R.W. Sharpe

Differential Geometry

Cartan's Generalization of Klein's Erlangen Program

Foreword by S.S. Chern

With 104 Illustrations

Springer

R.W. Sharpe
Department of Mathematics
University of Toronto
Toronto, M5S 1A1
Canada

Mathematics Subject Classification (1991): 53-01, 53A40, 53A20, 53B21

Library of Congress Cataloging-in-Publication Data
Sharpe, R.W. (Richard W.)
Differential geometry : Cartan's generalization of Klein's Erlangen program / by R.W. Sharpe.
p. cm. – (Graduate texts in mathematics ; 166)
Includes bibliographic references and index.
ISBN 0-387-94732-9 (hard : alk. paper
1. Geometry, Differential. I. Title. II. Series.
QA641.S437 1996
516.3′6–dc20 96-13757

Printed on acid-free paper.

Production managed by Robert Wexler; manufacturing supervised by Jacqui Ashri.
Photocomposed pages prepared by Sirovich, Warren, R.I. in TEX using Springer's svsing macro.
Printed and bound by R.R Donnelley and Sons, Harrisonburg, VA.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-94732-9 Springer-Verlag New York Berlin Heidelberg SPIN 10534001

Foreword

I am honored by Professor Sharpe's request to write a forward to his beautiful book.

In his preface he asks the innocent question, "Why is differential geometry the study of a connection on a principal bundle?" The answer is of course very simple; because Euclidean geometry studies a connection on a principal bundle, and all geometries are in a sense generalizations of Euclidean geometry.

In fact, let E^n be the Euclidean space of n dimensions. We call an orthonormal frame $x, e_1, \ldots, e_n$ ($n+1$ vectors), where x is the position vector and e_i have the scalar products

$$(e_i, e_j) = \delta_{ij}, \qquad 1 \le i, j \le n.$$

Then the space of all orthonormal frames is a principal fiber bundle with group $O(n)$ and base space E^n, the projection being defined by mapping $x, e_1, \ldots, e_n$ to x. The equations

$$de_i = \sum_{1 \le j \le n} \omega_{ij} e_j, \qquad 1 \le i \le n,$$

define the Maurer–Cartan forms ω_{ij}, with

$$\omega_{ij} + \omega_{ji} = 0, \qquad 1 \le i, j \le n.$$

They satisfy the Maurer–Cartan equations

$$d\omega_{ij} = \sum_{1 \le k \le n} \omega_{ik} \wedge \omega_{kj}, \qquad 1 \le i, j \le n.$$

This is Euclidean geometry by moving frames. The ω_{ij} define the parallelism or connection. The Maurer–Cartan equations say that the connection is flat. This formulation has a great generalization.

As in all disciplines, the development of differential geometry is tortuous. The basic notion is that of a manifold. This is a space whose coordinates are defined up to some transformation and have no intrinsic meaning. The notion is original, bold, and powerful. Naturally, it took some time for the concept to be absorbed and the technology to be developed. For example, the great mathematician Jacques Hadamard "felt insuperable difficulty ... in mastering more than a rather elementary and superficial knowledge of the theory of Lie groups," a notion based on that of a manifold [1]. Also, it took Einstein seven years to pass from his special relativity in 1908 to his general relativity in 1915. He explained the long delay in the following words: "Why were another seven years required for the construction of the general theory of relativity? The main reason lies in the fact that it is not so easy to free oneself from the idea that coordinates must have an immediate metrical meaning." [2]

On the technology side the breakthrough was achieved by the tensor analysis of Ricci calculus. The central theme was Riemannian geometry, which Riemann formulated in 1854. Its fundamental problem is the "form problem": To decide when two Riemannian metrics differ by a change in coordinates. This problem was solved by E. Christoffel and R. Lipschitz in 1870. Christoffel's solution introduces a covariant differentiation, which could be given an elegant geometrical setting through the parallelism of Levi–Civita. Tensor analysis is extremely effective and has dominated differential geometry for a century.

Another technical tool, which has not quite received the recognition it deserves, is the exterior differential calculus of Elie Cartan. This was introduced by Cartan in 1922, following the work of Frobenius and Darboux. All the exterior differential forms on a manifold form a ring. It depends only on the differentiable structure of the manifold and not on any additional structure such as a Riemannian metric or an affine connection. Topologically it leads to the de Rham theory. Less known is its effectiveness in treating local problems.

A fundamental question is the equivalence problem for G-structures: Given, on an n-dimensional manifold with coordinates u^i, a set of linear differential forms ω^i, a similar set ω^{*j} with coordinates u^{*j}, and a subgroup $G \subset Gl(n, \mathbf{R})$, determine the conditions under which there exist functions

$$u^{*j} = u^{*j}(u^1, \ldots, u^n), \qquad 1 \leq i, j \leq n,$$

such that after substitution the ω^{*j} differ from the ω^j by a transformation of G. The form problem in Riemannian geometry is the case $G = O(n)$.

The solution of the form problem by Cartan's method of equivalence leads automatically to the tensor analysis. Thus, the method of equivalence is more general. In the case $G = O(n)$, this leads to the Levi–Civita parallelism and the Riemannian geometry. In this way Euclidean geometry generalizes to Riemannian geometry. For a general G, the solution of the equivalence problem is not always easy (cf. the Preface), although it is proved that it can always be achieved in a finite number of steps. Philosophically nice problems have nice answers.

Klein geometry can be developed through the Maurer–Cartan equations. The generalization of the above discussion, from $O(n)$ to G, gives Cartan's generalized spaces, essentially a connection in a principal bundle.

A fundamental problem is the relation of the local geometry with the global properties of the spaces in question. Such a result is the so-called Chern–Weil theorem that the characteristic classes can be represented by differential forms constructed explicitly from the curvature. The simplest result is the Gauss–Bonnet formula.

I wish to take this occasion to mention some recent developments on Finsler geometry [3]. This is the geometry of a very simple integral and was discussed in problem 23 of Hilbert's Paris address in 1900. By a proper interpretation of the analytical results, Finsler geometry now assumes a very simple form showing it to be a family of geometries quite analogous to the Riemannian case.

Differential geometry offers an open vista of manifolds with structures, finite or infinite dimensional. There are also simple and difficult low-dimensional problems, of the garden variety. If one switches between the two, life is indeed very enjoyable.

It is a great mystery that the infinitesimal calculus is a source of such depth and beauty.

References

[1] J. Hadamard, *Psychology of Invention in the Mathematical Field*, Princeton University Press, Princeton, 1945, p. 115.

[2] A. Einstein, Autobiographical Notes, in *Albert Einstein: Philosopher Scientists*, p. 67, 2nd ed., 1949, in vol. 7 of the series The Library of Living Philosophers, Evanston, IL, edited by P.A. Schilpp.

[3] D. Bao and S.S. Chern, On a notable connection in Finsler geometry, *Houston J. Math.* **19** (1993), 135–180.

S.S. Chern
Mathematical Sciences Research Institute
Berkeley, CA 94720, USA

Preface

This book is a study of an aspect of Elie Cartan's contribution to the question "What is geometry?"

In the last century two great generalizations of Euclidean geometry appeared. The first was the discovery of the non-Euclidean geometries. These were organized into a coherent whole by Felix Klein, who recognized them as various examples of coset spaces G/H of Lie groups. In this book we refer to these latter as Klein geometries. The second generalization was Georg Riemann's discovery of what we now call Riemannian geometry. These two theories seemed largely incompatible with one other.[1]

In the early 1920s Elie Cartan, one of the pioneers of the theory of Lie groups, found that it was possible to obtain a common generalization of these theories, which he called *espaces généralizés* and we call Cartan geometries (see diagram).

$$\begin{tikzcd}
\textit{Euclidean Geometry} \arrow[r, "\textit{generalization}"] \arrow[d, "\textit{generalization}"] & \textit{Klein Geometries} \arrow[d, "\textit{generalization}"'] \\
\textit{Riemannian Geometry} \arrow[r, "\textit{generalization}"] & \textit{Cartan Geometries}
\end{tikzcd}$$

[1] The only relationship was the "accident" that some of the non-Euclidean geometries could be regarded as special cases of Riemannian geometries.

Looking at this diagram vertically, we can say that just as a Riemannian geometry may be regarded, locally, as modeled on Euclidean space but made "lumpy" by the introduction of a curvature, so a Cartan geometry may be regarded, locally, as modeled on one of the Klein geometries but made "lumpy" by the introduction of curvature appropriate to the model in question. Looking at the same diagram horizontally, a Cartan geometry may be regarded as a non-Euclidean analog of Riemannian geometry.

Cartan actually gave the first example of a Cartan geometry more than a decade earlier, in the remarkable *tour de force* [E. Cartan, 1910]. In that paper he considered the case of a two-dimensional distribution on a five-dimensional manifold. He showed that such a distribution determined, and was determined by, a Cartan geometry modeled on the homogeneous space G_2/H, where H is a certain nine-dimensional subgroup of the fourteen-dimensional exceptional Lie group G_2. This process of associating a Cartan geometry to a raw geometric entity (the distribution) is an example of "solving the equivalence problem" for the entity in question. Although the solution of an equivalence problem is not always a Cartan geometry, in many important cases it is. When it is, the invariants of the geometry (curvature, etc.) are a priori invariants of the raw geometric entity. We recommend [R.B. Gardner, 1989] for an account of the method of equivalence.

To be a little more precise, a Cartan geometry on M consists of a pair (P, ω), where P is a principal bundle $H \to P \to M$ and ω, the Cartan connection, is a differential form on P. The bundle generalizes the bundle $H \to G \to G/H$ associated to the Klein setting, and the form ω generalizes the Maurer–Cartan form ω_G on the Lie group G. In fact, the *curvature* of the Cartan geometry, defined as $d\omega + \frac{1}{2}[\omega, \omega]$, is the complete local obstruction to P being a Lie group.

One reason for the power of Cartan's method comes from the fact that these new geometries maintain the same intimate relation with Lie groups that one sees in the case of homogeneous spaces. This means, for example, that constructions in the theory of homogeneous spaces often generalize in a simple manner to the general "curved" case of Cartan geometries. It also means that the differential forms that appear are always related to components of the Maurer–Cartan form of the Lie group, a context in which their significance remains clear.

In the particular case of a Riemannian manifold M, Cartan's point of view offered a new and profound vantage point that is largely responsible for the modern insistence on "doing differential geometry on the bundle P of orthonormal frames over M."

The history of the study of Cartan geometries is somewhat troubled. First is the difficulty Cartan faced in trying to express notions for which there was no truly suitable language.[2] Next is the widely noted difficulty in reading Cartan.[3] In his paper [C. Ehresmann, 1950] Charles Ehresmann gave for the first time a rigorous global definition of a Cartan connection as a special case of a more general notion now called an Ehresmann connection (or more simply, a connection). For various reasons[4] the Ehresmann definition was taken as the definitive one, and Cartan's original notion went into a more or less total eclipse for a long time. The beautiful geometrical origin and insight connected with Cartan's view were, for many, simply lost. In short, although the Ehresmann definition gives us a good notion, it hides the real story about why it is so good. In this connection, the following quotation is interesting [S.S. Chern, 1979]:

> The physicist C.N. Yang wrote [C.N. Yang, 1977]: "That non-abelian gauge fields are conceptually identical to ideas in the beautiful theory of fibre bundles, developed by mathematicians without reference to the physical world, was a great marvel to me." In 1975 he mentioned to me: "This is both thrilling and puzzling, since you mathematicians dreamed up these concepts out of nowhere."

Far from arising "out of nowhere," the simple and compelling geometric origin of a connection on a principal bundle is that it is a generalization of the Maurer–Cartan form. Moreover, a study of the Cartan connection itself can illuminate and unify many aspects of differential geometry.

Novelties

Aside from the fact that one cannot find a fully developed, modern exposition of Cartan connections elsewhere, what is new or different in this book?

New Treatment

This book is written at a level that can be understood by a first- or second-year graduate student. In particular, we include the relevant theory of manifolds, distributions and Lie groups. For us, a manifold is, by definition, a

[2] This difficulty was resolved with the introduction of the notion of a principal bundle and of vector-valued forms on such a bundle.

[3] To paraphrase Robert Bryant, "You read the introduction to a paper of Cartan and you understand nothing. Then you read the rest of the paper and still you understand nothing. Then you go back and read the introduction again and there begins to be the faint glimmer of something very interesting."

[4] At that stage it was easier to read Ehresmann than Cartan. There was also the attraction of a more general and global notion.

locally Euclidean, paracompact Hausdorff space. This is the same as a locally Euclidean Hausdorff space each of whose components has a countable basis.[5] In particular, Lie groups are defined to be manifolds in this sense. The result of Yamabe and Kuranishi ([H. Yamabe, 1950]) that a connected subgroup of a Lie group is a Lie subgroup implies that *any* subgroup of a Lie group is a Lie group in the present sense. The discussion of submanifolds given in Chapter 1 is broad enough to include these subgroups as submanifolds.

In our coverage of bundle theory, we emphasize the abstract principal bundles rather than bundles of frames.[6] Of course, these two views are really equivalent. In the case of the "first-order" geometries, the equivalence is quite simple. However, in the case of "higher-order" geometries, the choice of the higher-order frames usually seems to be decided on a rather ad hoc basis and can be complicated. Here the bundle approach gives a real advantage, and the right choice of frames becomes clear (if needed) once the bundle is understood. Another important advantage of working with the bundles themselves is that they give a common language, facilitating comparison between geometries and emphasizing the relation to the model space. In this sense, comparing Cartan geometries is like comparing Klein geometries.

Chapter 3 contains a complete and economical development of the Lie group—Lie algebra correspondence based on the *fundamental theorem of non-abelian calculus*. One of the novelties here is the characterization of a Lie group as a manifold equipped with a Lie algebra-valued form on it satisfying certain properties. This characterization prepares the reader for the generalization to Cartan geometries in Chapter 5.

Finally, in Appendix B we explain how one manifold may roll without slipping or twisting on another in Euclidean space. We also show how this notion yields a differential system that contains both the Levi–Civita connection and the Ehresmann connection on the normal bundle for a submanifold of Euclidean space.

New Results

Let us move on to some results we believe are new. In Chapter 4 we introduce the fundamental property of Klein geometries characterizing the kernel of such a geometry. This result is used in Chapter 5 in an essential way to show the equivalence of the base and bundle definitions of Cartan geometries in the effective case. In Chapter 5 we introduce and classify Cartan space forms. These geometries generalize the classical Riemannian

[5]The usual definition requires a manifold to have a countable basis (cf., e.g., [Boothby, W. 1986, p. 6]).

[6]In much the same way, one might emphasize an abstract Lie group rather than a matrix group realizing it.

space forms.[7] One important ingredient of this classification is the property (apparently new) of a Cartan geometry called "geometric orientability." Another is the notion of "model mutation." Finally, in Chapter 7 we give a classification of the submanifolds of a Möbius geometry. This classification is more general than that of [A. Fialkow, 1944] in that ours allows the presence of umbilic points.

Prerequisites and Conventions

This book assumes very few prerequisites. The reader needs to be familiar with some basic ideas of group theory, including the notion of a group acting on a set. Results from the calculus of several variables, point set topology, and the theory of covering spaces are used in various places, and the long, exact sequence of homotopy theory is used once (at the end of Chapter 5). Aside from this, most of the material is developed *ab initio*. However, the reader is invited to shoulder some of the burden of the work in that essential use is made of a few of the exercises. These exercises are denoted by an asterisk to the right of the exercise number.

The numbering follows a single sequence throughout the book, with all items (definitions, theorems, figures, etc.) in a single stream. Thus 4.3.2 refers to Chapter 4, Section 3, item 2. For references to items occurring in the same chapter, we omit the chapter number, so that in Chapter 4, 4.3.2 becomes 3.2.

We use the following dictionary of symbols to denote the ends of various items:

symbol	*end of*
❀	definition
❏	exercise
■	proof
◆	example

Although it will often be convenient for us to write column vectors as row vectors, the reader should remember that all vectors are in fact column vectors.

[7]In fact, this notion is general enough to immediately allow a description of general symmetric spaces.

Limitations

The reader will find no mention here of some basic topics in differential geometry, such as Stokes' theorem, characteristic classes, and complex geometries. Also, our approach to Lie theory is "elementary" in that we do not discuss or use the classification theory of Lie groups, with its attendant study of roots, weights, and representations.

Originally, we had wished to include more than the three examples of Cartan geometries studied here; but in the end, the pressures of time, space, and energy limited this impulse. The three geometries we do study are not developed in complete analogy to each other. For example, the discussion of immersed curves in a Möbius geometry in terms of the normal forms given in Chapter 7 does of course have a Riemannian analog, but that is not studied in this book. And one may study subgeometries of projective geometries just as one studies subgeometries of Riemannian and Möbius geometries, but we do not do so here. We have also resisted the impulse to make a "dictionary" translating among the various versions of Cartan's view, Ehresmann's view,[8] and the view expressed in [L.P. Eisenhart, 1964]. In the end, however, for those who are interested in it, it should be abundantly clear how Cartan's view does illuminate the others.

Some Personal Remarks

An author often writes a book in order to sort out his or her own understanding of the subject. This is the circumstance in the present case. When I was an undergraduate, differential geometry appeared to me to be a study of curvatures of curves and surfaces in $\mathbf{R}^3$. As a graduate student I learned that it is the study of a connection on a principal bundle. I wondered what had become of the curves and surfaces, and I studied topology instead.

The reawakening of my interest in this subject began in 1987 when Tom Willmore very kindly wrote me a note thanking me for a preprint and mentioning his great interest in what is known as the Willmore conjecture (cf. **7**.6). This led me once again to look at principal bundles and connections. In particular, I wondered whether there was an intrinsically defined Ehresmann connection on a surface in S^3 that was invariant under the group of Möbius transformations of S^3. It turns out there is no such connection. However, after calculating normal forms for surfaces in the Möbius sphere S^3 (cf. [G. Cairns, R. Sharpe, and L. Webb, 1994]), it became clear to me that there must be some other kind of invariantly defined structure inherited on the surface from its embedding in S^3. (In Chapter 7 it is shown that a Cartan connection is defined in this situation, and, in fact, Cartan also knew this [E. Cartan, 1923].)

[8]See, however, the discussion in Appendix A dealing with the relationship between Cartan and Ehresmann connections.

During this time it began to seem strange to me that Ehresmann connections play such a prominent role in modern differential geometry. In some cases, such as the Levi–Civita connection, the connection is determined by the geometry. In many cases, however, one makes use of an arbitrary connection that one proves to exist by a general technique. This is the appropriate point of view for the construction of the characteristic classes of Chern and Pontryagin. There one may use any connection, since the aim is to obtain topological invariants for which the particular choice of connection does not matter. But these considerations seem to be at their base topological rather than differential geometric. My innocent question, left over from my undergraduate days, was "Why is differential geometry the study of a connection on a principal bundle?" And I began, rather impertinently, to ask this question at every opportunity, usually picking on some unsuspecting differential geometer who did not know me very well.

During one of these sessions, Min Oo remarked that Elie Cartan had considered connections with values in a Lie algebra larger than that of the fiber.[9] Later I read, and translated,[10] Cartan's book [E. Cartan, 1935]. I browsed through Cartan's collected works and through those of his successors and interpreters. It became clear to me that Cartan had a subtle and really wonderful idea, which gives a fully satisfying explanation for the modern, and approximately true, notion that differential geometry is the study of an Ehresmann connection on a principal bundle. There seems to be no treatment of these things in the standard texts on differential geometry. In the few books where the Cartan connections are mentioned at all (e.g., [J. Dieudonné 1974], [W.A. Poor, 1981], and [M. Spivak, 1979]), they make only a brief appearance, perhaps in the exercises or toward the end of the book, and one is left with the impression that the notion is only a quaint curiosity left over from bygone days. Six years ago I began to scribble some notes about these things and to talk about them; after a number of months had passed, I realized I was writing a book on the subject.

⋆ ⋆ ⋆

I would like to thank everyone who has had an influence on this book. In addition to those mentioned above, I am grateful to Bernard Kamte, Joe Repka, Qunfeng Yang, and my wife, Mary, for their comments on portions of the manuscript. I would also like to acknowledge my gratitude to

[9]See [E. Ruh, 1993] for a brief recent overview of Cartan connections and some of their applications.
[10]A copy of my translation, which is only a rough draft, can be found in the Mathematics Library at the University of Toronto.

the National Science and Engineering Research Council of Canada for its support of this project through grant #OGP0004621.

Finally, I would like to thank Velamir Jurdjevic for his encouragement over the years. It was Vel who suggested that, although it is perhaps impossible to catch all the errors before a book reaches print, the principal demand is that a book be interesting. As for the first part of his remark, the responsibility for any remaining errors lies with me. I will leave it to the reader to judge whether or not this principal demand has been met.

Richard Sharpe
Toronto, Canada
October 16, 1995

Contents

1

In the Ashes of the Ether: Differential Topology

It must be agreed that "hypergeometry" seems to have been devised in order to strike the imagination of people who have not enough mathematical knowledge to be aware of the true character of an algebraic construction expressed in geometric terms, for that is what "hypergeometry" really is. –R. Guénon, 1945

Is Euclidean geometry true? It has no meaning. ... One geometry cannot be more true than another; it can only be more convenient. –H. Poincaré, 1902

I attach special importance to the view of geometry which I have just set forth, because without it I should have been unable to formulate the theory of relativity. –A. Einstein, 1922

Although several mathematicians, especially C.F. Gauss, studied the notion of a smooth manifold in special cases, the idea of an abstract manifold of arbitrary finite dimension seems to be due to Riemann. Mathematicians were led to these notions only slowly. As the idea of a vector space of dimension higher than three became acceptable in the last century, algebraic geometers began to study the solutions of polynomial equations in many variables. For example, they studied the algebraic curves in the complex projective plane, which in the real sense is roughly the study of ordinary surfaces in four-dimensional space. At the same time, mathematical physicists interested themselves in six-dimensional space, the *state space* of a single particle with three position and three momentum variables; if N

particles are considered, the dimension of the state space jumps to $6N$. The seamless algebraic passage between the ordinary notions of space and their higher-dimensional analogs must have prepared the way for the acceptance of Riemann's abstract manifolds. But the willingness and even the necessity of regarding these developments as *truly geometrical* was slow to take hold and was fiercely resisted in some quarters.[1]

Riemann's work was followed with papers by Christoffel, Ricci, and Levi–Civita. Klein and Lie were interested in the study of Lie groups and their homogeneous spaces, which are again examples of manifolds, albeit very special ones. A real boost to the subject came with Einstein's discovery of general relativity in 1915. At this stage it became much clearer that the fourth dimension might be scientifically regarded as more than some geometrical fantasy. Einstein himself regarded the abstract four-manifold[2] as what remains of the "ether" in general relativity (cf. [A. Einstein, 1922]). In the early years these spaces were thought about only locally, as pieces of four-dimensional Euclidean space. Later the notion of "cosmology" appeared (cf. [D. Howard and J. Stachel, eds., 1989]) and began to show the influence of the global topology. Perhaps we may say that in studying smooth manifolds we are studying the possible shapes of the ether.

In the hierarchy of geometry (whose "spine" rises from homotopy theory through cell complexes, through topological and smooth manifolds to analytic varieties), the category of smooth manifolds and maps lies "halfway" between the global rigidity of the analytic category and the almost total flabbiness of the topological category. We might say that a smooth manifold possesses *full infinitesimal rigidity* governed by Taylor's theorem while at the same time having absolutely no rigidity relating points that are not "infinitesimally near" each other, as is seen by the existence of partitions of unity (cf. [W. Boothby, 1986], pp. 193–195). Smooth manifolds are sufficiently rigid to act as a support for the structures of differential geomtery while at the same time being sufficiently flexible to act as a model for many physical and mathematical circumstances that allow independent local perturbations. Perhaps the smooth "substance" may be regarded as a mathematical model for Aristotle's *materia prima* or the Hindu *prakriti.*

In this chapter we show how the differential calculus lives on after the death of a preferred coordinate system. In particular, we discuss some of the beautiful and elementary constructions surrounding the notion of a smooth

---

[1]Cf. [M. Monastyrski, 1987], pp. 22 and 64.

[2]Einstein admits [A. Einstein, 1922] his difficulty in conceiving space-time without a metric. In fact, he spent the rest of his life searching for a geometric structure on a four-manifold supplementing the metric to include electromagnetism within a unified field theory generalizing his general theory of relativity. The general question of "what geometry on a manifold supports physics?" remains vital to the present day. On the other hand, the context for all such structures seems to continue to include a smooth manifold.

manifold. Care is taken to show the relation between the forms of these constructions for the concrete situation of manifolds embedded in an ambient Euclidean space and those for an abstract, coordinate-independent, manifold.

§1. Smooth Manifolds

The definitive modern definition of a smooth manifold seems to have been given by Hassler Whitney ([H. Whitney, 1936]), in which a smooth manifold is presented as floppy pieces of Euclidean space glued together with a sort of differentiable glue. But in order to gain perspective, we start our more detailed discussion with *topological* manifolds.

Topological Manifolds

Definition 1.1. Let M be a paracompact[3] Hausdorff space. We call M *an n-dimensional topological manifold* (and write it M^n if we wish to denote the dimension) if for each point $p \in M$ there is an open set U in M containing p such that U is homeomorphic to an open subset of $\mathbf{R}^n$ by some homeomorphism φ.[4] Such a pair (U, φ) is called a *local coordinate system* or (in the maritime terminology) a *chart* on M. On the other hand, φ^{-1} is called a *local parameterization* of M. Often, however, speaking loosely, both φ and φ^{-1} are referred to as coordinate systems. ❀

We note that if M is an n-manifold, then so is every open subset of it; in particular, the components of M are also manifolds. On the other hand, manifolds are quite badly behaved with respect to quotients. Taking the quotient of a manifold by almost any equivalence relation on it leads out of the category of manifolds. Some notable exceptions to this are studied later under the name of *fiber bundles.*

Example 1.2 (the n-sphere). Let $M^n = S^n = \{x \in \mathbf{R}^{n+1} \mid x \cdot x = 1\}$. Set $U = \{x \in S^n \mid x_{n+1} > -1\}$. Then U is an open subset of S^n, as is its homeomorphic image $\rho(U)$, where $\rho: \mathbf{R}^{n+1} \to \mathbf{R}^{n+1}$ is the reflection in the hyperplane $x_{n+1} = 0$. The open sets U and $\rho(U)$ together form a covering for S^n. We can see that U (and hence $\rho(U)$) is homeomorphic to $\mathbf{R}^n$ by considering the stereographic projection φ from the point $-e_{n+1}$ given by

---

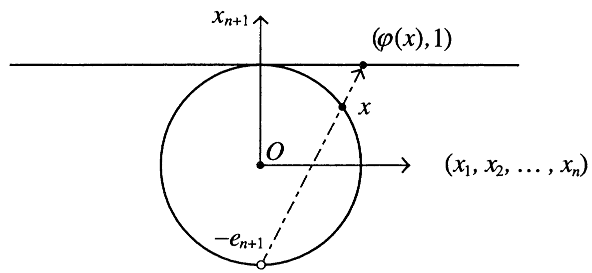
[3]This is equivalent to each component of M having a countable basis for its topology (cf. [J. Dugundiji, 1966], p. 241).

[4]We note that only the topological structure of $\mathbf{R}^n$ is brought into play here. In particular, neither the Euclidean structure nor the affine space structure has any role. Nevertheless, we shall continue to refer to $\mathbf{R}^n$ as *Euclidean space.*

$\varphi\colon U \to \mathbf{R}^n$ sending

$$(x_1, \ldots, x_{n+1}) \mapsto \frac{1}{1 + x_{n+1}}(x_1, \ldots, x_n).$$

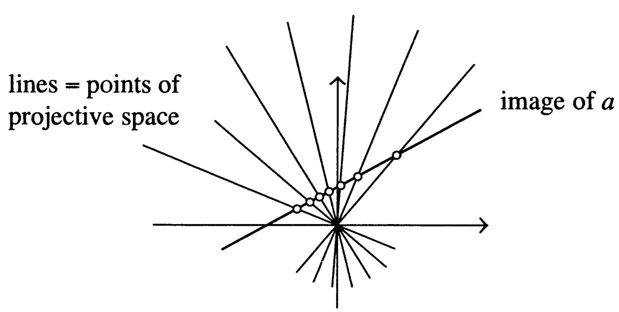
This situation is pictured below. ◆



Example 1.3 (Projective space). Let $P^n(\mathbf{R})$ = the set of one-dimensional subspaces of $\mathbf{R}^{n+1}$. The surjective map $p\colon \mathbf{R}^{n+1} - \{0\} \to P^n(\mathbf{R})$, sending $v \mapsto \langle v \rangle$ (= the line spanned by v), induces the quotient topology on $P^n(\mathbf{R})$.

Let us show that the topological space $P^n(\mathbf{R})$ is Hausdorff. Any two distinct points of $P^n(\mathbf{R})$ may be represented by vectors $v_1, v_2 \in S^n$ satisfying $v_1 \cdot v_2 > 0$. Choose $\varepsilon > 0$ to be less than half the distance from v_1 to v_2. Then no line through the origin meets both $B_\varepsilon(v_1)$ and $B_\varepsilon(v_2)$, where $B_\varepsilon(v)$ denotes the open ball in $\mathbf{R}^{n+1}$ of radius ε about v. It follows that $\langle v_1 \rangle, \langle v_2 \rangle \in P^n(\mathbf{R})$ lie in disjoint, open sets $p(B_\varepsilon(v_1)), p(B_\varepsilon(v_2)) \subset P^n(\mathbf{R})$.

If $a\colon \mathbf{R}^n \to \mathbf{R}^{n+1}$ is any affine map whose image does not contain the origin, then the composite map $pa\colon \mathbf{R}^n \to P^n(\mathbf{R})$ is injective and continuous.



Moreover, if $V \subset \mathbf{R}^n$ is open, then

$$p^{-1}(pa(V)) = \{\lambda v \in \mathbf{R}^{n+1} \mid \lambda \in \mathbf{R}^*, v \in a(V)\}$$

is open in $\mathbf{R}^{n+1}$. Thus, pa is a homeomorphism onto its image. It is called the *affine parameterization*, and its inverse is called the *affine coordinate system* for $P^n(\mathbf{R})$ arising from the affine map a.

Since every point of $P^n(\mathbf{R})$ lies in the image of some of the affine parameterization, it follows that $P^n(\mathbf{R})$ is a topological manifold.

A similar construction may be made for an arbitrary vector space V, yielding the projective space $P(V)$ canonically associated to V. ◆

Smooth Manifolds

The following definition comes into sharper focus if the meanings of the maritime terminology of "charts" and "atlas" are given due consideration.

Definition 1.4. If M^m is a topological manifold, then a (*smooth*) *atlas* on M is a collection $\mathcal{A} = \{(U_i, \varphi_i)\}$ of charts such that

(i) the U_is form an open covering of M, and

(ii) for each pair of charts (U, φ) and (V, ψ) in $\mathcal{A}$, the map

$$\Phi = \psi\varphi^{-1} \mid \varphi(U \cap V)\colon \varphi(U \cap V) \to \psi(U \cap V)$$

is a smooth[5] (i.e., C^∞) map between open sets in Euclidean space (*change of coordinates*; see Figure 1.6). For brevity, we may speak of the *chart* φ, leaving U = domain(φ) nameless. ❀

Example 1.5. Continuing Example 1.2 of the n-sphere, we have the following atlas consisting of two charts, $\mathcal{A} = \{(U, \varphi), (\rho(U), \varphi\rho)\}$. The composite $\varphi(\rho\varphi^{-1})\colon \mathbf{R}^{n+1} - \{0\} \to \mathbf{R}^{n+1} - \{0\}$ is easily seen to be smooth, so $\mathcal{A}$ is a smooth atlas. ◆

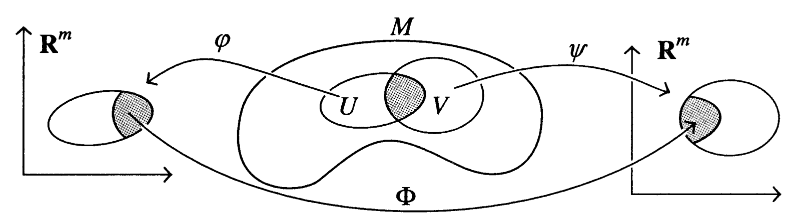


FIGURE 1.6. Change of coordinate system.

Exercise 1.7. show that the affine coordinate systems described in Example 1.3 constitute an atlas for $\mathbf{P}^n(\mathbf{R})$. ❑

---

[5]Here we use not only the topological structure of $\mathbf{R}^n$ but also its affine structure so that the notion of differentiation makes sense. The choice of origin and the Euclidean inner product are not involved.

The main point of introducing smooth atlases is to be able to unambiguously differentiate the composite appearing in Definition 1.4(ii) as often as we please. For this simple purpose the particular choice of atlas is not important, so we are led to make the definition that two atlases are *equivalent* if their union is also an atlas. This in turn leads to the following definition.

Definition 1.8. A *smooth structure* on a topological manifold is an equivalence class of atlases, and a *smooth manifold* is a topological manifold with a specified smooth structure. ❀

For example, the atlas for the n-sphere above endows it with a smooth structure called the *canonical smooth structure.*

Starting with a given atlas, we can enlarge it to the union of all the atlases equivalent to it. This will give a unique *maximal* atlas equivalent to the given one, and the smooth structure may be identified with the corresponding maximal atlas. In particular, if φ is a chart in the maximal atlas of a smooth manifold, and f is a local diffeomorphism of $\mathbf{R}^n$ with image$(\varphi) \subset$ domain(f), then $f\varphi$ is also a chart in the maximal atlas. This kind of procedure allows a great deal of flexibility in the choice of chart. For example, we can "tidy up" the chart φ with respect to a point $p \in U$ by following φ with a translation in $\mathbf{R}^n$ so that the new chart sends p to $0 \in \mathbf{R}^n$. We may further follow φ by some linear transformation of $\mathbf{R}^n$ to move the various directions into convenient positions. We shall always implicitly assume we are dealing with a maximal atlas, even if it is described by a particular choice of a nonmaximal one.

Orientability

Definition 1.9. Let (U, φ) and (V, ψ) be two charts for the smooth manifold M^n. We say these charts are *compatible* on $W \subset U \cap V$ if the change of coordinate mapping $\Phi = \psi\varphi^{-1} \mid \varphi(U \cap V)$ has positive Jacobian determinant at each point of $\varphi(W)$. We say they are *compatible* if they are compatible on $U \cap V$. ❀

Remark 1.10. Since the sign of the determinant of the Jacobian matrix $(\psi \circ \varphi^{-1})'(x)$ is constant on each component of $\varphi(U \cap V)$, it follows that if (U, φ) and (V, ψ) are compatible at $x \in U \cap V$, then they are compatible on the component of $U \cap V$ containing x. On the other hand, if they are not compatible at a point $x \in U \cap V$, then the charts (U, φ) and $(V, \phi\psi)$ are compatible, where ϕ is any linear transformation of $\mathbf{R}^n$ of negative determinant.

Definition 1.11. Let M be a smooth manifold. An atlas for M is *oriented* if any two charts in it are compatible. M is called *topologically orientable* if it has an oriented atlas. A maximal such atlas is called a *topological orientation* for M. ❀

Given an oriented atlas for M, we can always enlarge it to obtain a unique maximal, oriented atlas containing it. Thus, an oriented atlas for M determines an orientation of M.

Exercise 1.12. Show that a connected, orientable, smooth manifold M has exactly two orientations. ❏

The question of whether or not a given manifold is orientable can be determined by studying the loops on it. In preparation for this, we are going to study paths on M of the form

$$\sigma: (I, 0, 1) \to (M, p, q).$$

Definition 1.13. Let σ be a path on M^n. Suppose we are given a partition $0 = t_0 < t_1 < \ldots < t_k = 1$ and a family of charts (U_i, φ_i), $1 \le i \le k$, such that $\sigma([t_{i-1}, t_i]) \subset U_i$, $1 \le i \le k$. We call these charts *compatible along* σ if (U_i, φ_i) and (U_{i+1}, φ_{i+1}) are compatible at $\sigma(t_i) \in U_i \cap U_{i+1}$ for $1 \le i \le k-1$. ❀

Lemma 1.14. *For any path σ on M^n there exists a family of compatible charts along σ.*

Proof. Since I is compact and

$$\{\sigma^{-1}(U) \mid U \text{ a connected open set arising from a chart}\}$$

is an open cover of I, this covering has a Lebesgue number $\varepsilon > 0$ (cf. [J. Dugundji, 1966], p. 234). Choose an integer $k > 1/\varepsilon$, and set $t_i = i/k$, $0 \le i \le k$, so that $0 = t_0 < t_1 < \ldots < t_k = 1$ is a partition of I and $\sigma([t_{i-1}, t_i]) \subset U_i$, where U_i is the open set of some chart (U_i, φ_i).

We may assume inductively that (U_i, φ_i) and (U_{i+1}, φ_{i+1}) are compatible at $\sigma(t_i)$ for $i < s$. If (U_s, φ_s) and (U_{s+1}, φ_{s+1}) are compatible at $\sigma(t_s)$, we have the induction step. Otherwise we replace φ_{s+1} by $\phi\varphi_{s+1}$, where $\phi: \mathbf{R}^n \to \mathbf{R}^n$ is any linear map of negative determinant so that (U_s, φ_s) and $(U_{s+1}, \phi\varphi_{s+1})$ are compatible at $\sigma(t_s)$. Renaming $\phi\varphi_{s+1}$ as φ_{s+1} completes the inductive step. ■

Lemma 1.15. *Let $\sigma: (I, 0, 1) \to (M, p, q)$ be a path on M, and let (U_i, φ_i), $1 \le i \le k$, and (V_j, ψ_j), $1 \le j \le l$, be two compatible families of charts along σ. Then (U_1, φ_1) is compatible with (V_1, ψ_1) at $p \Rightarrow (U_k, \varphi_k)$ is compatible with (V_l, ψ_l) at q.*

Proof. Let $0 = s_0 < s_1 < \ldots < s_k = 1$ and $0 = t_0 < t_1 < \ldots < t_l = 1$ be the partitions of I corresponding to the compatible families of charts (U_i, φ_i), $1 \le i \le k$, and (V_j, ψ_j), $1 \le j \le l$, respectively. Let $S_i = [s_{i-1}, s_i]$ and $T_j = [t_{j-1}, t_j]$. The proof is by induction, where the inductive step is

Suppose that $S_i \cap T_j \neq \emptyset$ and (U_i, φ_i) is compatible with (V_j, ψ_j) on $\sigma(S_i \cap T_j)$. Then, either

(i) $S_{i+1} \cap T_j \neq \emptyset$ and (U_{i+1}, φ_{i+1}) is compatible with (V_j, ψ_j) on $\sigma(S_{i+1} \cap T_j)$,

or

(ii) $S_i \cap T_{j+1} \neq \emptyset$ and (U_i, φ_i) is compatible with (V_{j+1}, ψ_{j+1}) on $\sigma(S_i \cap T_{j+1})$.

Since (U_1, φ_1) is compatible with (V_1, ψ_1) at p, it will follow inductively that (U_k, φ_k) is compatible with (V_l, ψ_l) at q.

Let us verify the inductive step. If $s_i \le t_j$, then $s_i \in S_i \cap S_{i+1} \cap T_j$, so $S_{i+1} \cap T_j \neq \emptyset$. Now (U_i, φ_i) and (U_{i+1}, φ_{i+1}) are compatible at $\sigma(s_i)$, so it follows from the inductive hypothesis that (U_{i+1}, φ_{i+1}) is compatible with (V_j, ψ_j) at $\sigma(s_i)$. Hence, by Remark 1.9, (U_{i+1}, φ_{i+1}) is compatible with (V_j, ψ_j) on $\sigma(S_{i+1} \cap T_j)$. This verifies condition (i). Similarly, in the case $t_j \le s_i$, we get condition (ii). ■

Now we are ready for the notion of an orientation-preserving loop, which is the key to the question of the orientability of a manifold.

Definition 1.16. Let $\lambda: (I, 0, 1) \to (M, p, p)$ be a loop on M. We say that λ is *orientation preserving* if there is a compatible family of charts (U_i, φ_i), $1 \le i \le k$, along λ such that (U_1, φ_1) is compatible with (U_k, φ_k) at p. ❀

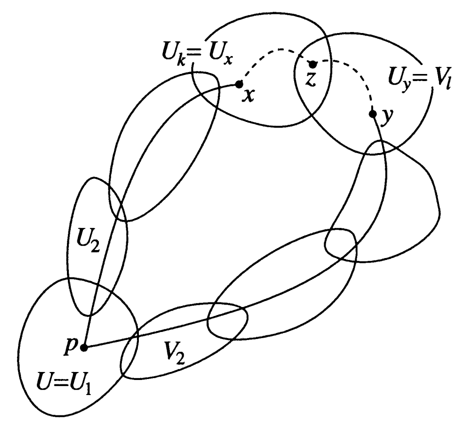
Note that, by Lemma 1.15, if a loop λ is orientation preserving, then *any* family of charts compatible along λ will have the property described in Definition 1.16. Now we are ready for our characterization of the orientability of manifolds in terms of loops.

Proposition 1.17. *Let M be a smooth manifold. Then*

$$M \text{ is orientable} \Leftrightarrow \text{every loop on } M \text{ is orientation preserving.}$$

Proof. $\Rightarrow$: Given a loop $\lambda: (I, 0, 1) \to (M, p, p)$ on M, we may select a family of charts, (U_i, φ_i), $1 \le i \le k$, from an oriented atlas for M as in the proof of Lemma 1.14. These charts will automatically be compatible since the atlas is oriented. In particular, (U_1, φ_1) is compatible with (U_k, φ_k) at p, so the loop λ is orientation preserving.

$\Leftarrow$: It suffices to consider the case when M is connected. Fix a point $p \in M$, and choose a connected chart (U, φ) with $p \in U$. For every point $x \in M$, choose a path $\sigma: (I, 0, 1) \to (M, p, x)$ joining p to x, and choose a family of connected charts (U_i, φ_i), $1 \le i \le k$, compatible along σ, starting with $(U_1, \varphi_1) = (U, \varphi)$. Set $(U_x, \varphi_x) = (U_k, \varphi_k)$. Then $\{(U_x, \varphi_x) \mid x \in M\}$ is an atlas for M. We claim that it is oriented. It suffices to show that any two charts (U_x, φ_x) and (U_y, φ_y) are compatible at every point $z \in U_x \cap U_y$. Suppose that (U_i, φ_i), $1 \le i \le k$, is the compatible family of charts used to obtain (U_x, φ_x) and that (V_i, ψ_i), $1 \le i \le l$, is the compatible family of charts used to obtain (U_y, φ_y).



Since U_x and U_y are connected, we can join x to z in U_x and y to z in U_y. These paths, together with the paths from p to x and y, give a loop based at z. Since this loop is orientation preserving and the family of charts along it corresponding to the open sets $U_x = U_k, \ldots, U_2, U_1 = U = V_1, V_2, \ldots, V_l = U_y$ is compatible, it follows that (U_x, φ_x) and (U_y, φ_y) are compatible at $z \in U_x \cap U_y$. ■

Corollary 1.18. *Let M be a connected, smooth manifold and let $p \in M$. Then*

$$M \text{ is orientable} \Leftrightarrow \text{every loop on } M \text{ based at } p \text{ is orientation preserving.}$$

Proof. In the proof of the proposition, the only loops we needed were ones that passed through the fixed point p. Thus it suffices to show that if, for a given orientation-preserving loop, we change the point on it that we regard as the base point, then the new loop is still orientation preserving. But this is clear. ■

Examples of Smooth Manifolds

The fundamental example of a smooth n-manifold is of course $\mathbf{R}^n$ itself with its atlas consisting of the identity map alone. More generally, any finite dimensional real vector space V carries a canonical smooth structure in the following manner. If $\dim(V) = n$, we take the atlas consisting of all linear isomorphisms $\varphi: V \to \mathbf{R}^n$. The collection of such maps is an atlas since for any two, φ and ψ, the change of coordinates is a *linear* map $\psi\varphi^{-1}: \mathbf{R}^n \to \mathbf{R}^n$ and hence smooth. Although we do not prove it here, it turns out that the canonical structure on V is, up to diffeomorphism (see Definition 1.20 ahead), the *unique* smooth structure except in the case with $\dim(V) = 4$. In the latter case it turns out that there are *infinitely* many smooth structures on V. (Cf. [D. Freed and K. Uhlenbeck, 1984], pp. 17–19, and [R. Kirby, 1989].)

We shall deal only with the canonical smooth structures on vector spaces. In particular, if V and W are finite-dimensional real vector spaces, then $\mathrm{Hom}(V, W)$ is a vector space whose dimension is $\dim(V) \cdot \dim(W)$ and hence has a canonical smooth structure. The special case when $V = W = \mathbf{R}^n$ is of particular interest to us. It is the space of real $n \times n$ matrices, which we denote by $M_n(\mathbf{R})$.

Example 1.19. Let M be an *open subset of* $\mathbf{R}^n$ (or more generally an open set in any finite-dimensional vector space). Then the inclusion map $M \subset \mathbf{R}^n$ is an atlas with one chart that provides a smooth structure on M. In particular,

$$Gl_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid \det(A) \neq 0\},$$

is an open set in the vector space $M_n(\mathbf{R})$ since $\det: M_n(\mathbf{R}) \to \mathbf{R}$ is a continuous map. (In fact, it is a polynomial map.) Thus $Gl_n(\mathbf{R})$ is, canonically, a smooth manifold. We can say even more. Both of the maps

$$\mu: Gl_n(\mathbf{R}) \times Gl_n(\mathbf{R}) \to Gl_n(\mathbf{R}) \quad (\textit{matrix multiplication}),$$

$$\iota: Gl_n(\mathbf{R}) \to Gl_n(\mathbf{R}) \quad (\textit{inversion})$$

are continuous (and even rational) maps. Thus $Gl_n(\mathbf{R})$ is an example of a *topological group*. ◆

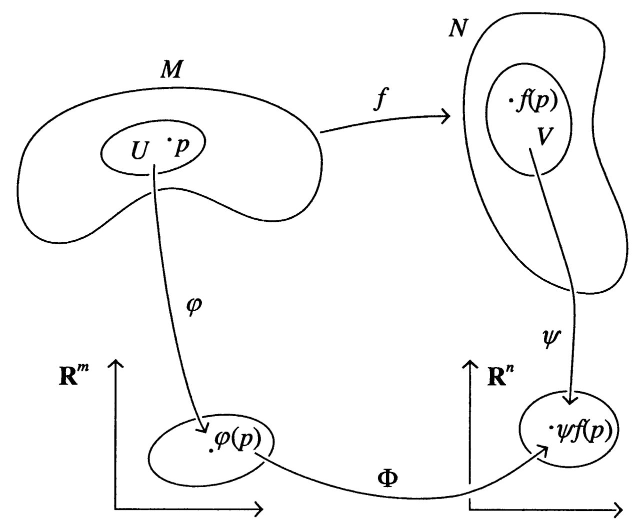
Now consider two smooth manifolds M^m and N^n. We can form their Cartesian product $M \times N$. If $(p, q) \in M \times N$, we may choose charts (U, φ) and (V, ψ) around p and q, respectively. Then $U \times V$ is open in $M \times N$, and $(U \times V, \varphi \times \psi)$ is a chart around (p, q). It is easy to verify that the collection of all such charts is an atlas for $M \times N$ and so determines a smooth structure on it. For example, the canonical smooth structure on $\mathbf{R}^{m+n}$ is the product of the canonical structures on the factors $\mathbf{R}^m \times \mathbf{R}^n$.

Two great questions in the study of smooth manifolds, which were largely answered in the 1960s, were "Does a given topological manifold necessarily have a smooth structure?" and "If a topological manifold has a smooth structure, is it unique?" The answers to both of these questions is generally *no*. For example, it is known that the spheres of dimension ≤ 6 have unique smooth structures, but there are 28 distinct smooth structures on the 7-sphere. For $n > 7$, there is generally more than one smooth structure on the n-sphere (cf. [A. Kosinski, 1993]).

Smooth Maps

Not only does a smooth structure allow us unambiguously to differentiate the composites Φ appearing in the definition of atlases, it also allows us to differentiate certain maps between manifolds. It will take us a while to see how this may be done, and the full story will only appear in §4. Here we prepare the way by describing the functions that we will eventually differentiate.

Definition 1.20. A map $f: M \to N$ between smooth manifolds is called *smooth* (or C^∞) if it is continuous and for each point $p \in M$ there is a chart (U, φ) on M with $p \in U$ and a chart (V, ψ) on M with $f(p) \in V$ such that the composite $\Phi = \psi f \varphi^{-1}$ is smooth. Φ is called the *coordinate expression* for f. The map f is called a *diffeomorphism* if it is smooth and bijective with a smooth inverse. ❀



It is clear that the smoothness of a map is independent of the choice of charts on M and N. What is more, we even have the notion of the rank of a smooth map at a point.

Definition 1.21. The *rank* of a smooth map $f: M \to N$ at a point $p \in M$, denoted by $\operatorname{rank}_p(f)$, is the rank of the Jacobian matrix

$$\Phi'(\varphi(p)) = (\psi f \varphi^{-1})'(\varphi(p)),$$

where φ and ψ are charts containing p and $f(p)$, respectively. ❀

Varying the choice of the charts merely left and right composes Φ with local diffeomorphisms of Euclidean space which, by the chain rule, left and right multiplies $\Phi'(\varphi(p))$ by invertible matrices, and this will not change the rank.

Differentiable topology is the study of the properties of smooth manifolds that are preserved by diffeomorphism.

Here is a criterion for a smooth homeomorphism to be a diffeomorphism.

Theorem 1.22. *Let $f: M \to N$ be a smooth bijection between smooth manifolds on the same dimension m. Then f is a diffeomorphism if and only if its rank at each point of M is m.*

Proof. Both the condition on the rank and the smoothness are local conditions, so it suffices to consider the case when M and N are open subsets of Euclidean space. Let $g: N \to M$ be the inverse function for f. Now if g is smooth, then applying the chain rule to $g(f(x)) = x$ yields $g'(f(x))f'(x) = I$ and hence $\det(g'(f(x)))\det(f'(x)) = 1$, so that $\det(f'(x)) \neq 0$ and thus $\operatorname{rank}_p f = m$. Conversely, if $\operatorname{rank}_p f = m$, then $\det(f'(p)) \neq 0$ and the inverse function theorem says that there is a unique local inverse h for f satisfying $h(f(p)) = p$ and that h is smooth. The uniqueness tells us that $g = h$ on their common domain, so that $g = f^{-1}$ is smooth. ■

Lie Groups

We now introduce some of the main players in the study of differential geometry; these are the Lie groups studied by Sophus Lie, Felix Klein, and Wilhelm Killing and developed by Elie Cartan. They constitute the basic symmetry groups of differential geometry, and of nature too.

Definition 1.23. A *Lie group* is a group G that is also a smooth manifold in a way that is compatible with the group structure in the sense that the maps

(i) (multiplication) $\mu: G \times G \to G$,

(ii) (inversion) $\iota: G \to G$

are both smooth. ❀

Example 1.24. Any finite-dimensional real vector space V is a smooth manifold in a canonical fashion and is an abelian group under vector addition. Since addition and subtraction of vectors are smooth, V is a Lie group. ◆

Example 1.25. We have already seen that $Gl_n(\mathbf{R}) = \{x \in M_n(\mathbf{R}) \mid \det(x) \neq 0\}$ is a topological group and that multiplication and inversion are the restriction of rational functions $M_n(\mathbf{R}) \times M_n(\mathbf{R}) \to M_n(\mathbf{R})$ and $M_n(\mathbf{R}) \to M_n(\mathbf{R})$. Hence they are smooth and $Gl_n(\mathbf{R})$ is a Lie group. ◆

More examples of Lie groups may be found at the end of this chapter, on page 63.

Definition 1.26. A *homomorphism* between Lie groups G and H is a smooth map $\varphi: G \to H$, which is also a homomorphism in the sense of group theory. ❀

Example 1.27. The exponential mapping $\exp: (\mathbf{R}, +) \to (\mathbf{R}^+, \times)$ is an isomorphism of Lie groups (i.e., it and its inverse are both homomorphisms). Although he did not have the larger concept of Lie groups, it was the existence of this isomorphism that made possible Napier's (1614) construction of his table of logarithms turning multiplication into addition. ◆

Smooth Maps of Constant Rank

Now we study the most important species of smooth maps, those of constant rank. For clarity, we begin with three local results that will soon be rephrased in terms of smooth manifolds.

Lemma 1.28. *Let $f: (\mathbf{R}^{r+m}, 0) \to (\mathbf{R}^r, 0)$ be a smooth map defined on a neighborhood of 0 which satisfies $f'(0) = (I, 0)$.*[6] *Then there is a diffeomorphism $J: (\mathbf{R}^{r+m}, 0) \to (\mathbf{R}^{r+m}, 0)$ defined on a neighborhood of 0 such that $f \circ J$ is a restriction of the canonical first factor projection mapping*

$$\begin{array}{rcl} \pi: \mathbf{R}^r \times \mathbf{R}^m & \to & \mathbf{R}^r. \\ (x,y) & \mapsto & x \end{array}$$

Proof. Define

$$\begin{array}{rcl} G: \mathbf{R}^r \times \mathbf{R}^m & \to & \mathbf{R}^r \times \mathbf{R}^m. \\ (x,y) & \mapsto & (f(x,y),y) \end{array}$$

Now G is smooth and $G'(0,0) = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$. Thus G is a local diffeomorphism of $(\mathbf{R}^r \times \mathbf{R}^m, 0)$ with inverse J, say. Since $\pi \circ G(x,y) = f(x,y)$, we have

$$f \circ J = \pi \circ G \circ J = \pi. \quad ■$$

[6]So, in particular, f has rank r at the origin.

Now we apply this lemma to study maps of constant rank.

Proposition 1.29. *Let $f\colon(\mathbf{R}^{r+m},0)\to(\mathbf{R}^{r+n},0)$ be a smooth map of constant rank r defined on a neighborhood of 0 that satisfies $f'(0)=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}$. Then there are diffeomorphisms, also defined on a neighborhood of 0, of the form $H\colon(\mathbf{R}^{r+n},0)\to(\mathbf{R}^{r+n},0)$ and $J\colon(\mathbf{R}^{r+m},0)\to(\mathbf{R}^{r+m},0)$ such that, on some neighborhood of 0, $H\circ f\circ J$ is a restriction of the canonical map*

$$\begin{array}{rcl}\pi\colon \mathbf{R}^r\times\mathbf{R}^m & \to & \mathbf{R}^r\times\mathbf{R}^n.\\ (x,y) & \mapsto & (x,0)\end{array}$$

Proof. Write f as $f=(f_1,f_2)\colon\mathbf{R}^{r+m}\to\mathbf{R}^r\times\mathbf{R}^n$, that is, f_1 is the first r components of f and f_2 is the last n components. Then $f_1'(0)=(I,0)$, so by Lemma 1.28 there is a local diffeomorphism of the form $J\colon(\mathbf{R}^{r+m},0)\to(\mathbf{R}^{r+m},0)$ defined on a neighborhood of 0 such that

$$(\mathbf{R}^r\times\mathbf{R}^m,0)\xrightarrow{f_1\circ J}(\mathbf{R}^r,0)$$

is a restriction of the canonical projection. Therefore, replacing f by $f\circ J$, we may assume that $f(x,y)=(x,f_2(x,y))$. It follows that

$$f'(x,y)=\begin{pmatrix} I_r & 0\\ \star & \left(\frac{\partial f_2}{\partial y}\right)\end{pmatrix}.$$

Now the condition that f has constant rank r, which is still true for our new f, means that the Jacobian matrix $(\partial f_2/\partial y)$ vanishes so that $f_2(x,y)$ is independent of the variable y. Thus, we may write $f_2(x,y)=h(x)$, say. Define $H\colon(\mathbf{R}^{r+n},0)\to(\mathbf{R}^{r+n},0)$ by $H(x,y)=(x,y-h(x))$. Then H is a local diffeomorphism and

$$Hf(x,y)=H(x,h(x))=(x,0). \qquad ■$$

Exercise 1.30. Show that in general, Proposition 1.28 becomes false if we insist that either H or J must be the identity. ❑

Now we put Proposition 1.29 in terms of an *arbitrary* smooth map of constant rank.

Theorem 1.31. *Let $f\colon M^m\to N^n$ be a smooth map with constant rank r. For each point $p\in M$, there are charts (U,φ) and (V,ψ) around p and $f(p)$, respectively, such that $\varphi(p)=0$, $\psi(f(p))=0$, $f(U)\subset V$, U and V are connected, and $\psi f\varphi^{-1}$ is a restriction of the canonical map*

$$\begin{array}{rcl}\mathbf{R}^r\times\mathbf{R}^{m-r} & \to & \mathbf{R}^r\times\mathbf{R}^{n-r}.\\ (x,y) & \mapsto & (x,0)\end{array}$$

Proof. Choose arbitrary charts (U,φ) and (V,ψ) around p and $f(p)$. After translation we may assume that $\varphi(p)=0$ and $\psi f(p))=0$. Then $\psi f\varphi^{-1}\colon(\mathbf{R}^m,0)\to(\mathbf{R}^n,0)$ is defined on a neighborhood of 0, and since the $n\times m$ matrix $(\psi f\varphi^{-1})'(0)$ has rank r, we may choose matrices $A\in Gl_n(\mathbf{R})$ and $B\in Gl_m(\mathbf{R})$ such that

$$(A\psi f(B\varphi)^{-1})_0'=A(\psi f\varphi^{-1})_0'B^{-1}=\begin{pmatrix} I_r & 0\\ 0 & 0\end{pmatrix}.$$

Replacing ψ and φ by $A\psi$ and $B\varphi$, respectively, we may assume that ψ and φ satisfy the hypotheses of Proposition 1.29. Thus there are diffeomorphisms

$$H\colon(\mathbf{R}^n,0)\to(\mathbf{R}^n,0)\quad\text{and}\quad J\colon(\mathbf{R}^m,0)\to(\mathbf{R}^m,0)$$

defined on a neighborhood of 0, such that $H\circ\psi\circ f\circ\varphi^{-1}\circ J\colon(\mathbf{R}^m,0)\to(\mathbf{R}^n,0)$ is a restriction of

$$\begin{array}{rcl}\mathbf{R}^r\times\mathbf{R}^{m-r} & \to & \mathbf{R}^r\times\mathbf{R}^{n-r}.\\ (x,y) & \mapsto & (x,0)\end{array}$$

Replacing ψ by $H\psi$ and φ by $J^{-1}\varphi$ finishes the proof, except for the conditions on U and V, which we leave to the reader. ■

Now let us mention the following two *extreme cases* of maps of constant rank.

Definition 1.32. Let $f\colon M^m\to N^n$ be a smooth map with constant rank. Then f is called an *immersion* if the rank is m, and a *submersion* if the rank is n. ❀

We apply Theorem 1.31 to these special cases to obtain the following result.

Corollary 1.33. *Let $f\colon M^m\to N^n$ be a smooth map. Then,*

(i) *f is an immersion*

$$\Leftrightarrow\left\{\begin{array}{l}\text{for each point } p\in M \text{ there are coordinate systems } (U,\varphi),\ (V,\psi)\\ \text{about } p \text{ and } f(p)\text{, respectively, such that the composite } \psi f\varphi^{-1}\\ \text{is a restriction of the coordinate inclusion } \iota\colon\mathbf{R}^m\to\mathbf{R}^m\times\mathbf{R}^{n-m},\end{array}\right.$$

(ii) *f is an submersion*

$$\Leftrightarrow\left\{\begin{array}{l}\text{for each point } p\in M \text{ there are coordinate systems } (U,\varphi),(V,\psi)\\ \text{about } p \text{ and } f(p)\text{, respectively, such that the composite } \psi f\varphi^{-1}\\ \text{is a restriction of the coordinate projection } \pi\colon\mathbf{R}^n\times\mathbf{R}^{m-n}\to\mathbf{R}^n.\end{array}\right.$$

Proof. In each case, $\Leftarrow$ is obvious, and $\Rightarrow$ comes from Theorem 1.31. ■

Exercise 1.34. Let M^m be a smooth manifold, and let $\iota: M \to \mathbf{R}^n$ be a smooth injective map. Show that ι is an immersion if and only if, for each point $p \in M$, there is an m-dimensional *coordinate* subspace $V \subset \mathbf{R}^n$ such that the composite with the orthogonal projection onto V, $\pi\iota: M \to V$, has rank m at p. ❑

Exercise 1.35. Show that in Corollary 1.33 the equivalences are still true if in (i) we prescribe the coordinate system (U, φ) and in (ii) we prescribe the coordinate system (V, φ). ❑

Proper Maps, Embeddings, and Weak Embeddings

We conclude this section with a discussion of some variations on the theme of well-behaved mappings. We begin with a discussion of the important notion of proper mappings.

Definition 1.36. Let $f: X \to Y$ be a continuous map with X, Y Hausdorff. Then f is called *proper* if $f^{-1}(K)$ is compact for every compact $K \subset Y$. (Note that if X is compact, then f is automatically proper.) ❀

Proposition 1.37. *Let X and Y be topological spaces that are Hausdorff and first countable. Let $f: X \to Y$ be a continuous proper injection. Then $f: X \to f(X)$ is a homeomorphism (where the topology on $f(X)$ is the subspace topology), and $f(X)$ is a closed subset of Y.*

Proof. Since $f: X \to f(X)$ is a continuous bijection, we need only see that it maps open sets to open sets, or equivalently, that it maps closed sets to closed sets. Let C be closed in X. Suppose that y lies in the closure of $f(C)$ in Y. Since Y is first countable, there is a sequence of points $x_1, x_2, \ldots \in C$ such that $f(x_j) = y_j \to y$. Now the set $K = \{y, y_1, y_2, \ldots\}$ is compact. Because f is proper, $f^{-1}(K)$ is also compact. Since X is first countable, the sequence $x_1, x_2, \ldots$ lying in $f^{-1}(K) \cap C$ has a convergent subsequence, converging to $x \in C$, say. Since Y is Hausdorff, the corresponding subsequence of the $y_1, y_2, \ldots$ must then converge to $f(x)$, and hence $y = f(x) \in f(C)$. Thus $f(C)$ is closed in Y and f maps closed sets to closed sets. In particular, taking $C = X$ shows $f(X)$ is a closed subset of Y. ■

Definition 1.38. An *embedding* is a one-to-one immersion $f: M \to N$ such that the mapping $f: M \to f(M)$ is a homeomorphism (where the topology on $f(M)$ is the subspace topology inherited from N). ❀

Proposition 1.39. *A proper one-to-one immersion is an embedding.*

Proof. This is a simple consequence of Proposition 1.37. ■

Definition 1.40. A *weak embedding* is a one-to-one immersion $f: M \to N$ such that, for every smooth map $g: S \to N$ with $g(S) \subset f(M)$, the induced map $\tilde{g}: S \to M$, defined by $g = f \circ \tilde{g}$, is smooth. ❀

Lemma 1.41. *If $f: M \to N$ is a weak embedding, there is a unique smooth structure on M for which this is true.*

Proof. Let M_1 be M with a possibly distinct smooth structure such that the map $f_1: M_1 \to N$ (which is the same as f) is also a weak embedding. Then the induced map $id: M_1 \to M$ is smooth. Similarly, the induced map $id: M \to M_1$ is smooth, so the two smooth structures are identical. ■

Exercise 1.42. Let $A \in Gl_{n+1}(\mathbf{R})$, and consider the mapping $\phi_A: P^n(\mathbf{R}) \to P^n(\mathbf{R})$ defined by $\phi_A(l) = Al$.

(a) Let an affine coordinate system on $P^n(\mathbf{R})$ arise from the affine map $f: \mathbf{R}^n \to \mathbf{R}^{n+1}$ given by $f(x_1, \ldots, x_n) = (1, x_1, \ldots, x_n)$. Show that in this coordinate system

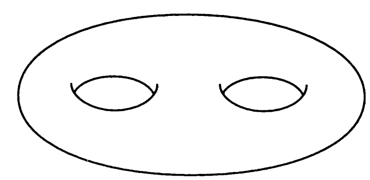
$$\phi_A(x) = \frac{ax + b}{cx + d}, \quad \text{where } A = \begin{pmatrix} d & c \\ b & a \end{pmatrix}, \text{ with } d \text{ a } 1 \times 1 \text{ block}$$

and a an $n \times n$ block.

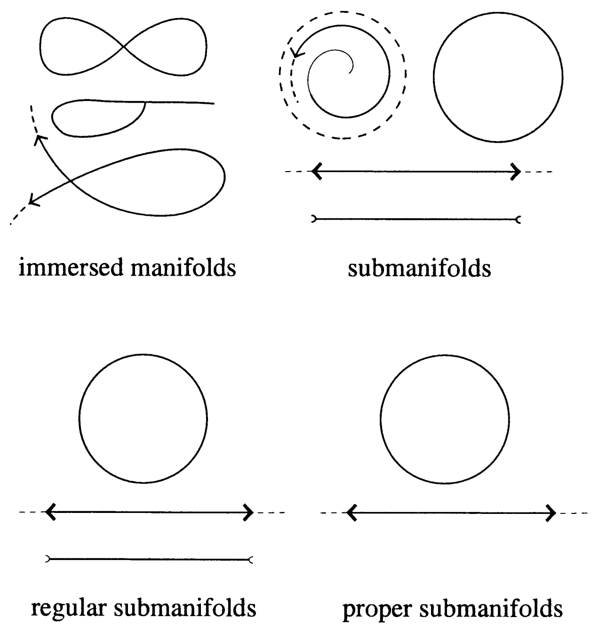
(b) Show that ϕ_A is a diffeomorphism. ❑

§2. Submanifolds

Originally, manifolds were regarded as subsets of Euclidean space; certainly these are the easiest ones to visualize.



In this section we study how one manifold may be situated inside another one. We give various definitions of increasing simplicity corresponding to the following picture.



Whereas all of these pictures intuitively represent one-dimensional manifolds in the plane, the first two kinds are not generally manifolds in the subspace topology. The first, an immersed submanifold, has one or more "bad points" (the *double point* in this picture) or odd convergent sequences in the subspace topology; the second can also have odd convergent sequences.

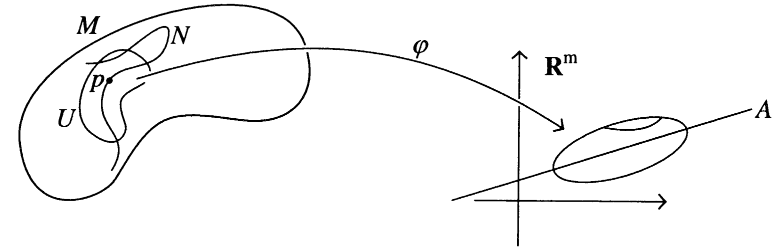
We study a general notion of submanifold which includes all but the immersed manifolds. The latter will be mentioned again at the end of this section.

Let M^m be a smooth manifold and let N^n be a subset. The idea is that N will be a submanifold if there are charts (U, φ) on M which, locally, *straighten out* N.

Definition 2.1. Let (U, φ) be a chart on M. The components of $N \cap U$ are called the *plaques* of N in this chart. The chart (U, φ) on M is said to *straighten out* a plaque W if φ restricts to a homeomorphism between W and an open set in some n-dimensional affine subspace[7] $A \subset \mathbf{R}^m$. In this case, the plaque is called a *flat* plaque of dimension n and the restriction $\varphi \mid W\colon W \to A$ is called the *plaque chart.* If N is covered by flat plaques, we say that N is *locally flat* in M. ❀

The picture corresponding to Definition 2.1 is the following one.

[7]An affine subspace of $\mathbf{R}^m$ is a set of the form $\{v + a \in \mathbf{R}^m \mid v \in V\}$, where V is a vector subspace of $\mathbf{R}^m$ and $a \in \mathbf{R}^m$.



Note that in this picture N meets U in two plaques, but here only one gets straightened out by this coordinate system. There is no limitation on the number of plaques in a chart; it may even be infinite.

Definition 2.2. Let M^m be a smooth manifold. A subset N of M is called a *smooth* (n-*dimensional*) *submanifold* if there is a covering $\{U_\alpha\}$ of $\bar{N}$ by open sets of M such that the components of $U_\alpha \cap N$ are all flat plaques of dimension n.[8] ❀

Of course we may always "tidy up" φ so that the affine subspace A is a coordinate subspace, say the coordinate subspace corresponding to the first n coordinates, and so that $\varphi(p) = 0$; but this is not always appropriate. The coordinate charts guaranteed by the definition tell us that "locally" the pair (M, N) looks like the pair $(\mathbf{R}^m, \mathbf{R}^n)$. This eliminates some of the phenomena of strange convergent sequences and the phenomena of limit points of nontrivial topology such as appear in Figure 2.3.

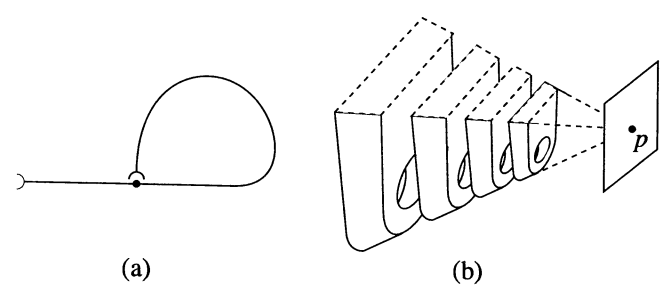


FIGURE 2.3. Odd limit points.

[8]This definition and the lemma and theorem that follow were inspired by a discussion in [P. Molino, 1988], pp. 11–12. We note, however, that the property that *every* plaque in $U \cap N$ is flat is not easily verified unless N has some other special property allowing us to assume that $U \cap N$ has just one plaque. The aim is to have a definition broad enough to include nonclosed subgroups of a Lie group and, more generally, the leaves of a foliation (cf. Chapter 2) for which this "one plaque" property may fail.

Figure 2.3(a) is an example of what is called an *immersed submanifold*,[9] which is the image of a manifold under a one-to-one immersion.

We are going to define a topology and a smooth structure on a submanifold N. The topology is called the submanifold topology; it generally has more open sets than the induced topology. Note that each plaque of a submanifold inherits the induced topology from M, and in this topology the flat plaques are homeomorphic to open sets in $\mathbf{R}^n$.

Definition 2.4. The *submanifold topology* on N is the one for which a set $U \subset N$ is open if it meets every flat plaque in an open set (in the induced topology on the plaque). ❀

Exercise 2.5. Suppose that N is a submanifold of the smooth manifold M. Show that if $\{W_\alpha\}$ is a given covering of N by flat plaques, then $U \subset N$ is open if and only if $U \cap W_\alpha$ is open for every index α. ❑

To study the submanifold topology, we shall use the following observation.

Lemma 2.6. *Suppose that N is a submanifold of the smooth manifold M. If $W \subset N$ is a flat plaque and $V \subset M$ is an open set, then $W \cap V$ is a union of at most countably many flat plaques.*

Proof. Let (U, φ) be a chart on M that straightens out W. This means that $\varphi(W) \subset \mathbf{R}^n$ is an open subset of an n-dimensional affine subspace A. Write $W \cap V = \cup W_\alpha$, where the W_α are the path components of $W \cap V$. Then each W_α is an open set in $W \cap V$ and hence in W as well. Hence each $\varphi(W_\alpha) \subset \mathbf{R}^n$ is open in A. Thus, W_α is a flat plaque straightened out by the chart $(U \cap V,\ \varphi \mid U \cap V)$ on M. Moreover, there can be at most countably many of the plaques W_α; otherwise the open set $\bigcup_\alpha \varphi(W_\alpha) \subset A$ would have uncountably many components, contradicting the fact that A, and therefore all its subspaces, have countable bases for their topologies. ■

Theorem 2.7. *Let M be a smooth manifold and N a submanifold equipped with the submanifold topology. Then*

(i) *N is a topological manifold,*

(ii) *the plaque charts provide a smooth atlas and hence a smooth structure for N,*

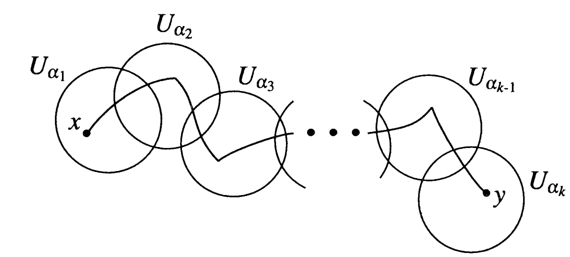
[9] Suppose $f\colon N \to M$ is a one-to-one immersion. It would clearly be a difficult matter, and perhaps even impossible, to reconstruct the manifold N and the immersion f merely from the immersed submanifold itself, that is, from the image set $f(N) \subset M$. The definition we give for submanifold is intermediate between the notion of an immersed submanifold and that of a regular submanifold.

(iii) *the inclusion map $N \subset M$ is a weak embedding.*

Proof. (i) To see that the inclusion $N \subset M$ is continuous, it suffices to show that any open set $V \subset M$ meets any plaque of N in an open subset of that plaque. But since the topology on a plaque is just the induced topology as a subset of M, this is clear. Since $N \subset M$ is continuous and M is Hausdorff, it follows that N is Hausdorff. The flat plaque charts show that N is locally Euclidean.

It remains to see that N is paracompact in the submanifold topology. For this it suffices to show that each component of N is a countable union of flat plaques. Without loss of generality, we may assume N is path connected. Since M is a paracompact Hausdorff space, so is $\bar{N}$, and we may choose a locally finite refinement of the open cover $\{U_\alpha\}_{\alpha\in I}$ of $\bar{N}$ (see Definition 2.2). Calling this locally finite refinement $\{U_\alpha\}_{\alpha\in I}$ again, we note that by Lemma 2.6 it still has the property that each component of $U_\alpha \cap N$ is a flat plaque.

Fix a base point $x \in N$. Now if $J = \{\alpha_1, \dots, \alpha_k\}$ is a sequence of elements of I, we let A_J denote the subset of N which may be joined to x by a piecewise smooth path on M that lies on N and passes successively through $U_{\alpha_1}, U_{\alpha_2}, \dots, U_{\alpha_k}$.



Of course, there may be no overlap between successive Us, in which case $A_J = \emptyset$.

First note that, because N is assumed path connected and the unit interval is compact, each point $y \in M$ may be joined to x by a path in A_J for *some* J; thus $\bigcup_J A_J = N$. Next note that $A_J \subset U_{\alpha_k} \cap N$ is a union of components of $U_{\alpha_k} \cap N$, each one of which is a flat plaque. If we can show that the number of components of A_J is at most countable, we will be finished since the local finiteness of the covering $\{U_\alpha\}_{\alpha\in I}$ implies that $\{J \mid \#J \le k \text{ and } A_J \ne \emptyset\}$ is finite and hence N is covered by at most countably many flat plaques.

We show by induction on $\#J$ that the number of components of A_J is at most countable. If $\#J = 1$, then clearly A_J has at most one component. Suppose that

$$J = \{\alpha_1, \dots, \alpha_k\}, \quad J' = \{\alpha_1, \dots, \alpha_{k-1}\},$$

and we already know that $A_{J'}$ has at most countably many components. By Lemma 2.6, each plaque of $A_{J'}$ meets at most countably many plaques in $U_{i_k} \cap N$. Thus, A_J has at most countably many components.

(ii) The smooth compatibility of the plaque charts is an immediate consequence of the smooth compatibility of the charts on M.

(iii) First note that the inclusion $N \subset M$ is obviously a one-to-one immersion. Now let $g: S \to M$ be a smooth map, with $g(S) \subset N$. To check that the induced map $f: S \to N$ is smooth, let $(W, \varphi \mid W)$ be a plaque chart. Then by definition,

$$f \mid f^{-1}(W) \text{ is smooth} \Leftrightarrow (\varphi \mid W)(f \mid f^{-1}(W)) \text{ is smooth.}$$

But $(\varphi \mid W) \circ (f \mid f^{-1}(W)) = (\varphi \mid W) \circ (g \mid g^{-1}(W)) = (\varphi g \mid g^{-1}(W))$ is smooth. ■

Regular Submanifolds

Now we pass to the definition of *regular* submanifolds, which are simpler and more important than the general submanifolds studied above. They intersect more neatly with coordinate charts of the ambient manifold; in particular, the various components of this intersection do not pile up. However, we pay for the simplicity of regular submanifolds by the exclusion of important examples, including the nonclosed subgroups of Lie groups and the leaves of some foliations.

Definition 2.8. An n-dimensional submanifold $N \subset M$ is *regular* if there is a covering $\{U_\alpha\}$ of N by open sets of M such that, for each α, $U_\alpha \cap N$ is a single flat plaque of dimension n. ❀

Proposition 2.9. *For a regular submanifold, the subspace topology is the same as the submanifold topology.*

Proof. Let $N \subset M$ be a regular submanifold. Since this inclusion map is continuous when N is given the submanifold topology, to show that the two topologies are the same, it suffices to show that given any subset $V \subset N$ open in the submanifold topology, there is an open set $V' \subset M$ with $N \cap V' = V$ so that V is also open in the subspace topology.

Since V is a union of flat plaques, it is sufficient to find V' in the case that V itself is a flat plaque. Choose a chart (U, φ) that straightens this plaque and contains no other plaques. We may assume $\varphi: (U, V) \to (\mathbf{R}^n \times \mathbf{R}^{m-n}, \mathbf{R}^n \times 0)$. Then $V' = \varphi^{-1}(\varphi(V) \times B_\varepsilon(0))$ will do the job, where $B_\varepsilon(0)$ is the open ε ball about the origin in $\mathbf{R}^{m-n}$. ■

Proper Submanifolds

Even more restrictive than regular submanifolds are the proper submanifolds.

Definition 2.10. A submanifold $N \subset M$ is called *proper* if it meets every compact subset of M in a compact subset of N (in the submanifold topology). Note that this is equivalent to saying that the inclusion map $N \subset M$ is proper. ❀

Theorem 2.11. *Let $N \subset M$ be a proper submanifold. Then M is regular.*

Proof. Fix $p \in N$. Take a chart (U, φ) for M, with $\bar{U}$ compact, that straightens out the plaque $W = U \cap N$ containing p. Now, as in the last paragraph of the proof of Proposition 2.9, we may assume that $\varphi(U) = \varphi(W) \times B_\varepsilon(0)$. We may choose a sequence of open sets $U_j = \varphi^{-1}(\varphi(W) \times B_{\varepsilon/j}(0))$, $j = 1, 2, \ldots$, so that $(U_j, \varphi \mid U_j)$ is also a chart straightening out the plaque W and $\bigcap_{j \geq 1} U_j = W$. We want to show that *some* U_i meets N in just the single plaque W, which will prove regularity. Suppose otherwise. Then we could find an *infinite* sequence of points $x_1, x_2, \ldots$ on $N - W$, with $x_i \in U_j \subset U$ each in a distinct plaque of U. Since $\bar{U}$ is compact, there is a convergent subsequence whose points, together with their limit, constitute a closed and therefore compact set $K \subset \bar{U} \cap N$. Repeating this argument with a slightly smaller plaque $W' \subset \bar{W}' \subset W$, we obtain $K \subset U \cap N$. But this is impossible since each point of K lies in a distinct plaque of U and these plaques constitute an open cover of K in the submanifold topology having no finite subcover, which contradicts the compactness of K. ■

Exercise 2.12. Show that neither the union nor the intersection of two proper submanifolds need be a submanifold. ❏

Submanifolds Described Implicitly

Here is a simple criterion for the smoothness of a submanifold described *implicitly.*

Theorem 2.13. *Let M^m and N^n be smooth manifolds with $m \geq n$, and let $f: M \to N$ be a smooth map. For $y \in N$, we set $M_y = f^{-1}(y)$. Assume that f has constant rank, r say, on a neighborhood of M_y. Then M_y is a smooth proper submanifold of M of dimension $m - r$.*

Proof. Let $p \in M_y$. Since the rank is constant on a neighborhood of p, by Theorem 1.31 there are coordinate charts (U, φ) and (V, ψ) around p and $f(p)$, respectively, such that $\varphi(p) = 0$, $\psi(f(p)) = 0$, and $\psi f \varphi^{-1}$ is a restriction of the canonical map

$$\begin{array}{rccc} \pi: & \mathbf{R}^r \times \mathbf{R}^{m-r} & \to & \mathbf{R}^r \times \mathbf{R}^{n-r}. \\ & (x,y) & \mapsto & (x,0) \end{array}$$

Thus $\varphi(U \cap M_y) \subset 0 \times \mathbf{R}^{m-r}$ in $\mathbf{R}^r \times \mathbf{R}^{m-r}$. Hence M_y has just one plaque in U and φ straightens out this plaque. This shows that M_y is a regular

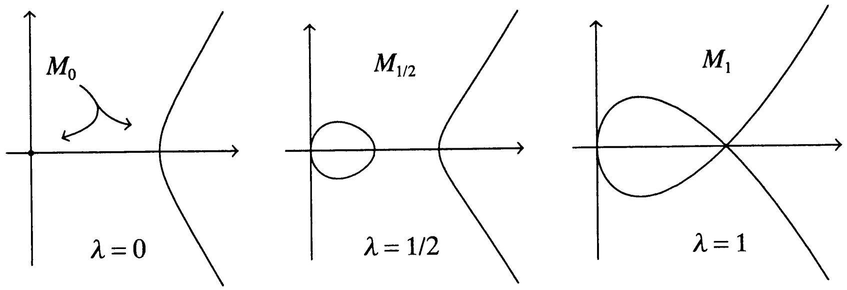


FIGURE 2.15. Various elliptic curves.

submanifold. Since M_y is closed and M is Hausdorff, it follows that $M_y \cap K$ is compact for every compact $K \subset M$. Thus M_y is proper. ■

Example 2.14. Let $M_\lambda = \{(x,y) \in \mathbf{R}^2 \mid y^2 = x(x-1)(x-\lambda)\}$, which is the Weierstrass normal form of the general cubic (elliptic) curve. λ is a parameter that we will take to be real. Thus M_λ is given implicitly as the zero set of the function $f\colon \mathbf{R}^2 \to \mathbf{R}$, where

$$f(x,y) = y^2 - x(x-1)(x-\lambda) = y^2 - x^3 + (\lambda+1)x^2 - \lambda x.$$

Calculating, we have $f'(x,y) = (-3x^2 + 2(\lambda+1)x - \lambda, 2y)$. The maximal possible rank for the derivative is 1, and we can find all points on M_λ where this fails by solving the system

$$\begin{aligned} x(x-1)(x-\lambda) &= y^2, \\ 3x^2 - 2(\lambda+1)x + 1 &= 0, \\ y &= 0. \end{aligned}$$

The only solutions are $(x,y,\lambda) = (0,0,0)$ and $(1,0,1)$, which shows that M_λ is a smooth submanifold unless $\lambda = 0$ or 1 and that, in these exceptional cases, $M_0 - \{(0,0)\}$ and $M_1 - \{(1,0)\}$ are also smooth submanifolds of $\mathbf{R}^2$. Figure 2.15 shows these singular points on M_0 and M_1, as well as one of the more typical cases of a nonsingular M_λ. ◆

Example 2.16. Let us reconsider the n-sphere

$$S^n = \{x \in \mathbf{R}^{n+1} \mid x \cdot x = 1\}.$$

Thus, $S^n = F^{-1}(\{1\})$, where $F\colon \mathbf{R}^{n+1} \to \mathbf{R}$ given by $F(x) = x \cdot x$. Calculating from first principles, $F'(x)v = 2x \cdot v$. Thus, F has rank 1 everywhere on S^n. It follows from Theorem 2.13 that S^n is a smooth submanifold of $\mathbf{R}^{n+1}$. ◆

Example 2.17. Let $O_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid AA^t = I\}$ = *the orthogonal group*. Define $F\colon M_n(\mathbf{R}) \to M_n(\mathbf{R})$ by $F(A) = AA^t$, so that $O_n(\mathbf{R}) = F^{-1}(I)$. Now F is smooth and we may calculate

$$\begin{aligned} F'(A)V &= \lim_{h\to 0} \frac{1}{h}\{F(A+hV) - F(A)\} \\ &= \lim_{h\to 0} \frac{1}{h}\{(A+hV)(A+hV)^t - AA^t\} \\ &= \lim_{h\to 0} \frac{1}{h}\{hVA^t + hAV^t + h^2VV^t\} \\ &= VA^t + AV^t. \end{aligned}$$

In the particular case that $A = I$, we have $\ker F'(I) = \mathfrak{o}_n(\mathbf{R})$, the vector space of all $n \times n$ skew-symmetric matrices. Note that $\dim \mathfrak{o}_n(\mathbf{R}) = \frac{1}{2}n(n-1)$. More generally we have an isomorphism $\mathfrak{o}_n(\mathbf{R}) \approx \ker F'(A)$, which sends $S \mapsto SA$ for any $A \in O_n(\mathbf{R})$. In particular, the rank of F is constant along $O_n(\mathbf{R})$, so $O_n(\mathbf{R})$ is a smooth submanifold of $M_n(\mathbf{R})$. ◆

Exercise 2.18. Note that the smooth map det: $O_n(\mathbf{R}) \to \mathbf{R}$ has values ± 1 (since $1 = \det(AA^t) = \det(A)\det(A^t) = \det(A)^2$), so it follows that $O_n(\mathbf{R})$ is the disjoint union of two open sets. Show that these are the components of $O_n(\mathbf{R})$. ❏

Exercise 2.19. Assume the situation of Theorem 2.12; let $W \subset N$ be a submanifold; and assume for each $x \in f^{-1}(W)$ that $f_*(T_xM) + T_{f(x)}W = T_{f(x)}N$. Show that $f^{-1}(W)$ is a submanifold of M. ❏

Submanifolds Described Parametrically

Now we pass to the dual situation of a manifold described *parametrically*.

Theorem 2.20. *Let M^m and N^n be smooth manifolds, and let $f\colon M \to N$ be a proper one-to-one immersion. Then $f(M)$ is a regular submanifold of N and the map $f\colon M \to f(M)$ is a diffeomorphism.*

Proof. We first show that $f(M)$ is a regular manifold. Choose $p \in M$, and set $q = f(p)$. Since f is an immersion, its rank is m everywhere. Then by Corollary 1.33 there are connected coordinate charts (U,φ) and (V,ψ) around p and q, respectively, such that $\varphi(p) = 0$, $\psi(q) = 0$, $f(U) \subset f(V)$, and $\psi f \varphi^{-1}$ is the restriction to $\varphi(U)$ of the canonical inclusion

$$\begin{aligned} \iota\colon \mathbf{R}^m &\to \mathbf{R}^m \times \mathbf{R}^{n-m}. \\ x &\mapsto (x,0) \end{aligned}$$

Thus $\psi(f(U))$ lies in an open subset of $\mathbf{R}^m \times 0 \subset \mathbf{R}^m \times \mathbf{R}^{n-m}$. Using the fact that f is proper, an argument similar to the proof of Theorem 2.11

allows us to shrink U and V so that $f(U) \cap V = f(M) \cap V$ is connected. Since (V, ψ) straightens out the unique plaque $f(M) \cap V$ in V, it follows that $f(M)$ is a regular submanifold.

By Theorem 2.7(iii), $f(M) \subset N$ is a weak embedding. Since $f: M \to N$ is smooth, it follows (cf. Definition 1.40) that the map $f: M \to f(M)$ is also smooth. Since both $f(M) \subset N$ and $f: M \to N$ are one-to-one immersions, it follows that $f: M \to f(M)$ is a bijective immersion. Then by Theorem 1.22, f is a diffeomorphism. ■

This theorem shows that the image of a proper embedding is a smooth submanifold. We remark that the embedding theorem of Whitney tells us that *every* smooth n-manifold admits a proper embedding as a smooth submanifold of Euclidean space of dimension $2n+1$ (cf. [J. Milnor, 1965]).

Immersed Manifolds and Immersed Submanifolds

Definition 2.21. An *immersed manifold* in M is the image of an immersion $f: N \to M$. (Note that the image is *not* in general a submanifold!) An *immersed submanifold* is the image of a one-to-one immersion $f: N \to M$. ❀

For example, the elliptic curve M_1 appearing in Example 2.14 may be regarded as an immersed manifold in $\mathbf{R}^2$. In such a case, it is still true that each point $p \in N$ has a neighborhood $U \subset N$ such that $f(U)$ is a submanifold of M diffeomorphic to U.

Exercise 2.22. Show that the curve $M_1 = \{(x, y) \in \mathbf{R}^2 \mid y^2 = x(x-1)^2\}$ is an immersed submanifold of $\mathbf{R}^2$ by parametrizing it by the slope of the line through $(x, y) \in M_1$ and $(1, 0)$. ❑

Exercise 2.23. Show that all of the following inclusions are proper inclusions:

$$\left\{\begin{matrix} \textit{immersed} \\ \textit{submanifolds} \end{matrix}\right\} \supset \left\{\begin{matrix} \textit{weakly} \\ \textit{embedded} \\ \textit{submanifolds} \end{matrix}\right\} \supset \{\textit{submanifolds}\}$$

$$\supset \left\{\begin{matrix} \textit{regular} \\ \textit{submanifolds} \end{matrix}\right\} \supset \left\{\begin{matrix} \textit{proper} \\ \textit{submanifolds} \end{matrix}\right\}.$$

(*Hint:* To show that the first and second inclusions are proper, consider Figure 2.3.) ❑

The following exercise provides a preview of the "second fundamental form" of a submanifold of $\mathbf{R}^n$.

Exercise 2.24.* Suppose that $M \subset \mathbf{R}^N$ is given implicitly in a neighborhood of $p \in M$ by the zero set of a function of maximal rank $F: \mathbf{R}^N \to \mathbf{R}^v$. The *tangent space* to M at x is defined to be the kernel of the linear surjection $F'(x): \mathbf{R}^N \to \mathbf{R}^v$ and is denoted $T_x(M)$. Define the *second fundamental form* of M at x by the formula

$$B_x: T_x(M) \times T_x(M) \to \mathbf{R}^N/T(M),$$
$$F'(x)(B_x(u, v)) = F''(x)(u, v).$$

It follows that the symmetric bilinear form $B_x(u, v) \in \mathbf{R}^N/T(M)$ is defined.

(i) Let $\sigma(t)$ be path on $M \subset \mathbf{R}^N$ through x and $v(t)$ a vector field tangent to M along σ. Show that at $x, \dot{v} = -B_x(\dot{\sigma}, v) \bmod T_x(M)$.

(ii) Use (i) to show that $B_x: T_x(M) \times T_x(M) \to \mathbf{R}^N/T_x(M)$ is independent of the choice of F.

(iii) Show that by using the Euclidean metric on $\mathbf{R}^N$ we can reinterpret B_x as a map $B_x: T_x(M) \times T_x(M) \to T_x(M)^\perp$ (orthogonal complement) and then define the "transpose"

$$B_x^t: T_x(M) \times T_x(M)^\perp \to T_x(M),$$
$$\langle u, B_x^t(v, w)\rangle = \langle B_x(u, v), w\rangle \quad \text{for all } u, v \in T_x(M), \ w \in T_x(M)^\perp.$$

Show that if $\sigma(t)$ is a path on $M \subset \mathbf{R}^N$ through x and $w(t)$ a vector field normal to M along σ, then at $x, \dot{w} = B_x^t(\dot{\sigma}, w) \bmod T_x(M)^\perp$.

(iv) Find an analog of (iii) that does not rely on using a Euclidean metric. ❑

In this formulation of the second fundamental form, it is clear that it has fairly strong invariance properties. The following exercises make this more precise.

Exercise 2.25. Suppose that $M \subset \mathbf{R}^N$ is given implicitly in a neighborhood of $p \in M$ as the zero set of a function of maximal rank $F: \mathbf{R}^N \to \mathbf{R}^v$, and let $\phi: \mathbf{R}^N \to \mathbf{R}^N$ be a diffeomorphism. Set $\tilde{F} = F \circ \phi$, and let B and $\tilde{B}$ be the second fundamental forms associated to $M = F^{-1}(0)$ and $\tilde{M} = \tilde{F}^{-1}(0)$, respectively. Show $\phi'_x(\tilde{B}(\tilde{u}, \tilde{v})) = B(u, v) + \phi''_x(\tilde{u}, \tilde{v}) \bmod T_{\phi(x)}(M)$ for all $\tilde{u}, \tilde{v} \in T_x(\tilde{M})$, where $u = \phi'_x(\tilde{u})$ and $v = \phi'_x(\tilde{v})$. ❑

Exercise 2.26. Let $M \subset \mathbf{R}^N$ and let $\phi: \mathbf{R}^N \to \mathbf{R}^N$ be any invertible affine map. Let B_x (respectively, $B_{\phi(x)}$) denote the second fundamental form of M at x (respectively, of $\phi(M)$ at $\phi(x)$). Show that

$$\phi'_x(B_x(u, v)) = B_{\phi(x)}(\phi'_x u, \phi'_x v) \bmod T_x(M) \quad \text{for all } u, v \in T_x(M).$$ ❑

Exercise 2.27. Assume the hypotheses of Exercise 2.26, and let $\rho: \mathbf{R}^N \to \mathbf{R}$ be a smooth function. Define $\varphi = \frac{1}{\rho}\phi: \mathbf{R}^N \to \mathbf{R}^N$, and show that

$$\varphi''_x(u, v) = -\frac{1}{\rho}\{\rho'_x(v)\phi'_x(u) + \rho'_x(u)\phi'_x(v)u - \phi''_x(u, v)\} \bmod \varphi(x)$$

$$\text{for all } u, v \in \mathbf{R}^N.$$

(Of course, we must avoid points where ρ vanishes.) Using the results of §4, deduce the analog of Exercise 2.26 for a submanifold $M \subset P^N(\mathbf{R})$ and an invertible projective transformation $\phi: P^n(\mathbf{R}) \to P^n(\mathbf{R})$. ❑

Exercise 2.28. Assume the hypotheses and notation of Exercises 2.25 and 2.27. Let $\rho(x) = x/(x \cdot x)$ (inversion). Show that, in this case, $\phi'_x(\tilde{B}_x(\tilde{u}, \tilde{v})) = B_{\phi(x)}(u, v) - 2(u \cdot v/\rho(x))\phi(x)$. ❑

Note that Exercise 2.28 shows that the "raw" second fundamental form fails to be invariant under inversion. Nevertheless, we can obtain an invariant second fundamental form by taking the "traceless" version of B, as indicated in the following exercise.

Exercise 2.29.* Let V be a Euclidean space (with inner product "$\cdot$") and W a vector space. Let $B: V \times V \to W$ be a symmetric form. Define Trace $B = \sum_{1 \le i \le \dim V} B(e_i, e_i) \in W$, where $\{e_i\}$ is an orthonormal basis for V. Define $B_0(u, v) = B(u, v) - (\text{Trace } B/\dim V) u \cdot v$.

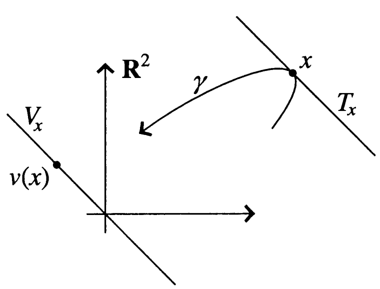
(a) Show that Trace B is independent of the choice of the basis.

(b) Assuming the hypotheses of Exercise 2.28, show that $\phi'_x(\tilde{B}_0(\tilde{u}, \tilde{v})) = B_0(u, v)$. ❑

We shall revisit some of the ideas of these exercises in Chapters 6 and 7.

§3. Fiber Bundles

Everyone knows the usefulness in advanced calculus of studying vector-valued functions. Consider, for example, the velocity field of a particle moving along a fixed curve γ in $\mathbf{R}^2$. This field is described by a function $v: \gamma \to \mathbf{R}^2$ defined along the curve. From another point of view, this same field may be regarded as a function whose value at $x \in \gamma$ lies in the one-dimensional subspace V_x of $\mathbf{R}^2$ that is parallel to the tangent line T_x to the curve at x.



We emphasize that $\{V_x \mid x \in \gamma\}$ is a *family* of subspaces parametrized by γ.

Fiber bundles formalize the notion of a "parametrized family of so-and-so's." The "so-and-so" is called the *fiber* and may be a vector space as in the example above (corresponding to a *vector bundle*) or a Lie group (corresponding to a *principal G bundle*). These two are the most important cases, although other fibers may also serve as necessary. *Sections* (or in the case of vector bundles, one often says *vector fields* or *tensors*), which generalize functions, are functions on the parameter space whose value at x lies in "the so-and-so at x." If the parametrized family is constant, then fields (sections) reduce to ordinary functions.

The justification for the degree of generality employed here will appear first in the next section, where it is shown that a smooth manifold has a *tangent bundle* intrinsically associated with it, and again in Chapter 5, where it is shown that a Cartan geometry has a *principal G bundle* intrinsically associated with it.

As with manifolds, it is possible to discuss both topological and smooth bundles. Of course, our principal interest is in the smooth case.

Topological Bundles and Smooth Bundles

Definition 3.1. Let F be a topological space and $\pi: E \to B$ a continuous map. We call the quadruple $\xi = (E, B, \pi, F)$ a (*locally trivial*) *fiber bundle with* (*abstract*) *fiber* F if, for each point $b \in B$, there is an open set $U \subset B$ containing p such that $\pi^{-1}(U)$ is homeomorphic to $U \times F$ by a homeomorphism φ such that the following diagram commutes.

$$\begin{tikzcd} \pi^{-1}(U) \arrow[rr, "\varphi"] \arrow[dr, "\pi"'] & & U \times F \arrow[dl, "\mathrm{proj}_1"] \\ & U & \end{tikzcd}$$

The pair (U, φ) is called a *chart* (or *local bundle coordinate system*). ❀

Note that $\pi^{-1}(b)$, the fiber over B, denoted F_b, is homeomorphic to F for all $b \in B$. B is called the *base space* and E the *total space* of the bundle. Given a bundle ξ, we sometimes denote by $B(\xi)$, $E(\xi)$, etc., the corresponding parts of the bundle. If the base, fiber, and total spaces are smooth manifolds, π is a smooth map, and the charts may always be chosen to be smooth maps, then we have a *smooth bundle*. Two bundles ξ_1 and ξ_2 over the same base B are said to be *isomorphic* if there is a homeomorphism (diffeomorphism) φ between the total spaces such that the following diagram commutes. φ is called a *bundle isomorphism* or, in the case when $\xi_1 = \xi_2$ a *bundle automorphism.*

$$\begin{tikzcd} E_2 \arrow[rr, "\varphi", "\approx"'] \arrow[dr, "\pi_1"'] & & E_1 \arrow[dl, "\pi_2"] \\ & B & \end{tikzcd}$$

Example 3.2. The simplest example of a bundle is the *product bundle* $E = B \times F$ with $\pi = \pi_1$ (projection on the first factor). This bundle is also called the *trivial bundle.* If E is a locally compact Hausdorff space, then a necessary and sufficient condition that a bundle $\xi = (E, B, \pi, F)$ be (isomorphic to) the trivial bundle is the existence of a *trivialization*, which is a continuous (smooth) map $t: E \to F$ that induces a homeomorphism (diffeomorphism) upon restriction to each fiber $\pi^{-1}(b)$. In this case the map $(p, t): E \to B \times F$ is a homeomorphism (diffeomorphism) commuting with the canonical projection to B. ◆

Example 3.3. Perhaps the simplest example of a nontrivial bundle is the *Möbius band* E pictured in Figure 3.4. It is clear that if we remove any point from the base B, we get an open set U homeomorphic to the interval $(0, 1)$ and $\pi^{-1}(U)$ is homeomorphic to $(0, 1) \times [0, 1]$. Thus it is a bundle over $B = S^1$ with fiber $I = [0, 1]$. But it is not a trivial bundle; if it were, the boundary would consist of two components, whereas in fact it has only one (Figure 3.4). ◆

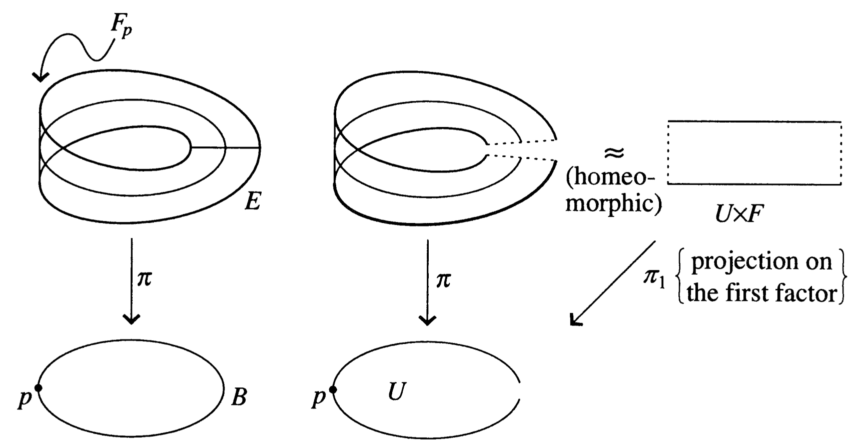


FIGURE 3.4. Möbius band.

Example 3.5. The *canonical line bundle* over projective space generalizes the Möbius band. This is the bundle $\pi: E \to P^n(\mathbf{R})$, where $E = \{(l, v) \in P^n(\mathbf{R}) \times \mathbf{R}^{n+1} \mid v \in l\}$ and $\pi(l, v) = l$. To see this is a bundle, let $a: \mathbf{R}^n \to \mathbf{R}^{n+1}$ be an affine map whose image does not contain the origin (cf. Example 1.3). Then $U_a = pa(\mathbf{R}^n)$ is open in $P^n(\mathbf{R})$ and $pa: \mathbf{R}^n \to U_a$ is a diffeomorphism. We define $\varphi_a^{-1}: U_a \times \mathbf{R} \to \pi^{-1}(U_a)$ by $(l, \lambda) \mapsto (l, \lambda a(pa)^{-1}(l))$. It is easily checked that φ_a^{-1} is a diffeomorphism that is linear on fibers, and that the pairs $\{(U_a, \varphi_a)\}$ constitute an atlas for the line bundle E.

In a similar fashion we may describe the canonical line bundle E over the projective space $P(V)$ associated to any vector space V. ◆

Example 3.6. By definition, every covering space $\pi: E \to B$ is a fiber bundle with discrete fiber. ◆

G Bundles

Now we consider the following, more refined notion of a bundle in which a *Lie group of symmetries* appears.

Definition 3.7. Let $\xi = (E, B, \pi, F)$ be a smooth fiber bundle, and suppose that G is a Lie group that acts smoothly on F as a group of diffeomorphisms. A G *atlas* for ξ is a collection $\mathcal{A} = \{(U_i, \varphi_i)\}$ of charts for ξ such that

(i) the U_i cover B,

(ii) for each pair of charts (U, φ) and (V, ψ) in $\mathcal{A}$ the map

$$\Phi = \psi\varphi^{-1}: (U \cap V) \times F \to (U \cap V) \times F,$$

called a *coordinate change*, has the form $\Phi(u, f) = (u, h(u)f)$, where $h: U \cap V \to G$ is a smooth map called a *transition function.* ❀

Note that in the case when the homomorphism $G \to \mathrm{Diff}(F)$ has a kernel H, then H is a closed[10] normal subgroup and the action factors through G/H, so the bundle may also be regarded as a G/H bundle.[11] Thus there is nothing lost if we make the restriction that G acts effectively on F, namely, that $H = 1$. In this case we may speak of an *effective* G bundle. Just as in the case of manifolds, we call two G atlases *equivalent* if their union is also a G atlas.

[10]Let $\Delta = \{(x, x) \in F \times F \mid x \in F\}$. Since the action is a smooth map $G \times F \to F$, the graph of this map, given by $\Gamma = \{(g, x, y) \in G \times F \times F \mid gx = y\}$, is closed and hence so is the intersection $\Gamma \cap G \times \Delta = H \times \Delta$.

[11]In Chapter 4 it will be shown that H and G/H are again Lie groups.

Definition 3.8. A G *structure* on the smooth bundle ξ is an equivalence class of G atlases on ξ, and a G *bundle* is a smooth bundle ξ with a specified G structure. ❀

Exercise 3.9 (Product bundles). Let $\xi_j = (E_j, B_j, \pi_j, F_j, G_j)$, $j = 1, 2$, be two "G" bundles. Show that $\xi_1 \times \xi_2 = (E_1 \times E_2,\ B_1 \times B_2,\ \pi_1 \times \pi_2,\ F_1 \times F_2,\ G_1 \times G_2)$ is a $G_1 \times G_2$ bundle in a canonical fashion. ❑

Exercise 3.10 (Induced, or pullback bundles). Let $\xi = (E, B, \pi, F, G)$ be a smooth G bundle and let $f\colon X \to B$ be a smooth map. Let $f^*(\xi) = (E_1, X, \pi_1, F, G)$, where $E_1 = \{(x, e) \in X \times E \mid f(x) = \pi(e)\}$ and $\pi_1(x, e) = x$. Show that it is a smooth G bundle. ❑

Definition 3.11. A G bundle is *flat* if all the transition functions $h\colon U \cap V \to G$ are constant. (This happens, in particular, if G is a discrete group.)[12] ❀

Exercise 3.12. Let B be a manifold with universal cover $\tilde{B}$. Show that a flat G-bundle over B with fiber F induces a trivial bundle over $\tilde{B}$ with a canonical trivialization and that the action of the group of covering transformations on the fiber is a representation $\pi_1(B, b) \to G$. Conversely, show that every representation $\pi_1(B, b) \to G$ arises in this way. ❑

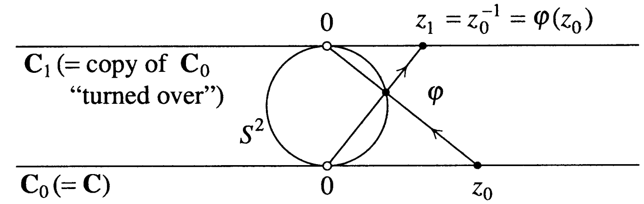
Given a G atlas, we can consider the union of all G atlases equivalent to it to obtain the unique maximal atlas equivalent to the given one. The

[12] The reader may wonder about the origin and importance of flat bundles in geometry. It often happens that one makes some locally defined geometrical construction on a manifold which depends on a choice in some discrete universe of possible choices. By the discreteness any choice can be extended smoothly over small neighborhoods, but globally one may not be able to do this. The appearance of this situation is a sure sign of the presence of a helpful flat bundle that resolves the global ambiguity. For example, at a nonumbilic point p on an oriented surface in 3-space, there are two principal tangent lines corresponding to the two principal curvatures. Although we may choose two unit vectors e_1 and e_2 in these two directions at p, there is no unique choice. Of course, we may assume that the basis (e_1, e_2) gives the correct orientation for the surface, but that still leaves the choice between $\pm(e_1, e_2)$. If we work only locally, we may just choose one of the two, say (e_1, e_2), and be done with it. It is a fact, however, that as we push our choice (e_1, e_2) along a loop enclosing a single (generic) umbilic point, it will come back to p as the other choice, $-(e_1, e_2)$. This indicates the presence of a flat bundle, in this case the flat line bundle L with discrete group ± 1, with the property that the fiber turns over every time we pass around a generic umbilic. Rather than regarding e_1 and e_2 as sections of the tangent bundle T of the surface, which would run into sign trouble as we pass around an umbilic, we regard them as sections of the tensor product $T \otimes L$. The rectification of the sign is built into L. And, moreover, rather than saying that we have now accommodated the frame by a clever construction, it is more true to say that the frame was *always* geometrically comprised of sections of $T \otimes L$ but we just didn't know it before.

G structure may be identified with this maximal atlas, and the remarks about tidying up the coordinate charts in the corresponding situation for smooth manifolds apply here as well. Two G bundles are *isomorphic* if they are isomorphic as bundles by an isomorphism identifying the G structures.

Example 3.13. A finite covering space $\pi\colon \tilde{M} \to M$ is a G bundle with discrete fiber whose group is the permutation group of the fiber. (If the fiber is not finite, the permutation group is not countable.) ◆

Example 3.14. We sketch a more complicated example of a G bundle, the Hopf fibration $S^1 \to S^3 \to S^2$. The Lie group $S^1 = \{\lambda \in \mathbf{C} \mid |\lambda| = 1\}$ acts smoothly on $S^3 = \{w = (w_0, w_1) \in \mathbf{C}^2 \mid |w_0|^2 + |w_1|^2 = 1\}$ by the formula $\lambda \cdot (w_0, w_1) = (\lambda w_0, \lambda w_1)$. We shall regard S^2 as the union of two copies of $\mathbf{C}$ with identifications, $S^2 = \mathbf{C}_0 \cup_\varphi \mathbf{C}_1$, where $\varphi\colon \mathbf{C}_0^* \to \mathbf{C}_1^*$ is given by $\varphi(z_0) = z_0^{-1}$. The map φ may be regarded as arising from two stereographic projections from the sphere of radius $1/2$ in $\mathbf{R}^3$ as in the following figure. (The complex planes are arranged so that their positive real axes are parallel.)



