mathrm{d}y^j$ as $\mathrm{d}y^j = \tilde{c}^j_l \alpha^l$, where $(\tilde{c}^j_l)$ is the inverse of the matrix (c^i_j), we have

$$\begin{aligned}
\mathrm{d}\alpha^i &= \mathrm{d}\left(c^i_j\, \mathrm{d}y^j\right) \\
&= \mathrm{d}c^i_j \wedge \mathrm{d}y^j \\
&= \mathrm{d}c^i_j \wedge \left(\tilde{c}^j_l \alpha^l\right) \\
&= \left(\tilde{c}^j_l\, \mathrm{d}c^i_j\right) \wedge \alpha^l \\
&\equiv \theta^i_l \wedge \alpha^l
\end{aligned}$$

with $\theta^i_l \in \Lambda^1(M)$, $i, l = 1, \ldots, k$. It turns out that this condition is also sufficient.

Theorem 4.5 (Frobenius) *Let* $\{\alpha^1, \dots, \alpha^k\}$ *be a set of independent* 1*-forms. In a neighborhood of each point there exist* k *independent functions* y^j *such that* $\alpha^i = c^i_j \, \mathrm{d}y^j$ *if and only if there exist* k^2 1*-forms*, $\theta^i_l \in \Lambda^1(M)$, *such that*

$$\mathrm{d}\alpha^i = \theta^i_l \wedge \alpha^l. \tag{4.21}$$

Proof (Sufficiency) If $\{\alpha^1, \dots, \alpha^k\}$ is a set of k independent 1-forms on a manifold of dimension k, the conclusion is trivial, since if $(x^1, \dots, x^k)$ is any coordinate system, then, $\alpha^i = c^i_j \, \mathrm{d}x^j$, where (c^i_j) is a non-singular matrix.

We now consider the case $n > k$. If $(x^1, \dots, x^n)$ is a local coordinate system on M, the k 1-forms α^i have the local expressions $\alpha^i = a^i_j \, \mathrm{d}x^j$. Let us assume that the 1-forms $\tilde{\alpha}^i \equiv \sum_{j=1}^{n-1} \tilde{a}^i_j \, \mathrm{d}x^j$, obtained from the α^i on setting $x^n = \text{const} \equiv (x^n)_0$, are independent, which can be achieved by relabeling the coordinates if necessary. Then the condition (4.21) implies that $\mathrm{d}\tilde{\alpha}^i = \tilde{\theta}^i_l \wedge \tilde{\alpha}^l$, considering x^n as a parameter. (Again, $\tilde{\alpha}^i$ is the pullback of α^i under the inclusion mapping $i : N \to M$, where N is the submanifold defined by $x^n = \text{const} \equiv (x^n)_0$ and, by abuse of notation, we denote by $x^1, \dots, x^{n-1}$ the restrictions to N (or pullbacks under i) of the coordinates of M.)

Assuming that the Theorem holds for $n - 1$ dimensions, there exist k independent functions, y^j, which depend parametrically on x^n, such that $\tilde{\alpha}^i = b^i_j \, \mathrm{d}y^j$, where (b^i_j) is a non-singular $k \times k$ matrix. Hence,

$$\alpha^i = b^i_j \, \mathrm{d}y^j + a^i \, \mathrm{d}x^n = b^i_j \big(\mathrm{d}y^j + \tilde{b}^j_l a^l \, \mathrm{d}x^n\big) \equiv b^i_j \big(\mathrm{d}y^j + b^j \, \mathrm{d}x^n\big), \tag{4.22}$$

where the a^i are functions and $(\tilde{b}^j_l)$ is the inverse of (b^i_j). Substituting the expression $\alpha^i = b^i_j(\mathrm{d}y^j + b^j \, \mathrm{d}x^n)$ into (4.21), we obtain

$$\mathrm{d}\big(\mathrm{d}y^i + b^i \, \mathrm{d}x^n\big) = {\theta'}^i_l \wedge \big(\mathrm{d}y^l + b^l \, \mathrm{d}x^n\big),$$

with ${\theta'}^i_l \in \Lambda^1(M)$; hence,

$$\mathrm{d}b^i \wedge \mathrm{d}x^n = {\theta'}^i_l \wedge \big(\mathrm{d}y^l + b^l \, \mathrm{d}x^n\big). \tag{4.23}$$

From this equation it follows that

$$\mathrm{d}b^i \wedge \mathrm{d}x^n \wedge \mathrm{d}x^n = {\theta'}^i_l \wedge \big(\mathrm{d}y^l + b^l \, \mathrm{d}x^n\big) \wedge \mathrm{d}x^n,$$

and, since $\mathrm{d}x^n \wedge \mathrm{d}x^n = 0$ [see (3.12)],

$${\theta'}^i_l \wedge \mathrm{d}y^l \wedge \mathrm{d}x^n = 0,$$

which implies that ${\theta'}^i_l = A^i_{lm} \, \mathrm{d}y^m + B^i_l \, \mathrm{d}x^n$, with $A^i_{lm}, B^i_l \in C^\infty(M)$ and $A^i_{lm} = A^i_{ml}$ (see (3.24) et seq.). Substituting into (4.23) we obtain

$$\mathrm{d}b^i \wedge \mathrm{d}x^n \wedge \mathrm{d}y^1 \wedge \cdots \wedge \mathrm{d}y^k = {\theta'}^i_l \wedge \big(\mathrm{d}y^l + b^l \, \mathrm{d}x^n\big) \wedge \mathrm{d}y^1 \wedge \cdots \wedge \mathrm{d}y^k = 0,$$

which means that b^i is function of $x^n, y^1, \dots, y^k$; $b^i = b^i(x^n, y^1, \dots, y^k)$.

Now we consider the system of ODEs made out of the functions $b^i(x^n, y^1, \ldots, y^k)$

$$\frac{\mathrm{d}y^i}{\mathrm{d}x^n} = -b^i\big(x^n, y^1, \ldots, y^k\big), \quad i = 1, 2, \ldots, k. \tag{4.24}$$

The general solution of this system must contain k arbitrary constants. Let $y'^1, \ldots, y'^k$ be independent functions such that $y'^j(x^n, y^1, \ldots, y^k) = \text{const}$ ($j = 1, 2, \ldots, k$) is a solution of the system (4.24). Then (by implicit differentiation as in the proof of Theorem 4.3) one finds that $\mathrm{d}y^i + b^i\,\mathrm{d}x^n = f^i_j\,\mathrm{d}y'^j$, where (f^i_j) is a non-singular matrix; therefore we have, finally

$$\alpha^i = b^i_j f^j_l\,\mathrm{d}y'^l \equiv c^i_l\,\mathrm{d}y'^l. \qquad \square$$

Exercise 4.6 Consider the distribution defined by the 1-forms $\alpha^1 \equiv \mathrm{d}y - z\,\mathrm{d}x$, $\alpha^2 \equiv \mathrm{d}z - F\,\mathrm{d}x$, where (x, y, z) are local coordinates of some manifold M, with $F \in C^\infty(M)$ [cf. (4.11)]. Find explicitly a set of four 1-forms $\theta^1_1, \theta^1_2, \theta^2_1, \theta^2_2$, such that $\mathrm{d}\alpha^i = \theta^i_j \wedge \alpha^j$. Hence, the distribution is locally completely integrable; its integral manifolds are one-dimensional submanifolds of M which represent the solutions of the ODE $y'' = F(x, y, y')$. (See, e.g., Example 4.15, below.)

It is convenient to notice that if the relations (4.21) hold, then, for $i = 1, 2, \ldots, k$,

$$\alpha^1 \wedge \alpha^2 \wedge \cdots \wedge \alpha^k \wedge \mathrm{d}\alpha^i = 0 \tag{4.25}$$

[cf. (4.17)], since, frequently, for a given set of 1-forms $\{\alpha^1, \ldots, \alpha^k\}$, it is simpler to verify that this condition is satisfied than to show the existence of 1-forms θ^i_j satisfying (4.21). Conversely, if the 1-forms $\alpha^1, \ldots, \alpha^k$ satisfy (4.25), then (4.21) holds. Indeed, locally there exist $(n-k)$ 1-forms $\alpha^{k+1}, \ldots, \alpha^n$ such that $\{\alpha^1, \ldots, \alpha^n\}$ is a basis for the 1-forms on M, hence, for $1 \leq i \leq k$, $\mathrm{d}\alpha^i = f^i_{\mu\nu}\alpha^\mu \wedge \alpha^\nu$, where μ, ν range from 1 to n and $f^i_{\mu\nu} = -f^i_{\nu\mu}$ are real-valued functions. Substituting this expression into (4.25) we obtain $\sum_{\mu,\nu=k+1}^n f^i_{\mu\nu}\alpha^1 \wedge \cdots \wedge \alpha^k \wedge \alpha^\mu \wedge \alpha^\nu = 0$, which implies that $f^i_{\mu\nu} = 0$ for $\mu, \nu > k$, provided that $k+2 \leq n$ (otherwise all the products $\alpha^1 \wedge \cdots \wedge \alpha^k \wedge \alpha^\mu \wedge \alpha^\nu$ are equal to zero); hence,

$$\begin{aligned}
\mathrm{d}\alpha^i &= \sum_{j,m=1}^k f^i_{jm}\alpha^j \wedge \alpha^m + 2\sum_{j=1}^k \sum_{\mu=k+1}^n f^i_{j\mu}\alpha^j \wedge \alpha^\mu \\
&= \sum_{m=1}^k \left(\sum_{j=1}^k f^i_{jm}\alpha^j + 2\sum_{\mu=k+1}^n f^i_{\mu m}\alpha^\mu \right) \wedge \alpha^m,
\end{aligned}$$

which is of the form (4.21). (In the cases where $k = n$ or $k = n-1$, the conclusion follows trivially.)

Example 4.7 Let us consider the set of 1-forms

$$\begin{aligned} \alpha^1 &= (xw - yz)\,\mathrm{d}x - xz\,\mathrm{d}y + x^2\,\mathrm{d}w, \\ \alpha^2 &= -z^2\,\mathrm{d}y + (xw - yz)\,\mathrm{d}z + xz\,\mathrm{d}w, \end{aligned} \tag{4.26}$$

where (x, y, z, w) are local coordinates of a manifold of dimension four. By a direct calculation one finds that

$$\begin{aligned} \mathrm{d}\alpha^1 &= x\,\mathrm{d}x \wedge \mathrm{d}w + y\,\mathrm{d}x \wedge \mathrm{d}z + x\,\mathrm{d}y \wedge \mathrm{d}z, \\ \mathrm{d}\alpha^2 &= z\,\mathrm{d}x \wedge \mathrm{d}w + w\,\mathrm{d}x \wedge \mathrm{d}z + z\,\mathrm{d}y \wedge \mathrm{d}z. \end{aligned}$$

Whereas it does not seem simple to determine if there exist 1-forms θ^i_j such that equations (4.21) are satisfied, it can be seen that

$$\begin{aligned} \alpha^1 \wedge \alpha^2 = (xw - yz)\big[&-z^2\,\mathrm{d}x \wedge \mathrm{d}y + xz\,\mathrm{d}x \wedge \mathrm{d}w - xz\,\mathrm{d}y \wedge \mathrm{d}z \\ &+ (xw - yz)\,\mathrm{d}x \wedge \mathrm{d}z - x^2\,\mathrm{d}z \wedge \mathrm{d}w\big], \end{aligned}$$

hence $\alpha^1 \wedge \alpha^2 \wedge \mathrm{d}\alpha^1 = 0 = \alpha^1 \wedge \alpha^2 \wedge \mathrm{d}\alpha^2$, and therefore the distribution given by α^1 and α^2 is completely integrable, at least locally.

Following the procedure employed in the proof of the Frobenius Theorem, we will start from the fact that a system of k 1-forms in k variables is locally completely integrable in a trivial manner; therefore, in this example, we have to reduce the number of variables from four to three and, afterwards, from three to two. The integration process will start, then, with two variables only.

Setting $w = w_0$, $z = z_0$ (constants), the 1-forms α^1 and α^2 reduce to the 1-forms in two variables, denoted by $\bar{\alpha}^1$ and $\bar{\alpha}^2$, that in matrix form are expressed as

$$\begin{pmatrix} \bar{\alpha}^1 \\ \bar{\alpha}^2 \end{pmatrix} = \begin{pmatrix} xw_0 - yz_0 & -xz_0 \\ 0 & -{z_0}^2 \end{pmatrix} \begin{pmatrix} \mathrm{d}x \\ \mathrm{d}y \end{pmatrix}. \tag{4.27}$$

(This is already of the form $\alpha^i = c^i_j\,\mathrm{d}y^j$.) With the aid of the matrix

$$\frac{-1}{{z_0}^2(xw_0 - yz_0)} \begin{pmatrix} -{z_0}^2 & xz_0 \\ 0 & xw_0 - yz_0 \end{pmatrix},$$

which is the inverse of the 2×2 matrix appearing in (4.27), we find that, when only w is kept constant (denoting by $\tilde{\alpha}^1, \tilde{\alpha}^2$ the corresponding forms), we have

$$\begin{pmatrix} \tilde{\alpha}^1 \\ \tilde{\alpha}^2 \end{pmatrix} = \begin{pmatrix} xw_0 - yz & -xz \\ 0 & -z^2 \end{pmatrix} \begin{pmatrix} \mathrm{d}x - \frac{x}{z}\,\mathrm{d}z \\ \mathrm{d}y - \frac{xw_0 - yz}{z^2}\,\mathrm{d}z \end{pmatrix} \tag{4.28}$$

[see (4.22)].

In order to express the column on the right-hand side of the foregoing equation in terms of the differentials of two functions only, we now consider the auxiliary system [cf. (4.24)]

$$\frac{\mathrm{d}x}{\mathrm{d}z} = \frac{x}{z},$$
$$\frac{\mathrm{d}y}{\mathrm{d}z} = \frac{xw_0 - yz}{z^2}.$$

The solution of the first equation is $x = c_1 z$, where c_1 is a constant, and using this expression in the second one we obtain the linear equation

$$\frac{\mathrm{d}y}{\mathrm{d}z} + \frac{y}{z} = \frac{c_1 w_0}{z},$$

whose solution is $yz = c_1 w_0 z + c_2$, where c_2 is a constant. Hence, the solution of the system is given by $u = \text{const}$, $v = \text{const}$, where

$$u \equiv \frac{x}{z}, \qquad v \equiv yz - xw_0.$$

In fact, the 1-forms in the column on the right-hand side of (4.28) are

$$\mathrm{d}x - \frac{x}{z}\,\mathrm{d}z = z\,\mathrm{d}u, \qquad \mathrm{d}y - \frac{xw_0 - yz}{z^2}\,\mathrm{d}z = w_0\,\mathrm{d}u + \frac{1}{z}\,\mathrm{d}v,$$

then, substituting into (4.28),

$$\begin{pmatrix} \tilde{\alpha}^1 \\ \tilde{\alpha}^2 \end{pmatrix} = \begin{pmatrix} xw_0 - yz & -xz \\ 0 & -z^2 \end{pmatrix} \begin{pmatrix} z & 0 \\ w_0 & \frac{1}{z} \end{pmatrix} \begin{pmatrix} \mathrm{d}u \\ \mathrm{d}v \end{pmatrix} = \begin{pmatrix} -yz^2 & -x \\ -z^2 w_0 & -z \end{pmatrix} \begin{pmatrix} \mathrm{d}u \\ \mathrm{d}v \end{pmatrix}.$$

Finally, eliminating the condition that w be a constant, one finds that the original 1-forms are given by

$$\begin{pmatrix} \alpha^1 \\ \alpha^2 \end{pmatrix} = \begin{pmatrix} -yz^2 & -x \\ -z^2 w & -z \end{pmatrix} \begin{pmatrix} \mathrm{d}u \\ \mathrm{d}v \end{pmatrix},$$

without the presence, in this case, of additional terms. This final expression is of the form (4.14) and the integral manifolds of the distribution defined by the 1-forms (4.26) are locally given by $x/z = \text{const}$, $yz - xw = \text{const}$

The constraints of a mechanical system given by a set of 1-forms $\alpha^1, \ldots, \alpha^k$ are *holonomic* if the distribution defined by $\{\alpha^1, \ldots, \alpha^k\}$ is completely integrable. Otherwise, the constraints are non-holonomic.

Exercise 4.8 Show that the constraints (4.13), as well as those of a disk rolling on a plane, are non-holonomic.

Even though the distribution defined by the two 1-forms (4.13) is not completely integrable, there exist one-dimensional integral manifolds of the distribution; a simple example is given by the curve $C^*x = \text{const}$, $C^*y = at$, $C^*\phi = 0$, $C^*\theta = -t$, $C^*\psi = \text{const}$, which therefore represents a possible motion of the sphere. The fact that the distribution is not completely integrable means that it does not have three-dimensional integral manifolds.

Involutive Distributions Given a distribution $\mathscr{D}$, we shall denote by $V_{\mathscr{D}}$ the set of vector fields $\mathbf{X}$ such that $\mathbf{X}_p \in \mathscr{D}_p$ for all $p \in M$; one finds that if $\mathbf{X}, \mathbf{Y} \in V_{\mathscr{D}}$, then $a\mathbf{X} + b\mathbf{Y}$ and $f\mathbf{X}$ also belong to $V_{\mathscr{D}}$ for any $a, b \in \mathbb{R}$ and any real-valued differentiable function f. That is, $V_{\mathscr{D}}$ is a submodule of $\mathfrak{X}(M)$. Conversely, if B is a submodule of $\mathfrak{X}(M)$ such that $\mathscr{D}_p \equiv \{\mathbf{X}_p \mid \mathbf{X} \in B\}$ has dimension l for all p, then $\mathscr{D}$ is a distribution of dimension l; we say that the distribution $\mathscr{D}$ is *involutive* if $[\mathbf{X}, \mathbf{Y}] \in V_{\mathscr{D}}$ for all $\mathbf{X}, \mathbf{Y} \in V_{\mathscr{D}}$.

For a distribution $\mathscr{D}$ defined by k independent 1-forms $\{\alpha^1, \ldots, \alpha^k\}$, $V_{\mathscr{D}}$ is formed by the vector fields $\mathbf{X}$ such that $\alpha^i(\mathbf{X}) = 0$, $i = 1, 2, \ldots, k$. If there exist 1-forms θ^i_l such that $\mathrm{d}\alpha^i = \theta^i_l \wedge \alpha^l$ [see (4.21)], then the distribution $\mathscr{D}$ is involutive since, if $\mathbf{X}, \mathbf{Y} \in V_{\mathscr{D}}$, using the definition (3.30) we have

$$
\begin{aligned}
2\,\mathrm{d}\alpha^i(\mathbf{X}, \mathbf{Y}) &= \mathbf{X}\big(\alpha^i(\mathbf{Y})\big) - \mathbf{Y}\big(\alpha^i(\mathbf{X})\big) - \alpha^i\big([\mathbf{X}, \mathbf{Y}]\big) \\
&= -\alpha^i\big([\mathbf{X}, \mathbf{Y}]\big).
\end{aligned}
$$

On the other hand, from (3.7)

$$
\begin{aligned}
2\,\mathrm{d}\alpha^i(\mathbf{X}, \mathbf{Y}) &= 2\big(\theta^i_l \wedge \alpha^l\big)(\mathbf{X}, \mathbf{Y}) \\
&= \theta^i_l(\mathbf{X})\alpha^l(\mathbf{Y}) - \theta^i_l(\mathbf{Y})\alpha^l(\mathbf{X}) = 0;
\end{aligned}
$$

therefore, $\alpha^i([\mathbf{X}, \mathbf{Y}]) = 0$, that is, $[\mathbf{X}, \mathbf{Y}] \in V_{\mathscr{D}}$.

Conversely, given an involutive distribution, $\mathscr{D}$, let $\mathbf{X}_1, \ldots, \mathbf{X}_l$ be independent vector fields such that $\{\mathbf{X}_{1p}, \ldots, \mathbf{X}_{lp}\}$ is basis of $\mathscr{D}_p$. If $\mathbf{X}_{l+1}, \ldots, \mathbf{X}_n$ are $n - l$ vector fields such that $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ is a basis of $\mathfrak{X}(M)$ and denoting by $\{\alpha^1, \ldots, \alpha^n\}$ its dual basis, the distribution $\mathscr{D}$ is defined by the 1-forms $\{\alpha^{l+1}, \ldots, \alpha^n\}$; that is, $\mathscr{D}_p = \{v_p \in T_pM \mid \alpha^i_p(v_p) = 0,\ i = l+1, \ldots, n\}$.

Since $\mathscr{D}$ is involutive, for $i > l$ and $1 \leq j, m \leq l$, we have

$$
\begin{aligned}
2\,\mathrm{d}\alpha^i(\mathbf{X}_j, \mathbf{X}_m) &= \mathbf{X}_j\big(\alpha^i(\mathbf{X}_m)\big) - \mathbf{X}_m\big(\alpha^i(\mathbf{X}_j)\big) - \alpha^i\big([\mathbf{X}_j, \mathbf{X}_m]\big) \\
&= -\alpha^i\big([\mathbf{X}_j, \mathbf{X}_m]\big) = 0,
\end{aligned}
$$

which, substituted into the identity

$$
\mathrm{d}\alpha^i = \big[\mathrm{d}\alpha^i(\mathbf{X}_j, \mathbf{X}_m)\big]\alpha^j \wedge \alpha^m = 2\sum_{j<m}\big[\mathrm{d}\alpha^i(\mathbf{X}_j, \mathbf{X}_m)\big]\alpha^j \wedge \alpha^m,
$$

yields, for $i > l$,

$$\begin{aligned} \mathrm{d}\alpha^i &= 2 \sum_{\substack{j<m \\ m>l}} \left[\mathrm{d}\alpha^i(\mathbf{X}_j, \mathbf{X}_m)\right]\alpha^j \wedge \alpha^m \\ &= \sum_{m>l} \left[2 \sum_{j<m} \mathrm{d}\alpha^i(\mathbf{X}_j, \mathbf{X}_m)\alpha^j\right] \wedge \alpha^m \\ &\equiv \sum_{m=l+1}^{n} \theta^i_m \wedge \alpha^m, \end{aligned}$$

which means that the distribution is completely integrable. Putting together the foregoing results, the Frobenius Theorem can be expressed in the following form.

Theorem 4.9 *Let $\mathscr{D}$ be a distribution on M. The distribution $\mathscr{D}$ is completely integrable in a neighborhood of each point if and only if $\mathscr{D}$ is involutive.*

The l-dimensional integral manifolds of a completely integrable distribution of dimension l defined by the l independent vector fields $\mathbf{X}_1, \ldots, \mathbf{X}_l$ are locally given by $y^i = \text{const}$, $i = 1, 2, \ldots, n-l$, where the y^i are $n-l$ functionally independent solutions of the l linear PDEs

$$\mathbf{X}_i y^j = 0, \quad i = 1, 2, \ldots, l; \;\; j = 1, 2, \ldots, n-l. \tag{4.29}$$

It may be noticed that these equations imply that $[\mathbf{X}_i, \mathbf{X}_j]y^m = 0$, for $i, j = 1, 2, \ldots, l$, $m = 1, 2, \ldots, n-l$. On the other hand, any vector field $\mathbf{Z}$ such that $\mathbf{Z}y^j = 0$ for $j = 1, 2, \ldots, n-l$ must be a linear combination of the $\mathbf{X}_i$, and therefore the Lie brackets $[\mathbf{X}_i, \mathbf{X}_j]$ must be linear combinations of the $\mathbf{X}_s$, which amounts to saying that the distribution must be involutive, as we already knew. The Frobenius Theorem ensures that the converse is also true; that is, if the distribution defined by the l vector fields $\mathbf{X}_1, \ldots, \mathbf{X}_l$ is involutive, then there exist locally $n-l$ functionally independent solutions y^j of (4.29), and the l-dimensional integral manifolds of the distribution are given by $y^j = \text{const}$.

Note that any distribution $\mathscr{D}$ of dimension one is involutive and, therefore, completely integrable, for if $\mathbf{X}$ is a vector field that at each point generates $\mathscr{D}_p$, then any pair of vector fields $\mathbf{Y}, \mathbf{Z} \in V_{\mathscr{D}}$ is of the form $\mathbf{Y} = f\mathbf{X}$ and $\mathbf{Z} = g\mathbf{X}$ [with $f, g \in C^\infty(M)$] and therefore

$$[\mathbf{Y}, \mathbf{Z}] = [f\mathbf{X}, g\mathbf{X}] = \big(f(\mathbf{X}g) - g(\mathbf{X}f)\big)\mathbf{X} \in V_{\mathscr{D}}.$$

The integral manifolds of $\mathscr{D}$ are the images of the integral curves of $\mathbf{X}$. For example, one finds that all the vector fields satisfying the condition $\alpha^i(\mathbf{X}) = 0$ for the two 1-forms (4.11) are of the form $f(\partial/\partial x + z\,\partial/\partial y + F\,\partial/\partial z)$, where f is an arbitrary real-valued function. Thus, the integral manifolds of the distribution defined by these two 1-forms are the images of the integral curves of the vector field $\partial/\partial x + z\,\partial/\partial y + F\,\partial/\partial z$.

Exercise 4.10 Let $\omega \in \Lambda^2(M)$ and let $\mathscr{D}_p \equiv \{v_p \in T_pM \mid v_p \lrcorner\, \omega_p = 0\}$. Show that if there exists a $\theta \in \Lambda^1(M)$ such that $\mathrm{d}\omega = \theta \wedge \omega$, then $\mathscr{D}$ is completely integrable.

4.3 Symmetries and Integrating Factors

In the previous section we have seen how to find, in principle, the integral manifolds of a completely integrable distribution. In this section we shall show how the knowledge of one-parameter groups of symmetries of a distribution allows us to find its integral manifolds.

Let φ_t be a (possibly local) one-parameter group of transformations on M and let α be a 1-form on M. We shall say that α is *invariant* under φ_t if for each value of t in the domain of φ_t there exists some function different from zero, χ_t, such that

$$\varphi_t{}^*\alpha = \chi_t \alpha. \tag{4.30}$$

Then, φ_t maps each integral manifold of α into another integral manifold. For instance, the one-parameter group of transformations on $\mathbb{R}^n$ given by $\varphi_t(x^1, \dots, x^n) = \mathrm{e}^t(x^1, \dots, x^n)$, i.e., $\varphi_t{}^* x^i = \mathrm{e}^t x^i$, leaves invariant any 1-form $\alpha = \alpha_i\, \mathrm{d}x^i$ whose components are homogeneous functions of the same degree k [that is, $\alpha_i(\lambda x^1, \dots, \lambda x^n) = \lambda^k \alpha_i(x^1, \dots, x^n)$] since

$$\varphi_t{}^*\big(\alpha_i\, \mathrm{d}x^i\big) = \mathrm{e}^{(k+1)t} \alpha_i\, \mathrm{d}x^i.$$

Condition (4.30) implies that $\pounds_{\mathbf{X}}\alpha = \nu\alpha$, where $\mathbf{X}$ is the infinitesimal generator of φ and ν is the partial derivative of χ_t with respect to t, evaluated at $t = 0$.

On the other hand, applying the relation (3.39) and the properties (3.27) and (3.23) we find that, for $\mathbf{X} \in \mathfrak{X}(M)$ and $\alpha \in \Lambda^1(M)$,

$$\begin{aligned} \alpha \wedge \pounds_{\mathbf{X}}\alpha &= \alpha \wedge \big[\mathbf{X} \lrcorner\, \mathrm{d}\alpha + \mathrm{d}(\mathbf{X} \lrcorner\, \alpha)\big] \\ &= -\mathbf{X} \lrcorner (\alpha \wedge \mathrm{d}\alpha) + (\mathbf{X} \lrcorner\, \alpha)\, \mathrm{d}\alpha - \mathrm{d}(\mathbf{X} \lrcorner\, \alpha) \wedge \alpha \\ &= -\mathbf{X} \lrcorner (\alpha \wedge \mathrm{d}\alpha) + (\mathbf{X} \lrcorner\, \alpha)^2\, \mathrm{d}\big[(\mathbf{X} \lrcorner\, \alpha)^{-1}\alpha\big], \end{aligned}$$

where we have assumed that $\mathbf{X} \lrcorner\, \alpha$ is different from zero; therefore if α is integrable (which, according to Theorem 4.3, implies that $\alpha \wedge \mathrm{d}\alpha = 0$), then

$$\mathrm{d}\big[(\mathbf{X} \lrcorner\, \alpha)^{-1}\alpha\big] = (\mathbf{X} \lrcorner\, \alpha)^{-2}\, \alpha \wedge \pounds_{\mathbf{X}}\alpha,$$

thus showing that $(\mathbf{X} \lrcorner\, \alpha)^{-1}$ is an integrating factor of α if and only if $\mathbf{X}$ is the infinitesimal generator of a (possibly local) one-parameter group of transformations that leaves α invariant.

Hence, if an integrable 1-form, α, is invariant under the group generated by $\mathbf{X}$ and $\mathbf{X} \lrcorner\, \alpha \neq 0$, then there exists locally a function y such that

$$\alpha = (\mathbf{X} \lrcorner\, \alpha)\, \mathrm{d}y. \tag{4.31}$$

This result implies that a nonzero integrable 1-form on a manifold of dimension greater than or equal to two possesses an infinite number of symmetries. For instance, if $\mathbf{X}$ is the infinitesimal generator of a one-parameter group of transformations that leaves α invariant and $\mathbf{Y}$ is any vector field such that $\mathbf{Y} \lrcorner \alpha = 0$, then $[(\mathbf{X}+\mathbf{Y}) \lrcorner \alpha]^{-1} = (\mathbf{X} \lrcorner \alpha)^{-1}$ is an integrating factor of α, and therefore $\mathbf{X}+\mathbf{Y}$ is the infinitesimal generator of another one-parameter group of transformations that leaves α invariant.

Example 4.11 Since the components of the 1-form

$$\alpha = 2z(y+z)\,\mathrm{d}x - 2xz\,\mathrm{d}y + \big[(y+z)^2 - x^2 - 2xz\big]\,\mathrm{d}z,$$

considered in Example 4.4, are homogeneous functions of degree 2, this 1-form is invariant under the one-parameter group of transformations given by $\varphi_t(x,y,z) = \mathrm{e}^t(x,y,z)$, whose infinitesimal generator is $\mathbf{X} = x\,\partial/\partial x + y\,\partial/\partial y + z\,\partial/\partial z$. Hence, an integrating factor for this 1-form is given by

$$(\mathbf{X} \lrcorner \alpha)^{-1} = \frac{1}{z[(y+z)^2 - x^2]}$$

and, in effect, one finds that

$$\frac{1}{z[(y+z)^2 - x^2]}\,\alpha = \mathrm{d}\ln\left|\frac{(y+z+x)z}{y+z-x}\right|$$

[cf. (4.20)].

Example 4.12 Another way of finding an integrating factor for the 1-form α considered in Example 4.11 consists of using a coordinate system adapted to the vector field $\mathbf{X} = x\,\partial/\partial x + y\,\partial/\partial y + z\,\partial/\partial z$, which generates a symmetry of α.

Following the steps given in Example 4.1, or by inspection, one finds that y/x and z/x are constant along the integral curves of $\mathbf{X}$ [i.e., $\mathbf{X}(y/x) = 0 = \mathbf{X}(z/x)$] and that $\mathbf{X}\ln|x| = 1$; hence, in terms of the coordinates (u, v, w) defined by

$$u \equiv \frac{y}{x}, \qquad v \equiv \frac{z}{x}, \qquad w \equiv \ln x,$$

we have $\mathbf{X} = \partial/\partial w$. In terms of the new coordinates, the 1-form α is given by

$$\begin{aligned}
\alpha &= 2\mathrm{e}^{2w} v(u+v)\,\mathrm{d}\mathrm{e}^w - 2\mathrm{e}^w v\,\mathrm{d}\big(u\mathrm{e}^w\big) + \mathrm{e}^{2w}\big[(u+v)^2 - 1 - 2v\big]\,\mathrm{d}\big(v\mathrm{e}^w\big) \\
&= \mathrm{e}^w\Big\{\big[(u+v)^2 v - v\big]\,\mathrm{d}w - 2v\,\mathrm{d}u + \big[(u+v)^2 - 1 - 2v\big]\,\mathrm{d}v\Big\} \\
&= \mathrm{e}^{3w} v\big[(u+v)^2 - 1\big]\left[\mathrm{d}w + \frac{\mathrm{d}v}{v} - \frac{2\,\mathrm{d}(u+v)}{(u+v)^2 - 1}\right] \\
&= \mathrm{e}^{3w} v\big[(u+v)^2 - 1\big]\,\mathrm{d}\left(w + \ln|v| + \ln\left|\frac{u+v+1}{u+v-1}\right|\right),
\end{aligned}$$

hence

$$\alpha = z\big[(y+z)^2 - x^2\big]\,\mathrm{d}\ln\left|\frac{(y+z+x)z}{y+z-x}\right|,$$

which coincides with the expression obtained above. (It should be pointed out, however, that if there exists an integrating factor for a 1-form α, then there exists an infinite number of integrating factors.)

Exercise 4.13 Show that the 1-form $\alpha = \alpha_i\,\mathrm{d}x^i$ is invariant under the one-parameter group of transformations generated by $\mathbf{X} = \partial/\partial x^1$ if and only if (assuming $\alpha_1 \neq 0$)

$$\frac{\partial}{\partial x^1}\left(\frac{\alpha_i}{\alpha_1}\right) = 0 \quad (i = 2, 3, \ldots, n).$$

If $\mathbf{X} = \partial/\partial x^1$ generates a one-parameter group of transformations that leaves invariant the 1-form $\alpha = \alpha_i\,\mathrm{d}x^i$, then $\alpha_1 = \mathbf{X}\lrcorner\alpha$ is an integrating factor of α. This means that locally there exists a function y such that

$$\alpha = \alpha_1\left(\mathrm{d}x^1 + \sum_{i=2}^{n}\frac{\alpha_i}{\alpha_1}\,\mathrm{d}x^i\right) = \alpha_1\,\mathrm{d}y$$

(cf. Example 4.12).

In a more general way, the set of 1-forms $\alpha^1, \ldots, \alpha^k$ is invariant under a (possibly local) one-parameter group of transformations, φ, if there exist k^2 functions Λ^i_j such that

$$\varphi_t{}^*\alpha^i = \Lambda^i_j\alpha^j; \tag{4.32}$$

therefore, there exist functions N^i_j such that

$$\pounds_{\mathbf{X}}\alpha^i = N^i_j\alpha^j, \tag{4.33}$$

where $\mathbf{X}$ is the infinitesimal generator of φ and N^i_j is the partial derivative with respect to t of Λ^i_j at $t = 0$.

If the system is completely integrable, then there exist k^2 functions c^i_j such that

$$\alpha^i = c^i_j\,\mathrm{d}y^j \tag{4.34}$$

[cf. (4.14)]. By analogy with (4.31), the functions c^i_j can be expressed in the form

$$c^i_j = \mathbf{X}_j\lrcorner\alpha^i, \tag{4.35}$$

in terms of k vector fields $\mathbf{X}_1, \ldots, \mathbf{X}_k$, which, however, are not uniquely determined by these relations. Combining (4.34) and (4.35) one finds that $c^i_j = \mathbf{X}_j\lrcorner\alpha^i =$

$c^i_m(\mathbf{X}_j y^m)$, which implies that

$$\mathbf{X}_j y^m = \delta^m_j. \tag{4.36}$$

Hence, $[\mathbf{X}_i, \mathbf{X}_j] y^m = 0$ or, by virtue of (4.34),

$$[\mathbf{X}_i, \mathbf{X}_j] \lrcorner \alpha^m = 0 \quad (i, j, m = 1, 2, \ldots, k). \tag{4.37}$$

Using several properties of the Lie derivative [(2.23), (2.38), and (2.39)] from (4.34) and (4.36) one finds that

$$\begin{aligned}
\pounds_{\mathbf{X}_j} \alpha^i &= \pounds_{\mathbf{X}_j} \left(c^i_m \mathrm{d} y^m \right) \\
&= \left(\mathbf{X}_j c^i_m \right) \mathrm{d} y^m + c^i_m \, \mathrm{d} \left(\mathbf{X}_j y^m \right) \\
&= \left(\mathbf{X}_j c^i_m \right) \tilde{c}^m_r \alpha^r .
\end{aligned}$$

Comparing with (4.33), one concludes that each of the k vector fields $\mathbf{X}_j$ defined by (4.35) is the infinitesimal generator of a one-parameter group of transformations that leaves invariant the system $\alpha^1, \ldots, \alpha^k$. Now, we shall show that the converse is also true.

Theorem 4.14 *Let $\alpha^1, \ldots, \alpha^k$ be a set of independent 1-forms that define a completely integrable distribution. Let $\mathbf{X}_1, \ldots, \mathbf{X}_k$ be vector fields that generate nontrivial symmetries of the distribution, i.e., the matrix (c^i_j) with $c^i_j \equiv \mathbf{X}_j \lrcorner \alpha^i$ is nonsingular, and let them satisfy the additional conditions $[\mathbf{X}_i, \mathbf{X}_j] \lrcorner \, \alpha^m = 0$. Then, locally, $\alpha^i = c^i_j \, \mathrm{d} y^j$, where $y^1, \ldots, y^k$ are real-valued functions.*

Proof According to the hypotheses, there exists a set of 1-forms θ^i_j such that $\mathrm{d}\alpha^i = \theta^i_j \wedge \alpha^j$ and

$$\begin{aligned}
\pounds_{\mathbf{X}_m} \alpha^r &= \mathbf{X}_m \lrcorner \mathrm{d}\alpha^r + \mathrm{d}\left(\mathbf{X}_m \lrcorner \alpha^r \right) \\
&= \mathbf{X}_m \lrcorner \left(\theta^r_s \wedge \alpha^s \right) + \mathrm{d} c^r_m \\
&= \left(\mathbf{X}_m \lrcorner \theta^r_s \right) \alpha^s - c^s_m \theta^r_s + \mathrm{d} c^r_m \\
&= N^r_{ms} \alpha^s ,
\end{aligned} \tag{4.38}$$

for some real-valued functions N^r_{ms} [cf. (4.33)]. The conditions $[\mathbf{X}_m, \mathbf{X}_j] \lrcorner \alpha^r = 0$ amount to

$$\begin{aligned}
0 = (\pounds_{\mathbf{X}_m} \mathbf{X}_j) \lrcorner \alpha^r &= \pounds_{\mathbf{X}_m} \left(\mathbf{X}_j \lrcorner \alpha^r \right) - \mathbf{X}_j \lrcorner \left(\pounds_{\mathbf{X}_m} \alpha^r \right) \\
&= \mathbf{X}_m c^r_j - \mathbf{X}_j \lrcorner N^r_{ms} \alpha^s = \mathbf{X}_m c^r_j - N^r_{ms} c^s_j ,
\end{aligned}$$

hence

$$N^r_{ms} = \tilde{c}^j_s \, \mathbf{X}_m c^r_j , \tag{4.39}$$

where $(\tilde{c}^i_j)$ is the inverse of the matrix (c^i_j). Thus, making use of (4.38) and (4.39)

$$\begin{aligned} \mathrm{d}\big(\tilde{c}^i_j\alpha^j\big) &= \mathrm{d}\tilde{c}^i_j \wedge \alpha^j + \tilde{c}^i_j\, \mathrm{d}\alpha^j \\ &= -\tilde{c}^i_r \tilde{c}^m_j\, \mathrm{d}c^r_m \wedge \alpha^j + \tilde{c}^i_j \theta^j_m \wedge \alpha^m \\ &= -\tilde{c}^i_r \tilde{c}^m_j \big[N^r_{ms}\alpha^s - \big(\mathbf{X}_m \lrcorner \theta^r_s\big)\alpha^s + c^s_m \theta^r_s\big] \wedge \alpha^j + \tilde{c}^i_j \theta^j_m \wedge \alpha^m \\ &= -\tilde{c}^i_r \tilde{c}^m_j \big(\tilde{c}^l_s\, \mathbf{X}_m c^r_l - \mathbf{X}_m \lrcorner \theta^r_s\big)\, \alpha^s \wedge \alpha^j \\ &= -\tilde{c}^i_r \tilde{c}^m_j \tilde{c}^l_s \big(\mathbf{X}_m c^r_l - c^p_l\, \mathbf{X}_m \lrcorner \theta^r_p\big)\, \alpha^s \wedge \alpha^j . \end{aligned} \tag{4.40}$$

On the other hand, from (3.30), we have

$$2\, \mathrm{d}\alpha^r(\mathbf{X}_m, \mathbf{X}_l) = \mathbf{X}_m\big(\mathbf{X}_l \lrcorner \alpha^r\big) - \mathbf{X}_l\big(\mathbf{X}_m \lrcorner \alpha^r\big) - [\mathbf{X}_m, \mathbf{X}_l] \lrcorner \alpha^r = \mathbf{X}_m c^r_l - \mathbf{X}_l c^r_m,$$

which must coincide with [see (3.7)]

$$2\, \mathrm{d}\alpha^r(\mathbf{X}_m, \mathbf{X}_l) = 2\big(\theta^r_p \wedge \alpha^p\big)(\mathbf{X}_m, \mathbf{X}_l) = c^p_l\, \mathbf{X}_m \lrcorner \theta^r_p - c^p_m\, \mathbf{X}_l \lrcorner \theta^r_p .$$

Thus, the expression inside the parentheses in (4.40) is symmetric in the subscripts m, l,

$$\mathbf{X}_m c^r_l - c^p_l\, \mathbf{X}_m \lrcorner \theta^r_p = \mathbf{X}_l c^r_m - c^p_m\, \mathbf{X}_l \lrcorner \theta^r_p ,$$

and therefore, by virtue of the skew-symmetry of the exterior product $\alpha^s \wedge \alpha^j$, we find that $\mathrm{d}(\tilde{c}^i_j \alpha^j) = 0$. □

Example 4.15 The distribution defined by the two 1-forms

$$\alpha^1 = \mathrm{d}y - z\, \mathrm{d}x, \alpha \qquad ^2 = \mathrm{d}z - (x-y)z^3\, \mathrm{d}x,$$

is invariant under the one-parameter groups generated by Stephani (1989, Sect. 7.5)

$$\mathbf{X}_1 = \frac{\partial}{\partial x} + \frac{\partial}{\partial y}, \qquad \mathbf{X}_2 = (x-y)\frac{\partial}{\partial x} + z(z-1)\frac{\partial}{\partial z}.$$

A direct computation shows that $[\mathbf{X}_1, \mathbf{X}_2] = 0$ and

$$\big(c^i_j\big) = \big(\mathbf{X}_j \lrcorner \alpha^i\big) = \begin{pmatrix} 1-z & -(x-y)z \\ -(x-y)z^3 & z(z-1) - (x-y)^2 z^3 \end{pmatrix}.$$

Hence $\det(c^i_j) = -z[(z-1)^2 + (x-y)^2 z^2]$ and therefore, except at the points where $z = 0$ or on the line $x = y$, $z = 1$, the conditions of Theorem 4.14 are satisfied.

A straightforward computation yields

$$\big(\tilde{c}^i_j\big) = \frac{1}{z[(z-1)^2 + (x-y)^2 z^2]} \begin{pmatrix} (x-y)^2 z^3 - z(z-1) & -(x-y)z \\ -(x-y)z^3 & z-1 \end{pmatrix}.$$

As shown in Theorem 4.14, the 1-forms $\tilde{c}^i_j\alpha^j$ must be, locally, exact; indeed, we find

$$\begin{aligned}
\tilde{c}^1_j\alpha^j &= \frac{1}{z[(z-1)^2+(x-y)^2z^2]} \\
&\quad \times \left\{z^2(z-1)\,\mathrm{d}x + \left[(x-y)z^3 - z(z-1)\right]\mathrm{d}y - (x-y)z\,\mathrm{d}z\right\} \\
&= \frac{1}{(z-1)^2+(x-y)^2z^2} \\
&\quad \times \left[z(z-1)\,\mathrm{d}(x-y) + (z-1)^2\,\mathrm{d}y + (x-y)^2z^2\,\mathrm{d}y - (x-y)\,\mathrm{d}z\right] \\
&= \mathrm{d}\left[y + \arctan\frac{(x-y)z}{z-1}\right]
\end{aligned}$$

and

$$\begin{aligned}
\tilde{c}^2_j\alpha^j &= \frac{1}{z[(z-1)^2+(x-y)^2z^2]}\left[(x-y)z^3\,\mathrm{d}x - (x-y)z^3\,\mathrm{d}y + (z-1)\,\mathrm{d}z\right] \\
&= \frac{z^2}{(z-1)^2+(x-y)^2z^2}\left[(x-y)\,\mathrm{d}(x-y) + \frac{z-1}{z^3}\,\mathrm{d}z\right] \\
&= \frac{1}{2}\,\mathrm{d}\ln\left[\left(1-\frac{1}{z}\right)^2 + (x-y)^2\right].
\end{aligned}$$

Thus, the integral manifolds of the distribution are given by

$$y + \arctan\frac{(x-y)z}{z-1} = \text{const}, \qquad \left(1-\frac{1}{z}\right)^2 + (x-y)^2 = \text{const}. \tag{4.41}$$

Symmetries of a Second-Order Ordinary Differential Equation The results derived above can be applied to the specific case of a second-order ODE.

The second-order ODE

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = F(x, y, \mathrm{d}y/\mathrm{d}x) \tag{4.42}$$

is equivalent to the system of first-order ODEs

$$\frac{\mathrm{d}x}{\mathrm{d}t} = 1, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = z, \qquad \frac{\mathrm{d}z}{\mathrm{d}t} = F(x, y, z),$$

which, as shown in Sect. 2.1, determines the integral curves of the vector field

$$\mathbf{A} = \frac{\partial}{\partial x} + z\frac{\partial}{\partial y} + F(x, y, z)\frac{\partial}{\partial z} \tag{4.43}$$

on a manifold M with local coordinates (x, y, z).

The vector field **A** is the only vector field on M that satisfies

$$\mathbf{A} \lrcorner \alpha^i = 0, \qquad \mathbf{A}x = 1, \tag{4.44}$$

where

$$\alpha^1 \equiv \mathrm{d}y - z\,\mathrm{d}x, \alpha \qquad ^2 \equiv \mathrm{d}z - F(x, y, z)\,\mathrm{d}x. \tag{4.45}$$

From the equation $(£_{\mathbf{X}}\mathbf{A}) \lrcorner \alpha^i = £_{\mathbf{X}}(\mathbf{A} \lrcorner \alpha^i) - \mathbf{A} \lrcorner (£_{\mathbf{X}}\alpha^i) = -\mathbf{A} \lrcorner (£_{\mathbf{X}}\alpha^i)$ it follows that the set formed by the two 1-forms α^1 and α^2 is invariant under the group generated by a vector field **X** if and only if

$$[\mathbf{X}, \mathbf{A}] = \lambda \mathbf{A} \tag{4.46}$$

(that is, $£_{\mathbf{X}}\mathbf{A} = \lambda \mathbf{A}$), for some real-valued function λ.

In some cases it is possible to find by inspection symmetry groups of a system of the form (4.45). For instance, the 1-forms $\mathrm{d}y - z\,\mathrm{d}x$ and $\mathrm{d}z - (3xz^3/y^2)\,\mathrm{d}x$ transform into multiples of themselves when x, y, z are replaced by ax, $a^{-2}y$, $a^{-3}z$, for $a \in \mathbb{R}$. This means that the system (4.45) with $F(x, y, z) = 3xz^3/y^2$ (which corresponds to the second-order equation $y^2 y'' = 3xy'^3$) is invariant under the one-parameter group of transformations $\varphi_t(x, y, z) = (\mathrm{e}^t x, \mathrm{e}^{-2t} y, \mathrm{e}^{-3t} z)$, whose infinitesimal generator is $\mathbf{X} = x\,\partial/\partial x - 2y\,\partial/\partial y - 3z\,\partial/\partial z$.

Writing

$$\mathbf{X} = \xi \frac{\partial}{\partial x} + \eta \frac{\partial}{\partial y} + \zeta \frac{\partial}{\partial z}, \tag{4.47}$$

one finds that (4.46) amounts to two PDEs (for the functions ξ, η, and ζ) whose solution is difficult to obtain. However, by imposing the condition that the functions ξ and η depend only on x and y (which corresponds to the so-called Lie point symmetries), a straightforward computation shows that the condition (4.46) is equivalent to

$$\zeta = \eta_x + z(\eta_y - \xi_x) - z^2 \xi_y \tag{4.48}$$

where the subscripts denote partial differentiation [cf. (2.21)] and

$$\xi F_x + \eta F_y + \zeta F_z = \zeta_x + z\zeta_y + F\zeta_z - (\xi_x + z\xi_y)F \tag{4.49}$$

[cf. (2.41)]. Substituting the relation (4.48) into (4.49), in order to eliminate ζ, one obtains a PDE for the two functions of two variables ξ and η [see also Hydon (2000, Sect. 3.2)].

Knowing one or several symmetry groups of the system (4.45) allows us to find the solutions of the ODE (4.42). For instance, if we have two suitable symmetries satisfying the conditions of Theorem 4.14, we can readily find the solutions of (4.42). Nevertheless, if we know only one vector field **X** that generates a nontrivial symmetry group of (4.45) (that is, **X** satisfies (4.46) but is not proportional to **A**), we can calculate the 1-form

$$\beta \equiv \mathbf{X} \lrcorner \big(\alpha^1 \wedge \alpha^2\big) \tag{4.50}$$

which is proportional to the differential of a *first integral* of (4.42), that is, $\beta = \mu\,\mathrm{d}\chi$, where μ is some function and $\mathbf{A}\chi = 0$. (If $\mathbf{X}$ is proportional to $\mathbf{A}$, then β is equal to zero.)

In order to demonstrate the preceding assertion it may be noticed that from (4.45) we have

$$\mathrm{d}(\alpha^1 \wedge \alpha^2) = F_z\,\mathrm{d}x \wedge (\alpha^1 \wedge \alpha^2)$$

(cf. Exercise 4.10), and from (4.33) it follows that

$$£_{\mathbf{X}}(\alpha^1 \wedge \alpha^2) = (N_1^1 + N_2^2)\,\alpha^1 \wedge \alpha^2.$$

Hence, using (3.39) and the previous relations,

$$\begin{aligned}
\mathrm{d}\beta &= \mathrm{d}\big[\mathbf{X} \lrcorner (\alpha^1 \wedge \alpha^2)\big] \\
&= £_{\mathbf{X}}(\alpha^1 \wedge \alpha^2) - \mathbf{X} \lrcorner \mathrm{d}(\alpha^1 \wedge \alpha^2) \\
&= (N_1^1 + N_2^2)\,\alpha^1 \wedge \alpha^2 - \mathbf{X} \lrcorner \big[F_z\,\mathrm{d}x \wedge (\alpha^1 \wedge \alpha^2)\big] \\
&= (N_1^1 + N_2^2)\,\alpha^1 \wedge \alpha^2 - F_z(\mathbf{X} \lrcorner \mathrm{d}x) \wedge (\alpha^1 \wedge \alpha^2) + F_z\,\mathrm{d}x \wedge \beta.
\end{aligned}$$

On the other hand, from (4.50) it follows that β is a combination of α^1 and α^2, therefore $\beta \wedge \mathrm{d}\beta = 0$, which is equivalent to the local existence of two functions μ and χ such that $\beta = \mu\,\mathrm{d}\chi$ (Theorem 4.3). Thus, from (4.50) and (4.44) we have $\mathbf{A} \lrcorner \beta = \mathbf{A} \lrcorner (\mathbf{X} \lrcorner (\alpha^1 \wedge \alpha^2)) = -\mathbf{X} \lrcorner (\mathbf{A} \lrcorner (\alpha^1 \wedge \alpha^2)) = 0$, i.e., $\mathbf{A} \lrcorner (\mu\,\mathrm{d}\chi) = \mu \mathbf{A}\chi = 0$, as claimed above.

The condition $\mathbf{A}\chi = 0$ means that each integral curve of $\mathbf{A}$ is contained in some surface $\chi = \text{const}$ (that is, in one of the level surfaces of χ). The definition (4.50) and the expression $\beta = \mu\,\mathrm{d}\chi$ give $(\mathbf{X} \lrcorner \alpha^1)\,\alpha^2 - (\mathbf{X} \lrcorner \alpha^2)\,\alpha^1 = \mu\,\mathrm{d}\chi$, therefore on a surface $\chi = \text{const}$, one of the 1-forms α^i is proportional to the other, hence on the submanifold $\chi = \text{const}$, any nonzero 1-form, γ, combination of the α^i, is integrable because it is a 1-form on a manifold of dimension two. As in the case of β, the 1-form γ is proportional to the differential of a first integral of (4.42) (since $\mathbf{A} \lrcorner \alpha^i = 0$), which is functionally independent of χ and these two first integrals give the integral curves of $\mathbf{A}$ or, equivalently, the solutions of (4.42).

Example 4.16 The second-order ODE $y'' = (x - y)y'^3$ corresponds to the system of 1-forms

$$\alpha^1 = \mathrm{d}y - z\,\mathrm{d}x, \alpha \qquad {}^2 = \mathrm{d}z - (x - y)z^3\,\mathrm{d}x,$$

which is invariant under the one-parameter group of translations $\varphi_t(x, y, z) = (x + t, y + t, z)$, whose infinitesimal generator is $\mathbf{X} = \partial/\partial x + \partial/\partial y$. The 1-form (4.50) is in this case

$$\mathbf{X} \lrcorner (\alpha^1 \wedge \alpha^2) = (1 - z)\,\mathrm{d}z - (x - y)z^3(\mathrm{d}x - \mathrm{d}y),$$

which is indeed integrable and is equivalent to

$$-\frac{z^3}{2}\,\mathrm{d}\left[(x-y)^2+\frac{1}{z^2}-\frac{2}{z}\right]=-\frac{z^3}{2}\,\mathrm{d}\left[(x-y)^2+\left(\frac{1}{z}-1\right)^2\right].$$

Hence, we can take $\chi=(x-y)^2+(\frac{1}{z}-1)^2$.

On the surface $\chi=\text{const}$ we have $1/z=1\pm\sqrt{c^2-(x-y)^2}$, where we have denoted by c^2 the value of that constant, and on that surface, the 1-form α^1 becomes

$$\begin{aligned}\mathrm{d}y-\frac{\mathrm{d}x}{1\pm\sqrt{c^2-(x-y)^2}}&=\frac{\pm\sqrt{c^2-(x-y)^2}}{1\pm\sqrt{c^2-(x-y)^2}}\left[\mathrm{d}y\mp\frac{\mathrm{d}(x-y)}{\sqrt{c^2-(x-y)^2}}\right]\\&=\frac{\pm\sqrt{c^2-(x-y)^2}}{1\pm\sqrt{c^2-(x-y)^2}}\,\mathrm{d}\left[y\mp\arcsin\left(\frac{x-y}{c}\right)\right].\end{aligned}$$

Hence, $y\mp\arcsin(\frac{x-y}{c})$ is another first integral of the equation, and therefore the solution is implicitly given by $y\mp\arcsin(\frac{x-y}{c})=\text{const}$.

The set of 1-forms considered in this example is the one already studied in Example 4.15. One can verify that, by eliminating z from (4.41), one obtains the solution given above.

An alternative procedure, applicable in the case where the symmetry is a Lie point symmetry, consists of using the rectification lemma in order to find a new coordinate system. This coordinate system frequently is denoted by (r,s), instead of (x,y), and it is such that the vector field $\xi\,\partial/\partial x+\eta\,\partial/\partial y$ takes the form $\partial/\partial r$ (which amounts to say that, in the new coordinates, ξ is equal to 1 and η is equal to 0). In that manner, from (4.48) one finds that ζ becomes equal to 0, while (4.49) reduces to $F_r=0$. As is well known, when F does not depend on one of the variables, the order of the equation can be reduced.

Example 4.17 The vector field $\mathbf{X}=\partial/\partial x+\partial/\partial y$ employed in Example 4.16 corresponds to a Lie point symmetry [i.e., it is of the form (4.47) with ζ given by (4.48)] and a coordinate system adapted to $\mathbf{X}$ is (r,s,w) with

$$r=x,\qquad s=x-y,\qquad w=z$$

(in the sense that $\mathbf{X}x=1$, $\mathbf{X}(x-y)=0$, and $\mathbf{X}z=0$; hence, in the new coordinate system, $\mathbf{X}=\partial/\partial r$). In terms of these coordinates, the ODE $y''=(x-y)y'^3$ takes the form $\mathrm{d}^2s/\mathrm{d}^2r=s(\mathrm{d}s/\mathrm{d}r-1)^3$, which does not contain the variable r. Hence, using the standard procedures, this last equation can be transformed into a first-order ODE, and finally one obtains the solution given above.

Chapter 5
Connections

5.1 Covariant Differentiation

The tangent space, T_xM, to a differentiable manifold M at a point x is a vector space different from the tangent space to M at any other point y, T_yM. In general, there is no natural way of relating T_xM with T_yM if $x \neq y$. This means that if v and w are two tangent vectors to M at two different points, e.g., $v \in T_xM$ and $w \in T_yM$, there is no natural way to compare or to combine them. However, in many cases it will be possible to define the parallel transport of a tangent vector from one point to another point of the manifold along a curve. Once this concept has been defined, it will be possible to determine the directional derivatives of any vector field on M; conversely, if we know the directional derivatives of an arbitrary vector field, the parallel transport of a vector along any curve in M is determined.

A *connection*, ∇, on M, is a rule to calculate the directional derivatives of the vector fields on M. If $\mathbf{X}$ and $\mathbf{Y}$ are two vector fields, $\nabla_{\mathbf{X}}\mathbf{Y}$ denotes the vector field whose value at each point $x \in M$ is equal to the directional derivative of $\mathbf{Y}$ in the direction of $\mathbf{X}_x$. In the following definition we copy the properties of the directional derivative of vector fields in $\mathbb{R}^n$.

Definition 5.1 Let M be a differentiable manifold. A connection on M assigns to each $\mathbf{X} \in \mathfrak{X}(M)$ an operator $\nabla_{\mathbf{X}}$ from $\mathfrak{X}(M)$ into itself, such that for all $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(M)$, $a, b \in \mathbb{R}$, and $f \in C^\infty(M)$,

$$\begin{aligned}
\nabla_{\mathbf{X}}(a\mathbf{Y} + b\mathbf{Z}) &= a\nabla_{\mathbf{X}}\mathbf{Y} + b\nabla_{\mathbf{X}}\mathbf{Z}, \\
\nabla_{\mathbf{X}}(f\mathbf{Y}) &= f\,\nabla_{\mathbf{X}}\mathbf{Y} + (\mathbf{X}f)\mathbf{Y}, \\
\nabla_{a\mathbf{X}+b\mathbf{Y}}\mathbf{Z} &= a\nabla_{\mathbf{X}}\mathbf{Z} + b\nabla_{\mathbf{Y}}\mathbf{Z}, \\
\nabla_{f\mathbf{X}}\mathbf{Y} &= f\,\nabla_{\mathbf{X}}\mathbf{Y}.
\end{aligned}$$

The vector field $\nabla_{\mathbf{X}}\mathbf{Y}$ is called the *covariant derivative* of $\mathbf{Y}$ with respect to $\mathbf{X}$.

If $(x^1, x^2, \ldots, x^n)$ is a coordinate system in some neighborhood, U, of M, any pair of vector fields $\mathbf{X}, \mathbf{Y}$ can be expressed in the form $\mathbf{X} = X^i(\partial/\partial x^i)$,

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_5,

$\mathbf{Y} = Y^j(\partial/\partial x^j)$. Using the properties established in the definition it follows that

$$\begin{aligned}\nabla_{\mathbf{X}}\mathbf{Y} &= \nabla_{X^i(\partial/\partial x^i)}\left(Y^j \frac{\partial}{\partial x^j}\right) \\ &= X^i \nabla_{\partial/\partial x^i}\left(Y^j \frac{\partial}{\partial x^j}\right) \\ &= X^i\left[\left(\frac{\partial Y^j}{\partial x^i}\right)\frac{\partial}{\partial x^j} + Y^j \nabla_{\partial/\partial x^i}\frac{\partial}{\partial x^j}\right].\end{aligned}$$

The covariant derivatives $\nabla_{\partial/\partial x^i}(\partial/\partial x^j)$ must be differentiable vector fields, which implies the existence of n^3 differentiable real-valued functions on U, Γ^k_{ji}, such that

$$\nabla_{\partial/\partial x^i}\frac{\partial}{\partial x^j} \equiv \Gamma^k_{ji}\frac{\partial}{\partial x^k}. \tag{5.1}$$

This set of functions characterizes the connection ∇ in the coordinate system chosen, since

$$\begin{aligned}\nabla_{\mathbf{X}}\mathbf{Y} &= X^i\left[\left(\frac{\partial Y^j}{\partial x^i}\right)\frac{\partial}{\partial x^j} + Y^j \Gamma^k_{ji}\frac{\partial}{\partial x^k}\right] \\ &= X^i\left(\frac{\partial Y^k}{\partial x^i} + \Gamma^k_{ji}Y^j\right)\frac{\partial}{\partial x^k}.\end{aligned} \tag{5.2}$$

This formula shows that in order to calculate $(\nabla_{\mathbf{X}}\mathbf{Y})_x$, the value of $\nabla_{\mathbf{X}}\mathbf{Y}$ at a point $x \in M$, we only need to know the value of $\mathbf{X}$ at that point (since only the components of $\mathbf{X}$ appear in (5.2), but not their partial derivatives) and the values of $\mathbf{Y}$ in a neighborhood of x at the points of some curve to which $\mathbf{X}_x$ is tangent (since the partial derivatives of the components Y^k only appear in the combination $X^i(x)(\partial/\partial x^i)_x Y^k = \mathbf{X}_x[Y^k]$). Hence it makes sense to define the covariant derivative of a vector field $\mathbf{Y}$ with respect to a tangent vector $v_x \in T_x M$ as the value at x of the covariant derivative of $\mathbf{Y}$ with respect to a vector field $\mathbf{X}$ such that $\mathbf{X}_x = v_x$. The expressions $\partial Y^k/\partial x^i + \Gamma^k_{ji}Y^j$, appearing in (5.2), are the components of a tensor field [of type $\binom{1}{1}$] traditionally denoted by $Y^k{}_{;i}$ and also by $\nabla_i Y^k$.

Exercise 5.2 Show that if $(x^1, \ldots, x^n)$ and $(x'^1, \ldots, x'^n)$ are two systems of coordinates on M, then the relation

$$\Gamma'^p_{sr} = \frac{\partial x^i}{\partial x'^r}\frac{\partial x^j}{\partial x'^s}\frac{\partial x'^p}{\partial x^k}\Gamma^k_{ji} + \frac{\partial x'^p}{\partial x^k}\frac{\partial^2 x^k}{\partial x'^r \partial x'^s}, \tag{5.3}$$

holds in the intersection of the domains of the two charts.

If a given manifold M can be covered by a single coordinate system $(x^1, \ldots, x^n)$ (as in the case of $\mathbb{R}^n$ with its natural coordinates), a connection can be defined by simply choosing n^3 arbitrary differentiable functions, Γ^i_{jk}, by means of (5.1) (see Examples 5.4 and 5.5), but if $\{(U_i, \phi_i)\}$ is a subatlas of M with more than one coordinate chart, the functions Γ^i_{jk} for each chart have to be related according to (5.3). As shown in the following chapter and in Appendix B, when a manifold has a metric tensor or the structure of a Lie group, there exists a naturally induced connection on the manifold.

Parallel Transport Let $C : I \to M$ be a differentiable curve. If $\mathbf{Y}$ is a vector field defined on the image of C, then its covariant derivative along C, $\nabla_{C'}\mathbf{Y}$, is the vector field on C such that $(\nabla_{C'}\mathbf{Y})_{C(t)} = \nabla_{C'_t}\mathbf{Y}$ for $t \in I$.

Definition 5.3 A vector field is *parallel* (to itself) *along* C if $\nabla_{C'}\mathbf{Y} = 0$ and a curve C is a *geodesic* if $\nabla_{C'}C' = 0$.

Since

$$C'_t = \frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\left(\frac{\partial}{\partial x^i}\right)_{C(t)}$$

[see (1.20)], making use of (5.2) we have

$$\begin{aligned}\nabla_{C'(t)}\mathbf{Y} &= \frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\left(\frac{\partial Y^k}{\partial x^i} + \Gamma^k_{ji}Y^j\right)\bigg|_{C(t)}\left(\frac{\partial}{\partial x^k}\right)_{C(t)} \\ &= \left(C'_t\big[Y^k\big] + \frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\Gamma^k_{ji}Y^j\right)\bigg|_{C(t)}\left(\frac{\partial}{\partial x^k}\right)_{C(t)};\end{aligned}$$

hence, $\mathbf{Y}$ is parallel along C if and only if its components satisfy the system of ODEs

$$\frac{\mathrm{d}(Y^k \circ C)}{\mathrm{d}t} + \frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\big(\Gamma^k_{ji} \circ C\big)\big(Y^j \circ C\big) = 0. \tag{5.4}$$

For a given a curve $C : I \to M$, these equations for $Y^k \circ C$ are linear; therefore there exists a unique solution defined on I for any initial condition $\mathbf{Y}(C(t_0))$ (see Fig. 5.1). Furthermore, the map $P_{t,t_0} : T_{C(t_0)}M \to T_{C(t)}M$, defined by $P_{t,t_0}(\mathbf{Y}_0) = \mathbf{Y}(C(t))$, where $\mathbf{Y}$ is parallel along C and $\mathbf{Y}(C(t_0)) = \mathbf{Y}_0$, is an isomorphism (Hochstadt 1964, Sect. 2.8) called *parallel transport* along of C from $C(t_0)$ to $C(t)$.

Example 5.4 Consider the connection on $\mathbb{R}^2$ defined by $\Gamma^1_{12} = 1$ and the other Γ^i_{jk} equal to zero, with respect to the basis associated with the natural coordinates $(x^1, x^2) = (x, y)$. The equations for the parallel transport of a vector field (5.4) are

$$\frac{\mathrm{d}Y^1}{\mathrm{d}t} + \frac{\mathrm{d}y}{\mathrm{d}t}Y^1 = 0, \qquad \frac{\mathrm{d}Y^2}{\mathrm{d}t} = 0,$$

where, by abuse of notation, we have written Y^1, Y^2, and y, in place of $Y^1 \circ C$, $Y^2 \circ C$, and $y \circ C$, respectively. From the second of these equations it follows that

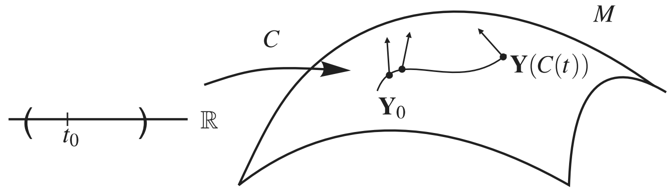

Fig. 5.1 The tangent vector $\mathbf{Y}_0$ is transported along the curve C

Y^2 is constant along any curve, whereas the first equation implies that $Y^1 \mathrm{e}^y$ is a constant; hence, under the parallel transport of a vector field $\mathbf{Y}$ along a curve C from $C(t_0)$ to $C(t)$, the components of $\mathbf{Y}$ with respect to the natural basis $\{\partial/\partial x^i\}$ are related by means of

$$\begin{pmatrix} Y^1(C(t)) \\ Y^2(C(t)) \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{y(C(t))-y(C(t_0))} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} Y^1(C(t_0)) \\ Y^2(C(t_0)) \end{pmatrix}.$$

The 2×2 matrix appearing in this last relation represents the isomorphism P_{t,t_0} mentioned above. In this example, the vector obtained by means of the parallel transport depends on the coordinates of the end-points $C(t_0)$ and $C(t)$, but not on the intermediate points. This is equivalent to the fact that under the parallel transport of an arbitrary vector along any closed curve one obtains the vector originally given at the initial point of the curve. (As we shall see, this corresponds to the fact that the curvature, defined in the following section, of the connection considered in this example is equal to zero.) (See Example 5.18.)

Example 5.5 Let us consider now the connection on $M = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ given by $\Gamma^1_{12} = \Gamma^1_{21} = \Gamma^2_{22} = -1/y = -\Gamma^2_{11}$, with the other Γ^i_{jk} being equal to zero. Equations (5.4) read

$$\frac{\mathrm{d}Y^1}{\mathrm{d}t} - \frac{1}{y}\left(\frac{\mathrm{d}x}{\mathrm{d}t}Y^2 + \frac{\mathrm{d}y}{\mathrm{d}t}Y^1\right) = 0, \qquad \frac{\mathrm{d}Y^2}{\mathrm{d}t} + \frac{1}{y}\left(\frac{\mathrm{d}x}{\mathrm{d}t}Y^1 - \frac{\mathrm{d}y}{\mathrm{d}t}Y^2\right) = 0. \quad (5.5)$$

This system can readily be solved employing the complex combination $Y^1 + \mathrm{i}Y^2$, in terms of which we have

$$\frac{\mathrm{d}(Y^1 + \mathrm{i}Y^2)}{\mathrm{d}t} = \frac{1}{y}\left(\frac{\mathrm{d}y}{\mathrm{d}t} - \mathrm{i}\frac{\mathrm{d}x}{\mathrm{d}t}\right)(Y^1 + \mathrm{i}Y^2),$$

and therefore

$$\left(\frac{Y^1 + \mathrm{i}Y^2}{y}\right)(C(t)) = \left(\frac{Y^1 + \mathrm{i}Y^2}{y}\right)(C(t_0)) \exp\left(-\mathrm{i}\int_{t_0}^{t} \frac{1}{y(t)}\frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{d}t\right),$$

which means that the isomorphism P_{t,t_0} defined above is represented by

$$\begin{pmatrix} Y^1(C(t)) \\ Y^2(C(t)) \end{pmatrix} = \frac{y(C(t))}{y(C(t_0))} \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix} \begin{pmatrix} Y^1(C(t_0)) \\ Y^2(C(t_0)) \end{pmatrix} \quad (5.6)$$

with

$$\Theta \equiv -\int_{t_0}^{t} \frac{1}{y(C(t))} \frac{\mathrm{d}(x \circ C)}{\mathrm{d}t} \, \mathrm{d}t.$$

(It should be clear that the use of complex variables is not essential, but only a convenience; one can verify directly that (5.6) is the solution of the system (5.5).)

Since Θ is the line integral of $y^{-1}\,\mathrm{d}x$, which is not an exact 1-form, Θ not only depends on the end-points of the curve, but on the entire (image of the) curve itself. This fact is equivalent to that after the parallel transport of a vector along a closed curve, the final vector may not coincide with the original one (see Example 1.28). Indeed, if C is a simple closed curve, using Green's theorem one finds that the angle Θ can also be expressed as the surface integral $-\iint_D \frac{\mathrm{d}x\,\mathrm{d}y}{y^2}$, where D is the region enclosed by C. For a closed curve, (5.6) reduces to

$$\begin{pmatrix} Y^1(C(t_0)) \\ Y^2(C(t_0)) \end{pmatrix}_{\text{final}} = \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix} \begin{pmatrix} Y^1(C(t_0)) \\ Y^2(C(t_0)) \end{pmatrix}_{\text{initial}},$$

so that the only effect of the parallel transport is similar to that of a rotation in the plane through the angle Θ (in this example, as in the rest of this chapter, we are not assuming the existence of a structure that allows us to define lengths of vectors or the angle between vectors). (Cf. Example 6.29.)

A *geodesic* C is a curve whose tangent vector field, C', is parallel along C. Hence, from (5.4), with $Y^j = \mathrm{d}(x^j \circ C)/\mathrm{d}t$ we obtain the *geodesic equations*

$$\frac{\mathrm{d}^2(x^k \circ C)}{\mathrm{d}t^2} + \left(\Gamma^k_{ji} \circ C\right) \frac{\mathrm{d}(x^j \circ C)}{\mathrm{d}t} \frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t} = 0. \tag{5.7}$$

By contrast with the equations for the parallel transport of a vector field (5.4) along a given curve, which are first-order linear equations for $Y^i \circ C$, the equations for the geodesics (5.7) are second-order equations for the functions $x^i \circ C$, which regularly are nonlinear.

Example 5.6 Considering the connection locally defined by

$$\Gamma^1_{11} = -\frac{2r}{1+r^2}, \qquad \Gamma^1_{22} = \frac{r(r^2-1)}{1+r^2}, \qquad \Gamma^2_{12} = \Gamma^2_{21} = \frac{1-r^2}{r(1+r^2)},$$

with all the other functions Γ^i_{jk} being equal to zero, with respect to the basis induced by the polar coordinates $(r,\theta) = (x^1, x^2)$ of the Euclidean plane, equations (5.7) take the form

$$\begin{aligned} \frac{\mathrm{d}^2 r}{\mathrm{d}t^2} - \frac{2r}{1+r^2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + \frac{r(r^2-1)}{1+r^2}\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 &= 0, \\ \frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \frac{2(1-r^2)}{r(1+r^2)} \frac{\mathrm{d}r}{\mathrm{d}t} \frac{\mathrm{d}\theta}{\mathrm{d}t} &= 0. \end{aligned} \tag{5.8}$$

The second of these equations amounts to $\frac{\mathrm{d}}{\mathrm{d}t}[\frac{r^2}{(1+r^2)^2}\frac{\mathrm{d}\theta}{\mathrm{d}t}] = 0$; therefore

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = L\frac{(1+r^2)^2}{r^2}, \tag{5.9}$$

where L is a constant. If $L = 0$, then θ is constant, and the first equation (5.8) reduces to

$$(1+r^2)\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{1+r^2}\frac{\mathrm{d}r}{\mathrm{d}t}\right) = 0.$$

Hence $\frac{1}{1+r^2}\frac{\mathrm{d}r}{\mathrm{d}t} = c$, where c is another constant and, therefore, $r = \tan c(t - t_0)$, which means that the (images of the) geodesics with $L = 0$ are straight lines passing through the origin.

When $L \neq 0$, substituting (5.9) into the first equation (5.8), we have

$$\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} - \frac{2r}{1+r^2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + L^2\frac{(r^2-1)(1+r^2)^3}{r^3} = 0.$$

Multiplying the previous equation by $(1+r^2)^{-2}\mathrm{d}r/\mathrm{d}t$, the result can be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{2}\frac{1}{(1+r^2)^2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + \frac{L^2(1+r^2)^2}{2r^2}\right] = 0.$$

Thus we have

$$\frac{1}{2}\frac{1}{(1+r^2)^2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + \frac{L^2(1+r^2)^2}{2r^2} = E, \tag{5.10}$$

where E is a constant. Equation (5.10) is an equation of separable variables that determines $r \circ C$, which substituted into (5.9) leads to $\theta \circ C$.

The image of C can be obtained by combining (5.9) and (5.10), which yields

$$\mathrm{d}\theta = \pm\frac{(1+r^2)\,\mathrm{d}r}{r^2\sqrt{\frac{2E}{L^2} - \frac{(1+r^2)^2}{r^2}}} = \pm\frac{(1+r^2)\,\mathrm{d}r}{r^2\sqrt{\frac{2E}{L^2} - 4 - \frac{(1-r^2)^2}{r^2}}}$$

and with the change of variable $\frac{1-r^2}{r} = \sqrt{\frac{2E}{L^2} - 4}\,\cos v$, this equation reduces to $\mathrm{d}\theta = \pm\mathrm{d}v$; hence,

$$1 - r^2 = \sqrt{\frac{2E}{L^2} - 4}\,r\cos(\theta - \theta_0)$$

or, in terms of the Cartesian coordinates,

$$\left(x+\sqrt{\frac{E}{2L^2}-1}\cos\theta_0\right)^2+\left(y+\sqrt{\frac{E}{2L^2}-1}\sin\theta_0\right)^2=\frac{E}{2L^2},$$

which corresponds to a circle enclosing the origin.

Exercise 5.7 Considering the connection given in Example 5.4, show that the geodesic starting at the point (x_0, y_0), with the initial velocity $a(\partial/\partial x)_{(x_0,y_0)} + b(\partial/\partial y)_{(x_0,y_0)}$, is given by $x = x_0 + a(1 - \mathrm{e}^{-bt})/b$, $y = y_0 + bt$.

Covariant Derivative of Tensor Fields The covariant derivative of a tensor field of type $\binom{0}{k}$, t, with respect to a vector field $\mathbf{X}$, denoted by $\nabla_{\mathbf{X}} t$, is defined by the relation

$$\begin{aligned}\mathbf{X}\big(t(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)\big) &= (\nabla_{\mathbf{X}} t)(\mathbf{Y}_1,\ldots,\mathbf{Y}_k)\\ &\quad+\sum_{i=1}^{k} t(\mathbf{Y}_1,\ldots,\mathbf{Y}_{i-1},\nabla_{\mathbf{X}}\mathbf{Y}_i,\mathbf{Y}_{i+1},\ldots,\mathbf{Y}_k),\end{aligned} \tag{5.11}$$

for $\mathbf{X}, \mathbf{Y}_1, \ldots, \mathbf{Y}_k \in \mathfrak{X}(M)$ [cf. (2.45)]. The covariant derivative of t with respect to $\mathbf{X}$ is also a tensor field of type $\binom{0}{k}$ since (see Sect. 1.4)

$$\begin{aligned}&(\nabla_{\mathbf{X}} t)(\mathbf{Y}_1,\ldots,f\mathbf{Y}_i,\ldots,\mathbf{Y}_k)\\ &\quad= \mathbf{X}(t(\mathbf{Y}_1,\ldots,f\mathbf{Y}_i,\ldots,\mathbf{Y}_k))\\ &\qquad-\sum_{\substack{j=1\\ j\neq i}}^{k} t(\mathbf{Y}_1,\ldots,f\mathbf{Y}_i,\ldots,\nabla_{\mathbf{X}}\mathbf{Y}_j,\ldots,\mathbf{Y}_k)\\ &\qquad- t\big(\mathbf{Y}_1,\ldots,\nabla_{\mathbf{X}}(f\mathbf{Y}_i),\ldots,\mathbf{Y}_k\big)\\ &\quad= \mathbf{X}(f\,t(\mathbf{Y}_1,\ldots,\mathbf{Y}_i,\ldots,\mathbf{Y}_k))\\ &\qquad-\sum_{\substack{j=1\\ j\neq i}}^{k} f\,t(\mathbf{Y}_1,\ldots,\mathbf{Y}_i,\ldots,\nabla_{\mathbf{X}}\mathbf{Y}_j,\ldots,\mathbf{Y}_k)\\ &\qquad- t(\mathbf{Y}_1,\ldots,f\,\nabla_{\mathbf{X}}\mathbf{Y}_i+(\mathbf{X}f)\mathbf{Y}_i,\ldots,\mathbf{Y}_k)\\ &\quad= f\,(\nabla_{\mathbf{X}} t)(\mathbf{Y}_1,\ldots,\mathbf{Y}_i,\ldots,\mathbf{Y}_k), \quad \text{for } f\in C^\infty(M).\end{aligned}$$

When $k = 0$, that is, when t is a function of M in $\mathbb{R}$, we define $\nabla_{\mathbf{X}} t \equiv \mathbf{X}t$. From the definition of $\nabla_{\mathbf{X}} t$ it follows that $\nabla_{f\mathbf{X}} t = f\,\nabla_{\mathbf{X}} t$ and that $\nabla_{a\mathbf{X}+b\mathbf{Y}} t = a\nabla_{\mathbf{X}} t + b\nabla_{\mathbf{Y}} t$, for $a, b \in \mathbb{R}$ and $f \in C^\infty(M)$.

Exercise 5.8 Show that

$$\begin{aligned}\nabla_{\mathbf{X}}(ft) &= f\,\nabla_{\mathbf{X}}t + (\mathbf{X}f)t,\\ \nabla_{\mathbf{X}}(at+bs) &= a\nabla_{\mathbf{X}}t + b\nabla_{\mathbf{X}}s,\\ \nabla_{\mathbf{X}}(t\otimes s) &= (\nabla_{\mathbf{X}}t)\otimes s + t\otimes(\nabla_{\mathbf{X}}s).\end{aligned}$$

If $(x^1,\ldots,x^n)$ is a local coordinate system and $\mathbf{Y} = Y^i(\partial/\partial x^i)$ is an arbitrary vector field, applying the foregoing definition we have

$$\begin{aligned}\big(\nabla_{\partial/\partial x^i}\,\mathrm{d}x^j\big)(\mathbf{Y}) &= \frac{\partial}{\partial x^i}\big(\mathrm{d}x^j(\mathbf{Y})\big) - \mathrm{d}x^j(\nabla_{\partial/\partial x^i}\mathbf{Y})\\ &= \frac{\partial Y^j}{\partial x^i} - \mathrm{d}x^j\left(\left(\frac{\partial Y^k}{\partial x^i} + \Gamma^k_{mi}Y^m\right)\frac{\partial}{\partial x^k}\right)\\ &= \frac{\partial Y^j}{\partial x^i} - \left(\frac{\partial Y^j}{\partial x^i} + \Gamma^j_{mi}Y^m\right)\\ &= -\Gamma^j_{mi}Y^m\\ &= -\Gamma^j_{mi}\,\mathrm{d}x^m(\mathbf{Y}),\end{aligned}$$

that is,

$$\nabla_{\partial/\partial x^i}\,\mathrm{d}x^j = -\Gamma^j_{mi}\,\mathrm{d}x^m. \tag{5.12}$$

Exercise 5.9 Show that if $\mathbf{X} = X^i(\partial/\partial x^i)$ and $t = t_{i\ldots j}\,\mathrm{d}x^i\otimes\cdots\otimes\mathrm{d}x^j$, then

$$\nabla_{\mathbf{X}}t = X^k\left(\frac{\partial t_{i\ldots j}}{\partial x^k} - \Gamma^m_{ik}t_{m\ldots j} - \cdots - \Gamma^m_{jk}t_{i\ldots m}\right)\mathrm{d}x^i\otimes\cdots\otimes\mathrm{d}x^j.$$

(The components $\partial t_{i\ldots j}/\partial x^k - \Gamma^m_{ik}t_{m\ldots j} - \cdots - \Gamma^m_{jk}t_{i\ldots m}$ are denoted by $t_{i\ldots j;k}$ or by $\nabla_k t_{i\ldots j}$.)

Exercise 5.10 Show that $\nabla_{\mathbf{X}}(\mathbf{Y}\lrcorner t) = (\nabla_{\mathbf{X}}\mathbf{Y})\lrcorner t + \mathbf{Y}\lrcorner(\nabla_{\mathbf{X}}t)$ for any tensor field t of type $\binom{0}{k}$ and $\mathbf{X},\mathbf{Y}\in\mathfrak{X}(M)$.

