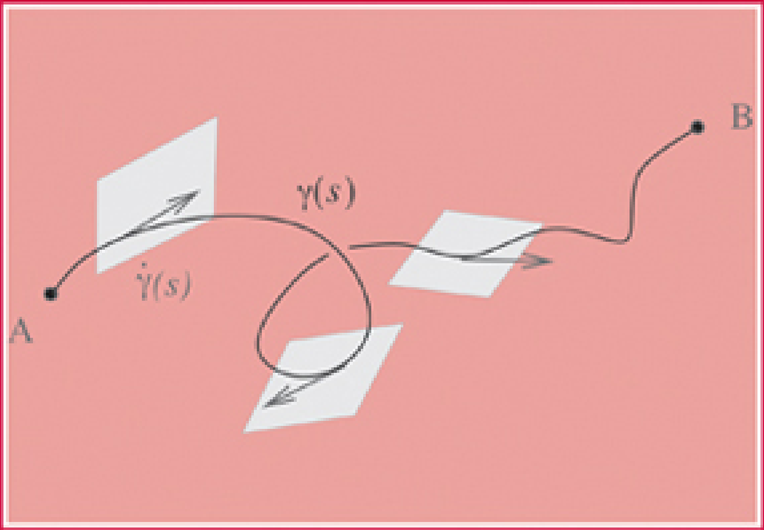
Encyclopedia of Mathematics and Its Applications 126

SUB-RIEMANNIAN GEOMETRY

General Theory and Examples

Ovidiu Calin
Der-Chen Chang



CAMBRIDGE

CAMBRIDGE
more information - www.cambridge.org/9780521897303

Sub-Riemannian Geometry

General Theory and Examples

Sub-Riemannian manifolds are manifolds with the Heisenberg principle built in. This comprehensive text and reference begins by introducing the theory of sub-Riemannian manifolds using a variational approach in which all properties are obtained from minimum principles, a robust method that is novel in this context. The authors then present examples and applications, showing how Heisenberg manifolds (step 2 sub-Riemannian manifolds) might in the future play a role in quantum mechanics similar to the role played by the Riemannian manifolds in classical mechanics.

Sub-Riemannian Geometry: General Theory and Examples is the perfect resource for graduate students and researchers in pure and applied mathematics, theoretical physics, control theory, and thermodynamics interested in the most recent developments in sub-Riemannian geometry.

Ovidiu Calin is an associate professor of mathematics at Eastern Michigan University and a former visiting assistant professor at the University of Notre Dame. He received his Ph.D. in geometric analysis from the University of Toronto in 2000. He has written several monographs and numerous research papers in the field of geometric analysis and has delivered research lectures in several universities in North America, Asia, the Middle East, and Eastern Europe.

Der-Chen Chang is a professor of mathematics at Georgetown University. He is a previous associate professor at the University of Maryland and a visiting professor at the Academia Sinica, among other institutions. He received his Ph.D. in Fourier analysis from Princeton University in 1987 and has authored several monographs and numerous research papers in the field of geometric analysis, several complex variables, and Fourier analysis.

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

FOUNDING EDITOR G.-C. ROTA
Editorial Board
R. Doran, P. Flajolet, M. Ismail, T.-Y. Lam, E. Lutwak

The titles below, and earlier volumes in the series, are available from booksellers or from Cambridge University Press at www.cambridge.org.

90 M. Lothaire *Algebraic Combinatorics on Words*
91 A. A. Ivanov and S. V. Shpectorov *Geometry of Sporadic Groups II*
92 P. McMullen and E. Schulte *Abstract Regular Polytopes*
93 G. Gierz et al. *Continuous Lattices and Domains*
94 S. Finch *Mathematical Constants*
95 Y. Jabri *The Mountain Pass Theorem*
96 G. Gasper and M. Rahman *Basic Hypergeometric Series, 2nd edn*
97 M. C. Pedicchio and W. Tholen (eds.) *Categorical Foundations*
98 M. E. H. Ismail *Classical and Quantum Orthogonal Polynomials in One Variable*
99 T. Mora *Solving Polynomial Equation Systems II*
100 E. Olivieri and M. E. Vares *Large Deviations and Metastability*
101 A. Kushner, V. Lychagin, and V. Rubtsov *Contact Geometry and Nonlinear Differential Equations*
102 L. W. Beineke, R. J. Wilson, and P. J. Cameron (eds.) *Topics in Algebraic Graph Theory*
103 O. Staffans *Well-Posed Linear Systems*
104 J. M. Lewis, S. Lakshmivarahan, and S. Dhall *Dynamic Data Assimilation*
105 M. Lothaire *Applied Combinatorics on Words*
106 A. Markoe *Analytic Tomography*
107 P. A. Martin *Multiple Scattering*
108 R. A. Brualdi *Combinatorial Matrix Classes*
109 J. Sakarovitch *Elements of Automata Theory*
110 M.-J. Lai and L. L. Schumaker *Spline Functions on Triangulations*
111 R. T. Curtis *Symmetric Generation of Groups*
112 H. Salzmann, T. Grundhöfer, H. Hähl, and R. Löwen *The Classical Fields*
113 S. Peszat and J. Zabczyk *Stochastic Partial Differential Equations with Lévy Noise*
114 J. Beck *Combinatorial Games*
115 I. Barreira and Y. Pesin *Nonuniform Hyperbolicity: Dynamics of Systems with Nonzero Lyapunov Exponents*
116 D. Z. Arov and H. Dym *J-Contractive Matrix Valued Functions and Related Topics*
117 R. Glowinski, J.-L. Lions, and J. He *Exact and Approximate Controllability for Distributed Parameter Systems*
118 A. A. Borovkov and K. A. Borovkov *Asymptotic Analysis of Random Walks*
119 M. Deza and M. Dutour Sikirić *Geometry of Chemical Graphs*
120 T. Nishiura *Absolute Measurable Spaces*
121 M. Prest *Purity, Spectra and Localisation*
122 S. Khrushchev *Orthogonal Polynomials and Continued Fractions: From Euler's Point of View*
123 H. Nagamochi and T. Ibaraki *Algorithmic Aspects of Graph Connectivity*
124 F. King *Hilbert Transforms: Volume 1*
125 F. King *Hilbert Transforms: Volume 2*

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Sub-Riemannian Geometry

General Theory and Examples

OVIDIU CALIN

Eastern Michigan University

DER-CHEN CHANG

Georgetown University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521897303

© Ovidiu Calin and Der-Chen Chang 2009

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2009

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Calin, Ovidiu.
Sub-Riemannian geometry : general theory and examples / Ovidiu Calin, Der-chen Chang.
p. cm. – (Encyclopedia of mathematics and its applications)
Includes bibliographical references and index.
ISBN 978-0-521-89730-3 (hardback)
1. Geometry, Riemannian. 2. Riemannian manifolds. 3. Geodesics (Mathematics)
4. Submanifolds. I. Chang, Der-chen E. II. Title. III. Title: Subriemannian geometry.
IV. Series.
QA649.C27 2009
516.3′73–dc22 2009000653

ISBN 978-0-521-89730-3 hardback

Cambridge University Press has no responsibility for the persistence or
accuracy of URLs for external or third-party Internet Web sites referred to in
this publication and does not guarantee that any content on such Web sites is,
or will remain, accurate or appropriate. Information regarding prices, travel
timetables, and other factual information given in this work are correct at
the time of first printing, but Cambridge University Press does not guarantee
the accuracy of such information thereafter.

Dedicated to
Professor Shing Tung Yau
on the occasion of his sixtieth birthday

“I don’t add hypotheses.
I derive everything from what is given.”
Isaac Newton

Contents

Preface

A few important discoveries in the field of thermodynamics in the 1800s made the first steps toward sub-Riemannian geometry. Carnot discovered the principle of an engine in 1824 involving two isotherms and two adiabatic processes, Jule studied adiabatic processes, and Clausius formulated the existence of the entropy in the second law of thermodynamics in 1854. In 1909 Carathéodory made the point regarding the relationship between the connectivity of two states by adiabatic processes and nonintegrability of a distribution, which is defined by the one-form of work. Chow proved the general global connectivity in 1934, and the same hypothesis was used by Hörmander in 1967 to prove the hypoellipticity of a sum of the squares of vector fields operators. However, the study of the invariants of a horizontal distribution, known as nonholonomic geometry, was initiated by the Romanian mathematician George Vranceanu in 1936.

The position of a ship on a sea is determined by three parameters: two coordinates x and y for the location and an angle to describe the orientation. Therefore, the position of a ship can be described by a point in a manifold. One can ask what is the shortest distance one should navigate to get from one position to another; this defines a Carnot–Carathéodory metric on the manifold $\mathbb{R}^2 \times \mathbb{S}^1$. In a similar way, a Carnot–Carathéodory metric can be defined on a general sub-Riemannian manifold. The study of sub-Riemannian geodesics is useful in determining the Carnot–Carathéodory distance between two points.

The study of the geometry of the Heisenberg group, which is the prototype of the sub-Riemannian geometry, was started by Gaveau in 1975. The understanding of the geometry of this group led Beals, Gaveau, and Greiner to characterize the fundamental solutions for heat-type subelliptic operators and Heisenberg sub-Laplacian operators in the 1990s. Meanwhile, many examples have been considered. Some of them have a behavior similar to the Heisenberg operator, but others do not. However, a unitary and general theory of these sub-Riemannian manifolds is still missing at the moment.

This book was written by the first author with the participation of the second. This work is mainly based on both the author's own recent research publications as well as a great deal of first author's unpublished work. It reflects the authors' best knowledge on the subject at the time it was written.