Now the map $\pi\colon S^3 \to S^2$ is given by

$$\pi(w) = \begin{cases} w_1/w_0 \in \mathbf{C}_1 & \text{if } w_0 \neq 0, \\ w_0/w_1 \in \mathbf{C}_0 & \text{if } w_1 \neq 0. \end{cases}$$

Note that if $w_0, w_1 \neq 0$, the two definitions agree. It is easy to verify that $\pi(u) = \pi(v) \Leftrightarrow u = \lambda v$ for some $\lambda \in S^1$. The trivializations are

$$\pi^{-1}(\mathbf{C}_0) \approx \mathbf{C}_0 \times S^1, \qquad \pi^{-1}(\mathbf{C}_1) \approx \mathbf{C}_1 \times S^1,$$

$$(w_0, w_1) \mapsto (w_1/w_0, w_0/|w_0|), \qquad (w_0, w_1) \mapsto (w_0/w_1, w_1/|w_1|),$$

and the coordinate change between these is seen to be

$$\mathbf{C}_0^* \times S^1 \to \mathbf{C}_1^* \times S^1,$$
$$(z, \lambda) \mapsto (1/z, \lambda z/|z|).$$

Since this is of the form $(z, \lambda) \mapsto (\varphi(z), h(z)\lambda)$, we have an S^1 bundle over S^2. The inverse image of a circle on S^2 in S^3 is a torus. Picturing

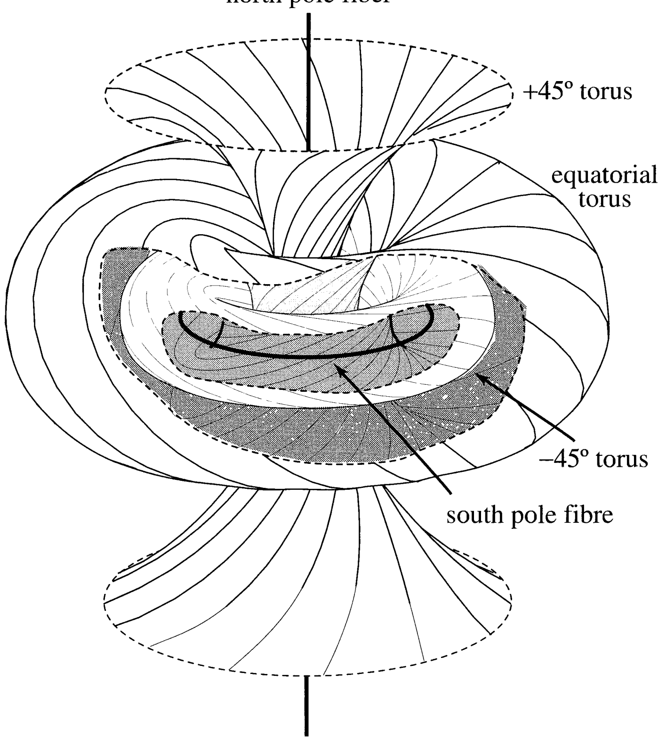


FIGURE 3.15. The Hopf fibration.

S^3 minus the north pole as $\mathbf{R}^3$ via stereographic projection, Figure 3.15 shows an accurate cutaway perspective view in $\mathbf{R}^3$ of the fibers on the three tori lying over the equator, and the ± 45 degree parallels of S^2. The fibers themselves correspond to the places where the ratio w_1/w_0 is constant and appear as circles in Figure 3.15. The fibers over the north and south poles, namely, the z-axis and the "central circle," are also shown. In fact, the subset of $\mathbf{R}^3$ consisting of $\mathbf{R}^3 - z\text{-}axis - central\ circle$ is "foliated" by the tori lying over the lines of latitude on the 2-sphere.

All but two of the fibers of the Hopf bundle lie on these tori in $\mathbf{R}^3$. On each torus the fibers run around "once in each direction." The base S^2 may be regarded as "the space of fibers" in the sense that π establishes a bijection between the set of fibers and the base. Moreover, the topology on the base indicates the mutual disposition of the fibers in the sense that two fibers are "near" each other when the corresponding points in B are "near" each other. Figure 3.16 shows how the "space of fibers" of the Hopf bundle fit together topologically to form the 2-sphere. One of the tori of Figure

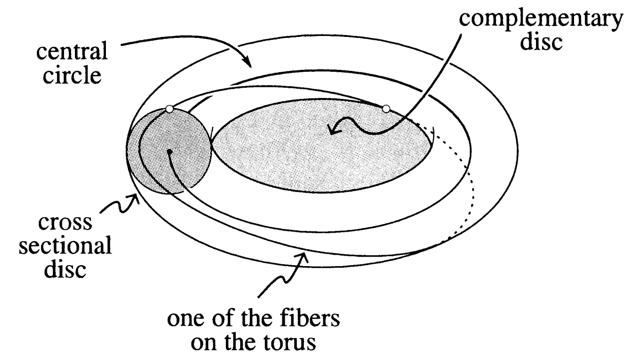


FIGURE 3.16. Topology of the space of fibers.

3.16 has been singled out. A cross-sectional disc for the corresponding solid torus meets every fiber inside the torus just once. The "complementary disc" meets each fiber outside the torus once. The fibers on the torus meet each disc once in its bounding circle and provide a diffeomorphism between these two circles. Thus the "space of fibers" is made up of two 2-discs glued together by a diffeomorphism of their boundaries, which is just the 2-sphere. ◆

Exercise 3.17. Show that the Möbius band is a G bundle with $G = \mathbf{Z}/2$. ❑

Construction of Bundles

Now that we have seen an interesting example of a G bundle, let us return to the definition to see what ingredients are required for the *construction* of a G bundle. It is clear from the definition that if we are given an effective G bundle, then we are given

(i) a Lie group G acting smoothly on a smooth manifold F,

(ii) an open covering $\{U_\alpha\}$ of a manifold B,

(iii) smooth maps $h_{\beta\alpha}: U_\alpha \cap U_\beta \to G$ related by the property that if $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$, then on this intersection we have $h_{\gamma\beta}h_{\beta\alpha} = h_{\gamma\alpha}$.

The latter property is a consequence of the existence of the effective G bundle and may be seen as follows. We may assume that the open sets U_α are indexed by the trivializations

$$\alpha: \pi^{-1}(U_\alpha) \to U_\alpha \times F.$$

The coordinate changes are $\beta\alpha^{-1}\colon (U_\alpha \cap U_\beta) \times F \to (U_\alpha \cap U_\beta) \times F$, where $\beta\alpha^{-1}(u, f) = (u, h_{\beta\alpha} f)$. Thus, the obvious identity $\gamma\beta^{-1}\beta\alpha^{-1} = \gamma\alpha^{-1}$ is equivalent to $h_{\gamma\beta} h_{\beta\alpha} = h_{\gamma\alpha}$.

It is quite clear that this procedure may be reversed in that if we are given the data of (i), (ii), and (iii), we can construct a G bundle by forming the disjoint union of the products $U_\beta \times F$ and dividing by the equivalence relation generated by making the identifications

$$\begin{array}{ccc} U_\alpha \times F & & U_\beta \times F \\ \cup & & \cup \\ (U_\alpha \cap U_\beta) \times F & \to & (U_\alpha \cap U_\beta) \times F \\ (u, f) & \mapsto & (u, h_{\beta\alpha} f). \end{array}$$

Exercise 3.18. Verify that this procedure does yield a smooth G bundle over B with fiber F and that if the data (i), (ii), and (iii) arose from a G bundle, this procedure reconstructs that G bundle (up to isomorphism). ❑

Principal Bundles

A special type of G bundle is the one for which the group G "is the same as" the fiber F in the sense that for some (and hence any) point $f_0 \in F$, the map $G \to F$ sending $g \to g f_0$ is a diffeomorphism. (For example, the Hopf bundle fits this description.) It then follows that the G bundle is effective, so the transition maps $h\colon U \cap V \to G$ are determined by the bundle. Note that the diffeomorphism $G \to F$ does not yield a canonical identification of G with the abstract fiber F, because the bijection will vary with the choice of $f_0 \in F$. Nevertheless, once we have chosen such an identification, we can reconstruct the bundle using G itself as the (abstract) fiber together with the left action of G on itself, as may be seen from the following diagram.

$$\begin{tikzcd}
 & (u, g f_0) \arrow[rr, mapsto] & & (u, h(u) g f_0) & \\
U \times F \supset & (U \cap V) \times F \arrow[rr] & & (U \cap V) \times F & \subset V \times F \\
U \times G \supset & (U \cap V) \times G \arrow[rr] \arrow[u] & & (U \cap V) \times G \arrow[u] & \subset V \times G \\
 & (u, g) \arrow[rr, mapsto] \arrow[uuu, mapsto, bend left] & & (u, h(u) g) \arrow[uuu, mapsto, bend right] &
\end{tikzcd}$$

Moreover, it further follows from this diagram that a principal G bundle has a *smooth right G action*, as may be seen by comparing the coordinate changes before and after we identify the fiber with G. The right G action commutes with the coordinate changes, which themselves involve only the left G action. This leads us to the following definition.

Definition 3.19. A *principal*[13] G bundle is a smooth fiber bundle $\xi = (P, B, \pi, F)$ together with a right action $P \times G \to P$ that is fiber preserving and acts simply transitively on each fiber. ❀

Example 3.20. A regular covering space $\pi\colon \tilde{M} \to M$ is a G bundle with discrete fiber whose group is $\pi_1(M, b)/N$, where $N = \text{Image } \pi_*\colon \pi_1(\tilde{M}, \tilde{b}) \to \pi_1(M, b)$. ◆

Exercise 3.21. Verify that a *principal* G bundle does in fact possess a canonical effective G bundle structure. [*Hint*: the coordinate changes must commute with the G action.] ❑

Exercise 3.22.* Let $\xi = (P, B, \pi, F)$ be a principal G bundle. Show that the action $P \times G \to P$ is *proper* in the sense that if A and B are compact subsets of P then $\{g \in G \mid (gA) \cap B \neq \emptyset\}$ is compact. ❑

Exercise 3.23.* Let $H \to P \to M$ be a principal bundle, and suppose that M is connected. Fix a component P_1 of P and an element $p_1 \in P_1$, and set $H_1 = \{h \in H \mid p_1 h \in P_1\}$.

(i) Show that H_1 is a codimension-zero subgroup of H.

(ii) Show that $P_1 H_1 \subset P_1$.

(iii) Show that $H_1 \to P_1 \to M$ is a principal H_1 bundle. ❑

The case of a principal G bundle is fundamental because of the inverse constructions that allow us to pass back and forth between effective G bundles with fiber F and principal G bundles. These passages may be described as follows. An effective G bundle with fiber F determines the transition functions $h\colon U \cap V \to G$, and hence, as above, we can construct a principal G bundle called the *associated G bundle*. Conversely, given a principal G bundle ξ and a smooth effective action of G on a manifold F, we can use the action to construct a G bundle with fiber F denoted by $\xi \times_G F$. These two constructions are inverse to each other. Therefore, we may say that the fiber of an effective G bundle may be regarded as a variable for which we may substitute any manifold on which G acts effectively.

Vector Bundles

A special case of a G bundle is a *real vector bundle*. In this case the fiber is a real finite-dimensional vector space V and the group is the general linear group $G = Gl(V)$. An *isomorphism* of vector bundles is an isomorphism of

[13]This term apparently arises because a principal bundle generalizes the so-called *principal group* of a Klein geometry (cf. Chapter 4).

bundles that is linear on the fibers. Since the action of G on V is transitive and effective, it follows that we may pass back and forth between the vector bundle and the principal G bundle ξ. What is more, given any smooth linear representation $\rho: G \to G(W)$, we can pass from the original vector bundle to the associated principal bundle ξ and then to the G bundle with fiber W denoted by $\xi \times_\rho W$. In fact, the fibers of the canonical map $\xi \times W \to \xi \times_\rho W$ are just the orbits of the left G action $G \times \xi \times W \to \xi \times W$ given by

$$g \cdot (x, w) = (xg^{-1},\ \rho(g)w).$$

Example 3.24 (Exterior power). If ξ is a vector bundle with fiber V, with group $Gl(V)$ acting on V by the standard representation, then $Gl(V)$ also act on $\Lambda^p(V)$, the pth exterior power. Thus, we have an associated bundle $\Lambda^p(\xi)$, called the *pth exterior power* of ξ. ◆

Exercise 3.25 (Whitney sum). Let ξ_1 and ξ_2 be two vector bundles over the same base B, and let $\Delta: B \to B \times B$ be the diagonal map. Then the Whitney sum of the bundles is $\xi_1 \oplus \xi_2 = \Delta^*(\xi_1 \times \xi_2)$. Show that the *Whitney sum* is associative and commutative (up to bundle isomorphism). ❑

Exercise 3.26 (Quotient Bundle). Let ξ_1 and ξ_2 be two vector bundles over the same base B, and let $\varphi: \xi_1 \to \xi_2$ be a smooth injection covering the identity on B and linear on fibers. Show that there is a canonical *quotient bundle* ξ_2/ξ_1 over B whose fibers are the quotients of the corresponding fibers of ξ_1 and ξ_2. ❑

Sections

Now let us consider sections (also called fields or tensors, in the case of vector bundles), the generalization of functions mentioned in the introduction to this section.

Definition 3.27. Let $\xi = (E, B, \pi, F)$ be a bundle. A *section* over $U \subset B$ is a continuous (or smooth, in the case of a smooth bundle) map $\sigma: U \to E$ such that $\pi\sigma = \mathrm{id}_B$. The space of all (smooth) sections of η is denoted by $\Gamma(\eta)$. ❀

In general, a field in an n-dimensional real vector bundle may be regarded as a generalization of a vector-valued function on the base, which is *locally*, of course, exactly what a field is in this case. In fact, the same may be said about sections on a general bundle. The advantage of the generalization is that, in many geometric circumstances, fields correspond more closely to the nature of things than do functions. The close relationship between fields and functions accounts for the fact that in the presence of a *trivial bundle with a canonical trivialization* $t: E \to F$, there is a tendency to identify the field $\sigma: B \to E$ with the corresponding function $t\sigma: B \to F$.

Let $\xi = (E, B, \pi, G)$ be a principal G bundle over B, and consider the bundle $\eta = \xi \times_G F$ with fiber F, associated to ξ by a smooth effective action ρ of G on a manifold F. There is an important bijective correspondence

$$\psi: A^0(E, \rho) \to \Gamma(\eta)$$

where

$$A^0(E, \rho) = \{f: E \to F \mid f \text{ is smooth, and } f(pg) = \rho(g^{-1})f(p)\}$$

and the section $\psi(f): B \to \eta$ is induced by the map $P \to P \times F$ sending $p \mapsto (p, f(p))$.

Exercise 3.28.* Verify that the above correspondence is well defined and bijective. ❑

§4. Tangent Vectors, Bundles, and Fields

Tangent vectors have a schizophrenic nature. On the one hand, they have a geometric aspect in which they appear as directions in space: if I stand in a manifold, I can move in a variety of directions, which may be described as the tangent vectors at my position. On the other hand, they have an analytical aspect in which they appear as "directional derivatives," which is to say as first-order partial differential operators that, when applied to a smooth function, give its rate of increase in the given direction. Although these are "the same" notion in some sense, they nevertheless have a different development. While "directions" strike the mind as a primarily geometric notion, differential operators lie in the domain of analysis and can, for example, be composed to yield partial differential operators of arbitrary order whose geometric significance is often unclear.[14]

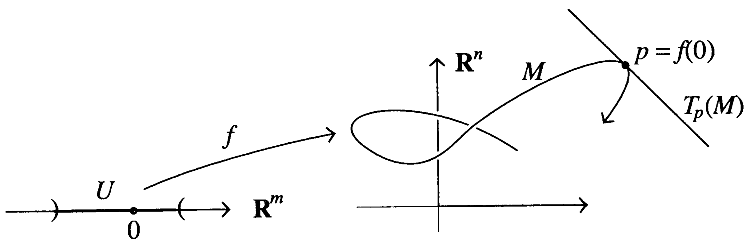
Geometric Vectors

We start with the geometric view of vectors. In courses on advanced calculus, we learn how, at least in certain cases, to find the tangent plane at a point p of a submanifold $M^m \subset \mathbf{R}^n$. There are two methods, depending on whether the submanifold is described parametrically or implicitly. Of

[14]In fact, we are simplifying matters somewhat. The schizophrenic nature of tangent vectors also includes a third personality, an algebraic one. If $\mathfrak{m}_x$ is the ideal of smooth functions on M which vanish at the point x, then the space of tangent vectors at x can also be identified with the dual space $(\mathfrak{m}_x/\mathfrak{m}_x^2)^*$. However, we do not pursue this aspect of tangent vectors here.

course, in cases where both methods apply, they lead to the same tangent plane which we shall denote by $T_p(M)_{\text{geometric}}$. Later we shall drop the term *geometric*.

The Parametric Method. Let $p \in M^m$, where M^m is a submanifold of $\mathbf{R}^n$. Assume that p lies in a neighborhood of M^m which is defined parametrically by a parameterization $f: (U, 0) \to (M^m, p)$, where U is an open set in $\mathbf{R}^m$ and $\operatorname{rank}_0(f) = n$. Then the tangent plane $T_p(M)_{\text{geometric}}$ is also described parametrically as the image of the affine map $A: \mathbf{R}^m \to \mathbf{R}^n$ given by $Av = f(0) + f'(0)v$.



Exercise 4.1. Let $\mathfrak{o}_n(\mathbf{R}) = \{X \in M_n(\mathbf{R}) \mid X^t = -X\}$ be the $\frac{1}{2}n(n-1)$-dimensional vector space of skew-symmetric matrices. Show that the *Cayley parameterization*

$$f: \mathfrak{o}_n(\mathbf{R}) \to M_n(\mathbf{R}), \quad \text{where} \quad f(X) = (I + X)(I - X)^{-1}$$

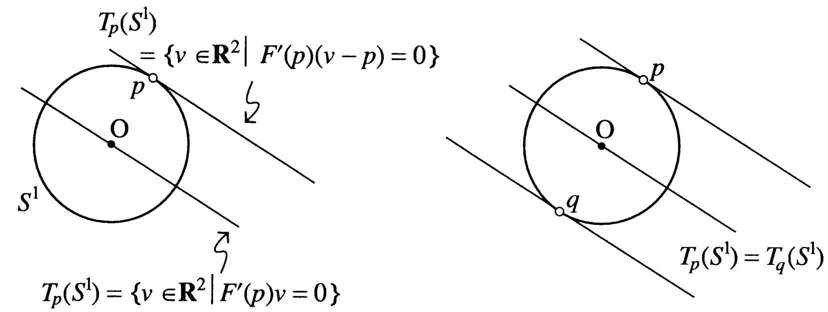
is a smooth embedding whose image is an open subset of $SO_n(\mathbf{R})$. ❑

Example 4.2. Let us find the tangent plane at the identity $I = f(0)$ of the submanifold described parametrically by f in Exercise 4.1. Calculating the derivative $f'(0)$ from first principles, we have

$$\begin{aligned}\lim_{h\to 0} \frac{1}{h}\{f(hY) - f(0)\} &= \lim_{h\to 0} \frac{1}{h}\{(I + hY)(I - hY)^{-1} - I\} \\ &= \lim_{h\to 0} \frac{1}{h}\{(I + hY) - (I - hY)\}(I - hY)^{-1} \\ &= 2Y.\end{aligned}$$

Thus, the tangent plane at the identity is (a translated copy of) $\mathfrak{o}_n(\mathbf{R})$ itself, parametrized by the affine map $A: \mathfrak{o}_n(\mathbf{R}) \to M_n(\mathbf{R})$ given by $A(Y) = I + 2Y$. ◆

The Implicit Method. Let $p \in M^m$, where M is a submanifold of $\mathbf{R}^n$. Assume that $p \in U \subset \mathbf{R}^n$, where $U \cap M$ is defined implicitly as the zero set of a smooth map $F: U \to \mathbf{R}^{n-m}$ such that $\operatorname{rank}_p(F) = n - m$. Then the



the two geometrical placements of tangent spaces

difficulty separating tangent vectors

FIGURE 4.4.

tangent plane $T_p(M)_{\text{geometric}}$ is also described implicitly as the zero set of the affine map $B: \mathbf{R}^m \to \mathbf{R}^n$ given by $Bv = F'(p)(v - p)$.

Example 4.3. Let us calculate the tangent plane of the n-sphere $S^n = \{x \in \mathbf{R}^{n+1} \mid x \cdot x = 1\}$ by this method. Here $F: \mathbf{R}^{n+1} \to \mathbf{R}$ is given by $F(x) = x \cdot x - 1$. Calculating the derivative of F from first principles yields $F'(x)v = 2x \cdot v$. Thus, $\operatorname{rank}_x(F) = 1$ for each $x \in S^n$, and the tangent plane is given by $T_x(S^n)_{\text{geometric}} = \{x + v \in \mathbf{R}^{n+1} \mid x \cdot v = 0\}$. ◆

These constructions of the tangent planes suffer from several disadvantages. The first one, which is not very serious, is that although the tangent plane can be regarded as a vector space, it is generally not a vector subspace of the ambient $\mathbf{R}^n$ (Figure 4.4). This can be fixed by translating each tangent plane $T_p(M)_{\text{geometric}}$ by $-p$ so that it becomes a subspace with p corresponding to the origin (cf. Figure 4.4). We continue to denote this by $T_p(M)_{\text{geometric}}$. The more substantial second objection is that vectors in distinct tangent planes may correspond to the same point in $\mathbf{R}^n$. This can be remedied by means of the following construction, which defines the *geometric tangent bundle* $T(M)_{\text{geometric}}$ of a smooth submanifold M of $\mathbf{R}^n$:

$$T(M)_{\text{geometric}} = \{(p, v) \in \mathbf{R}^n \times \mathbf{R}^n \mid p \in M, v \in T_p(M)\}.$$

It is easy to see that $T(M)_{\text{geometric}}$ is always a smooth submanifold of $\mathbf{R}^n \times \mathbf{R}^n$. For example, if M is given implicitly as

$$M = \{p \in \mathbf{R}^n \mid F(p) = 0\},$$

where $F: \mathbf{R}^n \to \mathbf{R}^{n-m}$ is smooth with $\operatorname{rank}_p(F) = n - m$ for all $p \in M$, then $T(M)$ is given implicitly as

$$T(M)_{\text{geometric}} = \{(p, v) \in \mathbf{R}^n \times \mathbf{R}^n \mid F(p) = 0, F'(p)v = 0\},$$

namely, as the zero set of the smooth function

$$(F, B)\colon \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}^{n-m} \times \mathbf{R}^{n-m},$$

where $B(p, v) = F'(p)v$. Let (x, y) denote the coordinates of the two factors $\mathbf{R}^n \times \mathbf{R}^n$. Calculating the derivative, we have

$$(F, B)'(p, v) = \begin{pmatrix} \left(\frac{\partial F_i}{\partial x_j}\right) & 0 \\ \left(\frac{\partial B_i}{\partial x_j}\right) & \left(\frac{\partial B_i}{\partial y_j}\right) \end{pmatrix},$$

which has rank $2(n-m)$ everywhere, since both main diagonal entries are copies of $F'(p)$, which has rank $n-m$.

Exercise 4.5. Show that if M is given implicitly by $F = 0$ as above, then the first factor projecton $\pi\colon T(M) \to M$ has a canonical vector bundle structure. ❏

Analytic Vectors

As interesting and geometrically explicit as the preceding results are, there still remains the major disadvantage that the definition of the tangent plane and that of tangent bundle seem to depend on the way in which M is embedded in $\mathbf{R}^n$. It is much more satisfying and useful to ask and answer the following question: "Is there an *intrinsic* way of associating a tangent plane and tangent bundle to a smooth manifold so that it is not only independent of a particular embedding, but also so that if we *do* have an embedding $M^m \to \mathbf{R}^n$, we can see that we get the same answer as above?" Using Whitney's embedding theorem it is possible, although awkward, to show directly that the above construction is what we seek. However, it is better now to abandon, temporarily, the geometric aspects of vectors and pass to the analytic description of tangent vectors as *directional derivatives.*

To a vector $v = (v_1, \ldots, v_n) \in \mathbf{R}^n$ we may associate the differential operator

$$D_v = \sum_{1 \le i \le n} v_i \frac{\partial}{\partial x_i},$$

called the *directional derivative in the direction* v. In the case that v is a unit vector in $\mathbf{R}^n$ and $f\colon \mathbf{R}^n \to \mathbf{R}$ is a smooth function, then the geometrical meaning of $D_v(f)(x)$ is that it gives the rate of change of f in the direction v. We codify the notion of a differential operator in the following definition.

Definition 4.6. A *derivation at* $p \in M^m$ is a real-valued operator D defined on smooth functions whose domain contains a neighborhood of p such that, for every pair f, g of such functions, D satisfies the following two properties:

(i) (Linearity) $D(\alpha f + \beta g) = \alpha D(f) + \beta D(g)$ for all $\alpha, \beta \in \mathbf{R}$,

(ii) (Leibnitz formula) $D(fg) = D(f)g(p) + f(p)D(g)$. ❀

The space of all derivations at a point p clearly constitutes a real vector space, which we call the *analytic tangent space* of M at p and denote by $T_p(M)_{\text{analytic}}$. Moreover, the two properties imply that $D(c) = 0$ for any constant function c. It is not clear at this stage that $T_p(M)_{\text{analytic}}$ is finite dimensional, although we shall soon see it is of dimension m. Later, when we have seen that this definition of the tangent space is in full agreement with the geometric definition, we shall drop the term *analytic.*

We need several simple results before we can compare analytic and geometric vectors. The first shows that smooth maps induce linear transformations between the analytic tangent spaces at corresponding points. In a word, it says that we can differentiate smooth maps.

Theorem 4.7.

(a) (*Derivative*) *If* $g\colon (M^m, x) \to (N^n, y)$ *is a smooth map, then* g *induces a linear map* $g_{*x}\colon T_x(M)_{\text{analytic}} \to T_y(N)_{\text{analytic}}$ *defined by* $(g_{*x}D)h = D(hg)$.

(b) (*Chain Rule*) *If* $g\colon (M^m, x) \to (N^n, y)$ *and* $f\colon (N^n, y) \to (P^p, z)$ *are smooth maps, then* $(fg)_{*x} = f_{*y}g_{*x}$.

Proof. Both statements are simple exercises. ■

Corollary 4.8. *If* $f\colon M \to N$ *is a diffeomorphism in some neighborhood of* $x \in M$, *then* $f_{*x}\colon T_x(M)_{\text{analytic}} \to T_{f(x)}(N)_{\text{analytic}}$ *is an isomorphism.*

Proof. Let $x \in U \subset M$ with U open and $f\colon U \to f(U)$ a diffeomorphism. Let g be the inverse diffeomorphism. Applying the chain rule to $gf = \mathrm{id}_U$ and $fg = \mathrm{id}_{f(U)}$, we get $\mathrm{id} = g_{*y}f_{*x}$ and $\mathrm{id} = f_{*x}g_{*y}$. Thus f_{*x} is an isomorphism. ■

The next result is a special case of Taylor's theorem and the result that follows it applies it to determine the derivations at a point in Euclidean space.

Lemma 4.9. *If* f *is a smooth function defined in a neighborhood of* $p \in \mathbf{R}^n$, *then*

$$f(x) = f(p) + \sum_{1 \le i \le n} (x_i - p_i) \left\{ \frac{\partial f}{\partial x_i}(p) + a_i(x) \right\},$$

where the functions $a_i(x)$ *are smooth and vanish at* p.

Proof. Write

$$f(x)-f(p) = \int_0^1 \frac{\partial f(p+t(x-p))}{\partial t}dt = \sum_{1\le i\le n}(x_i-p_i)\int_0^1 \frac{\partial f}{\partial x_i}(p+t(x-p))dt.$$

Now, integrating by parts, we get

$$\begin{aligned}\int_0^1 &\frac{\partial f}{\partial x_i}(p+t(x-p))dt \\ &= t\frac{\partial f}{\partial x_i}(p+t(x-p))\Big|_0^1 - \int_0^1 t\sum_{1\le j\le n}\frac{\partial^2 f}{\partial x_i \partial x_j}(p+t(x-p))(x_j-p_j)dt \\ &= \frac{\partial f}{\partial x_i}(x) + a_i(x),\end{aligned}$$

where the $a_i(x)$ are obviously smooth and vanish at p. ■

Proposition 4.10. *If U is an open subset of $\mathbf{R}^n$ and D is a derivation at $p \in U$, then*

$$D = \sum_{1\le i\le n} D(x_i)\frac{\partial}{\partial x_i}\Big|_p.$$

Proof. Write

$$f(x) = f(p) + \sum_{1\le i\le n}(x_i-p_i)\frac{\partial f}{\partial x_i}(p) + a_i(x)),$$

as in the lemma. Applying D to both sides yields

$$Df = 0 + \sum_{1\le i\le n} D(x_i-p_i)\left(\frac{\partial f}{\partial x_i}(p)+a_i(x)\right) + 0 = \sum_{1\le i\le n} D(x_i)\frac{\partial f}{\partial x_i}(p). \blacksquare$$

Corollary 4.11. $\dim T_p(\mathbf{R}^n)_{\text{analytic}} = n$.

Proof. The formula shows that $T_p(\mathbf{R}^n)_{\text{analytic}}$ is spanned by the derivations

$$D_i = \frac{\partial}{\partial x_i}\Big|_p.$$

These derivations are linearly independent since we obtain $\lambda_j = 0 (1 \le j \le n)$ when a putative relation $\sum \lambda_i D_i = 0$ is applied to the function x_j. ■

Corollary 4.12. *Let $f\colon M \to N$ be a smooth map. Then the rank of f at $p \in M$ is the rank of $f_{*p}\colon T_p(M)_{\text{analytic}} \to T_{f(p)}(N)_{\text{analytic}}$.*

Proof. This is a local question, so we may assume that M is an open set in $\mathbf{R}^m$ (with coordinates $x_1, \dots, x_m$) and N is an open set in $\mathbf{R}^n$ (with coordinates $x_1, \dots, x_n$). We show that in this case the rank of f_{*p} is the same as the rank of the Jacobian matrix of f at p. Set $D_i = \partial/\partial x_i$ and write

$$(f_{*p})D_i = \sum_{1\le j\le n} a_{ij}D_j, \qquad 1 \le i \le m.$$

Then $a_{ij} = ((f_{*p})D_i)x_j = D_i(x_j \circ f) = D_i(f_j) = (\partial f_j/\partial x_i)(p)$. Thus, the Jacobian matrix is, in fact, the matrix of the linear transformation f_{*p} in the basis D_i. ■

Corollary 4.13. *If $g\colon M \to N$ is a smooth map whose derivative $g_{*x}: T_p(M)_{\text{analytic}} \to T_{g(p)}(N)_{\text{analytic}}$ is a linear isomorphism for a given point p, then g is a local diffeomorphism (i.e., p has an open neighborhood U such that $g(U)$ is open and $g \mid U$ is a diffeomorphism onto its image).*

Proof. The isomorphism implies that M and N have the same dimension m and that f has rank m at p, so the inverse function theorem applies. ■

Identity of Geometric and Analytic Vectors

At this stage it is possible to compare the analytic vectors to the geometric ones. Thus, suppose we have a smooth submanifold of Euclidean space $\iota\colon M^m \subset \mathbf{R}^n$ given as the zero set of a smooth function $F = (F_1, \dots, F_{n-m})\colon \mathbf{R}^n \to \mathbf{R}^{n-m}$ of rank $n-m$. Then

$$T_p(M)_{\text{geometric}} = \ker F'(p)\colon \mathbf{R}^n \to \mathbf{R}^{n-m}.$$

Now the smooth inclusion $\iota\colon M^m \subset \mathbf{R}^n$ induces a linear map

$$\iota_{*p}\colon T_p(M^m)_{\text{analytic}} \to T_p(\mathbf{R}^n)_{\text{analytic}}.$$

If $D \in T_p(M^m)_{\text{analytic}}$, then $\iota_{*p}D \in T_p(\mathbf{R}^n)_{\text{analytic}}$, and by Proposition 4.10 we can write

$$\iota_{*p}D = \sum_{1\le i\le n} v_i \frac{\partial}{\partial x_i}, \quad \text{where } v_i = (\iota_{*p}D)x_i \in \mathbf{R}.$$

We claim that $(v_1, \dots, v_n) \in T_p(M)_{\text{geometric}}$. To see this, note that $F_j(\iota(x)) = 0$ for all $x \in M$ and for all j. Thus, for $D \in T_p(M)_{\text{analytic}}$, we have

$$\begin{aligned}0 &= D(F_j\iota) = (F_j\iota)_{*p}(D) = F_{j*\iota(p)}(\iota_{*p}(D)) \\ &= F_{j*\iota(p)}\left(\sum_{1\le i\le n} v_i\frac{\partial}{\partial x_i}\right) = \sum_{1\le i\le n} v_i \frac{\partial F_j}{\partial x_i}(p) \\ &\Rightarrow F'(p)v = 0 \Rightarrow v \in T_p(M)_{\text{geometric}}.\end{aligned}$$

This yields a linear map

$$T_p(M)_{\text{analytic}} \to T_p(M)_{\text{geometric}}$$
$$D \mapsto \big((\iota_{*p}D)x_1, \ldots, (\iota_{*p}D)x_n\big).$$

To show that the map is injective (and hence an isomorphism since the dimensions are the same), it suffices to show that if $(\iota_{*p}D)x_j = 0, 1 \leq j \leq n$ then $Df = 0$ for every smooth function f defined on a neighborhood of p in M. Now the value of Df is not altered if we restrict f to a smaller domain about p. But if the domain is small enough, then f extends to a neighborhood of p in $\mathbf{R}^m$, that is, $f = g\iota$ for g smooth on a neighborhood of p in $\mathbf{R}^m$. Thus,

$$Df = D(g\iota) = g_{*\iota(p)}\iota_{*p}(D) = (\iota_{*p}(D))g = \sum v_i \frac{\partial g}{\partial x_i}(p) = 0$$

since $v_i = (\iota_{*p}D)x_i = 0$. These remarks yield a canonical identification between $T_p(M^m)_{\text{analytic}}$ and $T_p(M)_{\text{geometric}}$ in the case $M \subset \mathbf{R}^m$.

The Tangent Bundle

Now let us return to the tangent bundle. We want to give a definition of $T(M)$ as a smooth manifold intrinsically associated to the smooth manifold M and also to show that it is a $Gl_m(\mathbf{R})$ bundle with fiber $\mathbf{R}^m$ under the standard action.

For each manifold M we define $T(M)_{\text{analytic}}$ as a set to be the *disjoint union* of the vector spaces $T_p(M)$, $p \in M$. Denote by π the canonical *projection map* $\pi: T(M)_{\text{analytic}} \to M$ sending the vector $v \in T_p(M)$ to the point $p \in M$. If $f: M \to N$ is a smooth map, then we define a map $f_*: T(M)_{\text{analytic}} \to T(N)_{\text{analytic}}$ by sending the vector $v \in T_p(M)$ to $f_{*p}v \in T_{f(p)}(N)$ such that the diagram commutes.

$$\begin{tikzcd}
T(M)_{\text{analytic}} \arrow[r, "f_*"] \arrow[d, "\pi_M"'] & T(N)_{\text{analytic}} \arrow[d, "\pi_N"] \\
M \arrow[r, "f"] & N
\end{tikzcd}$$

Now we put a topology on the sets $T(M)_{\text{analytic}}$ by modeling pieces of it on a special case of the geometric tangent bundle. We do this first for the case $M = U =$ open set in $\mathbf{R}^m$. As sets we have the identification

$$T(U)_{\text{analytic}} = T(U)_{\text{geometric}} = U \times \mathbf{R}^n,$$
$$\left(x, \sum_i v_i \frac{\partial}{\partial x_i}\right) \mapsto (x, (v_1, v_2, \ldots, v_n)).$$

Now the right-hand side comes equipped with a topology, a smooth structure, and even the structure of a (trivial) smooth $Gl_n(\mathbf{R})$ bundle. The idea is to pass this to $T(M)_{\text{analytic}}$.

Let $\mathcal{A} = \{(U, \varphi)\}$ be a smooth atlas for M. Then $\varphi: U \to \varphi(U) \subset \mathbf{R}^m$ is a homeomorphism onto its image and $\varphi_*: T(U)_{\text{analytic}} \to T(\varphi(U))_{\text{analytic}} = \varphi(U) \times \mathbf{R}^m$.

The Topology. A set $W \subset T(M)_{\text{analytic}}$ is open $\Leftrightarrow$ $\varphi_*(W \cap T(U)_{\text{analytic}})$ is open in $\varphi(U) \times \mathbf{R}^m$ for all $(U, \varphi) \in \mathcal{A}$.

The Atlas for the Smooth Structure. Take $\mathcal{T} = \{(T(U)_{\text{analytic}}, \varphi_*) \mid (U, \varphi) \in \mathcal{A}\}$.

The Bundle. Again take $\mathcal{B} = \{(T(U)_{\text{analytic}}, \varphi_*) \mid (U, \varphi) \in \mathcal{A}\}$.

Exercise 4.14.* Show that $T(M)_{\text{analytic}}$ is a smooth manifold of dimension $2m$ and that $(T(M)_{\text{analytic}}, M, \pi, \mathbf{R}^m)$ is a smooth $Gl_m(\mathbf{R})$ bundle that can be canonically identified with the geometric tangent bundle $T(M)_{\text{geometric}}$ whenever M is a submanifold of $\mathbf{R}^m$. ❑

From now on we drop the decorations *analytic* and *geometric* in reference to tangent vectors and bundles.

Since the action of $Gl_n(\mathbf{R})$ on $\mathbf{R}^n$ is faithful, we have available the *bundle* $\leftrightarrow$ *G bundle* correspondence, described on page 36, and so we may speak of the bundle associated to any representation of $Gl_n(\mathbf{R})$.

Exercise 4.15.* Let M and N be smooth manifolds and let $\pi_M: M \times N \to M$ and $\pi_N: M \times N \to N$ be the canonical projections. They induce maps $\pi_{M*}: T(M \times N) \to T(M)$ and $\pi_{N*}: T(M \times N) \to T(N)$, which we can combine into a map $\pi_{M*} \times \pi_{N*}: T(M \times N) \to T(M) \times T(N)$. Show that this map is a diffeomorphism. [*Hint:* Use the atlas for $M \times N$ to reduce the question to the case where M and N are open sets in Euclidean spaces.] ❑

Exercise 4.16.* Continuing the notation of the last exercise, we fix $p \in M$ and $q \in N$ and define

$$\begin{aligned} i_M:\ & M \to M \times N, & i_N:\ & N \to M \times N, \\ & x \mapsto (x, q) & & y \mapsto (p, y), \\ \Pi_M:\ & T_p(M) \times T_q(N) \to T_p(M), & \Pi_N:\ & T_p(M) \times T_q(N) \to T_q(N), \\ & (u, v) \longmapsto u, & & (u, v) \longmapsto v. \end{aligned}$$

Show that the following maps are inverse to each other:

$$T_{(p,q)}(M \times N) \underset{\iota_{M*p}\amalg_M + \iota_{N*q}\amalg}{\overset{\pi_{M*(p,q)} \times \pi_{N*(p,q)}}{\rightleftarrows}} T_p(M) \times T_q(N).$$

❑

Exercise 4.17. Show that a smooth manifold M of dimension n is orientable if and only if the nth exterior power $\lambda^n(T(M))$ is a trivial line bundle over M. [*Hint:* Choose a never-zero section $e \in \Gamma(\lambda^n(T(\mathbf{R}^n)))$. If $\lambda^n(T(M))$ is trivial, we may choose a never-zero section $\sigma \in \Gamma(\lambda^n(T(M)))$. Consider the charts (U, φ) such that, for each $x \in U$, $\sigma(x)$ and $e(\varphi(x))$ correspond up to a positive scalar factor under the isomorphism

$$\varphi_{*x}: \lambda^n(T_x(M)) \to \lambda^n(T_{\varphi(x)}(\mathbf{R}^n)).$$

Show these charts yield an orientation for M.] ❑

It sometimes happens that the tangent bundle $T(M)$ of a smooth manifold M may be described by a G atlas, where G is some closed subgroup of $Gl_n(\mathbf{R})$. In this case $T(M)$ acquires a G structure, so that it becomes a G bundle. In this case we say that *the manifold M* has a G structure. For example, every manifold has an $O_n(\mathbf{R})$ structure (cf. [W. Boothby, 1986], p. 195). In general, if H is a closed Lie subgroup of G and a G bundle has an atlas involving only H, we say that the H structure defined by such an atlas is a *reduction of the structure group from G to H*.

Here is a useful result generalizing part of the mean-value theorem.

Proposition 4.18. *Let $f: M \to N$ be a smooth map of smooth manifolds. Assume $f_*: T(M) \to T(N)$ is the zero map on each tangent space. Then f is constant on each component of M.*

Proof. It suffices to show that for any chart (V, ψ) on N, f is constant in each component of the open set $f^{-1}(V)$. Let p, q lie in one component of $f^{-1}(V)$. We show that $f(p) = f(q)$. Let $\sigma: (I, 0, 1) \to (f^{-1}(V), p, q)$ be a smooth path joining p and q. Now consider $F = \psi f \sigma: I \to \mathbf{R}^m$. By the chain rule we have $F'(t) = 0$ for all t, so by the ordinary mean-value theorem, each of the m components of F is constant. Thus, F is constant, and so $f(p) = \psi^{-1}F(0) = \psi^{-1}F(1) = f(q)$. ■

Next we generalize part of the fundamental theorem of calculus.

Proposition 4.19. *Let M and N be smooth manifolds with M connected and N having trivial tangent bundle, with trivialization $t: T(N) \to V$. Let $f, g: M \to N$ be smooth maps satisfying*

(i) *$f(p) = g(p)$ for some point $p \in M$,*

(ii) *$tf_* = tg_*$ as maps $T(M) \to V$.*

Then $f = g$.

Proof. It suffices to show that $f(q) = g(q)$ for all $q \in M$. But M is connected, so we may choose a smooth path $\sigma: (I, 0, 1) \to (M, p, q)$, and we are reduced to demonstrating the result for the case $M = \mathbf{R}$. Moreover, by covering $\sigma(I)$ with coordinate charts and using the compactness of I, we can write σ as a composite of finitely many paths, each one lying in a single coordinate chart. Thus we are reduced to considering the case $M = \mathbf{R}$ and $N = \mathbf{R}^m$. Now, if $tf_* = tg_*$ for one trivialization, then the same is true for any other trivialization, so writing f_i and g_i for the coordinates of f and g, we see that condition (ii) becomes

$$\frac{\partial f_i}{\partial t} = \frac{\partial g_i}{\partial t} \qquad \text{for all } i.$$

Thus, $f_i(t) = g_i(t) + C_i$ for some constant C_i. But by (i), all the C_i must vanish, and so $f = g$. ■

Vector Fields

A vector field on a smooth manifold M is a smooth section of the tangent bundle $T(M)$. In particular, it is a map X associating to each point $p \in M$ a vector $X_p \in T_p(M)$. Let us see what smoothness means in terms of local coordinates. Let $(T(U), \varphi_*)$ be a bundle coordinate chart corresponding to the chart (U, φ) on M. Setting $V = \varphi(U)$, we have the commutative diagram of smooth maps:

$$\begin{array}{ccccc} T(U) & \xrightarrow{\varphi_*} & T(V) & = V \times \mathbf{R}^m \\ \downarrow & & \downarrow & \\ U & \xrightarrow{\varphi} & V & \end{array}$$

The horizontal arrows are diffeomorphisms that identify the left and right sides. Thus, we may write a vector field X on U in these coordinates as a map

$$Y: V \to V \times \mathbf{R}^m,$$

$Y(x) = (x, a_1(x), \ldots, a_m(x))$. Clearly, Y (and hence X) is smooth if and only if each $a_m: V \to \mathbf{R}$ is smooth. We may say this is an invariant way by noting that

$$Y \text{ smooth} \Rightarrow Y(f) = \sum a_i(x) \frac{\partial f}{\partial x_i}(x) \text{ is smooth for all smooth } f: V \to \mathbf{R}.$$

Conversely,

$$Y(f) = \sum a_i(x) \frac{\partial f}{\partial x_i}(x) \text{ smooth for all smooth } f\colon V \to \mathbf{R}$$
$$\Rightarrow Y(x_j) = \sum a_i(x) \frac{\partial x_j}{\partial x_i}(x) = a_j(x) \text{ is smooth for all } j$$
$$\Rightarrow Y \text{ smooth.}$$

Proposition 4.20. *A section X of $T(M)$ is smooth $\Leftrightarrow$ $X(f)$ is smooth for all $f \in C^\infty(M)$.*

Proof. $\Rightarrow$ To check if $X(f)$ is smooth, we need only check that its restriction to each chart U is smooth. But by the remarks above, X smooth implies that $X(g)$ is smooth for all $g \in C^\infty(U)$. Now if $f \in C^\infty(M)$, then $f \mid U \in C^\infty(U)$, and so $X(f) \mid U = X(f \mid U)$ is smooth. Thus $X(f)$ is smooth.

$\Leftarrow$ Now assume that $X(f)$ is smooth for all $f \in C^\infty(M)$. If we can show that $X(g)$ is smooth for all $g \in C^\infty(U)$, then we will be done. But for each point $p \in U$ there is a function $h \in C^\infty(M)$ that is identically 1 on a neighborhood V of p and identically zero on a neighborhood of $M - U$ (cf. [W. Boothby, 1986], pp. 193–195, on partitions of unity). Thus, $hg \in C^\infty(M)$ so $X(hg)$ is smooth on M. But $X(g) = X(hg)$ on V, and so $X(g)$ is smooth on V. Since p was arbitrary, $X(g)$ is smooth on all of U. ∎

The mapping properties of vector fields are not very good. If $\varphi\colon M \to N$ is a smooth map and X is a vector field on M, then $\varphi_*(X)$ is not in general a vector field on N. It can generally be regarded only as a section of the pullback bundle $\varphi^*(T(N))$. If φ is injective, then $\varphi_*(X)$ may be regarded as a vector field on $\varphi(M)$. The following definition is the best we can do by way of a naturality property.

Definition 4.21. Let $\varphi\colon M \to N$ be a smooth injection and let X and Y be vector fields on M and N, respectively. We say that X and Y are φ *related* if $X(f\varphi) = (Y(f))\varphi$ for all $f \in C^\infty(N)$. ❀

Derivations

We may now assemble the properties of smooth vector fields X on M which say that

(1) $X\colon C^\infty(M) \to C^\infty(M)$ is an $\mathbf{R}$-linear map,

(2) $X(fg) = X(f)g + fX(g)$.

Property 2 merely restates the definition of a derivation at p for each $p \in M$. The only thing new here is that, for X smooth, $X(f)$ is also smooth. In fact, these properties say that the smooth vector fields on M are the derivations $\mathrm{Der}(M)$ on $C^\infty(M)$.

The derivations $\mathrm{Der}(M)$ constitute a vector space over $\mathbf{R}$ (of infinite dimension). But there is an additional structure, a multiplication called *bracket* and denoted $[X, Y]$, which turns $\mathrm{Der}(M)$ into a *Lie algebra.*[15] This operation is defined by $[X, Y] = XY - YX$ and satisfies

(1) (Bilinearity) $[\ ,\ ]\colon \mathrm{Der}(M) \times \mathrm{Der}(M) \to \mathrm{Der}(M)$ is $\mathbf{R}$-bilinear,

(2) (Skew symmetry) $[X, Y] = -[Y, X]$,

(3) (Jacobi identity) $[[X, Y], Z] + [[Y, Z], X] + [[Z, X], Y] = 0$.

These properties are all easy to show. In addition, there is the following naturality property.

Lemma 4.22. *Let $\varphi\colon M \to N$ be a smooth injection. Let X_1 and X_2 be vector fields on M, and let Y_1 and Y_2 be vector fields on N. If X_j is φ related to Y_j for $j = 1, 2$, respectively, then $[X_1, X_2]$ and $[Y_1, Y_2]$ are φ related.*

Proof. Since $(Y_j f)\varphi = X_j(f\varphi)$, we have

$$\begin{aligned}
([Y_1, Y_2]f)\varphi &= (Y_1 Y_2 f)\varphi - (Y_2 Y_1 f)\varphi \\
&= X_1((Y_2 f)\varphi) - X_2((Y_1 f)\varphi) \\
&= X_1(X_2(f\varphi)) - X_2(X_1(f\varphi)) \\
&= [X_1, X_2](f\varphi).
\end{aligned}$$
∎

We end the section by showing the local coordinate expression for the bracket of two vector fields. We have

$$X = \sum_{1 \le i \le n} a_i(x) \frac{\partial}{\partial x_i}, \qquad Y = \sum_{1 \le j \le n} b_j(x) \frac{\partial}{\partial x_j},$$

$$[X, Y] = \sum_{1 \le i,j \le n} \left\{ a_i \frac{\partial b_j}{\partial x_i} - b_i \frac{\partial a_j}{\partial x_i} \right\} \frac{\partial}{\partial x_j}.$$

Exercise 4.23. (i) Let $D \in \mathrm{Der}(M)$. Assume that $f \in C^\infty(M)$ vanishes on a neighborhood V of $p \in M$. Show that Df also vanishes on V.

(ii) Show that every derivation $D \in \mathrm{Der}(M)$ arises from a smooth vector field. ❑

[15] See Chapter **3**.2.6 for the formal definition.

§5. Differential Forms

Let us begin with a rough and ready description of p-forms for $p \leq 2$. The 0-forms (with values in a finite-dimensional vector space V) on a manifold M are just the V-valued functions on M. The 1-forms generalize the derivatives of functions on M. The 2-forms are used as a way of formalizing the necessary conditions on a 1-form for it to be the derivative of a function. Now let us give the formal definition of a 1-form.

Definition 5.1. Let M be a smooth manifold and V be a finite-dimensional vector space. A *smooth V-valued 1-form* on M is simply a smooth map $\omega: T(M) \to V$ that is linear on each fiber $T_x(M)$. When $V = \mathbf{R}$ (or $\dim(V) = 1$), we speak of a smooth 1-form on M. More generally, if E is a flat vector bundle over M, a smooth E-valued *1-form* on M is simply a smooth map $\omega: T(M) \to E$ restricting to a linear map $\omega_x: T_x(M) \to E_x$ on each fiber of $T(M)$.

Exterior Derivative of a Smooth Function

If $f: M \to \mathbf{R}$ is a smooth function defined on a smooth manifold, then a 1-form arises from differentiation as follows. The derivative of f is $f_*: T(M) \to T(\mathbf{R})$, and since $T(\mathbf{R})$ has a canonical trivialization

$$\underset{(t,\alpha\frac{\partial}{\partial t}|_t) \mapsto (t,\alpha)}{T(\mathbf{R}) \approx \mathbf{R} \times \mathbf{R},}$$

we may write f_* in the form $f_{*x}v = (f(x), f'(x)v)$. Then define the 1-form $df: T(M) \to \mathbf{R}$ by the formula $df(v) = f'(x)v$ for $v \in T_x(M)$. In exactly the same way, if $f: M \to N$ and $T(N)$ has a canonical trivialization $T(N) \approx N \times V$ (for example, when $N = V$ is a vector space), we can define a 1-form df with values in the vector space V, which is well defined (relative to the given trivialization). In every case, the form df is a smooth 1-form on M, which we call the *exterior derivative* of f.

Exercise 5.2. If we are given a trivialization $\theta: T(N) \to V$, then any smooth map $\tau: N \to Gl(V)$ determines a new trivialization by *twisting* the original according to $\theta^\tau: T(N) \to V$, where $\theta^\tau(p, v) = (p, \tau(p)\theta(v))$. Show that the exterior derivative computed relative to the trivialization θ is the same as that computed relative to the trivialization θ^τ if and only if τ is constant.[16] ❑

[16]This exercise shows that, in general, it is impossible to have a canonically defined exterior derivative for forms with values in a vector bundle E over M, since the differentiation is not independent of the coordinate changes. There is, however, an important exception to this. In the case of a *flat bundle* E, the coordinate changes are constant; in this case, we do have exterior derivatives.

Proposition 5.3. *If $v \in T_p(M)$ and $f: M \to \mathbf{R}$ is smooth, then $df(v) = v(f)$.*

Proof. Write $f_{*p}v = \alpha\frac{\partial}{\partial t}\big|_{f(p)}$. Applying both sides to the identity function $h: \mathbf{R} \to \mathbf{R}$, $h(t) = t$ yields $v(f) = v(hf) = (f_{*p}v)(h) = \alpha$. Thus, $f_{*p}v = v(f)\frac{\partial}{\partial t}\big|_{f(p)}$. So, by the definition of d, we get the result. ∎

Definition 5.4. If f is a smooth V-valued function on M and X is a smooth vector field on M, we define $X(f)|_p = f_*(X_p)$. ❀

Exercise 5.5. If e_i is any basis of V, and we write $f = \sum f_i e_i$, show that $X(f)|_p = \sum f_{i*}(X_p)e_i$. ❑

Interpretation of dx

The notion of the exterior derivative of a mapping allows us to throw some light on Leibnitz's differential notation for the calculus. For example, everyone is familiar with the expression for the derivative of $y = f(x)$ in the form $\frac{dy}{dx} = f'(x)$ or, written more symbolically as "differentials" by $dy = f'(x)dx$. The latter expression can be given the following interpretation. The smooth map $f: \mathbf{R} \to \mathbf{R}$ given by $y = f(x)$ has derivative $f_*: T(\mathbf{R}) \to T(\mathbf{R})$, which, if written in the canonical trivialization for $T(\mathbf{R})$ with coordinates (x, v), has the form $f_{*x}v = (f(x), f'(x)v)$. Thus, the exterior derivative df is given by $df_x(v) = f'(x)v$. In particular, if we regard $x: \mathbf{R} \to \mathbf{R}$ as the *identity function*, then $dx(v) = v$. Hence $df_x(v) = f'(x)dx(v)$, that is, $dy = f'(x)dx$. This identifies the Leibnitz dy with the exterior derivative of the mapping $y = f(x)$. The form dx has the interpretation as "the *general* infinitesimal change of x" in the sense that if we apply it to any "*specific* infinitesimal change of x" in the amount $\Delta x \in T_x(\mathbf{R})$, we get $dx(\Delta x) = \Delta x$. Similarly, dy is "the *general* allowable infinitesimal change in $y = f(x)$" in the sense that if we apply it to any "*specific* infinitesimal change of x" in the amount $\Delta x \in T_x(\mathbf{R})$, we get $dy(\Delta x) = f'(x)\Delta x$ for the resulting specific change in y.

Interpretation of Total Derivatives

Now let us look at the connection between the exterior derivative and the Leibnitz differential notation for smooth maps $f: \mathbf{R}^n \to \mathbf{R}$. Let (x, v) denote the coordinates for points in

$$\underset{\left(x,\sum v_i \frac{\partial}{\partial x_i}\big|_x\right) \mapsto (x,v)}{T(\mathbf{R}^n) \approx \mathbf{R}^n \times \mathbf{R}^n.}$$

We have the ith coordinate function $x_i: \mathbf{R}^n \to \mathbf{R}$, $x_i(x) = x_i$. As in the one-variable case, there is the convenient confusion between the function and the coordinate. The exterior derivative $dx_i: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ is given by $dx_i(v) = v_i$. Now write $f_{*x}v = (f(x), f'(x)v)$ so we see that the exterior derivative of f is given by

$$df_x(v) = f'(x)v = \sum_{1 \le i \le n} \frac{\partial f}{\partial x_i} v_i = \sum_{1 \le i \le n} \frac{\partial f}{\partial x_i} dx_i(v).$$

Suppressing the vector v yields $df_x = \sum_{1 \le i \le n} (\partial f / \partial x_i) dx_i$, which is the Leibnitz *total derivative*. For another discussion of the relationship between the old and new notation, see [M. Spivak, 1970], pp. 153–154.

Exterior Differential Forms

Now we come to the question of *higher derivatives*. As usual in modern differential geometry, we shall be concerned only with the *skew-symmetric part* of the higher derivatives. In essence, what we shall be doing is taking the partial derivatives with respect to the base (i.e., manifold) variables and skew-symmetrizing the result, thus forgetting about the part of the higher derivatives that vanish under this procedure. However, this will not be made explicit in our treatment. The part of the higher derivative that disappears has not been studied much in differential geometry since Elie Cartan showed how useful it is to consider only the skew-symmetric part, that is, the exterior derivative. The old masters did use the symmetric part, and more recently it seems to have found an application in probability theory (cf. [P.A. Meyer, 1989], [J.E. White, 1982], and [B.L. Foster, 1986]).

What are we going to differentiate? We are going to differentiate V-valued p-forms on M to obtain V-valued $(p+1)$-forms on M. These p-forms consist of the following generalizations of 1-forms.

Definition 5.6. Let M be a smooth manifold and V be a finite-dimensional vector space. A *smooth V-valued p-form*[17] on M is a smooth map $\omega: T(M) \oplus \ldots \oplus T(M) \to V$ whose restriction to any fiber $T_x(M) \oplus \ldots \oplus T_x(M) \to V$ is multilinear and totally skew-symmetric in its vector arguments for all $x \in M$. When $V = \mathbf{R}$ (or $\dim(V) = 1$), we speak of a smooth p-form on M. ❀

Although the definition is for general p, for the most part we shall need to work only with $p \le 2$ or 3. Note that, by definition, $A^0(M, V)$ is declared to be the V-valued functions on M. Also note that $A^p(M, V) = 0$ for $p > \dim M$.

[17] Just as in the case of 1-forms, we could consider forms with values in a flat vector bundle over M. Once again exterior differentiation is canonically defined (cf. Footnotes 10 and 11).

The p-forms make up an infinite-dimensional vector space denoted by $A^p(M, V)$. We set

$$A(M, V) = \bigoplus_{0 \le p \le m} A^p(M, V).$$

Note that a smooth map $f: M \to N$ induces a backward map $f^*: A^p(N, V) \to A^p(M, V)$ defined by

$$(f^*\omega)_x(X_1, \ldots, X_p) = \omega_{f(x)}(f_* X_1, \ldots, f_* X_p), \text{ where } X_1, \ldots, X_p \in T_x(M).$$

The *exterior* (or *wedge*) product of a V_1-valued p-form ω_1 with a V_2-valued q-form ω_2 is a $V_1 \otimes V_2$-valued $(p+q)$-form $\omega_1 \wedge \omega_2$. It is defined by[18]

$$\begin{aligned} &\omega_1 \wedge \omega_2(v_1, \ldots, v_{p+q}) \\ &= \sum_{\substack{\sigma \text{ runs over all} \\ (p,q) \text{ shuffles}}} (-1)^\sigma \omega_1(v_{\sigma(1)}, \ldots, v_{\sigma(p)}) \otimes \omega_2(v_{\sigma(p+1)}, \ldots, v_{\sigma(q)}). \end{aligned}$$

Exercise 5.7. Show that $\omega_1 \wedge \omega_2$ is a $(p+q)$-form. ❑

Exercise 5.8. Show that if $f: N \to M$ is a smooth map, then $f^*(\omega_1 \wedge \omega_2) = f^*(\omega_1) \wedge f^*(\omega_2)$. ❑

Now we are going to describe the totally skew maps at a point. Let U and V be finite-dimensional vector spaces of dimension u and v, respectively. Let $e_1, \ldots, e_u$ be a basis of U, $e_1^*, \ldots, e_u^*$ be the dual basis of U^*, and $f_1, \ldots, f_v$ be a basis of V. Let the symbol $I, J, \ldots$ denote the sequence $\{i_1, i_2, \ldots, i_p\}$, $\{j_1, j_2, \ldots, j_p\}, \ldots$, where $1 \le i_1 < i_2 < \ldots < i_p \le u$, etc. We use the abbreviations

$$e_I = (e_{i_1}, \ldots, e_{i_p}), \quad e_I^* = (e_{i_1}^*, \ldots, e_{i_p}^*), \quad \text{etc.}$$

Exercise 5.9. Show that

(i) $I \cap J \ne \emptyset \Rightarrow e_I^* \wedge e_J^* = 0$,

(ii) $I \cap J = \emptyset \Rightarrow e_I^* \wedge e_J^* = (-1)^\sigma e_K^*$, where K is the list of indices I, J written in increasing order, and σ is the shuffle permutation mapping $\{I, J\}$ to K. ❑

[18] σ is a shuffle permutation if $r < s \Rightarrow \sigma(r) < \sigma(s)$ whenever both r and s lie in $\{1, 2, \ldots, p\}$, or both lie in $\{p+1, \ldots, p+q\}$.

Lemma 5.10. *The vector space of all totally skew-symmetric maps*

$$\underbrace{U \times U \times \ldots \times U}_{p \text{ copies}} \to V$$

has dimension $\binom{u}{p} v$ *and has a basis given by* $e_I^* f_j$, *where* I *runs through all the possibilities described above and* $1 \leq j \leq v$.

Proof. First note that $e_I^*(e_J) = \delta_{IJ}$ ($= 1$ if $I = J$ and vanishes otherwise). Now given an arbitrary totally skew map $F: U \times U \ldots \times U \to V$, we can express it in terms of the basis for V as $F = \sum F_i f_i$, where $F_i: U \times U \times \ldots \times U \to \mathbf{R}$ are the various components of F in the basis of V. Clearly, the F_i are themselves totally skew maps. Thus it suffices to consider the case $V = \mathbf{R}$. The independence of the e_I^* is easily seen, for if $\sum a_I e_I^* = 0$, then evaluating on e_J shows $a_J = 0$ for all J. Conversely, evaluating $F - \sum F(e_I) e_I^*$ on any e_J always gives 0, so it follows from the multilinearity of the expression that the e_I^* span the space of totally skew maps. ■

Next we consider p-forms in the case $M = \mathbf{R}^m$. Since $T(M) = \mathbf{R}^m \times \mathbf{R}^m$, we have the canonical identification

$$\underbrace{T(M) \oplus \ldots \oplus T(M)}_{(p \text{ copies})} \approx \mathbf{R}^m \times (\underbrace{\mathbf{R}^m \oplus \ldots \oplus \mathbf{R}^m}_{(p \text{ copies})}).$$

Let us consider the *coordinate* function $x_i: \mathbf{R}^n \to \mathbf{R}$, with

$$dx_i: T(\mathbf{R}^n) \to \mathbf{R},$$
$$(x, v) \to v_i.$$

Now the $dx_i\big|_x$ form a basis for $T_x(\mathbf{R}^n)^*$, and thus

$$dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \ldots \wedge dx_{i_p}$$

(where I is the sequence $1 \leq i_1 < i_2 < \ldots < i_p \leq m$), restricts to $dx_I\big|_x$ to give a basis for the totally skew-symmetric $\mathbf{R}$-valued functions $T_x(\mathbf{R}^n)^{\oplus p} \to \mathbf{R}$. It follows that an $\mathbf{R}$-valued p-form on $\mathbf{R}^m$ has the form $\omega = \sum a_I(x) dx_I$, where the summation takes place over the sequences I described above and the coefficient functions $a_I(x)$ are smooth.

Exercise 5.11. Show that the p-form $\omega = \sum a_I(x) dx_I$ on $\mathbf{R}^m$ is smooth $\Leftrightarrow$ the coefficient functions $a_I(x)$ are all smooth. ❑

Exterior Differentiation of Forms

Now we define the *exterior derivative* $d: A^p(M, V) \to A^{p+1}(M, V)$ as follows. Choose a coordinate chart (U, φ) on M and a basis $f_1, \ldots, f_v$ of V. In terms of this coordinate system, we define

$$d\left(\sum a_{Ij} dx_I f_j\right) = \sum (da_{Ij}) \wedge dx_I f_j.$$

To see that this is well defined, (i.e., independent of the choice of coordinate chart (U, φ) and of the basis of V), we could, by brute force, try to compute the effect of changing these choices. However, it makes better sense to follow the more subtle route of showing that d, defined with respect to a coordinate system as above, satisfies several properties and that these properties characterize it.

Lemma 5.12.

(i) $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$.

(ii) $d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2$, *where* $k = \deg(\omega_1)$.

(iii) $d(d\omega) = 0$, *or, more briefly,* $d^2 = 0$.

Proof. (i) is obvious.

(ii) By virtue of (i), we may assume that $\omega_1 = f\, dx_I e_p$ and $\omega_2 = g\, dx_J h_q$, where $\{e_p\}$ is a basis of V_1 and $\{h_q\}$ is a basis of V_2. Then

$$\begin{aligned}
\omega_1 \wedge \omega_2 &= fg\, dx_I \wedge dx_J e_p \otimes h_q, \quad \text{so} \\
d(\omega_1 \wedge \omega_2) &= d(fg) dx_I \wedge dx_J e_p \otimes h_q \\
&= (g\, df + f\, dg) dx_I \wedge dx_J e_p \otimes h_q \\
&= g\, df \wedge dx_I \wedge dx_J e_p \otimes h_q + f\, dg \wedge dx_I \wedge dx_J e_p \otimes h_q \\
&= g\, df \wedge dx_I \wedge dx_J e_p \otimes h_q + (-1)^k f\, dx_I \wedge dg \wedge dx_J e_p \otimes h_q \\
&= d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2.
\end{aligned}$$

(iii) Again by virtue of (i) we may assume that $\omega = f\, dx_I e_p$. Then

$$d\omega = d(f\, dx_I e_p) = df \wedge dx_I e_p = \sum \frac{\partial f}{\partial x_i} dx_i \wedge dx_I e_p, \quad \text{and so}$$
$$d^2\omega = d\left(\sum \frac{\partial f}{\partial x_i} dx_i \wedge dx_I e_p\right) = \sum \frac{\partial^2 f}{\partial x_i \partial x_j} dx_j \wedge dx_i \wedge dx_I e_p = 0,$$

since the terms $\frac{\partial^2 f}{\partial x_i \partial x_j} dx_j \wedge dx_i$ are skew symmetric in the indices i and j. ■

Lemma 5.13. *If two vector-valued operators on forms satisfy the three properties of the previous lemma and agree on functions, then they are identical.*

Proof. Let d' be another operator on forms defined in an open set of $\mathbf{R}^n$ which satisfies the three properties of the previous lemma. To verify that $d = d'$, it suffices to show that $d\omega = d'\omega$ for $\omega = f\, dx_I e_p$. We have

$$\begin{aligned} d'(f\,dx_I e_p) &= (d'f)dx_I e_p + f\,d'(dx_I e_p) && \text{(by ii)} \\ &= (df)dx_I e_p + f\,d'(dx_I e_p) && \text{since } d' = d \text{ on functions} \\ &= d\omega + f d'(dx_I e_p) && \text{(by i).} \end{aligned}$$

Thus it suffices to show that $d'(dx_I) = 0$, where

$$dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \ldots \wedge dx_{i_p} = d'x_{i_1} \wedge d'x_{i_2} \wedge \ldots \wedge d'x_{i_p}$$

by equality on functions. Thus, $d'(dx_I) = d'(d'x_I) = 0$ inductively by (ii) and (iii). ■

Corollary 5.14. *There is a unique operator $d\colon A^p(M,V) \to A^{p+1}(M,V)$, defined for all p and for all V, that satisfies* (i), (ii), *and* (iii) *and reduces to $df = \sum \frac{\partial f}{\partial x_i} dx_i$ on functions.*

Proof. By the lemma, the ds defined in terms of the various coordinate charts must all agree on the intersections of these coordinate charts. ■

By a procedure like that above, we could also show that our d is the unique operator $d\colon A^p(M) \to A^{p+1}(M)$ satisfying (i), (ii), and (iii) and reducing to $df = \sum \frac{\partial f}{\partial x_i} dx_i$ on functions.

By Proposition 5.3, $df(X) = X(f)$ for V-valued functions on M. There are similar expressions for the derivative $d\omega$ of a p-form ω evaluated on $p+1$ vector fields $X_0, \ldots, X_p$. The most important case for us is the case $p = 1$. We demonstrate this case and leave the rest as an exercise.

Lemma 5.15. *Let ω be a V-valued 1-form on the smooth manifold M, and let X and Y be two vector fields on M. Then $d\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y])$.*

Proof. It suffices to verify this formula for 1-forms of the form $\omega = f\,dg$ since every 1-form is locally expressible as a sum of such terms.

{LHS}: $\quad d\omega = df \wedge dg$ so $d\omega(X,Y) = df(X)dg(Y) - df(Y)dg(X)$

$$= X(f)Y(g) - Y(f)X(g);$$

{RHS}: $\quad X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y])$

$$\begin{aligned} &= X(f\,dg(Y)) - Y(f\,dg(X)) - f\,dg([X,Y]) \\ &= X(fY(g)) - Y(fX(g)) - f[X,Y](g) \\ &= X(f)Y(g) + fXY(g) - Y(f)X(g) - fYX(g) - f[X,Y](g) \\ &= X(f)Y(g) + f[X,Y](g) - Y(f)X(g) - f[X,Y](g) \\ &= X(f)Y(g) - Y(f)X(g). \end{aligned}$$

Hence, {LHS} = {RHS}. ■

Exercise 5.16.* Show that if ω is a p-form, then

$$\begin{aligned} d\omega(X_0, \ldots, X_p) = &\sum_{0 \le i \le p} (-1)^i X_i(\omega(X_0, \ldots, \hat{X}_i, \ldots, X_p) \\ &+ \sum_{0 \le i < j \le p} (-1)^{i+j}(\omega([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_p) \end{aligned}$$

where a "hat" means "omit this entry." ❑

Proposition 5.17 (Naturality of d). *If $f\colon M \to N$ is a smooth map, then the following diagram commutes for all p.*

$$\begin{array}{ccc} A^p(N,V) & \xrightarrow{d} & A^{p+1}(N,V) \\ \downarrow{\scriptstyle f^*} & & \downarrow{\scriptstyle f^*} \\ A^p(M,V) & \xrightarrow{d} & A^{p+1}(M,V) \end{array}$$

Proof. It suffices to check the result for maps between Euclidean spaces. The proof is by induction on p. By linearity we may assume the p-form is $\omega = g\,dx_I$. For $p = 0$ we have

$$f^*(dg)(X) = dg(f_*X) = (f_*X)g = X(f^*(g)) = X(gf) = d(gf)(X),$$

and so $f^*(dg) = d(f^*(g))$.[19] Now assume the result holds in dimension $<p$. Writing $dx_I = dx_J \wedge dx_K$, with $|I| = p$ and $|J|$, $|K| < p$, we have

$$\begin{aligned} d(f^*\omega) &= d(f^*(g\,dx_J \wedge dx_K)) \\ &= d(f^*(g\,dx_J) \wedge f^*(dx_K)) \\ &= d(f^*(g\,dx_J)) \wedge f^*(dx_K) \pm f^*(g\,dx_J) \wedge d(f^*(dx_K)) \\ &= (f^*d(g\,dx_J)) \wedge f^*(dx_K) \pm f^*(g\,dx_J) \wedge f^*(d(dx_K)) \text{ (by induction)} \\ &= f^*(d(g\,dx_J) \wedge dx_K) \pm 0 \\ &= f^*(d(g\,dx_J \wedge dx_K)) \\ &= f^*(d\omega). \end{aligned}$$ ■

We remark that the naturality of the exterior derivative is perhaps its most important property. It says that d is a version of differentiation that is *independent of the coordinate system used.* Thus, it constitutes a basic ingredient in the search for "a theory of differential equations that is independent of the coordinate system used to describe them" (cf. the Introduction to Chapter 8). In fact, every differential equation may be expressed in terms of the operator d (cf., e.g., [S.S. Chern and C. Chevalley, 1952]).

[19] Note that $f^*(g) = gf$ (composition).

Change of Coefficients

If $\varphi\colon V_1 \to V_2$ is a linear map of vector spaces,[20] then it induces a linear map
$$\varphi_*\colon A^p(M, V_1) \to A^p(M, V_2)$$
defined by $\varphi_*(\omega)(X_1, \dots, X_p) = \varphi_*(\omega(X_1, \dots, X_p))$. It is clear that such a coefficient homomorphism always commutes with d.

Coefficient Multiplication

In the case that the coefficient vector space V has a multiplication given by $m\colon V \otimes V \to V$ (i.e., a bilinear but not necessarily commutative or associative map), then we obtain a corresponding map
$$\begin{aligned} m_*\colon A(M, V \otimes V) &\to A(M, V) \\ \omega &\mapsto m_*(\omega), \end{aligned}$$
where $(m_*(\omega))(X_1, \dots, X_p) = m(\omega(X_1, \dots, X_p))$.

This map turns $A(M, V)$ into an algebra with multiplication given by
$$A^p(M, V) \times A^q(M, V) \xrightarrow{\wedge} A^{p+q}(M, V \otimes V) \xrightarrow{m_*} A^{p+q}(M, V).$$

Exercise 5.18. $d(m_*(\omega_1 \wedge \omega_2)) = m_*(d\omega_1 \wedge \omega_2) + (-1)^p m_*(\omega_1 \wedge d\omega_2)$. ❑

Let us now consider the case where the coefficients consist of a graded associative algebra $\mathcal{B} = \bigoplus_{r \geq 0} \mathcal{B}_r$. As above we denote the multiplication in $\mathcal{B}$ by $m\colon \mathcal{B} \otimes \mathcal{B} \to \mathcal{B}$. Then the multiplication on forms is
$$A^p(M, \mathcal{B}_r) \times A^q(M, \mathcal{B}_s) \xrightarrow{\wedge} A^{p+q}(M, \mathcal{B}_r \otimes \mathcal{B}_s)) \xrightarrow{m_*} A^{p+q}(M, \mathcal{B}_{r+s})$$
and yields a *bigraded associative differential algebra*, as the following exercise will show.

Exercise 5.19. Let $\mathcal{B}$ be as above.

(i) Show that $(dm_*(\omega_1 \wedge \omega_2)) = m_*(d\omega_1 \wedge \omega_2) + (-1)^p m_*(\omega_1 \wedge d\omega_2)$.

(ii) Show that $m_*(m_*(\omega_1 \wedge \omega_2) \wedge \omega_3) = m_*(\omega_1 \wedge m_*(\omega_2 \wedge \omega_3))$.

(iii) If $\mathcal{B}$ is *graded commutative* (e.g., $\mathcal{B}$ is an exterior algebra or a Clifford algebra), show that $m_*(\omega_{p,r} \wedge \omega_{q,s}) = (-1)^{pr+qs} m_*(\omega_{q,s} \wedge \omega_{p,r})$, where $\omega_{p,r} \in A^p(M, \mathcal{B}_r)$, etc. ❑

[20] This may also be a morphism of flat vector bundles.

A particularly interesting case of coefficient multiplication occurs when $V = \mathfrak{g}$, a Lie algebra.[21] We denote the product here by $(\omega_1, \omega_2) \to [\omega_1, \omega_2]$. In this case we obtain a *differential graded Lie algebra* defined by the properties given in the following exercise.

Exercise 5.20.* Let $\omega_p, \omega_q, \omega_r \in A(M, \mathfrak{g})$ be forms of dimension p, q, and r, respectively. Show that the multiplication $[\ ,\ ]$ on $A(M, \mathfrak{g})$ satisfies the formulas

(i) $d[\omega_p, \omega_q] = [d\omega_p, \omega_q] + (-1)^p [\omega_p, d\omega_q]$,

(ii) $[\omega_q, \omega_p] = (-1)^{pq+1} [\omega_p, \omega_q]$,

(iii) $(-1)^{rp}[[\omega_p, \omega_q], \omega_r] + (-1)^{pq}[[\omega_q, \omega_r], \omega_p] + (-1)^{qr}[[\omega_r, \omega_p], \omega_q] = 0$. ❑

Note that if $\omega \in A^1(M, \mathfrak{g})$, then the mapping $T(M) \oplus T(M) \to \mathfrak{g}$ sending $(X, Y) \to [\omega(X), \omega(Y)]$ is skew symmetric and is therefore an element of $A^2(M, \mathfrak{g})$. The relation between $[\omega(X), \omega(Y)]$ and $[\omega, \omega](X, Y)$ is given by the following lemma.

Lemma 5.21.
$$[\omega_1, \omega_2](X, Y) = [\omega_1(X), \omega_2(Y)] + [\omega_2(X), \omega_1(Y)].$$
In particular,
$$[\omega(X), \omega(Y)] = \frac{1}{2}[\omega, \omega](X, Y).$$

Proof.
$$\begin{aligned} [\omega_1, \omega_2](X, Y) &= (m_*(\omega_1 \wedge \omega_2))(X, Y) \\ &= m((\omega_1 \wedge \omega_2)(X, Y)) \\ &= m(\omega_1(X) \otimes \omega_2(Y) - \omega_1(Y) \otimes \omega_2(X)) \\ &= [\omega_1(X), \omega_2(Y)] - [\omega_1(Y), \omega_2(X)] \\ &= [\omega_1(X), \omega_2(Y)] + [\omega_2(X), \omega_1(Y)] \end{aligned}$$
■

The proof of the following result has already been shown, but we state it explicitly for future reference.

Proposition 5.22. *Let $f\colon M \to N$ be a smooth map and let $\varphi\colon \mathfrak{h} \to \mathfrak{g}$ be a Lie algebra homomorphism. Then the map $\varphi_* f^*\colon A^*(N, \mathfrak{h}) \to A^*(M, \mathfrak{g})$ is a homomorphism of Lie algebras.*

It is the study of the deep meaning associated with the $\mathfrak{g}$-valued 1-forms and 2-forms on M—which may be regarded as elements of the graded Lie algebra $A^*(M, \mathfrak{g})$—which is the subject of this book.

[21] See Chapter 3 for the formal definition.

Basic and Semibasic Forms

The terms *basic* and *semibasic* have to do with forms on the total space of a bundle. They refer to how these forms are related to the fiber and base directions. Let $F \to E \xrightarrow{\pi} M$ be a fiber bundle over the manifold M. Then $\pi^*: A^p(M) \subset A^p(E)$.

Definition 5.23. Let $\omega \in A^p(E)$.

(i) ω is *basic* if it lies in the image of π^*.

(ii) ω is *semibasic* if $\omega(v_1, \dots, v_n) = 0$ whenever v_1 is tangent to a fiber. ❀

Lemma 5.24.

(i) *Basic forms are semibasic.*

(ii) *ω is semibasic $\Leftrightarrow$ each $p \in E$ has a neighborhood U on which there are basic forms ω_i such that $\omega = \sum a_i\omega_i$ for some functions $a_i: U \to \mathbf{R}$.*

Proof. (i) If ω is basic, then $\omega = \pi^*\eta$ for some form η on M. If v_1 is tangent to a fiber, then $\pi_* v_1 = 0$, so $\omega(v_1, \dots, v_n) = \pi^*\eta(v_1, \dots, v_n) = \eta(\pi_* v_1, \dots, \pi_* v_n) = 0$.

(ii) $\Leftarrow$: Since the forms ω_i are basic, they are also semibasic and their combination $\omega = \sum a_i\omega_i$ is also semibasic.

$\Rightarrow$: Fix $p \in E$. Choose a local trivialization ψ of E over a coordinate neighborhood (V, x) around $\pi(p) \in M$ so that $\psi: \pi^{-1}(V) \approx V \times F$. Write $p = (p_M, p_F)$ in this local trivialization. Then choose a coordinate neighborhood (W, y) around $p_F \in F$. Then $(x \circ \psi, y \circ \psi)$ is a local coordinate system on a neighborhood $\psi^{-1}(V \times W)$ of p. Abbreviate $(x \circ \psi, y \circ \psi)$ by (x, y). Write the semibasic p-form ω as $\omega = \sum a_{IJ} dx_I \wedge dy_J$. Let b_I and e_J be the bases dual to dx_I and dy_J. Since ω is semibasic, $a_{KL} = \omega(b_K \wedge e_L) = 0$ if $L \neq \emptyset$. Thus, $\omega = \sum a_{I\emptyset} dx_I$. ■

Lemma 5.25. *Let $H \to P \to M$ be a principal H bundle with right action $R_h: P \to P$, $h \in H$. Let η be a p-form on P. Then*

$$\omega \text{ is basic} \Leftrightarrow \omega \text{ is semibasic and right } H \text{ invariant.}$$

Proof. $\Rightarrow$: By Lemma 5.24(i), we need only show that ω is basic $\Rightarrow$ ω right H invariant. But ω is basic $\Rightarrow \omega = \pi^*\eta$ for some form η on $M \Rightarrow R_h^*\omega = R_h^*\pi^*\eta = (\pi R_h)^*\eta = \pi^*\eta = \omega$ (cf. Definition 3.19).

$\Leftarrow$: By Lemma 5.24(ii), we may write, locally, $\omega = \sum a_i\omega_i$ for some basic forms ω_i and functions $a_i: U \to \mathbf{R}$. We may assume that the ω_i are independent at each point. Now the H invariance implies that

$$\sum a_i\omega_i = R_h^*\left(\sum a_i\omega_i\right) = \sum R_h^* a_i R_h^*\omega_i = \sum (R_h^* a_i)\omega_i \Rightarrow R_h^* a_i = a_i.$$

Thus, $a_i(p) = a_i(ph)$, so the coefficients a_i are constant along the fibers and so are basic functions. Thus, ω itself is basic. ■

Some Lie Groups

Here is a description of some of the most common Lie groups. These examples are linear groups, which means that each occurs as a subgroup of $Gl_n(\mathbf{R})$ which we studied in Exercise 1.19. We also list their tangent spaces at the identity expressed as subspaces of the vector space $M_n(\mathbf{R})$. This means that these tangent spaces have been translated so as to pass through the origin. They are named by the corresponding Gothic letters.

The Positive General Linear Group (the identity component of $Gl_n(\mathbf{R})$

$$Gl_n^+(\mathbf{R}) = \{A \in Gl_n(\mathbf{R}) \mid \det A > 0\},$$
$$\mathfrak{gl}_n^+(\mathbf{R}) = M_n(\mathbf{R}).$$

The Projective General Linear Group

$$PGl_n(\mathbf{R}) = Gl_n(\mathbf{R})/\{\lambda I \in Gl_n(\mathbf{R}) \mid \lambda \in \mathbf{R}\},$$
$$\mathfrak{pgl}_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid \operatorname{Trace} A = 0\}.$$

The Special Linear Group

$$Sl_n(\mathbf{R}) = \{A \in Gl_n(\mathbf{R}) \mid \det A = 1\},$$
$$\mathfrak{sl}_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid \operatorname{Trace} A = 0\}.$$

The Projective Special Linear Group

$$PSl_n(\mathbf{R}) = Sl_n(\mathbf{R})/\{\lambda I \in Sl_n(\mathbf{R}) \mid \lambda \in \mathbf{R}\},$$
$$\mathfrak{psl}_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid \operatorname{Trace} A = 0\}.$$

The Orthogonal Group

$$O_n(\mathbf{R}) = \{A \in Gl_n(\mathbf{R}) \mid AA^t = 1\},$$
$$\mathfrak{o}_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid A + A^t = 0\}.$$

The Orthogonal Group of signature p, q

$$O_{p,q}(\mathbf{R}) = \left\{A \in Gl_{p+q}(\mathbf{R}) \mid A\Sigma_{p,q}A^t = \Sigma_{p,q}\right\},$$
$$\mathfrak{o}_{p,q}(\mathbf{R}) = \left\{A \in M_{p+q}(\mathbf{R}) \mid A\Sigma_{p,q} + \Sigma_{p,q}A^t = 0\right\},$$

where $\Sigma_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$.

The Special Orthogonal Group

$$SO_n(\mathbf{R}) = \{A \in Gl_n(\mathbf{R}) \mid AA^t = 1, \det A = 1\},$$
$$\mathfrak{so}_n(\mathbf{R}) = \{A \in M_n(\mathbf{R}) \mid A + A^t = 0\} = \mathfrak{o}_n(\mathbf{R}).$$

The Special Orthogonal Group of signature p, q

$$SO_{p,q}(\mathbf{R}) = \{A \in Gl_{p+q}(\mathbf{R}) \mid A\Sigma_{p,q}A^t = \Sigma_{p,q}, \det A = 1\},$$
$$\mathfrak{so}_{p,q}(\mathbf{R}) = \{A \in M_{p+q}(\mathbf{R}) \mid A\Sigma_{p,q} + \Sigma_{p,q}A^t = 0\},$$

where $\Sigma_{p,q} = \begin{pmatrix} I_p & 0 \\ 0 & -I_q \end{pmatrix}$.

The Lorentz Group

$$L_{n,1}(\mathbf{R}) = \{A \in Gl_{n+1}(\mathbf{R}) \mid A^t\Sigma A = \Sigma\},$$
$$\mathfrak{l}_{n,1}(\mathbf{R}) = \{A \in M_{n+1}(\mathbf{R}) \mid A^t\Sigma + \Sigma A = 0\},$$

where $\Sigma = \begin{pmatrix} 0 & 0 & -1 \\ 0 & I_{n-1} & 0 \\ -1 & 0 & 0 \end{pmatrix}$.

The Unitary Group

$$\begin{aligned} U_n(\mathbf{C}) &= \{A \in Gl_n(\mathbf{C}) \mid AA^* = I_n\}, \\ \mathfrak{u}_n(\mathbf{C}) &= \{A \in M_n(\mathbf{C}) \mid A + A^* = 0\}, \end{aligned} \quad \text{where } A^* = \bar{A}^t.$$

The Special Unitary Group

$$SU_n(\mathbf{C}) = \{A \in Gl_n(\mathbf{C}) \mid AA^* = I_n,\ \det A = 1\},$$
$$\mathfrak{su}_n(\mathbf{C}) = \{A \in M_n(\mathbf{C}) \mid A + A^* = 0,\ \operatorname{Trace} A = 0\}.$$

The Euclidean Group

$$Euc_n(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in Gl_{n+1}(\mathbf{R}) \mid v \in \mathbf{R}^n, A \in SO_n(\mathbf{R}) \right\},$$
$$\mathfrak{euc}_n(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in M_{n+1}(\mathbf{R}) \mid v \in \mathbf{R}^n, A \in \mathfrak{so}_n(\mathbf{R}) \right\}.$$

The Positive Affine Group

$$Aff_n^+(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in Gl_{n+1}^+(\mathbf{R}) \mid v \in \mathbf{R}^n, A \in Gl_n^+(\mathbf{R}) \right\},$$
$$\mathfrak{aff}_n^+(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in M_{n+1}(\mathbf{R}) \mid v \in \mathbf{R}^n, A \in \mathfrak{gl}_n(\mathbf{R}) \right\}.$$

Exercise 5.26. (i) Verify that each of these Lie groups is in fact a group and that each has the corresponding tangent space at the identity.

(ii) Show that $L_{n,1}(\mathbf{R}) \approx O_{n,1}(\mathbf{R})$. ❑

2

Looking for the Forest in the Leaves: Foliations

> *... leaves, leaves overhead and underfoot and in your face and in your eyes, endless leaves on endless trees.*
>
> –Ursula K. Le Guin, 1977

A foliation is, roughly speaking, a decomposition of a manifold M into p-dimensional submanifolds (the leaves of the foliation) that lie neatly side by side.[1] The simplest case ($p = 1$) consists of the integral curves of a nonzero vector field, or a line field. This case is studied in §1. The higher-dimensional analog of a line field is a p-dimensional distribution, that is, a p-dimensional subspace in each tangent space fitting together smoothly with each other. These are studied in §2. Here, however, in contrast to the vector field case, when $p > 1$ we cannot always find leaves tangent to the distribution at every point. There are integrability conditions that must be satisfied in order for this to be possible. These are studied in §3, and in §4 we show that they are sufficient for the existence of the leaves. In §5 we reformulate both the notion of distribution and the integrability conditions in terms of vector-valued differential forms. In §6 we give the modern definition of foliations, and in §7 we show that the leaves near a point on a given leaf are distributed in the same fashion no matter which point on the leaf one chooses. This leads us to introduce leaf holonomy in §8, which we use to study the space of leaves of a foliation, showing in particular that if all the leaves are compact and have trivial holonomy, then

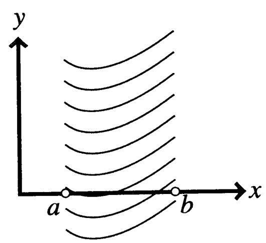
[1] For the formal definition, see §6.

they are all diffeomorphic and form the fibers of a smooth fiber bundle. We remark that we do not make use of §7 and §8 in the rest of the book.

Although the theory of foliations may appear at first sight to be a rather abstract subject, as we shall see it is a natural and useful generalization of the ideas of elementary calculus. In this chapter we present only a brief introduction to this subject which has undergone a vigorous development in the last two decades. For more information, we recommend the excellent book by Pierre Molino ([P. Molino, 1988]).

§1. Integral Curves

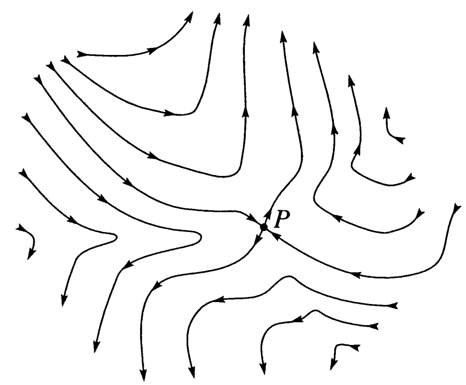
In this section we study a special case of how a submanifold may be obtained by specifying its tangent planes. The original example of this sort of thing is the *fundamental theorem of calculus*, which says that given the *slope* of a graph in $\mathbf{R}^2$ over each point of an interval $I = (a, b)$ on the x-axis, there is a unique graph, or integral curve, passing through each point over I. Here is a picture of some of these graphs for a given slope function.



On a general smooth manifold M, we may again obtain integral curves by arbitrarily prescribing smoothly varying tangent lines. But, in fact, in this section we will deal rather with vector fields on M (which do determine line fields at points of M where they do not vanish) and these yield *parameterized* integral curves. Parameterization of a leaf really only makes sense when the leaf has dimension one, so the generalization (in §2) of the work in this section to foliations consisting of higher-dimensional leaves will of necessity be only partial (but still very interesting). Given a smooth vector field X on a smooth manifold M, an integral curve for X through p is a parametrized curve $\sigma\colon ((-\varepsilon,\varepsilon),0) \to (M,p)$ whose tangent at the point $\sigma(t)$ is the vector $X_{\sigma(t)}$, that is,

$$X_{\sigma(t)} = \dot{\sigma}(t) \stackrel{\text{def}}{=} \sigma_{*t}\left(\frac{\partial}{\partial t}\right).$$

In the following figure we picture some of the integral curves of a vector field in $\mathbf{R}^2$. In this picture there is one point P (a critical point) where the vector field vanishes.



In this particular case, this point is the end of two integral curves and the beginning of two others. The parametrized integral curves "slow down" as they approach such a point and "speed up" as they leave it. Of course, a particle moving under the influence of such a velocity field will never pass through a critical point even if it were to reach such a point in finite time (which cannot happen for a smooth vector field; cf. Exercise 1.18. Its velocity would be zero there, and so it would sit there forever.

Our first aim in this section is to show that, locally, a smooth vector field always has an integral curve through any point p where it doesn't vanish. In fact, Theorem 1.2 on the linearization of vector fields given below shows even more: it shows that around any point $p \in M$, where $X_p \neq 0$, there is a local coordinate system (U, φ) such that the integral curves near p are given by $(\varphi_2, \ldots, \varphi_n) =$ constant. This means that locally, up to diffeomorphism, the integral curves are arranged exactly as the family of lines parallel to the x_1 axis in $\mathbf{R}^n$. To show this we need the following local existence theorem from the ordinary differential equations (cf., e.g., [L. Loomis and S. Sternberg, 1968], pp. 266–275).

Theorem 1.1. *Let $f(t,x) = (f_1(t,x), \ldots, f_n(t,x))$ be a smooth, $\mathbf{R}^n$-valued function defined on some open set $J \times U$ of $\mathbf{R} \times \mathbf{R}^n$, where $0 \in J$. Consider the system of equations for an unknown function $g = (g_1, \ldots, g_n)\colon \mathbf{R} \to U$ given by*

$$\begin{aligned} g_1'(t) &= f_1(t, g_1(t), \ldots, g_n(t)), \\ g_2'(t) &= f_2(t, g_1(t), \ldots, g_n(t)), \\ &\vdots \\ g_n'(t) &= f_n(t, g_1(t), \ldots, g_n(t)), \end{aligned}$$

with initial conditions $g_1(0) = c_1$, $g_2(0) = c_2, \ldots, g_n(0) = c_n$.

(i) (Existence and uniqueness). *For any $c = (c_1, \ldots, c_n) \in U$, there is an $\varepsilon > 0$ such that this system has a unique smooth solution $g\colon (-\varepsilon,\varepsilon) \to U$.*

(ii) (Smoothness in initial conditions). *For any $p \in U$, there is an open set V with $p \in V \subset U$, an $\varepsilon > 0$, and a smooth map $g\colon (-\varepsilon, \varepsilon) \times V \to U$ such that $g(\cdot, c)\colon (-\varepsilon, \varepsilon) \to U$ is the unique smooth solution of the system with initial condition c for any $c \in V$.*

Theorem 1.2 (Linearization of Vector Fields). *If X is a vector field defined on a smooth n-manifold M, then for each $p \in M$ where $X_p \neq 0$ it is possible to find a local coordinate system (U, φ) around p such that U is an open set of the form $[-\varepsilon, \varepsilon] \times [-\varepsilon, \varepsilon] \times \ldots \times [-\varepsilon, \varepsilon]$ (a "cube") in these coordinates, p is at the center of the cube, and $\varphi_*(X) = \partial/\partial x_1$.*[2]

Proof. Since this is a local result, we may assume that M is an open set $U \subset \mathbf{R}^n$, $p = 0$, X is never zero on U, and, after a change of basis in $\mathbf{R}^n$, $X_0 = e_1$. Write $X = \sum f_j(x) e_j$, where $f_j\colon U \to \mathbf{R}$ for each j. Now let $g\colon \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$ (defined on some neighborhood of $(0, c)$) be the solution of the system $(\partial g/\partial t)(t, c) = f(g(t, c))$, with initial condition $g(0, c) = c$ whose existence is guaranteed by Theorem 1.1. φ is going to be the (suitably restricted) inverse of

$$h(x_1, \ldots, x_n) = g(x_1, 0, x_2, \ldots, x_n).$$

In terms of h, the conditions on g become

$$\partial h/\partial x_1 = f(h) \quad \text{and} \quad h(0, x_2, \ldots, x_n) = (0, x_2, \ldots, x_n),$$

namely, $h_*(e_1) = X$ at every point and $h_*(e_j) = e_j$ $(2 \leq j \leq n)$ at the origin. Thus, the Jacobian matrix for h is the identity at the origin, and so h is a local diffeomorphism on some cube centered at the origin. The inverse φ of h has the required properties. ■

Exercise 1.3 (Application to PDE's). Consider the partial differential equation

$$a_1(x)\frac{\partial f}{\partial x_1} + a_2(x)\frac{\partial f}{\partial x_2} + \cdots + a_n(x)\frac{\partial f}{\partial x_n} = b(x),$$

where $x \in U$, an open set in $\mathbf{R}^n$, $a(x) = (a_1(x), \ldots, a_n(x))$ and $b(x)$ are smooth functions, and $a(x)$ is never zero on U. Show that about each point of U there is a coordinate system $y = (y_1, \ldots, y_n)$ such that in the y-coordinates the partial differential equation assumes the form

$$\frac{\partial f}{\partial y_1} = c(y).$$

Deduce the general solution of this equation. ❑

[2] These coordinates are sometimes called "flow box" coordinates.

Complete Vector Fields

If $\sigma_j\colon ((a_j, b_j), c_j) \to (M, p)$ $(j = 1, 2)$ are two integral curves through p, then after a translation of the parameter for σ_2 (say), we may assume that $c_1 = c_2$. Then by the uniqueness part of the theorem on differential equations, we see that $\sigma_1 = \sigma_2$ on $(a_1, b_1) \cap (a_2, b_2)$. Thus, we can define σ on $(a_1, b_1) \cup (a_2, b_2)$ to be σ_j on (a_j, b_j). Applying Zorn's lemma, we see that a maximal integral curve through a point p always exists.

There are four possible cases for a maximal integral curve σ. Up to a translation of parameters, it can only have domain of one of four types:

$$(0, a), \quad (0, \infty), \quad (-\infty, 0), \quad \text{or} \quad (-\infty, \infty).$$

Definition 1.4. An integral curve is *complete* if its domain is of the last type. A vector field is *complete* if all of its integral curves are complete. ❀

Exercise 1.5. Find four never-zero vector fields on the interval $(0, 1)$ whose integral curves exhibit the four possible types of domains. ❑

Proposition 1.6. *A vector field with compact support*[3] *is complete.*

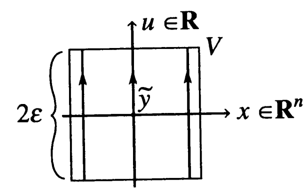
Proof. Suppose that X is a vector field on M with compact support. Let $\sigma(t)$ be a maximal integral curve of X, with domain (a, b). Let us assume that $b < \infty$ and deduce a contradiction. (Showing that $a = -\infty$ is handled in a similar manner.)

If X vanishes at a point $\sigma(t_0)$, then clearly $\sigma(t) = \sigma(t_0)$ for all $t \geq t_0$, and so $b = \infty$. Thus, we may assume that $X_{\sigma(t)}$ is never zero for $t \in (a, b)$. It follows that σ is a curve on the compact support of X, and hence there is a sequence $t_k \to b$ such that $\sigma(t_k)$ converges to a cluster point, y say, of the set $\{\sigma(t) \mid t \in (a, b)\}$. Now we pass to the product manifold $M \times \mathbf{R}$, and consider the vector field $Y = (X, \partial/\partial t)$ on it. We easily check that $\tau(t) = (\sigma(t), t)$, $t \in (a, b)$, is an integral curve for Y and that the projection on M of any integral curve for Y is an integral curve for X. Now Y is a never-zero field, so the linearization of vector fields (Theorem 1.1) applies to it, yielding

- a neighborhood V of the point $\tilde{y} = (y, b)$,
- flowbox coordinates $(x, u) \in \mathbf{R}^n \times \mathbf{R}$ for V in which
  - V is a cube of side 2ε,
  - $\tilde{y}$ corresponds to the point (0,b) (in flowbox coordinates),

[3] The *support* of a vector field on M is the closure of the subset of points of M where X does not vanish.

- $Y = \partial/\partial u$,
- the integral curves of Y have the form $x \times (-\varepsilon, \varepsilon)$ (in flowbox coordinates).



Since $\lim_{k\to\infty} \tau(t_k) = \tilde{y}$ the integral curve τ must meet V, and in fact it must meet the integral curve $0 \times (-\varepsilon, \varepsilon)$ (given in flowbox coordinates). It follows that the integral curve τ (and hence also the integral curve σ) is defined in the interval $(a, b+\varepsilon)$ which shows that b is not maximal, a contradiction to the finiteness of b. ■

Proposition 1.7. *Let M be a proper submanifold of N. Let X be a complete vector field on N such that for every $x \in M$, $X_x \in T_x(M)$. Then $X|_M$ is also complete.*

Proof. Let $\sigma: ((-\varepsilon, \varepsilon), 0) \to (M, p)$ be an integral curve on M for $X|_M$ through p. Then it is also an integral curve on N for X through $p \in N$. But since X is complete on N, σ extends to an integral curve $\sigma: (\mathbf{R}, 0) \to (N, p)$ on N for X through $p \in N$. If $\sigma(t)$ still lies on M for all $t \in \mathbf{R}$, then we're done. Therefore we may suppose that this fails for some $t > 0$ (reversing the sign of X if necessary). Set $t_0 = \sup\{t \in \mathbf{R} \mid \sigma([0,t]) \subset M\}$. If $\sigma(t_0) \in M$, then we have a contradiction, since then, by Theorem 1.1(i), for some $\delta > 0$ we must have $\sigma((t_0 - \varepsilon, t_0 + \varepsilon)) \subset M$ which conflicts with the maximality of t_0. If $\sigma(t_0) \notin M$, then the sequence $x_k = \sigma(t_0(1 - (1/n)))$ is a sequence of points on M which converges on N but not on M. But M is proper so this is also a contradiction. ■

Exercise 1.8. Show by example that the condition that the submanifold be proper is necessary for the conclusion to hold. ❑

Exercise 1.9. Show that the vector fields on $\mathbf{R}^2$ given by $X = y^2(\partial/\partial x)$, $Y = x^2(\partial/\partial y)$ are complete, but that $X + Y$ is not complete. ❑

Exercise 1.10. Show that *any* smooth vector field on a smooth manifold M is the sum of two complete vector fields. [*Hint*: you may assume that the constant function 1 on M may be written as the sum of two nonnegative smooth functions $1 = p_1 + p_2$ with the property that the support of each p_i $(i = 1, 2)$ is a disjoint union of compact sets.] ❑

One-Parameter Groups of Diffeomorphisms

Let us consider a smooth map $\varphi: \mathbf{R} \times M \to M$. For convenience we shall often abbreviate $\varphi(t, x)$ as $\varphi_t(x)$.

Definition 1.11. φ is a *one-parameter group of diffeomorphisms* (or a *flow*) if

(1) $\varphi_0(x) = x$ for all $x \in M$,

(2) $\varphi_s(\varphi_t(x)) = \varphi_{s+t}(x)$ for all $s, t \in \mathbf{R}$, and all $x \in M$. ❀

We remark that

(i) $\varphi_s(\varphi_{-s}(x)) = \varphi_0(x) = x$ so that each φ_s is in fact a diffeomorphism.

(ii) The map $\mathbf{R} \to \mathrm{Diff}(M)$ which sends $t \to \varphi_t$ is a group homomorphism.

Lemma 1.12. *Suppose that $\{U_\alpha\}$ is an open cover of M and $\varepsilon_\alpha > 0$.*

(i) *A one-parameter group is determined by its restriction to the sets $(-\varepsilon_\alpha, \varepsilon_\alpha) \times U_\alpha$,*

(ii) *Suppose that we are given smooth maps $\varphi_\alpha: (-\varepsilon_\alpha, \varepsilon_\alpha) \times U_\alpha \to M$ such that* 1.11(1) *holds and* 1.11(2) *holds where it makes sense, and that these maps agree along the overlaps of the domains. Then if either the cover $\{U_\alpha\}$ is finite or $\inf \varepsilon_\alpha > 0$ these maps are obtained as the restrictions of a unique one-parameter group of diffeomorphisms.*

Proof. (i) Suppose that φ and θ are two one-parameter families with the same restrictions to each $(-\varepsilon_\alpha, \varepsilon_\alpha) \times U_\alpha$. We claim that for each $x \in M$ we have $\varphi_t(x) = \theta_t(x)$ for all t. Clearly this is true for $t \in (-\varepsilon_\alpha, \varepsilon_\alpha)$, where $x \in U_\alpha$. Moreover the set $V_x = \{t \in \mathbf{R} \mid \varphi_t(x) = \theta_t(x)\}$ is closed. But if $t \in V_x$, then we have $\varphi_t(x) = \theta_t(x) \in U_\beta$ for some β. We then have, for all $s \in (-\varepsilon_\beta, \varepsilon_\beta)$,

$$\varphi_{s+t}(x) = \varphi_s(\varphi_t(x)) = \theta_s(\theta_t(x)) = \theta_{s+t}(x).$$

Thus V_x is open. Since $0 \in V_x$, V_x is not empty and hence $V_x = \mathbf{R}$.

(ii) Since a finite cover implies $\varepsilon = \inf \varepsilon_\alpha > 0$ we assume the latter. This implies that we have a well-defined smooth map $\varphi: (-\varepsilon, \varepsilon) \times M \to M$ defined by $\varphi|(-\varepsilon, \varepsilon) \times U_\alpha = \varphi_\alpha|(-\varepsilon, \varepsilon) \times U_\alpha$. Then we define $\varphi: \mathbf{R} \times M \to M$ by $\varphi_s(x) = \varphi^n_{s/n}(x)$ (the n-fold composition) for any $n > |s/\varepsilon|$.

Claim 1. φ is well defined. If $n > |s/\varepsilon|$ and k is a positive integer then $\varphi^{kn}_{s/kn}(x) = (\varphi^k_{s/kn})^n(x) = (\varphi_{s/n})^n(x)$. In particular, if m and n are both $> |s/\varepsilon|$, then $\varphi^n_{s/n}(x) = \varphi^{nm}_{s/nm}(x) = \varphi^m_{s/m}(x)$.

Claim 2. φ is smooth. Since $\varphi_s(x) = \varphi^n_{s/n}(x)$ it suffices to note that $\varphi_{s/n}(x)$ is smooth.

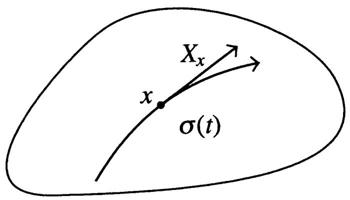
Claim 3. φ is a one-parameter group of diffeomorphisms,

$$\begin{aligned}\varphi_{s+t}(x) &= (\varphi_{(s+t)/n})^n(x) = (\varphi_{s/n}\varphi_{t/n})^n(x)\\ &= (\varphi_{s/n})^n(\varphi_{t/n})^n(x) \quad \text{(Since we can reorder for small parameters)}\\ &= \varphi_s(\varphi_t(x)).\end{aligned}$$

Claim 4. φ restricts to φ_α on $(-\varepsilon_\alpha, \varepsilon_\alpha) \times U_\alpha$. By definition φ restricts to φ_α on $(-\varepsilon, \varepsilon) \times U_\alpha$. But then the argument of (i) above shows φ restricts to φ_α on $(-\varepsilon_\alpha, \varepsilon_\alpha) \times U_\alpha$. ■

The Generator of a One-Parameter Group of Diffeomorphisms

Let us fix $x \in M$ and put $\sigma(t) = \varphi_t(x)$ so that $\sigma\colon (\mathbf{R}, 0) \to (M, x)$ is a curve on M. Now the tangent vector to this curve at x is $\sigma_*(D_0)$, where $D_s = \frac{\partial}{\partial t}\big|_{t=s}$. We set $X_x = \sigma_*(D_0)$.



Then X is a smooth vector field on M, called the *infinitesimal generator of* φ. It tells how each point begins to move under the action of the motion φ_t for "infinitesimal" t.

Exercise 1.13. Verify that the infinitesimal generator is a smooth vector field. [*Hint:* Show that $X_x(f) = (\partial/\partial t)f(\varphi(t,x))\big|_{t=0}$.] ❑

By Lemma 1.12 we see that for any $\varepsilon > 0$, knowledge of φ_t for $|t| < \varepsilon$ suffices to determine φ_t for all $t \in \mathbf{R}$. This may make it plausible that a knowledge of φ_t for t *infinitesimal* (i.e., knowledge of the infinitesimal generator) is enough to determine the flow φ_t. We develop this theme in the following proposition and its corollary.

Proposition 1.14. *Let φ_t be a flow with the vector field X as its infinitesimal generator. The curves $\sigma(t) = \varphi_t(x)$ are integral curves for the vector field X.*

Proof. We must show that $X_{\sigma(s)} = \sigma_*(D_s)$ for all s. Fix t and set $\sigma(s) = \varphi(s,x)$ and $\sigma_1(s) = \varphi(s, \varphi(t,x))$, which is $\varphi(s+t, x)$ by property (2) of Definition 1.11. Hence $\sigma_1(s) = \sigma(s+t)$. Thus,

$$X_{\sigma(t)} = \frac{\partial \sigma_1}{\partial s}\bigg|_{s=0} = \frac{\partial \sigma}{\partial s}\bigg|_{s=t} = \sigma_*(D_t). \qquad ■$$

Corollary 1.15. *Two one-parameter families of diffeomorphism with the same infinitesimal generator are equal.*

Proof. The vector field determines the integral curves, which determine the one-parameter families of diffeomorphisms. ■

Proposition 1.14 and its corollary imply that the following definition makes sense.

Definition 1.16. Let φ be a flow with infinitesimal generator X. Then we say that φ is the flow *generated* by X. ❀

There is a partial converse to Proposition 1.14:

Proposition 1.17. *Let X be a vector field on the manifold M. Assume any one of the following:*

(a) *M is compact; or*

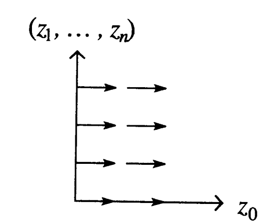
(b) *X has compact support; or*

(c) *X is complete.*

Then there is a unique one-parameter group of diffeomorphisms with X as generator.