5.2 Torsion and Curvature

The *torsion*, T, of the connection ∇ is the map from $\mathfrak{X}(M)\times\mathfrak{X}(M)$ into $\mathfrak{X}(M)$ given by

$$T(\mathbf{X},\mathbf{Y}) = \nabla_{\mathbf{X}}\mathbf{Y} - \nabla_{\mathbf{Y}}\mathbf{X} - [\mathbf{X},\mathbf{Y}], \quad \text{for } \mathbf{X},\mathbf{Y}\in\mathfrak{X}(M). \tag{5.13}$$

Clearly, T is skew-symmetric, $T(\mathbf{X},\mathbf{Y}) = -T(\mathbf{Y},\mathbf{X})$, and T is a tensor field since, making use of the result of Exercise 1.22, if $f\in C^\infty(M)$ we have

$$\begin{aligned} T(f\mathbf{X},\mathbf{Y}) &= \nabla_{f\mathbf{X}}\mathbf{Y} - \nabla_{\mathbf{Y}}(f\mathbf{X}) - [f\mathbf{X},\mathbf{Y}] \\ &= f\,\nabla_{\mathbf{X}}\mathbf{Y} - f\,\nabla_{\mathbf{Y}}\mathbf{X} - (\mathbf{Y}f)\mathbf{X} - f[\mathbf{X},\mathbf{Y}] + (\mathbf{Y}f)\mathbf{X} \\ &= f\,T(\mathbf{X},\mathbf{Y}). \end{aligned}$$

(The object T defined above does not satisfy the definition of a tensor field given in Sect. 1.4 since $T(\mathbf{X},\mathbf{Y})$ is not a function, but a vector field; however, T is equivalent to a tensor field of type $\binom{1}{2}$, $\tilde{T}$, defined by $\tilde{T}(\mathbf{X},\mathbf{Y},\alpha) \equiv \alpha(T(\mathbf{X},\mathbf{Y}))$, for any pair of vector fields $\mathbf{X},\mathbf{Y}$, and any covector field α.) A connection ∇ is *symmetric*, or *torsion-free*, if its torsion tensor is zero.

Exercise 5.11 Show that if $\mathbf{X} = X^i(\partial/\partial x^i)$ and $\mathbf{Y} = Y^j(\partial/\partial x^j)$ are two arbitrary differentiable vector fields, then $T(\mathbf{X},\mathbf{Y}) = X^iY^jT^k_{ij}(\partial/\partial x^k)$, where $T^k_{ij} = \Gamma^k_{ji} - \Gamma^k_{ij}$. Show that ∇ is symmetric if and only if $\Gamma^k_{ij} = \Gamma^k_{ji}$.

The *curvature* tensor, R, of the connection ∇ is a map that associates to each pair of vector fields an operator from $\mathfrak{X}(M)$ into itself, given by

$$R(\mathbf{X},\mathbf{Y}) = \nabla_{\mathbf{X}}\nabla_{\mathbf{Y}} - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}} - \nabla_{[\mathbf{X},\mathbf{Y}]}, \quad \text{for } \mathbf{X},\mathbf{Y} \in \mathfrak{X}(M). \tag{5.14}$$

It can readily be seen that $R(\mathbf{X},\mathbf{Y}) = -R(\mathbf{Y},\mathbf{X})$, $R(a\mathbf{X}+b\mathbf{Y},\mathbf{Z}) = aR(\mathbf{X},\mathbf{Z}) + bR(\mathbf{Y},\mathbf{Z})$, and $R(\mathbf{X},\mathbf{Y})(a\mathbf{Z}+b\mathbf{W}) = aR(\mathbf{X},\mathbf{Y})\mathbf{Z}+bR(\mathbf{X},\mathbf{Y})\mathbf{W}$. The curvature tensor is indeed a tensor field, since

$$\begin{aligned} R(\mathbf{X},\mathbf{Y})(f\mathbf{Z}) &= \nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}(f\mathbf{Z}) - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}(f\mathbf{Z}) - \nabla_{[\mathbf{X},\mathbf{Y}]}(f\mathbf{Z}) \\ &= f\,\nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{Z} + (\mathbf{X}f)\nabla_{\mathbf{Y}}\mathbf{Z} + (\mathbf{Y}f)\nabla_{\mathbf{X}}\mathbf{Z} + \big(\mathbf{X}(\mathbf{Y}f)\big)\mathbf{Z} \\ &\quad - f\,\nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}\mathbf{Z} - (\mathbf{Y}f)\nabla_{\mathbf{X}}\mathbf{Z} - (\mathbf{X}f)\nabla_{\mathbf{Y}}\mathbf{Z} - \big(\mathbf{Y}(\mathbf{X}f)\big)\mathbf{Z} \\ &\quad - f\,\nabla_{[\mathbf{X},\mathbf{Y}]}\mathbf{Z} - \big([\mathbf{X},\mathbf{Y}]f\big)\mathbf{Z} \\ &= f\,R(\mathbf{X},\mathbf{Y})\mathbf{Z} \end{aligned}$$

and

$$\begin{aligned} R(f\mathbf{X},\mathbf{Y})\mathbf{Z} &= \nabla_{f\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Y}}\nabla_{f\mathbf{X}}\mathbf{Z} - \nabla_{[f\mathbf{X},\mathbf{Y}]}\mathbf{Z} \\ &= f\,\nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{Z} - f\,\nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}\mathbf{Z} - (\mathbf{Y}f)\nabla_{\mathbf{X}}\mathbf{Z} \\ &\quad - f\,\nabla_{[\mathbf{X},\mathbf{Y}]}\mathbf{Z} + (\mathbf{Y}f)\nabla_{\mathbf{X}}\mathbf{Z} \\ &= f\,R(\mathbf{X},\mathbf{Y})\mathbf{Z}, \quad \text{for } \mathbf{X},\mathbf{Y},\mathbf{Z} \in \mathfrak{X}(M),\ f \in C^\infty(M). \end{aligned}$$

(As in the case of the torsion, R does not satisfy the definition of a tensor field given in Sect. 1.4; however, R is equivalent to the tensor field of type $\binom{1}{3}$ $\tilde{R}$ defined by $\tilde{R}(\mathbf{X},\mathbf{Y},\mathbf{Z},\alpha) \equiv \alpha(R(\mathbf{X},\mathbf{Y})\mathbf{Z})$.) A connection ∇ is *flat* if its curvature tensor is zero.

Exercise 5.12 Show that if $\mathbf{X} = X^i(\partial/\partial x^i)$, $\mathbf{Y} = Y^j(\partial/\partial x^j)$, and $\mathbf{Z} = Z^k(\partial/\partial x^k)$, then $R(\mathbf{X}, \mathbf{Y})\mathbf{Z} = X^i Y^j Z^k R^m{}_{kij}(\partial/\partial x^m)$, where

$$R^m{}_{kij} = \frac{\partial \Gamma^m_{kj}}{\partial x^i} - \frac{\partial \Gamma^m_{ki}}{\partial x^j} + \Gamma^m_{pi}\Gamma^p_{kj} - \Gamma^m_{pj}\Gamma^p_{ki}.$$

Exercise 5.13 Show that if $R(\mathbf{X}, \mathbf{Y})t \equiv \nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}t - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}t - \nabla_{[\mathbf{X},\mathbf{Y}]}t$, for any tensor field t of type $\binom{0}{k}$, then

$$\begin{aligned} R(\mathbf{X}, \mathbf{Y})(ft) &= f\,R(\mathbf{X}, \mathbf{Y})t, \\ R(\mathbf{X}, \mathbf{Y})(t \otimes s) &= \big(R(\mathbf{X}, \mathbf{Y})t\big) \otimes s + t \otimes \big(R(\mathbf{X}, \mathbf{Y})s\big). \end{aligned}$$

Exercise 5.14 Show that for $\mathbf{X}, \mathbf{Y}, \mathbf{Z}, \mathbf{W} \in \mathfrak{X}(M)$,

$$\begin{aligned} &R(\mathbf{X}, \mathbf{Y})\mathbf{Z} + R(\mathbf{Z}, \mathbf{X})\mathbf{Y} + R(\mathbf{Y}, \mathbf{Z})\mathbf{X} \\ &\quad = \nabla_{\mathbf{X}}\big(T(\mathbf{Y}, \mathbf{Z})\big) + \nabla_{\mathbf{Y}}\big(T(\mathbf{Z}, \mathbf{X})\big) + \nabla_{\mathbf{Z}}\big(T(\mathbf{X}, \mathbf{Y})\big) \\ &\qquad + T\big(\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]\big) + T\big(\mathbf{Y}, [\mathbf{Z}, \mathbf{X}]\big) + T\big(\mathbf{Z}, [\mathbf{X}, \mathbf{Y}]\big) \end{aligned} \tag{5.15}$$

and

$$\begin{aligned} &\nabla_{\mathbf{X}}\big(R(\mathbf{Y}, \mathbf{Z})\mathbf{W}\big) + \nabla_{\mathbf{Y}}\big(R(\mathbf{Z}, \mathbf{X})\mathbf{W}\big) + \nabla_{\mathbf{Z}}\big(R(\mathbf{X}, \mathbf{Y})\mathbf{W}\big) \\ &\quad = R(\mathbf{Y}, \mathbf{Z})\nabla_{\mathbf{X}}\mathbf{W} + R(\mathbf{Z}, \mathbf{X})\nabla_{\mathbf{Y}}\mathbf{W} + R(\mathbf{X}, \mathbf{Y})\nabla_{\mathbf{Z}}\mathbf{W} \\ &\qquad + R\big([\mathbf{Y}, \mathbf{Z}], \mathbf{X}\big)\mathbf{W} + R\big([\mathbf{Z}, \mathbf{X}], \mathbf{Y}\big)\mathbf{W} + R\big([\mathbf{X}, \mathbf{Y}], \mathbf{Z}\big)\mathbf{W}. \end{aligned} \tag{5.16}$$

The relations (5.16) are known as the *Bianchi identities*.

Parallel Transport in Terms of the Tangent Bundle Each curve $C : I \to M$, where I is an open interval of the real numbers, gives rise to curves in the tangent bundle of M, defined with the aid of the connection of M. Let $p = C(t_0)$, where t_0 is some point of I, and let $v_p \in T_pM$. As discussed in Sect. 5.1, the existence of a connection on M allows us to define an isomorphism $P_{t,t_0} : T_{C(t_0)}M \to T_{C(t)}M$ representing the parallel transport of tangent vectors to M along C. The curve $\overline{C}_{v_p}$ in the tangent bundle of M will be defined by $\overline{C}_{v_p}(t) \equiv P_{t,t_0}(v_p)$, so that $\overline{C}_{v_p}(t) \in T_{C(t)}M$ and, therefore, $\pi \circ \overline{C}_{v_p} = C$, where π is the canonical projection of TM on M. Furthermore, $\overline{C}_{v_p}(t_0) = v_p$ (see Fig. 5.2).

In terms of the coordinates $(q^i, \dot{q}^i)$ on TM, induced by a local coordinate system x^i on M (see Sect. 1.2), we have $q^i(\overline{C}_{v_p}(t)) = x^i(C(t))$ and the functions $\dot{q}^i(\overline{C}_{v_p}(t))$ satisfy [see (5.4) and (1.27)]

$$\frac{\mathrm{d}\dot{q}^k(\overline{C}_{v_p}(t))}{\mathrm{d}t} + \frac{\mathrm{d}q^i(\overline{C}_{v_p}(t))}{\mathrm{d}t}\,\Gamma^k_{ji}\big(C(t)\big)\,\dot{q}^j\big(\overline{C}_{v_p}(t)\big) = 0,$$

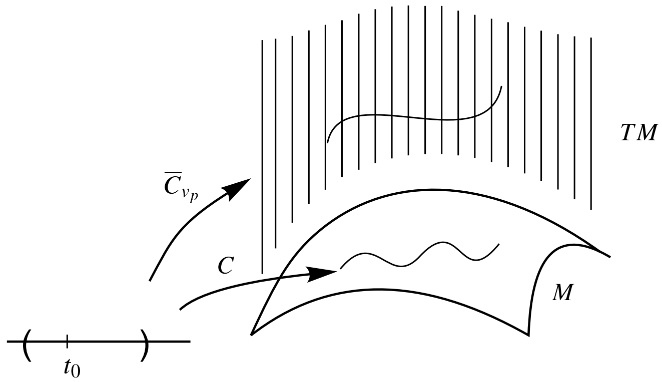

Fig. 5.2 The image of $\overline{C}_{v_p}$ is formed by the tangent vectors obtained by parallel transport of v_p along C

with $\dot{q}^i(\overline{C}_{v_p}(t_0)) = \dot{q}^i(v_p)$. According to the foregoing relations and (1.20), the tangent vector to $\overline{C}_{v_p}$ at $t = t_0$ is

$$\left.\frac{\mathrm{d}(q^i \circ \overline{C}_{v_p})}{\mathrm{d}t}\right|_{t_0} \left(\frac{\partial}{\partial q^i}\right)_{v_p} + \left.\frac{\mathrm{d}(\dot{q}^i \circ \overline{C}_{v_p})}{\mathrm{d}t}\right|_{t_0} \left(\frac{\partial}{\partial \dot{q}^i}\right)_{v_p}$$
$$= \left.\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\right|_{t_0} \left[\left(\frac{\partial}{\partial q^i}\right)_{v_p} - \Gamma^k_{ji}(p)\,\dot{q}^j(v_p)\left(\frac{\partial}{\partial \dot{q}^k}\right)_{v_p}\right].$$

The n real numbers $\mathrm{d}(x^i \circ C)/\mathrm{d}t|_{t=t_0}$, appearing on the right-hand side of the last expression, are the components of the tangent vector of C at $t = t_0$ and do not depend on v_p, while the n tangent vectors to *TM* at v_p,

$$\left(\frac{\partial}{\partial q^i}\right)_{v_p} - \Gamma^k_{ji}(p)\,\dot{q}^j(v_p)\left(\frac{\partial}{\partial \dot{q}^k}\right)_{v_p} \qquad (i = 1, 2, \ldots, n),$$

which do not depend on C, form a basis of an n-dimensional subspace of $T_{v_p}(TM)$, which is called the *horizontal subspace* of $T_{v_p}(TM)$. A curve in *TM* is a *horizontal curve* if at each point of the curve its tangent vector belongs to the horizontal subspace at that point. Thus, a horizontal curve σ in *TM* represents a parallel vector field along the curve $\pi \circ \sigma$ in M. It may be noticed that defining a connection on M is equivalent to defining the horizontal subspace of $T_{v_p}(TM)$ at each point v_p of *TM*. However, if v_p and w_p belong to T_pM, the horizontal subspaces of $T_{v_p}(TM)$ and $T_{w_p}(TM)$ are not independent of each other.

Exercise 5.15 A differentiable curve C in M defines a curve $t \mapsto C'_t$ in *TM*, such that $\pi(C'_t) = C(t)$. Show that C is a geodesic if and only if the curve $t \mapsto C'_t$ is horizontal.

The n vector fields $\mathbf{X}_i$ on TM locally given by

$$\mathbf{X}_i \equiv \frac{\partial}{\partial q^i} - \left(\pi^* \Gamma^k_{ji}\right) \dot{q}^j \frac{\partial}{\partial \dot{q}^k} \tag{5.17}$$

generate an n-dimensional distribution on TM and one readily finds that

$$[\mathbf{X}_i, \mathbf{X}_j] = -\dot{q}^k \left(\pi^* R^m{}_{kij}\right) \frac{\partial}{\partial \dot{q}^m}, \tag{5.18}$$

where the $R^m{}_{kij}$ are the components of the curvature tensor with respect to the basis $\{\partial/\partial x^i\}$ (see Exercise 5.12); hence, according to Frobenius' theorem, the distribution is locally integrable if and only if the curvature vanishes.

When the connection if flat, the integral manifold of the distribution defined by the vector fields (5.17) passing through v_p is formed by all tangent vectors to M obtained by the parallel transport of v_p along some curve in M passing through p.

Example 5.16 In the case of the connection considered in Example 5.4, the vector fields (5.17) are

$$\mathbf{X}_1 = \frac{\partial}{\partial q^1}, \qquad \mathbf{X}_2 = \frac{\partial}{\partial q^2} - \dot{q}^1 \frac{\partial}{\partial \dot{q}^1}.$$

One can readily verify that the Lie bracket of these vector fields is equal to zero, which implies that the connection is flat. One can also verify that the functions $\dot{q}^2$ and $\dot{q}^1 \mathrm{e}^{q^2}$ are two functionally independent solutions to the linear PDEs $\mathbf{X}_i f = 0$, $i = 1, 2$, and, therefore, the integral manifolds of the distribution generated by the horizontal vector fields $\mathbf{X}_i$ are given by

$$\dot{q}^2 = \text{const}, \qquad \dot{q}^1 \mathrm{e}^{q^2} = \text{const}.$$

According to the definition of the coordinates $\dot{q}^i$ [see (1.27)], this means that $\mathbf{Y} = Y^i \partial/\partial x^i$ is a parallel vector field if $Y^2 = \text{const}$, and $Y^1 \mathrm{e}^y = \text{const}$, which agrees with the result found in Example 5.4.

5.3 The Cartan Structural Equations

In order to represent a connection, or any tensor field, we can employ bases not induced by some coordinate system. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a set of differentiable vector fields defined on some open subset U of M such that, at each point $x \in U$, the tangent vectors $(\mathbf{e}_i)_x$ form a basis of $T_x M$, and let the set of 1-forms $\{\theta^1, \ldots, \theta^n\}$ be its dual basis (that is, $\theta^i(\mathbf{e}_j) = \delta^i_j$). If there exists a coordinate system $(x^1, \ldots, x^n)$ such that $\mathbf{e}_i = \partial/\partial x^i$ or, equivalently, $\theta^i = \mathrm{d}x^i$, we will say that the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is *holonomic*. A necessary and sufficient condition for a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ to be locally holonomic is that $[\mathbf{e}_i, \mathbf{e}_j] = 0$ or, equivalently, $\mathrm{d}\theta^i = 0$. As shown in this section and in the following chapters, when M possesses a connection, a metric, the structure of a Lie group, or some other structure, it is convenient to make use of nonholonomic bases adapted to the structure present.

As pointed out in Chap. 1, in some manifolds there are no coordinate systems covering all the points of the manifold (that is the case, e.g., of the circle S^1 and of the ordinary sphere S^2) and therefore, in those manifolds there are no global holonomic bases. For some manifolds, it may even be impossible to find nonholonomic bases defined globally (e.g., the sphere S^2; but in the case of the circle S^1 one can find a nowhere zero differentiable vector field). A manifold M is *parallelizable* if there exists a set of differentiable vector fields such that at every point of M they form a basis for the tangent space to M at that point. (As we shall see in Chap. 7, every Lie group is parallelizable.)

In the rest of this chapter, $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ will represent a local basis for the vector fields, holonomic or not. If ∇ is a connection on M, the *connection forms*, $\Gamma^i{}_j$, with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, are the n^2 1-forms defined by

$$\Gamma^i{}_j(\mathbf{X}) \equiv \theta^i(\nabla_\mathbf{X}\mathbf{e}_j), \tag{5.19}$$

for $\mathbf{X} \in \mathfrak{X}(M)$. From the properties that define a connection it follows that the $\Gamma^i{}_j$ are, in effect, linear differential forms. The definition (5.19) is equivalent to

$$\nabla_\mathbf{X}\mathbf{e}_i = \Gamma^j{}_i(\mathbf{X})\mathbf{e}_j, \tag{5.20}$$

for $\mathbf{X} \in \mathfrak{X}(M)$. Defining the n^3 functions $\Gamma^i{}_{jk}$ by

$$\Gamma^i{}_{jk} \equiv \Gamma^i{}_j(\mathbf{e}_k) \tag{5.21}$$

(i.e., $\Gamma^i{}_j = \Gamma^i{}_{jk}\theta^k$), one finds that (5.19) and (5.20) amount to

$$\nabla_{\mathbf{e}_i}\mathbf{e}_j = \Gamma^k{}_{ji}\mathbf{e}_k, \tag{5.22}$$

which is of the form (5.1), but now we are considering the possibility of dealing with a nonholonomic basis.

Exercise 5.17 Show that $\nabla_\mathbf{X}\theta^i = -\Gamma^i{}_j(\mathbf{X})\theta^j$.

The *torsion 2-forms*, T^i, with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, are defined by

$$T^i(\mathbf{X}, \mathbf{Y}) \equiv \frac{1}{2}\theta^i\big(T(\mathbf{X}, \mathbf{Y})\big). \tag{5.23}$$

Since the torsion is a tensor field satisfying the condition $T(\mathbf{X}, \mathbf{Y}) = -T(\mathbf{Y}, \mathbf{X})$, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{X}(M)$, each T^i is a 2-form and making use of the definitions (5.13), (5.11), (3.30), and (3.7) and the result of Exercise 5.17, we obtain

$$\begin{aligned} T^i(\mathbf{X}, \mathbf{Y}) &= \frac{1}{2}\theta^i\big(T(\mathbf{X}, \mathbf{Y})\big) \\ &= \frac{1}{2}\theta^i\big(\nabla_\mathbf{X}\mathbf{Y} - \nabla_\mathbf{Y}\mathbf{X} - [\mathbf{X}, \mathbf{Y}]\big) \end{aligned}$$

$$= \frac{1}{2}\big\{\mathbf{X}\big(\theta^i(\mathbf{Y})\big) - \big(\nabla_{\mathbf{X}}\theta^i\big)(\mathbf{Y}) - \mathbf{Y}\big(\theta^i(\mathbf{X})\big) + \big(\nabla_{\mathbf{Y}}\theta^i\big)(\mathbf{X}) - \theta^i\big([\mathbf{X},\mathbf{Y}]\big)\big\}$$

$$= \frac{1}{2}\big\{2\,\mathrm{d}\theta^i(\mathbf{X},\mathbf{Y}) + \Gamma^i{}_j(\mathbf{X})\theta^j(\mathbf{Y}) - \Gamma^i{}_j(\mathbf{Y})\theta^j(\mathbf{X})\big\}$$

$$= \big(\mathrm{d}\theta^i + \Gamma^i{}_j \wedge \theta^j\big)(\mathbf{X},\mathbf{Y}),$$

that is,

$$T^i = \mathrm{d}\theta^i + \Gamma^i{}_j \wedge \theta^j. \tag{5.24}$$

These equations are equivalent to the definition of the torsion tensor and are known as the *first Cartan structural equations*.

In a similar manner, defining the *curvature 2-forms*, $\mathscr{R}^i{}_j$, with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$, by

$$\mathscr{R}^i{}_j(\mathbf{X},\mathbf{Y}) \equiv \frac{1}{2}\theta^i\big(R(\mathbf{X},\mathbf{Y})\mathbf{e}_j\big), \tag{5.25}$$

the properties of the curvature tensor imply that each $\mathscr{R}^i{}_j$ is a 2-form and from (5.25), (5.14), (5.20), (5.19), (3.30), and (3.7) one finds that

$$\mathscr{R}^i{}_j(\mathbf{X},\mathbf{Y}) = \frac{1}{2}\theta^i(\nabla_{\mathbf{X}}\nabla_{\mathbf{Y}}\mathbf{e}_j - \nabla_{\mathbf{Y}}\nabla_{\mathbf{X}}\mathbf{e}_j - \nabla_{[\mathbf{X},\mathbf{Y}]}\mathbf{e}_j)$$

$$= \frac{1}{2}\theta^i\big(\nabla_{\mathbf{X}}\big(\Gamma^k{}_j(\mathbf{Y})\mathbf{e}_k\big) - \nabla_{\mathbf{Y}}\big(\Gamma^k{}_j(\mathbf{X})\mathbf{e}_k\big)\big) - \frac{1}{2}\Gamma^i{}_j\big([\mathbf{X},\mathbf{Y}]\big)$$

$$= \frac{1}{2}\big\{\mathbf{X}\big(\Gamma^i{}_j(\mathbf{Y})\big) + \Gamma^i{}_k(\mathbf{X})\,\Gamma^k{}_j(\mathbf{Y})$$

$$- \mathbf{Y}\big(\Gamma^i{}_j(\mathbf{X})\big) - \Gamma^i{}_k(\mathbf{Y})\,\Gamma^k{}_j(\mathbf{X}) - \Gamma^i{}_j\big([\mathbf{X},\mathbf{Y}]\big)\big\}$$

$$= \mathrm{d}\Gamma^i{}_j(\mathbf{X},\mathbf{Y}) + \big(\Gamma^i{}_k \wedge \Gamma^k{}_j\big)(\mathbf{X},\mathbf{Y}),$$

i.e.,

$$\mathscr{R}^i{}_j = \mathrm{d}\Gamma^i{}_j + \Gamma^i{}_k \wedge \Gamma^k{}_j. \tag{5.26}$$

These relations are known as the *second Cartan structural equations*.

If the components of the torsion and the curvature with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ are defined by means of $T(\mathbf{e}_i, \mathbf{e}_j) = T^k_{ij}\mathbf{e}_k$ and $R(\mathbf{e}_i, \mathbf{e}_j)\mathbf{e}_k = R^l{}_{kij}\mathbf{e}_l$, respectively (cf. Exercises 5.11 and 5.12), then the definitions (5.23) and (5.25) amount to

$$T^i = \frac{1}{2}T^i_{jk}\theta^j \wedge \theta^k \quad \text{and} \quad \mathscr{R}^i{}_j = \frac{1}{2}R^i{}_{jkl}\theta^k \wedge \theta^l. \tag{5.27}$$

In the domain of $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, knowing the torsion forms or the curvature forms is equivalent to knowing the torsion tensor or the curvature tensor, respectively. As can be seen in the following examples, the Cartan structural equations constitute a very convenient way of calculating the torsion and the curvature of a connection (further examples can be found in Chaps. 6 and 8, and in Appendix B).

Example 5.18 The connection considered in Example 5.4 corresponds to the connection forms $\Gamma^1{}_1 = \mathrm{d}y$, $\Gamma^1{}_2 = \Gamma^2{}_1 = \Gamma^2{}_2 = 0$, with respect to the holonomic basis $\mathbf{e}_i = \partial/\partial x^i$ (hence $\theta^i = \mathrm{d}x^i$). From (5.24), (3.36), and (5.26) we have

$$T^1 = \mathrm{d}(\mathrm{d}x) + \Gamma^1{}_j \wedge \mathrm{d}x^j = \mathrm{d}y \wedge \mathrm{d}x = -\theta^1 \wedge \theta^2,$$
$$T^2 = \mathrm{d}(\mathrm{d}y) + \Gamma^2{}_j \wedge \mathrm{d}x^j = 0,$$

which shows that the only components of the torsion different from zero are $T^1_{12} = -1 = -T^1_{21}$. On the other hand, $\mathscr{R}^i{}_j = 0$ and, therefore, the connection is flat (cf. Example 5.16).

In a similar way, the connection forms in Example 5.5, with respect to the holonomic basis $\mathbf{e}_i = \partial/\partial x^i$, are $\Gamma^1{}_1 = -y^{-1}\,\mathrm{d}y = \Gamma^2{}_2$, $\Gamma^1{}_2 = -y^{-1}\,\mathrm{d}x = -\Gamma^2{}_1$, so that from the first Cartan structural equations one finds that

$$\begin{aligned}
T^1 &= \mathrm{d}(\mathrm{d}x) + \Gamma^1{}_j \wedge \mathrm{d}x^j \\
&= -y^{-1}\,\mathrm{d}y \wedge \mathrm{d}x - y^{-1}\,\mathrm{d}x \wedge \mathrm{d}y = 0, \\
T^2 &= \mathrm{d}(\mathrm{d}y) + \Gamma^2{}_j \wedge \mathrm{d}x^j \\
&= y^{-1}\mathrm{d}x \wedge \mathrm{d}x - y^{-1}\,\mathrm{d}y \wedge \mathrm{d}y = 0,
\end{aligned}$$

and

$$\begin{aligned}
\mathscr{R}^1{}_1 &= \mathrm{d}\big(-y^{-1}\,\mathrm{d}y\big) + \Gamma^1{}_k \wedge \Gamma^k{}_1 = 0, \\
\mathscr{R}^1{}_2 &= \mathrm{d}\big(-y^{-1}\mathrm{d}x\big) + \Gamma^1{}_k \wedge \Gamma^k{}_2 \\
&= y^{-2}\,\mathrm{d}y \wedge \mathrm{d}x + y^{-2}\,\mathrm{d}y \wedge \mathrm{d}x + y^{-2}\,\mathrm{d}x \wedge \mathrm{d}y \\
&= -y^{-2}\theta^1 \wedge \theta^2, \\
\mathscr{R}^2{}_1 &= \mathrm{d}\big(y^{-1}\,\mathrm{d}x\big) + \Gamma^2{}_k \wedge \Gamma^k{}_1 \\
&= -y^{-2}\,\mathrm{d}y \wedge \mathrm{d}x - y^{-2}\,\mathrm{d}x \wedge \mathrm{d}y - y^{-2}\,\mathrm{d}y \wedge \mathrm{d}x \\
&= y^{-2}\theta^1 \wedge \theta^2, \\
\mathscr{R}^2{}_2 &= \mathrm{d}\big(-y^{-1}\,\mathrm{d}y\big) + \Gamma^2{}_k \wedge \Gamma^k{}_2 = 0.
\end{aligned} \tag{5.28}$$

Comparing (5.28) with (5.27) one finds that the only components different from zero of the curvature tensor are determined by $R^1{}_{212} = -y^{-2} = -R^2{}_{112}$.

Exercise 5.19 Compute the curvature of the connection given in Example 5.4 with the aid of (5.18).

Applying the operator of exterior differentiation, d, to the first Cartan structural equations and making use of the first as well as of the second structural equations we find the identities

$$\mathrm{d}T^i + \Gamma^i{}_j \wedge T^j = \mathscr{R}^i{}_j \wedge \theta^j. \tag{5.29}$$

Hence, if the torsion of the connection is equal to zero,

$$\mathscr{R}^i{}_j \wedge \theta^j = 0. \tag{5.30}$$

Similarly, applying d to the second Cartan structural equations we obtain the identities

$$\mathrm{d}\mathscr{R}^i{}_j = \mathscr{R}^i{}_k \wedge \Gamma^k{}_j - \Gamma^i{}_k \wedge \mathscr{R}^k{}_j. \tag{5.31}$$

Equations (5.29) and (5.31) are equivalent to (5.15) and (5.16), respectively; therefore (5.31) is an expression of the Bianchi identities.

Substituting the second equation (5.27) into (5.30) we obtain $R^i{}_{jkl}\theta^j \wedge \theta^k \wedge \theta^l = 0$, which amounts to the conditions $R^i{}_{[jkl]} = 0$ [see (3.24)] or, using the fact that $R^i{}_{jkl} = -R^i{}_{jlk}$, it follows that when the torsion is equal to zero, the components of the curvature satisfy

$$R^i{}_{jkl} + R^i{}_{klj} + R^i{}_{ljk} = 0. \tag{5.32}$$

The fact that the connection ∇ is flat is equivalent to the *local* existence of n linearly independent vector fields whose covariant derivatives are equal to zero. In effect, if $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ are vector fields such that $\nabla_{\mathbf{X}}\mathbf{Y}_i = 0$ for all $\mathbf{X} \in \mathfrak{X}(M)$, then writing $\mathbf{Y}_i = b_i^j\, \partial/\partial x^j$, from (5.2) it follows that

$$\frac{\partial b_i^j}{\partial x^k} + \Gamma^j_{mk} b_i^m = 0, \tag{5.33}$$

where the Γ^j_{mk} are the components of the connection with respect to the holonomic basis $\partial/\partial x^j$ given by some coordinate system. Applying $\partial/\partial x^l$ to the previous equation and using it again we find that

$$\frac{\partial}{\partial x^l}\frac{\partial b_i^j}{\partial x^k} = -\Gamma^j_{mk}\frac{\partial b_i^m}{\partial x^l} - b_i^m \frac{\partial \Gamma^j_{mk}}{\partial x^l} = \Gamma^j_{mk}\Gamma^m_{rl} b_i^r - b_i^m \frac{\partial \Gamma^j_{mk}}{\partial x^l};$$

therefore, the *integrability conditions* of equations (5.33), given by

$$\frac{\partial}{\partial x^l}\frac{\partial b_i^j}{\partial x^k} = \frac{\partial}{\partial x^k}\frac{\partial b_i^j}{\partial x^l},$$

are $(\partial \Gamma^j_{ml}/\partial x^k - \partial \Gamma^j_{mk}/\partial x^l + \Gamma^j_{rk}\Gamma^r_{ml} - \Gamma^j_{rl}\Gamma^r_{mk})\, b_i^m = 0$, or, simply, $R^j{}_{mkl}\, b_i^m = 0$ (see Exercise 5.12). The vector fields $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_n$ are linearly independent if and only if $\det(b_i^j) \neq 0$, which means that the matrix (b_i^j) has an inverse, so that $R^j{}_{mkl}\, b_i^m = 0$ amounts to $R^j{}_{mkl} = 0$.

Conversely, if $R^i{}_{mkl} = 0$, equations (5.33) are integrable and the integration constants appearing in the solution of this system of equations can be chosen in such a way that $\det(b_i^j)$ is different from zero and, according to the preceding derivation, the n vector fields given by $\mathbf{Y}_i = b_i^j\, \partial/\partial x^j$ are covariantly constant, $\nabla_{\mathbf{X}}\mathbf{Y}_i = 0$.

Thus, the curvature of a connection is equal to zero if and only if there exists locally an invertible matrix (b_i^j) such that

$$\Gamma^i_{jk} = -\tilde{b}^m_j \frac{\partial b^i_m}{\partial x^k}, \tag{5.34}$$

where $(\tilde{b}_i^j)$ denotes the inverse of the matrix (b_i^j) [see (5.33)] or, equivalently (noting that $0 = \partial \delta^i_j / \partial x^k = \partial (b^i_m \tilde{b}^m_j)/\partial x^k = b^i_m \, \partial \tilde{b}^m_j / \partial x^k + \tilde{b}^m_j \, \partial b^i_m / \partial x^k$)

$$\Gamma^i_{jk} = b^i_m \frac{\partial \tilde{b}^m_j}{\partial x^k}. \tag{5.35}$$

In terms of the connection 1-forms $\Gamma^i{}_j = \Gamma^i_{jk} \mathrm{d}x^k$ for the *holonomic* basis $\partial/\partial x^j$, equations (5.33) and (5.35) are equivalent to

$$\mathrm{d}b^j_i = -b^m_i \Gamma^j{}_m, \qquad \Gamma^i{}_j = b^i_m \mathrm{d}\tilde{b}^m_j, \tag{5.36}$$

respectively.

Exercise 5.20 Show that the matrix (b_i^j) is defined by (5.33) up to a multiplicative *constant* $n \times n$ matrix.

Example 5.21 With the aid of (5.26) one readily verifies that the connection 1-forms

$$\Gamma^1{}_1 = \Gamma^2{}_2 = \frac{u\,\mathrm{d}u + v\,\mathrm{d}v}{u^2 + v^2}, \qquad \Gamma^1{}_2 = -\Gamma^2{}_1 = \frac{v\,\mathrm{d}u - u\,\mathrm{d}v}{u^2 + v^2}, \tag{5.37}$$

where u, v is a coordinate system of a manifold M, correspond to a flat connection [without having to specify which are the vector fields appearing in (5.22)]. Assuming that these connection 1-forms correspond to the holonomic basis $\{\partial/\partial u, \partial/\partial v\}$, the components, Y^i, of a covariantly constant vector field $\mathbf{Y} = Y^1\,\partial/\partial u + Y^2\,\partial/\partial v$, are determined by $\mathrm{d}Y^i + Y^j \Gamma^i{}_j = 0$ [see (5.33)], that is,

$$\mathrm{d}Y^1 = -Y^1 \frac{u\,\mathrm{d}u + v\,\mathrm{d}v}{u^2 + v^2} - Y^2 \frac{v\,\mathrm{d}u - u\,\mathrm{d}v}{u^2 + v^2},$$
$$\mathrm{d}Y^2 = Y^1 \frac{v\,\mathrm{d}u - u\,\mathrm{d}v}{u^2 + v^2} - Y^2 \frac{u\,\mathrm{d}u + v\,\mathrm{d}v}{u^2 + v^2}.$$

By combining these equations one obtains

$$Y^1\,\mathrm{d}Y^1 + Y^2\,\mathrm{d}Y^2 = -\big[\big(Y^1\big)^2 + \big(Y^1\big)^2\big] \frac{u\,\mathrm{d}u + v\,\mathrm{d}v}{u^2 + v^2},$$

which implies that

$$\big(Y^1\big)^2 + \big(Y^1\big)^2 = \frac{\text{const}}{u^2 + v^2} \tag{5.38}$$

and

$$Y^2\,\mathrm{d}Y^1 - Y^1\,\mathrm{d}Y^2 = \left[\left(Y^1\right)^2 + \left(Y^1\right)^2\right] \frac{u\,\mathrm{d}v - v\,\mathrm{d}u}{u^2+v^2};$$

hence,

$$\arctan \frac{Y^1}{Y^2} = \arctan \frac{v}{u} + \text{const},$$

which leads to

$$\frac{Y^1}{Y^2} = \frac{v - cu}{cv + u}, \tag{5.39}$$

where c is a constant.

From (5.38) and (5.39) one finds that

$$Y^1 = \frac{K(v - cu)}{u^2+v^2}, \qquad Y^2 = \frac{K(cv + u)}{u^2+v^2},$$

where K is another arbitrary constant and therefore any covariantly constant vector field is a linear combination (with constant coefficients) of the vector fields

$$\frac{1}{u^2+v^2}\left(u\frac{\partial}{\partial u} - v\frac{\partial}{\partial v}\right), \qquad \frac{1}{u^2+v^2}\left(v\frac{\partial}{\partial u} + u\frac{\partial}{\partial v}\right).$$

Hence, as a consequence of the vanishing of the curvature, there exists a basis for the vector fields formed by covariantly constant vector fields.

5.4 Tensor-Valued Forms and Covariant Exterior Derivative

A k-form ω on M is a totally skew-symmetric $C^\infty(M)$-multilinear map of $\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)$ (k times) in $C^\infty(M)$. We can also define differential forms whose values are vector or tensor fields. For instance, the torsion tensor of a connection can be regarded as a 2-form whose values are vector fields, and the curvature tensor as a 2-form whose values are tensor fields of type $\binom{1}{1}$.

Definition 5.22 A *vector-valued* or *tensor-valued differential form of degree k* is a map, ω, from $\mathfrak{X}(M) \times \cdots \times \mathfrak{X}(M)$ (k times) in $\mathfrak{X}(M)$ or in $T^r_s(M)$, respectively, $C^\infty(M)$-multilinear and totally skew-symmetric:

$$\begin{aligned}
&\omega(\mathbf{X}_1, \ldots, f\mathbf{X}_i + g\mathbf{X}'_i, \ldots, \mathbf{X}_k) \\
&\quad = f\,\omega(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k) + g\,\omega(\mathbf{X}_1, \ldots, \mathbf{X}'_i, \ldots, \mathbf{X}_k), \\
&\omega(\mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_k) \\
&\quad = -\omega(\mathbf{X}_1, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_k),
\end{aligned}$$

for $\mathbf{X}_1, \ldots, \mathbf{X}_i, \mathbf{X}'_i, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$, $f, g \in C^\infty(M)$.

The vector-valued or tensor-valued k-forms can be added among themselves, multiplied by real numbers, by functions and by l-forms in the obvious way. The set of vector-valued k-forms will be denoted by $\mathfrak{X}(M) \otimes \Lambda^k(M)$, while the set of the k-forms whose values are tensor fields of type $\binom{r}{s}$ will be denoted by $T^r_s(M) \otimes \Lambda^k(M)$.

For instance, the map Π from $\mathfrak{X}(M)$ into $\mathfrak{X}(M)$ given by $\Pi(\mathbf{X}) = \mathbf{X}$ for $\mathbf{X} \in \mathfrak{X}(M)$ is a vector-valued 1-form as can be seen directly. The maps Θ and Ω defined by $\Theta(\mathbf{X}, \mathbf{Y}) \equiv \frac{1}{2} T(\mathbf{X}, \mathbf{Y})$ and $\Omega(\mathbf{X}, \mathbf{Y}) \equiv \frac{1}{2} R(\mathbf{X}, \mathbf{Y})$, where T is the torsion tensor of a connection ∇ on M and R is the curvature tensor, are 2-forms with values in $\mathfrak{X}(M)$ and in $T^1_1(M)$, respectively. ($\Omega(\mathbf{X}, \mathbf{Y})$ is the tensor field of type $\binom{1}{1}$ given by $\Omega(\mathbf{X}, \mathbf{Y})(\alpha, \mathbf{Z}) \equiv \frac{1}{2}\alpha(R(\mathbf{X}, \mathbf{Y})\mathbf{Z})$.)

If ω is a vector-valued k-form and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a set of independent vector fields, defining ω^i by

$$\omega^i(\mathbf{X}_1, \ldots, \mathbf{X}_k) \equiv \theta^i\big(\omega(\mathbf{X}_1, \ldots, \mathbf{X}_k)\big), \tag{5.40}$$

where $\{\theta^1, \ldots, \theta^n\}$ is the dual basis to $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, we have

$$\omega(\mathbf{X}_1, \ldots, \mathbf{X}_k) = \omega^i(\mathbf{X}_1, \ldots, \mathbf{X}_k)\, \mathbf{e}_i, \tag{5.41}$$

for $\mathbf{X}_1, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$. From the definition it can be seen that each ω^i is a k-form, so that a vector-valued k-form can be represented by n ordinary k-forms.

Definition 5.23 Let $\mathbf{X}$ be a vector field and let η be a k-form; the tensor product of $\mathbf{X}$ times η, denoted by $\mathbf{X} \otimes \eta$, is defined by

$$(\mathbf{X} \otimes \eta)(\mathbf{Y}_1, \ldots, \mathbf{Y}_k) \equiv \eta(\mathbf{Y}_1, \ldots, \mathbf{Y}_k)\mathbf{X}, \quad \text{for } \mathbf{Y}_1, \ldots, \mathbf{Y}_k \in \mathfrak{X}(M).$$

Clearly, $\mathbf{X} \otimes \eta$ is a vector-valued k-form.

Any vector-valued k-form, ω, can be expressed in terms of the k-forms ω^i defined above by means of

$$\omega = \mathbf{e}_i \otimes \omega^i. \tag{5.42}$$

Thus, for the vector-valued 1-form Π defined by $\Pi(\mathbf{X}) = \mathbf{X}$, from (5.40) we have, $\Pi^i(\mathbf{X}) = \theta^i(\Pi(\mathbf{X})) = \theta^i(\mathbf{X})$, and therefore $\Pi^i = \theta^i$ and $\Pi = \mathbf{e}_i \otimes \theta^i$. Similarly, $\Theta^i(\mathbf{X}, \mathbf{Y}) = \theta^i(\frac{1}{2} T(\mathbf{X}, \mathbf{Y})) = T^i(\mathbf{X}, \mathbf{Y})$; therefore, $\Theta^i = T^i$ and $\Theta = \mathbf{e}_i \otimes T^i$, where the T^i are the torsion 2-forms defined in (5.23).

In an analogous way, defining the tensor product of a tensor field t by a k-form η by

$$(t \otimes \eta)(\mathbf{X}_1, \ldots, \mathbf{X}_k) \equiv \eta(\mathbf{X}_1, \ldots, \mathbf{X}_k)t, \tag{5.43}$$

for $\mathbf{X}_1, \ldots, \mathbf{X}_k \in \mathfrak{X}(M)$, it follows that any tensor-valued k-form ω can be expressed in the form

$$\omega = \big(\mathbf{e}_i \otimes \mathbf{e}_j \otimes \cdots \otimes \theta^l \otimes \theta^m \otimes \cdots\big) \otimes \omega^{ij\ldots}_{lm\ldots}, \tag{5.44}$$

with

$$\omega^{ij\dots}_{lm\dots}(\mathbf{X}_1,\dots,\mathbf{X}_k) \equiv \big[\omega(\mathbf{X}_1,\dots,\mathbf{X}_k)\big]\big(\theta^i,\theta^j,\dots,\mathbf{e}_l,\mathbf{e}_m,\dots\big). \tag{5.45}$$

For instance, for the 2-form Ω with values in $T^1_1(M)$ defined above, using (5.25), we have $\Omega^i_j(\mathbf{X},\mathbf{Y}) = \Omega(\mathbf{X},\mathbf{Y})(\theta^i,\mathbf{e}_j) \equiv \frac{1}{2}\theta^i(R(\mathbf{X},\mathbf{Y})\mathbf{e}_j) = \mathscr{R}^i{}_j(\mathbf{X},\mathbf{Y})$; hence, $\Omega = (\mathbf{e}_i \otimes \theta^j)\mathscr{R}^i{}_j$.

The definition of the exterior derivative given in Chap. 3 cannot be applied for a vector-valued or a tensor-valued k-form, since now $\omega(\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\mathbf{X}_{k+1})$ is a vector-field or a tensor field and the expression $\mathbf{X}_i(\omega(\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\mathbf{X}_{k+1}))$ is not defined; in this case we can define the exterior differentiation in the following form.

Definition 5.24 Let M be a differentiable manifold with a connection ∇. If ω is a vector-valued or a tensor-valued k-form on M, its *covariant exterior derivative*, $\mathrm{D}\omega$, is given by

$$\begin{aligned}
&(k+1)\,\mathrm{D}\omega(\mathbf{X}_1,\dots,\mathbf{X}_{k+1})\\
&\quad= \sum_{i=1}^{k+1}(-1)^{i+1}\nabla_{\mathbf{X}_i}\big(\omega(\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\mathbf{X}_{k+1})\big)\\
&\qquad+\sum_{i<j}(-1)^{i+j}\omega\big([\mathbf{X}_i,\mathbf{X}_j],\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\widehat{\mathbf{X}_j},\dots,\mathbf{X}_{k+1}\big),
\end{aligned}$$

for $\mathbf{X}_1,\dots,\mathbf{X}_{k+1} \in \mathfrak{X}(M)$.

It can readily be seen that $\mathrm{D}\omega$ is totally skew-symmetric and that its values are of the same type as those of ω. The proof that $\mathrm{D}\omega$ is a $(k+1)$-form is completely analogous to that given for the exterior derivative of an ordinary differential form in Chap. 3. Clearly, $\mathrm{D}(a\omega_1 + b\omega_2) = a\,\mathrm{D}\omega_1 + b\,\mathrm{D}\omega_2$ for $a, b \in \mathbb{R}$.

If t is a vector or tensor field $\eta \in \Lambda^k(M)$, applying the definition above we have [see Exercise 5.8 and (3.28)]

$$\begin{aligned}
&(k+1)\mathrm{D}(t\otimes\eta)(\mathbf{X}_1,\dots,\mathbf{X}_{k+1})\\
&\quad= \sum_{i=1}^{k+1}(-1)^{i+1}\nabla_{\mathbf{X}_i}\big(\eta(\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\mathbf{X}_{k+1})t\big)\\
&\qquad+\sum_{i<j}(-1)^{i+j}\eta\big([\mathbf{X}_i,\mathbf{X}_j],\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\widehat{\mathbf{X}_j},\dots,\mathbf{X}_{k+1}\big)t\\
&\quad= \sum_{i=1}^{k+1}(-1)^{i+1}\big[\eta(\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\mathbf{X}_{k+1})\nabla_{\mathbf{X}_i}t\\
&\qquad+\mathbf{X}_i\big(\eta(\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\mathbf{X}_{k+1})\big)t\big]\\
&\qquad+\sum_{i<j}(-1)^{i+j}\eta\big([\mathbf{X}_i,\mathbf{X}_j],\mathbf{X}_1,\dots,\widehat{\mathbf{X}_i},\dots,\widehat{\mathbf{X}_j},\dots,\mathbf{X}_{k+1}\big)t
\end{aligned}$$

$$
\begin{aligned}
&= (k+1)\,\mathrm{d}\eta(\mathbf{X}_1,\ldots,\mathbf{X}_{k+1})\,t \\
&\quad + \sum_{i=1}^{k+1}(-1)^{i+1}\eta(\mathbf{X}_1,\ldots,\widehat{\mathbf{X}}_i,\ldots,\mathbf{X}_{k+1})\nabla_{\mathbf{X}_i}t.
\end{aligned}
$$

Using the identity $\mathbf{X}_i = \theta^j(\mathbf{X}_i)\mathbf{e}_j$, it follows that $\nabla_{\mathbf{X}_i}t = \theta^j(\mathbf{X}_i)\nabla_{\mathbf{e}_j}t$ and therefore

$$
\begin{aligned}
&\sum_{i=1}^{k+1}(-1)^{i+1}\eta(\mathbf{X}_1,\ldots,\widehat{\mathbf{X}}_i,\ldots,\mathbf{X}_{k+1})\nabla_{\mathbf{X}_i}t \\
&\quad = \sum_{i=1}^{k+1}(-1)^{i+1}\theta^j(\mathbf{X}_i)\,\eta\,(\mathbf{X}_1,\ldots,\widehat{\mathbf{X}}_i,\ldots,\mathbf{X}_{k+1})\nabla_{\mathbf{e}_j}t \\
&\quad = (k+1)\big(\theta^j\wedge\eta\big)(\mathbf{X}_1,\ldots,\mathbf{X}_{k+1})\nabla_{\mathbf{e}_j}t;
\end{aligned}
$$

hence,

$$
\mathrm{D}(t\otimes\eta) = t\otimes\mathrm{d}\eta + \nabla_{\mathbf{e}_i}t\otimes\big(\theta^i\wedge\eta\big). \tag{5.46}
$$

Applying this result, using the first and second Cartan structural equations, (5.24), and (5.20), we find that

$$
\begin{aligned}
\mathrm{D}\varPi &= \mathrm{D}\big(\mathbf{e}_i\otimes\theta^i\big) \\
&= \mathbf{e}_i\otimes\mathrm{d}\theta^i + \nabla_{\mathbf{e}_j}\mathbf{e}_i\otimes\big(\theta^j\wedge\theta^i\big) \\
&= \mathbf{e}_i\otimes\big(\theta^m\wedge\varGamma^i{}_m + T^i\big) + \varGamma^m{}_i(\mathbf{e}_j)\mathbf{e}_m\otimes\big(\theta^j\wedge\theta^i\big) \\
&= \mathbf{e}_i\otimes\big(\theta^m\wedge\varGamma^i{}_m + T^i\big) + \mathbf{e}_m\otimes\big(\varGamma^m{}_i\wedge\theta^i\big) \\
&= \mathbf{e}_i\otimes T^i,
\end{aligned}
$$

i.e.,

$$
\mathrm{D}\varPi = \varTheta.
$$

In a similar way one finds that the Bianchi identities amount to

$$
\mathrm{D}\varOmega = 0.
$$

By contrast with the usual exterior differentiation, if the connection is not flat, $\mathrm{D}^2 \neq 0$. In effect, making use of (5.46), (3.36), (3.35), (3.7), (5.24), and (5.14), we find that

$$
\begin{aligned}
\mathrm{D}^2(t\otimes\eta) &= \mathrm{D}\big[t\otimes\mathrm{d}\eta + \nabla_{\mathbf{e}_i}t\otimes\big(\theta^i\wedge\eta\big)\big] \\
&= t\otimes\mathrm{d}\,\mathrm{d}\eta + \nabla_{\mathbf{e}_i}t\otimes\big(\theta^i\wedge\mathrm{d}\eta\big) + \nabla_{\mathbf{e}_i}t\otimes\mathrm{d}\big(\theta^i\wedge\eta\big) \\
&\quad + \nabla_{\mathbf{e}_j}\nabla_{\mathbf{e}_i}t\otimes\big(\theta^j\wedge\theta^i\wedge\eta\big)
\end{aligned}
$$

$$
\begin{aligned}
&= \nabla_{\mathbf{e}_i} t \otimes (\theta^i \wedge \mathrm{d}\eta) + \nabla_{\mathbf{e}_i} t \otimes \big(\mathrm{d}\theta^i \wedge \eta - \theta^i \wedge \mathrm{d}\eta\big) \\
&\quad + \nabla_{\mathbf{e}_j} \nabla_{\mathbf{e}_i} t \otimes \big(\theta^j \wedge \theta^i \wedge \eta\big) \\
&= \nabla_{\mathbf{e}_i} t \otimes \big(\mathrm{d}\theta^i \wedge \eta\big) + \frac{1}{2}(\nabla_{\mathbf{e}_j} \nabla_{\mathbf{e}_i} t - \nabla_{\mathbf{e}_i} \nabla_{\mathbf{e}_j} t) \otimes \big(\theta^j \wedge \theta^i \wedge \eta\big) \\
&= \nabla_{\mathbf{e}_i} t \otimes \big[\big(\theta^j \wedge \Gamma^i{}_j + T^i\big) \wedge \eta\big] \\
&\quad + \frac{1}{2}\big(R(\mathbf{e}_j, \partial_i) t + \nabla_{[\mathbf{e}_j, \mathbf{e}_i]} t\big) \otimes \big(\theta^j \wedge \theta^i \wedge \eta\big),
\end{aligned}
$$

but $[\mathbf{e}_j, \mathbf{e}_i] = \nabla_{\mathbf{e}_j} \mathbf{e}_i - \nabla_{\mathbf{e}_i} \mathbf{e}_j - T(\mathbf{e}_j, \mathbf{e}_i) = [\Gamma^m{}_i(\mathbf{e}_j) - \Gamma^m{}_j(\mathbf{e}_i) - 2T^m(\mathbf{e}_j, \mathbf{e}_i)]\mathbf{e}_m$ [see (5.13), (5.20), and (5.23)]. Therefore, we have

$$
\begin{aligned}
&\frac{1}{2}\nabla_{[\mathbf{e}_j, \mathbf{e}_i]} t \otimes \big(\theta^j \wedge \theta^i \wedge \eta\big) \\
&\quad = \frac{1}{2}\nabla_{\mathbf{e}_m} t \otimes \big(\Gamma^m{}_i \wedge \theta^i \wedge \eta - \theta^j \wedge \Gamma^m{}_j \wedge \eta - 2T^m \wedge \eta\big) \\
&\quad = -\nabla_{\mathbf{e}_m} t \otimes \big[\big(\theta^j \wedge \Gamma^m{}_j + T^m\big) \wedge \eta\big],
\end{aligned}
$$

and, hence

$$
\mathrm{D}^2(t \otimes \eta) = \frac{1}{2} R(\mathbf{e}_i, \mathbf{e}_j) t \otimes \big(\theta^i \wedge \theta^j \wedge \eta\big).
$$

Exercise 5.25 Let ω be a vector-valued k-form given by $\omega = \mathbf{e}_i \otimes \omega^i$. Show that $\mathrm{D}\omega = \mathbf{e}_i \otimes (\mathrm{d}\omega^i + \Gamma^i{}_j \wedge \omega^j)$.

Chapter 6
Riemannian Manifolds

In many cases, the manifolds of interest possess a metric tensor which defines an inner product between tangent vectors at each point of the manifold. Some examples are the submanifolds of an Euclidean space and the space–time, in the context of special or general relativity.

6.1 The Metric Tensor

Definition 6.1 Let M be a differentiable manifold and let g be a symmetric tensor field of type $\binom{0}{2}$ on M, that is, $g_p(v_p, w_p) = g_p(w_p, v_p)$ for $v_p, w_p \in T_p M$. g is *positive definite* if for all $v_p \in T_p M$, we have $g_p(v_p, v_p) \geq 0$, and if $g_p(v_p, v_p) = 0$ implies $v_p = 0$ (that is, $g_p(v_p, v_p) > 0$ for all nonzero $v_p \in T_p M$); the tensor field g is *non-singular* if $g_p(v_p, w_p) = 0$ for all $w_p \in T_p M$ implies that $v_p = 0$.

If g is positive definite, then it is non-singular, for if $g_p(v_p, w_p) = 0$ for all $w_p \in T_p M$, we find, in particular, that $g_p(v_p, v_p) = 0$, which implies that $v_p = 0$.

Definition 6.2 A *Riemannian manifold* is a differentiable manifold M with a non-singular, symmetric differentiable tensor field of type $\binom{0}{2}$, called *the metric tensor* or *metric* of M. When the metric tensor is not positive definite, we also say that the manifold is *pseudo-Riemannian* (or *semi-Riemannian*).

In a Riemannian manifold, M, with a positive definite metric, g_p is an inner product on $T_p M$. The norm or length of a tangent vector $v_p \in T_p M$, $\|v_p\|$, is defined by $\|v_p\| = \sqrt{g_p(v_p, v_p)}$ and the length of a curve $C : [a, b] \to M$ is defined by

$$L_C \equiv \int_a^b \|C'_t\| \, \mathrm{d}t. \tag{6.1}$$

Let M be a Riemannian manifold and let $(x^1, \ldots, x^n)$ be a local coordinate system on M. The metric tensor is given by $g = g_{ij} \, \mathrm{d}x^i \otimes \mathrm{d}x^j$ with

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_6,

$g_{ij} = g(\partial/\partial x^i, \partial/\partial x^j)$. Since g is symmetric, we have $g_{ij} = g(\partial/\partial x^i, \partial/\partial x^j) = g(\partial/\partial x^j, \partial/\partial x^i) = g_{ji}$.

Example 6.3 The standard metric of the n-dimensional Euclidean space, $\mathbb{E}^n$, expressed in terms of Cartesian coordinates, $(x^1, \ldots, x^n)$, is

$$g = \mathrm{d}x^1 \otimes \mathrm{d}x^1 + \mathrm{d}x^2 \otimes \mathrm{d}x^2 + \cdots + \mathrm{d}x^n \otimes \mathrm{d}x^n, \tag{6.2}$$

that is, $(g_{ij}) = \mathrm{diag}(1, 1, \ldots, 1)$. This amounts to saying that at each point $p \in \mathbb{E}^n$, the tangent vectors $(\partial/\partial x^j)_p$ form an orthonormal basis of $T_p\mathbb{E}^n$.

Since $\{(\partial/\partial x^i)_p\}_{i=1}^n$ is a basis of T_pM, the condition $g_p(v_p, w_p) = 0$ for all $w_p \in T_pM$ is equivalent to $g_p(v_p, (\partial/\partial x^i)_p) = 0$, for $i = 1, 2, \ldots, n$; therefore, if $v_p = a^i(\partial/\partial x^i)_p$ is such that $g_p(v_p, w_p) = 0$ for all $w_p \in T_pM$, we have

$$g_{ij}(p)\, a^i = 0,$$

which is a homogeneous system of linear equations for the a^i. The tensor field g is non-singular if and only if $a^i = 0$ is the only solution of this system. Thus, g is non-singular if and only if the determinant of the matrix $(g_{ij}(p))$ is different from zero for all p in the domain of the coordinate system.

If $\mathbf{X}$ is a vector field on M, the contraction of $\mathbf{X}$ with g, $\mathbf{X} \lrcorner g$, is a tensor field of type $\binom{0}{1}$, that is, a covector field. If $\mathbf{X}$ is locally given by $\mathbf{X} = X^i(\partial/\partial x^i)$, we have

$$\mathbf{X} \lrcorner g = 2X^i g_{ij}\, \mathrm{d}x^j.$$

Since the determinant of the matrix (g_{ij}) never vanishes, the matrix (g_{ij}) has an inverse, whose entries are denoted by g^{ij}, that is,

$$g_{ij} g^{jk} = \delta_i^k. \tag{6.3}$$

Since the functions g_{ij} are differentiable, the functions g^{ij} are also differentiable. Furthermore, the symmetry of the components g_{ij} implies that $g^{ij} = g^{ji}$.

Hence, if α is a covector field, locally given by $\alpha = \alpha_i\, \mathrm{d}x^i$, there exists only one vector field $\mathbf{X}$ such that

$$\alpha = \frac{1}{2} \mathbf{X} \lrcorner g. \tag{6.4}$$

Indeed, in terms of the components of α and $\mathbf{X}$, the condition $\alpha = \frac{1}{2} \mathbf{X} \lrcorner g$ amounts to

$$\alpha_i = X^j g_{ji}, \tag{6.5}$$

therefore the components of $\mathbf{X}$ are determined by

$$X^j = \alpha_i g^{ij}. \tag{6.6}$$

Since the functions g_{ij} are differentiable, from (6.5) and (6.6) it follows that the vector field $\mathbf{X}$ is differentiable if and only if α is. Hence, in a Riemannian manifold, there exists a linear one-to-one correspondence between differentiable vector fields and 1-forms. (Owing to the form of expressions (6.5) and (6.6) this correspondence is an example of the operations called *raising and lowering of indices*.)

Definition 6.4 Let M be a Riemannian manifold and let $f \in C^\infty(M)$. The *gradient* of f, grad f, is the vector field on M such that

$$\mathrm{d}f = \frac{1}{2}(\operatorname{grad} f) \lrcorner\, g. \tag{6.7}$$

Then, from (2.42) and (1.45), for any vector field $\mathbf{X}$, we have

$$g(\operatorname{grad} f, \mathbf{X}) = \frac{1}{2}\big((\operatorname{grad} f) \lrcorner\, g\big)(\mathbf{X}) = \mathrm{d}f(\mathbf{X}) = \mathbf{X}f. \tag{6.8}$$

From the foregoing definition, (6.6) and (1.52) it follows that the gradient of f is locally given by

$$\operatorname{grad} f = g^{ij} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^j}. \tag{6.9}$$

Exercise 6.5 Show that $g^{ij} = g(\operatorname{grad} x^i, \operatorname{grad} x^j)$.

Let t and s be two tensor fields of type $\binom{0}{k}$ on M locally given by $t = t_{i_1 \ldots i_k}\, \mathrm{d}x^{i_1} \otimes \cdots \otimes \mathrm{d}x^{i_k}$ and $s = s_{j_1 \ldots j_k}\, \mathrm{d}x^{j_1} \otimes \cdots \otimes \mathrm{d}x^{j_k}$; the product $(t|s)$ will be defined by

$$(t|s) \equiv k!\, t_{i_1 \ldots i_k} s_{j_1 \ldots j_k} g^{i_1 j_1} \cdots g^{i_k j_k}. \tag{6.10}$$

Exercise 6.6 Show that the product $(\ |\ )$ is symmetric, bilinear and non-singular.

Exercise 6.7 Show that $g^{ij} = (\mathrm{d}x^i | \mathrm{d}x^j)$.

If M is a Riemannian manifold with a positive definite metric tensor g and $\psi : N \to M$ is a differentiable map from a manifold N into M such that for all $p \in N$, ψ_{*p} has maximal rank, that is, if $\psi_{*p} v_p = 0_{\psi(p)}$ implies $v_p = 0_p$, then $\psi^* g$ is a positive definite metric tensor in N since it is a symmetric tensor field of type $\binom{0}{2}$ and if $(\psi^* g)_p(v_p, w_p) = 0$ for all $w_p \in T_pN$, from the definition of $\psi^* g$, we have $g_{\psi(p)}(\psi_{*p} v_p, \psi_{*p} w_p) = 0$ for all $w_p \in T_pN$; in particular taking $w_p = v_p$ and using that g is positive definite it follows that $\psi_{*p} v_p = 0$, and one concludes that $v_p = 0_p$. A differentiable mapping satisfying the condition above is called an *immersion* (that is, for all $p \in N$, the rank of the linear mapping ψ_{*p} is equal to the dimension of N).

Example 6.8 The inclusion map $i : \mathrm{S}^2 \to \mathbb{R}^3$ is locally given by $i^* x = \sin\theta \cos\phi$, $i^* y = \sin\theta \sin\phi$, $i^* z = \cos\theta$, in terms of the usual coordinates (x, y, z) of $\mathbb{R}^3$ and of the spherical coordinates (θ, ϕ) of S^2. As can readily be verified, the Jacobian matrix of i with respect to these coordinate systems is given by

$$\begin{pmatrix} \cos\theta\cos\phi & -\sin\theta\sin\phi \\ \cos\theta\sin\phi & \sin\theta\cos\phi \\ -\sin\theta & 0 \end{pmatrix}$$

and its rank is equal to 2 in the whole domain of the spherical coordinates $(0 < \theta < \pi\,,\ 0 < \phi < 2\pi)$; therefore, the pullback of the Euclidean metric of $\mathbb{R}^3$, $g = \mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z$, under i is a positive definite metric for S^2. In fact, a straightforward computation yields

$$\begin{aligned} i^*g &= \mathrm{d}(\sin\theta\cos\phi) \otimes \mathrm{d}(\sin\theta\cos\phi) + \mathrm{d}(\sin\theta\sin\phi) \otimes \mathrm{d}(\sin\theta\sin\phi) \\ &\qquad + \mathrm{d}(\cos\theta) \otimes \mathrm{d}(\cos\theta) \\ &= \mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta\, \mathrm{d}\phi \otimes \mathrm{d}\phi \end{aligned} \tag{6.11}$$

and, as can be directly verified, it is positive definite at the points in the domain of the coordinate system.

Isometries. Killing Vector Fields Let M_1 and M_2 be two Riemannian manifolds with metric tensors g_1 and g_2, respectively. A diffeomorphism $\psi : M_1 \to M_2$ is an *isometry* if

$$\psi^* g_2 = g_1. \tag{6.12}$$

Two Riemannian manifolds M_1 and M_2 are *isometric* if there exists an isometry $\psi : M_1 \to M_2$.

Exercise 6.9 Show that the isometries of a manifold onto itself form a group under the composition.

Let φ be a one-parameter group of transformations on a Riemannian manifold, M, such that each transformation $\varphi_t : M \to M$ is an isometry; then, if $\mathbf{X}$ is the infinitesimal generator of φ, we have

$$\pounds_{\mathbf{X}} g = \lim_{t\to 0} \frac{{\varphi_t}^* g - g}{t} = 0. \tag{6.13}$$

The vector fields satisfying (6.13) are called *Killing vector fields*. The set of Killing vector fields of M will be denoted by $\mathfrak{K}(M)$.

Making use of the expression for the components of the Lie derivative of a tensor field (2.40), one finds that the components of a Killing vector field must satisfy the system of equations

$$X^k \frac{\partial g_{ij}}{\partial x^k} + g_{kj} \frac{\partial X^k}{\partial x^i} + g_{ik} \frac{\partial X^k}{\partial x^j} = 0. \tag{6.14}$$

Since the expression on the left-hand side of this equation is symmetric in the indices i and j, equations (6.14), known as the Killing equations, constitute a system of $n(n+1)/2$ *homogeneous, linear*, PDEs for the n components X^i of a Killing vector field, if n is the dimension of M. A solution of (6.13) is formed by n functions $X^1, X^2, \ldots, X^n$ of n variables. The linearity and homogeneity of the Killing equations imply that any linear combination with *constant coefficients* of solutions of these equations is also a solution (see the examples below).

From (6.14), it can be seen that if, for some specific value of the index m, the functions g_{ij} do not depend on x^m (i.e., $\partial g_{ij}/\partial x^m = 0$), then $\mathbf{X} = \partial/\partial x^m = \delta^i_m \partial/\partial x^i$ is a Killing vector field and conversely.

Exercise 6.10 Show that the set of Killing vector fields of a Riemannian manifold M, $\mathfrak{K}(M)$, is a Lie subalgebra of $\mathfrak{X}(M)$.

Example 6.11 Let us consider a Riemannian manifold, M, of dimension n such that, in some coordinate system, the components of the metric tensor are constant. (For instance, in Cartesian coordinates, the components of the metric tensor of an Euclidean space are $g_{ij} = \delta_{ij}$ and for the metric of the Minkowski space, (g_{ij}) is the matrix $\operatorname{diag}(1, 1, 1, -1)$ or its negative.) The Killing equations (6.14) then reduce to

$$\frac{\partial \xi_j}{\partial x^i} + \frac{\partial \xi_i}{\partial x^j} = 0, \quad \text{with } \xi_j \equiv g_{jk} X^k. \tag{6.15}$$

Hence,

$$\begin{aligned}\frac{\partial}{\partial x^k}\frac{\partial \xi_j}{\partial x^i} &= -\frac{\partial}{\partial x^k}\frac{\partial \xi_i}{\partial x^j} = -\frac{\partial}{\partial x^j}\frac{\partial \xi_i}{\partial x^k} = \frac{\partial}{\partial x^j}\frac{\partial \xi_k}{\partial x^i} = \frac{\partial}{\partial x^i}\frac{\partial \xi_k}{\partial x^j} = -\frac{\partial}{\partial x^i}\frac{\partial \xi_j}{\partial x^k} \\ &= -\frac{\partial}{\partial x^k}\frac{\partial \xi_j}{\partial x^i},\end{aligned}$$

and therefore $\partial^2 \xi_j/\partial x^k \partial x^i = 0$, which means that the components ξ_j must be of the form

$$\xi_j = a_{jk} x^k + b_j, \tag{6.16}$$

where the a_{ij} and b_i are constant. Substituting this expression into (6.15) one only obtains the condition $a_{ji} + a_{ij} = 0$; therefore, in a manifold of this class, locally, any Killing vector field is of the form

$$\begin{aligned}\mathbf{X} = X^k \frac{\partial}{\partial x^k} &= g^{kj} \xi_j \frac{\partial}{\partial x^k} = g^{kj}\left(a_{jl} x^l + b_j\right)\frac{\partial}{\partial x^k} \\ &= \frac{1}{2} a_{jl}\left(g^{kj} x^l - g^{kl} x^j\right)\frac{\partial}{\partial x^k} + b^k \frac{\partial}{\partial x^k},\end{aligned}$$

where $b^k \equiv g^{kj} b_j$. This means that the $n(n-1)/2$ vector fields

$$\mathbf{I}^{jl} \equiv \left(g^{kj} x^l - g^{kl} x^j\right)\frac{\partial}{\partial x^k}, \quad j < l, \tag{6.17}$$

together with the n vector fields

$$\mathbf{M}_k \equiv \frac{\partial}{\partial x^k} \tag{6.18}$$

form a basis of $\mathfrak{K}(M)$, which in this case has dimension $n(n+1)/2$. It turns out that, for an arbitrary n-dimensional Riemannian manifold, M, $\dim \mathfrak{K}(M) \leq n(n+1)/2$.

The integral curves of the Killing vector field $\frac{1}{2}a_{jl}\mathbf{I}^{jl}$ are determined by the system of *linear* ODEs

$$\frac{\mathrm{d}x^k}{\mathrm{d}t} = \frac{1}{2}a_{jl}\big(g^{kj}x^l - g^{kl}x^j\big) = g^{kj}a_{jl}x^l,$$

where $A \equiv (a_{ij})$ is an arbitrary real $n \times n$ skew-symmetric matrix. This system of equations can be expressed in matrix form:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \big(g^{-1}A\big)x,$$

where $g \equiv (g_{ij})$ is a symmetric (constant) matrix and x is a column matrix with entries $x^1, \ldots, x^n$ (or, more precisely, $x^1 \circ C, \ldots, x^n \circ C$, where C is an integral curve of $\frac{1}{2}a_{jl}\mathbf{I}^{jl}$). The solution of this matrix equation is (see, e.g., Hirsch and Smale 1974)

$$x(t) = \exp\big(tg^{-1}A\big)\, x(0),$$

where $\exp(tg^{-1}A) = \sum_{m=0}^{\infty}(tg^{-1}A)^m/m!$. That is, the column matrix $x(t)$ is related to $x(0)$ by means of the matrix $\exp(tg^{-1}A)$. One can readily verify that

$$\big(Ag^{-1}\big)^m g = g\big(g^{-1}A\big)^m, \quad m = 0, 1, 2, \ldots$$

and therefore, denoting by B^{t} the transpose of B, using the fact that $A^{\mathrm{t}} = -A$ and $g^{\mathrm{t}} = g$, we have

$$\begin{aligned}\big[\exp\big(tg^{-1}A\big)\big]^{\mathrm{t}} g \exp\big(tg^{-1}A\big) &= \big[\exp\big(-tAg^{-1}\big)\big]\, g \exp\big(tg^{-1}A\big) \\ &= g\big[\exp\big(-tg^{-1}A\big)\big] \exp\big(tg^{-1}A\big) \\ &= g,\end{aligned}$$

which means that, for all $t \in \mathbb{R}$, the matrix $\exp(tg^{-1}A)$ is *orthogonal* with respect to g. (Note that g is symmetric, but not necessarily diagonal.)

Example 6.12 The tensor field

$$g = y^{-2}(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y), \tag{6.19}$$

defines a positive definite metric on the *Poincaré half-plane* (or *hyperbolic plane*), $\mathbb{H}^2 \equiv \{(x, y) \in \mathbb{R}^2 \mid y > 0\}$. From (6.14), with x and y in place of x^1 and x^2, respectively, we obtain

$$-\frac{2}{y^3}X^2 + \frac{2}{y^2}\frac{\partial X^1}{\partial x} = 0, \tag{6.20a}$$

$$\frac{\partial X^2}{\partial x} + \frac{\partial X^1}{\partial y} = 0, \tag{6.20b}$$

$$-\frac{2}{y^3}X^2 + \frac{2}{y^2}\frac{\partial X^2}{\partial y} = 0. \tag{6.20c}$$

The last of these equations amounts to $\partial(y^{-1}X^2)/\partial y = 0$; hence, $X^2 = yh(x)$, where h is some real-valued function of one single variable.

On the other hand, the equality of the second partial derivatives $\partial^2 X^1/\partial x\,\partial y$ and $\partial^2 X^1/\partial y\,\partial x$, obtained from (6.20a) and (6.20b), is equivalent to $\mathrm{d}^2 h(x)/\mathrm{d}x^2 = 0$; hence, $h(x) = ax + b$, where a and b are two real constants. Therefore, $X^2 = axy + by$ and using again (6.20a) and (6.20b) one finds that $X^1 = \frac{1}{2}a(x^2 - y^2) + bx + c$, where c is another real constant. Thus, the general solution of the Killing equations (6.14) for the metric (6.19) has the form

$$\mathbf{X} = \frac{a}{2}\left(\left(x^2 - y^2\right)\frac{\partial}{\partial x} + 2xy\frac{\partial}{\partial y}\right) + b\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right) + c\frac{\partial}{\partial x}.$$

In other words, the vector fields

$$\left(x^2 - y^2\right)\frac{\partial}{\partial x} + 2xy\frac{\partial}{\partial y}, \qquad x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}, \qquad \frac{\partial}{\partial x}, \tag{6.21}$$

form a basis of $\mathfrak{K}(M)$. (In this case, as in the preceding example, $\mathfrak{K}(M)$ has the maximum dimension allowed by the dimension of M.)

Instead of the vector fields (6.21), we can choose the set

$$\begin{aligned} \mathbf{X}_1 &\equiv -2\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right), \\ \mathbf{X}_2 &\equiv -\frac{\partial}{\partial x}, \\ \mathbf{X}_3 &\equiv \left(x^2 - y^2\right)\frac{\partial}{\partial x} + 2xy\frac{\partial}{\partial y}, \end{aligned} \tag{6.22}$$

as a basis of $\mathfrak{K}(M)$. The Lie brackets among these vector fields are given by

$$[\mathbf{X}_1, \mathbf{X}_2] = 2\mathbf{X}_2, \qquad [\mathbf{X}_2, \mathbf{X}_3] = \mathbf{X}_1, \qquad [\mathbf{X}_3, \mathbf{X}_1] = 2\mathbf{X}_3, \tag{6.23}$$

which shows, in this particular case, that the Killing vector fields form a real Lie algebra (of dimension three in this example). The choice given by (6.22) has been made taking into account that with the group $\mathrm{SL}(2, \mathbb{R})$, formed by the 2×2 real matrices with determinant equal to 1, there is associated a Lie algebra that possesses a basis with relations identical to (6.23) (see Examples 7.16 and 7.60).

In fact, in this case it is not difficult to find the isometries generated by an arbitrary Killing vector field (that is, the one-parameter groups generated by these vector fields) and show that they are related with the group $\mathrm{SL}(2,\mathbb{R})$. To this end, it is convenient to make use of the complex variable

$$z \equiv x + \mathrm{i}y. \tag{6.24}$$

In the present case, any Killing vector field $\mathbf{X}$ can be expressed as

$$\begin{aligned}\mathbf{X} &= a^1\mathbf{X}_1 + a^2\mathbf{X}_2 + a^3\mathbf{X}_3 \\ &= \big(-2a^1x - a^2 + a^3(x^2-y^2)\big)\frac{\partial}{\partial x} + \big(-2a^1y + 2a^3xy\big)\frac{\partial}{\partial y},\end{aligned}$$

where the a^i are arbitrary real numbers. The integral curves of $\mathbf{X}$ are determined by the system of ODEs

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -2a^1x - a^2 + a^3(x^2-y^2), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = -2a^1y + 2a^3xy,$$

which amounts to the single equation

$$\begin{aligned}\frac{\mathrm{d}z}{\mathrm{d}t} &= \big(-2a^1x - a^2 + a^3(x^2-y^2)\big) + \mathrm{i}\big(-2a^1y + 2a^3xy\big) \\ &= a^3z^2 - 2a^1z - a^2.\end{aligned} \tag{6.25}$$

The form of the solution of this equation depends on the nature of the roots of the polynomial $a^3z^2 - 2a^1z - a^2$ or, equivalently, on the value of the discriminant $K \equiv -[(a^1)^2 + a^2a^3]$. If $K < 0$, the polynomial $a^3z^2 - 2a^1z - a^2$ has two different real roots, $\zeta_1 = (a^1 + \sqrt{-K})/a^3$, $\zeta_2 = (a^1 - \sqrt{-K})/a^3$ and from (6.25), according to the partial fractions method, we obtain

$$\int_0^t \mathrm{d}t = \int_{z(0)}^{z(t)} \frac{\mathrm{d}z}{2\sqrt{-K}}\left(\frac{1}{z-\zeta_1} - \frac{1}{z-\zeta_2}\right) = \frac{1}{2\sqrt{-K}} \ln \frac{z(t)-\zeta_1}{z(t)-\zeta_2}\,\frac{z(0)-\zeta_2}{z(0)-\zeta_1},$$

which amounts to the expression

$$z(t) = \frac{\alpha z(0) + \beta}{\gamma z(0) + \delta}, \tag{6.26}$$

where $\left(\begin{smallmatrix}\alpha & \beta \\ \gamma & \delta\end{smallmatrix}\right)$ is the matrix (dependent on the parameter t) belonging to $\mathrm{SL}(2,\mathbb{R})$ given by

$$\begin{pmatrix}\alpha & \beta \\ \gamma & \delta\end{pmatrix} = \cosh\sqrt{-K}\,t \begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix} - \frac{\sinh\sqrt{-K}\,t}{\sqrt{-K}} \begin{pmatrix}a^1 & a^2 \\ a^3 & -a^1\end{pmatrix}. \tag{6.27}$$

Note that we can multiply the coefficients α, β, γ, and δ appearing in (6.26) by a common nonzero, real or complex, factor λ, without altering the validity of (6.26).

Taking advantage of this freedom, it is convenient *to impose* the condition that the determinant of the matrix of the coefficients in the linear fractional transformation (6.26) be equal to 1. Nevertheless, this condition does not specify completely this matrix, because the determinant of its negative is also equal to 1. Thus, to each linear fractional transformation (6.26) there correspond two real matrices with determinant equal to 1.

In a similar way, in the cases where K is positive or equal to zero, one finds that the solution of (6.25) can be expressed in the form (6.26), with

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{cases} \cos\sqrt{K}\,t \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) - \frac{\sin\sqrt{K}\,t}{\sqrt{K}} \left(\begin{smallmatrix} a^1 & a^2 \\ a^3 & -a^1 \end{smallmatrix}\right) & \text{if } K > 0, \\ \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) - t \left(\begin{smallmatrix} a^1 & a^2 \\ a^3 & -a^1 \end{smallmatrix}\right) & \text{if } K = 0. \end{cases} \tag{6.28}$$

These matrices are also real and have determinant equal to 1 and, therefore, they also belong to $\mathrm{SL}(2,\mathbb{R})$. Note that, in all cases, the solution contains the traceless matrix

$$\begin{pmatrix} a^1 & a^2 \\ a^3 & -a^1 \end{pmatrix},$$

whose determinant is equal to K. One may notice that the expressions (6.28) can be obtained from (6.27) making use of the relationship between the hyperbolic functions of an imaginary argument and the trigonometric functions, or taking the limit as K goes to zero.

Even though one could express the solution (6.26) in terms of the original variables, x, y, it is more convenient to employ directly the formula (6.26), in part because the composition of linear fractional transformations is represented by matrix multiplication in the following sense. The composition of the linear fractional transformation $z \mapsto \frac{\alpha z+\beta}{\gamma z+\delta}$, which can be associated with the matrix $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$, followed by the map $z \mapsto \frac{az+b}{cz+d}$, associated with $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, is the linear fractional transformation $z \mapsto \frac{(a\alpha+b\gamma)z+a\beta+b\delta}{(c\alpha+d\gamma)z+c\beta+d\delta}$, which can be associated with the matrix product $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ (but also with any nonzero multiple of this product).

Not all the elements of the group $\mathrm{SL}(2,\mathbb{R})$ are of the form (6.27) or (6.28) (see Example 7.41); however, it can be directly verified that all the elements of this group give rise to isometries of (6.19).

Exercise 6.13 Show that if $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ is any matrix belonging to $\mathrm{SL}(2,\mathbb{R})$, then

$$\psi^* z = \frac{\alpha z + \beta}{\gamma z + \delta}, \tag{6.29}$$

with $z = x + \mathrm{i}y$, is an isometry of (6.19) [cf. (6.26)].

Exercise 6.14 Find the Killing vector fields of the *hyperbolic space* $\mathbb{H}^3 \equiv \{(x, y, z) \in \mathbb{R}^3 \mid z > 0\}$, which possesses the metric tensor

$$g = z^{-2}(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z).$$

Example 6.15 The metric induced on S^2, expressed in terms of the spherical coordinates $(x^1, x^2) = (\theta, \phi)$, has components $g_{11} = 1$, $g_{12} = 0$, $g_{22} = \sin^2\theta$ [see (6.11)], and therefore the Killing equations are

$$\begin{aligned} &2\frac{\partial X^1}{\partial\theta} = 0, \\ &\sin^2\theta\,\frac{\partial X^2}{\partial\theta} + \frac{\partial X^1}{\partial\phi} = 0, \\ &X^1\frac{\partial \sin^2\theta}{\partial\theta} + 2\sin^2\theta\,\frac{\partial X^2}{\partial\phi} = 0. \end{aligned} \tag{6.30}$$

The first of these equations amounts to $X^1 = F(\phi)$, where F is some real-valued function of one variable. Substituting this expression into the last equation of (6.30) we have $\partial(-X^2\tan\theta)/\partial\phi = F(\phi)$; hence, we have $-X^2\tan\theta = H(\phi) + G(\theta)$, where H is a primitive of F (i.e., $F = H'$) and G is some real-valued function of a single variable. Substitution of the expressions obtained above into the second equation of (6.30) yields

$$-\sin^2\theta\,\frac{\mathrm{d}(G(\theta)\,\cot\theta)}{\mathrm{d}\theta} + \frac{\mathrm{d}^2 H}{\mathrm{d}\phi^2} + H = 0.$$

Since the first term on the left-hand side of this last equation depends on θ only, while the last two terms depend on ϕ only, $\mathrm{d}^2H/\mathrm{d}\phi^2 + H = k$, where k is some constant and $\sin^2\theta\,\mathrm{d}(G(\theta)\,\cot\theta)/\mathrm{d}\theta = k$. The solutions of these equations are $H(\phi) = n^1\cos\phi + n^2\sin\phi + k$ and $G(\theta) = -k - n^3\tan\theta$, where n^1, n^2, and n^3 are real constants; thus, $X^1 = H'(\phi) = -n^1\sin\phi + n^2\cos\phi$ and $X^2 = -\cot\theta(H(\phi) + G(\theta)) = -\cot\theta(n^1\cos\phi + n^2\sin\phi) + n^3$. Therefore, the Killing vector fields of S^2, with the Riemannian structure induced by that of $\mathbb{R}^3$, are locally of the form

$$\begin{aligned} \mathbf{X} = n^1&\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right) \\ &+ n^2\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right) + n^3\frac{\partial}{\partial\phi}. \end{aligned} \tag{6.31}$$

Exercise 6.16 By means of the stereographic projection, each point of S^2 is put in correspondence with a point of the extended complex plane; in terms of the spherical coordinates of S^2, this mapping is given by $z = \mathrm{e}^{\mathrm{i}\phi}\cot(\theta/2)$ [see (1.3)]. Find the integral curves of (6.31) making use of the complex variable z and show that the isometries generated by the Killing vector fields (6.31) can be expressed in the form

(6.26) with $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ being a unitary complex matrix with determinant equal to 1 [that is, an element of the group SU(2)].