The main goal of Part I is to present a detailed analysis of the general theory of sub-Riemannian manifolds using Hamiltonian and Lagrangian formalism developed in the sub-Riemannian manifolds context. Other mathematical tools used are differential geometry, exterior differential systems, and the theory of elliptic functions.

Part II contains a rich collection of examples of sub-Riemannian manifolds of step 2 and higher, in which the computations can be done explicitly and a further precise study can be made. Each example involves different techniques, some of them involving elliptic integrals and hypergeometric functions. Some of these examples are computed here for the first time.

Why do we need a book on sub-Riemannian geometry? The authors believe the study of sub-Riemannian geometry helps with the understanding of subelliptic operators. A similar theory was developed between the Riemannian geometry and the elliptic operators. For instance, the heat kernel of a subelliptic operator depends on the geometry of the underlying horizontal distribution. It is known that for the case of bracket-generating distributions, any two points can be joined by piecewise horizontal curves. It is believed that the heat kernel is given by a path integral with respect to all horizontal curves joining the points x_0 and x in time t as in the formula $K(x_0, x; t) = \int_{\mathcal{PH}_{x_0,x;t}} e^{-S(\phi,t)}\, d\mathfrak{m}(\phi)$. Here $\mathcal{PH}_{x_0,x;t}$ denotes the space of horizontal curves between x_0 and x parameterized by $[0, t]$, $S(\phi, t)$ is the classical action along the horizontal curve $\phi \in \mathcal{PH}_{x_0,x;t}$, and $d\mathfrak{m}(\phi)$ is an analog of the Wiener measure along the horizontal distribution. The authors intend to return to these ideas in a future monograph.

An Overview for the Reader

The present work can be considered as a text for a course or seminars designed for graduate students interested in the most recent developments in sub-Riemannian geometry. It is useful for both pure and applied mathematicians and theoretical physicists working in the thermodynamics area. The goal of this book is to introduce the reader to the differential geometry of sub-Riemannian manifolds.

Scientific Outline

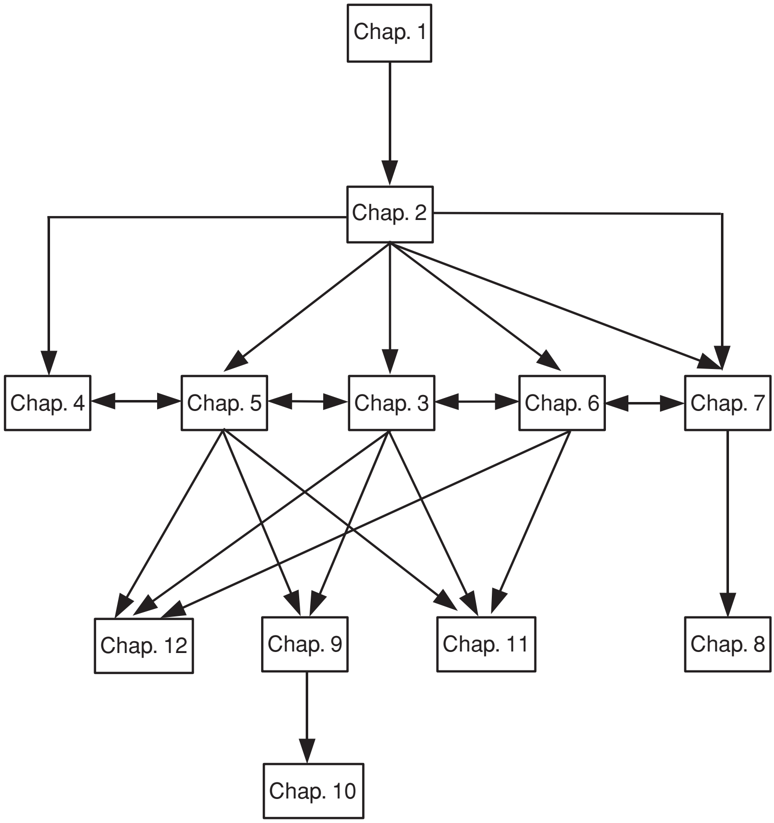
This book deals with the study of sub-Riemannian manifolds, which are manifolds with the Heisenberg principle built in. It is hoped that Heisenberg manifolds (step 2 sub-Riemannian manifolds) will play a role in quantum mechanics in the future, similar to the role played by the Riemannian manifolds in classical mechanics. Some people also speculate that superior-step sub-Riemannian manifolds may play a similar role in quantum field theory.

Therefore it is important to understand sub-Riemannian as well as Riemannian manifolds. However, the sub-Riemannian manifolds behave very differently than Riemannian ones, and we need new methods and insights of investigation.

Acknowledgments

The first author was partially supported by the NSF grant #0631541, while the second author was partially supported by the Hong Kong RGC grant #600607 and a competitive research grant at Georgetown University. Most of this material has been written and presented by both authors during the summers of 2006–2008 at the National Center for Theoretical Sciences, Mathematics Division at Hsinchu, Taiwan. Special thanks go to Professor Jin Yu for inviting the authors to NCTS and providing them with excellent research conditions.

Ann Arbor, MI
Washington, DC
July 4, 2008



The chapters diagram

Part I

General Theory

1

Introductory Chapter

1.1 Differentiable Manifolds

A *manifold* of dimension n is essentially a space that locally resembles the Euclidean space $\mathbb{R}^n$. Every point of the manifold has a neighborhood homeomorphic to an open set of $\mathbb{R}^n$, called a *chart*. The coordinates of the point are the coordinates induced by the chart. Since a point can be covered by several charts, these changes of the coordinates have to be correlated when changing from one chart to another. More precisely, we have the following definitions.

Definition 1.1.1. *Let M be a Hausdorff separated topological space. Then the pair (V, ψ) is called a chart (coordinate system) if $\psi : V \to \psi(V) \subset \mathbb{R}^n$ is a homeomorphism of the open set V in M onto an open set $\psi(V)$ of $\mathbb{R}^n$. The coordinate functions on V are defined as $x^j : V \to \mathbb{R}^n$ and $\psi(p) = (x^1(p), \dots, x^n(p))$; namely, $x^j = u^j \circ \psi$, where $u^j : \mathbb{R}^n \to \mathbb{R}$, $u^j(a_1, \dots, a_n) = a_j$, is the jth projection.*

Definition 1.1.2. *The space M is called a differentiable manifold if there is a collection of charts $\{(V_\alpha, \psi_\alpha)\}_\alpha$ such that*

(1) $V_\alpha \subset M,\ \bigcup_\alpha V_\alpha = M$ *(V_α covers M)*
(2) *if $V_\alpha \cap V_\beta \neq \emptyset$, the map*

$$\Phi_{\alpha\beta} = \psi_\alpha \circ \psi_\beta^{-1} : \psi_\beta(V_\alpha \cap V_\beta) \to \psi_\alpha(V_\alpha \cup V_\beta)$$

is smooth; i.e., the systems of coordinates overlap smoothly.

Since most of the computations in this book have a local character, we may consider that $M = \mathbb{R}^n$. The results are sometimes proved for this particular case; the extension to a general manifold case is left as an exercise for the reader.

1.2 Submanifolds

A *submanifold* is a subset of a manifold that behaves as a manifold. More precisely, we have the following definition.

Definition 1.2.1. *Consider a differentiable manifold M and let N be a subset of M. Let $f : N \to M$ be a smooth function such that*

(1) *f is one-to-one*
(2) *f is an immersion (f_* is one-to-one).*

The pair (N, f) is called a submanifold of M. A map with the properties (1) *and* (2) *is called an imbedding.*

Example 1.2.2. *Any inclusion is an imbedding. For instance, if we consider the standard inclusion $i : \mathbb{S}^2 \to \mathbb{R}^3$, then $\mathbb{S}^2$ becomes a submanifold of $\mathbb{R}^3$.*

Remark 1.2.3. *It is possible for f to be one-to-one without being an imbedding. For instance, $f : (-1, 1) \to \mathbb{R}$, $f(t) = t^3$ is one-to-one but does not have $f'(t)$ one-to-one.*

The fact that $(f_*)_p$ is one-to-one for all $p \in N$ makes possible the identification of the tangent spaces T_pN and $(f_*)_p(T_pN) \subset T_{f(p)}M$. Hence we can consider the tangent space T_pN as a subspace of the tangent space $T_{f(p)}M$.

A classical result dealing with imbedding was proved by Whitney (see [40]).

Theorem 1.2.4 (Whitney's Imbedding Theorem, 1937)**.** *Every n-dimensional manifold imbeds in $\mathbb{R}^{2n+1}$.*

1.3 Distributions

A *distribution* $\mathcal{D}$ of rank k on a manifold M assigns to each point p of M a k-dimensional subspace $\mathcal{D}_p$ of T_pM.

The distribution $\mathcal{D}$ is called *differentiable* if every point p has a neighborhood $\mathcal{V}$ and k differentiable vector fields on $\mathcal{V}$ denoted by $X_1, X_2, \ldots, X_k$, which form a basis of $\mathcal{D}_q$ for all $q \in \mathcal{V}$. We shall write $\mathcal{D} = span\{X_1, \ldots, X_k\}$ on $\mathcal{V}$. In future, by a distribution we will mean a differentiable distribution. k is called the *rank* of the distribution.

The distribution $\mathcal{D}$ is called *involutive* if $[X, Y] \in \mathcal{D}$ for any X, Y in $\mathcal{D}$. An *integral manifold* of the distribution $\mathcal{D}$ is a connected submanifold N of M such that

$$f_*(T_pN) = \mathcal{D}_p, \quad \forall p \in N,$$

where $f : N \to M$ is the imbedding map.