Proof. If such a one-parameter group exists then it is unique by the last corollary. Since (a) $\Rightarrow$ (b) $\Rightarrow$ (c), it is enough to prove existence assuming (c).

Since X is complete, we have an integral curve $\sigma\colon \mathbf{R}, 0 \to M, x$ through any point $x \in M$. We define $\varphi_t(x) = \sigma(t)$. This defines $\varphi_t(x)$ for all x and t, and it remains to show that φ_t is a one-parameter group of diffeomorphisms. Once again, we use the trick of passing to the vector field $Y = (\partial/\partial s, X)$ on $\mathbf{R} \times M$, and note that $\tau(t) \equiv F(t,s,x) = (s+t, \varphi_t(x))$ is an integral curve for Y through the point (s,x). Since Y is a never-zero field, we can linearize it on a neighborhood $U \times V$ of any point of $\mathbf{R} \times M$ as in the following figure



and then the integral curves are given locally in this coordinate system by translations

$$F(t, z_0, \dots, z_m) = (z_0 + t, z_1, \dots, z_m)$$

(where, of course, each z_j is a function of (s, x)). From this local description it is clear that

(i) F is smooth, and

(ii) $F(u, F(t, s, x)) = F(t + u, s, x)$ for s and t small.

Thus $\varphi_t(x)$ is smooth in (x, t). Moreover,

$$\begin{aligned}
(u + s + t, \varphi_u(\varphi_t(x))) &= F(u, s + t, \varphi_t(x)) \\
&= F(u, F(t, s, x)) \\
&= F(u + t, s, x) \\
&= (u + s + t, \varphi_{u+t}(x)) \quad \text{(for } s \text{ and } t \text{ small)},
\end{aligned}$$

and so φ satisfies property (ii) for s and t small, and hence for all s and t. Property (i) is automatic. ■

Exercise 1.18. Let X be a *smooth* vector field on $\mathbf{R}$ with $X_0 = 0$. Show that no integral curve can reach 0 in finite time. (Note that if X is complete, this follows from Proposition 1.17, since the one-parameter family of diffeomorphisms φ_t generated by X clearly satisfies $\varphi_t(0) = 0$ for all t. Also, since φ_t is a diffeomorphism, we can therefore never have $\varphi_t(y) = 0$ for $y \neq 0$.) Show that the same is true for a smooth vector field on a smooth manifold. ❑

§2. Distributions

Now we wish to generalize part of the study of integral curves to include leaves of larger dimension. We shall formalize the idea of "prescribing the tangent planes" in the notion of a *distribution.*

Definition 2.1. An r-dimensional *distribution* on M is a collection $\mathcal{D} = \{\mathcal{D}_p\}$ of r-dimensional subspaces $\mathcal{D}_p \subset T_pM$, one for each $p \in M$, that are smooth in the sense that they may be described on any sufficiently small open set $U \subset M$ as the span of r smooth vector fields $\{X_1, X_2, \dots, X_r\}$. These vector fields themselves are called a *local basis* for $\mathcal{D}$ on U. A distribution is said to be *involutive* (or *in involution*, or *integrable*) if for any point $p \in M$ there is a chart (U, φ) around p such that the vector fields

$$\varphi_*^{-1}\left(\frac{\partial}{\partial x_1}\right), \dots, \varphi_*^{-1}\left(\frac{\partial}{\partial x_r}\right)$$

form a basis of $\mathcal{D}$ on U. ❀

Definition 2.2. A connected r-dimensional submanifold N of M is called an *integral submanifold* for the r-dimensional distribution $\mathcal{D}$ if $\mathcal{D}_p = T_p(N)$ for each $p \in N$.[4] ❀

In the case of an involutive distribution, it is clear that the equations

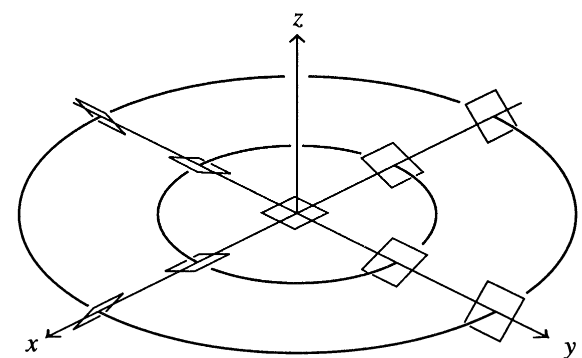
$$\varphi_{r+1}(p) = c_{r+1}, \dots, \varphi_n(p) = c_n$$

describe an $(n-r)$-parameter family of r-dimensional integral submanifolds for $\mathcal{D}$ (one for each choice of $c = (c_{r+1}, \dots, c_n)$) and that any point in U has such an integral submanifold passing through it.

§3. Integrability Conditions

In contrast to the case of a single vector field, not every distribution has an integral submanifold through every point.

Example 3.1. Consider the two-dimensional distribution on $\mathbf{R}^3$ consisting of planes normal to the vector field $n(x, y, z) = (y, -x, 1)$ as in the figure below.



We can see directly from this figure that there are no integral surfaces through 0 for this distribution. For if there were such a surface through the

[4]More generally, if $s \leq r$, an s-dimensional connected submanifold N of M is called an *integral* submanifold for $\mathcal{D}$ if $T_PN \subset \mathcal{D}_p$ for each $p \in N$. However, we won't use this generalization in this book.

origin, it would be tangent to the (x, y)-plane there, but a small loop on the surface about the z-axis could never actually close up since its z-coordinate would always be increasing as we pass counterclockwise around it. ◆

Suppose N is an integral submanifold for a distribution on M. Suppose that a local basis for the distribution is given on a neighborhood of $p \in N$ is by $X_1, \ldots, X_n$. Let $i: N \to M$ be the inclusion map. Since $X_j|_N$ is i related to X_j for each $j = 1, \ldots, n$, we have $i_*[X_i|_N, X_j|_N] = [X_i, X_j]|_N$. It follows that $[X_i, X_j]$ is also tangent to N at p, that is, it lies in $\operatorname{span}\{X_1, X_2, \ldots, X_n\}$ at p.

Definition 3.2. The system of vector fields $\{X_1, X_2, \ldots, X_n\}$ on $U \subset M$ is called (*algebraically*)[5] *involutive* (or *integrable*) if the bracket $[X_i, X_j]$ lies in $\operatorname{span}\{X_1, X_2, \ldots, X_n\}$ for each i and j. ❀

As we have just seen, algebraic involutivity is a necessary condition that a system of $n < m$ vector fields $\{X_1, X_2, \ldots, X_n\}$ on an m-manifold M have an integral submanifold through each point p. The existence of an integral manifold through each point is in turn a necessary condition for the distribution to be involutive. In the case of Example 3.1, a local basis for the distribution is given by

$$X_1 = \frac{\partial}{\partial x} - y\frac{\partial}{\partial z}, \quad X_2 = \frac{\partial}{\partial y} + x\frac{\partial}{\partial z},$$

so that

$$[X_1, X_2] = 2\frac{\partial}{\partial z} \notin \operatorname{span}\{X_1, X_2\}.$$

It follows that the distribution has no integral manifolds.

Exercise 3.3. Show that if $\{X_1, X_2, \ldots, X_r\}$ and $\{Y_1, Y_2, \ldots, Y_r\}$ are two local bases on an open set $U \subset M$ for the distribution $\mathcal{D}$, then one is algebraically involutive if and only if the other is. ❑

The exercise implies that the algebraic involutivity is a property of the distribution itself. Hence we see that a necessary condition that a distribution $\mathcal{D}$ be involutive is that it be algebraically involutive.

§4. The Frobenius Theorem

The content of the following theorem is that the condition of algebraic involutivity, which is necessary for the involutivity of the distribution, is

[5]This word is bracketed because we shall omit it as unnecessary once the Frobenius theorem (Theorem 4.1), is proved.

also a sufficient condition. This fact allows us to dispense henceforth with the adjective *algebraic* in the phrase *algebraically involutive.*

Theorem 4.1 (Frobenius). *Let M be an m-dimensional manifold and let $\mathcal{D}$ be an r-dimensional distribution on M. Then*

$$\mathcal{D} \text{ algebraically involutive} \Leftrightarrow \mathcal{D} \text{ involutive.}$$

Proof. (We follow [S.S. Chern and J. Wolfson, 1981]. For another interesting proof, cf. [W. Boothby, 1986], p. 161.)

$\Leftarrow$ This direction was shown above.

$\Rightarrow$ For $r = 1$ the theorem is implied by the linearization of vector fields. We proceed by induction on $r \geq 2$. The method is to use the induction hypothesis and the linearization of vector fields to successively improve a local basis of $\mathcal{D}$. The theorem is local, so we may assume $M = U =$ an open subset of $\mathbf{R}^m$, $p = 0 \in U$, and $\mathcal{D}$ is given on U by a local basis $\{X_1, X_2, \ldots, X_r\}$. At each step the vector fields and the open set U will change, but we will keep the same notation for them. Our goal is to produce a coordinate system $x = (x_1, \ldots, x_m)$ and a local basis for $\mathcal{D}$ on U of the form $\{X_1 = \partial/\partial X_1,\ X_2 = \partial/\partial x_2, \ldots, X_r = \partial/\partial x_r\}$.

Step 1. There is a local coordinate system $x = (x_1, \ldots, x_m)$ and a local basis for $\mathcal{D}$ on U of the form $\{X_1, X_2, \ldots, X_{r-1},\ X_r = \partial/\partial x_r\}$.

This is just the linearization of the vector field X_r.

Step 2. There is a local coordinate system $x = (x_1, \ldots, x_m)$ and a local basis for $\mathcal{D}$ on U of the form $\{X_1, X_2, \ldots, X_{r-1},\ X_r = \partial/\partial x_r\}$ such that $\{X_1, X_2, \ldots, X_{r-1}\}$ is in algebraic involution and $X_j(x_r) = 0$ for $1 \leq j \leq r - 1$.

Set

$$\begin{aligned} X_j' &= X_j - X_j(x_r)X_r \quad \text{for } 1 \leq j \leq r-1, \\ X_r' &= X_r. \end{aligned}$$

Note that

$$\begin{aligned} X_j'(x_r) &= 0 \quad \text{for } 1 \leq j \leq r-1, \\ X_r'(x_r) &= 1. \end{aligned}$$

Writing[6]

$$[X_i', X_j'] = a_{ij}(x)X_r \bmod\{X_1', X_2', \ldots, X_{r-1}'\} \quad \text{for } 1 \leq i, j \leq r-1,$$

[6]An equation of the form $A = B \bmod\{C_1, \ldots, C_s\}$ means that $A - B$ lies in the span of $C_1, \ldots, C_s$.

we see that $a_{ij}(x) = 0$ for $1 \le i, j \le r-1$ by evaluating both sides on x_r.
Thus $\{X_1', X_2', \ldots, X_{r-1}'\}$ is in algebraic involution.

Step 3. There is a local coordinate system $x = (x_1, \ldots, x_m)$ and a local basis for $\mathcal{D}$ on U of the form $\{X_1 = \partial/\partial x_1,\ X_2 = \partial/\partial x_2, \ldots, X_{r-1} = \partial/\partial x_{r-1},\ X_r = \sum_{1\le A\le m} \zeta_A \partial/\partial x_A\}$, where the final coefficients ζ_A, $r \le A \le m$ are functions of $x_r, x_{r+1}, \ldots, x_m$ alone.

By the induction hypothesis applied to the results of step 2, we can find a coordinate system $y = (y_1, \ldots, y_m)$ such that

$$\operatorname{span}\{X_1, X_2, \ldots, X_{r-1}\} = \operatorname{span}\{\partial/\partial y_1, \partial/\partial y_2, \ldots, \partial/\partial y_{r-1}\}.$$

It follows that $\{\partial/\partial y_1, \partial/\partial y_2, \ldots, \partial/\partial y_{r-1}, \partial/\partial x_r\}$ is a local basis for the distribution and, moreover, that $\partial x_r/\partial y_j = 0$ for $1 \le j \le r-1$.
Writing

$$[\partial/\partial y_j, \partial/\partial x_r] = b_j(x)\partial/\partial x_r \bmod\{\partial/\partial y_1, \ldots, \partial/\partial y_{r-1}\} \quad \text{for } 1 \le j \le r-1,$$

we see that $b_j(x) = 0$ for $1 \le j \le r-1$ by evaluating both sides on x_r. Thus we may write

$$\left[\frac{\partial}{\partial y_j}, \frac{\partial}{\partial x_r}\right] = \sum_{1\le k\le r-1} c_{jk}\frac{\partial}{\partial y_k} \quad \text{for } 1 \le j \le r-1.$$

Now, expressing the $\partial/\partial x_r$ in terms of the ys, we may write

$$\frac{\partial}{\partial x_r} = \sum_{1\le A\le m} \zeta_A \frac{\partial}{\partial y_A}.$$

Putting this in the previous formula yields

$$\sum_{1\le A\le m} \frac{\partial \zeta_A}{\partial y_j}\frac{\partial}{\partial y_A} = \sum_{1\le k\le r-1} c_{jk}\frac{\partial}{\partial y_r} \quad \text{for } 1 \le j \le r-1.$$

Comparing the two sides, we see that $\partial\zeta_A/\partial x_j = 0$ for $r \le A \le m$ and $1 \le j \le r-1$. It follows that $\xi_r, \ldots, \xi_m$ are functions of $x_r, \ldots, x_m$ alone.

Step 4. There is a local coordinate system $x = (x_1, \ldots, x_m)$ and a local basis for $\mathcal{D}$ on U of the form $\{X_1 = \partial/\partial x_1,\ X_2 = \partial/\partial x_2, \ldots, X_{r-1} = \partial/\partial x_{r-1},\ X_r = \sum_{r\le A\le m} \zeta_A(x_r, \ldots, x_m)(\partial/\partial x_A)\}$, where the coefficients ζ_A, $r \le A \le m$, are functions of $x_r, x_{r+1}, \ldots, x_m$ alone.

Set

$$X_j' = X_j \quad \text{for } 1 \le j \le r-1,$$
$$X_r' = X_r - \sum_{1\le A\le r-1} \zeta_A X_A.$$

Then $\{X_1', \ldots, X_r'\}$ is still a local basis for $\mathcal{D}$ and X_r' has the form

$$X_r' = \sum_{r\le A\le m} \zeta_A(x_r, \ldots, x_m)\frac{\partial}{\partial x_A}.$$

Step 5. There is a local coordinate system $x = (x_1, \ldots, x_m)$ and a local basis for $\mathcal{D}$ on U of the form $\{X_1 = \partial/\partial x_1,\ X_2 = \partial/\partial x_2, \ldots, X_r = \partial/\partial x_r\}$.

The vector field X_r is a vector field on $\mathbf{R}^{m-r}$, and by the linearization of vector fields we can find a change of the variables $x_r, \ldots, x_m$ that makes X_r' a coordinate field without affecting the other variables. This completes the proof. ■

The proof of the Frobenius theorem not only tells us that the leaves through any point exist, but also tells us something about their mutual disposition. It says that, locally, an integrable r-dimensional distribution on an m-dimensional manifold looks like the affine r-planes in $\mathbf{R}^r \times \mathbf{R}^{m-r}$ with "second coordinate" constant. More formally, we have the next corollary.

Corollary 4.2. *Let M be an m-dimensional manifold and let $\mathcal{D}$ be an involutive r-dimensional distribution on M. Then for each point $p \in M$ there is a local coordinate system (U, x) about p such that the integral manifold through q meets U in a set containing*

$$V(q) = \{q' \in U \mid x_j(q') = x_j(q) \quad \text{for } m-r < j \le m\}.$$

Moreover, if (U, x) and (V, y) are two such coordinate systems, then the coordinate changes $\Phi = xy^{-1}$ have the form

$$(\Phi_1(x_1, \ldots, x_m), \ldots, \Phi_{m-r}(x_1, \ldots, x_m),$$
$$\Phi_{m-r+1}(x_{m-r+1}, \ldots, x_m), \ldots, \Phi_m(x_{m-r+1}, \ldots, x_m)).$$

Proof. This is a simple consequence of step 5 in the proof of the theorem. ■

§5. The Frobenius Theorem in Terms of Differential Forms

For us, the most useful formulation of the Frobenius theorem is in terms of differential forms. This is the context in which a distribution is given, not in local terms as the span of smooth vector fields, but in global terms as the kernel of a V-valued 1-form.

Proposition 5.1. *Let ω be a smooth 1-form on M^m with values in the vector space V. Assume that $n = \dim \ker \omega_x$ is constant for $x \in M$. Then* $\ker \omega_x$ *is a distribution.*

Proof. We must show that, locally, $\ker \omega_x$ is spanned by n linearly independent, smooth vector fields. Choose a basis $\{e_1, \dots, e_q\}$ for V, and write $\omega = \sum \omega_i e_i$, where the ω_i are smooth 1-forms. Fix a point $p \in M$ around which we want to find a local basis of $\ker \omega_x$. Write $\omega_i = \sum a_{ij}(x) dx_j$ in a local coordinate system on U about p so that $\omega = \sum a_{ij}(x) dx_j e_i$, and set $A(x) = (a_{ij}(x))$, a $q \times m$ matrix. Now

$$\omega(v) = 0 \Leftrightarrow \sum a_{ij}(x) v_j = 0 \quad \text{for all } i$$

(where $v_j = dx_j(v)$), so

$$\dim \ker \omega_x = n \Leftrightarrow \operatorname{rank}(a_{ij}(x)) = m - n.$$

Thus we may assume, possibly after changing the basis of V and reordering the x_j that the first $(m-n) \times (m-n)$ block of $A(p) = (a_{ij}(p))$ is invertible. Thus, the same is true on some neighborhood of p, which we again call U, and because of the rank condition the last $(q - (m - n))$ rows are linearly dependent on the first $(m - n)$ rows. Thus, on U we clearly have $\ker \omega = \cap \ker \omega_j$, with the intersection taken over the range $1 \le j \le m - n$. Now set

$$\eta_i = \begin{cases} \omega_i, & 1 \le i \le m-n, \\ dx_i, & m-n < i \le m. \end{cases}$$

The η_i, $1 \le i \le m$, form a basis for the 1-forms on U. We write $\eta_i = \sum b_{ij}(x) dx_j$, $1 \le i \le m$. The vector fields $X_j = \sum c_{jk}(x)(\partial/\partial x_k)$ dual to the η_i are smooth since the coefficients are given by $(b_{ij}(x))\,(c_{jk}(x)) = I$, and $(b_{ij}(x))$ is invertible. Moreover, the smooth vector fields X_j, for $m, n < j \le m$, form a local basis for the distribution $\ker \omega_x$. ■

Exercise 5.2. Let M be a smooth manifold and $\omega: T(M) \to V$ a (smooth) trivialization of the tangent bundle. Show that for all $v \in V$, $\omega^{-1}(v)$ is a smooth vector field on M. If $p \in M$, let $c(t, p, v)$ (defined for all t on some neighborhood of zero) denote the integral curve of $\omega^{-1}(v)$ with $c(0, p, v) = p$. Show that $c(\alpha t, p, v) = c(t, p, \alpha v)$. Show that the map $\exp_\omega: T(M) \to M \times M$, which sends $(x, w) \to (x, c(1, x, \omega(w)))$, is defined on some neighborhood of the zero section and is a diffeomorphism of some neighborhood U of the zero section onto a neighborhood of the diagonal in $M \times M$. ❑

Now we are in a position to ask what form the integrability conditions for a distribution assume when the distribution is given as the kernel of a V-valued 1-form ω. We have the next result which is the differential form version of Frobenius' theorem.

Proposition 5.3. *Let ω be a smooth 1-form on M with values in the vector space V. Assume that $n = \dim \ker \omega_x$ is constant for $x \in M$, and let $\mathcal{D} = \{\ker \omega_x \mid x \in M\}$ be the distribution determined by ω. Then*

$$\mathcal{D} \text{ is integrable} \Leftrightarrow d\omega(X, Y) = 0 \text{ whenever } \omega(X) = \omega(Y) = 0.$$

Proof. Let us choose a local basis $X_1, \dots, X_r$ for $\mathcal{D}$ in a neighborhood U of a point of M.

$$\begin{aligned} \text{Then } \mathcal{D} \text{ integrable on } U &\Leftrightarrow [X_j, X_k] \in \operatorname{span}\{X_1, \dots, X_r\} \quad \text{for } 1 \le j, k \le r \\ &\Leftrightarrow \omega([X_j, X_k]) = 0 \quad \text{for } 1 \le j, k \le r. \end{aligned}$$

But since $\omega(X_j) = 0$, $1 \le j \le r$, we have (cf. Lemma 1.5.15) for $1 \le j, k \le r$

$$d\omega(X_j, X_k) = X_j(\omega(X_k)) - X_k(\omega(X_j)) - \omega([X_j, X_k]) = -\omega([X_j, X_k]),$$

and so

$$\mathcal{D} \text{ integrable on } U \Leftrightarrow d\omega(X_j, X_k) = 0 \quad \text{for } 1 \le j, k \le r,$$

which implies the result. ■

Corollary 5.4. *Let ω be a smooth 1-form on M with values in the vector space V. Let $W \subset V$ be a subspace, and assume that $n = \dim \omega_x^{-1}(W)$ is constant for $x \in M$. Then $\mathcal{D} = \{\omega_x^{-1}(W) \mid x \in M\}$ is a distribution and*

$$\mathcal{D} \textit{ integrable} \Leftrightarrow d\omega(X, Y) \in W \quad \textit{whenever } \omega(X) \textit{ and } \omega(Y) \textit{ lie in } W.$$

Proof. The distribution $\mathcal{D}$ is the kernel of the smooth 1-form ω mod W taking values in V/W. The equivalence of the integrability for condition for ω mod W and the one given above for $\mathcal{D}$ is clear. ■

Exercise 5.5. Let $\omega = (\omega_1, \dots, \omega_v)$ be a smooth 1-form on M with values in $\mathbf{R}^v$. Assume that $n = \dim \ker \omega_x$ is constant for $x \in M$, and let $\mathcal{D} = \{\ker \omega_x \mid x \in M\}$ be the distribution determined by ω. Show that

$$\mathcal{D} \text{ integrable} \Leftrightarrow d\omega_i \in I(\omega_1, \dots, \omega_v),$$

where $I(\omega_1, \dots, \omega_v)$ is the ideal in $A(M)$ generated by $\omega_1, \dots, \omega_v$. ❑

§6. Foliations

The modern study of integrable distributions is called the theory of foliations. The idea is to abstract the property contained in the corollary to the Frobenius theorem. A foliation on a manifold consists of the special coordinate systems guaranteed by this corollary, which clearly determine the integral manifolds. The integral manifolds themselves are called the leaves of the foliation.

Definition 6.1. Let M^m be a smooth manifold. A *q-codimensional foliated atlas* on M is an atlas $\mathcal{A}$ such that, if $(U, \varphi), (V, \psi) \in \mathcal{A}$, then the coordinate changes $\Phi = \psi\varphi^{-1}$ have the form

$$\begin{aligned} \Phi\colon \mathbf{R}^{m-q} \times \mathbf{R}^q &\to \mathbf{R}^{m-q} \times \mathbf{R}^q, \\ (x, y) &\mapsto (\Phi_1(x, y), \Phi_2(y)), \end{aligned}$$

namely, the last q-coordinates depend only on the last q variables. ❀

Definition 6.2. Two foliated atlases are *equivalent* if their union is a foliated atlas. A *foliation* on M is an equivalence class of foliated atlases. ❀

Note that every foliated atlas is equivalent to a maximal foliated atlas, which may be identified with the foliation. As we did for the definition of a smooth structure, we shall always implicitly assume that our atlases are maximal.

Given two charts (U, φ), (V, ψ) in a foliated atlas, the coordinate change $\Phi = \psi\varphi^{-1}$ sending $(x, y) \mapsto (\Phi_1(x, y), \Phi_2(y))$ determines the diffeomorphism $\Phi_2\colon \mathbf{R}^q \to \mathbf{R}^q$. We can use this diffeomorphism to replace the coordinate system φ on U by $\theta = (\mathrm{id} \times \Phi_2)\varphi$, also on U. The coordinate change between the charts (U, θ) and (V, ψ) is then the simpler diffeomorphism $(x, y) \to (\Phi_1(x, y), y)$.

Let M^m be a smooth manifold with a foliation. If (U, φ) is a foliated chart and

$$\varphi = (\varphi_1, \varphi_2)\colon U \to \mathbf{R}^{m-q} \times \mathbf{R}^q,$$

then for $x \in U$, set

$$\mathcal{D}_x = \varphi_{*x}^{-1}(0 \times T_{\varphi_2(x)}(\mathbf{R}^q)).$$

Exercise 6.3. Show that $\mathcal{D}_x$ is independent of the choice of foliated chart used to define it and determines an integrable distribution of codimension q. ❑

Definition 6.4. The integrable distribution guaranteed by Exercise 6.3 is called the *distribution associated to the foliation.* ❀

Definition 6.5. The *leaf* $\mathcal{L}$ through a point p of a foliation on M is defined to be the set of points on M which can be joined to p by a piecewise smooth path everywhere tangent to the distribution associated to the foliation. ❀

Proposition 6.6. *Every leaf $\mathcal{L}$ of a foliation of codimension q on M^m is a smooth submanifold of dimension $m - q$ and is an integral submanifold for the associated distribution $\mathcal{D}$.*

Proof. Let $x \in \mathcal{L}$ and (U, φ) be a chart of a foliated atlas with $x \in U$. Now $x \in \varphi^{-1}(\mathbf{R}^{m-q} \times y)$ for some $y \in \mathbf{R}^q$. Let W be the path component of $\varphi^{-1}(\mathbf{R}^{m-q} \times y)$ containing x. Since W is connected and is an integral manifold of the distribution $\mathcal{D}$, we have $x \in W \subset \mathcal{L}$. On the other hand, W is a flat plaque of $\mathcal{L}$ in the chart (U, φ). Thus, $\mathcal{L}$ is covered by flat plaques of dimension $m - q$, and so $\mathcal{L}$ satisfies the condition of Definition **1**.2.1 for a submanifold. Moreover, since $\mathcal{L}$ is a union of open plaques, each of which is an integral submanifold for the distribution $\mathcal{D}$, it follows that $\mathcal{L}$ is also an integral submanifold for $\mathcal{D}$. ■

Corollary 6.7. *Every leaf of a foliation is a maximal connected integral submanifold of the associated distribution $\mathcal{D}$.*

Proof. It suffices to show that every connected integral manifold of $\mathcal{D}$ containing x lies in the leaf containing x. Suppose that $N \subset M$ is such a submanifold. Let $\mathcal{L}$ be the leaf through x. Since N is connected, any point $y \in N$ may be joined to x by a smooth path on N. Since N is an integral submanifold for $\mathcal{D}$, this smooth path is everywhere tangent to $\mathcal{D}$, and hence $y \in \mathcal{L}$. Thus, $N \subset \mathcal{L}$. ■

Definition 6.8. A *product* (or *trivial*) foliation is one on a manifold $M = N^q \times \mathcal{L}$ for which the atlas is just the product of the atlases of the factors. ❀

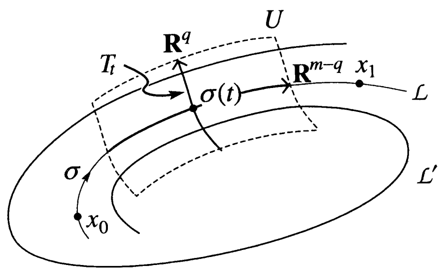
Example 6.9. If $F \to E \to B$ is a smooth bundle, then the local product structure determines a foliation on E, called the *vertical foliation*, for which the leaves are the fibers. ◆

§7. Leaf Holonomy

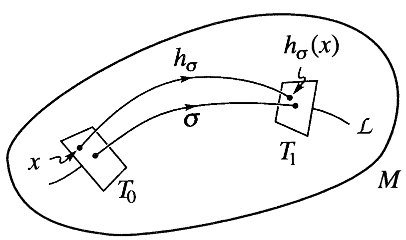
Now we show that for a fixed leaf $\mathcal{L}$, the way that the various nearby leaves arrange themselves about $\mathcal{L}$ near a given point $p \in \mathcal{L}$ is in fact independent of the choice of $p \in \mathcal{L}$. Of course, this arrangement does depend very much on the choice of $\mathcal{L}$ itself. This will lead to the notion of *holonomy*, which describes how the leaves near $\mathcal{L}$ "wind around" $\mathcal{L}$.

Sliding Along Leaves

The basic construction is the following. Fix a leaf $\mathcal{L}$ and two points on it $x_0, x_1 \in \mathcal{L}$. Next choose a continuous path $\sigma\colon (I, 0, 1) \to (\mathcal{L}, x_0, x_1)$. Now at any point of the path σ, say at $\sigma(t)$, we pick a chart (U, φ), $\varphi = (x, y): U \to \mathbf{R}^{m-q} \times \mathbf{R}^q$, of the foliation so that the plaque of $\mathcal{L}$ in U containing $\sigma(t)$ corresponds, under φ, to $\mathbf{R}^{m-q} \times 0$. We call $T_t = \varphi^{-1}(0 \times \mathbf{R}^q)$ a *transversal* at $\sigma(t)$. Each point of T_t determines the leaf through it, although distinct points may lie on the same leaf, as shown in the following diagram.



Now the coordinate changes of the last q coordinates do not depend on the first $m-q$ coordinates. It follows that on the overlap between two charts at $\sigma(a)$ and $\sigma(b)$, the coordinate change induces a diffeomorphism $\varphi_{b,a}: T_a \to T_b$.[7] Moreover, it is clear that $\varphi_{c,b}\varphi_{b,a} = \varphi_{c,a}$.[8] Since $\sigma(I)$ is compact, finitely many such coordinate systems suffice to cover it. By taking the composite of the corresponding finitely many diffeomorphisms, we can pass from one end of the path to the other to obtain a diffeomorphism $h_\sigma: T_0 \to T_1$ called the slide map as shown in the following diagram.



Exercise 7.1. Show that the germ of $h_\sigma: T_0 \to T_1$ at $\sigma(0)$ is independent of the particular coordinate systems used to define it and is even independent of the choice of $\sigma: (I, 0, 1) \to (\mathcal{L}, x_0, x_1)$ within its homotopy class. ❑

Thus, we have the following theorem.

[7]The diffeomorphism may not be defined on all of T_a nor have all of T_b as its image. Rather, it will be a diffeomorphism between *neighborhoods* of $\sigma(a)$ and $\sigma(b)$ in T_a and T_b, respectively. This means that it is not a diffeomorphism but the *germ* of a diffeomorphism.

[8]At least this is true on the (nonempty) domain where it makes sense.

Theorem 7.2. *Let $\sigma: (I, 0, 1) \to (\mathcal{L}, x_0, x_1)$ be a continuous path on a leaf of a foliation of M. Let T_0 and T_1 be two transversals of the foliation at x_0 and x_1, respectively. Then there exist open subsets T_0' and T_1' of T_0 and T_1, which are still transversals of the foliation at x_0 and x_1, respectively, and a diffeomorphism $\varphi: T_0' \to T_1'$ such that*

(i) *$\mathcal{L}_x = \mathcal{L}_{\varphi(x)}$ for all $x \in T_0'$, where $\mathcal{L}_x$ denotes the leaf through x,*

(ii) *φ depends only on the homotopy class of $\sigma: (I, 0, 1) \to (\mathcal{L}, x_0, x_1)$,*

(iii) *if σ_j corresponds to φ_j for $j = 1, 2$ and $\sigma_1(1) = \sigma_2(0)$, then the composite path $\sigma_1 \star \sigma_2$ (i.e., first follow σ_1, then follow σ_2) corresponds to the composite mapping $\varphi_2 \circ \varphi_1$.*

Corollary 7.3. *Let $\mathcal{L}$ be a closed leaf, and let T be a transversal. Then $T \cap \mathcal{L}$ is a discrete subset of T.*

Proof. It suffices to show that $K \cap \mathcal{L}$ is discrete for every compact set K in T. Since $\mathcal{L}$ is closed in M, it follows that $A = K \cap \mathcal{L}$ is closed. If A is not discrete, then there is a point $p \in K$ which is an accumulation point of A. Since A is closed, $p \in A$. Now by the theorem, if $q \in A$, there is a diffeomorphism between some neighborhood of p in T and some neighborhood of q in T mapping points of A to points of A. Thus, $q \in A$ is also an accumulation point of A. This means that *every* point of A is an accumulation point. Thus, A is a perfect set. But every perfect set is uncountable, so we see that $\mathcal{L}$ contains the uncountable set A, which is obviously discrete in the leaf topology. But $\mathcal{L}$ is paracompact, so any discrete set is at most countable,[9] which is a contradiction. ■

Corollary 7.4. *Every closed leaf is a regular submanifold of M.*

Proof. Let $\mathcal{L}$ be a closed leaf, $p \in \mathcal{L}$, and let T be a transversal at p arising from the chart (U, φ). Since $T \cap \mathcal{L}$ is a discrete subset of T, we can shrink the size of the chart (U, φ) so that the new reduced transversal meets the leaf in the single point p. This chart shows that $\mathcal{L}$ is a regular submanifold near p. Since p is arbitrary, $\mathcal{L}$ is a regular submanifold. ■

Exercise 7.5. Let $A \subset \mathbf{R}^n$ be a closed set with every point an accumulation point. Show that A is uncountable. [*Hint:* Since each point $p \in A$ is an accumulation point, for every $n > 0$ there is a point $q \neq p$ with $|q-p| < 1/n$. Fix $p \in A$ and take two distinct such qs for $n = 1$. Then repeat the argument for each of the qs in place of p for $n = 2$. Repeat this process to construct

[9]A discrete implies each point $p \in A$ has an open neighborhood U with $A - p$ in U^c.

inductively an uncountable number of convergent sequences with distinct limits.] ❑

Definition 7.6. Let us fix a leaf $\mathcal{L}$, a point $p \in \mathcal{L}$ and a transversal T to the foliation at p. Let $G(T,p)$ be the group of germs diffeomorphisms of (T,p) at p. Then our construction associates to each loop $\lambda: (I, \partial I) \to (\mathcal{L}, p)$ a germ of a diffeomorphism $\varphi(\lambda) \in G(T,p)$. In fact, because of properties (ii) and (iii) of Theorem 7.3, φ induces a group homomorphism $\varphi_*: \pi_1(\mathcal{L},p) \to G(T,p)$ called "the *leaf* (or *transverse*) *holonomy*." ❀

Exercise 7.7. Show that if T_0, T_1 are two transversals through the end points $x_0, x_1 \in \mathcal{L}$ of a path $\sigma: (I, 0, 1) \to (\mathcal{L}, x_0, x_1)$, then the canonical (germ of a) diffeomorphism $f: (T_0, x_0) \to (T_1, x_1)$ induces a commutative diagram:

$$\begin{tikzcd}
\pi_1(\mathcal{L}, x_0) \arrow[r, "c_\sigma"] \arrow[d, "\varphi_{0*}"'] & \pi_1(\mathcal{L}, x_1) \arrow[d, "\varphi_{1*}"] \\
G(T_0, x_0) \arrow[r, "c_f"] & G(T_1, x_1)
\end{tikzcd}$$

❑

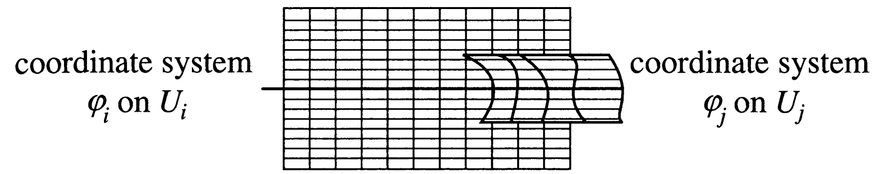
Of course, the reason for using *germs* of diffeomorphisms is to have well-defined maps, since there is no guarantee how the "Poincaré first return map" $\varphi(\lambda)$ will behave with respect to the transversal T on two counts. First, a loop may either increase or decrease the size of T (or even do something more complicated). Second, different loops may do different things. However, in the special case that the leaf has finitely generated fundamental group (in particular, if it is compact),[10] there is a kind of bound on the behavior of the holonomy in the sense that we may choose transversals $T_1 \subset T_2$ at p such that for some set of generators $\{g_1, \ldots, g_k\}$ of $\pi_1(\mathcal{L},p)$ we have $\varphi_*(g_i)T_1 \subset T_2$ for $1 \leq i \leq k$. If, in this special case, the holonomy is trivial, it follows easily that $\varphi_*(g_i)$ is defined on T_1 for all i and, moreover, is the identity there. It then follows that $\varphi_*(g)$ is defined and is the identity on T_1 for all $g \in \pi_1(\mathcal{L},p)$. This suggests the following.

Theorem 7.8. *Let M be a foliated manifold and let $\mathcal{L}$ be a compact leaf with trivial holonomy. Then there is a neighborhood U of $\mathcal{L}$ in M such that there exists a leaf-preserving diffeomorphism $\mathcal{L} \times T \to U$, where T is a transversal at $p \in \mathcal{L}$.*

---

[10]Here we use the fact that the fundamental group of a compact manifold is finitely generated. Cf. [J. Dugundji, 1966].

Proof. Since $\mathcal{L}$ is a compact leaf in M, each point of $\mathcal{L}$ lies in a coordinate neighborhood of the foliated atlas meeting $\mathcal{L}$ in just one plaque. Using the compactness of $\mathcal{L}$ again, we may refine this cover to a finite one consisting of foliated coordinate neighborhoods (R_i, φ_i), $1 \leq i \leq n$, meeting $\mathcal{L}$ in the plaques U_i, $1 \leq i \leq n$. Since the holonomy is trivial, we may assume once and for all that, for each $y \in \mathbf{R}^q$, the leaf containing the plaque $\varphi_i^{-1}(\mathbf{R}^{m-q} \times y)$ depends only on y, not on i. This means we have matched up all the plaques, as in the following figure.



The remaining problem is to match up all the transversals. Let us now choose smaller open sets $W_i \subset U_i$ that still form a cover of $\mathcal{L}$ and satisfy $\overline{W_i} \subset U_i$. To prove the theorem it is enough to construct inductively a sequence of open sets V_k of $\mathcal{L}$ and smooth maps $f_k: V_k \times \mathbf{R}^q \to M$, $1 \leq k \leq n$, such that

(i) $\bigcup_{1<i<k} W_i \subset V_k \subset \bigcup_{1<i<k} U_i$,

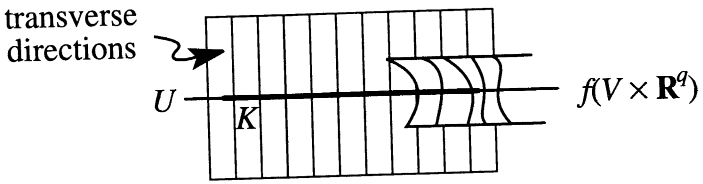
(ii) f_k is a leaf-preserving diffeomorphism onto its image which gives the canonical inclusion $V_k \subset \mathcal{L}$ upon restriction to $V_k \times 0$, and such that the leaf containing $f_k(V_k \times y)$ is independent of k.

Set $V_1 = R_1 \cap \mathcal{L}$, and let f_1 be the composite

$$V_1 \times \mathbf{R}^q \xrightarrow{\varphi_1 \times \mathrm{id}} \varphi_1(V_1) \times \mathbf{R}^q \to R_1 \subset M.$$

Next we describe the inductive step. Thus, we suppose we are given the open set V_i in $\mathcal{L}$ and the smooth map $f_i: V_i \times \mathbf{R}^q \to M$ satisfying the hypotheses (i) and (ii) above for $i = k$ and we wish to obtain them for $i = k+1$. We will do this by altering the definition of f_k on $(V_k \cap U_{k+1}) \times \mathbf{R}^q$ to make it match up with φ_{k+1}. As we shall see, in order to do this we will also have to shrink the size of the $\mathbf{R}^q$ factor in $V_k \times \mathbf{R}^q$ to an open disc, but we shall continue to denote the reduced factor by $\mathbf{R}^q$ since they are diffeomorphic.

Set $U_{k+1} = U$, $W_{k+1} = W$, $f_k = f$, and $V_k = V$. Choose K compact with $W \subset K \subset U$, and restrict the size of $\mathbf{R}^q$ so that $f((V \cap K) \times \mathbf{R}^q) \subset U$ as in the following picture.

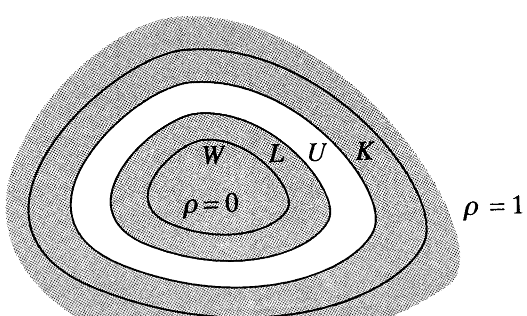


To alter f, we require a simple technical lemma.

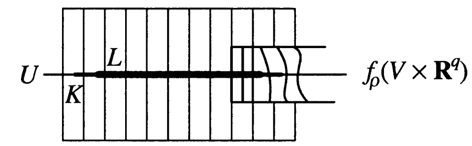
Sublemma 7.9. *Let $A \subset B$ be smooth manifolds and let $f: A \times \mathbf{R}^q \to B \times \mathbf{R}^q$ be a diffeomorphism onto its image of the form $f(x,y) = (\Phi(x,y), y)$ which satisfies $\Phi(x,0) = x$ for all $x \in A$. Let $\rho: A \times \mathbf{R}^q \to \mathbf{R}$ be any smooth function. Set $f_\rho(x,y) = (\Phi(x, y\rho(x,y)), y)$. Then for any compact set $K \subset A$ there is a neighborhood T of 0 in $\mathbf{R}^q$ such that $f_\rho: \operatorname{int}(K) \times T \to B \times \mathbf{R}^q$ is a diffeomorphism onto its image. In particular, if $\rho(x,y) = 0$ at some point, then $f_\rho(x,y) = (x,y)$ at that point.*

Proof. Since f and f_ρ agree along the coordinate plane $y = 0$ and f is an immersion there, f_ρ is also an immersion. It follows by continuity of the Jacobian matrix that f_ρ is also an immersion on some neighborhood of the coordinate plane $y = 0$ in $A \times \mathbf{R}^q$. But any such neighborhood contains some product neighborhood $K \times T_1$ with K compact and T_1 a transversal. Finally, we claim that since f_ρ is an embedding along $K \times 0$, it must also be an embedding along some neighborhood of the form $K \times T$; otherwise there would be sequences of *distinct* points $\{p_j\}$, $\{q_j\}$ with second coordinates tending to 0 satisfying $f_\rho(p_j) = f_\rho(q_j)$ for all j. Passing to a subsequence, we may assume the first coordinates converge; then they must converge to the same point in K. But since f_ρ is a local diffeomorphism, this means that we must have $p_j = q_j$ for j sufficiently large, which is a contradiction. ∎

Now we return to the proof of the theorem. Let L be any other compact set satisfying $W \subset L \subset \operatorname{int} K$.



We may choose $\rho: V \times \mathbf{R}^q \to \mathbf{R}$ with values in $[0,1]$ to be 1 on $(V - K) \times \mathbf{R}^q$ and 0 on $L \times \mathbf{R}^q$. Then we replace f by f_ρ as in the lemma and restrict the size of $\mathbf{R}^q$ again, and we have the following picture.



That is, for $(x,y) \in L \times \mathbf{R}^q$, we have $\varphi_{k+1} f_\rho(x,y) = (x,y)$. Thus, we can extend f_ρ over $(W \cup V) \times \mathbf{R}^q$ with the required properties. ∎

For the case of a compact leaf with trivial holonomy, we see the local triviality of the foliation near this leaf. If the leaf is noncompact, such an argument is unavailable. Nevertheless, if U is any relatively compact submanifold of a closed leaf $\mathcal{L}$ such that the holonomy of $\mathcal{L}$ vanishes on the image of $\pi_1(U)$ in $\pi_1(\mathcal{L})$, the foliation is once again trivial near this part of the leaf.

Here is an application of this result which amplifies an aspect of the fundamental theorem of calculus discussed in Proposition **1**.4.19.

Theorem 7.10. *Let M and N be smooth manifolds with M connected, and let $t: T(N) \to V$ be a trivialization of the tangent bundle of N. Suppose that f_n, $f: M \to N$ are smooth maps satisfying*

(i) $f^*t = f_n^*t$ *for all* $n = 1, 2, \ldots$,

(ii) $\lim_{n\to\infty} f_n(p) = f(p)$ *for some* $p \in M$.

Then $\lim_{n\to\infty} f_n(x) = f(x)$ *for all* $x \in M$.

Proof. Let $x \in M$ be an arbitrary point. Since M is connected we can always find a smooth immersion $\sigma: (I, 0, 1) \to (M, p, x)$. Now by replacing f_n and f by $f_n\sigma$ and $f\sigma$, respectively, and p and x by 0 and 1, respectively, we may assume that M has dimension one. Set $\eta = f^*t$ and note that the graphs $\mathcal{L}_n$ and $\mathcal{L}$ of f_n and f are leaves in the one-dimensional (and hence integrable) foliation on $I \times N$ given by $\pi_{1*}\eta - \pi_{2*}t = 0$. Clearly, each leaf passes from 0 to 1 in the first coordinate and all are homeomorphic to the interval $[0,1]$ by projection. By the previous result, we can find a neighborhood U of $\mathcal{L}$ in $[0,1] \times N$ and a leaf-preserving diffeomorphism $\mathcal{L} \times T \to U$, where T is a transversal at $p \in \mathcal{L}$. By (ii), for sufficiently large n, $\mathcal{L}_n \subset U$; so under the diffeomorphism $U \to \mathcal{L} \times T$, $\mathcal{L}_n$ corresponds to a leaf of the form $\mathcal{L} \times u$ for some $u \in T$. Then the result is clear. ∎

§8. Simple Foliations

In this final section we study the simple foliations, which are indeed one of the simplest kinds of foliations.

Definition 8.1. A foliation on M is called *simple* if its leaves are the level sets of a submersion $f\colon M \to N$. In particular, this means the level sets of f are required to be connected. ❀

Our aim is to characterize a simple foliation in terms of its *holonomy* and its *leaf space*. The leaf space is obtained as follows. If M is a foliated manifold, then there is an equivalence relation given on it by declaring

$$x \sim y \Leftrightarrow x \text{ and } y \text{ lie in the same leaf of } M.$$

Definition 8.2. The quotient space $M/\sim$ equipped with the quotient topology is called the *space of leaves*. We denote by $\pi\colon M \to M/\sim$ the canonical projection. ❀

The characterization of simple foliations is given in the following theorem.

Structure Theorem 8.3.

(A) *If a foliation is simple, then each leaf has trivial holonomy, and the leaf space is Hausdorff.*

(B) *Suppose a foliation has these properties:*

(i) *each leaf has trivial holonomy;*

(ii) *the leaf space is Hausdorff;*

(iii) *each leaf has a finitely generated fundamental group.*

Then the foliation is simple.

We begin the proof of this result with a technical lemma.

Lemma 8.4. *Let M be a foliated manifold and $S \subset M$.*

(i) *If S is a union of leaves, then so is $\bar{S}$.*

(ii) *If S is a union of leaves, then so is* $\operatorname{int}(S)$.

(iii) *Let U be an open set of M. Then the union V of all leaves meeting U is open.*

(iv) *The canonical projection $\pi\colon M \to M/\sim$ is an open mapping.*

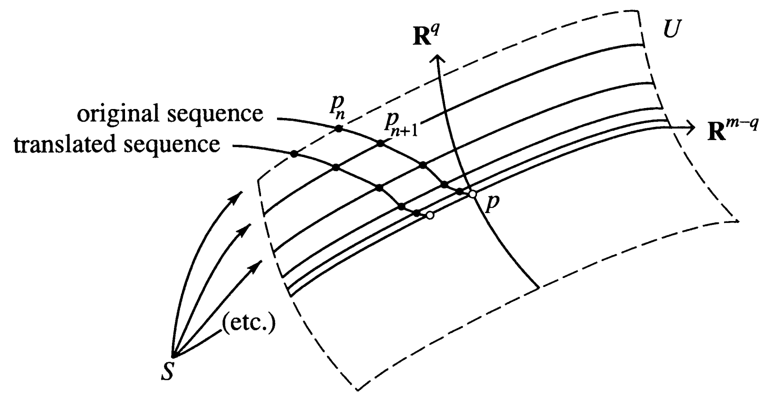
Proof. (i) Suppose the leaf $\mathcal{L}$ meets $\bar{S}$. We must show that $\mathcal{L} \subset \bar{S}$. Let $L = \mathcal{L} \cap \bar{S}$.

Step 1. It suffices to show that L is open in $\mathcal{L}$.

$L = \mathcal{L} \cap \bar{S}$ is nonempty and closed in $\mathcal{L}$, so if it is also open then it must be a component. But $\mathcal{L}$ is connected, so it has only one component. Thus $L = \mathcal{L}$ and hence $\mathcal{L} \subset \bar{S}$.

Step 2. L is open in $\mathcal{L}$.

Choose $p \in L = \mathcal{L} \cap \bar{S}$ and take a chart (U, φ) of the foliation. Then, as we can see from the following diagram,



if $p_1, p_2, \ldots$ is a sequence in S approaching p, since S is a union of leaves we can slide the whole sequence in any direction along the leaves (i.e., add some small vector to the leaf coordinates $\mathbf{R}^{m-q}$) to get sequences on S approaching points near p on the leaf $\mathcal{L}$.

(ii) Since $M - S$ is a union of leaves, so is $\overline{M - S} = M - \operatorname{int}(S)$, and hence so is $\operatorname{int}(S)$.

(iii) Since $\operatorname{int}(V)$ is a union of leaves and it contains U, it also contains V (i.e., $\operatorname{int}(V) \supset V$). But $\operatorname{int}(V) \subset V$, so $\operatorname{int}(V) = V$.

(iv) Let U be open in M. We must show that $\pi(U)$ is open. Now $\pi(U)$ is open $\Leftrightarrow \pi^{-1}\pi(U) = V$ is open, and V is the union of all leaves meeting U, which is open by part (iii). ■

Exercise 8.5. Give an example to show that if C is a closed set in M, then the union U of all leaves meeting C may not be closed. ❑

Proof of the structure theorem. (A) Let us first show that a foliation on M defined by a submersion $f\colon M \to N$ has trivial holonomy. If $h_\sigma\colon T_0 \approx T_1$ is the slide map along a curve σ tangent to the foliation in M, then we claim to have the following commutative diagram.

$$\begin{tikzcd} T_0 \arrow[rr, "h_\sigma"] \arrow[dr, "f"'] & & T_1 \arrow[dl, "f"] \\ & N & \end{tikzcd}$$

This is obviously true for transversals in a single chart; so it is also true in general, by the multiplicativity of the slide maps with respect to path multiplication. But f is injective on small transversals; thus it follows that for a *closed* curve, $h_\sigma = \mathrm{id}$, and so the holonomy is trivial. Now let us see that the leaf space is Hausdorff. The submersion $f\colon M \to N$ factors through the leaf space $M/\sim$, giving a continuous bijection $M/\sim \to N$. Since N is Hausdorff, it follows that $M/\sim$ is also Hausdorff.

Note that, for a foliation defined by a submersion, the continuous bijection $M/\sim \to N$ is actually a *homeomorphism* since submersions are open. This fact indicates how to proceed in the proof of (B): it suggests that we look for a smooth structure on the leaf space $M/\sim$ with respect to which the canonical projection is smooth and a submersion.

(B) Let us assume that M has a foliation with trivial holonomy and Hausdorff leaf space $M/\sim$. The latter condition implies in particular that each leaf is closed. Pick a leaf $\mathcal{L}$ and let $p \in \mathcal{L}$. Take a chart (U, φ) of the foliation around p. Let T be the transversal at p arising from this chart. Since $\pi_1(\mathcal{L})$ is finitely generated and the holonomy is trivial, we may assume that T is so small that it meets each leaf at most once. Thus $\pi \mid T\colon T \to M/\sim$ is injective. Now let V be a neighborhood of p in $\mathcal{L}$. Then, for T and V sufficiently small, the set $T \times V$ (in the product structure of the chart) is an open set in U. Since π is an open map, it follows that $\pi(T) = \pi(T \times V)$ is open in the leaf space. In fact, more generally it is clear that $\pi \mid T\colon T \to M/\sim$ is an open map and hence is a homeomorphism onto its image. This shows that $M/\sim$ is locally Euclidean of dimension equal to the codimension of the foliation. Moreover, two charts of the type we have constructed differ by the sliding maps, which are diffeomorphisms (where defined). Thus the atlas is smooth, and the projection map $\pi\colon M \to M/\sim$ is clearly a submersion: locally, it is just the projection map $T \times V \to T$. ■

Corollary 8.6. *A foliation whose leaves are all compact and all have trivial holonomy is simple. Moreover, if the ambient manifold M is connected, then the projection to the leaf space $\pi\colon M \to M/\sim$ provides M with the structure of a smooth fiber bundle.*

Proof. Since each leaf is compact, it follows that each leaf has a finitely generated fundamental group. Thus, by the structure theorem, to see that the foliation is simple, we need only show that the leaf space is Hausdorff. Let $\mathcal{L}_j$, $j = 1, 2$, be distinct leaves. Now, according to Theorem 7.8, we can find trivially foliated product neighborhoods U_j of $\mathcal{L}_j$, $j = 1, 2$, together with foliation-preserving diffeomorphisms $f_j\colon \mathcal{L}_j \times \mathbf{R}^q \to U_j$, $j = 1, 2$. Now the compact sets $\mathcal{L}_j$, $j = 1, 2$, can be separated by open sets and, again by the compactness of the $\mathcal{L}_j$, we may choose these open sets to be of the form $f_j(\mathcal{L}_j \times T_j) \subset U_j$. Clearly, these project to disjoint open sets in the leaf space separating the leaves $\mathcal{L}_1$ and $\mathcal{L}_2$. The proof that the projection to the leaf space equips M with the structure of a smooth fiber bundle follows easily from the existence of the trivially foliated product neighborhoods. ■

Exercise 8.7. Consider the components of $\pi^{-1}(y)$ for all $y \in \mathbf{R}$, where $\pi\colon \mathbf{R}^2 - (0,0) \to \mathbf{R}$ is given by projection on the second factor. Show that these components are the leaves of a foliation on $\mathbf{R}^2 - (0,0)$ such that every leaf is closed and has trivial holonomy. Show also that the leaf space is not Hausdorff. ❑

The phenomenon behind this example depends on the lack of "transverse completeness." The following result shows what can be done with a strong transversal completeness condition.

Proposition 8.8 ([C. Ehresmann, 1961]). *Let $\pi\colon M \to N$ be a submersion of connected manifolds such that, for each point $p \in N$, there are complete vector fields on M which project to complete vector fields on N whose values at p span $T_p(N)$. Then $\pi\colon M \to N$ is a smooth bundle.*

Proof. Fix $p \in N$ and choose complete vector fields $X_1, \ldots, X_n$ on M that project to complete vector fields $Y_1, \ldots, Y_n$ on N and such that $Y_1|_p, \ldots, Y_n|_p$ form a basis of $T_p(N)$. Let φ_t^j and ψ_t^j denote the one-parameter groups generated by X_j and Y_j, respectively. Since π maps the vector field X_j to the vector field Y_j, clearly it maps the integral curves of X_j to those of Y_j and we have $\pi \circ \varphi_t^j = \psi_t^j \circ \pi$. Now the map

$$(\mathbf{R}^n, 0) \to (N, p) \quad \text{given by} \quad (t_1, \ldots, t_n) \mapsto \psi_{t_1}^1 \circ \psi_{t_2}^2 \circ \cdots \circ \psi_{t_n}^n(p)$$

is a diffeomorphism on some neighborhood U of the origin in $\mathbf{R}^n$ to some neighborhood of p in N, and

$$\pi^{-1}(p) \times U \to M \quad \text{defined by} \quad (q, t_1, \ldots, t_n) \mapsto \varphi_{t_1}^1 \circ \varphi_{t_2}^2 \circ \cdots \circ \varphi_{t_n}^n(q)$$

is also a diffeomorphism onto its image and covers the previous map. Thus, it is a bundle chart. ■

3

The Fundamental Theorem of Calculus

Every time we integrate a function $\int_a^b f(x)dx$ we are concerned with a 1-form $f(x)dx$ on an interval $[a,b]$ with values in the Lie algebra of real numbers, and the integral is an element of the Lie group of real numbers. –common room conversation

The main theme of this chapter is the discussion of a *nonabelian* analog of the elementary fundamental theorem of calculus. We now sketch the main ideas followed in studying this theme.

The problem is to characterize the smooth maps $f: M \to G$, where M is a smooth manifold and G is a Lie group with Lie algebra $\mathfrak{g}$ (defined in §2). It turns out that the tangent bundle of G has a canonical trivialization $\omega_G: T(G) \to \mathfrak{g}$. This trivialization may be regarded as a 1-form on G with values in $\mathfrak{g}$ and is called the Maurer–Cartan form. Using this form, we can reinterpret the derivative $f_*: T(M) \to T(G)$ as the composite $\omega_G f_*: T(M) \to \mathfrak{g}$, which "forgets" the images $f(x) \in G$ and remembers only the linear part of the map. This composite $\omega_G f_* = f^*\omega_G$ is a $\mathfrak{g}$-valued 1-form on M called the *Darboux derivative* (cf. §5). It turns out that the Darboux derivative determines the map f up to translation by a constant element of G, the analog of the constant of integration of elementary calculus. The latter part of the chapter analyzes what conditions a $\mathfrak{g}$-valued 1-form ω on M must possess for it to be the Darboux derivative of *some* map $M \to G$. Such a map is an *indefinite integral* or *primitive* of ω. From another point of view, it is a *period mapping* determined by ω.[1] There are two conditions for the existence of a primitive, one local and one global. The local condition is an *integrability condition* (the *structural equation*), which is a special case of the integrability condition for distributions. The global condition is a *monodromy condition*, which is automatically satisfied for simply connected M. Taken together, this material constitutes the *nonabelian fundamental theorem of calculus* (Theorem 7.14) and is indeed fundamental for our later understanding of Cartan's geometries.

The second theme of this chapter is the study of the elementary theory of Lie groups. In particular, we obtain a full picture of the correspondence between Lie groups and Lie algebras.[2]

The third theme of this chapter is the characterization of a Lie group in terms of its Maurer–Cartan form (Theorem 8.7). It is this description that forms the basis for Cartan's generalization of Klein geometries. In particular, it will allow us to classify the Cartan space forms in Chapter 5.

§1. The Maurer–Cartan Form

In Euclidean space, parallel translation of vectors allows us to find a canonical trivialization of the tangent bundle. This notion is so fundamental that some authors *define* two vectors based at different points of $\mathbf{R}^n$ to be the *same* if there is a translation carrying one of them to the other. The possibility of doing this depends on the existence of a group—in this case, the group of translations—acting smoothly and simply transitively on $\mathbf{R}^n$.

Left and Right Translation

The circumstances mentioned above apply to *any* Lie group G, because the *left translation* $L_g: G \to G$ given by $L_g(a) = ga$ is a diffeomorphism (with inverse $L_{g^{-1}}$) so that the induced maps on the tangent spaces

$$L_{g^{-1}*}: T_g(G) \to T_e(G)$$

are all isomorphisms of vector spaces. This yields a canonical trivialization of the tangent bundle $T(G)$. In fact, we could equally well use *right translation* $R_h: G \to G$ given by $R_h(a) = ah$. It follows that there are two ways (which are generally distinct if the group is not abelian) of identifying the space of tangent vectors at any point $g \in G$ with the space of tangent vectors at the identity e. Here is an example showing how these identifications work for the general linear group.

---

[1]See Remark 8.12 at the end of the chapter.

[2]See the end of §2 for an outline of this.

Example 1.1. Regarding $Gl_n(\mathbf{R})$ as an open submanifold of the vector space $M_n(\mathbf{R})$, the "geometric" interpretation of $T(Gl_n(\mathbf{R}))$ discussed in §4 of Chapter 1 yields the identification $T(Gl_n(\mathbf{R})) = Gl_n(\mathbf{R}) \times M_n(\mathbf{R})$. Thus the tangent bundle is manifestly trivial in this way, but the trivialization uses the parallel translation in the vector space $M_n(\mathbf{R})$. We are going to use this trivialization to calculate the one we are really interested in, which arises from the group structure on $Gl_n(\mathbf{R})$. For $g \in Gl_n(\mathbf{R})$, we calculate the derivative $L_{g*}: Gl_n(\mathbf{R}) \times M_n(\mathbf{R}) \to Gl_n(\mathbf{R}) \times M_n(\mathbf{R})$ as follows. Let $(a, v) \in Gl_n(\mathbf{R}) \times M_n(\mathbf{R})$. Then

$$\frac{1}{t}\{L_g(a + tv) - L_g(a)\} = \frac{1}{t}L_g(tv) = gv.$$

Thus $L_{g*}(a, v) = (ga, gv)$. Similarly, $R_{g*}(a, v) = (ag, vg)$. ◆

The Map $\mu_*: T(G \times G) \to T(G)$

According to Exercise 1.4.16 we can identify the tangent bundle of $G \times G$ with $T(G) \times T(G)$ by means of the diffeomorphism $\pi_{1*} \times \pi_{2*}: T(G \times G) \to T(G) \times T(G)$. In the following proposition we use this identification to calculate the derivative of the multiplication $\mu: G \times G \to G$ (cf. Definition 1.1.23).

Proposition 1.2. *Define θ by the commutativity of the following diagram.*

$$\begin{tikzcd} T(G \times G) \arrow[rr, "\pi_{1*} \times \pi_{2*}", "\approx"'] \arrow[dr, "\mu_*"'] & & T(G) \times T(G) \arrow[dl, "\theta"] \\ & T(G) & \end{tikzcd}$$

Then θ satisfies $\theta((g, u) \times (h, v)) = (gh, R_{h*}u + L_{g*}v)$.

Proof. First note that the formula for θ makes sense in that since $u \in T_g(G)$, it follows that $R_{h*}u \in T_{gh}(G)$, and since $v \in T_h(G)$, it follows that $L_{g*}v \in T_{gh}(G)$. The spaces $T(G \times G)$, $T(G) \times T(G)$, and $T(G)$ are all vector bundles, and the three maps are bundle maps that cover the maps in the following diagram.

$$\begin{tikzcd} G \times G \arrow[rr, "\mathrm{id}"] \arrow[dr, "\mu"'] & & G \times G \arrow[dl, "\mu"] \\ & G & \end{tikzcd}$$

Thus, it suffices to verify the formula for $\theta|$, the restriction of θ to an arbitrary fiber, say the fiber over $(g,h) \in G \times G$. Since $\theta|$ is a linear map, it must be of the form

$$\begin{array}{rcl} \theta|\colon T_g(G) \times T_h(G) & \to & T_{gh}(G), \\ (u,v) & \mapsto & A(u)+B(v) \end{array}$$

where the maps A and B are linear (and may vary with g and h). To calculate A and B, we shall make use of the two maps

$$\begin{array}{rcl} \lambda_h\colon G & \to & G \times G, \\ g & \mapsto & (g,h) \end{array} \qquad \begin{array}{rcl} \rho_g\colon G & \to & G \times G. \\ h & \mapsto & (g,h) \end{array}$$

Now λ_h and ρ_g are connected by multiplication to left and right translations by the formulas $\mu \circ \lambda_h = R_h$ and $\mu \circ \rho_g = L_g$. The chain rule applied to the first equation yields the following commutative diagram (where $c_h\colon G \to G$ is the constant map with value h).

$$\begin{tikzcd}
 & T(G) \arrow[ddl, "R_{h*}"'] \arrow[d, "\lambda_{h*}"] \arrow[ddr, "\mathrm{id}_*\times c_{h*}"] & \\
 & T(G\times G) \arrow[dl, "\mu_*"'] \arrow[dr, "\approx"] & \\
T(G) & & T(G)\times T(G) \arrow[ll, "\theta"]
\end{tikzcd}
\qquad
\begin{tikzcd}
 & (g, u) \arrow[dl, mapsto, "R_{h*}"'] \arrow[dr, mapsto, "\mathrm{id}_*\times c_{h*}"] & \\
(gh, R_{h*}u) & \mu_* & (g, u) \times (h, 0) \arrow[ll, mapsto, "\theta"]
\end{tikzcd}$$

It follows that $A(u) = R_{h*}u$. Similarly, $B(v) = L_{g*}v$. ■

Maurer–Cartan Form

Let us continue our study of the trivialization of the tangent bundle of a Lie group determined by *left* translation. Let $\mathfrak{g} = T_e(G)$.

Definition 1.3. Let G be a Lie group. Then the *left-invariant Maurer–Cartan form* $\omega_G\colon T(G) \to \mathfrak{g}$ is defined by $\omega_G(v) = L_{g^{-1}*}(v)$ for $v \in T_g(G)$.[3]

[3]The left-invariant Maurer–Cartan form is the grandfather of all the left-invariant forms on G. Taking the pth exterior power yields a left-invariant $\lambda^p(\mathfrak{g})$-valued p-form on G, $\lambda^p(\omega_G)\colon \lambda^p(T(G)) \to \lambda^p(\mathfrak{g})$. Any left-invariant $\mathbf{R}$-valued p-form may be obtained from this one by composition with some linear map $\lambda^p(\mathfrak{g}) \to \mathbf{R}$. In particular, if $p = n \equiv \dim G$, then $\lambda^n(\mathfrak{g})$ has dimension 1 and $\lambda^n(\omega_G)$ is the Haar measure on G.

The term *left-invariant* refers to the fact that ω_G is invariant under left translation, which may be seen as follows. Since $v \in T_g(G)$ implies that $L_{h*}(v) \in T_{hg}(G)$, we have

$$(L_{h*}\omega_G)v = \omega_G(L_{h*}(v)) = L_{(hg)^{-1}*}(L_{h*}(v)) = L_{g^{-1}*}(v) = \omega_G(v).$$

To see that ω_G is a smooth form, we note that it may be written as a composite of smooth maps

$$\begin{array}{ccccccccc} T(G) & \xrightarrow{\pi\times\mathrm{id}} & G \times T(G) & \xrightarrow{\iota\times\mathrm{id}} & G \times T(G) & \xrightarrow{\sigma\times\mathrm{id}} & T(G) \times T(G) & \xrightarrow{\theta} & T(G). \\ (g,u) & \mapsto & g\times(g,u) & \mapsto & g^{-1}\times(g,u) & \mapsto & (g^{-1},0)\times(g,u) & \mapsto & (e, L_{g^{-1}*}u) \end{array}$$

Example 1.4. $G = \mathbf{R}$. The Maurer–Cartan form is exactly the form dx defined in Proposition 1.5.3. ◆

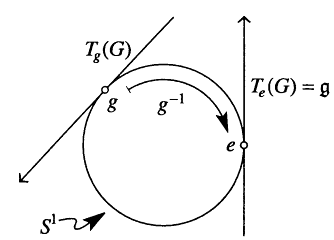
Example 1.5. $G = (\mathbf{R}^+, \cdot)$. Here $e = 1$ and $\omega(x,v) = L_{x^{-1}*}(v) = (1/x)v = (1/x)dx(v)$. Thus $\omega = (1/x)dx$, where dx is as above (but restricted to $\mathbf{R}^+$). ◆

Example 1.6. $G = S^1 = \{z \in \mathbf{C} \mid |z| = 1\}$, so $T(S^1) = \{(e^{i\theta}, ire^{i\theta}) \mid r, \theta \in \mathbf{R}\}$. In particular, $\mathfrak{g} = T_e(S^1) = \{(1, ri) \mid r \in \mathbf{R}\}$. The left action of S^1 on $T(S^1)$ is given by

$$\begin{array}{rcl} S^1 \times T(S^1) & \to & T(S^1). \\ (e^{i\varphi},(e^{i\theta},rie^{i\theta})) & \mapsto & (e^{i(\theta+\varphi)},rie^{i(\theta+\varphi)}) \end{array}$$

Let us calculate the Maurer–Cartan form. We have

$$\omega_G(e^{i\theta}, ire^{i\theta}) = L_{e^{-i\theta}*}(e^{i\theta}, ire^{i\theta}) = (1, ir).$$



This description of the Maurer–Cartan form is *extrinsic* since the group has been regarded as sitting in $\mathbf{C} = \mathbf{R}^2$. For an intrinsic view, see Example 1.10. ◆

Example 1.7. $G = Gl_n(\mathbf{R})$. The Maurer–Cartan form at a point $v \in T_g(G)$ is (cf. Example 1.1) $\omega(v) = L_{g^{-1}*}(g,v) = (e, g^{-1}v)$. Or, identifying the

$T_e(G)$ with $M_n(\mathbf{R})$, we have $\omega(v) = g^{-1}v$ for $v \in T_g(G)$. The classical way of writing the Maurer–Cartan form on $Gl_n(\mathbf{R})$ is $g^{-1}dg$. This has the following meaning. The factor g^{-1} is an abbreviation for $L_{g^{-1}*}$. Regard $g \in Gl_n(\mathbf{R})$ as "the general point," that is, g is the *identity map* on $Gl_n(\mathbf{R})$. Then dg is the identity map on the tangent bundle. If $v \in T_g(Gl_n(\mathbf{R}))$, then $g^{-1}dg(g,v) = g^{-1}(g,v) = (e, g^{-1}v)$; namely, $g^{-1}dg$ is the Maurer–Cartan form. Let us write this out explicitly in the case $n = 2$. (The case $n = 1$ appears above.) Let $g = (x_{ij})$ so that the x_{ij} are the coordinate functions on $Gl_2(\mathbf{R})$. Then $dg = (dx_{ij})$ and $g^{-1}dg = (x_{ij})^{-1}(dx_{ij})$. Writing out the matrices, we have

$$\begin{aligned}\omega &= \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}^{-1} \begin{pmatrix} dx_{11} & dx_{12} \\ dx_{21} & dx_{22} \end{pmatrix} \\ &= \begin{pmatrix} x_{22}/\Delta & -x_{12}/\Delta \\ -x_{21}/\Delta & x_{11}/\Delta \end{pmatrix} \begin{pmatrix} dx_{11} & dx_{12} \\ dx_{21} & dx_{22} \end{pmatrix},\end{aligned}$$

where $\Delta = \det g = x_{11}x_{22} - x_{12}x_{21}$. In its full glory, this is

$$\omega = \begin{pmatrix} (x_{22}dx_{11} - x_{12}dx_{21})/\Delta & (x_{22}dx_{12} - x_{12}dx_{22})/\Delta \\ (-x_{21}dx_{11} + x_{11}dx_{21})/\Delta & (-x_{21}dx_{12} - x_{11}dx_{22})/\Delta \end{pmatrix}.$$

Thus we have represented ω explicitly as a 2×2 matrix of 1-forms defined on $Gl_2(\mathbf{R})$ which is constructed from the coordinate functions x_{ij} (and their exterior derivatives) on $M_2(\mathbf{R})$ restricted to $Gl_2(\mathbf{R})$. Clearly, we can do the same for any $Gl_n(\mathbf{R})$.

Behavior of Maurer–Cartan Forms Under Homomorphism

Proposition 1.8. *Let $\varphi: G_1 \to G_2$ be a homomorphism of Lie groups. Then $\varphi^*\omega_2 = \varphi_{*e}\omega_1$, where the left-hand side is the pullback of the Maurer–Cartan form on G_2 to G_1 via φ, and the right-hand side is the Maurer–Cartan form on G_1 with values interpreted in $\mathfrak{g}_2 = T_e(G_2)$ via the derivative of φ at e, $\varphi_{*e}: T_e(G_1) \to T_2(G_2)$.*

Proof. Let $v \in T_g(G_1)$. Then

$$(\varphi^*\omega_2)v = \omega_2(\varphi_{*g}(v)) = L_{\varphi(g)^{-1}*}(\varphi_{*g}(v)) = \varphi_{*e}(L_{g^{-1}*}(v)) = \varphi_{*g}(\omega_1(v)).$$

All the equalities are obvious except perhaps for the third one. This equality, which depends on the fact that φ is a homomorphism, follows from the commutativity of the following diagram.

$$\begin{tikzcd} G_1 \arrow[r, "\varphi"] \arrow[d, "L_{g^{-1}}"'] & G_2 \arrow[d, "L_{\varphi(g)^{-1}}"] \\ G_1 \arrow[r, "\varphi"] & G_2 \end{tikzcd}$$

■

Corollary 1.9. *Let H be a Lie subgroup of G. Then $\omega_H = \omega_G|H$.*

Example 1.10. Consider the homomorphism $\exp: \mathbf{R} \to S^1$ given by $\exp(\theta) = e^{i\theta} (\theta \in \mathbf{R})$. According to Proposition 1.8, we have $\exp^* \omega_{S^1} = \exp_{*0} \omega_{\mathbf{R}^1}$. But we know that $\omega_{\mathbf{R}^1} = d\theta$ and that $\exp_{*0}$ is an isomorphism. Therefore, identifying the tangent spaces at the identity of $\mathbf{R}$ and S^1 via $\exp_{*0}$, we may write the equation simply as $\exp^* \omega_{S^1} = d\theta$. This may be further reinterpreted as follows. Instead of regarding θ as the identity map on $\mathbf{R}$, let us regard it as the "function" $\theta: S^1 \to \mathbf{R}$ sending $e^{i\theta} \mapsto \theta$. Of course, this is a "multivalued function" with many smooth branches differing from each other by constants, integral multiples of 2π. But, even though θ itself is not a function on S^1, nevertheless $d\theta: T(S^1) \to \mathbf{R}$ is a well-defined 1-form. With this interpretation of $d\theta$, the equation reads $\exp^* \omega_{S^1} = \exp^* d\theta$, i.e., $\omega_{S^1} = d\theta$. ◆

Example 1.11. $SO_n(\mathbf{R})$. Now $SO_n(\mathbf{R})$ is a Lie subgroup of $Gl_n(\mathbf{R})$. By Corollary 1.9, the Maurer–Cartan form on $SO_n(\mathbf{R})$ is just the restriction of the Maurer–Cartan form on $Gl_n(\mathbf{R})$. Thus, as above, it may be written as $g^{-1}dg$, and the explicit coordinate expression is the same as above. Similar statements hold for all Lie subgroups of $Gl_n(\mathbf{R})$. ◆

§2. Lie Algebras

Although the axioms for a Lie group G are quite simple, they are very strong. In this section we show how, by combining these axioms and using the various maps which they guarantee, we obtain a *multiplication* on the tangent space $T_e(G)$ which turns it into a (nonassociative) algebra, the *Lie algebra* of G. Our practice will be to designate the Lie algebra of G by $\mathfrak{g}$, of H by $\mathfrak{h}$, and so forth. The multiplication is constructed by identifying $\mathfrak{g}$ with the vector space of left-invariant vector fields on G which can be "multiplied" by the bracket operation. The Lie algebra $\mathfrak{g}$ should be regarded as an infinitesimal version of the group,[4] as it turns out to encode and control *almost* every aspect of the group G (cf. the "primer" at the end of this section).

Left-Invariant Vector Fields

Definition 2.1. A *left-invariant* vector field on the Lie group G is a vector field satisfying either of the properties in the following lemma. ❀

Lemma 2.2. *Let G be a Lie group and let X be a vector field on G. The following two properties of X are equivalent:*

[4] Indeed, Lie called it an "infinitesimal group."

(i) *$\omega_G(X)$ is a constant (as a $\mathfrak{g}$-valued function on G);*

(ii) *$L_{g*}X_a = X_{ga}$ for all $a, g \in G$.*

Proof.

$$\omega_G(X) \text{ is constant} \Leftrightarrow L_{a^{-1}*}(X_a) = L_{(ga)^{-1}*}(X_{ga})\ \forall a, g \in G$$
$$\text{(applying } L_{ga*} \text{ to both sides)} \Leftrightarrow L_{g*}(X_a) = X_{ga} \forall a, g \in G. \quad \blacksquare$$

Proposition 2.3.

(i) *The left-invariant vector fields form a vector subspace of the real vector space of all vector fields on G.*

(ii) *The mapping*

$$\begin{array}{ccc} \left\{ \begin{array}{c} \text{vector space of} \\ \text{left-invariant} \\ \text{vector fields on } G \end{array} \right\} & \longrightarrow & \mathfrak{g} \\ X & \longmapsto & X_e \end{array}$$

is a linear isomorphism.

(iii) *If X and Y are left-invariant vector fields, then so is $[X, Y]$.*

Proof. Part (i) is an obvious consequence of Lemma 2.2(i). For part (ii), note that the mapping is clearly linear. Also, if $X_e = 0$, then $0 = L_{g*}X_e = X_g$ by Lemma 2.2(ii). Moreover, if $v \in \mathfrak{g}$, we may define $X_g = L_{g*}v$, so that

$$L_{a*}X_g = L_{a*}L_{g*}v = L_{ag*}v = X_{ag}.$$

Thus, X is left invariant with $X_e = v$, and so the map is surjective. Finally, for part (iii) we note that the left invariance of X may be rephrased by saying that X is L_g related to itself for all $g \in G$, and the same may be said of Y. By Lemma **1**.4.22, the bracket $[X, Y]$ is also L_g related to itself for all $g \in G$; that is, $[X, Y]$ is left invariant. ■

Definition 2.4. If $v \in \mathfrak{g}$, the Lie algebra of G, then $v^\dagger$ denotes the corresponding left-invariant vector field on G. ❀

Lie Algebras

The vector space $\mathfrak{g}$ can be given a multiplication. Two vectors $u, v \in \mathfrak{g}$ determine two left-invariant vector fields $u^\dagger$ and $v^\dagger$ on G from which we may form the commutator $[u^\dagger, v^\dagger]$, which, by Proposition 2.3 (iii), is also left invariant. Then $[u, v] \in \mathfrak{g}$ is defined to be the value of the commutator at the origin. The vector space $\mathfrak{g}$, equipped with this multiplication, is called the *Lie algebra* of G.

Proposition 2.5. *The multiplication $[\ ,\ ]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfies the following properties.*

(i) (Skew symmetry) $[u, v] = -[v, u]$, $\forall u, v \in \mathfrak{g}$.

(ii) (Bilinearity) $[au + bv, w] = a[u, w] + b[v, w]$, $\forall u, v, w \in \mathfrak{g}$, *and* $\forall a, b \in \mathbf{R}$.

(iii) (Jacobi identity) $[[u, v], w] + [[v, w], u] + [[w, u], v] = 0$.

Proof. These are all immediate consequences of the properties of brackets of vector fields. ■

We now give the *abstract* notion of a Lie algebra which does not a priori arise as the Lie algebra of a Lie group (although cf. Theorem 7.20).

Definition 2.6. A *Lie algebra* consists of a finite-dimensional real[5] vector space $\mathfrak{g}$ together with a multiplication $[\ ,\ ]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ satisfying the three properties of Proposition 2.5. A *homomorphism* of Lie algebras is a linear map $\varphi: \mathfrak{g}_1 \to \mathfrak{g}_2$ that preserves the multiplication, namely

$$[\varphi(u), \varphi(v)]_{\mathfrak{g}_2} = \varphi([u, v]_{\mathfrak{g}_1}). \quad ❀$$

Definition 2.7. A *subalgebra* of a Lie algebra $\mathfrak{g}$ is a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ satisfying $[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}$. An *ideal* of a Lie algebra $\mathfrak{g}$ is a subalgebra $\mathfrak{h} \subset \mathfrak{g}$ satisfying $[\mathfrak{h}, \mathfrak{g}] \subset \mathfrak{h}$. ❀

Example 2.8. The Lie algebra of $Gl_n(\mathbf{R})$ is denoted by $\mathfrak{gl}_n(\mathbf{R})$. As a vector space it is just $T_e(Gl_n(\mathbf{R}))$, which we have identified with $M_n(\mathbf{R})$, the vector space of $n \times n$ matrices. To find the multiplication, we note that if $A = (a_{ij}) \in M_n(\mathbf{R})$, then the left-invariant vector field corresponding to

[5]The restriction to vector spaces over the *real* field is of course unnecessary. Lie algebras over the complex numbers or even arbitrary fields are commonly considered. The case of infinite-dimensional Lie algebras has also attracted much interest.

it is given by $A_g = (g, gA)$. As a derivation of functions (writing $g = (x_{ij})$), we have

$$A_g = \sum_{i,j,k} x_{ij} a_{jk} \frac{\partial}{\partial x_{ik}}.$$

Now let us calculate the commutator of two such vector fields corresponding to $A = (a_{ij})$ and $B = (b_{ij}) \in M_n(\mathbf{R})$. We have

$$\begin{aligned}\left[\sum_{i,j,k} x_{ij} a_{jk} \frac{\partial}{\partial x_{ik}}, \sum_{p,q,r} x_{pq} b_{qr} \frac{\partial}{\partial x_{pr}}\right] \\ = \sum_{\substack{i,j,k\\p,q,r}} x_{ij} a_{jk} \frac{\partial x_{pq}}{\partial x_{ik}} b_{qr} \frac{\partial}{\partial x_{pr}} - \sum_{\substack{i,j,k\\p,q,r}} x_{pq} b_{qr} \frac{\partial x_{ij}}{\partial x_{pr}} a_{jk} \frac{\partial}{\partial x_{ik}}.\end{aligned}$$

Now terms in the first sum on the right vanish unless $p = i$ and $q = k$. Similarly, terms in the second sum vanish unless $p = i$ and $r = j$. Thus, after this is done, we in the first sum we may replace p by i and q by k. After this is done, we exchange the indices $k \leftrightarrow r$. In the second sum we replace p by i and r by j. After this is done, we exchange the indices $q \leftrightarrow j$ and then replace the new q by r. Then the right-hand side becomes

$$\sum_{i,k} \left\{ \sum_j x_{ij} \sum_r (a_{jr} b_{rk} - b_{jr} a_{rk}) \right\} \frac{\partial}{\partial x_{ik}}.$$

Clearly, this is the left-invariant vector field corresponding to the matrix $AB - BA$. Thus, the multiplication in the Lie algebra of matrices $M_n(\mathbf{R})$ is given by the formula $[A, B] = AB - BA$. ❀

Proposition 2.9.

(i) *The derivative at the identity of a homomorphism of Lie groups $\varphi: G_1 \to G_2$ is a homomorphism of the corresponding Lie algebras $\varphi_{*e}: \mathfrak{g}_1 \to \mathfrak{g}_2$.*

(ii) *If $\varphi: G_1 \to G_2$ and $\psi: G_2 \to G_3$ are homomorphisms of Lie groups, then*
$$(\psi \circ \varphi)_{*e} = \psi_{*e} \circ \varphi_{*e}.$$

(iii) *If id: $G \to G$ is the identity map, then so is $id_{*e}: \mathfrak{g} \to \mathfrak{g}$.*

(iv) *If φ is an isomorphism of Lie groups, then φ_* is an isomorphism of Lie algebras.*

(v) *If G_1 is connected and φ_* is an isomorphism of Lie algebras, then* ker φ *is a discrete central subgroup of G.*

Proof. (i) Let $u, v \in \mathfrak{g}_1$. By Proposition 2.3(ii) we may uniquely extend these vectors to left-invariant vector fields X, Y on G_1 (i.e., such that $X_e = u$, $Y_e = v$). Set $u' = \varphi_*(u)$ and $v' = \varphi_*(v)$ in $\mathfrak{g}_2$ and extend these to left-invariant vector fields X' and Y' on G_2. Since φ is a homomorphism, it follows that $\varphi \circ L_g = L_{\varphi(g)} \circ \varphi$. Differentiating this equation yields $\varphi_* \circ L_{g*} = L_{\varphi(g)*} \circ \varphi_*$, so that

$$\varphi_*(X_g) = \varphi_*(L_{g*}u) = L_{\varphi(g)*}(\varphi_*(u)) = L_{\varphi(g)*}(u') = X'_{\varphi(g)}.$$

Thus, X and X' are φ related. Similarly, Y and Y' are φ related. It follows from Lemma **1**.4.22 that $[X, Y]$ and $[X', Y']$ are also φ-related. In particular, $\varphi_*([u, v]) = [\varphi_* u, \varphi_* v]$.

(ii) This is just the chain rule (Theorem **1**.4.7b).

(iii) This is obvious from the definition of the derivative.

(iv) From (ii) and (iii) we see that if the inverse of φ is ψ, then the inverse of φ_{*e} is ψ_{*e}.

(v) From (i) we have $\varphi_{*g} = L_{\varphi(g)*e} \circ \varphi_{*e} \circ L_{g*e}^{-1}$ for all g so that φ_{*g} is an isomorphism for all g. It follows that φ is a local diffeomorphism at each point and hence that $\varphi^{-1}(e)$ is discrete. ■

Corollary 2.10. *The Lie algebra of a subgroup H of a Lie group G is a subalgebra of the Lie algebra of G.*

Corollary 2.10 makes it possible to describe the structure of the Lie algebras of any subgroup of $Gl_n(\mathbf{R})$; for example, any of the subgroups given at the end of Chapter 1. In each case the bracket operation on the corresponding Lie algebra of matrices is just the commutator $[A, B] = AB - BA$.

One-Dimensional Lie Subgroups

One-dimensional Lie groups have quite special properties and are important, as they appear in large quantities as subgroups of general Lie groups.

Proposition 2.11. *Let G be a Lie group with Lie algebra $\mathfrak{g}$, and let $v \in \mathfrak{g}$. Then there is a unique homomorphism of Lie groups $\varphi: \mathbf{R} \to G$ such that $\varphi_{*e}(d/dt) = v$.*

Proof. By Proposition 2.3(ii), there is a left-invariant vector field X corresponding to v (i.e., so that $X_e = v$). Let $\varphi: ((a, b), 0) \to (G, e)$ be a maximal integral curve for X through zero. For $c \in (a, b)$, we consider the function $f_c: (a + c, b + c) \to G$ defined by $f_c(t) = \varphi(c)\varphi(t - c)$. Now we have

$$f_{c*}\left(\left.\frac{d}{dt}\right|_{t_0}\right) = L_{\varphi(c)*}\circ\varphi_*\left(\left.\frac{d}{dt}\right|_{t_0-c}\right) = L_{\varphi(c)*}(X_{\varphi(c)\varphi(t_0-c)}) = X_{f(t_0)}.$$

From this equation we can draw the following two conclusions.

First, the equation says that f_c is an integral curve for X through $f(c) = \varphi(c)$, so by uniqueness $f(t) = \varphi(t)$ for all t in the common domain. But then we may patch these integral curves together to get an integral curve with a larger domain of definition than that of φ. In particular, by taking $c > 0$ we see that b can be made bigger, and by taking $c < 0$ we see that a can be made smaller. These yield contradictions unless $a = -\infty$ and $b = \infty$. Thus $\varphi: (\mathbf{R}, 0) \to (G, e)$.

Second, the equation says that $\varphi(s)\varphi(t) = L_{\varphi(s)}\varphi((s+t)-s) = f_s(s+t) = \varphi(s+t)$, so that φ is a homomorphism. Since it is also smooth, it is a homomorphism of Lie groups. ■

Corollary 2.12. *Every left-invariant vector field on a Lie group is complete. (Of course, the same is also true for right-invariant vector fields.)*

Proof. The integral curve through e is a homomorphism $\varphi: \mathbf{R} \to G$; in particular, it is defined for all parameter values. Now consider the curve $L_g\varphi: \mathbf{R}, 0 \to G, g$. We have

$$(L_g\varphi)_*\left(\left.\frac{d}{dt}\right|_{t_0}\right) = L_{g*}\circ\varphi_*\left(\left.\frac{d}{dt}\right|_{t_0}\right) = L_{g*}(X_{\varphi(t_0)}) = X_{L_g\varphi(t_0)}$$

so that $L_g\varphi$ is an integral curve through g defined for all parameter values. Thus, the left-invariant vector field X is complete. ■

Example 2.13. Consider $G = \mathbf{R}^2/\mathbf{Z}^2$, where $\mathbf{Z}^2$ is the lattice of integer points in $\mathbf{R}^2$. This is a compact Lie group that is not simply connected. The universal cover is $\tilde{G} = \mathbf{R}^2$. The covering projection is $\pi: \tilde{G} \to G$, sending $(x, y) \mapsto (e^{2\pi i x}, e^{2\pi i y})$. It is clear that in $\tilde{G}$ a one-dimensional subgroup $\tilde{H}$ is just a one-dimensional subspace of $\mathbf{R}^2$; in particular, they are all closed subgroups of $\mathbf{R}^2$. However, the image $\pi(\tilde{H}) = H$ of such a subgroup (which can be any one-dimensional subgroup of G) is not necessarily closed. In fact, a necessary and sufficient condition for a subgroup $\pi(\tilde{H}) = H$ to be closed is that the line $\tilde{H}$ have rational slope in the standard coordinates of $\mathbf{R}^2$. ◆

Example 2.14. The group G in the preceding example sits as a closed (diagonal) subgroup in the simply connected group $Sl_3(\mathbf{C})$ via the homomorphism $(e^{ix}, e^{iy}) \to \mathrm{diag}(e^{ix}, e^{iy}, e^{-i(x+y)})$, so this latter group also contains one-dimensional subgroups that are not closed. ◆

Exercise 2.15.* Let P be a smooth manifold, H a Lie group, and $\mu: P \times H \to P$ a smooth right action. If $X \in \mathfrak{h}$ (the Lie algebra of H), we define a vector field $X^\dagger$ on P by $(X^\dagger)_p = \mu_{*(p,e)}(0, X)$. (This is the vector field that describes how each point $p \in P$ moves under the action of "the infinitesimal element $X \in H$.")

(a) Show that in the case where $P = G$ is a Lie group with $H \subset G$ a subgroup, and $\mu: G \times H \to G$ is the group multiplication, that $X^\dagger$ here agrees with that of Definition 2.4.

(b) Let Q be another smooth manifold and $\nu: Q \times H \to Q$ be another smooth right action. Suppose that $\phi: P \to Q$ is H equivariant, that is, the following diagram commutes.

$$\begin{array}{ccc} P \times H & \xrightarrow{\mu} & P \\ \phi \times \mathrm{id} \downarrow & & \downarrow \phi \\ Q \times H & \xrightarrow{\nu} & Q \end{array}$$

Show that the vector fields on P and Q corresponding to $X \in \mathfrak{h}$ and arising from the actions of H are ϕ related. ◆

A Primer on the Lie Group–Lie Algebra Correspondence

In this subsection we give a preview of certain facts concerning Lie groups and algebras whose proofs are consequences of the fundamental theorem of calculus, which we will supply later in the chapter, and Ado's theorem.

Given a Lie group G, the universal cover $\tilde{G}$ exists. By Corollary 8.11 $\tilde{G}$ has a unique[6] Lie group structure such that the projection map is a homomorphism. Moreover, the kernel of this homomorphism is a discrete central subgroup of the universal cover by Proposition 2.9(v). By Proposition 2.9 this homomorphism induces an isomorphism of the corresponding Lie algebras. The simplest example of this situation is the homomorphism $\mathbf{R} \to S^1$ sending $\theta \mapsto e^{i\theta}$.

The correspondence that associates to each Lie group its Lie algebra (Proposition 2.5) and to each homomorphism of Lie groups the corresponding homomorphism of Lie algebras (Proposition 2.9) is a *functor* from the category of Lie groups and homomorphisms to the category of Lie algebras and homomorphisms. It turns out that this functor is *onto* in the sense that every abstract Lie algebra is realized as the Lie algebra of some Lie group (Proposition 7.20), and every homomorphism of Lie algebras is the derivative of some homomorphism between some realization of the algebras (Proposition 7.15).

[6]It is not quite unique. There is no uniquely determined identity element. Any point lying over the identity element of G may serve as the identity. Once this choice is made, the rest of the Lie group structure is determined.

However, the functor is not injective, as we saw earlier, since a Lie group always has the same Lie algebra as its universal cover. Nevertheless, for connected Lie groups this is the *only* thing that can go wrong. Two connected Lie groups with the same Lie algebra have the same universal cover (Proposition 7.20). If we restrict the functor to the Lie groups G that are connected and simply connected, it becomes an equivalence of categories.

For subgroups, it may be shown [H. Yamabe, 1950] that if G is a Lie group and H is *any* subgroup, then there is a unique smooth structure on H such that the inclusion map is a weak embedding, and H is a Lie group with this smooth structure.

§3. Structural Equation

Let G be a Lie group with Lie algebra $\mathfrak{g}$. The *structural equation* for the Maurer–Cartan form leads to the first of two necessary conditions that a $\mathfrak{g}$-valued 1-form on a manifold M be the Darboux derivative (cf. §5) of a map $M \to G$.

Let us calculate the exterior derivative of the Maurer–Cartan form ω_G of a Lie group G. The calculation is merely an application of the general formula

$$d\omega_G(X,Y) = X(\omega_G(Y)) + Y(\omega_G(X)) - \omega_G([X,Y]),$$

derived in Chapter 1 (Lemma **1**.5.15), which holds for any 1-form and any pair of vector fields. In our case we take X and Y to be left-invariant fields so that $\omega_G(X)$ and $\omega_G(Y)$ are constant. Then the first two terms on the right vanish. As for the last term, since X and Y are left invariant, the bracket $[X,Y]$ is also left invariant, and so we have $\omega_G([X,Y]) = [X,Y]_e$. But $[X,Y]_e = [X_e, Y_e]$, where, by definition, the bracket on the right is the Lie bracket in $\mathfrak{g}$. Thus the last term is $-[\omega_G(X), \omega_G(Y)]$. This yields the following equation, known as the *structural equation*:

$$\boxed{d\omega_G(X,Y) + [\omega_G(X), \omega_G(Y)] = 0}$$

We have derived this equation for left-invariant fields X and Y. But it is a linear equation relating 2-forms, and hence it must hold for any pair of vectors $u, v \in T_g(G)$ that are the restrictions of left-invariant vector fields on G. But this is true for arbitrary vectors u and v, since by Proposition 2.3(ii) any vector may be extended by left translation to a left-invariant vector field of G. Thus, the structural equation holds for arbitrary vector fields X and Y.

The structural equation may also be written (Lemma **1**.5.21) as

$$d\omega_G + \frac{1}{2}[\omega_G, \omega_G] = 0.$$

The structural equation may be thought of as merely a formula for the exterior derivative of the Maurer–Cartan form. But it is much more than just that. In Theorem 8.7 we show that it provides a local characterization of a Lie group. Thus, it determines the structure of the group locally and can be regarded as a *fundamental defining property* of the Lie group.

If G is abelian, the second term in the structural equation vanishes and the equation reads $d\omega_G = 0$.

Example 3.1. $G = Gl_n(\mathbf{R})$. Let us check the structural equation using the explicit formula for the Maurer–Cartan form $\omega_G = (x_{ij})^{-1}(dx_{ij})$. We work in the differential algebra of $n \times n$ matrices over the ring of $\mathbf{R}$-valued differential forms on $M_n(\mathbf{R})$. We may write

$$(x_{ij})\omega_G = (dx_{ij}).$$

Taking the exterior derivative of this equation of matrices of 1-forms kills the right-hand side and leaves

$$(dx_{ij}) \wedge \omega_G + (x_{ij})d\omega_G = 0.$$

Now replace (dx_{ij}) by $(x_{ij})\omega_G$ to get

$$(x_{ij})\omega_G \wedge \omega_G + (x_{ij})d\omega_G = 0.$$

Since (x_{ij}) is invertible, we finally get

$$d\omega_G + \omega_G \wedge \omega_G = 0.$$

To compare this with the original structural equation, we write $\omega_G = (\omega_{ij})$ so that

$$\omega_G \wedge \omega_G = (\omega_{ij}) \wedge (\omega_{ij}) = (\Sigma_k \omega_{ik} \wedge \omega_{kj}).$$

Now

$$\Sigma_k \omega_{ik} \wedge \omega_{kj}(X,Y) = \Sigma_k \{\omega_{ik}(X)\omega_{kj}(Y) - \omega_{ik}(Y)\omega_{kj}(X)\}.$$

So[7]

$$(\Sigma_k \omega_{ik} \wedge \omega_{kj})(X,Y) = (\omega_{ik}(X))(\omega_{kj}(Y)) - (\omega_{ik}(Y))(\omega_{kj}(X))$$

or

$$\begin{aligned}\omega_G \wedge \omega_G(X,Y) &= \omega_G(X)\omega_G(Y) - \omega_G(Y)\omega_G(X)\\ &= [\omega_G(X), \omega_G(Y)].\end{aligned}$$

[7]Here $(\omega_{ij}(X))$ denotes the matrix with components $\omega_{ij}(X)$.

§4. Adjoint Action

The *adjoint action* is a bookkeeping device that allows us to compare the left and right translations. More precisely, it is the derivative of the conjugation action of G on itself. It will allow us (cf. Proposition 4.10) to pull back the Maurer–Cartan form via the multiplication and inversion maps μ and ι. Let us see how it arises.

A Lie group G acts on itself on the left by conjugation

$$\underset{(g,h)\mapsto \mathbf{Ad}(g)h=ghg^{-1}}{G\times G\to G.}$$

This smooth map induces, for each $g\in G$, the homomorphism of Lie groups

$$\underset{h\mapsto ghg^{-1}}{\mathbf{Ad}(g)\colon G\to G,}$$

called the *inner automorphism induced by* g. The map

$$\underset{g\mapsto \mathbf{Ad}(g)}{\mathbf{Ad}\colon G\to \mathrm{Aut}(G)}$$

is called the (*left*) *adjoint action of* G *on itself*. It is not necessary to concern ourselves with the question of whether $\mathrm{Aut}(G)$ is a manifold (it is!), so we shall put the map $\mathbf{Ad}$ in the background and concentrate on the maps $\mathbf{Ad}(g)$. For each $g\in G$ this map is a isomorphism of Lie groups, and so by Proposition 2.9 its derivative at the identity is an isomorphism of Lie algebras. It is this "infinitesimal" version of the adjoint action that interests us.

Definition 4.1. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Define $\mathrm{Ad}(g)=\mathbf{Ad}(g)_{*e}\in Gl(\mathfrak{g})$. The (*left*) *adjoint representation*[8] is the map $\mathrm{Ad}\colon G\to Gl(\mathfrak{g})$ sending $g\mapsto\mathrm{Ad}(g)$. Equivalently, we may speak of the (*left*) *adjoint action of* G *on* $\mathfrak{g}$ as the map $G\times\mathfrak{g}\to\mathfrak{g}$ sending $(g,v)\mapsto\mathrm{Ad}(g)v$. ❀

Proposition 4.2.

(a) $\mathrm{Ad}(g)\colon\mathfrak{g}\to\mathfrak{g}$ *is an isomorphism of Lie algebras, and the map* $\mathrm{Ad}\colon G\to Gl(\mathfrak{g})$ *is a homomorphism of Lie groups.*

(b) *If* X *is a left-invariant vector field on* G*, then so is* $\mathbf{Ad}(g)_*X$ *for all* $g\in G$.

(c) $R_g^*\omega_G=\mathrm{Ad}(g^{-1})\omega_G$. (*Note that* $\mathrm{Ad}(g^{-1})$ *acts on the values of* ω_G *and these values lie in* $\mathfrak{g}$.)

[8]If G is a Lie group and V is a (finite-dimensional) real or complex vector space, then a (finite dimensional) representation of G on V is a homomorphism of Lie groups $\rho\colon G\to Gl(V)$.

Proof. (a) This is just a special case of Proposition 2.9(i).

(b) We calculate

$$\begin{aligned}L_{h*}(\mathbf{Ad}(g)_*X)&=L_{h*}\circ L_{g*}\circ R_{g^{-1}*}(X)\\&=R_{g^{-1}*}\circ L_{h*}\circ L_{g*}(X)\\&=R_{g^{-1}*}(X)\\&=R_{g^{-1}*}\circ L_{g*}(X)=\mathbf{Ad}(g)_*X.\end{aligned}$$

(c) Let $v\in T_h(G)$ so that $R_{g*}(v)\in T_{gh}(G)$. Then

$$\begin{aligned}R_g^*\omega_G(v)&=\omega_G(R_{g*}v)=L_{(hg)^{-1}*}R_{g*}v=(L_{g^{-1}*}R_{g*})(L_{h^{-1}*}v)\\&=\mathrm{Ad}(g^{-1}\omega_G(v).\end{aligned}$$ ■

Example 4.3. $G=Gl_n(\mathbf{R})$. Let us compute the adjoint action in this case. We have $\mathbf{Ad}(g)=L_g\circ R_{g^{-1}}$. Thus $\mathbf{Ad}(g)_*X=L_{g*}\circ R_{g^{-1}*}X$. By Example 1.1, calculating the derivatives L_{g*} and $R_{g^{-1}*}$ shows us that $\mathbf{Ad}(g)_*\colon M_n(\mathbf{R})\to M_n(\mathbf{R})$ sends $v\mapsto gvg^{-1}$. ◆

Definition 4.4. For $\mathfrak{g}$ a Lie algebra,

(i) $\mathrm{Aut}_{\mathrm{Lie}}(\mathfrak{g})=\{T\in Gl(\mathfrak{g})\mid T[u,v]=[Tu,Tv],\ \forall u,v\in\mathfrak{g}\}$,

(ii) $\mathfrak{gl}_{\mathrm{Lie}}(\mathfrak{g})=\{T\in\mathfrak{gl}(\mathfrak{g})\mid T[u,v]=[Tu,v]+[u,Tv],\ \forall u,v\in\mathfrak{g}\}$. ❀

Exercise 4.5. Show that

(a) $\mathrm{Aut}_{\mathrm{Lie}}(\mathfrak{g})$ is a Lie group.

(b) the Lie algebra of $\mathrm{Aut}_{\mathrm{Lie}}(\mathfrak{g})$ is $\mathfrak{gl}_{\mathrm{Lie}}(\mathfrak{g})$.

(c) Let $\mathfrak{g}$ be the Lie algebra of G. Show the derivative of $\mathrm{Ad}\colon G\to\mathrm{Aut}_{\mathrm{Lie}}(\mathfrak{g})$ at e is the map $\mathrm{ad}\colon\mathfrak{g}\to\mathfrak{gl}_{\mathrm{Lie}}(\mathfrak{g})$ defined by $\mathrm{ad}(u)v=[u,v]$. (Hence G abelian $\Leftrightarrow$ $\mathbf{Ad}$ trivial $\Rightarrow$ Ad trivial $\Rightarrow$ ad trivial $\Rightarrow$ $[\ ,\ ]$ trivial.) ❏

Corollary 4.6.

(i) *The map* $\mathrm{Ad}\colon G\to Gl(\mathfrak{g})$ *takes its values in* $\mathit{Aut}_{\mathrm{Lie}}(\mathfrak{g})$.

(ii) $\mathbf{Ad}(\mathrm{Ad}(g))(ad)=ad\circ\mathrm{Ad}(g)$.

Proof. (i) This is just Proposition 4.2(a).
(ii) By (i) we have $\mathrm{Ad}(g)[u,v] = [\mathrm{Ad}(g)u, \mathrm{Ad}(g)v]$, or $\mathrm{Ad}(g)[u, \mathrm{Ad}(g^{-1})v] = [\mathrm{Ad}(g)u, v]$, so that $\mathbf{Ad}(\mathrm{Ad}(g))(\mathrm{ad}(u)) = \mathrm{ad}(\mathrm{Ad}(g)u)$. ■

Exercise 4.7.* Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let $H \subset G$ be a closed subgroup with Lie algebra $\mathfrak{h}$. Show that

(a) if H is normal in G, then $\mathfrak{h}$ is an ideal of $\mathfrak{g}$,

(b) if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$ and if H and G are connected, then H is normal in G,

(c) if H is normal in G, then $(\mathrm{Ad}(H) - I)\mathfrak{g} \subset \mathfrak{h}$. ❑

Exercise 4.8.*

(a) Show that $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{g})$ is an H module under the action

$$(h\varphi)(v,w) = \mathrm{Ad}(h)\varphi(\mathrm{Ad}(h^{-1})v, \mathrm{Ad}(h^{-1})w), \quad h \in H.$$

(b) Show that if H is connected, then

$$\mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}, \mathfrak{g}) \stackrel{\text{def}}{=} \{\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{g} \mid \mathrm{ad}(u)\varphi(v,w) = \varphi(\mathrm{ad}(u)v, w) + \varphi(v, \mathrm{ad}(u)w)\}$$

is the submodule of H invariant elements in $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}, \mathfrak{g})$ under the action given in (a).

(c) For the case of the Euclidean algebra $\mathfrak{g} = \mathfrak{euc}_n(\mathbf{R})$, $\mathfrak{h} = \mathfrak{so}_n(\mathbf{R})$, take the basis $e_i = E_i$, $e_{ij} = E_{ij} - E_{ji}$. Write $\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ as $\varphi = \sum a_{ijkl} e_i^* \wedge e_j^* \otimes e_{kl}$, where a_{ijkl} is skew symmetric in each of the first and last pairs of indices. Show that $\varphi \in \mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \Leftrightarrow a_{ijij} = c$ for some c and a_{ijkl} vanishes when $\{i,j\} \neq \{k,l\}$. [*Hint:* Treat the cases $n = 2$, $n = 3$, and $n > 3$ separately.] (This result will be basic for our study of constant-curvature Riemannian spaces in Chapter 6.) Show also that $\mathrm{Hom}_{O_n(\mathbf{R})}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) = \mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$. ❑

Exercise 4.9.* Show that the adjoint representation $\mathrm{Ad}: G \to Gl(\mathfrak{g})$ pulls back the Maurer–Cartan form $\omega_{Gl(\mathfrak{g})}$ according to $\mathrm{Ad}^*\omega_{Gl(\mathfrak{g})} = \mathrm{ad}(\omega_G)$. [Note: $\mathrm{ad}(\omega_G)$ is the $\mathfrak{gl}_{\mathrm{Lie}}(\mathfrak{g})$-valued form on G given by $\mathrm{ad}(\omega_G)v = \mathrm{ad}(\omega_G(v))$ for $v \in T(G)$. In more detail, this reads $(\mathrm{ad}(\omega_G)v)w = [\omega_G(v), w]$ for $v \in T(G)$ and $w \in \mathfrak{g}$.] ❑

Product and Quotient Rules

Using the adjoint action of G on $\mathfrak{g}$, we obtain the following fundamental formulas showing how the Maurer–Cartan form behaves with respect to multiplication and inversion.

Proposition 4.10. *Let $\mu: G \times G \to G$ be multiplication and $\iota: G \to G$ be the inverse. Then*[9]

(i) $\mu^*\omega_G = (\pi_2^*\mathrm{Ad}^{-1})(\pi_1^*\omega_G) + \pi_2^*\omega_G$, i.e.,
$\mu^*\omega_G(w) = \mathrm{Ad}^{-1}(h)(\omega_G(\pi_{1*}w)) + \omega_G(\pi_{2*}w)$ *for* $w \in T_{(g,h)}(G \times G)$,

(ii) $\iota^*\omega_G = -\mathrm{Ad}\,\omega_G$, i.e.,
$\iota^*\omega_G(v) = -\mathrm{Ad}(g)\omega_G(v)$ *for* $v \in T_g(G)$.