Example 6.17 The space $\mathbb{R}^3$ with its usual manifold structure and the metric tensor

$$\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y - \mathrm{d}z \otimes \mathrm{d}z, \tag{6.32}$$

where (x, y, z) are the natural coordinates of $\mathbb{R}^3$, is a pseudo-Riemannian manifold denoted by $\mathbb{R}^{2,1}$ (analogous to the Minkowski space, with two spatial and one temporal dimensions). Even though the metric tensor (6.32) is not positive definite, the metric tensor induced on the submanifold

$$M \equiv \left\{(x, y, z) \in \mathbb{R}^{2,1} \,|\, x^2 + y^2 - z^2 = -1,\ z > 0\right\}$$

is. This can be seen by noting first that the points of M can be put into a one-to-one correspondence with the points of the disk

$$\mathbb{D} \equiv \left\{(X, Y) \in \mathbb{R}^2 \,|\, X^2 + Y^2 < 1\right\}$$

by means of

$$X = \frac{x}{1+z}, \qquad Y = \frac{y}{1+z} \tag{6.33}$$

[cf. (1.4)] or, equivalently,

$$x = \frac{2X}{1 - X^2 - Y^2}, \qquad y = \frac{2Y}{1 - X^2 - Y^2}, \qquad z = \frac{1 + X^2 + Y^2}{1 - X^2 - Y^2}. \tag{6.34}$$

(The coordinates $(X, Y, 0)$ are those of the intersection of the plane $z = 0$ with the straight line joining the point $(x, y, z) \in M$ with the point $(0, 0, -1)$; see Fig. 6.1.) Making use of this correspondence, the variables (X, Y) can be used as coordinates of M and, in terms of these, the metric induced on M has the expression

$$\frac{4}{(1 - X^2 - Y^2)^2}(\mathrm{d}X \otimes \mathrm{d}X + \mathrm{d}Y \otimes \mathrm{d}Y). \tag{6.35}$$

A simple form of finding the Killing vector fields for the metric (6.35) consists of using the facts that the Killing vector fields of $\mathbb{R}^{2,1}$ are linear combinations of the six vector fields (6.17) and (6.18), with $(g^{ij}) = \mathrm{diag}(1, 1, -1)$, and that the only Killing vector fields of $\mathbb{R}^{2,1}$ tangent to the submanifold M are the linear combinations of the first three,

$$\mathbf{I}^{12} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}, \qquad \mathbf{I}^{13} = z\frac{\partial}{\partial x} + x\frac{\partial}{\partial z}, \qquad \mathbf{I}^{23} = z\frac{\partial}{\partial y} + y\frac{\partial}{\partial z}. \tag{6.36}$$

(Since M is defined by $x^2 + y^2 - z^2 = -1$ and that the derivative of $x^2 + y^2 - z^2$ along the direction of each of the fields (6.36) is equal to zero, it follows that these

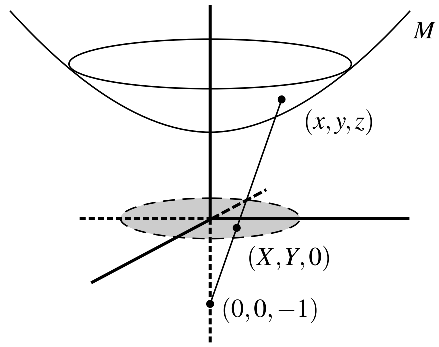

Fig. 6.1 Stereographic projection. The *points* $(x, y, z) \in M$, $(X, Y, 0)$ and $(0, 0, -1)$ lie on the *same straight line*

fields are tangent to M.) Using the relation (6.33) one finds that, on M,

$$\begin{aligned} \mathbf{I}^{12} &= Y\frac{\partial}{\partial X} - X\frac{\partial}{\partial Y}, \\ \mathbf{I}^{13} &= \frac{1}{2}\big(1 - X^2 + Y^2\big)\frac{\partial}{\partial X} - XY\frac{\partial}{\partial Y}, \\ \mathbf{I}^{23} &= -XY\frac{\partial}{\partial X} + \frac{1}{2}\big(1 + X^2 - Y^2\big)\frac{\partial}{\partial Y}, \end{aligned} \tag{6.37}$$

and by means of a direct computation it can be verified that these fields satisfy the Killing equations (6.14) for the metric (6.35).

Expression (6.35) can be regarded as that of the metric tensor of M in terms of the coordinates (X, Y) or as that of a metric tensor on $\mathbb{D}$. Formulas (6.33) and (6.34) then represent an isometry between M and $\mathbb{D}$. The vector fields (6.37) thus are also a basis for the Killing vector fields of $\mathbb{D}$.

On the other hand, the equation

$$x + \mathrm{i}y = \frac{(X + \mathrm{i}Y) + \mathrm{i}}{\mathrm{i}(X + \mathrm{i}Y) + 1} \tag{6.38}$$

establishes a correspondence between each point $(X, Y) \in \mathbb{D}$ and a point (x, y) of the Poincaré half-plane; this correspondence is one-to-one and, furthermore, an isometry. Consequently, there also exists a one-to-one correspondence between the Poincaré half-plane and the submanifold M of $\mathbb{R}^{2,1}$ defined above, and this correspondence is an isometry. Since all the Killing vector fields of the Poincaré half-plane and the isometries generated by them have been found in Example 6.12, by means of equations (6.33), (6.34), and (6.38) one can obtain all the Killing vector fields of M and the isometry groups generated by them [see also Lee (1997)].

Exercise 6.18 Show that, effectively, (6.38) establishes a one-to-one relation between the points (x, y) of the Poincaré half-plane (that is, $y > 0$) and the points $(X, Y) \in \mathbb{D}$ (with $X^2 + Y^2 < 1$) and that this relation is an isometry. Using the correspondence (6.38), show that the isometry of the Poincaré half-plane given by (6.29) amounts to the linear fractional transformation

$$\psi^* Z = \frac{aZ+b}{cZ+d},$$

where $Z \equiv X + \mathrm{i}Y$ and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -\mathrm{i} \\ -\mathrm{i} & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \mathrm{i} \\ \mathrm{i} & 1 \end{pmatrix}, \tag{6.39}$$

which is, therefore, an isometry of $\mathbb{D}$ (or of M). Show that the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ belongs to the group $\mathrm{SU}(1,1)$, which is formed by the complex 2×2 matrices, A, with determinant equal to 1, such that

$$A^\dagger \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{6.40}$$

(It can be verified that the relation (6.39) is an isomorphism of the group $\mathrm{SL}(2,\mathbb{R})$ onto $\mathrm{SU}(1,1)$.)

Conformal Mappings Besides the isometries, the transformations that preserve the metric up to a factor are also interesting. If M_1 and M_2 are two Riemannian manifolds, a differentiable mapping $\psi : M_1 \to M_2$ is a *conformal transformation* if there exists a positive function $\sigma \in C^\infty(M_1)$, such that $\psi^* g_2 = \sigma g_1$, where g_1 and g_2 are the metric tensors of M_1 and M_2, respectively.

Example 6.19 The inclusion mapping $i : \mathrm{S}^n \to \mathbb{R}^{n+1}$ identifies each point of the sphere S^n with the same point considered as a point of $\mathbb{R}^{n+1}$. The stereographic projection, $\phi : \mathrm{S}^n \setminus \{(0,0,\ldots,1)\} \to \mathbb{R}^n$ defined in Example 1.3 is a diffeomorphism and the composition $i \circ \phi^{-1}$ maps the points of $\mathbb{R}^n$ into the subset

$$\left\{ \left(a^1, \ldots, a^{n+1}\right) \in \mathbb{R}^{n+1} \mid \left(a^1\right)^2 + \cdots + \left(a^{n+1}\right)^2 = 1,\ a^{n+1} < 1 \right\},$$

i.e., the sphere with the north pole removed. In terms of the Cartesian coordinates $(y^1, y^2, \ldots, y^n)$ of $\mathbb{R}^n$ and $(x^1, x^2, \ldots, x^{n+1})$ of $\mathbb{R}^{n+1}$, the composition $i \circ \phi^{-1}$ is given by

$$\left(i \circ \phi^{-1}\right)^* x^j = \frac{2y^j}{1 + \mathbf{y}^2}, \quad \text{for } j = 1, 2, \ldots, n,$$

where $\mathbf{y}^2 \equiv (y^1)^2 + (y^2)^2 + \cdots + (y^n)^2$ [see (1.5)], and

$$\left(i \circ \phi^{-1}\right)^* x^{n+1} = \frac{\mathbf{y}^2 - 1}{\mathbf{y}^2 + 1}.$$

Hence, making use of the properties (1.52), (2.29), (2.30), and (2.32), one finds that

$$\left(i \circ \phi^{-1}\right)^* \left(\sum_{j=1}^{n+1} \mathrm{d}x^j \otimes \mathrm{d}x^j \right) = \frac{4}{(1+\mathbf{y}^2)^2} \sum_{j=1}^{n} \mathrm{d}y^j \otimes \mathrm{d}y^j$$

or, equivalently,

$$\left(\phi^{-1}\right)^* \left(i^* \sum_{j=1}^{n+1} \mathrm{d}x^j \otimes \mathrm{d}x^j \right) = \frac{4}{(1+\mathbf{y}^2)^2} \sum_{j=1}^{n} \mathrm{d}y^j \otimes \mathrm{d}y^j. \qquad (6.41)$$

The expression $\sum_{j=1}^{n} \mathrm{d}y^j \otimes \mathrm{d}y^j$ is the usual metric of $\mathbb{R}^n$, while $i^* \sum_{j=1}^{n+1} \mathrm{d}x^j \otimes \mathrm{d}x^j$ is the metric induced on S^n by the usual metric of $\mathbb{R}^{n+1}$ (see Example 6.8). Thus, ϕ^{-1} (and ϕ) is a conformal map. (Note, however, that ϕ is not defined on all of S^n.)

If $\mathbf{X}$ is the infinitesimal generator of a one-parameter group of conformal transformations of a Riemannian manifold M, then $\pounds_{\mathbf{X}} g = 2\chi g$, where χ is some function (the factor 2 is inserted for future convenience) and we say that $\mathbf{X}$ is a *conformal Killing vector field*. In terms of the components with respect to the natural basis induced by a coordinate system, $\mathbf{X}$ is a conformal Killing vector field if

$$X^k \frac{\partial g_{ij}}{\partial x^k} + g_{kj} \frac{\partial X^k}{\partial x^i} + g_{ik} \frac{\partial X^k}{\partial x^j} = 2\chi g_{ij}. \qquad (6.42)$$

When χ is a nonzero constant, $\mathbf{X}$ is called a *homothetic Killing vector field*.

Example 6.20 As in Example 6.11, we shall consider a Riemannian manifold such that, in some coordinate system, the components of the metric tensor are constant. Then equations (6.42) reduce to

$$\frac{\partial \xi_j}{\partial x^i} + \frac{\partial \xi_i}{\partial x^j} = 2\chi g_{ij}, \quad \text{with } \xi_j \equiv g_{jk} X^k \qquad (6.43)$$

[cf. (6.15)]. From (6.43) we obtain

$$\frac{\partial}{\partial x^k} \frac{\partial \xi_j}{\partial x^i} + \frac{\partial}{\partial x^k} \frac{\partial \xi_i}{\partial x^j} = 2 g_{ij} \frac{\partial \chi}{\partial x^k}. \qquad (6.44)$$

By cyclic permutations of the indices i, j, k in (6.44) we obtain two equations equivalent to that equation:

$$\frac{\partial}{\partial x^i} \frac{\partial \xi_k}{\partial x^j} + \frac{\partial}{\partial x^i} \frac{\partial \xi_j}{\partial x^k} = 2 g_{jk} \frac{\partial \chi}{\partial x^i},$$
$$\frac{\partial}{\partial x^j} \frac{\partial \xi_i}{\partial x^k} + \frac{\partial}{\partial x^j} \frac{\partial \xi_k}{\partial x^i} = 2 g_{ki} \frac{\partial \chi}{\partial x^j}.$$

Adding these two equations and subtracting equation (6.44) one finds that

$$\frac{\partial}{\partial x^i}\frac{\partial \xi_k}{\partial x^j} = g_{jk}\frac{\partial \chi}{\partial x^i} + g_{ki}\frac{\partial \chi}{\partial x^j} - g_{ij}\frac{\partial \chi}{\partial x^k}. \tag{6.45}$$

Applying $\partial/\partial x^m$ to both sides of this equation we now obtain

$$\frac{\partial}{\partial x^m}\frac{\partial}{\partial x^i}\frac{\partial \xi_k}{\partial x^j} - g_{jk}\frac{\partial}{\partial x^m}\frac{\partial \chi}{\partial x^i} = g_{ki}\frac{\partial}{\partial x^m}\frac{\partial \chi}{\partial x^j} - g_{ij}\frac{\partial}{\partial x^m}\frac{\partial \chi}{\partial x^k}.$$

Since the left-hand side is symmetric under the interchange of the indices i and m, the same must happen with the right-hand side, that is,

$$g_{ki}\frac{\partial}{\partial x^m}\frac{\partial \chi}{\partial x^j} - g_{ij}\frac{\partial}{\partial x^m}\frac{\partial \chi}{\partial x^k} = g_{km}\frac{\partial}{\partial x^i}\frac{\partial \chi}{\partial x^j} - g_{mj}\frac{\partial}{\partial x^i}\frac{\partial \chi}{\partial x^k}. \tag{6.46}$$

Multiplying both sides of (6.46) by g^{mj} we obtain

$$g_{ki}\nabla^2\chi + (n-2)\frac{\partial}{\partial x^i}\frac{\partial \chi}{\partial x^k} = 0, \tag{6.47}$$

where

$$\nabla^2\chi \equiv g^{mj}\frac{\partial}{\partial x^m}\frac{\partial \chi}{\partial x^j}$$

and n is the dimension of M. (The general definition of the Laplace operator, ∇^2, is given in Sect. 6.4 and it can be seen that the expression (6.113) derived there reduces to the one employed in the present case.) Multiplying (6.47) by g^{ki} one finds

$$(n-1)\nabla^2\chi = 0.$$

Thus, if $n \neq 1$, $\nabla^2\chi = 0$ and from (6.47) one concludes that, for $n \geq 3$,

$$\frac{\partial}{\partial x^i}\frac{\partial \chi}{\partial x^k} = 0.$$

Hence, for $n \geq 3$, χ must be of the form

$$\chi = c_j x^j + d, \tag{6.48}$$

where the c_j and d are arbitrary constants. Substituting (6.48) into (6.45) one finds that

$$\frac{\partial \xi_k}{\partial x^j} = g_{jk}c_i x^i + g_{ki}c_j x^i - g_{ij}c_k x^i + h_{kj}, \tag{6.49}$$

where the h_{kj} are constants. Using (6.48) and (6.49) it follows that (6.43) reduces to the equation

$$h_{ij} + h_{ji} = 2d g_{ij},$$

which relates the symmetric part of h_{ij} with d.

Finally, from (6.49) one finds that

$$\xi_k = g_{jk} c_i x^i x^j - \frac{1}{2} c_k g_{ij} x^i x^j + h_{kj} x^j + b_k,$$

where the b_k are arbitrary constants. Defining

$$a_{ij} \equiv \frac{1}{2}(h_{ij} - h_{ji}),$$

which satisfies the condition $a_{ij} = -a_{ji}$, we have

$$h_{ij} = \frac{1}{2}(h_{ij} + h_{ji}) + \frac{1}{2}(h_{ij} - h_{ji}) = d g_{ij} + a_{ij}.$$

Hence, when the dimension of M is greater than or equal to three, the general solution of (6.43) can be expressed in the form

$$\xi_k = g_{jk} c_i x^i x^j - \frac{1}{2} c_k g_{ij} x^i x^j + d g_{kj} x^j + a_{kj} x^j + b_k, \tag{6.50}$$

which contains $n + 1 + \frac{1}{2}n(n-1) + n = \frac{1}{2}(n+1)(n+2)$ arbitrary constants and reduces to (6.16) when $\chi = 0$ (that is, when $c_j = 0$ and $d = 0$). It can be verified directly that for $n = 1$ or 2 the expression (6.50) is also a solution of equations (6.42); however, when n is equal to 1 or 2, (6.50) does not contain all solutions of (6.42). In fact, when $n = 1$ any transformation of M into M with positive Jacobian is conformal (see also the comments at the end of this section regarding the case $n = 2$).

According to the preceding results, taking $n = 2$ and $g_{ij} = \delta_{ij}$, with $i, j = 1, 2$, the vector fields given by (6.50) generate *some* conformal transformations of the Euclidean plane onto itself (with x^1, x^2 being Cartesian coordinates). Fortunately, the transformations of this restricted class can be found explicitly in a relatively simple form making use of complex quantities. In fact, making $z \equiv x^1 + \mathrm{i}x^2$ one finds that in this case ($n = 2$, $g_{ij} = \delta_{ij}$) the integral curves of the vector field (6.50) are given by the equation

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{1}{2}(c_1 - \mathrm{i}c_2) z^2 + (d - \mathrm{i}a_{12}) z + b_1 + \mathrm{i}b_2,$$

where c_1, c_2, d, a_{12}, b_1, and b_2 are six arbitrary real constants [cf. (6.25)]. This equation can be integrated following a procedure similar to that employed in Example 6.12. The result is that $z(t)$ is related to $z(0)$ by means of a linear fractional transformation,

$$z(t) = \frac{\alpha z(0) + \beta}{\gamma z(0) + \delta} \tag{6.51}$$

[cf. (6.26)], where

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \cosh\left(\frac{1}{2}\sqrt{\Delta}\, t\right) I + \frac{\sinh(\frac{1}{2}\sqrt{\Delta}\, t)}{\sqrt{\Delta}} \begin{pmatrix} d - \mathrm{i}a_{12} & 2(b_1 + \mathrm{i}b_2) \\ -(c_1 - \mathrm{i}c_2) & -(d - \mathrm{i}a_{12}) \end{pmatrix},$$

with $\Delta \equiv (d - \mathrm{i}a_{12})^2 - 2(b_1 + \mathrm{i}b_2)(c_1 - \mathrm{i}c_2)$ [cf. (6.27)]. This matrix belongs to the $\mathrm{SL}(2, \mathbb{C})$ group, formed by the 2×2 complex matrices with determinant equal to 1.

As known from the complex variable theory, every analytic function, $f : \mathbb{C} \to \mathbb{C}$, is a conformal mapping; the linear fractional transformations (or *Möbius transformations*) (6.51) are distinguished because they are the only analytic one-to-one mappings of the extended complex plane (the complex plane plus the point at infinity) onto itself [see, e.g., Fisher (1999)].

Exercise 6.21 Show directly that any linear fractional transformation

$$\psi^* z = \frac{\alpha z + \beta}{\gamma z + \delta}$$

given by a matrix $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ belonging to the $\mathrm{SL}(2, \mathbb{C})$ group, with $z = x^1 + \mathrm{i}x^2$, is a conformal transformation of the Euclidean plane and find the corresponding conformal factor.

6.2 The Riemannian Connection

Theorem 6.22 *Let M be a Riemannian manifold. There exists a unique connection, ∇, the Riemannian or Levi-Civita connection, with vanishing torsion and such that $\nabla_{\mathbf{X}} g = 0$ for all $\mathbf{X} \in \mathfrak{X}(M)$; that is, there exists a unique connection on M such that*

$$[\mathbf{X}, \mathbf{Y}] = \nabla_{\mathbf{X}}\mathbf{Y} - \nabla_{\mathbf{Y}}\mathbf{X}, \tag{6.52}$$

$$\mathbf{X}\big(g(\mathbf{Y}, \mathbf{Z})\big) = g(\nabla_{\mathbf{X}}\mathbf{Y}, \mathbf{Z}) + g(\mathbf{Y}, \nabla_{\mathbf{X}}\mathbf{Z}), \tag{6.53}$$

for $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(M)$.

Proof Let $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(M)$. Assuming that such a connection exists, we have

$$\begin{aligned}
&\mathbf{X}\big(g(\mathbf{Y}, \mathbf{Z})\big) + \mathbf{Y}\big(g(\mathbf{Z}, \mathbf{X})\big) - \mathbf{Z}\big(g(\mathbf{X}, \mathbf{Y})\big) \\
&\quad = g(\nabla_{\mathbf{X}}\mathbf{Y}, \mathbf{Z}) + g(\mathbf{Y}, \nabla_{\mathbf{X}}\mathbf{Z}) + g(\nabla_{\mathbf{Y}}\mathbf{Z}, \mathbf{X}) + g(\mathbf{Z}, \nabla_{\mathbf{Y}}\mathbf{X}) \\
&\qquad - g(\nabla_{\mathbf{Z}}\mathbf{X}, \mathbf{Y}) - g(\mathbf{X}, \nabla_{\mathbf{Z}}\mathbf{Y}) \\
&\quad = g(\nabla_{\mathbf{X}}\mathbf{Y} + \nabla_{\mathbf{Y}}\mathbf{X}, \mathbf{Z}) + g(\mathbf{Y}, \nabla_{\mathbf{X}}\mathbf{Z} - \nabla_{\mathbf{Z}}\mathbf{X}) + g(\mathbf{X}, \nabla_{\mathbf{Y}}\mathbf{Z} - \nabla_{\mathbf{Z}}\mathbf{Y}) \\
&\quad = g\big(\nabla_{\mathbf{X}}\mathbf{Y} + \nabla_{\mathbf{X}}\mathbf{Y} + [\mathbf{Y}, \mathbf{X}], \mathbf{Z}\big) + g\big(\mathbf{Y}, [\mathbf{X}, \mathbf{Z}]\big) + g\big(\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]\big) \\
&\quad = 2g(\nabla_{\mathbf{X}}\mathbf{Y}, \mathbf{Z}) + g\big(\mathbf{Z}, [\mathbf{Y}, \mathbf{X}]\big) + g\big(\mathbf{Y}, [\mathbf{X}, \mathbf{Z}]\big) + g\big(\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]\big),
\end{aligned}$$

that is,

$$\begin{aligned}
2g(\nabla_{\mathbf{X}}\mathbf{Y}, \mathbf{Z}) = {}& \mathbf{X}\big(g(\mathbf{Y}, \mathbf{Z})\big) + \mathbf{Y}\big(g(\mathbf{Z}, \mathbf{X})\big) - \mathbf{Z}\big(g(\mathbf{X}, \mathbf{Y})\big) \\
& - g\big(\mathbf{Z}, [\mathbf{Y}, \mathbf{X}]\big) - g\big(\mathbf{Y}, [\mathbf{X}, \mathbf{Z}]\big) - g\big(\mathbf{X}, [\mathbf{Y}, \mathbf{Z}]\big).
\end{aligned} \tag{6.54}$$

Since g is non-singular, this relation defines $\nabla_{\mathbf{X}}\mathbf{Y}$. By construction, this connection has a vanishing torsion and satisfies $\nabla_{\mathbf{X}} g = 0$ for all $\mathbf{X} \in \mathfrak{X}(M)$. The explicit expression (6.54) shows its uniqueness. □

A connection, ∇, is a *metric connection* if $\nabla_{\mathbf{X}} g = 0$ for all $\mathbf{X} \in \mathfrak{X}(M)$. The Levi-Civita connection is the only metric connection whose torsion vanishes.

Making use of the relations $g_{ij} = g(\partial/\partial x^i, \partial/\partial x^j)$ and $[\partial/\partial x^i, \partial/\partial x^j] = 0$, from (6.54) we obtain

$$2g\left(\nabla_{\partial/\partial x^i}\frac{\partial}{\partial x^j}, \frac{\partial}{\partial x^k}\right) = \frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k};$$

hence, writing $\nabla_{\partial/\partial x^i}\, \partial/\partial x^j = \Gamma^l_{ji}\partial/\partial x^l$, it follows that

$$2\Gamma^l_{ji} g_{lk} = \frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k},$$

which leads to

$$\Gamma^l_{ji} = \frac{1}{2} g^{kl}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right). \tag{6.55}$$

This expression defines the so-called *Christoffel symbols*, which determine the Riemannian connection with respect to a holonomic basis. From (6.55) we find that $\Gamma^l_{ji} = \Gamma^l_{ij}$ (cf. Exercise 5.11); therefore, in a manifold of dimension n, there exist $n^2(n+1)/2$ independent Christoffel symbols.

From (6.55) and Exercise 5.12 it follows that if the components of the metric tensor in some *holonomic* basis are constant, then the curvature of the Riemannian connection is equal to zero.

Exercise 6.23 Show that if the g_{ij} are the components of the metric tensor with respect to a holonomic basis and the $\Gamma^i{}_j$ are the connection 1-forms for the Riemannian connection with respect to this basis, then $\mathrm{d}g_{ij} = g_{im}\Gamma^m{}_j + g_{jm}\Gamma^m{}_i$ and show that $g_{ij}\mathscr{R}^j{}_k = -g_{kj}\mathscr{R}^j{}_i$. (Note that this last equation amounts to $g_{ij}R^j{}_{klm} = -g_{kj}R^j{}_{ilm}$.)

A convenient way of computing the Christoffel symbols, especially in those cases where (g_{ij}) is diagonal, consists of using the fact that the geodesic equations (5.7) amount to the Euler–Lagrange equations for the Lagrangian

$$L = \frac{1}{2}(\pi^* g_{ij})\,\dot{q}^i\dot{q}^j, \tag{6.56}$$

where π is the canonical projection of the tangent bundle of M on M, the $\dot{q}^i$ are coordinates on *TM* induced by local coordinates x^i on M [see (1.28)], and the g_{ij} are the components of the metric tensor with respect to the holonomic basis $\partial/\partial x^i$.

In effect, the Euler–Lagrange equations (see Exercise 2.15)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial L}{\partial \dot{q}^k}\big(\widetilde{C}(t)\big)\right] - \frac{\partial L}{\partial q^k}\big(\widetilde{C}(t)\big) = 0, \quad k = 1, 2, \ldots, n,$$

where $\widetilde{C}$ is the curve in TM defined by $\widetilde{C}(t) = C'_t$, yield

$$\frac{\mathrm{d}}{\mathrm{d}t}\big[\big(\dot{q}^i\, \pi^* g_{ki}\big)\big(\widetilde{C}(t)\big)\big] - \left[\frac{1}{2}\dot{q}^i \dot{q}^j \frac{\partial(\pi^* g_{ij})}{\partial q^k}\right]\big(\widetilde{C}(t)\big) = 0.$$

Since $\pi(\widetilde{C}(t)) = C(t)$ and, according to (1.28) and (1.20), $\dot{q}^i(\widetilde{C}(t)) = C'_t[x^i] = \mathrm{d}(x^i \circ C)/\mathrm{d}t$, these equations are equivalent to (see Exercise 1.17)

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t} g_{ki}\big(C(t)\big)\right] - \frac{1}{2}\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\frac{\mathrm{d}(x^j \circ C)}{\mathrm{d}t}\frac{\partial g_{ij}}{\partial x^k}\big(C(t)\big) = 0$$

and to [see (1.20)]

$$\begin{aligned}
0 &= \frac{\mathrm{d}^2(x^i \circ C)}{\mathrm{d}t^2} g_{ki}\big(C(t)\big) + \frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\frac{\mathrm{d}(x^j \circ C)}{\mathrm{d}t}\frac{\partial g_{ki}}{\partial x^j}\big(C(t)\big) \\
&\quad - \frac{1}{2}\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\frac{\mathrm{d}(x^j \circ C)}{\mathrm{d}t}\frac{\partial g_{ij}}{\partial x^k}\big(C(t)\big) \\
&= \frac{\mathrm{d}^2(x^i \circ C)}{\mathrm{d}t^2} g_{ki}\big(C(t)\big) + \frac{1}{2}\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\frac{\mathrm{d}(x^j \circ C)}{\mathrm{d}t}\left(\frac{\partial g_{jk}}{\partial x^i} + \frac{\partial g_{ki}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^k}\right)\big(C(t)\big) \\
&= g_{km}\big(C(t)\big)\left[\frac{\mathrm{d}^2(x^m \circ C)}{\mathrm{d}t^2} + \Gamma^m_{ij}\big(C(t)\big)\frac{\mathrm{d}(x^i \circ C)}{\mathrm{d}t}\frac{\mathrm{d}(x^j \circ C)}{\mathrm{d}t}\right].
\end{aligned}$$

These equations are equivalent to the geodesic equations, since (g_{ij}) is non-singular.

Example 6.24 The tensor field

$$g = \frac{1}{1-kr^2}\,\mathrm{d}r \otimes \mathrm{d}r + r^2\big(\mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta\, \mathrm{d}\phi \otimes \mathrm{d}\phi\big), \tag{6.57}$$

with $k \in \mathbb{R}$, is a positive definite metric on an open subset of a manifold of dimension three defined by $r > 0$, $0 < \theta < \pi$, $0 < \phi < 2\pi$. In the case where k is positive, r is restricted by $0 < r < 1/\sqrt{k}$. The Lagrangian (6.56) becomes

$$L = \frac{1}{2}\left\{\frac{(\dot{q}^1)^2}{1 - k(q^1)^2} + \big(q^1\big)^2\big[\big(\dot{q}^2\big)^2 + \sin^2 q^2\big(\dot{q}^3\big)^2\big]\right\},$$

where $q^1 \equiv \pi^* r$, $q^2 \equiv \pi^* \theta$, $q^3 \equiv \pi^* \phi$. Substituting this expression into the Euler–Lagrange equations, one finds, for example,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\dot{q}^1}{1-k(q^1)^2}\big(\widetilde{C}(t)\big)\right] - \left\{\frac{kq^1(\dot{q}^1)^2}{[1-k(q^1)^2]^2} + q^1\left[\left(\dot{q}^2\right)^2 + \sin^2 q^2\left(\dot{q}^3\right)^2\right]\right\}\big(\widetilde{C}(t)\big) = 0,$$

that is,

$$\frac{1}{1-kr^2}\frac{\mathrm{d}^2 r}{\mathrm{d}t^2} + \frac{kr}{(1-kr^2)^2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 - r\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 - r\sin^2\theta\left(\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2 = 0,$$

where, in order to simplify the notation, we have written r, θ, and ϕ in place of $r \circ C$, $\theta \circ C$, and $\phi \circ C$, respectively. Comparing with the geodesic equations one obtains at once six of the Christoffel symbols:

$$\Gamma^1_{11} = \frac{kr}{1-kr^2}, \qquad \Gamma^1_{22} = -r\big(1-kr^2\big), \qquad \Gamma^1_{33} = -r\sin^2\theta\big(1-kr^2\big),$$

and $\Gamma^1_{ij} = 0$ for $i \neq j$.

Proceeding in this manner, one finds that the connection 1-forms $\Gamma^i{}_j = \Gamma^i_{jk}\,\mathrm{d}x^k$ are

$$\begin{aligned}
\Gamma^1{}_1 &= \frac{kr\,\mathrm{d}r}{1-kr^2}, & \Gamma^1{}_2 &= -r\big(1-kr^2\big)\,\mathrm{d}\theta, & \Gamma^1{}_3 &= -r\sin^2\theta\big(1-kr^2\big)\,\mathrm{d}\phi,\\
\Gamma^2{}_1 &= \frac{1}{r}\,\mathrm{d}\theta, & \Gamma^2{}_2 &= \frac{1}{r}\,\mathrm{d}r, & \Gamma^2{}_3 &= -\sin\theta\cos\theta\,\mathrm{d}\phi,\\
\Gamma^3{}_1 &= \frac{1}{r}\,\mathrm{d}\phi, & \Gamma^3{}_2 &= \cot\theta\,\mathrm{d}\phi, & \Gamma^3{}_3 &= \frac{1}{r}\,\mathrm{d}r + \cot\theta\,\mathrm{d}\theta,
\end{aligned} \tag{6.58}$$

and making use of the second Cartan structural equations one readily finds that the nonzero curvature forms are given by

$$\begin{aligned}
\mathscr{R}^1{}_2 &= kr^2\,\mathrm{d}r \wedge \mathrm{d}\theta,\\
\mathscr{R}^2{}_3 &= kr^2\sin^2\theta\,\mathrm{d}\theta \wedge \mathrm{d}\phi,\\
\mathscr{R}^3{}_1 &= \frac{k}{1-kr^2}\,\mathrm{d}\phi \wedge \mathrm{d}r,
\end{aligned}$$

and

$$\begin{aligned}
\mathscr{R}^2{}_1 &= \frac{k}{1-kr^2}\,\mathrm{d}\theta \wedge \mathrm{d}r,\\
\mathscr{R}^3{}_2 &= kr^2\,\mathrm{d}\phi \wedge \mathrm{d}\theta,\\
\mathscr{R}^1{}_3 &= kr^2\sin^2\theta\,\mathrm{d}r \wedge \mathrm{d}\phi
\end{aligned}$$

(see Exercise 6.23), which can be summarized by the expression $\mathscr{R}^i{}_j = kg_{jm}\,\mathrm{d}x^i \wedge \mathrm{d}x^m$, that is,

$$R^i{}_{jlm} = k\big(\delta^i_l g_{jm} - \delta^i_m g_{jl}\big) \tag{6.59}$$

[see (5.27)]. (A Riemannian manifold of dimension greater than two whose curvature is of the form (6.59) is said to be a *constant curvature manifold*.)

When $k = 0$, the curvature is equal to zero and (6.57) coincides with the usual metric of the Euclidean space of dimension three, in spherical coordinates. For $k = 1$, (6.57) coincides with the usual metric of the sphere S^3 (which, perhaps, can be more readily seen using, in place of r, the variable χ defined by $r = \sin\chi$).

Rigid Bases Besides the holonomic bases, $\{\partial/\partial x^i\}_{i=1}^n$, induced by coordinate systems, another important class of bases are the rigid ones. A set of basis vector fields $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, not necessarily holonomic, is a *rigid basis* if the components of the metric tensor, $g_{ij} \equiv g(\mathbf{e}_i, \mathbf{e}_j)$, are constant. From the property (6.53) and recalling that $\nabla_{\mathbf{e}_i}\mathbf{e}_j = \Gamma^l{}_{ji}\mathbf{e}_l$ [see (5.20) and (5.21)] it follows that

$$\begin{aligned} 0 = \mathbf{e}_i g_{jk} = \mathbf{e}_i\big(g(\mathbf{e}_j, \mathbf{e}_k)\big) &= g(\nabla_{\mathbf{e}_i}\mathbf{e}_j, \mathbf{e}_k) + g(\mathbf{e}_j, \nabla_{\mathbf{e}_i}\mathbf{e}_k) \\ &= g\big(\Gamma^l{}_{ji}\mathbf{e}_l, \mathbf{e}_k\big) + g\big(\mathbf{e}_j, \Gamma^l{}_{ki}\mathbf{e}_l\big) = g_{lk}\Gamma^l{}_{ji} + g_{jl}\Gamma^l{}_{ki}; \end{aligned}$$

hence, defining

$$\Gamma_{ijk} \equiv g_{il}\Gamma^l{}_{jk}, \tag{6.60}$$

we have

$$\Gamma_{kji} + \Gamma_{jki} = 0. \tag{6.61}$$

In this case the functions $\Gamma^i{}_{jk}$ or, equivalently, Γ_{ijk}, are called the *Ricci rotation coefficients*; owing to the skew-symmetry of Γ_{ijk} in the two first indices, in a manifold of dimension n, there are $n^2(n-1)/2$ independent Ricci rotation coefficients.

From the property (6.52) we see that

$$[\mathbf{e}_i, \mathbf{e}_j] = \nabla_{\mathbf{e}_i}\mathbf{e}_j - \nabla_{\mathbf{e}_j}\mathbf{e}_i = \big(\Gamma^k{}_{ji} - \Gamma^k{}_{ij}\big)\mathbf{e}_k, \tag{6.62}$$

that is, the Lie brackets of the basis fields give the skew-symmetric part in the last two indices of the Ricci rotation coefficients, $\Gamma^k{}_{[ij]} \equiv \frac{1}{2}(\Gamma^k{}_{ij} - \Gamma^k{}_{ji})$. These relations [alone or combined with the property (6.61)] allow us to calculate the Ricci rotation coefficients; in effect, noting that $2g(\nabla_{\mathbf{e}_i}\mathbf{e}_j, \mathbf{e}_k) = 2g(\Gamma^l{}_{ji}\mathbf{e}_l, \mathbf{e}_k) = 2\Gamma_{kji}$, from (6.54) we obtain

$$\begin{aligned} 2g(\nabla_{\mathbf{e}_i}\mathbf{e}_j, \mathbf{e}_k) &= -g\big(\mathbf{e}_k, [\mathbf{e}_j, \mathbf{e}_i]\big) - g\big(\mathbf{e}_j, [\mathbf{e}_i, \mathbf{e}_k]\big) - g\big(\mathbf{e}_i, [\mathbf{e}_j, \mathbf{e}_k]\big) \\ &= -g\big(\mathbf{e}_k, 2\Gamma^l{}_{[ij]}\mathbf{e}_l\big) - g\big(\mathbf{e}_j, 2\Gamma^l{}_{[ki]}\mathbf{e}_l\big) - g\big(\mathbf{e}_i, 2\Gamma^l{}_{[kj]}\mathbf{e}_l\big) \\ &= -2\Gamma_{k[ij]} - 2\Gamma_{j[ki]} - 2\Gamma_{i[kj]}; \end{aligned}$$

hence

$$\Gamma_{kji} = \Gamma_{k[ji]} - \Gamma_{j[ki]} - \Gamma_{i[kj]}. \tag{6.63}$$

(Note that (6.61) follows from (6.63).)

Alternatively, from the first Cartan structural equations, owing to the fact that the torsion of the Riemannian connection is equal to zero and that the exterior product of 1-forms is skew-symmetric, we have

$$\mathrm{d}\theta^i = \Gamma^i{}_{jk}\theta^j \wedge \theta^k = \Gamma^i{}_{[jk]}\theta^j \wedge \theta^k. \tag{6.64}$$

Therefore, the computation of the exterior derivative of the 1-forms θ^i also yields the skew-symmetric part in the last two indices of the Ricci rotation coefficients and by means of the relation (6.63) the value of each of the coefficients Γ_{ijk} can be obtained (see Examples 6.25, 6.37, 6.39, and 6.47).

The skew-symmetry of the Ricci coefficients, (6.61), is equivalent to the skew-symmetry of the connection 1-forms $\Gamma_{ij} \equiv g_{ik}\Gamma^k{}_j = \Gamma_{ijk}\theta^k$, with respect to a rigid basis, given by

$$\Gamma_{ij} = -\Gamma_{ji}, \tag{6.65}$$

which implies that the curvature 2-forms $\mathscr{R}_{ij} \equiv g_{ik}\mathscr{R}^k{}_j$ are also skew-symmetric

$$\mathscr{R}_{ij} = -\mathscr{R}_{ji}. \tag{6.66}$$

Indeed, from the second Cartan structural equations, using (6.65) and the anticommutativity of the exterior product of 1-forms, we have $\mathscr{R}_{ij} = \mathrm{d}\Gamma_{ij} + \Gamma_{ik} \wedge \Gamma^k{}_j = -\mathrm{d}\Gamma_{ji} - \Gamma_{ki} \wedge \Gamma^k{}_j = -\mathrm{d}\Gamma_{ji} - \Gamma^k{}_i \wedge \Gamma_{kj} = -\mathrm{d}\Gamma_{ji} - \Gamma_{jk} \wedge \Gamma^k{}_i = -\mathscr{R}_{ji}$ (cf. Exercise 6.23).

Example 6.25 By expressing the metric tensor (6.57) in the form

$$g = \frac{\mathrm{d}r}{\sqrt{1-kr^2}} \otimes \frac{\mathrm{d}r}{\sqrt{1-kr^2}} + r\,\mathrm{d}\theta \otimes r\,\mathrm{d}\theta + r\sin\theta\,\mathrm{d}\phi \otimes r\sin\theta\,\mathrm{d}\phi,$$

it follows that the 1-forms

$$\theta^1 \equiv \frac{\mathrm{d}r}{\sqrt{1-kr^2}}, \qquad \theta^2 \equiv r\,\mathrm{d}\theta, \theta \qquad {}^3 \equiv r\sin\theta\,\mathrm{d}\phi \tag{6.67}$$

form the dual basis of an orthonormal basis, that is, $g = g_{ij}\theta^i \otimes \theta^j$, with

$$(g_{ij}) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.68}$$

The exterior derivatives of the 1-forms (6.67) are

$$\begin{aligned} \mathrm{d}\theta^1 &= 0, \\ \mathrm{d}\theta^2 &= \mathrm{d}r \wedge \mathrm{d}\theta = \frac{\sqrt{1-kr^2}}{r}\theta^1 \wedge \theta^2, \\ \mathrm{d}\theta^3 &= \sin\theta\,\mathrm{d}r \wedge \mathrm{d}\phi + r\cos\theta\,\mathrm{d}\theta \wedge \mathrm{d}\phi \\ &= \frac{\sqrt{1-kr^2}}{r}\theta^1 \wedge \theta^3 + \frac{\cot\theta}{r}\theta^2 \wedge \theta^3, \end{aligned}$$

which, compared with (6.64) and using (6.68) lead to $\Gamma_{2[12]} = \frac{\sqrt{1-kr^2}}{2r} = \Gamma_{3[13]}$, $\Gamma_{3[23]} = \frac{\cot\theta}{2r}$, and the other $\Gamma_{i[jk]}$ are equal to zero. Substituting into (6.63) one finds that all the nonzero Ricci rotation coefficients are given by $\Gamma_{122} = -\sqrt{1-kr^2}/r$, $\Gamma_{233} = -\cot\theta/r$, and $\Gamma_{313} = \sqrt{1-kr^2}/r$. Hence,

$$\Gamma_{12} = -\sqrt{1-kr^2}\,\mathrm{d}\theta, \Gamma_{23} = -\cos\theta\,\mathrm{d}\phi, \Gamma_{31} = \sqrt{1-kr^2}\sin\theta\,\mathrm{d}\phi \tag{6.69}$$

[cf. (6.58)]. Employing now the second Cartan structural equations, one finds that

$$\mathscr{R}_{12} = k\theta^1 \wedge \theta^2, \qquad \mathscr{R}_{23} = k\theta^2 \wedge \theta^3, \qquad \mathscr{R}_{31} = k\theta^3 \wedge \theta^1,$$

which can be expressed in the form [see (6.68)] $\mathscr{R}_{ij} = \frac{k}{2}(g_{il}g_{jm} - g_{im}g_{jl})\,\theta^l \wedge \theta^m$. Thus, we find again that the components of the curvature tensor of the metric (6.57) are given by (6.59).

Exercise 6.26 Show that the curvature of the manifold $\mathbb{H}^n \equiv \{(x^1, x^2, \ldots, x^n) \in \mathbb{R}^n \mid x^n > 0\}$, with the metric tensor

$$g = (x^n)^{-2}(\mathrm{d}x^1 \otimes \mathrm{d}x^1 + \mathrm{d}x^2 \otimes \mathrm{d}x^2 + \cdots + \mathrm{d}x^n \otimes \mathrm{d}x^n),$$

called the hyperbolic space, is given by $R_{ijlm} = g_{im}g_{jl} - g_{il}g_{jm}$.

Geodesics of a Riemannian Manifold If $C : I \to M$ is a geodesic (that is, $\nabla_{C'}C' = 0$), then $g(C', C')$ is constant, since

$$C'[g(C', C')] = g(\nabla_{C'}C', C') + g(C', \nabla_{C'}C') = 2g(C', \nabla_{C'}C') = 0. \tag{6.70}$$

If M is a Riemannian manifold with a positive definite metric, $g(C', C')$ is the square of the length of the vector field tangent to C; therefore, in this case, the length of the tangent vector field of a geodesic is constant. In the case of a Riemannian manifold with a positive definite metric tensor, the geodesics are the curves that locally minimize length [see, e.g., do Carmo (1992), Lee (1997)].

The following theorem gives an alternative way of defining a Killing vector field, making use of the Riemannian connection [cf. (6.13)].

Theorem 6.27 *$\mathbf{X}$ is a Killing vector field if and only if*

$$g(\nabla_{\mathbf{Y}}\mathbf{X}, \mathbf{Z}) + g(\mathbf{Y}, \nabla_{\mathbf{Z}}\mathbf{X}) = 0,$$

for $\mathbf{Y}, \mathbf{Z} \in \mathfrak{X}(M)$.

Proof Making use of (2.45), (2.27), (6.53), and (5.13) with $T = 0$ one finds that

$$\begin{aligned}(\pounds_{\mathbf{X}}g)(\mathbf{Y}, \mathbf{Z}) &= \mathbf{X}(g(\mathbf{Y}, \mathbf{Z})) - g(\pounds_{\mathbf{X}}\mathbf{Y}, \mathbf{Z}) - g(\mathbf{Y}, \pounds_{\mathbf{X}}\mathbf{Z})\\ &= \mathbf{X}(g(\mathbf{Y}, \mathbf{Z})) - g([\mathbf{X}, \mathbf{Y}], \mathbf{Z}) - g(\mathbf{Y}, [\mathbf{X}, \mathbf{Z}])\end{aligned}$$

$$
\begin{aligned}
&= g(\nabla_{\mathbf{X}}\mathbf{Y}, \mathbf{Z}) + g(\mathbf{Y}, \nabla_{\mathbf{X}}\mathbf{Z}) \\
&\quad - g(\nabla_{\mathbf{X}}\mathbf{Y} - \nabla_{\mathbf{Y}}\mathbf{X}, \mathbf{Z}) - g(\mathbf{Y}, \nabla_{\mathbf{X}}\mathbf{Z} - \nabla_{\mathbf{Z}}\mathbf{X}) \\
&= g(\nabla_{\mathbf{Y}}\mathbf{X}, \mathbf{Z}) + g(\mathbf{Y}, \nabla_{\mathbf{Z}}\mathbf{X}),
\end{aligned}
$$

which leads to the desired result. □

Theorem 6.28 *If C is a geodesic and $\mathbf{X}$ is a Killing vector field, then $g(\mathbf{X}, C')$ is constant along C.*

Proof Making use of (6.53) and the definition of a geodesic we have

$$
C'\big[g(\mathbf{X}, C')\big] = g(\nabla_{C'}\mathbf{X}, C') + g(\mathbf{X}, \nabla_{C'}C') = g(\nabla_{C'}\mathbf{X}, C'),
$$

which is equal to zero according to Theorem 6.27. □

Example 6.29 In order to find the geodesics of Poincaré's half-plane we can take advantage of the existence of the three Killing vector fields (6.22). Making use of (6.19) and (6.22), according to Theorem 6.28 we obtain the three constants

$$
\begin{aligned}
c_1 &\equiv g(\mathbf{X}_1, C') = -2y^{-2}\left(x\frac{\mathrm{d}x}{\mathrm{d}t} + y\frac{\mathrm{d}y}{\mathrm{d}t}\right), \\
c_2 &\equiv g(\mathbf{X}_2, C') = -y^{-2}\frac{\mathrm{d}x}{\mathrm{d}t}, \\
c_3 &\equiv g(\mathbf{X}_3, C') = y^{-2}\left(\left(x^2 - y^2\right)\frac{\mathrm{d}x}{\mathrm{d}t} + 2xy\frac{\mathrm{d}y}{\mathrm{d}t}\right),
\end{aligned}
\tag{6.71}
$$

where, by abuse of notation, we have written x in place of $x \circ C$, and so on. By combining the first two equations one finds that $2c_2(x\,\mathrm{d}x/\mathrm{d}t + y\,\mathrm{d}y/\mathrm{d}t) = c_1\,\mathrm{d}x/\mathrm{d}t$, that is, $\mathrm{d}(c_2(x^2 + y^2) - c_1 x)/\mathrm{d}t = 0$; hence $c_2(x^2 + y^2) - c_1 x$ is a constant that, if $c_2 \neq 0$, is conveniently expressed by $c_2[R^2 - (c_1/2c_2)^2]$, so that we have $(x - c_1/2c_2)^2 + y^2 = R^2$, which corresponds to the upper part of a circle (since $y > 0$) whose center is on the x axis. When $c_2 = 0$, from (6.71) we see that x is a constant. Thus, the images of the geodesics for the metric (6.19) are half-circles with center on the x axis or vertical lines.

The parametrization of these curves can be obtained making use again of (6.71), which gives (for $c_2 \neq 0$)

$$
\mathrm{d}t = -\frac{1}{c_2 y^2}\,\mathrm{d}x = -\frac{\mathrm{d}x}{2Rc_2}\left[\frac{1}{x - (c_1/2c_2) + R} - \frac{1}{x - (c_1/2c_2) - R}\right]
$$

and, therefore,

$$
x(t) = \frac{(x_2 - x(0))x_1 + (x(0) - x_1)x_2\mathrm{e}^{-2Rc_2 t}}{x_2 - x(0) + (x(0) - x_1)\mathrm{e}^{-2Rc_2 t}},
$$

where $x_1 \equiv (c_1/2c_2) - R$, $x_2 \equiv (c_1/2c_2) + R$. (Note that when $t \to \pm\infty$, x tends to x_1 or x_2.) Substituting the expressions obtained above into (6.71) one finds that the radius R is related to the constants c_i through $R^2 = ({c_1}^2 + 4c_2c_3)/(4{c_2}^2)$. The expression for $y(t)$ can be obtained from the second equation in (6.71).

In addition to the three constants (6.71), equation (6.70) yields a fourth constant, $E \equiv \frac{1}{2} g(C', C')$, which turns out to be function of the c_i. In fact, making use of (6.19) and (6.71) one finds that

$$E = \frac{1}{2y^2}\left[\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)^2\right] = \frac{1}{8}\left({c_1}^2 + 4c_2c_3\right)$$

so that, if $c_2 \neq 0$, we have $E = \frac{1}{2} R^2 {c_2}^2$. Making use of the foregoing results, one can readily see that each geodesic of this manifold has infinite length since $L_C = \int_{-\infty}^{\infty} \sqrt{g_{C(t)}(C'_t, C'_t)}\,\mathrm{d}t = \sqrt{2E}\int_{-\infty}^{\infty} \mathrm{d}t$, and from equations (6.71) one also finds that given any point p of the half-plane $y > 0$ and any tangent vector at p, there exists a (unique) geodesic passing through p, where its tangent vector coincides with the given vector. For this reason, this manifold is *geodesically complete* [see also, e.g., do Carmo (1992), Lee (1997), and Conlon (2001)].

Example 6.30 Starting from the three basis Killing vector fields for the metric (6.11), given by (6.31), with the aid of Theorem 6.28 we have the three constant quantities

$$c_1 \equiv -\sin\phi \frac{\mathrm{d}\theta}{\mathrm{d}t} - \sin\theta\cos\theta\cos\phi\frac{\mathrm{d}\phi}{\mathrm{d}t},$$
$$c_2 \equiv \cos\phi \frac{\mathrm{d}\theta}{\mathrm{d}t} - \sin\theta\cos\theta\sin\phi\frac{\mathrm{d}\phi}{\mathrm{d}t},$$
$$c_3 \equiv \sin^2\theta \frac{\mathrm{d}\phi}{\mathrm{d}t}.$$

By combining these equations one obtains $c_1 \cos\phi + c_2 \sin\phi = -c_3 \cot\theta$ or, equivalently, $c_1 \sin\theta\cos\phi + c_2 \sin\theta\sin\phi + c_3\cos\theta = 0$. Taking into account the relation between the spherical and the Cartesian coordinates, one concludes that this last equation corresponds to the intersection of the sphere with the plane passing through the origin given by $c_1x + c_2y + c_3z = 0$; that is, the geodesics of S^2 are the intersections of the sphere with the planes passing through the origin.

Exercise 6.31 Show that the connection considered in Example 5.6 is the Levi-Civita connection corresponding to the metric tensor $g = (1 + r^2)^{-2}(\mathrm{d}r \otimes \mathrm{d}r + r^2\,\mathrm{d}\theta \otimes \mathrm{d}\theta)$. Since the components of the metric tensor in these coordinates do not depend on θ, $\partial/\partial\theta$ is a Killing vector field for this metric. Find the geodesics making use of Theorem 6.28 and of the fact that $g(C', C')$ is a constant for any geodesic C. Find all the Killing vector fields and the constants associated with them.

Exercise 6.32 Show that if the vector field $\mathbf{X}$ is the gradient of some function and $g(\mathbf{X}, \mathbf{X})$ is constant, then $\nabla_{\mathbf{X}}\mathbf{X} = 0$, i.e., the integral curves of $\mathbf{X}$ are geodesics.

(*Hint*: assuming that $\mathbf{X} = \mathrm{grad}\, f$, make use of the definition $[\mathbf{X}, \mathbf{Y}]f = \mathbf{X}(\mathbf{Y}f) - \mathbf{Y}(\mathbf{X}f)$ together with (6.8) to establish the equality $g(\mathbf{X}, [\mathbf{X}, \mathbf{Y}]) = \mathbf{X}[g(\mathbf{X}, \mathbf{Y})] - \mathbf{Y}[g(\mathbf{X}, \mathbf{X})]$, and then employ (6.52) and (6.53).)

The result stated in Exercise 6.32 is especially interesting for it relates the problem of writing down and solving the equations for the geodesics (5.7) (which may involve the computation of the functions Γ^i_{jk}) with that of solving the PDE

$$g(\mathrm{grad}\, W, \mathrm{grad}\, W) = \mathrm{const.} \tag{6.72}$$

It turns out that, locally, any geodesic is an integral curve of the gradient of a solution of (6.72); what is more remarkable and useful is that if one knows a *complete solution* of (6.72) (a concept defined in the next paragraph), then the geodesics can be found without having to solve the differential equations for the integral curves of $\mathrm{grad}\, W$.

A complete solution of (6.72) is a function satisfying (6.72) that depends on $n-1$ parameters a_i, where $n = \dim M$, in such a way that the partial derivatives of W with respect to the parameters a_i are (functionally) independent. In terms of a coordinate system x^i, equation (6.72) is equivalent to [see (6.9)]

$$g^{ij} \frac{\partial W}{\partial x^i} \frac{\partial W}{\partial x^j} = \mathrm{const.} \tag{6.73}$$

Differentiating this equation with respect to the parameter a_k one obtains

$$2g^{ij} \frac{\partial W}{\partial x^i} \frac{\partial}{\partial x^j} \frac{\partial W}{\partial a_k} = 0. \tag{6.74}$$

Since $g^{ij}(\partial W/\partial x^i)\, \partial/\partial x^j = \mathrm{grad}\, W$, equation (6.74) means that each of the $n-1$ partial derivatives $\partial W/\partial a_k$ is constant along the integral curves of $\mathrm{grad}\, W$; that is, if we define

$$b^k \equiv \frac{\partial W}{\partial a_k} \quad (k = 1, \ldots, n-1), \tag{6.75}$$

then the (images of the) integral curves of $\mathrm{grad}\, W$ (which are geodesics) are the intersection of the $n-1$ hypersurfaces given by $b^k = \mathrm{const}$. By suitably selecting the values of the $2n-2$ parameters a_k, b^k, we obtain the geodesic passing through a given point in any given direction (see Example 6.33).

Example 6.33 Considering again the Poincaré half-plane with the coordinates employed in Example 6.12, equation (6.73) takes the form

$$y^2 \left[\left(\frac{\partial W}{\partial x} \right)^2 + \left(\frac{\partial W}{\partial y} \right)^2 \right] = \mathrm{const.} \tag{6.76}$$

This PDE can be solved by the method of separation of variables, looking for a solution of the form $W = F(x) + G(y)$, substituting into (6.76) and denoting by c^2

the value of the constant on the right-hand side of the equation, one finds that

$$\left(\frac{\mathrm{d}F}{\mathrm{d}x}\right)^2 = \frac{c^2}{y^2} - \left(\frac{\mathrm{d}G}{\mathrm{d}y}\right)^2.$$

This equation will hold for all values of x and y only if each side of the equation is equal to a constant, which will be denoted by a^2; then (setting to zero the integration constants), $F(x) = ax$ and $G = \int \sqrt{c^2 y^{-2} - a^2}\,\mathrm{d}y$, so that $W = ax + \int \sqrt{c^2 y^{-2} - a^2}\,\mathrm{d}y$ is a solution of (6.76) depending on the parameter a and, since $\partial W/\partial a \neq 0$, this is a complete solution. Then, using the fact that $b \equiv \partial W/\partial a$ is constant along each geodesic one finds

$$b = x - a \int \frac{y\,\mathrm{d}y}{\sqrt{c^2 - a^2 y^2}} = x + \sqrt{\frac{c^2}{a^2} - y^2},$$

which represents a two-parameter family of arcs of circles.

Exercise 6.34 Making use of the procedure employed in the preceding example, find the geodesics of the metric $g = (1 + r^2)^{-2}(\mathrm{d}r \otimes \mathrm{d}r + r^2\,\mathrm{d}\theta \otimes \mathrm{d}\theta)$, considered in Exercise 6.31.

Exercise 6.35 Show that the geodesics of the metric $y^{-1}(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y)$ on $\{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ are cycloid arcs. (This problem corresponds to that of the *brachistochrone*, that is, to the problem of finding the curve along which a body slides in a uniform gravitational field to go from one given point to another, not directly below the first one, in the least time.)

The *eikonal equation*,

$$g(\operatorname{grad} S, \operatorname{grad} S) = n^2$$

or, in local coordinates,

$$g^{ij} \frac{\partial S}{\partial x^i} \frac{\partial S}{\partial x^j} = n^2, \tag{6.77}$$

where n is a real-valued function called the *refractive index*, arises in the study of geometrical optics. It can be derived from the Maxwell equations in the short-wavelength limit [see, e.g., Born and Wolf (1999)]. According to Exercise 6.32, the integral curves of grad S, which represent the light rays, are geodesics of the metric tensor $n^2 g$ (with the gradient of S calculated with this metric). Taking into account that the geodesics are the curves that locally minimize the length, defined by the corresponding metric g, and that the refractive index is inversely proportional to the velocity of light in the medium, it follows that the light rays are the curves that locally minimize the time required to go from one point to another. This is known as the *Fermat principle* (see also Sect. 8.4).

6.3 Curvature of a Riemannian Manifold

The algebraic properties of the curvature tensor of a Riemannian connection, as well as the definition of several tensor fields related to it, are more easily established making use of their components with respect to some basis [see, e.g., (5.27)]. It is convenient to define

$$R_{ijkl} \equiv g_{im} R^m{}_{jkl}$$

[cf. (6.5)], so that the skew-symmetry $R^i{}_{jkl} = -R^i{}_{jlk}$ [which is equivalent to $R(\mathbf{X}, \mathbf{Y}) = -R(\mathbf{Y}, \mathbf{X})$] amounts to

$$R_{ijkl} = -R_{ijlk}. \tag{6.78}$$

We have already seen that when the torsion is equal to zero, $R^i{}_{jkl} + R^i{}_{klj} + R^i{}_{ljk} = 0$ [see (5.32)], hence

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0, \tag{6.79}$$

and from (6.66) we have

$$R_{ijkl} = -R_{jikl} \tag{6.80}$$

(see also Exercise 6.23). Given that R is a tensor field, its components with respect to any basis satisfy (6.78)–(6.80).

As a consequence of the relations (6.78)–(6.80), the components of the curvature tensor also satisfy

$$R_{lkij} = R_{ijlk}. \tag{6.81}$$

In fact, from (6.79) and (6.78) we have

$$R_{ijkl} = -R_{iklj} - R_{iljk} = -R_{iklj} + R_{ilkj}. \tag{6.82}$$

On the other hand, from (6.80),

$$R_{ijkl} = \frac{1}{2}(R_{ijkl} - R_{jikl})$$

and expressing each of the terms on the right-hand side with the aid of (6.82), we find

$$R_{ijkl} = \frac{1}{2}(-R_{iklj} + R_{ilkj} + R_{jkli} - R_{jlki}).$$

Thus, exchanging i with k and j with l,

$$R_{klij} = \frac{1}{2}(-R_{kijl} + R_{kjil} + R_{lijk} - R_{ljik}),$$

which coincides with R_{ijkl}, by virtue of (6.78) and (6.80), as claimed above.

Since an object skew-symmetric in two indices has $n(n-1)/2$ independent components, relations (6.78) and (6.80) imply that, out of the n^4 components R_{lkij}, at most $[n(n-1)/2]^2$ are independent; while the relations (6.79), being totally skew-symmetric in the three indices j, k, l, for each value of the first index, constitute $n\binom{n}{3} = n^2(n-1)(n-2)/6$ restrictions. Therefore the number of independent components of the curvature tensor is

$$\frac{n^2(n-1)^2}{4} - \frac{n^2(n-1)(n-2)}{6} = \frac{n^2(n^2-1)}{12}. \tag{6.83}$$

Ricci Tensor, Conformal, and Scalar Curvature From the curvature tensor one can construct other tensor fields which can be conveniently defined in terms of components. The *Ricci tensor* is a tensor field of type $\binom{0}{2}$ with components R_{ij}, defined by

$$R_{ij} \equiv R^k{}_{ikj} = g^{kl} R_{likj}. \tag{6.84}$$

(This definition is not uniform; some authors adopt the definition $R_{ij} = R^k{}_{ijk}$, which amounts, by virtue of (6.78), to $-R^k{}_{ikj}$.) From (6.84) and (6.81) it follows that the Ricci tensor is symmetric:

$$R_{ij} = g^{kl} R_{kjli} = g^{lk} R_{kjli} = R_{ji}.$$

The *scalar curvature*, R, is the real-valued function locally defined by

$$R \equiv g^{ij} R_{ij}. \tag{6.85}$$

For a Riemannian manifold of dimension $n \geq 3$, the *Weyl tensor* or *conformal curvature tensor* is a tensor field with components defined by

$$\begin{aligned} C_{ijkl} \equiv R_{ijkl} &- \frac{1}{n-2}(g_{ik}R_{jl} - g_{jk}R_{il} + g_{jl}R_{ik} - g_{il}R_{jk}) \\ &+ \frac{1}{(n-1)(n-2)} R(g_{ik}g_{jl} - g_{il}g_{jk}). \end{aligned} \tag{6.86}$$

From (6.78)–(6.80), and the symmetry of R_{ij} and g_{ij} it follows that the components of the Weyl tensor (6.86) also satisfy the relations (6.78)–(6.80) and, additionally,

$$g^{kl} C_{kilj} = 0. \tag{6.87}$$

When $n = 3$, the Weyl tensor is identically zero, which amounts to saying that the components of the curvature tensor can be expressed in the form

$$R_{ijkl} = g_{ik}R_{jl} - g_{jk}R_{il} + g_{jl}R_{ik} - g_{il}R_{jk} - \frac{1}{2}R(g_{ik}g_{jl} - g_{il}g_{jk}) \tag{6.88}$$

[cf. (6.86)], so that the curvature tensor is completely determined by the Ricci tensor. The Ricci tensor in a manifold of dimension three, being symmetric, possesses six independent components, which coincides with the number of independent components of the curvature [see (6.83)].

Exercise 6.36 Show that for any Riemannian manifold of dimension three the components of the curvature tensor can be expressed in the form (6.88). (*Hint*: show that if the components C_{ijkl} satisfy (6.78)–(6.80), and (6.87), then $C_{ijkl} = 0$.)

When $n = 2$, the curvature has only one independent component [see (6.83)] and the curvature tensor (and, therefore, the Ricci tensor) is determined by the scalar curvature. In this case, the components of the curvature tensor are given by

$$R_{ijkl} = \frac{1}{2} R (g_{ik} g_{jl} - g_{il} g_{jk}). \tag{6.89}$$

Example 6.37 In the context of general relativity, the Schwarzschild metric, given locally by

$$\begin{aligned} g = {} & \left(1 - \frac{r_g}{r}\right)^{-1} \mathrm{d}r \otimes \mathrm{d}r + r^2 \big(\mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta \, \mathrm{d}\phi \otimes \mathrm{d}\phi\big) \\ & - \left(1 - \frac{r_g}{r}\right) c^2 \, \mathrm{d}t \otimes \mathrm{d}t \end{aligned} \tag{6.90}$$

in terms of a local coordinate system (r, θ, ϕ, t), where r_g is a constant and c is the velocity of light in vacuum, corresponds to the exterior gravitational field of a spherically symmetric distribution of matter. The constant r_g, called the gravitational radius, is related to M, the mass of the matter distribution, by $r_g = 2GM/c^2$, where G is the Newton gravitational constant. From (6.90) we see that the 1-forms

$$\begin{aligned} \theta^1 &= \left(1 - \frac{r_g}{r}\right)^{-1/2} \mathrm{d}r, \quad \theta^2 = r\,\mathrm{d}\theta, \\ \theta^3 &= r \sin\theta \, \mathrm{d}\phi, \quad \theta^4 = \left(1 - \frac{r_g}{r}\right)^{1/2} c\,\mathrm{d}t \end{aligned} \tag{6.91}$$

form the dual basis to a basis such that

$$(g_{ij}) = \mathrm{diag}(1, 1, 1, -1);$$

therefore, e.g., $\mathit{\Gamma}_{134} = \mathit{\Gamma}^1{}_{34}$, but $\mathit{\Gamma}_{412} = -\mathit{\Gamma}^4{}_{12}$ [see (6.60)]. Calculating the exterior derivative of each of the 1-forms (6.91), one finds that the connection 1-forms are

$$\begin{aligned} \mathit{\Gamma}_{12} &= -\left(1 - \frac{r_g}{r}\right)^{1/2} \mathrm{d}\theta, \quad \mathit{\Gamma}_{13} = -\left(1 - \frac{r_g}{r}\right)^{1/2} \sin\theta \, \mathrm{d}\phi, \\ \mathit{\Gamma}_{14} &= \frac{r_g}{2r^2} c \, \mathrm{d}t, \quad \mathit{\Gamma}_{23} = -\cos\theta \, \mathrm{d}\phi, \\ \mathit{\Gamma}_{24} &= 0, \qquad \mathit{\Gamma}_{34} = 0. \end{aligned}$$

The components of the curvature can be obtained making use of the second Cartan structural equations [cf. (5.26)]. One finds that the only nonvanishing components of the curvature 2-forms are given by [see (5.27)]

$$R_{1212} = -\frac{r_g}{2r^3}, \qquad R_{1313} = -\frac{r_g}{2r^3}, \qquad R_{1414} = -\frac{r_g}{r^3},$$
$$R_{2323} = \frac{r_g}{r^3}, \qquad R_{2424} = \frac{r_g}{2r^3}, \qquad R_{3434} = \frac{r_g}{2r^3}.$$

With these expressions and the aid of the properties (6.78) and (6.80) we can now compute the components of the Ricci tensor [see (6.84)]. Since in the present case $(g^{ij}) = \operatorname{diag}(1, 1, 1, -1)$ we have, for instance,

$$R_{11} = g^{ij} R_{i1j1} = R_{1111} + R_{2121} + R_{3131} - R_{4141} = R_{1212} + R_{1313} - R_{1414} = 0.$$

In a similar manner one finds that all the components of the Ricci tensor are equal to zero (for $r \neq 0$), i.e., $R_{ij} = 0$, which are the Einstein equations for the gravitational field in vacuum. Thus, for $r \neq 0$, the Schwarzschild metric (6.90) is a solution of the Einstein vacuum field equations.

Exercise 6.38 Calculate the Ricci tensor of the metric

$$g = \big[f(r)\big]^{-2} \, \mathrm{d}r \otimes \mathrm{d}r + r^2 \big(\mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta \, \mathrm{d}\phi \otimes \mathrm{d}\phi\big),$$

where f is a differentiable real-valued function of a single variable. Show that the Ricci tensor is proportional to the metric tensor, $R_{ij} = h g_{ij}$, where h is some real-valued function, if and only if

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{f^2 - 1}{r^2}\right) = 0$$

(cf. Example 6.24).

Example 6.39 The metric tensor of a Riemannian manifold of dimension two with a positive definite metric has the local expression

$$g = E \, \mathrm{d}x^1 \otimes \mathrm{d}x^1 + F\big(\mathrm{d}x^1 \otimes \mathrm{d}x^2 + \mathrm{d}x^2 \otimes \mathrm{d}x^1\big) + G \, \mathrm{d}x^2 \otimes \mathrm{d}x^2,$$

where E, F, and G are real-valued differentiable functions with $E > 0$ and $EG - F^2 > 0$, or, equivalently

$$g = E\left(\mathrm{d}x^1 + \frac{F}{E}\,\mathrm{d}x^2\right) \otimes \left(\mathrm{d}x^1 + \frac{F}{E}\,\mathrm{d}x^2\right) + \left(G - \frac{F^2}{E}\right) \mathrm{d}x^2 \otimes \mathrm{d}x^2.$$

Since in a manifold of dimension two any 1-form is (locally) integrable, there exist functions, μ and x'^1, such that

$$\mathrm{d}x^1 + \frac{F}{E}\,\mathrm{d}x^2 = \mu \, \mathrm{d}x'^1;$$

hence

$$g = E\mu^2\,\mathrm{d}x'^1 \otimes \mathrm{d}x'^1 + \left(G - \frac{F^2}{E}\right)\mathrm{d}x^2 \otimes \mathrm{d}x^2,$$

thus showing that it is possible to find locally systems of orthogonal coordinates. That is, we may assume that the metric tensor can be written (at least locally) in the form

$$\begin{aligned} g &= E\,\mathrm{d}x^1 \otimes \mathrm{d}x^1 + G\,\mathrm{d}x^2 \otimes \mathrm{d}x^2 \\ &= \big(\sqrt{E}\,\mathrm{d}x^1\big) \otimes \big(\sqrt{E}\,\mathrm{d}x^1\big) + \big(\sqrt{G}\,\mathrm{d}x^2\big) \otimes \big(\sqrt{G}\,\mathrm{d}x^2\big), \end{aligned} \tag{6.92}$$

so that

$$\theta^1 = \sqrt{E}\,\mathrm{d}x^1, \qquad \theta^2 = \sqrt{G}\,\mathrm{d}x^2 \tag{6.93}$$

is the dual basis of an orthonormal basis.

The exterior derivatives of the 1-forms (6.93) are

$$\mathrm{d}\theta^1 = -\frac{1}{2\sqrt{E}}\frac{\partial E}{\partial x^2}\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 = -\frac{1}{2E\sqrt{G}}\frac{\partial E}{\partial x^2}\theta^1 \wedge \theta^2$$

and

$$\mathrm{d}\theta^2 = \frac{1}{2\sqrt{G}}\frac{\partial G}{\partial x^1}\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 = \frac{1}{2G\sqrt{E}}\frac{\partial G}{\partial x^1}\theta^1 \wedge \theta^2.$$

Comparing with (6.64) one finds that $\Gamma^1{}_{12} - \Gamma^1{}_{21} = -\frac{1}{2E\sqrt{G}}\frac{\partial E}{\partial x^2}$, but since the dual basis of (6.93) is orthonormal (i.e., $g_{ij} = \delta_{ij}$), we have $\Gamma^1{}_{12} - \Gamma^1{}_{21} = \Gamma_{112} - \Gamma_{121}$ [see (6.60)]. This reduces to $-\Gamma_{121}$, since the skew-symmetry (6.61) implies that Γ_{112} is equal to zero, and hence we have $\Gamma_{121} = \frac{1}{2E\sqrt{G}}\frac{\partial E}{\partial x^2}$. In a similar way one obtains $\Gamma_{212} = \frac{1}{2G\sqrt{E}}\frac{\partial G}{\partial x^1}$, and therefore the connection 1-forms for the rigid basis (6.93) are determined by

$$\begin{aligned} \Gamma_{12} &= \Gamma_{121}\theta^1 + \Gamma_{122}\theta^2 = \Gamma_{121}\theta^1 - \Gamma_{212}\theta^2 \\ &= \frac{1}{2\sqrt{EG}}\left(\frac{\partial E}{\partial x^2}\,\mathrm{d}x^1 - \frac{\partial G}{\partial x^1}\,\mathrm{d}x^2\right). \end{aligned} \tag{6.94}$$

By virtue of the skew-symmetry $\mathscr{R}_{ij} = -\mathscr{R}_{ji}$ [cf. (6.66)], the curvature is determined by $\mathscr{R}_{12}$ and, according to the second Cartan structural equations (5.26), we have

$$\begin{aligned} \mathscr{R}_{12} &= \mathscr{R}^1{}_2 = \mathrm{d}\Gamma^1{}_2 + \Gamma^1{}_1 \wedge \Gamma^1{}_2 + \Gamma^1{}_2 \wedge \Gamma^2{}_2 = \mathrm{d}\Gamma^1{}_2 = \mathrm{d}\Gamma_{12} \\ &= -\frac{1}{2}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{EG}}\frac{\partial G}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{EG}}\frac{\partial E}{\partial x^2}\right)\right]\mathrm{d}x^1 \wedge \mathrm{d}x^2 \\ &= -\frac{1}{2\sqrt{EG}}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{EG}}\frac{\partial G}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{EG}}\frac{\partial E}{\partial x^2}\right)\right]\theta^1 \wedge \theta^2. \end{aligned}$$

On the other hand, $\mathscr{R}_{12} = \frac{1}{2} R_{12ij}\theta^i \wedge \theta^j = R_{1212}\theta^1 \wedge \theta^2 = \frac{1}{2} R\,\theta^1 \wedge \theta^2$ [see (5.27) and (6.89)], so that the scalar curvature is given by

$$R = -\frac{1}{\sqrt{EG}}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{EG}}\frac{\partial G}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{EG}}\frac{\partial E}{\partial x^2}\right)\right]. \tag{6.95}$$

If, instead of the rigid basis (6.93), we employ the holonomic basis

$$\theta^1 = \mathrm{d}x^1, \qquad \theta^2 = \mathrm{d}x^2, \tag{6.96}$$

the connection 1-forms can be obtained computing the Christoffel symbols (6.55) (with $g_{11} = E$, $g_{12} = 0$, $g_{22} = G$), which turn out to be

$$\begin{aligned} &\Gamma^1_{11} = \frac{1}{2E}\frac{\partial E}{\partial x^1}, \qquad \Gamma^1_{12} = \frac{1}{2E}\frac{\partial E}{\partial x^2}, \qquad \Gamma^1_{22} = -\frac{1}{2E}\frac{\partial G}{\partial x^1}, \\ &\Gamma^2_{11} = -\frac{1}{2G}\frac{\partial E}{\partial x^2}, \qquad \Gamma^2_{12} = \frac{1}{2G}\frac{\partial G}{\partial x^1}, \qquad \Gamma^2_{22} = \frac{1}{2G}\frac{\partial G}{\partial x^2}. \end{aligned} \tag{6.97}$$

Hence, the connection 1-forms for the holonomic basis (6.96), $\Gamma^i{}_j = \Gamma^i_{jk}\,\mathrm{d}x^k$, are

$$\begin{aligned} &\Gamma^1{}_1 = \frac{1}{2E}\,\mathrm{d}E, \quad \Gamma^1{}_2 = \frac{1}{2E}\left(\frac{\partial E}{\partial x^2}\,\mathrm{d}x^1 - \frac{\partial G}{\partial x^1}\,\mathrm{d}x^2\right), \\ &\Gamma^2{}_2 = \frac{1}{2G}\,\mathrm{d}G, \quad \Gamma^2{}_1 = -\frac{1}{2G}\left(\frac{\partial E}{\partial x^2}\,\mathrm{d}x^1 - \frac{\partial G}{\partial x^1}\,\mathrm{d}x^2\right) \end{aligned} \tag{6.98}$$

[cf. (6.94)]. The only independent curvature 2-form is then given by

$$\begin{aligned} \mathscr{R}_{12} &= g_{1i}\mathscr{R}^i{}_2 = E\mathscr{R}^1{}_2 = E\left(\mathrm{d}\Gamma^1{}_2 + \Gamma^1{}_1 \wedge \Gamma^1{}_2 + \Gamma^1{}_2 \wedge \Gamma^2{}_2\right) \\ &= \mathrm{d}\left(E\,\Gamma^1{}_2\right) - \frac{1}{2E}\,\mathrm{d}E \wedge E\,\Gamma^1{}_2 - \frac{1}{2G}\,\mathrm{d}G \wedge E\,\Gamma^1{}_2 \\ &= \sqrt{EG}\,\mathrm{d}\left(\frac{1}{\sqrt{EG}}E\,\Gamma^1{}_2\right) \\ &= \frac{\sqrt{EG}}{2}\,\mathrm{d}\left[\frac{1}{\sqrt{EG}}\left(\frac{\partial E}{\partial x^2}\,\mathrm{d}x^1 - \frac{\partial G}{\partial x^1}\,\mathrm{d}x^2\right)\right] \\ &= -\frac{\sqrt{EG}}{2}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{EG}}\frac{\partial G}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{EG}}\frac{\partial E}{\partial x^2}\right)\right]\mathrm{d}x^1 \wedge \mathrm{d}x^2. \end{aligned} \tag{6.99}$$

Taking (5.27) and (6.89) into account, we have

$$\mathscr{R}_{12} = R_{1212}\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 = \frac{1}{2}R\,g_{11}g_{22}\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 = \frac{1}{2}REG\,\mathrm{d}x^1 \wedge \mathrm{d}x^2,$$

so that from (6.99) we obtain again the expression (6.95) for the scalar curvature, as we should.

A *constant curvature* Riemannian manifold is a Riemannian manifold of dimension greater than two such that the components of the curvature tensor are of the form

$$R_{ijkl} = \frac{1}{n(n-1)} R\,(g_{ik}g_{jl} - g_{il}g_{jk}), \tag{6.100}$$

where R is the scalar curvature (cf. Example 6.25). From (5.27) it follows that the curvature 2-forms are $\mathscr{R}^i{}_j = R/[n(n-1)]\, g_{jl}\theta^i \wedge \theta^l$ (no matter what type of basis is used). Substituting this relation into the Bianchi identities (5.31) and using the first Cartan structural equations (5.24) (with the torsion equal to zero) one obtains $\mathrm{d}R \wedge \theta^i \wedge \theta^l = 0$ [the computation is simpler making use of a rigid basis, with the aid of the relation (6.65)]; then, since $n > 2$, it follows that R is constant. (It may be noticed that the curvature of a Riemannian manifold of dimension two is always of the form (6.100), but $\mathrm{d}R \wedge \theta^i \wedge \theta^l$ is necessarily equal to zero, because it is a 3-form; therefore, in this case, the Bianchi identities do not imply that R is a constant.)

Exercise 6.40 Show that the scalar curvature of the sphere [equation (6.11)] and of the Poincaré half-plane [equation (6.19)] is constant.

Apart from the fact that in a Riemannian manifold of dimension two with a positive definite metric one can always find orthogonal coordinates, where the metric tensor takes the "diagonal" form (6.92), it is also possible to find local coordinates where the metric tensor has the form (6.92) with $E = G$, i.e., any metric tensor of this class is locally conformally equivalent to a flat metric (and such a system of coordinates is not unique). This assertion can readily be proved making use of complex combinations of 1-forms. Writing

$$\begin{aligned} &E\,\mathrm{d}x^1 \otimes \mathrm{d}x^1 + G\,\mathrm{d}x^2 \otimes \mathrm{d}x^2 \\ &\quad = \frac{1}{2}\big[\big(\sqrt{E}\,\mathrm{d}x^1 + \mathrm{i}\sqrt{G}\,\mathrm{d}x^2\big) \otimes \big(\sqrt{E}\,\mathrm{d}x^1 - \mathrm{i}\sqrt{G}\,\mathrm{d}x^2\big) \\ &\qquad + \big(\sqrt{E}\,\mathrm{d}x^1 - \mathrm{i}\sqrt{G}\,\mathrm{d}x^2\big) \otimes \big(\sqrt{E}\,\mathrm{d}x^1 + \mathrm{i}\sqrt{G}\,\mathrm{d}x^2\big)\big] \end{aligned}$$

and taking into account that $\sqrt{E}\,\mathrm{d}x^1 + \mathrm{i}\sqrt{G}\,\mathrm{d}x^2$ is a (complexified) 1-form in two variables, it is locally integrable; that is, locally there exist complex-valued functions A, B such that $\sqrt{E}\,\mathrm{d}x^1 + \mathrm{i}\sqrt{G}\,\mathrm{d}x^2 = A\,\mathrm{d}B$ (though these functions are not unique, see the example below). Letting $B = y^1 + \mathrm{i}y^2$, with y^1, y^2 being real-valued functions, we obtain

$$\begin{aligned} E\,\mathrm{d}x^1 \otimes \mathrm{d}x^1 + G\,\mathrm{d}x^2 \otimes \mathrm{d}x^2 &= \frac{1}{2}(A\,\mathrm{d}B \otimes \overline{A}\,\mathrm{d}\overline{B} + \overline{A}\,\mathrm{d}\overline{B} \otimes A\,\mathrm{d}B) \\ &= |A|^2\big(\mathrm{d}y^1 \otimes \mathrm{d}y^1 + \mathrm{d}y^2 \otimes \mathrm{d}y^2\big), \end{aligned}$$

where the bar denotes complex conjugation, thus showing that the metric tensor is proportional to the flat metric $\mathrm{d}y^1 \otimes \mathrm{d}y^1 + \mathrm{d}y^2 \otimes \mathrm{d}y^2$.

For example, the standard metric of the sphere S^2 is locally given by $\mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta\, \mathrm{d}\phi \otimes \mathrm{d}\phi$ [see (6.11)], which is already of the form (6.92). An integrating factor for the 1-form $\mathrm{d}\theta + \mathrm{i}\sin\theta\, \mathrm{d}\phi$ can be found by inspection, namely

$$\begin{aligned}\mathrm{d}\theta + \mathrm{i}\sin\theta\, \mathrm{d}\phi &= \sin\theta(\csc\theta\, \mathrm{d}\theta + \mathrm{i}\, \mathrm{d}\phi)\\ &= \sin\theta\, \mathrm{d}\left(\ln\tan\frac{1}{2}\theta + \mathrm{i}\phi\right),\end{aligned}$$

giving a possible choice for the local coordinates y^1, y^2 (that is, $y^1 = \ln\tan\frac{1}{2}\theta$, $y^2 = \phi$). However, a more convenient choice is obtained on taking

$$\begin{aligned}\mathrm{d}\theta + \mathrm{i}\sin\theta\, \mathrm{d}\phi &= \sin\theta\, \mathrm{d}\ln\left(\mathrm{e}^{\mathrm{i}\phi}\tan\frac{1}{2}\theta\right)\\ &= \frac{\sin\theta}{\mathrm{e}^{\mathrm{i}\phi}\tan\frac{1}{2}\theta}\, \mathrm{d}\left(\mathrm{e}^{\mathrm{i}\phi}\tan\frac{1}{2}\theta\right)\\ &= 2\mathrm{e}^{-\mathrm{i}\phi}\cos^2\frac{1}{2}\theta\, \mathrm{d}\left(\mathrm{e}^{\mathrm{i}\phi}\tan\frac{1}{2}\theta\right);\end{aligned}$$

hence $\mathrm{d}\theta \otimes \mathrm{d}\theta + \sin^2\theta\, \mathrm{d}\phi \otimes \mathrm{d}\phi = 4\cos^4\frac{1}{2}\theta[\mathrm{d}(\tan\frac{1}{2}\theta\cos\phi) \otimes \mathrm{d}(\tan\frac{1}{2}\theta\cos\phi) + \mathrm{d}(\tan\frac{1}{2}\theta\sin\phi) \otimes \mathrm{d}(\tan\frac{1}{2}\theta\sin\phi)]$ (cf. Example 6.19).