N is called the *maximal integral manifold* of $\mathcal{D}$ if there is no other integral manifold of $\mathcal{D}$ that contains N. A distribution $\mathcal{D}$ that admits a unique maximal manifold through each point is called *integrable*. The classical theorem of

Frobenius states the relationship between the aforementioned two concepts (see, for instance, [52]).

Theorem 1.3.1 (Frobenius). *A distribution $\mathcal{D}$ is involutive if and only if it is integrable.*

In sub-Riemannian geometry the negation statement is used more often: *$\mathcal{D}$ is not involutive if and only if $\mathcal{D}$ is nonintegrable.*

1.4 Integral Curves of a Vector Field

A vector field X on a manifold M can be considered as a particular case of a distribution of rank 1. Since $[X, X] = 0$, the distribution is involutive and hence integrable. The integral manifold has dimension 1 and is called the *integral curve* of X. If t is the parameter along the integral curve $c(t)$, then for any parameter value t_0 the vector $X_{c(t_0)}$ is tangent to the curve $c(t)$ at $c(t_0)$. The existence of integral curves holds locally; i.e., for any $p_0 \in M$, there is $\epsilon > 0$ such that the integral curve $c(t)$ is defined on $(-\epsilon, \epsilon)$ and $c(0) = t_0$. This assertion can be shown as in the following: If $(x^1, \dots, x^m)$ is a local system of coordinates on M in a neighborhood $\mathcal{U}$ of p_0, then the integral curve $c(t)$ is a solution of the following ODE (ordinary differential equation) system:

$$\frac{dc^j(t)}{dt} = X^j\big(c(t)\big), \quad j = 1, \dots, m,$$

where $X = \sum_{i=1}^m X^i \frac{\partial}{\partial x_i}$ on $\mathcal{U}$ and $c^j(t) = x^j \circ c(t)$. The fundamental theorem of local existence and uniqueness of solutions of ODEs provides the proof of our assertion.

Let X be a vector field and define $\varphi_t(p) = c(t)$, where $c(t)$ is the integral curve of X passing through p at $t = 0$. The diffeomorphisms $\varphi_t : M \to M$ form a local one-parameter group of transformations of M; i.e.,

$$\varphi_{t+s}(p) = \varphi_t\big(\varphi_s(p)\big) = \varphi_s\big(\varphi_t(p)\big), \quad \forall t, s, t+s \in (-\epsilon, \epsilon).$$

One may show that the converse is also true; i.e., any local one-parameter group of diffeomorphisms generates locally a vector field. Sometimes φ_s is regarded as the flow in the direction of the vector field X.

We shall denote by $\Gamma(\mathcal{D})$ the set of vector fields tangent to the distribution $\mathcal{D}$. This notation agrees with the notation used for the sections of a subbundle.

Consider a noninvolutive distribution $\mathcal{D}$ and two vector fields X and Y tangent to the distribution. In the following we shall show how the one-parameter group of diffeomorphisms generated by $[X, Y]$ can be written in terms of the one-parameter group of diffeomorphisms associated with the vector fields X and Y. We shall start with an example.

Let $X = \partial_{x_1} + 2x_2\partial_{x_3}$ and $Y = \partial_{x_2} - 2x_1\partial_{x_3}$ be two vector fields on $\mathbb{R}^3$. Consider the ODE system satisfied by the integral curve $c(s) = \big(x_1(s), x_2(s), x_3(s)\big)$ of X:

$$\begin{aligned}\dot{x}_1(s) &= 1\\ \dot{x}_2(s) &= 0\\ \dot{x}_3(s) &= 2x_2(s)\end{aligned}$$

with the solution

$$x(s) = x(0) + s\big(1, 0, 2x_2(0)\big),$$

where $c(0) = \big(x_1(0), x_2(0), x_3(0)\big)$ is the initial point. Then the one-parameter group of diffeomorphisms associated with X is

$$\varphi_s(x) = x + s(1, 0, 2x_2).$$

In a similar way, the flow associated with the vector field Y is

$$\psi_s(x) = x + s(0, 1, -2x_1).$$

Since $[X, Y] = -4\partial_{x_3} \neq 0$, the flows φ_s and ψ_s do not commute. We shall compute next the difference $\varphi_s \circ \psi_s - \psi_s \circ \varphi_s$.

$$\begin{aligned}(\varphi_s \circ \psi_s)(x) &= \varphi_s(x_1, x_2 + s, x_3 - 2sx_1)\\ &= (x_1, x_2 + s, x_3 - 2sx_1) + s\big(1, 0, 2(x_2 + s)\big)\\ &= \big(x_1 + s, x_2 + s, x_3 + 2s(x_2 - x_1) + 2s^2\big),\end{aligned}$$

and

$$\begin{aligned}(\psi_s \circ \varphi_s)(x) &= \psi_s(x_1 + s, x_2, x_3 + 2sx_2)\\ &= (x_1 + s, x_2, x_3 + 2sx_2) + s\big(0, 1, -2(x_1 + s)\big)\\ &= \big(x_1 + s, x_2 + s, x_3 + 2s(x_2 - x_1) - 2s^2\big).\end{aligned}$$

We note that

$$\psi_s \circ \varphi_s(x) - \varphi_s \circ \psi_s(x) = s^2(0, 0, -4) = s^2[X, Y](x).$$

Let τ_s be the flow associated with the vector field $[X_1, X_2] = -4\partial_{x_3}$. We have

$$\tau_s(x) = x + (0, 0, -4s).$$

Then the preceding relation can also be written as

$$\psi_s \circ \varphi_s(x) - \varphi_s \circ \psi_s(x) = \tau_{s^2}(x) - x.$$

In particular, when $x = 0$ we get

$$\psi_s \circ \varphi_s(0) - \varphi_s \circ \psi_s(0) = \tau_{s^2}(0).$$

We shall prove these identities in a more general case.

Proposition 1.4.1. *Let* $(\psi_s)_s$ *and* $(\varphi)_s$ *be the one-parameter groups of diffeomorphisms associated with the vector fields* X *and* Y *on a manifold* M. *Then for any smooth function* $f \in \mathcal{F}(M)$ *we have*

$$f\big(\psi_t \circ \varphi_s(x)\big) - f\big(\varphi_s \circ \psi_t(x)\big) = ts[X, Y](f)(x) + o(s^2 + t^2).$$

Proof. Let $f \in \mathcal{F}(M)$ and consider the smooth function of two variables

$$u(t, s) = f\big(\psi_t \circ \varphi_s(x)\big) - f\big(\varphi_s \circ \psi_s(x)\big).$$

The Taylor expansion of u about $(0, 0)$ is

$$\begin{aligned} u(t, s) &= \sum_{n,m\geq 0} \partial_t^n \partial_s^m u(t, s)_{|_{t=s=0}} t^n s^m \\ &= u(0, 0) + \partial_t u(t, 0)_{|_{t=0}} t + \partial_s u(0, s)_{|_{s=0}} s + \partial_t^2 u(t, 0)_{|_{t=0}} t^2 \\ &\quad + \partial_s^2 u(0, s)_{|_{s=0}} s^2 + \partial_t \partial_s u(t, s)_{|_{t=s=0}} ts + o(s^2 + t^2). \end{aligned}$$

Since $\varphi_0(x) = x$ and $\psi_0(x) = x$, we have

$$\begin{aligned} u(0, 0) &= f\big(\psi_0 \circ \varphi_0(x)\big) - f\big(\varphi_0 \circ \psi_0(x)\big) = 0 \\ u(t, 0) &= f\big(\psi_t \circ \varphi_0(x)\big) - f\big(\varphi_0 \circ \psi_s(x)\big) = f\big(\psi_t(x)\big) - f\big(\psi_t(x)\big) = 0 \\ u(0, s) &= f\big(\psi_0 \circ \varphi_s(x)\big) - f\big(\varphi_s \circ \psi_0(x)\big) = f\big(\varphi_t(x)\big) - f\big(\varphi_t(x)\big) = 0, \end{aligned}$$

and then

$$\partial_t u(t, 0)_{|_{t=0}} = 0, \qquad \partial_s u(0, s)_{|_{s=0}} = 0, \qquad \partial_t^2 u(t, 0)_{|_{t=0}} = 0, \qquad \partial_s^2 u(0, s)_{|_{s=0}} = 0.$$

It follows that

$$u(t, s) = \partial_t \partial_s u(t, s)_{|_{t=s=0}}\, ts + o(s^2 + t^2). \tag{1.4.1}$$

It suffices to compute the mixed derivative at $t = s = 0$. Using the definition of a vector at a point we have

$$\partial_s f\big(\psi_s \circ \varphi_t(x)\big)_{|_{s=0}} = \partial_s f\big(\psi_s(\varphi_t(x))\big)_{|_{s=0}} = (Xf)(\varphi_t(x)) = g\big(\varphi_t(x)\big),$$

where $g = Xf$. Then

$$\partial_t \partial_s f\big(\psi_s(\varphi_t(x))\big)_{|_{t=s=0}} = \partial_t g\big(\varphi_t(x)\big)_{|_{t=0}} = (Yg)(x) = YX(f)(x).$$

Similarly we obtain

$$\partial_s \partial_t f\big(\varphi_t \circ \psi_s(x)\big)_{|_{t=s=0}} = XY(f)(x).$$

Using (1.4.1) yields

$$u(t, s) = ts[Y, X](f)(x) + o(s^2 + t^2).$$ ■

When $s = t$ we obtain the following consequence.