Proof. Recall from Proposition 1.2 the commutative diagram

$$\begin{tikzcd} T(G\times G) \arrow[rr, "\pi_{1*}\times\pi_{2*}", "\approx"'] \arrow[dr, "\mu_*"'] & & T(G)\times T(G) \arrow[dl, "\theta"] \\ & T(G) & \end{tikzcd}$$

where $\theta((g,u) \times (h,v)) = (gh, R_{h*}u + L_{g*}v)$. Let $w \in T(G \times G)$ correspond under $\pi_{1*} \times \pi_{2*}$ to the element $(g,u) \times (h,v) \in T(G) \times T(G)$. We calculate

$$\begin{aligned}
(\mu^*\omega_G)w &= \omega_G(\mu_* w) \\
&= \omega_G\theta((\pi_{1*} \times \pi_{2*})w) \\
&= \omega_G\theta((g,u) \times (h,v)) \\
&= \omega_G(gh, R_{h*}u + L_{g*}v) \\
&= L_{(gh)^{-1}*}(R_{h*}u + L_{g*}v) \\
&= L_{h^{-1}*}L_{g^{-1}*}R_{h*}u + L_{h^{-1}*}L_{g^{-1}*}L_{g*}v \\
&= L_{h^{-1}*}R_{h*}L_{g^{-1}*}u + L_{h^{-1}*}v \\
&= L_{h^{-1}*}R_{h*}\omega_G(u) + \omega_G(v) \\
&= \mathrm{Ad}(h^{-1})\omega_G(u) + \omega_G(v) \\
&= \mathrm{Ad}(h^{-1})\omega_G(\pi_{1*}w) + \omega_G(\pi_{2*}w) \\
&= \pi_1^*(\mathrm{Ad}(h^{-1})\omega_G)w + \pi_2^*(\omega_G)w,
\end{aligned}$$

[9] Of course, $\pi_i: G \times G \to G$ denotes projection to the ith factor, $\mathrm{Ad}^{-1}: G \to \mathrm{Aut}(\mathfrak{g})$ is the map sending $g \to \mathrm{Ad}(g)^{-1}$, and $\pi_2^*\mathrm{Ad}^{-1}$ acts on the *values* of $\pi^*\omega_G$.

which verifies (i). To see how (ii) is a consequence of (i), consider the composite map

$$\begin{array}{ccccccc} G & \xrightarrow{\Delta} & G \times G & \xrightarrow{\mathrm{id}\times\iota} & G \times G & \xrightarrow{\mu} & G. \\ g & \mapsto & (g,g) & \mapsto & (g,g^{-1}) & \mapsto & 1 \end{array}$$

Now, this composite is the constant map, so the pullback of the Maurer–Cartan form vanishes. Thus $((\mathrm{id}\times\iota)\Delta)^*\mu^*\omega_G = 0$, and so by (i),

$$\begin{aligned} 0 &= ((\mathrm{id}\times\iota)\Delta)^*(\pi_1^*(\mathrm{Ad}(g)\omega_G) + \pi_2^*\omega_G) \\ &= (\pi_1(\mathrm{id}\times\iota)\Delta)^*(\mathrm{Ad}(g)\omega_G) + (\pi_2(\mathrm{id}\times\iota)\Delta)^*\omega_G) \\ &= \mathrm{Ad}(g)\omega_G + \iota^*\omega_G. \end{aligned}$$

■

Corollary 4.11. *Let $f_1, f_2: M \to G$, and set $h(x) = f_1(x)f_2(x)^{-1}$. Then*

$$h^*\omega_G = \mathrm{Ad}(f_2(x))\{f_1^*\omega_G - f_2^*\omega_G\}.$$

Proof. Write $h(x)$ as the composite

$$\begin{array}{ccccccccc} M & \xrightarrow{\Delta} & M \times M & \xrightarrow{f_1\times f_2} & G \times G & \xrightarrow{\mathrm{id}\times\iota} & G \times G & \xrightarrow{\mu} & G. \\ x & \mapsto & (x,x) & \mapsto & (f_1(x),f_2(x)) & \mapsto & (f_1(x),f_2(x)^{-1}) & \mapsto & f_1(x)f_2(x)^{-1} \end{array}$$

Then

$$\begin{aligned} h^*\omega_G &= ((\mathrm{id}\times\iota)(f_1\times f_2)\Delta)^*\mu^*\omega_G \\ &= ((\mathrm{id}\times\iota)(f_1\times f_2)\Delta)^*(\pi_1^*(\mathrm{Ad}(f_2)\omega_G) + \pi_2^*\omega_G) \\ &= (\pi_1(\mathrm{id}\times\iota)(f_1\times f_2)\Delta)^*\mathrm{Ad}(f_2)\omega_G + (\pi_2(\mathrm{id}\times\iota)(f_1\times f_2)\Delta)^*\omega_G \\ &= f_1^*\mathrm{Ad}(f_2)\omega_G + f_2^*\iota^*\omega_G \\ &= f_1^*\mathrm{Ad}(f_2)\omega_G - f_2^*\mathrm{Ad}(f_2)\omega_G \\ &= \mathrm{Ad}(f_2)\{f_1^*(\omega_G) - f_2^*(\omega_G)\}. \end{aligned}$$

■

Exercise 4.12.* Let $f_1, f_2: M \to G$, and set $h(x) = f_1(x)f_2(x)$. Show that

$$h^*\omega_G = \mathrm{Ad}(f_2(x)^{-1})f_1^*\omega_G + f_2^*\omega_G.$$

❏

Exercise 4.13. Let H be a Lie group, V a vector space, and $\rho: H \to Gl(V)$ a representation. Let U be a manifold, X a vector field on U, and $h: U \to H$, $f: U \to V$ smooth functions. Show that

$$X(\rho(h^{-1})f) = \rho(h^{-1})X(f) - (\mathrm{Ad}(h)(h^*\omega_H(X)))f.$$

❏

Exercise 4.14.* Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let H be a Lie subgroup of G. Let θ be a $\mathfrak{g}$-valued 1-form on the manifold U and $k: U \to H$. Show that

$$d\{\mathrm{Ad}(k^{-1})\theta\} = \mathrm{Ad}(k^{-1})d\theta - [\mathrm{Ad}(k^{-1})\theta, k^*\omega_H]$$

(cf. Lemma **1.5.21** for the bracket of two $\mathfrak{g}$-valued 1-forms). ❏

The last two exercises are a good test of your ability to differentiate. Can you find a common generalization of them? We note that Exercise 4.14 is fundamental for the elementary properties of Cartan gauges in Chapter 5.

§5. The Darboux Derivative

Although we could have defined the Darboux derivative in §1, it is not until now that we can show anything of its usefulness.

Let

$$\begin{cases} G & \text{be a Lie group,} \\ \mathfrak{g} & \text{be the Lie algebra of } G, \\ \omega_G & \text{be the Maurer–Cartan form on } G, \\ M & \text{be a smooth manifold,} \\ f: M \to G & \text{be a smooth map.} \end{cases}$$

Definition 5.1. The (*left*) *Darboux derivative* of f is the $\mathfrak{g}$-valued 1-form $\omega_f = f^*(\omega_G)$ on M. ❀

The map f itself is called an *integral* or *primitive* of ω. It may also be called a *period map* associated to ω. It is obvious from the naturality of d that ω_f satisfies the analog of the structural equation, that is,

$$d\omega_f(X,Y) + [\omega_f(X), \omega_f(Y)] = 0.$$

Why is this 1-form ω_f called a derivative? In general, if $f: M \to N$, then we have called the induced map $f_*: T(M) \to T(N)$ "the derivative of f." However, this map is *not* exactly analogous to the usual derivative for functions on **R** since it has the original map f built into it. In the special case when $N = G$, a Lie group, we may follow f_* with the Maurer–Cartan form ω_G to obtain the map $\omega_G f_* = f^*\omega_G = \omega_f: T(M) \to T(N) \to \mathfrak{g}$. The precise effect of this composition is to "forget" the underlying map f and keep only the tangential information. Thus, ω_f is the precise analog of the usual derivative of functions on **R**.

The main property of the (left) Darboux derivative ω_f is that it determines f up to left translation by a constant element of G.

Theorem 5.2 (Uniqueness of the primitive). *Let M be a connected manifold and $f_1, f_2: M \to G$ be smooth maps such that $\omega_{f_1} = \omega_{f_2}$. Then there is an element $C \in G$ (the constant of integration) such that $f_2(x) = C \cdot f_1(x)$ for all $x \in M$.*

Proof. Consider the map $h\colon M \to G$ given by $h(x) = f_2(x)f_1(x)^{-1}$. According to Corollary 4.11, $h^*(\omega_G) = \mathrm{Ad}(f_1)(f_2^*\omega_G - f_1^*\omega_G)$. Since this latter expression vanishes by assumption and $h^*(\omega_G) = \omega_G h_*$, we see that $h_*\colon T(M) \to T(G)$ induces the zero map on each tangent space. It follows from Proposition **1**.4.18 that h is the constant function. Thus, we get $h(x) = C$ for all $x \in M$, where C is some fixed element of G. In particular, $f_2(x) = C \cdot f_1(x)$ for all $x \in M$. ∎

Corollary 5.3. (a) *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Then $\mathrm{Ad}\colon G \to Gl(\mathfrak{g})$ is the unique map $f\colon G \to Gl(\mathfrak{g})$ satisfying the two properties*

(i) $f(e) = e$

(ii) $f^*\omega_{Gl(\mathfrak{g})} = \mathrm{ad}(\omega_G)$

(b) *Let $g\colon I, 0 \to G$, e be a path on G. Then $\mathrm{Ad}(g)\colon I, e \to Gl(\mathfrak{g})$, e is the development, starting at the identity, of $\mathrm{ad}(g^*(\omega_G))$ on $Gl(\mathfrak{g})$.*

Proof. (a) First we note that Ad does indeed satisfy the two properties: the first is obvious and the second is Exercise 4.9. On the other hand, Theorem 5.2 guarantees the uniqueness of f.

(b) We have $(\mathrm{Ad}(g))^*\omega_{Gl(\mathfrak{g})} = g^*\mathrm{Ad}^*\omega_{Gl(\mathfrak{g})} = g^*(\mathrm{ad}(\omega_G)) = \mathrm{ad}(g^*\omega_G)$. ∎

Exercise 5.4. Show that if G is a connected Lie group and $\varphi\colon G \to G$ is a homomorphism of Lie groups that induces the identity map on the Lie algebras, then φ itself is the identity. ❑

Exercise 5.5. Show that the map $f\colon G \to Gl(\mathfrak{g})$ defined by $f(g) = \mathrm{Ad}(g^{-1})$ is the unique map satisfying (i) $f(e) = e$ and (ii) $f^*\omega_{Gl(\mathfrak{g})} = -\mathrm{ad}(\omega_G)$. ❑

§6. The Fundamental Theorem: Local Version

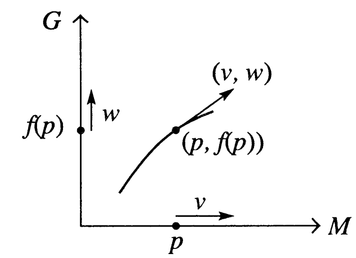
How can we characterize those $\mathfrak{g}$-valued 1-forms ω on M that are Darboux derivatives? In the last section we showed that a Darboux derivative satisfies the *structural equation*. Now we show that this necessary condition is also sufficient, at least locally. The beautiful proof of this is due to Elie Cartan.

Theorem 6.1. *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let ω be a $\mathfrak{g}$-valued 1-form on the smooth manifold M satisfying the structural equation $d\omega + \frac{1}{2}[\omega, \omega] = 0$. Then, for each point $p \in M$, there is a neighborhood U of p and a smooth map $f\colon U \to G$ such that $\omega \mid U = \omega_f$.*

Proof. The meaning of the Darboux derivative of a map f is that if we are sitting at a point $p \in M$ and we want to move in the direction $w \in T_p(M)$,

then, even though we don't know the image point $f(p)$, nevertheless the vector $\omega_p(w) \in \mathfrak{g}$, interpreted as a left-invariant vector field on G, tells us how the image point $f(x)$ should move no matter where it is situated.

The idea of Cartan's proof is to construct, locally, the *graph* of f in $M \times G$. If there is such an f, then a point $(p, f(p))$ on the graph can move only in certain directions if it is to remain on the graph. In fact, a tangent vector to the graph (v, w) must satisfy the condition $f_*(v) = w$.



This is equivalent to the condition $\omega_G(f_*(v)) = \omega_G(w)$, which in turn is just the condition $\omega_f(w) = \omega_G(v)$. That was for the case when f exists. If we are merely given the form ω, then the condition $\omega(w) = \omega_G(v)$ determines a distribution on $M \times G$. We shall show that this distribution is involutive and that any leaf is, locally, the graph of a possible f.

Now we pass to the formal proof. Let $\pi_G\colon G \times M \to G$ and $\pi_M\colon G \times M \to M$ denote the canonical projections. Let $\Omega = \pi_M^*(\omega) - \pi_G^*(\omega_G)$ and let $\mathcal{D} = \ker\ \Omega$ be the distribution defined by the kernel of Ω. Actually, we don't yet know that it is a distribution. By Proposition **2**.5.1, it suffices to show that it has constant rank.

Fix a point $(p, g) \in M \times G$. We are going to show

$$\pi_{M*(g,p)} \mid \mathcal{D}_{(g,p)}\colon \mathcal{D}_{(g,p)} \to T_p(M)$$

is an isomorphism. In particular, this will verify that $\ker\ \Omega$ has constant rank $=$ dim M. If $\pi_{M*}(v, w) = 0$ for some $(v, w) \in \mathcal{D}_{(g,p)} = \{(v, w) \in T_p(M) \times T_g(G) \mid \omega(v) = \omega_G(w)\}$, we have the implications $\pi_{M*}(v, w) = 0 \Rightarrow v = 0 \Rightarrow^{10} \omega_G(w) = 0 \Rightarrow w = 0 \Rightarrow (v, w) = 0$, and so we see that $\pi_{M*(g,p)} \mid \mathcal{D}$ is injective. Conversely, if $v \in T_p(M)$, then $(v, \omega_G^{-1}(\omega(v))_p) \in \mathcal{D}_{(g,p)}$, and so $\pi_{M*(g,p)} \mid \mathcal{D}$ is surjective.

Now we are going to show that $\mathcal{D}$ is integrable. Calculating the exterior derivative of Ω, we get

$$\begin{aligned} d\Omega &= d\pi_M^*(\omega) - d\pi_G^*(\omega_G) \\ &= \pi_M^*(d\omega) - \pi_G^*(d\omega_G) \end{aligned}$$

[10] This follows since $\omega(v) = \omega_G(w)$ for (v, w) in the distribution.

$$= \pi_M^*\left(-\frac{1}{2}[\omega,\omega]\right) - \pi_G^*\left(-\frac{1}{2}[\omega_G,\omega_G]\right)$$
$$= -\frac{1}{2}[\pi_M^*\omega, \pi_M^*\omega] + \frac{1}{2}[\pi_G^*\omega_G, \pi_G^*\omega_G].$$

Now make the replacement $\pi_M^*(\omega) = \pi_G^*(\omega_G) + \Omega$ so that[11]

$$d\Omega = -\frac{1}{2}[\pi_G^*\omega_G, \Omega] - \frac{1}{2}[\Omega, \pi_G^*\omega_G] - \frac{1}{2}[\Omega, \Omega].$$

Thus, $d\Omega(X,Y) = 0$ whenever $\Omega(X) = \Omega(Y) = 0$. Therefore, by Proposition **2**.5.3, the distribution $\ker \Omega$ is integrable.

Finally, we are going to construct, for each $p \in M$, a neighborhood U of p in M and a smooth map $f: U \to G$ such that $\omega \mid U = \omega_f$. Let $\mathcal{L}$ be a leaf through the point $(p,g) \in M \times G$. The derivative of the restriction of π_M to $\mathcal{L}$ induces the isomorphism $\pi_{M*}: \mathcal{D}_{(g,p)} \to T_p(M)$ studied above, and so $\pi_M \mid \mathcal{L}$ is a local diffeomorphism of a neighborhood of (p,g) to some neighborhood U of $p \in M$. Let $F: U \to \mathcal{L}$ be the inverse map. Since $\pi_M F = \mathrm{id}_U$, F must have the form $F(p) = (p, f(p))$, where $f: U \to G$. Now $F^*(\Omega) = \Omega F_* = 0$, since the image of F is tangent to the distribution on which Ω vanishes. Thus, we have

$$\begin{aligned} 0 &= F^*(\Omega) \\ &= F^*(\pi_M^*(\omega)) - F^*(\pi_G^*(\omega_G)) \\ &= (\pi_M F)^*\omega - (\pi_G F)^*\omega_G \\ &= \omega - f^*(\omega_G). \end{aligned}$$

That is, $\omega \mid U = \omega_f$. ■

Exercise 6.2. Let $\mathfrak{g}$ be a Lie algebra and $\mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Let ω be a $\mathfrak{g}$-valued 1-form on M such that $\omega^{-1}(\mathfrak{h})$ is a distribution $\mathcal{D}$ on M (i.e., $\dim \omega_x^{-1}(\mathfrak{h})$ is a constant, independent of x). Show that if $d\omega + \frac{1}{2}[\omega,\omega]$ takes values in $\mathfrak{h}$, then $\mathcal{D}$ is integrable. ❑

Exercise 6.3.* Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let $\mathfrak{h} \subset \mathfrak{g}$ be a Lie subalgebra of $\mathfrak{g}$. Show that there is a unique connected Lie subgroup $H \subset G$ with Lie algebra $\mathfrak{h}$. ❑

§7. The Fundamental Theorem: Global Version

Passing from the local existence to the global existence of the nonabelian integral requires the entry of a new idea. We assume that the reader is

[11] The reader should resist the urge to put the sum of the first two terms in the following expression equal to zero. In fact, the expression $[\omega_1, \omega_2]$ is *symmetric* in the two 1-forms ω_1, ω_2 (cf. Exercise **1**.5.20).

familiar with the notion of the fundamental group of a space, in the sense of algebraic topology. We shall also use the following facts.

(1) For M a smooth manifold, every continuous map $\lambda: (I,0,1) \to (M,p,q)$ is homotopic to a smooth map.

(2) If $\lambda_0, \lambda_1: (I,0,1) \to (M,p,q)$ are smooth maps that are continuously homotopic, then they are smoothly homotopic.

Thus, for a smooth manifold M, the fundamental group $\pi_1(M,b)$ can be regarded as the group of smooth homotopy classes of smooth loops on M.

We are going to show that a $\mathfrak{g}$-valued 1-form ω on M that satisfies the *structural equation* $d\omega + \frac{1}{2}[\omega,\omega] = 0$ determines a homomorphism

$$\Phi_\omega: \pi_1(M,b) \to G$$

called the *monodromy representation.* It will turn out that this representation is *trivial* if and only if ω is the Darboux derivative of some map $M \to G$. Thus, the monodromy will be the last obstruction to the global existence of a primitive for ω.

Development

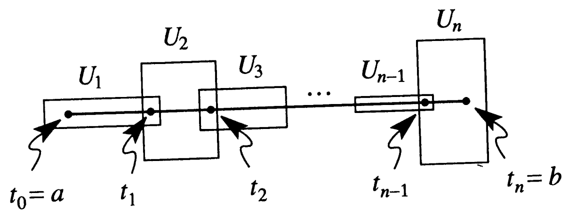
Let ω be a $\mathfrak{g}$-valued 1-form on M. The dim $M = 1$ case is quite exceptional in that, for it, the structural equation is always satisfied, there being no nonzero 2-forms on a one-dimensional manifold. Thus, the local existence theorem always applies to this case. Let us assume for definiteness that $M = I = [a,b]$. With this special assumption, we have the following *global existence theorem.*

Theorem 7.1. *Let ω be a smooth $\mathfrak{g}$-valued 1-form on $I = [a,b]$. Then there is a unique map*

$$f: (I,a) \to (G,g)$$

with Darboux derivative ω.

Proof. By Proposition 5.2, if f exists, it is unique. By the local fundamental theorem, Theorem 6.1, each point lies in an open interval on which ω is the Darboux derivative of a map into G. Since I is compact and these intervals form an open cover, it follows that I is covered by finitely many open intervals U_i, $1 \le i \le n$, on each of which ω is the Darboux derivative of a map into G. In fact, we may assume that $U_i \cap U_j = \emptyset$ unless $|i-j| \le 1$ and that $t_0 = a \in U_1$. Choose $t_i \in U_i \cap U_{i+1}$ for $1 \le i \le n$.



We inductively choose primitives $f_i\colon U_i \to G$ such that $f_1(a) = g$, $f_{i+1}(t_i) = f_i(t_i)$ for $1 \le i \le n-1$. By the uniqueness of the primitives, we have $f_{i+1} = f_i$ on $U_i \cap U_{i+1}$. Thus we may patch these maps together, setting $f(t) = f_i(t)$ for $t \in U_i$ to obtain a map $f\colon I \to G$. Clearly, $f(a) = g$ and $f^*(\omega_G) = \omega$. ■

Definition 7.2. The unique smooth map $f\colon I \to G$ satisfying $f(a) = g$ and $f^*(\omega_G) = \omega$ is called the *development of* ω *on* G *along* I *starting at* g. (It is, of course, merely a suggestive name for a primitive of ω.) ❀

Remark 7.3. It suffices that ω be merely piecewise smooth on I for there to be a unique development. It will be continuous everywhere and smooth where ω is smooth. Indeed, if $a = t_0 < t_1 < \cdots < t_n = b$ is a partition of I and $\omega \mid [t_{i-1}, t_i]$ is smooth for all $i = 1, \dots, n$, then we can choose developments $f_i\colon [t_{i-1}, t_i] \to G$ so that $f_1(a) = g$, $f_{i+1}(t_i) = f_i(t_i)$ for $1 \le i \le n-1$. Patching together, we set $f(t) = f_i(t)$ for $t \in [t_{i-1}, t_i]$ to obtain a continuous map $f\colon I \to G$ that satisfies $f^*(\omega_G) = \omega$ except when $t = t_i$, $1 \le i \le n-1$.

Next we pass to a more general situation.

Definition 7.4. Let $I = [a, b]$, let M be a manifold of arbitrary dimension, and let ω be a smooth $\mathfrak{g}$-valued 1-form on M. Given a piecewise smooth path $\sigma\colon (I, a, b) \to (M, p, q)$, let $\tilde{\sigma}\colon (I, a) \to (G, g)$ be the development of the $\mathfrak{g}$-valued 1-form $\sigma^*(\omega)$, which means $\tilde{\sigma}^*(\omega_G) = \sigma^*(\omega)$. We call $\tilde{\sigma}$ the *development of* ω *along* σ starting at g. ❀

Exercise 7.5.* Let I, M, and ω be as in Definition 7.4. Show the following:

(a) Let $\sigma\colon (I, a, b) \to (M, p, q)$ have development $\tilde{\sigma}\colon (I, a, b) \to (G, g, h)$. Then

  (i) $k\tilde{\sigma}\colon (I, a, b) \to (G, kg, kh)$ is also a development of σ.

  (ii) $\tilde{\sigma}^{-1}\colon (I, a, b) \to (G, h, g)$ is a development of $\sigma^{-1}\colon (I, a, b) \to (M, p, q)$. (Here inverse refers to the reverse path rather than inverse in the group G.)

(b) Set $J = [b, c]$, and let $\sigma\colon (I, a, b) \to (M, p, q)$ and $\rho\colon (I, b, c) \to (M, q, r)$ have developments $\tilde{\sigma}\colon (I, a, b) \to (G, g, h)$ and $\tilde{\rho}\colon (I, a, b) \to (G, h, k)$.

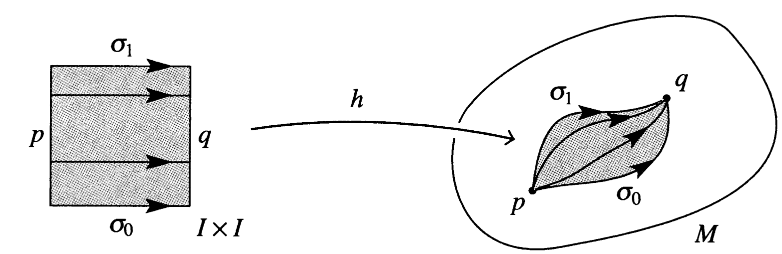
Then $\sigma \star \rho\colon (I, a, c) \to (M, p, r)$ has development $\tilde{\sigma} \star \tilde{\rho}\colon (I, a, c) \to (G, h, k)$. ❑

Exercise 7.6. Show that the endpoint $\tilde{\sigma}(1)$ of the development of a smooth path is invariant under smooth reparameterization of the path. ❑

In general, however, the endpoint $\tilde{\sigma}(1)$ of the development depends not just on q, but on the whole path σ joining p to q. Nevertheless, we now show that when the 1-form ω on M satisfies the structural equation $d\omega + \frac{1}{2}[\omega, \omega] = 0$, then $\tilde{\sigma}(1)$ depends only on the homotopy class of σ (and the starting point g).

Theorem 7.7. *Let M be a smooth manifold and let $\sigma_0, \sigma_1\colon (I, a, b) \to (M, p, q)$ be smooth homotopic paths. If ω is a smooth $\mathfrak{g}$-valued 1-form on M that satisfies the structural equation, then its developments, starting at g, along σ_0 and σ_1 have the same endpoints.*

Proof. Let $h\colon (I \times I, a \times I, b \times I) \to (M, p, q)$ be the homotopy joining σ_0 and σ_1.



Now $h^*\omega$ is a $\mathfrak{g}$-valued 1-form on $I \times I$ that satisfies the structural equation. By Theorem 6.1, each point of $I \times I$ lies in a square neighborhood on which $h^*\omega$ is the Darboux derivative of a map into G. Given a point $t \in I$, the compact set $I \times t$ is covered by finitely many of these open sets, and, by an argument like the one above for the interval, we obtain an open set of the form $I \times U$, with U an open interval, on which $h^*\omega$ is the Darboux derivative of some map into G. A similar analysis allows us to fit these maps together to get, finally, a map $H\colon I \times I \to G$ such that $H(a, a) = g$ and $H^*(\omega_G) = h^*\omega$. Since $h(b, u) = q$ is constant for $u \in I$, it follows that $h^*\omega = H^*(\omega_G)$ vanishes along the right edge of $I \times I$, and hence H is also constant along this right edge. In particular, $H(b, a) = H(b, b)$, and so the developments of ω along σ_0 and σ_1 both end at the same point in G. ■

Exercise 7.8. Show that Theorem 7.7 remains true for homotopic piecewise smooth paths. [*Hint:* Show that a piecewise smooth path has a reparameterization that is a smooth path.] ❑

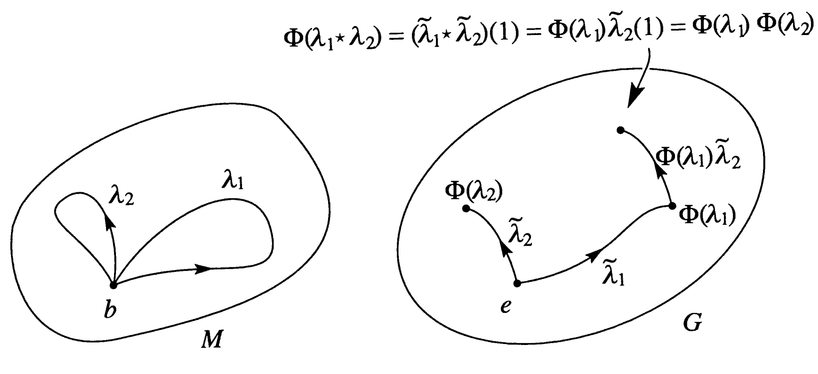
Monodromy

As before, M is a smooth manifold and ω is a $\mathfrak{g}$-valued 1-form on M that satisfies the structural equation. Let $I = [0,1]$, fix a base point $b \in M$, and let $\lambda\colon (I, \partial I) \to (M, b)$ be a smooth closed loop.[12] According to the last result, we may develop ω along λ starting at e, and the endpoint $\Phi(\lambda) = \tilde{\lambda}(1)$ of this development will be a point of G that, by Theorem 7.7, is independent of the specific loop λ chosen within a homotopy class. In this way the development yields a well-defined map, $\Phi_\omega\colon \pi_1(M,b) \to G$.

Definition 7.9. The map $\Phi_\omega\colon \pi_1(M,b) \to G$ described above is called the *monodromy representation* of ω. The image of Φ_ω denoted by $\Gamma \subset G$ is called the *group of periods* (*period group*), or the *monodromy group*. ❀

Proposition 7.10. *The monodromy representation is a group homomorphism.*

Proof. The reader can easily construct a proof by referring to the following picture.



In this diagram the left translation $\Phi(\lambda_1)\tilde{\lambda}_2$ is, of course, the development of λ_2 starting at $\Phi(\lambda_1)$. ■

Period Group Only Defined Up to Conjugacy

The fact that the fundamental group depends on the choice of base point is reflected in the fact that the period group is only defined up to conjugacy.

First we recall the effect of change of base point from b_0 to b_1 on the fundamental group of a path-connected space. If $\sigma\colon (I,0,1) \to (M, b_0, b_1)$ is a path joining the two points, then there is an isomorphism

[12] $\partial I = \{0,1\}$.

$$c_\sigma\colon \pi_1(M,b_1) \to \pi_1(M,b_0)$$

induced by the map on loops sending

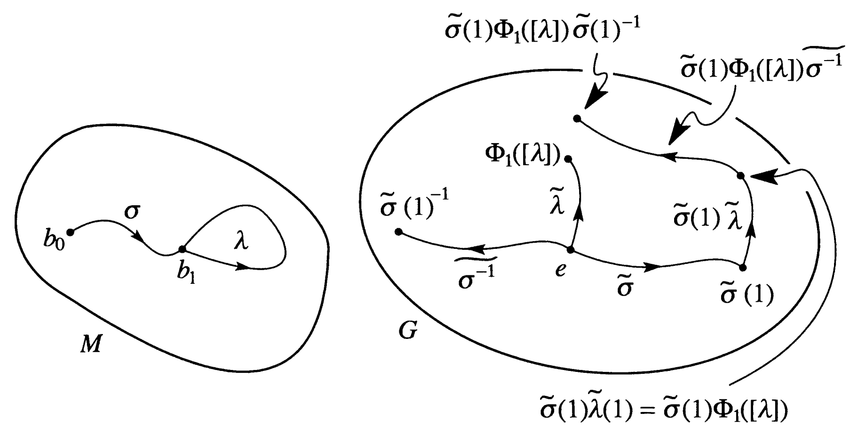
$$\lambda \mapsto \sigma \star \lambda \star \sigma^{-1}.$$

Let $\tilde{\sigma}\colon (I,0) \to (G,e)$ be the development of a $\mathfrak{g}$-valued 1-form ω on M along σ starting at e. Then we have the conjugation map $c_{\tilde{\sigma}(1)}\colon G \to G$ sending $g \mapsto \tilde{\sigma}(1) g \tilde{\sigma}(1)^{-1}$.

Theorem 7.11. *Let $\Phi_j\colon \pi_1(M,b_j) \to G$, $j = 0,1$, denote the monodromy representations with respect to two base points. Let $\sigma\colon (I,0,1) \to (M,b_0,b_1)$ be a path joining b_0 to b_1. Then the following diagram commutes.*

$$\begin{tikzcd}
\pi_1(M,b_0) \arrow[r, "\Phi_0"'] & G \\
\pi_1(M,b_1) \arrow[u, "c_\sigma"] \arrow[r, "\Phi_1"] & G \arrow[u, "c_{\tilde{\sigma}(1)}"']
\end{tikzcd}$$

Proof. Let $\lambda\colon (I,\partial I) \to (M,b_1)$ be a loop on M representing the element $[\lambda] \in \pi_1(M,b_1)$ and let $\tilde{\sigma}\colon (I,0) \to (G,e)$ and $\tilde{\lambda}\colon (I,0) \to (G,e)$ be the developments of σ and λ, respectively, starting at e.



The diagram shows that the development of $\sigma \star \lambda \star \sigma^{-1}$ is $\tilde{\sigma} \star (\tilde{\sigma}(1)\tilde{\lambda}) \star (\tilde{\sigma}(1)\Phi_1([\lambda])(\widetilde{\sigma^{-1}}))$, which ends at

$$\tilde{\sigma}(1)\Phi_1([\lambda])(\widetilde{\sigma^{-1}})(1) = \tilde{\sigma}(1)\Phi_1([\lambda])\tilde{\sigma}(1)^{-1} = c_{\tilde{\sigma}(1)}(\Phi_1([\lambda]).$$

The equation $(\widetilde{\sigma^{-1}})(1) = \tilde{\sigma}(1)^{-1}$ used here comes from Exercise 7.5(a). ■

Proposition 7.12. *Let M be a smooth connected manifold, let $b \in M$. and let $f\colon (M,b) \to (G,g)$ be a smooth map. Set $\omega = f^*\omega_G$. Then for any curve $\sigma\colon (I,0,1) \to (M,b,x)$, the development, starting at g, ends at $f(x)$.*

Proof. Since $(f\sigma)^*\omega_G = \sigma^* f^* \omega_G = \sigma^*\omega$, the development of ω along $\sigma\colon (I,0,1) \to (M,b,x)$ is just $f\sigma$, and $f\sigma(1) = f(x)$. ■

Corollary 7.13. *The monodromy representation of a Darboux derivative is trivial.*

Proof. Write $\omega = f^*\omega_G$, where $f\colon (M,b) \to (G,g)$. Replacing f by $L_{g^{-1}}f$ if necessary (which doesn't alter the Darboux derivative), we may assume that $f\colon (M,b) \to (G,e)$. For any loop $\lambda\colon (I,0,1) \to (M,b,b)$, the proposition says that the development, starting at e, of ω along λ ends at $f\lambda(1) = f(b) = e$. Thus $\Phi_\omega([\lambda]) = e$. ■

The Fundamental Theorem

Now we can assemble our results into the following statement, the fundamental theorem of nonabelian calculus.

Theorem 7.14 (Fundamental theorem of calculus). *Let G be a Lie group with Lie algebra $\mathfrak{g}$. Let M be a smooth connected manifold and let ω be a $\mathfrak{g}$-valued 1-form on M. Then*

$$\left\{ \begin{array}{c} \omega \text{ is the Darboux derivative} \\ \text{of some map } M \to G \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \text{(i)} \qquad d\omega + \tfrac{1}{2}[\omega,\omega] = 0; \\ \text{(ii) the monodromy representation} \\ \qquad \Phi_\omega\colon \pi_1(M,b) \to G \text{ is trivial.} \end{array} \right\}$$

Moreover, if these conditions are satisfied, the integral of ω is unique up to left translation by a constant element of G.

Proof. The last statement is just the uniqueness of the primitive, Theorem 5.2. For the implication $\Rightarrow$, (i) is shown in the discussion following Definition 5.1 and (ii) is Corollary 7.13.

Let us verify the implication $\Leftarrow$. Define $f\colon (M,b) \to (G,e)$ by declaring $f(x)$ to be the endpoint of the development of ω along any path from b to x. This definition makes f well defined under homotopies of the path by Theorem 7.7 and under general changes in the path by the very definition of monodromy, Definition 7.9. We need to check that $\omega = \omega_f$.

Note that the value of $f(x)$ may be obtained in two steps as follows. If $x_0 \in M$, we may choose a path from b to x_0 and then another from x_0 to x. Then if the development of the first path starts at e and ends at g_0, while the development of the second starts at e and ends at g, then the development of the composite path will start at e and end at $g_0 g$.

Now the local existence theorem guarantees that, for each point $x_0 \in M$, there is a connected open set U about x_0 and also a smooth map $f_U\colon U \to G$ satisfying $f_U^*\omega_G = \omega$. After a left translation of f_U by some element of G, we may assume that $f_U(x_0) = f(x_0)$. Now the development of ω, starting at $f(x_0)$, along any curve $\sigma\colon (I,0,1) \to (U,x_0,x)$ ends at $f_U(x)$. By the remarks in the last paragraph, it also ends at $f(x)$, so $f_U(x) = f(x)$ for all $x \in U$. Thus, $f^*\omega_G = f_U^*\omega_G = \omega$ on U. ■

Some Remarks About the Fundamental Theorem

The fundamental theorem certainly generalizes the usual fundamental theorem of elementary calculus concerning maps $f\colon (a,b) \to \mathbf{R}$. In that case, conditions (i) and (ii) are automatic, so the theorem reads that every smooth $\mathbf{R}$-valued 1-form on $[a,b]$ is the derivative of some smooth map f that is unique up to an additive constant.

It also generalizes the theorem that says that a vector field $v\colon U \to \mathbf{R}^n$ on a simply connected, open set U of $\mathbf{R}^n$ that satisfies the equation

$$\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} = 0$$

for all $i \le i, j \le n$, has a potential function. To see this, write $v = (v_1, \ldots, v_n)$, and set $\omega = \sum v_i dx_i$. Then ω is an $\mathbf{R}$-valued 1-form on U, and we have

$$d\omega = \sum_{1 \le i,j \le n} \frac{\partial v_j}{\partial x_i} dx_j \wedge dx_i = \sum_{1 \le i < j \le n} \left(\frac{\partial v_j}{\partial x_i} - \frac{\partial v_i}{\partial x_j} \right) dx_j \wedge dx_i = 0.$$

Moreover, since $\mathbf{R}$ is abelian, $[\omega,\omega] = 0$ and hence the structural equation holds. Thus, by the fundamental theorem, a function $V\colon U \to \mathbf{R}$ exists such that $V^*(dx) = \omega$. Now $(V^*(dx))e_i = dx(V_*(e_i)) = dx(\partial V/\partial x_i) = \partial V/\partial x_i$. Thus $\omega = V^*dx = \sum(\partial V/\partial x_i)dx_i$, that is, $v_i = \partial V/\partial x_i$, for all $1 \le i \le n$.

It is worthwhile to see (at least once) the coordinate expression of the structural equation for a nonabelian group. We shall do this for the group $Gl_n(\mathbf{R})$. Of course, the same expression then works for any subgroup. Let us assume that ω is an $M_n(\mathbf{R})$-valued 1-form on some open set U of $\mathbf{R}^m$. Thus, for $v \in T(U)$, we have $\omega(v) = (A_{ij}(v))$, where the entries A_{ij} are $\mathbf{R}$-valued 1-forms on U. We use the basis $\partial_p = \partial/\partial x_p$, $1 \le p \le m$, and set $A_p = (A_{ij}(\partial_p))$ so that $A_p\colon U \to M_n(\mathbf{R})$. Then

$$\begin{aligned} d\omega(\partial_p, \partial_q) &= \partial_p\omega(\partial_q) - \partial_q\omega(\partial_p) - \omega[\partial_p, \partial_q] \\ &= \partial_p\omega(\partial_q) - \partial_q\omega(\partial_p) \\ &= \partial_p A_q - \partial_q A_p \end{aligned}$$

and

$$\frac{1}{2}[\omega,\omega](\partial_p,\partial_q) = [\omega(\partial_p),\omega(\partial_q)] = [A_p, A_q].$$

Thus, the structural equation becomes

$$\frac{\partial A_q}{\partial x_p} - \frac{\partial A_p}{\partial x_q} + [A_p, A_q] = 0 \quad \text{for } 1 \le p, q \le m.$$

In terms of the entries A_{ijp} (where $A_p = \sum_{1\le i,j\le n} A_{ijp}e_{ij}$), the structural equation is

$$\frac{\partial A_{ijq}}{\partial x_p} - \frac{\partial A_{ijp}}{\partial x_q} + \sum_{1\le k\le n} (A_{ikp}A_{kjq} - A_{ikq}A_{kjp}) = 0$$

for $1 \le p, q \le m$ and $1 \le i, j \le n$.

The fundamental theorem tells us that whenever we see equations like these we may interpret the A_ps or A_{ijp}s as the coordinate expression of an $M_n(\mathbf{R})$-valued 1-form ω that on U is locally (and globally if U is simply connected) the Darboux derivative of a $Gl_n(\mathbf{R})$-valued function B on U. That is, $B^{-1}dB = \omega$. Moreover, if U is connected, then such a B is unique up to left translation.

This is an existence and uniqueness theorem for solutions of a certain partial differential equation. In coordinates it says that for U connected and simply connected, and provided the structural equations hold, the system

$$\frac{\partial B}{\partial x_p} = BA_p \quad \text{for } 1 \le p \le m$$

or, the same as above but in more detail, the system

$$\frac{\partial B_{ij}}{\partial x_p} = \sum_{1\le k\le n} B_{ik}A_{kjp} \quad \text{for } 1 \le p \le m \text{ and } 1 \le i, j \le n$$

has a solution,[13] which is unique up to left multiplication of the matrix B by a constant matrix.

The Lie Group–Lie Algebra Correspondence

We end this section with some applications of the fundamental theorem of calculus to the Lie group–Lie algebra correspondence outlined at the end of §2.

Proposition 7.15. *Let G and H be Lie groups with Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$. If G is connected and simply connected, then every homomorphism of Lie algebras $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ is induced by a unique homomorphism $\Phi\colon G \to H$.*

[13] The solution is a "nonabelian potential function."

Proof. The composition $\omega = \varphi\omega_G$ is an $\mathfrak{h}$-valued, left-invariant 1-form on G that satisfies the structural equation (since $d\omega + \frac{1}{2}[\omega,\omega] = d\varphi\omega_G + \frac{1}{2}[\varphi\omega_G, \varphi\omega_G] = \varphi d\omega_G + \frac{1}{2}\varphi[\omega_G,\omega_G] = 0$). Thus there is a unique smooth map $\Phi\colon (G,e) \to (H,e)$ satisfying $\Phi^*(\omega_H) = \omega$. We claim that Φ is a homomorphism. To see this, consider $f\colon (G,e) \to (H,e)$ defined by $f(g) = \Phi(g_0)^{-1}\Phi(g_0g)$ for some fixed g_0. Now $f = L_{\Phi(g_0)^{-1}} \circ \Phi \circ L_{g_0}$, and so

$$f^*\omega_H = L^*_{g_0}\Phi^* L^*_{\Phi(g_0)^{-1}}\omega_H = L^*_{g_0}\Phi^*\omega_H = L^*_{g_0}\omega = \omega.$$

Since f and Φ agree at the identity and have the same Darboux derivative, it follows that $f = \Phi$ and hence $\Phi(g) = \Phi(g_0)^{-1}\Phi(g_0g)$, that is, $\Phi(g_0)\Phi(g) = \Phi(g_0g)$ for all g and for all g_0. ■

In particular, connected and simply connected Lie groups with isomorphic Lie algebras are themselves isomorphic. More precisely, we have the following result.

Corollary 7.16. *Let G and H be connected and simply connected Lie groups with Lie algebra, $\mathfrak{g}$ and $\mathfrak{h}$ and let $\varphi\colon \mathfrak{g} \to \mathfrak{h}$ be an isomorphism of Lie algebras. Then φ is induced by a unique isomorphism of groups $\Phi\colon G \to H$.*

Proof. The existence and uniqueness of a homomorphism $\Phi\colon G \to H$ is guaranteed by Proposition 7.15. Since φ is an isomorphism, it follows that the kernel of Φ is a discrete central subgroup and so Φ is a covering map. Since H is simply connected, the covering must be trivial and hence Φ is an isomorphism. ■

We may also apply these ideas to obtain the correspondence between the representations of a Lie group and the representations of its Lie algebra.

Definition 7.17. A *representation* of a Lie algebra $\mathfrak{g}$ on a finite vector space V is a Lie algebra homomorphism $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$. ❀

Corollary 7.18. *Let G be a connected and simply connected Lie group with Lie algebra $\mathfrak{g}$. Let V be a finite-dimensional vector space and let $\rho\colon \mathfrak{g} \to \mathfrak{gl}(V)$ be a representation. Then there is a unique representation $R\colon G \to Gl(V)$ satisfying $R_{*e} = \rho$.*

Proof. Apply Corollary 7.16. ■

Finally, we want to show that every (real) Lie algebra is the Lie algebra of a Lie group. For this we use (the real version of) Ado's theorem, which we state without proof.

Theorem 7.19 (Ado's theorem). *Every real Lie algebra $\mathfrak{g}$ is isomorphic to a subalgebra of the Lie algebra $\mathfrak{gl}(V)$ for some finite-dimensional real, vector space V.*

Proof. Cf. [M. Postnikov, 1986] or [W. Fulton and J. Harris, 1991]. ∎

Theorem 7.20. *Let $\mathfrak{g}$ be an arbitrary, finite-dimensional, real Lie algebra. Then there is, up to isomorphism, a unique connected and simply connected Lie group G whose Lie algebra is isomorphic to $\mathfrak{g}$.*

Proof. By Theorem 7.19, we may regard the Lie algebra $\mathfrak{g}$ as a subalgebra $\mathfrak{g} \subset \mathfrak{gl}(V)$, where $\mathfrak{gl}(V)$ is the Lie algebra of the Lie group $Gl(V)$. By Exercise 6.3, there is a unique connected Lie group $G_0 \subset Gl(V)$ with Lie algebra $\mathfrak{g}$. Let G be the universal cover of G_0. Then by Corollary 8.11,[14] G is also a Lie group with Lie algebra $\mathfrak{g}$ and it is connected and simply connected. By Corollary 7.16, if G' is another connected and simply connected Lie group with Lie algebra $\mathfrak{g}$, then G and G' are isomorphic. ∎

§8. Monodromy and Completeness

In this, the final section of this chapter, we study the monodromy group from the viewpoint of covering spaces. In certain cases we relate the completeness of a Lie algebra-valued 1-form to the discreteness of the period group $\Gamma = \Phi_\omega(\pi_1(M,b)) \subset G$.

Monodromy from the Viewpoint of Covering Transformations

Let $\pi\colon \tilde{M} \to M$ be the universal cover of M, and let $\tilde{b} \in \tilde{M}$ be a choice of base point lying over b. Because π is a local diffeomorphism, the form $\pi^*(\omega)$ satisfies the structural equation when ω does. In this case, since $\tilde{M}$ is simply connected, its monodromy is trivial. Thus, by the fundamental theorem, there is a unique primitive $f\colon (\tilde{M},\tilde{b}) \to (G,e)$ for $\pi^*(\omega)$.

Proposition 8.1. *There is a unique homomorphism $\tilde{\Phi}\colon \mathcal{G}al(\tilde{M}/M) \to G$ satisfying $fT = \tilde{\Phi}(T)f$ for every $T \in \mathcal{G}al(\tilde{M}/M)$.*[15]

Proof. Every covering transformation $T \in \mathcal{G}al(\tilde{M}/M)$ preserves the form $\pi^*(\omega)$ (since $T^*(\pi^*(\omega)) = (\pi T)^*\omega = \pi^*(\omega)$), and so we have

$$(fT)^*\omega_G = T^*(f^*(\omega_G)) = T^*\pi^*(\omega) = \pi^*(\omega) = f^*\omega_G.$$

Thus, by the fundamental theorem, fT and f differ by left multiplication by a constant element $\tilde{\Phi}(T) \in G$, that is, $fT = \tilde{\Phi}(T)f$. In particular,

[14]We do not appeal to the statement of Theorem 7.20 in any other proof of this chapter.

[15]By $\mathcal{G}al(\tilde{M}/M) = \{T \in \mathrm{Diff}(M) \mid \pi\circ T = \pi\}$, we denote the group of covering transformations.

$fT(\tilde{b}) = \tilde{\Phi}(T)f(\tilde{b}) = \tilde{\Phi}(T)$. To see that $\tilde{\Phi}$ is a group homomorphism, note that

$$\tilde{\Phi}(ST) = f(ST(\tilde{b})) = \tilde{\Phi}(S)f(T(\tilde{b})) = \tilde{\Phi}(S)\tilde{\Phi}(T)f(\tilde{b}) = \tilde{\Phi}(S)\tilde{\Phi}(T). \quad ∎$$

Now we show that the isomorphism $\sigma\colon \mathcal{G}al(\tilde{M}/M) \to \pi_1(M,b)$ sending $T \mapsto [\pi(\lambda)]$ (where λ is any path on $\tilde{M}$ joining $\tilde{b}$ to $T(\tilde{b})$) relates the two versions of monodromy. We have the following result.

Proposition 8.2. *The following diagram commutes.*

$$\begin{tikzcd} \mathcal{G}al(\tilde{M},\tilde{b}) \arrow[d, "\sigma"'] \arrow[dr, "\tilde{\Phi}"] & \\ \pi_1(M,b) \arrow[r, "\Phi"] & G \end{tikzcd}$$

Proof. The equation $fT(\tilde{b}) = \tilde{\Phi}(T)$ identifies $\tilde{\Phi}(T)$ as the endpoint in G of the development of $\pi^*\omega$ along a path in $\tilde{M}$ joining $\tilde{b}$ to $T(\tilde{b})$. Clearly, development of ω along the image loop in M starting at b yields the same endpoint (in G). This fact is the commutativity of the diagram. ∎

Completeness

Next we introduce the important notion of a *complete* 1-form.

Definition 8.3. Let ω be a V-valued 1-form on the manifold M. A vector field X on M is said to be *ω-constant* if $\omega(X)$ is constant on M. The form ω is said to be *complete* if every ω-constant vector field X on M is complete. ❀

Example 8.4. Every V-valued 1-form on a compact manifold M is complete. This is a consequence of Proposition 2.1.6, which says that every vector field on a compact manifold is complete. ◆

Example 8.5. Let ω be a complete V-valued 1-form on the manifold M. Let N be a closed submanifold of M. Then $\omega \mid N$ is complete. Note that this example may be of small content since there may not be any ω-constant vector fields on N. ◆

Example 8.6. The Maurer–Cartan form on a Lie group G is complete. This follows since the ω constant vector fields are just the left-invariant fields, which we have shown to be complete (cf. Corollary 2.12). ◆

Characterization of Lie Groups

The final result, (Theorem 8.7), shows that in a certain sense the Maurer–Cartan form determines the group up to some covering. (It is this description of a Lie group that will generalize to the notion of a Cartan connection on a principal bundle. However, in this we are getting somewhat ahead of ourselves.)

Theorem 8.7. *Let M be a connected smooth manifold and let $\mathfrak{g}$ be a Lie algebra. Let ω be a $\mathfrak{g}$-valued 1-form on M that satisfies the conditions*

(i) $d\omega + \frac{1}{2}[\omega,\omega] = 0$,

(ii) $\omega: T(M) \to \mathfrak{g}$ *is an isomorphism on each fiber,*

(iii) ω *is complete.*

Then

(a) *The universal cover of M, $\pi: G \to M$, has, for an arbitrary choice $e \in G$, the structure of a Lie group with identity element e and Lie algebra $\mathfrak{g}$ whose Maurer–Cartan form is $\pi^*\omega$.*

(b) *The period group $\Gamma = \Phi_\omega(\pi_1(M,b) \subset G$ acts by left multiplication on G as the group of covering transformations for the cover $\pi: G \to M$.*

Proof. *The case $\pi_1(M) = 1$.*

Here we have $G = M$. This case involves five steps. The first shows how to construct a map $f: M \to Gl(\mathfrak{g})$ that will eventually be given by $f(g) = \mathrm{Ad}(g)$. The second step constructs, for an arbitrary choice of $e \in G = M$, a "multiplication" map $\mu: M \times M \to M$ satisfying the conditions $\mu(e,e) = e$ and $\mu^*\omega = (\pi_2^* f)^{-1}\pi_1^*\omega + \pi_2^*\omega$. The third step shows that f is a homomorphism with respect to the multiplication μ. The fourth shows that μ is associative. The fifth step constructs the inversion map $\iota: M \to M$, and shows that ι is inversion with with respect to μ.

Step 1. Corollary 5.3 shows that $f: M \to \mathrm{Aut}_{\mathrm{Lie}}(\mathfrak{g})$ should have Darboux derivative $\eta_f = \mathrm{ad}(\omega)$ ($\in \mathfrak{gl}_{\mathrm{Lie}}(\mathfrak{g})$) and satisfy $f(e) = e$. Thus, the form ω will determine f, provided $\eta = \mathrm{ad}(\omega)$ is integrable. But

$$d\eta + \frac{1}{2}[\eta,\eta] = \mathrm{ad}(d\omega) + \frac{1}{2}[\mathrm{ad}(\omega),\mathrm{ad}(\omega)] = \mathrm{ad}\left(d\omega + \frac{1}{2}[\omega,\omega]\right) = 0,$$

and since M is simply connected, it follows from the fundamental theorem, Theorem 7.14, that η is the Darboux derivative of a unique map $f: M \to Gl(\mathfrak{g})$ satisfying $f(e) = e$. In particular, $df = f \circ \mathrm{ad}(\omega)$ (where the product on the right is composition of elements of $\mathfrak{gl}(\mathfrak{g})$).

Step 2. We construct the multiplication map. Proposition 4.10 shows that $\mu: M \times M \to M$ should have Darboux derivative at $(x,y) \in M \times M$ given by

$$\eta = \pi_1^*(Ad(y^{-1})\omega) + \pi_2^*\omega = \pi_1^*(f(y)^{-1}\omega) + \pi_2^*\omega = (\pi_2^* f^{-1})\pi_1^*\omega + \pi_2^*\omega,$$

where $\pi_i: M \times M \to M$ denotes projection on the ith factor. Thus, the form ω will determine μ, provided η is integrable. Now

$$\begin{aligned}
d\eta + \frac{1}{2}[\eta,\eta] &= \pi_2^*(d(f^{-1}))(\pi_1^*(\omega)) + \pi_2^*(f^{-1})(\pi_1^*(d\omega)) + \pi_2^*(d\omega)\\
&\quad + \frac{1}{2}[\pi_2^*(f^{-1})(\pi_1^*(\omega)), \pi_2^*(f^{-1})(\pi_1^*(\omega))]\\
&\quad + [\pi_2^*(f^{-1})(\pi_1^*(\omega)), \pi_2^*\omega] + \frac{1}{2}[\pi_2^*\omega, \pi_2^*\omega]\\
&= \pi_2^*(d(f^{-1}))(\pi_1^*(\omega)) + \pi_2^*(f^{-1})(\pi_1^*(\underbrace{d\omega + \frac{1}{2}[\omega,\omega]}_{0}))\\
&\quad + \pi_2^*(\underbrace{d\omega + \frac{1}{2}[\omega,\omega]}_{0}) + [\pi_2^*(f^{-1})(\pi_1^*(\omega)), \pi_2^*\omega].
\end{aligned}$$

Since $d(f^{-1}) = -\mathrm{ad}\omega \circ f^{-1}$ and[16]

$$\begin{aligned}
\pi_2^*(-\mathrm{ad}\omega f^{-1})(\pi_1^*(\omega)) &= -\pi_2^*(\mathrm{ad}\omega)\pi_2^*(f^{-1})(\pi_1^*(\omega))\\
&= -[\pi_2^*\omega, \pi_2^*(f^{-1})(\pi_1^*(\omega))]\\
&= -[\pi_2^*(f^{-1})(\pi_1^*(\omega)), \pi_2^*\omega],
\end{aligned}$$

it follows that $d\eta + \frac{1}{2}[\eta,\eta] = 0$. Since M is simply connected, the fundamental theorem shows that η is the Darboux derivative of a unique map $\mu: M \times M \to M$ satisfying $\mu(e,e) = e$.

Step 3. We show that $f: M \to M$ is a homomorphism with respect to μ in the sense that $f(\mu(x,y)) = f(x)f(y)$ for all $x, y \in M$. We shall verify the commutativity of the following diagram.

$$\begin{tikzcd}
M \times M \arrow[r, "\mu"] \arrow[d, "f \times f"'] & M \arrow[d, "f"] \\
Gl(\mathfrak{g}) \times Gl(\mathfrak{g}) \arrow[r, "\mu_{Gl}"] & Gl(\mathfrak{g})
\end{tikzcd}$$

It is clear that this diagram commutes at (e,e). Since M is connected, by Proposition 1.4.19 it suffices to show that the Maurer–Cartan form ω_{Gl} on

[16]Note that by Exercise 1.5.20, the bracket is symmetric on 1-forms.

$Gl(\mathfrak{g})$, with values in $\mathfrak{gl}(\mathfrak{g})$, pulls back to the same form on $M \times M$ along the two possible routes. That is, we must show that

$$(f \circ \mu)^*\omega_{Gl} = (\mu_{Gl} \circ (f \times f))^*\omega_{Gl}.$$

Calculating the left-hand side at $(x, y) \in M \times M$, we have

$$\begin{aligned}(f \circ \mu)^*\omega_{Gl} &= \mu^* f^*\omega_{Gl} = \mu^*(\mathrm{ad}(\omega)) \quad (\text{where } \mathrm{ad}(\omega)v = [\omega, v])\\ &= \mathrm{ad}(\mu^*(\omega))\\ &= \mathrm{ad}(\pi_2^*(f^{-1}(\pi_1^*\omega) + \pi_2^*\omega)\\ &= \mathrm{ad}((f^{-1}(y)(\pi_1^*\omega) + \pi_2^*\omega)\\ &= \mathrm{Ad}(f^{-1}(y))(\pi_1^* ad\ \omega) + \pi_2^*\mathrm{ad}\ \omega \quad (\text{since } f^{-1}(y) \in \mathrm{Aut}_{\mathrm{Lie}}(\mathfrak{g})),\end{aligned}$$

while the right-hand side is

$$\begin{aligned}(\mu_{Gl} \circ (f \times f))^*\omega_{Gl} &= (f \times f)^*\mu_{Gl}^*\omega_{Gl}\\ &= (f \times f)^*\{\pi_2^*(\mathrm{Ad}^{-1})(\pi_1^*\omega_{Gl}) + \pi_2^*\omega_{Gl}\}\\ &= (\pi_2 \circ (f \times f))^*(\mathrm{Ad}^{-1})(\pi_1^*\omega_{Gl}) + (\pi_2 \circ (f \times f))^*\omega_{Gl}\\ &= (f \circ \pi_2)^*(\mathrm{Ad}^{-1})((f \circ \pi_1)^*\omega_{Gl}) + (f \circ \pi_2)^*\omega_{Gl}\\ &= (\mathrm{Ad}(f(y)^{-1})(\pi_1^* f^*\omega_{Gl}) + \pi_2^* f^*\omega_{Gl}\\ &= (\mathrm{Ad}(f(y)^{-1})(\pi_1^*\mathrm{ad}\ \omega) + \pi_2^*\mathrm{ad}\ \omega.\end{aligned}$$

This verifies that f is a homomorphism with respect to μ.

Step 4. Now we show that μ is associative. We shall verify the commutativity of the following diagram.

$$\begin{tikzcd} M \times M \times M \arrow[r, "\mu \times \mathrm{id}"] \arrow[d, "\mathrm{id} \times \mu"'] & M \times M \arrow[d, "\mu"] \\ M \times M \arrow[r, "\mu"] & M \end{tikzcd}$$

It is clear that this diagram commutes at (e, e, e). Since M is connected, it suffices to show that the form ω on M pulls back to the same form on $M \times M \times M$ along the two possible routes. That is, we must show that $(\mu \circ (\mu \times id))^*\omega = (\mu \circ (id \times \mu))^*\omega$. Calculating the left-hand side at $(x, y, z) \in M \times M \times M$, and using $\rho_i \colon M \times M \times M \to M$, $\rho_{ij} \colon M \times M \times M \to M \times M$ to denote projection to the ith or ijth factors, we have

$$\begin{aligned}(\mu \circ (\mu \times \mathrm{id}))^*\omega &= (\mu \times \mathrm{id})^*\mu^*\omega\\ &= (\mu \times \mathrm{id})^*(f(z)^{-1}\pi_1^*\omega + \pi_2^*\omega)\\ &= f(z)^{-1}(\pi_1 \circ (\mu \times \mathrm{id}))^*\omega + (\pi_2 \circ (\mu \times \mathrm{id}))^*\omega\end{aligned}$$

$$\begin{aligned}&= f(z)^{-1}(\mu \circ \rho_{12})^*\omega + \rho_3^*\omega\\ &= f(z)^{-1}\rho_{12}^*\mu^*\omega + \rho_3^*\omega\\ &= f(z)^{-1}\rho_{12}^*(f(y)^{-1}\pi_1^*\omega + \pi_2^*\omega) + \rho_3^*\omega\\ &= f(z)^{-1}f(y)^{-1}(\pi_1 \circ \rho_{12})^*\omega + f(z)^{-1}(\pi_2 \circ \rho_{12})^*\omega + \rho_3^*\omega\\ &= f(z)^{-1}f(y)^{-1}\rho_1^*\omega + f(z)^{-1}\rho_2^*\omega + \rho_3^*\omega.\end{aligned}$$

On the other hand,

$$\begin{aligned}(\mu \circ (\mathrm{id} \times \mu))^*\omega &= (\mathrm{id} \times \mu)^*\mu^*\omega\\ &= (\mathrm{id} \times \mu)^*(f(yz)^{-1}\pi_1^*\omega + \pi_2^*\omega)\\ &= f(yz)^{-1}(\pi_1 \circ (\mathrm{id} \times \mu))^*\omega + (\pi_2 \circ (\mathrm{id} \times \mu))^*\omega\\ &= f(yz)^{-1}\rho_1^*\omega + (\mu \circ \rho_{23})^*\omega\\ &= f(yz)^{-1}\rho_1^*\omega + \rho_{22}^*\mu^*\omega\\ &= f(yz)^{-1}\rho_1^*\omega + \rho_{23}^*(f(z)^{-1}\pi_1^*\omega + \pi_2^*\omega)\\ &= f(yz)^{-1}\rho_1^*\omega + f(z)^{-1}(\pi_1 \circ \rho_{23})^*\omega + (\pi_2 \circ \rho_{23})^*\omega\\ &= f(yz)^{-1}\rho_1^*\omega + f(z)^{-1}\rho_2^*\omega + \rho_3^*\omega\\ &= f(z)^{-1}f(y)^{-1}\rho_1^*\omega + f(z)^{-1}\rho_2^*\omega + \rho_3^*\omega.\end{aligned}$$

This verifies the associativity.

Step 5. Finally, we construct the inversion map. Proposition 4.10 shows that $\iota \colon M \to M$ must have Darboux derivative given by

$$\eta = -\mathrm{Ad}(x)\omega = -f\omega.$$

Thus the form ω will determine ι, provided we can show that $-f\omega$ is integrable. But since

$$\begin{aligned}d\eta + \frac{1}{2}[\eta, \eta] &= -df\ \omega - f\ d\omega + \frac{1}{2}f[\omega, \omega]\\ &= -f\ \mathrm{ad}(\omega)\omega - f\left\{-\frac{1}{2}[\omega, \omega]\right\} + \frac{1}{2}f[\omega, \omega]\\ &= -f[\omega, \omega] + f[\omega, \omega] = 0\end{aligned}$$

and M is simply connected, it follows from the fundamental theorem that η is the Darboux derivative of a unique map $\iota \colon M \to M$ satisfying $\iota(e) = e$. However, we must still verify that ι is the inverse map, that it satisfies $\mu(\iota(x), x) = e$. For this we shall verify the commutativity of the following diagram.

$$\begin{tikzcd} M \times M \arrow[r, "\iota \times \mathrm{id}"] & M \times M \arrow[d, "\mu"] \\ M \arrow[u, "\Delta"] \arrow[r] & e \in M \end{tikzcd}$$

It is clear that this diagram commutes at e. Since M is connected, for the general commutativity it suffices to show that the form ω on M pulls back to the same form on M along the two possible routes. Since it obviously pulls back to 0 via the lower route, we must show the same for the upper route, namely, that $(\mu(\iota \times \mathrm{id})\Delta)^*\omega = 0$. Calculating at $x \in M$, we have

$$\begin{aligned}
\mu \circ (\iota \times \mathrm{id}) \circ \Delta)^*\omega &= \Delta^*(\iota \times \mathrm{id})^*\mu^*\omega \\
&= \Delta^*(\iota \times \mathrm{id})^*(f(x)^{-1}\pi_1^*\omega + \pi_2^*\omega) \\
&= \Delta^*(f(x)^{-1}(\iota \times \mathrm{id})^*\pi_1^*\omega + (\iota \times \mathrm{id})^*\pi_2^*\omega) \\
&= \Delta^*(f(x)^{-1}(\pi_1 \circ (\iota \times \mathrm{id}))^*\omega + (\pi_2 \circ (\iota \times \mathrm{id}))^*\omega) \\
&= \Delta^*(f(x)^{-1}(\iota \circ \pi_1)^*\omega + \pi_2^*\omega) \\
&= \Delta^*(f(x)^{-1}\pi_1^*\iota^*\omega + \pi_2^*\omega) \\
&= \Delta^*(f(x)^{-1}\pi_1^*(-f(x)\omega) + \pi_2^*\omega) \\
&= \Delta^*(-\pi_1^*\omega + \pi_2^*\omega) \\
&= -(\pi_1\Delta)^*\omega + (\pi_2\Delta)^*\omega = -\omega + \omega = 0.
\end{aligned}$$

Thus, ι is an inverse for μ.

The case $\pi_1(M) \neq 1$.

Here the form ω on M pulls up to the form $\pi^*\omega$ on the universal cover G of M. Clearly, $\pi^*\omega$ satisfies the same three conditions on G that ω satisfies on M. Thus case 1 applies to the pair (G, ω), and we see that for any fixed choice of $e \in G$ there is a Lie group structure on G for which e is the identity and $\pi^*\omega$ is the Maurer–Cartan form. By Proposition 8.1, there is a unique injective homomorphism $\tilde{\Phi}$: $\mathcal{Gal}(G/M) \to G$ satisfying $\tilde{\Phi}(T)g = T(g)$ for all $T \in \mathcal{Gal}(G/M)$. Thus, the group of covering transformations is identified with the period group, a discrete subgroup $\Gamma \subset G$. ∎

Remark 8.8. Dropping condition (iii) in Theorem 8.7 still leaves us with the manifold M, which is *locally* a Lie group.

Remark 8.9. If we drop condition (i), we are left with what may be regarded as a "deformation" of a Lie group. For example, we might consider altering the form ω, which does satisfy (i), (ii), and (iii), by the addition of a small form η such that $\omega_t = \omega + t\eta$ still satisfies (ii) for $t \in [0, 1]$ so that ω_1 is a deformation of ω. The new form may also satisfy (iii), but it will not, in general, continue to satisfy (i), and so we will not even get a Lie group locally.

Remark 8.10. Except in special cases, we cannot weaken condition (ii) to the following

(ii)′ $\omega: T_x(M) \to \mathfrak{g}$ is an injection for all x

and still expect to obtain a discrete monodromy group. For example, we could take the case where $G = Sl_3(\mathbf{C})$ (so that $\mathfrak{g} = M_3(\mathbf{C})_0$, the 3×3 matrices of trace 0), $M = S^1$, and ω is the $\mathfrak{g}$-valued 1-form on S^1 which is the composite $\varphi\omega_M$, where ω_M is the Maurer–Cartan form on S^1, and $\varphi: \mathbf{R} \to \mathfrak{g}$ is the Lie algebra homomorphism given by $x \mapsto \mathrm{diag}(ix, \sqrt{2}ix, -(1+\sqrt{2})ix)$. Of course, locally, φ is the derivative of the locally defined map $S^1 \to T^2 \subset Sl_3(\mathbf{C})$ sending $e^{ix} \mapsto \mathrm{diag}(e^{ix}, e^{\sqrt{2}ix}, e^{-(1+\sqrt{2})ix})$. The form ω satisfies the structural equation automatically, since the dimension of M is 1, and since M is compact ω is complete. Moreover, condition (ii)′ is obviously satisfied. What is the monodromy group Γ of ω? Clearly, it is the subgroup generated by the element $\mathrm{diag}(1, e^{2\sqrt{2}i\pi}, e^{-2\sqrt{2}i\pi}) \in T^2$, and this subgroup is *not* a discrete subgroup of G.

Corollary 8.11. *Let G be a connected Lie group and let $\pi: \tilde{G} \to G$ denote the universal covering space. For any choice of $\tilde{e} \in \pi^{-1}(e)$, there is a unique Lie group structure on $\tilde{G}$ such that $\tilde{e}$ is the identity and π is a homomorphism of Lie groups. Moreover, the kernel of π is a discrete central subgroup of $\tilde{G}$.*

Proof. First we note that if such a Lie group structure exists on $\tilde{G}$, then (by Proposition 2.9) the Maurer–Cartan form on $\tilde{G}$ is $\pi^*\omega_G$. Thus, we are led to define a $\mathfrak{g}$-valued 1-form on $\tilde{G}$ by $\omega = \pi^*\omega_G$, and this form satisfies the structural equation since

$$\begin{aligned}
d\omega + \frac{1}{2}[\omega, \omega] &= d\pi^*\omega_G + \frac{1}{2}[\pi^*\omega_G, \pi^*\omega_G] \\
&= \pi^* d\omega_G + \frac{1}{2}\pi^*[\omega_G, \omega_G] \\
&= \pi^*\left(d\omega_G + \frac{1}{2}[\omega_G, \omega_G]\right) \\
&= 0.
\end{aligned}$$

Since π is a local diffeomorphism, not only is $\omega: T(\tilde{G}) \to \mathfrak{g}$ an isomorphism on each fiber but also ω is complete, since integral curves for ω_G on G will lift to integral curves for ω on $\tilde{G}$. By Theorem 8.7, given $\tilde{e} \in \pi^{-1}(e)$, $\tilde{G}$ has a unique Lie group structure with $\tilde{e}$ as the identity and for which ω is the Maurer–Cartan form. By Theorem 5.2, π is the unique map $\tilde{G}, \tilde{e} \to G, e$ pulling back ω from ω_G, and then by Proposition 7.15, π must be a homomorphism.

Finally, since $\ker \pi \subset \tilde{G}$ is a discrete subgroup of a connected group, the function

$$f_z: (\tilde{G}, \tilde{e}) \to (\ker \pi, \tilde{e}), \quad \text{where } z \in \ker \pi, \qquad g \mapsto g^{-1}zg$$

must be constant. Thus, $f_z(g) = f_z(\tilde{e}) = z$, and hence $\ker \pi$ is central in $\tilde{G}$. ∎

The Classical Period Mapping

Remark 8.12. This final remark constitutes a brief discussion of the classical period mapping and is addressed to those who know something about complex analysis. Let $M = M_g$ be a Riemann surface of genus g. Then $H^1(M_g, \mathbf{C}) = \mathbf{Z}^{2g}$ is spanned by g closed (i.e., $d\omega = 0$) holomorphic 1-forms $\omega_1, \ldots, \omega_g$ and their conjugates. Let $\omega = (\omega_1, \ldots, \omega_g)$. This is a closed 1-form on M_g with values in $\mathbf{C}^g = \mathbf{R}^{2g}$ which is the (abelian) Lie algebra of $\mathbf{C}^g$ (so that $[\omega, \omega] = 0$). Thus, ω satisfies the structural equation. Moreover, it is complete since M is compact. It is known from complex analysis that $\omega: T_p(M) \to \mathbf{C}^g$ is complex linear and injective for every $p \in M$ and that the period group is discrete. However, the period group must contain a basis for $\mathbf{C}^g$, for if it were merely to span a proper (complex) subspace V, then the composite

$$M \xrightarrow[\text{mapping}]{\text{period}} \mathbf{C}^g/L \xrightarrow{\text{projection}} \mathbf{C}^g/V$$

would give a nonconstant holomorphic function on M, which, by the maximum modulus principle, is impossible for M compact. Thus, L is a lattice in $\mathbf{C}^g$, and we get the map $M_g \to \mathbf{C}^g/L$, which is classically called the *period mapping*. The receiving group $\mathbf{C}^g/L$ is called the *Jacobian variety*.

4

Shapes Fantastic: Klein Geometries

Let us now do away with the concrete conception of space ... and regard it only as a manifoldness of n dimensions .... By analogy with the transformations of space we speak of transformations of the manifoldness; they also form groups. But there is no longer, as there is in space, one group distinguished above the rest by its signification; each group is of equal importance with every other. The following comprehensive problem then arises as a generalization of geometry:

"Given a manifoldness and a group of transformations of the same; to investigate the configurations belonging to the manifoldness with regard to such properties as are not altered by the transformation of the group." –Felix Klein, 1872

During the 19th century, non-Euclidean geometry made its appearance with the independent discovery by Gauss, Bolyai, and Lobachevski of hyperbolic geometry in the plane. This geometry is very close to Euclidean geometry, satisfying as it does all the axioms except for the existence of a unique line through a point parallel to a given line. Elliptic geometry, the geometry of antipodal pairs of points on the sphere, appeared and again was non-Euclidean in this strong sense. In fact, another geometry, spherical geometry, had been studied by navigators for centuries, without being considered as "non-Euclidean" since it has models in Euclidean 3-space and perhaps also because, considered in its two-dimensional aspect, it more violently violates the axioms of Euclidean geometry. In the course of a few decades, the number of these new geometries proliferated to include affine geometry, projective geometry, Möbius geometry, Lie sphere geometry, Laguerre geometry and so on. Each geometry has its own set of theorems, that is, its own theory. In addition to this there are relations among the geometries. For example a theorem that seems to belong to Euclidean geometry might still be true in projective geometry. Moreover, a purely projective proof, which could employ the larger group of projective symmetries, might well be simpler.[1]

It was Felix Klein's idea to bring order to these new geometries by means of the unifying notion of the *principal group*[2] of a geometry. He noticed that each geometry, M say, is a connected manifold and has a Lie group G of "motions" acting *transitively*[3] on it and, moreover, all the properties of figures studied in the geometry remain invariant under these motions. If, in addition, the action is *effective*,[4] we may speak of an *effective Klein geometry*. It is this group G of motions that Klein called the *principal group* of the geometry, or more briefly, the group of the geometry. In the case of Euclidean geometry, the properties studied are angle and length, and the group is the group of rigid motions; for projective geometry, the properties are concurrence of lines and collinearity of points and the group is the group of projective transformations, and so forth. Thus, Klein extended the notion of geometry by defining a geometry to consist of a Lie group G, a smooth manifold M, and a smooth, transitive, and effective action of G on M.[5] The study of such a geometry is the study of those properties of figures in M that remain invariant under the action of G.

[1] Associated to each of these geometries is not only a simple theory (a *goniometry*) in which one studies figures like triangles but also a *differential geometry* in which one studies the "figures" consisting of smooth curves and surfaces in the geometry. A great deal of effort was spent in studying these generalizations of differential geometry in the last decades of the 19th century and in the first four decades of this century.

[2] Klein called this group the *haugtgruppe* of the geometry. In his translation of Klein's Erlangen program [F. Klein, 1872], Haskell renders this as *principal group*. This has the two-fold advantage that the modern notion of a principal bundle generalizes this group and that the competing English term *fundamental group*, which many authors have used, is easily confused with Poincaré's completely accepted use of the latter term for $\pi_1(M)$.

[3] Let $G \times M \to M$ sending $(g, x) \to gx$ be a group action. Recall that the action is said to be *transitive* if, for any pair of points $x, y \in M$, there is an element $g \in G$ such that $gx = y$.

[4] An action is said to be *effective* if $gx = x$ for all $x \in M$ implies that g is the identity element of G.

[5] The formal definition of a Klein geometry is given in §3. We shall not insist on an effective action.

Klein's generalization of geometry allows us to shift the emphasis from the space M to the group G in the following manner. Let us fix a base point $x \in M$. Then there is a map $\pi: G \to M$ sending g to gx. By the transitivity of the action, this map is onto; but it is not one to one. The inverse image $\pi^{-1}(x) = \{x \in G \mid gx = x\} \stackrel{\text{def}}{=} H_x$ is clearly a closed subgroup of G. It is called the *stabilizer* of x. Moreover, if $y \in M$, it is clear that $\pi^{-1}(y) = \{g \in G \mid gx = y\} = g_0 H_x$, where g_0 is an arbitrary element of $\pi^{-1}(y)$. It follows that π induces a bijection $G/H_x \to M$. The condition that the action be effective translates into saying that the subgroup $N = \{g \in G \mid gp = p$ for all $p \in M\}$ is trivial.

Exercise 0.1. If we do not assume that the action is effective, show that this subgroup N is the largest subgroup of H_x that is normal in G. ❑

Thus, to summarize, we may say that instead of describing a geometry with base point as a pair (M, x) together with its principal group G, we could equally well describe it as the pair (G, H), where $H = H_x$ is a closed subgroup of G. Of course, all this depends on the choice of base point x. This base point may be regarded as "a constant of integration," which disappears in the infinitesimal description of Klein geometries.

Exercise 0.2. Show that different choices of x lead to conjugate subgroups H. ❑

It is clear that in these Klein geometries it is impossible to distinguish a point from any other by properties of the geometry since, by definition, the transitive action of G preserves these properties. Thus, the Klein geometries are fully homogeneous. We might say that they are "flat" or that they have no "lumps" in them.

Although we lend the name "Klein geometry" to the geometries studied in this chapter—and the justice of this should be clear from the quotation at the beginning of the chapter—Klein himself was interested in only the few cases that correspond most closely to our experience: projective geometry and its sons and daughters, the parabolic, Euclidean, and elliptic geometries. Thus, he would have felt our present point of view to be too general. It was that remarkable pioneer Wilhelm Killing whose passion it was to try to understand these geometries in general.

Our aim in this chapter is not to study the theories of individual Klein geometries, but rather to try to understand their nature in general.

§1. Examples of Planar Klein Geometries

Let us see how Klein's ideas, described in this chapter's Introduction, work in some cases of more or less familiar two-dimensional geometries.[6] Each of these planar geometries is merely a single example drawn from an infinite series of higher-dimensional analogs.

"School Geometry" (the Euclidean plane). In this case $M = \mathbf{R}^2$. The Lie group of *symmetries* or *congruences* or *rigid motions*[7] is

$$G = Euc_2(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ v & R(\theta) \end{pmatrix} \middle| R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\},$$

which acts on M by the formula

$$\begin{pmatrix} 1 & 0 \\ v & R(\theta) \end{pmatrix} \cdot x = R(\theta)x + v, \quad \text{where } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

An easy calculation shows that the stabilizer of the origin of $\mathbf{R}^2$ is the subgroup of rotations

$$H = SO_2(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & R(\theta) \end{pmatrix} \middle| \theta \in \mathbf{R} \right\}.$$

Then we have the basic identification $Euc_2(\mathbf{R})/SO_2(\mathbf{R}) \approx \mathbf{R}^2$. In fact, $Euc_2(\mathbf{R}) \approx SO_2(\mathbf{R}) \times \mathbf{R}^2$ (the semidirect product with respect to the standard representation of $SO_2(\mathbf{R})$ on $\mathbf{R}^2$).

Next follow the two non-Euclidean[8] geometries.

Hyperbolic Plane. In this case $M = \{z = x + iy \in \mathbf{C} \mid y > 0\}$. The group is the group of Möbius transformations $G = Sl_2(\mathbf{R}) = \{A \in M_2(\mathbf{R}) \mid \det A = 1\}$, which acts on M by the formula

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

It is easy to check that this is indeed an action on M and also that the stabilizer of the point $i \in M$ is the circle group

[6] A complete list of all 23 families of two-dimensional Klein geometries (13 of which are singleton families) may be found in [B. Komrakov, A. Churyumov, and B. Doubrov, 1993].

[7] Of course, in picking this group we have rather arbitrarily chosen to ignore the orientation-reversing congruences. In fact, there are actually two distinct interpretations of "school geometry."

[8] That is, non-Euclidean in the strong sense. See the introduction to this chapter.

$$SO_2(\mathbf{R}) = \{A \in M_2(\mathbf{R}) \mid \det A = 1 \text{ and } AA^t = I\}.$$

Elliptic Plane. This is the real projective space $\mathbf{P}^2(\mathbf{R})$, which may be regarded as the set of unoriented one-dimensional subspaces of three-dimensional Euclidean space $\mathbf{R}^3$, with coordinates x_0, x_1, x_2. (The coordinate charts of $\mathbf{P}^2(\mathbf{R})$ "are" the 2-planes in $\mathbf{R}^3$ that avoid the origin, cf. Example 1.3.) The group is $SO_3(\mathbf{R})$, whose action on the lines in $\mathbf{R}^3$ is induced by the standard action of $SO_3(\mathbf{R})$ on the ambient space $\mathbf{R}^3$. The stabilizer of a line, say the x_0-axis, is $O_2(\mathbf{R})$, which is embedded in $SO_3(\mathbf{R})$ via

$$A \mapsto \begin{pmatrix} \det A^{-1} & 0 \\ 0 & A \end{pmatrix}.$$

Complementing these, we have a series of six other geometries.

Geometry of Similarity. This geometry is a relative of Euclidean geometry that may also be regarded as a part of "school geometry." Here $M = \mathbf{R}^2$ and the group is the two-dimensional conformal group given by $G = H \times \mathbf{R}^2$ (semidirect product),

$$H = \left\{ A \in Gl_2^+(\mathbf{R}) \middle| AA^t = \begin{pmatrix} r^2 & 0 \\ 0 & r^2 \end{pmatrix} \text{ for some } r \in \mathbf{R} \right\}.$$

An element $A \in H$ has the form $A = \begin{pmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{pmatrix}$ acting by matrix multiplication on the plane.

Lorentz (or Minkovski) Plane. Here $M = \mathbf{R}^2$ and the group is the oriented two-dimensional Lorentz group given by $G = H \times \mathbf{R}^2$ (semidirect product), where

$$H = O_{1,1}^+(\mathbf{R}) = \{A \in Sl_2(\mathbf{R}) \mid A^t \Sigma A = \Sigma\}, \quad \text{where } \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

An element $A \in H$ has the form $A = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$ (a *boost*) acting by matrix multiplication on the plane.

Affine Plane. Here $M = \mathbf{R}^2$ and the group is the two-dimensional affine group given by

$$G = Aff_2^+(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in Gl_3^+(\mathbf{R}) \middle| A \in Gl_2^+(\mathbf{R}),\ v \in \mathbf{R}^2 \right\}.$$

The action on M is given by the formula $\begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \cdot x = Ax + v$, where x and v are column vectors. Again, it is easily verified that the stabilizer of the origin is

$$H = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \in Gl_3(\mathbf{R}) \middle| A \in Gl_2^+(\mathbf{R}) \right\} \approx Gl_2^+(\mathbf{R})$$

and $G \approx H \times \mathbf{R}^2$ (semidirect product).

Real Projective Plane. The space is the real projective space $\mathbf{P}^2(\mathbf{R})$ just as in the elliptic case, but the group is now the full general linear group $Gl_3(\mathbf{R})$. The group $Gl_3(\mathbf{R})$ acts on $\mathbf{R}^3$ in the standard way and so it also acts on the "lines" in $\mathbf{R}^3$. The stabilizer of the x_0-axis is H, where

$$H = \left\{ \begin{pmatrix} \lambda & v \\ 0 & A \end{pmatrix} \in Gl_3(\mathbf{R}) \middle| A \in Gl_2(\mathbf{R}),\ v \in \mathbf{R}^2,\ \lambda \in \mathbf{R}^* \right\}$$

and $N = \{\lambda I \in Gl_3(\mathbf{R}) \mid \lambda \in \mathbf{R}^*\}$ is the largest normal subgroup of G in H.

Möbius Plane. Next we consider the *Möbius plane.* Here the space M is the standard 2-sphere, $S^2 = \{x \in \mathbf{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$. The principal group is the oriented Lorentz group $L_{3,1}(\mathbf{R}) = \{A \in Sl_4(\mathbf{R}) \mid A^t\Sigma A = \Sigma\}$, where

$$\Sigma = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.$$

This time, however, it is more difficult to describe the action. The Lorentz group $G = L_{3,1}(\mathbf{R})$ acts on $\mathbf{R}^4$ so that the level sets of the quadratic form $-2x_0x_3 + x_1^2 + x_2^2$ are preserved. In particular, the *light cone* L, given by $-2x_0x_3 + x_1^2 + x_2^2 = 0$, is preserved. Now the light cone is homogeneous in the sense that if $x \in L$, then $\lambda x \in L$ for all $\lambda \in \mathbf{R}$. Thus, L is a union of lines through the origin, and the linear action of G on $\mathbf{R}^4$ restricts to give an action on this set of lines. Each such line meets the hyperplane $x_0 + x_3 = \sqrt{2}$ in exactly one point, so the set of lines in the light cone may be identified with the intersection

$$\begin{aligned} &L \cap \{x \in \mathbf{R}^4 \mid x_0 + x_3 = \sqrt{2}\} \\ &= \{x \in \mathbf{R}^4 \mid x_0 + x_3 = \sqrt{2}, -2x_0x_3 + x_1^2 + x_2^2 = 0\}. \end{aligned}$$

The change of variables

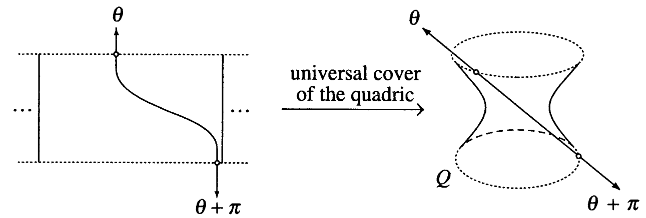
$$\begin{aligned} &x_0 = (1 + y_0)/\sqrt{2}, x_1 = y_1, x_2 = y_2, \\ &x_3 = (1 - y_0)/\sqrt{2}, \end{aligned}$$

identifies the set of lines in the light cone with the 2-sphere

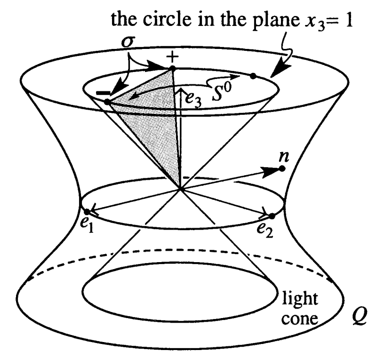
$$\{y \in \mathbf{R}^3 \mid y_1^2 + y_2^2 + y_3^2 = 1\}.$$

This yields the action of G on S^2.

Lie Sphere Plane (cf. [T.E. Cecil and S.-S. Chern, 1987] or [T.E. Cecil, 1980]). From one point of view, the space here is the quadric Q in $\mathbf{P}^3(\mathbf{R})$, given by $-x_0^2 + x_1^2 + x_2^2 - x_3^2 = 0$. From another point of view, it is the space Λ of lines on the quadric Q. Λ appears to be more important for Lie sphere geometry then Q (cf. the references cited above). However, dim $\Lambda = 3$, so this space is not a two-dimensional geometry, and we will restrict our remarks to Q, which does have dimension 2. Q is the *projective completion*[9] of a hyperboloid of 1-sheet in $\mathbf{R}^3$ given by $x_1^2 + x_2^2 - x_3^2 = 1$ and so is diffeomorphic to a torus. These facts can be pieced together from the following diagram.



The finite points of Q have the interpretation as oriented 0-spheres on S^1 as follows. For a finite point $n \in Q$, we have the oriented orthogonal plane $n^\perp$ meeting the circle on the light cone given by $x_3 = 1$ in two (distinct) oriented points (one plus, one minus), that is, in an oriented zero spehre σ. As n tends to infinity, the two points of the corresponding sphere σ move together to yield an unoriented point sphere in the limit.



[9]Briefly, given any set of polynomials $P_j(x) = 0$, $j = 1, \ldots, r$, where $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, we have, on the one hand, the locus $X \subset \mathbf{R}^n$ given by $X = \{x \in \mathbf{R}^n \mid P_j(x) = 0, j = 1, \ldots, r\}$. On the other hand, we may homogenize the polynomials P_j to get polynomials Q_j in the variables $y = (x_0, \ldots, x_n)$ by setting $Q_j(y) = x_0^{d_j} Q_j(x/x_0)$, where the powers d_j are chosen just larger enough to make the Qs polynomials. Then the equations $Q = 0$ have a locus $Y \subset P^n$, and $Y \cap \mathbf{R}^n = X$. Then Y is called the projective completion of X.

(A point sphere can be perturbed in a pencil of spheres containing two copies of each unoriented sphere, one for each orientation; a point sphere may be thought of as the moment of change of orientation in a pencil of infinitesimal oriented spheres.) The group is $G = O^+_{2,2}(\mathbf{R}) = \{A \in Sl_4(\mathbf{R}) \mid A^t\Sigma A = \Sigma\}$, where $\Sigma = \mathrm{diag}(-1,1,1,-1)$. G acts transitively on the quadric, and we take H to be the stabilizer of the line spanned by $(1,1,0,0)$ corresponding to the equatorial 0-sphere S^0 consisting of the two points $(\pm 1,0,0)$ in the figure above.

We mention one other plane geometry, which, although identical to one already considered, in its higher-dimensional incarnations becomes an independent geometry.

Complex Projective Line. Here $M = S^2$, the Riemann sphere (or the complex projective line), and the group G is the group of Möbius transformations of M, $G = Sl_2(\mathbf{C})$, with

$$H = \left\{ A \in Sl_2(\mathbf{C}) \middle| A = \begin{pmatrix} \star & 0 \\ \star & \star \end{pmatrix} \right\}.$$

Exercise 1.1. Show that $Sl_2(\mathbf{C})$ is isomorphic to the Lorentz group and that the complex projective line is the same as the Möbius plane. ❑

§2. Principal Bundles: Characterization and Reduction

The examples of the last section make it clear that the study of "non-Euclidean geometry" really is the study of the left coset spaces G/H, where H is a closed subgroup of the Lie group G. There are, however, some technical questions that begin to arise. For example, it is not so obvious that a geometry in Klein's sense is always a smooth manifold. The purpose of this section is to deal with this difficulty and also to show that the map $G \to G/H$ is a principal H bundle.[10] In fact, in anticipation of later need, we shall prove something still more general. We are going to characterize the principal H bundles $P \to M$ in terms of properties of the action $P \times H \to P$.

At the end of the section we give a brief study of the notion of a reduction of a principal bundle.

Definition 2.1. Let P be a smooth manifold, H a Lie group, and $P \times H \to P$ a smooth action.

[10] Such bundles constitute the first and motivating examples of principal bundles; the Hopf bundle $S^1 \to S^3 \to S^2$ studied in the first chapter is, of course, among these.

(i) The action is called *free* if $ph = p$ for some $p \in P \Rightarrow h = e$.

(ii) The action is called *proper* if A and B compact $\Rightarrow \{h \in H \mid Ah \cap B \neq \emptyset\}$ is compact. ❀

Proposition 2.2. *Let G be a Lie group and $H \subset G$ a closed subgroup. Then the action $G \times H \to G$ given by multiplication is free and proper.*

Proof. *Free action:* If $gh = g$, then multiplying by g^{-1} yields $h = e$.

Proper action: Let A and B be compact subsets of G, and let $K = \{h \in H \mid Ah \cap B \neq \emptyset\}$. We show that K is compact. Suppose that $h_1, h_2, \ldots$ is an infinite sequence in K. Then there are infinite sequences $a_1, a_2, \ldots$ in A and $b_1, b_2, \ldots$ in B such that $a_j h_j = b_j$ for all j. Since A is compact, we may pass to a subsequence $J \subset \{1, 2, \ldots\}$ such that $\{a_j\}_{j\in J}$ converges to a. Passing to a further subsequence $L \subset J$, we may also assume that $\{b_j\}_{j\in L}$ converges to b. Thus $\{h_j = (a_j)^{-1} b_j\}_{j\in L}$ converges to ba^{-1} and every sequence in K has a convergent subsequence. Thus K is compact. ■

Next we verify that right principal H bundles have these properties.

Proposition 2.3. *Let $\xi = (P, \pi, M, H)$ be a right principal H bundle. Then the action $P \times H \to P$ is free and proper.*

Proof. The fact that the bundle is locally a product $U \times H$, with the canonical right H action, shows that the action is free. The argument in Proposition 2.2 shows that, for any sequence $\{h_j\}$ in $K = \{h \in H \mid Ah \cap B \neq \emptyset\}$, there are sequences $\{a_j\}_{j\in L}$ in A and $\{b_j\}_{j\in L}$ in B converging to a and b, respectively, and such that $a_j h_j = b_j$ and for all j. Now $\pi(a_j) = \pi(a_j h_j) = \pi(b_j)$ for all j, so $\pi(a) = \pi(b)$. Thus $a = hb$ for some $h \in H$. We claim that $\lim\{h_j\}_{j\in L} = h$. To see this, note that for large j all of the terms in the sequences lie in a single coordinate chart. This allows us to assume we have the trivial bundle $P = M \times H$. Projecting on the H factor by ρ leads to the equation $\rho(a_j)h_j = \rho(b_j)$ in the group H. Thus $\{h_j = \rho(a_j)^{-1}\rho(b_j)\}_{j\in L}$ converges to $\rho(a)^{-1}\rho(b)$. Thus, every sequence in K has a convergent subsequence. ■

Now we complete this cycle of ideas by showing that these properties characterize principal bundles over a smooth manifold. We have the following result whose proof, being rather technical, is given in Appendix E.

Theorem 2.4. *Let P be a smooth manifold, H a Lie group, and $\mu\colon P \times H \to P$ a smooth, free, proper right action. Then*

(i) *P/H with the quotient topology is a topological manifold* ($\dim P/H = \dim P - \dim H$),

(ii) *P/H has a unique smooth structure for which the canonical projection $\pi: P \to P/H$ is a submersion,*

(iii) *$\xi = (P, \pi, P/H, H)$ is a smooth principal right H bundle.*

Example 2.5. It is important that the action be proper. For example, in the case of the right action $G \times H \to G$ by a subgroup $H \subset G$, if the subgroup is not closed, then the action is not proper and the quotient G/H need not be a manifold. The standard example of this is the case of the torus $G = T^2 = \mathbf{R}^2/\mathbf{Z}^2$ and an *irrational* subgroup H of G. To describe the meaning of this, we pass to the universal cover $G_1 = \mathbf{R}^2$ with the corresponding lattice $\mathbf{Z}^2 = \{(a,b) \in \mathbf{R}^2 \mid a, b \in \mathbf{Z}\}$ as kernel of the projection map $\mathbf{R}^2 \to \mathbf{R}^2/\mathbf{Z}^2$. Let $H_1 = \{(t, \sqrt{2}\,t) \mid t \in \mathbf{R}\}$. In G_1 the subgroup H_1 is a straight line, so it is closed. However, since $\sqrt{2}$ is irrational, its image H in G is a *dense* subgroup; in particular, H is not closed. To see how bad the quotient topology on G/H is, note that if $f: G/H \to \mathbf{R}$ is any continuous function, then its lift to G is also continuous. Since f is constant on a dense subset of G, it is constant everywhere. Thus there are no nonconstant continuous functions on G/H. In particular, it is not a manifold. ◆

Example 2.6. From the preceding example we might be led to wonder if the problem arises only from the fact that the torus T^2 fails to be simply connected. Might it be true that for simply connected groups G, subalgebras of $\mathfrak{g}$ may always be realized as closed subgroups of G? The answer is again no in general, for many simply connected groups contain the torus as a subgroup; and so the phenomenon of the previous example persists. For example, $Sl_3(\mathbf{C})$ is simply connected, and yet it contains the torus T^2 as the subgroup $\mathrm{diag}(e^{i\theta}, e^{i\varphi}, e^{-i(\theta+\varphi)})$, $\theta, \varphi \in \mathbf{R}$. ◆

Exercise 2.7. Show that if G is a Lie group and H is a closed normal subgroup of G, then G/H has a unique Lie group structure such that the projection map $\pi: G \to G/H$ is a homomorphism of Lie groups. ❑

Exercise 2.8. Let

$$\begin{aligned} \rho: X \times H &\to X \\ (x,h) &\mapsto x \cdot h \end{aligned}$$

be a right action of a topological group H on a Hausdorff space X. Show that ρ is proper as an action if and only if the map

$$\begin{aligned} X \times H &\to X \times X \\ (x,h) &\mapsto (x, x \cdot h) \end{aligned}$$

is proper. ❑

Exercise 2.9. Let

$$\begin{aligned} \rho: X \times H &\to X \\ (x,h) &\mapsto x \cdot h \end{aligned}$$

be a proper right action of a topological group H on a Hausdorff space X. Show that

$$\begin{aligned} X \times H \times H &\to X \times H \\ (x,h,k) &\mapsto (x \cdot k, k^{-1}h) \end{aligned}$$

is also a proper action. ❑

Exercise 2.10. Let W be a real n-dimensional vector space. Let $\mathrm{Gr}_s(W)$ denote the set of all s-dimensional subspaces of W. Show that $\mathrm{Gr}_s(W)$ can be given a unique smooth structure such that, if $V \in \mathrm{Gr}_s(W)$ and $W = U \oplus V$, then the canonical map $\mathrm{Hom}(V, U) \to \mathrm{Gr}_s(W)$ sending $f \mapsto \mathrm{graph}(f)$ is a smooth map onto an open subset. [Hint: Fix a basis and an inner product in W. Show that the map $O_n(\mathbf{R}) \to \mathrm{Gr}_s(W)$ sending A to the subspace of W spanned by the first s columns of A induces a bijection of sets $O_n(\mathbf{R})/(O_s(\mathbf{R}) \times O_{n-s}(\mathbf{R})) \to \mathrm{Gr}_s(W)$.] ❑

Definition 2.11. The manifold $\mathrm{Gr}_s(W)$ described in Exercise 2.10 is called the *Grassmannian of s-dimensional subspaces of W.* ❀

Reduction of Principal Bundles

An important procedure in working with principal bundles is the notion of a reduction of a principal bundle.

Definition 2.12. Let H be a Lie group and $H_0 \subset H$ a subgroup. Let $H \to P \to M$ be a principal H bundle over M. An H_0 *reduction* of this bundle is a submanifold $P_0 \subset P$ such that $P_0 \to M$ is an H_0 bundle and the action of H_0 on P_0 is the restriction of the action of H on P. ❀

The following lemma is a preparation for studying one way in which reductions arise.

Lemma 2.13. *Let*

$$\begin{aligned} \mu: G \times M &\to M \\ (g,x) &\mapsto g \cdot x \end{aligned}$$

be a smooth left action of a Lie group G on a connected smooth manifold M. Then every orbit $X \subset M$ of this action is a submanifold. Moreover, if the action is proper, then X is a proper submanifold.

Proof. Fix an orbit $X \subset M$ and choose $x_0 \in X$. Set $H = \{g \in G \mid gx_0 = x_0\}$. It will suffice to show that H is a closed subgroup of G, that the induced map

$$\begin{aligned} G/H &\to M \\ gH &\mapsto g{\cdot}x_0 \end{aligned}$$

is an injective immersion with image X, and that, if the original action is proper, then $G/H \to M$ is a proper embedding.

Step 1. H is a closed subgroup of G, and the induced map

$$\begin{aligned} G/H &\to M \\ gH &\mapsto g{\cdot}x_0 \end{aligned}$$

is injective with image X.

Since $\mu^{-1}(x_0)$ is closed and $H \times x_0 = \mu^{-1}(x_0) \cap (G \times x_0)$, it follows that $H \times x_0$ is closed in $G \times M$ and hence H is closed in G. The fact that the map is injective with image X is clear.

Step 2.

$$\begin{aligned} G/H &\to M \\ gH &\mapsto g{\cdot}x_0 \end{aligned}$$

is an immersion.

Define

$$\begin{aligned} \Psi: G &\to M \\ g &\mapsto g{\cdot}x_0 \end{aligned}$$

and note that $\Psi^{-1}(x_0) = H$. Set $V = \ker \Psi_{*e} \subset T_e(G) = \mathfrak{g}$, the Lie algebra of G, and let $\mathfrak{h}$ denote the Lie algebra of H. Since Ψ is constant on H, it follows that $\Psi_{*e}(T_e(H)) = 0$, and hence $\mathfrak{h} \subset V$. We are going to show that $V \subset \mathfrak{h}$. Since the left translations

$$\begin{aligned} L_g: G &\to G \\ h &\mapsto gh \end{aligned} \qquad \text{and} \qquad \begin{aligned} l_g: M &\to M \\ x &\mapsto gx \end{aligned}$$

are diffeomorphisms, and since $\Psi \circ L_g = l_g \circ \Psi$, it follows that

$$\Psi_{*g} \circ L_{g*e} = l_{g*x_0} \circ \Psi_{*e},$$

and hence $\ker \Psi_{*g} = L_{g*e}\ker(\Psi_{*e}) = L_{g*e}V$ for all $g \in G$. Thus on G, Ψ has constant rank $r = \dim \mathfrak{g} - \dim V$, and so by Theorem **1.1.31** the components of the level surfaces of Ψ foliate G with codimension r. The component of the identity in $\Psi^{-1}(x_0) = H$ is the identity component subgroup H_0 with Lie algera $\mathfrak{h}$. Thus $V \subset \mathfrak{h}$ and so $V = \mathfrak{h}$. Now consider the following commutative diagram.

$$\begin{tikzcd} T_g(G) \arrow[d] \arrow[dr, "\Psi_{*g}"] & \\ T_{gH}(G/H) \arrow[r] & T_{gx_0}(M) \end{tikzcd}$$

Since the two downward arrows in the diagram both have the kernel $L_{g*e}V$, it follows that the bottom map is injective, so

$$\begin{aligned} G/H &\to M \\ gH &\mapsto g{\cdot}x_0 \end{aligned}$$

is an immersion.

Step 3. If the original action is proper, then $G/H \to M$ is a proper embedding.

Suppose that $\mu: G \times M \to M$ is proper. Let $C \subset M$ be any compact set. Then

$$\Psi^{-1}(C) = \{g \in G \mid gx_0 \in C\},$$

and the latter is compact since C and $\{x_0\}$ are compact. Hence the image of $\Psi^{-1}(C)$ in G/H is also compact. But this image is just the inverse image of C under the map $G/H \to M$. ■

We now apply this lemma to obtain a simple but useful sufficient condition for the existence of a reduction.

Proposition 2.14. *Suppose that $H \to P \to M$ is a smooth principal H bundle. Let Q be a manifold equipped with a smooth, proper, right H action. Let $f: P \to Q$ be a smooth equivariant map (i.e., $f(ph) = f(p)h$ for all $h \in H$). Fix a point $q_0 \in Q$, and set $H_0 = \{h_0 \in H \mid q_0h_0 = q_0\}$. Suppose that $q_0 \in Q$ lies in the image under f of each fiber of P. Then*

(i) *$P_0 = f^{-1}(q_0)$ is an H_0 reduction of P.*

Fix another point $q_1 \in Q$, and set $H_1 = \{h_1 \in H \mid q_1h_1 = q_1\}$. Suppose that $q_1 \in Q$ also lies in the image under f of each fiber of P, so that by (i) *$P_1 = f^{-1}(q_1)$ is also an H_0 reduction of P. Then*

(ii) *there exists an element $h \in H$ such that $H_1 = h^{-1}H_0h$ and, for any such element h, we have $P_1 = P_0h$.*

Proof. (i) First we show that P_0 is a submanifold of P. For this it suffices to show that f has constant rank. Since q_0 lies in the image under f of each fiber of P, it follows that the image of f is the orbit $X = q_0H \subset Q$. By Lemma 2.13, since the action $Q \times H \to Q$ is proper, all the orbits, and in particular X, are proper submanifolds of Q. Since f takes values in X, it follows that $\operatorname{rank}_p f \leq \dim X$. We show equality by verifying that, for each fiber, $\operatorname{rank}(f \mid \text{fiber}) = \dim X$. But the map that f restricts to on any fiber is, up to change of coordinates, the map $H \to X$ sending $h \to q_0h$, which has rank $= \dim X$ by the lemma. Thus f has constant rank and P_0 is a submanifold of P.

Next we note that

$$h \in H_0 \Leftrightarrow q_0 h = q_0 \Leftrightarrow f^{-1}(q_0 h) = f^{-1}(q_0) \Leftrightarrow f^{-1}(q_0)h = f^{-1}(q_0)$$
$$\Leftrightarrow P_0 h = P_0,$$

so we see that P_0 is stable under the induced action of H_0.

Let p, p' lie in the same fiber of P_0. Then there is an $h \in H$ such that $p' = ph \in P_0$. Thus $q_0 = f(ph) = f(p)h = q_0 h$, and so $h \in H_0$. It follows that the induced action of H_0 on P_0 is transitive on each fiber of P_0. Since $P_0 = f^{-1}(q_0)$ is closed in P and H_0 is closed in H, the induced action of H_0 on P_0 is proper and free. By Theorem 2.4, P_0 is a principal H_0 bundle over M and is therefore an H_0 reduction of P.

(ii) This part is easy, and so we leave it to the reader. ■

Exercise 2.15. Suppose that $H \to P \to M$ is a smooth principal H bundle. Let $\rho: H \to Gl(V)$ be a representation. Give sufficient conditions for the following subsets $S \subset P$ to be reductions.

(i) Let $f: P \to V$ transform according to $f(ph) = \rho(h^{-1})f(p)$. Let $v_0 \in V$, and take $S = f^{-1}(v_0)$.

(ii) Let $\mathrm{Gr}_n(V)$ = the Grassmannian of n-dimensional subspaces of V and $\tilde{\rho}$ denote the standard action of H on $\mathrm{Gr}_n(V)$ induced by ρ. Let $f: P \to \mathrm{Gr}_n(V)$ transform according to $f(ph) = \tilde{\rho}(h^{-1})f(p)$. Let $V_0 \in \mathrm{Gr}_n(V)$, and take $S = f^{-1}(V_0)$.

§3. Klein Geometries

Now we are in a position to give the formal definition for Klein geometries, which we have roughly described in the introduction to this chapter. However, to flesh out the context in which they appear, we first prove the following result.

Proposition 3.1. *Let G be a Lie group and $H \subset G$ a closed subgroup. Then there is a unique maximal normal subgroup K of G lying in H. Moreover, K is a closed Lie subgroup of H, the left action of G on G/H induces a left action of G/K on G/H, and there is a diffeomorphism $(G/K)/(H/K) \to G/H$ commuting with the canonical left G/K actions.*

Proof. Let K be the group generated by all the normal subgroups of G that lie in H. Then K is clearly a normal subgroup of G that lies in H and is, moreover, the unique maximal normal subgroup of G that lies in H. Since the closure of K is also a normal subgroup of G that lies in H, it follows that K itself is closed. Now apply the general result of Kuranishi and Yamabe (cf. [H. Yamabe, 1950]) that a subgroup of a Lie group is a Lie group[11] to see that K is a Lie group. Finally, by Theorem 2.4, the canonical map $\pi: G/K \to G/H$ is a principal fiber bundle with fiber H/K. Thus π induces a diffeomorphism $(G/K)/(H/K) \to G/H$ commuting with the canonical left G/K actions. ■

Definition 3.2. A *Klein geometry* is a pair (G, H), where G is a Lie group and $H \subset G$ a closed subgroup such that G/H is connected. G is called the *principal group* of the geometry. The *kernel* of a Klein geometry (G/H) is the largest subgroup K of H that is normal in G. A Klein geometry (G, H) is *effective* if $K = 1$ and *locally effective* if K is discrete. A Klein geometry is *geometrically oriented* if G is connected. The connected coset space $M = G/H$ is called the *space of the Klein geometry* or sometimes, by abuse of notation, merely the *Klein geometry.*[12] A Klein geometry is called *primitive* if the identity component $H_e \subset H$ is maximal among the proper closed connected subgroups of G. A Klein geometry is called *reductive* if there is an Ad H-module decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$, where $\mathfrak{g}$ and $\mathfrak{h}$ are the Lie algebras of G and H, respectively. ❀

We use the remainder of this subsection to discuss some aspects of this definition.

If (G, H) is a Klein geometry with kernel K, then, for any closed subgroup $N \subset K$ that is normal in G, by Proposition 3.1 the pair $(G/N,\ H/N)$ is also a Klein geometry with space $(G/N)/(H/N) \approx G/H$. Of course, these geometries are all ineffective except when $N = K$. This leads to the following.

Definition 3.3. Let (G, H) be a Klein geometry with kernel K. Then the Klein geometry $(G/K,\ H/K)$ is called the *associated effective Klein geometry.* ❀

Exercise 3.4. Let K_e be the identity component of K. Show that $(G/K_e, H/K_e)$ is a locally effective Klein geometry with kernel K/K_e. ❑

The reader may be wondering why all this fuss about ineffective Klein geometries. Proposition 3.1 seems to indicate that if we are interested in the geometry of the coset space G/H we need only consider the *effective case.* Why not *define* Klein geometries to be effective? The first point is that doing this would eliminate the subtle phenomenon of spin. While we do not deal with this notion in this book, we can keep it within the scope of

[11] What they show is that a path-connected subgroup of a Lie group is a Lie subgroup. Since we allow uncountably many components for our manifolds, this is enough to ensure that any subgroup of a Lie group is a Lie subgroup.

[12] Note that the Euclidean plane and the affine plane have the same *space* (i.e., $\mathbf{R}^2$) even though the geometries are distinct.

our definition by allowing our Klein geometries to be merely *locally effective* rather than demanding effectiveness. We may say that the geometrically important cases of Klein geometries are the locally effective ones. The second point to be made is that ineffective geometries, even ones that are not locally effective, do turn up, so we might as well have a language that allows us to speak of them. For example, of the planar geometries given in §1, only four were given with an effective description (Euclidean geometry, hyperbolic geometry, geometry of similarity, and Lorentz geometry). We could have done this for the others, too, but then we would have had to describe the kernels.

The space G/H of a geometrically oriented Klein geometry need not be topologically orientable in the sense of Definition **1**.1.9. For example, the effective version of the real projective plane has connected principal group $G = PSl_3(\mathbf{R})$ and so is geometrically oriented, but the projective plane is not topologically orientable. On the other hand, consider the subgeometry (G, H) of the geometry of similarity, where H is generated by $SO_2(\mathbf{R})$ and the dilation

$$\begin{aligned} \mathbf{R}^2 &\to \mathbf{R}^2 \\ v &\mapsto 2v \end{aligned}$$

and $G = H \times \mathbf{R}^2$ (semidirect product). The space here is $\mathbf{R}^2$, which is topologically orientable, but since G has infinitely many components the Klein geometry (G/H) is not geometrically orientable. Note that G acts on the space $\mathbf{R}^2$ by transformations preserving the topological orientation.

The real meaning of the notion of a geometrically oriented geometry will become clear only in §4 of Chapter 5 when we extend the definition to include Cartan geometries.

The following result shows that every Klein geometry (G, H) determines a geometrically oriented geometry with the same space but with a possibly smaller principal group.

Proposition 3.5. *Let (G, H) be a pair of Lie groups, where H is a closed subgroup of G, and suppose that G/H is connected. Let G_e be the identity component of G, and set $H_0 = H \cap G_e$. Then*

(i) $G = G_e \cdot H$,

(ii) $G/H = G_e/H_0$.