6.4 Volume Element, Divergence, and Duality of Differential Forms

Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \subset \mathfrak{X}(M)$ be an orthonormal basis for the vector fields on M, that is, $g(\mathbf{e}_i, \mathbf{e}_j) = \pm\delta_{ij}$. There exists an n-form η on M, called a *volume element*, such that $\eta(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1/n!$. In fact, if $\{\theta^1, \ldots, \theta^n\}$ is the dual basis to $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ we have

$$\eta = \theta^1 \wedge \theta^2 \wedge \cdots \wedge \theta^n. \tag{6.101}$$

If $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ is any other orthonormal basis of vector fields, then $\eta(\mathbf{e}'_1, \ldots, \mathbf{e}'_n) = \pm 1/n!$. We say that $\{\mathbf{e}'_1, \ldots, \mathbf{e}'_n\}$ is positively or negatively oriented according to η, if $\eta(\mathbf{e}'_1, \ldots, \mathbf{e}'_n)$ is greater or less than zero, respectively. The n-form $-\eta$ is another volume element defining the opposite orientation to that defined by η.

If the 1-forms θ^i are given locally by $\theta^i = c^i_j\, \mathrm{d}x^j$, then

$$\begin{aligned}\eta &= \left(c^1_i\, \mathrm{d}x^i\right) \wedge \left(c^2_j\, \mathrm{d}x^j\right) \wedge \cdots \wedge \left(c^n_k\, \mathrm{d}x^k\right)\\ &= c^1_i c^2_j \cdots c^n_k\, \mathrm{d}x^i \wedge \mathrm{d}x^j \wedge \cdots \wedge \mathrm{d}x^k\\ &= c^1_i c^2_j \cdots c^n_k \varepsilon^{ij\ldots k}\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n\\ &= \det\left(c^i_j\right) \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n.\end{aligned}$$

On the other hand, we have the matrix relation

$$\pm\delta_{ij} = \left(\theta^i \middle| \theta^j\right) = c^i_k c^j_l \, g^{kl};$$

hence, the determinants of these matrices are related by

$$\left[\det\left(c^i_j\right)\right]^2 \det\left(g^{kl}\right) = \pm 1,$$

i.e.,

$$\det\left(c^i_j\right) = \pm\sqrt{\left|\det(g_{ij})\right|},$$

where the sign is positive or negative according to whether the basis $\{\partial/\partial x^1, \ldots, \partial/\partial x^n\}$ is positively or negatively oriented according to η. Thus,

$$\eta = \pm\sqrt{\left|\det(g_{ij})\right|}\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n. \tag{6.102}$$

For some manifolds there does not exist a nowhere vanishing n-form defined on all of M; such manifolds are called *non orientable*. A manifold M of dimension n is *orientable* if there exists a nowhere vanishing n-form defined at all points of M. (This property does not depend on the existence of a Riemannian structure on M, but is a topological property of the manifold.) If M is an orientable Riemannian manifold, then there exists a volume element defined at all points of M.

Example 6.41 As shown in Example 6.39, in a Riemannian manifold, M, of dimension two, with a positive definite metric, there exist systems of orthogonal coordinates, in which the metric tensor has the diagonal form (6.92). Using the Christoffel symbols (6.97) one finds that the equations for the parallel transport of a vector (5.4) are

$$\frac{\mathrm{d}Y^1}{\mathrm{d}t} + \frac{1}{2E}\left[\frac{\partial E}{\partial x^1}\frac{\mathrm{d}x^1}{\mathrm{d}t}Y^1 + \frac{\partial E}{\partial x^2}\left(\frac{\mathrm{d}x^2}{\mathrm{d}t}Y^1 + \frac{\mathrm{d}x^1}{\mathrm{d}t}Y^2\right) - \frac{\partial G}{\partial x^1}\frac{\mathrm{d}x^2}{\mathrm{d}t}Y^2\right] = 0,$$

$$\frac{\mathrm{d}Y^2}{\mathrm{d}t} + \frac{1}{2G}\left[\frac{\partial G}{\partial x^2}\frac{\mathrm{d}x^2}{\mathrm{d}t}Y^2 + \frac{\partial G}{\partial x^1}\left(\frac{\mathrm{d}x^1}{\mathrm{d}t}Y^2 + \frac{\mathrm{d}x^2}{\mathrm{d}t}Y^1\right) - \frac{\partial E}{\partial x^2}\frac{\mathrm{d}x^1}{\mathrm{d}t}Y^1\right] = 0,$$

or, equivalently,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left(\sqrt{E}\,Y^1 + \mathrm{i}\sqrt{G}\,Y^2\right) \\
&\quad = \frac{\mathrm{i}}{2\sqrt{EG}}\left(\frac{\partial E}{\partial x^2}\frac{\mathrm{d}x^1}{\mathrm{d}t} - \frac{\partial G}{\partial x^1}\frac{\mathrm{d}x^2}{\mathrm{d}t}\right)\left(\sqrt{E}\,Y^1 + \mathrm{i}\sqrt{G}\,Y^2\right).
\end{aligned}$$

Then,

$$\begin{aligned}&\big(\sqrt{E}\,Y^1 + \mathrm{i}\sqrt{G}\,Y^2\big)\big(C(t)\big)\\&\quad = \big(\sqrt{E}\,Y^1 + \mathrm{i}\sqrt{G}\,Y^2\big)\big(C(t_0)\big)\exp\frac{\mathrm{i}}{2}\int_{t_0}^{t}\frac{1}{\sqrt{EG}}\left(\frac{\partial E}{\partial x^2}\frac{\mathrm{d}x^1}{\mathrm{d}t} - \frac{\partial G}{\partial x^1}\frac{\mathrm{d}x^2}{\mathrm{d}t}\right)\mathrm{d}t.\end{aligned}$$

The value of the line integral appearing in the preceding formula only depends on the endpoints of the curve, $C(t_0)$ and $C(t)$, if and only if the 1-form

$$\gamma \equiv \frac{1}{2\sqrt{EG}}\left(\frac{\partial E}{\partial x^2}\,\mathrm{d}x^1 - \frac{\partial G}{\partial x^1}\,\mathrm{d}x^2\right)$$

is exact [cf. (6.94)]. By means of a direct computation we see that

$$\mathrm{d}\gamma = -\frac{1}{2}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{EG}}\frac{\partial G}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{EG}}\frac{\partial E}{\partial x^2}\right)\right]\mathrm{d}x^1\wedge\mathrm{d}x^2 = \frac{R}{2}\eta, \qquad (6.103)$$

where R is the scalar curvature and $\eta = \theta^1\wedge\theta^2 = \sqrt{EG}\,\mathrm{d}x^1\wedge\mathrm{d}x^2$ is a volume element [see (6.95), (6.93), and (6.101)]. Hence, if $R \neq 0$, the 1-form γ is not closed.

In particular, if C is a simple closed curve, with $C(t_0) = C(t_1)$, then

$$\begin{pmatrix}(\sqrt{E}\,Y^1) & (C(t_1))\\ (\sqrt{G}\,Y^2) & (C(t_1))\end{pmatrix} = \begin{pmatrix}\cos\Theta & -\sin\Theta\\ \sin\Theta & \cos\Theta\end{pmatrix}\begin{pmatrix}(\sqrt{E}\,Y^1) & (C(t_0))\\ (\sqrt{G}\,Y^2) & (C(t_0))\end{pmatrix}, \qquad (6.104)$$

with

$$\Theta \equiv \oint_C \gamma = \int_{t_0}^{t_1}\frac{1}{2\sqrt{EG}}\left(\frac{\partial E}{\partial x^2}\frac{\mathrm{d}x^1}{\mathrm{d}t} - \frac{\partial G}{\partial x^1}\frac{\mathrm{d}x^2}{\mathrm{d}t}\right)\mathrm{d}t. \qquad (6.105)$$

The functions $\sqrt{E}\,Y^1$ and $\sqrt{G}\,Y^2$ appearing in (6.104) are the components of the vector field $\mathbf{Y}$ with respect to the *orthonormal basis* $\mathbf{e}_1 = (1/\sqrt{E})\,\partial/\partial x^1$, $\mathbf{e}_2 = (1/\sqrt{G})\,\partial/\partial x^2$ [cf. (6.93)], and therefore (6.104) represents a rotation through the angle Θ at $T_{C(t_0)}M$. That is, the parallel transport of any vector along a closed curve C only rotates the original vector through the angle Θ, with Θ being independent of the vector chosen and of the point of the curve taken as the initial point [see (6.105)]. (Example 5.5 is a particular case of the present example, with $E = G = 1/y^2$.)

The fact that the parallel transport of a vector along a closed curve corresponds to a rotation is to be expected, because the Riemannian connection is compatible with the metric tensor, so that during the parallel transport of a vector, its length does not vary. The parallel transport of vectors along a curve is a linear transformation (see Sect. 5.1) and the only linear transformations of a space with inner product into itself that preserve the inner product are rotations or reflections.

Divergence of a Vector Field If η is a volume element and $\mathbf{X} \in \mathfrak{X}(M)$, the Lie derivative of η with respect to $\mathbf{X}$ is also an n-form and therefore there exists a real-valued function, $\operatorname{div} \mathbf{X}$, the *divergence* of $\mathbf{X}$, such that

$$\pounds_{\mathbf{X}} \eta = (\operatorname{div} \mathbf{X}) \eta. \tag{6.106}$$

The definition of the divergence of a vector field does not depend on the orientation, since, if η is substituted by $-\eta$ into (6.106), the value of $\operatorname{div} \mathbf{X}$ does not change. Using (3.39) and taking into account that $\mathrm{d}\eta = 0$, because it is an $(n+1)$-form, the definition (6.106) amounts to

$$(\operatorname{div} \mathbf{X}) \eta = \mathrm{d}(\mathbf{X} \lrcorner\, \eta). \tag{6.107}$$

Using the local expression of the volume element (6.102) and equations (2.23), (2.37)–(2.39), and (3.26), we find that

$$\begin{aligned}
\pounds_{\mathbf{X}} \eta &= \pm \Bigg[X^k \frac{\partial \sqrt{\det(g_{ij})}}{\partial x^k} \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n \\
&\quad + \sqrt{\det(g_{ij})} \left(\frac{\partial X^1}{\partial x^k} \mathrm{d}x^k \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n + \cdots + \mathrm{d}x^1 \wedge \cdots \wedge \frac{\partial X^n}{\partial x^k} \mathrm{d}x^k \right) \Bigg] \\
&= \pm \left(X^k \frac{\partial \sqrt{\det(g_{ij})}}{\partial x^k} + \sqrt{\det(g_{ij})} \frac{\partial X^k}{\partial x^k} \right) \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n \\
&= \left[\frac{1}{\sqrt{\det(g_{ij})}} \frac{\partial}{\partial x^k} \left(\sqrt{\det(g_{ij})}\, X^k \right) \right] \eta
\end{aligned}$$

and, comparing with (6.106), we obtain the well-known expression

$$\operatorname{div} \mathbf{X} = \frac{1}{\sqrt{\det(g_{ij})}} \frac{\partial}{\partial x^k} \left(\sqrt{\det(g_{ij})}\, X^k \right) \tag{6.108}$$

for the divergence of a vector field in terms of its components and those of the metric tensor with respect to a coordinate system $(x^1, \ldots, x^n)$.

Exercise 6.42 Show that if $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthonormal basis and $\mathbf{X} = X^i \mathbf{e}_i$ is a differentiable vector field, then

$$\operatorname{div} \mathbf{X} = \mathbf{e}_k X^k + \Gamma^i{}_{ki} X^k, \tag{6.109}$$

where $\Gamma^i{}_{jk}$ are the Ricci rotation coefficients for the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. (*Hint*: employ (6.101), (3.39), (6.64), (3.27), and (6.61).)

Using the expression (6.108) or (6.109) it follows that for $f \in C^\infty(M)$, $\mathbf{X} \in \mathfrak{X}(M)$,

$$\operatorname{div}(f\mathbf{X}) = f \operatorname{div} \mathbf{X} + \mathbf{X} f \tag{6.110}$$

or, equivalently [see (6.8)],

$$\operatorname{div}(f\mathbf{X}) = f \operatorname{div}\mathbf{X} + g(\operatorname{grad} f, \mathbf{X}). \tag{6.111}$$

From (6.108) or (6.109) it also follows that $\operatorname{div}(\mathbf{X}+\mathbf{Y}) = \operatorname{div}\mathbf{X} + \operatorname{div}\mathbf{Y}$.

Exercise 6.43 Show that $\operatorname{div}[\mathbf{X}, \mathbf{Y}] = \mathbf{X}(\operatorname{div}\mathbf{Y}) - \mathbf{Y}(\operatorname{div}\mathbf{X})$.

The *Laplacian* of a differentiable function $f \in C^\infty(M)$, denoted by $\nabla^2 f$ or by Δf, is defined as the divergence of its gradient

$$\nabla^2 f \equiv \operatorname{div} \operatorname{grad} f. \tag{6.112}$$

From the expressions (6.9) and (6.108) it follows that

$$\nabla^2 f = \frac{1}{\sqrt{\det(g_{ij})}} \frac{\partial}{\partial x^k}\left(\sqrt{\det(g_{ij})}\, g^{kl} \frac{\partial f}{\partial x^l}\right). \tag{6.113}$$

Exercise 6.44 Show that under a conformal rescaling of the metric of a Riemannian manifold of dimension two, $g' = \mathrm{e}^{2u} g$, the scalar curvatures of g and g' are related by

$$R - 2\Delta u = \mathrm{e}^{2u} R'. \tag{6.114}$$

In the case of the sphere S^2, with its standard metric, the scalar curvature is $R = 2$; the existence of a solution to (6.114) with $R' = 0$ would mean that the standard metric of S^2 is conformally flat (that is, conformally equivalent to a flat metric). However, the PDE $\Delta u = 1$ has no solution on S^2 (cf. Example 6.19 and the example at the end of Sect. 6.3). This fact can be proved by integrating both sides of this equation on S^2, making use of the natural area element of S^2, so that the integral of 1 yields the total area of S^2, i.e., 4π. Meanwhile the integral of Δu, being the integral of a divergence on a surface without boundary, is equal to zero.

Duality of Differential Forms Let $\alpha \in \Lambda^k(M)$ and $\beta \in \Lambda^{n-k}(M)$; the exterior product $\alpha \wedge \beta$ is an n-form and, therefore, there exists a function $f \in C^\infty(M)$ such that $\alpha \wedge \beta = f\,\eta$. If $(x^1, \ldots, x^n)$ is a local coordinate system positively oriented according to η, α, and β are given by $\alpha = \alpha_{i_1 \ldots i_k}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k}$ and $\beta = \beta_{j_{k+1} \cdots j_n}\, \mathrm{d}x^{j_{k+1}} \wedge \cdots \wedge \mathrm{d}x^{j_n}$; then we have

$$\begin{aligned}
\alpha \wedge \beta &= \alpha_{i_1 \ldots i_k} \beta_{j_{k+1} \ldots j_n}\, \mathrm{d}x^{i_1} \wedge \cdots \wedge \mathrm{d}x^{i_k} \wedge \mathrm{d}x^{j_{k+1}} \wedge \cdots \wedge \mathrm{d}x^{j_n} \\
&= \alpha_{i_1 \ldots i_k} \beta_{j_{k+1} \ldots j_n} \varepsilon^{i_1 \ldots i_k j_{k+1} \ldots j_n}\, \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n \\
&= \alpha_{i_1 \ldots i_k} \beta_{j_{k+1} \ldots j_n} \varepsilon^{i_1 \ldots i_k j_{k+1} \ldots j_n} \left|\det(g_{ij})\right|^{-1/2} \eta,
\end{aligned}$$

that is,

$$f = \left|\det(g_{ij})\right|^{-1/2} \varepsilon^{i_1 \ldots i_k j_{k+1} \ldots j_n} \alpha_{i_1 \ldots i_k} \beta_{j_{k+1} \ldots j_n}.$$

Noting that

$$\varepsilon_{i_1\ldots i_n} g^{i_1 j_1} \cdots g^{i_n j_n} = \det\bigl(g^{ij}\bigr)\varepsilon^{j_1\cdots j_n}$$
$$= \bigl[\det(g_{ij})\bigr]^{-1}\varepsilon^{j_1\cdots j_n},$$

we also have

$$f = \bigl|\det(g_{ij})\bigr|^{1/2}\varepsilon_{l_1\ldots l_n} g^{i_1 l_1} \cdots g^{i_k l_k} g^{j_{k+1} l_{k+1}} \cdots g^{j_n l_n} \alpha_{i_1\ldots i_k} \beta_{j_{k+1}\cdots j_n}.$$

Therefore, there exists a unique $(n-k)$-form, denoted by $*\alpha$, such that $f = (*\alpha|\beta)$; in fact, $*\alpha$ is given locally by

$$*\alpha = \frac{|\det(g_{ij})|^{1/2}}{(n-k)!}\varepsilon_{l_1\ldots l_n} g^{i_1 l_1} \cdots g^{i_k l_k} \alpha_{i_1\ldots i_k}\, \mathrm{d}x^{l_{k+1}} \wedge \cdots \wedge \mathrm{d}x^{l_n}.$$

The uniqueness of $*\alpha$ comes from the fact that the product $(\ |\ )$ is non-singular.

The mapping $* : \Lambda^k(M) \to \Lambda^{n-k}(M)$, given by $\alpha \mapsto *\alpha$, is called the *star* or *Hodge operator*, and $*\alpha$ is called the *dual form* of α. From its local expression we see that the Hodge operator is an isomorphism of $\Lambda^k(M)$ onto $\Lambda^{n-k}(M)$, that is, any $(n-k)$-form is the dual of a unique k-form in such a way that

$$*(f\omega_1 + g\omega_2) = f * \omega_1 + g * \omega_2 \quad \text{for } \omega_1, \omega_2 \in \Lambda^k(M) \text{ and } f, g \in C^\infty(M).$$

Exercise 6.45 Let $\mathbf{X}$ be a vector field and let $\alpha = \frac{1}{2}\mathbf{X}\lrcorner g$. Show that $*\alpha = \mathbf{X}\lrcorner\,\eta$.

Note that if the orientation is reversed, that is, if η is replaced by $-\eta$, then $*\alpha$ changes sign. Owing to this behavior $*\alpha$ is said to be a *pseudotensor field*.

6.5 Elementary Treatment of the Geometry of Surfaces

The theory of Riemannian manifolds started with the study of (two-dimensional) surfaces in $\mathbb{R}^3$. Here we shall present only an introductory study of surfaces, as an example of the usefulness of the formalism already given. We shall be interested mainly in two-dimensional submanifolds of a three-dimensional Riemannian manifold, which may not be the Euclidean space. More detailed treatments can be found, e.g., in do Carmo (1992), Oprea (1997), and O'Neill (2006).

Let M be a Riemannian manifold of dimension three, with a positive definite metric and let Σ be a submanifold of M of dimension two, with the metric induced by that of M. Among other things, we want to relate the intrinsic properties of Σ (that is, the properties of Σ as a Riemannian manifold on its own) with the behavior of a unit normal vector field to Σ.

Let $p \in \Sigma$ and let $\mathbf{n}$ be a unit normal vector field to Σ defined in a neighborhood of p. The *shape operator* (or *Weingarten map*) of Σ at p, S_p, is defined by

$$S_p(v_p) \equiv -\nabla_{v_p}\mathbf{n}, \quad \text{for } v_p \in T_p\Sigma. \tag{6.115}$$

Roughly speaking, $S_p(v_p)$ measures how quickly Σ bends in the direction of v_p.

It can readily be seen that, as a consequence of (6.53), $S_p(v_p)$ also belongs to $T_p\Sigma$, since

$$g_p\big(S_p(v_p), \mathbf{n}_p\big) = -g_p(\nabla_{v_p}\mathbf{n}, \mathbf{n}_p) = -\frac{1}{2}v_p\big[g(\mathbf{n}, \mathbf{n})\big] = -\frac{1}{2}v_p[1] = 0.$$

By virtue of the properties of a connection, S_p is a linear map. The *Gaussian curvature* and the *mean curvature* of Σ at p are defined as

$$K(p) = \det S_p \quad \text{and} \quad H(p) = \frac{1}{2}\operatorname{tr} S_p,$$

respectively.

The shape operator S_p is also *symmetric*, in the sense that

$$g_p\big(S_p(v_p), w_p\big) = g_p\big(v_p, S_p(w_p)\big),$$

for all $v_p, w_p \in T_p\Sigma$.

In order to prove that S_p is symmetric, we shall use the fact that for each point $p \in \Sigma$ one can find an orthonormal set of vector fields $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, defined in some neighborhood of p, such that at the points of Σ, $\mathbf{e}_1$, and $\mathbf{e}_2$ span the tangent space to Σ and, therefore, restricted to Σ we see that $\mathbf{e}_3$ is a unit normal vector field to Σ. Then the vector fields $\mathbf{e}_1$ and $\mathbf{e}_2$, restricted to Σ, form an orthonormal basis for the vector fields on Σ. Thus,

$$i^*\theta^3 = 0, \tag{6.116}$$

where $i : \Sigma \to M$ is the inclusion map; therefore, $i^*\,\mathrm{d}\theta^3 = \mathrm{d}(i^*\theta^3) = 0$, and the equation $\mathrm{d}\theta^3 = \Gamma^3{}_{ij}\theta^i \wedge \theta^j = \Gamma_{3ij}\theta^i \wedge \theta^j$ [see (6.64)] yields

$$i^*(\Gamma_{312} - \Gamma_{321}) = 0. \tag{6.117}$$

Making use of (5.22) and the skew-symmetry (6.61) we have, letting the lowercase Greek indices $\mu, \nu, \ldots$ take the values 1, 2,

$$\nabla_{\mathbf{e}_\mu}\mathbf{e}_3 = \Gamma^i{}_{3\mu}\mathbf{e}_i = \Gamma^\nu{}_{3\mu}\mathbf{e}_\nu + \Gamma^3{}_{3\mu}\mathbf{e}_3 = \Gamma^\nu{}_{3\mu}\mathbf{e}_\nu. \tag{6.118}$$

By comparing this equation with the definition of the shape operator, (6.115), one finds that, with respect to the orthonormal basis of $T_p\Sigma$ formed by $(\mathbf{e}_1)_p$ and $(\mathbf{e}_2)_p$, S_p is represented by the 2×2 matrix

$$\big(-\Gamma^\nu{}_{3\mu}\big(i(p)\big)\big) = \big(\Gamma_{3\nu\mu}\big(i(p)\big)\big). \tag{6.119}$$

Hence, (6.117) means that this matrix is symmetric and, therefore, S_p is symmetric, as claimed above.

The symmetry of S_p implies the existence of two linearly independent eigenvectors, whose directions are called the *principal curvature directions* of Σ at p and the corresponding eigenvalues are called principal curvatures at p.

The 1-forms

$$\phi^\mu \equiv i^* \theta^\mu \tag{6.120}$$

constitute the dual basis of the orthonormal basis formed by the restriction of $\mathbf{e}_1$ and $\mathbf{e}_2$ to Σ. Making use of the properties of the pullback, the first Cartan structural equations, and (6.116), we have

$$\mathrm{d}\phi^\mu = i^*\big(\mathrm{d}\theta^\mu\big) = -i^*\big(\Gamma^\mu{}_i \wedge \theta^i\big) = -\big(i^*\Gamma^\mu{}_\nu\big) \wedge \phi^\nu,$$

which implies that the 1-forms $i^*\Gamma^\mu{}_\nu$ are the connection 1-forms associated with the orthonormal basis (6.120).

In a similar manner, we can compute the pullback under the inclusion map of the second Cartan structural equations. We start by computing $i^*\mathscr{R}^\nu{}_\mu$, for $\mu, \nu = 1, 2$. Owing to the skew-symmetry (6.66), this reduces to a computation, e.g., of $i^*\mathscr{R}^1{}_2$ and, making use of the fact that $\Gamma^1{}_1$ and $\Gamma^2{}_2$ are equal to zero by virtue of the skew-symmetry (6.65), we find

$$i^*\mathscr{R}^1{}_2 = \mathrm{d}\big(i^*\Gamma^1{}_2\big) + \big(i^*\Gamma^1{}_j\big) \wedge \big(i^*\Gamma^j{}_2\big) = \mathrm{d}\big(i^*\Gamma^1{}_2\big) + \big(i^*\Gamma^1{}_3\big) \wedge \big(i^*\Gamma^3{}_2\big). \tag{6.121}$$

On the other hand, applying the second Cartan structural equations to compute the curvature of Σ, if we denote by $\Omega^\mu{}_\nu$ the curvature 2-forms of Σ with respect to the basis (6.120), we have $\Omega^1{}_2 = \mathrm{d}(i^*\Gamma^1{}_2) + (i^*\Gamma^1{}_\mu) \wedge (i^*\Gamma^\mu{}_2) = \mathrm{d}(i^*\Gamma^1{}_2)$. Hence, (6.121) amounts to the relation

$$i^*\mathscr{R}^1{}_2 = \Omega^1{}_2 + \big(i^*\Gamma^1{}_3\big) \wedge \big(i^*\Gamma^3{}_2\big) \tag{6.122}$$

or, equivalently [see (5.27) and (6.116)],

$$\begin{aligned}\big(i^*R^1{}_{212}\big)\phi^1 \wedge \phi^2 = \Omega^1{}_{212}\phi^1 \wedge \phi^2 + \big[\big(i^*\Gamma^1{}_{31}\big)\big(i^*\Gamma^3{}_{22}\big) \\ - \big(i^*\Gamma^1{}_{32}\big)\big(i^*\Gamma^3{}_{21}\big)\big]\phi^1 \wedge \phi^2.\end{aligned}$$

Hence, taking into account (6.119),

$$i^*R^1{}_{212} = \Omega^1{}_{212} - \det S = \Omega^1{}_{212} - K. \tag{6.123}$$

When the curvature of M is equal to zero, as in the case of the Euclidean space with its standard metric, equation (6.123) gives $K = \Omega^1{}_{212}$; that is, the Gaussian curvature, K, defined above in terms of the (*extrinsic*) behavior of the unit normal vector field to Σ, is equal to the *intrinsic* curvature of Σ, defined by the Riemannian connection of Σ. This result is the famous Gauss' *Theorema Egregium*, which states that the Gaussian curvature of Σ depends only on the metric induced on the surface. Thus, according to (6.95), if the metric induced on Σ is expressed in the form $E\,\mathrm{d}x^1 \otimes \mathrm{d}x^1 + G\,\mathrm{d}x^2 \otimes \mathrm{d}x^2$, the Gaussian curvature of Σ is given by

$$K = -\frac{1}{2\sqrt{EG}}\left[\frac{\partial}{\partial x^1}\left(\frac{1}{\sqrt{EG}}\frac{\partial G}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(\frac{1}{\sqrt{EG}}\frac{\partial E}{\partial x^2}\right)\right].$$

Since $\mathscr{R}^3{}_3$ identically vanishes, we are left with $\mathscr{R}^3{}_\mu$ only and we obtain

$$\begin{aligned}
i^*\mathscr{R}^3{}_\mu &= \mathrm{d}\big(i^*\Gamma^3{}_\mu\big) + \big(i^*\Gamma^3{}_j\big) \wedge \big(i^*\Gamma^j{}_\mu\big) \\
&= \mathrm{d}\big(i^*\Gamma^3{}_\mu\big) + \big(i^*\Gamma^3{}_\nu\big) \wedge \big(i^*\Gamma^\nu{}_\mu\big) \\
&= \mathrm{d}\big[\big(i^*\Gamma^3{}_{\mu\nu}\big)\phi^\nu\big] + \big(i^*\Gamma^3{}_{\nu\rho}\big)\phi^\rho \wedge \big(i^*\Gamma^\nu{}_{\mu\sigma}\big)\phi^\sigma \\
&= \mathrm{d}\big(i^*\Gamma^3{}_{\mu\nu}\big) \wedge \phi^\nu + \big(i^*\Gamma^3{}_{\mu\nu}\big)\big(i^*\Gamma^\nu{}_{\rho\sigma}\big)\phi^\rho \wedge \phi^\sigma \\
&\quad + \big(i^*\Gamma^3{}_{\nu\rho}\big)\big(i^*\Gamma^\nu{}_{\mu\sigma}\big)\phi^\rho \wedge \phi^\sigma .
\end{aligned}$$

Denoting by $\{\mathbf{X}_1, \mathbf{X}_2\}$ the dual basis of that formed by the 1-forms (6.120), the last equation amounts to

$$\begin{aligned}
i^*R^3{}_{\mu 12} = {}& \mathbf{X}_1\big(i^*\Gamma^3{}_{\mu 2}\big) - \mathbf{X}_2\big(i^*\Gamma^3{}_{\mu 1}\big) \\
& + \big(i^*\Gamma^3{}_{\mu\nu}\big)\big(i^*\Gamma^\nu{}_{12}\big) - \big(i^*\Gamma^3{}_{\mu\nu}\big)\big(i^*\Gamma^\nu{}_{21}\big) \\
& + \big(i^*\Gamma^3{}_{\nu 1}\big)\big(i^*\Gamma^\nu{}_{\mu 2}\big) - \big(i^*\Gamma^3{}_{\nu 2}\big)\big(i^*\Gamma^\nu{}_{\mu 1}\big).
\end{aligned} \tag{6.124}$$

These equations are known as the *Codazzi–Mainardi equations* [cf. Oprea (1997, Sect. 3.4)].

Exercise 6.46 Assuming that the curvature of M is equal to zero, show that the Codazzi–Mainardi equations (6.124) are equivalent to the symmetry

$$(\nabla_{\mathbf{X}_1} S)(X_2) = (\nabla_{\mathbf{X}_2} S)(X_1),$$

where ∇ denotes the Riemannian connection of Σ and S is considered as a tensor field on Σ of type $\binom{1}{1}$.

Example 6.47 The *catenoid* is a well-known example of a *minimal surface*, that is, a surface with mean curvature equal to zero. This is a surface of revolution obtained by revolving a catenary, and can be defined by means of the parametrization

$$x = \cosh u \cos v, \qquad y = \cosh u \sin v, \qquad z = u.$$

This means that u and v can be considered as local coordinates on Σ, so that the inclusion, $i : \Sigma \to \mathbb{R}^3$, is given by

$$\begin{aligned}
i^*x &= \cosh u \cos v, \\
i^*y &= \cosh u \sin v, \\
i^*z &= u,
\end{aligned} \tag{6.125}$$

where (x, y, z) is the natural coordinate system of $\mathbb{R}^3$.

At each point $p \in \Sigma$, the tangent space $T_p\Sigma$ is generated by $(\partial/\partial u)_p$, $(\partial/\partial v)_p$, and, according to (1.24) and (6.125),

$$\begin{aligned}
i_{*p}\left(\frac{\partial}{\partial v}\right)_p &= \left(\frac{\partial}{\partial v}\right)_p \big[i^* x^j\big]\left(\frac{\partial}{\partial x^j}\right)_{i(p)} \\
&= -\cosh u(p)\sin v(p)\left(\frac{\partial}{\partial x}\right)_{i(p)} + \cosh u(p)\cos v(p)\left(\frac{\partial}{\partial y}\right)_{i(p)} \\
&= \left(-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}\right)_{i(p)}
\end{aligned}$$

and, similarly,

$$\begin{aligned}
i_{*p}\left(\frac{\partial}{\partial u}\right)_p &= \sinh u(p)\cos v(p)\left(\frac{\partial}{\partial x}\right)_{i(p)} \\
&\quad + \sinh u(p)\sin v(p)\left(\frac{\partial}{\partial y}\right)_{i(p)} + \left(\frac{\partial}{\partial z}\right)_{i(p)} \\
&= \left(\frac{x\sqrt{x^2+y^2-1}}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{y\sqrt{x^2+y^2-1}}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right)_{i(p)}.
\end{aligned}$$

The latter equality is valid only where $\sinh u \geq 0$, that is, only for $z \geq 0$.

One readily finds that the vector fields

$$-y\frac{\partial}{\partial x} + x\frac{\partial}{\partial y}, \qquad \frac{x\sqrt{x^2+y^2-1}}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{y\sqrt{x^2+y^2-1}}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y} + \frac{\partial}{\partial z}$$

are orthogonal to each other and their norms are equal to $\sqrt{x^2+y^2}$. Looking for a vector field orthogonal to these vector fields one obtains the orthonormal basis

$$\begin{aligned}
\mathbf{e}_1 &= -\frac{y}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x} + \frac{x}{\sqrt{x^2+y^2}}\frac{\partial}{\partial y}, \\
\mathbf{e}_2 &= \frac{x\sqrt{x^2+y^2-1}}{x^2+y^2}\frac{\partial}{\partial x} + \frac{y\sqrt{x^2+y^2-1}}{x^2+y^2}\frac{\partial}{\partial y} + \frac{1}{\sqrt{x^2+y^2}}\frac{\partial}{\partial z}, \\
\mathbf{e}_3 &= \frac{x}{x^2+y^2}\frac{\partial}{\partial x} + \frac{y}{x^2+y^2}\frac{\partial}{\partial y} - \frac{\sqrt{x^2+y^2-1}}{\sqrt{x^2+y^2}}\frac{\partial}{\partial z},
\end{aligned} \tag{6.126}$$

whose dual basis is

$$\begin{aligned}
\theta^1 &= \frac{-y\,\mathrm{d}x + x\,\mathrm{d}y}{\sqrt{x^2+y^2}}, \\
\theta^2 &= \frac{\sqrt{x^2+y^2-1}}{x^2+y^2}(x\,\mathrm{d}x + y\,\mathrm{d}y) + \frac{\mathrm{d}z}{\sqrt{x^2+y^2}},
\end{aligned} \tag{6.127}$$

$$\theta^3 = \frac{x\,\mathrm{d}x + y\,\mathrm{d}y}{x^2+y^2} - \sqrt{\frac{x^2+y^2-1}{x^2+y^2}}\,\mathrm{d}z.$$

The computation of the exterior derivative of these 1-forms is simplified by the fact that $\mathrm{d}(x^2+y^2) = 2(x\,\mathrm{d}x + y\,\mathrm{d}y)$. A straightforward computation yields

$$\begin{aligned} \mathrm{d}\theta^1 &= -\frac{1}{x^2+y^2}\,\theta^1\wedge\theta^3 - \frac{\sqrt{x^2+y^2-1}}{x^2+y^2}\,\theta^1\wedge\theta^2, \\ \mathrm{d}\theta^2 &= \frac{1}{x^2+y^2}\,\theta^2\wedge\theta^3, \\ \mathrm{d}\theta^3 &= \frac{1}{(x^2+y^2)\sqrt{x^2+y^2-1}}\,\theta^2\wedge\theta^3. \end{aligned} \tag{6.128}$$

(Note that these equations imply that the three 1-forms θ^i are integrable or, equivalently, that the pairs of vector fields $\{\mathbf{e}_1, \mathbf{e}_2\}$, $\{\mathbf{e}_2, \mathbf{e}_3\}$, and $\{\mathbf{e}_3, \mathbf{e}_1\}$ generate integrable distributions.)

Comparison with (6.64), using the fact that $\Gamma_{ijk} = -\Gamma_{jik}$, shows that the only nonzero Ricci rotation coefficients for the orthonormal basis (6.126) are given by

$$\begin{aligned} \Gamma_{121} &= \frac{\sqrt{x^2+y^2-1}}{x^2+y^2}, \\ \Gamma_{232} &= \Gamma_{311} = -\frac{1}{x^2+y^2}, \\ \Gamma_{233} &= -\frac{1}{(x^2+y^2)\sqrt{x^2+y^2-1}}. \end{aligned} \tag{6.129}$$

Hence, with respect to the orthonormal basis $\{\mathbf{X}_1, \mathbf{X}_2\}$, dual to $\{\phi^1, \phi^2\}$, the shape operator is represented by the matrix

$$\frac{1}{\cosh^2 u}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore, the mean curvature of the catenoid is indeed equal to zero, while its Gaussian curvature is $K = -1/(\cosh^4 u)$.

Exercise 6.48 Consider the helicoid, which is a surface in $\mathbb{R}^3$ that can be defined by

$$\begin{aligned} i^*x &= av\cos u, \\ i^*y &= av\sin u, \\ i^*z &= bu, \end{aligned}$$

where a, b are positive real constants. Construct an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ such that, at the points of the surface, $\mathbf{e}_1$ and $\mathbf{e}_2$ span the tangent space. Find the shape operator making use of the first Cartan structural equations and show that the mean curvature is zero.

Exercise 6.49 Let $p \in \Sigma$ and let $\mathbf{n}$ be a unit vector field defined on a neighborhood of p in M, such that at the points of Σ, $\mathbf{n}$ is normal to the surface. Show that

$$H(p) = \frac{1}{2}(\operatorname{div}\mathbf{n})(p).$$

(*Hint*: the conclusion can readily be obtained making use of (6.109) and (6.119).)

Chapter 7
Lie Groups

7.1 Basic Concepts

A Lie group is a group that possesses, in addition to the algebraic structure of a group, a differentiable manifold structure compatible with its algebraic structure in the sense that the group operations are differentiable functions.

Definition 7.1 Let G be a group which is a differentiable manifold. We say that G is a *Lie group* if the map from $G \times G$ into G given by $(g_1, g_2) \mapsto g_1 g_2$ and the map from G into G given by $g \mapsto g^{-1}$, where g^{-1} is the inverse of g, are differentiable. The dimension of the group is the dimension of the manifold.

Roughly speaking, if G is a Lie group, there exist locally coordinates labeling the elements of the group in such a way that the coordinates of the product $g_1 g_2$ are differentiable functions of the coordinates of g_1 and g_2. The coordinates of g^{-1} must be differentiable functions of those of g. In this context, the coordinates are also called *group parameters*.

Example 7.2 The space $\mathbb{R}^n$ where the group operation is the usual sum of n-tuples, with its usual differentiable manifold structure (see Sect. 1.1), is a Lie group of dimension n. In fact, if $(x^1, \ldots, x^n)$ is the natural coordinate system of $\mathbb{R}^n$, we have $x^i(gg') = x^i(g) + x^i(g')$ and $x^i(g^{-1}) = -x^i(g)$, which shows that the coordinates of gg' are differentiable functions of the coordinates of g and g', while the coordinates of g^{-1} are differentiable functions of the coordinates of g.

Example 7.3 Let $\mathrm{GL}(n, \mathbb{R})$ be the group of non-singular $n \times n$ real matrices, where the group operation is the usual matrix multiplication. Each $g \in \mathrm{GL}(n, \mathbb{R})$ is a matrix (a^i_j) and the n^2 functions $x^i_j : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R}$, defined by $x^i_j(g) = a^i_j$, can be used as coordinates in all of $\mathrm{GL}(n, \mathbb{R})$. The atlas containing this chart of coordinates defines a differentiable manifold structure for $\mathrm{GL}(n, \mathbb{R})$. Since $x^i_j(gg') = x^i_k(g) x^k_j(g')$ and

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_7,

$x^i_j(g^{-1})$ is a differentiable function of the $x^i_j(g)$ (specifically,

$$x^l_k\big(g^{-1}\big) = \frac{1}{(n-1)!\det g}\varepsilon_{ki_2\ldots i_n}\varepsilon^{lj_2\ldots j_n}x^{i_2}_{j_2}(g)\cdots x^{i_n}_{j_n}(g),$$

where

$$\begin{aligned}\varepsilon_{i_1i_2\ldots i_n} &= \varepsilon^{i_1i_2\ldots i_n}\\ &= \begin{cases}1 & \text{if } (i_1,\ldots,i_n) \text{ is an even permutation of } (1,2,\ldots,n),\\ -1 & \text{if } (i_1,\ldots,i_n) \text{ is an odd permutation of } (1,2,\ldots,n),\\ 0 & \text{if one of the values of the indices appears repeated,}\end{cases}\end{aligned}$$

i.e., $x^i_j(g^{-1})$ is a rational function of the $x^i_j(g)$ and the denominator in the preceding expression does not vanish because g is a non-singular matrix), $\mathrm{GL}(n,\mathbb{R})$ is a Lie group of dimension n^2. The group $\mathrm{GL}(n,\mathbb{R})$ is Abelian only when $n=1$ and $\mathrm{GL}(1,\mathbb{R})$ can be identified with $\mathbb{R}\setminus\{0\}$ with the usual multiplication.

Example 7.4 Any pair of real numbers a, b, with $a \neq 0$, defines an *affine motion* of $\mathbb{R}$, given by $x \mapsto ax+b$. One can readily verify that these transformations form a group under the composition. It is convenient to note that

$$\begin{pmatrix} a & b\\ 0 & 1\end{pmatrix}\begin{pmatrix} x\\ 1\end{pmatrix} = \begin{pmatrix} ax+b\\ 1\end{pmatrix},$$

which shows that the affine motions of $\mathbb{R}$ can be represented by the 2×2 real matrices of the form $\left(\begin{smallmatrix} a & b\\ 0 & 1\end{smallmatrix}\right)$, with $a \neq 0$, which form a group with the usual matrix multiplication. By associating the matrix $\left(\begin{smallmatrix} a & b\\ 0 & 1\end{smallmatrix}\right)$ to the transformation $x \mapsto ax+b$, the composition of two transformations of this class is associated with the product of the corresponding matrices.

The coordinate system (x^1, x^2) defined by

$$g = \begin{pmatrix} x^1(g) & x^2(g)\\ 0 & 1\end{pmatrix} \tag{7.1}$$

covers the entire group and, therefore, defines a differentiable manifold structure (the image of the entire group under this chart of coordinates is $\{(x,y)\in\mathbb{R}^2 \mid x \neq 0\}$, which is an open subset of $\mathbb{R}^2$). The product of two matrices, g and g', of the form (7.1) is another matrix of the same form with

$$x^1(gg') = x^1(g)x^1(g'), \qquad x^2(gg') = x^1(g)x^2(g') + x^2(g) \tag{7.2}$$

and

$$x^1\big(g^{-1}\big) = \frac{1}{x^1(g)}, \qquad x^2\big(g^{-1}\big) = -\frac{x^2(g)}{x^1(g)}. \tag{7.3}$$

The differentiability of these expressions implies that we are dealing with a Lie group (recall that $x^1(g)$ cannot be equal to zero). It may be noticed that this group is not connected (the set of matrices with $x^1 > 0$ is separated from the set of matrices with $x^1 < 0$, but is simply connected (a closed curve is shrinkable to a point).

Example 7.5 The group $\mathrm{SL}(2,\mathbb{R})$ is formed by the 2×2 real matrices with determinant equal to 1, with the usual operation of matrix multiplication. Any element of this group in a neighborhood of the identity is of the form $\left(\begin{smallmatrix} a & b \\ c & (1+bc)/a \end{smallmatrix}\right)$, with $a \neq 0$; therefore we can define the local coordinate system (x^1, x^2, x^3) by

$$g = \begin{pmatrix} x^1(g) & x^2(g) \\ x^3(g) & \frac{1+x^2(g)x^3(g)}{x^1(g)} \end{pmatrix}, \quad x^1(g) \neq 0. \tag{7.4}$$

Calculating the product of two elements of this group, g and g', expressed in the form (7.4), we find that

$$\begin{aligned} x^1(gg') &= x^1(g)x^1(g') + x^2(g)x^3(g'), \\ x^2(gg') &= x^1(g)x^2(g') + x^2(g)\frac{1+x^2(g')x^3(g')}{x^1(g')}, \\ x^3(gg') &= x^3(g)x^1(g') + \frac{1+x^2(g)x^3(g)}{x^1(g)}x^3(g') \end{aligned} \tag{7.5}$$

[assuming that $x^1(gg') \neq 0$, so that gg' is also of the form (7.4)]. Calculating the inverse of the matrix in (7.4) one has

$$\begin{aligned} x^1\left(g^{-1}\right) &= \frac{1+x^2(g)x^3(g)}{x^1(g)}, \\ x^2\left(g^{-1}\right) &= -x^2(g), \\ x^3\left(g^{-1}\right) &= -x^3(g). \end{aligned} \tag{7.6}$$

Taking into account that x^1 does not vanish in the domain of the coordinate system defined in (7.4), the expressions (7.5) and (7.6) are differentiable functions. The coordinate system x^i does not cover all of the set $\mathrm{SL}(2,\mathbb{R})$, but together with the coordinates (y^1, y^2, y^3) given by

$$g = \begin{pmatrix} y^1(g) & y^2(g) \\ \frac{y^1(g)y^3(g)-1}{y^2(g)} & y^3(g) \end{pmatrix}, \quad y^2(g) \neq 0, \tag{7.7}$$

it forms a subatlas that defines a differentiable manifold structure for $\mathrm{SL}(2,\mathbb{R})$. Considering the possible products of matrices of the form (7.4) by matrices of the form (7.7), the result must be of the form (7.4) or (7.7), which leads to expressions similar to (7.5), showing that the mapping $(g, g') \mapsto gg'$ is differentiable. Similarly, expressing the inverse of a matrix of the form (7.4) or (7.7) in the form (7.4) or (7.7), one obtains differentiable expressions analogous to (7.6), leading one to conclude that $\mathrm{SL}(2,\mathbb{R})$ is a Lie group of dimension three.

Example 7.6 The group SU(2) is formed by the complex unitary 2×2 matrices with determinant equal to 1, with the usual operation of matrix multiplication. It can readily be seen that any element of SU(2) is of the form $\left(\begin{smallmatrix} a+\mathrm{i}b & c+\mathrm{i}d \\ -c+\mathrm{i}d & a-\mathrm{i}b \end{smallmatrix}\right)$, with $a^2+b^2+c^2+d^2=1$, which means that the set SU(2) can be identified with S^3, the sphere of radius 1 in $\mathbb{R}^4$. In a neighborhood of the identity we can define the coordinate system (x^1, x^2, x^3) in such a manner that if $g = \left(\begin{smallmatrix} a+\mathrm{i}b & c+\mathrm{i}d \\ -c+\mathrm{i}d & a-\mathrm{i}b \end{smallmatrix}\right)$, then $x^1(g) \equiv -d$, $x^2(g) \equiv -c$, $x^3(g) \equiv -b$, with $a = \sqrt{1 - \sum_{i=1}^3 [x^i(g)]^2}$. That is,

$$g = \begin{pmatrix} h(g) - \mathrm{i}x^3(g) & -x^2(g) - \mathrm{i}x^1(g) \\ x^2(g) - \mathrm{i}x^1(g) & h(g) + \mathrm{i}x^3(g) \end{pmatrix}, \quad \text{with } h \equiv \sqrt{1 - \sum_{i=1}^{3} \left(x^i\right)^2} > 0. \tag{7.8}$$

It may be noticed that $x^1(e) = x^2(e) = x^3(e) = 0$, where e is the identity of the group, and that the coordinate system (x^1, x^2, x^3) covers almost one half of SU(2), corresponding to $a > 0$.

Calculating the product of two matrices of the form (7.8), it can readily be seen that

$$\begin{aligned} x^1(gg') &= x^1(g)h(g') + h(g)x^1(g') + x^2(g)x^3(g') - x^3(g)x^2(g'), \\ x^2(gg') &= x^2(g)h(g') + h(g)x^2(g') + x^3(g)x^1(g') - x^1(g)x^3(g'), \\ x^3(gg') &= x^3(g)h(g') + h(g)x^3(g') + x^1(g)x^2(g') - x^2(g)x^1(g'), \end{aligned} \tag{7.9}$$

and

$$x^i\left(g^{-1}\right) = -x^i(g). \tag{7.10}$$

These expressions are differentiable functions for $h(g), h(g') > 0$. As in the previous example, it is necessary to consider additional coordinate systems in order to cover the whole group and it can be verified that SU(2) is a Lie group of dimension three.

Example 7.7 Let SE(2) be the group of all the isometries of the Euclidean plane that preserve the orientation (translations and rigid rotations), with the group operation being the composition. Using Cartesian coordinates in the plane, each element g of this group can be characterized by three real numbers $x(g)$, $y(g)$, and $\theta(g)$, where $(x(g), y(g))$ are the coordinates of the image of the origin under the transformation g and $\theta(g)$ is the angle between the new x axis and the original one. (Here we are considering active transformations; the points of the plane move under the transformation, with the coordinate axes fixed.) It can readily be seen that if gg' is the transformation obtained by applying g after having applied g', then

$$\begin{aligned} x(gg') &= x(g) + x(g')\cos\theta(g) - y(g')\sin\theta(g), \\ y(gg') &= y(g) + x(g')\sin\theta(g) + y(g')\cos\theta(g), \\ \theta(gg') &= \theta(g) + \theta(g') \end{aligned} \tag{7.11}$$

(these formulas can be obtained taking into account the fact that under the transformation g, a point of the plane, with Cartesian coordinates (a, b), is mapped into the point with coordinates $(x(g) + a\cos\theta(g) - b\sin\theta(g), y(g) + a\sin\theta(g) + b\cos\theta(g))$ and then calculating the effect of the composition of two transformations g and g'). Similarly, for g^{-1}, the inverse transformation of g,

$$\begin{aligned} x\left(g^{-1}\right) &= -x(g)\cos\theta(g) - y(g)\sin\theta(g), \\ y\left(g^{-1}\right) &= x(g)\sin\theta(g) - y(g)\cos\theta(g), \\ \theta\left(g^{-1}\right) &= -\theta(g). \end{aligned} \tag{7.12}$$

In order for (x, y, θ) to be a coordinate system it is necessary to restrict the values of θ, for instance, imposing the condition $-\pi < \theta(g) < \pi$ (so that the image of this chart is an open subset of $\mathbb{R}^3$ and θ is single-valued); hence, this chart of coordinates will not cover all of the group, but, as in the two previous examples, introducing additional coordinate systems in a similar way, it can be verified that SE(2) is a Lie group of dimension three.

Equations (7.11) and (7.12) can also be obtained associating with each $g \in \mathrm{SE}(2)$ the matrix

$$\rho(g) \equiv \begin{pmatrix} \cos\theta(g) & -\sin\theta(g) & x(g) \\ \sin\theta(g) & \cos\theta(g) & y(g) \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.13}$$

Then it can be verified that $\rho(gg') = \rho(g)\rho(g')$. By virtue of this relation, the map $g \mapsto \rho(g)$ is a *matrix representation* of the group SE(2). In general, if G is any group, a matrix representation of G is a map, ρ, that assigns to each element $g \in G$ a non-singular square matrix, $\rho(g)$, in such a way that $\rho(gg') = \rho(g)\rho(g')$, for any pair of elements $g, g' \in G$.

Exercise 7.8 Let G be a group which is a differentiable manifold. Show that G is a Lie group if and only if the map from $G \times G$ into G given by $(g_1, g_2) \mapsto g_1 g_2^{-1}$ is differentiable.

Definition 7.9 Let G be a Lie group. A *Lie subgroup* of G is a subgroup of G which is a submanifold of G.

Thus, in Example 7.7, the set H formed by the elements with $\theta = 0$ is, clearly, a submanifold of G. Using (7.11) and (7.12) it can readily be verified that H is a subgroup of G; hence, H is a Lie subgroup of SE(2) (which corresponds to the rigid translations of the plane).

Exercise 7.10 Show that the set $\mathrm{SL}(n, \mathbb{R})$, formed by the real $n \times n$ matrices with determinant equal to 1, is a Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$.

7.2 The Lie Algebra of the Group

In this section we shall show that each Lie group possesses an associated Lie algebra whose properties reflect those of the group.

Let G be a Lie group. For $g \in G$, L_g denotes the map from G onto G defined by

$$L_g(g') = gg', \quad \text{for } g' \in G$$

(sometimes called the *left translation by* g). Similarly, $R_g : G \to G$ (the *right translation by* g) is defined by

$$R_g(g') = g'g, \quad \text{for } g' \in G.$$

From the definition of a Lie group it follows that L_g and R_g are differentiable maps and, furthermore, are diffeomorphisms since $(L_g)^{-1} = L_{g^{-1}}$ and $(R_g)^{-1} = R_{g^{-1}}$ for all $g \in G$.

Exercise 7.11 Show that $L_{g_1 g_2} = L_{g_1} \circ L_{g_2}$, $R_{g_1 g_2} = R_{g_2} \circ R_{g_1}$, and $R_{g_1} \circ L_{g_2} = L_{g_2} \circ R_{g_1}$ for $g_1, g_2 \in G$.

Definition 7.12 Let $\mathbf{X}$ be a vector field on G. We say that $\mathbf{X}$ is *left-invariant* if ${L_g}^*\mathbf{X} = \mathbf{X}$ for all $g \in G$; analogously, $\mathbf{X}$ is *right-invariant* if ${R_g}^*\mathbf{X} = \mathbf{X}$ for all $g \in G$.

In other words, $\mathbf{X}$ is left-invariant if and only if $\mathbf{X}$ is L_g-related with itself for all $g \in G$ (see Fig. 7.1); therefore, $\mathbf{X}$ is left-invariant if and only if [see (1.40)]

$$(\mathbf{X}f) \circ L_g = \mathbf{X}(f \circ L_g), \quad \text{for all } g \in G \text{ and } f \in C^\infty(G).$$

From this expression we see that if $\mathbf{X}$ and $\mathbf{Y}$ are two left-invariant vector fields, then the linear combination $a\mathbf{X} + b\mathbf{Y}$, for $a, b \in \mathbb{R}$, and the Lie bracket $[\mathbf{X}, \mathbf{Y}]$ are left-invariant (see Sect. 1.3). This means that the left-invariant vector fields form a Lie subalgebra of $\mathfrak{X}(G)$. Of course, something analogous holds for the right-invariant vector fields. (Clearly, if G is Abelian, the left-invariant vector fields coincide with the right-invariant ones.) The *Lie algebra* of G, denoted by $\mathfrak{g}$, is the Lie algebra of the left-invariant vector fields on G.

If $\mathbf{X} \in \mathfrak{g}$, then, for $g \in G$, we have $\mathbf{X} = {L_{g^{-1}}}^*\mathbf{X}$; hence, $\mathbf{X}_g = ({L_{g^{-1}}}^*\mathbf{X})_g$, that is [see (2.24)],

$$\mathbf{X}_g = L_{g*e}\mathbf{X}_e, \tag{7.14}$$

where e denotes the identity element of G. Therefore, a left-invariant vector field is uniquely defined by its value at the identity.

From the foregoing formula it follows that each tangent vector ξ to G at the identity ($\xi \in T_eG$) defines a left-invariant vector field $\mathbf{X}$, given by

$$\mathbf{X}_g = L_{g*e}\xi. \tag{7.15}$$

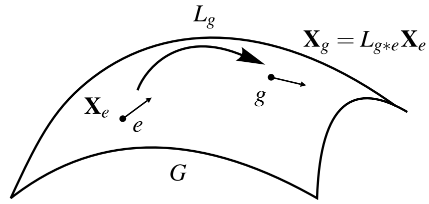

Fig. 7.1 A left-invariant vector field on G

The vector field thus defined belongs effectively to $\mathfrak{g}$, since if $f \in C^\infty(G)$ and $g, g' \in G$, using (1.30), (7.15), the chain rule (1.25), the fact that $L_{gg'} = L_g \circ L_{g'}$, and (1.23), we find

$$\begin{aligned}\big((\mathbf{X}f) \circ L_g\big)(g') &= (\mathbf{X}f)(gg') \\ &= \mathbf{X}_{gg'}[f] \\ &= (L_{gg'*e}\xi)[f] \\ &= \big(L_{g*g'}(L_{g'*e}\xi)\big)[f] \\ &= (L_{g*g'}\mathbf{X}_{g'})[f] \\ &= \mathbf{X}_{g'}[f \circ L_g] \\ &= \big(\mathbf{X}(f \circ L_g)\big)(g').\end{aligned}$$

Thus, there exists a one-to-one correspondence between the Lie algebra of G and T_eG. Using this correspondence the bracket of any pair of elements ξ and $\zeta \in T_eG$ is defined by means of

$$[\xi, \zeta] \equiv [\mathbf{X}, \mathbf{Y}]_e, \tag{7.16}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are the left-invariant vector fields such that $\xi = \mathbf{X}_e$ and $\zeta = \mathbf{Y}_e$. With this bracket, T_eG becomes a Lie algebra isomorphic to the Lie algebra of the group.

Exercise 7.13 Show that, effectively, T_eG is a Lie algebra.

The existence of this isomorphism between the left-invariant vector fields and the tangent vectors at the identity shows that the dimension of the Lie algebra of G coincides with the dimension of G.

Example 7.14 Let us consider $\mathbb{R}^n$ with the structure of Lie group defined in Example 7.2, and let $(x^1, \ldots, x^n)$ be the natural coordinates of $\mathbb{R}^n$. Then, $(x^i \circ L_g)(g') = x^i(gg') = x^i(g) + x^i(g')$, for $g, g' \in \mathbb{R}^n$, that is, $x^i \circ L_g = x^i(g) + x^i$. Any $\xi \in T_e\mathbb{R}^n$ is of the form $\xi = a^i(\partial/\partial x^i)_e$ with $a^i \in \mathbb{R}$. The left-invariant vector field corresponding to ξ is given in these coordinates by [see (7.15) and (1.24)]

$$\mathbf{X}_g = L_{g*e}\xi = a^i L_{g*e}\left(\frac{\partial}{\partial x^i}\right)_e = a^i \left(\frac{\partial}{\partial x^i}\right)_e \left[x^j \circ L_g\right]\left(\frac{\partial}{\partial x^j}\right)_g$$
$$= a^i \left(\frac{\partial}{\partial x^i}\right)_e \left[x^j(g) + x^j\right]\left(\frac{\partial}{\partial x^j}\right)_g = a^i \delta_i^j \left(\frac{\partial}{\partial x^j}\right)_g = a^i \left(\frac{\partial}{\partial x^i}\right)_g,$$

that is, $\mathbf{X} = a^i(\partial/\partial x^i)$. For $\zeta = b^i(\partial/\partial x^i)_e$, the corresponding left-invariant vector field is then $\mathbf{Y} = b^i(\partial/\partial x^i)$; therefore, $[\mathbf{X}, \mathbf{Y}] = 0$ and $[\xi, \zeta] = [\mathbf{X}, \mathbf{Y}]_e = 0$, that is, the Lie algebra of this group is Abelian. (In fact, as shown in Sects. 7.3 and 7.5, the Lie algebra of a group is Abelian if and only if the group is Abelian.)

Example 7.15 Let $\mathrm{GL}(n, \mathbb{R})$ be the Lie group with the coordinates x^i_j defined in Example 7.3. We then have $(x^i_j \circ L_g)(g') = x^i_j(gg') = x^i_k(g)x^k_j(g')$, for $g, g' \in \mathrm{GL}(n, \mathbb{R})$, that is, $x^i_j \circ L_g = x^i_k(g)x^k_j$. Any tangent vector to $\mathrm{GL}(n, \mathbb{R})$ at the identity is of the form $\xi = a^i_j(\partial/\partial x^i_j)_e$, with $a^i_j \in \mathbb{R}$, and the corresponding left-invariant vector field is

$$\begin{aligned}
\mathbf{X}_g = L_{g*e}\xi &= a^i_j L_{g*e}\left(\frac{\partial}{\partial x^i_j}\right)_e \\
&= a^i_j \left(\frac{\partial}{\partial x^i_j}\right)_e \left[x^k_m \circ L_g\right]\left(\frac{\partial}{\partial x^k_m}\right)_g \\
&= a^i_j \left(\frac{\partial}{\partial x^i_j}\right)_e \left[x^k_p(g)x^p_m\right]\left(\frac{\partial}{\partial x^k_m}\right)_g \\
&= a^i_j x^k_i(g)\left(\frac{\partial}{\partial x^k_j}\right)_g,
\end{aligned}$$

that is,

$$\mathbf{X} = a^i_j x^k_i \frac{\partial}{\partial x^k_j}. \tag{7.17}$$

Thus, the left-invariant vector fields on $\mathrm{GL}(n, \mathbb{R})$ are in a one-to-one correspondence with the real $n \times n$ matrices. If $A \equiv (a^i_j)$, we will denote by $\mathbf{X}_A$ the vector field (7.17).

If $\zeta = b^i_j(\partial/\partial x^i_j)_e$ is another element of $T_e\mathrm{GL}(n, \mathbb{R})$, then the corresponding left-invariant vector field is $\mathbf{X}_B = b^i_j x^k_i(\partial/\partial x^k_j)$, where $B \equiv (b^i_j)$. A direct computation yields $[\mathbf{X}_A, \mathbf{X}_B] = (a^i_m b^m_j - b^i_m a^m_j)x^k_i(\partial/\partial x^k_j)$ and, since $x^k_i(e) = \delta^k_i$, we have

$$[\xi, \zeta] = [\mathbf{X}_A, \mathbf{X}_B]_e = \left(a^i_m b^m_j - b^i_m a^m_j\right)\left(\frac{\partial}{\partial x^i_j}\right)_e.$$

Noting that $a^i_m b^m_j - b^i_m a^m_j$ are the entries of the matrix $[A, B] \equiv AB - BA$, we conclude that $[\mathbf{X}_A, \mathbf{X}_B] = \mathbf{X}_{[A,B]}$.

In other words, associating to each element of $T_e\mathrm{GL}(n,\mathbb{R})$ the matrix formed by its components with respect to the basis $\{(\partial/\partial x^i_j)_e\}$, the matrix associated with the bracket of a pair of elements of $T_e\mathrm{GL}(n,\mathbb{R})$, is the commutator of the corresponding matrices. Furthermore, $\mathbf{X}_{aA+bB} = a\mathbf{X}_A + b\mathbf{X}_B$, for $a, b \in \mathbb{R}$. For these reasons, the Lie algebra of the group $\mathrm{GL}(n,\mathbb{R})$, denoted by $\mathfrak{gl}(n,\mathbb{R})$, is identified with the space of $n \times n$ matrices, where the bracket is given by the commutator.

Example 7.16 For $\mathrm{SL}(2,\mathbb{R})$ with the coordinates defined by (7.4), equations (7.5) amount to

$$\begin{aligned}
\left(x^1 \circ L_g\right) &= x^1(g)x^1 + x^2(g)x^3, \\
\left(x^2 \circ L_g\right) &= x^1(g)x^2 + x^2(g)\frac{1+x^2x^3}{x^1}, \\
\left(x^3 \circ L_g\right) &= x^3(g)x^1 + \frac{1+x^2(g)x^3(g)}{x^1(g)}x^3;
\end{aligned}$$

hence, if the left-invariant vector fields $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$ are such that $(\mathbf{X}_i)_e = (\partial/\partial x^i)_e$, then, for instance, taking into account that $x^1(e) = 1$, $x^2(e) = 0 = x^3(e)$,

$$\begin{aligned}
(\mathbf{X}_1)_g &= L_{g*e}\left(\frac{\partial}{\partial x^1}\right)_e = \left(\frac{\partial}{\partial x^1}\right)_e \left[x^j \circ L_g\right]\left(\frac{\partial}{\partial x^j}\right)_g \\
&= x^1(g)\left(\frac{\partial}{\partial x^1}\right)_g + x^2(g)\left(-\frac{1+x^2x^3}{(x^1)^2}\right)(e)\left(\frac{\partial}{\partial x^2}\right)_g + x^3(g)\left(\frac{\partial}{\partial x^3}\right)_g \\
&= x^1(g)\left(\frac{\partial}{\partial x^1}\right)_g - x^2(g)\left(\frac{\partial}{\partial x^2}\right)_g + x^3(g)\left(\frac{\partial}{\partial x^3}\right)_g,
\end{aligned}$$

i.e.,

$$\mathbf{X}_1 = x^1\frac{\partial}{\partial x^1} - x^2\frac{\partial}{\partial x^2} + x^3\frac{\partial}{\partial x^3}, \tag{7.18}$$

and in a similar way one finds that

$$\begin{aligned}
\mathbf{X}_2 &= x^1\frac{\partial}{\partial x^2}, \\
\mathbf{X}_3 &= x^2\frac{\partial}{\partial x^1} + \frac{1+x^2x^3}{x^1}\frac{\partial}{\partial x^3}.
\end{aligned} \tag{7.19}$$

It should be noticed that these expressions are local [valid only in the domain of the coordinates (x^1, x^2, x^3)], but that, in all cases, the left-invariant vector fields are globally defined (even if, as in the case of $\mathrm{GL}(n,\mathbb{R})$, G is not connected). Among other things, this means that any Lie group is a parallelizable manifold.

Since $\mathrm{SL}(2,\mathbb{R})$ is a Lie subgroup of $\mathrm{GL}(2,\mathbb{R})$, the left-invariant vector fields (7.18) and (7.19) must be expressible in the form (7.17), in terms of the coordinates x^i_j. From (7.4) we find that the inclusion $i : \mathrm{SL}(2,\mathbb{R}) \to \mathrm{GL}(2,\mathbb{R})$ is given locally by

$$i^* x^1_1 = x^1, \qquad i^* x^1_2 = x^2, \qquad i^* x^2_1 = x^3, \qquad i^* x^2_2 = \frac{1 + x^2 x^3}{x^1},$$

and hence, using (1.24),

$$\begin{aligned}
i_{*e}(\mathbf{X}_1)_e &= i_{*e}\left(\frac{\partial}{\partial x^1}\right)_e = \left(\frac{\partial}{\partial x^1_1}\right)_e - \left(\frac{\partial}{\partial x^2_2}\right)_e, \\
i_{*e}(\mathbf{X}_2)_e &= i_{*e}\left(\frac{\partial}{\partial x^2}\right)_e = \left(\frac{\partial}{\partial x^1_2}\right)_e, \\
i_{*e}(\mathbf{X}_3)_e &= i_{*e}\left(\frac{\partial}{\partial x^3}\right)_e = \left(\frac{\partial}{\partial x^2_1}\right)_e.
\end{aligned}$$

Thus, the matrices associated with the vector fields (7.18) and (7.19), in the sense defined in the preceding example, are

$$\mathbf{X}_1 \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathbf{X}_2 \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{X}_3 \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{7.20}$$

The matrices (7.20) have trace equal to 0 as a consequence of the fact that $\mathrm{SL}(2,\mathbb{R})$ is formed by matrices with determinant equal to 1. The group $\mathrm{SL}(2,\mathbb{R})$ corresponds to the submanifold of $\mathrm{GL}(2,\mathbb{R})$ defined by the equation $x^1_1 x^2_2 - x^1_2 x^2_1 = 1$; therefore, if $\mathbf{X}_e = a^i_j (\partial/\partial x^i_j)_e$ is tangent to this submanifold, $0 = \mathbf{X}_e[x^1_1 x^2_2 - x^1_2 x^2_1]$ (see Sect. 4.2). Hence, taking into account that $x^i_j(e) = \delta^i_j$, we obtain $a^1_1 + a^2_2 = 0$, that is, $\mathrm{tr}\,(a^i_j) = 0$.

Exercise 7.17 Show that if $\xi = a(\partial/\partial x)_e + b(\partial/\partial y)_e + c(\partial/\partial\theta)_e$ is a tangent vector to the group of isometries of the plane at the identity, expressed in the coordinates defined in Example 7.7, then $\mathbf{X} = (a\cos\theta - b\sin\theta)(\partial/\partial x) + (a\sin\theta + b\cos\theta)(\partial/\partial y) + c(\partial/\partial\theta)$ is the element of the Lie algebra of the group such that $\xi = \mathbf{X}_e$.

Exercise 7.18 Show that the Lie algebra of $\mathrm{SO}(n) \equiv \{A \in \mathrm{GL}(n,\mathbb{R}) \mid \det A = 1, AA^{\mathrm{t}} = I\}$ can be identified with the set of skew-symmetric $n \times n$ matrices.

Exercise 7.19 Show that if $\mathbf{X}_1$, $\mathbf{X}_2$, and $\mathbf{X}_3$ are the left-invariant vector fields on $\mathrm{SU}(2)$ such that $(\mathbf{X}_i)_e = \frac{1}{2}(\partial/\partial x^i)_e$, where the x^i are the coordinates defined in (7.8), then

$$\mathbf{X}_i = \frac{1}{2}\left(h\frac{\partial}{\partial x^i} - \sum_{j,k=1}^{3} \varepsilon_{ijk} x^j \frac{\partial}{\partial x^k}\right). \tag{7.21}$$

(The factor $1/2$ included in the definition of the $\mathbf{X}_i$ is introduced in order for them to coincide with the elements of the basis of $\mathfrak{su}(2)$ [the Lie algebra of SU(2)] commonly employed.)

Exercise 7.20 Making use of the formulas (7.2), find the left-invariant vector fields on the group of affine motions of $\mathbb{R}$, considered in Example 7.4, whose values at the identity are $(\partial/\partial x^1)_e$ and $(\partial/\partial x^2)_e$. Show that $[(\partial/\partial x^1)_e, (\partial/\partial x^2)_e] = (\partial/\partial x^2)_e$.

The Structure Constants Let $\{\xi_1, \xi_2, \ldots, \xi_n\}$ be a basis of T_eG, since the bracket $[\xi_i, \xi_j]$ belongs to T_eG, and then there exists a set of real numbers c_{ij}^k $(i, j, k = 1, 2, \ldots, n)$ such that $[\xi_i, \xi_j] = c_{ij}^k \xi_k$. Denoting by $\mathbf{X}_i$ the element of $\mathfrak{g}$ corresponding to ξ_i, we have $[\mathbf{X}_i, \mathbf{X}_j] = c_{ij}^k \mathbf{X}_k$. The scalars c_{ij}^k are called the *structure constants* of G with respect to the basis $\{\mathbf{X}_i\}$. The skew-symmetry of the bracket and the Jacobi identity imply that the structure constants must satisfy the relations

$$c_{ij}^k = -c_{ji}^k \tag{7.22}$$

and

$$c_{ij}^m c_{mk}^l + c_{jk}^m c_{mi}^l + c_{ki}^m c_{mj}^l = 0, \tag{7.23}$$

respectively.

Exercise 7.21 Calculate the structure constants of $\mathbb{R}^n$, the group of isometries of the plane, $SL(2, \mathbb{R})$, and SU(2).

The fact that the values of the structure constants depend on the basis of $\mathfrak{g}$ chosen means, among other things, that it is possible to obtain some simplification in the expressions for the structure constants by conveniently choosing the basis of $\mathfrak{g}$. A simple example is given by considering the Lie algebras of dimension two. If $\{\mathbf{X}_1, \mathbf{X}_2\}$ is a basis of $\mathfrak{g}$ (or of any real Lie algebra of dimension two), we necessarily have $[\mathbf{X}_1, \mathbf{X}_1] = 0 = [\mathbf{X}_2, \mathbf{X}_2]$ and $[\mathbf{X}_1, \mathbf{X}_2] = -[\mathbf{X}_2, \mathbf{X}_1]$, so that the only relevant bracket is $[\mathbf{X}_1, \mathbf{X}_2]$, which must be of the form $a\mathbf{X}_1 + b\mathbf{X}_2$ with $a, b \in \mathbb{R}$. (Note that when the dimension of the algebra is 2, the Jacobi identity is identically satisfied as a consequence of the skew-symmetry of the bracket, and therefore there are no restrictions on the values of a and b.)

It is necessary to analyze separately the following two cases:

(i) both coefficients are zero, $a = b = 0$,
(ii) at least one coefficient is different from zero.

In the first case the algebra is Abelian and $c_{ij}^k = 0$ with respect to any basis. In the second case, assuming, for instance, $b \neq 0$, owing to the bilinearity and the skew-symmetry of the bracket it follows that the set $\{\mathbf{X}'_1, \mathbf{X}'_2\}$, with $\mathbf{X}'_1 \equiv b^{-1}\mathbf{X}_1$, $\mathbf{X}'_2 \equiv a\mathbf{X}_1 + b\mathbf{X}_2$, is a basis of $\mathfrak{g}$ such that $[\mathbf{X}'_1, \mathbf{X}'_2] = \mathbf{X}'_2$ (cf. Exercise 7.20).

Thus, for any Lie algebra of dimension two we have $c_{ij}^k = 0$ (the algebra is Abelian) or it is possible to choose a basis for which the only structure constants different from zero are

$$c_{12}^2 = 1, \qquad c_{21}^2 = -1 \tag{7.24}$$

[cf. Erdmann and Wildon (2006, Chap. 3)].

Lie Group Homomorphisms Let G and H be two Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$, respectively, if $\phi : G \to H$ is a *Lie group homomorphism*, that is, ϕ is differentiable and $\phi(gg') = \phi(g)\phi(g')$, then to each $\mathbf{X} \in \mathfrak{g}$ there corresponds a unique left-invariant vector field on H, which will be denoted by $\phi_*\mathbf{X}$, such that $\mathbf{X}$ and $\phi_*\mathbf{X}$ are ϕ-related, i.e.,

$$\phi_{*g}\mathbf{X}_g = (\phi_*\mathbf{X})_{\phi(g)}, \quad \text{for } g \in G. \tag{7.25}$$

Indeed, the condition $\phi(gg') = \phi(g)\phi(g')$ can be written in the form $\phi(L_g g') = L_{\phi(g)}(\phi(g'))$, that is,

$$\phi \circ L_g = L_{\phi(g)} \circ \phi, \quad \text{for } g \in G; \tag{7.26}$$

therefore, if $\phi_*\mathbf{X}$ is the left-invariant vector field on H such that

$$(\phi_*\mathbf{X})_e \equiv \phi_{*e}\mathbf{X}_e, \tag{7.27}$$

then using (7.14), (7.27), the chain rule, and (7.26) we have

$$\begin{aligned}(\phi_*\mathbf{X})_{\phi(g)} &= L_{\phi(g)*e}(\phi_*\mathbf{X})_e = L_{\phi(g)*e}\phi_{*e}\mathbf{X}_e = (L_{\phi(g)} \circ \phi)_{*e}\mathbf{X}_e \\ &= (\phi \circ L_g)_{*e}\mathbf{X}_e = \phi_{*g}L_{g*e}\mathbf{X}_e = \phi_{*g}\mathbf{X}_g.\end{aligned}$$

Since $\mathbf{X}$ and $\phi_*\mathbf{X}$ are ϕ-related, the map $\mathbf{X} \mapsto \phi_*\mathbf{X}$ from $\mathfrak{g}$ into $\mathfrak{h}$ is a Lie algebra homomorphism.

Example 7.22 The mapping

$$\begin{pmatrix} a & b \\ 0 & 1/a \end{pmatrix} \mapsto \begin{pmatrix} a^2 & ab \\ 0 & 1 \end{pmatrix},$$

from the group G formed by the upper triangular 2×2 real matrices with determinant equal to 1 into the group H of the affine motions of $\mathbb{R}$ (see Example 7.4) is a (two-to-one) Lie group homomorphism. In fact, making use of the coordinate systems (y^1, y^2) and (x^1, x^2), on G and H, respectively, defined by

$$g = \begin{pmatrix} y^1(g) & y^2(g) \\ 0 & 1/y^1(g) \end{pmatrix}, \quad g \in G$$

and

$$g = \begin{pmatrix} x^1(g) & x^2(g) \\ 0 & 1 \end{pmatrix}, \quad g \in H$$

(as in Example 7.4), the mapping $\phi : G \to H$ defined above is given by

$$\phi^* x^1 = \left(y^1\right)^2, \qquad \phi^* x^2 = y^1 y^2 \tag{7.28}$$

and is differentiable.

Proceeding as in the examples above, one finds that a basis of $\mathfrak{g}$ is formed by the vector fields

$$\mathbf{X}_1 = y^1 \frac{\partial}{\partial y^1} - y^2 \frac{\partial}{\partial y^2}, \qquad \mathbf{X}_2 = y^1 \frac{\partial}{\partial y^2} \tag{7.29}$$

on G. With the aid of (1.24), (7.28), and (7.29) one finds that, for $g \in G$,

$$\begin{aligned}
\phi_{*g}(\mathbf{X}_1)_g &= (\mathbf{X}_1)_g\left[\phi^* x^i\right]\left(\frac{\partial}{\partial x^i}\right)_{\phi(g)} \\
&= 2\left(y^1(g)\right)^2\left(\frac{\partial}{\partial x^1}\right)_{\phi(g)} = \left(2x^1\frac{\partial}{\partial x^1}\right)_{\phi(g)}, \\
\phi_{*g}(\mathbf{X}_2)_g &= (\mathbf{X}_2)_g\left[\phi^* x^i\right]\left(\frac{\partial}{\partial x^i}\right)_{\phi(g)} \\
&= \left(y^1(g)\right)^2\left(\frac{\partial}{\partial x^2}\right)_{\phi(g)} = \left(x^1\frac{\partial}{\partial x^2}\right)_{\phi(g)}.
\end{aligned}$$

On the other hand, a similar, direct computation shows that the vector fields $x^1\,\partial/\partial x^1$ and $x^1\,\partial/\partial x^2$ form a basis of $\mathfrak{h}$ (see Exercise 7.20) and, therefore, in this case the mapping $\mathbf{X} \mapsto \phi_*\mathbf{X}$ is an isomorphism of Lie algebras.

It may be remarked that the mapping ϕ considered in this example is not injective nor surjective; however, the mapping $\phi_* : \mathfrak{g} \to \mathfrak{h}$ is one-to-one. Also note that even in those cases where ϕ is not surjective, the vector field $\phi_*\mathbf{X}$, as any left-invariant vector field on a Lie group, is defined at all the points of H.

Example 7.23 Since the determinant of a product of $n \times n$ matrices is equal to the product of their determinants, the mapping $\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R} \setminus \{0\}$ is a Lie group homomorphism, considering $\mathbb{R} \setminus \{0\}$ as a group with the multiplication; the differentiability of the mapping is evident from its explicit expression in terms of the coordinates x_i^j of $\mathrm{GL}(n, \mathbb{R})$, $\det = \frac{1}{n!}\varepsilon_{i_1\ldots i_n}\varepsilon^{j_1\cdots j_n} x_{j_1}^{i_1}\cdots x_{j_n}^{i_n}$. Any $n \times n$ matrix $A = (a_j^i)$ defines an element $\mathbf{X}_A$ of $\mathfrak{gl}(n, \mathbb{R})$ in such a way that $(\mathbf{X}_A)_e = a_j^i(\partial/\partial x_j^i)_e$ (see Example 7.15), and $\det_{*e}(\mathbf{X}_A)_e = a_j^i \det_{*e}(\partial/\partial x_j^i)_e = a_j^i(\partial/\partial x_j^i)_e(x \circ \det)(\partial/\partial x)_1$, where x is the natural coordinate of $\mathbb{R} \setminus \{0\}$ (i.e., $x = \mathrm{id}$). By means of a simple calculation, taking into account that $x_i^j(e) = \delta_i^j$, one obtains

$$\begin{aligned}
\left(\frac{\partial}{\partial x_j^i}\right)_e (x \circ \det) &= \left(\frac{\partial}{\partial x_j^i}\right)_e \det = \left(\frac{\partial}{\partial x_j^i}\right)_e \frac{1}{n!}\varepsilon_{i_1\ldots i_n}\varepsilon^{j_1\ldots j_n} x_{j_1}^{i_1}\cdots x_{j_n}^{i_n} \\
&= \frac{1}{(n-1)!}\varepsilon_{i i_2\ldots i_n}\varepsilon^{j j_2\ldots j_n}\delta_{j_2}^{i_2}\cdots\delta_{j_n}^{i_n} = \delta_i^j;
\end{aligned}$$

thus,

$$(\det_* \mathbf{X}_A)_e = \det_{*e} a^i_j \left(\frac{\partial}{\partial x^i_j}\right)_e = a^i_i \left(\frac{\partial}{\partial x}\right)_1 = \operatorname{tr} A \left(\frac{\partial}{\partial x}\right)_1.$$

Hence, just like $\mathbf{X}_A$ is identified with the matrix A, its image under $\det_*$ is identified with the trace of A. According to the previous results, it follows that the mapping $A \mapsto \operatorname{tr} A$ is a Lie algebra homomorphism, which simply amounts to saying that the trace of the commutator of any two matrices is equal to zero (the Lie algebra of $\mathbb{R} \setminus \{0\}$, as any Lie algebra of dimension one, is Abelian), $\operatorname{tr}[A, B] = [\operatorname{tr} A, \operatorname{tr} B] = 0$ and that the trace is a linear mapping.

Two well-known examples of Lie group homomorphisms are the following. Let G be the additive group of the real numbers and let H be the group of the complex numbers of modulus equal to 1 with the usual multiplication (identifiable with the unit circle S^1). Then, the map $x \mapsto \mathrm{e}^{\mathrm{i}x}$ is an infinite-to-one homomorphism (the kernel of this homomorphism is formed by all the integral multiples of 2π). The second example corresponds to the two-to-one homomorphism between SU(2) and SO(3). In order to give explicitly this homomorphism it is convenient to make use of the Pauli matrices

$$\sigma^1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma^2 \equiv \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma^3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which form a basis for the *real* vector space formed by the traceless Hermitian 2×2 complex matrices. Furthermore, the Pauli matrices satisfy

$$\sigma^i \sigma^j = \delta^{ij} I + \mathrm{i} \sum_{k=1}^{3} \varepsilon^{ijk} \sigma^k, \tag{7.30}$$

where I denotes the 2×2 unit matrix.

If $g \in \mathrm{SU}(2)$, then, for $i = 1, 2, 3$, $g^{-1}\sigma^i g$ is also a traceless Hermitian 2×2 complex matrix and therefore there exist real numbers, a^i_j (which depend on g) such that

$$g^{-1}\sigma^i g = a^i_j \sigma^j, \quad i = 1, 2, 3. \tag{7.31}$$

As we shall show now, (a^i_j) belongs to SO(3), i.e., (a^i_j) is an orthogonal 3×3 real matrix with determinant equal to 1. To this end, we calculate

$$\begin{aligned} g^{-1}\sigma^i \sigma^j g &= g^{-1}\left(\delta^{ij} I + \mathrm{i} \sum_{k=1}^{3} \varepsilon^{ijk} \sigma^k\right) g \\ &= \delta^{ij} I + \mathrm{i} \sum_{k=1}^{3} \varepsilon^{ijk} g^{-1} \sigma^k g \end{aligned}$$

$$= \delta^{ij} I + \mathrm{i} \sum_{k=1}^{3} \varepsilon^{ijk} a^k_m \sigma^m.$$

On the other hand,

$$\begin{aligned} g^{-1}\sigma^i\sigma^j g &= g^{-1}\sigma^i g\, g^{-1}\sigma^j g \\ &= a^i_k \sigma^k a^j_m \sigma^m \\ &= a^i_k a^j_m \left(\delta^{km} I + \mathrm{i} \sum_{r=1}^{3} \varepsilon^{kmr} \sigma^r \right). \end{aligned}$$

Using the fact that $\{I, \sigma^1, \sigma^2, \sigma^3\}$ is linearly independent, it follows that

$$a^i_k a^j_m \delta^{km} = \delta^{ij}, \qquad \sum_{k=1}^{3} \varepsilon^{ijk} a^k_r = \varepsilon^{kmr} a^i_k a^j_m.$$

The first of these equations means that (a^i_j) is an orthogonal matrix, and from the second one we have

$$\sum_{k=1}^{3} \varepsilon^{ijk} a^k_r a^n_s \delta^{rs} = \varepsilon^{kmr} a^i_k a^j_m a^n_s \delta^{rs}$$

and since $a^k_r a^n_s \delta^{rs} = \delta^{kn}$, we obtain $\varepsilon^{ijn} = \varepsilon^{kmr} a^i_k a^j_m a^n_r$, which means that $\det(a^i_j) = 1$ (see Example 7.23). Thus, we have a map $\phi : \mathrm{SU}(2) \to \mathrm{SO}(3)$ given by $\phi(g) = (a^i_j)$, with (a^i_j) defined by (7.31).