Corollary 1.4.2. *In the hypothesis of Proposition 1.4.1 we have*

$$f\big(\psi_s \circ \varphi_s(x)\big) - f\big(\varphi_s \circ \psi_s(x)\big) = s^2[Y, X](f)(x) + o(s^2).$$

Lemma 1.4.3. *If $(\tau_s)_s$ is the one-parameter group of diffeomorphisms associated with the vector field Z, then*

$$\tau_s(x) = x + sZ(x) + o(s^2).$$

Proof. It follows from

$$\lim_{s\to 0} \frac{\tau_s(x) - \tau_0(x)}{s - 0} = Z(x)$$

and $\tau_0(x) = x$. ■

In the following we shall consider $M = \mathbb{R}^m$ and choose $f = x^i$ to be the ith coordinate function. Then Corollary 1.4.2 becomes

$$\big(\psi_s \circ \varphi_s(x)\big)^i - \big(\varphi_s \circ \psi_s(x)\big)^i = s^2[Y, X]^i(x) + o(s^2), \quad i = 1, \dots, m.$$

In vectorial notation we have

$$\psi_s \circ \varphi_s(x) - \varphi_s \circ \psi_s(x) = s^2[Y, X](x) + o(s^2). \tag{1.4.2}$$

Using Lemma 1.4.3 yields

$$\psi_s \circ \varphi_s(x) - \varphi_s \circ \psi_s(x) = \tau_{s^2}(x) - x + o(s^2),$$

where τ_s is the one-parameter group of diffeomorphisms of $[Y, X]$.

Denote $q = \varphi_s \circ \psi_s(x)$. Then

$$\psi_s \circ \varphi_s(x) = \psi_s \circ \varphi_s \circ \psi_s^{-1} \circ \varphi_s^{-1}(q)$$

and (1.4.2) becomes

$$\psi_s \circ \varphi_s \circ \psi_s^{-1} \circ \varphi_s^{-1}(q) - q = s^2[Y, X] + o(s^2).$$

We arrive at the following result.

Proposition 1.4.4. *Let $[\psi_s, \varphi_s] := \psi_s \circ \varphi_s \circ \psi_s^{-1} \circ \varphi_s^{-1}$. Then*

$$\begin{aligned} [\psi_s, \varphi_s](q) &= q + s^2[Y, X](q) + o(s^2) \\ &= \tau_{s^2}(q) + o(s^2). \end{aligned}$$

If $X, Y \in \Gamma(\mathcal{D})$ and $[X, Y] \notin \Gamma(\mathcal{D})$, then we can move in the $[X, Y]$ direction by just going along the integral curves of X and Y. This is the main idea of the proof of Chow's theorem of connectivity by horizontal curves. In other words, if a creature lives in a universe where it is constrained to move only along a noninvolutive distribution, then it can move in any direction just by taking tangent paths to the distribution (see Fig. 1.1).

The commutator in local coordinates. Given two tangent vector fields U and V to the differentiable manifold M, their commutator vector field is defined by

$$[U, V] = UV - VU = \nabla_U V - \nabla_V U.$$

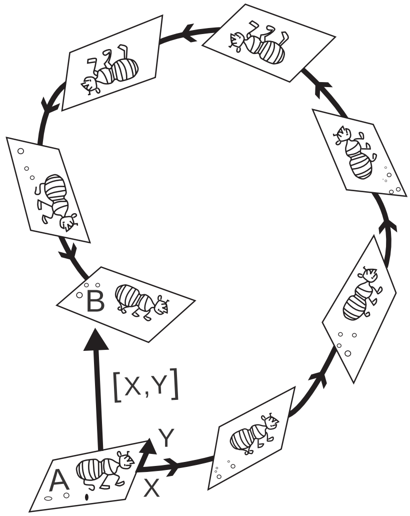


Figure 1.1. The ant can go in the $[X, Y]$ direction by just walking along the integral curves of the noncommuting vector fields X and Y

If $U = \sum_i U^i \partial_{x_i}$ and $V = \sum_i V^i \partial_{x_i}$ are the representations in a local chart $(x_1, \dots, x_n)$, then the commutator in local coordinates becomes

$$[U, V] = UV - VU = \Big(U^i\, \partial_{x_i}(V^j) - V^i\, \partial_{x_i}(U^j)\Big)\partial_{x_j},$$

with summation in the repeated indices. The reader can verify the following properties of the commutator:

(1) The commutator is skew-symmetric: $[U, V] = -[V, U]$.
(2) Jacobi's identity is satisfied:

$$[U, [V, W]] + [V, [W, U]] + [W, [U, V]] = 0.$$

(3) For any smooth functions f and g on M we have

$$[fU, hV] = fh[U, V] + f\,(Uh)\,V - h\,(Vf)\,U.$$

Geometrical interpretation of a vanishing commutator. Let φ_t and ϕ_s be the one parameter groups of diffeomorphisms associated with the vector fields U and V. Then $[U, V] = 0$ if and only if $\varphi_t(\phi_s(p)) = \phi_s(\varphi_t(p))$; i.e., if starting at any point p and going in an arc s along the integral curve of V and then an arc t along the integral curve of U, we end up at the same point as if we performed the procedure in the reverse way.

1.5 Independent One-Forms

We review here a few basic notions regarding one-forms, which will be useful in the future presentation. One reason for studying them is that a distribution can also be defined in terms of one-forms.

Let M be a differentiable manifold. A one-form on M is a section through the cotangent bundle T^*M, i.e., a smooth assignment $M \ni p \to T_p^*M$. If $(x_1, \dots, x_n)$ are the coordinates on an open domain $U \subset M$, a one-form ω can be written in local coordinates as $\omega = \sum_{i=1}^n \omega_i(x)dx_i$, where $\omega_i(x)$ are smooth functions of x. Since all the computations in this section have a local character, we may assume $M = \mathbb{R}^n$.

Consider two one-forms

$$\omega_1 = \sum_{j=1}^n \omega_1^i \, dx_i, \qquad \omega_2 = \sum_{j=1}^n \omega_2^i \, dx_i$$

on $\mathbb{R}^n$ and let

$$S_i = ker\, \omega_{i\,|p} = \{X \in T_pM;\ \omega_{i\,|p}(X) = 0\}, \quad i \in \{1, 2\}$$

be the $(n-1)$-dimensional vectorial subspaces of T_pM defined by the preceding one-forms at p.

Definition 1.5.1. *The spaces S_1 and S_2 are called transversal if they are not parallel. We shall write in this case $S_1 \nparallel S_2$.*

Let $\langle\ ,\ \rangle$ be the natural inner product of $\mathbb{R}^n$. If $X = \sum_{k=1}^n X^k \partial_{x_k} \in ker\, \omega_i$, then we can write

$$\begin{aligned} 0 = \omega_i(X) &= \sum_{j=1}^n \omega_i^j \, dx_j(X) \\ &= \sum_{j=1}^n \omega_i^j X^j = \langle \nu_i, X \rangle, \end{aligned}$$

and hence $\nu_i = \sum_{i=1}^n \omega_i^j \partial_{x_j}$ is a normal vector field to the space $S_i = ker\, \omega_i$.

Definition 1.5.2. *Two one-forms ω_1 and ω_2 are called functionally independent if*

$$rank \begin{pmatrix} \omega_1^i \\ \omega_2^j \end{pmatrix}_{1 \le i,j \le n} = 2.$$

k one-forms $\omega_1, \dots, \omega_k$ are called functionally independent if

$$rank \begin{pmatrix} \omega_1^{i_1} \\ \vdots \\ \omega_k^{i_n} \end{pmatrix}_{1 \le i_1, \dots, i_n \le n} = k;$$

i.e., the coefficients matrix has maximum rank.

Remark 1.5.3. *Definition* 1.5.2 *does not depend on the choice of the basis of one-forms. If $\omega = \sum \omega^i dx_i = \sum \widetilde{\omega}^j d\widetilde{x}_j$ is the representation of the one-form in two local systems of coordinates, then $\omega^i = \widetilde{\omega}^j d\widetilde{x}_j\big(\partial_{x_i}\big) = \widetilde{\omega}^j(\partial\widetilde{x}_j/\partial x_i)$, and hence $rank\, \widetilde{\omega}_p^j = rank\, \omega_p^j$.*

Proposition 1.5.4. (1) *The spaces $ker\ \omega_1$ and $ker\ \omega_2$ are transversal at p if and only if ω_1 and ω_2 are linearly independent at p.*

(2) $\bigcap_{j=1}^{k} ker\ \omega_j \neq \emptyset$ *if and only if the one-forms $\omega_1, \ldots, \omega_k$ are functionally independent.*

Proof.

(1) Let $S_i = ker\ \omega_i$. Then the spaces $S_1 \not\parallel S_2$ if and only if the normal vectors are not parallel, i.e., $\nu_1 \not\parallel \nu_2$, or $(\omega_1^1, \ldots, \omega_1^n)$ and $(\omega_2^1, \ldots, \omega_2^n)$ are not proportional. This means that there is a 2×2 nondegenerate minor matrix

$$\det \begin{pmatrix} \omega_1^{i_1} & \omega_1^{i_2} \\ \omega_2^{j_1} & \omega_2^{j_2} \end{pmatrix} \neq 0,$$

and therefore the rank of the coefficients matrix is 2. Hence ω_1 and ω_2 are functionally independent.

(2) We leave this as an exercise for the reader. ∎

Example 1.5.1. *The following two one-forms on $\mathbb{R}^4_{(x_1,x_2,y_1,y_2)}$*

$$\omega_1 = dy_1 - x_1 dx_2$$
$$\omega_2 = dy_2 - \frac{1}{2} x_1^2 dx_2$$

are functionally independent.