Proof. (i) It suffices to show that $G \subset G_e \cdot H$. Let $g \in G$. Since G/H is connected, there is a path in G/H joining $gH \in G/H$ to $eH \in G/H$. Since the map $G \to G/H$ is a bundle, we can lift this path to a path joining $g \in G$ to some element $h \in H$. Thus, $gh^{-1} \in G_e$ and $g = (gh^{-1})h \in G_e \cdot H$.

(ii) By (i) the map $j: G_e \to G/H$ is surjective. Suppose $g_1, g_2 \in G_e$. Then $g_2^{-1}g_1 \in G_e$ and

$$j(g_1) = j(g_2) \Leftrightarrow g_1 \in g_2H \Leftrightarrow g_2^{-1}g_1 \in H.$$

Thus, j induces a diffeomorphism $G_e/H_0 \to G/H$. ■

Exercise 3.6. Let (G, H) be a pair of Lie groups, where H is a closed subgroup of G. Let G_e be the identity component of G, and set $H_0 = H \cap G_e$.

(i) Show that the set of products $G_e \cdot H = \{gh \in G \mid g \in G_e,\ h \in H\}$ is a subgroup of G.

(ii) Show that the inclusion $G_e \subset G$ induces a smooth inclusion $G_e/H_0 \subset G/H$ whose image is a component of G/H. ❑

We remark that the bundle $H_0 \to G_e \to G_e/H_0$ is an example of a reduction of the bundle $H \to G \to G/H$.

Lie-Theoretic Properties of Klein Geometries

As motivation for the bundle definition of Cartan geometries given in Definition **5**.3.1, we draw the reader's attention to the following data associated to a Klein geometry (G, H):

(a) the smooth manifold $M = G/H$;

(b) the principal H bundle $H \subset G \to G/H$;

(c) the Maurer-Cartan form $\omega_G: T(G) \to \mathfrak{g}$ satisfying

  (i) ω_G is a linear isomorphism on each fiber,

  (ii) $R_h^*\omega_G = \mathrm{Ad}(h^{-1})\omega_G$ for all $h \in H$ (Proposition **3**.4.2, part c)[13],

  (iii) $\omega_G(X^\dagger) = X$ for all $X \in \mathfrak{h}$ (cf. Exercise **3**.2.15).

The general usage in modern literature is to use the term *homogeneous space* for the coset space G/H, where nothing is assumed about the Lie groups $H \subset G$, except that H is closed in G (cf. e.g., [M. Golubitski, 1972]). Older literature uses the term *Klein space.*

Geometrical Isomorphism and Mutation

Definition 3.7. Klein geometries (G_1, H_1) and (G_2, H_2) are called *geometrically isomorphic* if there is a Lie group isomorphism $\varphi: G_1 \to G_2$ such that $\varphi(H_1) = H_2$. ❀

In particular, the pairs (G, H) and (G, gHg^{-1}) are geometrically isomorphic. A useful generalization of isomorphism is the notion of *mutation.*

---

[13] Actually, this is true for all $h \in G$, but R_h arises from right H action on the principal bundle

Definition 3.8. Klein geometries (G_1, H_1) and (G_2, H_2) are *mutants* (of each other) if there is an isomorphism of Lie groups $\varphi: H_1 \to H_2$ such that the induced algebra isomorphism $\varphi_{*e}: \mathfrak{h}_1 \to \mathfrak{h}_2$ extends to a linear isomorphism $\lambda: \mathfrak{g}_1 \to \mathfrak{g}_2$, which is an H module map (i.e., a map satisfying $\lambda(\mathrm{Ad}(h)v) = \mathrm{Ad}(\varphi(h))(\lambda(v))$). The pair (φ, λ) is called a *mutation* from (G_1, H_1) to (G_2, H_2). ❀

Exercise 3.9. Show that $(Euc_n(\mathbf{R}), SO_n(\mathbf{R}))$, $(SO_{n+1}(\mathbf{R}), SO_n(\mathbf{R}))$, and $(L_{n,1}(\mathbf{R}), SO_n(\mathbf{R}))$ are all mutants of each other. (They are Euclidean space $\mathbf{R}^n$, elliptic space S^n, and hyperbolic space H^n, respectively.) ❏

Locally Klein Geometries

Here is a generalization of the notion of a Klein geometry.

Definition 3.10. Let (G, H) be a homogeneous space, and let $\Gamma \subset G$ be a discrete subgroup such that Γ acts effectively by left multiplication as a group of covering transformations on the space G/H with $\Gamma \setminus G/H$ connected. Then the triple (Γ, G, H) is called a *locally Klein geometry*. It is called *geometrically oriented* if G is connected. The double coset space $\Gamma \setminus G/H$ is called the *space* of the locally Klein geometry or, by abuse of notation, merely a *locally Klein geometry*. ❀

In the case that Γ is the identity, this reduces to the Klein geometry (G, H).

Note that in the definition of a locally Klein geometry (Γ, G, H) it is not assumed that (G, H) is a Klein geomtery, that is, it is not assumed that G/H is connected. A consequence of the following exercise is that we can always replace (Γ, G, H) by a locally Klein geometry (Γ_0, G_0, H) with the same space, and for which (G_0, H) is a Klein geometry.

Exercise 3.11. Suppose that (Γ, G, H) is a locally Klein geometry. Set $G_0 = G_e \cdot H$, where G_e is the identity component of G, and $\Gamma_0 = G_0 \cap \Gamma$. Show that G_0 is a Lie group, G_0/H is connected, and the inclusion $G_0 \subset G$ induces a diffeomorphism $\phi: \Gamma_0 \setminus G_0 \to \Gamma \setminus G$. ❏

Just as in the case of Klein geometries, locally Klein geometries also have an associated principal bundle. This fact depends on the following lemma.

Lemma 3.12. *Let Γ and H be Lie groups with free commuting left and right proper actions (respectively) on the smooth manifold X. Then*

$$(\Gamma \setminus X) \times H \to \Gamma \setminus X \text{ is proper} \Leftrightarrow \Gamma \times (X/H) \to X/H \text{ is proper.}$$

Proof. By symmetry it suffices to prove only the implication $\Rightarrow$.

Step 1. The action
$$\begin{array}{ccc}(\Gamma \times H) \times X \to & X \\ ((g,h),x) & \mapsto gxh^{-1}\end{array}$$
is proper.

Let $A, B \subset X$ be compact, and let $C = \{(g,h) \in \Gamma \times H \mid gAh \cap B \neq \emptyset\}$. We must show that C is compact. It suffices to show it is sequentially compact.

Let $(g_i, h_i) \in \Gamma \times H$, $i = 1, 2, \ldots$ be a sequence in C. Let A' and B' denote the images (which are compact) of A and B, respectively, in $\Gamma \setminus X$. Since $(\Gamma \setminus X) \times H \to \Gamma \setminus X$ is a proper action, the set $C' = \{h \in H \mid A'h \cap B' \neq \emptyset\}$ is compact. Moreover it is clear that $h_i \in C'$, $i = 1, 2, \ldots$. Hence $\{h_i\}$ has a convergent subsequence. Replacing $(g_i, h_i) \in C$, $i = 1, 2, \ldots$ by the corresponding subsequence, we may assume that

(i) $\lim h_i = h_\infty \in C'$.

Since $(g_i, h_i) \in C$ we may choose points $a_i \in A$ and $b_i \in B$ such that $g_i a_i h_i = b_i$, $i = 1, 2, \ldots$. Since A and B are compact $\{a_i\}$ and $\{b_i\}$ have convergent subsequences. Replacing $(g_i, h_i) \in C$, $i = 1, 2, \ldots$ by a corresponding subsequence, we obtain, in addition to (i),

(ii) $\lim a_i = a_\infty \in A$

(iii) $\lim b_i = b_\infty \in B$.

Now let $K \subset X$ be a compact neighborhood of $a_\infty h_\infty$ so that $a_i h_i \in K$, $i \geq N$. Since $\Gamma \times X \to X$ is a proper action, the set

$$C'' = \{g \in \Gamma \mid gK \cap \{b_\infty\} \neq \emptyset\} = \{g \in \Gamma \mid g^{-1} b_\infty \in K\}$$

is compact. Now $\lim g_i^{-1} b_\infty = \lim g_i^{-1} b_i = \lim a_i h_i = a_\infty h_\infty \in K$. Thus $g_i^{-1} b_\infty \in K$, $i \geq M$ and hence $g_i \in C''$, $i \geq M$. Thus $\{g_i\}$ has a convergent subsequence, and we may pass to this subsequence to obtain, in addition to (i), (ii), and (iii):

(iv) $\lim g_i = g_\infty \in \Gamma$

It follows that the original sequence $(g_i, h_i) \in \Gamma \times H$, $i = 1, 2, \ldots$ has a convergent subsequence, and hence C is sequentially compact.

Step 2. The action $\Gamma \times (X/H) \to X/H$ is proper.

Let $A, B \subset X/H$ be compact. We must show that $\{g \in \Gamma \mid gA \cap B \neq \emptyset\}$ is compact. If we write $A = \bigcup_{1 \leq i \leq a} A_i$, $B = \bigcup_{1 \leq j \leq b} B_j$, with A_i and B_j compact, then it suffices to show that $\{g \in \Gamma \mid gA_i \cap B_j \neq \emptyset\}$ is compact for all i, j. Thus it suffices to assume that A and B are small sets (i.e., subordinate to some open covering of X/H). Now, by Theorem 2.4, the map $X \to X/H$ is a principal H bundle, so there is an open covering of

X/H such that over each open set of this cover the bundle $X \to X/H$ is trivial. We assume that A and B are subordinate to this cover.

Applying a local section to A and B, we can obtain compact sets A' and B' in X with images A and B in X/H. Thus, $\{(g,h) \in \Gamma \times H \mid gA'h^{-1} \cap B' \neq \emptyset\}$ is compact. Therefore, the image of this set under the canonical projection $\Gamma \times H \to \Gamma$ is also compact; but this image is clearly $\{g \in \Gamma \mid gA \cap B \neq \emptyset\}$. ■

Corollary 3.13. *Let (Γ, G, H) be a locally Klein geometry. Then the map $\Gamma \setminus G \to \Gamma \setminus G/H$ is a principal H bundle.*

Proof. Apply Lemma 3.12, taking $X = G$, and use the left and right multiplication actions of Γ and H, respectively, on G. Since Γ and H are closed subgroups, these actions are proper. Since (Γ, G, H) is locally Klein, it follows that the action $\Gamma \times (G/H) \to G/H$ is free and proper. So by the lemma, $(\Gamma \setminus G) \times H \to \Gamma \setminus G$ is also free and proper. The result then follows from Theorem 2.4. ■

We remark that, according to Exercise 3.11, the principal bundle $\Gamma \setminus G \to \Gamma \setminus G/H$ is the same as the principal bundle $\Gamma_0 \setminus G_0 \to \Gamma_0 \setminus G_0/H$.

Example 3.14. Consider the case of the hyperbolic plane $M = Sl_2(\mathbf{R})/SO_2(\mathbf{R})$. It is known that any closed, orientable surface M_g of genus $g > 1$ can appear as the quotient $\Gamma \setminus M$ for some choice of $\Gamma \subset Sl_2(\mathbf{R})$. Let Γ_g be such a subgroup. Then the quotient $M_g = \Gamma_g \setminus M$ is a locally Klein geometry in the sense of the definition above. ◆

We note that if (Γ, G, H) is a locally Klein geometry, then the Maurer–Cartan form $\omega_G : T(G) \to \mathfrak{g}$, because of its left invariance, induces a form $\omega_{\Gamma \setminus G} : T(\Gamma \setminus G) \to \mathfrak{g}$.

Exercise 3.15. Show that the form $\omega_{\Gamma\setminus G}$ has properties analogous to those of the Maurer–Cartan form given on page 153. ❑

Lie Algebras of a Klein Geometry

For each Klein geometry (G, H), we have the corresponding pair of Lie algebras $(\mathfrak{g}, \mathfrak{h})$. If (G, H) is effective, then $\mathfrak{h}$ contains no nontrivial ideal of $\mathfrak{g}$ (cf. Exercise **3**.4.7).

Definition 3.16. An *infinitesimal Klein geometry*, or more briefly, a *Klein pair*, is a pair of Lie algebras $(\mathfrak{g}, \mathfrak{h})$ where $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$. The kernel $\mathfrak{k}$ of $(\mathfrak{g}, \mathfrak{h})$ is the largest ideal of $\mathfrak{g}$ contained in $\mathfrak{h}$. If $\mathfrak{k} = 0$, we say the pair $(\mathfrak{g}, \mathfrak{h})$ is *effective*. If there is an $\mathfrak{h}$-module decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$, we say $(\mathfrak{g}, \mathfrak{h})$ is *reductive*. ❀

Let us consider the correspondence

$$\begin{array}{ccc}\{\text{Klein geometries}\} & \to & \{\text{Klein pairs}\}. \\ (G,H) & \mapsto & (\mathfrak{g},\mathfrak{h})\end{array}$$

This correspondence is not surjective, since Remark **3**.8.10 gives an example of an effective Klein pair $(\mathfrak{g}, \mathfrak{h})$ for which $\mathfrak{h}$ cannot be realized as a closed subgroup in any realization of $\mathfrak{g}$. We could define the closure of $\bar{\mathfrak{h}}$ of $\mathfrak{h}$ to be the Lie algebra of the closure of the realization of $\mathfrak{h}$ in the connected and simply connected realization of $\mathfrak{g}$. Then it is clear that the image of the correspondence consists of the Klein pairs $(\mathfrak{g}, \mathfrak{h})$ where $\mathfrak{h}$ is closed (i.e., $\bar{\mathfrak{h}} = \mathfrak{h}$). It is easy to see that the Lie algebra $\mathfrak{k}$ of the kernel K of a Klein geometry (G, H) is the kernel of the associated Klein pair $(\mathfrak{g}, \mathfrak{h})$.

Exercise 3.17. Let $(\mathfrak{g}, \mathfrak{h})$ be an effective Klein pair with $\mathfrak{h}$ closed.

(a) Show that every realization is locally effective.

(b) Show that there is an effective realization.

(c) Study the relationship among the collection of all realizations of $(\mathfrak{g}, \mathfrak{h})$. ❑

One-Dimensional Effective Geometries

It is a remarkable fact that the real line supports exactly *three* effective Klein geometries. These are as follows.

Definition 3.18. (i) *The Euclidean line* is $(\mathbf{R}, 0)$ with Klein pair $(\mathfrak{g}, \mathfrak{h}) = (\mathbf{R}, 0)$.

(ii) *The affine line* is $(Aff_1^+(\mathbf{R}), Gl_1^+(\mathbf{R}))$, with Klein pair $(\mathfrak{g}, \mathfrak{h}) = (\mathfrak{aff}_1(\mathbf{R}), \mathfrak{gl}_1(\mathbf{R}))$, where

$$Aff_1^+(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix} \middle| c \in \mathbf{R}, d \in \mathbf{R}^+ \right\}, \; Gl_1^+(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix} \middle| d \in \mathbf{R}^+ \right\},$$

$$\mathfrak{aff}_1^+(\mathbf{R}) = \left\{ \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix} \middle| c \in \mathbf{R}, d \in \mathbf{R}^+ \right\}, \; \mathfrak{gl}_1^+(\mathbf{R}) = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} \middle| d \in \mathbf{R}^+ \right\}.$$

(iii) *The universal cover of the projective line* is $(\tilde{G}, \tilde{H})$ with Klein pair $(\mathfrak{g}, \mathfrak{h})$, where

$\tilde{G} = \tilde{Sl}_2(\mathbf{R}) =$ the universal cover of $G = Sl_2(\mathbf{R})$,
$\tilde{H} =$ the identity component of the inverse image in $\tilde{G}$ of $H \subset G$,

and where

$$H = \left\{ \begin{pmatrix} d^{-1} & b \\ 0 & d \end{pmatrix} \middle| d \in \mathbf{R}^*, b \in \mathbf{R} \right\}.$$

The Lie algebras of G and H (and of $\tilde{G}$ and $\tilde{H}$) are

$$\mathfrak{g} = \left\{ \begin{pmatrix} -d & b \\ c & d \end{pmatrix} \middle| b, c, d \in \mathbf{R} \right\}, \; \mathfrak{h} = \left\{ \begin{pmatrix} -d & b \\ 0 & d \end{pmatrix} \middle| b, d \in \mathbf{R} \right\}. \qquad ❀$$

The proof of this depends on the following classification of Klein pairs.

Proposition 3.19. *Let $(\mathfrak{g}, \mathfrak{h})$ be an effective Klein pair with dim $\mathfrak{g}/\mathfrak{h} =$ 1. Then $(\mathfrak{g}, \mathfrak{h})$ is isomorphic to one of the three Klein pairs described in Definition* 3.18.

■

Proof. See Appendix C.

We use this proposition to prove the following slightly more general result than that stated at the start of this subsection.

Proposition 3.20.

(i) *There are exactly three isomorphism classes of effective Klein geometries on the real line: the Euclidean line, the affine line, and the universal cover of the real projective line.*

(ii) *There are two families of isomorphism classes of effective Klein geometries on the circle.*

(a) *One is the family of Euclidean circles $(\mathbf{R}/(l\mathbf{Z}), 0)$ with the same Klein pair as the Euclidean line. The parameter $l \in \mathbf{R}^+$ is the length of the circle.*

(b) *The other is the family of finite covers of the projective line $(PSl_2(\mathbf{R})^{(n)}, H^{(n)})$ with the same Klein pair as the projective line. The parameter $n \in \mathbf{Z}^+$ (the degree) indexes the subgroup*

$$n\mathbf{Z} \cong n\pi_1(PSl_2(\mathbf{R})) \subset \pi_1(PSl_2(\mathbf{R})) \cong \mathbf{Z}$$

determining the covering $PSl_2(\mathbf{R})^{(n)} \to PSl_2(\mathbf{R})$, and $H^{(n)}$ is the component of the identity of the preimage of H in $PSl_2(\mathbf{R})^{(n)}$.

Proof. Suppose (G, H) is a one-dimensional effective Klein geometry with Klein pair $(\mathfrak{g}, \mathfrak{h})$. Then by Proposition 3.19, $(\mathfrak{g}, \mathfrak{h})$ must be the Euclidean, affine, or projective Klein pair.

First assume that $(\mathfrak{g}, \mathfrak{h}) = (\mathbf{R}, 0)$. Then the universal cover of G is $\mathbf{R}$; in particular, G is abelian, and so H is a normal subgroup. Since the geometry is effective, $H = 0$. The preimage of H in the universal cover is a discrete subgroup that must then take the form $l\mathbf{Z} \subset \mathbf{R}$. Thus $G = \mathbf{R}/(l\mathbf{Z})$. This is the real line if $l = 0$ and is the circle otherwise.

Next assume that $(\mathfrak{g}, \mathfrak{h})$ is the affine Klein pair. Then $\mathfrak{g}$ is realized by the connected and simply connected group $Aff_1^+(\mathbf{R})$. As a simple calculation shows, $Aff_1^+(\mathbf{R})$ has no center and hence $Aff_1^+(\mathbf{R})$ is the unique realization of $\mathfrak{g}$. Now $\mathfrak{h} \subset \mathfrak{g}$ has $Gl_1^+(\mathbf{R})$ as its unique connected realization in $Aff_1^+(\mathbf{R})$. Another simple calculation shows that $Gl_1^+(\mathbf{R})$ is a maximal subgroup in $Aff_1^+(\mathbf{R})$ and hence $H = Gl_1^+(\mathbf{R})$. Thus $(G, H) = (Aff_1^+(\mathbf{R}), Gl_1^+(\mathbf{R}))$. In particular, $G/H = \mathbf{R}$. Thus, there are no affine Klein geometries on the circle and just one on the line.

Finally, assume that $(\mathfrak{g}, \mathfrak{h})$ is the projective Klein pair. Then $\mathfrak{g}$ is realized by the connected and simply connected universal cover of $Sl_2(\mathbf{R})$ denoted by $\tilde{S}l_2(\mathbf{R})$. In fact, the projection $Sl_2(\mathbf{R}) \to Sl_2(\mathbf{R})/\{\pm I\} = PSl_2(\mathbf{R})$ is a 2-fold covering so that the composite $\tilde{S}l_2(\mathbf{R}) \to Sl_2(\mathbf{R}) \to PSl_2(\mathbf{R})$ displays $\tilde{S}l_2(\mathbf{R})$ as the universal cover of $PSl_2(\mathbf{R})$. Now the kernel of the covering homomorphism $\tilde{S}l_2(\mathbf{R}) \to PSl_2(\mathbf{R}))$ is $\pi_1(PSl_2(\mathbf{R})) = \mathbf{Z}$ and is a central subgroup. On the other hand, $PSl_2(\mathbf{R})$ has no center, and so $\mathbf{Z}$ is the whole center of $\tilde{S}l_2(\mathbf{R})$. Now if G is an arbitrary connected realization of $\mathfrak{g}$, there is a covering homomorphism $p: G \to PSl_2(\mathbf{R})$ determined by the subgroup $p_*\pi_1(G) = n\mathbf{Z} \subset \pi_1(PSl_2(\mathbf{R})) = \mathbf{Z}$, that is, by the integer n. Thus, we may write $G = PSl_2(\mathbf{R})^{(n)}$. Since the pair (G, H) is effective, $H \cap \text{center}(G) = \{e\}$. Thus, H is isomorphic to its image H_0 in $PSl_2(\mathbf{R})$, and hence H_0 contains the connected subgroup of $PSl_2(\mathbf{R})$ with Lie algebra $\mathfrak{h}$. This latter subgroup is isomorphic to the image of

$$\left\{ \begin{pmatrix} d^{-1} & b \\ 0 & d \end{pmatrix} \middle| b \in \mathbf{R},\ d \in \mathbf{R}^+ \right\} \subset Sl_2(\mathbf{R}).$$

under the projection map $Sl_2(\mathbf{R}) \to PSl_2(\mathbf{R})$.

However, a third simple calculation shows that this subgroup is a maximal subgroup of $Sl_2(\mathbf{R})$. Thus, it is isomorphic under the projection map to $H_0 \subset PSl_2(\mathbf{R})$. In fact, since H_0 is connected and simply connected, there is a canonical homomorphism $H_0 \subset PSl_2(\mathbf{R})^{(n)}$ lifting the inclusion for each n and hence we have the following diagram.

$$\begin{array}{ccccc}
H_0 & \subset & \widetilde{PSl_2}(\mathbf{R}) & \to & \mathbf{R} \\
\approx\downarrow & & \downarrow & & \downarrow \\
H_0 & \subset & PSl_2(\mathbf{R})^{(n)} & \to & S^1 \\
\approx\downarrow & & \downarrow & & \downarrow \text{ n-fold cover} \\
H_0 & \subset & PSl_2(\mathbf{R}) & \to & S^1
\end{array}$$

From this it follows that there is just one isomorphism class of projective geometries on the real line, while for the circle there is one for each positive integer. ■

Exercise 3.21. Verify the three "simple calculations" referred to in the previous proof. ❏

Exercise 3.22. (a) Show that the Euclidean circles (Definition 3.18(ii), part a) are isomorphic as Klein geometries if and only if they have the same length.

(b) Show that two projective circles (Definition 3.18(ii), part b) are isomorphic as Klein geometries if and only if they have the same degree n. ❑

Exercise 3.23. Show that the geometries in Definition 3.18(i) and (ii) are reductive but that in 3.18(iii) is not reductive. ❑

§4. A Fundamental Property

In this section we prove a basic technical result about the connection between a Klein geometry and its associated infinitesimal geometry.

Theorem 4.1. *Let (G, H) be a Klein geometry with kernel K and associated Klein pair $(\mathfrak{g}, \mathfrak{h})$ with kernel $\mathfrak{k}$. Then $K = \{h \in H \mid Ad(h)v - v \in \mathfrak{k}$ for all $v \in \mathfrak{g}\}$.*

Corollary 4.2 (Fundamental property of effective Klein geometries). *If (G, H) is an effective Klein geometry, N is a subgroup of H, and $\mathfrak{n}$ is the Lie algebra of N, then*

$$N = \{h \in H \mid Ad(h)v - v \in \mathfrak{n} \text{ for all } v \in \mathfrak{g}\} \Rightarrow N = \{e\}.$$

Our proof of this result and its corollary depend on the following two lemmas.

Lemma 4.3. *Let $\mathfrak{n} \subset \mathfrak{h} \subset \mathfrak{g}$ be Lie algebras and $N \subset H$ Lie groups realizing the inclusion $\mathfrak{n} \subset \mathfrak{h}$. Assume N is normal in H, and let*

$$N' = \{h \in H \mid Ad(h)v - v \in \mathfrak{n} \quad \text{for all } v \in \mathfrak{g}\}.$$

Then N' is also a normal subgroup of H.

Proof. (i) Clearly $e \in N'$.

(ii) Next, $\alpha \in N' \Rightarrow \mathrm{Ad}(\alpha)v - v \in \mathfrak{n}$ for all $v \in \mathfrak{g} \Rightarrow v - \mathrm{Ad}(\alpha^{-1})v \in \mathrm{Ad}(\alpha^{-1})\mathfrak{n}$, for all $v \in \mathfrak{g}$; but N is normal in H. Thus $\mathrm{Ad}(\alpha^{-1})\mathfrak{n} \subset \mathfrak{n}$. Hence $\alpha^{-1} \in N'$.

(iii)

$$\begin{aligned}\alpha, \beta \in N' \Rightarrow \mathrm{Ad}(\alpha\beta)v - v &= \mathrm{Ad}(\alpha)(\mathrm{Ad}(\beta)v - v) + (\mathrm{Ad}(\alpha)v - v)\\ &\in \mathrm{Ad}(\alpha)\mathfrak{n} + \mathfrak{n} = \mathfrak{n} \text{ for all } v \in \mathfrak{g}.\end{aligned}$$

Hence $\alpha\beta \in N'$. Thus N' is a group. Finally, we verify it is normal in H.

(iv)

$$\begin{aligned}\alpha \in N', h \in H \Rightarrow \mathrm{Ad}(h^{-1}\alpha h)v - v &= \mathrm{Ad}(h^{-1})\{\mathrm{Ad}(\alpha)(\mathrm{Ad}(h)v) - (\mathrm{Ad}(h)v)\}\\ &\in \mathrm{Ad}(h^{-1})\mathfrak{n} = \mathfrak{n} \text{ for all } v \in \mathfrak{g}.\end{aligned}$$ ■

Lemma 4.4. *Let $\mathfrak{h} \subset \mathfrak{g}$ be an embedding of Lie algebras and H realize $\mathfrak{h}$. Define a sequence of subgroups of H inductively by*

$$\begin{aligned}
N_0 &= H,\\
N_1 &= \{h \in H \mid Ad(h)v - v \in \mathfrak{n}_0 \text{ for all } v \in \mathfrak{g}\},\\
&\qquad \text{where } \mathfrak{n}_0 = \text{Lie algebra of } N_0 = \mathfrak{h},\\
N_2 &= \{h \in H \mid Ad(h)v - v \in \mathfrak{n}_1 \text{ for all } v \in \mathfrak{g}\},\\
&\qquad \text{where } \mathfrak{n}_1 = \text{Lie algebra of } N_1,\\
&\dots\\
N_k &= \{h \in H \mid Ad(h)v - v \in \mathfrak{n}_{k-1} \text{ for all } v \in \mathfrak{g}\},\\
&\qquad \text{where } \mathfrak{n}_{k-1} = \text{Lie algebra of } N_{k-1}.
\end{aligned}$$

Then $N_0 \supset N_1 \supset N_2 \supset \dots \supset N_k \supset \dots$ are all Lie groups that are closed and normal in H and, after finitely many steps, the sequence stabilizes at a group N_∞ whose Lie algebra $\mathfrak{n}_\infty$ is an ideal in $\mathfrak{g}$ and satisfies

$$N_\infty = \{h \in H \mid Ad(h)v - v \in \mathfrak{n}_\infty \text{ for all } v \in \mathfrak{g}\}.$$

Proof. Applying Lemma 4.3 inductively, we see that each of the groups N_k is normal in H. Also, if we assume $N_j \supset N_{j+1}$ (which holds for $j = 0$), then $\mathfrak{n}_j \supset \mathfrak{n}_{j+1}$, so that

$$\begin{aligned}n \in N_{j+2} &\Rightarrow \mathrm{Ad}(n)v - v \in \mathfrak{n}_{j+1} \text{ for all } v \in \mathfrak{g}\\ &\Rightarrow \mathrm{Ad}(n)v - v \in \mathfrak{n}_j \text{ for all } v \in \mathfrak{g}\\ &\Rightarrow n \in N_{j+1}, \text{ and hence } N_{j+1} \supset N_{j+2}.\end{aligned}$$

Now set $N_\infty = \cap N_j$. Since each is N_j normal in H, so is N_∞. We note that the sequence of Lie algebras $\mathfrak{n}_j$ and hence also the identity components of the N_js form a decreasing sequence which must become constant after finitely many steps. It follows from their definition that the N_js themselves must also stabilize in the same finite number of steps, and so N_∞ is again a Lie group. Clearly

$$N_\infty = \{h \in H \mid \mathrm{Ad}(h)v - v \in \mathfrak{n}_\infty \text{ for all } v \in \mathfrak{g}\}.$$ ■

Corollary 4.5. *Let H be a closed subgroup of G realizing the inclusion $\mathfrak{h} \subset \mathfrak{g}$. Assume that G is connected. Then in addition to the result of Lemma 4.4, N_∞ is normal in G.*

Proof. We continue from the end of the proof of Lemma 4.4. Writing the condition

$$N_\infty = \{h \in H \mid \mathrm{Ad}(h)v - v \in N_\infty \text{ for all } v \in \mathfrak{g}\}$$

infinitesimally, we see the Lie algebra $\mathfrak{n}_\infty$ of N_∞ satisfies $\{\mathfrak{n}_\infty, \mathfrak{g}\} \subset \mathfrak{g}$, so that $\mathfrak{n}_\infty$ is an ideal of $\mathfrak{g}$. By Exercise **3**.4.7 the identity component of N_∞ is a normal subgroup of G. For $n \in N_\infty$, we have

$$\mathrm{Ad}(n)(\mathrm{Ad}(g)v) - \mathrm{Ad}(g)v \in \mathfrak{n}_\infty,$$

so that

$$\mathrm{Ad}(g^{-1}ng)v - v = \mathrm{Ad}(g^{-1})\{\mathrm{Ad}(n)(\mathrm{Ad}(g)v) - \mathrm{Ad}(g)v\} \in \mathrm{Ad}(g^{-1})\mathfrak{n}_\infty = \mathfrak{n}_\infty.$$

(The latter equality comes from the fact that the identity component of N_∞ is already known to be normal in G.) Thus, $g^{-1}ng \in N_\infty$. It follows that N_∞ itself is normal in G. ■

Proof of Theorem 4.1. It suffices to show that $K = N_\infty$. The inclusion $K \supset N_\infty$ follows from the maximality of K (cf. Definition 3.2). We need to show that $K \subset N_\infty$. The fact that $K \subset N_0 = H$ is automatic. Let us assume inductively that $K \subset N_j$. Thus $\mathfrak{k} \subset \mathfrak{n}_j$. Now the fact that K is normal in G implies, by Exercise **3**.4.7, that $\mathrm{ad}(k)v - v \in \mathfrak{k}$ for all $k \in K$ and for all $v \in \mathfrak{g}$. Thus,

$$\begin{aligned} K &\subset \{h \in H \mid \mathrm{ad}(h)v - v \in \mathfrak{k} \text{ for all } v \in \mathfrak{g}\} \\ &\subset \{h \in H \mid \mathrm{ad}(h)v - v \in \mathfrak{n}_j \text{ for all } v \in \mathfrak{g}\} = N_{j+1}. \end{aligned}$$

It follows by induction that $K \subset N_\infty$, and hence $K = N_\infty$. ■

Proof of Corollary 4.2. The second half of the preceding argument shows that any subgroup N satisfying the condition $N \subset \{h \in H \mid \mathrm{ad}(h)v - v \in \mathfrak{n}$ for all $v \in \mathfrak{g}\}$ lies in N_∞. But $N_\infty = 1$, since it is a normal subgroup of G lying in H. ■

§5. The Tangent Bundle of a Klein Geometry

Let us consider a bundle chart (U, ψ) for the principal H bundle $\pi: G \to G/H$. Thus we have the diffeomorphism $\psi: U \times H \to \pi^{-1}(U) \subset G$. Clearly, this yields a commutative diagram.

$$\begin{tikzcd} T(\pi^{-1}(U)) \arrow[dr, "\pi_*"'] & T(U \times H) \arrow[l, "\psi_*"', "\approx"] \arrow[r, "\pi_{U*} \times \pi_{H*}", "\approx"'] \arrow[d, "\pi_{U*}"] & T(U) \times T(H) \arrow[dl, "\mathrm{proj}_1"] \\ & T(U) & \end{tikzcd}$$

The top right-hand diffeomorphism is the one described in Exercise **1**.4.15. From this diagram it follows that, if $x = \pi(g)$, the diagram

$$\begin{tikzcd} T_g(gH) \arrow[r, "\omega_H", "\approx"'] \arrow[d] & \mathfrak{h} \arrow[d, phantom, "\cap"] \\ T_g(G) \arrow[r, "\omega_G", "\approx"'] \arrow[d] & \mathfrak{g} \arrow[d] \\ T_x(G/H) \arrow[r, "\varphi_g", "\approx"'] & \mathfrak{g}/\mathfrak{h} \end{tikzcd}$$

is commutative,[14] with exact columns. Since $\mathfrak{h}$ is only a subalgebra of $\mathfrak{g}$ (not an ideal), the quotient space $\mathfrak{g}/\mathfrak{h}$ is merely a vector space (not a Lie algebra). The linear isomorphism φ_g is the unique map making the diagram commute.

We can identify the tangent space of G/H at x with $\mathfrak{g}/\mathfrak{h}$, but the identification φ_g depends on the choice of $g \in G$ over x. In fact, since $\pi R_h = \pi$, the relation $R_h^*\omega_H = \mathrm{Ad}(h)^{-1}\omega_H$ implies that $\varphi_{gh} = \mathrm{Ad}(h)^{-1}\varphi_g$. It follows that the identification of $T_x(G/H)$ with $\mathfrak{g}/\mathfrak{h}$ is determined only up to the adjoint action of H on $\mathfrak{g}/\mathfrak{h}$. This fact accounts for the frequent occurrence of the adjoint action in the sequel.

Proposition 5.1. $T(G/H) \approx G \times_H \mathfrak{g}/\mathfrak{h}$ *(as vector bundles over* G/H*).*

Proof. Define a map $\varphi: G \times \mathfrak{g} \to T(G/H)$ by $\varphi(g \times v) = (\pi(g), \pi_* L_{g*}v)$, where $\pi: G \to G/H$ is the canonical projection. Clearly, φ is smooth, surjective, and linear for fixed g. Moreover, for $v \in T_e(H)$, $L_{g*}v \in T_g(gH) \in \ker \pi_*$. Thus, $\varphi(g \times v) = (g, 0)$. There is also another fact about φ. We have

$$\begin{aligned} \varphi(gh, \mathrm{Ad}(h^{-1})v) &= (\pi(gh), \pi_* L_{gh*}\mathrm{Ad}(h^{-1})v) = (\pi(g), \pi_*(L_{gh*}L_{h^{-1}*}R_{h*}v)) \\ &= (\pi(g), L_{g*}\pi_*(R_{h*}v)) = (\pi(g), L_{g*}(\pi R_h)_* v) \\ &= (\pi(g), L_{g*}(\pi_*(v)) = \varphi(g, v). \end{aligned}$$

Thus, φ induces a smooth quotient mapping $\bar{\varphi}: G \times_H \mathfrak{g}/\mathfrak{h} \to T(G/H)$. Moreover, this map is injective since $\varphi(g \times v) = \varphi(g' \times v')$ implies $g' = gh$ for some $h \in H$ and $\pi_* L_{gh*}v' = \pi_* L_{g*}v$. The latter means that $\pi_* L_{h*}v' = \pi_* v$ and hence that $v' - \mathrm{Ad}(h^{-1})v \in \mathfrak{h}$. Thus,

[14] Of course, the top map is not really ω_H except in the case when $g \in H$. However, for any choice of $gh \in gH$, the composite $T_g(gH) \xrightarrow{L_{(gh)^{-1}*}} T_h(hH) = T_h(H) \xrightarrow{\omega_H} \mathfrak{h}$ is the same, so we may sensibly denote this map by ω_H.

$$g' \times v' = gh \times \mathrm{Ad}(h^{-1})v \,(\mathrm{mod}\ \mathfrak{h}) = g \times v \quad \text{in } G \times_H \mathfrak{g}/\mathfrak{h}.$$

Thus, $\bar{\varphi}$ is a vector bundle map covering the identity on the base G/H, which is linear on fibers. ■

This proposition gives an identification of each tangent space $T_g(G/H)$ with $\mathfrak{g}/\mathfrak{h}$. This identification is canonical up to the adjoint action of H on $\mathfrak{g}/\mathfrak{h}$. Since tangential information involves only derivatives of the first order, we can make the following rough division of Klein geometries into two types according to whether or not H is faithfully represented by the adjoint action on the tangent space.

Definition 5.2. A Klein geometry G/H is of *first order* if the representation $\mathrm{Ad}_{\mathfrak{g}/\mathfrak{h}}: H \to Gl(\mathfrak{g}/\mathfrak{h})$ is faithful (i.e., injective). Otherwise, the geometry is said to have *higher order.* ❀

As one may see in Appendix C, the classification of primitive effective pairs $(\mathfrak{g}, \mathfrak{h})$ breaks into two cases of quite different nature depending on whether or not $ad_{\mathfrak{g}/\mathfrak{h}}: \mathfrak{h} \to Gl(\mathfrak{g}/\mathfrak{h})$ is injective.

§6. The Meteor Tracking Problem

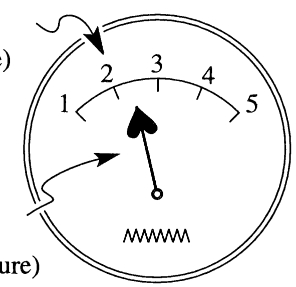
In order to motivate the need for the notion of a *gauge*, we begin by describing the *meteor tracking problem* in a Klein geometry $M = G/H$.

Suppose we live in a small open set $U \subset M$ and a meteor flashes through U. We wish to describe the motion of the meteor. We assume the meteor is *rigid* and *sufficiently complicated.* By the term *rigid* in the given geometry, we mean that for any two of its positions, where it has the configurations[15] X_0 and X_1 say, there is an element of G carrying X_0 to X_1. By *sufficiently complicated* we mean that the subgroup of G that fixes the body pointwise (its stabilizer) is the identity. It follows that if $X(t)$ is the configuration of its points at time t, then there is a unique path $g(t)$ in G, the group of "rigid motions" for X, such that $X(t) = g(t)X(0)$. One way to describe the motion would be to specify the path $g(t)$ itself. However, it is more useful for us, and closer to what is actually observed, to describe the motion differently. We first describe the motion of *one* of its points $q \in X(0)$, which we take to be eH, by a path $q(t) = g(t)H$ on U, and then describe the motion of the rest of the body as turning about this one point as it moves along. To describe this turning, we need the notion of a gauge.

[15]If the "abstract" meteor is the point set X, then a configuration X_0 of these points in M is a map $X_0: X \to M$. This is what Cartan originally had in mind when using the term "moving frame."

What Is a Gauge?

The following figure depicts the anatomy of a generic gauge.[16]



We see that this structure consists of two parts:

(1) a part that is fixed but arbitrary, the numbered marks (the "choice of gauge")

(2) a part that is moving and that we wish to measure by comparison to the fixed part. In our example of the meteor tracking problem, the moving part is the meteor itself. What is the fixed part?

Let us assume that the open set U is small enough so that there is a smooth section $\sigma: U \to G$. Of course, once there is one σ there will be many others also, and if σ_1 and σ_2 are two such sections, then they will differ by a map $h: U \to H$, namely,

$$\sigma_2(u) = \sigma_1(u)h(u).$$

The "choice of gauge" is merely a choice of one of these sections. The section is the gauge and the relation between σ_1 and σ_2 above is called a gauge transformation. A choice of gauge can be regarded as a choice of motions, varying smoothly with u, which maps the base point to u. The value of a gauge at a point may be called a *frame* at that point and the gauge itself may be called a *moving frame.* From this point of view the principal bundle $G \to G/H$ is a bundle of frames.

Now let us see how we may use the gauge $\sigma: U \to G$ to track the meteor. Since both $\sigma(q(t))$ and $g(t)$ lie over $q(t)$, they must differ by an element $k(t) \in H$. Thus, in the presence of a gauge, the motion may be described by giving $q(t)$ and $k(t)$. The latter describes the way the meteor turns

[16]Hermann Weyl's original use of the term *gauge* was in the restricted sense of the gauge of a railway track and refers to a scale factor. This appears as the special case $H = \mathbf{R}^+$. It is a happy accident of language that the term *gauge* also has the more generic interpretation given here.

about the point $q(t)$. We emphasize that there is no intrinsic geometric meaning attached to the choice of gauge, or indeed to the function $k(t)$ alone. Together, however, they complete the description of the motion of the meteor.

§7. The Gauge View of Klein Geometries

In this section we are going to formulate various aspects of the Klein geometry $M = G/H$ from the gauge point of view. These remarks will serve as the motivation for the base definition of Cartan geometries given in §1 of the next chapter, which studies Cartan's generalization of Klein geometries.

Gauge Picture of Bundle Charts

Fix a bundle coordinate chart (U, ψ) for the principal H bundle $\pi\colon G \to G/H$. Thus we have the diffeomorphism $\psi\colon U \times H \to \pi^{-1}(U)$. Note that, since such a chart is by definition right H equivariant, specifying ψ is equivalent to specifying a section σ over U. More precisely, setting $\sigma(u) = \psi(u, e)$, we have $\psi(u, h) = \sigma(u)h$.

Let us study the coordinate change resulting from a change of sections over U. Suppose that σ_1 and σ_2 are two sections of $G \to G/H$ over U. Thus, there is a unique smooth map $k\colon U \to H$ such that $\sigma_2(u) = \sigma_1(u)k(u)$ for $u \in U$. If ψ_1 and ψ_2 are the corresponding trivializations of $\pi^{-1}(U)$, we have the diagram

$$\begin{array}{ccccc} U \times H & \xrightarrow{\psi_1} & \pi^{-1}(U) & \xleftarrow{\psi_2} & U \times H. \\ (u,h) & \mapsto & \sigma_1(u)h = \sigma_2(u)\hat{h} & \leftarrow & (u,\hat{h}) \end{array}$$

It follows that $\psi_2^{-1}\psi_1(u, h) = (u, \sigma_2(u)^{-1}\sigma_1(u)h) = (u, k(u)^{-1}h)$. Thus, it is possible to reconstruct the bundle $G \to G/H$ if we are given only the transition functions (or gauge transformations) $k\colon U \to H$.

Gauge Picture of Maurer–Cartan Form

Let us consider now the shape taken by the Maurer–Cartan form on a coordinate chart given by a section σ of $G \to G/H$ over an open set U. The coordinate chart corresponding to σ is

$$\begin{array}{ccc} \psi : U \times H & \to & \pi^{-1}(U). \\ (u,h) & \mapsto & R_h\sigma(u) \end{array}$$

We calculate $\psi^*(\omega_G)$.

Proposition 7.1. *The following diagram commutes, where*

$$\varphi(v, y) = Ad(h)^{-1}\omega_G(v) + \omega_H(y).$$

$$\begin{tikzcd} T_{(u,h)}(U \times H) \arrow[d, "\pi_{U*} \times \pi_{H*}"'] \arrow[dr, "\psi^*\omega_G"] & \\ T_u(U) \times T_h(H) \arrow[r, "\varphi"'] & \mathfrak{g} \end{tikzcd}$$

Proof. By Exercise **1**.4.15, the vertical map is an isomorphism with inverse

$$\begin{array}{ccc} T_u(U) \times T_h(H) & \to & T_{(u,h)}(U \times H), \\ (v,y) & \mapsto & i_{h*}(v) + j_{u*}(y) \end{array}$$

where $i_h\colon U \to U \times H$ sends $x \to (x, h)$ and $j_u\colon H \to U \times H$ sends $k \to (u, k)$. Thus it suffices to verify the commutativity of

$$\begin{tikzcd} T_{(u,h)}(U \times H) \arrow[dr, "\psi^*\omega_G"] & \\ T_u(U) \arrow[u, "i_{h*}"] \arrow[r, "\varphi|T_u(U) = \mathrm{Ad}(h^{-1})\omega_G"'] & \mathfrak{g} \end{tikzcd} \qquad \begin{tikzcd} T_{(u,h)}(U \times H) \arrow[dr, "\psi^*\omega_G"] & \\ T_h(H) \arrow[u, "j_{u*}"] \arrow[r, "\varphi|T_h(H) = \omega_H = \omega_G|T_h(H)"'] & \mathfrak{g} \end{tikzcd}$$

For the first diagram, we have

$$\begin{aligned} (\psi^*\omega_G) \circ i_{h*} &= (\psi i_h)^*\omega_G = (R_h\sigma)^*\omega_G = \sigma^* R_h^*\omega_G \\ &= \sigma^*(\mathrm{Ad}(h^{-1})\omega_G) = \mathrm{Ad}(h^{-1})\sigma^*\omega_G, \end{aligned}$$

and for the second we have

$$(\psi^*\omega_G) \circ j_{u*} = (\psi j_u)^*\omega_G = (L_{\sigma(u)})^*\omega_G = \omega_G. \qquad \blacksquare$$

Infinitesimal Gauge

Let $\sigma\colon U \to G$ be a gauge, namely, a section of the principal bundle $G \to G/H$ over an open set U in G/H. Then σ pulls back the Maurer–Cartan form ω_G to a $\mathfrak{g}$-valued 1-form θ on U:

$$\sigma^*(\omega_G) = \theta.$$

The structural equation for ω_G also pulls back to yield

$$d\theta + \frac{1}{2}[\theta, \theta] = 0.$$

We note that not only does σ determine θ but the reverse is *almost* true as well, since the fundamental theorem of calculus says that θ determines σ up to left multiplication by a fixed element of G. Indeed, for an effective Klein geometry, if σ is a section, then $g\sigma$ is a section $\Leftrightarrow g = e$.[17] For this reason we may refer to both σ and θ as gauges on M. If we wish to distinguish them, we shall refer to θ as an *infinitesimal* gauge. In fact, it will be the infinitesimal form of the gauge that will be of the most use to us exactly because it is "independent of the base point." We may further note that since σ is a section, by the diagrams in §5 identifying (noncanonically) the tangent space of G/H with $\mathfrak{g}/\mathfrak{h}$, the composite mapping

$$\bar{\theta}: T_u(U) \xrightarrow{\ \theta\ } \mathfrak{g} \xrightarrow[\ ]{\text{canonical projection}} \mathfrak{g}/\mathfrak{h}$$

is an isomorphism for every $u \in U$.

The infinitesimal gauge θ can be roughly interpreted as assigning to each tangent vector $v \in T_u(M)$ the infinitesimal motion $I + \varepsilon\theta(v) \in G$ (where ε is infinitesimal and we are thinking of G as a matrix group so that the addition makes sense) of M whose effect on u itself is to move it to $u + \varepsilon v$.

By varying the trivialization ψ, or equivalently by varying the section σ, we change the infinitesimal gauge θ. We may see the variation explicitly in the following way. Let σ_1 and σ_2 be two sections over U. Then, as we saw above, there is a smooth map $h: U \to H$ such that $\sigma_2(u) = \sigma_1(u)h(u)$ for $u \in U$. According to the product rule (Proposition **3**.4.10), differentiating this formula yields

$$\sigma_2^*\omega_G = \mathrm{Ad}(h^{-1})\sigma_1^*\omega_G + h^*\omega_H,$$

that is,

$$\theta_2 = \mathrm{Ad}(h^{-1})\theta_1 + h^*\omega_H.$$

More explicitly, if $v \in T_u(U)$, then $\theta_2(v) = \mathrm{Ad}(h^{-1})\theta_1(v) + \omega_H(h_*(v))$.

We call such a variation of the infinitesimal gauge θ an *(infinitesimal) change of gauge*, and the two infinitesimal gauges are said to be *(infinitesimally) gauge equivalent*. We denote this symbolically by

$$\theta_1 \Rightarrow_h \theta_2.$$

It is quite clear from the "integral" version of the gauge that this is an equivalence relation, at least among gauges with a common domain of definition.

---

[17] If $g\sigma$ is also a section, then $x = \pi(g\sigma(x))$ and hence $gx = x$ for all $x \in U$. Thus $f^{-1}gf \in H$ for all $f \in \pi^{-1}(U) \subset G$. This shows that $\{f \in G \mid f^{-1}gf \in H\}$ contains an open subset of G. But it is also an analytic subset of G, so it must be equal to G, at least for G connected. Thus $g \in \cap_{f\in G} fHf^{-1}$, and this latter is a normal subgroup of G in H and is therefore trivial.

Exercise 7.2. Let $\hat{\omega}_G$ be the *right*-invariant Maurer–Cartan form on G, and set $\hat{\theta} = \sigma^*\hat{\omega}_G$. Show that this (right) infinitesimal gauge transforms according to $\hat{\theta}_2 = \hat{\theta}_1 + \mathrm{Ad}(\sigma_1)h^*\hat{\omega}_H$, where $\sigma_2 = \sigma_1 h$. ❑

To prepare for later work with Cartan geometries, we make the following definition.

Definition 7.3. A *gauge symmetry* of a Klein geometry (G, H) is a smooth map $b: G \to G$ such that

(i) $b(gh) = b(g)h$ for all $h \in H$, and

(ii) $b(g) \in gHg^{-1}$ for all $g \in G$. ❀

Exercise 7.4. Show that a gauge symmetry is just a bundle automorphism of the principal bundle $H \to G \to G/H$. ❑

Exercise 7.5. Show that a gauge symmetry $b: G \to G$ determines and is determined by a smooth map $f: G/H \to G$ such that $f(g) \in gHg^{-1}$ for all $g \in G$. ❑

5

Shapes High Fantastical: Cartan Geometries

> *In the wake of the movement of ideas which followed the general theory of relativity, I was led to introduce the notion of new geometries, more general than Riemannian geometry, and playing with respect to the different Klein geometries the same role as the Riemannian geometries play with respect to Euclidean space. The vast synthesis that I realized in this way depends of course on the ideas of Klein formulated in his celebrated Erlangen programme while at the same time going far beyond it since it includes Riemannian geometry, which had formed a completely isolated branch of geometry, within the compass of a very general scheme in which the notion of group still plays a fundamental role.*
>
> –E. Cartan, 1939

The universe appears to be a nice mixture of homogeneity and nonhomogeneity. At almost any location and scale, one is likely to see nothing (i.e., homogeneity) except for some concentrations ("lumps") of *something* at some great distance. If we move to the center of one of these concentrations, the nonhomogeneity may become more apparent; but if we then change our scale, the nonhomogeneity disappears and we are left with virtual homogeneity again. In a broad sense this is the fractal nature of the universe. Human life on earth is quite exceptional in this regard. We appear to be located at a position and scale where the nonhomogeneity is quite manifest. Nevertheless, we have not been so completely drowned in nonhomogeneity that we were prevented from discovering Euclidean geometry, a totally homogeneous idealization of our circumstance. Until this century, physics was regarded as the study of events in the *amphitheater* (as it were) of Euclidean geometry.[1] The present century has seen the appearance of "*lumpy geometry*" in physics, both in Einstein's theory of general relativity as well as in the gauge theories of electromagnetism and of weak and strong interactions. Here geometry sheds its passive appearance as backdrop and assumes increasingly the role of actor. Indeed the question arises, "Is geometry in its various forms the *only* actor?"

The new geometries in the quotation from Cartan at the head of this chapter refer to his "*espaces généralisés*," or what we are calling here "Cartan geometries." If Klein geometries represent *perfect homogeneity*, then Cartan geometries represent *a perfect mixture of homogeneity and nonhomogeneity*. Riemannian geometry (see Chapter 6) had its origins in 1854 in Riemann's celebrated talk.[2] It can be regarded as a nonhomogeneous version of Euclidean space. Cartan's generalization of Klein's geometries do for them what Riemann's generalization does for Euclidean space, that is, it adds "lumps." If a geometric theory is going to be of any use in a seamless description of a nonhomogeneous universe, it had better be "lumpy." Physical gauge theories describe fermions ("particles") as complex-valued functions on a principal bundle over space-time, and bosons ("forces") as connections on this principal bundle (cf. [Y.I. Manin, 1989]). Thus, it may become physically interesting to understand these notions in their proper geometrical context.

Cartan geometries are modeled on, and named by, Klein geometries. For example, we refer to a Cartan geometry modeled on Möbius geometry as a Möbius geometry. This usage is justified, for instance, by the fact (cf. Theorem 5.1) that if the Cartan geometry is flat, then it is locally the same as the model space.

There are, in fact, *two forms* of Cartan geometries, the local form and the global form.[3] The two forms are equivalent only when the model Klein geometry is effective.[4]

Our first definition is of local character and is called the *base definition*. Here the geometrical side is least apparent and analysis comes to the fore. This appears to be the version most preferred by the physicists because of

---

[1]However, Gauss had his suspicions about the truth of this. After his discovery of hyperbolic geometry, he realized that it could equally well serve as the amphitheater of events. In his capacity as director of the project to survey the Kingdom of Hanover, he was aware that the measurement of earthly triangles would not settle the issue. This story is told in Chapter 9 of [W.K. Bühler, 1981]. It was Lobachevski's proposal to attempt to settle the question experimentally by studying stellar triangles.

[2]An English translation of this talk appears in [M. Spivak, 1975], p. 135.

[3]See Appendix A for the relation to Ehresmann connections.

[4]Although this is a very important case, it does not exhaust the interesting possibilities. For example, spin geometries are based on ineffective models.

the local and analytical character of measurement. The second definition is global and is called the *principal bundle definition*. Roughly speaking, it presents a Cartan geometry as "a deformation of a Lie group which fixes the cosets of a closed subgroup."

In §1 we use the gauge idea to generalize the notion of a Klein geometry (G, H) to that of the local version of "a Cartan geometry modeled on (G, H)." We also introduce the curvature 2-form, which is a measure of the failure of the structural equation. In §2 we see how, for effective geometries, a Cartan geometry determines a *principal bundle* and a *Cartan connection* on it. In §3 we give the global definition of a Cartan geometry in terms of the data obtained in §2. At the same time this definition generalizes the Lie-theoretic properties of the Klein geometries given on page 153. Section 3 continues with the study of a number of concepts related to Cartan geometry, including the tangent bundle, the curvature function, Bianchi identity, tensors, differentiation, and special geometries. Section 4 introduces the notion of developing a curve in a Cartan geometry as a curve in the model space. In the case of closed curves, this leads to the notion of the holonomy group. This group is a subgroup of the group G of the model. In the case of a complete flat geometry, it is just the monodromy group. It is one measure of the failure of the geometry to be Klein. Also in this section, the notion of geometric orientation is generalized to Cartan geometries to prepare the way for the classification in §5 of geometrically oriented locally Klein geometries among Cartan geometries. In §6 we study the *Cartan space forms*, a class of geometries generalizing the Riemannian space forms. These again turn out to be locally Klein, but the model may change. The ideas in this section revolve around the notion of *model mutation*, which refers to the possibility of altering the model Klein geometry on which a given Cartan geometry is modeled. It is a change that may alter the curvature but in which there is no loss of information about the original geometry. Finally, in the short §7, we apply the classification of Cartan space forms to the case of symmetric spaces. These latter can be defined for any reductive model.

§1. The Base Definition of Cartan Geometries

In Chapter 4 we showed how a Klein geometry M gives rise to a gauge, or rather an equivalence class of gauges, on each sufficiently small open set. We are now going to generalize this notion to that of a geometry *modeled* on a Klein geometry. But first there is the question of what data to take as the model geometry. The apparently simplest way would be to use a Klein geometry (G, H) itself as the model. This approach, however, would result in later difficulties related to the fact that the influence of the model is essentially local. For this reason we take the following definition.

Definition 1.1. A *model geometry* for a Cartan geometry consists of

(i) an effective[5] infinitesimal Klein geometry $(\mathfrak{g}, \mathfrak{h})$,

(ii) a Lie group H realizing $\mathfrak{h}$,

(iii) a representation, denoted $\mathrm{Ad}\colon H \to Gl_{\mathrm{Lie}}(\mathfrak{g})$, extending $\mathrm{Ad}_{\mathfrak{h}}\colon H \to Gl_{\mathrm{Lie}}(\mathfrak{h})$.

The kernel K of the representation $\mathrm{Ad}\colon H \to Gl_{\mathrm{Lie}}(\mathfrak{g})$ is called the *kernel of the model geometry.* If K is trivial, the model geometry is called *effective.* A model geometry is called *primitive* is the subalgebra $\mathfrak{h}$ is a maximal subalgebra of $\mathfrak{g}$. It is called *reductive* if there is an H module decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$. ❀

We note that the hypothesis of effectiveness in Definition 1.1(i) implies that the kernel K is a discrete subgroup of H. It is clear that an effective Klein geometry (G, H), or more generally a Klein geometry with discrete kernel, canonically determines a model geometry. From now on we assume that we have chosen once and for all a fixed model geometry.

Cartan Gauge

Definition 1.2. Consider a model geometry $(\mathfrak{g}, \mathfrak{h})$ with group H. A *Cartan gauge* with this model on a smooth manifold M is a pair (U, θ_U), where U is an open set of M and θ_U (which we may abbreviate by θ) is a $\mathfrak{g}$-valued 1-form on U satisfying the *regularity* condition that

$$\bar{\theta}_U\colon T_u(U) \xrightarrow{\theta_U} \mathfrak{g} \xrightarrow[\text{projection}]{\text{canonical}} \mathfrak{g}/\mathfrak{h}$$

is a linear isomorphism for each $u \in U$. (One usually assumes that U is a coordinate neighborhood in M, although this is not strictly necessary.) ❀

Definition 1.3. If M is a smooth manifold, then a *Cartan atlas* on M is a collection $\mathcal{A} = \{(U_\alpha, \theta_\alpha)\}$ of Cartan gauges with model $(\mathfrak{g}, \mathfrak{h})$ and group H such that

(i) (Covering) the U_α's form an open cover of M,

(ii) (Compatibility) if (U, θ_U), $(V, \theta_V) \in \mathcal{A}$, then there exists a smooth map $k\colon U \cap V \to H$ such that[6] $\theta_V = \mathrm{Ad}(k^{-1})\theta_U + k^*\omega_H$ on $U \cap V$.

[5]It is possible to omit the effectiveness condition here, but there seems to be no gain in doing so.

[6]This equation is an abbreviation. It means that for each $x \in U \cap V$ we have $(\theta_V)_x = \mathrm{Ad}(k(x)^{-1})(\theta_U)_x + (k^*)_x\omega_H$.

(This compatibility relation is also called *gauge equivalence.*) ❀

Let us take a moment to study the compatibility relation between θ_U and θ_V. As in the last chapter, we denote this relation by writing $\theta_U \Rightarrow_k \theta_V$. The following lemma is not unexpected since it is obvious for the original ("integral") version of gauges given in the last chapter.

Lemma 1.4. *Suppose that (θ_i, U) are Cartan gauges for $i = 1, 2, 3$. Then*

(i) $\theta_1 \Rightarrow_{id} \theta_1$,

(ii) $\theta_1 \Rightarrow_k \theta_2$ *implies* $\theta_2 \Rightarrow_{k^{-1}} \theta_1$,

(iii) $\theta_1 \Rightarrow_h \theta_2$ *and* $\theta_2 \Rightarrow_k \theta_3$ *imply* $\theta_1 \Rightarrow_{hk} \theta_3$.

Proof. (i) is obvious.

(ii) $\theta_2 = \mathrm{Ad}(k^{-1})\theta_1 + k^*(\omega_h)$ implies

$$\begin{aligned}\theta_1 &= \mathrm{Ad}(k)(\theta_2 - k^*(\omega_H))\\ &= \mathrm{Ad}(k)\ \theta_2 - \mathrm{Ad}(k)(k^*(\omega_H)).\end{aligned}$$

Now by the quotient rule, Corollary **3**.4.11, $\mathrm{Ad}(k)(k^*(\omega_H)) = -(k^{-1})^*\omega_H$, so

$$\theta_1 = \mathrm{Ad}(k)\theta_2 + (k^{-1})^*\omega_H, \quad \text{i.e., } \theta_2 \Rightarrow_{k^{-1}} \theta_1.$$

(iii) $\theta_2 = \mathrm{Ad}(h^{-1})\theta_1 + h^*(\omega_H)$ and $\theta_3 = \mathrm{Ad}(k^{-1})\theta_2 + k^*(\omega_H)$ imply

$$\begin{aligned}\theta_3 &= \mathrm{Ad}(k^{-1})(\mathrm{Ad}(h^{-1})\theta_1 + h^*(\omega_H)) + k^*(\omega_H)\\ &= \mathrm{Ad}((hk)^{-1})\theta_1 + \mathrm{Ad}(k^{-1})h^*(\omega_H) + k^*(\omega_H).\end{aligned}$$

Now by the product rule, Proposition **3**.4.10, applied to the composite

$$\begin{array}{ccccccc} U & \xrightarrow{\Delta} & U \times U & \xrightarrow{h\times k} & H \times H & \xrightarrow{\mu} & H,\\ u & \mapsto & (u,u) & \mapsto & (h,k) & \mapsto & hk\end{array}$$

we have

$$\begin{aligned}(hk)^*\omega_H &= ((h \times k)\Delta)^*\mu^*\omega_H\\ &= ((h \times k)\Delta)^*(\pi_1^*\mathrm{Ad}(k^{-1})\omega_H + \pi_2^*\omega_H)\\ &= (\pi_1(h \times k)\Delta)^*\mathrm{Ad}(k^{-1})\omega_H + (\pi_2(h \times k)\Delta)^*\omega_H\\ &= \mathrm{Ad}(k^{-1})h^*\omega_H + k^*\omega_H\end{aligned}$$

and so

$$\begin{aligned}\theta_3 &= \mathrm{Ad}(hk)^{-1})\theta_1 + \mathrm{Ad}(k^{-1})h^*(\omega_H)) + k^*(\omega_H)\\ &= Ad((hk)^{-1})\theta_1 + (hk)^*\omega_H.\end{aligned}$$

■

Exercise 1.5. Suppose that $\mathrm{Ad}\colon H \to Gl(\mathfrak{g}/\mathfrak{h})$ is surjective. Let (U, θ) be a Cartan gauge and let $x\colon U \to \mathfrak{g}/\mathfrak{h}$ be any local coordinate system. Show that each point $p \in U$ lies in a neighborhood $V \subset U$ such that the Cartan gauge $(V, \theta \mid V)$ is gauge equivalent to a gauge (V, ψ) with $\psi = dx \bmod \mathfrak{h}$. ❑

Just as in the case of manifolds, bundles, and foliations, we call two atlases *equivalent* if their union is also an atlas, and we note that there is a unique maximal atlas equivalent to a given one.

Definition 1.6. A *Cartan structure* on a smooth manifold M consists of an equivalence class of Cartan atlases on M. A *Cartan geometry*[7] is a smooth manifold M together with a specified Cartan structure. A Cartan geometry is *effective* if the model is effective. ❀

Definition 1.7. Let M_1 and M_2 be two Cartan geometries with the same model geometry. A diffeomorphism $\varphi\colon M_1 \to M_2$ is called an *isomorphism of geometries* (or a *geometric isomorphism*) if for any Cartan gauge (U, θ_U) on M_2, the gauge $(\varphi^{-1}(U), \varphi^*\theta_U)$ induced on M_1 is compatible with the Cartan atlas on M_1. ❀

Curvature

We note that in defining the notion of a Cartan geometry, we have not mentioned the structural equation. This is not an oversight. Rather, it constitutes Elie Cartan's basic insight as the means for generalizing the Klein geometries. In particular, for any Cartan gauge (U, θ_U), the $\mathfrak{g}$-valued 2-form $\Theta_U = d\theta_U + \frac{1}{2}[\theta_U, \theta_U]$ on U need not vanish. As we shall presently see, Θ_U is a measure of the nonhomogeneity—the lumpiness—of a Cartan geometry.

Definition 1.8. The form $\Theta_U = d\theta_U + \frac{1}{2}[\theta_U, \theta_U]$ on U associated to a Cartan gauge (θ, U) is called the *curvature* with respect to this gauge. ❀

Of course, like the gauge θ_U, the curvature is not intrinsically defined. Let us see how it transforms as we alter the gauge.

Lemma 1.9. *Suppose that (θ_i, U) are Cartan gauges on U for $i = 1, 2$. Then*

$$\theta_1 \Rightarrow_h \theta_2 \quad \textit{implies} \quad \Theta_2 = Ad(h^{-1})\Theta_1.$$

Proof. Calculating the exterior derivative of the equation

[7] Later in this chapter we shall give another definition (3.1, page 184) of a Cartan geometry equivalent to this one only in the effective case. The later definition is the definitive one.

$$\theta_2 = \mathrm{Ad}(h^{-1})\theta_1 + h^*\omega_H,$$

we get, using the formula for $d\{\mathrm{Ad}(h^{-1})\theta_1\}$ in Exercise **3**.4.14 and the symmetry of the bracket on 1-forms (Exercise **1**.5.20(ii)),

$$d\theta_2 = \mathrm{Ad}(h^{-1})d\theta_1 - \frac{1}{2}[\mathrm{Ad}(h^{-1})\theta_1, h^*\omega_H] - \frac{1}{2}[h^*\omega_H, \mathrm{Ad}(h^{-1})\theta_1] + h^*(d\omega_H).$$

It follows that

$$\begin{aligned}
\Theta_2 &= d\theta_2 + \frac{1}{2}[\theta_2, \theta_2] \\
&= \mathrm{Ad}(h^{-1})d\theta_1 - \frac{1}{2}[\mathrm{Ad}(h^{-1}\theta_1, h^*\omega_H] - \frac{1}{2}[h^*\omega_H, \mathrm{Ad}(h^{-1})\theta_1] + h^*d\omega_H \\
&\quad + \frac{1}{2}[\mathrm{Ad}(h^{-1})\theta_1 + h^*\omega_H, \mathrm{Ad}(h^{-1})\theta_1 + h^*\omega_H] \\
&= \mathrm{Ad}(h^{-1})\left\{d\theta_1 + \frac{1}{2}[\theta_1, \theta_1]\right\} + h^*\left\{d\omega_H + \frac{1}{2}[\omega_H, \omega_H]\right\} \\
&= \mathrm{Ad}(h^{-1})\Theta_1 \quad \text{since} \quad d\omega_H + \frac{1}{2}[\omega_H, \omega_H] = 0
\end{aligned}$$

by the structural equation given in §3 of Chapter 3. ■

A simple consequence of this is that the vanishing of the curvature is an intrinsic condition, independent of the choice of gauge.

Definition 1.10. A Cartan geometry whose curvature vanishes at every point is called *flat*. ❀

The analog for a Cartan geometry of the structural equation for a Lie group is flatness. While the strucural equation always holds for a Lie group, not all Cartan geometries are flat.

Example 1.11. The simplest examples of Cartan geometries are, of course, the Klein geometries G/H themselves equipped with the Cartan gauges described in Chapter 4. We showed there that these charts were all regular and compatible. This is called the canonical Cartan geometry on G/H. ◆

Example 1.12. The next simplest examples are the open subsets of a Klein geometry. These clearly inherit the structure of a Cartan geometry by restriction. In general, these will not be Klein geometries. Nevertheless, they still satisfy the Maurer–Cartan equation, so they are flat Cartan geometries. Moreover, if such a flat Klein geometry is not simply connected, each of its covering spaces will have the structure of a Cartan geometry induced on it by the covering map. These geometries are again flat. ◆

Example 1.13. The final example we give now is the case of a locally Klein geometry $\Gamma \setminus G/H$, where Γ is a subgroup of G that acts on G/H by left multiplication as a group of covering transformations. The transformations in Γ act as geometric isomorphisms on G/H. Since the projection $G/H \to \Gamma\backslash G/H$ is a covering map, it is a local diffeomorphism, and we can transport the Cartan atlas from G/H to $\Gamma \backslash G/H$. The result is a Cartan atlas on $\Gamma \backslash G/H$, and the covering projection is locally a geometric isomorphism. In particular, the curvature on $\Gamma \backslash G/H$ that vanishes so it is a flat Cartan geometry. ◆

Since all of these examples are flat, the reader may well be wondering if there are any nonflat examples. Of course, there are such examples. In particular, in §6 we will study cases where the curvature is nonzero but, in a certain sense, constant. More generally, it often happens that a submanifold of a Klein geometry can be canonically equipped with the structure of a Cartan geometry, and these are seldom flat. Examples of this are given in Chapters 6 and 7.

Finally, we define the notion of a gauge symmetry of a Cartan atlas.

Definition 1.14. Suppose that $\mathcal{A} = \{(U_\alpha, \theta_\alpha) \mid \alpha \in A\}$ is a maximal atlas for a Cartan geometry on the manifold M. A *gauge symmetry* of the Cartan geometry is a permutation $b: A \to A$ such that

(i) $U_{b(\alpha)} = U_\alpha$, and

(ii) b satisfies the compatibility condition: if $U_\alpha = U_\beta$ and $\theta_\alpha \Rightarrow_h \theta_\beta$, then $\theta_{b(\alpha)} \Rightarrow_h \theta_{b(\beta)}$. ❀

The full meaning of this rather subtle notion of symmetry will become evident only after we study, in §2, the principal bundle associated to a Cartan geometry. In any case it is clear from its definition that a gauge symmetry does not alter the geometry in any way.

§2. The Principal Bundle Hidden in a Cartan Geometry

In the case of a Klein geometry $M = G/H$, there is a canonical principal bundle η over M given by $H \to G \to M$ with the property that the Cartan gauges are all obtained by pulling back the Maurer–Cartan form on G via sections of η. It is natural to ask whether a similar bundle exists canonically for *any* Cartan geometry. The answer is that it does, provided the model geometry is effective, and in this section we describe it. The construction depends on the fundamental property of Klein geometries given in §4 of Chapter 4.

Proposition 2.1. *Let U support a Cartan geometry modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Let K be the kernel and let θ_j $(j = 1, 2)$ be two compatible Cartan*

gauges on U. Then $\theta_1 \Rightarrow_k \theta_2$ for a smooth function $k: U \to H$ that is unique up to multiplication with a smooth function $l: U \to K$. In particular, if the Cartan geometry is effective, then k is unique.[8]

Proof. If k_1 and k_2 are two such ks, then by Lemma 1.4,

$$\theta_1 \Rightarrow_{k_1 k_2^{-1}} \theta_1.$$

Thus it suffices to show that $\theta \Rightarrow_k \theta$ implies k takes values in K. That is, we must show that every solution k to the equation

$$\theta = \mathrm{Ad}(k^{-1})\theta + k^*\omega_H, \quad \text{where } k: U \to H, \tag{2.2}$$

takes values in K. Although this may look like a nasty problem in differential equations, it turns out that it can be solved by means of linear algebra by using the fundamental property of Klein geometries. Recall that we showed in Lemma **4.4.4** that the series of groups

$$\begin{aligned}
N_0 &= H, \text{ with Lie algebra } \mathfrak{n}_0 = \mathfrak{h},\\
N_1 &= \{h \in H \mid \mathrm{Ad}(h)v - v \in \mathfrak{n}_0 \text{ for all } v \in \mathfrak{g}\}, \text{ with Lie algebra } \mathfrak{n}_1,\\
N_2 &= \{h \in H \mid \mathrm{Ad}(h)v - v \in \mathfrak{n}_1 \text{ for all } v \in \mathfrak{g}\}, \text{ with Lie algebra } \mathfrak{n}_2,\\
&\cdots\\
N_k &= \{h \in H \mid \mathrm{Ad}(h)v - v \in \mathfrak{n}_{k-1} \text{ for all } v \in \mathfrak{g}\}, \text{ with Lie algebra } \mathfrak{n}_k,\\
&\cdots
\end{aligned}$$

are all Lie groups that are closed and normal in H. They satisfy $N_0 \supset N_1 \supset N_2 \supset \ldots \supset N_k \supset \ldots$, and this chain stabilizes after finitely many steps at a group N_∞ whose Lie algebra $\mathfrak{n}_\infty$ is an ideal in $\mathfrak{g}$; moreover, $N_\infty = \{h \in H \mid \mathrm{Ad}(h)v - v \in \mathfrak{n}_\infty \ \forall v \in \mathfrak{g}\}$.

In the present case, since $\mathfrak{n}_\infty \subset \mathfrak{h}$, it follows (cf. Definition 1.1(i)) that $\mathfrak{n}_\infty = 0$. Thus, N_∞ is discrete and

$$\begin{aligned}
N_\infty &= \{h \in H \mid \mathrm{Ad}(h)v - v = 0 \text{ for all } v \in \mathfrak{g}\}\\
&= \ker(\mathrm{Ad}: H \to Gl(\mathfrak{g})) = K.
\end{aligned}$$

We now show by induction that k takes its values in N_s, for $s = 0, 1, \ldots$. Since $N_0 = H$, we see that $k: U \to N_s$ for $s = 0$. Assume inductively that $k: U \to N_s$. Fix $u \in U$. Then Eq. (2.2), rewritten as $\mathrm{Ad}(k^{-1})\theta - \theta = -\omega_H k_*$, says that for any $v \in \theta(T_u(U))$ we have

$$\mathrm{Ad}(k(u)^{-1})v - v \in \mathrm{image}(\omega_H k_{*u}) \subset \mathfrak{n}_s.$$

[8] We remark that this result is true even in cases where K is not discrete. Thus, it applies to the more general notion of a Cartan geometry with model as described in footnote 5.

But since N_s is normal in H, by Exercise **3**.4.7(c), the same conclusion also holds if $v \in \mathfrak{h}$.

By the regularity condition on θ, $\mathfrak{g} = \mathfrak{h} + \theta(T_u(U))$, which implies

$$\mathrm{Ad}(k(u)^{-1})v - v \in \mathfrak{n}_s$$

for every $v \in \mathfrak{g}$. Thus, $k(u) \in N_{s+1}$. But u was arbitrary here, so in fact $k\colon U \to N_{s+1}$. It follows that $k\colon U \to N_\infty = K$. ∎

The Principal Bundle of an Effective Cartan Geometry

Now we are ready to pass to a description of the principal H bundle associated to an effective Cartan geometry. Let $\mathcal{U} = \{U\}$ be a cover of M by sufficiently small, connected open sets (i.e., so that each is contained in the domain of a Cartan gauge and the intersection of any two of them is connected). The principal bundle we seek is obtained by glueing together the products $U \times H$ for all $U \in \mathcal{U}$. For each $U \in \mathcal{U}$, we choose a representative connection 1-form θ_U. Now if $U_1, U_2 \in \mathcal{U}$ have corresponding forms θ_1, θ_2, then along $U_1 \cap U_2$ there is a gauge equivalence $\theta_1 \Rightarrow_k \theta_2$ given, according to Proposition 2.1, by a uniquely determined smooth map $k\colon U_1 \cap U_2 \to H$. We can glue $U_1 \times H$ to $U_2 \times H$ along the common $(U_1 \cap U_2) \times H$ by making the identification $(u, h) \leftrightarrow (u, k(u)^{-1}h)$ as in Figure 2.2. It is easily checked that these identifications are compatible along the intersections of three open sets, so they fit together to give a right principal H bundle P. Of course, the right action of H in a coordinate patch is just right multiplication on the second factor, which clearly commutes with the identification maps.

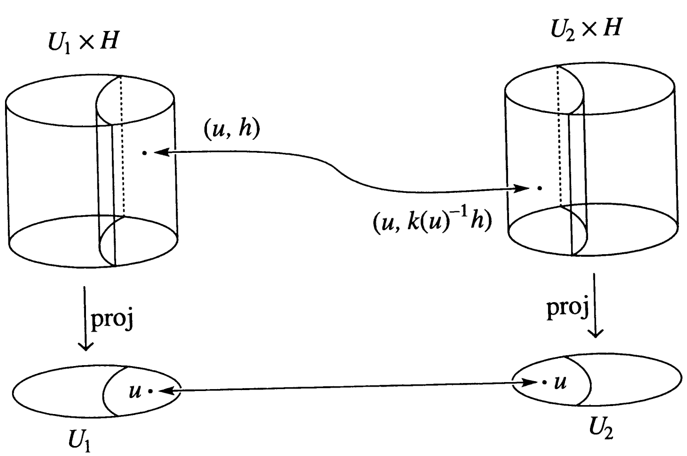


FIGURE 2.2.

By this construction, an effective Cartan geometry uniquely determines a right principal H bundle $\pi\colon P \to M$.

Right Multiplication and Ad

As we know, each vector $v \in \mathfrak{h}$ uniquely determines a left-invariant vector field V on the Lie group H whose value at e is v. Such a vector field extends over $U \times H$ as $v^\dagger = (0, V)$. Because the field V is left invariant and the bundle P is made up of identification maps that alter the second factor only by left multiplication, the vector fields $v^\dagger$ on the coordinate patches $U \times H$ fit together to yield a well-defined vector field $v^\dagger$ on P itself.

Lemma 2.3. *Let $R_k\colon P \to P$ denote right multiplication by $k \in H$. Then $(R_k)_*(v^\dagger) = (Ad(k^{-1})v)^\dagger$ for all $v \in \mathfrak{h}$.*

Proof. Since both sides are invariantly defined, we need only check the formula on a coordinate patch of the form $U \times H$. In this coordinate patch, right multiplication takes the form $R_k = \mathrm{id} \times r_k$, where $r_k\colon H \to H$ is just right multiplication on H. Let $l_k\colon H \to H$ denote left multiplication on H. Thus, in our coordinate patch we have

$$\begin{aligned}(R_k)_*(v^\dagger) &= (\mathrm{id} \times r_k)_*((0, V)) = (0, r_{k*}V)\\ &= (0, r_{k*}l_{k^{-1}*}V) = (\mathrm{Ad}(k^{-1})v)^\dagger,\end{aligned}$$

where we use the fact that V is a left-invariant vector field to write $V = l_{k^{-1}*}V$. ∎

The Cartan Connection

In addition to the bundle P just constructed, we also get a $\mathfrak{g}$-valued 1-form ω on P, called the *Cartan connection*, arising as follows. Given a gauge (U, θ), we have the canonical linear isomorphism

$$\begin{array}{cccccc} \omega\colon T_{(u,h)}(U \times H) & \xrightarrow{\text{canonical}} & T_u(U) \times T_h(H) & \to & T_u(U) \times \mathfrak{h} \to \mathfrak{g}. \\ (v,y) & \mapsto & (v, \omega_H(y)) & \mapsto & \mathrm{Ad}(h^{-1})\theta(v) + \omega_H(y) \end{array}$$

Let us verify that as we vary the gauge, these isomorphisms fit together smoothly to give a 1-form ω on P. We do this by comparing them by a transition function of the form

$$\begin{array}{ccc} f = (f_1, f_2)\colon U \times H & \to & U \times H. \\ (u,h) & \mapsto & (u, k(u)^{-1}h) \end{array}$$

The derivative of this map is the vertical left-hand map in the following diagram.

$$\begin{tikzcd} T_{(u,h)}(U\times H) \arrow[r] \arrow[d] & \mathfrak{g} \arrow[d, equal] \\ T_{(u,k(u)^{-1}h)}(U\times H) \arrow[r] & \mathfrak{g} \end{tikzcd} \quad \text{where} \quad \begin{tikzcd} (v,y_1) \arrow[r, mapsto] \arrow[d] & \mathrm{Ad}(h^{-1})\theta_1(v)+\omega_H(y_1) \arrow[d, dashed, "?"] \\ (v,y_2) \arrow[r, mapsto] & \mathrm{Ad}(h^{-1}k)\theta_2(v)+\omega_H(y_2) \end{tikzcd}$$

Since $f_2(u,h) = k(u)^{-1}h$ and $f_{2*}(v,y_1) = y_2$, it follows from Proposition **3**.4.10 ii) and Exercise **3**.4.12 that

$$\omega_H(y_2) = \omega_H(f_{2*}(v,y_1)) = f_2^*(\omega_H)(v,y_1) = -\mathrm{Ad}(h^{-1}k)(k^*\omega_H)v + \omega_H(y_1).$$

This fact, together with the formula $\theta_2 = \mathrm{Ad}(k^{-1})\theta_1 + k^*\omega_H$, verifies the arrow conjectured in the diagram above.

Properties of the Cartan Connection

Proposition 2.4. *The Cartan connection ω on P with values in $\mathfrak{g}$ has the properties*

(i) *for each point $p \in P$, the linear map $\omega_p: T_p(P) \to \mathfrak{g}$ is an isomorphism,*

(ii) *$(R_h)^*\omega = Ad(h^{-1})\omega$,*

(iii) *$\omega(v^\dagger) = v$ for all $v \in \mathfrak{h}$.*

Proof. (i) Since $\dim P = \dim M + \dim H = \dim \mathfrak{g}/\mathfrak{h} + \dim \mathfrak{h} = \dim \mathfrak{g}$, it suffices to show that $\omega_p: T_p(P) \to \mathfrak{g}$ is a monomorphism for each p. Thus, it suffices to show that for $v \in T_u(U)$ and $y \in T_h(H)$, the equation $\mathrm{Ad}(h^{-1})\theta(v) + \omega_H(y) = 0$ implies $(v,y) = 0$. Now $\mathrm{Ad}(h^{-1})\theta(v) = -\omega_H(y)$ lies in $\mathfrak{h}$, and since $\mathfrak{h}$ is $\mathrm{Ad}(H)$ stable, this means that $\theta(v) \in \mathfrak{h}$. Thus $v = 0$ by the regularity property for Cartan gauges. But then $\omega_H(y) = 0$, and hence $y = 0$ since ω_H is injective.

(ii) We must verify the commutativity of the following diagram.

$$\begin{tikzcd} T_{(u,h)}(U\times H) \arrow[r, "\omega"] \arrow[d, "R_{k*}"'] & \mathfrak{g} \arrow[d, "Ad(k^{-1})"'] \\ T_{(u,hk)}(U\times H) \arrow[r, "\omega"] & \mathfrak{g} \end{tikzcd} \quad \text{where} \quad \begin{tikzcd} (v,y) \arrow[r, mapsto] \arrow[d] & \mathrm{Ad}(h^{-1})\theta(v)+\omega_H(y) \arrow[d, dashed, "\mathrm{Ad}(k^{-1})"'] \\ (v,R_{k*}y) \arrow[r, mapsto] \arrow[ur, phantom, "?"] & \mathrm{Ad}(hk)^{-1}\theta(v)+\omega_H(R_{k*}y) \end{tikzcd}$$

But $\mathrm{Ad}(hk)^{-1}\theta(v) + \omega_H(R_{k*}y) = \mathrm{Ad}(k^{-1})(\mathrm{Ad}(h^{-1})\theta(v) + \omega_H(y))$.

(iii) The vector field $v^\dagger$ in a chart is given by $(0,V)$, where V is the left-invariant vector field on H corresponding to v. Thus, $\omega(0,V) = \mathrm{Ad}(h^{-1})\theta(0) + \omega_H(V) = v$. ■

Note that the Cartan connection ω on P gives a canonical parallelization of the space P and is the analog of the Maurer–Cartan form that appears on $P = G$ in the case of a Klein geometry $M = G/H$.

Proposition 2.5. *Let $\sigma: U \to P$ be any section over U.*

(i) *$\theta = \sigma^*\omega$ is a Cartan gauge on U compatible with the Cartan geometry on M.*

(ii) *Let Ω be the $\mathfrak{g}$-valued 2-form on P given by $\Omega = d\omega + \frac{1}{2}[\omega,\omega]$. Then $\sigma^*\Omega = \Theta$, the curvature computed from the gauge $\sigma^*\omega$.*

Proof. (i) To see this we note that it is automatically true for any of the sections used to define P and that the others differ from these by gauge transformations.

(ii) $\Theta = d\theta + \frac{1}{2}[\theta,\theta] = \sigma^*(d\omega + \frac{1}{2}[\omega,\omega]) = \sigma^*(\Omega)$. ■

Proposition 2.6. *There is a canonical one-to-one correspondence between gauge symmetries of an effective Cartan geometry and bundle automorphisms of the bundle P.*

Proof. Let $b: P \to P$ be a bundle automorphism and ω the Cartan connection on P. Then b determines a gauge symmetry on the maximal atlas $\mathcal{A}$ in the following fashion. Given a connected open set $U \subset M$ over which P is trivial, and a section $\sigma: U \to P$, we may compose σ with b to obtain another section $b\sigma: U \to P$. Thus, we obtain two Cartan gauges, $(U, \sigma^*(\omega))$ and $(U, \sigma^*(b^*\omega))$. We define the gauge symmetry by sending $(U, \sigma^*(\omega)) \to (U, \sigma^*(b^*\omega))$. Compatibility condition (i) for gauge symmetries is automatic, and condition (ii) follows also, for if $\theta_\alpha \Rightarrow_h \theta_\beta$, then writing $\theta_\alpha = \sigma_\alpha^*(\omega)$ and $\theta_\beta = \sigma_\beta^*(\omega)$, we have $\sigma_\beta = \sigma_\alpha h$ and hence $b\sigma_\beta = b(\sigma_\alpha h) = b(\sigma_\alpha)h$ since b is a bundle map; thus, $(b\sigma_\alpha)^*\omega \Rightarrow_h (b\sigma_\beta)^*\omega$, which is condition (ii).

Conversely, suppose that we are given a gauge symmetry b permuting the charts in a maximal atlas for a Cartan geometry. Then we may construct a bundle automorphism of the bundle P as follows. Given a chart $(U_\alpha, \theta_\alpha) \in \mathcal{A} = \{(U_\alpha, \theta_\alpha) \mid \alpha \in \mathcal{A}\}$, we have the associated chart $(U_{b(\alpha)}, \theta_{b(\alpha)}) \in \mathcal{A}$ given by the gauge symmetry. Now $U_\alpha = U_{b(\alpha)} = U$, say. These two charts determine two trivializations (which are bundle maps), $\psi_\alpha: U \times H \to P$ and $\psi_{b(\alpha)}: U \times H \to P$. We define the bundle map over U by $\psi_{b(\alpha)}\psi_\alpha^{-1}$. Moreover, compatibility condition (ii) for gauge symmetries ensures that if $(U_\beta, \theta_\beta) \in \mathcal{A}$ is another chart with $U_\beta = U_\alpha$, then the same gauge transformation h giving $\theta_\alpha \Rightarrow_h \theta_\beta$ also identifies their images under the gauge symmetry $\theta_{b(\alpha)} \Rightarrow_h \theta_{b(\beta)}$. This means that $\psi_\beta(x,k) = \psi_\alpha(x, h(x)k)$ and $\psi_{b(\beta)}(x,k) = \psi_{b(\alpha)}(x, h(x)k)$, and hence $\psi_{b(\alpha)}\psi_\alpha^{-1} = \psi_{b(\beta)}\psi_\beta^{-1}$. Thus, the bundle maps we have defined over our open sets U fit together to give a uniquely determined bundle map on P.

Finally, we leave it as an easy task for the reader to verify that the two correspondences we have described are actually inverse to each other. ∎

§3. The Bundle Definition of a Cartan Geometry

Motivated either by the discussion above or by the Lie-theoretic properties of Klein geometries mentioned on p. 153, we are led to the following alternate definition of a Cartan geometry.

Definition 3.1. A *Cartan geometry* $\xi = (P, \omega)$ on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H consists of the following data:

(a) a smooth manifold M;

(b) a principal right H bundle P over M;

(c) a $\mathfrak{g}$-valued 1-form ω on P satisfies the following conditions:

(i) for each point $p \in P$, the linear map $\omega_p : T_p(P) \to \mathfrak{g}$ is an isomorphism;

(ii) $(R_h)^*\omega = \mathrm{Ad}(h^{-1})\omega$ for all $h \in H$;

(iii) $\omega(X^\dagger) = X$ for all $X \in \mathfrak{h}$.

By abuse of notation, we also speak of a *Cartan geometry* M. The $\mathfrak{g}$-valued 2-form on P given by $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$ is called the *curvature*. If $\rho : \mathfrak{g} \to \mathfrak{g}/\mathfrak{h}$ is the canonical projection, then $\rho(\Omega)$ is called the *torsion*. If Ω takes values in the subalgebra $\mathfrak{h}$, we say the geometry is *torsion free*, or *without torsion*. The geometry is called *complete* if the form ω is complete, that is, if all the ω-constant vector fields are complete.[9] We say a Cartan geometry is *effective*, *primitive* or *reductive*, respectively, if the model geometry is effective, primitive or reductive. ❀

A geometry in the global sense of Definition 3.1 determines one in the local sense of Definition 1.6, but the definitions are equivalent only when the model is effective. The discussion in §2 culminating in Proposition 2.5 of the previous section shows that every effective Cartan geometry in the original sense corresponds to an effective Cartan geometry in the new sense. Proposition 2.6 shows that the converse is true and that the two correspondences are inverse to each other. We shall henceforth take the present definition as the definitive one.

[9]It would be very interesting to have a definition of completeness in terms of M somewhat along the lines of Riemannian geometry, namely, something like completeness of geodesics. Cf. [B. Kamté, 1995].

Definition 3.2. Let (P_1, ω_1) and (P_2, ω_2) be Cartan geometries on M_1 and M_2 respectively, modeled on the pair $(\mathfrak{g}, \mathfrak{h})$ with group H. Let $f : M_1 \to M_2$ be an immersion covered by an H bundle map $\tilde{f} : P_1 \to P_2$ with the property that $\tilde{f}^*\omega_2 = \omega_1$. Then f is called a *local isomorphism of geometries*, or a *local geometric isomorphism*. If in addition f is a diffeomorphism then it is called *an isomorphism of geometries*, or a *geometric*[10] *isomorphism*. ❀

Exercise 3.3. If (P, ω) is a Cartan geometry and $b : P \to P$ is any bundle automorphism, show that $(P, b^*(\omega))$ is also a Cartan geometry. ❑

Note, in particular, that if $\xi = (P, \omega)$ is a Cartan geometry on M, then it is geometrically equivalent to the geometry $b^*\xi = (P, b^*\omega))$, where $b : P \to P$ is *any* bundle map. It turns out that the two Cartan connections will be equal if and only if the b is the identity (see Theorem 3.5 ahead), but it is clear that the geometries must be regarded as "the same" from the base definition of the geometry. The two Cartan connections ω and $b^*\omega$ are called *gauge equivalent connections*.

Exercise 3.4. Verify that two effective Cartan geometries are isomorphic in the original sense of Definition 1.7 if and only if their corresponding geometries are isomorphic in the sense of Definition 3.2. ❑

We have the following uniqueness theorem for the bundle map covering an isomorphism of geometries.

Theorem 3.5. *Suppose that $\varphi : M_1 \to M_2$ is an isomorphism of effective Cartan geometries, and let $f_j : P_1 \to P_2$ $(j = 1, 2)$ be two H bundle maps covering φ and satisfying $f_j^*\omega_2 = \omega_1$ $(j = 1, 2)$. Then $f_1 = f_2$.*

Proof. Setting $f = f_1^{-1} f_2$ (where $f_1^{-1} : P_2 \to P_1$ is the inverse of f_1), we obtain an H bundle map $f : P_1 \to P_1$ satisfying $f^*\omega_1 = \omega_1$ and covering the identity map. Thus, it suffices to show that such a map is the identity. We set $\omega = \omega_1$ and $P = P_1$. Define $\psi : P \to H$ by $f(p) = p\psi(p)$. Let us factor f as

$$f : P \xrightarrow{\Delta} P \times P \xrightarrow{\mathrm{id} \times \psi} P \times H \xrightarrow{\mu} P,$$

where μ is the right multiplication map. Thus

$$\begin{aligned}
\omega &= (\mu \circ (\mathrm{id} \times \psi) \circ \Delta)^*\omega \\
&= \Delta^*(\mathrm{id} \times \psi)^*\mu^*\omega \\
&= \Delta^*(\mathrm{id} \times \psi)^*\{\pi_1^*\mathrm{Ad}(\psi^{-1})\omega + \pi_2^*\omega_H\} \\
&= \mathrm{Ad}(\psi^{-1})\omega + \psi^*\omega_H.
\end{aligned}$$

[10]Cartan uses the term *holoédrique* instead of *geometric*.

It follows that $\mathrm{Ad}(\psi^{-1})\omega - \omega = -\psi^*\omega_H$. But then, proceeding as in the last paragraph of Proposition 2.1, we may show by induction that ψ takes its values in N_s, for $s = 0, 1, 2, \ldots$. Since $N_0 = H$, we see that $\psi: P \to N_s$ for $s = 0$. Assume inductively that $\psi: P \to N_s$. Fix $p \in P$. Then the equation says that for $v \in \omega_p(T_p(P)) = \mathfrak{g}$ we have $\mathrm{Ad}(\psi(p)^{-1})v - v \in \mathrm{image}(\omega_H \psi_{*x}) \subset \mathfrak{n}_s$. Thus, $\psi(p) \in N_{j+1}$ by definition. But p was arbitrary here, so in fact $\psi: P \to N_{j+1}$. It follows by induction that $\psi: P \to N_\infty$. Since N_∞ is a normal subgroup of G that lies in H, it follows that $N_\infty \subset K$. Since we have assumed that the geometry is effective, $\psi(p) = 1$ for all $p \in P$. ■

The following two exercises express obvious facts about the hierarchy of Cartan geometries on bundles intermediate between P and M.

Exercise 3.6. Let (P, ω) be a Cartan geometry modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H and let $B \subset H$ be a closed subgroup with Lie algebra $\mathfrak{b}$. Show that (P, ω) may also be regarded as a Cartan geometry on P/B modeled on $(\mathfrak{g}, \mathfrak{b})$ with group B. ❑

Exercise 3.7. Let P be an H bundle over M and let $B \subset H$ be a closed subgroup. Assume that (P, ω) is a Cartan geometry on P/B modeled on $(\mathfrak{g}, \mathfrak{b})$ with group B. Show that necessary and sufficient conditions on the form ω so that (P, ω) is a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H are that condition (c), parts (ii) and (iii) hold for all elements of H and $\mathfrak{h}$, respectively, and not just those of B and $\mathfrak{b}$. ❑

Exercise 3.8. Resolve the following paradox. We have described above a procedure that yields a one-to-one correspondence:

$$\left\{ \begin{matrix} \text{Equivalence classes of} \\ \text{Cartan atlases on } M \end{matrix} \right\} \underset{\beta}{\overset{\alpha}{\rightleftarrows}} \left\{ \begin{matrix} \text{Cartan geometries} \\ (P, \omega) \text{ on } M \end{matrix} \right\}$$

Let (P, ω) be a Cartan geometry. If $b: P \to P$ is any *nontrivial* bundle automorphism, set $\eta = b^*(\omega)$. Then, by Theorem 3.5, $\eta \neq \omega$, and by Exercise 3.4, (P, η) is also a Cartan geometry. Now, given an open set $U \subset M$ and a section $\sigma: U \to P$, the forms ω and η pull back by σ to yield forms $\sigma^*\omega$ and $\sigma^*(\eta) = \sigma^*(b^*(\omega)) = (b\sigma)^*\omega$, which are compatible in the sense of Definition 1.3 since σ and $b\sigma$ are both sections over U (cf. Proposition 2.5). Thus $\beta(P, \eta) = \beta(P, \omega)$ and hence $(P, \eta) = \alpha\beta(P, \eta) = \alpha\beta(P, \omega) = (P, \omega)$. This is in apparent contradiction to the fact that $\eta \neq \omega$. ❑

As one might expect from the strong form of the invariance of the curvature in the base definition, there is an analogously strong invariance in the bundle definition. We have the following result.

Lemma 3.9. *Let (P, ω) be a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Assume $\psi: P \to H$ is a smooth map. Define $f: P \to P$ by $f(p) = R_{\psi(p)}p$. Then $f^*\Omega = \mathrm{Ad}(\psi(p))\Omega$.*

Proof. By the calculation in Theorem 3.5, we have $f^*\omega = \mathrm{Ad}(\psi^{-1})\omega + \psi^*\omega_H$. Thus, by a calculation entirely analogous to the one in Lemma 1.9, we obtain $f^*\Omega = \mathrm{Ad}(\psi(p))\Omega$. ■

Corollary 3.10. *The curvature form $\Omega(u, v)$ vanishes whenever u or v is tangent to the fiber.*

Proof. We may suppose that $u, v \in T_p(P)$ are independent and that v is tangent to the fiber. We may choose arbitrarily any map $\psi: (P, p) \to (H, e)$ such that $\psi_{*p}(v) = -\omega_p(v)$. (Since v is tangent to the fiber, $\psi_{*p}(v) \in \mathfrak{h} = T_e(H)$.) Define $f: P \to P$ by $f(q) = q \cdot \psi(q)$. By the calculation in Theorem 3.5, and by the lemma, at p we have

$$f^*\omega = \mathrm{Ad}(\psi^{-1})\omega + \psi^*\omega_H = \omega + \psi^*\omega_H \quad \text{and} \quad f^*\Omega = \Omega,$$

so that $\omega_p(f_*v) = \omega_p(v) + \omega_H \psi_{*p}(v) = \omega_p(v) - \omega_p(v) = 0$, that is, $f_*v = 0$. Thus, $\Omega(u, v) = (f^*\Omega)(u, v) = \Omega(f_*u, f_*v) = \Omega(f_*u, 0) = 0$. ■

Corollary 3.11. *The curvature form Ω may be regarded as a 2-form on the pullback of the tangent bundle of M to P.*

Proof. By Corollary 3.10, the curvature may be regarded as a 2-form on the quotient bundle $T(P)/\ker \pi_*$ that is canonically isomorphic to $\pi^*(T(M))$. ■

We also obtain various foliations on the principal bundle P, as seen in the following exercise.

Exercise 3.12.* Let M be a manifold and $\xi = (P, \omega)$ a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H.

(a) Let V be any vector subspace of the Lie algebra $\mathfrak{h}$. Show that $\omega^{-1}(V)$ is an integrable distribution on P if and only if V is a subalgebra of $\mathfrak{h}$.

(b) If ξ is torsion free and V is a vector subspace of $\mathfrak{g}$ containing $\mathfrak{h}$, show that $\omega^{-1}(V)$ is an integrable distribution on P if and only if V is a subalgebra of $\mathfrak{g}$. ❑

The manifold P may be regarded as some sort of "lumpy Lie group"[11] that is homogeneous in the H direction. Moreover, ω may be regarded

[11] Of course, it is not a group.

as a "lumpy" version of the Maurer–Cartan form. As we shall see in the following result, ω restricts to the Maurer–Cartan form on the fibers and hence satisfies the structural equation in the fiber directions; but when $\Omega \neq 0$, we lose the "rigidity" that would otherwise have been provided by the structural equation in the base directions and that would have as a consequence that, locally, P would be a Lie group with ω its Maurer–Cartan form. (This subject is studied in detail again in §5.) Thus, the curvature measures this loss of rigidity.

Lemma 3.13. *Each fiber F of the principal bundle P is canonically identified with H up to left multiplication by some element of H. The Maurer–Cartan form ω_H on H induces a canonical form ω_F on F that agrees with the restriction of the Cartan connection ω on P to F.*

Proof. By the very definition of a principal bundle, any fiber F of P may be canonically identified with H up to left multiplication by some element of H. This means that the left-invariant Maurer–Cartan form ω_H on H canonically determines a "Maurer–Cartan" form ω_F on the fiber F. Again by definition, the vector field $v^\dagger$ on P, for $v \in \mathfrak{h}$, restricts to a vector field on each fiber, which, under the canonical identification of F with H (always modulo left multiplication), corresponds to a left-invariant vector field on H. Then the condition $\omega(v^\dagger) = v$ for all $v \in \mathfrak{h}$ implies $\omega_F = \omega \mid F$. ■

Exercise 3.14. Use this result to get a second proof that the curvature form vanishes when restricted to any fiber. ❏

Tangent Bundle of a Cartan Geometry

In §4.5, we saw that the tangent bundle of a Klein geometry G/H can be expressed as a vector bundle associated to the principal bundle $H \to G \to G/H$ via the representation

$$\mathrm{Ad}_{\mathfrak{g}/\mathfrak{h}}: H \to \mathrm{End}(\mathfrak{g}/\mathfrak{h}),$$

so that $T(G/H) \approx G \times_H \mathfrak{g}/\mathfrak{h}$. We also saw the related fact that for each element $g \in G$ there is a canonical linear isomorphism $\varphi_g: T_{gH}(G/H) \to \mathfrak{g}/\mathfrak{h}$ such that $\varphi_{gh} = \mathrm{Ad}(h^{-1})\varphi_g$. These relationships continue to hold for Cartan geometries. This is expressed in the next result.

Theorem 3.15. *Let (P,ω) be a Cartan geometry on M modeled on $(\mathfrak{g},\mathfrak{h})$ with group H. Then there is a canonical bundle isomorphism $T(M) \approx P \times_H \mathfrak{g}/\mathfrak{h}$. Moreover, for each point $p \in P$ with $\pi(p) = x$, there is a canonical linear isomorphism $\varphi_p: T_x(M) \to \mathfrak{g}/\mathfrak{h}$ such that $\varphi_{ph} = Ad(h^{-1})\varphi_p$.*

Proof. Consider the following diagram. The columns are short, exact sequences and the two upper rows are isomorphisms, so there is a canonical isomorphism across the bottom making the diagram commute.

$$\begin{tikzcd}
T_p(pH) \arrow[r, "\omega_H", "\approx"'] \arrow[d] & \mathfrak{h} \arrow[d, phantom, "\cap"] \\
T_p(P) \arrow[r, "\omega", "\approx"'] \arrow[d] & \mathfrak{g} \arrow[d, "\rho"] \\
T_x(M) \arrow[r, dashed, "\varphi_p", "\approx"'] & \mathfrak{g}/\mathfrak{h}
\end{tikzcd}$$

Moreover, if $v \in T_x(M)$, we may write $v = \pi_{*p}(u) = \pi_{*ph}(R_{h*}u)$ for some $u \in T_p(P)$. Thus,

$$\begin{aligned}
\varphi_{ph}(v) &= \varphi_{ph}(\pi_{*ph}(R_{h*}u)) \\
&= \rho(\omega_{ph}(R_{h*}u)) \\
&= \rho(\mathrm{Ad}(h^{-1})\omega_p(u)) \\
&= \mathrm{Ad}(h^{-1})\rho(\omega_p(u)) \\
&= \mathrm{Ad}(h^{-1})\varphi_p(\pi_{*p}(u)) \\
&= \mathrm{Ad}(h^{-1})\varphi_p(v).
\end{aligned}$$

It follows that we may define a smooth bundle map

$$\begin{aligned}
q: P \times \mathfrak{g} &\to T(M). \\
(p,w) &\mapsto (\pi(p), \varphi_p^{-1}(\rho(w))
\end{aligned}$$

Note that

$$\begin{aligned}
q(ph, \mathrm{Ad}(h^{-1})w) &= (\pi(ph), \varphi_{ph}^{-1}(\rho(\mathrm{Ad}(h^{-1})w))) \\
&= (\pi(p), (\mathrm{Ad}(h)\varphi_{ph})^{-1}(\rho(w)) \\
&= (\pi(p), \varphi_p^{-1}(\rho(w)) \\
&= q(p,w).
\end{aligned}$$

Thus, we get a canonical smooth bundle map $\bar{q}: P \times_H \mathfrak{g}/\mathfrak{h} \to T(M)$. This is a vector bundle equivalence since it is an isomorphism on fibers and induces the identity map on the base M. ■

Corollary 3.16. *Let (P,ω) be a Cartan geometry on M modeled on $(\mathfrak{g},\mathfrak{h})$ with group H. The vector fields X on M are in bijective correspondence with functions $f: P \to \mathfrak{g}/\mathfrak{h}$ transforming according to the adjoint representation (i.e., $f(ph) = \mathrm{Ad}_{\mathfrak{g}/\mathfrak{h}}(h^{-1})f(p)$). The correspondence is given by*

$$X \mapsto f_X = \{p \in P \mapsto \varphi_p(X_{\pi(p)}) \in \mathfrak{g}/\mathfrak{h}\}.$$

Proof. f_X transforms correctly since

$$f_X(ph) = \varphi_{ph}(X_{\pi(ph)}) = \mathrm{Ad}(h^{-1})\varphi_p(X_{\pi(p)}) = \mathrm{Ad}(h^{-1})f_X(p).$$

Conversely, a function $f\colon P \to \mathfrak{g}/\mathfrak{h}$ transforming according to the adjoint representation arises from the vector field X given by $X_x = \varphi_p^{-1}(f(p))$, where $p \in P$ is any point lying over $x \in M$. The choice of p does not matter since $\varphi_{ph}^{-1}(f(ph)) = (\mathrm{Ad}(h^{-1})\varphi_p)^{-1}(\mathrm{Ad}(h^{-1})f(p)) = \varphi_p^{-1}(f(p))$. ■

Definition 3.17. Let (P,ω) be a Cartan geometry on M modeled on $(\mathfrak{g},\mathfrak{h})$ with group H. A vector bundle $E \to P$ is called a *geometric vector bundle* if it is given in the form $E = P \times_H V$ for some representation $\rho\colon H \to Gl(V)$. ❀

Example 3.18. Theorem 3.15 shows that $T(M)$ is always a geometric bundle. ◆

Proposition 3.19. *Let M be a connected manifold. Let (P,ω) be a Cartan geometry on M modeled on $(\mathfrak{g},\mathfrak{h})$ with groups H. Fix $x \in M$ and fix $p \in P$ lying over x. Then*

$$M \text{ is orientable} \Leftrightarrow \left\{\begin{array}{l} Ad_{\mathfrak{g}/\mathfrak{h}}(h) \in Gl^+(\mathfrak{g}/\mathfrak{h}) \text{ for every} \\ h \in H \text{ such that } ph \in P \text{ lies in the} \\ \text{same path components as } p \in P. \end{array}\right.$$

Proof. Recall that M is orientiable if and only if every loop on M is orientation preserving (Proposition **1**.1.17). Let $\lambda\colon (I,0,1) \to (M,x,x)$ be a loop on M. By Proposition **1**.1.14, we may choose a partition of I, $0 = t_0 < t_1 < \cdots < t_k = 1$, and a compatible family of charts (U_i,ψ_i), $1 \le i \le k$, such that $\lambda([t_{i-1},t_i]) \subset U_i$, $1 \le i \le k$. Let $\sigma\colon (I,0,1) \to (P,p,hp)$ be any lift of λ. We consider the linear isomorphisms

$$\mathbf{R}^n \xrightarrow{\psi^{-1}_{i*\lambda(t)}} T_{\sigma(t)}(M) \xrightarrow{\varphi_{\sigma(t)}} \mathfrak{g}/\mathfrak{h}, \quad t \in [t_{i-1},t_i],\ 1 \le i \le k,$$

and a basis for $\mathfrak{g}/\mathfrak{h}$ so that this composite has positive determinant for $t = 0$. The compatibility of the family of charts together with the continuity of these maps implies that these maps have positive determinant for all $t \in I$. Then the commutativity of the diagram

$$\begin{tikzcd}
\mathbf{R}^n \arrow[d, "\psi_{k*x}\psi_{1*x}^{-1}"'] & T_x(M) \arrow[l, "\psi_{1*x}"'] \arrow[r, "\varphi_p"] \arrow[d, equal] & \mathfrak{g}/\mathfrak{h} \arrow[d, "\varphi_{ph}\varphi_p^{-1} = \mathrm{Ad}(h^{-1})"] \\
\mathbf{R}^n & T_x(M) \arrow[l, "\psi_{k*x}"'] \arrow[r, "\varphi_{ph}"] & \mathfrak{g}/\mathfrak{h}
\end{tikzcd}$$

shows that the loop λ is orientation preserving (i.e., $\psi_{k*x}\psi_{1*x}^{-1}$ has positive determinant) if and only if $\mathrm{Ad}(h^{-1})$ has positive determinant.

Thus, if M is orientable, then a path on P joining p to ph projects to an orientable loop on M, and hence $\mathrm{Ad}(h)$ has positive determinant.