Combining (7.31) and (7.30), making use of the fact that $\operatorname{tr} \sigma^i = 0$ and $\operatorname{tr} I = 2$, and the linearity of the trace, we have

$$\operatorname{tr}\left(g^{-1}\sigma^i g \sigma^k\right) = \operatorname{tr}\left(a^i_j \sigma^j \sigma^k\right) = a^i_j \operatorname{tr}\left(\delta^{jk} I + \mathrm{i} \sum_{m=1}^{3} \varepsilon^{jkm} \sigma^m \right) = 2 a^i_k,$$

that is, $a^i_j = \frac{1}{2} \operatorname{tr}(g^{-1}\sigma^i g \sigma^j)$. In terms of the natural coordinates x^i_j of $\mathrm{GL}(3, \mathbb{R})$, $a^i_j = x^i_j(\phi(g))$, hence

$$x^i_j\left(\phi(g)\right) = \frac{1}{2} \operatorname{tr}\left(g^{-1}\sigma^i g \sigma^j\right). \tag{7.32}$$

With the aid of the explicit expression (7.32) we can verify that ϕ is a group homomorphism. Indeed, for $g, g' \in \mathrm{SU}(2)$, making use of (7.32) we obtain

$$\begin{aligned} x^i_j\left(\phi(gg')\right) &= \frac{1}{2} \operatorname{tr}\left(g'^{-1} g^{-1} \sigma^i g g' \sigma^j\right) = \frac{1}{2} \operatorname{tr}\left(g'^{-1} x^i_k\left(\phi(g)\right) \sigma^k g' \sigma^j\right) \\ &= x^i_k\left(\phi(g)\right) \frac{1}{2} \operatorname{tr}\left(g'^{-1} \sigma^k g' \sigma^j\right) = x^i_k\left(\phi(g)\right) x^k_j\left(\phi(g')\right). \end{aligned}$$

The fact that ϕ is two-to-one is equivalent to saying that there exist only two elements of SU(2) that are mapped by ϕ to the identity of SO(3) (i.e., $\ker\phi$ consists of exactly two elements). If $g \in \mathrm{SU}(2)$ is such that $\phi(g)$ is the identity of SO(3), then from (7.31) we have $g^{-1}\sigma^i g = \sigma^i$, which amounts to $\sigma^i g = g\sigma^i$, for $i = 1, 2, 3$. These equations imply that g is a multiple of I, and from the condition $\det g = 1$ one concludes that $g = \pm I$.

Lie Subgroups If H is a Lie subgroup of G, the left-invariant vector fields of H, being defined at e, can be extended to all of G as left-invariant vector fields on G, using (7.14). In this manner, the Lie algebra of H can be regarded as a Lie subalgebra of the Lie algebra of G. Conversely, if $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, $\mathfrak{h}$ defines a distribution in G which is involutive and, according to the Frobenius Theorem, completely integrable. Let H be the maximal integral manifold of this distribution containing e. Since $\mathfrak{h}$ is formed by left-invariant vector fields, for $h \in H$, $L_{h^{-1}}(H)$ is also an integral manifold of the distribution that contains the identity; therefore, $L_{h^{-1}}(H) \subset H$, which implies that H is a Lie subgroup of G.

In particular, any $\mathbf{X} \in \mathfrak{g}$ different from zero generates a Lie subalgebra of dimension one, $\mathfrak{h} = \{\mathbf{Y} \in \mathfrak{g} \,|\, \mathbf{Y} = a\mathbf{X}, a \in \mathbb{R}\}$, and the integral manifold of the distribution generated by $\mathbf{X}$ containing the identity (which in this case is the image of a curve) is a one-parameter subgroup of G (see Sect. 7.4).

Example 7.24 The matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

form a basis of a Lie subalgebra of $\mathfrak{gl}(2, \mathbb{R})$. Substituting these matrices into (7.17) one obtains the two left-invariant vector fields on $\mathrm{GL}(2, \mathbb{R})$,

$$\mathbf{X}_1 \equiv x^1_1 \frac{\partial}{\partial x^1_1} + x^2_1 \frac{\partial}{\partial x^2_1} - x^1_2 \frac{\partial}{\partial x^1_2} - x^2_2 \frac{\partial}{\partial x^2_2}, \qquad \mathbf{X}_2 \equiv x^1_1 \frac{\partial}{\partial x^1_2} + x^2_1 \frac{\partial}{\partial x^2_2}.$$

It is convenient to simplify the notation, using (x, y, z, w) in place of $(x^1_1, x^1_2, x^2_1, x^2_2)$, so that the vector fields above are

$$\mathbf{X}_1 = x\frac{\partial}{\partial x} + z\frac{\partial}{\partial z} - y\frac{\partial}{\partial y} - w\frac{\partial}{\partial w}, \qquad \mathbf{X}_2 = x\frac{\partial}{\partial y} + z\frac{\partial}{\partial w}.$$

In order to find the integral manifolds of the distribution defined by $\mathbf{X}_1$ and $\mathbf{X}_2$ in the underlying manifold of $\mathrm{GL}(2, \mathbb{R})$, we look for two functionally independent solutions of the system of linear PDEs $\mathbf{X}_1 f = 0$, $\mathbf{X}_2 f = 0$.

Following the procedure employed in Example 4.1 [see also (4.9)] one finds that the functions xy, xw, and yz satisfy $\mathbf{X}_1 f = 0$ (that is, they are constant along the integral curves of $\mathbf{X}_1$). Similarly, by inspection, x and z are constant along the integral curves of $\mathbf{X}_2$, and therefore $xw - zy$ also satisfies $\mathbf{X}_2 f = 0$. Hence, $xw - zy$ and z/x are two functionally independent solutions of the system of equations $\mathbf{X}_1 f = 0$,

$\mathbf{X}_2 f = 0$, which means that the integral manifolds of the distribution under consideration are given by $xw - zy = \text{const}$, $z/x = \text{const}$ Since $x(e) = 1 = w(e)$, $y(e) = 0 = z(e)$, the integral manifold passing through the identity of $\mathrm{GL}(2, \mathbb{R})$ is given by $xw - yz = 1$, $z = 0$, which corresponds to matrices of the form $\left(\begin{smallmatrix} x & y \\ 0 & 1/x \end{smallmatrix}\right)$. These matrices form, effectively, a subgroup of $\mathrm{GL}(2, \mathbb{R})$ (cf. Example 7.22).

Alternatively, one can find first two independent 1-forms that, contracted with $\mathbf{X}_1$ and $\mathbf{X}_2$, yield zero. A possible choice is given by the 1-forms

$$\alpha^1 = (xw - yz)\,\mathrm{d}x - xz\,\mathrm{d}y + x^2\,\mathrm{d}w, \alpha \qquad ^2 = -z^2\,\mathrm{d}y + (xw - yz)\,\mathrm{d}z + xz\,\mathrm{d}w,$$

which can be written in the form

$$\alpha^1 = x\,\mathrm{d}(xw - yz) + x^2 y\,\mathrm{d}\left(\frac{z}{x}\right), \qquad \alpha^2 = z\,\mathrm{d}(xw - yz) + x^2 w\,\mathrm{d}\left(\frac{z}{x}\right)$$

(cf. Example 4.7). Hence, we find again that the integral manifolds sought for are given by $xw - yz = \text{const}$, $z/x = \text{const}$

Exercise 7.25 Verify that the matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

form a basis of a Lie subalgebra of $\mathfrak{gl}(2, \mathbb{R})$ and identify the corresponding subgroup of $\mathrm{GL}(2, \mathbb{R})$.

7.3 Invariant Differential Forms

In this section we shall see explicitly that from a given Lie algebra one can find a Lie group. This process is simplified by the use of differential forms.

Let G be a Lie group and let ω be a differential form on G, we say that ω is *left-invariant* if $L_g{}^*\omega = \omega$ for all $g \in G$. If ω is left-invariant, $\mathrm{d}\omega$ is also left-invariant, since, according to (3.38), $L_g{}^*\,\mathrm{d}\omega = \mathrm{d}L_g{}^*\omega = \mathrm{d}\omega$. Given two left-invariant differential forms, ω_1 and ω_2, the combinations $a\omega_1 + b\omega_2$ and $\omega_1 \wedge \omega_2$ also are left-invariant differential forms, for $a, b \in \mathbb{R}$ [see (2.30) and (3.25)]. Thus, the set of all the left-invariant differential forms is a subalgebra of the algebra of forms of G, which is closed under the operator of exterior differentiation.

Exercise 7.26 Show that a 0-form on G, that is, a differentiable function $f : G \to \mathbb{R}$, is left-invariant if and only if f is constant.

As in the case of a left-invariant vector field, a left-invariant differential form is determined by its value at the identity. Therefore, the set of left-invariant k-forms forms a vector subspace of $\Lambda^k(G)$ of dimension $\binom{n}{k}$.

Exercise 7.27 Let α be a left-invariant 1-form. Show that $\alpha_g = \alpha_e \circ L_{g^{-1}*g}$, for $g \in G$.

According to the formula $\alpha_g = \alpha_e \circ L_{g^{-1}*g}$, established in Exercise 7.27, if the x^i form a local coordinate system on a neighborhood of the identity, the value of any left-invariant 1-form α at the identity can be expressed in the form $\alpha_e = a_i\, \mathrm{d}x^i_e$, where the a_i are some real numbers. Hence, for any other point g in the domain of the coordinate system we have [see (1.49) and (1.24)]

$$\begin{aligned}
\alpha_g &= \alpha_g\left(\left(\frac{\partial}{\partial x^i}\right)_g\right)\mathrm{d}x^i_g = \alpha_e\left(L_{g^{-1}*g}\left(\frac{\partial}{\partial x^i}\right)_g\right)\mathrm{d}x^i_g \\
&= \left(a_j\, \mathrm{d}x^j_e\right)\left[\left(\frac{\partial}{\partial x^i}\right)_g \left[x^k \circ L_{g^{-1}}\right]\left(\frac{\partial}{\partial x^k}\right)_e\right]\mathrm{d}x^i_g \\
&= a_j\left[\left(\frac{\partial}{\partial x^i}\right)_g \left[x^j \circ L_{g^{-1}}\right]\right]\mathrm{d}x^i_g. \qquad (7.33)
\end{aligned}$$

Example 7.28 By combining (7.5) and (7.6) we find that

$$\begin{aligned}
x^1 \circ L_{g^{-1}} &= \frac{1 + x^2(g)x^3(g)}{x^1(g)}x^1 - x^2(g)x^3, \\
x^2 \circ L_{g^{-1}} &= \frac{1 + x^2(g)x^3(g)}{x^1(g)}x^2 - x^2(g)\frac{1 + x^2x^3}{x^1}, \\
x^3 \circ L_{g^{-1}} &= -x^3(g)x^1 + x^1(g)x^3.
\end{aligned}$$

Hence, according to (7.33) one readily sees that the left-invariant 1-forms ω^i on $\mathrm{SL}(2, \mathbb{R})$, whose values at the identity are $\mathrm{d}x^i_e$, are given locally by

$$\begin{aligned}
\omega^1 &= \frac{1 + x^2x^3}{x^1}\,\mathrm{d}x^1 - x^2\,\mathrm{d}x^3, \\
\omega^2 &= \frac{x^2(1 + x^2x^3)}{(x^1)^2}\,\mathrm{d}x^1 + \frac{1}{x^1}\,\mathrm{d}x^2 - \frac{(x^2)^2}{x^1}\,\mathrm{d}x^3, \\
\omega^3 &= -x^3\,\mathrm{d}x^1 + x^1\,\mathrm{d}x^3.
\end{aligned}$$

The exterior derivative of ω^i is a left-invariant 2-form and therefore can be written as a linear combination of $\{\omega^1 \wedge \omega^2, \omega^2 \wedge \omega^3, \omega^3 \wedge \omega^1\}$. A straightforward computation shows that

$$\mathrm{d}\omega^1 = -\omega^2 \wedge \omega^3, \qquad \mathrm{d}\omega^2 = -2\omega^1 \wedge \omega^2, \qquad \mathrm{d}\omega^3 = 2\omega^1 \wedge \omega^3.$$

As shown below, the coefficients in this linear combinations are related to the structure constants of the group with respect to the dual basis to $\{\omega^1, \omega^2, \omega^3\}$ [see (7.35)].

Note that these relations, being equalities between left-invariant 2-forms, hold globally on the whole group manifold, not only in the domain of the coordinate system employed.

Exercise 7.29 Find a basis for the left-invariant 1-forms of $\mathbb{R}^n$ and of $\mathrm{GL}(n, \mathbb{R})$.

Exercise 7.30 Show that if $\phi : G \to H$ is a homomorphism of Lie groups and ω is a left-invariant k-form on H, then $\phi^*\omega$ is left-invariant on G.

Exercise 7.31 A differential form ω on G is right-invariant if ${R_g}^*\omega = \omega$, for all $g \in G$. Show that if $\iota : G \to G$ is the inversion mapping, $\iota(g) \equiv g^{-1}$, then $\iota^*\omega$ is right-invariant if and only if ω is left-invariant.

The Maurer–Cartan Equations Let $\{\omega^1, \ldots, \omega^n\}$ be a basis for the 1-forms on G and let $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ be a basis for the vector fields on G such that $\omega^i(\mathbf{X}_j) = \delta^i_j$ (that is, these bases are dual to each other); then the elements of each of these bases are left-invariant if and only if the elements of the other are. This follows from ${L_g}^*[\omega^i(\mathbf{X}_j)] = ({L_g}^*\omega^i)({L_g}^*\mathbf{X}_j)$ and ${L_g}^*\delta^i_j = \delta^i_j$, wherefore $({L_g}^*\omega^i)({L_g}^*\mathbf{X}_j) = \delta^i_j$. This relation and the fact that for a given basis there exists only one dual basis prove the assertion above.

If $\{\omega^1, \ldots, \omega^n\}$ is a basis of the space of left-invariant 1-forms, then the exterior products $\omega^j \wedge \omega^k$ with $j < k$ form a basis for the left-invariant 2-forms. Since $\mathrm{d}\omega^i$ is a left-invariant 2-form, $\mathrm{d}\omega^i$ should be a linear combination (with constant coefficients) of the products $\omega^j \wedge \omega^k$ with $j < k$. In fact, if $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ is the dual basis to $\{\omega^1, \ldots, \omega^n\}$, the components of $\mathrm{d}\omega^i$ are given by [see (3.30)]

$$\begin{aligned}
\mathrm{d}\omega^i(\mathbf{X}_j, \mathbf{X}_k) &= \frac{1}{2}\Big\{\mathbf{X}_j\big(\omega^i(\mathbf{X}_k)\big) - \mathbf{X}_k\big(\omega^i(\mathbf{X}_j)\big) - \omega^i\big([\mathbf{X}_j, \mathbf{X}_k]\big)\Big\} \\
&= -\frac{1}{2}\omega^i\big([\mathbf{X}_j, \mathbf{X}_k]\big) = -\frac{1}{2}\omega^i\big(c^l_{jk}\mathbf{X}_l\big) \\
&= -\frac{1}{2}c^i_{jk},
\end{aligned}$$

where the c^i_{jk} are the structure constants of the group with respect to the basis $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$; therefore

$$\mathrm{d}\omega^i = -\frac{1}{2}c^i_{jk}\omega^j \wedge \omega^k. \tag{7.34}$$

These relations are known as the *Maurer–Cartan equations*. Taking into account the skew-symmetry of the structure constants in the two subscripts [equation (7.22)] and that of the exterior product of 1-forms, it follows that

$$\mathrm{d}\omega^i = -\sum_{j<k} c^i_{jk}\omega^j \wedge \omega^k. \tag{7.35}$$

Exercise 7.32 Show that $\mathrm{d}^2\omega^i = 0$ amounts to $c^m_{ij} c^l_{mk} + c^m_{jk} c^l_{mi} + c^m_{ki} c^l_{mj} = 0$.

Employing the Maurer–Cartan equations it is possible to determine *locally* the group G starting from its structure constants. For instance, if $c^k_{ij} = 0$, that is, if $\mathfrak{g}$ is Abelian, equations (7.34) give $\mathrm{d}\omega^i = 0$, which means that, locally, there exist n functions x^i such that $\omega^i = \mathrm{d}x^i$ (that is, being closed, the ω^i are locally exact). The functions x^i form a local coordinate system, since the ω^i are linearly independent. Since, in addition, the ω^i are left-invariant, $L_g{}^* \mathrm{d}x^i = \mathrm{d}x^i$, but $L_g{}^* \mathrm{d}x^i = \mathrm{d}(L_g{}^* x^i)$; therefore $\mathrm{d}(L_g{}^* x^i - x^i) = 0$, for $g \in G$, and this implies that $L_g{}^* x^i - x^i$ is a constant (which may depend on g), $a^i(g)$. Thus

$$L_g{}^* x^i = x^i + a^i(g),$$

hence, for $g' \in G$ such that g' and gg' belong to the domain of the coordinates x^i,

$$x^i(gg') = (x^i \circ L_g)(g') = (L_g{}^* x^i)(g') = x^i(g') + a^i(g). \tag{7.36}$$

In particular, if $g' = e$, from the previous equation we obtain

$$x^i(g) = x^i(e) + a^i(g),$$

so that $a^i(g) = x^i(g) - x^i(e)$ and substituting into (7.36)

$$x^i(gg') = x^i(g') + x^i(g) - x^i(e), \tag{7.37}$$

and therefore $x^i(gg') = x^i(g'g)$, that is, G is Abelian.

If we define $y^i \equiv x^i - x^i(e)$, then equation (7.37) amounts to $y^i(gg') = y^i(g') + y^i(g)$, which is identical to the relation found in the case of the additive group $\mathbb{R}^n$ (see Example 7.8); however, since the coordinates x^i (and the y^i) may not cover all of G, this does not imply that G be isomorphic to $\mathbb{R}^n$ globally, but only locally. As pointed out above, the structure constants determine the group G only locally. However, the structure constants define a unique simply connected Lie group, which is a covering group of any other Lie group with the given structure constants [see, e.g., Warner (1983), Sattinger and Weaver (1986)].

A second example of the determination of the group from its structure constants is given by the Lie algebra of dimension two given by (7.24); in this case the Maurer–Cartan equations yield

$$\mathrm{d}\omega^1 = 0, \qquad \mathrm{d}\omega^2 = -\omega^1 \wedge \omega^2. \tag{7.38}$$

The first of these equations implies that ω^1 is locally exact, that is, there exists locally a function x^1 such that

$$\omega^1 = \mathrm{d}x^1. \tag{7.39}$$

Substituting this expression into the second equation of (7.38) we have

$$\mathrm{d}\omega^2 + \mathrm{d}x^1 \wedge \omega^2 = 0,$$

which amounts to $d(e^{x^1}\omega^2) = 0$; therefore, locally, there exists a function x^2 such that

$$\omega^2 = e^{-x^1} dx^2. \tag{7.40}$$

The functions x^1 and x^2 form a local coordinate system in some neighborhood in G ($dx^1 \wedge dx^2 = e^{x^1}\omega^1 \wedge \omega^2 \neq 0$).

From the condition $\omega^i = {L_g}^*\omega^i$, for $g \in G$, and equations (7.39) and (7.40) we have

$$\begin{aligned} dx^1 &= {L_g}^* dx^1 = d\big({L_g}^* x^1\big) = d\big(x^1 \circ L_g\big), \\ e^{-x^1} dx^2 &= {L_g}^*\big(e^{-x^1} dx^2\big) = e^{-(x^1 \circ L_g)} d\big(x^2 \circ L_g\big), \end{aligned}$$

which leads to

$$x^1 \circ L_g = x^1 + a^1(g), \tag{7.41}$$

where $a^1(g)$ is a constant (which may depend on g), and $d(x^2 \circ L_g) = e^{x^1 \circ L_g - x^1} dx^2 = e^{a^1(g)} dx^2$; therefore

$$x^2 \circ L_g = e^{a^1(g)} x^2 + a^2(g), \tag{7.42}$$

where $a^2(g)$ is another constant. Evaluating both sides of (7.41) and (7.42) at e we obtain

$$x^1(g) = x^1(e) + a^1(g), \qquad x^2(g) = e^{a^1(g)} x^2(e) + a^2(g),$$

so that $a^1(g) = x^1(g) - x^1(e)$ and $a^2(g) = x^2(g) - e^{x^1(g) - x^1(e)} x^2(e)$. Substituting these expressions into (7.41) and (7.42), and evaluating at g', we then obtain

$$\begin{aligned} x^1(gg') &= x^1(g') + x^1(g) - x^1(e), \\ x^2(gg') &= e^{x^1(g) - x^1(e)}\big[x^2(g') - x^2(e)\big] + x^2(g). \end{aligned} \tag{7.43}$$

Equivalently, defining the coordinates

$$y^1 \equiv e^{[x^1 - x^1(e)]/2}, \qquad y^2 \equiv [x^2 - x^2(e)]\, e^{-[x^1 - x^1(e)]/2},$$

equations (7.43) become

$$y^1(gg') = y^1(g) y^1(g'), \qquad y^2(gg') = y^1(g) y^2(g') + \frac{y^2(g)}{y^1(g')}.$$

It can readily be seen that these equations correspond to the group formed by the upper triangular 2×2 real matrices with determinant equal to 1, with the usual matrix multiplication, if the elements of this group are expressed in the form

$$g = \begin{pmatrix} y^1(g) & y^2(g) \\ 0 & 1/y^1(g) \end{pmatrix}$$

(cf. Example 7.24).

Alternatively, we can make $y^1 \equiv \mathrm{e}^{x^1 - x^1(e)}$, $y^2 \equiv x^2 - x^2(e)$, so that equations (7.43) take the form

$$y^1(gg') = y^1(g)y^1(g'), \qquad y^2(gg') = y^1(g)y^2(g') + y^2(g), \tag{7.44}$$

which coincide with the equations corresponding to the group formed by the 2×2 matrices of the form $\left(\begin{smallmatrix} y^1(g) & y^2(g) \\ 0 & 1 \end{smallmatrix}\right)$, $y^1(g) > 0$, with the usual operation of matrix multiplication (see Example 7.4).

Exercise 7.33 Verify that the structure constants $c^1_{13} = -c^1_{31} = 1, c^2_{23} = -c^2_{32} = k$, with all the others being equal to zero, define a Lie algebra of dimension three, and find the local expressions for the operation of the corresponding group or groups, by integrating the Maurer–Cartan equations.

Exercise 7.34 Verify that the structure constants $c^2_{12} = -c^2_{21} = 1, c^2_{13} = -c^2_{31} = 1, c^3_{13} = -c^3_{31} = 1$, with all the others being equal to zero, define a Lie algebra of dimension three, and find the local expressions for the operation of the corresponding group or groups, by integrating the Maurer–Cartan equations.

Invariant Forms on Subgroups of $\mathrm{GL}(n, \mathbb{R})$ In the case of the Lie subgroups of $\mathrm{GL}(n, \mathbb{R})$, there exists a particularly simple form of finding a basis for the left-invariant or the right-invariant 1-forms. As shown in Example 7.15, the vector fields $x^k_i (\partial/\partial x^k_j)$, where the x^i_j are the natural coordinates on $\mathrm{GL}(n, \mathbb{R})$, form a basis for the left-invariant vector fields of $\mathrm{GL}(n, \mathbb{R})$. Hence, if H is a Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$ and $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p\}$, where $p = \dim H$, is a basis of $\mathfrak{h}$, the Lie algebra of H, then there exist real numbers $(\lambda_a)^i_j$, with $a = 1, 2, \dots, p$ and $i, j = 1, 2, \dots, n$, such that

$$\mathbf{X}_a = (\lambda_a)^i_j x^k_i \frac{\partial}{\partial x^k_j}, \quad a = 1, \dots, \dim H \tag{7.45}$$

(see, e.g., Example 7.16) and the $n \times n$ matrices $\lambda_a \equiv ((\lambda_a)^i_j)$ satisfy the commutation relations $[\lambda_a, \lambda_b] \equiv \lambda_a \lambda_b - \lambda_b \lambda_a = c^r_{ab} \lambda_r$, where the c^r_{ab} are the structure constants of the basis $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p\}$ (see Example 7.15). The 1-forms $(\iota^* x^i_j)\, \mathrm{d}x^j_k$, where ι is the inversion map, $\iota(g) \equiv g^{-1}$, form the dual basis to $\{x^k_i (\partial/\partial x^k_j)\}$ and, therefore, are left-invariant on $\mathrm{GL}(n, \mathbb{R})$; hence, the restriction of $(\iota^* x^i_j)\, \mathrm{d}x^j_k$ to H is equal to $(\lambda_a)^i_k \omega^a$, where the 1-forms ω^a are left-invariant on H and form the dual basis to (7.45).

In effect, from (7.45), using that $(\iota^* x^i_j) x^j_k = \delta^i_k$, we have

$$\left[\left(\iota^* x^i_j\right) \mathrm{d}x^j_k\right](\mathbf{X}_a) = \left(\iota^* x^i_j\right)(\lambda_a)^m_k x^j_m = (\lambda_a)^i_k$$

and, on the other hand,

$$\left[(\lambda_b)^i_k \omega^b\right](\mathbf{X}_a) = (\lambda_b)^i_k \delta^b_a = (\lambda_a)^i_k.$$

Since at each point of H, the vector fields $\mathbf{X}_a$ form a basis of the tangent space to H, it follows that the restriction of $(\iota^* x^i_j)\,\mathrm{d}x^j_k$ to H coincides with $(\lambda_b)^i_k \omega^b$. Thus, we have proved the following.

Theorem 7.35 *If H is a Lie subgroup of* $\mathrm{GL}(n,\mathbb{R})$ *and* $j : H \to \mathrm{GL}(n,\mathbb{R})$ *denotes the inclusion map, then* $j^*((\iota^* x^i_j)\,\mathrm{d}x^j_k) = (\lambda_a)^i_k \omega^a$, *where the* ω^a *are the left-invariant* 1*-forms on* H *that form the dual basis to* (7.45).

Expressed in matrix form, this theorem shows that if g represents an arbitrary element of H, then

$$g^{-1}\,\mathrm{d}g = \lambda_a \omega^a, \tag{7.46}$$

where $\mathrm{d}g$ is the matrix whose entries are the differentials of the entries of g.

Example 7.36 The basis of $\mathfrak{sl}(2,\mathbb{R})$ [the Lie algebra of $\mathrm{SL}(2,\mathbb{R})$] given by (7.18) and (7.19) is of the form (7.45), where the λ_a are the matrices given in (7.20). Making use of the expression (7.4) one readily finds that

$$\begin{aligned} g^{-1}\,\mathrm{d}g &= \begin{pmatrix} \frac{1+x^2x^3}{x^1} & -x^2 \\ -x^3 & x^1 \end{pmatrix} \begin{pmatrix} \mathrm{d}x^1 & \mathrm{d}x^2 \\ \mathrm{d}x^3 & \mathrm{d}\left(\frac{1+x^2x^3}{x^1}\right) \end{pmatrix} \\ &= \left(\frac{1+x^2x^3}{x^1}\,\mathrm{d}x^1 - x^2\,\mathrm{d}x^3\right)\lambda_1 \\ &\quad + \left(\frac{x^2(1+x^2x^3)}{(x^1)^2}\,\mathrm{d}x^1 + \frac{\mathrm{d}x^2}{x^1} - \frac{(x^2)^2}{x^1}\,\mathrm{d}x^3\right)\lambda_2 + \left(x^1\,\mathrm{d}x^3 - x^3\,\mathrm{d}x^1\right)\lambda_3. \end{aligned}$$

According to (7.46), the coefficients of the matrices λ_a are the left-invariant 1-forms that form the dual basis to (7.18) and (7.19), and they coincide with the left-invariant 1-forms obtained in Example 7.28. (See also Examples B.1, B.6, and B.8.)

Exercise 7.37 Find the basis of the left-invariant 1-forms for the group formed by the real 2×2 matrices of the form $\left(\begin{smallmatrix} x & y \\ 0 & z \end{smallmatrix}\right)$, with $xz \neq 0$, dual to the basis of left-invariant vector fields corresponding to the matrices

$$\lambda_1 \equiv \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad \lambda_2 \equiv \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \lambda_3 \equiv \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

In a similar manner one finds that the vector fields

$$\dot{\mathbf{X}}_a = (\lambda_a)^i_j x^j_k \frac{\partial}{\partial x^i_k}, \qquad a = 1, \ldots, \dim H,$$

form a basis for the right-invariant vector fields on H and its dual basis, $\{\dot{\omega}^a\}$, is such that $j^*(x^i_j\,\mathrm{d}(\iota^* x^j_k)) = -(\lambda_a)^i_k \dot{\omega}^a$ or, equivalently, $j^*((\iota^* x^j_k)\,\mathrm{d}x^i_j) = (\lambda_a)^i_k \dot{\omega}^a$. In terms of matrices, we have $g\,\mathrm{d}g^{-1} = -\lambda_a \dot{\omega}^a$, which amounts to $(\mathrm{d}g)g^{-1} = \lambda_a \dot{\omega}^a$. Comparing with (7.46) it follows that $\dot{\omega}^a = -\iota^* \omega^a$ (cf. Exercise 7.31).

7.4 One-Parameter Subgroups and the Exponential Map

In any Lie group, the one-parameter subgroups are particularly important. A set of elements of G, $\{g_t\}$, with $t \in \mathbb{R}$, is a *one-parameter subgroup* of G if $g_t g_s = g_{t+s}$ and if g_t depends differentiably on the parameter t. This implies that $g_0 = e$ and that $g_{-t} = (g_t)^{-1}$. The map from $\mathbb{R}$ into G given by $t \mapsto g_t$ is then a (differentiable) curve in G starting at the identity. The following result relates the one-parameter groups of G with the left-invariant or right-invariant vector fields and is a particular case of the relation between subgroups and subalgebras mentioned at the end of Sect. 7.2.

Theorem 7.38 *Let $\{g_t\}$, with $t \in \mathbb{R}$, be a one-parameter subgroup of G; then the curve $t \mapsto g_t$ is the integral curve starting at e of some left-invariant (or right-invariant) vector field.*

Proof Let ξ be the tangent vector to the curve $\sigma(t) \equiv g_t$ at $t = 0$, $\xi = \sigma'_0$, that is,

$$\xi[f] = \frac{\mathrm{d}}{\mathrm{d}t}\big(f(g_t)\big)\Big|_{t=0} \quad \text{for } f \in C^\infty(G). \tag{7.47}$$

Similarly, the tangent vector to this curve at $t = s$, σ'_s, is such that

$$\sigma'_s[f] = \frac{\mathrm{d}}{\mathrm{d}t}\big(f(g_t)\big)\Big|_{t=s};$$

but $g_t = g_s g_{t-s}$ and making the change of variable $u = t - s$, we have

$$\begin{aligned}\sigma'_s[f] &= \frac{\mathrm{d}}{\mathrm{d}t}\big(f(g_s g_{t-s})\big)\Big|_{t=s} = \frac{\mathrm{d}}{\mathrm{d}u}\big(f(g_s g_u)\big)\Big|_{u=0}\\ &= \frac{\mathrm{d}}{\mathrm{d}u}\big((f \circ L_{g_s})(g_u)\big)\Big|_{u=0} = \xi[f \circ L_{g_s}] = (L_{g_s *e}\xi)[f]\end{aligned}$$

that is, $\sigma'_s = L_{g_s *e}\xi = \mathbf{X}_{g_s} = \mathbf{X}_{\sigma(s)}$, where $\mathbf{X}$ is the left-invariant vector field such that $\mathbf{X}_e = \xi$ [see (7.15)]. Thus showing that σ is an integral curve of $\mathbf{X}$.

Alternatively, from the previous expressions we also have

$$\begin{aligned}\sigma'_s[f] &= \frac{\mathrm{d}}{\mathrm{d}u}\big(f(g_u g_s)\big)\Big|_{u=0} = \frac{\mathrm{d}}{\mathrm{d}u}\big((f \circ R_{g_s})(g_u)\big)\Big|_{u=0}\\ &= \xi[f \circ R_{g_s}] = (R_{g_s *e}\xi)[f],\end{aligned}$$

which means that $\sigma'_s = R_{g_s *e}\xi$, which is the value at $\sigma(s)$ of the right-invariant vector field whose value at the identity is ξ. □

Conversely, given a left-invariant or right-invariant vector field on G (or, equivalently, given $\xi \in T_e G$) there exists a one-parameter subgroup of G, $\{g_t\}$, such that $t \mapsto g_t$ is the integral curve starting at e of the given vector field (or, equivalently, ξ is the tangent vector to the curve $t \mapsto g_t$ at $t = 0$). Indeed, if $\mathbf{X}$ is any vector field on G, according to the existence and uniqueness theorems for systems of ODEs (see Sect. 2.1), there exists an integral curve of $\mathbf{X}$ starting at e, defined in some neighborhood I_e of 0 for which we have: if $t \mapsto \varphi(g,t)$ is the integral curve of $\mathbf{X}$ starting at g, then

$$\varphi(g, t+s) = \varphi\big(\varphi(g,t), s\big), \tag{7.48}$$

for all those values of t and s for which both sides of the equation are defined [see (2.6)]. In the case of a left-invariant vector field on G, $\varphi(g,t)$ is defined for all $t \in \mathbb{R}$. In fact, if C is an integral curve of $\mathbf{X}$ that starts at e [i.e., $C(t) = \varphi(e,t)$], then, for $g \in G$, $\tilde{C}(t) \equiv (L_g \circ C)(t)$ is defined in the same neighborhood I_e of 0 as C and $\tilde{C}(0) = L_g(C(0)) = L_g(e) = ge = g$. The tangent vector to $\tilde{C}$ at $t = 0$ is $L_{g*e} C'(0) = L_{g*e}\mathbf{X}_e = \mathbf{X}_g$, since $\mathbf{X} \in \mathfrak{g}$ [see (7.14)]; therefore $\tilde{C}$ is an integral curve of $\mathbf{X}$ starting at g. Since g is arbitrary, from (7.48) we see that $\varphi(g,t)$ is defined for all $t \in \mathbb{R}$. On the other hand, $\tilde{C}(t) = \varphi(g,t)$, that is, $\varphi(g,t) = (L_g \circ C)(t) = g\,C(t) = g\,\varphi(e,t)$. Taking $g = \varphi(e,s)$ in this equation and using (7.48), we then obtain

$$\varphi\big(\varphi(e,s), t\big) = \varphi(e,s)\varphi(e,t) = \varphi(e, t+s), \tag{7.49}$$

which means that the elements of G defined by $g_t = \varphi(e,t)$ form a one-parameter subgroup of G. It can readily be verified, basically replacing L_g by R_g in the foregoing derivation, that (7.49) also holds if $\mathbf{X}$ is right-invariant.

The element $\varphi(e,1)$ is denoted by $\exp \mathbf{X}$ and the map from $\mathfrak{g}$ into G given by $\exp \mathbf{X} = \varphi(e,1)$ is called the *exponential map*. It can readily be seen that $\exp t\mathbf{X} = \varphi(e,t)$ and that $\exp(s+t)\mathbf{X} = (\exp s\mathbf{X})(\exp t\mathbf{X})$. The element $\exp(\mathbf{X}+\mathbf{Y})$ may not coincide with $(\exp \mathbf{X})(\exp \mathbf{Y})$ unless $[\mathbf{X}, \mathbf{Y}] = 0$. The exponential map, $\exp : \mathfrak{g} \to G$, is not always injective or onto; in some groups there exist elements that are not the exponential of some $\mathbf{X} \in \mathfrak{g}$ (see the examples given below). Nevertheless, any element of a group G belonging to the connected component of the identity (that is, it can be joined with the identity by means of a curve in G), can be expressed as the product of exponentials, $\exp \mathbf{X}_1 \exp \mathbf{X}_2 \cdots \exp \mathbf{X}_k$ (see Example 7.41).

Exercise 7.39 Show that $\exp t\mathbf{X} = \varphi(e,t)$. (*Hint*: consider the curve $\tau(s) \equiv \varphi(e, st)$, with t fixed, and calculate τ'_0.)

Example 7.40 Consider again the group formed by the 2×2 real matrices of the form $g = \left(\begin{smallmatrix} x^1(g) & x^2(g) \\ 0 & 1 \end{smallmatrix}\right)$, $x^1(g) \neq 0$ (see Example 7.4). Each element of this group can be identified with a point of $\mathbb{R}^2$ excluding the x^2 axis, which allows us to see that this group is not a connected set, but has two components (identified with the right and left half-planes). Using (7.44), it can readily be verified that the vector fields $\mathbf{X}_1 \equiv x^1\,\partial/\partial x^1$, and $\mathbf{X}_2 \equiv x^1\,\partial/\partial x^2$ form a basis of the Lie algebra of this group (see Exercise 7.20). Any element of this algebra can be expressed in the form $\mathbf{X} = a x^1\,\partial/\partial x^1 + b x^1\,\partial/\partial x^2$, with $a, b \in \mathbb{R}$; then $\exp t\mathbf{X}$ corresponds to the solution of the system of equations [see (2.4)]

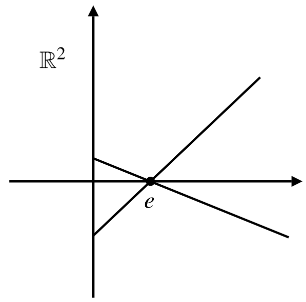

Fig. 7.2 The matrix group considered in Example 7.40 is identified with the Cartesian plane with the vertical axis removed. The *point* $(1, 0)$ corresponds to the identity and the one-parameter subgroups correspond to the intersections of the *straight lines* passing through $(1, 0)$ with the half–plane $x^1 > 0$

$$\frac{\mathrm{d}x^1}{\mathrm{d}t} = ax^1, \qquad \frac{\mathrm{d}x^2}{\mathrm{d}t} = bx^1,$$

where, by abuse of notation, we have written x^i in place of $x^i \circ C$, with the initial condition $x^1(0) = 1$, $x^2(0) = 0$ (so that the integral curve of $\mathbf{X}$ starts at e). Then, it can readily be seen that, if $a \neq 0$, $x^1(t) = \mathrm{e}^{at}$, then $x^2(t) = b(\mathrm{e}^{at} - 1)/a$, that is,

$$\exp t\mathbf{X} = \begin{pmatrix} \mathrm{e}^{at} & \frac{b}{a}(\mathrm{e}^{at} - 1) \\ 0 & 1 \end{pmatrix}, \quad a \neq 0.$$

When $a = 0$ one obtains $x^1(t) = 1$, $x^2(t) = bt$. Eliminating the parameter t from the foregoing expressions, one finds that $ax^2 = b(x^1 - 1)$, which is the equation of a straight line passing through the point (1, 0), which corresponds to the identity (see Fig. 7.2). Since $x^1(t) = \mathrm{e}^{at} > 0$, in this case the image of the exponential map is one half of G (the connected component of the identity).

Example 7.41 In terms of the parametrization of the group $\mathrm{SL}(2, \mathbb{R})$ given by (7.4), in a neighborhood of the identity, any left-invariant vector field can be expressed in the form $\mathbf{X} = a^i \mathbf{X}_i$, with $a^i \in \mathbb{R}$ and $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ being the basis of $\mathfrak{sl}(2, \mathbb{R})$ given by (7.18) and (7.19). The integral curve of $\mathbf{X}$ starting at e corresponds to the solution of the system of equations

$$\begin{aligned} \frac{\mathrm{d}x^1}{\mathrm{d}t} &= a^1x^1 + a^3x^2, \\ \frac{\mathrm{d}x^2}{\mathrm{d}t} &= -a^1x^2 + a^2x^1, \\ \frac{\mathrm{d}x^3}{\mathrm{d}t} &= a^1x^3 + a^3\frac{1 + x^2x^3}{x^1}, \end{aligned} \tag{7.50}$$

where, for simplicity, we have written x^i instead of $x^i \circ C$, with the initial condition $x^1(0) = 1$, $x^2(0) = 0 = x^3(0)$. By combining equations (7.50) one finds that each of the functions x^i and $(1 + x^2x^3)/x^1$ satisfies the equation

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2} = \left[\left(a^1\right)^2 + a^2a^3\right] f,$$

whose solution is

$$f(t) = \begin{cases} a\cos\sqrt{K}t + b\sin\sqrt{K}t & \text{if } K \equiv -[(a^1)^2 + a^2a^3] > 0, \\ a\cosh\sqrt{-K}t + b\sinh\sqrt{-K}t & \text{if } K < 0, \\ a + bt & \text{if } K = 0. \end{cases}$$

Hence, using again (7.50) and the initial conditions, one obtains

$$\exp\left(ta^i\mathbf{X}_i\right) = \begin{cases} \cos\sqrt{K}t\, I + \frac{\sin\sqrt{K}t}{\sqrt{K}} A & \text{if } K > 0, \\ \cosh\sqrt{-K}t\, I + \frac{\sinh\sqrt{-K}t}{\sqrt{-K}} A & \text{if } K < 0, \\ I + tA & \text{if } K = 0, \end{cases} \tag{7.51}$$

where I is the 2×2 identity matrix and $A \equiv \left(\begin{smallmatrix} a^1 & a^2 \\ a^3 & -a^1 \end{smallmatrix}\right)$ [cf. (6.27) and (6.28)]. (Note that $K = \det A$.) Even though the foregoing expressions were obtained making use of a local coordinate system, it turns out that equations (7.51) are globally valid [that is, for any value of t and for any $\mathbf{X} \in \mathfrak{sl}(2, \mathbb{R})$].

For any real number $a < 0$ and $a \neq -1$, the matrix $\left(\begin{smallmatrix} a & 0 \\ 0 & 1/a \end{smallmatrix}\right)$ belongs to $\mathrm{SL}(2, \mathbb{R})$, but cannot be expressed in the form $\exp\mathbf{X}$, as can be seen from (7.51), even though the set $\mathrm{SL}(2, \mathbb{R})$ is connected.

In fact, noting that the trace of the matrix A is equal to zero, from (7.51) we find that the trace of the exponential of any element of the Lie algebra of $\mathrm{SL}(2, \mathbb{R})$ belongs to the interval $[-2, 2]$, if $K > 0$; to the interval $[2, \infty)$, if $K < 0$; and is equal to 2, if $K = 0$. On the other hand, the trace of the matrix given above is equal to $a + 1/a$, which is less than -2 for $a < 0$, and $a \neq -1$.

Now,

$$\begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -a & 0 \\ 0 & -1/a \end{pmatrix}$$
$$= \exp\left(\pi(\mathbf{X}_2 - \mathbf{X}_3)\right) \exp\left(\ln|a|\mathbf{X}_1\right).$$

Example 7.42 In the case of the group $\mathrm{GL}(n, \mathbb{R})$, any left-invariant vector field is of the form $x_i^k a_j^i (\partial/\partial x_j^k)$ [see (7.17)], where (a_j^i) is some $n \times n$ matrix; therefore, denoting by A the matrix (a_j^i) and by $\mathbf{X}_A$ the vector field $x_i^k a_j^i (\partial/\partial x_j^k)$, as in Example 7.15, $\exp t\mathbf{X}_A$ is the solution of the system of linear ODEs

$$\frac{\mathrm{d}x_j^k}{\mathrm{d}t} = x_i^k a_j^i,$$

with the initial condition, $x^k_j|_{t=0} = \delta^k_j$. Noting that this system of equations can be written as the matrix equation

$$\frac{\mathrm{d}g}{\mathrm{d}t} = gA,$$

where $g \equiv (x^i_j)$, with $g|_{t=0} = I$ (the $n \times n$ identity matrix), it can readily be seen that the solution is given by

$$g(t) = \sum_{m=0}^{\infty} \frac{(tA)^m}{m!}.$$

Hence we have

$$\exp t\mathbf{X}_A = \sum_{m=0}^{\infty} \frac{(tA)^m}{m!}. \tag{7.52}$$

The series appearing in this equation is defined as the exponential of the matrix tA and is denoted by $\exp tA$ or e^{tA} (cf. Example 6.11). Thus, the exponential of any $\mathbf{X} \in \mathfrak{gl}(n, \mathbb{R})$ can be expressed by means of the series (7.52) which only involves the components of $\mathbf{X}_e$ with respect to the natural basis $(\partial/\partial x^i_j)$:

$$\exp t\mathbf{X}_A = \exp tA. \tag{7.53}$$

In particular, $\mathrm{GL}(1, \mathbb{R})$ is the group $\mathbb{R} \setminus \{0\}$ with the operation of multiplication, and therefore for this group the exponential is precisely the usual exponential function.

This result can be applied to the calculation of the exponential for any Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$. For instance, the basis of the Lie algebra of the group $\mathrm{SL}(2, \mathbb{R})$ given by (7.18) and (7.19), corresponds to the matrices (7.20), so that an arbitrary linear combination $a^i\mathbf{X}_i$ corresponds to the matrix $\begin{pmatrix} a^1 & a^2 \\ a^3 & -a^1 \end{pmatrix}$, which will be denoted by A as in Example 7.41. It can readily be seen that, for $m = 0, 1, 2, \ldots$, $A^{2m} = (-K)^m I$ and $A^{2m+1} = (-K)^m A$, where $K = \det A = -[(a^1)^2 + a^2a^3]$. Therefore, the series (7.52) becomes

$$\begin{aligned} \sum_{m=0}^{\infty} \frac{t^m}{m!} A^m &= \sum_{m=0}^{\infty} \frac{t^{2m}}{(2m)!} A^{2m} + \sum_{m=0}^{\infty} \frac{t^{2m+1}}{(2m+1)!} A^{2m+1} \\ &= \sum_{m=0}^{\infty} \frac{(-1)^m K^m t^{2m}}{(2m)!} I + \sum_{m=0}^{\infty} \frac{(-1)^m K^m t^{2m+1}}{(2m+1)!} A, \end{aligned}$$

which coincides with the result (7.51), as can be seen making use of the series expansions of the functions sin, cos, sinh, and cosh.

Exercise 7.43 Show that for $\mathbb{R}^n$, with the group operation being the usual sum, $\exp a^i(\partial/\partial x^i) = (a^1, a^2, \ldots, a^n)$. Thus, in this case, the exponential map is one-to-one and onto.

Exercise 7.44 Using the notation and the results of Exercise 7.19, show that in the case of the group SU(2) the exponential map is given by

$$\exp t a^i \mathbf{X}_i = \begin{pmatrix} \cos\frac{Kt}{2} - \mathrm{i}\frac{a^3}{K}\sin\frac{Kt}{2} & \frac{-\mathrm{i}a^1 - a^2}{K}\sin\frac{Kt}{2} \\ \frac{-\mathrm{i}a^1 + a^2}{K}\sin\frac{Kt}{2} & \cos\frac{Kt}{2} + \mathrm{i}\frac{a^3}{K}\sin\frac{Kt}{2} \end{pmatrix}$$

$$= \cos\frac{Kt}{2} I - \frac{\mathrm{i}}{K}\sin\frac{Kt}{2}\begin{pmatrix} a^3 & a^1 - \mathrm{i}a^2 \\ a^1 + \mathrm{i}a^2 & -a^3 \end{pmatrix} \tag{7.54}$$

with $K \equiv \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2}$. Thus, for this group, the exponential map is onto but not injective.

The results of Examples 7.40–7.42, and of Exercises 7.43 and 7.44 do not depend on having considered left-invariant vector fields; the same results are obtained employing right-invariant vector fields. For a given group, the value of $\exp \mathbf{X}$ only depends on $\mathbf{X}_e$.

Theorem 7.45 *Let G and H be Lie groups, and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. If $\phi : G \to H$ is a homomorphism of Lie groups, then for $\mathbf{X} \in \mathfrak{g}$, we have $\phi(\exp t\mathbf{X}) = \exp t(\phi_* \mathbf{X})$, where $\phi_* \mathbf{X}$ is the left-invariant vector field on H such that $(\phi_* \mathbf{X})_e = \phi_{*e} \mathbf{X}_e$.*

Proof Let $\gamma_t \equiv \phi(\exp t\mathbf{X})$; then γ_t is a one-parameter subgroup of H. Therefore, according to the preceding results, $\gamma_t = \exp t\mathbf{Y}$, where $\mathbf{Y} \in \mathfrak{h}$ is such that $\mathbf{Y}_e$ is the tangent vector to the curve $t \mapsto \gamma_t = \phi(\exp t\mathbf{X})$ at $t = 0$, which amounts to ϕ_{*e} applied to the tangent vector to the curve $t \mapsto \exp t\mathbf{X}$ at $t = 0$ [see (1.26)]. Therefore $\mathbf{Y}_e = \phi_{*e}\mathbf{X}_e = (\phi_* \mathbf{X})_e$, thus showing that $\mathbf{Y} = \phi_* \mathbf{X}$. □

Applying this theorem and some of the results established above we have the following proposition, which turns out to be very useful. Among the consequences of the following theorem is that the exponential map in $\mathrm{GL}(n, \mathbb{R})$ can only yield matrices with positive determinant and, therefore, is not onto.

Theorem 7.46 *Let A be an arbitrary $n \times n$ matrix, then* $\det \mathrm{e}^A = \mathrm{e}^{\mathrm{tr} A}$.

Proof For any $n \times n$ matrix, A, $\mathrm{e}^A = \exp \mathbf{X}_A$ (see Example 7.42) and since $\det : \mathrm{GL}(n, \mathbb{R}) \to \mathbb{R} \setminus \{0\} \simeq \mathrm{GL}(1, \mathbb{R})$ is a homomorphism of Lie groups, $\det \mathrm{e}^A = \det(\exp \mathbf{X}_A) = \exp(\det_* \mathbf{X}_A)$. On the other hand, $\det_* \mathbf{X}_A$ is a left-invariant vector field on $\mathrm{GL}(1, \mathbb{R})$ and, according to the results of Example 7.42, $\exp(\det_* \mathbf{X}_A)$ coincides with the usual exponential of the component of $(\det_* \mathbf{X}_A)_1$ with respect to the natural basis $(\partial/\partial x)_1$. But, from Example 7.23, $(\det_* \mathbf{X}_A)_1 \equiv \det_{*e}(\mathbf{X}_A)_e = \det_{*e} a^i_j (\partial/\partial x^i_j)_e = \mathrm{tr}\, A\, (\partial/\partial x)_1$, and therefore, $\det \mathrm{e}^A = \mathrm{e}^{\mathrm{tr} A}$. □

7.5 The Lie Algebra of the Right-Invariant Vector Fields

The set of the right-invariant vector fields on G forms a Lie algebra over $\mathbb{R}$ that will be denoted by $\dot{\mathfrak{g}}$. Each right-invariant vector field, $\dot{\mathbf{X}}$, is determined by its value at the identity,

$$\dot{\mathbf{X}}_g = R_{g*e}\dot{\mathbf{X}}_e; \tag{7.55}$$

therefore there exists a one-to-one correspondence between $\dot{\mathfrak{g}}$ and T_eG, and the dimension of $\dot{\mathfrak{g}}$ is the same as that of T_eG. Making use of this correspondence we can define a second bracket on T_eG, which will be denoted by $[\ ,\ ]^\bullet$. If $\xi, \zeta \in T_eG$,

$$[\xi, \zeta]^\bullet \equiv [\dot{\mathbf{X}}, \dot{\mathbf{Y}}]_e \tag{7.56}$$

where $\dot{\mathbf{X}}$ and $\dot{\mathbf{Y}}$ are the right-invariant vector fields such that $\xi = \dot{\mathbf{X}}_e$ and $\zeta = \dot{\mathbf{Y}}_e$.

Example 7.47 In the case of the group of affine motions of $\mathbb{R}$, with the coordinates defined in Example 7.4, the right translations are given by

$$x^1 \circ R_g = x^1(g)x^1, \qquad x^2 \circ R_g = x^2(g)x^1 + x^2$$

[see (7.2)]. Making use of (7.55) we find that the right-invariant vector fields $\dot{\mathbf{X}}_1$ and $\dot{\mathbf{X}}_2$, whose values at the identity are $(\partial/\partial x^1)_e$ and $(\partial/\partial x^2)_e$, respectively, are given by

$$(\dot{\mathbf{X}}_1)_g = \left(\frac{\partial}{\partial x^1}\right)_e \left[x^i \circ R_g\right] \left(\frac{\partial}{\partial x^i}\right)_g = x^1(g)\left(\frac{\partial}{\partial x^1}\right)_g + x^2(g)\left(\frac{\partial}{\partial x^2}\right)_g$$

that is, $\dot{\mathbf{X}}_1 = x^1\,\partial/\partial x^1 + x^2\,\partial/\partial x^2$ and, similarly, $\dot{\mathbf{X}}_2 = \partial/\partial x^2$. Thus, $[\dot{\mathbf{X}}_1, \dot{\mathbf{X}}_2] = -\dot{\mathbf{X}}_2$ and therefore $[(\partial/\partial x^1)_e, (\partial/\partial x^2)_e]^\bullet = -(\partial/\partial x^2)_e$ (cf. Exercise 7.20).

The following theorem relates the bracket (7.56) with that induced by the left-invariant vector fields, defined in Sect. 7.2.

Theorem 7.48 *Let* $\xi, \zeta \in T_eG$; *then* $[\xi, \zeta]^\bullet = -[\![\xi, \zeta]\!]$.

Proof Let $\dot{\mathbf{X}}$ and $\dot{\mathbf{Y}}$ be the right-invariant vector fields such that $\dot{\mathbf{X}}_e = \xi$, $\dot{\mathbf{Y}}_e = \zeta$, let $\mathbf{X} \in \mathfrak{g}$ such that $\mathbf{X}_e = \xi$, and let g_t be the one-parameter subgroup $g_t = \exp t\mathbf{X}$ defined in the preceding section. For $g' \in G$ arbitrary, the tangent vector to the curve $\gamma(t) \equiv L_{g_t}(g') = g_t g' = R_{g'}(g_t)$ at $t = 0$ satisfies

$$\begin{aligned}\gamma'_0[f] &= \frac{\mathrm{d}}{\mathrm{d}t} f\big(\gamma(t)\big)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}(f \circ R_{g'})(g_t)\Big|_{t=0} \\ &= \xi[f \circ R_{g'}] = (R_{g'*e}\xi)[f] = \dot{\mathbf{X}}_{g'}[f]\end{aligned}$$

[see (7.47) and (1.23)]. This means that $\dot{\mathbf{X}}$ is the infinitesimal generator of the one-parameter group of transformations L_{g_t} (i.e., a right-invariant vector field is the infinitesimal generator of a group of left translations); hence, according to (2.27)

$$[\dot{\mathbf{X}}, \dot{\mathbf{Y}}] = \pounds_{\dot{\mathbf{X}}} \dot{\mathbf{Y}} = \lim_{t \to 0} \frac{L_{g_t}{}^* \dot{\mathbf{Y}} - \dot{\mathbf{Y}}}{t} \tag{7.57}$$

and, from the definition (7.56), equation (7.55), and the chain rule, we have

$$\begin{aligned}
[\xi, \zeta]^{\bullet} &= \lim_{t \to 0} \frac{(L_{g_t}{}^* \dot{\mathbf{Y}})_e - \zeta}{t} = \lim_{t \to 0} \frac{L_{g_t^{-1} * g_t} \dot{\mathbf{Y}}_{g_t} - \zeta}{t} \\
&= \lim_{t \to 0} \frac{L_{g_t^{-1} * g_t} (R_{g_t * e} \zeta) - \zeta}{t} = \lim_{t \to 0} \frac{(L_{g_t^{-1}} \circ R_{g_t})_{*e} \zeta - \zeta}{t} \\
&= \lim_{t \to 0} \frac{(R_{g_t} \circ L_{g_t^{-1}})_{*e} \zeta - \zeta}{t} = \lim_{t \to 0} \frac{R_{g_t * g_t^{-1}} \mathbf{Y}_{g_t^{-1}} - \zeta}{t} \\
&= \lim_{t \to 0} \frac{(R_{g_t^{-1}}{}^* \mathbf{Y})_e - \mathbf{Y}_e}{t},
\end{aligned} \tag{7.58}$$

where $\mathbf{Y}$ is the left-invariant vector field such that $\zeta = \mathbf{Y}_e$.

On the other hand, the tangent vector to the curve $\delta(t) \equiv R_{g_t^{-1}}(g') = g' g_t^{-1} = g' g_{-t} = L_{g'}(g_{-t})$ at $t = 0$ satisfies

$$\begin{aligned}
\delta_0'[f] &= \frac{\mathrm{d}}{\mathrm{d}t} f\big(\delta(t)\big)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} (f \circ L_{g'})(g_{-t})\Big|_{t=0} = -\frac{\mathrm{d}}{\mathrm{d}u} (f \circ L_{g'})(g_u)\Big|_{u=0} \\
&= -\xi[f \circ L_{g'}] = -(L_{g' * e} \xi)[f] = -\mathbf{X}_{g'}[f],
\end{aligned}$$

that is, $-\mathbf{X}$ is the infinitesimal generator of the one-parameter group of transformations $R_{g_t^{-1}}$. Therefore, returning to the last expression in (7.58)

$$\lim_{t \to 0} \frac{(R_{g_t^{-1}}{}^* \mathbf{Y})_e - \mathbf{Y}_e}{t} = (\pounds_{-\mathbf{X}} \mathbf{Y})_e = [-\mathbf{X}, \mathbf{Y}]_e = -[\mathbf{X}, \mathbf{Y}]_e = -[\xi, \zeta]. \tag{7.59}$$

□

Making use of part of the steps of the proof of the previous theorem it can readily be seen that the following result also holds.

Theorem 7.49 *The Lie bracket of a right-invariant vector field with a left-invariant vector field vanishes.*

Proof Let $\dot{\mathbf{X}} \in \dot{\mathfrak{g}}$ and $\mathbf{Y} \in \mathfrak{g}$; then, proceeding as in (7.57) we have

$$[\dot{\mathbf{X}}, \mathbf{Y}] = \pounds_{\dot{\mathbf{X}}} \mathbf{Y} = \lim_{t \to 0} \frac{L_{g_t}{}^* \mathbf{Y} - \mathbf{Y}}{t} = 0,$$

since $\mathbf{Y}$ is left-invariant. □

If G is an Abelian Lie group, a vector field on G is left-invariant if and only if is right-invariant, then, according to the previous theorem, if $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$, we have $[\mathbf{X}, \mathbf{Y}] = 0$, since $\mathbf{Y}$ also belongs to $\dot{\mathfrak{g}}$. Thus, the Lie algebra of any Abelian group is Abelian.

Theorem 7.50 *If G is a connected Lie group and $\mathbf{Y}$ is a vector field on G such that $[\mathbf{X}, \mathbf{Y}] = 0$, for all $\mathbf{X} \in \mathfrak{g}$, then $\mathbf{Y}$ is right-invariant.*

Proof Let $\mathbf{X} \in \mathfrak{g}$ and let $g_t = \exp t\mathbf{X}$, then

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(R_{g_t}{}^*\mathbf{Y})\Big|_{t=s} &= \frac{\mathrm{d}}{\mathrm{d}t}\big(R_{g_s}{}^*(R_{g_{t-s}}{}^*\mathbf{Y})\big)\Big|_{t-s=0} = R_{g_s}{}^*\frac{\mathrm{d}}{\mathrm{d}u}(R_{g_u}{}^*\mathbf{Y})\Big|_{u=0} \\ &= R_{g_s}{}^*(\pounds_{\mathbf{X}}\mathbf{Y}) = R_{g_s}{}^*\big([\mathbf{X}, \mathbf{Y}]\big) = 0, \end{aligned}$$

since $\mathbf{X}$ is the infinitesimal generator of the one-parameter group of transformations R_{g_t} and by hypothesis $[\mathbf{X}, \mathbf{Y}] = 0$ for $\mathbf{X} \in \mathfrak{g}$. Hence, $R_{g_t}{}^*\mathbf{Y}$ does not depend on t, but $R_{g_0} = \mathrm{id}$, so that $R_{g_t}{}^*\mathbf{Y} = \mathbf{Y}$ or $R_{\exp t\mathbf{X}}{}^*\mathbf{Y} = \mathbf{Y}$, which means that $\mathbf{Y}$ is right-invariant, at least under the transformations corresponding to elements of G of the form $\exp t\mathbf{X}$, but if G is connected, any $g \in G$ is a product of exponentials, $g = \exp \mathbf{X}_1 \cdots \exp \mathbf{X}_k$ [see, e.g., Warner (1983, Chap. 3)]. Therefore $R_g{}^*\mathbf{Y} = \mathbf{Y}$ for all $g \in G$. □

Exercise 7.51 Show that if G is connected, a 1-form α is left-invariant if and only if $\pounds_{\dot{\mathbf{X}}}\alpha = 0$ for all $\dot{\mathbf{X}} \in \dot{\mathfrak{g}}$.

7.6 Lie Groups of Transformations

The Lie groups more commonly encountered arise as groups of transformations on some manifold or some vector space. For instance, the isometries generated by the Killing vector fields of a Riemannian manifold form a Lie group (see, e.g., Examples 6.12, 6.17, and Exercise 6.16). The orthogonal and the unitary groups correspond to the linear transformations that preserve the inner product of a vector space. Further examples are given below and in Sects. 8.5 and 8.6.

Definition 7.52 Let G be a Lie group and let M be a differentiable manifold. We say that G is a *Lie group of transformations* on M or that G *acts on the right* on M, if to each $g \in G$ there is associated a transformation from M onto itself in such a way that if xg denotes the image of $x \in M$ under the transformation defined by g, then the following conditions hold:

(1) The map from $G \times M$ in M given by $(g, x) \mapsto xg$ is differentiable.
(2) $x(g_1 g_2) = (xg_1)g_2$, for $g_1, g_2 \in G$ and $x \in M$.

We say that G *acts on the left* on M when condition (2) is replaced by $(g_1 g_2)x = g_1(g_2 x)$ (now we write gx instead of xg for the image of x under the transformation defined by g). From (2) it follows that $xe = x$ for all $x \in M$. In some cases, we shall also write $R_g(x)$ for xg.

It is said that G acts *freely* on M if the existence of some $x \in M$ such that $xg = x$ implies that $g = e$; the group G acts *effectively* on M if $xg = x$ for all $x \in M$ implies that $g = e$. In other words, G acts freely on M if the only transformation with fixed points is the one corresponding to e, whereas G acts effectively on M if the identity transformation of M only corresponds to e.

Exercise 7.53 Show that G acts effectively on M if and only if distinct elements of G define distinct transformations on M.

For $x \in M$, the *orbit* of x is formed by the images of x under all the elements of G, that is, the orbit of x is the set $\{xg \mid g \in G\}$. The group G acts *transitively* on M (or the action of G on M is transitive) if the orbit of any point $x \in M$ coincides with the whole manifold M. For instance, the group SE(2) acts transitively on the plane (see Example 7.7), while the orbits in $\mathbb{R}^3$ of the group of rotations about the origin, SO(3), are spheres and the action is not transitive (however, SO(3) does act transitively on each sphere centered at the origin).

Let G be a Lie group that acts on the *right* on a manifold M. Each $x \in M$ defines a differentiable map $\Phi_x : G \to M$, given by $\Phi_x(g) = xg$ for $g \in G$. For $\mathbf{X} \in \mathfrak{g}$, $g_t = \exp t\mathbf{X}$ is a one-parameter subgroup of G and, therefore, the transformations R_{g_t}, from M onto M, defined by $R_{g_t}(x) \equiv xg_t$, form a one-parameter group of transformations on M (see Sect. 2.1) whose infinitesimal generator will be denoted by $\mathbf{X}^+$. Hence, $\mathbf{X}^+_x$ is the tangent vector at $t = 0$ to the curve $t \mapsto R_{g_t}(x) = xg_t = \Phi_x(g_t)$; therefore $\mathbf{X}^+_x$ is the image under Φ_{x*e} of the tangent vector to the curve $t \mapsto g_t$ at $t = 0$, which is $\mathbf{X}_e$. Thus [see (1.26)]

$$\mathbf{X}^+_x = \Phi_{x*e}\mathbf{X}_e \tag{7.60}$$

[cf. (7.15)]. Since the Jacobian is a linear transformation, we have $(a\mathbf{X} + b\mathbf{Y})^+ = a\mathbf{X}^+ + b\mathbf{Y}^+$, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$ and $a, b \in \mathbb{R}$. As we shall see, the mapping $\mathbf{X} \mapsto \mathbf{X}^+$ is not only linear, but also a Lie algebra homomorphism (Theorem 7.61).

It may be noticed that, by virtue of the definition of the vector field $\mathbf{X}^+$, its integral curve starting at $x \in M$ is given by $t \mapsto \Phi_x(\exp t\mathbf{X})$.

Exercise 7.54 Show that if G acts freely on M and $\mathbf{X}^+$ vanishes at some point, then $\mathbf{X} = 0$.

Exercise 7.55 Show that if G acts effectively on M and $\mathbf{X}^+ = 0$ (the vector field whose value is zero everywhere), then $\mathbf{X} = 0$.

Exercise 7.56 Let $\phi : G \to H$ be a homomorphism of Lie groups. Show that for $h \in H$ and $g \in G$, the equation $hg \equiv h\phi(g)$, where $h\phi(g)$ is the product of two elements of H, defines an action of G on the right on H. Show that if $\mathbf{X} \in \mathfrak{g}$, then the vector field $\mathbf{X}^+$ on H is the vector field $\phi_*\mathbf{X}$ defined in Sect. 7.2.

Exercise 7.57 Show that if a Lie group G acts on the right on a manifold M with $xg \equiv \phi(g, x)$, then G acts on the left on M by means of $gx \equiv \phi(g^{-1}, x)$. In other words, the action of a group on the right can be turned into an action on the left, replacing g by g^{-1}. (*Hint*: show that $\phi(g_1 g_2, x) = \phi(g_2, \phi(g_1, x))$.)

Example 7.58 The group SO(3), formed by the orthogonal 3×3 matrices, with determinant equal to 1, corresponds to the rotations about the origin in $\mathbb{R}^3$. The natural action (on the right) of any element $g \in \mathrm{SO}(3)$ on a point $(a_1, a_2, a_3) \equiv \mathbf{a} \in \mathbb{R}^3$, is given by the matrix product $\mathbf{a}g$; that is, $R_g(\mathbf{a}) = \mathbf{a}g$. Since SO(3) is a subgroup of $\mathrm{GL}(3, \mathbb{R})$, we can make use of the coordinates x^i_j defined on the latter to parameterize the elements of SO(3). Denoting by x_i the usual coordinates of $\mathbb{R}^3$ (i.e., $x_i(\mathbf{a}) = a_i$) we have $(x_i \circ \Phi_{\mathbf{a}})(g) = x_i(\mathbf{a}g) = x_j(\mathbf{a}) x^j_i(g)$, that is,

$$x_i \circ \Phi_{\mathbf{a}} = x_j(\mathbf{a}) x^j_i. \tag{7.61}$$

The group SO(3) corresponds to the submanifold of $\mathrm{GL}(3, \mathbb{R})$ defined by the equations

$$x^k_i x^l_j \delta_{kl} = \delta_{ij}, \qquad \frac{1}{3!} \varepsilon_{ijk} \varepsilon^{lmn} x^i_l x^j_m x^k_n = 1.$$

Hence, if $\mathbf{X}_e = a^i_j (\partial / \partial x^i_j)_e$ is a tangent vector to SO(3) at the identity, from the first of these equations it follows that $0 = \mathbf{X}_e[x^k_i x^l_j \delta_{kl}] = \delta^k_i a^l_j \delta_{kl} + a^k_i \delta^l_j \delta_{kl} = a^i_j + a^j_i$ (since $x^i_j(e) = \delta^i_j$), whereas from the second equation one obtains $a^i_i = 0$ (see Example 7.23). Thus, $\mathfrak{so}(3)$, the Lie algebra of SO(3), corresponds to the skew-symmetric 3×3 matrices. A basis for the skew-symmetric 3×3 matrices is formed by the matrices

$$S_1 \equiv \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$
$$S_2 \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tag{7.62}$$
$$S_3 \equiv \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

which satisfy the commutation relations

$$[S_i, S_j] = \sum_{k=1}^{3} \varepsilon_{ijk} S_k. \tag{7.63}$$

The definition of the matrices S_i is summarized by the formula $(S_i)^j_k = -\varepsilon_{ijk}$.

Hence, the value of an arbitrary element $\mathbf{X} \in \mathfrak{so}(3)$ at the identity can be expressed in the form $\mathbf{X}_e = (b^k S_k)^i_j (\partial / \partial x^i_j)_e$, where b^1, b^2, b^3 are some real numbers. Then, from (7.60) and (7.61) one obtains

$$\mathbf{X}_{\mathbf{a}}^{+} = \Phi_{\mathbf{a}*e}\mathbf{X}_e = \left(b^k S_k\right)^i_j \Phi_{\mathbf{a}*e}\left(\frac{\partial}{\partial x^i_j}\right)_e = b^k (S_k)^i_j \left(\frac{\partial}{\partial x^i_j}\right)_e [x_m \circ \Phi_{\mathbf{a}}]\left(\frac{\partial}{\partial x_m}\right)_{\mathbf{a}}$$

$$= b^k (S_k)^i_j \left(\frac{\partial}{\partial x^i_j}\right)_e \left[x_l(\mathbf{a}) x^l_m\right]\left(\frac{\partial}{\partial x_m}\right)_{\mathbf{a}} = -\sum_{l=1}^{3} b^k \varepsilon_{klm} x_l(\mathbf{a})\left(\frac{\partial}{\partial x_m}\right)_{\mathbf{a}}.$$

Hence, $\mathbf{X}^{+} = b^k \mathbf{S}_k^{+}$, with

$$\mathbf{S}_k^{+} \equiv -\sum_{l=1}^{3} \varepsilon_{klm} x_l \frac{\partial}{\partial x_m}, \tag{7.64}$$

and one can readily verify that $[\mathbf{S}_i^{+}, \mathbf{S}_j^{+}] = \sum_{k=1}^{3} \varepsilon_{ijk} \mathbf{S}_k^{+}$ [cf. (7.63)]. As shown in Example 6.11, the vector fields (7.64) are Killing vector fields for the standard metric of $\mathbb{R}^3$ [see (6.17)]. This is related to the fact that the rotations are isometries of $\mathbb{R}^3$.

Example 7.59 An arbitrary point $(a, b, c) \in \mathbb{R}^3$ can be identified with the matrix $\left(\begin{smallmatrix} a & b+c \\ b-c & -a \end{smallmatrix}\right)$. Making use of this one-to-one correspondence between the points of $\mathbb{R}^3$ and the real 2×2 matrices of trace zero, one can define an action on the right of the group $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}^3$ in the following way. For $g \in \mathrm{SL}(2, \mathbb{R})$ and $(a, b, c) \in \mathbb{R}^3$, $(a, b, c)g$ is the point of $\mathbb{R}^3$ corresponding to the matrix

$$g^{-1}\begin{pmatrix} a & b+c \\ b-c & -a \end{pmatrix} g. \tag{7.65}$$

It can readily be seen that (7.65) defines an action on the right on $\mathbb{R}^3$ which is not effective, because if g is the negative of the 2×2 identity matrix, one obtains the identity transformation. Nor is it free, because $(0, 0, 0)g = (0, 0, 0)$ for all $g \in G$. In fact, this action is a linear representation of $\mathrm{SL}(2, \mathbb{R})$ on $\mathbb{R}^3$ and any linear representation is not a free action.

From the results of Example 7.41 we have $\exp t\mathbf{X}_1 = \left(\begin{smallmatrix} e^t & 0 \\ 0 & e^{-t} \end{smallmatrix}\right)$ [see (7.51)] and substituting $g = \exp t\mathbf{X}_1$ into (7.65) one obtains the matrix

$$\begin{pmatrix} e^{-t} & 0 \\ 0 & e^t \end{pmatrix}\begin{pmatrix} a & b+c \\ b-c & -a \end{pmatrix}\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix} = \begin{pmatrix} a & e^{-2t}(b+c) \\ e^{2t}(b-c) & -a \end{pmatrix}$$

which corresponds to the point $(a, b\cosh 2t - c\sinh 2t, -b\sinh 2t + c\cosh 2t) \in \mathbb{R}^3$ (and is, therefore, $(a, b, c)\exp t\mathbf{X}_1$). Now, in terms of the natural coordinates of $\mathbb{R}^3$, the tangent vector to the curve $t \mapsto (a, b\cosh 2t - c\sinh 2t, -b\sinh 2t + c\cosh 2t)$ at $t = 0$ is

$$(\mathbf{X}_1^{+})_{(a,b,c)} = -2c\left(\frac{\partial}{\partial y}\right)_{(a,b,c)} - 2b\left(\frac{\partial}{\partial z}\right)_{(a,b,c)};$$

hence, $\mathbf{X}_1^{+} = -2z(\partial/\partial y) - 2y(\partial/\partial z)$. As pointed out above, the procedure followed in this example also gives us the integral curves of the vector fields $\mathbf{X}_i^{+}$ or, equivalently, the one-parameter groups of transformations generated by these vector fields.

In a similar way, using $\exp t\mathbf{X}_2 = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ and $\exp t\mathbf{X}_3 = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$ [see (7.51)], one finds that $\mathbf{X}_2^+ = (z-y)(\partial/\partial x) + x(\partial/\partial y) + x(\partial/\partial z)$ and $\mathbf{X}_3^+ = (y+z)(\partial/\partial x) - x(\partial/\partial y) + x(\partial/\partial z)$. It can readily be verified that the structure constants for $\{\mathbf{X}_1^+, \mathbf{X}_2^+, \mathbf{X}_3^+\}$ are equal to those of the basis of $\mathrm{SL}(2,\mathbb{R})$ given by (7.18) and (7.19).

The action on $\mathbb{R}^3$ defined by (7.65) is not transitive; in fact, each surface $x^2 + y^2 - z^2 = \text{const}$, is invariant under this action, which follows from the fact that $\mathbf{X}_i^+[x^2 + y^2 - z^2] = 0$, for $i = 1, 2, 3$, or by noting that $x^2 + y^2 - z^2 = -\det\begin{pmatrix} x & y+z \\ y-z & -x \end{pmatrix}$, and that the determinant is invariant under any similarity transformation such as (7.65). (Cf. Example 6.17; note that the vector fields $\mathbf{X}_i^+$ obtained here are related with the vector fields (6.36) by means of $\mathbf{X}_1^+ = -2\mathbf{I}^{23}$, $\mathbf{X}_2^+ = \mathbf{I}^{13} - \mathbf{I}^{12}$, and $\mathbf{X}_3^+ = \mathbf{I}^{12} + \mathbf{I}^{13}$.)

As pointed out at the beginning of this section, the isometries generated by the Killing vector fields of a Riemannian manifold can be associated with a Lie group of transformations on this manifold.

Example 7.60 As claimed in Exercise 6.13, the group $\mathrm{SL}(2,\mathbb{R})$ acts isometrically on the Poincaré half-plane. Taking into account that the inverse of the matrix $g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$ is $\begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}$, we define an action of $\mathrm{SL}(2,\mathbb{R})$ on the right on M, the Poincaré half-plane, by

$$(a+\mathrm{i}b)g = \frac{\delta(a+\mathrm{i}b) - \beta}{-\gamma(a+\mathrm{i}b) + \alpha}$$

[cf. (6.29)], identifying a point $(a, b) \in M$ with the complex number $a + \mathrm{i}b$.

Since

$$\begin{aligned}(a+\mathrm{i}b)g &= \frac{\delta(a+\mathrm{i}b) - \beta}{-\gamma(a+\mathrm{i}b) + \alpha} = \frac{\delta(a+\mathrm{i}b) - \beta}{-\gamma(a+\mathrm{i}b) + \alpha} \cdot \frac{-\gamma(a-\mathrm{i}b) + \alpha}{-\gamma(a-\mathrm{i}b) + \alpha} \\ &= \frac{-\gamma\delta(a^2+b^2) + (\alpha\delta + \beta\gamma)a - \alpha\beta + \mathrm{i}(\alpha\delta - \beta\gamma)b}{\gamma^2(a^2+b^2) - 2\alpha\gamma a + \alpha^2},\end{aligned}$$

expressing the transformation in terms of pairs of real numbers instead of complex variables, with the aid of the condition $\alpha\delta - \beta\gamma = 1$, one finds that the mapping $\Phi_{(a,b)} : G \to M$ is given by

$$\Phi_{(a,b)}(g) = \frac{1}{\gamma^2(a^2+b^2) - 2\alpha\gamma a + \alpha^2}\big(-\gamma\delta\big(a^2+b^2\big) + (1+2\beta\gamma)a - \alpha\beta, b\big).$$

Hence, using the natural coordinates (x, y) on the Poincaré half-plane (as in Example 6.12) and the local coordinates (x^1, x^2, x^3) on $\mathrm{SL}(2,\mathbb{R})$, defined by $x^1(g) = \alpha$, $x^2(g) = \beta$, $x^3(g) = \gamma$, and $\delta = [1 + x^2(g)x^3(g)]/x^1(g)$ (as in Example 7.5), one finds that

$$\begin{aligned}(\mathbf{X}_1^+)_{(a,b)} &\equiv \Phi_{(a,b)*e}\left(\frac{\partial}{\partial x^1}\right)_e \\ &= \left(\frac{\partial}{\partial x^1}\right)_e [x\circ\Phi_{(a,b)}]\left(\frac{\partial}{\partial x}\right)_{(a,b)} + \left(\frac{\partial}{\partial x^1}\right)_e [y\circ\Phi_{(a,b)}]\left(\frac{\partial}{\partial y}\right)_{(a,b)} \\ &= -2\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)_{(a,b)}\end{aligned}$$

and, similarly,

$$(\mathbf{X}_2^+)_{(a,b)} \equiv \Phi_{(a,b)*e}\left(\frac{\partial}{\partial x^2}\right)_e = -\left(\frac{\partial}{\partial x}\right)_{(a,b)},$$

$$(\mathbf{X}_3^+)_{(a,b)} \equiv \Phi_{(a,b)*e}\left(\frac{\partial}{\partial x^3}\right)_e = \left(\left(x^2-y^2\right)\frac{\partial}{\partial x} + 2xy\frac{\partial}{\partial y}\right)_{(a,b)}.$$

Comparing these expressions with (6.22) we find that the vector fields $\mathbf{X}_i^+$ coincide with the Killing vector fields of the Poincaré half-plane obtained in Example 6.12 [cf. (6.22)]. The relations (6.23) show that the mapping $\mathbf{X}_i \mapsto \mathbf{X}_i^+$ is a Lie algebra homomorphism.

Theorem 7.61 *Let G be a Lie group that acts on the right on a manifold M and let $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$; then* $[\mathbf{X}^+, \mathbf{Y}^+] = [\mathbf{X}, \mathbf{Y}]^+$.

Proof Since $\mathbf{X}^+$ is, by definition, the infinitesimal generator of the one-parameter group of transformations on M denoted by R_{g_t}, we have

$$[\mathbf{X}^+, \mathbf{Y}^+] = \pounds_{\mathbf{X}^+}\mathbf{Y}^+ = \lim_{t\to 0}\frac{R_{g_t}{}^*\mathbf{Y}^+ - \mathbf{Y}^+}{t}.$$

The value of $R_{g_t}{}^*\mathbf{Y}^+$ at a point $x \in M$ is given by [see (2.24), (7.60), and (1.25)]

$$(R_{g_t}{}^*\mathbf{Y}^+)_x = R_{{g_t}^{-1}*xg_t}\mathbf{Y}^+_{xg_t} = R_{{g_t}^{-1}*xg_t}(\Phi_{xg_t*e}\mathbf{Y}_e) = (R_{{g_t}^{-1}}\circ\Phi_{xg_t})_{*e}\mathbf{Y}_e.$$

On the other hand, for $g' \in G$,

$$(R_{{g_t}^{-1}}\circ\Phi_{xg_t})(g') = R_{{g_t}^{-1}}(xg_tg') = xg_tg'g_t^{-1} = (\Phi_x\circ R_{{g_t}^{-1}}\circ L_{g_t})(g'),$$

hence

$$(R_{g_t}{}^*\mathbf{Y}^+)_x = (\Phi_x\circ R_{{g_t}^{-1}}\circ L_{g_t})_{*e}\mathbf{Y}_e = \Phi_{x*e}(R_{{g_t}^{-1}*g_t}\mathbf{Y}_{g_t}),$$

since $\mathbf{Y}$ is left-invariant [see (7.14)]. Thus

$$\begin{aligned}[\mathbf{X}^+, \mathbf{Y}^+]_x &= \lim_{t\to 0}\frac{(R_{g_t}{}^*\mathbf{Y}^+)_x - \mathbf{Y}^+_x}{t} \\ &= \lim_{t\to 0}\frac{\Phi_{x*e}(R_{{g_t}^{-1}*g_t}\mathbf{Y}_{g_t}) - \Phi_{x*e}\mathbf{Y}_e}{t}\end{aligned}$$

$$= \Phi_{x*e}\left[\lim_{t\to 0}\frac{(R_{g_t}{}^*\mathbf{Y})_e - \mathbf{Y}_e}{t}\right] = \Phi_{x*e}(\pounds_{\mathbf{X}}\mathbf{Y})_e$$

$$= \Phi_{x*e}\big([\mathbf{X},\mathbf{Y}]\big)_e = [\mathbf{X},\mathbf{Y}]^+_x$$

[cf. (7.57) and (7.59)], that is,

$$[\mathbf{X}^+, \mathbf{Y}^+] = [\mathbf{X},\mathbf{Y}]^+.$$

From the foregoing results one concludes that the map $\mathbf{X} \mapsto \mathbf{X}^+$, from $\mathfrak{g}$ into $\mathfrak{X}(M)$, is a Lie algebra homomorphism. □

It can readily be verified in a similar way that if G is a Lie group that acts on the left on a manifold M, then a proposition analogous to Theorem 7.61 holds: if $\dot{\mathbf{X}}, \dot{\mathbf{Y}} \in \dot{\mathfrak{g}}$ then $[\dot{\mathbf{X}}^+, \dot{\mathbf{Y}}^+] = [\dot{\mathbf{X}}, \dot{\mathbf{Y}}]^+$, with $\dot{\mathbf{X}}^+_x = \Phi_{x*e}\dot{\mathbf{X}}_e$ [cf. (7.60)], where now Φ_x is the map from G into M given by $\Phi_x(g) = gx$. The vector field $\dot{\mathbf{X}}^+$ is the infinitesimal generator of the one-parameter group of transformations L_{g_t}, from M onto M, defined by $L_{g_t}(x) = g_t x$, where $\{g_t\}$ is the one-parameter subgroup generated by $\dot{\mathbf{X}}$.

Exercise 7.62 Show that the group, G, of affine motions of $\mathbb{R}$ (see Example 7.4) acts on the right on $\mathbb{R}^3$ by means of

$$(a,b,c)g \equiv \left(\frac{a + x^2(g)}{x^1(g)}, \left[x^1(g)\right]^2 b, \left[x^1(g)\right]^3 c\right),$$

for $g \in G$, $(a,b,c) \in \mathbb{R}^3$, where (x^1, x^2) are the coordinates on G defined in Example 7.4, and show that the vector fields on $\mathbb{R}^3$ induced by this action are linear combinations of

$$-x\frac{\partial}{\partial x} + 2y\frac{\partial}{\partial y} + 3z\frac{\partial}{\partial z} \quad \text{and} \quad \frac{\partial}{\partial x},$$

where (x,y,z) are the natural coordinates of $\mathbb{R}^3$. (The 1-forms $\alpha^1 \equiv \mathrm{d}y - z\,\mathrm{d}x$ and $\alpha^2 \equiv \mathrm{d}z - (2y^{-1}z^2 + y^2)\,\mathrm{d}x$ are invariant under the action of this group, in the sense defined in Sect. 4.3.)