Example 1.5.2. *The following three one-forms on $\mathbb{R}^4_{(x_1,x_2,y_1,y_2)}$*

$$\begin{aligned} \omega_1 &= dx_1 + x_1 dx_2 + y_2 dy_1 + y_1 dy_2 \\ \omega_2 &= \qquad\quad dx_2 + x_2^2 dy_1 + y_2 dy_2 \\ \omega_3 &= \qquad\qquad\qquad\quad dy_1 + y_1^2 dy_2 \end{aligned}$$

are functionally independent.

1.6 Distributions Defined by One-Forms

Codimension 1. The simplest case is when the distribution is defined by only one one-form ω as $\mathcal{D} = ker\ \omega$. We note that for any $f \neq 0$ the distribution is still given by $\mathcal{D} = ker\ f\omega$, and therefore the one-form is unique up to a multiplicative nonvanishing function.

Codimension 2. Consider the case of a distribution defined by two functionally independent one-forms ω_1 and ω_2. Then we define the distribution by

$$\mathcal{D}_{(\omega_1,\omega_2)} = ker\ \omega_1 \cap ker\ \omega_2.$$

The following result shows the invariance of the distribution under some algebraic operations with one-forms.

Proposition 1.6.1. *Let ω_1 and ω_2 be two functionally independent one-forms. Let $a, b, \alpha,$ and β be real-valued functions with $a\beta \neq b\alpha$, and let $\widetilde{\omega}_1 = a\omega_1 + b\omega_2$ and $\widetilde{\omega}_2 = \alpha\omega_1 + \beta\omega_2$. Then*

$$\mathcal{D}_{(\omega_1,\omega_2)} = \mathcal{D}_{(\widetilde{\omega}_1,\widetilde{\omega}_2)}.$$

Proof. It is easy to see that $\widetilde{\omega}_1$ and $\widetilde{\omega}_2$ are functionally independent. The conclusion is equivalent to

$$ker\ \omega_1 \cap ker\ \omega_2 = ker(a\omega_1 + b\omega_2) \cap ker(\alpha\omega_1 + \beta\omega_2).$$

This will be shown by double inclusion.

Let $X \in ker\ \omega_1 \cap ker\ \omega_2$. Then $\omega_1(X) = 0$ and $\omega_2(X) = 0$ and obviously $(a\omega_1 + b\omega_2)(X) = 0$ and $(\alpha\omega_1 + \beta\omega_2)(X) = 0$; i.e., $X \in ker(a\omega_1 + b\omega_2) \cap ker(\alpha\omega_1 + \beta\omega_2)$.

Let $Y \in ker(a\omega_1 + b\omega_2) \cap ker(\alpha\omega_1 + \beta\omega_2)$. Then

$$a\omega_1(Y) + b\omega_2(Y) = 0$$
$$\alpha\omega_1(Y) + \beta\omega_2(Y) = 0.$$

Since $a\beta \neq b\alpha$, this homogeneous system has only the zero solution; i.e., $\omega_1(Y) = 0$ and $\omega_2(Y) = 0$ and therefore $Y \in ker\omega_1 \cap ker\omega_2$. ∎

The preceding result is very useful when dealing with a system of two functionally independent nonholonomic constraints, i.e., constraints given by one-forms. One may want to do some transformations that preserve the distribution and at the same time make the nonholonomic constraints more simple. We shall do this in the next example.

Example 1.6.1. *Let $\omega_1 = dy_1 - x_1 dx_2$ and $\omega_2 = dy_2 - \frac{1}{2}x_1^2 dx_2$ be the functionally independent one-forms given by Example 1.5.1. Consider*

$$\begin{aligned}\widetilde{\omega}_1 &= \omega_1\\ \widetilde{\omega}_2 &= \omega_2 - \frac{1}{2}x_1\omega_1\\ &= dy_2 - \frac{1}{2}x_1 dy_1.\end{aligned}$$

We have $a = 1$, $b = 0$, $\alpha = -\frac{1}{2}x_1$, $\beta = 1$ and the hypothesis $a\beta \neq b\alpha$ is satisfied. The distribution generated by ω_1 and ω_2 is the same as the distribution generated by $dy_1 - x_1 dx_2$ and $dy_2 - \frac{1}{2}x_1 dy_1$. We note that in this case the coefficients are linear, while in the initial case a coefficient was quadratic.

Codimension k. In the case of k functionally independent one-forms $\omega_1, \ldots, \omega_k$, the distribution is defined by

$$\mathcal{D}_{(\omega_1,\ldots,\omega_k)} = \bigcap_{j=1}^{k} ker\ \omega_j.$$

Since the forms are functionally independent, then *dim* $\mathcal{D}_{(\omega_1,\dots,\omega_k)} = n - k$. The number of the forms is the codimension of the distribution. One may prove a similar result as in the case of two one-forms.

Proposition 1.6.2. *Let* $A = (A_i^j)$ *be a matrix with the entries functions such that* $\det A_p \neq 0$ *at every* p. *Let* $\Omega = (\omega_1, \dots, \omega_k)$ *and consider* $\widetilde{\Omega} = A\Omega$*; i.e.,* $\widetilde{\omega}_j = \sum_p A_j^p \omega_p$. *Then* $\mathcal{D}_\Omega = \mathcal{D}_{A\Omega}$.

1.7 Integrability of One-Forms

The integrable factors for a one-form are used in the proof of the second law of thermodynamics (see Chapter 3).

Definition 1.7.1. *A nowhere vanishing function* $f : M \to \mathbb{R}$ *is called an integrating factor for the one-form* ω *if* $d(f\omega) = 0$. *The one-form* ω *is called integrable if it has an integrating factor.*

Example 1.7.2. *The one-form* $\omega = xdy$ *is integrable. An integrating factor is* $f(x) = \frac{1}{x}$ *since*

$$d(f\omega) = d^2y = 0.$$

One may notice that all the integrating factors are of the form $f(x) = \frac{c}{x}$ *with* $c \in \mathbb{R}$. *This is obtained by solving the equation* $d(f\omega) = 0$*:*

$$\begin{aligned} df \wedge \omega + f\, d\omega = 0 &\Longleftrightarrow \\ (f_x\, dx + f_y\, dy) \wedge xdy + f\, dx \wedge dy = 0 &\Longleftrightarrow \\ (xf_x + f)dx \wedge dy = 0 &\Longleftrightarrow \\ xf_x + f = 0. & \end{aligned}$$

The method of separation of variables leads to the preceding expression of $f(x)$.

Example 1.7.3. *Consider the one-form* $\omega = xdy - ydx$. *An integrating factor is* $f(x) = \frac{1}{x^2+y^2}$. *This follows from the fact that in polar coordinates* (r, ϕ) *we have* $\omega = r^2 d\phi$.

Example 1.7.4. *The one-form* $\omega = dt - xdy$ *is not integrable. Suppose* ω *has a nonzero integrating factor* f. *Then the equation* $d(f\omega) = 0$ *becomes*

$$\begin{aligned} df \wedge \omega + f\, d\omega = 0 &\Longleftrightarrow \\ (f_x dx + f_y dy + f_t dt) \wedge (dt - xdy) - f(dx \wedge dy) = 0 &\Longleftrightarrow \\ -(xf_x + f)dx \wedge dy + (f_y + xf_t)dy \wedge dt + f_x dx \wedge dt = 0, & \end{aligned}$$

and equating the coefficients to zero yields

$$xf_x + f = 0, \qquad f_y + xf_t = 0, \qquad f_x = 0.$$

From the first and the last equations we get $f = 0$, *which is a contradiction.*

Let $\mathcal{D} = ker\ \omega$ be the distribution defined by ω. The integrability relationship between $\mathcal{D}$ and ω is provided by the following result.

Proposition 1.7.5. *The distribution $\mathcal{D}$ is integrable if and only if the one-form ω is integrable.*

Proof. "$\Longleftarrow$" If ω is integrable, then there is an integral factor f such that $d(f\omega) = 0$. By Poincaré's lemma, there is a function h such that locally we have $f\omega = dh$. Let $h(p) = c$. We shall show that $h^{-1}(c)$ is a locally integrable manifold of the distribution $\mathcal{D}$ that passes through p. Let $X \in \Gamma(\mathcal{D})$ be a vector field. Then

$$X(h) = dh(X) = f\omega(X) = 0,$$

which means that h is constant along the integral curve of X that passes through p. Then locally X is tangent to $h^{-1}(c)$ and hence the surface $h^{-1}(c)$ is tangent to the distribution $\mathcal{D}$. The submanifold condition $dh \neq 0$ is satisfied since $f \neq 0$.

"$\Longrightarrow$" An integral manifold can be written locally as $h^{-1}(c)$, $dh \neq 0$. Then $ker\ (dh) = \mathcal{D} = ker\ \omega$, and hence the one-forms dh and ω are proportional; i.e., there is a nonvanishing function f such that $f\omega = dh$. Then

$$d(f\omega) = d^2h = 0,$$

so ω is integrable. ∎

The following result deals with equivalent integrability conditions for one-forms.

Proposition 1.7.6. *Let ω be a one-form on $\mathbb{R}^3$. Then the following conditions are equivalent:*

(1) *ω is integrable*
(2) *there is a one-form θ such that $d\omega = \theta \wedge \omega$*
(3) $\omega \wedge d\omega = 0$
(4) $d\omega_{|ker\ \omega} = 0$
(5) *the distribution $ker\ \omega$ is involutive.*

Proof.

(1) $\Longrightarrow$ (2) Let ω be an integrable one-form. If f is an integrable factor, the relation $d(f\omega) = 0$ becomes

$$fd\omega = -df \wedge \omega,$$

which is $d\omega = \theta \wedge \omega$ with $\theta = \frac{-df}{f}$.