Conversely, suppose $\mathrm{Ad}(h)$ has positive determinant whenever p can be joined to ph by a path. Then if λ is any loop on M based at x, we can lift λ to a path joining p and ph for some $h \in H$. Since $\mathrm{Ad}(h)$ has positive determinant, it follows that λ is orientation preserving. ■

Definition 3.20. A Cartan geometry $\xi = (P,\omega)$ on M, modeled on $(\mathfrak{g},\mathfrak{h})$ with group H, is a *first-order* geometry if $\mathrm{Ad}\colon H \to Gl(\mathfrak{g}/\mathfrak{h})$ is injective. Otherwise it is a *higher-order* geometry. ❀

Exercise 3.21.* Let $\xi = (P,\omega)$ be a first-order Cartan geometry on M, modeled on $(\mathfrak{g},\mathfrak{h})$ with group H. Fix a basis $(\bar{e}_1,\dots,\bar{e}_n)$ for $\mathfrak{g}/\mathfrak{h}$. Call the bases of $T_x(M)$ of the form $(\varphi_p^{-1}(\bar{e}_1),\dots,\varphi_p^{-1}(\bar{e}_n))$, where $p \in P$ lies over x, *admissible frames*. Let Q be the set of admissible frames over M with right H action given by

$$(\varphi_p^{-1}(\bar{e}_1),\dots,\varphi_p^{-1}(\bar{e}_n)) \cdot h = (\varphi_{ph}^{-1}(\bar{e}_1),\dots,\varphi_{ph}^{-1}(\bar{e}_n)).$$

Show that Q is a principal right H bundle over M and that the map $P \to Q$ sending

$$p \mapsto (\varphi_p^{-1}(\bar{e}_1),\dots,\varphi_p^{-1}(\bar{e}_n))$$

is a bundle isomorphism. ❑

Curvature Function

The curvature form Ω on the principal bundle P determines a certain function K on P called the curvature function.

Definition 3.22. The *curvature function* $K\colon P \to \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{g})$ is defined by the formula

$$K(p)(X_1,X_2) = \Omega_p(\omega_p^{-1}X_1,\omega_p^{-1}X_2).$$

Lemma 3.23. *The curvature function is well defined and satisfies the invariance property*

$$K(ph)(X_1,X_2) = Ad(h^{-1})(K(p)(Ad(h)X_1,Ad(h)X_2)).$$

Proof. First consider p fixed and suppose that $Y_j = X_j + V_j$, for $j = 1,2$, for some $V_j \in \mathfrak{h}$. Then $\Omega_p(\omega_p^{-1}Y_1,\omega_p^{-1}Y_2) = \Omega_p(\omega_p^{-1}X_1,\omega_p^{-1}X_2)$ by Corollary 3.10, since $\omega_p^{-1}V_j$ is tangent to the fiber. Thus, K takes values in

$\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{g})$. The invariance property follows from those of ω and Ω. ■

Exercise 3.24. Let M be a manifold; let (P_i, ω_i), $i = 1, 2$, be two geometrically equivalent Cartan geometries on M with curvature functions K_i, and let $b\colon P_1 \to P_2$ be a geometrical equivalence. Show that $K_2(b(p)) = K_1(p)$. ❑

As in Exercise 3.28 of Chapter 1, we may interpret the curvature function as a section, *the curvature section*, of the vector bundle, for $p = 2$, over M given by $C^p(M) = P \times_H \mathrm{Hom}(\lambda^p(\mathfrak{g}/\mathfrak{h}), \mathfrak{g})$. The action of H on $\mathrm{Hom}(\lambda^p(\mathfrak{g}/\mathfrak{h}), \mathfrak{g})$ is given by $h \cdot \psi = \mathrm{Ad}(h)\psi(\lambda^p(\mathrm{Ad}(h^{-1})))$.

Exercise 3.25. Show that the two bundles $C^p(M)$ associated to two geometrically equivalent Cartan geometries on M are canonically isomorphic and that, for $p = 2$, this isomorphism identifies the two curvature sections. ❑

Exercise 3.26. Show that a Cartan geometry is torsion free if and only if the curvature function takes values in the subrepresentation $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{g})$. ❑

Exercise 3.27. Show that $K(p)(X, Y) = [X, Y] - \omega_p([\omega^{-1}(X), \omega^{-1}(Y)])$. This identity interprets the curvature function as measuring the difference between the Lie algebra bracket and the bracket of the corresponding vector fields on P. ❑

Bianchi Identity

Within the graded Lie algebra $A(P, \mathfrak{g})$ of $\mathfrak{g}$-valued exterior differential forms on *any* manifold P, various identities arise automatically. Here are some of them.

Lemma 3.28. *Let $\alpha, \beta \in A(P, \mathfrak{g})$ be homogeneous*[12] *elements. Then*

$$[\alpha, [\beta, \beta]] = \begin{cases} 2[[\alpha, \beta], \beta] & \text{if } \deg \beta \text{ is odd,} \\ 0 & \text{if } \deg \beta \text{ is even.} \end{cases}$$

Proof. The graded Jacobi identity (Exercise **1**.5.20(iii)) yields

$$(-1)^{ba}[[\beta, \beta], \alpha] + (-1)^{b^2}[[\beta, \alpha], \beta] + (-1)^{ab}[[\alpha, \beta], \beta] = 0,$$

where $a = \deg\ \alpha$ and $b = \deg\ \beta$. By graded commutativity (Exercise **1**.5.20(ii)), this is

[12]That is, each lies entirely within some grade of $A(P, \mathfrak{g})$.

$$(-1)^{ba}(-1)^{2ba+1}[\alpha, [\beta, \beta]] + (-1)^{b^2}(-1)^{ba+1}[[\alpha, \beta], \beta] + (-1)^{ab}[[\alpha, \beta], \beta] = 0,$$

which yields the result. ■

Corollary 3.29. $[\alpha, [\alpha, \alpha]] = 0$ *for every homogeneous element* $\alpha \in A(P, \mathfrak{g})$.

Proof. We may clearly assume that deg α is even, so that

$$\begin{aligned} [\alpha, [\alpha, \alpha]] &= 2[[\alpha, \alpha], \alpha] && \text{(by (3.28))} \\ &= -2[\alpha, [\alpha, \alpha]] && \text{(by graded commutativity).} \end{aligned}$$ ■

The following *Bianchi identity* is a formal consequence of the above identities and hence is true of the "curvature" associated to any element of degree 1 in $A(P, \mathfrak{g})$.[13]

Lemma 3.30 (The Bianchi identity). $d\Omega = [\Omega, \omega]$.

Proof. Taking the exterior derivative of $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$, we get (cf. Exercise **1**.5.20(i)

$$\begin{aligned} d\Omega &= 0 + \frac{1}{2}\{[d\omega, \omega] - [\omega, d\omega]\} \\ &= \frac{1}{2}\{[\Omega - \frac{1}{2}[\omega, \omega], \omega] - [\omega, \Omega - \frac{1}{2}[\omega, \omega]]\} \\ &= \frac{1}{2}\{[\Omega, \omega] - \frac{1}{2}[[\omega, \omega], \omega] - [\omega, \Omega] + \frac{1}{2}[\omega, [\omega, \omega]]\} \\ &= \frac{1}{2}\{[\Omega, \omega] - [\omega, \Omega]\} && \text{(by the corollary 3.29)} \\ &= [\Omega, \omega] && \text{(by graded commutativity).} \end{aligned}$$ ■

Exercise 3.31. Show that $[\Omega, \Omega] + [[\Omega, \omega], \omega] = \frac{1}{2}[\Omega, [\omega, \omega]]$ by taking the exterior derivative of the Bianchi identity. (In fact, there is a whole sequence of derived identities obtained by successive differentiation of the Bianchi identity.) ❑

The original Bianchi identity of Riemannian geometry is usually expressed as *two* identities, called the first and second Bianchi identities. In that case the Lie algebra $\mathfrak{g}$ is the Lie algebra of rigid motions of Euclidean space which decomposes canonically as a direct sum of a translation part and a rotation part. The two classical identities correspond to the two projections. See Chapter 6, section 2, for more details.

[13]More generally, it is true of the curvature associated to any element of degree 1 in any differential graded Lie algebra.

Tensors

Definition 3.32. Let $\xi = (P,\omega)$ be a Cartan geometry on M modeled on $(\mathfrak{g},\mathfrak{h})$ with group H. Let V be a vector space and $\rho: H \to Gl(V)$ a representation. A *tensor of type* (V,ρ) is a function $f: P \to V$ transforming according to the formula $R_h^* f = \rho(h^{-1})f$. ❀

We have already seen several examples of tensors. For example, the curvature function of a Cartan geometry is a tensor of type $(\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{g}), \mathrm{Hom}(\lambda^2(\mathrm{Ad}_{\mathfrak{g}/\mathfrak{h}}),\mathrm{Ad}_{\mathfrak{g}}))$. Tensors of type (V,ρ) may of course be interpreted as sections of the vector bundle $E(\rho) = P \times_H (V,\rho)$ (cf. Corollary 3.16 and Exercise **1**.3.28, for example).

How does a tensor of type (V,ρ) appear in a gauge? The following definition replies to this question.

Definition 3.33. Let (U,θ) be a gauge of a Cartan geometry (P,ω) corresponding to the section $\sigma: U \to P$ (i.e., $\sigma^*\omega = \theta$). If $f: P \to V$ is a tensor of type (V,ρ), then $\phi = f\sigma: U \to V$ is called (*the expression of*) *the tensor in the gauge* (U,θ). ❀

Lemma 3.34. *Let $\phi: U \to V$ be the expression of a tensor of type (V,ρ) in the gauge (U,θ). Let $\theta \Rightarrow_h \theta'$, $h: U \to H$, denote a change of gauge. Then $\rho(h^{-1})\phi: U \to V$ is the expression of the same tensor in the gauge (U,θ').*

Proof. The change of gauge determined by $h: U \to H$ replaces the section $\sigma: U \to P$ by $\sigma h: U \to P$, so that $\phi = f\sigma$ gets replaced by $\phi' = f(\sigma h) = \rho(h^{-1})f(\sigma) = \rho(h^{-1})\phi$. ∎

Universal Covariant Derivative

Suppose V is a vector space and $f \in A^0(P,V)$ (i.e., f is a function on P with values in V). Then the universal covariant derivative $\tilde{D}$ is, roughly speaking, just the derivative of f with respect to the ω-constant vector fields.[14] More precisely, if $X \in \mathfrak{g}$, then

$$\tilde{D}_X f = \omega^{-1}(X)f, \quad \text{so } \tilde{D}_X: A^0(P,V) \to A^0(P,V). \tag{3.35}$$

Since this expression is linear in X, we may regard $\tilde{D}$ itself as the adjoint operator

$$\tilde{D}: A^0(P,V) \to A^0(P, V \otimes \mathfrak{g}^*) \ \text{ defined by } \iota_{X*}(\tilde{D}f) = \tilde{D}_X f.$$

(Note that ι_{X*} is the coefficient homomorphism induced by

[14]E. Cartan studied this notion in the context of projective and conformal geometry in his papers [E. Cartan, 1934, 1935, and 1937].

$$\begin{aligned} \iota_X: V \otimes \mathfrak{g}^* &\to V. \\ v\otimes\eta &\mapsto \eta(X)v \end{aligned}$$

For the details of this, see the subsection concerning change of coefficients in §**1**.5.) The universal covariant derivative is the grandfather of all "geometric" differential operators (i.e., operators with geometric meaning) in a Cartan geometry.[15] For example, as we shall see below, the universal covariant derivative, which exists in any Cartan geometry, gives rise to the usual covariant derivative if the geometry is reductive (cf. Definition 3.42). In Appendix D we show how the classical operators in the plane (i.e., div, curl, and the Cauchy-Riemann operator) arise from it via representation theory.

Suppose that $f: P \to V$ is a tensor of type (V,ρ). It is interesting to inquire after the transformation properties of $\tilde{D}\rho$ to see if it also may be interpreted as a tensor of some type.

Definition 3.36. For any representation $\rho: H \to Gl(V)$, we denote by

$$\begin{aligned} A^q(P,(V,\rho)) \ (&= A^q(P,\rho)) \\ &= \{\eta: \lambda^q(T(P)) \to V \mid R_h^*\eta = \rho(h^{-1})\eta, \quad \text{for all } h \in H\} \end{aligned}$$

the associated space of functions (if $q = 0$) or forms (if $q > 0$). $A^q(P,V,\rho))$ is called the *space of q-forms on P transforming according to the representation ρ*. ❀

(We shall deal mostly with the case of functions in our use of this definition.)

Exercise 3.37. Show that there is a canonical isomorphism $\phi: A^q(P,\rho) \approx A^0(P, \rho \otimes \lambda^q(\mathrm{Ad}^*))$. (Note that an example of this correspondence is given in the case $q = 2$ by $K = \phi(\Omega)$; cf. Definition 3.22.) ❑

Lemma 3.38. $\tilde{D}: A^0(P,\rho) \to A^0(P, \rho \otimes \mathrm{Ad}^*)$.

Proof. Let $f \in A^0(P,\rho)$; then for any $p \in P$, $v \in \mathfrak{g}$, we have

$$\begin{aligned} \iota_{X*}(R_h^*(\tilde{D}f))(p) &= \iota_{X*}((\tilde{D}f)(ph)) \\ &= (\tilde{D}_X f)(ph) \\ &= \omega_{ph}^{-1}(X)f. \end{aligned}$$

Now the equation $R_h^*\omega = \mathrm{Ad}(h^{-1})\omega$ may be read as $\omega_{ph}R_{h*} = \mathrm{Ad}(h^{-1})\omega_p$ for each p, and hence $\omega_{ph}^{-1} = R_{h*}\omega_p^{-1}Ad(h)$. Thus,

[15]Cartan's book [E. Cartan, 1938] (English translation, [E. Cartan, 1966]) is devoted to the study of differential operators in Riemannian and Lorentzian geometries from this point of view. For example, he is able to derive the Dirac operator.

$$\iota_{X*}(R_h^*(\tilde{D}f))(p) = (R_{h*}(\omega_p^{-1}(\mathrm{Ad}(h)X)))f.$$

Since for any vector field Y on P we have

$$(R_{h*}Y)f = f_*(R_{h*}Y) = (fR_h)_*Y = \rho(h^{-1})f_*Y = \rho(h^{-1})Yf,$$

taking $Y = \omega_p^{-1}(\mathrm{Ad}(h)X)$, we get

$$\iota_{X*}(R_h^*(\tilde{D}f))(p) = \rho(h^{-1})(\omega_p^{-1}(\mathrm{Ad}(h)X))f = \rho(h^{-1})\tilde{D}_{\mathrm{Ad}(h)X}f. \qquad \blacksquare$$

Note that even if the representation (V, ρ) is irreducible, the same need not be true of $(V \otimes \mathfrak{g}^*,\ \rho \otimes \mathrm{Ad}^*)$. For example, it may be that the latter decomposes as a sum of several representations $V \otimes \mathfrak{g}^* = W_1 \oplus \ldots \oplus W_k$. In that case $\tilde{D}$ may be broken up correspondingly as a sum $\tilde{D} = \tilde{D}_1 + \cdots + \tilde{D}_k$ of several first-order differential operators where the $\tilde{D}_j$ are the projections of $\tilde{D}$ on the various summands. This is the way in which the Cauchy–Riemann operator, div, and curl are obtained in Appendix D.

The following result calculates $\tilde{D}f$ in the fiber direction.

Lemma 3.39. *For $X \in \mathfrak{h}$ and $f \in A^0(P, \rho)$, we have $\iota_{X*}(\tilde{D}f) = -\rho_*(X)f$, where $\rho_*: \mathfrak{h} \to \mathrm{End}(V)$ is the derivative at the identity of the representation $\rho: H \to Gl(V)$.*

Proof.

$$\begin{aligned}(\iota_{X*}(\tilde{D}f))(p) &= \omega_p^{-1}(X)f = f_*(\omega_p^{-1}(X)) \\ &= \left.\frac{d}{dt}\right|_{t=0} f(p\exp(tX)) = \left.\frac{d}{dt}\right|_{t=0} \rho(\exp(-tX))f(p) \\ &= -\rho_*(X)f(p). \qquad \blacksquare\end{aligned}$$

Exercise 3.40. Exercise 3.37 allows us to extend the definition of the universal covariant derivative to forms on P with values in (V, ρ) to yield a map $\tilde{D}: A^q(P, \rho) \to A^q(P, \rho \otimes \mathfrak{g}^*)$. Suppose η is any q form on P. Show $\tilde{D}\eta = 0$ if and only if η may be expressed as a linear combination, with constant coefficients, of exterior products of the components of the Cartan connection ω with respect to some basis of $\mathfrak{g}$. ❑

Because of the canonical identification $A^0(P, \rho) = A^0(M, E)$, where E is the vector bundle $P \times_H (V, \rho)$, the universal covariant derivative associated to (V, ρ) may equivalently be regarded as a linear first-order differential operator $\tilde{D}: A^0(M, E) \to A^0(M, F)$, where $F = P \times_H (V \otimes \mathfrak{g}^*,\ \rho \otimes \mathrm{Ad}^*)$. We note that this operator gives the derivatives of a section of E not in a direction given by a tangent vector of M but in a direction given by a tangent vector of P. Another way to say this is that to describe how a section changes from point to point it is not enough to give the two nearby points of M; we must give two nearby frames of M (infinitesimal *meteors* in the parlance of §4.6).

Exercise 3.41. Show that, if P is connected, then

$$\ker(\tilde{D}: A^0(M, E) \to A^0(M, F)) = V^H$$

where

$$V^H = \{v \in V \mid \rho(h)v = v \quad \text{for all } h \in H\}. \qquad ❑$$

This exercise shows that, apart from the constant functions, there are no tensors on M that are "generalized covariant constant" functions.

Finally, we remark that if the representation $V \otimes \mathfrak{g}^*$ decomposes as a sum of representations $V \otimes \mathfrak{g}^* = W_1 \oplus \cdots \oplus W_k$, then the bundle F decomposes correspondingly as a sum $F = F_1 \oplus \cdots \oplus F_k$, where $F_i = P \times_H W_i$. Moreover, the constituent operators in the decomposition $\tilde{D} = \tilde{D}_1 + \cdots + \tilde{D}_k$ may be regarded as first-order linear differential operators $\tilde{D}_k: A^p(M, E) \to A^p(M, F_i)$.

Covariant Derivative in a Reductive Geometry

The covariant derivative D (see Definition 3.47) generalizes the vector gradient in Euclidean space (see Exercise 3.50). However, the existence of a covariant derivative requires that the geometry be reductive.[16]

Definition 3.42. A Cartan geometry modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H is *reductive* if there is an H module decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ (cf. Definition 4.3.2).[17] ❀

Let us assume now that the geometry we are considering is reductive in this sense. Any form with values in $\mathfrak{g}$ will decompose into an $\mathfrak{h}$ component and a $\mathfrak{p}$ component. This is true in particular for a Cartan connection $(\omega = \omega_{\mathfrak{h}} + \omega_{\mathfrak{p}})$ and any gauge $(\theta = \theta_{\mathfrak{h}} + \theta_{\mathfrak{p}})$. In addition, the universal covariant derivative also decomposes as $\tilde{D}_X = \tilde{D}_{\mathfrak{h}X} + \tilde{D}_{\mathfrak{p}X}$.

By Lemma 3.39, for $X \in \mathfrak{h}$ and $f \in A^0(P, \rho)$, $\iota_{X*}(\tilde{D}f) = -\rho(X)f$, namely, $\tilde{D}_{\mathfrak{h}} = -\rho$. That is, $\tilde{D}_{\mathfrak{h}}$ merely tells us how f transforms under

[16]The reductive geometries have, therefore, a much richer structure than the nonreductive ones.

[17]This use of the term *reductive* conflicts with another use of it in the theory of Lie algebras, which calls a Lie algebra reductive if it is the sum of a semisimple ideal and an abelian ideal. The two notions are loosely related in the following sense. Given a Klein pair $(\mathfrak{g}, \mathfrak{h})$ for which the representation $\mathrm{ad}: \mathfrak{h} \to \mathfrak{gl}(\mathfrak{g}/\mathfrak{h})$ is injective and $\mathfrak{h}$ is reductive in the Lie algebra sense, then $(\mathfrak{g}, \mathfrak{h})$ is reductive in the sense of Definition 3.42 (cf., e.g., [J. Humphreys, 1972], pp. 31 and 102, and [W. Fulton and J. Harris, 1991], p. 131).

H, which we already know, so this projection is not very interesting. The component $\tilde{D}_{\mathfrak{p}}$, on the other hand, is very interesting.

Definition 3.43. In a reductive geometry the operator $\tilde{D}_{\mathfrak{p}}$ is called (the *bundle version* of) the *covariant derivative*. A function $f: P \to V$ is called *covariant constant* or *parallel* if $\tilde{D}_{\mathfrak{p}} f = 0$. ❀

Since sections of geometric vector bundles over M can be interpreted as functions on P, it follows that this notion of parallel applies also to these sections. We wish now to go into more detail on this point in order to reinterpret the covariant derivative (as well as the notion of parallel) at the level of sections. The key idea is to use the reductive hypothesis again to get a canonical way of lifting a vector field on M to a vector field on P.

Definition 3.44. The distribution on P given by $\omega_{\mathfrak{h}} = 0$ is called the *horizontal distribution* and vectors in it are called *horizontal vectors.* ❀

Note that a function $f: P \to V$ is parallel if and only if it is constant in the horizontal directions. We now relate these horizontal directions to the directions in the base M.

Lemma 3.45. *The projection $\pi: P \to M$ induces an isomorphism between the vector space of horizontal vectors at $p \in P$ and $T_{\pi(p)}(M)$.*

Proof. From the exactness of the columns in the following diagram and from its commutativity

$$\begin{tikzcd}
T_p(pH) \arrow[r, "\omega_H", "\approx"'] \arrow[d] & \mathfrak{h} \arrow[d, phantom, "\cap"] \\
T_p(P) \arrow[r, "\omega", "\approx"'] \arrow[d] & \mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p} \arrow[d, "\rho"] \\
T_x(M) \arrow[r, "\varphi_p", "\approx"'] & \mathfrak{g}/\mathfrak{h}
\end{tikzcd}$$

(see the proof of Theorem 3.15), the lemma is immediate. ∎

Definition 3.46. Let (P, ω) be a reductive Cartan geometry on M. If X is a vector field on M, then the *horizontal lift* of X, denoted by $\tilde{X}$, is the (unique) vector field on M such that $\pi_{*p}(\tilde{X}) = X_{\pi(p)}$ for all $p \in P$ and $\omega_{\mathfrak{h}}(\tilde{X}) = 0$. ❀

Definition 3.47. Let (P, ω) be a reductive Cartan geometry on M, $E = P \times_H (V, \rho)$, and $X \in \Gamma(T(M))$. Let $\psi: \Gamma(E) \approx A^0(P, \rho)$ be the usual identification of sections of vector bundles associated to P with functions on P (cf. Exercise 1.3.28). The (*base version of the*) *covariant derivative* $D_X: \Gamma(E) \to \Gamma(E)$ is defined by the equation $\psi(D_X f) = \tilde{X}(\psi(f))$, where $f \in \Gamma(E)$. ❀

Proposition 3.48. *Let $k: M \to \mathbf{R}$, $X, Y \in \Gamma(T(M))$, and $f, g \in \Gamma(E)$; then the covariant derivative has the following properties:*

(i) $D_X(f+g) = D_X f + D_X g$;

(ii) $D_{X+Y} f = D_X f + D_X f$;

(iii) $D_{kX} f = k D_X f$;

(iv) $D_X(kf) = X(k) f + k D_X f$.

Proof. (i) and (ii) are straightforward from the $\mathbf{R}$–linearity of ψ and of the $\mathbf{R}$–bilinearity of the derivation $\tilde{X}$. For (iii) we note that $\pi^*(k)\tilde{X}$ is the horizontal lift of kX, where $\pi: P \to M$ is the canonical projection. Thus,

$$\psi(D_{kX} f) = \tilde{D}_{\pi^*(k)\tilde{X}}(\psi(f)) = \pi^*(k)\tilde{D}_{\tilde{X}}(\psi(f)) = \pi^*(k)\psi(D_X f) = \psi(k D_X f)$$

whence (iii). Finally,

$$\begin{aligned}
\psi(D_X(kf)) &= \tilde{X}(\psi(kf)) \\
&= \tilde{X}(\pi^*(k)\psi(f)) \\
&= \tilde{X}(\pi^*(k))\psi(f) + \pi^*(k)\tilde{X}(\psi(f)) \\
&= (\pi^* X(k))\psi(f) + \pi^*(k)\psi(D_X f) \\
&= \psi(X(k) f) + \psi(k D_X f),
\end{aligned}$$

which, by the linearity and injectivity of ψ, yields (iv). ∎

Although we see some of the properties of D from Proposition 3.48, it is still not clear how to calculate it in terms of a gauge. The following result remedies this.

Proposition 3.49. *Let*

$$\begin{cases}
(U, \theta) \text{ be a gauge for a reductive Cartan geometry,} \\
X \text{ be a vector field on } U, \\
\phi \text{ be the expression in the gauge } (U, \theta) \text{ of a tensor of type } (V, \rho).
\end{cases}$$

Then $D_X \phi = X(\phi) - \rho_*(\theta_{\mathfrak{h}}(X))\psi$ is the expression in the gauge (U, θ) of the covariant derivative of the tensor expressed by ϕ in the gauge (U, θ).

Proof. Let $\sigma: U \to P_U = \pi^{-1}(U)$ be the section corresponding to the gauge (U, θ) and let $\Phi: P \to V$ be the tensor of type (V, ρ) of which ϕ is the expression. If $\sigma_*(X)$ were horizontal, then our job would be easy since then it would be the horizontal lift of X and we would have

$$\begin{aligned}(D_X\phi)(x) &= (D_X\Phi)\sigma(x) = \tilde{X}_{\sigma(x)}(\Phi) = \Phi_*(\tilde{X}_{\sigma(x)}) = \Phi_*(\sigma_*(X_x)) \\ &= (\Phi\circ\sigma)_*(X_x) = \phi_*(X_x) = X_x(\phi).\end{aligned}$$

This does indeed yield the formula claimed when $\sigma_*(X)$ is horizontal since in that case the expression $\theta_{\mathfrak{h}}(X)$ vanishes. The general case will come from making a change of gauge so that $\sigma'_*(X)$ is horizontal in the new gauge σ'. We have $\theta' = \mathrm{Ad}(h^{-1})\theta + h^*(\omega_H)$. Applying this to X and taking the $\mathfrak{h}$ components yields

$$\theta'_{\mathfrak{h}}(X) = \mathrm{Ad}(h^{-1})\theta_{\mathfrak{h}}(X) + \omega_H(h_*X).$$

We may choose $h\colon U \to H$ so that at a fixed but arbitrarily chosen point x we have $h(x) = e$ and $h_*X = -\theta_{\mathfrak{h}}(X)$. Then, at x, we have

$$\begin{aligned}D_X\phi' &= X(\phi') = X(\rho(h^{-1})\phi) = X(\rho(h^{-1}))\phi + \rho(h^{-1})X(\phi) \\ &= -(\rho(h)_*(X))\phi + X(\phi).\end{aligned}$$

The left- (and therefore right-) hand side of this equation is $D_X\Phi$ expressed in the gauge (U, θ'). By Lemma 3.34, and because $h(x) = e$, this is the same, at x, as the expression for $D_X\Phi$ in the gauge (U, θ). ■

Exercise 3.50. Show that in Euclidean space the covariant derivative of a vector field is just the usual vector gradient. [*Hint*: Apply Proposition 3.49 using the gauge given by $\theta(e_i) = \begin{pmatrix} 0 & 0 \\ e_i & 0 \end{pmatrix} \in \mathfrak{euc}_n(\mathbf{R})$, where e_i is the standard column vector.] ❑

For simplicity, we have expressed the covariant derivative as a map $D_X\colon \Gamma(E) \to \Gamma(E)$, where X is a vector field on X. However, we note that this operator is local in the sense that the value of D_Xf at x depends only on the value of X at x and the behavior of f on a neighborhood of x in M. (In fact, only the zero- and first-order terms of the Taylor series of f at x play a role in the contribution of f to this formula. Cartan would say that the formula depends only on the values of f in a *first-order neighborhood* of x in M.) These and similar facts may be invoked to justify the following definition and exercise.

Definition 3.51. A section $\eta \in \Gamma(E(\rho))$ is called *parallel* along a curve $\sigma\colon I \to M$ if $D_{\sigma(t)}\eta = 0$ for every $t \in I$. ❀

Exercise 3.52. Show that a function $f\colon P \to (V, \rho)$ is parallel in the sense of Definition 3.43 if and only if the corresponding section $\eta \in \Gamma(E(\rho))$ is parallel along every curve in M. ❑

Special Geometries

In general, the values of the curvature form will span, as a vector space, the whole of the Lie algebra $\mathfrak{g}$. In special cases, however, this span may not be the whole of $\mathfrak{g}$. For example, in the case of a torsion free geometry, the span will lie in $\mathfrak{h}$. However, as the following lemma shows, the span cannot be an arbitrary vector subspace of $\mathfrak{g}$.

Lemma 3.53. *Let $V \subset \mathfrak{g}$ be the vector subspace spanned by the values of the curvature form Ω. Then V is an H submodule of $\mathfrak{g}$.*

Proof. It suffices to show that the set of values of Ω is stable under the adjoint action of H. Let $v = \Omega_p(X_p, Y_p)$. Then $\mathrm{Ad}(h^{-1})v = \mathrm{Ad}(h^{-1})(\Omega_p(X_p, Y_p)) = (R_h^*\Omega_p)(X_p, Y_p) = \Omega_{ph}(R_{h*}X_p, R_{h*}Y_p) =$ a value of Ω. ■

In particular, if $V \subset \mathfrak{h}$, so that the geometry is torsion free, then V is an ideal in $\mathfrak{h}$.

Definition 3.54. Let $V \subset \mathfrak{g}$ be an H submodule. A Cartan geometry of $\mathfrak{g}$-*curvature type* V is a geometry whose curvature form takes values in V. ❀

Assume now that our geometry is torsion free and the adjoint action of H on $\mathfrak{h}$ is irreducible. In this case, there are no special geometries arising from $\mathfrak{g}$ curvature-type conditions. Nevertheless, the representation of H on $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ may have nontrivial submodules V so that one may distinguish various cases according to how K relates to these submodules.

The following definition carries the same idea as Definition 3.54 but the context is different.

Definition 3.55. Let $V \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ be an H submodule. A Cartan geometry of *curvature type* V is a geometry whose curvature function K takes values in V. ❀

By Exercise 3.4.8(b), $\mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is a submodule and, for H connected, it is the submodule of H invariant elements. If the submodule is nontrivial, this leads to a class of special geometries, the constant-curvature geometries, studied in §4.

When H is a compact group, it is known from the representation theory of Lie groups that $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ decomposes as a direct sum of irreducible submodules. In this case $C^2(M)$ and its section K will split up correspondingly, and we may speak of the various "component curvatures" associated to K. For example, in the general Riemannian case $C^2(M)$ splits into three pieces corresponding to the scalar, the Ricci, and the Weyl curvatures (cf. Table 2.5 on page 236). In the special case when M has dimension 4, the Weyl curvature splits up further into a self-dual and an anti-self-dual part.

When H is noncompact, even though $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h})$ may not in general decompose into irreducibles, it may still have nontrivial submodules. For example, the composite mapping

$$\begin{aligned}\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h}) &\approx \lambda^2(\mathfrak{g}/\mathfrak{h})^*\otimes\mathfrak{h} \xrightarrow{id\otimes\mathrm{ad}} \lambda^2(\mathfrak{g}/\mathfrak{h})^*\otimes\mathrm{End}(\mathfrak{g}/\mathfrak{h})\\ &\approx \underset{t^*\wedge u^*\otimes v\otimes w^*}{\lambda^2(\mathfrak{g}/\mathfrak{h})^*\otimes\mathfrak{g}/\mathfrak{h}\otimes(\mathfrak{g}/\mathfrak{h})^*} \underset{\mapsto}{\to} \underset{(t^*(v)u^*-u^*(v)t^*)\otimes w^*}{(\mathfrak{g}/\mathfrak{h})^*\otimes(\mathfrak{g}/\mathfrak{h})^*}\end{aligned}$$

is an H module homomorphism, and so its kernel is a submodule.

Definition 3.56. The kernel of the composite mapping above is called the *normal submodule of* $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h})$. ❀

Exercise 3.57. Show that if $\mathrm{End}(\mathfrak{g}/\mathfrak{h})$ is given the H module structure $h\cdot\phi = \mathrm{Ad}(h)\phi\mathrm{Ad}(h^{-1})$, then the map $\mathrm{ad}\colon\mathfrak{h}\to\mathrm{End}(\mathfrak{g}/\mathfrak{h})$ and the canonical map $\mathrm{End}(\mathfrak{g}/\mathfrak{h}) \to \mathfrak{g}/\mathfrak{h}\otimes(\mathfrak{g}/\mathfrak{h})^*$ are H module maps. Use these facts to verify that the composite mapping above is an H module map. ❀

The normal submodule will be useful in defining *normal* geometries. In some cases—for example, in conformal geometries—"normal" means that K takes values in the normal submodule. In other cases, such as projective geometry, it means that K takes values in a submodule *analogous* to the normal submodule. In general, however, a normal geometry is defined in a somewhat ad hoc manner so that the Cartan geometry will be uniquely determined by a given set of geometric data that may seem at the outset to have no necessary connection with a Cartan geometry at all but determines one via Cartan's *method of equivalence*. Several examples of this are given in Chapters 6, 7, and 8 on specific geometries.

§4. Development, Geometric Orientation, and Holonomy

In this section we study properties related to paths in a Cartan geometry. In particular, we introduce the notion of geometric orientation[18] in this context. Its importance for us is in connection with the classification of locally Klein geometries given in Theorem 5.4. Then we discuss issues connected with the development of paths in a Cartan geometry as paths on the model space. In the last part of this section, we apply the notion of development to introduce and to discuss some issues surrounding the holonomy group of a Cartan geometry.

[18]Our notion of a geometric orientation differs from the notion of the same name discussed in [E. Cartan, 1941].

Recall the notion of development given in Definition 3.7.4 in connection with the global version of the fundamental theorem. In that case we assumed we had

$$\begin{cases}\text{a Lie group } G \text{ with Lie algebra } \mathfrak{g},\\ \text{a manifold } P \text{ equipped with a } \mathfrak{g}\text{-valued 1-form } \omega,\\ \text{a piecewise smooth path } \sigma\colon(I,a,b)\to(P,p,q), \text{ where } I=[a,b].\end{cases}$$

We showed that, given $g\in G$, there is always a unique path $\tilde\sigma\colon(I,a)\to(G,g)$ such that $\tilde\sigma^*\omega_G=\sigma^*\omega$. This path on G was called the development of ω starting at g.

Recall that $\mathrm{ad}\,\omega$ takes values in the Lie algebra $\mathrm{End}(\mathfrak{g})$ of $Gl(\mathfrak{g})$ (cf. Exercise 3.4.9).

Lemma 4.1. *Let λ be a path on P and let*

(i) *$\tilde\lambda\colon(I,a)\to(G,e)$ be the development of λ via ω,*

(ii) *$\hat\lambda\colon(I,a)\to(Gl(\mathfrak{g}),e)$ be the development of λ via $\mathrm{ad}\,\omega$.*

Then $\hat\lambda(t)=\mathrm{Ad}(\tilde\lambda(t))$ for $t\in I$.

Proof. Since, by Proposition 3.1.8, the homomorphism $\mathrm{Ad}\colon G\to Gl(\mathfrak{g})$ has the property that $\mathrm{Ad}^*\omega_{Gl(\mathfrak{g})}=\mathrm{Ad}_{*e}\omega_G$, it follows that

$$\begin{aligned}\mathrm{Ad}(\tilde\lambda)^*\omega_{Gl(\mathfrak{g})} &= \tilde\lambda^*\mathrm{Ad}^*\omega_{Gl(\mathfrak{g})} = \tilde\lambda^*\mathrm{Ad}_{*e}\omega_G = \mathrm{Ad}_{*e}(\tilde\lambda^*\omega_G)\\ &= \mathrm{Ad}_{*e}(\lambda^*\omega) = \lambda^*\mathrm{Ad}_{*e}(\omega) = \lambda^*\mathrm{ad}\,\omega = \hat\lambda^*\omega_{Gl(\mathfrak{g})}.\end{aligned}$$

Since $\hat\lambda(0)=e=\mathrm{Ad}(\tilde\lambda(0))$, it follows that $\mathrm{Ad}(\tilde\lambda)=\hat\lambda$. ■

Geometric Orientation

Let $\xi=(P,\omega)$ be a Cartan geometry on a manifold M with model $(\mathfrak{g},\mathfrak{h})$ and (not necessarily connected) group H. We approach the notion of a geometric orientation for ξ indirectly through the notion of the geometric orientation-preserving subgroup of H.

Definition 4.2. Fix a point $p\in P$ lying over x. An element $h\in H$ is called *geometrically orientation preserving* with respect to the base point p if there is a path λ from $p\in P$ to $ph\in P$ whose development, via ad ω, yields a path $\hat\lambda$ on $Gl(\mathfrak{g})$ joining the identity e to $\mathrm{Ad}(h)$. The set of geometrically orientation-preserving elements of H is denoted by H_{or} (cf. Propositions 4.4 and 4.5).

We will need the following result.

Lemma 4.3. *Let $p, q \in P$, let σ be a path joining p to q, and let $\hat{\sigma}: (I, a) \to (Gl(\mathfrak{g}), e)$ be its development, via ad ω, on $Gl(\mathfrak{g})$. We claim that, for any $h \in H$, the development of $R_h\sigma$ on $Gl(\mathfrak{g})$, via ad ω, and starting at the identity, is the composite*

$$(I, 0) \xrightarrow{\hat{\sigma}} (Gl(\mathfrak{g}), e) \underset{g \mapsto \mathrm{Ad}(h)^{-1} g \mathrm{Ad}(h)}{\xrightarrow{\mathbf{Ad}(\mathrm{Ad}(h)^{-1})}} (Gl(\mathfrak{g}), e).$$

Proof. We calculate

$$\begin{aligned}
(\mathbf{Ad}(\mathrm{Ad}(h)^{-1}) \circ \hat{\sigma})^* \omega_{Gl(\mathfrak{g})} &= \hat{\sigma}(\mathbf{Ad}(\mathrm{Ad}(h)^{-1}))^* \omega_{Gl(\mathfrak{g})} \\
&= \hat{\sigma}^*(\mathrm{Ad}(\mathrm{Ad}(h)^{-1}))\omega_{Gl(\mathfrak{g})} \quad (\text{by } \mathbf{3}.1.8) \\
&= \mathrm{Ad}(\mathrm{Ad}(h^{-1}))\hat{\sigma}^* \omega_{Gl(\mathfrak{g})} = \mathrm{Ad}(\mathrm{Ad}(h^{-1}))\sigma^*(\mathrm{ad}\,\omega) \\
&\qquad (\text{by definition of } \hat{\sigma}) \\
&= \sigma^*(\mathrm{Ad}(\mathrm{Ad}(h^{-1}))\mathrm{ad}\,\omega) = \sigma^*(\mathrm{ad}(\mathrm{Ad}(h^{-1})\omega)) \quad (\text{by } \mathbf{3}.3.5(\mathrm{ii})) \\
&= \sigma^*(\mathrm{ad}(R_h^*\omega)) = \sigma^* R_h^* \mathrm{ad}(\omega) \\
&= (R_h\sigma)^* \mathrm{ad}(\omega).
\end{aligned}$$

By the uniqueness part of the fundamental theorem of calculus ($\mathbf{3}$.5.2), $\mathbf{Ad}(Ad(h)^{-1}) \circ \hat{\sigma}$ is the development of $R_h\sigma$. ■

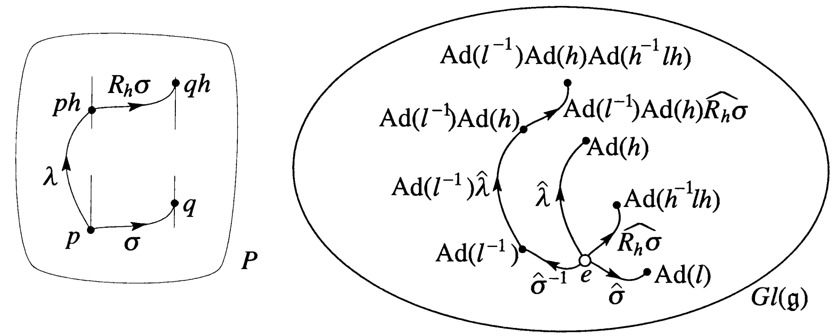
Proposition 4.4. *For P connected, the subset $H_{or} \subset H$ of elements preserving the geometric orientation does not depend on the choice of p.*

Proof. Fix $p, q \in P$. Let $h \in H_{or}$ preserve the geometric orientation with respect to p. We wish to show that it also preserves the geometric orientation with respect to q.

Let λ be a path joining p to ph and σ be a path joining p to q. Let the developments of λ and σ on $Gl(\mathfrak{g})$, via ad ω, and starting at the identity, be denoted by $\hat{\lambda}: (I, a, b) \to (Gl(\mathfrak{g}), e, \mathrm{Ad}(h))$ and $\hat{\sigma}: (I, a, b) \to (Gl(\mathfrak{g}), e, \mathrm{Ad}(\Gamma^1))$, respectively.

By Lemma 4.3, the development of $R_h\sigma$, via ad ω, and starting at the identity, ends at $\mathbf{Ad}(\mathrm{Ad}(h)^{-1})\mathrm{Ad}(l) = \mathrm{Ad}(h^{-1}lh)$.

Using this fact, the following diagram, and Exercise $\mathbf{3}$.7.5, we see the path $\sigma^{-1} \star \lambda \star (R_h\sigma)$ develops, via ad ω_G, to a path on $Gl(\mathfrak{g})$ joining e to $\mathrm{Ad}(\Gamma^1)\mathrm{Ad}(h) \cdot \mathrm{Ad}(h^{-1}lh) = \mathrm{Ad}(h)$.



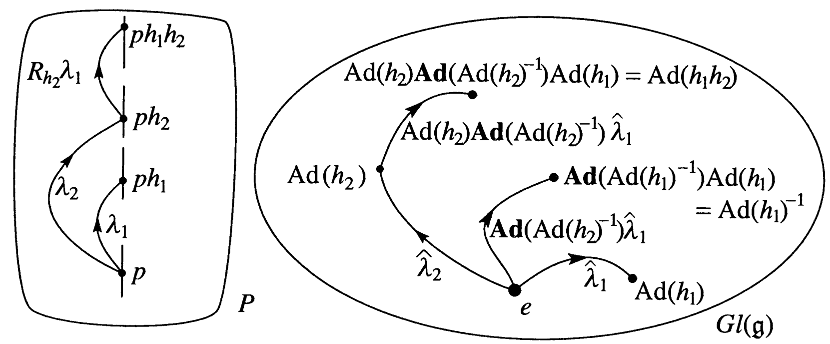
This shows that h preserves the geometric orientation with respect to q. ■

Proposition 4.5. *For P connected, we have $H_e \subset H_{or} \lhd H$.*

Proof. If h lies in the identity component H_e of H, then there is a path $h(t)$ on H with $h(0) = e$ and $h(1) = h$. This path then yields a path $ph(t)$ on P joining p and ph. Since this path lies on the fiber pH, where ω restricts to the Maurer–Cartan form of H, it follows that $ph(t)$ develops to $h(t)$ on H and hence to $\mathrm{Ad}(h(t))$ on $Gl(\mathfrak{g})$. Thus it joins e to $\mathrm{Ad}(h)$, showing that $h \in H_{or}$.

Next we show that H_{or} is a subgroup of H by showing that it is closed under multiplication and inverse.

If $h_1, h_2 \in H_{or}$, then there are paths λ_1 and λ_2 joining p to ph_1 and ph_2, respectively, whose developments $\hat{\lambda}_1$ and $\hat{\lambda}_2$ on $Gl(\mathfrak{g})$, via ad ω, join the identity to $\mathrm{Ad}(h_1)$ and $\mathrm{Ad}(h_2)$, respectively. Moreover, by Lemma 4.3, $R_{h_2}\lambda_1$ develops to $\mathbf{Ad}(\mathrm{Ad}(h_2)^{-1})\hat{\lambda}_1$.



Then the path $\lambda_2 \star (R_{h_2}\lambda_1)$ joins p to ph_1h_2 and develops, via ad ω, to the path $\hat{\lambda}_2 \star \mathrm{Ad}(h_2)\mathbf{Ad}(\mathrm{Ad}(h_2)^{-1})\hat{\lambda}_1$ on $Gl(\mathfrak{g})$ joining e to $\mathrm{Ad}(h_1h_2)$. Thus, $h_1h_2 \in H_{or}$.

If $h \in H_{or}$, then there is a path λ joining p to ph whose development $\hat{\lambda}$ on $Gl(\mathfrak{g})$ via ad ω, joins the identity to $\mathrm{Ad}(h)$. Then λ^{-1} is a path joining

q $(= ph)$ to qh^{-1} $(= p)$, and the development of λ^{-1}, via ad ω, starting at the identity on $Gl(\mathfrak{g})$ is (cf. Exercise **3**.7.5) $\mathrm{Ad}(h)^{-1}\hat{\lambda}^{-1}$, which ends at $\mathrm{Ad}(h)^{-1}$. Thus, $h^{-1} \in H_{or}$.

To show that $H_{or} \subset H$ is normal, let $h \in H_{or}$ and $k \in H$. Choose a path λ, joining p to ph, so that it develops, via ad ω, to a path $\hat{\lambda}$ on $Gl(\mathfrak{g})$ joining e to $\mathrm{Ad}(h)$. By Lemma 4.3, $R_k\lambda$, which joins pk to $phk = pk(k^{-1}hk)$, develops, via ad ω, to the composite

$$I, 0 \xrightarrow{\hat{\lambda}} Gl(\mathfrak{g}), e \xrightarrow{\mathbf{Ad}(\mathrm{Ad}(k)^{-1})} Gl(\mathfrak{g}), e$$

which ends at $\mathbf{Ad}(\mathrm{Ad}(k)^{-1})\mathrm{Ad}(h) = \mathrm{Ad}(k^{-1}hk)$. Thus, $k^{-1}hk \in H_{or}$. ■

While in general it seems difficult to describe the geometric orientation-preserving subgroup H_{or} of a Cartan geometry with more precision than that given in Proposition 4.5, there is an important case in which $H_{or} = H_e$.

Proposition 4.6. *Let (P, ω) be a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Let G be a Lie group realizing $\mathfrak{g}$ and containing H as a subgroup. Assume that*

(i) *$G_e \cap H = H_e$, and*

(ii) *$\mathrm{Ad}_{\mathfrak{g}}: G \to Gl(\mathfrak{g})$ is injective.*

Then $H_{or} = H_e$.

Proof. Since $H_{or} \supset H_e$ by Proposition 4.5, it suffices to show the reverse inclusion. Let $h \in H_{or}$ and fix $p \in P$. Then there is a path λ joining p to ph whose development, via ad ω, yields a path $\hat{\lambda}$ on $Gl(\mathfrak{g})$ joining the identity e to $\mathrm{Ad}_{\mathfrak{g}}(h)$. We can also develop λ via ω, to yield a path $\tilde{\lambda}$ on G joining the identity e to an element $g \in G$. By Lemma 4.1, $\mathrm{Ad}_{\mathfrak{g}}(\tilde{\lambda}) = \hat{\lambda}$ so, in particular, $\mathrm{Ad}_{\mathfrak{g}}(g) = \mathrm{Ad}_{\mathfrak{g}}(h)$. Thus, $g = h$, since $\mathrm{Ad}_{\mathfrak{g}}$ is injective. It follows that $h \in G_e \cap H = H_e$ and hence $H_{or} \subset H_e$. ■

Corollary 4.7. *Suppose that H is a Lie group and $\rho: H \to Gl(V)$ is an injective homomorphism. Regarding V as an abelian Lie algebra, we set $\mathfrak{g} = \mathfrak{h} \times_\rho V$ (i.e., $[h, v] = \rho_*(h)v$ for $h \in \mathfrak{h}$, $v \in V$). Let (P, ω) be a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Then $H_{or} = H_e$ whenever P is connected.*

Proof. First note that $\mathfrak{g}$ is the Lie algebra of the Lie group $G = H \times_\rho V$. Clearly, $H \subset G$. By the proposition, it suffices to verify conditions (i) and (ii).

(i) Let $(h, v) \in H \times_\rho V$ and $(k, w) \in \mathfrak{h} \times_\rho V$. We have

$$\mathrm{Ad}(h, v)(k, w) = (\mathrm{ad}(h)k, \rho(h)(w - \rho_*(k)v)).$$

Then

$$\mathrm{Ad}(h, v) = \mathrm{id}_{\mathfrak{g}} \Rightarrow \begin{cases} \mathrm{Ad}(h)k = k, \\ \rho(h)w = w, \\ \rho(h)\rho_*(k)v = 0, \end{cases} \quad \text{for all } (k, w) \in \mathfrak{g} \Rightarrow \begin{cases} \mathrm{Ad}(h) = \mathrm{id}_{\mathfrak{h}}, \\ \rho(h) = \mathrm{id}_V, \\ v = 0, \end{cases}$$

and so $\mathrm{Ad}_{\mathfrak{g}}: G \to Gl(\mathfrak{g})$ is injective because ρ is injective.

(ii) As a topological space, G is just the topological product of H and V. Since V is contractible, it follows that $G_e = H_e \times V$ and hence $G_e \cap H = H_e$. ■

Definition 4.8. Let $\xi = (P, \omega)$ be a Cartan geometry on a manifold M with model $(\mathfrak{g}, \mathfrak{h})$ and (not necessarily connected) group H. ξ is called *geometrically oriented* if $H_{\mathrm{or}} = H$. The principal covering $M_{\mathrm{or}} = P/H_{\mathrm{or}} \to M$ (with group H/H_{or}) is called the *geometric orientation cover* of M. ξ is called *geometrically orientable* if there is a reduction of the bundle $P \to M$ to the subgroup H_{or}. A *geometric orientation* is the choice of such a reduction. ❀

Exercise 4.9. Show that a Cartan geometry $\xi = (P, \omega)$ on M is geometrically orientiable if and only if the principal covering space $M_{\mathrm{or}} \to M$ is trivial. Show that the geometric orientation cover of M is geometrically oriented. ❑

Proposition 4.10. *Locally Klein geometries are geometrically oriented.*

Proof. Consider a locally Klein geometry $M = \Gamma \setminus G/H$. As a Cartan geometry, this has principal bundle $P = \Gamma \setminus G$ and Cartan connection $\omega = \omega_{\Gamma\setminus G}$. Fix the point $p = \Gamma e \in P = \Gamma \setminus G$. Let $h \in H$ be arbitrary. Since G is connected, we may choose a path $\tilde{\lambda}: (I, 0, 1) \to (G, e, h)$. Let its projection to P be $\lambda: (I, 0, 1) \to (P, p, ph)$. Since $\pi^*\omega_{\Gamma\setminus G} = \omega_G$, it follows that the development of λ on G via $\lambda^*\omega_{\Gamma\setminus G}$ is the same as the development of $\tilde{\lambda}$ on G via ω_G, which is just $\tilde{\lambda}$ itself, which ends at h. Thus, the development of λ on $Gl(\mathfrak{g})$ via ad $\omega_{\Gamma\setminus G}$ is the development of $\tilde{\lambda}$ on $Gl(\mathfrak{g})$ via ad ω_G, which is $\hat{\lambda} = \mathrm{Ad}(\tilde{\lambda})$ ending at $\mathrm{Ad}(h)$. Thus, $h \in H_{\mathrm{or}}$ for any $h \in H$. It follows that locally Klein geometries are geometrically oriented. ■

Lemma 4.11. *The Cartan geometry $\xi = (P, \omega)$ on M_{or} with model $(\mathfrak{g}, \mathfrak{h})$ and group H_{or} is orientable with a canonical orientation.*

Proof. By the very definition of H_{or}, the geometry is geometrically orientable and is itself a geometric orientation.

Development of Paths on M

Let $\xi = (P, \omega)$ be a Cartan geometry on M modelled on the Klein geometry (G, H). Here we are ultimately interested in the development of a path on M as a path on G/H. However, we begin by considering again the development of paths on P. If ξ is flat, so that the structural equation holds for ω then the discussion of monodromy given in §7 of Chapter 3 applies directly. In the case of a general Cartan geometry the structural equation fails so the development of a path on P depends on more than just the homotopy class of this path. However, there is a remnant of the structural equation still in force. According to Lemma 3.13, on each fiber pH of P the Cartan connection can be identified with the Maurer–Cartan form of H. This means that the structural equation continues to hold in the fiber directions even in the most general Cartan geometries. The effect of this will be to maintain some homotopy invariance in the fiber direction in the development construction. Our study of this phenomenon will be based on the following observation.

Lemma 4.12. *Let $\xi = (P, \omega)$ be a Cartan geometry modeled on the Klein geometry (G, H). Let*

$$\begin{cases} \sigma: (I, a, b) \to (P, p, q) \\ h: I \to H \end{cases}$$

be piecewise smooth maps, where $I = [a, b]$. Then $(\sigma h)^\sim = \tilde{\sigma} h$, where the development $(\sigma h)^\sim$ starts at $h(a)$ and the development $\tilde{\sigma}$ starts at $e \in G$.

Proof. First note that $(\sigma h)^\sim(a) = h(a) = \tilde{\sigma}(a)h(a)$, so the two curves on G start at the same place. Let us show that the Darboux derivatives of $\tilde{\sigma} h$ and $(\sigma h)^\sim$ are the same.

For $\tilde{\sigma} h$ we can use Exercise **3**.4.12 to calculate the Darboux derivative as

$$(\tilde{\sigma} h)^* \omega_G = \mathrm{Ad}(h^{-1}) \tilde{\sigma}^* \omega_G + h^* \omega_G.$$

Now consider $(\sigma h)^\sim$. By the definition of development, we have $(\sigma h)^{\sim *} \omega_G = (\sigma h)^* \omega$. We cannot use Exercise **3**.4.12 to calculate $(\sigma h)^* \omega$, since σh takes values in P, which is not a Lie group. However, a similar calculation is possible. Factoring σh as

$$I \xrightarrow{\Delta} I \times I \xrightarrow{\sigma \times h} P \times H \xrightarrow{\mu} P,$$

we find

$$\begin{aligned}
(\sigma h)^* \omega &= (\mu \circ (\sigma \times h) \circ \Delta)^* \omega \\
&= ((\sigma \times h) \circ \Delta)^* \mu^* \omega \\
&= ((\sigma \times h) \circ \Delta)^* (\mathrm{Ad}(h^{-1}) \pi_P^* \omega + \pi_H^* \omega_H) \\
&= \mathrm{Ad}(h^{-1})(\pi_P \circ (\sigma \times h) \circ \Delta)^* \omega + \pi_H \circ (\sigma \times h) \circ \Delta)^* \omega_H \\
&= \mathrm{Ad}(h^{-1}) \sigma^* \omega + h^* \omega_H \\
&= \mathrm{Ad}(h^{-1}) \tilde{\sigma}^* \omega_G + h^* \omega_G.
\end{aligned}$$

Thus, $(\tilde{\sigma} h)^* \omega_G = (\sigma h)^{\sim *} \omega_G$ and hence $\tilde{\sigma} h = (\sigma h)^\sim$. ■

The first consequence of Lemma 4.12 is that we can extend the notion of development of paths on P to that of development of paths on M. Whereas a path on P develops a path on G, a path on M develops a path on a model space G/H. This is studied in Proposition 4.13 and Definition 4.15.

Proposition 4.13. *Let $\xi = (P, \omega)$ be a Cartan geometry on the manifold M modeled on the Klein geometry (G, H). Let $\sigma: (I, a, b) \to (M, x, y)$, where $I = [a, b]$. Let $\hat{\sigma}: (I, a, b) \to (P, p, q)$ be any lift, and let $\tilde{\hat{\sigma}}: (I, a) \to (G, e)$ be its development on G. Then its image $\tilde{\sigma} = \pi \tilde{\hat{\sigma}}: (I, a) \to (G/H, e)$ in G/H is a curve that is independent of the choice of the lift $\hat{\sigma}$.*

Proof. If $\hat{\sigma}: (I, a, b) \to (P, p, q)$ is a lift of σ developing to $\tilde{\hat{\sigma}}: (I, a) \to (G, e)$, then any other lift has the form $\hat{\sigma} h$ for some map $h: I \to H$ and by Lemma 4.12, $\hat{\sigma} h$ develops to $\tilde{\hat{\sigma}} h$. Since $\pi \tilde{\hat{\sigma}} h = \pi \tilde{\hat{\sigma}}$, the two developments on G/H are the same. ■

Exercise 4.14.* In the context of Proposition 4.13, consider a reductive model of the form $(G, H) = (H \times_\rho V, H)$, where $\rho: H \to Gl(V)$ is a representation and we regard V as a commutative Lie group. Then $G/H = V$ and $\mathfrak{g} = \mathfrak{h} \times_\rho V$ (i.e., $[h, v] = \rho(h)v$ for $h \in \mathfrak{h}$, $v \in V$). Now the curve $\sigma: (I, a, b) \to (M, x, y)$ may be lifted to a horizontal curve $\hat{\sigma}: (I, a, b) \to (P, p, q)$, namely, a curve such that $\hat{\sigma}^* \omega$ takes values in V.[19] Show that the development of σ on V is the curve $\tilde{\sigma}: (I, a) \to (V, 0)$ satisfying $\hat{\sigma}^* \omega = d\tilde{\sigma}$. ❑

Since any curve on M has a lift to P, the following definition makes sense.

Definition 4.15. Let $\xi = (P, \omega)$ be a Cartan geometry on the manifold M modeled on the Klein geometry (G, H). Let $\sigma: (I, a, b) \to (M, x, y)$, where $I = [a, b]$. Then the *development of σ on G/H* is the projection to G/H of the development on G of any lift to P of σ. ❀

For a concrete interpretation of this notion of development, see §3 in Appendix B.

[19]Of course, V is its own Lie algebra.

This definition has a certain conceptual power. For example, if we have a notion of *straight line*[20] on G/H, we can define the corresponding notion of a *geodesic* on M by calling a path on M a geodesic if it develops to a straight line on G/H. See Definitions **6**.2.6 and **8**.3.4 for examples in Riemannian and projective geometries. In general, however, there is no notion of a straight line in G/H. There is, however, the following more general notion.

Definition 4.16. Let $\xi = (P, \omega)$ be a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. A *generalized circle* is a curve on M that is the projection of an integral curve of an ω constant vector field on P. ❀

Exercise 4.17. (i) Show that the generalized circles in a Klein geometry $M = G/H$ are the projections to M of the left G translates of the one-parameter subgroups of G.

(ii) Show that the generalized circles in the Euclidean plane are the circles and the straight lines. Determine the generalized circles in Euclidean 3-space. ❏

Holonomy

A second consequence of Lemma 4.12 is that we obtain an analog of the notion of the monodromy of a loop on P called the *holonomy* of a loop on M. In general, only the loops on M whose classes lie in Im $\pi_*: \pi_1(P, p) \to \pi_1(M, x)$ have well-defined holonomies attached to them. (But see §3 of Appendix A.)

Definition 4.18. Let $\xi = (P, \omega)$ be a Cartan geometry on M modeled on the Klein geometry (G, H). Fix a point $p \in P$ lying over $x \in M$. Let $\lambda: (I, \partial I) \to (M, x)$ (where $I = [a, b]$) represent an element of Im $\pi_*: \pi_1(P, p) \to \pi_1(M, x)$. Let $\tilde{\lambda}: (I, \partial I) \to (P, p)$ be any lift, and let $\tilde{\tilde{\lambda}}: (I, a) \to (G, e)$ be its development on G. Then $\tilde{\tilde{\lambda}}(a) \in G$ is called the *holonomy of the loop* λ with respect to p. The *holonomy group of* ξ *with respect to* p is the set $\Phi(p) \subset G$ of all such holonomies. ❀

This definition is justified in the following exercise.

Exercise 4.19.

(i) Show that the holonomy of a loop with respect to p is independent of the choice of lift.

[20] In a reductive model G/H where $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$, a straight line is (up to left translation) the image of a one-parameter subgroup whose generator lies in $\mathfrak{p}$. In the absence of this property, there is no general notion of a straight line in a homogeneous space. The best one can do in general is speak about generalized circles, that is, the images of arbitrary one-parameter subgroups.

(ii) Show that if the holonomy of the loop λ with respect to p is g, then its holonomy with respect to ph is $h^{-1}gh$. In particular, $\Phi(ph) = h^{-1}\Phi(p)h$.

(iii) Show that $\Phi(p) \subset G$ is a subgroup.

(iv) Show that if $p, q \in P$ lie over $x, y \in M$, respectively, and $\sigma: (I, 0, 1) \to (P, p, q)$, then $\Phi(y) = g^{-1}\Phi(x)g$, where g is the endpoint of a development (starting at $e \in G$) of σ.

(v) Let $\rho \subset$ Im $\pi_*: \pi_1(P, p) \to \pi_1(M, x)$ be a subgroup. Show that the holonomies corresponding to the loops in M with classes in ρ form a subgroup $\Phi_\rho(p) \subset \Phi(p)$.

(vi) Show that for a flat geometry $\Phi(p)$ is the monodromy group of (P, ω) of Definition **3**.7.9. ❏

Definition 4.20. The subgroup $\Phi_0(p) \subset \Phi(p)$ corresponding to the null-homotopic loops on M is called the *restricted holonomy group*. ❀

Exercise 4.21. Show that

(i) $\Phi_0(p)$ is a connected subgroup of G,

(ii) $\Phi_0(p)$ is a normal subgroup of $\Phi(p)$,

(iii) $\Phi_0(p)$ is trivial for a flat geometry,

(iv) there is a canonical epimorphism $\pi_1(M, x) \to \Phi(p)/\Phi_0(p)$. (This could justly be called the monodromy representation of the Cartan geometry.) ❏

Example 4.22. One interpretation of the meaning of a torsion free geometry is that infinitesimal loops have no translation part to their holonomy. However, this doesn't necessarily hold true for actual loops. The example of a sphere rolling without slipping or twisting on a plane given in Appendix B models the development, and clearly rolling a sphere once along an equator gives a pure translation as its development. ◆

§5. Flat Cartan Geometries and Uniformization

In this section we study in detail the question, *when is a connected Cartan geometry a locally Klein geometry?* Locally, the only obstruction is the obvious one: the curvature. This is our first result. For the corresponding global statement, there are two reasons why mere flatness is not enough. First, any open subset of a Klein geometry is flat but is not geometrically equivalent to a Klein geometry. The missing ingredient here is completeness. Second, by Proposition 4.10, locally Klein geometries are geometrically orientable. For these reasons the argument in the global case is more involved than the local one and uses the "characterization of Lie groups" given in Chapter 3.

Theorem 5.1. *Let M be a flat, effective Cartan geometry modeled on $(\mathfrak{g},\mathfrak{h})$ with group H. Then each point of M has a neighborhood U which is canonically isomorphic as a Cartan geometry to an open subset of the model Klein geometry G/H, where G is any Lie group realizing $\mathfrak{g}$.*

Proof. Fix a point $x \in M$. Let (U,θ) be a Cartan gauge on M, with U connected, simply connected, and containing x. Flatness means the structural equation holds, so by the fundamental theorem of calculus, there is a map $f: U \to G$, unique up to left translation, satisfying $f^*(\omega_G) = \theta$. Fix $u \in U$, and set $g = f(u)$, $x = \pi(g)$. Then we have the following commutative diagram.

$$\begin{tikzcd}
 & T_g(gH) \arrow[r, "\approx"] \arrow[d] & \mathfrak{h} \arrow[d, phantom, "\cap"] \\
 & T_g(G) \arrow[r, "(\omega_G)_g", "\approx"'] \arrow[d, "\pi_{*g}"] & \mathfrak{g} \arrow[d, "\mathrm{proj}"] \\
T_u(U) \arrow[ur, "f_{*u}"] & T_x(G/H) \arrow[r, dashed, "\varphi_g", "\approx"'] & \mathfrak{g}/\mathfrak{h}
\end{tikzcd}$$

By the regularity condition on θ, it follows that the composite

$$\mathrm{proj}\ \theta_u = \mathrm{proj}\ \omega_G f_{*u} = \varphi_g(\pi f)_{*u}: T_u(U) \to \mathfrak{g}/\mathfrak{h}$$

is an isomorphism. Thus, the map $\rho = \pi f: U \to G/H$ is an immersion, and hence a local diffeomorphism.

Shrinking U if necessary, we get an injective immersion $\iota: U \to V \subset G/H$ onto an open set V. Clearly, this is a geometrical isomorphism onto its image since $\sigma: V \to G$ defined by $\sigma(\iota(v)) = f(v)$ is a section, over V, of $\pi: G \to G/H$, and the gauge $\sigma^*(\omega_G)$ on G/H pulls back under ι to yield $\iota^*(\sigma^*(\omega_G)) = (\sigma\iota)^*(\omega_G) = f^*(\omega_G) = \theta$. ∎

We also have the following companion result giving uniqueness.

Theorem 5.2. *Let $U \subset G/H$ be a connected and simply connected open subset and let (P,ω) be the Cartan geometry induced on U by this inclusion. Let $f: U \to G/H$ be a local geometric isomorphism. Then there is an element $g \in G$ such that $f(x) = gx$ for all $x \in U$.*

Proof. Let (P_f,ω_f) be the Cartan geometry on U induced by f. We have the following commutative diagrams.

$$\begin{tikzcd}
P \arrow[r, phantom, "\overset{\tilde{\iota}}{\subset}"] \arrow[d] & G \arrow[d, "\pi"] \\
U \arrow[r, phantom, "\overset{\iota}{\subset}"] & G/H
\end{tikzcd}
\qquad
\begin{tikzcd}
P_f \arrow[r, phantom, "\overset{\tilde{f}}{\subset}"] \arrow[d] & G \arrow[d, "\pi"] \\
U \arrow[r, phantom, "\overset{f}{\subset}"] & G/H
\end{tikzcd}$$

Since f is a local geometric isomorphism, there is a bundle map $k: P \to P_f$ covering the identity map on U and such that $k^*\omega_f = \omega$. Then

$$\tilde{\iota}^*\omega_G = \omega = k^*\omega_f = k^*\tilde{f}^*\omega_G = (\tilde{f}k)^*\omega_G$$

so that by Theorem **3**.5.2, $\tilde{\iota}$ and $\tilde{f}k$ differ by left translation by an element $g \in G$. Since k covers the identity, it follows that $f(x) = gx$ for all $x \in U$. ∎

Geometrically Oriented, Complete, Flat Cartan Geometries

Now we pass to a study of geometrically oriented, complete, flat Cartan geometries. Every locally Klein geometry $\Gamma \setminus G/H$ is not only flat in its canonical Cartan structure but is complete since the ω-constant vector fields on G are just the left-invariant vector fields which are complete. It is also geometrically oriented by Proposition 4.10. Our aim is to show the converse, namely that every geometrically oriented, complete, flat Cartan geometry arises in this way.

The following theorem generalizes results found, for example, in [J.H.C. Whitehead, 1932] and [C. Ehresmann, 1936].

Theorem 5.3. *Let $\xi = (P,\omega)$ be a complete, flat, connected, geometrically oriented Cartan geometry, with model $(\mathfrak{g},\mathfrak{h})$ and group H, on a manifold M. Then there is a connected Lie group G with Lie algebra $\mathfrak{g}$ containing H as a closed subgroup and a subgroup $\Gamma \subset G$ such that $(\Gamma \setminus G, \omega_{\Gamma\setminus G})$ is a locally Klein geometry on $\Gamma \setminus G/H$ and such that ξ is geometrically isomorphic to it. In particular, $M = \Gamma \setminus G/H$.*

Proof. Let $\pi: G_0 \to P$ denote the universal cover of P, and fix $e \in G_0$. Let $p = \pi(e) \in P$. Theorem **3**.8.7 applies to give us the following two pieces of data:

(a) there is a unique Lie group structure on G_0 with Lie algebra $\mathfrak{g}$ such that e is the identity and $\pi^*(\omega) = \omega_{G_0}$;

(b) the group of covering transformations of $\pi: G_0 \to P$ is a subgroup $\Gamma_0 \subset G_0$ acting on G_0 by left translations. Thus, $P = \Gamma_0 \setminus G_0$.

Using this data, we define $K = \bigcup_{h\in H}\{k \in \pi^{-1}(ph) \mid \mathrm{Ad}(k) = \mathrm{Ad}(h)\}$. We are going to show that

(i) K is a closed subgroup of G_0 (see steps 2 and 3 below),

(ii) the map $\phi: K \to H$ defined by $\pi(k) = p\phi(k)$ is a covering epimorphism of Lie groups (step 4),

(iii) the kernel of ϕ is a normal subgroup Z of G_0 and $\Gamma_0 \cap K = Z$ (step 6).

In the presence of these three facts, we can set $G = G_0/Z$, $\Gamma = \Gamma_0/Z$. Then $H = K/Z \subset G$ is a closed subgroup, $P = \Gamma_0 \setminus G_0 = \Gamma \setminus G$, and $\omega = \omega_{\Gamma\setminus G}$. Thus, ξ is geometrically isomorphic to $(\Gamma \setminus G, \omega_{\Gamma\setminus G})$ and, by Lemma **4**.3.12, Γ acts as a group of covering transformations on the space G/H with quotient M. Now we proceed to verify (i), (ii), and (iii).

Step 1. ϕ is surjective.

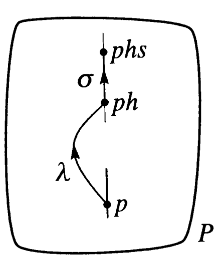
Proof of step 1: We show that for every $h \in H$ there is an element $k \in \pi^{-1}(ph) \subset G_0$ such that $\mathrm{Ad}(k) = \mathrm{Ad}(h)$. It will follow that $k \in K$ and, since $\pi(k) = ph$, $\phi(k) = h$.

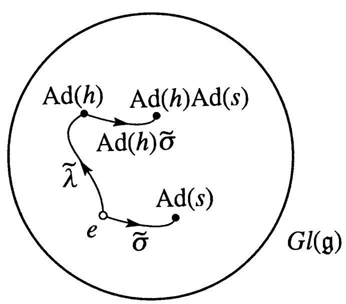
Since the geometry is geometrically oriented, we may choose a path $\lambda: I, 0, 1 \to P, p, ph$ such that it develops on $Gl(\mathfrak{g})$, via $\mathrm{ad}(\omega)$, to a path $\tilde{\lambda}: I, 0, 1 \to Gl(\mathfrak{g}), I, \mathrm{Ad}(h)$. The lift of λ to G_0 joins the identity to some element k on G_0 by a path whose development on $Gl(\mathfrak{g})$, via $\mathrm{ad}(\omega_G)$, is, since $\pi^*(\omega) = \omega_G$, once again $\tilde{\lambda}$. But by Corollary **3**.5.3, this development is just the restriction to the path of the adjoint representation of G_0. Thus $\mathrm{Ad}(k) = \mathrm{Ad}(h)$.

Step 2. K is a union of path components of $\pi^{-1}(pH)$ and hence is a closed submanifold of G_0.

Proof of Step 2: Suppose that k and l lie in the same path component of $\pi^{-1}(pH)$. It is sufficient to show that $k \in K \Rightarrow l \in K$.

Write $\pi(k) = ph$ and $\pi(l) = phs$, where $h, s \in H$. Then there is a path $\hat{\sigma}$ on $\pi^{-1}(pH)$ joining k to l and covering a path σ on P joining ph and phs. Since $k \in K$, there is a path $\hat{\lambda}$ on G_0 from e to k covering a path λ from p to ph on P such that $\mathrm{Ad}(k) = \mathrm{Ad}(h)$. Then the path $\hat{\lambda} \star \hat{\sigma}$ joins e to l and projects to a path $\lambda \star \sigma$ on P joining p to phs. This path develops (see the figure below) via $\mathrm{ad}(\omega)$ to a path joining e to $\mathrm{Ad}(hs)$ on $Gl(\mathfrak{g})$. Then, as in step 1, $\mathrm{Ad}(l) = \mathrm{Ad}(hs)$, and so $l \in K$.





Finally, we note that since $pH \subset P$ is a regular submanifold and $\pi: G_0 \to P$ is a covering space, it follows that $\pi^{-1}(pH)$ is a regular submanifold of G_0. Thus K, being a union of some of the components of $\pi^{-1}(pH)$, is also a regular submanifold and hence is closed.

Step 3. K is a group and $K = \{g \in G \mid \bar{R}_g = R_h \text{ for some } h \in H\}$, where $\bar{R}_g: P \to P$ is induced by $R_g: G_0 \to G_0$.

Proof of Step 3: First we show the inclusion $\subset$.

Let $k \in K$. Since $P = \Gamma_0 \setminus G_0$, and the left and right actions commute, it follows that $R_k: G \to G$ induces a map on P, which we denote by $\bar{R}_k: P \to P$. Now

$$\pi^*(\bar{R}_k^*\omega) = R_k^*\pi^*\omega = R_k^*\omega_{G_0} = \mathrm{Ad}(k^{-1})\omega_{G_0} = \mathrm{Ad}(h^{-1})\omega_{G_0}$$
$$= \mathrm{Ad}(h^{-1})\pi^*\omega = \pi^*(\mathrm{Ad}(h^{-1})\omega) = \pi^*(R_h^*\omega).$$

Since π^* is injective, it follows that $\bar{R}_k^*\omega = R_h^*\omega$, so that $\bar{R}_k$ and R_h are equal if they agree at a single point. But we also know that $\bar{R}_k p = \pi(R_k e) = \pi(k) = ph = R_h p$, so that $\bar{R}_k = R_h$ on P.

Now we show the inclusion $\supset$.

Let g satisfy $\bar{R}_g = R_h$ for some $h \in H$. On the one hand,

$$\mathrm{Ad}(g^{-1})\omega_{G_0} = R_g^*\pi^*\omega = \pi^*\bar{R}_g^*\omega = \pi^*R_h^*\omega = \pi^*\mathrm{Ad}(h^{-1})\omega = \mathrm{Ad}(h^{-1})\omega_{G_0}$$

and hence $\mathrm{Ad}(g) = \mathrm{Ad}(h)$. On the other hand,

$$\bar{R}_g = R_h \Rightarrow \bar{R}_g p = R_h p \Rightarrow \pi(eg) = ph \Rightarrow g \in \pi^{-1}(ph).$$

Thus, $g \in K$.

Now we show that K is a group. Clearly, $e \in K$. If $k \in K$, then $\bar{R}_k = R_h$ for some $h \in H$, so $\bar{R}_{k^{-1}} = \bar{R}_k^{-1} = R_h^{-1} = R_{h^{-1}}$ and hence $k^{-1} \in K$. Finally, if $k_1, k_2 \in K$, then $\bar{R}_{k_i} = R_{h_i}$, $h_i \in H$ for $i = 1, 2$. Thus, $\bar{R}_{k_1k_2} = \bar{R}_{k_2}\bar{R}_{k_1} = R_{k_2}R_{h_1} = R_{h_1h_2}$ and hence $k_1k_2 \in K$.

Step 4. For all elements $k \in K$ and $g \in G_0$, $\pi(gk) = \pi(g)\phi(k)$. Also, $\phi: K \to H$ is a covering epimorphism. In particular, the Lie algebra of K is the subalgebra $\mathfrak{h} \subset \mathfrak{g}$.

Proof of Step 4: We have $\pi(gk) = \pi(R_k g) = \bar{R}_k\pi(g) = R_h\pi(g) = \pi(g)h$. Taking $g = e$, we see by definition that $\phi(k) = h$ and we have verified the formula. Now we apply this formula to show that ϕ is a homomorphism. Let $k_1, k_2 \in K$. Then

$$\pi(e)\phi(k_1k_2) = \pi(k_1k_2) = \pi(k_1)\phi(k_2) = \pi(e)\phi(k_1)\phi(k_2).$$

Thus, $\phi(k_1k_2) = \phi(k_1)\phi(k_2)$.

Since $\pi: G_0 \to P$ is a covering map and ϕ is the restriction of π to a union of path components over pH ($\approx H$), it follows that ϕ is also a covering map.

Step 5. $\Gamma_0 \cap K = \{\gamma \in \Gamma_0 \mid \mathrm{Ad}(\gamma) = e\}$.

Proof of Step 5: $\gamma \in \Gamma_0 \Rightarrow \pi(\gamma e) = \pi(e) = p$. Thus, $\Gamma_0 \subset \pi^{-1}(p) \subset \pi^{-1}(pH)$. Hence $\gamma \in \Gamma_0 \cap K \Leftrightarrow \gamma \in \Gamma_0$ and $\mathrm{Ad}(\gamma) = e$.

Step 6. Let Z be the kernel of ϕ. Then $Z = \Gamma_0 \cap K$ and is a normal subgroup of G_0.

Proof of Step 6: Consider the right action of $z \in Z$ on $\Gamma_0 \setminus G_0$. This action is of course trivial, but it means that for each $g \in G_0$ there is an element $\gamma \in \Gamma_0$ such that $\gamma g = gz$. This shows that $gzg^{-1} \in \Gamma_0$ for all $g \in G_0$. Let N be the subgroup of G_0 generated by the set $\{gzg^{-1} \mid z \in Z,\, g \in G_0\} \subset \Gamma_0$. Since this generating set is stable under conjugation by elements of G, it follows that N is a normal subgroup of G that lies in Γ_0.

We have $Z \subset N \subset \Gamma_0$. Also, for $z \in Z$,

$$\begin{aligned}\mathrm{Ad}(gzg^{-1}) &= \mathrm{Ad}(g)\mathrm{Ad}(z)\mathrm{Ad}(g^{-1})\\ &= \mathrm{Ad}(g)\mathrm{Ad}(g^{-1})\\ &= e.\end{aligned}$$

So by step 5, $N \subset \Gamma_0 \cap K$. Hence $Z \subset N \subset \Gamma_0 \cap K$.

On the other hand, let $g \in \Gamma_0 \cap K$. Then $g \in \Gamma_0 \Rightarrow \pi(e) = \pi(g) = \pi(e)\phi(g)$. Thus, $\phi(g) = e$ and hence $g \in \ker \phi = Z$. It follows that $\Gamma_0 \cap K \subset Z$ and $Z \subset N \subset \Gamma_0 \cap K \subset Z$. Thus, $Z = N = \Gamma_0 \cap K$ is normal in G_0. ■

If we have two locally Klein geometries $(\Gamma_j \backslash G/H)$, $j = 1, 2$, for which the subgroups Γ_j, $j = 1, 2$, are conjugate in G, say $c\Gamma_1 c^{-1} = \Gamma_2$, then there is a geometric isomorphism relating these geometries induced by $L_c\colon G \to G$. The following result shows that this is the only geometric isomorphism between two such locally Klein geometries.

Proposition 5.4 (Geometric rgidity). *If $(\Gamma_j \setminus G, H)$, $j = 1, 2$, are two locally Klein geometries that are geometrically isomorphic in their canonical structures as Cartan geometries modeled on the Klein geometry (G, H), then the subgroups Γ_1 and Γ_2 are conjugate in G by some element $c \in G$ and the geometric isomorphism is induced by the left translation $L_c\colon G \to G$.*

Proof. Let $\pi\colon \tilde{G} \to G$ be the universal covering group, and consider the following commutative diagram

$$\begin{tikzcd} \pi^{-1}(\Gamma_1) = \tilde{\Gamma}_1 \subset \tilde{G} \arrow[r, "\tilde{f}"] \arrow[d, "\pi"] & \tilde{G} \supset \tilde{\Gamma}_2 = \pi^{-1}(\Gamma_2) \arrow[d, "\pi"] \\ \Gamma_1 \subset G \arrow[d, "\pi_1"] & G \supset \Gamma_1 \arrow[d, "\pi_2"] \\ \Gamma_1 \backslash G \arrow[r, "f"] & \Gamma_2 \backslash G \end{tikzcd}$$

where f is the H bundle map satisfying $f^*\omega_{\Gamma_2\backslash G} = \omega_{\Gamma_1\backslash G}$ and $\tilde{f}$ is any lift of f to the universal covers. Since $\omega_{\Gamma_2\backslash G}$ and $\omega_{\Gamma_1\backslash G}$ both lift to $\omega_{\tilde{G}}$ on $\tilde{G}$, and $\tilde{f}$ is just f locally, it follows that $\tilde{f}^*\omega_{\tilde{G}} = \omega_{\tilde{G}}$ and hence by Theorem **3**.5.2, $\tilde{f}$ is just left multiplication by some element $\tilde{c} \in \tilde{G}$.

For all $\tilde{\gamma}_1 \in \tilde{\Gamma}_1$, we have

$$\pi_2 \circ \pi \circ \tilde{f}(\tilde{\gamma}_1) = f \circ \pi_1 \circ \pi(\tilde{\gamma}_1) = f \circ \pi_1 \circ \pi(e) = \pi_2 \circ \pi \circ \tilde{f}(e),$$

and so it follows that for all $\tilde{\gamma}_1 \in \tilde{\Gamma}_1$, there is a $\tilde{\gamma}_2 \in \tilde{\Gamma}_2$ such that $\tilde{f}(\tilde{\gamma}_1) = \tilde{\gamma}_2\tilde{f}(e)$. Recalling that $\tilde{f}$ is left multiplication yields $\tilde{c}\tilde{\gamma}_1 = \tilde{\gamma}_2\tilde{c}$ or $\tilde{c}\tilde{\gamma}_1\tilde{c}^{-1} \in \tilde{\Gamma}_2$. Thus, $\tilde{c}\tilde{\Gamma}_1\tilde{c}^{-1} \subset \tilde{\Gamma}_2$. The same argument applied to f^{-1} yields the reverse inclusion so that $\tilde{c}\tilde{\Gamma}_1\tilde{c}^{-1} = \tilde{\Gamma}_2$. Since the kernel of $\pi\colon \tilde{G} \to G$ is a central subgroup Z and $\tilde{\Gamma}_j/Z = \Gamma_j$, it follows that $c\Gamma_1 c^{-1} = \Gamma_2$, where $c = \pi(\tilde{c})$. ■

Moduli Space of Complete, Flat Geometries

Suppose that $\Gamma \subset G$ is a subgroup. It acts as a group of covering transformations on G/H if and only if the following two conditions hold.

(i) (Free action) For every $g \in \Gamma$, $g^{-1}\Gamma g \cap H = e$.

(ii) (Proper action) For every compact set $K \subset G$, $\{\gamma \in \Gamma \mid \gamma K \cap KH \neq \emptyset\}$ is finite.

In light of Theorems 5.3 and Proposition 5.4, we see that the "moduli space" of geometric isomorphism classes of complete, flat Cartan geometries modeled on a given Klein geometry (G, H) may be identified with the set of G conjugacy classes of subgroups $\Gamma \subset G$ satisfying (i) and (ii).

§6. Cartan Space Forms

In this section we introduce and study Cartan geometries of constant curvature. The possibility for the existence of nontrivial (i.e., nonflat) examples of these is governed by the H module $\mathrm{Hom}_H(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$, which depends only on the model. If this module vanishes, there are no nontrivial examples (cf. Exercise 6.7). A Cartan form is a complete, torsion free, geometrically oriented, constant-curvature geometry, and our main result is that these geometries are all locally Klein geometries of the form $\Gamma \setminus G'/H$, where the Lie algebra of G' may *not* be the model algebra.

Mutation

The appearance of G' in the preceding description is the result of *model mutation.* Mutation replaces a geometry modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H

by one modeled on $(\mathfrak{g}', \mathfrak{h})$ with group H. We formulate the precise relation we require for mutation between models in the following definition which is essentially the same as Definition 4.3.8.

Definition 6.1. Let $(\mathfrak{g}, \mathfrak{h})$ and $(\mathfrak{g}', \mathfrak{h})$ be two models with group H. A *mutation map* is an $\mathrm{Ad}(H)$ module isomorphism $\lambda\colon \mathfrak{g} \to \mathfrak{g}'$ satisfying

(i) $\lambda \mid \mathfrak{h} = id_{\mathfrak{h}}$,

(ii) for all $u, v \in \mathfrak{g}$, $[\lambda(u), \lambda(v)] = \lambda([u,v]) \bmod \mathfrak{h}$.

We say the model $(\mathfrak{g}', \mathfrak{h})$ with group H is a *mutant* of the model $(\mathfrak{g}, \mathfrak{h})$ with group H. ❀

We mention that in [K. De Paepe, 1996] it is shown that, for primitive models of higher order, mutations are always trivial in the sense that the map λ is always a Lie algebra isomorphism. Thus, in the primitive case, mutation is entirely a phenomenon of first-order geometries.

Example 6.2. Take $H = SO_n(\mathbf{R})$, $\mathfrak{h} = \mathfrak{so}_n(\mathbf{R})$. Now define $\mathfrak{g}^-$, $\mathfrak{g}^0$, $\mathfrak{g}^+$ as follows:

$$\mathfrak{g}^- = \left\{ \begin{pmatrix} 0 & v^t \\ v & A \end{pmatrix} \middle| v \in \mathbf{R}^n, A \in \mathfrak{h} \right\},$$

$$\mathfrak{g}^0 = \left\{ \begin{pmatrix} 0 & 0 \\ v & A \end{pmatrix} \middle| v \in \mathbf{R}^n, A \in \mathfrak{h} \right\},$$

$$\mathfrak{g}^+ = \left\{ \begin{pmatrix} 0 & -v^t \\ v & A \end{pmatrix} \middle| v \in \mathbf{R}^n, A \in \mathfrak{h} \right\}.$$

These are all mutants of each other, since we may define

$$\mathfrak{g}^0 \xrightarrow{\lambda^+} \mathfrak{g}^+ \quad \text{sending} \quad \begin{pmatrix} 0 & 0 \\ v & A \end{pmatrix} \mapsto \begin{pmatrix} 0 & -v^t \\ v & A \end{pmatrix}$$

and

$$\mathfrak{g}^0 \xrightarrow{\lambda^-} \mathfrak{g}^- \quad \text{sending} \quad \begin{pmatrix} 0 & 0 \\ v & A \end{pmatrix} \mapsto \begin{pmatrix} 0 & v^t \\ v & A \end{pmatrix}.$$

It is easily verified that these are mutation maps. ◆

Mutation of models gives rise to mutation of geometries. This is described in the following result.

Proposition 6.3. *Let M be a manifold and let $\xi = (P, \omega)$ be a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Let $\lambda\colon \mathfrak{g} \to \mathfrak{g}'$ be a mutation of models, and set $\omega' = \lambda\omega$. Then $\xi' = (P, \omega')$ is a Cartan geometry on M modeled on $(\mathfrak{g}', \mathfrak{h})$ with group H. Moreover,*

(i) *the curvatures of these geometries are related by $\Omega' = \lambda\Omega + \frac{1}{2}([\omega', \omega'] - \lambda[\omega, \omega])$,*

(ii) *the geometry ξ' is complete if and only if the geometry ξ is complete.*

Under the additional hypothesis that ξ is torsion free, we have

(iii) *ξ' is torsion free,*

(iv) *λ induces a bijection*

$$\left\{ \mathfrak{l} \subset \mathfrak{g} \,\middle|\, \begin{matrix} \text{i) } \mathfrak{l} \text{ is a subalgebra} \\ \text{ii) } \mathfrak{h} \subset \mathfrak{l} \end{matrix} \right\} \longleftrightarrow \left\{ \mathfrak{l} \subset \mathfrak{g}' \,\middle|\, \begin{matrix} \text{i) } \mathfrak{l} \text{ is a subalgebra} \\ \text{ii) } \mathfrak{h} \subset \mathfrak{l} \end{matrix} \right\}.$$

Proof. Let us first verify that ω' is a Cartan connection. (We refer to Definition 3.1, page 184.)

Properties (a) and (b) are clear.

(c) (i) Since $\lambda\colon \mathfrak{g} \to \mathfrak{g}'$ is a linear isomorphism, it follows that $\omega' = \lambda\omega\colon T_p(P) \to \mathfrak{g}_2$ is a linear isomorphism for each $p \in P$.

(c) (ii) Using the fact that $\lambda\colon \mathfrak{g} \to \mathfrak{g}'$ is an H module isomorphism, we see that

$$(R_h)^*\omega' = (R_h)^*\lambda\omega = \lambda(R_h)^*\omega = \lambda\,\mathrm{Ad}(h^{-1})\omega = \mathrm{Ad}(h^{-1})\omega'.$$

(c) (iii) If $X \in \mathfrak{h}$, let $X^\dagger$ be the corresponding vector field on P so that $\omega(X^\dagger) = X$. Since $\lambda \mid \mathfrak{h} = \mathrm{id}_{\mathfrak{h}}$, it follows that $X^\dagger = \lambda(X)^\dagger$. Thus, $\omega'(\lambda(X)^\dagger) = \omega'(X^\dagger) = \lambda\omega(X^\dagger) = \lambda(X)$.

(i) We calculate the curvature of the mutated geometry:

$$\begin{aligned}
\Omega' &= d\omega' + \frac{1}{2}[\omega', \omega'] \\
&= d(\lambda\omega) + \frac{1}{2}[\omega', \omega'] \\
&= \lambda(d\omega) + \frac{1}{2}[\omega', \omega'] \\
&= \lambda\left(d\omega + \frac{1}{2}[\omega, \omega]\right) + \frac{1}{2}([\omega', \omega'] - \lambda[\omega, \omega]) \\
&= \lambda\Omega + \frac{1}{2}([\omega', \omega'] - \lambda[\omega, \omega]).
\end{aligned}$$

(ii) The equivalence of the completeness of the geometries ξ_1 and ξ_2 is a consequence of the fact that the ω_1–constant vector fields are the same as the ω_2–constant fields.

(iii) By the definition of mutation, we have $\frac{1}{2}([\omega', \omega'] - \lambda[\omega, \omega]) \in \mathfrak{h}$. Thus, $\Omega \in \mathfrak{h} \Leftrightarrow \lambda\Omega \in \mathfrak{h} \Leftrightarrow \Omega' \in \mathfrak{h}$.

(iv) Suppose that $\mathfrak{h} \subset \mathfrak{l} \subset \mathfrak{g}$ for some Lie subalgebra $\mathfrak{l}$ different from $\mathfrak{h}$ and $\mathfrak{g}$. Then, by Exercise 3.12(a), $\omega^{-1}(\mathfrak{l})$ is an integrable distribution on P. But then

$$(\omega')^{-1}(\lambda(\mathfrak{l})) = (\lambda\omega)^{-1}(\lambda(\mathfrak{l})) = \omega^{-1}(\mathfrak{l}),$$

so that the same integrable distribution may be described as $(\omega')^{-1}(\lambda(\mathfrak{l}))$. Then $\mathfrak{h} \subset \lambda(\mathfrak{l}) \subset \mathfrak{g}'$, and by Exercise 3.12(b) we see that $\lambda(\mathfrak{l})$ is a subalgebra. Since $\lambda: \mathfrak{g} \to \mathfrak{g}'$ is an isomorphism, it follows that it induces an injective mapping

$$\left\{ \mathfrak{l} \subset \mathfrak{g} \,\middle|\, \begin{array}{l} \text{i) } \mathfrak{l} \text{ is a subalgebra} \\ \text{ii) } \mathfrak{h} \subset \mathfrak{l} \end{array} \right\} \subset \left\{ \mathfrak{l} \subset \mathfrak{g}' \,\middle|\, \begin{array}{l} \text{i) } \mathfrak{l} \text{ is a subalgebra} \\ \text{ii) } \mathfrak{h} \subset \mathfrak{l} \end{array} \right\}.$$

Finally, since $\lambda^{-1}: \mathfrak{g}' \to \mathfrak{g}$ is also a model mutation, this inclusion must be a bijection. ■

Note that in mutating the model geometry, no information is lost. We may read the mutation data "backwards" and so recover the original geometry. One may say that mutant geometries are the same geometry with different presentations. The interest in mutation lies in the possibility of changing, and perhaps simplifying, the curvature.

Mutation for Reductive Models

In the case of a reductive model, the notion of mutation takes on a special significance, which can already be seen in Example 6.2 involving reductive models. Recall that a Cartan geometry with model pair $(\mathfrak{g}, \mathfrak{h})$ is *reductive* if there is an H module decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$.

Lemma 6.4. *Let $(\mathfrak{g}, \mathfrak{h})$ be a reductive model pair, with group H. Write the H module decomposition as $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$. Then there is, up to isomorphism, a unique mutant $(\mathfrak{g}', \mathfrak{h})$ with $\mathfrak{g}' = \mathfrak{h} \oplus \mathfrak{p}'$ and $[\mathfrak{p}', \mathfrak{p}'] = 0$.*

Proof. *Existence.* Set $\mathfrak{p}' = \mathfrak{p}$ as an $\mathfrak{h}$ (and H) module but with multiplication on $\mathfrak{p}'$ given by $[\mathfrak{p}', \mathfrak{p}'] = 0$. It is then easy to verify that $\mathfrak{g}' = \mathfrak{h} \oplus \mathfrak{p}'$ is a Lie algebra and that the canonical map $\mathfrak{g} \to \mathfrak{g}'$ is a mutation.

Uniqueness. This part is easy. ■

Since, in the reductive case, we can always find a mutation such that $[\mathfrak{p}, \mathfrak{p}] = 0$, and since we lose no information by passing to a mutation, it follows that for reductive Cartan geometries we exclude nothing if we assume that $[\mathfrak{p}, \mathfrak{p}] = 0$. This allows us to completely forget about the larger Lie algebra $\mathfrak{g}$ and retain only the subalgebra $\mathfrak{h}$ and the H module $\mathfrak{p}$.

If, in addition to being reductive, the geometry is of first order (i.e., $\mathrm{Ad}: H \to Gl(\mathfrak{g}/\mathfrak{h})$ is injective, cf. Definition 3.20), we may use the isomorphism $\varphi_p: T(M) \to \mathfrak{g}/\mathfrak{h} \approx \mathfrak{p}$ to reinterpret the bundle P as a bundle of frames (cf. Exercise 3.21) with group $H \subset Gl_n(\mathbf{R})$, and we can forget about $\mathfrak{p}$ too. This leads to the notion of an Ehresmann connection (see also Appendix A).

Uniformization of Constant-Curvature Geometries

Constant-curvature Cartan geometries are generalizations of the classical Riemannian space forms, namely, geometries of constant sectional curvature that are, locally, Euclidean space, the round spheres, and hyperbolic space. The generalization we consider includes, for example, products of the classical space forms. The main point is that every constant-curvature Cartan geometry is locally a mutation of a flat Cartan geometry. In particular, each of the three classical types of Riemannian space forms is locally a mutation of each of the other two.

Definition 6.5. Let M be a connected manifold and let $\xi = (P, \omega)$ be a Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Let Ω be the curvature 2-form. We say that ξ has *constant curvature* if $\Omega_p(X_p, Y_p)$ is independent of $p \in P$ whenever the vector fields X and Y are ω–constant vector fields. A *Cartan space form* is a complete, torsion free, geometrically oriented Cartan geometry of constant curvature. ❀

The constant-curvature condition may also be expressed by saying that the curvature function $K: P \to \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is constant, since

$$\begin{aligned} K(p) = \text{constant} &\Leftrightarrow \left\{ \begin{array}{l} \Omega_p(\omega_p^{-1}(u), \omega_p^{-1}(v)) \\ \text{is independent of } p \\ \text{for all } u, v \in \mathfrak{g} \end{array} \right\} \\ &\Leftrightarrow (P, \omega) \text{ has constant curvature.} \end{aligned}$$

This may be restated a bit more computationally as follows.

Lemma 6.6. *Let M be a connected manifold and let $\xi = (P, \omega)$ be a torsion free Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H and curvature 2-form Ω. Let $\{e_I\}$ be a basis of $\mathfrak{h}$ and let $\{e_i\}$ complete this to a basis of $\mathfrak{g}$. Let θ_i denote the e_ith component of ω with respect to this basis. Then we may write $\Omega = \Sigma_{ijI} a_{ijI} \theta_i \wedge \theta_j e_I$, where $a_{ijI} = -a_{jiI}: P \to \mathbf{R}$ and*

$$(P, \omega) \textit{ has constant curvature} \Leftrightarrow \textit{the functions } a_{ijI} \textit{ are all constant.}$$

Proof. It is clear (cf. Corollary 3.10) that (P, ω) has constant curvature $\Leftrightarrow$ $\Omega_p(\omega_p^{-1} e_s, \omega_p^{-1} e_t)$ is independent of p for all $1 \le s, t \le n$. But

$$\begin{aligned} \Omega_p(\omega_p^{-1} e_s, \omega_p^{-1} e_t) &= \Sigma_{ijI} a_{ijI} \theta_i \wedge \theta_j (\omega^{-1} e_s, \omega^{-1} e_t) e_I \\ &= \Sigma_{ijI} a_{ijI} (\delta_{is}\delta_{jt} - \delta_{it}\delta_{js}) e_I \\ &= \Sigma_I (a_{stI} - a_{tsI}) e_I = 2\Sigma a_{stI} e_I, \end{aligned}$$

which finishes the proof. ■

The following exercise shows that the existence of nontrivial examples of Cartan space forms modeled on the Klein pair $(\mathfrak{g}, \mathfrak{h})$ depends on the nontriviality of $\mathrm{Hom}_H(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$.

Exercise 6.7. Suppose that $\mathrm{Hom}_H(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is trivial and that K is constant along each fiber of P. Deduce that the geometry is flat (cf. Exercise **3**.4.8). ❑

The fundamental fact about the constant-curvature Cartan geometries is that, at least locally, *they are all mutations of flat geometries.* The first step toward proving this is to find the appropriate mutation of the Lie algebra $\mathfrak{g}$.

Proposition 6.8. *Let $K \in \mathrm{Hom}(\lambda^2(\mathfrak{g}), \mathfrak{h})$ be the (value of the) curvature function of a torsion free, constant-curvature Cartan geometry, so that $K(\mathfrak{g}, \mathfrak{h}) = 0$. Let $\mathfrak{g}'$ be $\mathfrak{g}$ equipped with the multiplication given by $[u, v]' = [u, v] - K(u, v)$. Then $\mathfrak{g}'$ is a Lie algebra, and for all $u, v \in \mathfrak{g}$ and $h \in H$, $[Ad(h)u, Ad(h)v]' = Ad(h)[u, v]'$.*

Proof. The bilinearity and skew symmetry of the bracket $[\ ,\ ]'$ are obvious, but we need to check the Jacobi identity, $[[u, v]', w]' + [[w, u]', v]' + [[v, w]', u]' = 0$. Since $[u, v]' = [u, v] - K(u, v)$, this reads, in terms of $[\ ,\ ]$, as

$$\begin{aligned}&[[u, v] - K(u, v), w] - K([u, v] - K(u, v), w)\\ &\qquad + \text{(cyclic permutations of } u,\ v, \text{ and } w) = 0. \end{aligned} \tag{6.9}$$

Now the bracket $[\ ,\ ]$ already satisfies the Jacobi identity. The terms $K(K(u, v), w)$ are of the form $K(\mathfrak{h}, \mathfrak{g})$ and so vanish by Corollary 3.10. Thus, the Jacobi identity for $[\ ,\ ]'$ is equivalent to the identity

$$-[K(u, v), w] - K([u, v], w) + \text{(cyclic permutations of } u,\ v, \text{ and } w) = 0. \tag{6.10}$$

We show this is a consequence of the Bianchi identity $d\Omega = [\Omega, \omega]$ as follows. Let X_0, X_1, X_2 be three ω–constant vector fields on P. Then we may explicitly calculate the two sides of the Bianchi identity

$$d\Omega(X_0, X_1, X_2) = [\Omega, \omega](X_0, X_1, X_2).$$

By Exercise **1**.5.16, the left-hand side is

$$\sum_{i=0}^{2}(-1)^i X_i(\Omega(X_0, \ldots, \hat{X}_i, \ldots, X_2))$$

$$+ \sum_{0 \le i < j \le 2} (-1)^{i+j}\Omega([X_i, X_j], X_0, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_2).$$

The first sum vanishes by the constancy of Ω on ω–constant vector fields, and the second sum is simply

$$-\Omega([X_0, X_1], X_2) + \Omega([X_0, X_2], X_1) - \Omega([X_1, X_2], X_0).$$

The right-hand side is

$$\begin{aligned}&\sum_{(2,1) \text{ shuffles } \sigma} (-1)^\sigma[\Omega(X_{\sigma(0)}, X_{\sigma(1)}), \omega(X_{\sigma(2)})]\\ &= [\Omega(X_0, X_1), \omega(X_2)] - [\Omega(X_0, X_2), \omega(X_1)] + [\Omega(X_1, X_2), \omega(X_0)].\end{aligned}$$

Thus the Bianchi identity reduces to

$$\begin{aligned}&\Omega([X_0, X_1], X_2) + [\Omega(X_0, X_1), \omega(X_2)]\\ &\qquad + \text{(cyclic permutations of 0, 1, and 2)} = 0.\end{aligned} \tag{6.11}$$

Set $\omega(X_0) = u$, $\omega(X_1) = v$, $\omega(X_2) = w$. Calculating again, we have

$$\begin{aligned}d\omega(X_0, X_1) &= X_0(\omega(X_1)) - X_1(\omega(X_0)) - \omega([X_0, X_1]\\ &= 0 - 0 - \omega([X_0, X_1]),\end{aligned}$$

and so

$$\begin{aligned}K(u, v) = \Omega(X_0, X_1) &= d\omega(X_0, X_1) + [\omega(X_0), \omega(X_1)]\\ &= -\omega([X_0, X_1]) + [u, v].\end{aligned}$$

Thus, $\Omega([X_0, X_1], X_2) = K(\omega([X_0, X_1]), \omega(X_2)) = K([u, v] - K(u, v), w) = K([u, v], w)$. This identity allows us to rewrite Eq. (6.11) in terms of K, yielding

$$K([u, v], w) + [K(u, v), w] + \text{(cyclic permutations of } u,\ v, \text{ and } w) = 0,$$

which is exactly the identity in Eq. (6.10) required of $[\ ,\ ]'$ for it to satisfy the Jacobi identity.

Finally, we show $[\mathrm{Ad}(h)u, \mathrm{Ad}(h)v]' = \mathrm{Ad}(h)[u, v]'$. Recall Lemma 3.23, which says that

$$K(\mathrm{Ad}(h)u, \mathrm{Ad}(h)v) = \mathrm{Ad}(h)K(u, v)$$

Thus

$$\begin{aligned}[Ad(h)u, Ad(h)v]' &= [Ad(h)u, Ad(h)v] - K(Ad(h)u, Ad(h)v)\\ &= Ad(h)[u, v] - Ad(h)K(u, v)\\ &= Ad(h)[u, v]'.\end{aligned}$$ ∎

The following result describes all the Cartan space forms with a connected model group H.

Theorem 6.12. *Let M be a connected manifold and let $\xi = (P, \omega)$ be a Cartan space form on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Suppose one of the following conditions holds.*[21]

(i) *H is connected.*

(ii) *M is simply connected.*

Then ξ is a mutation of a locally Klein geometry $\Gamma \setminus G'/H$, where G' has Lie algebra $\mathfrak{g}'$ as described in Proposition 6.8.

Proof. The "identity" map $\iota: \mathfrak{g} \to \mathfrak{g}'$ is a linear isomorphism. Since $K(\mathfrak{h}, \mathfrak{g}) = 0$, the restriction of ι to $\mathfrak{h}$ is an isomorphism of Lie algebras. This same identity also ensures that $[u, v]' = [u, v]$ whenever $u \in \mathfrak{h}$. It follows that the identity map $\iota: \mathfrak{g} \to \mathfrak{g}'$ commutes with the adjoint action of the identity component $H_e \subset H$. By model mutation we may regard M as equipped with the geometry of vanishing curvature modeled on $(\mathfrak{g}', \mathfrak{h})$ with group H. Either of conditions (i) or (ii) implies that the geometry is geometrically oriented. Then the classification of complete, flat geometries, Theorem 5.3, applies to tell us that ξ is locally Klein and of the form $\Gamma \setminus G'/H$. ■

Exercise 6.13. (a) Suppose that (P, ω) is a constant-curvature Cartan geometry, and let (U, θ) be a compatible gauge with curvature Θ. Let e_i, $1 \leq i \leq n$, be a fixed basis for $\mathfrak{g}/\mathfrak{h}$ and let $e_i(x) \in T_x(U)$ be the corresponding smooth vector fields on U defined by $e_i = \bar{\theta}(e_i(x))$, $1 \leq i \leq n$. Show that

(i) for each $1 \leq i, j \leq n$, $\Theta_x(e_i(x), e_j(x))$ is independent of $x \in U$;

(ii) for any $h \in H$ and any $x \in U$,

$$\mathrm{Ad}(h)\Theta_x(\bar{\theta}^{-1}(u), \bar{\theta}^{-1}(v)) = \Theta_x(\bar{\theta}^{-1}(\mathrm{Ad}(h), u), \bar{\theta}^{-1}(\mathrm{Ad}(h)v)).$$

[*Hint* for (ii): Relate $\Omega_p(\omega_p^{-1}(u), \omega_p^{-1}(v))$ and $\Omega_{ph}(\omega_{ph}^{-1}(u), \omega_{ph}^{-1}(v))$].

(b) Show that when the adjoint representation $\mathrm{Ad}: H \to \mathrm{End}(\mathfrak{g}/\mathfrak{h})$ is faithful, then the converse of (a) holds. That is, in this case, show that if a Cartan geometry (P, ω) is covered by gauges satisfying (i) and (ii), then it must have constant curvature. [*Hint*: Show first that if (i) and (ii) hold for one gauge and one basis $\{e_i\}$, they hold for all equivalent gauges and all bases.] ❑

[21] I do not know whether or not the property of being geometrically orientable is unchanged under mutation. If it is, then Theorem 6.12 is true without the necessity of these conditions.

§7. Symmetric Spaces

One of the classical definitions of a Riemannian locally symmetric space is a Riemannian geometry in which the curvature is covariant constant. This definition continues to make sense for any reductive Cartan geometry using the covariant derivative of Definition 3.43. As usual, for reductive geometries we assume that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is a fixed H module decomposition.

Definition 7.1. Let M be a reductive Cartan geometry modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. M is called a *locally symmetric* Cartan geometry if the curvature function K satisfies $D_{\mathfrak{p}}K = 0$. ❀

The following result shows that the locally symmetric spaces are "geometries extended from flat geometries."

Proposition 7.2. *Let M be a locally symmetric Cartan geometry (P, ω) modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. Then there is a reduction of the bundle P to a connected principal bundle P_1 with group $H_1 \subset H$ such that $(P_1, \omega_1 = \omega \mid P_1)$ is a constant-curvature Cartan geometry modeled on the Klein pair $(\mathfrak{h}_1 \oplus \mathfrak{p}, \mathfrak{h}_1)$ with group H_1. Moreover, if (P, ω) is complete, then so is (P_1, ω_1).*

Proof. Fix $p_0 \in P$, let $P_0 = \{p \in P \mid K(p) = K(p_0)\}$, and let $H_0 = \{h \in H \mid h \cdot K(p_0) = K(p_0)\}$. Then H_0 is a closed subgroup of H. Let $\mathfrak{h}_0$ be its Lie algebra. Since K is constant along horizontal paths in P (i.e., paths tangent to the distribution $\omega^{-1}(\mathfrak{p})$) and horizontal paths can join any point to any fiber, the function K takes on the value $K(p_0)$ on each fiber. Since $K(ph) = h^{-1} \cdot K(p)$, K is equivariant. Thus, by Proposition 4.2.14, $H_0 \to P_0 \to M$ is a principal bundle. The fact that K is constant along paths in P tangent to the distribution $\omega^{-1}(\mathfrak{p})$ implies that, for each $p \in P$, we have $\omega_p(T_p(P_0)) \supset \mathfrak{p}$. Since the fiber of P_0 is H_0, it follows that $\omega_p(T_p(P_0)) \supset \mathfrak{h}_0$. The equality $\dim P_0 = \dim M + \dim H_0$ implies $\omega_p(T_p(P_0)) = \mathfrak{h}_0 \oplus \mathfrak{p}$. Thus, $\omega_0 = \omega \mid P_0$ is an $(\mathfrak{h}_0 \oplus \mathfrak{p})$-valued 1-form on P_0. Since the form $\omega: T_p(P) \to \mathfrak{g}$ satisfies conditions (i), (ii), and (iii) of Definition 3.1, it follows that ω_0 does as well, and so it is a Cartan connection on P_0. Let us calculate the curvature of this geometry. It is

$$\Omega_0 = d\omega_0 + \frac{1}{2}[\omega_0, \omega_0] = \left(d\omega + \frac{1}{2}[\omega, \omega]\right) \mid P_0 = \Omega \mid P_0.$$

Since the curvature function $K_0 = K \mid P_0 =$ constant, the geometry (P_0, ω_0) has constant curvature.

Finally, fix a component P_1 of P_0 and an element $p_1 \in P_1$, and set $H_1 = \{h \in H \mid p_1 h \in P_1\}$. By Exercise 1.3.23, H_1 has codimension zero in H and $H_1 \to P_1 \to M$ is a principal H_1 bundle. Setting $\omega_1 = \omega_0 \mid P_1$, we see that (P_1, ω_1) is a connected, constant-curvature Cartan geometry on M modeled on $(\mathfrak{h}_1 \oplus \mathfrak{p}, \mathfrak{h}_1)$ with group H_1.