Further examples are given in Examples 8.29–8.32.

The Adjoint Representation Any Lie group acts on its Lie algebra (on the left) by means of linear transformations in the following way. For $g \in G$, the map from G onto G, $L_{g^{-1}} \circ R_g = R_g \circ L_{g^{-1}}$ is a diffeomorphism and for $\mathbf{X} \in \mathfrak{g}$, the vector field $(L_{g^{-1}} \circ R_g)^*\mathbf{X}$, denoted by $\mathrm{Ad}\,g(\mathbf{X})$, also belongs to $\mathfrak{g}$. Indeed,

$$\mathrm{Ad}\,g(\mathbf{X}) = (L_{g^{-1}} \circ R_g)^*\mathbf{X} = R_g{}^*(L_{g^{-1}}{}^*\mathbf{X}) = R_g{}^*\mathbf{X};$$

therefore if g' is an arbitrary element of G, using the fact that R_g commutes with $L_{g'}$, we have

$$\begin{aligned} L_{g'}{}^{*}\operatorname{Ad} g(\mathbf{X}) &= L_{g'}{}^{*}(R_g{}^{*}\mathbf{X}) = R_g{}^{*}(L_{g'}{}^{*}\mathbf{X}) \\ &= R_g{}^{*}\mathbf{X} = \operatorname{Ad} g(\mathbf{X}). \end{aligned}$$

Since

$$\begin{aligned} \operatorname{Ad} g(a\mathbf{X}_1 + b\mathbf{X}_2) &= R_g{}^{*}(a\mathbf{X}_1 + b\mathbf{X}_2) = aR_g{}^{*}\mathbf{X}_1 + bR_g{}^{*}\mathbf{X}_2 \\ &= a \operatorname{Ad} g(\mathbf{X}_1) + b \operatorname{Ad} g(\mathbf{X}_2), \end{aligned}$$

for all $\mathbf{X}_1, \mathbf{X}_2 \in \mathfrak{g}$ and $a, b \in \mathbb{R}$, $\operatorname{Ad} g$ is a linear map of $\mathfrak{g}$ into itself.

Considering now two arbitrary elements $g_1, g_2 \in G$, and using $R_{g_1 g_2} = R_{g_2} \circ R_{g_1}$, we obtain

$$\begin{aligned} \operatorname{Ad}(g_1 g_2)(\mathbf{X}) &= R_{g_1 g_2}{}^{*}\mathbf{X} = R_{g_1}{}^{*}(R_{g_2}{}^{*}\mathbf{X}) \\ &= \operatorname{Ad} g_1\big(\operatorname{Ad} g_2(\mathbf{X})\big), \quad \text{for } \mathbf{X} \in \mathfrak{g}. \end{aligned}$$

That is, $\operatorname{Ad}(g_1 g_2) = (\operatorname{Ad} g_1) \circ (\operatorname{Ad} g_2)$, for $g_1, g_2 \in G$, which means that the map $g \mapsto \operatorname{Ad} g$ from G in the set of linear transformations of $\mathfrak{g}$ into itself is a *linear representation* of G called the *adjoint representation* of G.

Exercise 7.63 Show that $[\operatorname{Ad} g(\mathbf{X}), \operatorname{Ad} g(\mathbf{Y})] = \operatorname{Ad} g([\mathbf{X}, \mathbf{Y}])$, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$ and $g \in G$; that is, $\operatorname{Ad} g$ is a Lie algebra homomorphism from $\mathfrak{g}$ into itself.

Recalling that the Lie algebra of G can be identified with $T_e G$ (identifying each $\mathbf{X} \in \mathfrak{g}$ with $\mathbf{X}_e \in T_e G$) we can find the effect of $\operatorname{Ad} g$ by expressing $[\operatorname{Ad} g(\mathbf{X})]_e$ in terms of $\mathbf{X}_e$. Making use of the definition of the pullback of a vector field we have

$$\big[\operatorname{Ad} g(\mathbf{X})\big]_e = \big[(R_g \circ L_{g^{-1}})^{*}\mathbf{X}\big]_e = (R_{g^{-1}} \circ L_g)_{*e}\mathbf{X}_e.$$

Hence, $\operatorname{Ad} g$ is represented by the Jacobian $(R_{g^{-1}} \circ L_g)_{*e}$.

Example 7.64 In the case of the group of affine motions of $\mathbb{R}$, with the coordinates employed in Example 7.4, we have

$$\begin{aligned} \big[x^1 \circ (R_{g^{-1}} \circ L_g)\big](g') &= x^1\big(gg'g^{-1}\big) = x^1(gg')x^1\big(g^{-1}\big) = x^1(g'), \\ \big[x^2 \circ (R_{g^{-1}} \circ L_g)\big](g') &= x^2\big(gg'g^{-1}\big) = x^1(gg')x^2\big(g^{-1}\big) + x^2(gg') \\ &= -x^2(g)x^1(g') + x^1(g)x^2(g') + x^2(g), \end{aligned}$$

i.e., $x^1 \circ (R_{g^{-1}} \circ L_g) = x^1$ and $x^2 \circ (R_{g^{-1}} \circ L_g) = -x^2(g)x^1 + x^1(g)x^2 + x^2(g)$. Thus,

$$\begin{aligned} (R_{g^{-1}} \circ L_g)_{*e}\left(\frac{\partial}{\partial x^1}\right)_e &= \left(\frac{\partial}{\partial x^1}\right)_e \big[x^i \circ (R_{g^{-1}} \circ L_g)\big]\left(\frac{\partial}{\partial x^i}\right)_e \\ &= \left(\frac{\partial}{\partial x^1}\right)_e - x^2(g)\left(\frac{\partial}{\partial x^2}\right)_e \end{aligned}$$

and, similarly,

$$(R_{g^{-1}} \circ L_g)_{*e}\left(\frac{\partial}{\partial x^2}\right)_e = x^1(g)\left(\frac{\partial}{\partial x^2}\right)_e.$$

Hence, with respect to the basis $\{(\partial/\partial x^i)_e\}$ of T_eG, $\operatorname{Ad} g$ is represented by the 2×2 matrix

$$\begin{pmatrix} 1 & 0 \\ -x^2(g) & x^1(g) \end{pmatrix}$$

and one readily verifies that $\operatorname{Ad}(g_1g_2) = (\operatorname{Ad} g_1)\circ(\operatorname{Ad} g_2)$ amounts to the multiplication table of the group (7.2). It may be noticed that, in the present case, the correspondence $g \mapsto \operatorname{Ad}(g)$ is one-to-one.

In the case where G is an Abelian group, we have $(R_{g^{-1}} \circ L_g)(g') = gg'g^{-1} = g'$, that is, $R_{g^{-1}} \circ L_g = \mathrm{id}$, and therefore, $\operatorname{Ad}(g)$ is the identity map for all $g \in G$.

Exercise 7.65 Let $\mathbf{X}$ be an element of the Lie algebra of $\mathrm{GL}(n,\mathbb{R})$ such that $\mathbf{X}_e = a^i_j(\partial/\partial x^i_j)_e$, where the x^i_j are the natural coordinates of $\mathrm{GL}(n,\mathbb{R})$. Show that

$$\big[\operatorname{Ad} g(\mathbf{X})\big]_e = x^k_i(g)a^i_j x^j_m\big(g^{-1}\big)\left(\frac{\partial}{\partial x^k_m}\right)_e.$$

Hence, associating the matrix $A = (a^i_j)$ with $\mathbf{X} \in \mathfrak{gl}(n,\mathbb{R})$, as in Example 7.15, the matrix associated with $\operatorname{Ad} g(\mathbf{X})$ is gAg^{-1}.

Exercise 7.66 Let G be a Lie group that acts on the right on M. Show that $[\operatorname{Ad} g(\mathbf{X})]^+ = R_g{}^*\mathbf{X}^+$, for $\mathbf{X} \in \mathfrak{g},\ g \in G$.

Exercise 7.67 Show that $g(\exp t\mathbf{X})g^{-1} = \exp[t \operatorname{Ad} g(\mathbf{X})]$, for $\mathbf{X} \in \mathfrak{g},\ g \in G$. (*Hint*: show that $\gamma(t) \equiv g(\exp t\mathbf{X})g^{-1}$ is the integral curve of $\operatorname{Ad} g(\mathbf{X})$ starting at e.)

Chapter 8
Hamiltonian Classical Mechanics

In this chapter we start by showing that any finite-dimensional differentiable manifold M possesses an associated manifold, denoted by T^*M, called the cotangent bundle of M, which has a naturally defined nondegenerate 2-form, which allows us to define a Poisson bracket between real-valued functions defined on T^*M. We then apply this structure to classical mechanics and geometrical optics, emphasizing the applications of Lie groups and Riemannian geometry. Here we will have the opportunity of making use of all of the machinery introduced in the previous chapters.

8.1 The Cotangent Bundle

Let M be a differentiable manifold of dimension n. The *cotangent bundle* of M, denoted by T^*M, is the set of all covectors at all points of M, that is, $T^*M = \bigcup_{p\in M} T_p^*M$. The *canonical projection*, π, from T^*M onto M is the mapping that sends each element of T^*M to the point of M at which it is attached; that is, if $\alpha_p \in T_p^*M$, then $\pi(\alpha_p) = p$, and therefore, $\pi^{-1}(p) = T_p^*M$.

The set T^*M acquires, in a natural way, the structure of differentiable manifold induced by that of M. If (U, ϕ) is a chart on M and $p \in U$, any covector $\alpha_p \in T_p^*M$ can be expressed as a linear combination of the covectors $\{\mathrm{d}x^i_p\}_{i=1}^n$ with real coefficients, which depend on α_p, that is,

$$\alpha_p = p_i(\alpha_p)\,\mathrm{d}x^i_p, \tag{8.1}$$

with $p_i(\alpha_p) \in \mathbb{R}$ [cf. (1.27)]. Then, from (1.49),

$$p_i(\alpha_p) = \alpha_p\left(\left(\frac{\partial}{\partial x^i}\right)_p\right). \tag{8.2}$$

Let now $\overline{\phi} : \pi^{-1}(U) \to \mathbb{R}^{2n}$ be given by

$$\begin{aligned}\overline{\phi}(\alpha_p) &= \big(x^1(p), \ldots, x^n(p), p_1(\alpha_p), \ldots, p_n(\alpha_p)\big)\\ &= \big(x^1\big(\pi(\alpha_p)\big), \ldots, x^n\big(\pi(\alpha_p)\big), p_1(\alpha_p), \ldots, p_n(\alpha_p)\big),\end{aligned} \tag{8.3}$$

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2_8,

for $\alpha_p \in \pi^{-1}(U)$; it can readily be seen that $(\pi^{-1}(U), \overline{\phi})$ is a chart on T^*M. If $\{(U_i, \phi_i)\}$ is a C^∞ subatlas on M, then $\{(\pi^{-1}(U_i), \overline{\phi_i})\}$ is a C^∞ subatlas on T^*M that defines a structure of differentiable manifold for T^*M, in such a way that the projection π is differentiable.

Exercise 8.1 A covector field α on M can be regarded as a mapping from M into T^*M, $p \mapsto \alpha_p$, such that $\pi \circ \alpha = \mathrm{id}_M$. Show that the map $p \mapsto \alpha_p$ is differentiable if and only if α is differentiable (in the sense defined in Sect. 1.4).

The Fundamental 1-Form Let $\alpha_p \in T_p^*M$. Since π is a differentiable map from T^*M on M, which sends α_p into p, the Jacobian $\pi_{*\alpha_p}$ is a linear transformation of $T_{\alpha_p}(T^*M)$ into T_pM; hence, the composition $\alpha_p \circ \pi_{*\alpha_p}$ is a linear transformation from $T_{\alpha_p}(T^*M)$ in $\mathbb{R}$; that is, $\alpha_p \circ \pi_{*\alpha_p} \in T^*_{\alpha_p}(T^*M)$. Thus, the mapping θ defined by

$$\theta_{\alpha_p} \equiv \alpha_p \circ \pi_{*\alpha_p} \tag{8.4}$$

is a covector field on T^*M.

If (U, ϕ) is a chart on M, defining $q^i \equiv x^i \circ \pi = \pi^* x^i$, from (8.3) we obtain

$$\overline{\phi}(\alpha_p) = \big(q^1(\alpha_p), \ldots, q^n(\alpha_p), p_1(\alpha_p), \ldots, p_n(\alpha_p)\big). \tag{8.5}$$

Hence, $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ is a coordinate system on T^*M; the tangent vectors $(\partial/\partial q^i)_{\alpha_p}, (\partial/\partial p_i)_{\alpha_p}$, $i = 1, 2, \ldots, n$, form a basis of $T_{\alpha_p}(T^*M)$ and the covector field θ is given locally by [see (1.50)]

$$\theta = \theta\left(\frac{\partial}{\partial q^i}\right) \mathrm{d}q^i + \theta\left(\frac{\partial}{\partial p_i}\right) \mathrm{d}p_i.$$

Using the definition of θ we see that the real-valued functions appearing in the last equation are given by

$$\left[\theta\left(\frac{\partial}{\partial q^i}\right)\right](\alpha_p) = \theta_{\alpha_p}\left(\left(\frac{\partial}{\partial q^i}\right)_{\alpha_p}\right) = (\alpha_p \circ \pi_{*\alpha_p})\left(\left(\frac{\partial}{\partial q^i}\right)_{\alpha_p}\right),$$

and using the expression for the Jacobian (1.24) and (8.2), it follows that

$$\begin{aligned}\left[\theta\left(\frac{\partial}{\partial q^i}\right)\right](\alpha_p) &= \alpha_p\left(\left(\frac{\partial}{\partial q^i}\right)_{\alpha_p}[x^j \circ \pi]\left(\frac{\partial}{\partial x^j}\right)_p\right) \\ &= \alpha_p\left(\left(\frac{\partial}{\partial q^i}\right)_{\alpha_p}[q^j]\left(\frac{\partial}{\partial x^j}\right)_p\right) \\ &= \alpha_p\left(\left(\frac{\partial}{\partial x^i}\right)_p\right) = p_i(\alpha_p),\end{aligned}$$

that is, $\theta((\partial/\partial q^i)) = p_i$.

Similarly,

$$\left[\theta\left(\frac{\partial}{\partial p_i}\right)\right](\alpha_p) = \alpha_p\left(\left(\frac{\partial}{\partial p_i}\right)_{\alpha_p}\left[q^j\right]\left(\frac{\partial}{\partial x^j}\right)_p\right) = \alpha_p(0) = 0.$$

Therefore

$$\theta = p_i\,\mathrm{d}q^i. \tag{8.6}$$

This expression shows that θ is differentiable; that is, θ is a 1-form on T^*M, called the *fundamental* 1-*form* of T^*M.

Exercise 8.2 Show that, for $f \in C^\infty(M)$,

$$\pi^*\left(\frac{\partial f}{\partial x^i}\right) = \frac{\partial(\pi^* f)}{\partial q^i}.$$

If M_1 and M_2 are two differentiable manifolds and $\psi : M_1 \to M_2$ is a diffeomorphism, we define $\overline{\psi} : T^*M_1 \to T^*M_2$ by

$$\overline{\psi}(\alpha_p) \equiv \alpha_p \circ \left(\psi^{-1}\right)_{*\psi(p)} \quad \text{for } \alpha_p \in T_p^*(M_1). \tag{8.7}$$

Denoting by π_2 the projection from T^*M_2 on M_2 and similarly for π_1, since $(\psi^{-1})_{*\psi(p)}$ maps $T_{\psi(p)}M_2$ onto T_pM_1, we have

$$\pi_2\left(\overline{\psi}(\alpha_p)\right) = \psi(p) = \psi\left(\pi_1(\alpha_p)\right), \quad \text{for } \alpha_p \in T_p^*(M_1),$$

that is,

$$\pi_2 \circ \overline{\psi} = \psi \circ \pi_1. \tag{8.8}$$

Exercise 8.3 Show that if $\psi_1 : M_1 \to M_2$ and $\psi_2 : M_2 \to M_3$ are two diffeomorphisms, then $\overline{(\psi_2 \circ \psi_1)} = \overline{\psi}_2 \circ \overline{\psi}_1$.

Theorem 8.4 *Let $\psi : M_1 \to M_2$ be a diffeomorphism and let θ_1 and θ_2 be the fundamental* 1*-forms of T^*M_1 and T^*M_2, respectively*; *then $\theta_1 = \overline{\psi}^*\theta_2$.*

Proof Taking $v \in T_{\alpha_p}(T^*M_1)$ and applying the chain rule to (8.8), we have

$$\pi_{2*\overline{\psi}(\alpha_p)}(\overline{\psi}_{*\alpha_p}v) = \psi_{*\pi_1(\alpha_p)}(\pi_{1*\alpha_p}v) = \psi_{*p}(\pi_{1*\alpha_p}v), \tag{8.9}$$

hence, using the definition (2.28) of the pullback, (8.4), (8.9), (8.7), and the chain rule, (1.25),

$$\begin{aligned}(\overline{\psi}^*\theta_2)_{\alpha_p}(v) &= \theta_{2\,\overline{\psi}(\alpha_p)}(\overline{\psi}_{*\alpha_p}v)\\ &= \overline{\psi}(\alpha_p)\left[\pi_{2*\overline{\psi}(\alpha_p)}(\overline{\psi}_{*\alpha_p}v)\right]\\ &= \overline{\psi}(\alpha_p)\left[\psi_{*p}(\pi_{1*\alpha_p}v)\right]\end{aligned}$$

$$= \big(\alpha_p \circ (\psi^{-1})_{*\psi(p)}\big)\big[\psi_{*p}(\pi_{1*\alpha_p} v)\big]$$
$$= \big(\alpha_p \circ (\psi^{-1} \circ \psi)_{*p}\big)(\pi_{1*\alpha_p} v)$$
$$= \alpha_p(\pi_{1*\alpha_p} v) = \theta_{1\alpha_p}(v),$$

that is, $\theta_1 = \overline{\psi}^* \theta_2$. □

The foregoing theorem can also be proved making use of the local expression (8.6). Denoting by q^i, p_i and q'^i, p'_i the coordinates induced on T^*M_1 and T^*M_2 by systems of coordinates x^i and x'^i on M_1 and M_2, respectively, making use of the definition of the pullback of a function, (1.8), together with (8.2), (8.7), and (1.24), we have for $\alpha_p \in T_p^*M_1$

$$(\overline{\psi}^* p'_i)(\alpha_p) = p'_i\big(\overline{\psi}(\alpha_p)\big) = \overline{\psi}(\alpha_p)\left(\left(\frac{\partial}{\partial x'^i}\right)_{\psi(p)}\right)$$
$$= \big(\alpha_p \circ (\psi^{-1})_{*\psi(p)}\big)\left(\left(\frac{\partial}{\partial x'^i}\right)_{\psi(p)}\right)$$
$$= \alpha_p\left(\left(\frac{\partial}{\partial x'^i}\right)_{\psi(p)}\big[x^k \circ \psi^{-1}\big]\left(\frac{\partial}{\partial x^k}\right)_p\right)$$
$$= \left.\frac{\partial(x^k \circ \psi^{-1})}{\partial x'^i}\right|_{\psi(p)} p_k(\alpha_p),$$

that is,

$$\overline{\psi}^* p'_i = \pi_1{}^*\left(\psi^* \frac{\partial(x^k \circ \psi^{-1})}{\partial x'^i}\right) p_k, \tag{8.10}$$

hence, using (8.6), (8.8), and (8.10),

$$\overline{\psi}^*\theta_2 = \overline{\psi}^*\big(p'_i \,\mathrm{d}q'^i\big) = (\overline{\psi}^* p'_i)\,\mathrm{d}\big(\overline{\psi}^* \pi_2{}^* x'^i\big) = (\overline{\psi}^* p'_i)\,\mathrm{d}\big(\pi_1{}^* \psi^* x'^i\big)$$
$$= (\overline{\psi}^* p'_i)\pi_1{}^*\left(\frac{\partial(x'^i \circ \psi)}{\partial x^l}\,\mathrm{d}x^l\right)$$
$$= \pi_1{}^*\left(\psi^* \frac{\partial(x^k \circ \psi^{-1})}{\partial x'^i}\,\frac{\partial(x'^i \circ \psi)}{\partial x^l}\right) p_k \,\mathrm{d}q^l = \big(\pi_1{}^*\delta^k_l\big) p_k \,\mathrm{d}q^l = \theta_1.$$

If φ_t is a flow on M, then according to Exercise 8.3, $\overline{\varphi_t} \circ \overline{\varphi_s} = \overline{\varphi_t \circ \varphi_s} = \overline{\varphi_{t+s}}$; therefore, $\overline{\varphi_t}$ is a flow on T^*M with $\pi \circ \overline{\varphi_t} = \varphi_t \circ \pi$. From this relation it follows that if $\mathbf{X}$ and $\overline{\mathbf{X}}$ are the infinitesimal generators of φ_t and $\overline{\varphi_t}$, respectively, then

$$\pi_{*\alpha_p}\overline{\mathbf{X}}_{\alpha_p} = \mathbf{X}_p, \quad \alpha_p \in T_p^*M, \tag{8.11}$$

that is, $\overline{\mathbf{X}}$ and $\mathbf{X}$ are π-related (see Exercise 2.8).

8.2 Hamiltonian Vector Fields and the Poisson Bracket

The exterior derivative of the fundamental 1-form of T^*M is called the *fundamental 2-form* of T^*M; from (8.6) we obtain the local expression

$$\mathrm{d}\theta = \mathrm{d}p_i \wedge \mathrm{d}q^i, \tag{8.12}$$

in terms of the coordinates q^i, p_i induced by a chart of coordinates on M. The fundamental 2-form of T^*M induces an identification between differentiable vector fields and 1-forms on T^*M, associating to each vector field $\mathbf{X} \in \mathfrak{X}(T^*M)$ the 1-form $\mathbf{X} \lrcorner\, \mathrm{d}\theta$.

If the vector field $\mathbf{X}$ is locally given by $\mathbf{X} = A^i(\partial/\partial q^i) + B_i(\partial/\partial p_i)$, then

$$\begin{aligned}\mathbf{X} \lrcorner\, \mathrm{d}\theta &= \mathbf{X} \lrcorner \left(\mathrm{d}p_i \wedge \mathrm{d}q^i\right) = (\mathbf{X} \lrcorner\, \mathrm{d}p_i)\, \mathrm{d}q^i - \left(\mathbf{X} \lrcorner\, \mathrm{d}q^i\right) \mathrm{d}p_i \\ &= B_i\, \mathrm{d}q^i - A^i\, \mathrm{d}p_i. \end{aligned} \tag{8.13}$$

From this expression one concludes that the map from $\mathfrak{X}(T^*M)$ into $\Lambda^1(T^*M)$, given by $\mathbf{X} \mapsto \mathbf{X} \lrcorner\, \mathrm{d}\theta$, is $C^\infty(T^*M)$-linear, one-to-one, and onto.

If $\mathbf{X}$ is a vector field on T^*M, we say that $\mathbf{X}$ is *Hamiltonian* if the 1-form $\mathbf{X} \lrcorner\, \mathrm{d}\theta$ is exact; that is, $\mathbf{X}$ is a Hamiltonian vector field if there exists some real-valued function $f \in C^\infty(T^*M)$ such that

$$\mathbf{X} \lrcorner\, \mathrm{d}\theta = -\mathrm{d}f \tag{8.14}$$

(the minus sign is introduced for convenience); $\mathbf{X}$ is *locally Hamiltonian* if $\mathbf{X} \lrcorner\, \mathrm{d}\theta$ is closed. Since every exact differential form is closed, all Hamiltonian vector fields are locally Hamiltonian. In order to emphasize the difference between the Hamiltonian vector fields and the locally Hamiltonian ones, the former are also called *globally Hamiltonian*.

Lemma 8.5 *Let* $\mathbf{X}$ *be a vector field on* T^*M. $\mathbf{X}$ *is locally Hamiltonian if and only if* $\pounds_{\mathbf{X}}\, \mathrm{d}\theta = 0$.

Proof The conclusion follows from the identity (3.39) and the fact that $\mathrm{d}^2 = 0$

$$\pounds_{\mathbf{X}}\, \mathrm{d}\theta = \mathbf{X} \lrcorner\, \mathrm{d}(\mathrm{d}\theta) + \mathrm{d}(\mathbf{X} \lrcorner\, \mathrm{d}\theta) = \mathrm{d}(\mathbf{X} \lrcorner\, \mathrm{d}\theta). \qquad \square$$

This result means that if φ_t is the flow generated by a vector field $\mathbf{X}$ on T^*M, then $\varphi_t^*(\mathrm{d}\theta) = \mathrm{d}\theta$ if and only if $\mathbf{X}$ is locally Hamiltonian. Any map $\psi : T^*M_1 \to T^*M_2$ such that $\psi^*(\mathrm{d}\theta_2) = \mathrm{d}\theta_1$ is referred to as a *canonical transformation* or *symplectomorphism*. Hence, $\mathbf{X} \in \mathfrak{X}(T^*M)$ is locally Hamiltonian, if and only if it is the infinitesimal generator of a local one-parameter group of canonical transformations.

According to Theorem 8.4, any diffeomorphism $\psi : M_1 \to M_2$ gives rise to a canonical transformation $\overline{\psi}$, which satisfies the stronger condition $\theta_1 = \overline{\psi}^*\theta_2$.

Exercise 8.6 Show that the set of canonical transformations of T^*M onto itself forms a group with the operation of composition.

Theorem 8.7 *The locally Hamiltonian vector fields form a Lie subalgebra of* $\mathfrak{X}(T^*M)$. *The Lie bracket of two locally Hamiltonian vector fields is globally Hamiltonian.*

Proof Let $\mathbf{X}$ and $\mathbf{Y}$ be locally Hamiltonian vector fields; using Lemma 8.5, we have

$$\pounds_{a\mathbf{X}+b\mathbf{Y}}\, \mathrm{d}\theta = a\pounds_{\mathbf{X}}\, \mathrm{d}\theta + b\pounds_{\mathbf{Y}}\, \mathrm{d}\theta = 0, \quad \text{for } a, b \in \mathbb{R},$$

therefore $a\mathbf{X} + b\mathbf{Y}$ is locally Hamiltonian. Furthermore, since $[\mathbf{X}, \mathbf{Y}] = \pounds_{\mathbf{X}}\mathbf{Y}$, using (2.27), (2.44), Lemma 8.5, and (3.39), we have

$$\begin{aligned}
[\mathbf{X}, \mathbf{Y}] \lrcorner\, \mathrm{d}\theta &= (\pounds_{\mathbf{X}}\mathbf{Y}) \lrcorner\, \mathrm{d}\theta \\
&= \pounds_{\mathbf{X}}(\mathbf{Y} \lrcorner\, \mathrm{d}\theta) - \mathbf{Y} \lrcorner\, \pounds_{\mathbf{X}}\, \mathrm{d}\theta \\
&= \pounds_{\mathbf{X}}(\mathbf{Y} \lrcorner\, \mathrm{d}\theta) \\
&= \mathbf{X} \lrcorner\, \mathrm{d}(\mathbf{Y} \lrcorner\, \mathrm{d}\theta) + \mathrm{d}\big(\mathbf{X} \lrcorner\, (\mathbf{Y} \lrcorner\, \mathrm{d}\theta)\big) \\
&= \mathrm{d}\big(\mathbf{X} \lrcorner\, (\mathbf{Y} \lrcorner\, \mathrm{d}\theta)\big).
\end{aligned} \tag{8.15}$$

□

With each differentiable function $f \in C^\infty(T^*M)$ there exists an associated Hamiltonian vector field, $\mathbf{X}_{\mathrm{d}f}$, defined by

$$\mathbf{X}_{\mathrm{d}f} \lrcorner\, \mathrm{d}\theta = -\mathrm{d}f. \tag{8.16}$$

From the local expression (1.52), $\mathrm{d}f = (\partial f/\partial q^i)\, \mathrm{d}q^i + (\partial f/\partial p_i)\, \mathrm{d}p_i$, and (8.13), it follows that

$$\mathbf{X}_{\mathrm{d}f} = \frac{\partial f}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial}{\partial p_i}. \tag{8.17}$$

The set of the globally Hamiltonian vector fields is a Lie subalgebra of the Lie algebra of locally Hamiltonian fields; in fact, if $\mathbf{X}_{\mathrm{d}f}$ and $\mathbf{X}_{\mathrm{d}g}$ are two globally Hamiltonian vector fields, any linear combination of them, $a\mathbf{X}_{\mathrm{d}f} + b\mathbf{X}_{\mathrm{d}g}$, and their Lie bracket, $[\mathbf{X}_{\mathrm{d}f}, \mathbf{X}_{\mathrm{d}g}]$, are also globally Hamiltonian since

$$\begin{aligned}
(a\mathbf{X}_{\mathrm{d}f} + b\mathbf{X}_{\mathrm{d}g}) \lrcorner\, \mathrm{d}\theta &= a\mathbf{X}_{\mathrm{d}f} \lrcorner\, \mathrm{d}\theta + b\mathbf{X}_{\mathrm{d}g} \lrcorner\, \mathrm{d}\theta \\
&= -a\, \mathrm{d}f - b\, \mathrm{d}g = -\mathrm{d}(af + bg), \quad \text{for } a, b \in \mathbb{R},
\end{aligned}$$

and from (8.15) and (8.16),

$$\begin{aligned}
[\mathbf{X}_{\mathrm{d}f}, \mathbf{X}_{\mathrm{d}g}] \lrcorner\, \mathrm{d}\theta &= \mathrm{d}\big(\mathbf{X}_{\mathrm{d}f} \lrcorner\, (\mathbf{X}_{\mathrm{d}g} \lrcorner\, \mathrm{d}\theta)\big) \\
&= -\mathrm{d}(\mathbf{X}_{\mathrm{d}f} \lrcorner\, \mathrm{d}g) \\
&= -\mathrm{d}(\mathbf{X}_{\mathrm{d}f}\, g).
\end{aligned} \tag{8.18}$$

Definition 8.8 Let $f, g \in C^\infty(T^*M)$; the *Poisson bracket* of f and g, denoted by $\{f, g\}$, is defined by

$$\{f, g\} \equiv \mathbf{X}_{\mathrm{d}f}\, g. \tag{8.19}$$

By virtue of (8.16) and the definition (8.19), the relation (8.18) is equivalent to

$$[\mathbf{X}_{\mathrm{d}f}, \mathbf{X}_{\mathrm{d}g}] = \mathbf{X}_{\mathrm{d}\{f,g\}}. \tag{8.20}$$

From (8.17) and (8.19) one finds that the Poisson bracket is locally given by

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i}. \tag{8.21}$$

Exercise 8.9 Show that $\psi : T^*M_1 \to T^*M_2$ is a canonical transformation if and only if $\psi^*\{f, g\} = \{\psi^* f, \psi^* g\}$, for $f, g \in C^\infty(T^*M_2)$.

Theorem 8.10 *The space $C^\infty(T^*M)$ is a Lie algebra over $\mathbb{R}$ with the Poisson bracket.*

Proof Let $f, g \in C^\infty(T^*M)$, from (8.19) and (8.16) it follows that the Poisson bracket of f and g is given by

$$\{f, g\} = \mathbf{X}_{\mathrm{d}f}\, g = \mathbf{X}_{\mathrm{d}f} \lrcorner\, \mathrm{d}g = -\mathbf{X}_{\mathrm{d}f} \lrcorner (\mathbf{X}_{\mathrm{d}g} \lrcorner\, \mathrm{d}\theta) = 2\,\mathrm{d}\theta(\mathbf{X}_{\mathrm{d}f}, \mathbf{X}_{\mathrm{d}g}). \tag{8.22}$$

From this expression it is clear that the Poisson bracket is skew-symmetric and bilinear. Furthermore, for $f, g, h \in C^\infty(T^*M)$, from (8.19) and (8.20) we have

$$\begin{aligned}\big\{\{f, g\}, h\big\} &= \mathbf{X}_{\mathrm{d}\{f,g\}} h = [\mathbf{X}_{\mathrm{d}f}, \mathbf{X}_{\mathrm{d}g}] h \\ &= \mathbf{X}_{\mathrm{d}f}(\mathbf{X}_{\mathrm{d}g} h) - \mathbf{X}_{\mathrm{d}g}(\mathbf{X}_{\mathrm{d}f} h) = \mathbf{X}_{\mathrm{d}f}\{g, h\} - \mathbf{X}_{\mathrm{d}g}\{f, h\} \\ &= \big\{f, \{g, h\}\big\} - \big\{g, \{f, h\}\big\} \\ &= -\big\{\{g, h\}, f\big\} - \big\{\{h, f\}, g\big\}. \end{aligned}$$

□

Exercise 8.11 Making use of (8.19), (8.20), (8.22), (3.31), and the fact that the fundamental 2-form is closed, show that the Poisson bracket satisfies the Jacobi identity.

From the foregoing results we conclude that the map $f \mapsto \mathbf{X}_{\mathrm{d}f}$, from $C^\infty(T^*M)$ into $\mathfrak{X}(T^*M)$ is a homomorphism of Lie algebras whose kernel is formed by the constant functions.

The fundamental 2-form of T^*M is analogous to a Riemannian metric in the sense that both are non-singular tensor fields of type $\binom{0}{2}$, with the only difference that a 2-form is skew-symmetric, while a Riemannian metric is symmetric. The globally Hamiltonian vector field $\mathbf{X}_{\mathrm{d}f}$ is analogous to the gradient of a function f [compare (6.7) with (8.16)], and for this reason the vector field $\mathbf{X}_{\mathrm{d}f}$ is also denoted by sgrad f. In the same way, according to Lemma 8.5, the locally Hamiltonian vector fields are the analog of the Killing vector fields. Another difference comes from the fact that, while the gradient of a function f is orthogonal to the level surfaces of f, $\mathbf{X}_{\mathrm{d}f}$ is tangent to these surfaces ($\mathbf{X}_{\mathrm{d}f} f = 0$).

The Canonical Lift of a Vector Field to the Cotangent Bundle As shown at the end of the preceding section (p. 204), any vector field $\mathbf{X} \in \mathfrak{X}(M)$ gives rise to a vector field, $\overline{\mathbf{X}}$, on T^*M, which will be called the *canonical lift* of $\mathbf{X}$ to T^*M. The vector field $\overline{\mathbf{X}}$ is globally Hamiltonian; in effect, if $\overline{\varphi_t}$ is the flow generated by $\overline{\mathbf{X}}$, according to Theorem 8.4 we have $\overline{\varphi_t}^*\theta = \theta$ and, therefore, the Lie derivative of θ with respect to $\overline{\mathbf{X}}$ is zero. On the other hand, $\pounds_{\overline{\mathbf{X}}}\theta = \overline{\mathbf{X}} \lrcorner \mathrm{d}\theta + \mathrm{d}(\overline{\mathbf{X}} \lrcorner \theta)$; hence

$$\overline{\mathbf{X}} \lrcorner \mathrm{d}\theta = -\mathrm{d}(\overline{\mathbf{X}} \lrcorner \theta), \tag{8.23}$$

which shows that, indeed, $\overline{\mathbf{X}}$ is globally Hamiltonian [cf. (8.14)].

We shall denote by $f_{\mathbf{X}}$ the function of T^*M in $\mathbb{R}$ appearing on the right-hand side of (8.23), that is,

$$f_{\mathbf{X}} \equiv \overline{\mathbf{X}} \lrcorner \theta. \tag{8.24}$$

Then from the definition of θ and (8.11) we have

$$\begin{aligned} f_{\mathbf{X}}(\alpha_p) &= (\overline{\mathbf{X}} \lrcorner \theta)(\alpha_p) = \theta_{\alpha_p}(\overline{\mathbf{X}}_{\alpha_p}) = (\alpha_p \circ \pi_{*\alpha_p})\overline{\mathbf{X}}_{\alpha_p} \\ &= \alpha_p(\pi_{*\alpha_p}\overline{\mathbf{X}}_{\alpha_p}) = \alpha_p(\mathbf{X}_p), \quad \text{for } \alpha_p \in T_p^*M. \end{aligned} \tag{8.25}$$

Hence, if $\mathbf{X} \in \mathfrak{X}(M)$ is locally given by $\mathbf{X} = X^i(\partial/\partial x^i)$, and using (8.2) we obtain

$$\begin{aligned} f_{\mathbf{X}}(\alpha_p) &= \alpha_p\left(X^i(p)\left(\frac{\partial}{\partial x^i}\right)_p\right) = p_i(\alpha_p)X^i(p) \\ &= p_i(\alpha_p)X^i\big(\pi(\alpha_p)\big) = \big[p_i\big(X^i \circ \pi\big)\big](\alpha_p), \end{aligned}$$

that is,

$$f_{\mathbf{X}} = p_i\big(X^i \circ \pi\big) = p_i\big(\pi^* X^i\big), \tag{8.26}$$

which shows that, in terms of a coordinate system $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ induced by a coordinate system on M, $f_{\mathbf{X}}$ is a homogeneous function of degree 1 in the p_i (the $\pi^* X^i$ are functions of the q^j only).

Exercise 8.12 Show that if $\mathbf{X} \in \mathfrak{X}(M)$ is locally given by $\mathbf{X} = X^i(\partial/\partial x^i)$, then

$$\overline{\mathbf{X}} = \big(\pi^* X^i\big)\frac{\partial}{\partial q^i} - p_j\, \pi^*\left(\frac{\partial X^j}{\partial x^i}\right)\frac{\partial}{\partial p_i}, \tag{8.27}$$

where $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ is the coordinate system on T^*M induced by the coordinates $(x^1, \ldots, x^n)$ on M.

Exercise 8.13 Show that a vector field $\mathbf{X}$ on T^*M satisfies $\pounds_{\mathbf{X}}\theta = 0$ only if $\mathbf{X}$ is the canonical lift of some vector field on M.

Exercise 8.14 Show that if $\mathbf{X} = (\partial/\partial x^l)$, then $f_{\mathbf{X}} = p_l$ and $\overline{\mathbf{X}} = (\partial/\partial q^l)$. (In the case where $M = \mathbb{E}^3$ and the x^i are Cartesian coordinates, $\partial/\partial x^l$ is the infinitesimal generator of translations in the x^l direction.)

Exercise 8.15 Show that if $\mathbf{X} = x^k(\partial/\partial x^l) - x^l(\partial/\partial x^k)$, then $f_{\mathbf{X}} = q^k p_l - q^l p_k$ and $\overline{\mathbf{X}} = q^k(\partial/\partial q^l) - q^l(\partial/\partial q^k) + p_k(\partial/\partial p_l) - p_l(\partial/\partial p_k)$. (In the case where $M = \mathbb{E}^3$ and the x^i are Cartesian coordinates, $x^k(\partial/\partial x^l) - x^l(\partial/\partial x^k)$ is the infinitesimal generator of rotations in the x^k–x^l plane; see Example 7.58.)

Let $\mathbf{X}$ and $\mathbf{Y}$ be vector fields on M and let $\overline{\mathbf{X}}$ and $\overline{\mathbf{Y}}$ be their canonical lifts to T^*M. Since $\overline{\mathbf{X}}$ and $\overline{\mathbf{Y}}$ are π-related to $\mathbf{X}$ and $\mathbf{Y}$, respectively [see (8.11)], the Lie bracket $[\overline{\mathbf{X}}, \overline{\mathbf{Y}}]$ is π-related with $[\mathbf{X}, \mathbf{Y}]$ (see Sect. 1.3); therefore, for $\alpha_p \in T_p^*M$, making use of (8.25), we have

$$\begin{aligned} f_{[\mathbf{X},\mathbf{Y}]}(\alpha_p) &= \alpha_p\big([\mathbf{X}, \mathbf{Y}]_p\big) = \alpha_p\big(\pi_{*\alpha_p}[\overline{\mathbf{X}}, \overline{\mathbf{Y}}]_{\alpha_p}\big) \\ &= \theta_{\alpha_p}\big([\overline{\mathbf{X}}, \overline{\mathbf{Y}}]_{\alpha_p}\big) = \big([\overline{\mathbf{X}}, \overline{\mathbf{Y}}] \lrcorner \theta\big)(\alpha_p), \end{aligned}$$

i.e.,

$$f_{[\mathbf{X},\mathbf{Y}]} = [\overline{\mathbf{X}}, \overline{\mathbf{Y}}] \lrcorner \theta. \tag{8.28}$$

An alternative expression for the function $f_{[\mathbf{X},\mathbf{Y}]}$ is obtained as follows, using the properties of the Lie derivative (2.27) and (2.44), and using $\pounds_{\overline{\mathbf{X}}}\theta = 0$. From (8.28) we have

$$\begin{aligned} f_{[\mathbf{X},\mathbf{Y}]} &= [\overline{\mathbf{X}}, \overline{\mathbf{Y}}] \lrcorner \theta = (\pounds_{\overline{\mathbf{X}}}\overline{\mathbf{Y}}) \lrcorner \theta = \pounds_{\overline{\mathbf{X}}}(\overline{\mathbf{Y}} \lrcorner \theta) - \overline{\mathbf{Y}} \lrcorner \pounds_{\overline{\mathbf{X}}}\theta \\ &= \pounds_{\overline{\mathbf{X}}}(\overline{\mathbf{Y}} \lrcorner \theta) = \pounds_{\overline{\mathbf{X}}} f_{\mathbf{Y}} = \overline{\mathbf{X}} f_{\mathbf{Y}}. \end{aligned} \tag{8.29}$$

On the other hand, comparing (8.16) and (8.23) one finds that $\overline{\mathbf{X}}$ is the Hamiltonian vector field corresponding to the function $f_{\mathbf{X}}$; hence, according to (8.19), $\overline{\mathbf{X}} f_{\mathbf{Y}} = \{f_{\mathbf{X}}, f_{\mathbf{Y}}\}$, so that (8.29) amounts to

$$f_{[\mathbf{X},\mathbf{Y}]} = \{f_{\mathbf{X}}, f_{\mathbf{Y}}\}, \tag{8.30}$$

which together with (8.25) means that the map $\mathbf{X} \mapsto f_{\mathbf{X}}$ from $\mathfrak{X}(M)$ into $C^\infty(T^*M)$ is a Lie algebra homomorphism.

Furthermore, from (8.23) and (8.24) we have

$$\overline{[\mathbf{X}, \mathbf{Y}]} \lrcorner \mathrm{d}\theta = -\mathrm{d}\big(\overline{[\mathbf{X}, \mathbf{Y}]} \lrcorner \theta\big) = -\mathrm{d} f_{[\mathbf{X},\mathbf{Y}]} \tag{8.31}$$

and, since $\overline{\mathbf{X}}$ is the Hamiltonian vector field corresponding to $f_{\mathbf{X}}$, from (8.18) and (8.29) it follows that

$$[\overline{\mathbf{X}}, \overline{\mathbf{Y}}] \lrcorner \mathrm{d}\theta = -\mathrm{d}(\overline{\mathbf{X}} f_{\mathbf{Y}}) = -\mathrm{d} f_{[\mathbf{X},\mathbf{Y}]}. \tag{8.32}$$

Comparing (8.31) and (8.32), and using the fact that $\mathrm{d}\theta$ is non-singular, we conclude that

$$\overline{[\mathbf{X}, \mathbf{Y}]} = [\overline{\mathbf{X}}, \overline{\mathbf{Y}}], \tag{8.33}$$

which implies that the map $\mathbf{X} \mapsto \overline{\mathbf{X}}$ is also a Lie algebra homomorphism.

Exercise 8.16 Prove (8.30) and (8.33), making use of the explicit local expressions (8.21), (8.26), and (8.27).

Symplectic Manifolds The cotangent bundle of a differentiable manifold is an example of a symplectic manifold. A *symplectic manifold* is a differentiable manifold M endowed with a closed nondegenerate 2-form ω; that is, $\mathrm{d}\omega = 0$ and for each $p \in M$, $v_p \lrcorner \omega_p = 0$ implies $v_p = 0$. The 2-form ω is called a *symplectic form* and it is said that it defines a symplectic structure on M. In the case of the cotangent bundle of a manifold, the fundamental 2-form $\mathrm{d}\theta$ is a symplectic form that is not only closed, but exact.

In any symplectic manifold one can define the notion of a Hamiltonian vector field and the Poisson bracket by simply substituting into (8.16) and (8.22) the corresponding symplectic form ω in place of $\mathrm{d}\theta$. The fact that ω is nondegenerate requires that the dimension of a symplectic manifold be even, and the Darboux Theorem ensures that in a neighborhood of any point of a symplectic manifold there is a coordinate system $(q^1, \ldots, q^n, p_1, \ldots, p_n)$, such that

$$\omega = \mathrm{d}p_i \wedge \mathrm{d}q^i \tag{8.34}$$

[cf. (8.12)] (see, e.g., Crampin and Pirani 1986; Woodhouse 1997; Berndt 2001). Any local coordinate system (q^i, p_i) in which the symplectic form ω has the expression (8.34) is called a *canonical coordinate system*. A symplectic manifold possesses an infinite number of local canonical coordinate systems. A (*passive*) canonical transformation is a coordinate transformation that relates two systems of canonical coordinates. According to (8.12), the coordinates defined on T^*M by (8.5), induced by any coordinate system $(x^1, \ldots, x^n)$ on M, are canonical, considering T^*M as a symplectic manifold with the symplectic structure given by the fundamental 2-form; however, there is an infinite number of canonical coordinate systems that are not obtained in this manner (see Examples 8.17, 8.20, and 8.34 below).

Example 8.17 A simple well-known example of a canonical transformation is given by

$$\begin{aligned} p &= \sqrt{2m\omega_0 P}\cos Q, \\ q &= \sqrt{\frac{2P}{m\omega_0}}\sin Q, \end{aligned} \tag{8.35}$$

where m and ω_0 are constants. One readily verifies that

$$\begin{aligned}\mathrm{d}p \wedge \mathrm{d}q &= 2\Big(-P^{1/2}\sin Q\,\mathrm{d}Q + \frac{1}{2}P^{-1/2}\cos Q\,\mathrm{d}P\Big)\\ &\qquad \wedge \Big(P^{1/2}\cos Q\,\mathrm{d}Q + \frac{1}{2}P^{-1/2}\sin Q\,\mathrm{d}P\Big)\\ &= \mathrm{d}P \wedge \mathrm{d}Q,\end{aligned}$$

which means that (8.35) is a canonical transformation. (This canonical transformation is useful in connection with the problem of a one-dimensional harmonic oscillator.)

Example 8.18 The area element of the sphere $\mathrm{S}^2 \equiv \{(x, y, z) \in \mathbb{R}^3 \,|\, x^2 + y^2 + z^2 = 1\}$ is (locally) given in terms of the usual spherical coordinates by $\sin\theta\,\mathrm{d}\theta \wedge \mathrm{d}\phi$. This 2-form is closed (any 3-form on S^2 is equal to zero) and, as can readily be seen, non-singular. With this 2-form, S^2 is a symplectic manifold in such a way that all the rotations about the origin in $\mathbb{R}^3$ are symplectomorphisms. Since $\sin\theta\,\mathrm{d}\theta \wedge \mathrm{d}\phi = \mathrm{d}\phi \wedge \mathrm{d}\cos\theta$, the functions $p_1 = \phi$ and $q^1 = \cos\theta$ form a *local* canonical coordinate system for this symplectic manifold. By contrast with the canonical 2-form of a cotangent bundle, the area element of S^2 is not exact.

8.3 The Phase Space and the Hamilton Equations

Now we will consider a mechanical system whose configuration space is a differentiable manifold, M, of finite dimension (that is, we are considering a mechanical system with a finite number of degrees of freedom, without constraints or with holonomic constraints). According to Newton's laws, the configuration of the system at some instant is not enough to determine its configuration at some other instant; however, usually, the evolution of the system is fixed by the configuration and the momentum of the system at some instant.

The momentum of the system corresponds to a covector α_p, at the point p of M that represents the configuration of the system at that instant; therefore, each point of T^*M determines a state of the system. When M is a configuration space, T^*M is called the *phase space*. If $\alpha_p \in T_p^*M$ represents the state of the system, there exists a unique curve in T^*M passing through α_p describing the evolution of the state of the system.

If the external conditions of the system do not vary with time, we define a map $\varphi_t : T^*M \to T^*M$ by the condition that $\varphi_t(\alpha_p)$ be the state of the system a time t after the system was at the state α_p. Then, $\varphi_{t_1} \circ \varphi_{t_2} = \varphi_{t_2} \circ \varphi_{t_1} = \varphi_{t_1+t_2}$ and φ_0 is the identity mapping. It will be assumed that the φ_t form a one-parameter group of diffeomorphisms whose infinitesimal generator is a Hamiltonian vector field $\mathbf{X}_{\mathrm{d}H}$, where $H \in C^\infty(T^*M)$ is called the *Hamiltonian* of the system. Hence, the curves in the phase space T^*M that represent the evolution of the system are the integral curves of $\mathbf{X}_{\mathrm{d}H}$ and each φ_t is an active canonical transformation, that is, ${\varphi_t}^*(\mathrm{d}\theta) = \mathrm{d}\theta$. From (8.17) we have

$$\mathbf{X}_{\mathrm{d}H} = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i},$$

and therefore the integral curves of $\mathbf{X}_{\mathrm{d}H}$ are given by the equations [see (2.4)]

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial q^i}, \tag{8.36}$$

which are known as the *Hamilton equations*. (As in previous cases, by abuse of notation, we have written q^i and p_i instead of $q^i \circ C$ and $p_i \circ C$, respectively.)

Usually, the configuration space of a mechanical system is a Riemannian manifold, with a metric tensor related to the kinetic energy. In many cases, the configuration space is a submanifold of a product of Euclidean spaces, and its metric is induced by the usual metric of the Euclidean space. For instance, the configuration space of the system formed by two point particles of masses m_1 and m_2 free to move in the Euclidean plane is $\mathbb{E}^2 \times \mathbb{E}^2$, and the metric

$$g = m_1(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y) + m_2(\mathrm{d}\tilde{x} \otimes \mathrm{d}\tilde{x} + \mathrm{d}\tilde{y} \otimes \mathrm{d}\tilde{y}), \tag{8.37}$$

where (x, y) and $(\tilde{x}, \tilde{y})$ are Cartesian coordinates of m_1 and m_2, respectively, is such that $E_{\mathrm{K}} = \frac{1}{2}g(C'_t, C'_t)$ is the kinetic energy of the system if C is the curve in M such that $C(t)$ is the configuration of the system at time t.

In the case of a system formed by a block of mass m_1 sliding under the influence of gravity on a wedge of mass m_2 that lies on a horizontal table, with both blocks restricted to movement in a vertical plane, the configuration space, M, can be viewed as the two-dimensional submanifold of $\mathbb{E}^2 \times \mathbb{E}^2$ defined by $y - (x - \tilde{x})\tan\theta = 0$, and $\tilde{y} = 0$, where θ is the angle of the wedge, (x, y) and $(\tilde{x}, \tilde{y})$ are Cartesian coordinates of the block and the wedge, respectively (see Fig. 8.1). More precisely, if $i : M \to \mathbb{E}^2 \times \mathbb{E}^2$ denotes the inclusion map, then we have $i^*(y - (x - \tilde{x})\tan\theta) = 0$ and $i^*\tilde{y} = 0$. Defining the coordinates (x^1, x^2) on M by

$$x^1 \equiv i^*x, \qquad x^2 \equiv i^*\tilde{x},$$

we have $i^*y = (x^1 - x^2)\tan\theta$ and $i^*\tilde{y} = 0$. Thus, the metric induced on M by (8.37) is given locally by

$$\begin{aligned} i^*g &= i^*\big[m_1(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y) + m_2(\mathrm{d}\tilde{x} \otimes \mathrm{d}\tilde{x} + \mathrm{d}\tilde{y} \otimes \mathrm{d}\tilde{y})\big] \\ &= m_1\big[\mathrm{d}x^1 \otimes \mathrm{d}x^1 + \tan^2\theta\big(\mathrm{d}x^1 - \mathrm{d}x^2\big) \otimes \big(\mathrm{d}x^1 - \mathrm{d}x^2\big)\big] \\ &\quad + m_2\,\mathrm{d}x^2 \otimes \mathrm{d}x^2, \end{aligned} \tag{8.38}$$

so that the kinetic energy of this mechanical system is $E_{\mathrm{K}} = \frac{1}{2}(i^*g)(C'_t, C'_t)$, where C is the curve in M such that $C(t)$ represents the configuration of the system at time t.

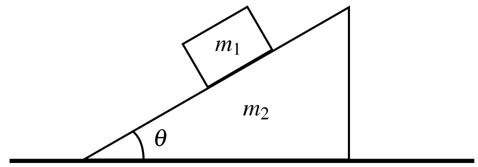

Fig. 8.1 The block can slide on the wedge-shaped block, which lies on a horizontal surface

In many elementary examples, the Hamiltonian function corresponds to the total energy (but not always) and is given by $H = \frac{1}{2}(\pi^* g^{ij}) p_i p_j + \pi^* V$, where (g^{ij}) is the inverse of the matrix (g_{ij}) formed by the components of the metric tensor of M with respect to a coordinate system $(x^1, \ldots, x^n)$ [see, e.g., (8.38)]. The p_i are part of the coordinates on T^*M induced by the x^i, and V is some real-valued function defined on M, which corresponds to the potential energy. The standard procedure to find a Hamiltonian makes use of a Lagrangian, which can readily be constructed provided that the forces are derivable from a potential. Alternatively, a Hamiltonian can be proposed starting from the equations of motion (see Sect. 8.7).

Example 8.19 The cotangent bundle of a manifold may have various symplectic forms, apart from the fundamental 2-form (8.12). In fact, the interaction of a charged particle with a static magnetic field can be accounted for by making use of a suitable symplectic form on the cotangent bundle of the configuration space. We shall consider a point particle of mass m and electric charge e in the three-dimensional Euclidean space, in the presence of a static magnetic field, which is represented by a vector field $\mathbf{B}$ on $\mathbb{E}^3$. If η is a volume element on $\mathbb{E}^3$, the 2-form $\mathbf{B} \lrcorner\, \eta$ is closed because the divergence of $\mathbf{B}$ vanishes, according to the basic equations of electromagnetism [see (6.107)]. Hence, the 2-form

$$\omega = \mathrm{d}\theta + \frac{e}{c} \pi^* (\mathbf{B} \lrcorner\, \eta), \tag{8.39}$$

where θ is the fundamental 1-form of $T^*\mathbb{E}^3$ and c is the speed of light in vacuum, is closed and, as can readily be verified, is always nondegenerate; therefore, ω is a symplectic 2-form. (The magnetic field is a pseudovector field; $\mathbf{B}$ is multiplied by a factor -1 when the orientation is reversed, so that the last term in (8.39) does not depend on which one of the two volume forms, or orientations, of $\mathbb{E}^3$ one chooses.)

Making use of the local expressions (8.6) and (6.102) we have

$$\omega = \mathrm{d}p_i \wedge \mathrm{d}q^i + \frac{e}{2c} \pi^* \big(B^k \sqrt{\det(g_{ij})}\, \varepsilon_{kij} \big)\, \mathrm{d}q^i \wedge \mathrm{d}q^j; \tag{8.40}$$

hence, for any $f \in C^\infty(T^*\mathbb{E}^3)$ the Hamiltonian vector field $\mathbf{X}_{\mathrm{d}f}$ is given locally by

$$\mathbf{X}_{\mathrm{d}f} = \frac{\partial f}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i} \frac{\partial}{\partial p_i} - \frac{e}{c} \pi^* \big(B^k \sqrt{\det(g_{ij})}\, \varepsilon_{kij} \big) \frac{\partial f}{\partial p_i} \frac{\partial}{\partial p_j} \tag{8.41}$$

and the Poisson bracket has the expression

$$\{f, g\} = \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i} - \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{e}{c}\pi^*\big(B^k\sqrt{\det(g_{ij})}\,\varepsilon_{kij}\big)\frac{\partial f}{\partial p_i}\frac{\partial g}{\partial p_j}$$

[cf. (8.17) and (8.21)].

Since $\mathrm{d}(\mathbf{B}\lrcorner\,\eta) = 0$, there exists, at least locally, a 1-form, α, on $\mathbb{E}^3$, such that $\mathbf{B}\lrcorner\,\eta = \mathrm{d}\alpha$. Writing $\alpha = A_i\,\mathrm{d}x^i$, where the A_i are some real-valued functions defined on $\mathbb{E}^3$, locally, we have

$$\omega = \mathrm{d}p_i \wedge \mathrm{d}q^i + \frac{e}{c}\pi^*\big(\mathrm{d}A_i \wedge \mathrm{d}x^i\big) = \mathrm{d}\left(p_i + \frac{e}{c}\pi^* A_i\right) \wedge \mathrm{d}q^i,$$

which shows that $(q^i, p_i + \frac{e}{c}\pi^* A_i)$ are canonical coordinates. The 1-form α is not uniquely defined by $\mathbf{B}$; if we define $\alpha' \equiv \alpha + \mathrm{d}\xi$, where ξ is an arbitrary (differentiable) real-valued function, we have $\mathrm{d}\alpha' = \mathrm{d}(\alpha + \mathrm{d}\xi) = \mathrm{d}\alpha$, and therefore, $\mathbf{B}\lrcorner\,\eta = \mathrm{d}\alpha = \mathrm{d}\alpha'$. Hence, if we write $\alpha' = A'_i\,\mathrm{d}x^i$, it follows that q^i, $p_i + \frac{e}{c}\pi^* A'_i$ is another system of canonical coordinates. (It is said that α' and α are related to each other by a *gauge transformation*.)

Thus, in the case of the interaction with a magnetic field, we can employ the coordinates q^i, p_i appearing in the equations above, which are not canonical (see Example 8.30, below), or we can make use of the coordinates q^i, P_i, with $P_i \equiv p_i + \frac{e}{c}\pi^* A_i$, which are canonical but depend on the choice of the *vector potential* A_i. It should be clear, however, that we are dealing with just *one* symplectic structure, which can be expressed in terms of various coordinate systems [cf. Woodhouse (1997, Sect. 2.6)].

The Hamiltonian function is given by

$$H = \frac{1}{2}\big(\pi^* g^{ij}\big) p_i p_j, \tag{8.42}$$

as in the case of a free particle. Assuming that the x^i are Cartesian coordinates on $\mathbb{E}^3$ (thus, $g_{ij} = m\delta_{ij}$), substituting (8.42) into (8.41) one obtains

$$\mathbf{X}_{\mathrm{d}H} = \frac{1}{m}\delta^{ij} p_j \frac{\partial}{\partial q^i} - \frac{e}{c}\big(\pi^* B^k\big)\varepsilon_{kij}\frac{1}{m}\delta^{il} p_l \frac{\partial}{\partial p_j},$$

and therefore the integral curves of $\mathbf{X}_{\mathrm{d}H}$ are given by

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{1}{m}\delta^{ij} p_j, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{e}{c}\big(\pi^* B^k\big)\varepsilon_{kji}\frac{1}{m}\delta^{jl} p_l$$

which are equivalent to the elementary expression of the Lorentz force, $\mathrm{d}\mathbf{p}/\mathrm{d}t = (e/c)\mathbf{v} \times \mathbf{B}$, with $\mathbf{p} = m\mathbf{v}$.

Example 8.20 As pointed out in Example 8.19, the interaction of a charged particle with a magnetic field can be accounted for by means of the symplectic 2-form $\omega =$

$\mathrm{d}\theta + (e/c)\pi^*(\mathbf{B} \lrcorner \eta)$ on T^*M, where e is the electric charge of the particle. In the case where the particle is moving in the three-dimensional Euclidean space in the presence of a uniform magnetic field $\mathbf{B} = B\,(\partial/\partial x^3)$, where the x^i are Cartesian coordinates and B is a constant,

$$\begin{aligned}\omega &= \mathrm{d}p_1 \wedge \mathrm{d}q^1 + \mathrm{d}p_2 \wedge \mathrm{d}q^2 + \mathrm{d}p_3 \wedge \mathrm{d}q^3 + \frac{eB}{c}\,\mathrm{d}q^1 \wedge \mathrm{d}q^2 \\ &= \mathrm{d}\left(p_1 - \frac{eB}{2c}q^2\right) \wedge \mathrm{d}q^1 + \mathrm{d}\left(p_2 + \frac{eB}{2c}q^1\right) \wedge \mathrm{d}q^2 + \mathrm{d}p_3 \wedge \mathrm{d}q^3.\end{aligned}$$

Hence, $(q^1, q^2, q^3, p_1 - \frac{eB}{2c}q^2, p_2 + \frac{eB}{2c}q^1, p_3)$ is a system of canonical coordinates. A straightforward computation shows that the coordinates (q'^i, p'_i) defined by

$$q^1 = q'^1 + q'^2, \qquad q^2 = \frac{c}{eB}(p'_1 - p'_2), \qquad q^3 = q'^3,$$

$$p_1 - \frac{eB}{2c}q^2 = \frac{1}{2}(p'_1 + p'_2), \qquad p_2 + \frac{eB}{2c}q^1 = \frac{eB}{2c}\left(q'^2 - q'^1\right), \qquad p_3 = p'_3,$$

are also canonical, i.e., $\omega = \mathrm{d}p'_i \wedge \mathrm{d}q'^i$. In terms of these coordinates, the Hamiltonian (8.42) takes the form

$$H = \frac{1}{2m}(p'_1)^2 + \frac{m}{2}\left(\frac{eB}{mc}\right)^2 \left(q'^1\right)^2 + \frac{1}{2m}(p'_3)^2.$$

The first two terms on the right-hand side of this last expression constitute the usual Hamiltonian of a one-dimensional harmonic oscillator (of angular frequency eB/mc) and, since the *canonical* coordinates q'^2, p'_2, and q'^3 do not appear in H (i.e., are *ignorable* or *cyclic* variables), p'_2, q'^2, and p'_3 are constants of motion [see (8.36)].

If f is a differentiable real-valued function defined on T^*M, the rate of change of f along a curve C followed by the system in its time evolution is given by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(f \circ C)\bigg|_{t=t_0} &= C'_{t_0}[f] = (\mathbf{X}_{\mathrm{d}H})_{C(t_0)}[f] = (\mathbf{X}_{\mathrm{d}H}f)\big(C(t_0)\big) \\ &= \{H, f\}\big(C(t_0)\big). \end{aligned} \tag{8.43}$$

Hence, f is a *constant of motion* if and only if $\{H, f\} = 0$. The Hamiltonian H is a constant of motion since $\{H, H\} = 0$.

If f and g are both constants of motion, it is clear that $\{H, af + bg\} = 0$ for $a, b \in \mathbb{R}$. By virtue of the Jacobi identity, we have, in addition,

$$\big\{H, \{f, g\}\big\} = -\big\{f, \{g, H\}\big\} - \big\{g, \{H, f\}\big\} = 0,$$

hence the set of constants of motion is a Lie subalgebra of $C^\infty(T^*M)$.

Constants of Motion and Symmetries If $f \in C^\infty(T^*M)$ is a constant of motion, then H is invariant under the (possibly local) one-parameter group of canonical transformations generated by $\mathbf{X}_{\mathrm{d}f}$ since

$$\pounds_{\mathbf{X}_{\mathrm{d}f}} H = \mathbf{X}_{\mathrm{d}f} H = \{f, H\} = -\{H, f\} = 0. \tag{8.44}$$

Conversely, if H is invariant under a one-parameter group of *canonical* transformations then there exists, locally, a constant of motion associated with this symmetry. In effect, if $\mathbf{X}$ is the infinitesimal generator of a one-parameter group of canonical transformations, then $\mathrm{d}(\pounds_{\mathbf{X}}\theta) = \pounds_{\mathbf{X}}\,\mathrm{d}\theta = 0$; hence, there exists locally a real-valued function (defined up to an additive constant), F, such that

$$\pounds_{\mathbf{X}}\theta = \mathrm{d}F,$$

that is, $\mathbf{X} \lrcorner\, \mathrm{d}\theta + \mathrm{d}(\mathbf{X} \lrcorner\, \theta) = \mathrm{d}F$ or, equivalently, $\mathbf{X} \lrcorner\, \mathrm{d}\theta = -\mathrm{d}(\mathbf{X} \lrcorner\, \theta - F)$, which explicitly shows that $\mathbf{X}$ is locally Hamiltonian [see (8.16)] and that it corresponds to the function

$$\chi \equiv \mathbf{X} \lrcorner\, \theta - F,$$

which is a constant of motion, as follows from $0 = \mathbf{X}H = \{\chi, H\} = -\{H, \chi\} = -\mathbf{X}_{\mathrm{d}H}\chi$. The function F can be chosen equal to zero if and only if $\mathbf{X}$ is the canonical lift of a vector field on M (see Exercise 8.13) and in that case the expression for the function χ reduces to (8.24). (Here we are restricting ourselves to constants of motion that do not depend explicitly on the time; the most general case is considered in Sect. 8.7.)

Example 8.21 Let us consider a system formed by two point particles of masses m_1 and m_2 in the three-dimensional Euclidean space, whose positions are represented by the vectors $\mathbf{r}_1 = (x^1, x^2, x^3)$ and $\mathbf{r}_2 = (x^4, x^5, x^6)$. The configuration space for this system has dimension six and can be identified with $\mathbb{R}^3 \times \mathbb{R}^3$. Denoting by $(q^1, \ldots, q^6, p_1, \ldots, p_6)$ the coordinates on T^*M induced by $(x^1, \ldots, x^6)$ in the form defined in Sect. 8.1, the Hamiltonian has the expression

$$H = \frac{1}{2}\big(\pi^* g^{ij}\big) p_i p_j + V = \frac{1}{2m_1}\sum_{i=1}^{3}(p_i)^2 + \frac{1}{2m_2}\sum_{i=4}^{6}(p_i)^2 + V, \tag{8.45}$$

where (g^{ij}) is the inverse of the matrix (g_{ij}) formed by the components of the metric tensor of M with respect to the coordinate system $(x^1, \ldots, x^6)$ [cf. (8.37)] and V is the potential energy of the system.

If the particles do not interact with objects external to the system, in the absence of velocity-dependent forces (such as the magnetic force), V must be a function of $|\mathbf{r}_1 - \mathbf{r}_2|$ only (more precisely, $V = v(r)$, where v is a real-valued function of a single variable and $r \equiv [(q^4 - q^1)^2 + (q^5 - q^2)^2 + (q^6 - q^3)^2]^{1/2}$ is the distance between the particles). This means, for instance, that the Hamiltonian function (8.45) is invariant under the simultaneous translations of the particles in the x-direction, that is, H is invariant under the one-parameter group of transformations

$$\varphi_t^* q^1 = q^1 + t, \varphi \qquad {}_t^* q^4 = q^4 + t,$$
$$\varphi_t^* q^i = q^i, \quad \text{if } i \neq 1, 4, \qquad \varphi_t^* p_i = p_i,$$

as can readily be verified. Clearly, this is a one-parameter group of canonical transformations (in fact, $\varphi_t{}^* \, \mathrm{d}q^i = \mathrm{d}q^i$, $\varphi_t{}^* \, \mathrm{d}p_i = \mathrm{d}p_i$) and its infinitesimal generator is $\partial/\partial q^1 + \partial/\partial q^4$, which is globally Hamiltonian, $(\partial/\partial q^1 + \partial/\partial q^4) \lrcorner \, \mathrm{d}\theta = -\mathrm{d}(p_1 + p_4)$. Therefore $p_1 + p_4$ is a constant of motion, which is associated with the invariance of H under translations in the x direction and corresponds to the x component of the linear momentum *of the system* (in fact, $(\partial/\partial q^1 + \partial/\partial q^4) H = 0$). In a similar way, $p_2 + p_5$ and $p_3 + p_6$ are constants of motion that represent the y and z components of the *total* linear momentum, respectively.

Since the distance between the two particles, $|\mathbf{r}_2 - \mathbf{r}_1|$, is also invariant under rotations of the system, it is to be expected that there exist constants of motion associated with this symmetry; however, in order to find a constant of motion it is necessary that H be invariant under a one-parameter group of transformations acting *on the phase space* and that these transformations be canonical. The infinitesimal generators of the rotations about the x, y, and z axes in the configuration space are

$$\mathbf{X}_i = \sum_{k=1}^{3} \varepsilon_{ijk} \left(x^j \frac{\partial}{\partial x^k} + x^{j+3} \frac{\partial}{\partial x^{k+3}} \right), \quad i = 1, 2, 3. \tag{8.46}$$

The canonical lifts $\overline{\mathbf{X}_i}$ to T^*M of the vector fields (8.46) are given by

$$\overline{\mathbf{X}_i} = \sum_{k=1}^{3} \varepsilon_{ijk} \left(q^j \frac{\partial}{\partial q^k} + q^{j+3} \frac{\partial}{\partial q^{k+3}} \right) + \sum_{j=1}^{3} \varepsilon_{ijk} \left(p_j \frac{\partial}{\partial p_k} + p_{j+3} \frac{\partial}{\partial p_{k+3}} \right)$$

(see Exercises 8.12 and 8.15). As shown in Sect. 8.2, these vector fields are globally Hamiltonian and correspond to the functions [see (8.26) and (8.46)]

$$L_i \equiv \sum_{k=1}^{3} \varepsilon_{ijk} \left(q^j p_k + q^{j+3} p_{k+3} \right), \quad i = 1, 2, 3.$$

One can readily verify that, in effect, $\overline{\mathbf{X}_i} H = 0$, and therefore the functions L_i are constants of motion, which represent the Cartesian components of the *total* angular momentum. (Further examples are given in Sects. 8.4–8.6.)

Exercise 8.22 Consider the system formed by a block and a wedge discussed at the beginning of this section (p. 212). Assuming that the potential energy is given by $V = i^*(m_1 g \, y)$, where g is the acceleration of gravity, show that the Hamiltonian is given by

$$H = \frac{m_2(p_1)^2 + m_1[(p_2)^2 + \tan^2\theta(p_1+p_2)^2]}{2m_1[m_2 + (m_1+m_2)\tan^2\theta]} + m_1 g \tan\theta\big(q^1 - q^2\big).$$

Show that the canonical lift of $\partial/\partial x^1 + \partial/\partial x^2$ is the infinitesimal generator of a one-parameter group of canonical transformations that leave the Hamiltonian invariant, and find the corresponding constant of motion. (Note that the one-parameter group of diffeomorphisms generated by $\partial/\partial x^1 + \partial/\partial x^2$ corresponds to translations of the mechanical system as a whole.) Show that $(m_1 q^1 + m_2 q^2) m_1 g \tan\theta + \frac{1}{2}[(p_1)^2 - (p_2)^2]$ is a constant of motion, but that it is not associated with the canonical lift of a vector field on M (and therefore, it corresponds to a so-called *hidden symmetry*; see also Sect. 8.5).

8.4 Geodesics, the Fermat Principle, and Geometrical Optics

If M is a Riemannian manifold, one can consider the Hamiltonian function $H(\alpha_p) \equiv \frac{1}{2}(\alpha_p|\alpha_p)$ [see (6.10)]. In terms of the coordinates (q^i, p_i) of T^*M, induced by a coordinate system x^i on M, this Hamiltonian has the local expression

$$H = \frac{1}{2}\big(\pi^* g^{ij}\big) p_i p_j, \tag{8.47}$$

where (g^{ij}) is the inverse of the matrix (g_{ij}) formed by the components of the metric tensor of M with respect to the natural basis induced by the coordinates x^i (hence, H is a differentiable function). There exist several examples where there appear Hamiltonians of this form. In the theory of relativity (special or general), if M represents the space–time, the Hamiltonian function (8.47) corresponds to a particle subject to the gravitational field (represented by the metric tensor g). Other important examples, to be considered below, are those of a free rigid body (see Sect. 8.6), geometrical optics, and the Jacobi principle. Since the Hamiltonian (8.47) is defined starting from the metric tensor, it is to be expected that it shows a simple behavior under an isometry.

Theorem 8.23 *Let M be a Riemannian manifold and let $H \in C^\infty(T^*M)$ be defined by $H(\alpha_p) = \frac{1}{2}(\alpha_p|\alpha_p)$; then the diffeomorphism $\psi : M \to M$ is an isometry if and only if $\overline{\psi}^* H = H$.*

Proof Let $g = g_{ij}\, \mathrm{d}x^i \otimes \mathrm{d}x^j$ be the metric tensor of M and let $g' \equiv \psi^* g$; then

$$\begin{aligned} g'_{ij}\, \mathrm{d}x^i \otimes \mathrm{d}x^j &= (\psi^* g_{ij})\, \mathrm{d}\big(\psi^* x^i\big) \otimes \mathrm{d}\big(\psi^* x^j\big) \\ &= (\psi^* g_{ij}) \frac{\partial(x^i \circ \psi)}{\partial x^k} \frac{\partial(x^j \circ \psi)}{\partial x^l}\, \mathrm{d}x^k \otimes \mathrm{d}x^l, \end{aligned}$$

which amounts to the relation

$$\psi^* g^{ij} = g'^{kl} \frac{\partial(x^i \circ \psi)}{\partial x^k} \frac{\partial(x^j \circ \psi)}{\partial x^l}. \tag{8.48}$$

Using (8.47), (8.8), (8.10), and (8.48) we have

$$\begin{aligned}
\overline{\psi}^* H &= \frac{1}{2}\big((\pi \circ \overline{\psi})^* g^{ij}\big)(\overline{\psi}^* p_i)(\tilde{\psi}^* p_j) \\
&= \frac{1}{2}\pi^* \left[\big(\psi^* g^{ij}\big) \left(\psi^* \frac{\partial(x^k \circ \psi^{-1})}{\partial x'^i} \right) \left(\psi^* \frac{\partial(x^l \circ \psi^{-1})}{\partial x'^j} \right) \right] p_k p_l \\
&= \frac{1}{2}\big(\pi^* g'^{kl}\big) p_k p_l,
\end{aligned}$$

which coincides with H if and only if $g' = g$, that is, $\psi^* g = g$. □

If $\mathbf{X} \in \mathfrak{X}(M)$ is a Killing vector field, $\mathbf{X}$ is the infinitesimal generator of a local one-parameter group of isometries φ_t of M; according to Theorems 8.4 and 8.23, the transformations $\overline{\varphi_t}$ are canonical and leave invariant the Hamiltonian (8.47), which is equivalent to the existence of a constant of motion associated with the vector field $\mathbf{X}$. Since the infinitesimal generator of $\overline{\varphi_t}$ is the Hamiltonian vector field associated with the function $f_{\mathbf{X}}$ [see (8.23) and (8.24)], the function $f_{\mathbf{X}} = \overline{\mathbf{X}} \lrcorner\, \theta$ (or, in local form, $f_{\mathbf{X}} = (\pi^* X^i) p_i$, where the X^i are the components of $\mathbf{X}$) is a constant of motion.

Exercise 8.24 Show that, conversely, if $\mathbf{X} \in \mathfrak{X}(M)$ and $f_{\mathbf{X}} = \overline{\mathbf{X}} \lrcorner\, \theta$ is a constant of motion for the system with Hamiltonian (8.47), then $\mathbf{X}$ is a Killing vector field.

Exercise 8.25 Show that the Hamilton equations corresponding to the Hamiltonian (8.47) yield the geodesic equations (more precisely: the projection on M of the integral curves of $\mathbf{X}_{\mathrm{d}H}$ are the geodesics of M) and that if $\mathbf{X}$ is a Killing vector field, the value of $f_{\mathbf{X}}$ along an integral curve of $\mathbf{X}_{\mathrm{d}H}$ coincides with the value of $g(\mathbf{X}, C')$ on the corresponding geodesic C (see Theorem 6.28).

Jacobi's Principle Many of the examples considered in classical mechanics correspond to Hamiltonian functions of the form

$$H = \frac{1}{2}\big(\pi^* g^{ij}\big) p_i p_j + \pi^* V, \tag{8.49}$$

in terms of the coordinates (q^i, p_i) on T^*M, induced by a coordinate system x^i on the configuration space M, where (g^{ij}) is the inverse of the matrix formed by the components of a metric tensor on M, and V is a function of M in $\mathbb{R}$ (this means that the potential energy only depends on the configuration). The *Jacobi principle* states that the orbits followed in M are the geodesics of the metric $(E - V) g_{ij}\, \mathrm{d}x^i \otimes \mathrm{d}x^j$, where E is the (constant) value of H determined by the initial conditions.

In effect, assuming that $\mathrm{d}H|_{H=E} \neq 0$, we define the auxiliary Hamiltonian

$$h \equiv \frac{H - \pi^* V}{E - \pi^* V}, \tag{8.50}$$

and we find that $\mathrm{d}h = (E - \pi^* V)^{-2}[(E - \pi^* V)\,\mathrm{d}H + (H - E)\,\mathrm{d}\pi^* V]$; therefore, $\mathrm{d}h|_{H=E} = (E - \pi^* V)^{-1}\,\mathrm{d}H|_{H=E}$, and then we have $\mathbf{X}_{\mathrm{d}H}|_{H=E} = (E - \pi^* V)\mathbf{X}_{\mathrm{d}h}|_{H=E}$. Thus, at the points of the submanifold $H = E$, the vector fields $\mathbf{X}_{\mathrm{d}H}$ and $\mathbf{X}_{\mathrm{d}h}$ are collinear, and therefore their integral curves differ only in the parametrization. (Note that $H = E$ amounts to $h = 1$.) Whereas the integral curves of $\mathbf{X}_{\mathrm{d}H}$ are parameterized by the time, t, the parameter of the integral curves of $\mathbf{X}_{\mathrm{d}h}$ is another variable, τ, which is related to t as follows. If C is an integral curve of $\mathbf{X}_{\mathrm{d}H}$ on the hypersurface $H = E$, then $\tau = I(t)$ with

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (E - \pi^* V) \circ C$$

[cf. (2.14)]. Indeed, the curve $\sigma(\tau) \equiv C(I^{-1}(\tau))$ is a reparametrization of C that is an integral curve of $\mathbf{X}_{\mathrm{d}h}$, since, for $f \in C^\infty(T^*M)$,

$$\begin{aligned}
(\mathbf{X}_{\mathrm{d}h} f)\big(\sigma(\tau)\big) &= \left[\frac{1}{E - \pi^* V}\mathbf{X}_{\mathrm{d}H} f\right]\Big(C\big(I^{-1}(\tau)\big)\Big) \\
&= \frac{1}{(E - \pi^* V)(C(I^{-1}(\tau)))} \left.\frac{\mathrm{d}(f \circ C)}{\mathrm{d}t}\right|_{I^{-1}(\tau)} \\
&= \frac{1}{(\mathrm{d}I/\mathrm{d}t)|_{I^{-1}(\tau)}} \left.\frac{\mathrm{d}(f \circ \sigma \circ I)}{\mathrm{d}t}\right|_{I^{-1}(\tau)} \\
&= \frac{1}{(\mathrm{d}I/\mathrm{d}t)|_{I^{-1}(\tau)}} \left.\frac{\mathrm{d}(f \circ \sigma)}{\mathrm{d}\tau}\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{I^{-1}(\tau)} = \frac{\mathrm{d}(f \circ \sigma)}{\mathrm{d}\tau}.
\end{aligned}$$

From (8.49) and (8.50) we obtain the equivalent expression

$$h = \frac{1}{2}\pi^*\left(\frac{g^{ij}}{E - V}\right)p_i p_j \tag{8.51}$$

[cf. (8.47)], whose orbits in the configuration space are the geodesics corresponding to the metric $(E - V)g_{ij}\,\mathrm{d}x^i \otimes \mathrm{d}x^j$ (see Exercise 8.25).

Combining the foregoing result with the findings at the end of Sect. 6.2 (pp. 140–141) we conclude that the orbits in the configuration space of a system with a Hamiltonian function of the form (8.49) correspond to the intersections of the hypersurfaces $b^k = \text{const}$, where $b^k = \partial W/\partial a_k$ and W is a complete solution of

$$\frac{g^{ij}}{E - V}\frac{\partial W}{\partial x^i}\frac{\partial W}{\partial x^j} = \text{const}$$

which depends on the parameters a_k [cf. (6.73)]. This equation amounts to

$$g^{ij}\frac{\partial W}{\partial x^i}\frac{\partial W}{\partial x^j} = \text{const}\,(E - V) \tag{8.52}$$

and, choosing the arbitrary constant appearing in (8.52) as 2, equation (8.52) becomes the *Hamilton–Jacobi equation* for the Hamilton characteristic function, W,

$$\frac{1}{2}g^{ij}\frac{\partial W}{\partial x^i}\frac{\partial W}{\partial x^j} + V = E \tag{8.53}$$

[cf. (8.49)]. (Note that this equation involves quantities defined on M, not on T^*M.)

Example 8.26 In terms of the parabolic coordinates (u, v) on $\mathbb{E}^2$, defined by $x = u^2 - v^2$, $y = 2uv$, with $v > 0$, where (x, y) are Cartesian coordinates, the usual metric of the Euclidean plane is given locally by

$$\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y = 4\big(u^2 + v^2\big)(\mathrm{d}u \otimes \mathrm{d}u + \mathrm{d}v \otimes \mathrm{d}v);$$

therefore, the Hamiltonian function for the two-dimensional Kepler problem, which corresponds to the potential $V = -k/r$, where k is a positive constant and r is the distance from the particle to a fixed center of force (placed at the origin), is

$$H = \frac{1}{2m}\frac{{p_u}^2 + {p_v}^2}{4(u^2 + v^2)} - \frac{k}{u^2 + v^2},$$

where, by abuse of notation, we are using the same symbols for the coordinates u, v and for their pullbacks under π. Thus, equation (8.53) takes the form

$$\frac{1}{8m(u^2 + v^2)}\left[\left(\frac{\partial W}{\partial u}\right)^2 + \left(\frac{\partial W}{\partial v}\right)^2\right] - \frac{k}{u^2 + v^2} = E.$$

Using the method of separation of variables we look for a complete solution of the form $W = F(u) + G(v)$ and we obtain (cf. Example 6.33)

$$\left(\frac{\mathrm{d}F}{\mathrm{d}u}\right)^2 - 4mk - 8mEu^2 = a,$$

$$\left(\frac{\mathrm{d}G}{\mathrm{d}v}\right)^2 - 4mk - 8mEv^2 = -a,$$

where a is a separation constant. In this problem, the constant E, which represents the total energy, can be positive, negative, or zero. The simplest case corresponds to $E = 0$, and we find that

$$W = \sqrt{4mk + a}\,u + \sqrt{4mk - a}\,v;$$

thus, equating $\partial W/\partial a$ to a constant b, say, we obtain

$$\frac{u}{2\sqrt{4mk + a}} - \frac{v}{2\sqrt{4mk - a}} = b.$$

This equation corresponds to a two-parameter family of parabolas with foci at the origin (the parameters a and b determine the orientation of the axis of the parabola and its focal distance). The cases $E < 0$ and $E > 0$ are dealt with in a similar manner and, as is well known, one obtains ellipses and hyperbolas, respectively, with one focus at the origin.

Geometrical Optics The formalism of the Hamiltonian mechanics is applicable to geometrical optics, in which it is assumed that the light travels along curves (the light rays). At each point of an isotropic medium, which is assumed to be a Riemannian manifold M (usually the three-dimensional Euclidean space), the speed of light does not depend on the direction of the ray and is expressed as c/n, where c is the speed of light in vacuum and n is a real-valued function defined on M, known as the refractive index.