(2) $\Longrightarrow$ (3) Since $d\omega = \theta \wedge \omega$ we have

$$\omega \wedge d\omega = \omega \wedge \theta \wedge \omega = -(\omega \wedge \omega) \wedge \theta = 0.$$

(3) $\Longrightarrow$ (4) Let $X_1, X_2 \in ker\,\omega$ and $X_3 \notin ker\,\omega$ such that $\{X_1, X_2, X_3\}$ are linearly independent. Then

$$0 = (\omega \wedge d\omega)(X_1, X_2, X_3) = \omega(X_1)d\omega(X_2, X_3) - \omega(X_2)d\omega(X_1, X_3) \\ + \omega(X_3)d\omega(X_1, X_2) = \omega(X_3)d\omega(X_1, X_2).$$

Since $X_3 \notin ker\,\omega$, then $\omega(X_3) \neq 0$ and hence $d\omega(X_1, X_2) = 0$ for all $X_1, X_2 \in ker\,\omega$.

(4) $\Longrightarrow$ (5) We have

$$0 = d\omega(X_1, X_2) = X_1\omega(X_2) - X_2\omega(X_1) - \omega([X_1, X_2]) = -\omega([X_1, X_2]).$$

Hence $[X_1, X_2] \in ker\,\omega$ for all $X_1, X_2 \in ker\,\omega$; i.e., $ker\,\omega$ is an involutive distribution.

(5) $\Longrightarrow$ (1) Since $ker\,\omega$ is involutive, by Frobenius' theorem it is integrable. Applying Proposition 1.7.5 it follows that ω is integrable. ■

Remark 1.7.7. *If $\omega = Adx + Bdy + Cdz$ is a one-form on $\mathbb{R}^3$, the integrability condition $\omega \wedge d\omega = 0$ becomes*

$$A\Big(\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}\Big) - B\Big(\frac{\partial C}{\partial x} - \frac{\partial A}{\partial z}\Big) + C\Big(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\Big) = 0.$$

Definition 1.7.8. *A constraint on the velocity of a curve given by a one-form ω is called nonholonomic*[1] *if ω is nonintegrable.*

A nonholonomic constraint can be written as

$$\omega(\dot{c}) = \sum \omega^i dx_i(\dot{c}) = \sum \omega_i \dot{c}_i = 0.$$

Example 1.7.1. *The one-form $\omega = dx - xdy$ on $\mathbb{R}^2$ is not integrable, so*

$$\omega(\dot{c}) = \dot{c}_1 - c_1\dot{c}_2 = 0$$

is a nonholonomic constraint.

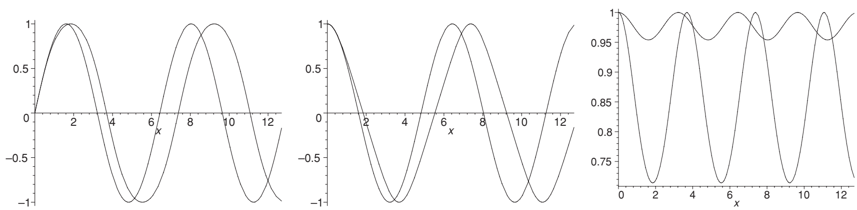
Example 1.7.2. *The one-form $\omega = 2xdx - dy$ on $\mathbb{R}^2$ is integrable, so*

$$\omega(\dot{c}) = 2c_1\dot{c}_1 - \dot{c}_2 = 0$$

is a holonomic constraint.

The literature of nonholonomic geometry deals with the concepts of *rheonomic* (flowing) and *scleronomic* (ridgid) nonholonomic constraints. A rheonomic condition means that the constraint depends directly on the time parameter t. All our constraints in this book will be independent of time; i.e., they are *scleronomic nonholonomic* (see [28]).

[1] In Greek *holos* means integer.



(a) $sn(z, k)$ (b) $cn(z, k)$ (c) $dn(z, k)$

Figure 1.2. The graphs of functions $sn(z, k)$, $cn(z, k)$, and $dn(z, k)$, for $k = 0.3$ and $k = 0.7$

1.8 Elliptic Functions

We shall provide in the following the definitions of the elliptic functions used in the next chapters. For a detailed description the reader may consult reference [53].

The integral

$$z = \int_0^w \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}, \quad |k| < 1,$$

is called an *elliptic integral of the first kind*. The integral exists if w is real and $|w| < 1$. Using the substitution $t = \sin\theta$ and $w = \sin\phi$,

$$z = \int_0^\phi \frac{d\theta}{\sqrt{1 - k^2 \sin^2\theta}}.$$

If $k = 0$, then $z = \sin^{-1} w$ or $w = \sin z$. By analogy, this integral is denoted by $sn^{-1}(w; k)$, where $k \neq 0$. The number k is called the *modulus*. Thus

$$z = sn^{-1} w = \int_0^w \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}.$$

The function $w = sn\, z$ is called a *Jacobian elliptic function*.

By analogy with the trigonometric functions, it is convenient to define other elliptic functions (see Fig. 1.2).

$$cn\, z = \sqrt{1 - sn^2 z}, \qquad dn\, z = \sqrt{1 - k^2\, sn^2 z}.$$

A few properties of these functions are

$$sn(0) = 0, \qquad cn(0) = 1, \qquad dn(0) = 1,$$

$$sn(-z) = sn(z), \qquad cn(-z) = cn(z),$$

$$\frac{d}{dz} sn\, z = cn\, z\, dn\, z, \qquad \frac{d}{dz} cn\, z = -sn\, z\, dn\, z, \qquad \frac{d}{dz} dn\, z = -k^2 sn\, z\, cn\, z,$$

$$-1 \le cn\, z \le 1, \qquad -1 \le sn\, z \le 1, \qquad 0 \le dn\, z \le 1.$$

Let

$$K = K(k) = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2\theta}}$$

be the complete Jacobi integral. Then, as real functions, the elliptic functions sn and cn are periodic functions of the principal period $4K$.

1.9 Exterior Differential Systems

Let Ω^p be the space of p-forms on a connected, open set $U \subseteq \mathbb{R}^m$. The p-form $\omega \in \Omega^p$ can be written as

$$\omega = \sum \omega_{i_1 \dots i_p} dx_{i_1} \wedge \cdots \wedge dx_{i_p},$$

where the coefficients $\omega_{i_1 \dots i_p}$ are differentiable functions on U and skew-symmetric in the indices $i_1, \dots, i_p$. We shall denote by Ω the graded algebra with components Ω^p; i.e., $\Omega = \bigoplus_{p=0}^m \Omega^p$. The space Ω is sometimes called the *sheaf of differentiable forms* on U. The multiplication $\wedge : \Omega \times \Omega \to \Omega$ is the usual wedge product of two forms. Ω^0 denotes the set of differentiable functions on U. In the following we shall review a few notions regarding ideals (see [70]).

An *ideal* $\mathcal{I}$ of Ω is a subset with the following properties:

(1) $\forall \theta, \theta' \in I$ and $f \in \Omega^0$, then $\theta + \theta' \in \mathcal{I}$ and $f\theta \in \mathcal{I}$.
(2) $\forall \theta \in \mathcal{I}$ and $\forall \omega \in \Omega$, then $\theta \wedge \omega \in \mathcal{I}$.

An ideal $\mathcal{I}$ of Ω is called *finitely generated* if it has a finite number of generators $\theta_1, \dots, \theta_k$; i.e., any form $\theta \in \mathcal{I}$ can be written as

$$\theta = \omega_1 \wedge \theta_1 + \cdots + \omega_k \wedge \theta_k,$$

where $\omega_i \in \Omega$ with $\deg \omega_i = \deg \theta - \deg \theta_i$.

An ideal $\mathcal{I}$ of Ω is called *homogeneous* if $\bigoplus_{p=0}^m \mathcal{I}^p = \mathcal{I}$, where $\mathcal{I}^p = \mathcal{I} \cap \Omega^p$. With these introductions we can define the concept of a system of forms, which will be useful in the analysis of distributions.

Definition 1.9.1. *Let U be an open set of $\mathbb{R}^m$. An exterior differential system on the set U is a homogeneous, finitely generated ideal $\mathcal{I}$ of Ω.*

In the following we shall introduce the notion of integral manifold of an exterior differential system.

Definition 1.9.2. (1) *A point $x \in U$ is called an integral point of the exterior differential system $\mathcal{I}$ if $f(x) = 0$ for any $f \in \mathcal{I}^0$; i.e., all the functions of $\mathcal{I}$ vanish at x.*

(2) *Let $x \in U$ be an integral point of I. A vector $v \in T_x U$ is called an integral vector of the system $\mathcal{I}$ if $\theta_x(v) = 0$ for any $\theta \in \mathcal{I}^1 = \mathcal{I} \cap \Omega^1$.*

(3) *Let V be a subspace of dimension k of the vector space $T_x U$, where $x \in U$ is an integral point. V is called an integral k-plane of the exterior differential system $\mathcal{I}$ if V has a basis $\{v_1, v_2, \dots, v_k\}$ such that for any $\theta \in \mathcal{I}^r$ and any subindex $\{i_1, \dots, i_r\} \subset \{i_1, \dots, i_k\}$ we have $\theta_x(v_{i_1}, \dots, v_{i_r}) = 0$, for all $r \le k$.*

(4) *Any subspace S of V of dimension $s \le k$ is an integral s-plane of the system $\mathcal{I}$.*

The definition makes sense, since, as we shall show later, this definition does not depend on the basis of the integral k-plane. If $\{v_1, \dots, v_k\}$ is a basis of the subspace V of $T_x U$, then the fact that V is an integral k-plane can be written as

$$\sum_j \theta_{j_1\dots j_r}(x) v_{i_1}^{j_1} \dots v_{i_r}^{j_r} = 0. \tag{1.9.3}$$

Let $\{w_1, \dots, w_k\}$ be another basis such that $w_j = \sum_p a_{jp} v_p$. Componentwise, we have $w_j^\ell = \sum_p a_{jp} v_p^\ell$ and then

$$\begin{aligned} \sum_j \theta_{j_1\dots j_r}(x) w_{i_1}^{j_1} \dots w_{i_r}^{j_r} &= \sum_j \theta_{j_1\dots j_r}(x) \sum_{p_1} a_{i_1 p_1} v_{p_1}^{j_1} \cdots \sum_{p_r} a_{i_r p_r} v_{p_r}^{j_r} \\ &= \sum_{p_1\dots p_r} \underbrace{\Big(\sum_j \theta_{j_1\dots j_r}(x) v_{i_1}^{j_1} \cdots v_{i_r}^{j_r} \Big)}_{=0 \text{ by } (1.9.3)} a_{i_1 p_1} \cdots a_{i_r p_r} \\ &= 0, \end{aligned}$$

which is the relation (1.9.3) for the basis $\{v_1, \dots, w_1, \dots, w_k\}$.