The completeness of P implies the completeness of P_0 since the ω_0–constant vector fields on P_0 are restrictions of ω–constant vector fields on P. ■

Corollary 7.3. *Let $\xi = (P, \omega)$ be a complete, reductive, locally symmetric Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H. If M is connected and simply connected, then it has the form G_1/H_1, where $H_1 \subset H_0$ is a closed subgroup with Lie algebra $\mathfrak{h}_1$ and G_1 is a Lie group with Lie algebra $\mathfrak{h}_0 \oplus \mathfrak{p}$, with bracket given by $[u, v]' = [u, v] - K_0(u, v)$.*

Proof. The long, exact homotopy sequence

$$\ldots \to \underbrace{\pi_1(M)}_{0} \to \pi_0(H_1) \to \underbrace{\pi_0(P_1)}_{0} \to \ldots$$

shows that H_1 is connected. The corollary follows by applying Theorem 6.12 to the conclusion of Proposition 7.2. ■

6
Riemannian Geometry

A Riemannian geometry consists of a smooth manifold M together with a Riemannian metric, that is, a smooth function $q_M: T(M) \to \mathbf{R}$, which restricts to a positive, definite, quadratic form on each tangent space.[1] Euclidean geometry of dimension n, denoted by E^n, consists of the pair $(\mathbf{R}^n, \langle -, - \rangle)$, where $\langle -, - \rangle$ is the usual inner product on $\mathbf{R}^n$. It may be regarded as the simplest example of a Riemannian geometry by defining

$$q: T(M) = \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R} \quad \text{by } q(x, v) = \langle v, v \rangle.$$

This example makes it clear how Riemannian geometry is a generalization of Euclidean geometry.

A deeper study of Riemannian geometry shows that *more* structure can be canonically associated to a manifold with a Riemannian metric, including the bundle of orthonormal frames along with its tautological 1-forms and the Levi–Civita connection. This construction is reviewed in §3. The extra canonical structure just mentioned determines a torsion free Cartan geometry on M modeled on Euclidean space. Conversely, any Cartan geometry on M modeled on Euclidean space determines, up to a constant scalar factor, a Riemannian metric on M. These facts provide a more profound justification for regarding a Riemannian manifold as a generalization of Euclidean space and show the equivalence between Riemann's original idea and Cartan's version of the same thing.

---

[1]Riemann actually had something more general in mind. Cf. [S.-S. Chern, 1996].

In §1 we study, in an *ad hoc* manner, the representation theory associated to the model Euclidean space. In §2 we define the Cartan geometries modeled on Euclidean space, giving a brief picture of the corresponding special geometries and ending with a discussion of the geodesics. In §3 we show how these geometries "geometrize" Riemannian metrics as described in the preceding paragraph. In §4 we study some of the special geometries, the Riemannian space forms. In §5 we pass to a study of subgeometries, second fundamental form, Ehresmann connection on the normal bundle, and a complete set of invariants for a submanifold. In §6 we apply these notions to introduce isoparametric submanifolds of Riemannian space form. We end that section by proving Cartan's formula relating the principal curvatures of an isoparametric hypersurface in a space form.

§1. The Model Euclidean Space

There are really two versions of Euclidean geometry, the oriented and the unoriented models. We describe the unoriented version here; the oriented one is obtained from it by replacing $O_n(\mathbf{R})$ by $SO_n(\mathbf{R})$. Let $G = Euc_n(\mathbf{R})$ be the group of rigid motions of Euclidean n space. Let $H = O_n(\mathbf{R}) \subset G$ denote the subgroup fixing the origin. In more detail,

$$G = \left\{ \begin{pmatrix} 1 & 0 \\ v & A \end{pmatrix} \in M_{n+1}(\mathbf{R} \mid A \in O_n(\mathbf{R}) \right\},$$
$$H = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix} \in M_{n+1}(\mathbf{R}) \mid A \in O_n(\mathbf{R}) \right\},$$

with Lie algebras $\mathfrak{g} = \mathfrak{euc}_n(\mathbf{R})$ and $\mathfrak{h} = \mathfrak{o}_n(\mathbf{R})$ given by

$$\mathfrak{g} = \left\{ \begin{pmatrix} 0 & 0 \\ v & A \end{pmatrix} \in M_{n+1}(\mathbf{R}) \mid A + A^t = 0, v \in \mathbf{R}^n \right\},$$
$$\mathfrak{h} = \left\{ \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix} \in M_{n+1}(\mathbf{R}) \mid A + A^t = 0 \right\}.$$

Definition 1.1. The group $Euc_n(\mathbf{R}) = G$ is called the *Euclidean group* in dimension n. The pair (G, H), with $H \approx O_n(\mathbf{R})$, described above is called the *Euclidean model* dimension n. ❀

The model geometry is reductive since the subalgebra $\mathfrak{h} \subset \mathfrak{g}$ has an $\mathrm{Ad}(H)$ invariant complement

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & 0 \\ v & 0 \end{pmatrix} \in M_{n+1}(\mathbf{R}) \mid v \in \mathbf{R}^n \right\},$$

and so $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is an H module decomposition, and the adjoint action of H on $\mathfrak{p} \approx \mathfrak{g}/\mathfrak{h}$ is given by $\mathrm{ad}(A)v = Av$. Let $e_i \in \mathfrak{p}$ denote the standard basis and let $e_{ij} \in \mathfrak{h}$ denote the unique elements satisfying $\mathrm{ad}(e_{ij})e_k = \delta_{jk}e_i - \delta_{ik}e_j$ such that $\mathrm{ad}(e_{pq})$ corresponds to $e_p \otimes e_q^* - e_q \otimes e_p^*$ under the canonical isomorphism $\mathrm{End}(\mathfrak{g}/\mathfrak{h}) \approx (\mathfrak{g}/\mathfrak{h}) \otimes (\mathfrak{g}/\mathfrak{h})^*$. Note that $\{e_{ij} \mid i < j\}$ is the standard basis for $\mathfrak{h}$.

The structure of the H modules $\mathfrak{h}$ and $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ determine the possible types of the (torsion free) special Cartan geometries modeled on Euclidean space. The remainder of this section is devoted to the study of these models.

Exercise 1.2. Show that $\mathfrak{h}$ is a simple Lie algebra for $n \neq 4$, and that for $n = 4$, $\mathfrak{h} \approx \mathfrak{so}(3) \oplus \mathfrak{so}(3)$. [*Hint:* For $n = 4$, consider the matrices

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}.]$$ ❑

It follows from this exercise that for $n \neq 4$ there are no special types of torsion free Cartan geometry on M modeled on Euclidean space arising from the ideals of $\mathfrak{h}$. However, there is still the possibility of special types of geometries arising from the decomposition of the H module $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$. Note that this module vanishes for $n = 1$.

Definition 1.3. The *Ricci homomorphism* is the composite mapping

$$\begin{aligned} \mathrm{Ricci} : \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) &\overset{\text{canonical}}{\approx} \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathfrak{h} \\ &\xrightarrow{\mathrm{id}\otimes\mathrm{ad}} \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathrm{End}(\mathfrak{g}/\mathfrak{h}) \xrightarrow{\phi} (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*, \end{aligned}$$

where $\phi(v_1^* \wedge v_2^* \otimes \varphi) = v_1^* \otimes (v_2^* \circ \varphi) - v_2^* \otimes (v_1^* \circ \varphi)$. ❀

Proposition 1.4. *Let $n \geq 2$.*

(i) *Ricci is an H module homomorphism satisfying*

$$Ricci(e_i^* \wedge e_j^* \otimes e_{pq}) = \delta_{jp}e_i^* \otimes e_q^* - \delta_{jq}e_i^* \otimes e_p^* - \delta_{ip}e_j^* \otimes e_q^* + \delta_{iq}e_j^* \otimes e_p^*$$

so that, in particular,

(a) *when $i \neq k \neq j$, $Ricci(e_i^* \wedge e_k^* \otimes e_{kj}) = e_i^* \otimes e_j^* + \delta_{ij}e_k^* \otimes e_k^*$.*

Moreover

(b) *for* $n = 2$, *Ricci is injective.*

(c) *for* $n > 2$, *Ricci is surjective.*

(ii) *Let* $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ *denote the subspace of elements satisfying the "first Bianchi identity" given by:*
$\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ *if and only if*

$$\Sigma_{\sigma \in A_3}[\varphi(\nu_{(\sigma(1)} \wedge \nu_{\sigma(2)}), \nu_{\sigma(3)}] = 0, \quad \text{for all } \nu_1, \nu_2, \nu_3 \in \mathfrak{g}/\mathfrak{h}$$

(*where* A_3 *is the alternating group of order* 3). *Then* $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ *is an* H *submodule of* $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$.
Moreover

(a) *let* $\varphi = \Sigma a_{ijpq} e_i^* \wedge e_j^* \otimes e_{pq} \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ *where* a_{ijpq} *is skew symmetric in the first two and last two indices. Then*

$$\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B \Leftrightarrow a_{ijkq} + a_{jkiq} + a_{kijq} = 0 \quad \text{for all } i, j, k, q$$

(b) *for* $n = 2$, $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B = \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$

(c) *for* $n > 2$, $\mathit{Ricci}(\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B) = S^2(\mathfrak{g}/\mathfrak{h})^*$ (*the symmetric elements in* $(\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$).

(iii) *Let* $n > 2$ *and set* $b_{ij} = \sum_{1 \le k \le n}(e_i^* \wedge e_k^* \otimes e_{kj} + e_j^* \wedge e_k^* \otimes e_{ki})$, $1 \le i, j \le n$. *Then*

$$\mathit{Ricci}(b_{ij}) = (n-2)(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*) + 2\delta_{ij} \Sigma_k e_k^* \otimes e_k^*$$

The elements b_{ij} *are linearly independent, and the vector space they span is an* H *submodule* $R_n \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$. *The Ricci homomorphism induces an isomorphism* $R_n \to S^2(\mathfrak{g}/\mathfrak{h})^*$ *and a canonical* H *module decomposition into irreducible pieces*

$$\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B \approx R_n \otimes W_n$$

where $W_n = \ker \mathit{Ricci} \mid R_n$.

(iv) *The submodule* $\mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ *has dimension* 1 *and generator* $\Sigma_{ik} e_i^* \wedge e_k^* \otimes e_{ki} = \frac{1}{2}\Sigma_i b_{ii}$ *and its image in* $S^2(\mathfrak{g}/\mathfrak{h})^*$ *is* $\frac{1}{2}(3n-4)\Sigma_k e_k^* \otimes e_k^*$.

(v) *There is an* H *module decomposition* $R_n \approx R_{0n} \otimes \mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ *corresponding under the Ricci homomorphism to the decomposition*

$$S^2(\mathfrak{g}/\mathfrak{h})^* \approx \langle \Sigma_{1 \le k \le n} e_k^* \otimes e_k^* \rangle \oplus \{w \in S^2(\mathfrak{g}/\mathfrak{h})^* \mid \mathrm{Trace}\ w = 0\}.$$

Proof. (i) We leave it to the reader to check that each of the maps composing *Ricci* is an H module homomorphism. Next, note that under *Ricci*, the element $e_i^* \wedge e_j^* \otimes e_{pq}$ is mapped according to

$$\begin{aligned} e_i^* \wedge e_j^* \otimes e_{pq} &\mapsto e_i^* \wedge e_j^* \otimes \mathrm{ad}(e_{pq}) \\ &\mapsto e_i^* \wedge e_j^* \otimes \Sigma_t(\delta_{qt} e_p - \delta_{pi} e_q) e_t^* \\ &\mapsto \Sigma_t e_j^*(\delta_{qt} e_p - \delta_{pt} e_q) e_i^* \otimes e_t^* - \Sigma_t e_i^*(\delta_{qt} e_p - \delta_{pt} e_q) e_j^* \otimes e_t^* \\ &= \delta_{jp} e_i^* \otimes e_q^* - \delta_{jq} e_i^* \otimes e_p^* - \delta_{ip} e_j^* \otimes e_q^* + \delta_{iq} e_j^* \otimes e_p^* \end{aligned}$$

and in particular,

(a) for $i \ne j = p \ne q$, $\mathit{Ricci}(e_i^* \wedge e_j^* \otimes e_{jq}) = e_i^* \otimes e_q^* + \delta_{iq} e_j^* \otimes e_j^*$.

(b) For $n = 2$, $\dim \lambda^2(\mathfrak{g}/\mathfrak{h}) = \dim \mathfrak{h} = 1$ so that $e_1^* \wedge e_2^* \otimes e_{21}$ is clearly a basis for $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ and, by the formula (a), $\mathit{Ricci}(e_1^* \wedge e_2^* \otimes e_{21}) = e_1^* \otimes e_1^* + e_2^* \otimes e_2^* \ne 0$.

(c) On the other hand, for $n > 2$, it is clear that *Ricci* is surjective from the formula for $\mathit{Ricci}(e_i^* \wedge e_k^* \otimes e_{kj})$ given in (a).

(ii) The condition that $\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ implies that for all $\nu_1, \nu_2, \nu_3 \in \mathfrak{g}/\mathfrak{h}$,

$$\Sigma_{\sigma \in A_3} \mathrm{Ad}(h)[\varphi(\mathrm{Ad}(h^{-1})\nu_{\sigma(1)} \wedge \mathrm{Ad}(h^{-1})\nu_{\sigma(2)}), \mathrm{Ad}(h^{-1})\nu_{\sigma(3)}] = 0$$

which is

$$\Sigma_{\sigma \in A_3}[\mathrm{Ad}(h)(\varphi(\mathrm{Ad}(h^{-1})\nu_{\sigma(2)})), \nu_{\sigma(3)}] = 0$$

and so

$$\Sigma_{\sigma \in A_3}[(\mathrm{Ad}(h)\varphi)(\nu_{\sigma(1)} \wedge \nu_{\sigma(2)}), \nu_{\sigma(3)}] = 0.$$

Thus $\mathrm{Ad}(h)\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ and so the latter is an H submodule of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$.

(a) The condition that $\varphi \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}, \mathfrak{h})_B$ may clearly be written as

$$\Sigma_{\sigma \in A_3}[\varphi(e_{i_{\sigma(1)}} \wedge e_{i_{\sigma(2)}}), e_{i_{\sigma(3)}}] = 0, \quad \text{for all } i_1, i_2, i_3$$

which, since $\varphi(e_{i_{\sigma(1)}} \wedge e_{i_{\sigma(2)}}) = 2\Sigma_{pq}(a_{i_{\sigma(1)} i_{\sigma(2)} pq}) e_{pq}$, becomes

$$0 = 2\Sigma_{\sigma \in A_3, pq} a_{i_{\sigma(1)} i_{\sigma(2)} pq}[e_{pq}, e_{i_{\sigma(3)}}] = -4\Sigma_{\sigma \in A_3, p} a_{i_{\sigma(1)} i_{\sigma(2)} i_{\sigma(3)} p} e_p.$$

Since the e_p are independent, this is equivalent to (ii)(a).

(b) Let us fix $i \ne k \ne j$ and note that by (ii)(a)

$$\frac{1}{2} e_i^* \wedge e_k^* \otimes e_{kj} + 0 e_k^* \wedge e_k^* \otimes e_{ij} - \frac{1}{2} e_k^* \wedge e_i^* \otimes e_{kj} = e_i^* \wedge e_j^* \otimes e_{ji} \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B.$$

Thus, for $n = 2$, the basis $e_1^* \wedge e_2^* \otimes e_{21}$ for $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ lies in $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ which verifies (ii)(b).

(c) On the other hand, for $n > 2$ the formula in (i) verifies (ii)(b).

(iii) Again we leave it to the reader to verify that R_n is an H submodule. Using the formula in (i)(a) we calculate

$$\begin{aligned}
Ricci(b_{ij}) &= \Sigma_{k\neq i,j} Ricci(e_i^* \wedge e_k^* \otimes e_{kj} + e_j^* \wedge e_k^* \otimes e_{ki}) \\
&= \Sigma_{k\neq i,j}(e_i^* \otimes e_j^* + e_j^* \otimes e_i^* + 2\delta_{ij} e_k^* \otimes e_k^*) \\
&= \begin{cases} \Sigma_{k\neq i,j}(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*) & \text{if } i \neq j \\ 2\Sigma_{k\neq i}(e_i^* \otimes e_i^* + e_k^* \otimes e_k^*) & \text{if } i = j \end{cases} \\
&= \begin{cases} (n-2)(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*) & \text{if } i \neq j \\ 2(n-1)e_i^* \otimes e_i^* + 2\Sigma_{k\neq i} e_k^* \otimes e_k^* & \text{if } i = j \end{cases} \\
&= \begin{cases} (n-2)(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*) & \text{if } i \neq j \\ 2(n-2)e_i^* \otimes e_i^* + 2\Sigma_k e_k^* \otimes e_k^* & \text{if } i = j \end{cases} \\
&= (n-2)(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*) + 2\delta_{ij}\Sigma_k e_k^* \otimes e_k^*.
\end{aligned}$$

It is clear from this formula that the elements $Ricci(b_{ij})$, $i \leq i, j \leq n$ are linearly independent and span the symmetric submodule $S^2(\mathfrak{g}/\mathfrak{h})^* \subset (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$. It follows immediately that the Ricci homomorphism induces an isomorphism $R_n \to S^2(\mathfrak{g}/\mathfrak{h})^*$ and that $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B \approx R_n \otimes W_n$ is an isomorphism of H modules.

(iv) We refer to Exercise **3**.4.8(c) for the proof that $\mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ has dimension one with the given generator. It is clear that $\Sigma_{ik} e_i^* \wedge e_k^* \otimes e_{ki} = \frac{1}{2}\Sigma_i b_{ii}$ and, using the formula in (iii), it is easy to determine its image in $S^2(\mathfrak{g}/\mathfrak{h})^*$ is as stated.

(v) The decomposition

$$\begin{aligned} S^2(\mathfrak{g}/\mathfrak{h})^* &\approx \langle \Sigma_{1\leq k\leq n} e_k^* \otimes e_k^* \rangle \\ &\quad \oplus \{w \in S^2(\mathfrak{g}/\mathfrak{h})^* \mid w = \Sigma_{ij} a_{ij} e_i^* \otimes e_j^* \text{ with } \Sigma_i a_{ii} = 0\} \end{aligned}$$

is just the decomposition of $n \times n$ symmetric matrices as a sum of multiples of the identity matrix and trace zero matrices. This decomposition is clearly stable under conjugation by any orthogonal matrix. The first summand is irreducible because its dimension is one. For the irreducibility of the second summation, see Exercise 1.6. Since the Ricci homomorphism restricts to an isomorphism $R_n \approx S^2(\mathfrak{g}/\mathfrak{h})^*$ that, by (iv), identifies $\mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ with $\langle \Sigma_{1\leq k\leq n} e_k^* \otimes e_k^* \rangle$, we may define R_{0n} to be the submodule of R_n corresponding to the "trace zero" submodule of $S^2(\mathfrak{g}/\mathfrak{h})^*$. Then the decomposition $R_n \approx \mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \oplus R_{0n}$ is clear. ∎

Definition 1.5. The submodules of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ described in Proposition 1.4 have the following names:

(i) $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ is the *Bianchi* submodule.

(ii) R_n is the *Ricci* submodule.

(iii) W_n is the *Weyl* submodule.

(iv) R_{0n} is the *traceless Ricci* submodule.

(v) $\mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is the *scalar* submodule. ❀

Exercise 1.6. Let $SO_n(\mathbf{R})$ act on the space of trace-free symmetric $n \times n$ matrices $\mathrm{Sym}_n(\mathbf{R})_0$ by conjugation. Show that this representation is irreducible. [*Hint:* Suppose there is a nonzero subrepresentation $V \subset \mathrm{Sym}_n(\mathbf{R})_0$. Each nonzero element of V is conjugate to a diagonal matrix of the form $\mathrm{diag}(\alpha, \beta, \ldots)$, where $\alpha \neq \beta$ and neither is zero. Then this diagonal matrix is in turn conjugate to the same matrix with α and β interchanged. The difference, which also lies in V, is a multiple of a matrix of the form $\mathrm{diag}(1, -1, 0, \ldots, 0)$. It then easily follows that $V = \mathrm{Sym}_n(\mathbf{R})_0$.] ❑

It is known that as an $O_n(\mathbf{R})$ module, W_n is irreducible (cf. [I.M. Singer and J.A. Thorpe, 1969]).

Exercise 1.7. Show that, as an $SO_n(\mathbf{R})$ module, W_n can decompose into at most two submodules and, if it does so, these submodules are isomorphic and will be interchanged by the action of any orientation-reversing element of $O_n(\mathbf{R})$. ❑

Note that when $n = 4$, $W_4 = W^+ \otimes W^-$ as an $SO_4(\mathbf{R})$ module (cf. [I.M. Singer and J.A. Thorpe, 1969]). W_n is irreducible as an $SO_n(\mathbf{R})$ module for $n \neq 4$. The existence of the decomposition $W_4 = W^+ \oplus W^-$ in dimension 4 is at the root of Donaldson's important work relating differential geometry to the structure of smooth four-dimensional manifolds (cf. [S. Donaldson and P.B. Kronheimer, 1990]).

In summary, the decomposition into three irreducible $O_n(\mathbf{R})$ modules

$$\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B = \mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \oplus R_n \oplus W_n$$

provides a corresponding decomposition of $L \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$ as $L = s(L) + r_0(L) + W(L)$, where the notation is chosen so that s, r_0, and W will correspond to the scalar, traceless Ricci, and Weyl curvatures respectively.

We end this section with the following criterion for the equality of elements of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$.

Proposition 1.8.

(i) *$\varphi \in Hom(\lambda^2(\mathfrak{p}), \mathfrak{h})_B$ satisfies*

$$\langle ad(\varphi(X \wedge Y))Z, W\rangle = \langle ad(\varphi(Z \wedge W))X, Y\rangle \quad \textit{for all } X, Y, Z, W \in \mathfrak{p}.$$

(ii) *Let $\varphi_1, \varphi_2 \in Hom(\lambda^2(\mathfrak{p}), \mathfrak{h})_B$. Then*

$$\varphi_1 = \varphi_2 \Leftrightarrow \begin{cases} \langle \mathrm{ad}(\varphi_1(v_1 \wedge v_2))v_1, v_2\rangle = \langle \mathrm{ad}(\varphi_2(v_1 \wedge v_2))v_1, v_2\rangle \\ \textit{for all } v_1, v_2 \in \mathfrak{p}. \end{cases}$$

Proof. (i) This is a consequence of the Bianchi identity. Adding the equations

$$\begin{aligned}
\langle \mathrm{ad}(\varphi(X\wedge Y))Z,W\rangle + \langle \mathrm{ad}(\varphi(Y\wedge Z))X,W\rangle + \langle \mathrm{ad}(\varphi(Z\wedge X))Y,W\rangle &= 0\\
-(\langle \mathrm{ad}(\varphi(W\wedge X))Y,Z\rangle + \langle \mathrm{ad}(\varphi(X\wedge Y))W,Z\rangle + \langle \mathrm{ad}(\varphi(Y\wedge W))X,Z\rangle) &= 0\\
-(\langle \mathrm{ad}(\varphi(Z\wedge W))X,Y\rangle + \langle \mathrm{ad}(\varphi(W\wedge X))Z,Y\rangle + \langle \mathrm{ad}(\varphi(X\wedge Z))W,Y\rangle) &= 0\\
\langle \mathrm{ad}(\varphi(Y\wedge Z))W,X\rangle + \langle \mathrm{ad}(\varphi(Z\wedge W))Y,X\rangle + \langle \mathrm{ad}(\varphi(W\wedge Y))Z,X\rangle &= 0
\end{aligned}$$

and using the skew symmetry of $\langle \mathrm{ad}(\varphi(X\wedge Y))Z,W\rangle$ in X, Y and Z, W yields

$$2\langle \mathrm{ad}(\varphi(X\wedge Y))Z,W\rangle - 2\langle \mathrm{ad}(\varphi(Z\wedge W))X,Y\rangle = 0 \quad \text{for all } X,Y,Z,W\in\mathfrak{p}.$$

(ii) It suffices to prove this result for the case φ_1 ($=\varphi$ say) and $\varphi_2 = 0$. Since $\Rightarrow$ is automatic, we prove $\Leftarrow$. Suppose

$$\langle \mathrm{ad}(\varphi(v_1\wedge v_2))v_1, v_2\rangle = 0 \quad \text{for all } v_1,v_2\in\mathfrak{p}. \tag{$*$}$$

Then $\langle \mathrm{ad}(\varphi(v_1\wedge (v_2+v_4)))v_1, v_2+v_4\rangle = 0$ for all $v_1,v_2,v_4\in\mathfrak{p}$. So by $(*)$,

$$\langle \mathrm{ad}(\varphi(v_1\wedge v_2))v_1, v_4\rangle + \langle \mathrm{ad}(\varphi(v_1\wedge v_4))v_1, v_2\rangle = 0 \quad \text{for all } v_1,v_2,v_4\in\mathfrak{p}$$

or, by (i),

$$\langle \mathrm{ad}(\varphi(v_1\wedge v_2))v_1, v_4\rangle = 0 \quad \text{for all } v_1,v_2,v_4\in\mathfrak{p}. \tag{$**$}$$

Thus, $\langle \mathrm{ad}(\varphi((v_1+v_3)\wedge v_2))(v_1+v_3), v_4\rangle = 0$ for all $v_1,v_2,v_3,v_4\in\mathfrak{p}$. So by $(**)$,

$$\langle \mathrm{ad}(\varphi(v_1\wedge v_2))v_3, v_4\rangle + \langle \mathrm{ad}(\varphi(v_3\wedge v_2))v_1, v_4\rangle = 0 \quad \text{for all } v_1,v_2,v_3,v_4\in\mathfrak{p}$$

and so, by the Bianchi identity, $\langle \mathrm{ad}(\varphi(v_1\wedge v_2))v_3, v_4\rangle = 0$ for all $v_1, v_2, v_3, v_4\in\mathfrak{p}$. Thus, $\varphi = 0$. ■

§2. Euclidean and Riemannian Geometry

Definition 2.1. Let M be a smooth manifold. A *Euclidean geometry* on M is a Cartan geometry on M modeled on Euclidean space. A *Riemannian geometry*[2] on M is a torsion free Euclidean geometry. ❀

We continue the notation at the beginning of §1. Let us assume that (P,ω) is a Euclidean geometry on M with model $(\mathfrak{g},\mathfrak{h})$ and group H. Since the Lie algebra $\mathfrak{g}$ decomposes as an H module as $\mathfrak{g} = \mathfrak{h}\oplus\mathfrak{p}$, the Cartan connection $\omega\colon T(P)\to\mathfrak{g}$ and the curvature $\Omega\colon A^2(P)\to\mathfrak{g}$ decompose correspondingly as $\omega = \omega_{\mathfrak{h}}\oplus\omega_{\mathfrak{p}}$ and $\Omega = \Omega_{\mathfrak{h}}\oplus\Omega_{\mathfrak{p}}$. These may also be expressed as

[2]Of course, this is not the usual definition of a Riemannian geometry, which is a manifold together with a Riemnanian metric on it. The equivalence (up to scale) of these definitions will be shown in §3.

$$\omega = \begin{pmatrix} 0 & 0\\ \omega_{\mathfrak{p}} & \omega_{\mathfrak{h}}\end{pmatrix} = \sum_{1\le i\le n}\omega_i e_i + \sum_{1\le i<j\le n}\omega_{ij}e_{ij}, \quad \text{or } \omega = \begin{pmatrix} 0 & 0\\ (\omega_i) & (\omega_{ij})\end{pmatrix},$$

$$\Omega = \begin{pmatrix} 0 & 0\\ \Omega_{\mathfrak{p}} & \Omega_{\mathfrak{h}}\end{pmatrix} = \sum_{1\le i\le n}\Omega_i e_i + \sum_{1\le i<j\le n}\Omega_{ij}e_{ij}, \quad \text{or } \Omega = \begin{pmatrix} 0 & 0\\ (\Omega_i) & (\Omega_{ij})\end{pmatrix}.$$

Definition 2.2. In the case of a Riemannian geometry, the form $\omega_{\mathfrak{h}}$ is called the *Levi-Civita connection.* The form $\omega_{\mathfrak{p}}$ has been called the *solder form.* ❀

Exercise 2.3. Verify that the Levi-Civita connection on the principal bundle P satisfies the following properties:

(i) $R_h^*\omega_{\mathfrak{h}} = \mathrm{Ad}(h^{-1})\omega_{\mathfrak{h}}$;

(ii) $\omega_{\mathfrak{h}}(X^\dagger) = X$ for every $X\in\mathfrak{h}$. ❏

The properties of $\omega_{\mathfrak{h}}$ given in this exercise were adopted by Ehresmann as the definition of what is today called an Ehresmann connection on the principal bundle P with any group as fiber.

Definition 2.4. Let $H\to P\to M$ be an arbitrary principal bundle over M and let $\mathfrak{h}$ be the Lie algebra of H. An *Ehresmann connection* of P is an $\mathfrak{h}$-valued 1-form γ satisfying the conditions

(i) $R_h^*\gamma = \mathrm{Ad}(h^{-1})\gamma$, and

(ii) $\gamma(X^\dagger) = X$ for every $X\in\mathfrak{h}$.

The *curvature* of the Ehresmann connection γ is the two form $d\gamma + \frac{1}{2}[\gamma,\gamma]$. ❀

Ehresmann connections are studied in [S. Kobayashi and K. Nomizu, 1963] and enjoy wide usage since in important cases—Riemannian geometry, for example—all of the geometry may be simply expressed in terms of an Ehresmann connection. See Appendix A for a discussion of Ehresmann connections and their relationship with Cartan connections.

Special Geometries

We are going to study the most important case, the Riemannian geometries[3] on M. Thus, we assume that $\Omega_{\mathfrak{p}} = 0$ or, equivalently, $\Omega_i = 0$, $1\le i\le n$.

[3]The more general Euclidean geometries are apparently not very important. The Lorentzian analog of the extreme case of a Euclidean geometry on a manifold M with $\Omega_{\mathfrak{h}} = 0$ was investigated by Einstein from 1929–1932 in the hope that a

The Bianchi identity is $d\Omega = [\Omega, \omega]$ or

$$\begin{pmatrix} 0 & 0 \\ 0 & d\Omega_{\mathfrak{h}} \end{pmatrix} = \left[\begin{pmatrix} 0 & 0 \\ 0 & \Omega_{\mathfrak{h}} \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ \omega_{\mathfrak{p}} & \omega_{\mathfrak{h}} \end{pmatrix} \right] = \begin{pmatrix} 0 & 0 \\ \Omega_{\mathfrak{h}} \wedge \omega_{\mathfrak{p}} & [\Omega_{\mathfrak{h}}, \omega_{\mathfrak{h}}] \end{pmatrix}.$$

This identity decomposes into

(*first Bianchi identity*) $\quad 0 = \Omega_{\mathfrak{h}} \wedge \omega_{\mathfrak{p}}$ or $0 = \sum_j \Omega_{ij} \wedge \omega_j$,

(*second Bianchi identity*) $\quad d\Omega_{\mathfrak{h}} = [\Omega_{\mathfrak{h}}, \omega_{\mathfrak{h}}]$, or
$d\Omega_{ij} = \sum_k (\Omega_{ik} \wedge \omega_{kj} - \omega_{ik} \wedge \Omega_{kj})$.

The curvature function $K: P \to \mathrm{Hom}(\lambda^2(\mathfrak{p}), \mathfrak{h})$ is defined by

$$K(p)(v_1, v_2) = \Omega_p(\omega_p^{-1}(v_1), \omega_p^{-1}(v_2)) \quad \text{for } v_1, v_2 \in \mathfrak{p}.$$

Writing $\Omega_{ij} = \Sigma_{kl} a_{ijkl} \omega_k \wedge \omega_l$ (which we may do since Ω is semibasic), the first Bianchi identity becomes $0 = \Sigma_{jkl} a_{ijkl} \omega_k \wedge \omega_l \wedge \omega_j$, or $a_{ijkl} + a_{iklj} + a_{iljk} = 0$ for all i, j, k, l. Thus K takes values in $\mathrm{Hom}(\lambda^2(\mathfrak{p}), \mathfrak{h})_B$. It follows from the results of §1 that in a Riemannian geometry the curvature function decomposes as

$$K = s(K) + r_0(K) + W(K).$$

Table 2.5 lists the usual terminology.

TABLE 2.5

$s(K)$	$r_0(K)$	$s(K) + r_0(K)$	$W(K)$
scalar curvature	traceless Ricci curvature	Ricci curvature	Weyl curvature

$W(K) = r_0(K) = 0$	$W(K) = 0$	$r_0(K) = 0$
constant curvature if $n > 2$ (cf. §4)	conformally flat (cf. Theorem **7**.3.9)	Einstein manifold

Geodesics

The notion of a straight line in Euclidean space immediately extends to the notion of a geodesic in a Riemannian manifold as follows.

Definition 2.6. A *geodesic* in a Riemannian manifold M is a curve on M whose development in Euclidean space is a straight line. ❀

---

unified field theory might be based on it [E. Cartan and A. Einstein, 1979]. Since the holonomy of such a geometry is purely translational, and the translations form a normal subgroup of all the group of all motions, there is an *absolute parallelism* on M in the sense that there is a canonical trivialization of the tangent bundle of M inducing isometries between the tangent spaces at any two points of M.

Note that we speak here of an *unparametrized* geodesic. One may of course choose the arclength parameterization for a geodesic, which is also the parameterization by the arclength of the development.

Proposition 2.7. *Let (U, θ) be a Riemannian gauge on M. Let I denote the interval (a, b). Then the geodesics $c: I \to M$ are the solutions of the* ODE.

$$\frac{\theta_1(\dot c)^{\cdot} + P_1(\dot c)}{\theta_1(\dot c)} = \cdots = \frac{\theta_n(\dot c)^{\cdot} + P_n(\dot c)}{\theta_n(\dot c)}, \quad \text{where } P_i(v) = \sum_{1 \le j \le n} \theta_{ij}(v)\theta_j(v).$$

Proof. We must show that the given differential equation is a necessary and sufficient condition for the regular curve $c(t)$ to develop to a straight line in the model space $\mathbf{R}^n$. Consider the associated function $\theta(\dot c): I \to \mathfrak{g}$, which develops to give a curve on G

$$\tilde c: I, a \to G, e$$

satisfying $\omega_G(\dot{\tilde c}(t)) = \theta(\dot c(t))$. Now the projection of this curve to the model space $\mathbf{R}^n$ is $\tilde c(t)e_0$, where $e_0 \in \mathbf{R}^{n+1}$. Then $\tilde c(t)e_0$ lies in a straight-line segment in $\mathbf{R}^n$ if and only if $\dot{\tilde c}(t)e_0$ and $\ddot{\tilde c}(t)e_0$ are linearly dependent. Since $c(t)$ is regular, it is a geodesic if and only if there is a function $\lambda(t)$ such that $\ddot{\tilde c}e_0 = \lambda \dot{\tilde c}e_0$ or, equivalently, $\tilde c^{-1}\ddot{\tilde c}e_0 = \lambda \tilde c^{-1}\dot{\tilde c}e_0$.

But we have $\tilde c^{-1}\dot{\tilde c} = \omega_G(\dot{\tilde c}) = \theta(\dot c)$. Since $(\tilde c^{-1})^{\cdot} = -\tilde c^{-1}\dot{\tilde c}\tilde c^{-1}$, we also have

$$\theta(\dot c)^{\cdot} = -\tilde c^{-1}\dot{\tilde c}\tilde c^{-1}\dot{\tilde c} + \tilde c^{-1}\ddot{\tilde c} = -\theta(\dot c)^2 + \tilde c^{-1}\ddot{\tilde c}.$$

With these identities, the equation for geodesics becomes $\theta(\dot c)^{\cdot} e_0 + \theta(\dot c)^2 e_0 = \lambda\theta(\dot c)e_0$. If the gauge is written as

$$\theta = \begin{pmatrix} 0 & 0 \\ \theta_i & \theta_{ij} \end{pmatrix},$$

the geodesic equation may be expressed as

$$\frac{\theta_i(\dot c)^{\cdot} + \sum_{1 \le j \le n} \theta_{ij}(\dot c)\theta_j(\dot c)}{\theta_i(\dot c)} = \lambda \quad \text{for } i = 1, \ldots, n. \qquad \blacksquare$$

Exercise 2.8. Show that the equations for a geodesic parametrized by arclength are $\theta_i(\dot c)^{\cdot} + P_i(\dot c) = 0$, $1 \le i \le n$. ❑

§3. The Equivalence Problem for Riemannian Metrics

We begin with the classical definition of a Riemannian metric.

Definition 3.1. A *Riemannian metric* on a manifold M is a smooth function $q_M: T(M) \to \mathbf{R}$ whose restrictions $q_x: T_x(M) \to \mathbf{R}$ are all nondegenerate quadratic forms. ❀

What is the relation between this notion and that of a Riemannian geometry as given in Definition 2.1? It turns out that these two notions are, up to a constant scale factor, the same. The following proposition shows that a Euclidean geometry on a manifold always determines a Riemannian metric up to scale. The converse will be treated in Theorem 3.4. Taken together, these results "geometrize" or "solve the equivalence problem" for Riemannian metrics in the sense that they show that two manifolds equipped with Riemannian metrics are isometric up to scale if and only if the associated Riemannian geometries are geometrically isomorphic.

Proposition 3.2. *A Riemannian geometry on M determines a Riemannian metric on M up to constant scale factor.*

Proof. As we noted in §1, the adjoint action of $H = SO_n(\mathbf{R})$ on $\mathfrak{g}$ induces the standard action on $\mathfrak{g}/\mathfrak{h} \approx \mathbf{R}^n$ given by $\mathrm{ad}(A)v = Av$. This action preserves the standard quadratic form q on $\mathfrak{g}/\mathfrak{h}$, and in fact q is, up to scale, the only quadratic form on $\mathfrak{g}/\mathfrak{h}$ preserved by H. We may use the isomorphisms $\varphi_p: T_x(M) \to \mathfrak{g}/\mathfrak{h}$ (where $p \in \pi^{-1}(x)$) to transport q to a quadratic form q_p on $T_x(M)$ defined by $q_p(v) = q(\varphi_p(v))$. Since $\varphi_{ph} = \mathrm{Ad}(h^{-1})\varphi_p$, it follows that

$$q_{ph}(v) = q(\varphi_{ph}(v)) = q(\mathrm{Ad}(h^{-1})\varphi_p(v)) = q(\varphi_p(v)) = q_p(v).$$

Thus, the quadratic form q_p on $T_x(M)$ is independent of the choice of $p \in \pi^{-1}(x)$, and we have found a canonical (up-to-scale) Riemannian metric q_M on M. To see that q_M is smooth, consider the following diagrams.

$$\begin{tikzcd} T_p(P) \arrow[r, "\omega_p"] \arrow[d, "\pi_{*p}"'] & \mathfrak{g} \arrow[d] \\ T_x(M) \arrow[r, "\varphi_p"] \arrow[dr, "q_M"'] & \mathfrak{g}/\mathfrak{h} \arrow[d, "q"] \\ & \mathbf{R} \end{tikzcd} \qquad \begin{tikzcd} T(P) \arrow[r, "\omega"] \arrow[d, "\pi_*"'] & \mathfrak{g} \arrow[d] \\ T(M) \arrow[dr, "q_M"'] & \mathfrak{g}/\mathfrak{h} \arrow[d, "q"] \\ & \mathbf{R} \end{tikzcd}$$

FIGURES 3.3(a) and (b)

The diagram on the left commutes by the definition of q_x and of φ_p, and it implies the commutativity of the diagram on the right. However, the upper composite $T(P) \to \mathbf{R}$ in the diagram on the right is clearly smooth, and since π_* is a submersion, it follows that q_M is smooth. ∎

The proof of the converse of Proposition 3.2 depends on the following lemma, which will also be useful in other contexts.

Cartan's Lemma 3.4. (a) *Let V be an n-dimensional vector space, and let $\theta_i \in V^*$, $i = 1, \ldots, n$, be a basis. Let $\mu_i \in \lambda^2(V^*)$, $i = 1, \ldots, n$, be arbitrary. Then there exists a unique collection of elements $\theta_{ij} \in V^*$, $i, j = 1, \ldots, n$, satisfying*

(i) $\theta_{ij} + \theta_{ji} = 0$,

(ii) $\mu_i + \Sigma_j \theta_{ij} \wedge \theta_j = 0$.

(b) *Suppose that $\theta_i \in A^1(U)$, where U is some open set in $\mathbf{R}^n$, and that $\mu_i \in A^2(U)$. Then the forms θ_{ij} guaranteed by the lemma are smooth, that is, $\theta_{ij} \in A^1(U)$.*

Proof. (a) *Existence.* Write $\mu_i = \sum_{j,k} A_{ijk}\theta_j \wedge \theta_k$, where we may assume $A_{ijk} + A_{ikj} = 0$, so the As are uniquely determined. Set

$$\theta_{ij} = -\sum_k (A_{jik} + A_{ikj} - A_{kji})\theta_k.$$

Then

$$\begin{aligned} &\text{(i)} \quad \theta_{ij} + \theta_{ji} = -\Sigma_k(A_{jik} + A_{ikj} - A_{kji} + A_{ijk} + A_{jki} - A_{kij})\theta_k = 0, \\ &\text{(ii)} \quad \Sigma_j \theta_{ij} \wedge \theta_j = -\Sigma_{jk} A_{ikj}\theta_k \wedge \theta_j - \Sigma_{jk}(A_{jik} - A_{kji})\theta_k \wedge \theta_j \\ &\qquad = -\mu_i - \Sigma_{j<k}(A_{jik} - A_{kji} - A_{kij} + A_{jki})\theta_k \wedge \theta_j \\ &\qquad = -\mu_i. \end{aligned}$$

Uniqueness. If there were two sets of 1-forms satisfying (i) and (ii), their differences γ_{ij} would satisfy (i) and (ii) with $\mu_i = 0$. Writing $\gamma_{ij} = \Sigma_k B_{ijk}\theta_k$, (i) implies $B_{ijk} + B_{jik} = 0$ and (ii) implies $0 = \Sigma_j \gamma_{ij} \wedge \theta_j = \Sigma_{jk} B_{ijk}\theta_k \wedge \theta_j = \Sigma_{j<k}(B_{ijk} - B_{ikj})\theta_k \wedge \theta_j$, so $B_{ijk} = B_{ikj}$. But skew symmetry in the first two indices together with symmetry in the last two indices imply that all the Bs vanish.

(b) This is automatic from the formula for θ_{ij} appearing in the proof of (a). ∎

We apply this lemma to obtain the next result.

Theorem 3.5. *Let (M, g) be a smooth manifold equipped with a Riemannian metric g. There is exactly one torsion free Cartan geometry on M whose associated metric is, up to scale, g.*

Proof. Let $e(x) = (e_1(x), e_2(x), \ldots, e_n(x))$ be any choice of orthonormal frame field on an open set $U \subset M$, and let $\theta_1, \ldots, \theta_n \in A^1(U)$ be the dual 1-forms. Lemma 3.4 assures us that there are unique forms (θ_{ij}) on U such that $d(\theta_i) + (\theta_{ij}) \wedge (\theta_i) = 0$. We set

$$\theta = \begin{pmatrix} 0 & 0 \\ \theta_i & \theta_{ij} \end{pmatrix} \in A^1(U, \mathfrak{g}).$$

Thus, θ is a $\mathfrak{g}$-valued 1-form on U, and since $e(x)$ is a frame, θ has the property that the composite

$$T_x(U) \xrightarrow{\theta_x} \mathfrak{g} \xrightarrow{\text{proj.}} \mathfrak{g}/\mathfrak{h}$$

is an isomorphism. Thus, θ is a Cartan gauge. It is clear that the Riemannian metric induced on U from the standard Euclidean metric on $\mathfrak{g}/\mathfrak{h}$ is the original metric $g \mid U$. The whole of M is covered by such Cartan gauges, and it remains only to show that a different choice of orthonormal frame fields yields an equivalent gauge. Suppose that $f(x) = (f_1(x), f_2(x), \dots, f_n(x))$ is another orthonormal frame field with dual 1-forms $\psi_1, \dots, \psi_n \in A^1(U)$. Now the two frame fields must be related by an orthogonal transformation $h: U \to O(n)$, so that $f_i(x) = \Sigma_j h_{ij} e_j$, and $\psi_i(x) = \Sigma_j h_{ij}\theta_j$. But under the change of gauge $h: U \to O(n)$, we have

$$\begin{aligned} \theta \Rightarrow_h &\operatorname{Ad}(h^{-1})\theta + h^*\omega_H \\ &= \begin{pmatrix} 1 & 0 \\ 0 & h^{-1} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \theta_i & \star \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & h \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & \star \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ h^{-1}(\theta_i) & \star \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \psi_i & \star \end{pmatrix}, \end{aligned}$$

and since this new gauge must also be torsion free, by the Cartan lemma it must be the unique torsion free gauge corresponding to the coframe $\psi_1, \dots, \psi_n$. Thus, θ and ψ are gauge equivalent. In this way we have constructed a Riemannian geometry on M whose associated Riemannian metric is, up to scale, g. Moreover, the construction shows that any Riemannian gauge whose associated metric is g will be gauge equivalent to the ones constructed above. ■

Exercise 3.6. Show that the bundle P associated to the Cartan geometry described in the proof of Theorem 3.5 may be identified, with the aid of the isomorphisms $\varphi_p: T_x(M) \to \mathfrak{p}$, with the $O_n(\mathbf{R})$ bundle of orthonormal frames on M. If P is regarded as the bundle of orthonormal frames on M, show that the solder form $\omega_\mathfrak{p}$ is determined by P alone. (For this reason, the analog of $\omega_\mathfrak{p}$ on the bundle of orthonormal frames is sometimes referred to as the *tautological* form.) Cf. also Exercise **5**.3.21 and §2 of Appendix A. ❑

Exercise 3.7. Let (U, x) be a coordinate neighborhood in a Riemannian geometry, and suppose that in this coordinate system

$$g = \sum_{1 \le i,j \le n} g_{ij}\, dx_i \otimes dx_j.$$

Choosing an orthonormal frame field $(e_1, e_2, \dots, e_n)$ on U determines the dual coframe field $(\theta_1, \theta_2, \dots, \theta_n)$. Let

$$\theta = \theta_\mathfrak{p} + \theta_\mathfrak{h} \in A^1(U, \mathfrak{g}),$$

where

$$\theta_\mathfrak{p} = \begin{pmatrix} 0 & 0 \\ \theta_i & 0 \end{pmatrix} \quad \text{and} \quad \theta_\mathfrak{h} = \begin{pmatrix} 0 & 0 \\ 0 & \theta_{ij} \end{pmatrix},$$

be the corresponding unique torsion free gauge on U guaranteed by Theorem 3.5. Recall that the covariant derivative $D_Y X$ is determined by

$$\theta_\mathfrak{p}(D_X Y) = X_x(\theta_\mathfrak{p}(Y)) + [\theta_\mathfrak{h}(X), \theta_\mathfrak{p}(Y)]$$

(cf. Proposition **5**.3.49).

(a) Show that $D_X e_i = \sum_{1 \le k \le n} \theta_{ki}(X) e_k$.

(b) Show that $X(g(Y, Z)) = g(D_X Y, Z) + g(Y, D_X Z)$. [*Hint:* Use the fact that the linear map $\theta_\mathfrak{p}: T_u(U) \to \mathfrak{p}$ is an isometry to write

$$g(D_X Y, Z) = X(\theta_\mathfrak{p}(Y)) \cdot \theta_\mathfrak{p}(Z) + [\theta_\mathfrak{h}(X), \theta_\mathfrak{p}(Y)] \cdot \theta_\mathfrak{p}(Z),$$

and then use the skew symmetry of $\theta_\mathfrak{h}(X)$.]

(c) Writing $D_{\partial_i}\partial_j = \sum_{1 \le k \le n} \gamma_{ij}^k \partial_k$, where $\partial_s = \partial/\partial x_s$, $1 \le s \le n$, calculate $\partial_k g(\partial_i, \partial_j)$ in terms of the γs using the formula in (b). Show that

$$\gamma_{ij}^h = \sum_{1 \le k \le n} g^{kh} \frac{1}{2}\{g_{jk,i} + g_{ki,j} - g_{ij,k}\},$$

where $g_{jk,i} = \partial_i g_{jk}$, and so forth. ❑

§4. Riemannian Space Forms

In this section we study the special geometries for which $W(K) = r_0(K) = 0$. Our first aim is to justify the terminology in Table 2.5, referring to these geometries as having constant sectional curvature. In these geometries the curvature function takes values in $\operatorname{Hom}_\mathfrak{h}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$. By Exercise **3**.4.6(c), this is a one-dimensional vector space, so the curvature must have the form

$$\Omega = \begin{pmatrix} 0 & 0 \\ 0 & (c\theta_i \wedge \theta_j) \end{pmatrix}.$$

Thus, the curvature function is $K(p) = \sum c(p) e_i^* \wedge e_j^* \otimes e_{ij}$, where $c = c(p)$ is some function on P. The case of $n = \dim \mathfrak{g}/\mathfrak{h} = 2$ is somewhat special since $\operatorname{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) = \operatorname{Hom}_\mathfrak{h}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$, so it is no condition at all to say the curvature takes values in this submodule. However, if we avoid this case, we can prove that the curvature function is constant as follows.

Lemma 4.1. *Suppose the curvature function for a Riemannian geometry takes values in* $\mathrm{Hom}_{\mathfrak{g}}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h})$. *If* $n \geq 3$, *then the curvature function is constant.*

Proof. This is an application of the Bianchi identity $d\Omega = [\Omega, \omega]$, which in this case reads

$$d\begin{pmatrix} 0 & 0 \\ 0 & c(\theta_i \wedge \theta_j) \end{pmatrix} = \left[\begin{pmatrix} 0 & 0 \\ 0 & c(\theta_i \wedge \theta_j) \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ (\theta_i) & (\theta_{ij}) \end{pmatrix}\right]$$
$$= \begin{pmatrix} 0 & 0 \\ 0 & c\sum_k(\theta_i \wedge \theta_k \wedge \theta_{kj} - \theta_{ik} \wedge \theta_k \wedge \theta_j) \end{pmatrix}.$$

Since the geometry is torsion free, $d\theta_i = -\sum \theta_k \wedge \theta_{ki}$, so the Bianchi identity reads, for each i, j,

$$d(c\theta_i \wedge \theta_j) = c(-\theta_i \wedge d\theta_j + d\theta_i \wedge \theta_j) = cd(\theta_i \wedge \theta_j).$$

Hence $(dc) \wedge \theta_i \wedge \theta_j = 0$ for all i, j and so, if $n \geq 3$, then $dc = 0$ and c is constant. ■

A *Riemannian space form* is a complete Riemannian geometry with constant sectional curvature. (We take here the definition of completeness described in Definition **5**.3.1. This implies the completeness of geodesics, which is the classical notion of completeness.) By Proposition 4.5, a Riemannian space form is a Cartan space form whose model is Euclidean space (cf. §4 of Chapter 5). Of course this case is simpler than the general one in that the coefficients of the curvature function are not only constant (cf. Lemma **5**.6.7), they are all equal. For $n > 2$ this follows from Lemma 4.1, while for $n = 2$ there is only one coefficient. We are going to describe the universal cover of a Riemannian space form.

Proposition 4.2. *Let* M^n, $n \geq 2$, *be a Riemannian space form of curvature* c. *Let* $\tilde{M}$ *be the universal cover of* M.

(i) *If* $c = 0$, $\tilde{M} = Euc_n(\mathbf{R})/O_n(\mathbf{R})$ (Euclidean n-space)

(ii) *If* $c > 0$, $\tilde{M} = O_{n+1}(\mathbf{R})/O_n(\mathbf{R})$ (the n-sphere)

(iii) *If* $c < 0$, $\tilde{M} = O_{1,n}(\mathbf{R})/O_n(\mathbf{R})$ (hyperbolic n-space).

Proof. Since the property of having a constant curvature function is local, the universal cover $\tilde{M}$ also has constant curvature. Moreover the property of completeness for M passes to completeness for the universal cover. Thus $\tilde{M}$ is also a Riemannian space form. If $c = 0$ then $\tilde{M}$ is flat and hence it

is the model space $Euc_n(\mathbf{R})/O_n(\mathbf{R})$. Suppose that $c \neq 0$. Then we have model mutations

$$\mathfrak{euc}_n(\mathbf{R}) \to \mathfrak{o}_{n+1}(\mathbf{R}) \text{ if } c > 0 \quad\text{and}\quad \mathfrak{euc}_n(\mathbf{R}) \to \mathfrak{o}_{1,n}(\mathbf{R}) \text{ if } c < 0$$

$$\begin{pmatrix} 0 & 0 \\ v & A \end{pmatrix} \mapsto \begin{pmatrix} 0 & -cv^t \\ cv & A \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ v & A \end{pmatrix} \mapsto \begin{pmatrix} 0 & cv^t \\ cv & A \end{pmatrix}$$

By Proposition **5**.6.4(i) and (ii) both of these mutated geometries are flat and complete. Since they are also simply connected they are geometrically orientable (Exercise **5**.4.9) and so, by Theorem **5**.5.4, they are as described in (ii) and (iii) above. ■

Sectional Curvature

In this subsection we relate the previous results of this section to the classical notion of a manifold with constant sectional curvature.

Exercise 4.3. Let M be a Riemannian manifold and let $V \subset T_x(M)$ be a 2-plane. Choose

(i) an orthonormal basis $e, f \in V$,

(ii) a point $p \in P$ lying over x,

(iii) lifts $\tilde{e}, \tilde{f} \in T_p(P)$ lying over e and f.

Show that the number $R(e, f) = \langle \mathrm{ad}(K(p)(\omega_p(\tilde{e}), \omega_p(\tilde{f})))\omega_p(\tilde{e}), \omega_p(\tilde{f})\rangle$ depends only on the (unoriented) 2-plane V. In particular, it is independent of the choices in (i), (ii), and (iii). ❑

Let $\mathrm{Gr}_2(T(M))$ denote the Grassman bundle of oriented 2-planes in the tangent bundle of M. Exercise 4.3 justifies the following definition.

Definition 4.4. *Sectional curvature* is the function $R\colon \mathrm{Gr}_2(T(M)) \to \mathbf{R}$ defined in Exercise 4.3. A Riemannian geometry is said to have *constant sectional curvature* c at $x \in M$ if $R(V) = c$ for every 2-plane $V \subset T_x(M)$. A Riemannian geometry has *constant sectional curvature* c if R is the constant function with value c on $\mathrm{Gr}_2(T(M))$. ❀

Proposition 4.5. *The following are equivalent.*

(i) *A Riemannian geometry has constant sectional curvature* c *at* $x \in M$.

(ii) $K(p) = c\sum_{i<j} e_i^* \wedge e_j^* \otimes e_{ij}$ $(c \in \mathbf{R})$ *on the fiber over* x.

Proof. Let $\varphi = c\sum_{i<j} e_i^* \wedge e_j^* \otimes e_{ij} = \frac{1}{2}c\sum_{ij} e_i^* \wedge e_j^* \otimes e_{ij} \in \mathrm{Hom}_{\mathfrak{h}}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})_B$. Then for orthonormal vectors $v_1 = \sum_k a_k e_k$, $v_2 = \sum_l b_l e_l$, we have

$$
\begin{aligned}
&\langle \mathrm{ad}(\varphi(v_1 \wedge v_2))v_2, v_1\rangle \\
&\quad = \sum_{klst} a_k b_l a_s b_t \langle \mathrm{ad}(\varphi(e_k \wedge e_l))e_t, e_s\rangle \\
&\quad = \frac{1}{2} c \sum_{ijklst} a_k b_l a_s b_t \langle \mathrm{ad}(e_i^* \wedge e_j^*(e_k \wedge e_l) \otimes e_{kl})e_t, e_s\rangle \\
&\quad = \frac{1}{2} c \sum_{ijklst} a_k b_l a_s b_t (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) \langle \mathrm{ad}(e_{kl})e_t, e_s\rangle \\
&\quad = \frac{1}{2} c \sum_{ijklst} a_k b_l a_s b_t (\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) \langle \delta_{lt}e_k - \delta_{kt}e_l, e_s\rangle \\
&\quad = \frac{1}{2} c \sum_{ijklst} a_k b_l a_s b_t \delta_{ik}\delta_{jl}\delta_{lt}\delta_{ks} - \frac{1}{2} c \sum_{ikjlst} a_k b_l a_s b_t \delta_{ik}\delta_{jl}\delta_{kt}\delta_{ls} \\
&\qquad - \frac{1}{2} c \sum_{ijklst} a_k b_l a_s b_t \delta_{il}\delta_{jk}\delta_{lt}\delta_{ks} + \frac{1}{2} c \sum_{ijklst} a_k b_l a_s b_t \delta_{il}\delta_{jk}\delta_{kt}\delta_{ls} \\
&\quad = \frac{1}{2} c \sum_{ij} a_i b_j a_i b_j - \frac{1}{2} c \sum_{ij} a_i b_j a_j b_i - \frac{1}{2} c \sum_{ij} a_j b_i a_j b_i + \frac{1}{2} c \sum_{ij} a_j b_i a_i b_j \\
&\quad = c \sum_{ij} a_i^2 b_j^2 - c \sum_{ij} a_i b_i a_j b_j \\
&\quad = c \sum_i a_i^2 \sum_j b_j^2 - c \sum_i a_i b_i \sum_j a_j b_j \\
&\quad = c.
\end{aligned}
$$

Thus (i) is equivalent to the statement:

$$\langle \mathrm{ad}(K(p)(\omega_p(\tilde{e}), \omega_p(\tilde{f})))\omega_p(\tilde{e}), \omega_p(\tilde{f})\rangle = c$$

for any orthonormal vectors $\tilde{e}, \tilde{f} \in \omega_p^{-1}(\mathfrak{p})$. This, in turn, is equivalent to the statement: $\langle \mathrm{ad}(K(p)(v_1 \wedge v_2)v_1, v_2\rangle = c$ for every pair of orthonormal vectors $v_1, v_2 \in \mathfrak{p}$. But by Proposition 1.8(ii), $\mathrm{ad}(K(p))$ is determined by this formula, so we must have $\mathrm{ad}(K(p)) = c\varphi$. ∎

§5. Subgeometry of a Riemannian Geometry

Suppose we are given a Riemannian manifold N^{n+r} and an immersion $f: M^n \to N^{n+r}$ with normal bundle ν (cf. Exercise 5.7). Our main aim in this section is to describe (P_f, ω_f), the "locally ambient geometry of N along f" (Definition 5.5). This is a geometric entity that contains all the geometry of the map f and in which the role of the ambient manifold N is greatly reduced. It is an example of "a locally ambient Riemannian geometry (P, ω) of codimension r on M" described in Definition 5.2. In

Proposition 5.8 we show that when N is the Euclidean space $\mathbf{R}^{n+r}$ the map f is determined up to an isometry of $\mathbf{R}^{n+r}$ by the locally ambient geometry of N along f. In Proposition 5.9, 5.12, and 5.14 we show how to fracture a locally ambient Riemannian geometry on M into three pieces consisting of

(i) the induced Riemannian geometry on M (the *tangential* part of (P, ω)),

(ii) an Ehresmann connection γ on the normal bundle ν (the *normed* part of (P, ω)),

(iii) a ν-valued form B (the second fundamental form) on $T(M)$ (the *glueing* data).

The remainder of the section studies the principal curvatures associated to the second fundamental form when the normal bundle is *relatively flat* (Definition 5.23).

Model Algebra

We begin with some notation describing the algebra for the model consisting of Euclidean $(n+r)$-dimensional space and the standard n-dimensional subspace. The corresponding Lie algebras are

$$\mathfrak{g} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ \star & \star & \star \\ \star & \star & \star \end{pmatrix} \right\}, \quad \mathfrak{h} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \star & \star \\ 0 & \star & \star \end{pmatrix} \right\},$$

$$\mathfrak{a} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ \star & \star & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}, \quad \mathfrak{b} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \star & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\},$$

where the diagonal blocks are 1×1, $n \times n$, and $r \times r$. Of course, we must bear in mind that not all possible entries are allowed. Any matrix in $\mathfrak{g}$ must have the form

$$\begin{pmatrix} 0 & 0 & 0 \\ \star & A & -B^t \\ \star & B & D \end{pmatrix},$$

where A and D are skew-symmetric matrices. As in §1 we have the $O_{n+r}(\mathbf{R})$ module decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$. We also note the $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ decompositions $\mathfrak{a} = \mathfrak{b} \oplus \mathfrak{t}$ and $\mathfrak{p} = \mathfrak{t} \oplus \mathfrak{u}$, where

$$\mathfrak{t} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ \star & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}, \quad \mathfrak{u} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \star & 0 & 0 \end{pmatrix} \right\}.$$

(In particular, we may identify $\mathfrak{t}$ with $\mathfrak{a}/\mathfrak{b}$ and $\mathfrak{u}$ with $\mathfrak{g}/(\mathfrak{h} + \mathfrak{a})$.)

Geometry of N Localized Along f

Recall that we are studying an immersion $f\colon M \to N$ into a Riemannian manifold. The description of the geometry of N localized along f is based on the tangent reduction P_τ along f. This is a reduction of the pullback bundle $f^*(P)$ given by

$$P_\tau = \{(x,p) \in f^*(P) \mid \varphi_p(f_*(T_x(M)) = \mathfrak{t}\}.$$

P_τ sits in the following commutative diagram

$$\begin{array}{ccccc} P_\tau & \subset & f^*(P) & \xrightarrow[\tilde{f}]{} & P \\ \downarrow & & \downarrow & & \downarrow \\ M & \overset{\mathrm{id}}{=} & M & \xrightarrow{f} & N \end{array}$$

It is a principal bundle over M with group $H_\tau = O(n) \times O(r)$ whose Lie algebra is

$$\mathfrak{h}_\tau = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & \star & 0 \\ 0 & 0 & \star \end{pmatrix} \right\}.$$

If we set $\omega_\tau = \tilde{f}^*(\omega) \mid P_\tau$, then since

$$\begin{aligned} \mathfrak{t} &= \varphi_p(f_*(T(M))) = \varphi_p(f_*(\pi_*(T_p(P_\tau)))) \\ &= \omega_p(\tilde{f}_*(T(P_\tau)))\operatorname{mod}\mathfrak{h} = \tilde{f}^*(\omega)(T_p(P_\tau))\operatorname{mod}\mathfrak{h}, \end{aligned}$$

it follows that ω_τ mod $\mathfrak{h}$ takes values in $\mathfrak{t}$ and hence has the form

$$\omega_\tau = \begin{pmatrix} 0 & 0 & 0 \\ \theta & \alpha & -\beta^t \\ 0 & \beta & \gamma \end{pmatrix}.$$

Moreover, the restriction of the curvature form to P_τ has the shape

$$\begin{aligned} \Omega_\tau &= d\omega_\tau + \omega_\tau \wedge \omega_\tau \\ &= \begin{pmatrix} 0 & 0 & 0 \\ d\theta + \alpha\wedge\theta & d\alpha + \alpha\wedge\alpha - \beta^t\wedge\beta & \star \\ 0 & d\beta + \beta\wedge\alpha + \gamma\wedge\beta & d\gamma + \gamma\wedge\gamma - \beta\wedge\beta^t \end{pmatrix}. \end{aligned} \tag{5.1}$$

This leads us to make the following definition, which we will relate to (P_τ, ω_τ) in Lemma 5.4.

Definition 5.2. Let M be a smooth manifold of dimension n. A *locally ambient Riemannian geometry on M of codimension* r is a pair (P, ω) where

(i) P is a principal $O(n) \times O(r)$ bundle over M, and

(ii) ω is an $(\mathfrak{h} \oplus \mathfrak{t})$-valued form on P satisfying the following conditions:

(a) $\omega\colon T(P_\tau) \to \mathfrak{h} \oplus \mathfrak{t}$ is injective at each point, and its composite with the canonical projection $\mathfrak{h} \oplus \mathfrak{t} \to \mathfrak{h}_\tau \oplus \mathfrak{t}$ is surjective at each point;

(b) $R_h^*\omega = \operatorname{Ad}(h^{-1})\omega$ for all $h \in H_\tau$;

(c) $\omega_\tau(X^\dagger) = X$ for each $X \in \mathfrak{h}_\tau$.

(P, ω) is called *torsion free* if, in the notation of Eq. (5.1), $d\theta + \alpha \wedge \theta = 0$. Two locally ambient geometries (P_1, ω_1) and (P_2, ω_2) on M are said to be *equivalent* if there is an $O_r(\mathbf{R}) \times O_r(\mathbf{R})$ bundle map $b\colon P_1 \to P_2$ such that $b^*(\omega_2) = \omega_1$. ❀

Definition 5.3. The *normal bundle* associated to the locally ambient Riemannian geometry (P, ω) is the r-dimensional vector bundle

$$\nu = P \times_{O_n(\mathbf{R}) \times O_r(\mathbf{R})} \mathbf{R}^r$$

where the action of $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ on $\mathbf{R}^r$ is the standard second factor action. It has a canonical (up to constant scale) metric $q_\nu\colon \nu \to \mathbf{R}$ on it. ❀

The point of Definition 5.2 is that it describes the data associated to the immersion $f\colon M \to N$ we have been discussing.

Lemma 5.4. *(P_τ, ω_τ) is a locally ambient geometry on M of codimension r.*

Proof. We verify the conditions of Definition 5.2. (i) is obvious, as is the fact that ω_τ is an $(\mathfrak{h} \oplus \mathfrak{t})$-valued form on P_τ. Condition (ii) (b) is an automatic consequence of the formula on P that $R_h^*\omega = \operatorname{Ad}(h^{-1})\omega$ for all $h \in H$. Condition (ii) (c) comes from the fact that the inclusion $P_\tau \subset f^*(P)$ is equivariant with respect to the two actions of H_τ; hence, for $X \in \mathfrak{h}_\tau$, the vector field $X^\dagger$ on $f^*(P)$ restricts to the vector field $X^\dagger$ on P_τ. The injectivity of condition (ii) (a) comes from the facts that $\omega_\tau = \tilde{f}^*(\omega) \mid P_\tau$, f is an immersion, and ω is injective. The rest of (ii) (a) comes from combining (ii) (c) with the fact, shown just before Definition 5.2, that $\omega_\tau(T_p(P_\tau)) = \mathfrak{t}$ mod $\mathfrak{h}$ for all $p \in P_\tau$. ∎

We signal the importance of this example of a locally ambient geometry by giving the following definition.

Definition 5.5. Let N be a Riemannian manifold and let $f\colon M \to N$ be an immersion. *The locally ambient geometry of N along f* (or *along M* if f is an inclusion) is the pair (P_τ, ω_τ). ❀

Exercise 5.6. Let $\phi: N \to N$ is an isometry of the Riemannian manifold N. If $f: M \to N$ is an immersion, show that the locally ambient geometry of N along f is equivalent to the locally ambient geometry of N along ϕf. ❑

Exercise 5.7. Show that the normal bundle of an immersion $f: M \to N$ given by $\nu = \{(x, v) \in M \times T(N) \mid v \in M, v \in f_*(T_x(M))^\perp\}$, is canonically isomorphic to the normal bundle, in the sense of Definition 5.3, of the locally ambient geometry (P_τ, ω_τ) associated to f as in Lemma 5.4. ❑

Now we can justify the terminology "ambient geometry localized along f" by showing that, at least in the case of an ambient Euclidean space, it does indeed encapsulate *all* the geometry of f.

Proposition 5.8. (i) *Let $f: M^n \to \mathbf{R}^{n+r}$ (Euclidean space) be an immersion. If (P_τ, ω_τ) is the ambient geometry localized along f, then $d\omega_\tau + \omega_\tau \wedge \omega_\tau = 0$ and the monodromy representation Φ_{ω_τ} is trivial.*

(ii) *Conversely, let M^n be a connected n-dimensional manifold and let (P, ω) be a locally ambient geometry on M of codimension r satisfying $d\omega + \omega \wedge \omega = 0$ with trivial monodromy. Then there is an immersion $f: M^n \to \mathbf{R}^{n+r}$, determined up to left multiplication by an isometry of $\mathbf{R}^{n+r}$, such that the Euclidean geometry of $\mathbf{R}^{n+r}$ localized along f is equivalent to the geometry (P, ω).*

Proof. This is just an application of the fundamental theorem of calculus (cf. Theorem **3**.7.14). ■

Fracturing the Local Geometry

Now let (P, ω) be a locally ambient geometry on M of codimension r. The next step in our investigation is to decompose (P, ω) into the three pieces mentioned in the introduction to this section. First we show that the bundle P itself decomposes as the fiber product of principal bundles determined by the tangent bundle of M and the associated normal vector bundle ν.

Proposition 5.9. *Let $P_{tan} = P/O_r(\mathbf{R})$ and $P_{nor} = P/O_n(\mathbf{R})$. Then*

(i) *P_{tan} is a principal $O_n(\mathbf{R})$ bundle and P_{nor} is a principal $O_r(\mathbf{R})$ bundle,*

(ii) *$P = P_{tan} \times_M P_{nor}$ (fiber product),*

(iii) *P_{tan} and P_{nor} are principal bundles associated to the tangent of M and the normal bundle ν over M arising from (P, ω). Moreover, if (P, ω) is the locally ambient geometry of an immersion f, then ν is the normal bundle of f.*

Proof. (i) Since the actions of $O_n(\mathbf{R})$ and of $O_r(\mathbf{R})$ on P commute, we see that $O_n(\mathbf{R})$ acts smoothly on P_{tan}. Since the action of $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ is proper on P and transitive on the fibers of $P \to M$, it follows that the action of $O_n(\mathbf{R})$ is proper on P_{tan} and transitive on the fibers of $P_{\text{tan}} \to M$. Hence P_{tan} is a right principal $O_n(\mathbf{R})$ bundle over M. The case of $P_{\text{nor}} \to M$ is similar.

(ii) The two smooth bundle maps (defined by projection) $O_r(\mathbf{R}) \to P \to P_{\text{tan}}$ and $O_n(\mathbf{R}) \to P \to P_{\text{nor}}$ induce a smooth map into the fiber product $P \to P_{\text{tan}} \times_M P_{\text{nor}}$. This map covers the identity on M and is obviously an $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ bundle map.

(iii) Since the standard action of $O_n(\mathbf{R})$ on $\mathbf{R}^n$ is faithful, the principal bundle $P_{\text{tan}} \to M$ is determined by its associated vector bundle. The case of $P_{\text{nor}} \to M$ is similar.

As for the identification of the associated vector bundles, we deal with the case of the normal bundle only and leave the similar case of the tangent bundle to the reader. Define $P \times \mathfrak{u} \to \nu$ by sending $(p, v) \to \varphi_p^{-1}(v)$. Note that since φ_p is an isometry and for $p \in P$, $\varphi_p(T_{\pi(p)}(M)) = \mathfrak{t}$, it follows that $\varphi_p(\nu_{\pi(p)}) = \mathfrak{u}$, and hence $\varphi_p^{-1}(\mathfrak{u}) = \nu_{\pi(p)}$. Since every $h \in O_n(\mathbf{R})$ acts trivially on $\mathfrak{u}$ we have $\varphi_{ph}^{-1}(v) = \varphi_p^{-1}(\mathrm{Ad}(h)v) = \varphi_p^{-1}(v)$. Thus, the map induces a smooth bundle map $P_{\text{nor}} \times \mathfrak{u} = P/O_n(\mathbf{R}) \times \mathfrak{u} \to \nu$. But we also have, for every $k \in O_r(\mathbf{R})$, $\varphi_{pk}^{-1}(v) = \varphi_p^{-1}(\mathrm{Ad}(k)v)$. It follows that (p, v) and $(pk, \mathrm{Ad}(k^{-1})v)$ have the same image in ν. Thus, we get a further induced map $P_{\text{nor}} \times_{O(r)} \mathfrak{u} \to \nu$. This map is vector bundle map covering the identity on M, and the fibers have the same dimensions. Thus, it is a bundle equivalence. ■

Exercise 5.10. Let (P, ω) be a locally ambient geometry on M. Show that $P \times_{\mathrm{Ad}} \mathfrak{t}$ may be canonically identified with the tangent bundle of M. ❑

Writing

$$\omega = \begin{pmatrix} 0 & 0 & 0 \\ \theta & \alpha & -\beta^t \\ 0 & \beta & \gamma \end{pmatrix}, \tag{5.11}$$

the next step is to see that the parts α and γ of the form ω determine, and are determined by, certain induced forms (to which we give the same names) on P_{tan} and P_{nor}.

Proposition 5.12. *Let (P, ω) be a locally ambient geometry on M of codimension r.*

(i) *The forms α and θ (respectively, γ) appearing in the block decomposition of ω are the pullback under the canonical projection $P \to P_{tan}$ (respectively, P_{nor}) of unique forms, which we denoted again by α and θ (respectively, γ) on P_{tan} (respectively, P_{nor}).*

(ii) *Let γ be the form on P_{nor} guaranteed by* (i). *Then γ is an Ehresmann connection on P_{nor}.*

(iii *Let*

$$\omega_M = \begin{pmatrix} 0 & 0 \\ \theta & \alpha \end{pmatrix}$$

be the form on P_{tan} guaranteed by (i). *Then (P_{tan}, ω_M) is a Euclidean geometry on M. Moreover, if (P, ω) arises from an immersion, then the Riemannian metric on M associated to (P_{tan}, ω_M) is the metric induced on M from N. Finally, (P_{tan}, ω_M) is torsion free if and only if (P, ω) is torsion free.*

Proof. (i) We deal with γ only, the cases of α and θ being similar. The transformation law in Definition 5.2(b) implies that for $h \in O_n(\mathbf{R})$, $R_h^*\gamma = \gamma$; moreover, 5.2(c) says that $\omega(X^\dagger) = X$ for $X \in \mathfrak{h}_\tau$, so in particular $\alpha(X^\dagger) = 0$ for $X \in \mathfrak{a}$. Thus, γ is basic for the principal bundle $P \to P/O_n(\mathbf{R}) = P_{\text{nor}}$, and since $\pi^*: A^1(P/O_n(\mathbf{R})) \to A^1(P)$ is injective, there is a unique form on P_{nor} pulling back to γ.

(ii) The transformation law in 5.2(b) implies that for $h \in O_r(\mathbf{R})$, $R_h^*\gamma = \operatorname{Ad}(h^{-1})\gamma$; moreover, 5.2(c) says that $\omega(X^\dagger) = X$ for $X \in \mathfrak{h}_\tau$, so in particular $\gamma(X^\dagger) = X$ for $X \in \mathfrak{u}$. Since the projection $P \to P/O_n(\mathbf{R})$ is $O_r(\mathbf{R})$ equivariant and $\pi^*: A^1(P/O_n(\mathbf{R}), \mathfrak{u}) \to A^1(P, \mathfrak{u})$ is injective, it follows that the formulas $R_h^*\gamma = \operatorname{Ad}(h^{-1})\gamma$ for $h \in O_r(\mathbf{R})$ and $\gamma(X^\dagger) = X$ for $X \in \mathfrak{o}(r)$ also hold for the version of γ on P_{nor}. Thus, γ satisfies the conditions for an Ehresmann connection (cf. Definition 2.4) on P_{nor}.

(iii) To see that ω_M is a trivialization of the tangent bundle of P_{tan}, since $\dim P_{\text{tan}} = \dim \mathfrak{a}$, it suffices to show that $\omega_M: T(P_{\text{tan}}) \to \mathfrak{a}$ is surjective, or, equivalently, that the form $\pi^*(\omega_M): T(P) \to \mathfrak{a}$ is surjective. But since $\pi^*(\omega_M)$ is just the canonical projection of ω on $\mathfrak{a}$, by condition 5.2(a) it is surjective. The proof that ω_M satisfies the conditions

(a) $R_h^*\omega_M = Ad(h^{-1})\omega_M$ for $h \in O_n(\mathbf{R})$ and

(b) $\omega_M(X^\dagger) = X$ for $X \in \mathfrak{b}$

is similar to the proof of (ii). Thus, ω_M is a Euclidean geometry on M.

Suppose that (P, ω) arises from an immersion $f: M \to N$. Let (P_N, ω_N) denote the Riemannian geometry on N and let $\varphi_{p_N}^N: T_y(N) \to \mathfrak{g}/\mathfrak{h}$ be the usual isomorphism, where $p_N \in P_N$ lies over $y \in N$. Since it is clear that $\varphi_p(v) = \varphi_{\bar{f}(p)}^N \circ f_{*p}(v)$, for any $v \in T_x(M)$ and for any $p \in P_{\text{tan}}$ covering x, the induced Riemannian metric $q_M: T(M) \to R$ arising from the immersion is given by

$$\begin{aligned} q_M(v) &= q_N(f_*(v)) \\ &= \langle \varphi_p^N(f_*(v)), \varphi_p^N(f_*(v)) \rangle_{\mathfrak{g}/\mathfrak{h}} \\ &= \langle \varphi_p(v), \varphi_p(v) \rangle_{\mathfrak{a}/\mathfrak{b}} \end{aligned}$$

where φ_p refers to the geometry $(P_{\text{tan}}, \omega_M)$.

The last statement of (iii) is obvious. ■

Exercise 5.13. Suppose that $G \to P \to M$ is any principal bundle and $H \subset G$ is a closed normal subgroup. Then G acts on P and P/H so that the principal bundle map $\pi: P \to P/H$ is G equivariant. Let $X \in \mathfrak{g}$. Show that the two versions of $X^\dagger$, one on P and one on P/H, correspond under π_*. (This is a point we glossed over in the proof of Proposition 5.12(i) and (ii).) ❑

At this stage we have two of the three pieces of the localized geometry (P, ω), namely the Riemannian geometry on M (corresponding to α) and the Ehresmann connection on the normal bundle (corresponding to γ). The remaining piece, which corresponds to β, is the second fundamental form.

The Second Fundamental Form

Proposition 5.14. *Let (P, ω) be a locally ambient geometry on M of codimension r and let v be the normal bundle associated to it. The form β on P given in Eq.* (5.11) *determines, and is determined by, the bilinear symmetric form $B \in Hom(T(M) \oplus T(M), v)$, which is given, independently of the choice of p lying over x, by*

$$\varphi_p(B_x(u, v)) = \beta(\omega_p^{-1}(\varphi_p(u)))\varphi_p(v), \quad \textit{where } u, v \in T_x(M).$$

Moreover, there are symmetric matrix-valued functions $h(i): P_\tau \to M_n(\mathbf{R})$ $(= End(\mathfrak{t}))$ defined by the equation $\beta_i = \theta^t h(i)$, where β_i is the ith row of β. These functions are related to B by the formula

$$\varphi_p(B(u, v)) = \begin{pmatrix} \varphi_p(u)^t h(1) \varphi_p(v) \\ \vdots \\ \varphi_p(u)^t h(r) \varphi_p(v) \end{pmatrix}.$$

Proof. First we show that the expression $\beta(\omega_p^{-1}(w_1))w_2$ is symmetric in $w_1, w_2 \in \mathfrak{t}$. The structural equation for ω_τ in the $(3,1)$ block is $\beta \wedge \theta = 0$. This implies that the ith row of β may be expressed as $\beta_i = \theta^t h(i)$, where $h(i)$ is an $r \times r$ matrix-valued function on P_τ. (We may also regard $h(i)$ as taking values in $\operatorname{End}(\mathfrak{t})$; cf. Lemma 5.21 ahead.) In fact, the condition $\beta_i \wedge \theta = 0$ for all i implies that $h(i)$ is a *symmetric* matrix. Thus, the expression

$$\beta(\omega_p^{-1}(w_1))w_2 = \begin{pmatrix} \theta^t(\omega_p^{-1}(w_1))h(1) \\ \vdots \\ \theta^t(\omega_p^{-1}(w_1))h(r) \end{pmatrix} w_2 = \begin{pmatrix} w_1^t h(1) w_2 \\ \vdots \\ w_1^t h(r) w_2 \end{pmatrix} \tag{5.15}$$

is symmetric in w_1 and w_2.

Now consider the function $b \in C(P_\tau, (S^2(\mathfrak{t}^*) \otimes \mathfrak{u}, \rho)) = S^2(\mathrm{Ad}^*) \otimes \mathrm{Ad}$, where ρ, given by

$$b(p)(w_1, w_2) = \beta(\omega_p^{-1}(w_1))w_2, \quad \text{where } w_i \in \mathfrak{t} \text{ for } i = 1, 2.$$

Let us show that the function b has the transformation properties indicated. For $h = \mathrm{diag}(A, D) \in O_n(\mathbf{R}) \times O_r(\mathbf{R})$, we have

$$\begin{aligned} b(ph)(w_1, w_2) &= \beta(\omega_{ph}^{-1}(w_1))w_2 \\ &= \beta(R_{h*}\omega_p^{-1}(\mathrm{Ad}(h)w_1))w_2 \\ &= (R_h^*\beta)(\omega_p^{-1}(\mathrm{Ad}(h)w_1))w_2 \\ &= D^{-1}\beta(\omega_p^{-1}(\mathrm{Ad}(h)w_1))Aw_2 \\ &= \mathrm{Ad}(h^{-1})b(p)(\mathrm{Ad}(h)w_1, \mathrm{Ad}(h)w_2) = (\rho(h^{-1})b(p))(w_1, w_2). \end{aligned}$$

Clearly, β determines and is determined by b.

Note that $b(p)(\varphi_p(v_1), \varphi_p(v_2)) \in \mathfrak{u}$. Since $\varphi_p^{-1}(\mathfrak{u}) = \nu_x =$ the normal space at $x \in M$, it follows that the formula defining $B_x(v_1, v_2)$ makes sense; we must also show that it is independent of the choice of p lying over x. But this follows immediately from the transformation laws for b and for φ. Finally, we note that since φ_p is an isomorphism for each p, B and b determine each other.

As for the last formula, we have

$$\begin{aligned} \varphi_p(B(u, v)) = b(\varphi_p(u), \varphi_p(v)) &= \beta(\omega_p^{-1}(\varphi_p(u)))\varphi_p(v) \\ &= \begin{pmatrix} \varphi_p(u)^t h(1) \varphi_p(v) \\ \vdots \\ \varphi_p(u)^t h(r) \varphi_p(v) \end{pmatrix}. \end{aligned}$$

■

With Proposition 5.14 we have completed the decomposition of a locally ambient geometry into its three parts as mentioned in the introduction to this section. Moreover it should be clear that the original locally ambient geometry can be unambiguously reconstructed from these parts.

The following exercise describes the mean curvature associated to a locally ambient geometry.

Exercise 5.16. Let $e_1, e_2, \ldots, e_n$ be an orthonormal basis for $T_x(M)$. Show that the normal vector $\mathrm{Tr}\, B = \sum_i B(e_i, e_i)$ is independent of the choice of this basis. (The normal vector $\frac{1}{n}\mathrm{Tr}\, B$ is called the *mean curvature vector at* $x \in M$. Its length is called the *mean curvature* at x.) ❑

Now we pass to a study of the Weingarten maps associated to a locally ambient geometry.

Definition 5.17. Let (P, ω) be a locally ambient geometry on M of codimension r. The *second fundamental form* associated to (P, ω) is the bilinear symmetric form $B \in \mathrm{Hom}(T(M) \oplus T(M), \nu)$ given in Proposition 5.14. Given a normal vector $X \in \nu_x$, $x \in M$, the corresponding *Weingarten map* or *shape operator* $L_X \in \mathrm{End}(T_x(M))$ is defined by $\langle X, B(u, v)\rangle_\nu = \langle u, L_X(v)\rangle_M$. ❀

Lemma 5.18. *L_X is a self-adjoint map depending linearly on X.*

Proof. The calculation

$$\begin{aligned} \langle u, L_X(v)\rangle = \langle X, B(u, v)\rangle &= \langle X, B(v, u)\rangle \\ &= \langle v, L_X(u)\rangle = \langle L_X(u), v\rangle \end{aligned}$$

shows that L_X is self-adjoint. For the linearity in X, we have

$$\begin{aligned} \langle u, L_{aX+bY}(v)\rangle &= \langle aX + bY, B(u, v)\rangle = a\langle X, B(u, v)\rangle + b\langle Y, B(u, v)\rangle \\ &= a\langle u, L_X(v)\rangle + b\langle u, L_Y(v)\rangle = \langle u, aL_X(v) + bL_Y(v)\rangle, \end{aligned}$$

and so $L_{aX+bY} = aL_X + bL_Y$. ■

Since the endomorphisms L_X are self-adjoint, all their eigenvalues are real. This leads to the following definition.

Definition 5.19. The eigenvalues of $L_X \in \mathrm{End}(T_x(M))$ are called the *principal curvatures* associated to the normal vector $X \in \nu_x(M)$. ❀

Definition 5.20. The principal curvatures of $M \subset N$ are *constant* if, for any curve $\sigma: I \to M$ and any parallel normal field X on M defined along σ, the set of eigenvalues of L_X is independent of the point on σ where they are calculated. ❀

Next we compare the Weingarten maps to the matrices $h(i)$.

Lemma 5.21.

(i) $L_X(v) = \sum_{1 \le k \le r} \varphi_p(X)_k \varphi_p^{-1}(h(k)\varphi_p(v))$.

(ii) $h(i) = \varphi_p \circ L_{\varphi_p^{-1}(e_i)} \circ \varphi_p^{-1} \in End(\mathfrak{t})$.

Proof. We have

$$\begin{aligned}\langle u, L_X(v)\rangle_M &= \langle X, B(u,v)\rangle_\nu = \langle \varphi_p(X), \varphi_p(B(u,v))\rangle_{\mathfrak{u}} \\ &= \langle \varphi_p(X), b(p)(\varphi_p(u), \varphi_p(v))\rangle_{\mathfrak{u}} \\ &= \left\langle \varphi_p(X), \begin{pmatrix} \varphi_p(u)^t h(1)\varphi_p(v) \\ \vdots \\ \varphi_p(u)^t h(r)\varphi_p(v)\end{pmatrix}\right\rangle_{\mathfrak{u}}.\end{aligned}$$

Thus,

$$\begin{aligned}\langle \varphi_p(u), \varphi_p(L_X(v))\rangle_{\mathfrak{t}} &= \sum_{1\le k\le r} \varphi_p(X)_k \varphi_p(u)^t h(k)\varphi_p(v) \\ &= \langle \varphi_p(u), \sum_{1\le k\le r} \varphi_p(X)_k h(k)\varphi_p(v)\rangle_{\mathfrak{t}}.\end{aligned}$$

This yields (i). Taking $X = \varphi_p^{-1}(e_i)$ (where e_i, $1 \le i \le n$, is the standard basis of $\mathfrak{t}$) in (i), we get $L_X(v) = \varphi_p^{-1}(h(i)\varphi_p(v))$, which is (ii). ■

Corollary 5.22. *The Weingarten maps all commute with each other if and only if all the $h(i)$ commute with each other.*

Principal Curvatures for Relatively Flat Normal Bundles

Definition 5.23. Let N be a Riemannian manifold. An immersion $f: M \to N$ has a *relatively flat* normal bundle if $\beta \wedge \beta^t = 0$ or, equivalently, $d\gamma + \gamma \wedge \gamma = \Omega$, where Ω is the curvature entry on P_τ of the $(3,3)$ block of the locally ambient geometry. ❀

Example 5.24. A codimension-1 submanifold of Euclidean space is automatically relatively flat since $\Omega = 0$ in a flat geometry, and $\gamma = 0$ since γ is a 1×1 skew matrix. More generally, a codimension-r submanifold of Euclidean space is relatively flat if and only if the Ehresmann connection on the normal bundle (cf. Definition 2.4) is flat (i.e., has curvature zero). ◆

Proposition 5.25. *The following statements are equivalent:*

(i) *$M \subset N$ has a relatively flat normal bundle.*

(ii) *$h(i)h(j) = h(j)h(i)$ for all $1 \le i, j \le r$.*

(iii) *The Weingarten maps L_X all commute with each other.*

Proof. (i) $\Leftrightarrow$ (ii). By definition, (i) is equivalent to $\beta \wedge \beta^t = 0$. Thus, this equivalence is the same as

$$(\beta \wedge \beta^t)_{ij} = 0 \Leftrightarrow h(i) \text{ and } h(j) \text{ commute.}$$

We have

$$\begin{aligned}(\beta\wedge\beta^t)_{ij} &= \beta_i \wedge \beta^t_j = \theta^t h(i) \wedge (\theta^t h(j))^t \quad \text{(cf. Proposition 5.14)} \\ &= \theta^t h(i) \wedge h(j)\theta = \sum_{krl} \theta_k h(i)_{kr} \wedge h(j)_{rl}\theta_l \\ &= \sum_{k<l}\left(\sum_r (h(i)_{kr}h(j)_{rl} - h(i)_{kr}h(j)_{rl})\right)\theta_k\wedge\theta_l \\ &= \sum_{k<l} [h(i), h(j)]_{kl}\theta_k \wedge \theta_l.\end{aligned}$$

Since $h(i)$ and $h(j)$ are symmetric matrices, it follows that $[h(i), h(j)]$ is skew symmetric. Hence,

$$(\beta\wedge\beta^t)_{ij} = 0 \Leftrightarrow [h(i), h(j)]_{kl} = 0 \ \text{ for all } k < l \Leftrightarrow [h(i), h(j)] = 0.$$

The equivalence (ii) $\Leftrightarrow$ (iii) is just Corollary 5.22. ■

Corollary 5.26. *Suppose $M \subset N$ has a relatively flat normal bundle. Define*

$$\Lambda_x = \{\lambda \in \nu_x^* \mid L_X v = \lambda(X)v \textit{ for some } v \in T_x(M) - \{0\} \textit{ and all } X \in \nu_x\},$$

and set

$$T_x(M)_\lambda = \{v \in T_x(M) \mid L_X v = \lambda(X)v \ \textit{ for all } X \in \nu_x\}.$$

Then $T_x(M) = \bigoplus_{\lambda\in\Lambda_x} T_x(M)_\lambda$ is an orthogonal decomposition of the tangent space.

Proof. Since the collection of all Weingarten maps is a commutative subset of $\mathrm{End}(T_x(M))$, we know from linear algebra that they are simultaneously diagonalizable (cf., e.g., [N. Jacobson, 1953], pp. 132–133). This gives a decomposition $T_x(M) = \oplus E_i$ such that each E_i is an eigenspace for every Weingarten map. Let $\lambda_i(X)$ denote the eigenvalue of L_X on E_i. Then the linearity of L_X in X implies that $\lambda_i \in \nu_x^*$. Thus, $\lambda_i \in \Lambda_x$ and $E_i \subset T_x(M)_{\lambda_i}$. Since $T_x(M)_\lambda \cap T_x(M)_\mu = 0$ for $\lambda \ne \mu$ and $T_x(M) = \oplus E_i \subset \oplus T_x(M)_{\lambda_i}$, the direct-sum decomposition follows. To see that it is orthogonal, choose

$$u \in T_x(M)_\mu,\ v \in T_x(M)_\lambda,\ \lambda \ne \mu.$$

Then there is a vector $X \in \nu_x$ with $\lambda(X) \ne \mu(X)$ so that

$$\begin{aligned}(\lambda(X) - \mu(X))\langle u, v\rangle &= \langle \lambda(X)u, v\rangle - \langle u, \mu(X)v\rangle \\ &= \langle L_X u, v\rangle - \langle u, L_X v\rangle = 0.\end{aligned}$$

Thus, $\langle u, v\rangle = 0$. ■

Definition 5.27. Suppose that $f: M \to N$ has a relatively flat normal bundle. Referring to Corollary 5.26 we have the following:

(i) The elements of Λ_x are called the *dual principal curvatures*[4] at x.

(ii) The eigenspaces $T_x(M)_\lambda \subset T_x(M)$, $\lambda \in \Lambda_x$ given in the proof of Corollary 5.26 are called the *principal subspaces* (or *directions* if they are one-dimensional) of $T_x(M)$ at x.

(iii) The decomposition $T_x(M) = \oplus T_x(M)_\lambda$ is called *the principal curvature decomposition.*

(iv) The normal vector $X_\lambda \in T_x(M)^\perp$ given by $\langle X_\lambda, v\rangle = \lambda(v)$ for all $v \in T_x(M)^\perp$ is called the *mean curvature vector associated to* $\lambda \in \Lambda$ (or to the principal subspace $T_x(M)_\lambda \subset T_x(M)$). ❀

Exercise 5.28.* Let $M \subset N$ have a relatively flat normal bundle.

(i) Show that $B(u,v) = \langle u,v\rangle X_\lambda$ for all $u, v \in T_x(M)_\lambda$.

(ii) Assume that Λ_x contains a constant number of elements for $x \in U$, a simply connected open set in M. Show that there is a canonical identification among the Λ_x, $x \in U$, and that under this identification the X_λ are smooth normal vector fields along U.

(iii) Deduce from (ii) that, for each $\lambda \in \Lambda_x$, the $T_x(M)_\lambda$ yield a distribution along U.

(iv) Show that the mean curvature vector is

$$\frac{1}{n}\sum_{\lambda\in\Lambda} n_\lambda X_\lambda, \quad \text{where } n_\lambda = \dim T(M)_\lambda,$$

and the mean curvature μ is given by $\mu^2 = (1/n^2)\sum_{\kappa,\lambda\in\Lambda} n_\kappa n_\lambda \langle X_\kappa, X_\lambda\rangle$. ❑

This exercise allows us to identify the elements in different Λ_x for x varying in a small, open set when $\#\Lambda_x$ is constant. In fact, more is true in this case.

Proposition 5.29. *Suppose the immersion $f: M \to N$ has a relatively flat normal bundle and the function $\#\Lambda_x: M \to Z$ is constant. Then*

(i) *$n_\lambda = \dim T_x(M)_\lambda$ is independent of x.*

Choose any orthogonal decomposition $\mathfrak{t} = \bigoplus_{\lambda\in\Lambda}\mathfrak{t}_\lambda$, where $\dim \mathfrak{t}_\lambda = n_\lambda$.

(ii) *There is a canonical reduction of $O_n(\mathbf{R}) \times O_r(\mathbf{R}) \to P_\tau \to M$ to a principal bundle*

$$(\times_{\lambda\in\Lambda} O_{n_\lambda}(R)) \times O_r(\mathbf{R}) \to P_\Lambda \to M$$

[4] In the case of a hypersurface, the numerical values of the elements of Λ_x on the unique (up to sign) unit normal vector are what are usually known as the principal curvatures.

given by

$$P_\Lambda = \{p \in P_\tau \mid \varphi_p(T_x(M)_\lambda) = \mathfrak{t}_\lambda \textit{ for all } \lambda \in \Lambda, \textit{ where } x = \pi(p)\}.$$

(iii) *Regarding $h(k)$ as taking values in $End(\mathfrak{t})$, we may also describe P_Λ as*

$$P_\Lambda = \{p \in P_\tau \mid h(k) \mid \mathfrak{t}_k = \lambda(\varphi_p^{-1}(e_k))id_{\mathfrak{t}_\lambda} \textit{ for all } \lambda \in \Lambda \textit{ and } 1 \le k \le r\}.$$

Set $a_{si} = \lambda_i(\varphi_p^{-1}(e_s))$, $1 \le i \le n$, $n+1 \le s \le n+r$, where $\lambda_i \in \Lambda$ is determined by the condition $e_i \in \mathfrak{t}_{\lambda_i}$.

(iv) *The $\times_{\lambda\in\Lambda} O_{n_\lambda}(R)$ bundle $P_\Lambda/(1 \times O_r(\mathbf{R})) \to M$ is a reduction of the bundle*

$$O_n(\mathbf{R}) \to P_{tan} \to M.$$

(v) *The $(2,2)$ block of the curvature function for P_{tan} (i.e., for the submanifold M) is given by*

$$K_{2,2} + \sum_{1\le i,j\le n} \langle X_i, X_j\rangle e_i^* \wedge e_j^* \otimes e_{ij},$$

where X_i is the normal vector given by $\langle X_i, -\rangle = \lambda$, where i is determined by the condition $e_i \in \mathfrak{t}_\lambda$. $K_{2,2}$ is the $(2,2)$ block of the locally ambient curvature function restricted to P_Λ. It is $O_r(\mathbf{R})$ invariant and so is basic for the bundle $P_\Lambda \to P_\Lambda/(1 \times O_r(\mathbf{R})) = P_{tan}$. (In particular, the multiplicities of the Xs correspond to the multiplicities of the λs.)

Proof. (i) This follows immediately from Exercise 5.28(iii).

(ii) We shall only verify that the fiber is as stated. Let $p \in P_\Lambda$ and $h \in O_n(\mathbf{R}) \times O_r(\mathbf{R})$. Then

$$\begin{aligned} ph \in P_\Lambda &\Leftrightarrow \varphi_{ph}(T_x(M)_\lambda) = \mathfrak{t}_\lambda \text{ for all } \lambda \in \Lambda \\ &\Leftrightarrow \mathrm{Ad}(h^{-1})\varphi_p(T_x(M)_\lambda) = \mathfrak{t}_\lambda \text{ for all } \lambda \in \Lambda \\ &\Leftrightarrow \mathrm{Ad}(h^{-1})\mathfrak{t}_\lambda = \mathfrak{t}_\lambda \text{ for all } \lambda \in \Lambda \\ &\Leftrightarrow h \in (\times_{\lambda\in\Lambda} O_{n_\lambda}(R)) \times O_r(\mathbf{R}). \end{aligned}$$

(iii) Assume that $p \in P_\Lambda$. If $v \in \mathfrak{t}_\lambda$, then

$$h(k)v = \varphi_p \circ L_{\varphi_p^{-1}(e_k)} \circ \varphi_p^{-1}(v) = \varphi_p(\lambda(\varphi_p^{-1}(e_k))\varphi_p^{-1}(v)) = \lambda(\varphi_p^{-1}(e_k))v,$$

and so

$$P_\Lambda \subset \{p \in P_\tau \mid h(k)|\mathfrak{t}_\lambda = \lambda(\varphi_p^{-1}(e_k))\mathrm{id}_{\mathfrak{t}_\lambda} \text{ for all } \lambda \in \Lambda \text{ and } 1 \le k \le r\}.$$

The reverse inclusion is similar.

(iv) The inclusion $P_\Lambda \subset P_\tau$ induces an inclusion $P_\Lambda/O_r(\mathbf{R}) \subset P_\tau/O_r(\mathbf{R}) = P_{\tan}$.

(v) We have

$$h(i) = \begin{pmatrix} a_{i1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & a_{in} \end{pmatrix} \quad \text{so } \beta = \begin{pmatrix} a_{11}\theta_1 & \cdots & a_{1n}\theta_n \\ \vdots & \ddots & \vdots \\ a_{n1}\theta_1 & \cdots & a_{nn}\theta_n \end{pmatrix}$$

and

$$\beta^t \wedge \beta = \left(\left(\sum_{n+1 \le s \le n+r} a_{si} a_{sj} \right) \theta_i \wedge \theta_j \right).$$

Now $a_{si} = \lambda_i(\varphi_p^{-1}(e_s)) = \langle X_i, \varphi_p^{-1}(e_s)\rangle$ is the sth component of X_i in the orthonormal basis $\varphi_p^{-1}(e_s)$, $n+1 \le s \le n+r$, for $T_x(M)^\perp$. Thus,

$$\sum_{n+1 \le s \le n+r} a_{si} a_{sj} = \langle X_i, X_j \rangle.$$

From Eq. (5.1), the (2,2) block of the curvature on P_Λ is $\Omega_{22} = d\alpha + \alpha \wedge \alpha - \beta^t \wedge \beta$, from which the result easily follows. ■

Definition 5.30. The reduction of $O_n(\mathbf{R}) \to P_{\tan} \to M$ described in Proposition 5.29(iv) is called the *principal curvature reduction*. ❀

Definition 5.31. Let $P \to M$ be a principal bundle and $Q \subset P$ be a reduction of it. Let $\mathcal{L}$ be a distribution on P. We say that the *reduction is compatible with the distribution* if, for each point $q \in Q$, we have $\mathcal{L}_q \subset T_q(Q)$. ❀

In general, there is no reason why the principal curvature reduction should be compatible with the Levi–Civita connection.

Example 5.32. Consider the case of an umbilic-free surface M in 3-space with no flat points. The curvature reduction $Q \subset P_{\tan}$ is a principal bundle with group $O_1(\mathbf{R}) \times O_1(\mathbf{R}) = Z_2 \times Z_2$ and so is a covering space of the surface. Compatibility with the Levi–Civita connection would mean that the Levi–Civita connection form α would vanish on Q. Choosing a local section of Q would yield a gauge in which α vanishes, and hence so would the curvature. This would imply that M is flat. ◆

For a submanifold $M \subset N$ with a relatively flat normal bundle, the condition of constant principal curvatures has a special significance in that the vector fields X_λ are parallel.

Lemma 5.33. *Let $M \subset N$ be a submanifold of a Riemannian manifold. Then the following conditions are equivalent:*

(i) *The principal curvature sections $\lambda \in \Lambda$ of the normal bundle are parallel along any curve.*

(ii) *The principal curvatures of M are constant.*

(iii) *The normal vector fields X_λ, $\lambda \in \Lambda$, are parallel along any curve.*

Proof. Let σ be a curve on M and let Z be its tangent field. Let X be any parallel normal field along σ (i.e., normal to M) and let $\lambda \in \Lambda$.

(i) $\Leftrightarrow$ *(ii)*

λ is parallel along $\sigma \Leftrightarrow D_Z\lambda = 0$ along σ

$$\Leftrightarrow \left\{ \begin{array}{c} Z(\lambda(X)) = 0 \text{ along } \sigma \\ \text{for all } X \text{ parallel along } \sigma \end{array} \right\} (\text{since } Z(\lambda(X)) = (D_Z\lambda)(X) + \lambda(\underbrace{D_Z X}_{0}))$$

$$\Leftrightarrow \left\{ \begin{array}{c} \lambda(X) \text{ constant along } \sigma \\ \text{for all } X \text{ parallel along } \sigma \end{array} \right\}.$$

(ii) $\Leftrightarrow$ *(iii)*

$$\left\{ \begin{array}{c} \lambda(X) \text{ constant along } \sigma \\ \text{for all } X \text{ parallel along } \sigma \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{c} \langle X_\lambda, X\rangle \text{ constant along } \sigma \\ \text{for all } X \text{ parallel along } \sigma \end{array} \right\}$$

$$\Leftrightarrow X_\lambda \text{ parallel along } \sigma.$$ ■

§6. Isoparametric Submanifolds

Elie Cartan [E. Cartan, 1938] uses the term *isoparametric* ("same parameters") to refer to hypersurfaces $M^n \subset Q_c^{n+1}$ having constant principal curvatures (these are the "parameters" of the submanifold) in a space Q_c^{n+1} of constant curvature c.

Definition 6.1. A submanifold $M \subset N$ of a Riemannian manifold is *isoparametric*[5] if it has constant principal curvatures and its normal bundle is relatively flat. ❀

Isoparametric submanifolds of Riemannian space forms have been studied extensively in the past decade or so (see [C.-L. Terng, 1993] and the references given there). Cartan showed that isoparametric hypersurfaces of Euclidean space are spheres, hyperplanes, or a product of the two. He also

[5]In the more usual definition, one assumes that N is a Riemannian space form and that the normal bundle of $M \subset N$ is flat. However, our definition generalizes this one since the flatness follows from the commuting of the Weingarten maps in this case; cf. [C.-L. Terng, 1993].

showed that in hyperbolic space isoparametric hypersurfaces are totally umbilic, while in spheres there are many examples. Thorbergson [G. Thorbergson, 1991] has shown that any isoparametric submanifold $(\mathbf{R}^{n+r}, M^n)$ of codimension $r \geq 3$ in Euclidean space is isometric to a pair $(T_e(G/H), \mathrm{Ad}(H)v)$, where $v \in T_e(G/H)$, for some symmetric space G/H. For hypersurfaces in spheres ([H. Ozeki and M. Takeuchi, 1975 and 1976]) and for codimension 2 in Euclidean space [D. Ferus, H. Karcher, and H. Munzner, 1981], there are examples of isoparametric submanifolds that are not locally symmetric.

We shall study only the case of hypersurfaces. Let $M^n \subset Q_c^{n+1}$ be a hypersurface in a Riemannian space form of curvature c. Let us choose a section of the principal curvature reduction (Proposition 5.29(iv)) so that the 1-forms that make up the Cartan connection on P pull back to M to yield a gauge of the form

$$\begin{pmatrix} 0 & 0 & 0 \\ \theta & \omega & -h\theta \\ 0 & \theta^t h & 0 \end{pmatrix},$$

where $h = h(1) = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, the diagonal matrix of principal curvatures. The structural equation is then

$$\begin{pmatrix} 0 & 0 & 0 \\ d\theta + \omega \wedge \theta & d\omega + \omega \wedge \omega - h\theta \wedge \theta^t h & -d(h\theta) - \omega \wedge (h\theta) \\ 0 & \star & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & c\theta \wedge \theta^t & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{6.2}$$

whence, in particular,

$$\begin{aligned} 0 &= -d(h\theta) - \omega \wedge h\theta \\ &= -(dh) \wedge \theta - hd\theta - \omega \wedge h\theta \\ &= -(dh) \wedge \theta + h\omega \wedge \theta - \omega \wedge h\theta. \end{aligned} \tag{6.3}$$

Umbilic Hypersurfaces

Our first application of Eq. (6.3) is to the case when all the eigenvalues of h are equal.

Definition 6.4. Let $M^n \subset Q_c^{n+1}$ be a hypersurface in a Riemannian space form of curvature c. M is called an *umbilic hypersurface* if the principal curvatures of M are all equal at each point. ❀

Proposition 6.5. *If $n > 1$, the principal curvature of a connected umbilic hypersurface $M^n \subset Q_c^{n+1}$ is constant on M. Moreover, M is a Riemannian space form of curvature $c + \lambda^2$.*

Proof. In this case h is a multiple of the identity, so $h\omega \wedge \theta = \omega \wedge h\theta$. Thus, Eq. (6.3) becomes $(d\lambda I) \wedge \theta = 0$, or $d\lambda \wedge \theta_i = 0$ for $1 \leq i \leq n$. Thus, there exist functions a_i such that $d\lambda = a_i\theta_i$ for each i. Since $n \geq 2$, it follows that $a_2\theta_2 - a_1\theta_1 = 0$. Hence, $a_2 = a_1 = 0$ and $d\lambda = 0$. Thus, λ is constant. The induced geometry has Cartan connection $\begin{pmatrix} 0 & 0 \\ \theta & \omega \end{pmatrix}$ with curvature form

$$\begin{pmatrix} 0 & 0 \\ d\theta + \omega \wedge \theta & d\omega + \omega \wedge \omega \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & (\lambda^2 + c)\theta \wedge \theta^t \end{pmatrix},$$

where the last equality comes from the structural equation (6.2). ■

Cartan's Formula for Isoparametric Hypersurfaces

We return to the case where M is an isoparametric hypersurface in a space. Then the diagonal matrix h is constant over M. Equation (6.3) now reads $h\omega \wedge \theta = \omega \wedge h\theta$ or, in coordinate form,

$$\sum_j (\lambda_i - \lambda_j)\omega_{ij} \wedge \theta_j = 0 \quad \text{for } i = 1, \ldots, n. \tag{6.6}$$

Let us write $\omega_{ij} = \sum_k g_{ijk}\theta_k$ (since the θs constitute a basis). Since $\omega_{ij} = -\omega_{ji}$ it follows that $g_{ijk} = -g_{jik}$. Putting this in Eq. (6.6) yields

$$\sum_{jk} (\lambda_i - \lambda_j) g_{ijk}\theta_k \wedge \theta_j = 0 \quad \text{for } i = 1, \ldots, n,$$

or

$$(\lambda_i - \lambda_j) g_{ijk} = (\lambda_i - \lambda_k) g_{ikj} \quad \text{for all } i, j, k = 1, \ldots, n. \tag{6.7}$$

Lemma 6.8. *Let $a_{ijk} = (\lambda_i - \lambda_j) g_{ijk}$. Then a_{ijk} is fully symmetric in all its indices and it vanishes if two indices are equal.*

Proof. Equation (6.7) shows that a_{ijk} is symmetric in the last two indices. Moreover, both g_{ijk} and $(\lambda_i - \lambda_j)$ are skew symmetric in the indices i and j, so a_{ijk} is also symmetric in the first two indices. It follows that a_{ijk} is symmetric in all the indices. Since a_{ijk} obviously vanishes if the first two indices are equal, by symmetry it must vanish if any two indices are equal. ■

Corollary 6.9. *If either* (i) $\lambda_i \neq \lambda_j = \lambda_k$ *or* (ii) $\lambda_i = \lambda_j \neq \lambda_k$, *then* $g_{ijk} = 0$.