Since c/n is the velocity of the light at each point of M, if the curve $C : [a, b] \to M$ represents a light ray, the time spent by the light going from point $C(a)$ to $C(b)$ along C is

$$\frac{1}{c}\int_a^b n\big(C(t)\big)\|C_t'\|\,\mathrm{d}t = \frac{1}{c}\int_a^b \sqrt{n^2\big(C(t)\big)g(C_t', C_t')}\,\mathrm{d}t, \tag{8.54}$$

where g is the metric tensor of M. The variable t appearing in the last integral does not need to be the time, since the integral (8.54) is invariant under changes of parameter. This invariance is similar to that of the integral (6.1), which gives the length of a curve. In fact, comparing (8.54) with (6.1), one finds that the integral in (8.54) represents the length of C defined by the metric tensor $n^2 g$.

According to Fermat's principle, given two points of M, the path followed by the light going from one point to the other is that for which the time required is minimum or a stationary value. This implies that the light rays are the geodesics of the metric $n^2 g$. Hence, the light rays are the projections on M of the integral curves of $\mathbf{X}_{\mathrm{d}H}$, with the Hamiltonian function, H, locally given by

$$H = \frac{c}{2}\pi^*\left(\frac{g^{ij}}{n^2}\right)p_i p_j \tag{8.55}$$

[cf. (8.47) and (8.51); the constant factor c inserted in (8.55) is introduced for later convenience]. From (8.55) and the Hamilton equations (8.36) one deduces that if σ is an integral curve of $\mathbf{X}_{\mathrm{d}H}$,

$$\frac{\mathrm{d}q^i(\sigma(t))}{\mathrm{d}t} = c\left[\pi^*\left(\frac{g^{ij}}{n^2}\right)p_j\right]\big(\sigma(t)\big);$$

equivalently, if $C \equiv \pi \circ \sigma$, i.e., C is the projection of σ on M,

$$\frac{\mathrm{d}x^i(C(t))}{\mathrm{d}t} = c\left(\frac{g^{ij}}{n^2}\right)\big(C(t)\big)p_j\big(\sigma(t)\big). \tag{8.56}$$

Hence, the square of the norm of C_t' is

$$\|C_t'\|^2 = g_{ij}\big(C(t)\big)c^2\left(\frac{g^{ik}}{n^2}\right)\big(C(t)\big)p_k\big(\sigma(t)\big)\left(\frac{g^{jl}}{n^2}\right)\big(C(t)\big)p_l\big(\sigma(t)\big)$$
$$= c^2\left(\frac{g^{kl}}{n^4}\right)\big(C(t)\big)p_k\big(\sigma(t)\big)p_l\big(\sigma(t)\big) = \frac{c^2}{n^2(C(t))}\frac{2H(\sigma(t))}{c}. \tag{8.57}$$

Since $\{H, H\} = 0$, $H(\sigma(t))$ is some constant and if t represents the time, $\|C_t'\|$ must be c/n; therefore, from (8.57) it follows that the constant value of H along any integral curve of $\mathbf{X}_{\mathrm{d}H}$ must be $c/2$ and, therefore,

$$\big(\pi^* g^{ij}\big) p_i p_j = \pi^* n^2. \tag{8.58}$$

Thus, in order for the integral curves of $\mathbf{X}_{\mathrm{d}H}$ to be parameterized by the time, the only possible value of the Hamiltonian function (8.55) is $c/2$; in other words, any initial condition and any integral curve of $\mathbf{X}_{\mathrm{d}H}$ must lie on the hypersurface $H = c/2$.

The existence of a condition of the form (8.58), which implies that not any point of T^*M is acceptable as an initial condition, is not a unique feature of geometrical optics; in the theory of relativity (special or general) a particle subject, at most, to a gravitational field moves along a geodesic of the space–time, which is a pseudo-Riemannian manifold M. Therefore, we can choose the Hamiltonian function $H = \frac{1}{2}(\pi^* g^{ij}) p_i p_j$. Then, if the integral curves of $\mathbf{X}_{\mathrm{d}H}$ are parameterized by the proper time of the particle, we have $|(\pi^* g^{ij}) p_i p_j| = m^2c^2$, where m is the rest mass of the particle [cf. (8.58)].

As we shall show now, the Snell law follows from (8.55). Assuming that M is the Euclidean space of dimension three, making use of Cartesian coordinates (x, y, z), the components of the metric tensor are $g_{ij} = \delta_{ij}$; then, from (8.55) and the Hamilton equations it follows that in a region where n is a constant, the light rays are straight lines. If we assume that the plane $z = 0$ is the boundary between two regions with distinct (constant) refractive indices n_1 and n_2; $n = n_1$ for $z > 0$ and $n = n_2$ for $z < 0$ (the function n is then discontinuous at $z = 0$, which can be avoided, assuming that n changes smoothly from of n_2 to n_1 around $z = 0$), then the Hamiltonian function (8.55) does not depend on x nor y, therefore p_x and p_y are constants of motion [see (8.36)]. From (8.56) it follows that if a light ray forms an angle θ with the z axis then ${p_x}^2 + {p_y}^2 = ({p_x}^2 + {p_y}^2 + {p_z}^2)\sin^2\theta = (\pi^* n^2)\sin^2\theta$, where we have made use of (8.58). Since ${p_x}^2 + {p_y}^2$ is constant, it follows that

$$n_1 \sin\theta_1 = n_2 \sin\theta_2, \tag{8.59}$$

where θ_1 and θ_2 are the angles made by the light ray with the z axis in the regions $z > 0$ and $z < 0$, respectively. The fact that p_x and p_y are constant implies that the incident ray, the refracted ray, and the z axis are coplanar. Equation (8.59) is the usual expression of Snell's law.

Making use again of the result of Exercise 6.32, the light rays in a homogeneous medium can be obtained from a complete solution of the PDE

$$g^{ij}\frac{\partial S}{\partial x^i}\frac{\partial S}{\partial x^j} = \text{const}\, n^2 \tag{8.60}$$

[cf. (6.73) and (8.52)]. If the arbitrary constant that appears in this equation is set equal to 1, the resulting equation is the *eikonal equation* (and S is called the *eikonal*). The light rays are orthogonal to the surfaces $S = \text{const.}$ (which represent the wave fronts).

8.5 Dynamical Symmetry Groups

As we have shown, each constant of motion that does not depend explicitly on the time corresponds to a possibly local one-parameter group of canonical transformations that leave the Hamiltonian invariant; now we shall study in some detail the action of an arbitrary Lie group on a symplectic manifold that leaves invariant a given Hamiltonian. Usually, attention is restricted to actions by symplectomorphisms; in many elementary examples, one has a Lie group of transformations acting in an arbitrary manner on a manifold M, which does not need to possess any additional structure, and then this action is lifted to the cotangent bundle of M. In the other cases, one has to consider directly the action by symplectomorphisms of a Lie group on a symplectic manifold. We shall begin with the simplest case, assuming that we have a Lie group that acts on a configuration space; we will only have to put together several results obtained above.

Lifted Actions Let G be a Lie group that acts on the right on a differentiable manifold M; that is, each $g \in G$ defines a diffeomorphism $R_g : M \to M$ (with $R_g(p) = pg$, see Sect. 7.6). The diffeomorphism R_g, in turn, gives rise to a diffeomorphism $\overline{R_g} : T^*M \to T^*M$, defined by (8.7), which is a canonical transformation; moreover, $\overline{R_g}^*\theta = \theta$ (see Theorem 8.4). For $p \in M$ and $\alpha_p \in T^*M$, each $\mathbf{X} \in \mathfrak{g}$ defines a curve $t \mapsto \exp t\mathbf{X}$, in G; a curve $t \mapsto R_{\exp t\mathbf{X}}(p)$, in M, and a curve $t \mapsto \overline{R_{\exp t\mathbf{X}}}(\alpha_p)$, in T^*M. The tangent vectors to these curves at $t = 0$ are $\mathbf{X}_e$, $\mathbf{X}^+_p$, and $\overline{\mathbf{X}^+}_{\alpha_p}$, respectively (with $\overline{\mathbf{X}^+}$ being the canonical lift of $\mathbf{X}^+$) (see Sect. 7.6).

As a consequence of the fact that $\overline{R_g}^*\theta = \theta$ for all $g \in G$, in particular, $\overline{R_{\exp t\mathbf{X}}}^*\theta = \theta$, for all $\mathbf{X} \in \mathfrak{g}$; therefore, $\pounds_{\overline{\mathbf{X}^+}}\theta = 0$, which means that the vector field $\overline{\mathbf{X}^+}$ is globally Hamiltonian

$$\overline{\mathbf{X}^+} \lrcorner\, \mathrm{d}\theta = -\mathrm{d} f_{\mathbf{X}^+}, \tag{8.61}$$

where

$$f_{\mathbf{X}^+} \equiv \overline{\mathbf{X}^+} \lrcorner\, \theta \tag{8.62}$$

[see (8.23) and (8.24)]. This expression together with the relation $(a\mathbf{X}+b\mathbf{Y})^+ = a\mathbf{X}^+ + b\mathbf{Y}^+$, for $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$ and $a, b \in \mathbb{R}$, imply that the mapping $\mathbf{X} \mapsto f_{\mathbf{X}^+}$, from $\mathfrak{g}$ into $C^\infty(T^*M)$, is $\mathbb{R}$-linear (i.e., $f_{(a\mathbf{X}+b\mathbf{Y})^+} = af_{\mathbf{X}^+} + bf_{\mathbf{Y}^+}$). By combining Theorem 7.61 and (8.30) we also have

$$f_{[\mathbf{X},\mathbf{Y}]^+} = f_{[\mathbf{X}^+,\mathbf{Y}^+]} = \{f_{\mathbf{X}^+}, f_{\mathbf{Y}^+}\}, \tag{8.63}$$

and therefore the mapping $\mathbf{X} \mapsto f_{\mathbf{X}^+}$ is a Lie algebra homomorphism.

Summarizing, given the action on the right of an arbitrary Lie group G on M, the mapping $(g,\alpha_p) \mapsto \overline{R_g}(\alpha_p)$ defines an action of G on the right on T^*M so that each $\overline{R_g}$ is a canonical transformation (or symplectomorphism); the vector fields induced on T^*M by this action are *globally* Hamiltonian, with $\overline{\mathbf{X}^+}$ corresponding to the function $f_{\mathbf{X}^+}$ in such a way that the map from $\mathfrak{g}$ into $C^\infty(T^*M)$ given by $\mathbf{X} \mapsto f_{\mathbf{X}^+}$ is a Lie algebra homomorphism.

Now, if a given Hamiltonian, H, is invariant under the transformations $\overline{R_g}$ (i.e., $\overline{R_g}^*H = H$ for all $g \in G$), then $\overline{R_{\exp t\mathbf{X}}}^*H = H$ for all $\mathbf{X} \in \mathfrak{g}$, which implies that $0 = \pounds_{\overline{\mathbf{X}^+}}H = \overline{\mathbf{X}^+}H = \{f_{\mathbf{X}^+}, H\}$ [see (8.19)], i.e., each function $f_{\mathbf{X}^+}$ is a constant of motion. The constants of motion obtained in this manner are homogeneous functions of first degree in the variables p_i [see (8.26)]. However, in many cases of interest, some constants of motion are not homogeneous functions of first degree in the p_i, and, therefore, are not associated with the action of a group on M; their existence is a consequence of groups of canonical transformations on T^*M that leave the Hamiltonian invariant, which do not come from a group that acts on M (see Examples 8.29–8.32, below).

Hidden Symmetries Let G be a Lie group that acts on the right on T^*M in such a way that for each $g \in G$ the transformation $R_g : T^*M \to T^*M$, given by $R_g(x) = xg$, is a canonical transformation, that is, $R_g{}^*(\mathrm{d}\theta) = \mathrm{d}\theta$. Then the vector field $\mathbf{X}^+$ on T^*M associated with the vector field $\mathbf{X} \in \mathfrak{g}$ is locally Hamiltonian (see Lemma 8.5); hence, for each $\mathbf{X} \in \mathfrak{g}$ there exists locally a function $\mu_\mathbf{X} \in C^\infty(T^*M)$, defined up to an additive constant, such that

$$\mathbf{X}^+ \lrcorner\, \mathrm{d}\theta = -\mathrm{d}\mu_\mathbf{X} \tag{8.64}$$

[cf. (8.16)]. As we shall see, under certain conditions, it will be possible to choose the functions $\mu_\mathbf{X}$ in such a way that the mapping $\mathbf{X} \mapsto \mu_\mathbf{X}$ is a Lie algebra homomorphism.

Starting from the relations $(a\mathbf{X}+b\mathbf{Y})^+ = a\mathbf{X}^+ + b\mathbf{Y}^+$ and $[\mathbf{X},\mathbf{Y}]^+ = [\mathbf{X}^+,\mathbf{Y}^+]$, valid for every pair of elements $\mathbf{X}, \mathbf{Y}$ of the Lie algebra of G, with $a, b \in \mathbb{R}$, it follows that

$$\begin{aligned} \mathrm{d}\mu_{a\mathbf{X}+b\mathbf{Y}} &= -(a\mathbf{X}+b\mathbf{Y})^+ \lrcorner\, \mathrm{d}\theta = -(a\mathbf{X}^+ + b\mathbf{Y}^+) \lrcorner\, \mathrm{d}\theta \\ &= a\,\mathrm{d}\mu_\mathbf{X} + b\,\mathrm{d}\mu_\mathbf{Y} = \mathrm{d}(a\mu_\mathbf{X} + b\mu_\mathbf{Y}) \end{aligned}$$

and, similarly, using (8.18) and (8.19), we find

$$\begin{aligned} \mathrm{d}\mu_{[\mathbf{X},\mathbf{Y}]} &= -[\mathbf{X},\mathbf{Y}]^{+} \lrcorner \mathrm{d}\theta = -[\mathbf{X}^{+},\mathbf{Y}^{+}] \lrcorner \mathrm{d}\theta \\ &= \mathrm{d}\{\mu_{\mathbf{X}},\mu_{\mathbf{Y}}\}, \quad \text{for } \mathbf{X},\mathbf{Y} \in \mathfrak{g},\ a,b \in \mathbb{R}. \end{aligned}$$

This means that the functions $\mu_{a\mathbf{X}+b\mathbf{Y}} - a\mu_{\mathbf{X}} - b\mu_{\mathbf{Y}}$ and $\mu_{[\mathbf{X},\mathbf{Y}]} - \{\mu_{\mathbf{X}},\mu_{\mathbf{Y}}\}$ are constant. Using the freedom in the definition of each of the functions $\mu_{\mathbf{X}}$, they can always be chosen in such a way that $\mu_{a\mathbf{X}+b\mathbf{Y}} = a\mu_{\mathbf{X}} + b\mu_{\mathbf{Y}}$ for $\mathbf{X},\mathbf{Y} \in \mathfrak{g},\ a,b \in \mathbb{R}$. In fact, if $\{\mathbf{X}_1,\ldots,\mathbf{X}_m\}$ is a basis of $\mathfrak{g}$ and $\mu_{\mathbf{X}_i}$ is a function such that $\mathbf{X}_i^{+} \lrcorner \mathrm{d}\theta = -\mathrm{d}\mu_{\mathbf{X}_i}$, then for $\mathbf{X} \in \mathfrak{g}$, given by $\mathbf{X} = a^i \mathbf{X}_i$, we define $\mu_{\mathbf{X}}$ by $\mu_{\mathbf{X}} \equiv a^i \mu_{\mathbf{X}_i}$; then we have

$$\mathrm{d}\mu_{\mathbf{X}} = \mathrm{d}\big(a^i \mu_{\mathbf{X}_i}\big) = a^i\, \mathrm{d}\mu_{\mathbf{X}_i} = -a^i \mathbf{X}_i^{+} \lrcorner \mathrm{d}\theta = -\big(a^i \mathbf{X}_i\big)^{+} \lrcorner \mathrm{d}\theta = -\mathbf{X}^{+} \lrcorner \mathrm{d}\theta$$

and, as can readily be verified, $\mu_{a\mathbf{X}+b\mathbf{Y}} = a\mu_{\mathbf{X}} + b\mu_{\mathbf{Y}}$.

In what follows we shall assume that we have a set of functions $\mu_{\mathbf{X}}$ satisfying (8.64) with $\mu_{a\mathbf{X}+b\mathbf{Y}} - a\mu_{\mathbf{X}} - b\mu_{\mathbf{Y}}$ equal to zero. However, it will not always be possible to simultaneously make $\mu_{[\mathbf{X},\mathbf{Y}]} - \{\mu_{\mathbf{X}},\mu_{\mathbf{Y}}\}$ also equal to zero for all $\mathbf{X},\mathbf{Y} \in \mathfrak{g}$.

As pointed out already, the difference $\mu_{[\mathbf{X},\mathbf{Y}]} - \{\mu_{\mathbf{X}},\mu_{\mathbf{Y}}\}$ is a constant function whose value, denoted by $c(\mathbf{X},\mathbf{Y})$, depends on $\mathbf{X}$ and $\mathbf{Y}$ (hence, we can consider c as a real-valued function defined on $\mathfrak{g} \times \mathfrak{g}$, i.e., $c : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$). Then we have $c(\mathbf{X},\mathbf{Y}) = -c(\mathbf{Y},\mathbf{X})$ and since

$$\begin{aligned} \mu_{[a\mathbf{X}+b\mathbf{Y},\mathbf{Z}]} - \{\mu_{a\mathbf{X}+b\mathbf{Y}},\mu_{\mathbf{Z}}\} &= \mu_{a[\mathbf{X},\mathbf{Z}]+b[\mathbf{Y},\mathbf{Z}]} - \{a\mu_{\mathbf{X}} + b\mu_{\mathbf{Y}},\mu_{\mathbf{Z}}\} \\ &= a\mu_{[\mathbf{X},\mathbf{Z}]} + b\mu_{[\mathbf{Y},\mathbf{Z}]} - a\{\mu_{\mathbf{X}},\mu_{\mathbf{Z}}\} - b\{\mu_{\mathbf{Y}},\mu_{\mathbf{Z}}\}, \end{aligned}$$

it follows that $c(a\mathbf{X}+b\mathbf{Y},\mathbf{Z}) = a\,c(\mathbf{X},\mathbf{Z}) + b\,c(\mathbf{Y},\mathbf{Z})$, for $\mathbf{X},\mathbf{Y},\mathbf{Z} \in \mathfrak{g},\ a,b \in \mathbb{R}$. In other words, the map c, from $\mathfrak{g} \times \mathfrak{g}$ into $\mathbb{R}$, is skew-symmetric and bilinear. [In the language of cohomology of Lie algebras, c is a cochain; see, e.g., Jacobson (1979, Chap. III).]

Exercise 8.27 Show that $c([\mathbf{X},\mathbf{Y}],\mathbf{Z}) + c([\mathbf{Y},\mathbf{Z}],\mathbf{X}) + c([\mathbf{Z},\mathbf{X}],\mathbf{Y}) = 0$. (This means that c is a closed cochain.)

Theorem 8.28 *There exists a set of functions* $\mu'_{\mathbf{X}}$ *such that* $\mathrm{d}\mu'_{\mathbf{X}} = -\mathbf{X}^{+} \lrcorner\, \mathrm{d}\theta$, $\mu'_{a\mathbf{X}+b\mathbf{Y}} = a\mu'_{\mathbf{X}} + b\mu'_{\mathbf{Y}}$, *and* $\mu'_{[\mathbf{X},\mathbf{Y}]} = \{\mu'_{\mathbf{X}},\mu'_{\mathbf{Y}}\}$, *for* $\mathbf{X},\mathbf{Y} \in \mathfrak{g}$ *and* $a,b \in \mathbb{R}$, *if and only if there exists* $h \in \mathfrak{g}^*$ *such that* $c(\mathbf{X},\mathbf{Y}) = h([\mathbf{X},\mathbf{Y}])$.

Proof If such a set of functions $\mu'_{\mathbf{X}}$ exists, the condition $\mathrm{d}\mu'_{\mathbf{X}} = -\mathbf{X}^{+} \lrcorner \mathrm{d}\theta$ implies that the difference $\mu_{\mathbf{X}} - \mu'_{\mathbf{X}}$ is a constant whose value, denoted by $h(\mathbf{X})$, may depend on $\mathbf{X}$. Then

$$\begin{aligned} & h(a\mathbf{X}+b\mathbf{Y}) - ah(\mathbf{X}) - bh(\mathbf{Y}) \\ &\quad = \mu_{a\mathbf{X}+b\mathbf{Y}} - a\mu_{\mathbf{X}} - b\mu_{\mathbf{Y}} - \mu'_{a\mathbf{X}+b\mathbf{Y}} + a\mu'_{\mathbf{X}} + b\mu'_{\mathbf{Y}} \\ &\quad = 0, \end{aligned}$$

for $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$ and $a, b \in \mathbb{R}$; that is, h is linear; hence, $h \in \mathfrak{g}^*$ and

$$\begin{aligned} c(\mathbf{X}, \mathbf{Y}) &= \mu_{[\mathbf{X},\mathbf{Y}]} - \{\mu_{\mathbf{X}}, \mu_{\mathbf{Y}}\} \\ &= \mu'_{[\mathbf{X},\mathbf{Y}]} + h\big([\mathbf{X}, \mathbf{Y}]\big) - \big\{\mu'_{\mathbf{X}} + h(\mathbf{X}), \mu'_{\mathbf{Y}} + h(\mathbf{Y})\big\} \\ &= \mu'_{[\mathbf{X},\mathbf{Y}]} - \{\mu'_{\mathbf{X}}, \mu'_{\mathbf{Y}}\} + h\big([\mathbf{X}, \mathbf{Y}]\big) \\ &= h\big([\mathbf{X}, \mathbf{Y}]\big), \end{aligned}$$

for $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$.

Conversely, if there exists $h \in \mathfrak{g}^*$ such that $c(\mathbf{X}, \mathbf{Y}) = h([\mathbf{X}, \mathbf{Y}])$, we define $\mu'_{\mathbf{X}} \equiv \mu_{\mathbf{X}} - h(\mathbf{X})$, for $\mathbf{X} \in \mathfrak{g}$; since $h(\mathbf{X}) \in \mathbb{R}$ and h is linear we have $\mathrm{d}\mu'_{\mathbf{X}} = \mathrm{d}\mu_{\mathbf{X}} = -\mathbf{X}^+ \lrcorner \mathrm{d}\theta$. Furthermore,

$$\begin{aligned} &\mu'_{a\mathbf{X}+b\mathbf{Y}} - a\mu'_{\mathbf{X}} - b\mu'_{\mathbf{Y}} \\ &\quad = \mu_{a\mathbf{X}+b\mathbf{Y}} - a\mu_{\mathbf{X}} - b\mu_{\mathbf{Y}} - h(a\mathbf{X} + b\mathbf{Y}) + ah(\mathbf{X}) + bh(\mathbf{Y}) \\ &\quad = 0, \end{aligned}$$

and

$$\begin{aligned} \mu'_{[\mathbf{X},\mathbf{Y}]} - \{\mu'_{\mathbf{X}}, \mu'_{\mathbf{Y}}\} &= \mu_{[\mathbf{X},\mathbf{Y}]} - h\big([\mathbf{X}, \mathbf{Y}]\big) - \big\{\mu_{\mathbf{X}} - h(\mathbf{X}), \mu_{\mathbf{Y}} - h(\mathbf{Y})\big\} \\ &= c(\mathbf{X}, \mathbf{Y}) - h\big([\mathbf{X}, \mathbf{Y}]\big) \\ &= 0, \end{aligned}$$

for $\mathbf{X}, \mathbf{Y} \in \mathfrak{g}$ and $a, b \in \mathbb{R}$. □

A necessary condition for the existence of an $h \in \mathfrak{g}^*$ such that $c(\mathbf{X}, \mathbf{Y}) = h([\mathbf{X}, \mathbf{Y}])$ is obtained making use of the Jacobi identity and the linearity of h, i.e.,

$$\begin{aligned} &c\big([\mathbf{X}, \mathbf{Y}], \mathbf{Z}\big) + c\big([\mathbf{Y}, \mathbf{Z}], \mathbf{X}\big) + c\big([\mathbf{Z}, \mathbf{X}], \mathbf{Y}\big) \\ &\quad = h\Big(\big[[\mathbf{X}, \mathbf{Y}]\mathbf{Z}\big] + \big[[\mathbf{Y}, \mathbf{Z}], \mathbf{X}\big] + \big[[\mathbf{Z}, \mathbf{X}], \mathbf{Y}\big]\Big) \\ &\quad = 0, \end{aligned}$$

which is always satisfied (see Exercise 8.27). For some Lie algebras (e.g., the semisimple Lie algebras) this condition is also sufficient [see, e.g., Jacobson (1979, Chap. III).].

Example 8.29 Let G be the additive group $\mathbb{R}^2$ and let $M = \mathbb{R}^2$. For each $g = (a, b) \in \mathbb{R}^2$ we define $R_g : T^*M \to T^*M$ by

$$\begin{aligned} {R_g}^* q^1 &= q^1 + a + Kbp_2, \\ {R_g}^* q^2 &= q^2 + K\left(bp_1 - \frac{1}{2}b^2\right), \\ {R_g}^* p_1 &= p_1 - b, \\ {R_g}^* p_2 &= p_2, \end{aligned} \tag{8.65}$$

where (q^1, q^2, p_1, p_2) are the canonical coordinates on T^*M associated with the natural coordinates of $\mathbb{R}^2$, and K is an arbitrary real constant. These expressions define an action of G on T^*M such that each R_g is a canonical transformation. In fact,

$$\begin{aligned} R_g{}^*\big(\mathrm{d}p_1 \wedge \mathrm{d}q^1 + \mathrm{d}p_2 \wedge \mathrm{d}q^2\big) &= \mathrm{d}p_1 \wedge \big(\mathrm{d}q^1 + Kb\,\mathrm{d}p_2\big) + \mathrm{d}p_2 \wedge \mathrm{d}\big(q^2 + Kb\,\mathrm{d}p_1\big) \\ &= \mathrm{d}p_1 \wedge \mathrm{d}q^1 + \mathrm{d}p_2 \wedge \mathrm{d}q^2. \end{aligned}$$

In order to find the vector fields induced on T^*M by the action of G, we note that, according to (8.65), for $\alpha_p \in T^*M$, the mapping $\Phi_{\alpha_p} : G \to T^*M$, defined by $\Phi_{\alpha_p}(g) \equiv R_g(\alpha_p)$, is given by the expressions

$$\begin{aligned} \Phi_{\alpha_p}{}^* q^1 &= q^1(\alpha_p) + x^1 + Kp_2(\alpha_p)x^2, \\ \Phi_{\alpha_p}{}^* q^2 &= q^2(\alpha_p) + K\left(p_1(\alpha_p)x^2 - \frac{1}{2}\big(x^2\big)^2\right), \\ \Phi_{\alpha_p}{}^* p_1 &= p_1(\alpha_p) - x^2, \\ \Phi_{\alpha_p}{}^* p_2 &= p_2(\alpha_p), \end{aligned} \tag{8.66}$$

where (x^1, x^2) are the natural coordinates on G, that is, if $g = (a, b)$, then $x^1(g) = a$ and $x^2(g) = b$. Making use of (7.60) and (8.66) we find that

$$\begin{aligned} \Phi_{\alpha_p *e}\left(\frac{\partial}{\partial x^1}\right)_e &= \left(\frac{\partial}{\partial q^1}\right)_{\alpha_p}, \\ \Phi_{\alpha_p *e}\left(\frac{\partial}{\partial x^2}\right)_e &= Kp_2(\alpha_p)\left(\frac{\partial}{\partial q^1}\right)_{\alpha_p} + Kp_1(\alpha_p)\left(\frac{\partial}{\partial q^2}\right)_{\alpha_p} - \left(\frac{\partial}{\partial p_1}\right)_{\alpha_p}. \end{aligned}$$

Thus, if $\mathbf{X}_1$ and $\mathbf{X}_2$ are the left-invariant vector fields on G such that $(\mathbf{X}_i)_e = (\partial/\partial x^i)_e$, the corresponding vector fields on T^*M are

$$\mathbf{X}_1^+ = \frac{\partial}{\partial q^1}, \qquad \mathbf{X}_2^+ = Kp_2\frac{\partial}{\partial q^1} + Kp_1\frac{\partial}{\partial q^2} - \frac{\partial}{\partial p_1}.$$

These vector fields turn out to be globally Hamiltonian; indeed, $\mathbf{X}_1^+ \lrcorner\, \mathrm{d}\theta = -\mathrm{d}p_1$ and $\mathbf{X}_2^+ \lrcorner\, \mathrm{d}\theta = -\mathrm{d}(q^1 + Kp_1p_2)$. Hence,

$$\mu_{\mathbf{X}_1} = p_1 + \text{const}, \qquad \mu_{\mathbf{X}_2} = q^1 + Kp_1p_2 + \text{const},$$

and one finds that $\{\mu_{\mathbf{X}_1}, \mu_{\mathbf{X}_2}\} = \mathbf{X}_1^+\mu_{\mathbf{X}_2} = -1$. (Note that $\mu_{\mathbf{X}_2}$ is not a homogeneous function of degree 1 in the p_i and, therefore, it cannot come from a group of canonical transformations on T^*M induced by a group of transformations on M.) However, $[\mathbf{X}_1, \mathbf{X}_2] = 0$ (G is an Abelian group) so that if we want to have a linear mapping $\mathbf{X} \mapsto \mu_{\mathbf{X}}$, then $\mu_{[\mathbf{X}_1,\mathbf{X}_2]} = 0$, which cannot coincide with $\{\mu_{\mathbf{X}_1}, \mu_{\mathbf{X}_2}\}$, no matter how we choose the arbitrary constants contained in $\mu_{\mathbf{X}_1}$ and $\mu_{\mathbf{X}_2}$. (Note that $c(\mathbf{X}_1, \mathbf{X}_2) = \mu_{[\mathbf{X}_1,\mathbf{X}_2]} - \{\mu_{\mathbf{X}_1}, \mu_{\mathbf{X}_2}\} = 1$, but, since $[\mathbf{X}_1, \mathbf{X}_2] = 0$, there does not exist $h \in \mathfrak{g}^*$ such that $c(\mathbf{X}_1, \mathbf{X}_2) = h([\mathbf{X}_1, \mathbf{X}_2])$, in accordance with Theorem 8.28.)

The Hamiltonian function

$$H = \frac{{p_1}^2 + {p_2}^2}{2m} + \frac{1}{mK} q^2,$$

which corresponds to a particle of mass m in a uniform field (e.g., a uniform gravitational field with acceleration $1/m^2 K$), is invariant under the action (8.65). [In fact, one can readily verify that $\mathbf{X}_i^+ p_2 = 0$ and $\mathbf{X}_i^+ (q^2 + \frac{1}{2} K {p_1}^2) = 0$, for $i = 1, 2$; therefore any function of only p_2 and $q^2 + \frac{1}{2} K {p_1}^2$ is invariant under the action (8.65).]

Example 8.30 Now we shall start by specifying a Hamiltonian function and we shall find a group of canonical transformations that leave the Hamiltonian invariant. Taking $M = \mathbb{R}^3$, we shall consider the Hamiltonian function

$$H = \frac{1}{2m} \left({p_1}^2 + {p_2}^2 + {p_3}^2 \right), \tag{8.67}$$

which corresponds to a particle of mass m and electric charge e in a uniform magnetic field $\mathbf{B} = B\,(\partial/\partial x^3)$, where B is a constant, provided that we use the symplectic 2-form $\omega = \mathrm{d}p_i \wedge \mathrm{d}q^i + (eB/c)\, \mathrm{d}q^1 \wedge \mathrm{d}q^2$, on T^*M; $(q^1, q^2, q^3, p_1, p_2, p_3)$ are the coordinates on T^*M induced by the natural coordinates of $\mathbb{R}^3$, and c is the speed of light in vacuum (see Example 8.19).

The vector fields $\partial/\partial q^i$, $i = 1, 2, 3$, on T^*M satisfy $\partial H/\partial q^i = 0$ and are globally Hamiltonian since

$$\frac{\partial}{\partial q^1} \lrcorner \omega = -\mathrm{d}\left(p_1 - \frac{eB}{c} q^2 \right),$$

$$\frac{\partial}{\partial q^2} \lrcorner \omega = -\mathrm{d}\left(p_2 + \frac{eB}{c} q^1 \right),$$

$$\frac{\partial}{\partial q^3} \lrcorner \omega = -\mathrm{d}p_3.$$

Hence, the functions

$$K_1 \equiv p_1 - \frac{eB}{c} q^2, \qquad K_2 \equiv p_2 + \frac{eB}{c} q^1, \qquad K_3 \equiv p_3$$

are constants of motion. (Note that K_1 and K_2 are not homogeneous functions of degree 1 in the p_i.) The (globally) Hamiltonian vector fields corresponding to them are

$$\mathbf{X}_{\mathrm{d}K_1} = \frac{\partial}{\partial q^1}, \qquad \mathbf{X}_{\mathrm{d}K_2} = \frac{\partial}{\partial q^2}, \qquad \mathbf{X}_{\mathrm{d}K_3} = \frac{\partial}{\partial q^3};$$

therefore [see (8.19)]

$$\{K_1, K_2\} = \mathbf{X}_{\mathrm{d}K_1} K_2 = \frac{eB}{c},$$

$$\{K_2, K_3\} = \mathbf{X}_{\mathrm{d}K_2} K_3 = 0,$$
$$\{K_3, K_1\} = \mathbf{X}_{\mathrm{d}K_3} K_1 = 0,$$

which implies that the Lie brackets of the vector fields $\mathbf{X}_{\mathrm{d}K_i}$ are all equal to zero [see (8.20)].

Proceeding as in Sect. 6.1, we can find the one-parameter group of diffeomorphisms, φ_s, generated by an arbitrary linear combination

$$a^i\,\mathbf{X}_{\mathrm{d}K_i} = a^1 \frac{\partial}{\partial q^1} + a^2 \frac{\partial}{\partial q^2} + a^3 \frac{\partial}{\partial q^3}.$$

The result can be expressed in the form $\varphi_s{}^* q^i = q^i + a^i s$, $\varphi_s{}^* p_i = p_i$ $(i = 1, 2, 3)$. Thus, the vector fields $\mathbf{X}_{\mathrm{d}K_i}$ are induced by the action of the additive group $\mathbb{R}^3$ on T^*M given by

$$R_g{}^* q^i = q^i + a^i, \qquad R_g{}^* p_i = p_i \quad (i = 1, 2, 3),$$

for $g = (a^1, a^2, a^3) \in \mathbb{R}^3$. One can readily verify that these transformations are canonical (in fact, $R_g{}^*\,\mathrm{d}q^i = \mathrm{d}q^i$ and $R_g{}^*\,\mathrm{d}p_i = \mathrm{d}p_i$; hence $R_g{}^*\omega = \omega$), give an action of $\mathbb{R}^3$ on T^*M, and leave invariant the Hamiltonian (8.67). As in Example 8.29, if $B \neq 0$, it is impossible to find a Lie algebra homomorphism from the *Abelian Lie algebra* of $\mathbb{R}^3$ into $C^\infty(T^*M)$, associated with this action.

It may be noticed that the Hamiltonian (8.67) also satisfies

$$\left(q^1 \frac{\partial}{\partial q^2} - q^2 \frac{\partial}{\partial q^1} + p_1 \frac{\partial}{\partial p_2} - p_2 \frac{\partial}{\partial p_1}\right) H = 0.$$

The vector field appearing on the left-hand side of the last equation is the canonical lift of the infinitesimal generator of rotations about the x^3 axis (see Exercise 8.15) and is globally Hamiltonian

$$\begin{aligned}&\left(q^1 \frac{\partial}{\partial q^2} - q^2 \frac{\partial}{\partial q^1} + p_1 \frac{\partial}{\partial p_2} - p_2 \frac{\partial}{\partial p_1}\right) \lrcorner \left(\mathrm{d}p_i \wedge \mathrm{d}q^i + \frac{eB}{c}\,\mathrm{d}q^1 \wedge \mathrm{d}q^2\right)\\ &\quad = -\mathrm{d}\big\{q^1 p_2 - q^2 p_1 + (eB/2c)\big[\big(q^1\big)^2 + \big(q^2\big)^2\big]\big\}.\end{aligned}$$

Thus,

$$\begin{aligned}L_3 &\equiv q^1 p_2 - q^2 p_1 + \frac{eB}{2c}\big[\big(q^1\big)^2 + \big(q^2\big)^2\big]\\ &= q^1\left(p_2 + \frac{eB}{2c} q^1\right) - q^2\left(p_1 - \frac{eB}{2c} q^2\right)\end{aligned}$$

is a constant of motion. Note that, as shown in Example 8.20, $(q^1, q^2, q^3, p_1 - \frac{eB}{2c} q^2, p_2 + \frac{eB}{2c} q^1, p_3)$ is a set of canonical coordinates for the symplectic structure considered here, but the constants of motion K_1 and K_2 *do not* coincide with $p_1 - \frac{eB}{2c} q^2$ and $p_2 + \frac{eB}{2c} q^1$, respectively, for $B \neq 0$.

Exercise 8.31 Show that

$$\{L_3, K_1\} = -K_2, \qquad \{L_3, K_2\} = K_1, \qquad \{L_3, K_3\} = 0.$$

Example 8.32 The Hamiltonian function

$$H = \frac{1}{2m}\sum_{i=1}^{n}(p_i)^2 + \frac{m{\omega_0}^2}{2}\sum_{i=1}^{n}\left(q^i\right)^2 \tag{8.68}$$

corresponds to an isotropic harmonic oscillator in n dimensions; m is the mass of the oscillator and ω_0 is its angular frequency. The q^i and p_i appearing in (8.68) are the canonical coordinates induced by a set of Cartesian coordinates x^i on the n-dimensional Euclidean space. Defining the complex (row) vector

$$\mathbf{b} \equiv \left(-\mathrm{i}p_1 + m\omega_0 q^1, \ldots, -\mathrm{i}p_n + m\omega_0 q^n\right),$$

the Hamiltonian function (8.68) can be expressed in the form

$$H = \frac{1}{2m}\mathbf{b}\mathbf{b}^\dagger, \tag{8.69}$$

where $\mathbf{b}^\dagger$ is the Hermitian adjoint of $\mathbf{b}$ (obtained transposing and conjugating the row $\mathbf{b}$).

Let $\mathrm{SU}(n)$ be the group of unitary complex $n \times n$ matrices with determinant equal to 1; then for $U \in \mathrm{SU}(n)$ the Hamiltonian function (8.69) is invariant under the transformation

$$\mathbf{b} \mapsto \mathbf{b}U, \tag{8.70}$$

since by virtue of the unitarity of U, we have $\mathbf{b}\mathbf{b}^\dagger \mapsto \mathbf{b}U(\mathbf{b}U)^\dagger = \mathbf{b}UU^\dagger\mathbf{b}^\dagger = \mathbf{b}\mathbf{b}^\dagger$. Furthermore, for each $U \in \mathrm{SU}(n)$, the transformation (8.70) is canonical as can be seen noting that

$$p_i = \mathrm{i}\frac{b_i - \overline{b_i}}{2}, \qquad q^i = \frac{b_i + \overline{b_i}}{2m\omega_0}, \tag{8.71}$$

where the b_i are the components of $\mathbf{b}$, the bar denotes complex conjugation, and

$$\mathrm{d}p_i \wedge \mathrm{d}q^i = \frac{\mathrm{i}}{4m\omega_0}\sum_{i=1}^{n}(\mathrm{d}b_i - \mathrm{d}\overline{b_i}) \wedge (\mathrm{d}b_i + \mathrm{d}\overline{b_i}) = \frac{\mathrm{i}}{2m\omega_0}\sum_{i=1}^{n}\mathrm{d}b_i \wedge \mathrm{d}\overline{b_i}.$$

Equation (8.70) amounts to $R_U{}^*b_i = b_j U_i^j$, where $U = (U_i^j)$, so that

$$R_U{}^*\left(\mathrm{d}p_i \wedge \mathrm{d}q^i\right) = \frac{\mathrm{i}}{2m\omega_0}\sum_{i=1}^{n}R_U{}^*\,\mathrm{d}b_i \wedge R_U{}^*\,\mathrm{d}\overline{b_i} = \frac{\mathrm{i}}{2m\omega_0}\sum_{i=1}^{n}U_i^j\,\overline{U_i^k}\,\mathrm{d}b_j \wedge \mathrm{d}\overline{b_k}$$

$$= \frac{\mathrm{i}}{2m\omega_0}\sum_{j=1}^{n}\mathrm{d}b_j \wedge \mathrm{d}\overline{b_j} = \mathrm{d}p_i \wedge \mathrm{d}q^i.$$

(As in previous examples, the use of complex quantities, such as the b_i, simplifies the computations, but is not essential.)

Thus, SU(n) acts on the right on the phase space by means of canonical transformations and, therefore, the fields $\mathbf{X}_i^+$ on the phase space, induced by this action, are Hamiltonian, at least locally. In what follows we shall consider in more detail the case with $n = 2$, showing that the vector fields $\mathbf{X}^+$ are actually globally Hamiltonian and that, in contrast to Examples 8.29 and 8.30, the functions $\mu_{\mathbf{X}}$ can be chosen in such a way that the map $\mathbf{X} \mapsto \mu_{\mathbf{X}}$ is a homomorphism of Lie algebras.

Substituting the matrix $U = \exp ta^i \mathbf{X}_i$ given by (7.54) into (8.70) and using (8.71) we find that (cf. Example 7.59)

$$R_{\exp ta^i \mathbf{X}_i}{}^* q^1 = q^1 \cos(Kt/2) + \left[a^2 q^2 + \frac{1}{m\omega_0}\left(a^3 p_1 + a^1 p_2\right)\right] \frac{1}{K} \sin(Kt/2),$$

$$R_{\exp ta^i \mathbf{X}_i}{}^* q^2 = q^2 \cos(Kt/2) - \left[a^2 q^1 - \frac{1}{m\omega_0}\left(a^1 p_1 - a^3 p_2\right)\right] \frac{1}{K} \sin(Kt/2),$$

$$R_{\exp ta^i \mathbf{X}_i}{}^* p_1 = p_1 \cos(Kt/2) + \left[a^2 p_2 - m\omega_0\left(a^3 q^1 + a^1 q^2\right)\right] \frac{1}{K} \sin(Kt/2),$$

$$R_{\exp ta^i \mathbf{X}_i}{}^* p_2 = p_2 \cos(Kt/2) - \left[a^2 p_1 + m\omega_0\left(a^1 q^1 - a^3 q^2\right)\right] \frac{1}{K} \sin(Kt/2),$$

and calculating the tangent vector to the curve given by these expressions at $t = 0$ we obtain the vector field

$$\begin{aligned} \left(a^i \mathbf{X}_i\right)^+ = {} & \frac{1}{2}\left[a^2 q^2 + \frac{a^3 p_1 + a^1 p_2}{m\omega_0}\right] \frac{\partial}{\partial q^1} + \frac{1}{2}\left[-a^2 q^1 + \frac{a^1 p_1 - a^3 p_2}{m\omega_0}\right] \frac{\partial}{\partial q^2} \\ & + \frac{1}{2}\left[a^2 p_2 - m\omega_0\left(a^3 q^1 + a^1 q^2\right)\right] \frac{\partial}{\partial p_1} \\ & + \frac{1}{2}\left[-a^2 p_1 - m\omega_0\left(a^1 q^1 - a^3 q^2\right)\right] \frac{\partial}{\partial p_2}, \end{aligned} \tag{8.72}$$

which is globally Hamiltonian; its contraction with $\mathrm{d}\theta$ gives $-\mathrm{d}(a^i \mu_{\mathbf{X}_i})$, where

$$\begin{aligned} \mu_{\mathbf{X}_1} &\equiv \frac{1}{2m\omega_0}\left(p_1 p_2 + m^2 {\omega_0}^2 q^1 q^2\right), \\ \mu_{\mathbf{X}_2} &\equiv \frac{1}{2}\left(p_1 q^2 - p_2 q^1\right), \\ \mu_{\mathbf{X}_3} &\equiv \frac{1}{4m\omega_0}\left\{(p_1)^2 - (p_2)^2 + m^2 {\omega_0}^2\left[\left(q^1\right)^2 - \left(q^2\right)^2\right]\right\}. \end{aligned} \tag{8.73}$$

Recall that the functions $\mu_{\mathbf{X}_i}$ are not uniquely defined by (8.72); as we shall show below, with the choice (8.73) one obtains a Lie algebra homomorphism. Note also that, out of these three constants of motion, only $\mu_{\mathbf{X}_2}$ is a homogeneous function of degree 1 of the p_i, and therefore it is the only one associated with a group of transformations acting on the configuration space; see (8.26).

It can readily be verified directly that these functions satisfy the relations

$$\{\mu_{\mathbf{X}_i}, \mu_{\mathbf{X}_j}\} = \sum_{k=1}^{3} \varepsilon_{ijk}\, \mu_{\mathbf{X}_k}, \tag{8.74}$$

which correspond to the relations $[\mathbf{X}_i, \mathbf{X}_j] = \sum_{k=1}^{3} \varepsilon_{ijk} \mathbf{X}_k$ satisfied by the basis of $\mathfrak{su}(2)$ given by (7.21).

A convenient way of calculating the Poisson bracket on the left-hand side of (8.74) consists of employing the definition (8.19), which yields $\{\mu_{\mathbf{X}_i}, \mu_{\mathbf{X}_j}\} = \mathbf{X}_i^+ \mu_{\mathbf{X}_j}$, noting that the vector field $\mathbf{X}_i^+$ is the coefficient of a^i on the right-hand side of (8.72). For instance, from (8.72) and (8.73) we obtain

$$\begin{aligned}
&\{\mu_{\mathbf{X}_1}, \mu_{\mathbf{X}_2}\} \\
&\quad = \frac{1}{2}\left(\frac{p_2}{m\omega_0}\frac{\partial}{\partial q^1} + \frac{p_1}{m\omega_0}\frac{\partial}{\partial q^2} - m\omega_0 q^2 \frac{\partial}{\partial p_1} - m\omega_0 q^1 \frac{\partial}{\partial p_2}\right)\frac{1}{2}\big(p_1 q^2 - p_2 q^1\big) \\
&\quad = \mu_{\mathbf{X}_3}.
\end{aligned}$$

The results established in the preceding paragraphs, in connection with Lie groups that act on the cotangent bundle of a manifold by means of canonical transformations, also apply if in place of the cotangent bundle of a manifold one considers any symplectic manifold, replacing the fundamental 2-form $\mathrm{d}\theta$ by the corresponding symplectic form.

Example 8.33 The rotations about the origin in $\mathbb{R}^3$, which form the group SO(3), leave invariant the sphere S^2 as well as its area element, which will be denoted by ω. The 2-form ω defines a symplectic structure for S^2 (see Example 8.18) and, by virtue of the invariances already mentioned, the vector fields $\mathbf{X}^+$ induced by the action of SO(3) on $\mathbb{R}^3$ are tangent to S^2 and are, at least locally, Hamiltonian. In fact, expressing the vector fields $\mathbf{S}_k^+$, given in Example 7.58, in terms of the spherical coordinates one finds that

$$\begin{aligned}
\mathbf{S}_1^+ &= \sin\phi \frac{\partial}{\partial\theta} + \cot\theta \cos\phi \frac{\partial}{\partial\phi}, \\
\mathbf{S}_2^+ &= -\cos\phi \frac{\partial}{\partial\theta} + \cot\theta \sin\phi \frac{\partial}{\partial\phi}, \\
\mathbf{S}_3^+ &= -\frac{\partial}{\partial\phi},
\end{aligned} \tag{8.75}$$

which shows that these vector fields are tangent to the sphere and contracting them with $\omega = \sin\theta\, \mathrm{d}\theta \wedge \mathrm{d}\phi$, we obtain $-\mathrm{d}(\sin\theta\cos\phi)$, $-\mathrm{d}(\sin\theta\sin\phi)$, and $-\mathrm{d}\cos\theta$, respectively [cf. (8.49)], thus showing explicitly that the $\mathbf{S}_k^+$ are locally Hamiltonian. [Since the spherical coordinates are not defined globally on S^2, from the previous computations we cannot conclude that the $\mathbf{S}_k^+$ are globally Hamiltonian. For instance, in the domain of the spherical coordinates, $\omega = \mathrm{d}(\phi\, \mathrm{d}\cos\theta)$, but ω is not an exact 2-form.]

In the present case, from (7.63) it follows that $S_k = \frac{1}{2}\sum_{i,j=1}^{3} \varepsilon_{kij}[S_i, S_j]$, therefore, $\mathbf{S}_k^+ = \frac{1}{2}\sum_{i,j=1}^{3} \varepsilon_{kij}[\mathbf{S}_i^+, \mathbf{S}_j^+]$, and since according to Theorem 8.7 the Lie bracket of two locally Hamiltonian vector fields is globally Hamiltonian, the vector fields $\mathbf{S}_k^+$ are globally Hamiltonian. It may be noticed that $\mathbf{S}_k^+ \lrcorner\, \omega = -\mathrm{d}\hat{x}^k$, where $\hat{x}^k$ is the restriction of the Cartesian coordinate x^k to S^2 (that is, $\hat{x}^k = i^* x^k$, where $i : \mathrm{S}^2 \to \mathbb{R}^3$ is the inclusion map). Finally, making use of the expressions (8.75) one finds that

$$\{\hat{x}^i, \hat{x}^j\} = \mathbf{S}_i^+ \hat{x}^j = \varepsilon_{ijk}\hat{x}^k.$$

Example 8.34 The so-called Kepler problem corresponds to the motion of a particle in a central force field with potential energy $V = -k/r$, where k is a positive constant and r is the distance from the particle to the center of force. Assuming that the motion of the particle takes place in the three-dimensional Euclidean space, the Hamiltonian function expressed in terms of the canonical coordinates induced by a set of Cartesian coordinates is

$$H = \frac{1}{2M}\left({p_1}^2 + {p_2}^2 + {p_3}^2\right) - \frac{k}{\sqrt{(q^1)^2 + (q^2)^2 + (q^3)^2}}. \tag{8.76}$$

The Hamiltonian (8.76) is invariant under the transformations on the phase space induced by the rotations about the origin in the Euclidean space, which implies the conservation of the angular momentum, $L_i = \varepsilon_{ijk} q^j p_k$, with summation over repeated indices (see Exercise 8.15). But, as is well known, the so-called Runge–Lenz vector

$$\mathbf{A} \equiv \mathbf{p} \times (\mathbf{r} \times \mathbf{p}) - \frac{mk}{r}\mathbf{r}, \tag{8.77}$$

where $\mathbf{r}$ is the position vector of the particle, is also a constant of motion (that is, the functions $A_i = p_j p_j q^i - p_j q^j p_i - mkq^i/\sqrt{q^k q^k}$ are constants of motion). Since the A_i are not homogeneous functions of first degree in the p_j, the existence of these constants of motion is not related to transformation groups acting on the configuration space $M = \mathbb{E}^3$.

Of course, in principle we can find the (possibly local) one-parameter group of transformations generated by each vector field $\mathbf{X}_{\mathrm{d}A_i}$, which must be formed by canonical transformations that leave H invariant. However, it is possible to relate the Kepler problem with other problems in such a way that the conservation of the A_i becomes obvious. In this example we shall restrict ourselves to the trajectories in phase space on which H has the constant value $E = 0$. In order to identify the canonical transformations associated with the conservation of the A_i, we introduce the new coordinates

$$Q^i = a_0 \frac{p_i}{p_j p_j}, \quad i = 1, 2, 3,$$

where a_0 is a constant with dimensions of linear momentum times length. Then we obtain $p_i = a_0 Q^i/(Q^j Q^j)$, and a straightforward computation shows that

$$\begin{aligned}
p_i\,\mathrm{d}q^i &= a_0 \frac{Q^i}{Q^j Q^j}\,\mathrm{d}q^i \\
&= \mathrm{d}\left(a_0 \frac{Q^i q^i}{Q^j Q^j}\right) - \frac{a_0}{Q^j Q^j} q^i\,\mathrm{d}Q^i + a_0 Q^i q^i \frac{2Q^k\,\mathrm{d}Q^k}{(Q^j Q^j)^2} \\
&= \mathrm{d}\left(a_0 \frac{Q^i q^i}{Q^j Q^j}\right) + P_i\,\mathrm{d}Q^i,
\end{aligned}$$

where

$$P_i \equiv a_0 \frac{2Q^k q^k Q^i - Q^k Q^k q^i}{(Q^j Q^j)^2}$$

so that (Q^i, P_i) are canonical coordinates [cf. (8.12)]. Then we have

$$P_i P_i = \frac{{a_0}^2 q^i q^i}{(Q^j Q^j)^2}.$$

By combining the foregoing expressions we also have

$$P_i = \frac{1}{a_0}\left(2p_k q^k p_i - p_k p_k q^i\right). \tag{8.78}$$

We now introduce the auxiliary Hamiltonian

$$h \equiv 2mk^2 {a_0}^2 \left({a_0}^2 - 2m Q^i Q^i H\right)^{-2}$$

which satisfies

$$\mathrm{d}h = 8m^2 k^2 {a_0}^2 \left({a_0}^2 - 2m Q^i Q^i H\right)^{-3} \left(Q^i Q^i\,\mathrm{d}H + 2H Q^i\,\mathrm{d}Q^i\right).$$

Hence, on the hypersurface $H = 0$,

$$\mathrm{d}h|_{H=0} = \frac{8m^2k^2}{{a_0}^4} Q^i Q^i\,\mathrm{d}H\Big|_{H=0} = \frac{4mk}{{a_0}^2} r\,\mathrm{d}H\Big|_{H=0},$$

which means that on this hypersurface the integral curves of $\mathbf{X}_{\mathrm{d}h}$ only differ in parametrization from those of $\mathbf{X}_{\mathrm{d}H}$.

In terms of the new canonical coordinates, the auxiliary Hamiltonian h is given by

$$h = \frac{1}{2m} P_i P_i, \tag{8.79}$$

which has the form of the usual Hamiltonian for a free particle of mass m moving in the three-dimensional Euclidean space [cf. (8.49)] and therefore is invariant under a group of canonical transformations isomorphic to the group of rigid motions of $\mathbb{E}^3$. Equivalently, the six functions P_i and $\varepsilon_{ijk} Q^j P_k$ are constants of the motion (but only along the integral curves of $\mathbf{X}_{\mathrm{d}H}$ lying on the hypersurface $H = 0$). Making use of (8.78) one readily verifies that, on the hypersurface $H = 0$, we have $P_i = -2A_i/a_0$ and $\varepsilon_{ijk} Q^j P_k = \varepsilon_{ijk} q^j p_k$, thus explaining the conservation of angular momentum and of the Runge–Lenz vector.

A problem in geometrical optics closely related to the Kepler problem of classical mechanics is the so-called *Maxwell's fish-eye*, which is characterized by a refractive index of the form

$$n = \frac{a}{b + r^2}, \tag{8.80}$$

where a and b are real constants (with dimensions of length squared) and r is the distance from a given point O. (As usual in this context, we assume that light propagates in three-dimensional Euclidean space.) The spherical symmetry of the function (8.80) implies that the corresponding Hamiltonian (8.55) is invariant under the canonical transformations induced on $T^*\mathbb{E}^3$ by the rigid rotations about O. This invariance leads to the conservation of the components of the "angular momentum" $L_i = \varepsilon_{ijk} q^j p_k$, where the q^i, p_i are the coordinates induced by a Cartesian coordinate system with origin O. (In fact, the L_i are conserved if the refractive index is any function of r only.)

Exercise 8.35 Show that the specific form of the refractive index (8.80) implies that the Cartesian components of the vector

$$\mathbf{r} \times (\mathbf{p} \times \mathbf{r}) - \frac{a}{2} \frac{\mathbf{p}}{\sqrt{\mathbf{p} \cdot \mathbf{p}}} \tag{8.81}$$

[cf. (8.77)] are also conserved, that is, the functions $q^j q^j p_i - q^j p_j q_i - a p_i/(2\sqrt{p_j p_j})$ are constants of motion.

Making use of the conservation of the vector (8.81) one can readily show that the vector $\mathbf{p}$ traces a conic with one of its foci at the origin and that the light rays are circles or arcs of circles. In a similar manner, making use of conservation of the Laplace–Runge–Lenz vector (8.77), in the case of the Kepler problem one finds that the orbits are conics with one of the foci at the origin and the momentum traces circles or arcs of circles [see, e.g., Goldstein (1980, Chap. 3)].

8.6 The Rigid Body and the Euler Equations

A nice application of the formalism developed in this chapter and the previous ones is found in the study of the rigid body motion. As we shall show, by restricting ourselves to the motion of a rigid body with a fixed point, the configuration space can be identified with the group of rotations in the three-dimensional Euclidean space, SO(3).

In order to study the motion of a rigid body with a fixed point, it is convenient to consider an orthonormal basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ fixed in the body, with the orientation of the canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ of $\mathbb{R}^3$. Then, the configuration of the rigid body can be represented by means of a real 3×3 matrix whose *columns* are the components of $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ with respect to the canonical basis. This matrix is orthogonal, as a consequence of the fact that the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$ is orthonormal, and its determinant is equal to 1, by virtue of the assumed orientation of the basis. In this manner we have a one-to-one correspondence between the configurations of the rigid body and the orthogonal 3×3 matrices with determinant 1; thus, the configuration space of a rigid body with a fixed point can be identified with the underlying manifold of the group SO(3).

Using the definitions given in Example 7.58, one finds, for instance, that

$$\exp tS_3 = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

as can readily be verified by noting that the matrices

$$\gamma_t \equiv \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

form a one-parameter subgroup of $\mathrm{GL}(3,\mathbb{R})$ and calculating γ'_0 one obtains S_3; therefore $\gamma_t = \exp tS_3$ (see Sect. 7.4).

According to the definition given above, if the rigid body is initially at the configuration represented by $g \in \mathrm{SO}(3)$, then $(\exp tS_3)g$ represents the configuration obtained by rotating the body about the $\mathbf{e}_3$ axis through an angle t. Note that if the configuration of the rigid body is represented by the matrix whose *rows* are the components of $\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$ with respect to the canonical basis, then the configuration obtained by rotating the body about the $\mathbf{e}_3$ axis through an angle t corresponds to $g \exp(-tS_3)$.

According to the results established in the proof of Theorem 7.48, the infinitesimal generator of the one-parameter group of transformations $\varphi_t(g) = (\exp tS_3)g$ is the right-invariant vector field whose value at the identity corresponds to the matrix S_3, which will be denoted by $\dot{\mathbf{S}}_3$. Hence, $\dot{\mathbf{S}}_3$ is the infinitesimal generator of rotations of the rigid body about the $\mathbf{e}_3$ axis. In a similar way, the right-invariant vector field $\dot{\mathbf{S}}_k$, whose value at the identity corresponds to the matrix S_k, is the infinitesimal generator of the rotations of the body about the $\mathbf{e}_k$ axis.

On the other hand, for $g \in \mathrm{SO}(3)$, the matrix $g(\exp tS_k)$ corresponds to the configuration of the rigid body that, being originally in the configuration represented by g, has been rotated through an angle t about the $\mathbf{e}'_k$ axis fixed in the body. This implies that the left-invariant vector field $\mathbf{S}_k$, whose value at the identity corresponds to the matrix S_k, is the infinitesimal generator of rotations about the $\mathbf{e}'_k$ axis.

The vector fields $\mathbf{S}_k$ and $\dot{\mathbf{S}}_k$ on the configuration space $M = \mathrm{SO}(3)$ define Hamiltonian vector fields $\overline{\mathbf{S}_k}$ and $\overline{\dot{\mathbf{S}}_k}$ on the phase space $T^*\mathrm{SO}(3)$ (their canonical lifts), which correspond to the functions

$$L_i \equiv \overline{\mathbf{S}_i} \lrcorner \theta \quad \text{and} \quad K_i \equiv \overline{\dot{\mathbf{S}}_i} \lrcorner \theta, \tag{8.82}$$

respectively [see (8.24)]. In terms of the notation used in (8.24), $L_i = f_{\mathbf{S}_i}$ and $K_i = f_{\dot{\mathbf{S}}_i}$. Then, according to (8.30) and (7.63), the Poisson brackets for the functions L_i are given by

$$\{L_i, L_j\} = \{f_{\mathbf{S}_i}, f_{\mathbf{S}_j}\} = f_{[\mathbf{S}_i, \mathbf{S}_j]} = \sum_{k=1}^{3} \varepsilon_{ijk} f_{\mathbf{S}_k} = \sum_{k=1}^{3} \varepsilon_{ijk} L_k. \tag{8.83}$$

Since the left-invariant vector field $\mathbf{S}_i$ is the infinitesimal generator of rotations about the $\mathbf{e}'_i$ axis fixed in the body, the function L_i corresponds to the ith component of the angular momentum of the rigid body with respect to the axes fixed in the body. Similarly, the function K_i corresponds to the ith component of the angular momentum of the rigid body with respect to the canonical basis of $\mathbb{R}^3$ (the axes "fixed in space"). From Theorem 7.48 and (7.63) it follows that $[\dot{\mathbf{S}}_i, \dot{\mathbf{S}}_j] = -\sum_{k=1}^{3} \varepsilon_{ijk} \dot{\mathbf{S}}_k$, and using again (8.30) it follows that the Poisson brackets for the functions K_i are given by

$$\{K_i, K_j\} = -\sum_{k=1}^{3} \varepsilon_{ijk} K_k. \tag{8.84}$$

Finally, from Theorem 7.49 we see that the Lie bracket of each of the vector fields $\mathbf{S}_i$ with each of the fields $\dot{\mathbf{S}}_j$ vanishes; hence

$$\{L_i, K_j\} = 0. \tag{8.85}$$

If (x^1, x^2, x^3) is a local coordinate system for SO(3), the vector fields $\dot{\mathbf{S}}_k$ can be expressed in the form

$$\dot{\mathbf{S}}_k = M^i_k \frac{\partial}{\partial x^i}, \tag{8.86}$$

where the M^i_k are real-valued functions defined on the domain of the coordinates x^i. From (8.82) one concludes that the components K_i of the angular momentum of the rigid body with respect to the axes fixed in space are given in terms of the canonical coordinates q^i, p_i induced by the x^i by means of

$$K_i = \left(\pi^* M^j_i\right) p_j \tag{8.87}$$

[see (8.26)] (since $q^i = \pi^* x^i$, the only effect of π^* on the expressions for the M^j_i is replacing the variables x^i by q^i). The 1-forms $\dot{\omega}^i$ that form the dual basis to $\{\dot{\mathbf{S}}_i\}$ are right-invariant and have the local expression

$$\dot{\omega}^k = \tilde{M}^k_i \, \mathrm{d}x^i, \tag{8.88}$$

where $(\tilde{M}^i_j)$ is the inverse of the matrix (M^i_j) (i.e., $\tilde{M}^i_j M^j_k = \delta^i_k$).

The functions $\tilde{M}^k_i$ relate the angular velocity of the body with respect to the axes fixed in the space with the velocities $\mathrm{d}x^j(g(t))/\mathrm{d}t$. If $t \mapsto g(t)$ is a differentiable curve in SO(3) that represents the configuration of a rigid body as a function of time t, then the tangent vector to that curve is given locally by $(\mathrm{d}x^i(g(t))/\mathrm{d}t)(\partial/\partial x^i)_{g(t)}$ and can also be expressed as a linear combination of the tangent vectors $(\dot{\mathbf{S}}_i)_{g(t)}$ in the form $\Omega^i(t)(\dot{\mathbf{S}}_i)_{g(t)}$. Since $\dot{\mathbf{S}}_i$ is the infinitesimal generator of rotations about $\mathbf{e}_i$, $\Omega^i(t)$ is the angular velocity of the body about the $\mathbf{e}_i$ axis. From the equality $(\mathrm{d}x^i(g(t))/\mathrm{d}t)(\partial/\partial x^i)_{g(t)} = \Omega^i(t)(\dot{\mathbf{S}}_i)_{g(t)}$ and (8.86), there results

$$\Omega^i(t) = \tilde{M}^i_j\big(g(t)\big)\frac{\mathrm{d}x^j(g(t))}{\mathrm{d}t} \tag{8.89}$$

or, by abuse of notation,

$$\Omega^i = \tilde{M}^i_j \frac{\mathrm{d}x^j}{\mathrm{d}t}. \tag{8.90}$$

In a similar way, the basis $\{\mathbf{S}_i\}$ of $\mathfrak{so}(3)$ and its dual, $\{\omega^i\}$, have expressions of the form

$$\mathbf{S}_k = M'^i_k \frac{\partial}{\partial x^i}, \qquad \omega^k = \tilde{M}'^k_i\,\mathrm{d}x^i, \tag{8.91}$$

where the M'^i_k are real-valued functions defined on the domain of the coordinates x^i and $(\tilde{M}'^i_j)$ is the inverse of the matrix (M'^i_j). The components of the angular momentum of the rigid body with respect to the axes fixed in the body are given by

$$L_i = \big(\pi^* M'^j_i\big)p_j, \tag{8.92}$$

and

$$\Omega'^i = \tilde{M}'^i_j \frac{\mathrm{d}x^j}{\mathrm{d}t} \tag{8.93}$$

is the component of the angular velocity of the body about the $\mathbf{e}'_i$ axis.

Exercise 8.36 Show that the relations (8.83) are equivalent to the Maurer–Cartan equations for the left-invariant 1-forms ω^i.

Euler Angles A commonly employed coordinate system for SO(3) is that formed by the Euler angles, though there are several slightly different forms of defining them. Following the convention of parameterizing a rotation $g \in \mathrm{SO}(3)$ by means of the three angles $\phi(g)$, $\theta(g)$, and $\psi(g)$ in such a way that

$$g = \big(\exp\phi(g)S_3\big)\big(\exp\theta(g)S_1\big)\big(\exp\psi(g)S_3\big), \tag{8.94}$$

the configuration corresponding to g is obtained rotating the body first about the $\mathbf{e}'_3$ axis by an angle $\phi(g)$, continuing with a rotation by $\theta(g)$ about the $\mathbf{e}'_1$ axis and, finally, with a rotation by $\psi(g)$ about $\mathbf{e}'_3$. Since these rotations are made about the axes fixed in the body, according to the discussion at the beginning of this section, each of these rotations multiplies by the right those applied first. In order to have a coordinate chart, the values of the Euler angles are restricted by $0 < \phi < 2\pi$, $0 < \theta < \pi$,$0 < \psi < 2\pi$.

The explicit form of the functions M^i_j, $\tilde{M}^i_j$, M'^i_j, and $\tilde{M}'^i_j$, can be conveniently obtained with the aid of Theorem 7.35. Calculating the product $g^{-1}\,\mathrm{d}g$, from (8.94) we obtain

$$\begin{aligned} g^{-1}\,\mathrm{d}g &= \big(\mathrm{e}^{-\psi S_3}\mathrm{e}^{-\theta S_1}\mathrm{e}^{-\phi S_3}\big)\,\mathrm{d}\big(\mathrm{e}^{\phi S_3}\mathrm{e}^{\theta S_1}\mathrm{e}^{\psi S_3}\big) \\ &= \mathrm{e}^{-\psi S_3}\mathrm{e}^{-\theta S_1} S_3\,\mathrm{d}\phi\,\mathrm{e}^{\theta S_1}\mathrm{e}^{\psi S_3} + \mathrm{e}^{-\psi S_3} S_1\,\mathrm{d}\theta\,\mathrm{e}^{\psi S_3} + S_3\,\mathrm{d}\psi. \end{aligned} \tag{8.95}$$

On the other hand, we can see that, for instance,

$$\mathrm{e}^{-\psi S_3} S_1 \mathrm{e}^{\psi S_3} = (\cos\psi) S_1 - (\sin\psi) S_2. \tag{8.96}$$

Indeed, denoting by $R(\psi)$ the left-hand side of (8.96), differentiating with respect to ψ and using (7.63) one finds that $\mathrm{d}R/\mathrm{d}\psi = -\mathrm{e}^{-\psi S_3} S_3 S_1 \mathrm{e}^{\psi S_3} + \mathrm{e}^{-\psi S_3} S_1 S_3 \mathrm{e}^{\psi S_3} = -\mathrm{e}^{-\psi S_3} S_2 \mathrm{e}^{\psi S_3}$. In a similar manner one obtains $\mathrm{d}^2 R/\mathrm{d}\psi^2 = \mathrm{e}^{-\psi S_3}[S_3, S_2]\mathrm{e}^{\psi S_3} = -\mathrm{e}^{-\psi S_3} S_1 \mathrm{e}^{\psi S_3} = -R$; therefore, $R = (\cos\psi)A + (\sin\psi)B$, where A and B are matrices that do not depend on ψ. Evaluating R and $\mathrm{d}R/\mathrm{d}\psi$ at $\psi = 0$ we have $R(0) = S_1 = A$ and $(\mathrm{d}R/\mathrm{d}\psi)(0) = -S_2 = B$, thus showing the validity of (8.96).

Now making use of (8.96) and the relations similar to it obtained by cyclic permutations of the indices, from (8.95) one arrives at the expression

$$\begin{aligned} g^{-1}\,\mathrm{d}g &= \mathrm{e}^{-\psi S_3}(\cos\theta\, S_3 + \sin\theta\, S_2)\mathrm{e}^{\psi S_3}\,\mathrm{d}\phi + (\cos\psi\, S_1 - \sin\psi\, S_2)\,\mathrm{d}\theta + S_3\,\mathrm{d}\psi \\ &= \big[\cos\theta\, S_3 + \sin\theta(\cos\psi\, S_2 + \sin\psi\, S_1)\big]\,\mathrm{d}\phi + (\cos\psi\, S_1 - \sin\psi\, S_2)\,\mathrm{d}\theta \\ &\quad + S_3\,\mathrm{d}\psi \\ &= (\sin\theta\sin\psi\,\mathrm{d}\phi + \cos\psi\,\mathrm{d}\theta)S_1 + (\sin\theta\cos\psi\,\mathrm{d}\phi - \sin\psi\,\mathrm{d}\theta)S_2 \\ &\quad + (\cos\theta\,\mathrm{d}\phi + \mathrm{d}\psi)S_3, \end{aligned} \tag{8.97}$$

where the coefficient of the matrix S_i is the 1-form ω^i [see (7.46)] and comparing with (8.91) we obtain the matrix $(\tilde{M}'^i_j)$. Then, it is easy to calculate the dual basis to $\{\omega^i\}$, and the result is

$$\begin{aligned} \mathbf{S}_1 &= \csc\theta\sin\psi\frac{\partial}{\partial\phi} + \cos\psi\frac{\partial}{\partial\theta} - \cot\theta\sin\psi\frac{\partial}{\partial\psi}, \\ \mathbf{S}_2 &= \csc\theta\cos\psi\frac{\partial}{\partial\phi} - \sin\psi\frac{\partial}{\partial\theta} - \cot\theta\cos\psi\frac{\partial}{\partial\psi}, \\ \mathbf{S}_3 &= \frac{\partial}{\partial\psi}. \end{aligned} \tag{8.98}$$

(The last of these equations also follows directly from the definition of the Euler angles, taking into account that $\mathbf{S}_3$ generates rotations about the $\mathbf{e}'_3$ axis.) It should be noticed that the expressions on the right-hand side of (8.98) are not defined at $\theta = 0$, but θ does not vanish in the domain of the coordinate system ϕ, θ, ψ. In fact, a straightforward computation gives $\omega^1 \wedge \omega^2 \wedge \omega^3 = \sin\theta\,\mathrm{d}\theta \wedge \mathrm{d}\phi \wedge \mathrm{d}\psi$, but the left-invariant 3-form $\omega^1 \wedge \omega^2 \wedge \omega^3$ is everywhere different from zero. Then, using (8.91), (8.92), and (8.98) we find that

$$\begin{aligned} L_1 &= \csc\theta \sin\psi \, p_\phi + \cos\psi \, p_\theta - \cot\theta \sin\psi \, p_\psi, \\ L_2 &= \csc\theta \cos\psi \, p_\phi - \sin\psi \, p_\theta - \cot\theta \cos\psi \, p_\psi, \\ L_3 &= p_\psi, \end{aligned} \tag{8.99}$$

where, by abuse of notation, the variables $\pi^*\phi$, $\pi^*\theta$, and $\pi^*\psi$ have been denoted by ϕ, θ, and ψ, respectively; that is, in (8.99), the Euler angles are regarded as variables defined on the phase space $T^*\mathrm{SO}(3)$.

The 1-forms $\dot{\omega}^i$ can readily be obtained by means of the relation $\dot{\omega}^i = -\iota^*\omega^i$ (see Exercise 7.31) using the fact that $\iota^*\phi = -\psi$, $\iota^*\theta = -\theta$, and $\iota^*\psi = -\phi$ [see (8.94)]; in this way we obtain, for instance,

$$\begin{aligned} \dot{\mathbf{S}}_1 &= \csc\theta \sin\phi \frac{\partial}{\partial\psi} + \cos\phi \frac{\partial}{\partial\theta} - \cot\theta \sin\phi \frac{\partial}{\partial\phi}, \\ \dot{\mathbf{S}}_2 &= -\csc\theta \cos\phi \frac{\partial}{\partial\psi} + \sin\phi \frac{\partial}{\partial\theta} + \cot\theta \cos\phi \frac{\partial}{\partial\phi}, \\ \dot{\mathbf{S}}_3 &= \frac{\partial}{\partial\phi}, \end{aligned} \tag{8.100}$$

and, therefore,

$$\begin{aligned} K_1 &= \csc\theta \sin\phi \, p_\psi + \cos\phi \, p_\theta - \cot\theta \sin\phi \, p_\phi, \\ K_2 &= -\csc\theta \cos\phi \, p_\psi + \sin\phi \, p_\theta + \cot\theta \cos\phi \, p_\phi, \\ K_3 &= p_\phi. \end{aligned} \tag{8.101}$$

Dynamics of a Rigid Body If the curve $t \mapsto g(t)$ in SO(3) corresponds to the motion of a rigid body with a fixed point, from the elementary definition of the kinetic energy of a particle, it follows that the kinetic energy of the rigid body is given by $E_\mathrm{K} = \frac{1}{2} I_{ij} \Omega'^i \Omega'^j$, where $\Omega'^i(t)$ is the component of the angular velocity of the body about the $\mathbf{e}'_i$ axis and the constants $I_{ij} = I_{ji}$ are the components of the *inertia tensor* of the body with respect to the basis $\{\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3\}$. From (8.93) and (8.91) it follows that

$$E_\mathrm{K} = \frac{1}{2} I_{ij} \big(\tilde{M}'^i_{\ k} \, \mathrm{d}x^k \otimes \tilde{M}'^j_{\ l} \, \mathrm{d}x^l\big)(g'_t, g'_t) = \frac{1}{2}\big(I_{ij}\omega^i \otimes \omega^j\big)(g'_t, g'_t), \tag{8.102}$$

where g'_t is the tangent vector to the curve $t \mapsto g(t)$.

The tensor field $I_{ij}\omega^i \otimes \omega^j$, appearing in (8.102), is symmetric and positive definite (excluding the case where the rigid body is formed by point particles aligned on an axis passing through the fixed point of the body); therefore it is a metric tensor for the manifold SO(3). Furthermore, since the I_{ij} are constant and the ω^i are left-invariant 1-forms, $I_{ij}\omega^i \otimes \omega^j$ is a *left-invariant metric*, i.e.,

$L_g{}^*(I_{ij}\omega^i \otimes \omega^j) = I_{ij}\omega^i \otimes \omega^j$ for $g \in \mathrm{SO}(3)$. In other words, for each $g \in \mathrm{SO}(3)$, the transformation L_g, from SO(3) onto itself, is an isometry and the right-invariant vector fields on SO(3) are Killing vector fields for this metric, regardless of the values of the components of the inertia tensor I_{ij}. Making use of the local expression for the ω^k given in (8.91), we can also write the metric tensor in the standard form,

$$I_{ij}\omega^i \otimes \omega^j = I_{ij}\tilde{M}'^i_{\ k}\tilde{M}'^j_{\ l}\,\mathrm{d}x^k \otimes \mathrm{d}x^l = g_{kl}\,\mathrm{d}x^k \otimes \mathrm{d}x^l, \tag{8.103}$$

where $g_{kl} \equiv I_{ij}\tilde{M}'^i_{\ k}\tilde{M}'^j_{\ l}$.

It should be stressed that all the rigid bodies with a fixed point possess the same configuration space [the underlying manifold of the group SO(3)], but the metric on this manifold is given by the inertia tensor of the body.

The vector fields $\mathbf{S}_i$ form a rigid basis with respect to the metric $I_{ij}\omega^i \otimes \omega^j$. Comparing equations (7.63) and (6.62), and using (6.63) it follows that the connection 1-forms for the corresponding Riemannian connection, with respect to this basis, are

$$\Gamma_{ij} = -\frac{1}{2}\sum_{m=1}^{3}(I_{im}\varepsilon_{mjk} - I_{jm}\varepsilon_{mik} - I_{km}\varepsilon_{mij})\omega^k.$$

(See also Appendix B.)

In the particular case where $I_{ij} = I\delta_{ij}$, where I is a constant (which corresponds to the so-called *spherical top*), the left-invariant vector fields $\mathbf{S}_i$ are also Killing vector fields. Indeed, we have $\pounds_{\mathbf{S}_i}(I\delta_{jk}\omega^j \otimes \omega^k) = I\delta_{jk}[(\pounds_{\mathbf{S}_i}\omega^j) \otimes \omega^k + \omega^j \otimes \pounds_{\mathbf{S}_i}\omega^k]$; on the other hand, from (3.39), (7.63), the Maurer–Cartan equations, and (3.27), $\pounds_{\mathbf{S}_i}\omega^j = \mathbf{S}_i \lrcorner \mathrm{d}\omega^j + \mathrm{d}(\mathbf{S}_i \lrcorner \omega^j) = \mathbf{S}_i \lrcorner (-\frac{1}{2}\varepsilon_{jkl}\omega^k \wedge \omega^l) = \varepsilon_{ijl}\omega^l$; hence,

$$\pounds_{\mathbf{S}_i}\big(I\delta_{jk}\omega^j \otimes \omega^k\big) = I(\varepsilon_{ikm} + \varepsilon_{imk})\omega^m \otimes \omega^k = 0.$$

In this case, the connection 1-forms for the basis formed by the $\mathbf{S}_i$ are $\Gamma^i{}_j = -\frac{1}{2}I\varepsilon_{ijk}\omega^k$ and from the second Cartan structural equations (5.18) one finds that $\mathscr{R}^i{}_j = \frac{1}{4}I\omega^i \wedge \omega^j$, or, equivalently, $R_{ijkl} = \frac{1}{4}I(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk})$, which corresponds to a space of constant curvature [see (6.100)]. With this metric, SO(3) is locally isometric to the sphere S^3. (However, SO(3) and S^3 are globally distinct; whereas S^3 is simply connected, SO(3) is not.) Using, for instance, the expressions given by (8.97) for the 1-forms ω^i in terms of the Euler angles one finds that the metric $I\delta_{ij}\omega^i \otimes \omega^j$ is

$$I\big[\mathrm{d}\theta \otimes \mathrm{d}\theta + \mathrm{d}\phi \otimes \mathrm{d}\phi + \mathrm{d}\psi \otimes \mathrm{d}\psi + \cos\theta(\mathrm{d}\phi \otimes \mathrm{d}\psi + \mathrm{d}\psi \otimes \mathrm{d}\phi)\big].$$

Exercise 8.37 Show that if $(I_{ij}) = \mathrm{diag}(I_1, I_1, I_3)$ (a symmetric top), then $\mathbf{S}_3$ is a Killing vector field of the metric $I_{ij}\omega^i \otimes \omega^j$.

Going back to the general case, from (8.102) and (8.103) it follows that the kinetic energy of a rigid body with a fixed point can also be expressed in the form

$$T = \frac{1}{2}\left(\pi^* g^{ij}\right) p_i p_j, \tag{8.104}$$

where (g^{ij}) is the inverse of the matrix $g_{ij} \equiv I_{kl} \tilde{M}'^k_i \tilde{M}'^l_j$ [cf. (8.47)]. (In contrast to E_K, T is a function defined on the phase space.) According to (8.92), we also have

$$T = \frac{1}{2} I^{ij} L_i L_j, \tag{8.105}$$

where (I^{ij}) denotes the inverse of the matrix (I_{ij}). The standard Hamiltonian function for a rigid body is the sum of its kinetic and potential energies. If the axes of the coordinate system fixed in the body are principal axes of the inertia tensor [with respect to which the matrix (I_{ij}) is diagonal, $(I_{ij}) = \text{diag}(I_1, I_2, I_3)$ and, therefore, $(I^{ij}) = \text{diag}(1/I_1, 1/I_2, 1/I_3)$], the Hamiltonian is then

$$H = \frac{{L_1}^2}{2I_1} + \frac{{L_2}^2}{2I_2} + \frac{{L_3}^2}{2I_3} + \pi^* V, \tag{8.106}$$

where I_1, I_2, I_3 are the so-called *principal moments of inertia* [see (8.105)] and V corresponds to the potential energy. From (8.43), together with (8.106), (8.83), and the properties of the Poisson bracket that follow from the definition (8.19), we find that

$$\begin{aligned}\frac{\mathrm{d}L_1}{\mathrm{d}t} &= \{H, L_1\} \\ &= \frac{1}{I_1}\{L_1, L_1\}L_1 + \frac{1}{I_2}\{L_2, L_1\}L_2 + \frac{1}{I_3}\{L_3, L_1\}L_3 + \{\pi^* V, L_1\} \\ &= -\frac{1}{I_2}L_2 L_3 + \frac{1}{I_3} L_2 L_3 + \{\pi^* V, L_1\}.\end{aligned}$$

The functions L_i appearing in (8.105) and (8.106) are generating functions of the rotations of the rigid body about the axes fixed in the body [see (8.82)]. However, according to its elementary definition, the angular momentum should depend linearly on the angular velocity. From the Hamilton equations (8.36) and (8.104) one finds that

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i} = \frac{\partial T}{\partial p_i} = \left(\pi^* g^{ij}\right) p_j.$$

Hence, making use of (8.92) and (8.93),

$$L_i = \left(\pi^* M'^j_i\right) p_j = \pi^*\left(M'^j_i g_{jk}\right) \frac{\mathrm{d}q^k}{\mathrm{d}t} = \pi^*\left(M'^j_i g_{jk} M'^k_l\right) \Omega'^l = I_{il} \Omega'^l.$$

Thus, with respect to the principal axes of the inertia tensor, $L_i = I_i \Omega'^i$ (without sum on i), and we have

$$I_1 \frac{\mathrm{d}\Omega'^1}{\mathrm{d}t} - (I_2 - I_3)\Omega'^2\Omega'^3 = N'_1, \tag{8.107}$$

where $N'_1 \equiv \{\pi^* V, L_1\}$. In an analogous way we obtain

$$\begin{aligned} I_2 \frac{\mathrm{d}\Omega'^2}{\mathrm{d}t} - (I_3 - I_1)\Omega'^1\Omega'^3 &= N'_2, \\ I_3 \frac{\mathrm{d}\Omega'^3}{\mathrm{d}t} - (I_1 - I_2)\Omega'^1\Omega'^2 &= N'_3, \end{aligned} \tag{8.108}$$

with $N'_i \equiv \{\pi^* V, L_i\}$. Since L_i is a generating function of rotations about the $\mathbf{e}'_i$ axis (fixed in the body), the functions N'_i correspond to the components of the torque with respect to the axes fixed in the body. Equations (8.107) and (8.108) are known as the Euler equations.

When $V = 0$, the torque is equal to zero and the Hamiltonian (8.106) reduces to the kinetic energy T, which is given by (8.105) or by (8.104); therefore, the Euler equations (8.107) and (8.108) with $N'_i = 0$ amount to the equations for the geodesics of the metric $I_{ij}\omega^i \otimes \omega^j$.