The main problem of the theory of exterior differential systems is to study their integral manifolds. In the following we shall present two definitions of a submanifold of $\mathbb{R}^m$, which are equivalent to Definition 1.2.1 for the case $M = \mathbb{R}^m$.

Definition 1.9.3. *A subset $M \subset \mathbb{R}^m$ is called a k-dimensional differential manifold of $\mathbb{R}^m$ if for every point $x \in M$, there is an open neighborhood $U \subset \mathbb{R}^m$ and differentiable functions $f_i : U \to \mathbb{R}$, $i = 1, \dots, m-k$, such that*

(1) $M \cap U = \{x \in U; f_1(x) = \cdots = f_{m-k}(x) = 0\}$
(2) *rank* $J_f(x) = m - k$, *where*

$$J_f(x) = \frac{\partial(f_1, \dots, f_{m-k})}{\partial(x_1, \dots, x_m)}$$

is the Jacobian of $f = (f_1, \dots, f_{m-k})$.

Condition (2) can be restated by saying that the Jacobian of f has maximum rank.

Example 1.9.1. *If $m = k + 1$ then f has only one component and the Jacobian becomes the gradient of f. The manifold can be written in this case as*

$$\mathcal{H}^{m-1} = \{x \in \mathbb{R}^m; f(x) = 0, \nabla f \neq 0\}$$

and it is called a hypersurface of $\mathbb{R}^m$. In the particular case when $f(x) = a_0 + \sum_{i=1}^m a_i x_i$, with $a_j \in \mathbb{R}$ not all zero, we obtain a hyperplane. When $f(x) = \sum_{i=1}^m (x_i)^2 - 1$ we obtain a hypersphere in $\mathbb{R}^m$.

Example 1.9.2. *Let $m = 3$ and $k = 1$ and consider the manifold defined by the equations*

$$f_1(x_1, x_2, x_3) = x_1 + x_2 + x_3 = 0$$
$$f_2(x_1, x_2, x_3) = x_2 - x_3 = 0.$$

Since $J_f = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ has the rank equal to $2 = m - k$, this system of functions define a manifold of dimension $k = 1$ of $\mathbb{R}^3$. This would be more clear if we solve the system in the variables x_2 and x_3 as

$$x_2 = -\frac{1}{2}x_1, \qquad x_3 = -\frac{1}{2}x_1.$$

Letting $x_1 = t \in \mathbb{R}$ we obtain the parametric equations of the manifold

$$x_1 = t$$
$$x_2 = -\frac{1}{2}t$$
$$x_3 = -\frac{1}{2}t, \quad t \in \mathbb{R},$$

which define a line in $\mathbb{R}^3$. This procedure can be carried out for any manifold M of $\mathbb{R}^3$.

Let M be a manifold of dimension k in $\mathbb{R}^m$ given by the equations

$$f_1(x_1, \ldots, x_m) = 0$$
$$\vdots$$
$$f_{m-k}(x_1, \ldots, x_m) = 0.$$

By the Implicit Function Theorem this system can be solved locally with respect to $m - k$ of the variables x_i, which will be denoted by t_i, such that we have

$$x_1 = \varphi_1(t_1, \ldots, t_k)$$
$$\vdots$$
$$x_m = \varphi_m(t_1, \ldots, t_k)$$

with φ_i differentiable functions and with the Jacobian J_φ of rank k. These equations hold locally; i.e., the coordinates $t_1, \ldots, t_k$ belong to an open set $U \subseteq \mathbb{R}^k$.

We shall often use the preceding parametric representation for a manifold. Sometimes M is regarded as a submanifold of $\mathbb{R}^m$ to emphasize that M inherits topological and differential structures of $\mathbb{R}^m$. If $(t_1, \ldots, t_k)$ are the coordinates on an open subset $U \subseteq \mathbb{R}^k$, then $\varphi : U \to \mathbb{R}^m$ is an immersion since the rank of the Jacobian is maximum on U.

Definition 1.9.4. *A manifold M of dimension k contained in the open set $U \subset \mathbb{R}^m$ is called an integral manifold for the exterior differential system $\mathcal{I}$ if for any $x \in M$ the tangent plane at x, $T_x M$, is an integral k-plane of the system $\mathcal{I}$.*

Remark 1.9.5. *If S is a submanifold of the integral manifold M, then S is an integral manifold of $\mathcal{I}$.*

In order to present the main theorem about integral manifolds we need the following definition. Let $U \subset \mathbb{R}^k$ and $V \subset \mathbb{R}^m$ be two open sets and consider a differentiable function $F = (F^1, \dots, F^m) : U \to V$. Denote by $t_1, \dots, t_k$ the coordinates on U and by $x_1, \dots, x_m$ the coordinates on $\mathbb{R}^m$. Let $\theta = \sum \theta_{i_1 \dots i_p}(x) dx_1 \wedge \cdots \wedge dx_p$ be a p-form on V. The pullback $F^*(\theta)$ of θ through F is a p-form on U obtained from θ by substituting x_i by $F^i(t)$ and differentials dx_i by $dF^i = \sum_{j=1}^k \frac{\partial F^i}{\partial t^j} dt_j$. This means

$$\begin{aligned}
F^*(\theta) &= \sum \theta_{i_1 \dots i_p}\big(F(t)\big) dF^{i_1} \wedge \cdots \wedge dF^{i_p} \\
&= \sum \theta_{i_1 \dots i_p}\big(F(t)\big) \sum_{j_1=1}^{k} \frac{\partial F^{i_1}}{\partial t_{j_1}} dt_{j_1} \wedge \cdots \wedge \sum_{j_p=1}^{k} \frac{\partial F^{i_p}}{\partial t_{j_p}} dt_{j_p} \\
&= \sum \theta_{i_1 \dots i_p}\big(F(t)\big) \frac{\partial F^{i_1}}{\partial t_{j_1}} dt_{j_1} \cdots \frac{\partial F^{i_p}}{\partial t_{j_p}} dt_{j_p}\, dt_{j_1} \wedge \cdots \wedge dt_{j_p} \qquad (1.9.4) \\
&= \sum_{I,J} \theta_I\big(F(t)\big) \frac{\partial F^I}{\partial t_J} dt_J,
\end{aligned}$$

where $I = (i_1, \dots i_p)$ and $J = (j_1, \dots, j_p)$ are multi-indices.

Theorem 1.9.6. *Let M be a manifold of $\mathbb{R}^m$ of dimension k given locally by the parametric equations $x_i = \varphi_i(t_1, \dots, t_k)$, $i = 1, \dots, m$. Then M is an integral manifold for the exterior differential system $\mathcal{I}$ if and only if $\varphi^*(\theta) = 0$ for any form $\theta \in \mathcal{I}$, where $\varphi = (\varphi_1, \dots, \varphi_m)$.*

Proof. The coordinates tangent vector fields at $x = \varphi(t)$ are

$$v_j = \left(\frac{\partial \varphi_1}{\partial t_j}, \dots, \frac{\partial \varphi_m}{\partial t_j} \right) \in T_x M, \quad j = 1, \dots, k.$$

M is an integral manifold for the system $\mathcal{I}$ if for any point $x \in M$ the vector space $T_x M$ is an integral k-plane of the system $\mathcal{I}$. Then for any form $\theta \in \mathcal{I}^p$, we have

$$\begin{aligned}
\theta_x(v_{i_1}, \dots, v_{i_p}) &= 0 \Longleftrightarrow \\
\sum \theta_{j_1 \dots j_p}(x)\, v_{i_1}^{j_1} \dots v_{i_p}^{j_p} &= 0 \Longleftrightarrow \\
\sum \theta_{j_1 \dots j_p}\big(\varphi(t)\big) \frac{\partial \varphi^{j_1}}{\partial t_{i_1}} \dots \frac{\partial \varphi^{j_p}}{\partial t_{i_p}} &= 0 \Longleftrightarrow \\
\varphi^*(\theta) &= 0,
\end{aligned}$$

where we used (1.9.4). ∎

Definition 1.9.7. *An exterior system $\mathcal{I}$ is called closed if $d\mathcal{I} \subset \mathcal{I}$; i.e., for all $\theta \in \mathcal{I}$ we have $d\theta \in \mathcal{I}$.*

Proposition 1.9.8. *Let $\mathcal{I}$ be an exterior differential system. Then*

$$\overline{\mathcal{I}} = \mathcal{I} + d\mathcal{I} = \{\theta + d\omega;\ \theta, \omega \in \mathcal{I}\}$$

is a closed exterior differential system, called the closure of $\mathcal{I}$.

Proof. We shall show first that $\overline{\mathcal{I}}$ is an ideal of Ω.

Let $\eta, \eta' \in \overline{\mathcal{I}}$, with $\eta = \theta + d\omega$ and $\eta' = \theta' + d\omega'$, where $\theta, \omega, \theta', \omega' \in \mathcal{I}$. Then we have

$$\eta + \eta' = (\theta + \theta') + d(\omega + \omega') \in \mathcal{I} + d\mathcal{I} = \overline{\mathcal{I}};$$

i.e., the sum of any two forms of $\overline{\mathcal{I}}$ belongs to $\overline{\mathcal{I}}$.

Now we shall show that $\omega \wedge \theta \in \overline{\mathcal{I}}$ for any $\omega \in \Omega$ and $\theta \in \overline{\mathcal{I}}$.

If $\theta \in \mathcal{I}$, then using that $\mathcal{I}$ is an ideal of Ω, we have $\omega \wedge \theta \in \mathcal{I} \subset \mathcal{I} + d\mathcal{I} = \overline{\mathcal{I}}$.

If $\theta \in d\mathcal{I}$, i.e., $\theta = d\xi$ with $\xi \in \mathcal{I}$, we have for any $\omega \in \Omega^p$

$$(-1)^p \omega \wedge \theta = (-1)^p \omega \wedge d\xi = d(\omega \wedge \xi) - d\omega \wedge \xi \in \mathcal{I} + d\mathcal{I} = \mathcal{I},$$

since $d\omega \wedge \xi \in \mathcal{I}$ and $\omega \wedge \xi \in \mathcal{I}$. Hence $\overline{\mathcal{I}}$ is an ideal of the graded algebra Ω.

We shall show next that the ideal $\overline{\mathcal{I}}$ is finitely generated. Let $\{\theta_1, \ldots, \theta_k\}$ be a system of generators for the ideal $\mathcal{I}$. Then $\{\theta_1, \ldots, \theta_k, d\theta_1, \ldots, d\theta_k\}$ is a system of generators for the ideal $\mathcal{I} + d\mathcal{I} = \overline{\mathcal{I}}$.

Since $\bigoplus_{p=0}^m (\mathcal{I}^p + d\mathcal{I}^p) = \bigoplus_{p=0}^m \mathcal{I}^p + \bigoplus_{p=0}^m d\mathcal{I}^p = \mathcal{I} + d \bigoplus_{p=0}^m \mathcal{I}^p = \mathcal{I} + d\mathcal{I}$, it follows that the ideal $\mathcal{I} + d\mathcal{I}$ is homogeneous.

In order to show the closeness of $\overline{\mathcal{I}}$, we use the involutivity of the exterior derivative:

$$d(\overline{\mathcal{I}}) = d(\mathcal{I} + d\mathcal{I}) = d\mathcal{I} + d^2\mathcal{I} = d\mathcal{I} \subset \mathcal{I} + d\mathcal{I} = \overline{\mathcal{I}}. \qquad \blacksquare$$

The following result deals with the integral manifolds of the closure of an exterior differential system.

Theorem 1.9.9. *A manifold M is an integral manifold for the exterior differential system $\mathcal{I}$ if and only if it is an integral manifold for the system $\overline{\mathcal{I}} = \mathcal{I} + d\mathcal{I}$.*

Proof. "$\Longrightarrow$" Let M be an integral manifold for the system $\mathcal{I}$ defined locally by the parametric equations

$$x_i = \varphi_i(t_1, \ldots, t_k), \quad i = 1, \ldots, m.$$

By Theorem 1.9.6, we have $\varphi^*(\theta) = 0$ for any form $\theta \in \mathcal{I}$. Since d and φ^* commute,

$$\varphi^*(d\theta) = d\varphi^*(\theta) = 0.$$

Then for all $\theta, \eta \in \mathcal{I}$ we have $\varphi^*(\theta + d\eta) = 0$; i.e., $\varphi^*(\omega) = 0$, $\forall \omega \in \mathcal{I} + d\mathcal{I}$. Using Theorem 1.9.6 it follows that M is an integral manifold for $\overline{\mathcal{I}} = \mathcal{I} + d\mathcal{I}$.

"$\Longleftarrow$" Assume M is an integral manifold of dimension k for the system $\mathcal{I} + d\mathcal{I}$. Let $x \in M$. Then the tangent plane T_xM is an integral k-plane for $\mathcal{I} + d\mathcal{I}$; i.e., any form of the type $\theta + d\eta$ with $\theta, \eta \in \mathcal{I}$ will vanish on T_xM. In particular, for $\eta = 0$, we obtain that the forms of $\mathcal{I}$ vanish on T_xM; i.e., T_xM is an integral k-plane for the system $\mathcal{I}$. Since x was chosen arbitrarily in M, it follows that M is an integral manifold of the system $\mathcal{I}$. ■

In general, the system $\mathcal{I} + d\mathcal{I}$ has fewer integral planes than $\mathcal{I}$. For instance, if $\mathcal{I}$ is generated by the form $\omega = x_1 dx_2$ on $\mathbb{R}^2$, the plane $\{x_1 = 0\}$ is an integral 2-plane. However, the form $d\omega = dx_1 \wedge dx_2$ does not vanish on any vector, so the system $\mathcal{I} + d\mathcal{I}$ generated by ω and $d\omega$ does not have integral planes. In this case the system $\mathcal{I}$ does not have any integral manifolds; i.e., it is not integrable.

Theorem 1.9.9 reduces the problem of finding the integral manifolds of the system $\mathcal{I}$ to the same problem for the closure $\mathcal{I} + d\mathcal{I}$.

The following definition says that a system is called *integrable* if it has an integral manifold of maximal dimension through each point.

Definition 1.9.10. *Let $\mathcal{I}$ be an exterior differential system on the open set $U \subset \mathbb{R}^m$ generated by the functionally independent one-forms $\theta_1, \dots, \theta_k$. The system $\mathcal{I}$ is called integrable on U if for any $x \in U$ there is an integral manifold M_x of $\mathcal{I}$ passing through x such that $\dim M_x = k$.*

The system $\mathcal{I}$ is called nonintegrable if it is not integrable. In sub-Riemannian geometry we deal with nonintegrable exterior differential systems.

1.10 Formulas Involving Lie Derivative

Let U be an open subset of $\mathbb{R}^m$ and $\omega \in \Omega^p$ be a p-form on U. If X is a vector field on U, then the Lie derivative of ω with respect to X is a p-form on U, i.e., $L_X\omega \in \Omega^p$, defined by

$$\Big(L_X\omega\Big)(Y_1, \dots, Y_p) = X\omega(Y_1, \dots, Y_p) - \sum_{i=1}^{p} \omega\big(Y_1, \dots, [X, Y_i], \dots, Y_p\big), \tag{1.10.5}$$

where Y_j are vector fields on U. In particular, when ω is a one-form, we have

$$\Big(L_X\omega\Big)(Y) = X\omega(Y) - \omega\big([X, Y]\big).$$

When $\omega = f \in \Omega^0$ is a function, we have

$$L_X f = X(f).$$

The *exterior derivative* $d : \Omega = \bigoplus_{p=0}^{m} \to \Omega$ is defined by

$$d\omega(X_1, \dots, X_{r+1}) = \sum_{i=1}^{r+1} (-1)^{i+1} X_i \omega(X_1, \dots, \widehat{X}_i, \dots, X_{r+1})$$
$$+ \sum_{i<j} (-1)^{i+j} \omega\big([X_i, X_j], X_1, \dots, \widehat{X}_i, \dots, \widehat{X}_j, \dots X_{r+1}\big),$$

where $\widehat{X}_i$ means that X_i is missing from the argument. In the case when ω is a one-form,

$$d\omega(X_1, X_2) = X_1\omega(X_2) - X_2\omega(X_1) - \omega\Big([X_1, X_2]\Big).$$

Proposition 1.10.1. *The operator d satisfies the following properties:*

(1) $df(X) = X(f), \ \forall f \in \Omega^0$
(2) $d(\alpha\omega + \beta\eta) = \alpha d\omega + \beta d\eta, \ \forall \alpha, \beta \in \mathbb{R}, \omega, \eta \in \Omega$
(3) $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^p \omega \wedge d\eta, \ \forall \omega \in \Omega^p$
(4) $d^2\omega = 0, \ i.e.\ d(d\omega) = 0$
(5) $d(\phi^*\omega) = \phi^*(d\omega),$

where $\phi : V \to U$ is a smooth map and $\phi^*\omega(X_1, \dots, X_p) = \omega(\phi_* X_1, \dots, \phi_* X_p)$, where ϕ_* is the tangent application given in local coordinates as the Jacobian of ϕ.

The proof of the proposition is left to the reader.

The next definition is introducing the concept of *interior multiplication.*

Definition 1.10.2. *Let $\omega \in \Omega^p$ and X be a vector field on the open domain U in $\mathbb{R}^m$. Then the $(p-1)$-form $i_X\omega$ defined by*

$$i_X\omega(X_1, \dots, X_{p-1}) = \begin{cases} 0 & \textit{if } \ p = 0 \\ \omega(X, X_1, \dots, X_{p-1}) & \textit{if } \ p \geq 1 \end{cases}$$

is called the interior multiplication of X with ω.

The relation among the interior multiplication i_X, the exterior derivative d, and the Lie derivative L_X is given by the following magic decomposition result.

Theorem 1.10.3 (Cartan). *For $\omega \in \Omega^p$ we have the decomposition*

$$L_X\omega = i_X(d\omega) + d(i_X\omega).$$