Finally, we consider the case of a *symmetric rigid body* with a fixed point in a uniform gravitational field. Choosing, as usual, $I_1 = I_2$ and taking the fixed point of the body as the origin of the coordinate systems fixed in space and in the body, from (8.106) and (8.99) one finds that

$$H = \frac{1}{2I_1}\left[{p_\theta}^2 + \frac{(p_\phi - \cos\theta\, p_\psi)^2}{\sin^2\theta} \right] + \frac{{p_\psi}^2}{2I_3} + mgl\cos\theta, \tag{8.109}$$

where m is the mass of the body and l is the distance from the fixed point to the center of mass. Since H does not depend on ϕ, ψ, and t, the Hamilton equations (8.36) imply that

$$p_\phi = \text{const}, \qquad p_\psi = \text{const}, \qquad H = \text{const}(\equiv E) \tag{8.110}$$

(i.e., K_3 and L_3 are constants of motion; cf. Exercise 8.37). On the other hand

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{I_1}, \tag{8.111}$$

therefore, from (8.109)–(8.111) one obtains the separated equation

$$\frac{I_1}{2}\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 + \frac{(p_\phi - \cos\theta\, p_\psi)^2}{2I_1\sin^2\theta} + mgl\cos\theta = E' \equiv E - \frac{{p_\psi}^2}{2I_3}, \tag{8.112}$$

which is usually obtained by means of the Lagrangian formalism [see, e.g., Goldstein (1980, Sects. 5–7)].

8.7 Time-Dependent Formalism

In the foregoing sections we have restricted ourselves to the case where the Hamiltonian is a real-valued function defined on the phase space. The Hamilton equations (8.36) then constitute an *autonomous* system of ODEs (that is, the right-hand sides of the Hamilton equations (8.36) do not depend explicitly on t). However, the formalism can readily be extended to the more general case where the Hamiltonian depends explicitly on time, and even in those cases where a given Hamiltonian does not depend explicitly on time, it is convenient to consider canonical transformations that lead to a new Hamiltonian that may depend on time.

Throughout this section, P will denote a $2n$-dimensional differentiable manifold, which in many cases will be the cotangent bundle of some n-dimensional differentiable manifold. We begin by noticing that for a given 2-form Ω on $P \times \mathbb{R}$, of the form $\Omega = \mathrm{d}p_i \wedge \mathrm{d}q^i - \mathrm{d}H \wedge \mathrm{d}t$, where (q^i, p_i, t) is a local system of coordinates on $P \times \mathbb{R}$ and $H \in C^\infty(P \times \mathbb{R})$, there exists a *unique* vector field $\mathbf{A} \in \mathfrak{X}(P \times \mathbb{R})$ such that $\mathbf{A}t = 1$ and $\mathbf{A} \lrcorner\, \Omega = 0$. In fact, a straightforward computation shows that these two conditions imply that

$$\mathbf{A} = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i}\frac{\partial}{\partial p_i} + \frac{\partial}{\partial t}$$

and therefore the integral curves of $\mathbf{A}$ are determined by the Hamilton equations

$$\frac{\mathrm{d}(q^i \circ C)}{\mathrm{d}s} = \frac{\partial H}{\partial p_i} \circ C, \qquad \frac{\mathrm{d}(p_i \circ C)}{\mathrm{d}s} = -\frac{\partial H}{\partial q^i} \circ C, \qquad \frac{\mathrm{d}(t \circ C)}{\mathrm{d}s} = 1. \tag{8.113}$$

The last equation, which amounts to $\mathbf{A}t = 1$, means that the integral curves of $\mathbf{A}$ are parametrized by t, which represents the time. In what follows t will denote the natural coordinate of $\mathbb{R}$, but q^i, p_i need not be coordinates on P (for instance, the functions x and $r = \sqrt{x^2 + y^2}$, where (x, y) are the natural coordinates of $\mathbb{R}^2$, form a coordinate system on $\mathbb{R} \times \mathbb{R}$ (that covers, e.g., the half-plane $y > 0$); whereas x is the natural coordinate on the first copy of $\mathbb{R}$, r is not a coordinate on the second copy).

The local expression of Ω considered above follows from $\mathrm{d}\Omega = 0$ and the condition that at each point $x \in P \times \mathbb{R}$, the linear mapping from $T_x(P \times \mathbb{R})$ into $T_x^*(P \times \mathbb{R})$, given by $v_x \mapsto v_x \lrcorner\, \Omega_x$, has rank $2n$. The kernel of this mapping has dimension one and is generated by $\mathbf{A}_x$. (Actually, the Darboux Theorem guarantees the local existence of $2n$ functionally independent functions, P_i, Q^i, such that $\Omega = \mathrm{d}P_i \wedge \mathrm{d}Q^i$. By means of a canonical transformation (see below) one can take Ω to the desired form.)

Symmetries and Constants of Motion As shown in Sect. 8.3, in the case where the time evolution of a mechanical system is defined by a Hamiltonian vector field, $\mathbf{X}_{\mathrm{d}H}$, on T^*M, a function $f \in C^\infty(T^*M)$ is a constant of motion if and only if the vector field $\mathbf{X}_{\mathrm{d}f}$ generates a one-parameter group of canonical transformations that leave the Hamiltonian invariant (i.e., $\mathbf{X}_{\mathrm{d}f} H = 0$). Even for such mechanical systems, there exist constants of motion that depend explicitly on time (see, e.g., Example 8.38 below), which therefore are not related with symmetries in the framework developed in the preceding sections. As we shall show, whenever the evolution equations can be expressed in the form of the Hamilton equations, *any* constant of motion (that may explicitly depend on time) is associated with a one-parameter group of symmetries.

We will say that the vector field $\mathbf{X}$ on $P \times \mathbb{R}$ is a *symmetry* of Ω if $\pounds_{\mathbf{X}}\Omega = 0$. As a consequence of the formula $\pounds_{\mathbf{X}}\Omega = \mathbf{X} \lrcorner \, \mathrm{d}\Omega + \mathrm{d}(\mathbf{X} \lrcorner \, \Omega)$, and the fact that $\mathrm{d}\Omega = 0$, we see that $\mathbf{X}$ is a symmetry of Ω if and only if the 1-form $\mathbf{X} \lrcorner \, \Omega$ is closed (cf. Lemma 8.5). Thus, if $\mathbf{X}$ is a symmetry of Ω there exists, locally, a function $\chi \in C^\infty(P \times \mathbb{R})$ such that $\mathbf{X} \lrcorner \, \Omega = -\mathrm{d}\chi$. Then, the function χ is a *constant of motion*, i.e., $\mathbf{A}\chi = 0$. Indeed,

$$\mathbf{A}\chi = \mathbf{A} \lrcorner \, \mathrm{d}\chi = -\mathbf{A} \lrcorner (\mathbf{X} \lrcorner \, \Omega) = \mathbf{X} \lrcorner (\mathbf{A} \lrcorner \, \Omega) = 0.$$

The vector field $\mathbf{A}$ satisfies the symmetry condition $\pounds_{\mathbf{A}}\Omega = 0$, but no nontrivial constant of motion is associated with $\mathbf{A}$, since $\mathbf{A} \lrcorner \, \Omega = 0$.

Conversely, given a constant of motion, χ, there exists a vector field $\mathbf{X}$, defined up to the addition of a multiple of $\mathbf{A}$, such that $\mathbf{X} \lrcorner \, \Omega = -\mathrm{d}\chi$ (then $\mathbf{X}$ is a symmetry of Ω). In fact, writing

$$\mathbf{X} = A^i \frac{\partial}{\partial q^i} + B_i \frac{\partial}{\partial p_i} + C \frac{\partial}{\partial t}$$

the condition $\mathbf{X} \lrcorner \, \Omega = -\mathrm{d}\chi$ amounts to

$$\begin{aligned} -\mathrm{d}\chi &= \mathbf{X} \lrcorner \big(\mathrm{d}p_i \wedge \mathrm{d}q^i - \mathrm{d}H \wedge \mathrm{d}t\big) \\ &= (\mathbf{X}p_i)\, \mathrm{d}q^i - \big(\mathbf{X}q^i\big)\, \mathrm{d}p_i - (\mathbf{X}H)\, \mathrm{d}t + (\mathbf{X}t)\, \mathrm{d}H \\ &= B_i\, \mathrm{d}q^i - A^i\, \mathrm{d}p_i - (\mathbf{X}H)\, \mathrm{d}t + C\, \mathrm{d}H, \end{aligned}$$

that is,

$$\frac{\partial \chi}{\partial q^i} = -B_i - C\frac{\partial H}{\partial q^i}, \qquad \frac{\partial \chi}{\partial p_i} = A^i - C\frac{\partial H}{\partial p_i}, \qquad \frac{\partial \chi}{\partial t} = A^i \frac{\partial H}{\partial q^i} + B_i \frac{\partial H}{\partial p_i}.$$

From the first two equations we find that

$$A^i = \frac{\partial \chi}{\partial p_i} + C\frac{\partial H}{\partial p_i}, \qquad B_i = -\frac{\partial \chi}{\partial q^i} - C\frac{\partial H}{\partial q^i},$$

and substituting into the last equation we obtain

$$\frac{\partial \chi}{\partial t} = \frac{\partial \chi}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial \chi}{\partial q^i}\frac{\partial H}{\partial p_i},$$

which is equivalent to the assumed condition $\mathbf{A}\chi = 0$. Thus, we have

$$\mathbf{X} = \frac{\partial \chi}{\partial p_i}\frac{\partial}{\partial q^i} - \frac{\partial \chi}{\partial q^i}\frac{\partial}{\partial p_i} + C\mathbf{A}, \tag{8.114}$$

where C is an arbitrary real-valued differentiable function [cf. (8.17)].

Thus, *any* constant of motion (which possibly depends explicitly on the time) is associated with a symmetry of Ω, but, in contrast to the result derived in Sect. 8.3, we do not necessarily have $\mathbf{X}H = 0$ (cf. Example 3.17).

Example 8.38 The vector field

$$\mathbf{A} = \frac{p_1}{m}\frac{\partial}{\partial q^1} + \frac{p_2}{m}\frac{\partial}{\partial q^2} - mg\frac{\partial}{\partial p_2} + \frac{\partial}{\partial t}, \tag{8.115}$$

where m and g are constants, is the only vector field that satisfies the conditions $\mathbf{A}t = 1$ and $\mathbf{A}\lrcorner\,\Omega = 0$, with Ω given by

$$\Omega = \mathrm{d}p_1 \wedge \mathrm{d}q^1 + \mathrm{d}p_2 \wedge \mathrm{d}q^2 - \mathrm{d}\left(\frac{{p_1}^2 + {p_2}^2}{2m} + mgq^2\right) \wedge \mathrm{d}t, \tag{8.116}$$

i.e., $H = ({p_1}^2 + {p_2}^2)/(2m) + mgq^2$ is a Hamiltonian function (which represents the total energy of a particle of mass m in a uniform gravitational field, with g being the acceleration of gravity) for the equations of motion defined by $\mathbf{A}$.

Even though $(\partial/\partial q^2)H = mg \neq 0$, one can verify that $\pounds_{\partial/\partial q^2}\Omega = 0$; in fact, one finds that $(\partial/\partial q^2)\lrcorner\,\Omega = -\mathrm{d}(p_2 + mgt)$, i.e., $\partial/\partial q^2$ is a symmetry of Ω. Hence, according to the discussion above, $\chi \equiv p_2 + mgt$ is a constant of motion, which explicitly depends on time. (Note that H does not explicitly depend on t in the coordinate system employed here.)

If $\mathbf{X}$ is a symmetry of Ω, then

$$(\pounds_{\mathbf{X}}\mathbf{A})\lrcorner\,\Omega = \pounds_{\mathbf{X}}(\mathbf{A}\lrcorner\,\Omega) - \mathbf{A}\lrcorner(\pounds_{\mathbf{X}}\Omega) = 0,$$

and therefore $\pounds_{\mathbf{X}}\mathbf{A}$ must be proportional to $\mathbf{A}$.

For a given vector field $\mathbf{A}$, the set of constants of motion is a vector space over $\mathbb{R}$ (with the usual operations of sum of functions and multiplication by scalars) which can be made into a Lie algebra by defining

$$\{\chi_1, \chi_2\} \equiv \mathbf{X}_1\chi_2, \tag{8.117}$$

where $\mathbf{X}_1$ is a vector field such that $\mathbf{X}_1\lrcorner\,\Omega = -\mathrm{d}\chi_1$ [cf. (8.19)]. As we have shown, the vector field $\mathbf{X}_1$ is defined up to the addition of a multiple of $\mathbf{A}$, but this ambiguity has no effect on the definition (8.117) since $\mathbf{A}\chi_2 = 0$. The bracket $\{\chi_1, \chi_2\}$ is indeed a constant of motion because $\mathbf{A}\{\chi_1, \chi_2\} = \mathbf{A}\mathbf{X}_1\chi_2 = [\mathbf{A}, \mathbf{X}_1]\chi_2 = 0$, since $[\mathbf{A}, \mathbf{X}_1] = -[\mathbf{X}_1, \mathbf{A}]$ is proportional to $\mathbf{A}$.

Furthermore, if $\mathbf{X}_2$ is a vector field such that $\mathbf{X}_2\lrcorner\,\Omega = -\mathrm{d}\chi_2$, $[\mathbf{X}_1, \mathbf{X}_2]$ is also a symmetry of Ω (see Exercise 2.30) and

$$[\mathbf{X}_1, \mathbf{X}_2]\lrcorner\,\Omega = (\pounds_{\mathbf{X}_1}\mathbf{X}_2)\lrcorner\,\Omega = \pounds_{\mathbf{X}_1}(\mathbf{X}_2\lrcorner\,\Omega) = -\pounds_{\mathbf{X}_1}\mathrm{d}\chi_2 = -\mathrm{d}\{\chi_1, \chi_2\}$$

[cf. (8.18)]. We also have

$$\{\chi_1, \chi_2\} = \mathbf{X}_1 \chi_2 = -\mathbf{X}_1 \lrcorner (\mathbf{X}_2 \lrcorner \Omega) = 2\Omega(\mathbf{X}_1, \mathbf{X}_2),$$

which implies the skew-symmetry of the bracket [cf. (8.22)].

Exercise 8.39 Show that the bracket (8.117) satisfies the Jacobi identity.

From (8.114) and (8.117) one obtains the local expression

$$\{\chi_1, \chi_2\} = \frac{\partial \chi_1}{\partial p_i}\frac{\partial \chi_2}{\partial q^i} - \frac{\partial \chi_1}{\partial q^i}\frac{\partial \chi_2}{\partial p_i}$$

[cf. (8.21)].

Canonical Transformations The coordinates q^i, p_i, as well as the Hamiltonian H, in terms of which the 2-form Ω has the form $\mathrm{d}p_i \wedge \mathrm{d}q^i - \mathrm{d}H \wedge \mathrm{d}t$, are not defined uniquely by Ω. There is an infinite number of sets $\{Q^i, P_i, K\}$ such that (Q^i, P_i, t) is a coordinate system on $P \times \mathbb{R}$ and $\Omega = \mathrm{d}P_i \wedge \mathrm{d}Q^i - \mathrm{d}K \wedge \mathrm{d}t$ (which implies that the integral curves of $\mathbf{A}$ are determined by equations of the form (8.113), with $\{Q^i, P_i, K, t\}$ in place of $\{q^i, p_i, H, t\}$). Indeed, the equality

$$\mathrm{d}p_i \wedge \mathrm{d}q^i - \mathrm{d}H \wedge \mathrm{d}t = \mathrm{d}P_i \wedge \mathrm{d}Q^i - \mathrm{d}K \wedge \mathrm{d}t$$

is equivalent, e.g., to

$$\mathrm{d}\big(p_i\,\mathrm{d}q^i - H\,\mathrm{d}t - P_i\,\mathrm{d}Q^i + K\,\mathrm{d}t\big) = 0.$$

In turn, this is locally equivalent to the existence of a function F such that

$$p_i\,\mathrm{d}q^i - H\,\mathrm{d}t - P_i\,\mathrm{d}Q^i + K\,\mathrm{d}t = \mathrm{d}F; \tag{8.118}$$

cf. Example 3.16. (Note that we consider these transformations as coordinate transformations, that is, as passive transformations that do not affect the points of the manifold $P \times \mathbb{R}$.) If q^i, Q^i, and t are functionally independent, they can be used as local coordinates on $P \times \mathbb{R}$, and from (8.118) it follows that

$$p_i = \frac{\partial F}{\partial q^i}, \qquad P_i = -\frac{\partial F}{\partial Q^i}, \qquad K - H = \frac{\partial F}{\partial t}. \tag{8.119}$$

The function F is a *generating function* of the canonical transformation.

Example 8.40 The coordinate transformation

$$q^1 = Q^1 + Q^2, \qquad q^2 = \frac{c}{eB}(P_1 - P_2),$$
$$p_1 = \frac{1}{2}(P_1 + P_2), \qquad p_2 = \frac{eB}{2c}\big(Q^2 - Q^1\big),$$

with $K = H$, where e, B, and c are nonzero constants, is canonical. In fact, one readily verifies that

$$\mathrm{d}p_1 \wedge \mathrm{d}q^1 + \mathrm{d}p_2 \wedge \mathrm{d}q^2 = \mathrm{d}P_1 \wedge \mathrm{d}Q^1 + \mathrm{d}P_2 \wedge \mathrm{d}Q^2; \tag{8.120}$$

moreover,

$$p_1\,\mathrm{d}q^1 + p_2\,\mathrm{d}q^2 - P_1\,\mathrm{d}Q^1 - P_2\,\mathrm{d}Q^2 = \mathrm{d}\left[\frac{eB}{2c}q^2\left(Q^2 - Q^1\right)\right]$$

but the set $\{q^1, q^2, Q^1, Q^2, t\}$ is not functionally independent and, therefore, the relations (8.119) make no sense. However, (8.120) also follows from

$$\begin{aligned} &p_1\,\mathrm{d}q^1 - q^2\,\mathrm{d}p_2 + Q^1\,\mathrm{d}P_1 + Q^2\,\mathrm{d}P_2 \\ &\quad = \mathrm{d}\left[\frac{1}{2}\left(q^1 - \frac{2c}{eB}p_2\right)P_1 + \frac{1}{2}\left(q^1 + \frac{2c}{eB}p_2\right)P_2\right] \end{aligned}$$

and (among other choices) the set $\{q^1, p_2, P_1, P_2, t\}$ is functionally independent. Therefore, using (q^1, p_2, P_1, P_2, t) as local coordinates on $P \times \mathbb{R}$, the coordinate transformation considered here can be reproduced from the generating function $F = \frac{1}{2}(q^1 - \frac{2c}{eB}p_2)P_1 + \frac{1}{2}(q^1 + \frac{2c}{eB}p_2)P_2$, appearing on the right-hand side of the last equation. One can readily verify that the relations

$$p_1 = \frac{\partial F}{\partial q^1}, \qquad q^2 = -\frac{\partial F}{\partial p_2}, \qquad Q^1 = \frac{\partial F}{\partial P_1}, \qquad Q^2 = \frac{\partial F}{\partial P_2},$$

are equivalent to the given coordinate transformation.

Alternative Hamiltonians It is not widely known that for a given vector field $\mathbf{A} \in \mathfrak{X}(P \times \mathbb{R})$, with $\mathbf{A}t = 1$, there exists an infinite number of closed 2-forms of rank $2n$, Ω, such that $\mathbf{A} \lrcorner\, \Omega = 0$, which are not multiples of one another (except in the case where $\dim P = 2$; see below). For instance, one readily finds that the vector field (8.115), considered in Example 8.38, contracted with the closed 2-form

$$\Omega' = \mathrm{d}p_2 \wedge \mathrm{d}q^1 + \mathrm{d}p_1 \wedge \mathrm{d}q^2 - \mathrm{d}\left(\frac{p_1 p_2}{m} + mgq^1\right) \wedge \mathrm{d}t, \tag{8.121}$$

yields zero, but Ω' cannot be written as some real-valued function multiplied by the 2-form Ω given by (8.116). In fact, by means of a straightforward computation, one readily verifies that Ω and Ω' can be expressed as

$$\begin{aligned} \Omega &= \mathrm{d}p_1 \wedge \mathrm{d}\left(q^1 + \frac{p_1 p_2}{m^2 g}\right) + \mathrm{d}(p_2 + mgt) \wedge \mathrm{d}\left(q^2 + \frac{{p_1}^2 + {p_2}^2}{2m^2 g}\right), \\ \Omega' &= \mathrm{d}(p_2 + mgt) \wedge \mathrm{d}\left(q^1 + \frac{p_1 p_2}{m^2 g}\right) + \mathrm{d}p_1 \wedge \mathrm{d}\left(q^2 + \frac{{p_1}^2 + {p_2}^2}{2m^2 g}\right). \end{aligned} \tag{8.122}$$

The vector field $\partial/\partial q^2$, which is a symmetry of Ω, is also a symmetry of Ω', but now $(\partial/\partial q^2) \lrcorner\, \Omega' = -\mathrm{d}p_1$, which implies that p_1 is a constant of motion.

The only way in which the vector field $\mathbf{A}$ satisfies the relations $\mathbf{A} \lrcorner\, \Omega = 0$ and $\mathbf{A} \lrcorner\, \Omega' = 0$, considered in most textbooks on analytical mechanics [e.g., Goldstein (1980, Sect. 9-1)], is the trivial one, where Ω' differs from Ω at most by a *constant* factor.

For a given vector field $\mathbf{A} \in \mathfrak{X}(P \times \mathbb{R})$ such that $\mathbf{A}t = 1$, the *local* existence of an infinite number of closed 2-forms of rank $2n$, Ω, such that $\mathbf{A} \lrcorner\, \Omega = 0$ can be demonstrated in the following way. Let $\chi^1, \chi^2, \ldots, \chi^{2n}$ be $2n$ functionally independent constants of motion (i.e., $\mathbf{A}\chi^i = 0$, $i = 1, 2, \ldots, 2n$); then $\Omega \equiv \mathrm{d}\chi^1 \wedge \mathrm{d}\chi^2 + \mathrm{d}\chi^3 \wedge \mathrm{d}\chi^4 + \cdots + \mathrm{d}\chi^{2n-1} \wedge \mathrm{d}\chi^{2n}$ is closed, has rank $2n$ (as a consequence of the assumed functional independence of the χ^i), and we have

$$\begin{aligned} \mathbf{A} \lrcorner\, \Omega &= \left(\mathbf{A}\chi^1\right) \mathrm{d}\chi^2 - \left(\mathbf{A}\chi^2\right) \mathrm{d}\chi^1 + \cdots + \left(\mathbf{A}\chi^{2n-1}\right) \mathrm{d}\chi^{2n} - \left(\mathbf{A}\chi^{2n}\right) \mathrm{d}\chi^{2n-1} \\ &= 0, \end{aligned}$$

since, by hypothesis, $\mathbf{A}\chi^i = 0$, for $i = 1, 2, \ldots, 2n$. The ordered set of constants of motion $\{\chi^1, \chi^2, \ldots, \chi^{2n}\}$ is not unique; we can simply make permutations of the functions χ^i [as in (8.122)] or we can replace $\chi^1, \chi^2, \ldots, \chi^{2n}$ by any functionally independent set of functions of them.

Conversely, if Ω is a closed 2-form of rank $2n$ such that $\mathbf{A} \lrcorner\, \Omega = 0$, according to the Darboux Theorem, Ω is locally of the form $\mathrm{d}p_i \wedge \mathrm{d}q^i$, with the set $\{q^i, p_i\}$ being functionally independent. Then, from $\mathbf{A} \lrcorner\, \Omega = 0$ it follows that the q^i, p_i are constants of motion.

The Case dim $P = 2$ In the special case where $\dim P = 2$, locally there exist essentially only two functionally independent constants of motion, χ^1, χ^2; any other two functionally independent constants of motion, χ'^1, χ'^2, must be functions of χ^1, χ^2 only, hence

$$\mathrm{d}\chi'^1 \wedge \mathrm{d}\chi'^2 = \frac{\partial(\chi'^1, \chi'^2)}{\partial(\chi^1, \chi^2)}\, \mathrm{d}\chi^1 \wedge \mathrm{d}\chi^2.$$

Furthermore, the Jacobian determinant appearing in the last equation must be a function of χ^1 and χ^2 only and, therefore, is a constant of motion (cf. Exercise 3.18). Thus, when $\dim P = 2$, the 2-form Ω is not unique, but is defined up to multiplicative constant of motion (see Example 8.44, below).

We can give another proof of the assertion above, which allows us to find the 2-forms Ω explicitly, without assuming that we know explicitly all the constants of motion.

In terms of an *arbitrary* coordinate system (x, y, t) on $P \times \mathbb{R}$, the vector field $\mathbf{A}$ can be written as

$$\mathbf{A} = f\frac{\partial}{\partial x} + g\frac{\partial}{\partial y} + \frac{\partial}{\partial t}, \tag{8.123}$$

where $f, g \in C^\infty(P \times \mathbb{R})$ are given, and any 2-form on $P \times \mathbb{R}$ has the local expression

$$\Omega = 2\Omega_{12}\,\mathrm{d}x \wedge \mathrm{d}y + 2\Omega_{13}\,\mathrm{d}x \wedge \mathrm{d}t + 2\Omega_{23}\,\mathrm{d}y \wedge \mathrm{d}t,$$

for some $\Omega_{ij} \in C^\infty(P \times \mathbb{R})$.

The condition $\mathbf{A} \lrcorner \Omega = 0$ amounts to

$$0 = \mathbf{A} \lrcorner \Omega = 2\big[(-g\Omega_{12} - \Omega_{13})\,\mathrm{d}x + (f\Omega_{12} - \Omega_{23})\,\mathrm{d}y + (f\Omega_{13} + g\Omega_{23})\,\mathrm{d}t\big];$$

hence, $\Omega_{13} = -g\Omega_{12}$ and $\Omega_{23} = f\Omega_{12}$, that is,

$$\Omega = 2\Omega_{12}\big[\mathrm{d}x \wedge \mathrm{d}y + (f\,\mathrm{d}y - g\,\mathrm{d}x) \wedge \mathrm{d}t\big], \tag{8.124}$$

where only the function Ω_{12} remains unspecified. The rank of Ω can only be 0 or 2; therefore, if $\Omega_{12} \neq 0$, the rank of Ω is equal to 2. Finally, from the condition $\mathrm{d}\Omega = 0$ one readily finds that the function Ω_{12} has to satisfy the linear PDE

$$\left(f\frac{\partial}{\partial x} + g\frac{\partial}{\partial y} + \frac{\partial}{\partial t}\right)\Omega_{12} = -\left(\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y}\right)\Omega_{12}. \tag{8.125}$$

The non-uniqueness of Ω comes from the fact that (8.125) has infinitely many solutions; if Ω_{12} and Ω'_{12} are two solutions of (8.125), setting $\nu \equiv \Omega'_{12}/\Omega_{12}$, one finds that $\mathbf{A}\nu = 0$, i.e., ν is a first integral.

Once Ω_{12} satisfies (8.125), the 2-form (8.124) can be written in the "canonical form" $\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t$, introducing two auxiliary functions $\phi, \psi \in C^\infty(P \times \mathbb{R})$ such that

$$\frac{\partial}{\partial x}\big[(f - \psi)\,\Omega_{12}\big] + \frac{\partial}{\partial y}\big[(g - \phi)\,\Omega_{12}\big] = 0. \tag{8.126}$$

This condition guarantees the local existence of a function $H \in C^\infty(P \times \mathbb{R})$ such that

$$2\Omega_{12}\big[(f - \psi)\,\mathrm{d}y - (g - \phi)\,\mathrm{d}x\big] = -\mathrm{d}H + \text{terms proportional to } \mathrm{d}t. \tag{8.127}$$

Thus

$$\Omega = 2\Omega_{12}(\mathrm{d}x - \psi\,\mathrm{d}t) \wedge (\mathrm{d}y - \phi\,\mathrm{d}t) - \mathrm{d}H \wedge \mathrm{d}t.$$

Since Ω and $\mathrm{d}H \wedge \mathrm{d}t$ are closed forms, $2\Omega_{12}(\mathrm{d}x - \psi\,\mathrm{d}t) \wedge (\mathrm{d}y - \phi\,\mathrm{d}t)$ is closed and by virtue of the Darboux Theorem, there exist functions p, q such that $2\Omega_{12}(\mathrm{d}x - \psi\,\mathrm{d}t) \wedge (\mathrm{d}y - \phi\,\mathrm{d}t) = \mathrm{d}p \wedge \mathrm{d}q$, so that $\Omega = \mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t$. These results are summarized in the following proposition.

Proposition 8.41 *Let P be a differentiable manifold of dimension two. Given a vector field $\mathbf{A}$ on $P \times \mathbb{R}$ such that $\mathbf{A}t = 1$, locally there exist infinitely many rank 2, closed 2-forms Ω such that $\mathbf{A} \lrcorner \Omega = 0$. Any pair of such 2-forms, Ω, Ω', are related by $\Omega' = \nu\Omega$, where ν is a real-valued function satisfying $\mathbf{A}\nu = 0$ (i.e., ν is constant along the integral curves of $\mathbf{A}$). For each Ω, locally there exist coordinates*

(q, p, t) on $P \times \mathbb{R}$, where t is the natural coordinate of $\mathbb{R}$, and some function $H \in C^\infty(P \times \mathbb{R})$, defined up to canonical transformations, such that $\Omega = \mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t$.

Example 8.42 The system of first-order ODEs

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{ty - x}{t - 1}$$

corresponds to the linear second-order ODE

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} - \frac{t}{t-1}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{x}{t-1} = 0$$

and to the integral curves of the vector field

$$\mathbf{A} = y\frac{\partial}{\partial x} + \frac{ty - x}{t-1}\frac{\partial}{\partial y} + \frac{\partial}{\partial t}$$

[cf. (8.123)], that is, $f = y$ and $g = (ty - x)/(t-1)$. Therefore, the component Ω_{12} must satisfy the PDE [see (8.125)]

$$\left(y\frac{\partial}{\partial x} + \frac{ty - x}{t-1}\frac{\partial}{\partial y} + \frac{\partial}{\partial t}\right)\Omega_{12} = -\frac{t}{t-1}\Omega_{12}.$$

A solution to this equation is

$$\Omega_{12} = \frac{1}{2(t-1)\,\mathrm{e}^t}$$

(the factor $1/2$ is included for later convenience).

Condition (8.126) is satisfied choosing $\phi = ty/(t-1)$, $\psi = 0$, and substituting these expressions into (8.127) we find that, up to an additive function of t only,

$$H = \frac{1}{(1-t)\,\mathrm{e}^t}\left(\frac{y^2}{2} + \frac{x^2}{2(t-1)}\right)$$

and

$$\begin{aligned}
&2\Omega_{12}(\mathrm{d}x - \psi\,\mathrm{d}t) \wedge (\mathrm{d}y - \phi\,\mathrm{d}t) \\
&\quad = \frac{1}{(1-t)\,\mathrm{e}^t}\left(\mathrm{d}y + \frac{ty}{1-t}\,\mathrm{d}t\right) \wedge \mathrm{d}x \\
&\quad = \mathrm{d}\left(\frac{y}{(1-t)\,\mathrm{e}^t}\right) \wedge \mathrm{d}x.
\end{aligned}$$

Hence, we can take $p = y/[(1-t)\mathrm{e}^t]$ and $q = x$. Further examples can be found in Torres del Castillo and Rubalcava-García (2006) and Torres del Castillo (2009).

Exercise 8.43 Show that the vector fields

$$e^t \frac{\partial}{\partial x} + e^t \frac{\partial}{\partial y}, \qquad t\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$$

are symmetries of the 2-form Ω found in Example 8.42. Find the first integrals associated with these symmetries and their Poisson bracket. Show explicitly that Ω is proportional to the exterior product of the differentials of these two first integrals. (Note that all this can be done making use of the original coordinates (x, y, t).)

Example 8.44 In the case of a one-dimensional harmonic oscillator, the standard Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{m{\omega_0}^2 q^2}{2},$$

where q, p are canonical coordinates induced by a coordinate x on the configuration space, m is the mass of the oscillator and ω_0 is its angular frequency. The expression $H = (p/\sqrt{2m})^2 + (\sqrt{m/2}\,\omega_0 q)^2$ corresponds to the square of the distance from the origin to the point $(p/\sqrt{2m}, \sqrt{m/2}\,\omega_0 q)$ of $\mathbb{R}^2$. Then, the analogs of the polar coordinates are

$$P = \sqrt{H}, \qquad Q = \arctan \frac{\sqrt{m/2}\,\omega_0 q}{p/\sqrt{2m}} = \arctan \frac{m\omega_0 q}{p}$$

and one finds that

$$P\,\mathrm{d}P \wedge \mathrm{d}Q = \mathrm{d}\left(\frac{p}{\sqrt{2m}}\right) \wedge \mathrm{d}\left(\sqrt{\frac{m}{2}}\,\omega_0 q\right) = \frac{\omega_0}{2}\,\mathrm{d}p \wedge \mathrm{d}q.$$

Hence,

$$\begin{aligned}\mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t &= \frac{2}{\omega_0} P\,\mathrm{d}P \wedge \mathrm{d}Q - \mathrm{d}\big(P^2\big) \wedge \mathrm{d}t \\ &= \frac{2P}{\omega_0}(\mathrm{d}P \wedge \mathrm{d}Q - \omega_0\,\mathrm{d}P \wedge \mathrm{d}t).\end{aligned}$$

The factor $2P/\omega_0 = 2\sqrt{H}/\omega_0$ appearing on the right-hand side of the last equation is a constant of motion and the function $\omega_0 P$ is a Hamiltonian for the new canonical coordinates (Q, P).

Thus, in accordance with Proposition 8.41, the 2-form $\Omega' = \mathrm{d}P \wedge \mathrm{d}Q - \omega_0\,\mathrm{d}P \wedge \mathrm{d}t$ differs from $\Omega = \mathrm{d}p \wedge \mathrm{d}q - \mathrm{d}H \wedge \mathrm{d}t$ by a factor which is a constant of motion. Finally, it may be noticed that

$$\Omega' = \mathrm{d}P \wedge \mathrm{d}(Q - \omega_0 t).$$

Hence, both P and $Q - \omega_0 t$ are constants of motion.

Appendix A
Lie Algebras

Definition A.1 A *Lie algebra*, L, over a field $\mathbb{K}$, is a vector space over $\mathbb{K}$ which possesses a mapping from $L \times L$ into L, usually denoted by $[\ ,\ ]$, such that

(i) it is bilinear

$$[u, av + bw] = a[u, v] + b[u, w], \tag{A.1}$$

$$[au + bv, w] = a[u, w] + b[v, w], \tag{A.2}$$

for $u, v, w \in L$, $a, b \in \mathbb{K}$,

(ii) it is skew-symmetric

$$[u, v] = -[v, u], \tag{A.3}$$

for $u, v \in L$ (by virtue of (A.3), the linearity of the bracket on the second argument (A.1) implies its linearity on the first argument (A.2), and vice versa),

(iii) it satisfies the Jacobi identity

$$\big[u, [v, w]\big] + \big[v, [w, u]\big] + \big[w, [u, v]\big] = 0, \tag{A.4}$$

for $u, v, w \in L$. A Lie algebra is Abelian if $[u, v] = 0$ for $u, v \in L$.

Let L be a Lie algebra of finite dimension (that is, L is a vector space of finite dimension), and let $\{e_i\}_{i=1}^n$ be a basis of L. Owing to the bilinearity of the bracket, the value of $[u, v]$, for $u, v \in L$ arbitrary, is determined by the values of $[e_i, e_j]$ ($i, j = 1, \ldots, n$), for if $u = u^i e_i$ and $v = v^j e_j$, we have $[u, v] = [u^i e_i, v^j e_j] = u^i v^j [e_i, e_j]$.

Since $[e_i, e_j]$ must belong to L, $[e_i, e_j] = c_{ij}^k e_k$, where c_{ij}^k ($i, j, k = 1, \ldots, n$) are n^3 scalars, called the *structure constants* of L. The values of the structure constants are not independent, since the bracket must be skew-symmetric, and it satisfies the Jacobi identity, which imposes the following relations among the c_{ij}^k:

$$c_{ij}^k = -c_{ji}^k \quad \text{and} \tag{A.5}$$

$$c_{ij}^m c_{km}^l + c_{jk}^m c_{im}^l + c_{ki}^m c_{jm}^l = 0. \tag{A.6}$$

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2,

Exercise A.2 Let V be a vector space and let $\mathfrak{gl}(V)$ be the set of the linear maps from V to V with the usual sum and multiplication by scalars, and with the bracket given by $[A, B] \equiv AB - BA$. Show that $\mathfrak{gl}(V)$ is a Lie algebra. If V is of finite dimension and $\{e_i\}_{i=1}^n$ is a basis of V, the linear transformations ϕ^i_j defined by $\phi^i_j(e_k) \equiv \delta^i_k e_j$, form a basis of $\mathfrak{gl}(V)$. Show that $[\phi^j_i, \phi^l_k] = (\delta^r_i \delta^j_k \delta^l_s - \delta^l_i \delta^r_k \delta^j_s)\phi^s_r$.

Definition A.3 Let L be a Lie algebra. A *subalgebra*, M, of L is a subset of L which is a Lie algebra with the operations inherited from L.

Since most of the properties that define a Lie algebra are automatically satisfied by any subset of a given algebra (for instance, the bilinearity and skew-symmetry of the bracket), it suffices to employ the criterion given by the following theorem in order to show that some subset is or is not a subalgebra.

Theorem A.4 *Let L be a Lie algebra and let $M \subset L$. M is a subalgebra of L if and only if for $u, v \in M$ and $a \in \mathbb{K}$, the elements $u + v$, au and $[u, v]$ belong to M.*

The proof of this theorem is immediate and is left to the reader.

Definition A.5 Let L be a Lie algebra and M a subalgebra of L. M is an *ideal* of L if for $u \in M$ and $v \in L$, $[u, v] \in M$.

L itself and $\{0\}$ are ideals of L, and if L is Abelian, then any subalgebra of L is invariant.

Definition A.6 A Lie algebra, L, is *simple* if it is not Abelian and does not possess other ideals apart from L and $\{0\}$. L is *semisimple* if the only Abelian ideal contained in L is $\{0\}$.

For example, the set of globally Hamiltonian vector fields of a symplectic manifold is an ideal of the Lie algebra of the locally Hamiltonian vector fields (see Sect. 8.2).

Definition A.7 Let L_1 and L_2 be two Lie algebras over the same field $\mathbb{K}$. A map $f : L_1 \to L_2$ is a *Lie algebra homomorphism* if

(i) f is a linear transformation (i.e., $f(au + bv) = af(u) + bf(v)$, for $u, v \in L_1$, $a, b \in \mathbb{K}$) and
(ii) $f([u, v]) = [f(u), f(v)]$, for $u, v \in L_1$.

If, in addition, f is bijective we say that f is a *Lie algebra isomorphism*.

Exercise A.8 Let $f : L_1 \to L_2$ be a Lie algebra homomorphism. Show that $\operatorname{Ker} f \equiv \{u \in L_1 \mid f(u) = 0\}$ is an ideal of L_1.

Appendix B
Invariant Metrics

Any Lie group can be turned into a Riemannian manifold in such a way that all the left translations L_g (or the right translations R_g) are isometries. Let G be a Lie group and let $\{\omega^1, \ldots, \omega^n\}$ be a basis for the left-invariant 1-forms; if (a_{ij}) is any (constant) non-singular symmetric $n \times n$ matrix, then

$$a_{ij}\omega^i \otimes \omega^j \tag{B.1}$$

is a metric tensor on G, which is a *left-invariant metric* since $L_g{}^*(a_{ij}\omega^i \otimes \omega^j) = a_{ij}\omega^i \otimes \omega^j$, for all $g \in G$. If (a_{ij}) is positive definite, the metric (B.1) is also positive definite. If, in place of the 1-forms ω^i we employ right-invariant 1-forms, in an analogous manner we obtain a *right-invariant metric*. A metric on G is *bi-invariant* if it is left-invariant and right-invariant simultaneously.

From the results of Sect. 7.5 it follows that the right-invariant vector fields are Killing vector fields for any left-invariant metric (see Exercise 7.51). For a bi-invariant metric, the right-invariant vector fields, and the left-invariant vector fields are Killing vector fields.

Example B.1 The 2×2 real matrices of the form $\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$, with $x > 0$, form a Lie subgroup of $\mathrm{GL}(2, \mathbb{R})$. Making use of Theorem 7.35, from the equation

$$\begin{aligned}
\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{d}x & \mathrm{d}y \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} x^{-1} & -yx^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathrm{d}x & \mathrm{d}y \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} x^{-1}\,\mathrm{d}x & x^{-1}\,\mathrm{d}y \\ 0 & 0 \end{pmatrix} \\
&= x^{-1}\mathrm{d}x \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + x^{-1}\mathrm{d}y \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

it follows that

$$\omega^1 \equiv x^{-1}\,\mathrm{d}x, \omega \qquad ^2 \equiv x^{-1}\,\mathrm{d}y,$$

form a basis for the left-invariant 1-forms. Using the fact that the inversion mapping, $\iota(g) = g^{-1}$, is given by $\iota^* x = x^{-1}$, $\iota^* y = -yx^{-1}$ [see (7.3)], one finds that the basis

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2,

of the right-invariant 1-forms $\dot{\omega}^i = -\iota^*\omega^i$, is

$$\dot{\omega}^1 = x^{-1}\,\mathrm{d}x, \qquad \dot{\omega}^2 = -yx^{-1}\,\mathrm{d}x + \mathrm{d}y,$$

and the dual basis is given by

$$\dot{\mathbf{X}}_1 = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}, \qquad \dot{\mathbf{X}}_2 = \frac{\partial}{\partial y} \tag{B.2}$$

(cf. Example 7.47). Thus, $\dot{\mathbf{X}}_1$ and $\dot{\mathbf{X}}_2$ are Killing vector fields for the metric $a_{ij}\omega^i \otimes \omega^j = x^{-2}\big[a_{11}\,\mathrm{d}x \otimes \mathrm{d}x + a_{12}(\mathrm{d}x \otimes \mathrm{d}y + \mathrm{d}y \otimes \mathrm{d}x) + a_{22}\,\mathrm{d}y \otimes \mathrm{d}y\big]$, no matter what the values are of the constants a_{11}, a_{12}, and a_{22}. In particular, taking $a_{ij} = \delta_{ij}$, we obtain the metric

$$x^{-2}(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y), \tag{B.3}$$

which is the metric of Poincaré's half-plane [see (6.19)] and possesses three linearly independent Killing vector fields (see Example 6.12).

Exercise B.2 Show that if G is connected, the metric $a_{ij}\omega^i \otimes \omega^j$ is also right-invariant if and only if

$$a_{im}c^m_{jk} + a_{jm}c^m_{ik} = 0, \tag{B.4}$$

where the c^i_{jk} are the structure constants of G with respect to the basis $\{\omega^i\}$.

Exercise B.3 Find a basis for the left-invariant 1-forms and its dual basis for the group formed by the 3×3 matrices of the form

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \qquad x, y, z \in \mathbb{R},$$

which is related to the Heisenberg group [see, e.g., Baker (2002), Sect. 7.7]. Determine the structure constants of the group in this basis. Is it possible to find a bi-invariant metric?

Since the coefficients a_{ij} in (B.1) are constant, the dual basis $\{\mathbf{X}_i\}$ to $\{\omega^i\}$ is a rigid basis with respect to the metric $a_{ij}\omega^i \otimes \omega^j$; thus, comparing $[\mathbf{X}_i, \mathbf{X}_j] = c^k_{ij}\mathbf{X}_k$ with (6.62) one finds that $c^k_{ij} = \Gamma^k{}_{ji} - \Gamma^k{}_{ij} \equiv 2\Gamma^k{}_{[ji]}$, where the $\Gamma^i{}_{jk}$ are the Ricci rotation coefficients for the basis $\{\mathbf{X}_i\}$. Using the identity (6.63), we obtain

$$\Gamma_{ijk} = \frac{1}{2}\big(a_{im}c^m_{kj} - a_{jm}c^m_{ki} - a_{km}c^m_{ji}\big). \tag{B.5}$$

The foregoing expression is simplified if the metric (B.1) is bi-invariant because in that case the last two terms on the right-hand side of (B.5) cancel [see (B.4)], leaving

$$\Gamma_{ijk} = \frac{1}{2}a_{im}c^m_{kj}, \tag{B.6}$$

so that the connection and curvature forms in this basis are

$$\Gamma^i{}_j = \frac{1}{2} c^i_{kj} \omega^k \quad \text{and} \quad \mathscr{R}^i{}_j = \frac{1}{4} c^m_{jk} c^i_{ml} \omega^k \wedge \omega^l, \tag{B.7}$$

respectively [see (5.26), (7.34), and (A.6)]. Hence, the components of the curvature with respect to the basis $\{\mathbf{X}_i\}$ are

$$R^i{}_{jkl} = \frac{1}{4}\left(c^m_{jk} c^i_{ml} - c^m_{jl} c^i_{mk}\right) = -\frac{1}{4} c^i_{mj} c^m_{kl}. \tag{B.8}$$

It may be noticed that in the expressions (B.7) the matrix (a_{ij}) does not appear and, furthermore, that they make sense *independently* of choosing a metric on the group. It can be directly verified that, with respect to a basis for the left-invariant 1-forms, $\{\omega^1, \ldots, \omega^n\}$, the connection 1-forms (B.7) define a connection with torsion equal to zero. Hence, in any Lie group there exists a torsion-free connection, defined in a natural way, without having to specify a Riemannian metric.

From (B.6) it follows that, if the metric (B.1) is bi-invariant, the coefficients Γ_{ijk} are totally skew-symmetric, since, in general, $\Gamma_{ijk} = -\Gamma_{jik}$, while from the relation $c^m_{kj} = -c^m_{jk}$ it follows that $\Gamma_{ijk} = -\Gamma_{ikj}$. Combining these formulas one finds that $\Gamma_{ijk} = -\Gamma_{kji}$. If the dimension of G is two, then the total skew-symmetry of the Ricci rotation coefficients implies that they are equal to zero and, since (a_{im}) must be invertible, $c^m_{kj} = 0$ and, therefore, G must be Abelian.

If the dimension of G is three, the skew-symmetry of Γ_{ijk} implies that $\Gamma_{ijk} = b\,\varepsilon_{ijk}$, where b is some constant. Then, from (B.6), we have

$$c^m_{kj} = 2a^{im} b\,\varepsilon_{ijk}, \tag{B.9}$$

where (a^{im}) is the inverse of the matrix (a_{im}); therefore

$$\frac{1}{4} c^i_{mj} c^m_{kl} = b^2 a^{pi}\,\varepsilon_{pjm} a^{qm}\,\varepsilon_{qlk} = b^2 a^{pi} \det\left(a^{rs}\right) (a_{lp} a_{kj} - a_{lj} a_{kp})$$

and from (B.8) we obtain

$$R_{ijkl} = b^2 \det\left(a^{rs}\right) (a_{ik} a_{jl} - a_{il} a_{jk}), \tag{B.10}$$

which means that G is a constant curvature space (see Examples B.6 and B.8). For any value of b, the structure constants (B.9) satisfy the Jacobi identity (A.6). It can be noticed that in this case, if the six vector fields $\mathbf{X}_i$ and $\dot{\mathbf{X}}_i$ $(i = 1, 2, 3)$ are linearly independent, then they form a basis for the Killing vector fields of G, since the maximum dimension of the Lie algebra of the Killing vector fields of a Riemannian manifold of dimension n is $n(n+1)/2$.

Exercise B.4 Show that for any Lie group, G, the left-invariant vector fields $\mathbf{X}_i$, and the right-invariant vector fields $\dot{\mathbf{X}}_i$ are linearly independent if and only if the *center* of the Lie algebra of G is $\{0\}$; that is, if and only if zero is the only element of $\mathfrak{g}$ whose Lie bracket with all the elements of the algebra is equal to zero.

Exercise B.5 Show that if $a_{ij}\omega^i \otimes \omega^j$ is a bi-invariant metric on G, where $\{\omega^i\}$ is a basis for the left-invariant 1-forms, then $\nabla_{\mathbf{Y}}\mathbf{Z} = \frac{1}{2}[\mathbf{Y},\mathbf{Z}]$ and $R(\mathbf{Y},\mathbf{Z})\mathbf{W} = -\frac{1}{4}[[\mathbf{Y},\mathbf{Z}],\mathbf{W}]$, for $\mathbf{Y},\mathbf{Z},\mathbf{W} \in \mathfrak{g}$, where ∇ denotes the Riemannian connection associated with the bi-invariant metric and R is its curvature tensor. Show that the integral curves of any left-invariant vector field are geodesics.

If c^i_{jk} denote the structure constants of an arbitrary Lie algebra, then the constants

$$g_{ij} = -c^k_{im}c^m_{jk} \tag{B.11}$$

form a symmetric matrix, $g_{ji} = -c^k_{jm}c^m_{ik} = -c^m_{ik}c^k_{jm} = g_{ij}$. Furthermore, making use of (B.11), and the identities $c^m_{ij}c^l_{mk} + c^m_{jk}c^l_{mi} + c^m_{ki}c^l_{mj} = 0$ and $c^i_{jk} = -c^i_{kj}$ [see (A.5) and (A.6)] one finds that

$$\begin{aligned} g_{im}c^m_{jk} + g_{jm}c^m_{ik} &= -c^s_{ir}c^r_{ms}c^m_{jk} - c^s_{jr}c^r_{ms}c^m_{ik} \\ &= c^s_{ir}\big(c^m_{ks}c^r_{mj} + c^m_{sj}c^r_{mk}\big) - c^s_{jr}c^r_{ms}c^m_{ik} \\ &= \big(-c^s_{ki}c^m_{rs} - c^s_{rk}c^m_{is}\big)c^r_{mj} + c^s_{ir}c^m_{sj}c^r_{mk} - c^s_{jr}c^r_{ms}c^m_{ik} \\ &= 0. \end{aligned}$$

Hence, if G is a connected Lie group and $\{\omega^i\}$ is a basis for the left-invariant 1-forms, the tensor field $g_{ij}\omega^i \otimes \omega^j$, with the g_{ij} defined by (B.5), is bi-invariant (see Exercise B.2). However, the matrix (g_{ij}) can be singular, and therefore $g_{ij}\omega^i \otimes \omega^j$ does not need to be a Riemannian metric on G. It can be shown that the matrix (g_{ij}), defined in (B.11), is invertible if and only if the Lie algebra is *semisimple* (that is, it does not have Abelian proper ideals) [see, e.g., Sattinger and Weaver (1986, Chap. 9)].

It may be noticed that the components of the Ricci tensor associated with the curvature tensor (B.8) are given by $R_{ij} = \frac{1}{4}g_{ij}$, with g_{ij} defined by (B.11).

Example B.6 Let us consider the group $G = \mathrm{SU}(2)$ with the parametrization given by the Euler angles, ϕ, θ, ψ,

$$g = \big(\exp\phi(g)\mathbf{X}_3\big)\big(\exp\theta(g)\mathbf{X}_1\big)\big(\exp\psi(g)\mathbf{X}_3\big), \tag{B.12}$$

where $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ is the basis of $\mathfrak{su}(2)$ given in Exercise 7.19 [cf. (8.94)]. From (7.54) it follows that (B.12) is equivalent to

$$\begin{aligned} g &= \begin{pmatrix} \mathrm{e}^{\mathrm{i}\phi/2} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\phi/2} \end{pmatrix} \begin{pmatrix} \cos\theta/2 & \mathrm{i}\sin\theta/2 \\ \mathrm{i}\sin\theta/2 & \cos\theta/2 \end{pmatrix} \begin{pmatrix} \mathrm{e}^{\mathrm{i}\psi/2} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\psi/2} \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{e}^{\mathrm{i}(\phi+\psi)/2}\cos\theta/2 & \mathrm{i}\,\mathrm{e}^{\mathrm{i}(\phi-\psi)/2}\sin\theta/2 \\ \mathrm{i}\,\mathrm{e}^{\mathrm{i}(\psi-\phi)/2}\sin\theta/2 & \mathrm{e}^{-\mathrm{i}(\phi+\psi)/2}\cos\theta/2 \end{pmatrix}, \end{aligned} \tag{B.13}$$

where, by abuse of notation, we have simply written ϕ, θ, ψ, in place of $\phi(g), \theta(g)$, and $\psi(g)$, respectively. As in Example B.1, we can make use of Theorem 7.35 to find the basis of the left-invariant 1-forms, dual to the basis $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$. Since the structure constants for the basis $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ are the same as those of the basis $\{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3\}$ of $\mathfrak{so}(3)$, equations (8.95)–(8.100) hold if $\mathbf{S}_i$ is replaced by $\mathbf{X}_i$; hence the set

$$\begin{aligned} \omega^1 &= \sin\theta \sin\psi \, \mathrm{d}\phi + \cos\psi \, \mathrm{d}\theta, \\ \omega^2 &= \sin\theta \cos\psi \, \mathrm{d}\phi - \sin\psi \, \mathrm{d}\theta, \\ \omega^3 &= \cos\theta \, \mathrm{d}\phi + \mathrm{d}\psi \end{aligned} \tag{B.14}$$

is the dual basis to $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$. Making use of the fact that $[\mathbf{X}_i, \mathbf{X}_j] = \sum_{k=1}^{3} \varepsilon_{ijk} \mathbf{X}_k = \delta^{kl} \varepsilon_{ijk} \mathbf{X}_l$, we have $c^l_{ij} = \delta^{kl} \varepsilon_{ijk}$; therefore, from (B.11), $g_{ij} = -\delta^{pk} \varepsilon_{imp} \delta^{qm} \varepsilon_{jkq} = -\delta^{pk} (\delta_{jp}\delta_{ki} - \delta_{ji}\delta_{kp}) = 2\delta_{ij}$, which is an invertible matrix and $g_{ij}\omega^i \otimes \omega^j = 2\delta_{ij}\omega^i \otimes \omega^j$. From (B.14) we then have

$$\begin{aligned} g_{ij}\omega^i \otimes \omega^j = 2\big[&\mathrm{d}\phi \otimes \mathrm{d}\phi + \mathrm{d}\theta \otimes \mathrm{d}\theta + \mathrm{d}\psi \otimes \mathrm{d}\psi \\ &+ \cos\theta (\mathrm{d}\phi \otimes \mathrm{d}\psi + \mathrm{d}\psi \otimes \mathrm{d}\phi)\big]. \end{aligned} \tag{B.15}$$

According to the foregoing results, we may conclude that the metric (B.15) is bi-invariant. As we shall show below, this metric is essentially the usual metric of the sphere S^3.

The underlying manifold of the group SU(2) can be identified with the sphere S^3 in the following manner. All the elements of SU(2) are of the form

$$\begin{pmatrix} x + \mathrm{i}y & z + \mathrm{i}w \\ -z + \mathrm{i}w & x - \mathrm{i}y \end{pmatrix}, \tag{B.16}$$

where x, y, z, w are real numbers such that $x^2 + y^2 + z^2 + w^2 = 1$. Hence, there is a one-to-one correspondence between the elements of SU(2) and the points of $\mathrm{S}^3 \equiv \{(x, y, z, w) \in \mathbb{R}^4 \,|\, x^2 + y^2 + z^2 + w^2 = 1\}$. From the expressions (B.13) and (B.16), separating the real and imaginary parts, one obtains a local expression for the inclusion of SU(2), or S^3, in $\mathbb{R}^4$ ($i : \mathrm{SU}(2) \to \mathbb{R}^4$), namely

$$\begin{aligned} i^*x &= \cos\tfrac{1}{2}\theta \cos\tfrac{1}{2}(\phi + \psi), & i^*y &= \cos\tfrac{1}{2}\theta \sin\tfrac{1}{2}(\phi + \psi), \\ i^*z &= -\sin\tfrac{1}{2}\theta \sin\tfrac{1}{2}(\phi - \psi), & i^*w &= \sin\tfrac{1}{2}\theta \cos\tfrac{1}{2}(\phi - \psi). \end{aligned}$$

The pullback under i of the usual metric of $\mathbb{R}^4$ is then

$$\begin{aligned} &i^*(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z + \mathrm{d}w \otimes \mathrm{d}w) \\ &\quad = \frac{1}{4}\big[\mathrm{d}\phi \otimes \mathrm{d}\phi + \mathrm{d}\theta \otimes \mathrm{d}\theta + \mathrm{d}\psi \otimes \mathrm{d}\psi + \cos\theta (\mathrm{d}\phi \otimes \mathrm{d}\psi + \mathrm{d}\psi \otimes \mathrm{d}\phi)\big], \end{aligned}$$

which, except for a factor 1/8, coincides with the metric (B.15). This means that the metric (B.15), which, as we have shown, is the metric of a constant curvature space, is essentially the standard metric of S^3 (which is, clearly, a constant curvature space). Moreover, the left-invariant vector fields $\mathbf{S}_i$ [given by (8.98)] and the right-invariant vector fields $\dot{\mathbf{S}}_i$ [given by (8.100)] of SU(2), are Killing vector fields for the metric (B.15) and, therefore, for S^3. Thus, the Lie algebra of the Killing vector fields of S^3, which is $\mathfrak{so}(4)$ [the Lie algebra of SO(4)], possesses the basis $\{\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, -\dot{\mathbf{S}}_1, -\dot{\mathbf{S}}_2, -\dot{\mathbf{S}}_3\}$, which satisfies the relations

$$\begin{aligned} [\mathbf{S}_i, \mathbf{S}_j] &= \sum_{k=1}^{3} \varepsilon_{ijk} \mathbf{S}_k, \\ \big[(-\dot{\mathbf{S}}_i), (-\dot{\mathbf{S}}_j)\big] &= \sum_{k=1}^{3} \varepsilon_{ijk} (-\dot{\mathbf{S}}_k), \\ \big[\mathbf{S}_i, (-\dot{\mathbf{S}}_j)\big] &= 0; \end{aligned} \tag{B.17}$$

hence, $\mathfrak{so}(4)$ is the direct sum of two copies of $\mathfrak{su}(2)$:

$$\mathfrak{so}(4) = \mathfrak{su}(2) \oplus \mathfrak{su}(2). \tag{B.18}$$

Each $g \in \mathrm{SU}(2)$ can be regarded as a point of S^3 (by expressing g in the form (B.16) and taking the corresponding x, y, z, w as the coordinates of a point of S^3), and for any $g_1 \in \mathrm{SU}(2)$, both L_{g_1} and R_{g_1} are isometries for the metric (B.15). Hence, if $(g_1, g_2) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$, the mapping $g \mapsto L_{g_1} R_{g_2} g = g_1 g g_2 = R_{g_2} L_{g_1} g$, from SU(2) onto SU(2), can be seen as an isometric map from S^3 onto S^3. In fact, it turns out that any isometry of S^3 that does not change the orientation is obtained in this manner, with g_1 and g_2 determined up to sign; if $(g_1, g_2) \in \mathrm{SU}(2) \times \mathrm{SU}(2)$, then $(-g_1, -g_2)$ also belongs to $\mathrm{SU}(2) \times \mathrm{SU}(2)$ and $L_{g_1} R_{g_2} = L_{-g_1} R_{-g_2}$. From the preceding discussion it also follows that any rotation about the origin in $\mathbb{R}^4$ can be represented in the form

$$\begin{pmatrix} x' + \mathrm{i}y' & z' + \mathrm{i}w' \\ -z' + \mathrm{i}w' & x' - \mathrm{i}y' \end{pmatrix} = g_1 \begin{pmatrix} x + \mathrm{i}y & z + \mathrm{i}w \\ -z + \mathrm{i}w & x - \mathrm{i}y \end{pmatrix} g_2 \tag{B.19}$$

[cf. (7.65)] with $g_1, g_2 \in \mathrm{SU}(2)$ determined up to sign. [This result is the counterpart of (B.18).]

Exercise B.7 Show that from (B.19) it follows directly that the transformation $(x, y, z, w) \mapsto (x', y', z', w')$ belongs to SO(4).

Example B.8 The functions $\alpha, \beta, \gamma : \mathrm{SL}(2, \mathbb{R}) \to \mathbb{R}$ defined by

$$g = \left(\exp \frac{1}{2}\alpha(g)\mathbf{X}_1\right)\left(\exp \frac{1}{2}\beta(g)(\mathbf{X}_2 + \mathbf{X}_3)\right)\left(\exp \frac{1}{2}\gamma(g)\mathbf{X}_1\right), \tag{B.20}$$

where $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ is the basis of $\mathfrak{sl}(2, \mathbb{R})$ given in Example 7.16, form a local coordinate system for $\mathrm{SL}(2, \mathbb{R})$, alternative to that defined by (7.4). From (B.20) and (7.51) we then have

$$
\begin{aligned}
g &= \begin{pmatrix} e^{\alpha/2} & 0 \\ 0 & e^{-\alpha/2} \end{pmatrix} \begin{pmatrix} \cosh\beta/2 & \sinh\beta/2 \\ \sinh\beta/2 & \cosh\beta/2 \end{pmatrix} \begin{pmatrix} e^{\gamma/2} & 0 \\ 0 & e^{-\gamma/2} \end{pmatrix} \\
&= \begin{pmatrix} e^{(\alpha+\gamma)/2}\cosh\beta/2 & e^{(\alpha-\gamma)/2}\sinh\beta/2 \\ e^{-(\alpha-\gamma)/2}\sinh\beta/2 & e^{-(\alpha+\gamma)/2}\cosh\beta/2 \end{pmatrix},
\end{aligned}
\tag{B.21}
$$

where we have written α, β, γ instead of $\alpha(g)$, $\beta(g)$, $\gamma(g)$ [cf. (B.13)]. The dual basis to $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$, expressed in terms of the coordinates α, β, γ, can be obtained making use of (B.20) and Theorem 7.35, which leads to [see (7.20)]

$$
\begin{aligned}
g^{-1}\,dg &= \frac{1}{2}\exp\left(-\frac{1}{2}\gamma\mathbf{X}_1\right)\exp\left(-\frac{1}{2}\beta(\mathbf{X}_2+\mathbf{X}_3)\right) \\
&\quad \cdot \lambda_1 \exp\left(\frac{1}{2}\beta(\mathbf{X}_2+\mathbf{X}_3)\right)\exp\left(\frac{1}{2}\gamma\mathbf{X}_1\right)d\alpha \\
&\quad + \frac{1}{2}\exp\left(-\frac{1}{2}\gamma\mathbf{X}_1\right)(\lambda_2+\lambda_3)\exp\left(\frac{1}{2}\gamma\mathbf{X}_1\right)d\beta + \frac{1}{2}\lambda_1\,d\gamma \\
&= \frac{1}{2}(\cosh\beta\,d\alpha + d\gamma)\,\lambda_1 + \frac{1}{2}e^{-\gamma}(\sinh\beta\,d\alpha + d\beta)\,\lambda_2 \\
&\quad + \frac{1}{2}e^{\gamma}(-\sinh\beta\,d\alpha + d\beta)\,\lambda_3,
\end{aligned}
$$

and thus

$$
\begin{aligned}
\omega^1 &= \frac{1}{2}(\cosh\beta\,d\alpha + d\gamma), \\
\omega^2 &= \frac{1}{2}e^{-\gamma}(\sinh\beta\,d\alpha + d\beta), \\
\omega^3 &= \frac{1}{2}e^{\gamma}(-\sinh\beta\,d\alpha + d\beta).
\end{aligned}
\tag{B.22}
$$

On the other hand, from (7.20) we find that $[\lambda_1, \lambda_2] = 2\lambda_2$, $[\lambda_2, \lambda_3] = \lambda_1$, $[\lambda_3, \lambda_1] = 2\lambda_3$ (i.e., the structure constants that are different from zero are given by $c^2_{12} = 2 = c^3_{31}$, $c^1_{23} = 1$) and from (B.11) it follows that

$$
(g_{ij}) = \begin{pmatrix} -8 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & -4 & 0 \end{pmatrix};
\tag{B.23}
$$

therefore, using (B.22) and (B.23),

$$
\begin{aligned}
g_{ij}\omega^i \otimes \omega^j &= -4\left(2\omega^1 \otimes \omega^1 + \omega^2 \otimes \omega^3 + \omega^3 \otimes \omega^2\right) \\
&= -2\left[d\alpha \otimes d\alpha + d\beta \otimes d\beta + d\gamma \otimes d\gamma \right. \\
&\quad \left. + \cosh\beta(d\alpha \otimes d\gamma + d\gamma \otimes d\alpha)\right]
\end{aligned}
\tag{B.24}
$$

is a pseudo-Riemannian bi-invariant metric on $\mathrm{SL}(2,\mathbb{R})$ and, with this metric, $\mathrm{SL}(2,\mathbb{R})$ is a constant curvature space. (Note that $g_{ij}\omega^i \otimes \omega^j = -8\omega^1 \otimes \omega^1 - 2(\omega^2+\omega^3)\otimes(\omega^2+\omega^3) + 2(\omega^2-\omega^3)\otimes(\omega^2-\omega^3)$, which explicitly shows that this metric is pseudo-Riemannian.) In a similar manner to the case of SU(2), considered in the foregoing example, $\mathrm{SL}(2,\mathbb{R})$ with the metric (B.24) can be identified with a submanifold of $\mathbb{R}^4$, provided that in the latter we introduce a pseudo-Riemannian flat metric [cf. Conlon (2001), Sect. 10.7].

Indeed, any element of $\mathrm{SL}(2,\mathbb{R})$ is of the form

$$\begin{pmatrix} x+w & y+z \\ z-y & x-w \end{pmatrix}, \tag{B.25}$$

where x, y, z, w are real numbers with $x^2+y^2-z^2-w^2=1$. This means that the underlying manifold of $\mathrm{SL}(2,\mathbb{R})$ can be identified with the hyperboloid $N \equiv \{(x,y,z,w)\in\mathbb{R}^4 \,|\, x^2+y^2-z^2-w^2=1\}$. Comparing (B.21) with (B.25), one finds the following local expression for the inclusion of $\mathrm{SL}(2,\mathbb{R})$ in $\mathbb{R}^4$:

$$i^*x = \cosh\frac{1}{2}\beta\cosh\frac{1}{2}(\alpha+\gamma), \qquad i^*y = \sinh\frac{1}{2}\beta\sinh\frac{1}{2}(\alpha-\gamma),$$

$$i^*z = \sinh\frac{1}{2}\beta\cosh\frac{1}{2}(\alpha-\gamma), \qquad i^*w = \cosh\frac{1}{2}\beta\sinh\frac{1}{2}(\alpha+\gamma),$$

hence, the metric induced on $\mathrm{SL}(2,\mathbb{R})$, or on N, by the pseudo-Riemannian metric $\mathrm{d}x\otimes\mathrm{d}x + \mathrm{d}y\otimes\mathrm{d}y - \mathrm{d}z\otimes\mathrm{d}z - \mathrm{d}w\otimes\mathrm{d}w$ of $\mathbb{R}^4$ is

$$\begin{aligned} &i^*(\mathrm{d}x\otimes\mathrm{d}x + \mathrm{d}y\otimes\mathrm{d}y - \mathrm{d}z\otimes\mathrm{d}z - \mathrm{d}w\otimes\mathrm{d}w) \\ &\quad = -\frac{1}{4}\big[\mathrm{d}\alpha\otimes\mathrm{d}\alpha + \mathrm{d}\beta\otimes\mathrm{d}\beta + \mathrm{d}\gamma\otimes\mathrm{d}\gamma + \cosh\beta(\mathrm{d}\alpha\otimes\mathrm{d}\gamma + \mathrm{d}\gamma\otimes\mathrm{d}\alpha)\big] \end{aligned}$$

and coincides, except for a factor 1/8, with the metric (B.24). Then, owing to the bi-invariance of (B.24), the left-invariant vector fields of $\mathrm{SL}(2,\mathbb{R})$, together with the right-invariant ones are Killing vector fields for the metric (B.24) and for the metric induced on N. On the other hand, N and the metric $\mathrm{d}x\otimes\mathrm{d}x + \mathrm{d}y\otimes\mathrm{d}y - \mathrm{d}z\otimes\mathrm{d}z - \mathrm{d}w\otimes\mathrm{d}w$ are invariant under the linear transformations of $\mathbb{R}^4$ into $\mathbb{R}^4$ represented by the real 4×4 matrices, A, with determinant equal to 1, such that

$$A^{\mathrm{t}}\begin{pmatrix} 1 &&& \\ & 1 && \\ && -1 & \\ &&& -1 \end{pmatrix} A = \begin{pmatrix} 1 &&& \\ & 1 && \\ && -1 & \\ &&& -1 \end{pmatrix}, \tag{B.26}$$

which form the group SO(2, 2), whose dimension is six. Thus, in an analogous way to (B.18), we have

$$\mathfrak{so}(2,2) = \mathfrak{sl}(2,\mathbb{R}) \oplus \mathfrak{sl}(2,\mathbb{R}). \tag{B.27}$$

Since for any $g \in \mathrm{SL}(2,\mathbb{R})$, the transformations L_g and R_g are isometries of the metric (B.24), if $(g_1, g_2) \in \mathrm{SL}(2,\mathbb{R}) \times \mathrm{SL}(2,\mathbb{R})$, the transformation $g \mapsto L_{g_1} R_{g_2} g = g_1 g g_2$, from $\mathrm{SL}(2,\mathbb{R})$ onto $\mathrm{SL}(2,\mathbb{R})$, is an isometry and can be identified with an isometric transformation from N onto N. That is, using (B.25), the expression

$$\begin{pmatrix} x'+w' & y'+z' \\ z'-y' & x'-w' \end{pmatrix} = g_1 \begin{pmatrix} x+w & y+z \\ z-y & x-w \end{pmatrix} g_2 \tag{B.28}$$

gives an isometric transformation from N onto N, for any pair of elements $g_1, g_2 \in \mathrm{SL}(2,\mathbb{R})$, and it turns out that any transformation belonging to $\mathrm{SO}(2,2)$ can be represented in this manner with g_1 and g_2 determined up to a common sign.

Harmonic Maps The harmonic mapping equations constitute a generalization of the geodesic equations (5.7). In their general form, given two Riemannian manifolds, N and M, of dimensions n and m, respectively, a differentiable map $\phi : N \to M$ is *harmonic* if

$$\frac{1}{\sqrt{|h|}} \frac{\partial}{\partial y^\alpha} \left(\sqrt{|h|}\, h^{\alpha\beta} \frac{\partial(\phi^* x^k)}{\partial y^\beta} \right) + \left(\phi^* \Gamma^k_{ji}\right) h^{\alpha\beta} \frac{\partial(\phi^* x^j)}{\partial y^\alpha} \frac{\partial(\phi^* x^i)}{\partial y^\beta} = 0, \tag{B.29}$$

where $(h^{\alpha\beta})$ is the inverse of the matrix $(h_{\alpha\beta})$, formed by the components of the metric tensor of N with respect to a local coordinate system $(y^1, \ldots, y^n)$, $h \equiv \det(h_{\alpha\beta})$, $(x^1, \ldots, x^m)$ is a coordinate system on M and the Γ^k_{ji} are the Christoffel symbols corresponding to the metric tensor of M in the coordinate system x^i [see, e.g., Hélein (2002)]. When $N = \mathbb{R}$, with $y^1 = t$ and $h_{11} = 1$, equations (B.29) reduce to the equations of the geodesics (5.7). When $M = \mathbb{R}$, with its usual metric, equations (B.29) reduce to the Laplace equation, $\nabla^2 \phi = 0$ [see (6.113)].

An interesting fact is that in the case of a harmonic map $\phi : N \to G$, where G is a Lie group that admits a bi-invariant metric, equations (B.29) amount to

$$\frac{1}{\sqrt{|h|}} \frac{\partial}{\partial y^\alpha} \left[\sqrt{|h|}\, h^{\alpha\beta} \left(\phi^* \omega^k\right) \left(\frac{\partial}{\partial y^\beta} \right) \right] = 0, \tag{B.30}$$

where the ω^k are left-invariant 1-forms on G. In effect, the 1-forms ω^k can be expressed locally in the form

$$\omega^k = M^k_i \, \mathrm{d}x^i, \tag{B.31}$$

with each $M^k_i \in C^\infty(G)$. Then

$$\frac{\partial}{\partial x^i} = M^k_i \mathbf{X}_k, \tag{B.32}$$

where the $\mathbf{X}_k$ are the left-invariant fields that form the dual basis to $\{\omega^k\}$. Using the properties of a connection [see (5.1)], from Exercise B.5 it follows that the Christoffel symbols for the bi-invariant metric of G with respect to the coordinate system x^i are given by

$$\begin{aligned}\Gamma^i_{jk}\frac{\partial}{\partial x^i} &= M^s_k \nabla_{\mathbf{X}_s}\big(M^r_j \mathbf{X}_r\big)\\ &= M^s_k \mathbf{X}_s\big(M^r_j\big)\mathbf{X}_r + M^s_k M^r_j \nabla_{\mathbf{X}_s}\mathbf{X}_r\\ &= \left(\frac{\partial}{\partial x^k}M^r_j\right)\mathbf{X}_r + \frac{1}{2}M^s_k M^r_j[\mathbf{X}_s,\mathbf{X}_r].\end{aligned}$$

Since the Christoffel symbols Γ^i_{jk} are symmetric in the indices j, k, while $M^s_k M^r_j[\mathbf{X}_s,\mathbf{X}_r]$ is antisymmetric in these indices, using (B.32), it follows that

$$\Gamma^i_{jk} = \widetilde{M}^i_r \frac{\partial}{\partial x^{(k}} M^r_{j)}, \tag{B.33}$$

where $(\widetilde{M}^i_j)$ is the inverse of the matrix (M^i_j), and the parentheses denote symmetrization on the enclosed indices [e.g., $t_{(ij)} = \frac{1}{2}(t_{ij}+t_{ji})$].

Thus, from (B.31), (1.23), and (1.24) we have

$$\begin{aligned}&h^{\alpha\beta}\frac{\partial}{\partial y^\alpha}\left[\big(\phi^*\omega^k\big)\left(\frac{\partial}{\partial y^\beta}\right)\right]\\ &\quad= h^{\alpha\beta}\frac{\partial}{\partial y^\alpha}\left[\big(\phi^* M^k_i\big)\frac{\partial(\phi^* x^i)}{\partial y^\beta}\right]\\ &\quad= h^{\alpha\beta}\big(\phi^* M^k_i\big)\frac{\partial}{\partial y^\alpha}\frac{\partial(\phi^* x^i)}{\partial y^\beta} + h^{\alpha\beta}\frac{\partial(\phi^* x^i)}{\partial y^\beta}\frac{\partial(\phi^* x^j)}{\partial y^\alpha}\,\phi^*\left(\frac{\partial M^k_i}{\partial x^j}\right).\end{aligned}$$

Using the fact that $(h^{\alpha\beta})$ is symmetric, from (B.33) we then have

$$\begin{aligned}&h^{\alpha\beta}\frac{\partial}{\partial y^\alpha}\left[\big(\phi^*\omega^k\big)\left(\frac{\partial}{\partial y^\beta}\right)\right]\\ &\quad= \big(\phi^* M^k_s\big)\left[h^{\alpha\beta}\frac{\partial}{\partial y^\alpha}\frac{\partial(\phi^* x^s)}{\partial y^\beta} + h^{\alpha\beta}\frac{\partial(\phi^* x^i)}{\partial y^\beta}\frac{\partial(\phi^* x^j)}{\partial y^\alpha}\phi^*\Gamma^s_{ij}\right],\end{aligned}$$

which shows the equivalence of (B.29) and (B.30) in the case where M is a Lie group with a bi-invariant metric.

As pointed out previously, when $N=\mathbb{R}$ with the usual metric, the equations for a harmonic map reduce to the geodesic equations. Hence, the equations for a geodesic, C, of a group G with a bi-invariant metric, can be expressed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\big(C^*\omega^k\big)\left(\frac{\partial}{\partial t}\right)\right] = 0$$

[see (B.30)]; therefore $(C^*\omega^k)(\partial/\partial t) = a^k$, where each a^k is a real constant. That is, $\omega^k(C'_t) = a^k$, which amounts to $C'_t = a^k\mathbf{X}_k(C(t))$. Thus, in this case, a geodesic is an integral curve of some left-invariant vector field (cf. Exercise B.5).

Taking into account that, when G is some subgroup of $\mathrm{GL}(p,\mathbb{R})$, a basis for the left-invariant 1-forms can be found from the relation $g^{-1}\mathrm{d}g = \lambda_a \omega^a$ [see (7.46)], where the λ_a are constant matrices that form a basis for a representation of the Lie algebra of G, it follows that equations (B.30) amount to the matrix equation

$$\frac{1}{\sqrt{|h|}}\frac{\partial}{\partial y^\alpha}\left(\sqrt{|h|}\,h^{\alpha\beta}g^{-1}\frac{\partial g}{\partial y^\beta}\right) = 0, \tag{B.34}$$

where it is understood that g is an arbitrary element of G, parameterized in terms of the y^α through the map $\phi : N \to G$.

Each Killing vector field of a Riemannian manifold, M, gives rise to a conserved quantity, constant of motion, or first integral of the geodesic equations (Theorem 6.28). This result can be extended to the equations for the harmonic maps: with each Killing vector field of a Riemannian manifold M and each harmonic map $\phi : N \to M$ one obtains a vector field on N whose divergence is equal to zero. (Such vector fields are called *conserved currents*.)

This assertion can be proved using (B.29), (1.23), and (1.24), denoting by K^i the components of a Killing vector field with respect to the coordinate system x^i and by g_{ij} the components of the metric tensor of M,

$$\begin{aligned}
&\frac{1}{\sqrt{|h|}}\frac{\partial}{\partial y^\alpha}\left[\sqrt{|h|}\,h^{\alpha\beta}\phi^*\big(g_{ik}K^i\big)\frac{\partial(\phi^* x^k)}{\partial y^\beta}\right] \\
&\quad = h^{\alpha\beta}\frac{\partial(\phi^* x^k)}{\partial y^\beta}\frac{\partial}{\partial y^\alpha}\phi^*\big(g_{ik}K^i\big) - \phi^*\big(g_{ik}K^i\big)\big(\phi^*\Gamma^k_{js}\big)h^{\alpha\beta}\frac{\partial(\phi^* x^j)}{\partial y^\alpha}\frac{\partial(\phi^* x^s)}{\partial y^\beta} \\
&\quad = h^{\alpha\beta}\frac{\partial(\phi^* x^k)}{\partial y^\beta}\frac{\partial(\phi^* x^s)}{\partial y^\alpha}\,\phi^*\left[\frac{\partial(g_{ik}K^i)}{\partial x^s} - \Gamma^i_{ks}g_{ij}K^j\right] \\
&\quad = 0,
\end{aligned} \tag{B.35}$$

where the last equality follows from (6.14) and (6.55), and the fact that the factor $h^{\alpha\beta}[\partial(\phi^* x^k)/\partial y^\beta][\partial(\phi^* x^s)/\partial y^\alpha]$ is symmetric in the indices k, s. The left-hand side of this equality is the divergence of the vector field

$$\mathbf{J} \equiv h^{\alpha\beta}\phi^*\big(g_{ik}K^i\big)\frac{\partial(\phi^* x^k)}{\partial y^\beta}\frac{\partial}{\partial y^\alpha}$$

[cf. (6.108)].

As pointed out at the beginning of this appendix, the left-invariant and the right-invariant vector fields are Killing vector fields for a Lie group with a bi-invariant metric; therefore, the relation (B.35) holds if the K^i are the components with respect to the coordinate system x^i of a left-invariant or right-invariant vector field, when M is a Lie group with a bi-invariant metric. In fact, the m relations (B.30), applicable in the case where M is a Lie group with a bi-invariant metric, are particular cases of (B.35).

Exercise B.9 Show that for each value of the index p, the functions $K^i = g^{ij} M_j^p$ are components of a Killing vector field with respect to the coordinate system x^i, where the M_j^i are the functions defined in (B.31). (In fact, they are components of a left-invariant vector field.) Show that the relations (B.30) follow from (B.35), using these m Killing vector fields.

References

Baker, A. (2002). *Matrix Groups: An Introduction to Lie Group Theory* (Springer, London).

Berndt, R. (2001). *An Introduction to Symplectic Geometry* (American Mathematical Society, Providence).

Boothby, W.M. (2002). *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd ed. (Academic Press, Amsterdam).

Born, M. and Wolf, E. (1999). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, 7th ed. (Cambridge University Press, Cambridge).

Conlon, L. (2001). *Differentiable Manifolds*, 2nd ed. (Birkhäuser, Boston).

Crampin, M. and Pirani, F.A.E. (1986). *Applicable Differential Geometry*, London Mathematical Society Lecture Notes, Vol. 59 (Cambridge University Press, Cambridge).

do Carmo, M. (1994). *Differential Forms and Applications* (Springer, Berlin).

do Carmo, M. (1992). *Riemannian Geometry* (Birkhäuser, Boston).

Erdmann, K. and Wildon, M.J. (2006). *Introduction to Lie Algebras* (Springer, London).

Fisher, S.D. (1999). *Complex Variables*, 2nd ed. (Dover, New York).

Goldstein, H. (1980). *Classical Mechanics*, 2nd ed. (Addison-Wesley, Reading).

Guillemin, V. and Pollack, A. (1974). *Differential Topology* (Prentice-Hall, Englewood Cliffs) (American Mathematical Society, Providence, reprinted 2010).

Hélein, F. (2002). *Harmonic Maps, Conservation Laws and Moving Frames*, 2nd ed. (Cambridge University Press, Cambridge).

Hirsch, M. and Smale, S. (1974). *Differential Equations, Dynamical Systems, and Linear Algebra* (Academic Press, New York).

Hochstadt, H. (1964). *Differential Equations: A Modern Approach* (Holt, Rinehart and Winston, New York) (Dover, New York, reprinted 1975).

Hydon, P.E. (2000). *Symmetry Methods for Differential Equations* (Cambridge University Press, Cambridge).

Isham, C.J. (1999). *Modern Differential Geometry for Physicists*, 2nd ed. (World Scientific, Singapore).

Jacobson, N. (1979). *Lie Algebras* (Dover, New York).

Lee, J.M. (1997). *Riemannian Manifolds: An Introduction to Curvature* (Springer, New York).

Lee, J.M. (2002). *Introduction to Smooth Manifolds* (Springer, New York).

O'Neill, B. (2006). *Elementary Differential Geometry*, 2nd ed. (Academic Press, New York).

Oprea, J. (1997). *Differential Geometry and Its Applications* (Prentice-Hall, Upper Saddle River).

Sattinger, D.H. and Weaver, O.L. (1986). *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics* (Springer, New York).

Sneddon, I.N. (2006). *Elements of Partial Differential Equations* (Dover, New York).

Stephani, H. (1989). *Differential Equations: Their Solution Using Symmetries* (Cambridge University Press, Cambridge).

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2,

Torres del Castillo, G.F. (2009). The Hamiltonian description of a second-order ODE, J. Phys. A, Math. Theor. **42**, 265202.

Torres del Castillo, G.F. and Rubalcava-García, I. (2006). Hamiltonians and Lagrangians of non-autonomous one-dimensional mechanical systems, Rev. Mex. Fís. **52**, 429.

Warner, F.W. (1983). *Foundations of Differentiable Manifolds and Lie Groups* (Springer, New York).

Woodhouse, N.M.J. (1997). *Geometric Quantization*, 2nd ed. (Oxford University Press, New York).

Index

G.F. Torres del Castillo, *Differentiable Manifolds*,
DOI 10.1007/978-0-8176-8271-2, © Springer Science+Business Media, LLC 2012

V

W