Proof. (i) $\lambda_j = \lambda_k \Rightarrow a_{jki} = 0 \Rightarrow a_{ijk} = 0 \Rightarrow g_{ijk} = 0$ (using $\lambda_i \neq \lambda_j$). The other case is similar. ■

Next, consider the exterior derivative of the equation $\omega_{ij} = \sum_k g_{ijk}\theta_k$. It is

$$\begin{aligned} d\omega_{ij} &= \sum_k dg_{ijk} \wedge \theta_k + \sum_k g_{ijk} d\theta_k \\ &= \sum_k dg_{ijk} \wedge \theta_k - \sum_{kl} g_{ijk}\omega_{kl} \wedge \theta_l \\ &= \sum_k dg_{ijk} \wedge \theta_k - \sum_{klr} g_{ijk} g_{klr} \theta_r \wedge \theta_l, \end{aligned}$$

where we use the $(2,1)$ block of the structural equation (6.2) for the second equality. From the $(2,2)$ block of the same equation, we get, for all $1 \leq i, j \leq n$,

$$c\theta_i \wedge \theta_j = \sum_k dg_{ijk} \wedge \theta_k - \sum_{klr} g_{ijk} g_{klr} \theta_r \wedge \theta_l + \sum_{klr} g_{irk} g_{rjl} \theta_k \wedge \theta_l - \lambda_i\lambda_j \theta_i \wedge \theta_j.$$

The only terms in $\sum_k dg_{ijk} \wedge \theta_k$ that can contain $\theta_i \wedge \theta_j$ are $dg_{iji} \wedge \theta_i + dg_{ijj} \wedge \theta_j$. If $\lambda_i \neq \lambda_j$, these terms vanish by Corollary 6.9, and equating the coefficients of $\theta_i \wedge \theta_j$ will yield

$$c = -\sum_k g_{ijk}(g_{kji} - g_{kij}) + \sum_r (g_{iri} g_{rjj} - g_{irj} g_{rji}) - \lambda_i \lambda_j$$

or

$$\lambda_i \lambda_j + c = \sum_k (-g_{ijk} g_{kji} + g_{ijk} g_{kij} - g_{ikj} g_{kji}) \quad (\text{since } g_{iri} = 0). \tag{6.10}$$

Lemma 6.11. *If $\lambda_i \neq \lambda_j$, the only terms giving a nonzero contribution to the sum in Eq.* (6.10) *are those for which the index k satisfies $\lambda_k \neq \lambda_i$ and $\lambda_k \neq \lambda_j$.*

Proof. Consider a term $-g_{ijk} g_{kji} + g_{ijk} g_{kij} + g_{iki} g_{kjj} - g_{ikj} g_{kji}$ for which $\lambda_k = \lambda_i \neq \lambda_j$. By Corollary 6.9, we have $g_{ijk} = g_{kjj} = g_{ikj} = 0$, so the term vanishes. The case of $\lambda_k \neq \lambda_i = \lambda_j$ is similar. ∎

Thus, the only nonzero terms in the sum in Eq. (6.10) arise from indices k for which all three of λ_i, λ_j, and λ_k are distinct. In particular, each of the gs appearing in these terms of the sum may be expressed in terms of the as (Lemma 6.8), and we get

$$\begin{aligned} \lambda_i \lambda_j + c = - \sum_{\substack{\text{all } k \text{ such that} \\ \lambda_k \neq \lambda_i, \lambda_j}} a_{ijk}^2 \Big(&-\frac{1}{\lambda_i - \lambda_j} \frac{1}{\lambda_k - \lambda_j} \\ &+ \frac{1}{\lambda_i - \lambda_j} \frac{1}{\lambda_k - \lambda_i} - \frac{1}{\lambda_i - \lambda_k} \frac{1}{\lambda_k - \lambda_j} \Big). \end{aligned}$$

This formula quickly simplifies to yield *Cartan's formula*,

$$\lambda_i \lambda_j + c = -2 \sum_{\substack{\text{all } k \text{ such that} \\ \lambda_k \neq \lambda_i, \lambda_j}} \frac{a_{ijk}^2}{(\lambda_i - \lambda_k)(\lambda_j - \lambda_k)}.$$

Proposition 6.12. *If an isoparametric hypersurface in a space form of curvature c has only two distinct principal curvatures, λ_1 and λ_2 say, then $\lambda_1 \lambda_2 = -c$.*

Proof. The right-hand side of Cartan's formula vanishes. ∎

In [E. Cartan, 1938] other conclusions are drawn from this formula. In particular, it is shown that a connected isoparametric hypersurface in $M^m \subset Q_c^{m+1}$ has at most two eigenvalues if $c \leq 0$.

7
Möbius Geometry

Hermann Weyl defined a conformal structure on a manifold M as an equivalence class of Riemannian metrics, where two metrics are equivalent if they differ by a factor that is a smooth positive function on M. Suppose M is equipped with a conformal structure and $u, v \in T_x M$. Then the angle between u and v makes sense, as does the ratio of the lengths of u and v, but the lengths themselves have no meaning.

In Elie Cartan's view, a "conformal structure" on M is a Cartan geometry on M modeled on the Klein geometry whose space is the n-sphere and whose principal group is the full group of conformal (i.e., angle preserving) transformations. This group turns out to be the projective Lorentz group in dimension $n+2$. This is the geometry that we call the Möbius geometry.

Weyl also considered a structure on M intermediate between the two described above which amounts to a Cartan geometry on M modeled on $\mathbf{R}^n$ with its full group of conformal transformations as the principal group, which is the group of Euclidean similarities. This is the geometry we call the Weyl geometry. If we regard S^n as being $\mathbf{R}^n \cup \{\infty\}$ (via stereographic projection), then the subgroup of the projective Lorentz group which fixes ∞ is the principal group of the Weyl model.

A recent survey and overview of many results about conformal geometry and its generalizations from a somewhat different point of view than ours may be found in [M.A. Akivis and V.V. Goldberg, 1996].

In §1 we study the Möbius and Weyl model geometries. In §2 we pass to the consideration of Cartan geometries with these models. In particular, we show that every Möbius (or Weyl) geometry on a manifold M determines a conformal structure on M. We also study some basic properties of Möbius and Weyl geometries, including special curvatures and normal geometries. In §3 we solve the equivalence problem for conformal structures. That is, for manifolds of dimension >2, we show that to a given conformal structure there corresponds a unique normal Möbius geometry giving rise to the original conformal structure as in §2 (Proposition 3.1). Since a Riemannian structure determines a conformal structure, which in turn determines a normal Möbius geometry, the curvature of this latter geometry is a conformal invariant of the original Riemannian geometry. This turns out to be the Weyl curvature (Table **6**.2.5) of the Riemannian geometry, and so we have a conceptual proof for the conformal invariance of this part of the Riemannian curvature (Theorem 3.8). The section ends by solving the equivalence problem for Weyl structures (Definition 3.11) in terms of Weyl geometries (Proposition 3.14). The following figure summarizes the relationships among these geometries in dimensions >2.

$$\left\{\begin{matrix}\text{conformal structure}\\ \textit{or}\\ \text{normal}\\ \text{Möbius geometry}\end{matrix}\right\} \xrightarrow{\text{generalization}} \left\{\begin{matrix}\text{Weyl structure}\\ \textit{or}\\ \text{Weyl geometry}\\ \textit{or}\\ \text{special curvature}\\ \text{Möbius geometry}\end{matrix}\right\} \xrightarrow{\text{generalization}} \left\{\begin{matrix}\text{general}\\ \text{Möbius}\\ \text{geometry}\end{matrix}\right\}$$

Conversely (again for dimension >2), an arbitrary Möbius geometry canonically determines a normal Möbius geometry, but of course this correspondence is not one to one.

In §4 we study immersions of submanifolds in normal Möbius geometries, showing how the ambient geometry localizes along the immersion to give data (a "locally ambient geometry"; cf. Definition 4.8 and Lemma 4.14) determining it. We then study how the localized geometry decomposes into tangential, normal, and gluing data[1] (see Propositions 4.18, 4.20, and 4.21), and we characterize such data in simple cases. If the submanifold is not a surface, then these three data determine the locally ambient geometry from which they arise.

For a surface immersed in $\mathbf{R}^n$, we apply this to study the relation between Riemannian and Möbius geometries. Although the conformal structure inherited by this surface is necessarily flat, the (nonnormal) Möbius geometry it inherits is not flat. When $n = 3$, its curvature may be identified with the Willmore form $(H^2 - K)dA$. In Riemannian geometry, the Gauss curvature K is regarded as intrinsic (since it can be calculated from the metric induced on the surface) while the mean curvature is regarded as extrinsic. Thus, the Willmore form, while appearing to be *extrinsic* data from the Riemannian perspective, is seen to be *intrinsic* data from the perspective of the Möbius geometry[2] induced on M. Section 4 ends with a brief comparison with Aaron Fialkow's work on the classification of submanifolds.

[1] I.e., second fundamental form data.

In §5 we classify the immersions of vertex free curves in normal geometries of two and three dimensions. The method here is alternative to, and less general than, the one given in §4, but the form of the answer is simpler.

In §6 we use the method of §5 to study the class of umbilic free surfaces in the conformal three sphere. This gives an approach that is an alternate to the one given in §4 and avoids the difficulty with the exceptional case of surfaces referred to above.

Throughout this chapter, $(G, H) = (\text{Möb}_n(\mathbf{R}), \text{Möb}_n(\mathbf{R})_0)$ denotes the Klein geometry of the Möbius group and its subgroup (see Definition 1.7); $(\mathfrak{g}, \mathfrak{h})$ denotes the corresponding pair of Lie algebras (see Exercise 1.13); M_0 will denote (a space canonically identified with) the homogeneous space G/H (see Definition 1.2); and M itself will denote the space of an arbitrary Cartan geometry modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H.

§1. The Möbius and Weyl Models

In this section we study the standard model of Möbius n space which, as we will see, may be identified with the n-sphere S^n together with its group G consisting of the group of all conformal transformations. To describe the group G and its action, we will regard S^n as the space of one-dimensional lightlike subspaces of $(n+2)$-dimensional Lorentz space (cf. Lemma 1.3). Thus, G will be related to the Lorentz group.

Lorentz Space and Möbius Space

Lorentz space is $\mathbf{L} = \mathbf{R}^{n+2}$ (with basis $e_0, e_1, \ldots, e_{n+1}$) equipped with the indefinite quadratic form $q_L\colon \mathbf{L} \to \mathbf{R}$ given by

$$q_L(x) = -2x_0x_{n+1} + \sum_{1\le k\le n} x_k^2.$$

In the given basis, the matrix for this form is

$$\Sigma_{n+1,1} = \begin{pmatrix} 0 & 0 & -1 \\ 0 & I_n & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

[2] It was this fact that originally kindled the author's interest in Cartan's *espaces généralisés*.

Lemma 1.1. *The affine map $\psi\colon \mathbf{R}^{n+1} \to \mathbf{L}$ sending*

$$(y_0, y_1, \ldots, y_n) \mapsto \left(\frac{1}{\sqrt{2}}(1+y_0), y_1, \ldots, y_n, \frac{1}{\sqrt{2}}(1-y_0)\right)$$

is an affine isometric embedding of Euclidean space into Lorentz space.

Proof. The derivative of ψ at $x \in \mathbf{R}^{n+1}$ is

$$\psi_{*x}\colon T_x(\mathbf{R}^{n+1}) \to T_{\varphi(x)}(\mathbf{L}),$$
$$\nu = (\nu_0, \nu_1, \ldots, \nu_n) \mapsto \left(\frac{1}{\sqrt{2}}\nu_0, \nu_1, \ldots, \nu_n, \frac{-1}{\sqrt{2}}\nu_0\right).$$

In particular,

$$\begin{aligned} q_L(\psi_{*x}(v)) &= q_L\left(\frac{1}{\sqrt{2}}\nu_0, \nu_1, \ldots, \nu_n, -\frac{1}{\sqrt{2}}\nu_0\right) \\ &= 2\left(\frac{1}{\sqrt{2}}\nu_0\right)\left(\frac{1}{\sqrt{2}}\nu_0\right) + \sum_{1\le k\le n} \nu_k^2 = q_0(\nu), \end{aligned}$$

where q_0 is the standard metric on $\mathbf{R}^{n+2}$. ∎

Definition 1.2. A vector $v \in \mathbf{L}$ is called *lightlike* if $q_L(v) = 0$. A subspace of $\mathbf{L}$ is called *lightlike* if it consists of lightlike vectors. The set of lightlike points in $\mathbf{P}(\mathbf{L})$ (the projective space of $\mathbf{L}$; cf. Example 1.1.3) is denoted by M_0 and is called the *projective model of Möbius n space.* ❀

Lemma 1.3. *The map $\psi\colon \mathbf{R}^{n+1} \to \mathbf{L}$ induces a canonical diffeomorphism $\bar{\psi}\colon S^n \to M_0$.*

Proof. According to Example 1.1.3, the affine map $\psi\colon \mathbf{R}^{n+1} \to \mathbf{L}$ given in Lemma 1.1 defines an affine coordinate system $\bar{\psi}\colon \mathbf{R}^{n+1} \to P(\mathbf{L})$ for $P(\mathbf{L})$. This coordinate system includes all points of $\mathbf{P}(\mathbf{L})$ except for those corresponding to the one-dimensional subspaces lying on the hyperplane $x_0 + x_{n+1} = 0$, and these latter are never lightlike. Thus, the space of lightlike lines in $\mathbf{L}$ lies in the image of $\bar{\psi}$. In the coordinate system $\bar{\psi}$, the condition of being lightlike is

$$y_0^2 - 1 + \sum_{1\le k\le n} y_k^2 = 0 \quad\text{or}\quad \sum_{0\le k\le n} y_k^2 = 1.$$ ∎

The Canonical Line Bundle

The canonical (real) line bundle over $\mathbf{P}(\mathbf{L})$ (cf. Example 1.3.5) may be restricted to yield a line bundle E over the subspace of lightlike lines, $M_0 \subset \mathbf{P}(\mathbf{L})$, which we again call the canonical line bundle. More precisely, we have the following definition.

Definition 1.4. The *canonical line bundle* over M_0 is $E = \{(v, l) \in \mathbf{L}\times M_0 \mid v \in l\}$. ❀

Möbius Group

Consider the Lorentz group

$$L_{n+1,1}(\mathbf{R}) = \{g \in Gl_{n+2}(\mathbf{R}) \mid g^t \Sigma_{n+1,1} g = \Sigma_{n+1,1}\}.$$

It is clear that $L_{n+1,1}(\mathbf{R})$ is a Lie group that acts smoothly on the projective space $\mathbf{P}(\mathbf{L})$ of lines in $\mathbf{L}$ and that this in turn induces a smooth action on the subspace of lightlike lines $M_0 \subset \mathbf{P}(\mathbf{L})$. However, the induced action on M_0 is not effective since $\pm I$ acts trivially.

Exercise 1.5. Show that the action of $L_{n+1,1}(\mathbf{R})$ on $\mathbf{P}(\mathbf{L})$ is transitive on the Möbius space of lightlike lines $M_0 \subset \mathbf{P}(\mathbf{L})$. ❑

Lemma 1.6. *The action of $L_{n+1,1}(\mathbf{R})$ on M_0^n has kernel $\pm I$.*

Proof. As we mentioned above, the kernel K obviously contains $\pm I$ so we need only show the reverse inclusion. Now $L_{n+1,1}(\mathbf{R})$ acts on the canonical line bundle E (cf. Definition 1.4) according to the formula $g(v, l) = (gv, gl)$. If g fixes $l \in M_0$, then for each $v \in l$ we have $gv = \lambda_l v$ for some $\lambda_l \in \operatorname{Spec}(g)$, the set of eigenvalues of g. Thus, $g \in K$ implies that $gv = \lambda_l v$ (where l is the span of v) for every lightlike vector v. Since M_0 $(= S^n)$ is connected and λ_l depends continuously on $l \in S^n$ and $\operatorname{Spec}(g)$ is discrete, it follows that λ_l is independent of l, say $\lambda_l = \lambda$. Thus, $gv = \lambda v$ for every lightlike vector $v \in \mathbf{L}$. Since lightlike vectors given by

$$e_0, e_{n+1}, e_j + \frac{1}{\sqrt{2}}(e_0 + e_{n+1}) \quad\text{for}\quad 1 \le j \le n$$

span $\mathbf{L}$, it follows that $g = \lambda I$. Finally, we note that

$$\lambda^2 \Sigma_{n+1,1} = g^t \Sigma_{n+1,1} g = \Sigma_{n+1,1} \Rightarrow \lambda^2 = 1 \Rightarrow \lambda = \pm 1.$$ ∎

Definition 1.7. The group $\text{Möb}_n(\mathbf{R}) = L_{n+1,1}(\mathbf{R})/\{\pm I\}$ is called the *Möbius group* in n dimensions. The pair (G, H), where $H = \text{Möb}_n(\mathbf{R})_0 = \{h \in \text{Möb}_n(\mathbf{R}) \mid h[e_0] = [e_0]\}$, is called the *Möbius model* in dimension n. Here $[e_0]$ denotes the class of e_0 in $P(\mathbf{L})$. ❀

It is the group $H = \text{Möb}_n(\mathbf{R})_0$ that will be crucial for an understanding of Möbius geometries, so we need to have a way of representing its elements. These may be expressed as matrices, as the following result shows.

Lemma 1.8. *Let $\tilde{H}$ be*

$$\left\{ \begin{pmatrix} z & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & q & r \\ 0 & 1 & q^t \\ 0 & 0 & 1 \end{pmatrix} \middle| z \in \mathbf{R}^+, a \in O_n(\mathbf{R}), q^t \in \mathbf{R}^n, r = \frac{1}{2}qq^t \right\}.$$

(i) *$\tilde{H}$ is a subgroup of $L_{n+1,1}(\mathbf{R})$.*

(ii) *Every element of $\tilde{H}$ has a unique decomposition as a block diagonal times a block upper triangular matrix as above.*

(iii) *$\tilde{H}$ acts effectively on M_0.*

(iv) *The canonical projection $L_{n+1,1}(\mathbf{R}) \to$ Möb$_n(\mathbf{R})$ induces an isomorphism $\tilde{H} \to$ Möb$_n(\mathbf{R})_0$ so that $H =$ Möb$_n(\mathbf{R})_0$ may be regarded as the matrix group $\tilde{H}$.*

Proof. Parts (i) and (ii) are easy. Part (iii) is a consequence of the fact that $\tilde{H} \cap \{\pm I\} = \{I\}$. Let us deal with part (iv). It is clear that $L_{n+1,1}(\mathbf{R}) \to L_{n+1,1}(\mathbf{R})/\{\pm I\}$ induces a homomorphism $\tilde{H} \to$ Möb$_n(\mathbf{R})_0$ and that by (iii) this has trivial kernel. We must see that it is surjective. Any element of Möb$_n(\mathbf{R})_0$ may be covered by some element $g \in L_{n+1,1}(\mathbf{R})$, and we must have $g[e_0] = [e_0]$. Thus, g must have the form

$$g = \begin{pmatrix} \star & \star & \star \\ 0 & \star & \star \\ 0 & 0 & \star \end{pmatrix},$$

where the zero in the $(3,2)$ block arises as a consequence of the fact that $g^t \Sigma_{n+1,1} g = \Sigma_{n+1,1}$. Multiplying g by $-I$ if necessary (which doesn't alter its image in Möb$_n(\mathbf{R})_0$), we may assume g has the form

$$g = \begin{pmatrix} z & \star & \star \\ 0 & s & \star \\ 0 & 0 & z^{-1} \end{pmatrix},$$

where $z \in \mathbf{R}^+$. It is then an easy matter to use again the fact that g satisfies $g^t \sum_{n+1,1} g = \sum_{n+1,1}$ to deduce that $g \in \tilde{H}$. ■

Exercise 1.9. Prove the assertion of the last sentence in Lemma 1.8. ❑

Corollary 1.10. *Each of G and H has two components. H is generated by its identity component H_e and the element* diag$(1, -1, 1, \ldots, 1) \in H$.

Proof. The proof of Lemma 1.8 provides a diffeomorphism $H \to \mathbf{R}^+ \times O_n(\mathbf{R}) \times \mathbf{R}^n$, sending (in the notation of Lemma 1.8) $h \mapsto (z, a, q^t)$. Now $O_n(\mathbf{R})$ has two components while $\mathbf{R}^+$ and $\mathbf{R}^n$ are connected. Thus, H has two components. Since diag$(1, -1, 1, \ldots, 1)$ lies in the nonidentity component of H, it, together with H_e, generates H. On the other hand, for $n > 1$, $G/H = S^n$ is connected and simply connected. Thus, the long exact sequence of homotopy groups $1 = \pi_1(G/H) \to \pi_0(H) \to \pi_0(G) \to \pi_0(G/H) = 1$ shows that H and G have the same number of components (cf. [D. Husemoller, 1966], p. 92). We leave the case $n = 1$ for the reader. ■

By Lemma 1.12 and Corollary 3.7 ahead, we see that the group G may also be described as the group of all conformal transformations of the standard n-sphere (cf. also [M. Berger, 1987], Theorem 18.10.4). The precise definition of conformal transformations is as follows.

Definition 1.11. A diffeomorphism $\phi: S^n \to S^n$ is called a *conformal transformation* of the standard n-sphere if, for each $x \in S^n$, $\phi_{*x}: T_x(S^n) \to T_{\phi(x)}(S^n)$ preserves the canonical Riemannian metric q_0 up to scale. ❀

Lemma 1.12. *Möb$_n(\mathbf{R})$ acts on S^n as a group of conformal transformations.*

Proof. The action of $L_{n+1,1}(\mathbf{R})$ on S^n as defined above may also be described by the formula $g\psi(x) = \mu(x)\psi(gx)$, where $x \in S^n$ and $\mu(x) \in \mathbf{R}^*$ is chosen so that $\mu(x)\psi(gx) \in$ Im ψ. Differentiating, we see that for any $v \in T_x(S^n)$, we have

$$g\psi_{*x}(v) = \mu_{*x}(v)\psi(gx) + \mu(x)\psi_{*gx}(g_{*x}(v)).$$

Since $g_{*x}(v) \in T_{gx}(S^n)$, it follows that gx and $g_{*x}(v)$ are perpendicular in $\mathbf{R}^{n+1}$. Hence by Lemma 1.1, it follows that $\psi(gx)$ and $\psi_{*gx}(g_{*x}(v))$ are perpendicular in $\mathbf{L}$. Since we also know that $\psi(gx)$ is lightlike, it follows that

$$\begin{aligned}
q_0(v) &= q_L(\psi_{*x}(v)) \quad \text{(by Lemma 1.1)} \\
&= q_L(g\psi_{*x}(v)) \quad \text{(by the definition of } L_{n+1,1}(\mathbf{R})\text{)} \\
&= q_L(\mu_{*x}(v)\psi(gx) + \mu(x)\psi_{*gx}(g_{*x}(v))) \quad \text{(by the formula above)} \\
&= q_L(\mu_{*x}(v)\psi(gx)) + 2(\mu_{*x}(v)\psi(gx))^t\Sigma_{n+1,1}(\mu(x)\psi_{*gx}(g_{*x}(v))) \\
&\quad + q_L(\mu(x)\psi_{*gx}(g_{*x}(v))) \\
&= q_L(\mu(x)\psi_{*gx}(g_{*x}(v))) \quad \text{(by the remarks above)} \\
&= \mu(x)^2 q_L(\psi_{*gx}(g_{*x}(v))) \\
&= \mu(x)^2 q_0(g_{*x}(v)) \quad \text{(by Lemma 1.1).}
\end{aligned}$$

Thus, the action of $L_{n+1,1}(\mathbf{R})$, and hence of Möb$_n(\mathbf{R})$, preserves q_0 up to scale. ■

Lie Algebras for the Möbius Model

Exercise 1.13.* (i) The pair (G, H) has Lie algebras $(\mathfrak{g}, \mathfrak{h})$, where

$$\mathfrak{g} = \left\{ \begin{pmatrix} z & q & 0 \\ p & s & q^t \\ 0 & p^t & -z \end{pmatrix} \right\}, \quad \mathfrak{h} = \left\{ \begin{pmatrix} z & q & 0 \\ 0 & s & q^t \\ 0 & 0 & -z \end{pmatrix} \right\}, \quad \text{where } s^t = -s,$$

and $\mathfrak{g}$ has blocks down the main diagonal of size 1×1, $n \times n$, and 1×1.

(ii) $\mathfrak{g}$ decomposes (as a vector space) as $\mathfrak{g} = \mathfrak{p} \oplus \mathfrak{s} \oplus \mathfrak{z} \oplus \mathfrak{q}$, so that $\mathfrak{h} = \mathfrak{s} \oplus \mathfrak{z} \oplus \mathfrak{q}$, where

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ p & 0 & 0 \\ 0 & p^t & 0 \end{pmatrix} \right\}, \quad \mathfrak{s} = \left\{ \begin{pmatrix} 0 & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\},$$

$$\mathfrak{z} = \left\{ \begin{pmatrix} z & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -z \end{pmatrix} \right\}, \quad \mathfrak{q} = \left\{ \begin{pmatrix} 0 & q & 0 \\ 0 & 0 & q^t \\ 0 & 0 & 0 \end{pmatrix} \right\}.$$

(iii) Show that $\mathfrak{q}, \mathfrak{z} \oplus \mathfrak{q}$ and $\mathfrak{s} \oplus \mathfrak{q}$ are ideals of $\mathfrak{h}$.

(iv) Show that for $n \neq 4$ the ideals in (iii) are the only ideals of $\mathfrak{h}$.

(v) Show that for $n = 4$ we have $\mathfrak{s} = \mathfrak{so}(4) = \mathfrak{so}(3) \oplus \mathfrak{so}(3) = W_+ \oplus W_-$ as Lie algebras. Show also that the complete list of ideals of $\mathfrak{h}$ in this case is

$$\mathfrak{q}, \mathfrak{z} \oplus \mathfrak{q}, \mathfrak{s} \oplus \mathfrak{q}, W_+ \oplus \mathfrak{q}, W_- \oplus \mathfrak{q}, \mathfrak{z} \oplus W_+ \oplus \mathfrak{q}, \quad \text{and } \mathfrak{z} \oplus W_- \oplus \mathfrak{q}. \qquad \square$$

Exercise 1.14. Setting $\mathfrak{g}_{-1} = \mathfrak{p}$, $\mathfrak{g}_0 = \mathfrak{s} \oplus \mathfrak{z}$, and $\mathfrak{g}_1 = \mathfrak{q}$, show that $[\mathfrak{g}_a, \mathfrak{g}_b] \subset \mathfrak{g}_{a+b}$ for all a and b (where $\mathfrak{g}_a = 0$ if $a \neq -1, 0, 1$). $\square$

Adjoint Action

It is a simple computation to show that the adjoint action of $h \in H$ on $\mathfrak{g}/\mathfrak{h}$ is given by $\mathrm{Ad}(h)v = sz^{-1}v$, where

$$h = \begin{pmatrix} z & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & q & r \\ 0 & 1 & q^t \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and } r = \frac{1}{2}qq^t.$$

Quadratic Form Invariant up to Scale on $\mathfrak{g}/\mathfrak{h}$

Lemma 1.15. *There is a quadratic form $q_{\mathfrak{g}/\mathfrak{h}}\colon \mathfrak{g}/\mathfrak{h} \to \mathbf{R}$ which is $\mathrm{Ad}(H)$ invariant up to scale. Such a quadratic form is unique up to scale.*

$$\begin{tikzcd} L\text{-}\{0\} \arrow[r] & E \arrow[d] & G \arrow[l] \arrow[d] \arrow[ll, bend right, "\alpha"'] \\ S^n \arrow[u, "\psi"] \arrow[r, "\bar\psi"', "\text{diffeo.}"] & M_0 & G/H \arrow[l, "\text{diffeo.}"] \end{tikzcd} \quad \text{with maps} \quad \begin{tikzcd} ge_0,\ w \arrow[r, mapsto] & (w, [w]),\ (ge_0, [ge_0]) & g \arrow[l, mapsto] \arrow[d, mapsto] \arrow[ll, bend right] \\ \psi(x) & (v, l) \arrow[d, mapsto] & \\ x \arrow[u, mapsto] \arrow[r, mapsto] & {[\psi(x)]},\ [ge_0] & gH \arrow[l, mapsto] \end{tikzcd}$$

FIGURE 1.16.

Proof. Define

$$q_{\mathfrak{g}/\mathfrak{h}} \begin{pmatrix} \star & \star & 0 \\ p & \star & \star \\ 0 & p^t & \star \end{pmatrix} = q_0(p) = \sum_{1 \le k \le n} p_k^2.$$

The formula above for the adjoint action implies that $q_{\mathfrak{g}/\mathfrak{h}}(\mathrm{Ad}(h)v) = z^{-2} q_{\mathfrak{g}/\mathfrak{h}}(v)$. As for the uniqueness, it is well known (and, in any case, easily shown) that q_0 is unique up to scale among the quadratic forms invariant under the subgroup $O_n(\mathbf{R}) \subset H$. Hence it is also unique up to scale under the larger group. ■

Identifying Möbius Space with G/H

The Möbius model may be described in three ways: as the homogeneous space G/H; as the space M_0 of lightlike lines in $\mathbf{P}(\mathbf{L})$; and as the n-sphere S^n. These descriptions are related in Figure 1.16, where E is the canonical line bundle over M_0. Moreover, there is a left action of G on every space in this diagram. (The action on S^n is of course defined via the map $\bar\psi$.) It is clear that all the maps are G-equivariant except for ψ. As usual, we let $N = e_0 \in \mathbf{R}^{n+1}$ be the "north pole" of S^n. Since $\bar\psi(N) = [e_0]$, it follows that the composite map $G/H \to S^n$ sends $gH \to gN$.

Lie-Theoretic Interpretation of the Quadratic Form q_0

Lemma 1.17. *Let $w \in T_e(G/H)$, and write it as $w = \pi_{e*}(v)$, where*

$$v = \begin{pmatrix} \star & \star & 0 \\ p & \star & \star \\ 0 & p^t & \star \end{pmatrix} \in \mathfrak{g} = T_eG.$$

Then under the identification $G/H \to S^n$ given by Figure 1.16, v corresponds to the tangent vector $u = (0, p_1, \ldots, p_n) \in T_N(S^n) \subset T_N(\mathbf{R}^{n+1})$.

Proof. The map $\alpha\colon G \to \mathbf{L}$ sending $g \mapsto ge_0$ has derivative $\alpha_{*e}\colon \mathfrak{g} = T_eG \to \mathbf{L}$ sending $v \mapsto ve_0$. The derivative of the commutative diagram in Figure 1.16 yields the following commutative diagram.

$$\begin{tikzcd}
T_{\psi(N)}(L) \arrow[r] & T_{(e_0,[e_0])}(E) \arrow[d] & T_e(G) \arrow[l] \arrow[ll, bend right, "\alpha_{*e}"'] \arrow[d, "\pi_{*e}"'] & \ni v \arrow[d] \\
T_N(S^n) \arrow[u, "\psi_{*N}"] \arrow[r, "\overline{\psi}_{*N}"] & T_{[e_0]}(M_0) & T_{[e]}(G/H) \arrow[l] & \ni w
\end{tikzcd}$$

For the proof of the lemma, it suffices to show that the elements $v \in T_e(G) = \mathfrak{g}$ and $\psi_{*N}(u) \in T_{e_0}(\mathbf{L}) = \mathbf{L}$ correspond under α_{*e}. But by the definition of ψ (Lemma 1.1) and the definition of α,

$$\psi_{*N}(u) = (0, p_1, \ldots, p_n, 0) = \begin{pmatrix} 0 \\ p \\ 0 \end{pmatrix},$$

$$\alpha_{*e}(v) = \begin{pmatrix} 0 & 0 & 0 \\ p & 0 & 0 \\ 0 & p^t & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ p \\ 0 \end{pmatrix}. \qquad \blacksquare$$

The diffeomorphism $G/H \to S^n$ sending $gH \mapsto gN$ determines a quadratic form up to scale on T_xS^n for each $x \in S^n$ as follows. Fix $g \in G$ such that $x = gN$, and define the quadratic form as the composite

$$T_x(S^n) \overset{\approx}{\leftarrow} T_{[g]}(G/H) \xrightarrow{\varphi_g} \mathfrak{g}/\mathfrak{h} \xrightarrow{q_{\mathfrak{g}/\mathfrak{h}}} \mathbf{R}.$$

Of course the choice of g is not unique. It may be varied by right multiplication by any element

$$h = \begin{pmatrix} z & \star & \star \\ 0 & \star & \star \\ 0 & 0 & z^{-1} \end{pmatrix} \in H,$$

which will replace $q_{\mathfrak{g}/\mathfrak{h}}\varphi_g$ by $q_{\mathfrak{g}/\mathfrak{h}}\varphi_{gh} = q_{\mathfrak{g}/\mathfrak{h}}(\mathrm{Ad}(h^{-1})\varphi_g) = z^2 q_{\mathfrak{g}/\mathfrak{h}}(\varphi_g)$.

Moreover, the following diagram is commutative, where the vertical arrows are induced by the left action of $\gamma \in G$.

$$\begin{tikzcd}
T_x(S^n) \arrow[d] & T_{[g]}(G/H) \arrow[l, "\approx"'] \arrow[d] \arrow[dr, "\varphi_g"] & \\
& & \mathfrak{g}/\mathfrak{h} \\
T_{\gamma x}(S^n) & T_{[\gamma g]}(G/H) \arrow[l, "\approx"] \arrow[ur, "\varphi_{\gamma g}"'] &
\end{tikzcd}$$

From this it follows that the quadratic form we have constructed on $T(S^n)$ is invariant up to scale under the action of G.

Lemma 1.18. *Up to scale, the quadratic form on $T(S^n)$ constructed above is the standard one.*

Proof. By Lemma 1.12 and the remarks above, both q_0 and the form constructed above are invariant up to scale under the action of G on S^n. Since this action is transitive, it suffices to prove the two forms agree at N. But from Lemma 1.17, we see that

$$u = (0, p) = (0, p_1, \ldots, p_n) \in T_N(S^n) \subset T_N(\mathbf{R}^{n+1}) \mapsto \begin{pmatrix} \star & \star & 0 \\ p & \star & \star \\ 0 & p^t & \star \end{pmatrix}$$
$$\in \mathfrak{g}/\mathfrak{h} \mapsto q_{\mathfrak{g}/\mathfrak{h}}(p) = \sum_{1 \le i \le n} p_i^2 = q_0(v). \qquad \blacksquare$$

Ricci Homomorphism

For the next definitions, and for later purposes, we once again need the *Ricci homomorphism*, denoted by *Ricci*, which in the Möbius context is given as the composite of the following maps.

$$\begin{aligned}
\mathit{Ricci}\colon \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{s}) \approx \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathfrak{s} &\overset{\mathrm{id}\otimes\mathrm{ad}}{\longrightarrow} \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathrm{End}(\mathfrak{g}/\mathfrak{h}) \\
\approx \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathfrak{g}/\mathfrak{h} \otimes (\mathfrak{g}/\mathfrak{h})^* &\overset{\mathrm{contraction}\otimes\mathrm{ad}}{\longrightarrow} (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^* \\
t^* \wedge u^* \otimes v \otimes w^* &\mapsto (u^*(v)t^* - t^*(v)u^*) \otimes w^*
\end{aligned} \tag{1.19}$$

Exercise 1.20. Show that $\mathfrak{q}$ and $\mathfrak{q} \oplus \mathfrak{s}$ are H submodules of $\mathfrak{h}$ under the adjoint action. It follows that we may regard $\mathfrak{s}$ as an H module via the canonical isomorphism $\mathfrak{s} \approx (\mathfrak{q} \oplus \mathfrak{s})/\mathfrak{q}$. Show that *Ricci* is an H module map. ❑

Definition 1.21. In Möbius geometry the *normal* H submodule of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is the kernel of homomorphism

$$\mathit{Ricci} : \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{s}) \to (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*.$$

The following exercise is a partial analog of Proposition 6.1.4. Such an analog is possible since the adjoint action of H on $\mathfrak{g}/\mathfrak{h}$ is the same as in the Riemannian case except for an additional scaling factor (see the subsection preceding Lemma 1.16).

Exercise 1.22.* Let $n \geq 2$, and fix an orthogonal equinormal basis $e_i \in \mathfrak{g}/\mathfrak{h}$, $1 \leq i \leq n$ (i.e., orthogonal and of equal length with respect to the quadratic form of Lemma 1.15). Let $e_{pq} \subset \mathfrak{s}$, $1 \leq p < q \leq n$ be the unique (and standard) basis such that $\mathrm{ad}(e_{pq})$ corresponds to $e_p \otimes e_q^* - e_q \otimes e_p^*$ under the canonical isomorphism $\mathrm{End}(\mathfrak{g}/\mathfrak{h}) \approx (\mathfrak{g}/\mathfrak{h}) \otimes (\mathfrak{g}/\mathfrak{h})^*$.

(i) Show that *Ricci* is an H module homomorphism satisfying

$$Ricci(e_i^* \wedge e_j^* \otimes e_{pq}) = \delta_{jp} e_i^* \otimes e_q^* - \delta_{jp} e_i^* \otimes e_p^* - \delta_{ip} e_j^* \otimes e_q^* + \delta_{iq} e_j^* \otimes e_p^*$$

so that, in particular when $i \neq k \neq j$,

$$Ricci(e_i^* \wedge e_k^* \otimes e_{kj}) = e_i^* \otimes e_j^* + \delta_{ij} e_k^* \otimes e_k^*.$$

(ii) Show that

(a) for $n = 2$, *Ricci* is injective so that the normal submodule is trivial,

and

(b) for $n > 2$, *Ricci* is surjective.

(iii) Let $n > 2$ and set $\varphi = \sum_{ijpq} a_{ijpq} e_i^* \wedge e_j^* \otimes e_{pq} \in \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ where a_{ijpq} is skew symmetric in the first two indices and the last two indices. Assume in addition that $a_{kijk} = a_{kjik}$ for all indices. This latter condition would be a consequence of the Bianchi identities were they to hold (cf. the proof of **I** in Proposition 3.1 ahead). Show

$$Ricci(\varphi) = \sum_{ijk} a_{kijk} e_i^* \otimes e_j^*.$$

(iv) Let $n > 2$ and set $b_{ij} = \sum_{1 \leq k \leq n} (e_i^* \wedge e_k^* \otimes e_{kj} + e_j^* \wedge e_k^* \otimes e_{ki})$, $1 \leq i, j \leq n$. Show that

$$Ricci(b_{ij}) = (n-2)(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*) + 2\delta_{ij} \sum_k e_k^* \otimes e_k^*.$$

Show the elements b_{ij} are linearly independent, and span a submodule $R_n \subset \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ (the analogue in Möbius geometry of the Ricci submodule, cf. Definition **6**.1.5(ii)). Show that R_n has trivial intersection with $W_n = \ker Ricci$ (the analogue of the Weyl submodule). ❑

The Weyl Model

Definition 1.23. The Weyl model is the Euclidean geometry of similarity. Its groups (G_N, H_N) are given by

$$G_N = \left\{ \begin{pmatrix} z & 0 \\ p & s \end{pmatrix} \middle| \, z \in \mathbf{R}^+, p \in \mathbf{R}^n, s \in O_n(\mathbf{R}) \right\},$$

$$H_N = \left\{ \begin{pmatrix} z & 0 \\ 0 & s \end{pmatrix} \middle| \, z \in \mathbf{R}^+, s \in O_n(\mathbf{R}) \right\}$$ ❀

Exercise 1.24. Show that the Lie algebras of the Weyl model are

$$\mathfrak{g}_N = \left\{ \begin{pmatrix} z & 0 \\ p & s \end{pmatrix} \middle| \, z \in \mathbf{R}, p \in \mathbf{R}^n, s \in \mathfrak{o}_n(\mathbf{R}) \right\},$$

$$\mathfrak{h}_N = = \left\{ \begin{pmatrix} z & 0 \\ 0 & s \end{pmatrix} \middle| \, z \in \mathbf{R}, s \in \mathfrak{o}_n(\mathbf{R}) \right\}.$$ ❑

The Euclidean group of similarities G_N acts on $\mathbf{R}^n$ according to the formula

$$\begin{pmatrix} z & 0 \\ p & s \end{pmatrix} \cdot x = z^{-1}(sx + p).$$

If we identify $\mathbf{R}^n$ with $S^n - N$ via stereographic projection (where N is the north pole of S^n), then G_N and H_N are identified with the subgroups fixing N of the corresponding Möbius groups G and H. Thus, the Weyl model may be regarded as a subgeometry of the Möbius model. In particular, the inclusion $\mathfrak{g}_N \subset \mathfrak{g}$ is given by

$$\begin{pmatrix} z & 0 \\ p & s \end{pmatrix} \mapsto \begin{pmatrix} z & 0 & 0 \\ p & s & 0 \\ 0 & p^t & -z \end{pmatrix}.$$

§2. Möbius and Weyl Geometries

In this section we study some elementary aspects of Möbius geometry. We show that a Möbius geometry on M determines a conformal metric on M. We also investigate the various possibilities of special curvatures including normal geometries and Weyl geometries.

Definition 2.1. Let M be a smooth manifold. A *Möbius geometry on* M is a Cartan geometry on M modeled on the Möbius model. ❀

Let us begin by studying the conformal metric canonically associated to a Möbius geometry.

The Conformal Metric Determined by a Möbius Geometry

Definition 2.2. Let M be a smooth n-manifold. Two metrics q_1, q_2: $T(M) \to \mathbf{R}$ are called *conformally equivalent* if there is a positive smooth function $\lambda: M \to \mathbf{R}$ such that $q_2 = \lambda q_1$. A conformal equivalence class of metrics $\mathcal{G}$ on M is called a *conformal metric.* If $x \in M$, then $\mathcal{G}_x$ denotes the restrictions of the elements of $\mathcal{G}$ to $T_x(M)$. ❀

Lemma 2.3. *Let M be a smooth manifold. A Möbius geometry (P, ω) on M determines a conformal metric on M.*

Proof. As usual, we have the isomorphism $\varphi_p: T_x(M) \to \mathfrak{g}/\mathfrak{h}$ for each $p \in P$ situated above $x \in M$. Using a fixed choice, say $q_{\mathfrak{g}/\mathfrak{h}}$, of one of the quadratic forms given in Lemma 1.15, we can construct, on each open set $U \subset M$ over which there is a section $\sigma: U \to P$, a metric defined by

$$q_x(v) = q_{\mathfrak{g}/\mathfrak{h}}(\varphi_{\sigma(x)}(v)) \quad \text{for each } v \in T_x(M).$$

Now two such sections differ by a change of gauge of the form

$$h = \begin{pmatrix} z & \star & \star \\ 0 & \star & \star \\ 0 & 0 & z^{-1} \end{pmatrix} : U \to H,$$

and as in the remarks following Lemma 1.17, the corresponding metrics differ by a factor of z^2, where $z = z(h)$ is as in the expression for h above. From this it is also clear that, by an appropriate change of gauge in this way, we can obtain *any* metric on U conformally equivalent to q_x. Thus, we have a well-defined conformal metric on U.

The only remaining point is to show that there is a globally defined metric on M restricting to the local conformal metrics we have described above. For this we use the notion of a partition of unity (cf. [W.M. Boothby, 1986], pp. 193–198). Choose sections $\sigma_\alpha: U_\alpha \to P$, where the U_α cover M, and let $q_\alpha: T(U_\alpha) \to \mathbf{R}$ be the corresponding metrics. Since M is paracompact, we may find a locally finite refinement of the cover $\{U_\alpha\}$ and restrict the σ_αs to it. Thus, without loss of generality, we may assume that $\{U_\alpha\}$ itself is locally finite. Then we may choose a partition of unity p_α subordinate to $\{U_\alpha\}$ and set $q = \sum_\alpha p_\alpha q_\alpha$. Then q is a smooth metric on M, and on each U_α it is a finite sum of conformally equivalent local metrics of the type described above and is therefore conformally equivalent to any one of them. ■

Example 2.4. Consider the model geometry G/H, which by Lemma 1.3 we have identified with the unit sphere $S^n \subset \mathbf{R}^{n+1}$. According to Lemma 1.18, the conformal metric constructed by the procedure of Lemma 2.3 is the class of the standard metric on S^n. ◆

Special Curvatures

In Möbius geometry, the Cartan connection and its curvature take values in $\mathfrak{g}$, and so they may in general be written in block form as

$$\omega = \begin{array}{c} \begin{matrix} 1 & n & 1 \end{matrix} \\ \begin{pmatrix} \varepsilon & v & 0 \\ \theta & \alpha & v^t \\ 0 & \theta^t & -\varepsilon \end{pmatrix} \end{array}$$

(with blocks of size $1 \times 1, n \times n$ and 1×1 down the main diagonal) and

$$\begin{aligned} \Omega &= \begin{pmatrix} E & Y & 0 \\ \Theta & A & Y^t \\ 0 & \Theta^t & -E \end{pmatrix} \\ &= \begin{pmatrix} d\varepsilon + v \wedge \theta & dv + \varepsilon \wedge v + v \wedge \alpha & 0 \\ d\theta + \theta \wedge \varepsilon + \alpha \wedge \theta & d\alpha + \theta \wedge v + \alpha \wedge \alpha + v^t \wedge \theta^t & \star \\ 0 & \star & \star \end{pmatrix}. \end{aligned} \tag{2.5}$$

We may say that the torsion free Möbius geometries that are "farthest" from the Riemannian geometries are those for which the curvature has no Riemannian part, that is, no rotational part. These are the geometries with $\Theta = 0$ and $A = 0$. However, the following result shows that nontrivial examples of such geometries can exist only in dimensions ≤ 3.

Proposition 2.6. *Suppose the curvature takes values in the ideal $\mathfrak{z} \oplus \mathfrak{q}$ (cf. Exercise 1.13) so that it has the form*

$$\Omega = \begin{pmatrix} E & Y & 0 \\ 0 & 0 & Y^t \\ 0 & 0 & -E \end{pmatrix}.$$

Then

(i) $\dim M \geq 3 \Rightarrow E = 0$,

(ii) $\dim M \geq 4 \Rightarrow Y = 0$.

Proof. Since E is semibasic, we may write it as

$$E = \sum_{i<j} f_{ij} \theta_i \wedge \theta_j.$$

Recall the Bianchi identity $d\Omega = [\Omega, \omega]$ (Lemma **5**.3.30). The $(2,1)$ block of this identity is

$$0 = \theta \wedge E$$

or

$$0 = \theta_k \wedge E \quad \text{for } k = 1, \dots, n = \dim M.$$

Thus, for all k,

$$0 = \sum_{i<j} f_{ij} \theta_k \wedge \theta_i \wedge \theta_j.$$

For k fixed and $k \neq i < j \neq k$, the forms $\theta_k \wedge \theta_i \wedge \theta_j$ are independent. Thus, $f_{ij} = 0$ whenever there is a $k \neq i, j$. But for $n \geq 3$, such a k exists for every pair (i, j). Thus $E = 0$.

Now consider the $(2,2)$ block of the Bianchi identity. It reads

$$0 = Y^t \wedge \theta^t - \theta \wedge Y$$

or

$$0 = Y_i \wedge \theta_j - \theta_i \wedge Y_j \quad \text{for } i, j = 1, \ldots, n.$$

Thus, for all i, j, we have $Y_i \wedge \theta_j \wedge \theta_i = \theta_i \wedge Y_j \wedge \theta_i = 0$. This means that the 3-form $Y_i \wedge \theta_i$ has a factor θ_j for all j. For $n \geq 4$, this means that $Y_i \wedge \theta_i = 0$ for all i. Thus, we may write $Y_i = \phi_i \wedge \theta_i$ for some 1-forms ϕ_i. Putting this in $0 = Y_i \wedge \theta_j - \theta_i \wedge Y_j$ yields

$$\phi_i \wedge \theta_i \wedge \theta_j - \theta_i \wedge \phi_j \wedge \theta_j = 0$$

or

$$(\phi_i + \phi_j) \wedge \theta_i \wedge \theta_j = 0.$$

From this we see that for any three distinct indices i, j, k, we have

$\phi_i + \phi_j$ is a linear combination of θ_i and θ_j,
$\phi_j + \phi_k$ is a linear combination of θ_j and θ_k,
$\phi_i + \phi_k$ is a linear combination of θ_i and θ_k.

We may solve these equations for ϕ_i. The solution shows that, for any three distinct indices i, j, k, the form ϕ_i is a linear combination of θ_i, θ_j, and θ_k. For $n \geq 4$ and l any index different from i, this process will yield a linear expression for ϕ_i that is free of θ_l. This shows that ϕ_i can only be a multiple of θ_i. But then $Y_i = \phi_i \wedge \theta_i = 0$. ∎

Normal Möbius Geometries

Somewhat more subtle in their definition than the subclasses of geometries considered above are the *normal geometries*.

Definition 2.7. A Möbius geometry is called *normal* if

(i) it is of type $\mathfrak{s} \oplus \mathfrak{q}$ (i.e., $\Theta = 0$ and $E = 0$), and

(ii) the $\mathfrak{s}$ component $K_{\mathfrak{s}}\colon P \to \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{s})$ of curvature function takes values in the normal submodule described in Definition 1.21. ❀

One of the consequences of the proof of Proposition 3.1 ahead is that, in the presence of condition (i), condition (ii) may be regarded as saying that the block v of the form ω (see Eq. (2.5)) is determined in a certain way by the other blocks of ω.

We shall need the explicit formula for $K_{\mathfrak{s}}$ in terms of the block A in the curvature form Ω of Eq. (2.5). Let $e_i \in \mathfrak{g}/\mathfrak{h}$ and $e_{pq} \in \mathfrak{s}$ be as in Exercise 1.22. Since A is semibasic, we may express it as

$$A = \sum_{\substack{i<j \\ p<q}} A_{ijpq} \theta_i \wedge \theta_j \otimes e_{pq}.$$

Lemma 2.8. $K_{\mathfrak{s}} = \sum_{r<s, p<q} A_{rspq} e_r^* \wedge e_s^* \otimes e_{pq}$.

Proof. Since $K(u,v) = \Omega(\omega^{-1}(u), \omega^{-1}(v))$, for $r < s$ we have

$$\begin{aligned} K_{\mathfrak{s}}(e_r, e_s) &= A(\omega^{-1}(e_r), \omega^{-1}(e_s)) \\ &= \sum_{\substack{i<j \\ p<q}} A_{ijpq} (\theta_i \wedge \theta_j)(\omega^{-1}(e_r), \omega^{-1}(e_s)) \otimes e_{pq} \\ &= \sum_{\substack{i<j \\ p<q}} A_{ijpq} (\delta_{ir}\delta_{js} - \delta_{jr}\delta_{is}) e_{pq} = \sum_{p<q} A_{rspq} e_{pq}. \end{aligned}$$

Thus,

$$K_{\mathfrak{s}} = \sum_{r<s} K_{\mathfrak{s}}(e_r, e_s) e_r^* \wedge e_s^* = \sum_{\substack{r<s \\ p<q}} A_{rspq} e_r^* \wedge e_s^* \otimes e_{pq}.$$ ∎

Three-Dimensional Normal Möbius Geometry

Proposition 2.9. *Let (P, ω) be a Möbius geometry on the three-dimensional manifold M. Then (P, ω) is normal $\Leftrightarrow$ the curvature of (P, ω) takes its values in $\mathfrak{q}$.*

Proof. $\Rightarrow$ By Definition 2.7(i), the connection and curvature take the form

$$\omega = \begin{pmatrix} \varepsilon & v & 0 \\ \theta & \alpha & v^t \\ 0 & \theta^t & -\varepsilon \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} 0 & Y & 0 \\ 0 & A & Y^t \\ 0 & 0 & 0 \end{pmatrix}.$$

To show that the curvature takes its values in $\mathfrak{q}$ is to show that $A = 0$, which we do now.

Since the space of the geometry has dimension 3, it follows that $\theta_2 \wedge \theta_3$, $\theta_3 \wedge \theta_1$, and $\theta_1 \wedge \theta_2$ constitute a basis for the semibasic 2-forms on P. Since curvature is semibasic, we may express A in the form $A = (a_1 * *)\theta_2 \wedge \theta_3 + (* a_2 *)\theta_3 \wedge \theta_1 + (* * a_3)\theta_1 \wedge \theta_2$, where a_1 is the first column of $(a_1 * *)$, and so forth. The Bianchi identity in the $(2,1)$ block has the form $A \wedge \theta = 0$. Mutiplying this out yields

$$(a_1\theta_1 \wedge \theta_2 \wedge \theta_3, 0, 0) + (0, a_2\theta_1 \wedge \theta_2 \wedge \theta_3, 0) + (0, 0, a_3\theta_1 \wedge \theta_2 \wedge \theta_3) = 0,$$

and hence $a_1 = a_2 = a_3 = 0$. After a brief reflection, it follows that A must have the form

$$A = \begin{pmatrix} 0 & b_3\theta_1 \wedge \theta_2 & b_2\theta_1 \wedge \theta_3 \\ -b_3\theta_1 \wedge \theta_2 & 0 & b_1\theta_2 \wedge \theta_3 \\ -b_2\theta_1 \wedge \theta_3 & -b_1\theta_2 \wedge \theta_3 & 0 \end{pmatrix}.$$

Thus, $K_\mathfrak{s} = b_1 e_2^* \wedge e_e^* \otimes e_{23} + b_2 e_3^* \wedge e_1^* \otimes e_{31} + b_3 e_1^* \wedge e_2^* \otimes e_{12}$. Applying the Ricci homomorphism (cf. Exercise 1.22(i)), we get

$$\begin{aligned} -Ricci(K_\mathfrak{s}) &= b_1(e_2^* \otimes e_2^* + e_3^* \otimes e_3^*) + b_2(e_2^* \otimes e_2^* + e_3^* \otimes e_3^*) \\ &\quad + b_3(e_1^* \otimes e_1^* + e_2^* \otimes e_2^*) \\ &= (b_2 + b_3) e_1^* \otimes e_1^* + (b_3 + b_1) e_2^* \otimes e_2^* + (b_1 + b_2) e_3^* \otimes e_3^*. \end{aligned}$$

Since the geometry is normal, $Ricci(K_\mathfrak{s}) = 0$ and hence $b_1 = b_2 = b_3 = 0$. Thus $A = 0$.

$\Leftrightarrow$ Since the curvature of (P, ω) takes its values in $\mathfrak{q}$, condition 2.7(i) holds and $A = 0$. Thus, we need verify only 2.7(ii). But by Lemma 2.8, $A = 0 \Rightarrow K_\mathfrak{s} = 0 \Rightarrow Ricci(K_\mathfrak{s}) = 0$. ∎

Weyl Geometry

In 1918, H. Weyl introduced a generalization of Riemannian geometry in an attempt to formulate a unified field theory (cf. [H. Weyl, 1952]). The fundamental idea was that, since there is no standard of length in the universe, the final length of a ruler moved along a path might well depend on the path. From this ingredient he constructed a theory unifying gravitation and electromagnetism. Although Einstein rejected Weyl's unreliable ruler as physically wrong, and the theory fell into disrepute, its mathematical structure has nevertheless resurfaced as a fundamental part of the "standard theory" of particle physics, where the unreliable length has been reinterpreted as an unreliable phase [S.-S. Chern, 1977].

We shall introduce Weyl's geometry, not as Weyl originally did,[3] but as a particular case of a Cartan geometry.

Definition 2.10. Let M be a smooth manifold. A *Weyl geometry* on M is a Cartan geometry modeled on the Weyl model. ❀

In a Weyl geometry, the Cartan connection and its curvature take values in $\mathfrak{g}_N$, and so they may in general be written in block form as

$$\omega = \begin{pmatrix} \varepsilon & 0 \\ \theta & \alpha \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} E & 0 \\ \Theta & A \end{pmatrix}.$$

Theorem 2.11. *A torsion free Möbius geometry (P, ω) on the surface M with connection and curvature given by*

$$\omega = \begin{pmatrix} \varepsilon & v_1 & v_2 & 0 \\ \theta_1 & 0 & -\alpha & v_1 \\ \theta_2 & \alpha & 0 & v_2 \\ 0 & \theta_1 & \theta_2 & -\varepsilon \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} E & Y_1 & Y_2 & 0 \\ 0 & 0 & -A_{21} & Y_1 \\ 0 & A_{21} & 0 & Y_2 \\ 0 & 0 & 0 & E \end{pmatrix},$$

[3] Cf. Definition 3.11 for a definition equivalent to Weyl's.

for which E and A never vanish simultaneously, canonically determines a Weyl geometry $(P_{\mathrm{Weyl}}, \omega_{\mathrm{Weyl}})$ on M of curvature

$$\Omega_{\mathrm{Weyl}} = \begin{pmatrix} d\varepsilon & 0 & 0 \\ 0 & 0 & -d\alpha \\ 0 & d\alpha & 0 \end{pmatrix}.$$

Proof. Every entry in the curvature form is a semibase 2-form. Since $\theta_1 \wedge \theta_2$ is a basis for the semibasic 2-forms, it follows that each entry is a multiple of it, so we may write

$$\Omega = \begin{pmatrix} e & u_1 & u_2 & 0 \\ 0 & 0 & -a & u_1 \\ 0 & a & 0 & u_2 \\ 0 & 0 & 0 & -e \end{pmatrix} \theta_1 \wedge \theta_2.$$

Now the element

$$k = \begin{pmatrix} 1 & q_1 & q_2 & r \\ 0 & 1 & 0 & q_1 \\ 0 & 0 & 1 & q_2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in H, \quad \text{where } r = \frac{1}{2}(q_1^2 + q_2^2), \tag{2.12}$$

transforms the curvature according to the formula

$$\begin{aligned} R_k^* \Omega &= k^{-1} \Omega k \\ &= \begin{pmatrix} 1 & -q_1 & -q_2 & r \\ 0 & 1 & 0 & -q_1 \\ 0 & 0 & 1 & -q_2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} e & u_1 & u_2 & 0 \\ 0 & 0 & -a & u_1 \\ 0 & a & 0 & u_2 \\ 0 & 0 & 0 & -e \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} 1 & q_1 & q_2 & r \\ 0 & 1 & 0 & q_1 \\ 0 & 0 & 1 & q_2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \theta_1 \wedge \theta_2 \\ &= \begin{pmatrix} e & u_1 - q_2 a + q_1 e & u_2 + q_1 a + q_2 e & 0 \\ 0 & 0 & -a & \star \\ 0 & a & 0 & \star \\ 0 & 0 & 0 & -e \end{pmatrix} \theta_1 \wedge \theta_2. \end{aligned}$$

The equations

$$\begin{aligned} u_1 - q_2 a + q_1 e &= 0, \\ u_2 + q_1 a + q_2 e &= 0 \end{aligned}$$

may be solved uniquely for q_1 and q_2 provided that

$$\det \begin{pmatrix} -a & e \\ e & a \end{pmatrix} = -(a^2 + e^2) \neq 0,$$

that is, provided that E and A never vanish simultaneously. Since we are assuming this, it follows not only that in any fiber there are points for which $Y_1 = Y_2 = 0$, but also that changing such a point by right multiplication by any element $k \neq I$ of the form given in Eq. (2.12) will never preserve this property. On the other hand, elements of the form

$$k = \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & \cos b & -\sin b & 0 \\ 0 & \sin b & \cos b & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix}$$

always preserve this property. Set

$$P_{\text{Weyl}} = \{p \in P \mid Y_1 = Y_2 = 0\}.$$

Then by Proposition **4**.2.14, P_{Weyl} is a reduction of P to the Weyl group H_N. The form

$$\omega = \begin{pmatrix} \varepsilon & v_1 & v_2 & 0 \\ \theta_1 & 0 & -\alpha & v_1 \\ \theta_2 & \alpha & 0 & v_2 \\ 0 & \theta_1 & \theta_2 & -\varepsilon \end{pmatrix}$$

restricts to give a form on P_{Weyl}, but of course it need not happen that $v_1 = v_2 = 0$. Nevertheless, the part of this form given by

$$\omega_{\text{Weyl}} = \begin{pmatrix} \varepsilon & 0 & 0 \\ \theta_1 & 0 & -\alpha \\ \theta_2 & \alpha & 0 \end{pmatrix}$$

on P_{Weyl} determines a Weyl geometry on M with curvature

$$\begin{aligned} \Omega_{\text{Weyl}} = d\omega_{\text{Weyl}} + \omega_{\text{Weyl}} \wedge \omega_{\text{Weyl}} &= \begin{pmatrix} d\varepsilon & 0 & 0 \\ d\theta_1 & 0 & -d\alpha \\ d\theta_2 & d\alpha & 0 \end{pmatrix} \\ &+ \begin{pmatrix} \varepsilon & 0 & 0 \\ \theta_1 & 0 & -\alpha \\ \theta_2 & \alpha & 0 \end{pmatrix} \wedge \begin{pmatrix} \varepsilon & 0 & 0 \\ \theta_1 & 0 & -\alpha \\ \theta_2 & \alpha & 0 \end{pmatrix} \\ &= \begin{pmatrix} d\varepsilon & 0 & 0 \\ d\theta_1 + \theta_1 \wedge \varepsilon - \alpha \wedge \theta_2 & 0 & -d\alpha \\ d\theta_2 + \theta_2 \wedge \varepsilon + \alpha \wedge \theta_1 & d\alpha & 0 \end{pmatrix} = \begin{pmatrix} d\varepsilon & 0 & 0 \\ 0 & 0 & -d\alpha \\ 0 & d\alpha & 0 \end{pmatrix}. \quad \blacksquare \end{aligned}$$

§3. Equivalence Problems for a Conformal Metric

In this section we study two equivalence problems associated with conformal metrics. The first is the equivalence problem for a conformal metric itself. The solution here is a certain normal Möbius geometry. The second is the equivalence problem for a "Weyl structure," which is a conformal metric together with a certain compatible family of 1-forms. Here the solution is a Weyl geometry.

The Equivalence Problem for a Conformal Metric

As we saw in Lemma 2.3, a Möbius geometry on M^n uniquely determines a conformal metric on M. Here we study the converse and show that to a conformal metric $\mathcal{G}$ on M^n, $n \geq 3$, there corresponds a unique normal Möbius geometry on M giving rise to $\mathcal{G}$ by the process described in Lemma 2.3. (The case $n = 2$ is exceptional in this regard; cf. the discussion given in the course of the proof below of Proposition 3.1.)

This result is an immediate consequence of the following proposition.

Proposition 3.1. *To each conformal metric $\mathcal{G}$ on the open set $U \subset \mathbf{R}^n$, $n > 2$, there corresponds a unique normal Möbius geometry on U whose associated conformal metric is the original one.*

Proof. We first find a Cartan gauge on U. Let q be a representative metric in the given conformal class of metrics on U. We may choose, with respect to q, an orthonormal basis of tangent vector fields $e_i(x)$ along U. Let $\theta_i: T(U) \to \mathbf{R}$, $i = 1, \ldots, n$, be the dual 1-forms on U. Then

$$q(v) = \sum_i \theta_i(v)^2.$$

Conversely, every decomposition of q as a sum of squares arises in this way.

A gauge on U modeled on Möbius geometry may be written in block form as

$$\omega = \begin{pmatrix} \varepsilon & v & 0 \\ \theta & \alpha & v^t \\ 0 & \theta^t & -\varepsilon \end{pmatrix}.$$

We claim it is no restriction to consider only gauges satisfying $\varepsilon = 0$, which we shall assume in the sequel. To justify this, consider a change of gauge of the form $h: U \to H$, where

$$h = \begin{pmatrix} 1 & r & \frac{1}{2}rr^t \\ 0 & 1 & r^t \\ 0 & 0 & 1 \end{pmatrix}.$$

This will effect the replacement $\omega \to \mathrm{Ad}(h^{-1})\omega + h^*\omega_H$, or

$$\begin{pmatrix} \varepsilon & v & 0 \\ \theta & \alpha & v^t \\ 0 & \theta^t & -\varepsilon \end{pmatrix} \mapsto \begin{pmatrix} \varepsilon - r\theta & \star & 0 \\ \theta & \star & \star \\ 0 & \theta^t & -\varepsilon + r\theta \end{pmatrix}.$$

Since the θ_i, $i = 1, \ldots, n$, is a basis for the 1-forms on U, the form $r\theta = \sum_i r_i\theta_i$ is an arbitrary 1-form on U. Thus, we may take it to be equal to ε so that the $(1,1)$ entry of the new gauge vanishes.

The curvature of our gauge (with $\varepsilon = 0$) has the form

$$d\omega + \omega \wedge \omega = \begin{pmatrix} v \wedge \theta & dv + v \wedge \alpha & 0 \\ d\theta + \alpha \wedge \theta & d\alpha + \alpha \wedge \alpha + \theta \wedge v + v^t \wedge \theta^t & \star \\ 0 & \star & \star \end{pmatrix}$$
$$= \begin{pmatrix} E & Y & 0 \\ T & A & \star \\ 0 & \star & \star \end{pmatrix}, \tag{3.2}$$

where the components denoted by stars are determined by symmetry from other components already displayed. We aim to choose ω so that T and E vanish and so that "A has no Ricci curvature."

According to Cartan's lemma **6**.3.3, given forms θ_i: $T(U) \to \mathbf{R}$, $i = 1, \ldots, n$, there is a unique skew symmetric matrix-valued form $\alpha = (\alpha_{ij})$: $T(U) \to \mathfrak{o}_n(\mathbf{R}) = \mathrm{Skew}_n(\mathbf{R})$ such that the torsion $T \equiv d\theta + \alpha \wedge \theta = 0$. Thus, the condition that $T = 0$ forces the choice of α.

We may write $v = \theta^t l$, where $l = (l_{ij})$ is some $n \times n$ matrix-valued function on U. We claim that the symmetry of l is necessary and sufficient for the vanishing of E. In fact,

$$E = v \wedge \theta = \theta^t l \wedge \theta = \sum_{1 \le i,j \le n} l_{ij} \theta_i \wedge \theta_j$$
$$= \sum_{1 \le i < j \le n} (l_{ij} - l_{ji}) \theta_i \wedge \theta_j.$$

Thus, $E = 0 \Leftrightarrow l^t = l$.

Finally, we ask if we can choose v (i.e., l) so that $K_{\mathfrak{s}}$ lies in the kernel of the Ricci homomorphism *Ricci*: $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{s}) \to (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$. There are two parts to this.

I. We show that no matter what v is, $Ricci(K_{\mathfrak{s}})$ lies in the symmetric submodule $S^2(\mathfrak{g}/\mathfrak{h})^* \subset (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$. (Since the $\mathfrak{s}$ here corresponds to $\mathfrak{h}$ in Proposition **6**.1.4, the symmetry of the values of the Ricci homomorphism follows from that discussion. However, for the convenience of the reader, we reprove this result in the present context.)

II. For $n \ge 3$, there is a unique choice of l such that $K_{\mathfrak{s}} \in \ker Ricci$.

Proof of I. $Ricci(K_{\mathfrak{s}})$ lies in the symmetric submodule of $(\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$. Since $T = 0$, the Bianchi identity corresponding to this block reads $A \wedge \theta = 0$, or $\sum_s A_{ks} \wedge \theta_s = 0$ for $1 \le k, s \le n$. Evaluating this on $e_i \wedge e_j \wedge e_k$ yields

$$0 = \sum_s A_{ks} \wedge \theta_s (e_i \wedge e_j \wedge e_k)$$
$$= 2 \sum_s \{A_{ks}(e_i, e_j)\delta_{sk} + A_{ks}(e_j, e_k)\delta_{si} + A_{ks}(e_k, e_i)\delta_{sj}\}$$
$$= 2\{0 + A_{ki}(e_j, e_k) + A_{kj}(e_k, e_i)\}.$$

Thus, $A_{ki}(e_j, e_k) = A_{kj}(e_i, e_k)$, which shows that

$$K_{\mathfrak{s}} = \sum_{\substack{i<j \\ p<q}} A_{pq}(e_i, e_j) e_i^* \wedge e_j^* \otimes e_{pq}$$

satisfies the conditions of Exercise 1.22(iii), and so

$$Ricci(K_{\mathfrak{s}}) = \sum_{i,j,k} A_{kj}(e_i, e_n) e_i^* \otimes e_j^*$$

is symmetric. ■

Proof of II. Let us see first how the curvature block A depends on l. The structural Eq. (3.2) implies (where E_{pq} is the matrix with 1 in position (p, q) and zeroes elsewhere)

$$A - (d\alpha + \alpha \wedge \alpha) = \theta \wedge v + v^t \wedge \theta^t$$
$$= \theta \wedge (\theta^t l) + (l^t \theta) \wedge \theta^t$$
$$= \sum_{j,p,q} \theta_p \wedge \theta_j l_{jq} \otimes E_{pq} + \sum_{i,p,q} l_{ip} \theta_i \wedge \theta_q \otimes E_{pq}$$
$$= \sum_{i,j,p,q} \delta_{ip} l_{jq} \theta_i \wedge \theta_j \otimes E_{pq} + \sum_{i,j,p,q} \delta_{jq} l_{ip} \theta_i \wedge \theta_j \otimes E_{pq}.$$

In the second sum we switch $i \leftrightarrow j$, $p \leftrightarrow q$ and reorder the wedge product to get

$$A - (d\alpha + \alpha \wedge \alpha) = \sum_{i,j,p,q} \delta_{ip} l_{jq} \theta_i \wedge \theta_j \otimes E_{pq}$$
$$- \sum_{i,j,p,q} \delta_{ip} l_{jq} \theta_i \wedge \theta_j \otimes E_{qp}$$
$$= \sum_{i,j,p,q} \delta_{ip} l_{jq} \theta_i \wedge \theta_j \otimes e_{pq},$$

where $e_{pq} = E_{pq} - E_{qp}$. Thus, we see that the terms in $K_{\mathfrak{s}}$ involving l are

$$\sum_{i,j,p,q} \delta_{ip} l_{jq} e_i^* \wedge e_j^* \otimes e_{pq} = -\sum_{i,j,q} l_{jq} e_j^* \wedge e_i^* \otimes e_{iq}$$
$$= -\frac{1}{2} \sum_{i,j,q} l_{jq} (e_j^* \wedge e_i^* \otimes e_{iq} + e_q^* \wedge e_i^* \otimes e_{ij})$$
$$= -\frac{1}{2} \sum_{j,q} l_{jq} b_{jq},$$

where we have used the symmetry of l and the definition of b_{ij} given in Exercise 1.22(iv). Both of these ingredients are used again to show that the terms in $Ricci(K_{\mathfrak{s}})$ involving l are

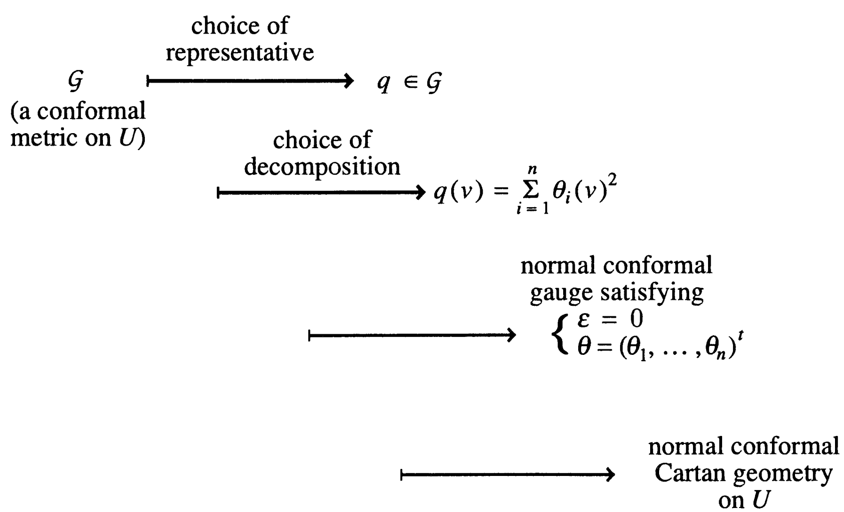
$$\begin{aligned} -\frac{1}{2}\sum_{j,q} l_{jq} Ricci(b_{jq}) &= -\frac{1}{2}\sum_{j,q} l_{jq}\big((n-2)(e_j^* \otimes e_q^* + e_q^* \otimes e_j^*) \\ &\quad + 2\delta_{jq}\Sigma_k e_k^* \otimes e_k^*\big) \\ &= -\frac{1}{2}(n-2)\Sigma_{j,q} l_{jq}(e_j^* \otimes e_q^* + e_q^* \otimes e_j^*) \\ &\quad - \sum_j l_{jj}\Sigma_k e_k^* \otimes e_k^* \\ &= -(n-2)\sum_{j,q} l_{jq} e_j^* \otimes e_q^* - \mathrm{Trace}(l)\Sigma_k e_k^* \otimes e_k^* \end{aligned}$$

It is easily seen that for $n > 2$ this is an arbitrary element of the symmetric submodule of $(\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$ and that l is determined by this expression. Thus, in that case there is a unique value of l for which $Ricci(K_{\mathfrak{s}}) = 0$. On the other hand, when $n = 2$, the expression reduces to

$$-\mathrm{Trace}(l)\sum_i e_i^* \otimes e_i^*,$$

in which case we do not get an arbitrary element of the symmetric submodule of $(\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$, nor is l determined by this element.

Let us review what we have done so far. We have described a procedure that may be represented schematically as in the following diagram.



It is clear from the construction that the conformal class of metric obtained from the resulting geometry is $\mathcal{G}$ again. Moreover, the procedure we have

described is completely determined except for the choices in the first two stages. Let us investigate the effect of varying these two choices. If $\tilde{\theta}_i$ is another coframe decomposing another choice $q' \in \mathcal{G}$, then we may write $\theta_i = \sum b_{ij}\tilde{\theta}_j$, where at each point $(b_{ij}) = za^{-1}$, a positive scalar multiple of z of a rotation a^{-1}. Let us set

$$h = \begin{pmatrix} z & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & z^{-1} \end{pmatrix}.$$

Then $h: U \to H$ is a gauge transformation and

$$\tilde{\omega} = \mathrm{Ad}(h^{-1})\omega + h^*\omega_H$$

is also a normal gauge (since normality is a gauge-invariant notion). Moreover, $\tilde{\theta} = \mathrm{Ad}(h^{-1})\theta \bmod \mathfrak{h}$, so $\tilde{\omega}$ is the unique normal gauge associated to the decomposition of q' by $\tilde{\theta}$ as in the construction above. In particular, varying the two choices merely replaces the gauge ω by the equivalent gauge $\tilde{\omega}$ and has no effect on the corresponding normal Möbius geometry. ■

Exercise 3.3. In the derivation of step **I** of Proposition 3.1 it is assumed that $e_i \wedge e_j \wedge e_k \neq 0$, which requires $n > 2$. Verify that $A_{ki}(e_j, e_k) = A_{kj}(e_i, e_k)$ even when $n = 2$. ❑

Corollary 3.4. *Let (P, ω) be a Möbius geometry on M^n, $n > 2$, of type $\mathfrak{s} \oplus \mathfrak{q}$. Then there is a unique replacement for the block v of ω (cf. Eq. (2.5)) so that the resulting geometry on M is normal.*

Proof. Let $\mathcal{G}$ be the conformal structure on M determined by the given Möbius geometry. By Proposition 3.1 there is a unique geometry on M of type $\mathfrak{s} \oplus \mathfrak{q}$ with *Ricci* $K_{\mathfrak{s}}$ in the normal submodule. Moreover, the construction of this geometry given in Proposition 3.1 shows that, except for the block v, all the blocks of its Cartan connection are determined by $\mathcal{G}$ and the condition that the geometry have type $\mathfrak{s} \oplus \mathfrak{q}$. It follows that, except for the block v, the geometry (P, ω) is the one constructed in Proposition 3.1 and that there is a unique choice of the block v yielding a normal geometry. ■

Corollary 3.5. *Let $U \subset S^n$ $(n > 2)$ be any connected and simply connected open subset and let $f: U \to S^n$ be any smooth immersion pulling back the canonical metric q_0 on S^n to a metric f^*q_0 conformally equivalent to the induced metric $q_0 \mid U$. Then there is an element $g \in G$ such that $f(x) = gx$ for all $x \in U$. In particular, f is injective and extends to a conformal automorphism of the whole sphere.*

Proof. Since $n > 2$, there are canonical normal Möbius geometries associated to the conformal metrics on U and S^n. By the theorem, f must be a local geometric isomorphism of these geometries. So, by Theorem **5**.5.2, f is the restriction of the action by left translation of an element $g \in G$. ■

Exercise 3.6. Show that this corollary is false if $n = 2$. [*Hint:* the Riemann mapping theorem says that *any* connected and simply connected proper open subset $U \subset \mathbf{C}$ is diffeomorphic to the open unit disc by a holomorphic and therefore conformal mapping.]

Corollary 3.7. *The group of conformal diffeomorphisms of S^n $(n > 2)$ is the Möbius group G.*

Proof. By Lemma 1.12 we know that G is a subgroup of the conformal diffeomorphisms of S^n. By Corollary 3.5 we know that, for $n > 2$, any conformal diffeomorphism is an element of G. ■

This corollary also holds for $n = 2$, although the proof given is no longer valid.

The Möbius Geometry Associated to a Riemannian Geometry

In this subsection we shall identify the Weyl curvature, introduced in Table **6**.2.5 for a Riemannian manifold M^n with metric q, with part of the curvature for the normal Möbius geometry associated to the conformal class of metric determined by q. This will show not only that the Weyl curvature of a Riemannian manifold is an invariant of the conformal class of the metric but that for $n \geq 4$ its vanishing is a necessary and sufficient condition for the metric to be locally conformally flat. More precisely, we have the following result.

Theorem 3.8. *Let M^n be a smooth manifold. If q is a Riemannian metric on M, let $W(q)$ denote the Weyl curvature of the Riemannian manifold (M, q).*

(i) *If q_1 and q_2 are two Riemannian metrics on M differing by a smooth positive factor, then $W(q_1) = W(q_2)$.*

(ii) *If $W(q) = 0$ and $n \geq 4$, then on each sufficiently small open set U about any point in M, there is a flat Riemannian metric on U that differs from q by a smooth positive factor.*

Proof. This theorem is local, so we work with a gauge on an open set $U \subset M$. We may write the Riemannian gauge and (torsion free) curvature as

$$\omega_{\text{Rie}} = \begin{pmatrix} 0 & 0 \\ \theta & \alpha \end{pmatrix}$$

and

$$\Omega_{\text{Rie}} = \begin{pmatrix} 0 & 0 \\ d\theta + \alpha \wedge \theta & d\alpha + \alpha \wedge \alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & d\alpha + \alpha \wedge \alpha \end{pmatrix}.$$

The corresponding normal Möbius gauge and curvature for the conformal metric determined by q according to the procedure in Proposition 3.1 are

$$\omega_{\text{Möb}} = \begin{pmatrix} 0 & v & 0 \\ \theta & \alpha & v^t \\ 0 & \theta^t & 0 \end{pmatrix},$$

$$\Omega_{\text{Möb}} = \begin{pmatrix} v \wedge \theta & dv + v \wedge \alpha & 0 \\ d\theta + \alpha \wedge \theta & d\alpha + \alpha \wedge \alpha + \theta \wedge v + v^t \wedge \theta^t & \star \\ 0 & \star & \star \end{pmatrix}$$
$$= \begin{pmatrix} 0 & dv + v \wedge \alpha & 0 \\ 0 & d\alpha + \alpha \wedge \alpha + \theta \wedge v + v^t \wedge \theta^t & \star \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that we are taking Riemannian and Möbius gauges with the same choice of θ. Since in each case the condition that the geometry be torsion free is the same (i.e., $d\theta + \theta \wedge \alpha = 0$), it follows from Cartan's Lemma **6**.3.4, that the same α appears in both gauges. Moreover, since $n \geq 3$, Corollary 3.4 shows that the block v is determined by the condition that the Möbius geometry is normal. We claim that the identity

$$d\alpha + \alpha \wedge \alpha = (d\alpha + \alpha \wedge \alpha + \theta \wedge v + v^t \wedge \theta^t) - (\theta \wedge v + v^t \wedge \theta^t)$$

is the decomposition of the Riemannian curvature into its Weyl curvature and its Ricci curvature (cf. Proposition **6**.1.3(iii) and Table **6**.2.5). To see this, note first that by the Definition 2.7 of a normal connection, the part K_5 of the Möbius curvature function corresponding to $d\alpha + \alpha \wedge \alpha + \theta \wedge v + v^t \wedge \theta^t$ must lie in the Weyl submodule (the kernel of the Ricci homomorphism). Now consider the other term, $\theta \wedge v + v^t \wedge \theta^t$, and write $v = \theta^t l$, where $l = (l_{ij})$ is symmetric. The part of the curvature function corresponding to this term is, as in the proof of **II** in Proposition 3.1,

$$\sum_{i,j,p,q} \delta_{ip} l_{jq} e_i^* \wedge e_j^* \otimes e_{pq} = -\frac{1}{2} \sum_{j,q} l_{jq} b_{jq}$$

which, by Exercise 1.22(iv), lies in the Ricci submodule. In particular, $W(q)$ is the curvature of the associated normal Möbius geometry and hence depends only on the conformal class of q, proving (i). On the other hand, when $n \geq 4$, Proposition 2.6 assures us that when $W(q) = 0$, then $\Omega_{\text{Möb}} = 0$, so the Möbius geometry is flat. ■

The following exercise shows that even if $W(q) = 0$ and $n \geq 4$, q need not be *globally* conformally equivalent to a flat metric.

Exercise 3.9. Let $\tilde{M} = \mathbf{R}^n - \{0\}$ be equipped with the standard metric induced from Euclidean n space $\mathbf{R}^n$. Define an action $\mathbf{Z} \times \tilde{M} \to \tilde{M}$ by the formula $n \cdot x = 2^n x$.

(i) Show that this is an action by conformal covering transformations so that $M = \tilde{M}/\mathbf{Z}$ acquires a conformal structure $\mathcal{G}$.

(ii) Show that M is diffeomorphic to the product $S^1 \times S^n$.

(iii) Show that M is a locally flat conformal structure with holonomy group $\mathbf{Z}$.

(iv) Deduce that M is not complete.

(v) Deduce that there is no global flat Riemannian metric on M in the conformal class $\mathcal{G}$. ❑

On the other hand, we have the following result.

Proposition 3.10. *Let M^n be a simply connected, complete, locally flat Möbius geometry. Then it is the model geometry.*

Proof. This is a special case of Theorem 5.5.3. ■

The Equivalence Problem for a Weyl Structure

A Weyl structure is a refinement of the notion of a conformal metric. This is the structure actually used by Weyl (see [H. Weyl, 1952]) in his attempt to unify the forces of gravity and electromagnetism. We are going to show that a Weyl structure is the same as a torsion free Weyl geometry.

Definition 3.11. A *Weyl structure* on M is a conformal metric $\mathcal{G}$ on M together with a function $F: \mathcal{G} \to A^1(M)$ satisfying $F(e^\lambda q) = F(q) - d\lambda$ for every smooth function $\lambda: M \to \mathbf{R}$. ❀

Lemma 3.12. *A Weyl geometry on M determines a canonical Weyl structure on M.*

Proof. Let $U \subset M$ be a small open set and let ω be an arbitrary gauge on U for the Weyl geometry on M. Write

$$\omega = \begin{pmatrix} \varepsilon & 0 \\ \theta & \alpha \end{pmatrix},$$

$$\Omega = d\omega + \omega \wedge \omega = \begin{pmatrix} d\varepsilon & 0 \\ d\theta + \theta \wedge \varepsilon + \alpha \wedge \theta & d\alpha + \alpha \wedge \alpha \end{pmatrix}.$$

Define $q = \sum_{1 \le j \le n} \theta_j^2$ to be the Riemannian metric on M associated to the gauge ω. Then we define $F(q) = -2\varepsilon$. To see that this is a Weyl structure on M, we need to see that every expression of the form $e^\lambda q$ (for any $\lambda: M \to \mathbf{R}$) appears as the Riemannian metric on M associated to some gauge and that $F(e^\lambda q) = F(q) - d\lambda$.

If $\tilde{\omega}$ is another gauge on U for the given Weyl geometry, then it differs from ω by a unique change of gauge of the form

$$h = \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix} : U \to H_N,$$

and the gauges themselves are related by the formula

$$\begin{aligned} \tilde{\omega} &= \mathrm{Ad}(h^{-1})\omega + h^*\omega_{H_N} \\ &= \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon & 0 \\ \theta & \alpha \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix}^{-1} \begin{pmatrix} dz & 0 \\ 0 & da \end{pmatrix} \\ &= \begin{pmatrix} \varepsilon + z^{-1}dz & 0 \\ za^{-1}\theta & a^{-1}\alpha a + a^{-1}da \end{pmatrix}. \end{aligned} \tag{3.14}$$

Moreover, any such choice of h will yield a valid change of gauge.

It is clear from this that the Riemannian metric associated to $\tilde{\omega}$ is $z^2 q$ and that

$$F(z^2 q) = -2\varepsilon - 2z^{-1}dz = F(q) - 2z^{-1}dz.$$

Writing $e^\lambda = z^2$ so that $d\lambda = 2z^{-1}dz$ then yields the result. ■

Now we solve the equivalence problem for Weyl structures.

Proposition 3.14. *Let F by a Weyl structure on the smooth manifold M. Then there is a unique torsion free Weyl geometry on M giving rise, as in Lemma 3.12, to this Weyl structure.*

Proof. Fix $q \in \mathcal{G}$. On a sufficiently small open set U about any point of M, we may express q as $q = \sum \theta_i^2$ for some coframe $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$. We take for the gauge associated to this coframe the expression

$$\omega = \begin{pmatrix} -\frac{1}{2}F(q) & 0 \\ \theta & \alpha \end{pmatrix},$$

where (as usual, by Cartan's Lemma 6.3.4) α is the unique form such that the torsion, $d\theta - \frac{1}{2}\theta \wedge F(g) + \alpha \wedge \theta$, vanishes. Suppose we change our choice of q to $\tilde{q} = z^2 q$ and the choice of coframe θ to $\tilde{\theta} = za^{-1}\theta$, where $z: U \to \mathbf{R}^+$ and $a: U \to O_n(\mathbf{R})$. Setting $h = \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix} : U \to H_N$, we calculate

$$\begin{aligned} &\mathrm{Ad}(h^{-1})\omega + h^*\omega_{H_N} \\ &= \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix}^{-1} \begin{pmatrix} -\frac{1}{2}F(q) & 0 \\ \theta & \star \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix} + \begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix}^{-1} \begin{pmatrix} dz & 0 \\ 0 & \star \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} -\frac{1}{2}F(q) + z^{-1}dz & 0 \\ za^{-1}\theta & \star \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{1}{2}F(z^2q) & 0 \\ za^{-1}\theta & \star \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}F(\tilde{q}) & 0 \\ \tilde{\theta} & \star \end{pmatrix}.$$

Since a gauge transformation will not alter the torsion, the vanishing of the torsion in the new gauge is automatic. Thus, the $(2,2)$ entry of this gauge is the correct one.

In summary, we see that altering the choice of metric and coframe has no effect on the Weyl geometry. It merely replaces the given gauge by one equivalent to it. Thus, the torsion free Weyl geometry is uniquely determined by the Weyl structure. Moreover, it is clear from the discussion that the Weyl structure determined by this geometry according to Lemma 3.12 is just the original one. ∎

§4. Submanifolds of Möbius Geometry

Suppose we are given a *torsion free* Möbius geometry N^{n+r} and an immersion $f: M^n \to N^{m+r}$. After establishing some algebraic notation, we explain how to obtain the "locally ambient geometry of N along f" (this notion is formalized in Definition 4.8). The main technical point is to find a certain reduction of the restriction to M of the ambient principal bundle to a principal bundle $H_\lambda \to P_\lambda \to M$ with group $H_\lambda \approx \text{Möb}_m(\mathbf{R})_0 \times O_r(\mathbf{R})$. This is achieved by taking a subbundle of the tangent reduction along which a certain collection of traces vanishes. The locally ambient geometry of N along f is the bundle P_λ together with the restriction to it of the Cartan connection for N. This is analogous to the Riemannian case in that this geometric entity contains, in principle, all of the geometry of the map f. In Proposition 4.17 we show that, when N is the Möbius sphere, this datum does indeed determine the immersion $f: M \to N$ up to a Möbius transformation of N. Thus, in this case, the locally ambient geometry of N along M is a complete invariant for the map f.

The next step in studying the locally ambient Möbius geometry on M is to see how it fractures into the tangential part (the "intrinsic" Möbius geometry on M, which is not necessarily normal), the normal[4] part (the Ehresmann connection on the normal bundle ν), and the "gluing" data (the second fundamental form). This decomposition depends on factoring the bundle P_λ as a fiber product $P_\lambda \approx P_{\tan} \times_M P_{\text{nor}}$. The tangential and normal parts of the locally ambient geometry on M are essentially given by the restrictions $\omega \mid P_{\tan}$ and $\omega \mid P_{\text{nor}}$ of the Cartan connection ω. These

[4]Note the two different uses of the word *normal* here. The meaning should always be clear from the context.

pieces exist for any torsion free, locally ambient Möbius geometry on M. In Corollary 4.30, we show that when dim $M \neq 2$ and the locally ambient geometry is normal, then it can be reconstructed from these pieces. A different approach is needed when M is a surface; this is discussed in §6 when codimension is one.

Since the "intrinsic" Möbius geometry $(P_{\tan}, \omega \mid P_{\tan})$ on M is not necessarily normal, it is *not* the geometry on M corresponding to the conformal metric induced on it by the conformal metric associated (by Lemma 2.3) to the Möbius geometry on N. Nevertheless, the former determines the latter (cf. Proposition 4.20(iii)). This means that more data about the geometry of the immersion $M \to N$ are packed into the "intrinsic" geometry of M than is possible with a normal geometry.

The discussion of submanifolds given here parallels the modern treatment of the Riemannian case given in the last chapter. We remark that a more classical treatment, analogous to the classical treatment of the Riemannian case, may be found in [A. Fialkow, 1944]. See page 315ff for a brief look at Fialkow's method. Whereas the present treatment is quite general, Fialkow's treatment requires the absence of umbilic points.

Model Algebra

The models for the Möbius geometries on N^{m+r} and M^n involve the Klein pairs of Lie algebras $(\mathfrak{g}, \mathfrak{h})$ and $(\mathfrak{a}, \mathfrak{b})$, respectively, where

$$\mathfrak{g} = \left\{ \begin{pmatrix} \star & \star & \star & 0 \\ \star & \star & \star & \star \\ \star & \star & \star & \star \\ 0 & \star & \star & \star \end{pmatrix} \right\}, \quad \mathfrak{h} = \left\{ \begin{pmatrix} \star & \star & \star & 0 \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & 0 & 0 & \star \end{pmatrix} \right\},$$
$$\mathfrak{a} = \left\{ \begin{pmatrix} \star & \star & 0 & 0 \\ \star & \star & 0 & \star \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \star \end{pmatrix} \right\}, \quad \mathfrak{b} = \left\{ \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & 0 & \star \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \star \end{pmatrix} \right\},$$

and $\mathfrak{g}$ has blocks down the main diagonal of size 1×1, $n \times n$, $r \times r$, and 1×1. Of course, we must bear in mind that not all possible entries are allowed; a matrix in any of these Lie algebras must have the form

$$\begin{pmatrix} z & p & q & 0 \\ u & a & -b^t & p^t \\ v & b & c & q^t \\ 0 & u^t & v^t & -z \end{pmatrix},$$

where a and c are skew-symmetric matrices. Let us set

$$\mathfrak{z} = \left\{ \begin{pmatrix} \star & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \star \end{pmatrix} \right\}, \quad \mathfrak{s} = \left\{ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \star & \star & 0 \\ 0 & \star & \star & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\},$$

$$\mathfrak{o}(n) = \left\{ \begin{pmatrix} \star & 0 & 0 & 0 \\ 0 & \star & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}, \quad \mathfrak{o}(r) = \left\{ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \star & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\},$$

$$\mathfrak{t} = \left\{ \begin{pmatrix} 0 & 0 & 0 & 0 \\ \star & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \star & 0 & 0 \end{pmatrix} \right\}, \quad \mathfrak{u} = \left\{ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \star & 0 & 0 & 0 \\ 0 & 0 & \star & 0 \end{pmatrix} \right\},$$

$$\mathfrak{q} = = \left\{ \begin{pmatrix} 0 & \star & \star & 0 \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}.$$

Then we have the canonical $\mathfrak{z}\oplus\mathfrak{o}(n)\oplus\mathfrak{o}(r)$ module decomposition $\mathfrak{g}/\mathfrak{h} = \mathfrak{t}\oplus\mathfrak{u}$.

The First Reduction

Recall that we are studying an immersion $f\colon M^n \to N^{n+r}$, where N is equipped with a Möbius geometry (P,ω). The description of the geometry of N localized along f is more complicated than in the case of Riemannian geometry, although it begins in the same way with the tangent reduction P_τ along f. This is the reduction of the pullback bundle $f^*(P)$ given by

$$P_\tau = \{(x,p) \in f^*(P) \mid \varphi_p(f_*(T_x(M))) = \mathfrak{a}/\mathfrak{b} \subset \mathfrak{g}/\mathfrak{h}\}.$$

The reduction P_τ sits in the commutative diagram

$$\begin{tikzcd} P_\tau \arrow[d] & \subset & f^*(P) \arrow[r, "\tilde{f}"'] \arrow[d] & P \arrow[d] \\ M & \overset{\mathrm{id}}{=} & M \arrow[r, "f"] & N \end{tikzcd}$$

The tangent reduction P_τ has group $H_\tau = \{h \in H \mid \mathrm{Ad}(h^{-1})\mathfrak{a}/\mathfrak{b} = \mathfrak{a}/\mathfrak{b}\}$. An easy calculation (starting from the description of H given in Lemma 1.8) shows that H_τ consists of all elements of H having the form

$$\begin{pmatrix} 1 & p & q & s \\ 0 & 1 & 0 & p^t \\ 0 & 0 & 1 & q^t \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix}$$

where $p^t \in \mathbf{R}^n$, $q^t \in \mathbf{R}^n$, $s = \frac{1}{2}(pp^t + qq^t) \in \mathbf{R}$, $z \in \mathbf{R}^+$, $a \in O_n(\mathbf{R})$, $c \in O_r(\mathbf{R})$. The Lie algebra $\mathfrak{h}_\tau$ of H_τ consists of elements of the form

$$\begin{pmatrix} \star & \star & \star & \star \\ 0 & \star & 0 & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & 0 & \star \end{pmatrix}.$$

If we set $\omega_\tau = \tilde{f}^*(\omega) \mid P_\tau$, then we have the following.

Lemma 4.1. $\tilde{f}^*(\omega)(T(P_\tau)) = \mathfrak{a} \bmod \mathfrak{h}$.

Proof. $\mathfrak{a}/\mathfrak{b} = \varphi_p(f_*(T_{\pi(p)}M)) = \varphi_p(f_*\pi_*(T_pP_\tau)) = \omega_p(\tilde{f}_*(T_pP\tau)) \bmod \mathfrak{h} = \tilde{f}^*(\omega)(T(P_\tau)) \bmod \mathfrak{h}$. ■

Corollary 4.2. *$\omega_\tau \bmod \mathfrak{h}$ takes values in $\mathfrak{a}/\mathfrak{b} \subset \mathfrak{g}/\mathfrak{h}$ and hence ω_τ has the form*

$$\omega_\tau = \begin{pmatrix} \varepsilon & \upsilon & \nu & 0 \\ \theta & \alpha & -\beta^t & \upsilon^t \\ 0 & \beta & \delta & \nu^t \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}.$$

The Second Fundamental Form on the First Reduction

To continue we need to find a reduction of P_τ involving some of the "normal" information encoded in the second fundamental form. We note that this procedure has no analog in Riemannian geometry, where the tangential information alone is sufficient to describe all of the reduction needed.

The appearance of an aspect of the second fundamental form for the very definition of the second bundle reduction requires that we introduce it right away, not in the final form in which it will eventually appear, but in its raw form as a collection of matrices.

The form ω_τ transforms under $k \in H_\tau$ according to

$$R_k^*\omega_\tau = \mathrm{Ad}(k^{-1})\omega_\tau = k^{-1} \begin{pmatrix} \varepsilon & \upsilon & \nu & 0 \\ \theta & \alpha & -\beta^t & \upsilon^t \\ 0 & \beta & \gamma & \nu^t \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix} k,$$

where

$$k = \begin{pmatrix} 1 & p & q & s \\ 0 & 1 & 0 & p^t \\ 0 & 0 & 1 & q^t \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix}, \text{ with } s = \frac{1}{2}(pp^t + qq^t). \quad (4.3)$$

For any $X \in \mathfrak{h}_\tau$, we have $\omega_\tau(X^\dagger) = X$ and hence $\beta(X^\dagger) = 0$. Since the value of $X^\dagger$ at some point $p \in P_\tau$ runs through all the fiber directions as X varies in $\mathfrak{h}_\tau$, it follows that the form β vanishes on each fiber and is therefore semibasic. This implies that the ith row of β may be expressed as $\beta_i = \theta^t h(i)$, where $h(i)$, $1 \le i \le r$, is an $n \times n$ matrix-valued function on P_τ. These $h(i)$ constitute the raw version of the second fundamental form[5] on M.

Lemma 4.4. *The vanishing of the* $(3,1)$ *curvature block is equivalent to the symmetry of* $h(i)$ *for all* i.

Proof. The vanishing of the $(3,1)$ curvature block is equivalent to $\beta \wedge \theta = 0$, or, for all i,

$$\begin{aligned} 0 = \beta_i \wedge \theta = \theta^t h(i) \wedge \theta &= \sum_{p,q} \theta_p h(i)_{pq} \wedge \theta_q \\ &= \sum_{p<q} (h(i)_{pq} - h(i)_{qp}) \theta_p \wedge \theta_q, \end{aligned}$$

which is equivalent to the symmetry of $h(i)$ for all i. ∎

Since we are dealing here only with the case when the ambient geometry is torsion free, then in particular the $(3,1)$ block of the curvature *does* vanish. Thus, the following expression, Eq. (4.5), is symmetric in v_1 and v_2:

$$\begin{aligned} b(p)(v_1, v_2) = \beta(\omega_p^{-1}(v_1)) v_2 &= \begin{pmatrix} \theta^t(\omega_p^{-1}(v_1)) h(1) \\ \cdots \\ \theta^t(\omega_p^{-1}(v_1)) h(r) \end{pmatrix} v_2 \\ &= \begin{pmatrix} v_1^t h(1) v_2 \\ \cdots \\ v_1^t h(r) v_2 \end{pmatrix}, \quad \text{where } v_1, v_2 \in \mathfrak{t}. \end{aligned} \tag{4.5}$$

To proceed further, we need to know how the $h(i)$ transform under the action of $k \in H_\tau$.

Lemma 4.6. *For* k *as in Eq.* (4.3), *and for each* j,

$$R_k^* h(j) = z^{-1} z^{-1} a^{-1} \left\{ \sum_i c_{ij} (h(i) + q_i I) \right\} a.$$

Proof. Since $\beta_i = \theta^t h(i)$, we first need to calculate how θ and β on P_τ change under the action of $k \in H_\tau$. Using the formula $R_k \omega = \mathrm{Ad}(k^{-1})\omega$ with k as in Eq. (4.3), we easily calculate that

$$R_k^* \beta = c^{-1}(\beta + q^t \theta^t) a \quad \text{and} \quad R_k^* \theta = z a^{-1} \theta.$$

[5] Of course, it is only in the Riemannian case that there is, properly speaking, a *first* fundamental form. It is the Riemannian metrics itself.

On the one hand,

$$R_k^* \beta = c^{-1}(\beta + q^t \theta^t) a = c^{-1} \begin{pmatrix} \beta_1 + q_1 \theta^t \\ \cdots \\ \beta_r + q_r \theta^t \end{pmatrix} a = c^{-1} \begin{pmatrix} \theta^t (h(1) + q_1 I) a \\ \cdots \\ \theta^t (h(r) + q_r I) a \end{pmatrix}.$$

On the other hand,

$$\begin{aligned} R_k^* \beta = R_k^* \begin{pmatrix} \theta^t h(1) \\ \cdots \\ \theta^t h(r) \end{pmatrix} &= \begin{pmatrix} R_k^* \theta^t R_k^* h(1) \\ \cdots \\ R_k^* \theta^t R_k^* h(r) \end{pmatrix} \\ &= \begin{pmatrix} (z a^{-1} \theta)^t R_k^* h(1) \\ \cdots \\ (z a^{-1} \theta)^t R_k^* h(r) \end{pmatrix} = \begin{pmatrix} \theta^t z a R_k^* h(1) \\ \cdots \\ \theta^t z a R_k^* h(r) \end{pmatrix}. \end{aligned}$$

Combining these yields, for each j,

$$\begin{aligned} \theta^t z a R_k^* h(j) &= j\text{th row of } c^{-1} \begin{pmatrix} \theta^t (h(1) + q_1 l) a \\ \cdots \\ \theta^t (h(r) + q_r l) a \end{pmatrix} \\ &= \sum_i c_{ij} \theta^t (h(i) + q_i I) a, \end{aligned}$$

which, since θ has linearly independent entries, yields the result. ∎

The Second Reduction P_λ

From Lemma 4.6 we see that, for any point $p \in P_\tau$, there is a $k \in H_\tau$ of the form

$$k = \begin{pmatrix} 1 & 0 & q & s \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & q^t \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{where } s = \frac{1}{2} q q^t,$$

such that, at $pk \in P_\tau$, all of the symmetric matrices $h(j)$ have trace zero. (Just take $q_j = -\text{trace } h(j)$ for all j.) Moreover, given any point $p \in P_\tau$ at which trace $h(j)(p) = 0$ for all j, then in the notation of Eq. (4.3), for $k \in H_\tau$,

$$\text{trace } h(j)(pk) = 0 \quad \text{for all } j \Leftrightarrow q_j = 0 \text{ for all } j.$$

This, together with the smoothness of the function trace $h(j)$, is sufficient (cf. Proposition 4.2.14) for the equations trace $h(j) = 0$, $1 \le j \le r$, to give a further reduction of the principal bundle $H_\tau \to P_\tau \to M$ to a principal bundle $H_\lambda \to P_\lambda \to M$, where

$$H_\lambda = \left\{ \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p^t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{ where } s = \frac{1}{2} p p^t \right\}.$$

Note that H_λ is isomorphic to a product of the two subgroups

$$H_\lambda \approx \text{Möb}_m(\mathbf{R})_0 \times O_r(\mathbf{R})$$

corresponding to the decomposition of group elements given by

$$\begin{aligned} g &= \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p^t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \left(\begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p^t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) \\ &\quad \times \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{where } s = \frac{1}{2}pp^t). \end{aligned} \tag{4.7}$$

It follows that the Lie algebra of H_λ is given by

$$\mathfrak{h}_\lambda = \mathfrak{b} \oplus \mathfrak{o}(r) = \left\{ \begin{pmatrix} \star & \star & 0 & 0 \\ 0 & \star & 0 & \star \\ 0 & 0 & \star & 0 \\ 0 & 0 & 0 & \star \end{pmatrix} \right\}.$$

We set $\omega_\lambda = \tilde{f}^*(\omega) \mid P_\lambda$. The pair $(P_\lambda, \omega_\lambda)$ is our candidate for the invariant describing the immersion f (cf. Lemma 4.14).

Geometry of N Localized Along f

The preceding discussion leads us to the following definition, which is justified in Lemma 4.14.

Definition 4.8. Let M be a smooth manifold of dimension n. A *locally ambient Möbius geometry on M of codimension r* is a pair (P, ω), where

(i) P is a principal $H_\lambda = \text{Möb}_n(\mathbf{R})_0 \times O_r(\mathbf{R})$ bundle over M,

(ii) ω is an $(\mathfrak{h} + \mathfrak{a})$-valued form on P satisfying the conditions

(a) $\omega: T(P) \to \mathfrak{h} + \mathfrak{a}$ is injective at each point; $\omega^{-1}(\mathfrak{h})$ is tangent to the fiber at each point; and the composite of ω and the canonical H_λ module projection $\mathfrak{h} + \mathfrak{a} \to \mathfrak{a} \oplus \mathfrak{o}(r)$ are surjective at each point.

(b) $R_k^*\omega = \text{Ad}(k^{-1})\omega$ for all $k \in H_\lambda$.

(c) $\omega(X^\dagger) = X$ for each $X \in \mathfrak{h}_\lambda$ $(= \mathfrak{b} \oplus \mathfrak{o}(r))$.

Just as for the case of a submanifold, we may write

$$\omega = \begin{pmatrix} \varepsilon & \upsilon & \nu & 0 \\ \theta & \alpha & -\beta^t & \upsilon^t \\ 0 & \beta & \gamma & \nu^t \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}.$$

(d) The ith row $\beta(i)$ of β has the form $\beta(i) = \theta^t h(i)$, where $h(i)$ is an $n \times n$ symmetric matrix of trace zero. (This is the condition that the ambient geometry be torsion free along M in the direction normal to M.)

The *curvature* of (P, ω) is $d\omega + \omega \wedge \omega =$

$$\begin{pmatrix} d\varepsilon + \upsilon \wedge \theta & d\upsilon + \varepsilon \wedge \upsilon + \upsilon \wedge \alpha + \nu \wedge \beta & d\nu + \varepsilon \wedge \nu - \upsilon \wedge \beta^t + \nu \wedge \gamma & 0 \\ d\theta + \theta \wedge \varepsilon + \alpha \wedge \theta & d\alpha + \theta \wedge \upsilon + \alpha \wedge \alpha - \beta^t \wedge \beta + \upsilon^t \wedge \theta^t & \star & \star \\ 0 & d\beta + \beta \wedge \alpha + \gamma \wedge \beta + \nu^t \wedge \theta^t & d\gamma - \beta \wedge \beta^t + \gamma \wedge \gamma^t & \star \\ 0 & \star & 0 & \star \end{pmatrix}.$$

(P, ω) is called *torsion free* if $d\theta + \varepsilon \wedge \theta + \alpha \wedge \theta = 0$.

(P, ω) is called *normal* if it is torsion free, of type $\mathfrak{s} \oplus \mathfrak{q}$, and $K_{\mathfrak{s}}$ lies in the kernel of the Ricci homomorphism restricted to $\text{Hom}(\lambda^2(\mathfrak{a}/\mathfrak{b}), \mathfrak{s}) \subset \text{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{s})$. ($K_{\mathfrak{s}}$ is the part of the curvature function corresponding to the $(2,2)$, $(2,3)$, $(3,2)$, and $(3,3)$ curvature blocks.)

A *gauge* on $U \subset M$ for a locally ambient geometry (P, ω) is a form $\sigma^*(\omega)$ where $\sigma: U \to P$ is any section.

Two local ambient geometries (P_1, ω_1) and (P_2, ω_2) on M are said to be *equivalent* if there is an $\text{Möb}_n(\mathbf{R}) \times O_r(\mathbf{R})$ bundle map $b: P_1 \to P_2$ such that $b^*(\omega_2) = \omega_1$. ❀

We note that since $\omega(X^\dagger) = X$ for each $X \in \mathfrak{h}_\lambda$, it follows that $\nu(X^\dagger) = 0$ for $X \in \mathfrak{h}_\lambda$. Thus ν vanishes on all vectors tangent to the fiber of P, namely, ν is semibasic. We also note without proof that an effective locally ambient geometry may be described by its gauges just as for Cartan geometries (cf. Chapter 5, §2).

Proposition 4.9. *Let M^n be a manifold of dimension $n > 0$ and let*

$$\omega = \begin{pmatrix} \varepsilon & \upsilon & \nu & 0 \\ \theta & \alpha & -\beta^t & \upsilon^t \\ 0 & \beta & \gamma & \nu^t \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}$$

be a gauge for a locally ambient Möbius geometry of codimension r on M. Regard α, β, γ, and ε as fixed. Let us write the forms υ and ν in terms of the basis θ arising from the gauge as $(\upsilon\ \nu) = \theta^t l$ where l is an $n \times (n + r)$ matrix.

(a) (i) *If $n \neq 2$ there is a unique choice of ν such that $\text{Ricci}(K_{\mathfrak{s}})$ takes values in the span of $e_i^* \otimes e_j^*$ for $1 \leq i, j \leq n$.*

(ii) *If $n = 2$ then ν has no influence on $Ricci(K_{\mathfrak{s}})$.*

(b) (i) *If $n > 2$ there is a unique choice of v such that $Ricci(K_{\mathfrak{s}})$ takes values in the span of $e_i^* \otimes e_j^*$ for $1 \le i \le n$, $n+1 \le j \le n+r$.*

(ii) *If $n = 2$ only the trace of v $(= l_{11} + l_{22})$ has an influence on $Ricci(K_{\mathfrak{s}})$. It contributes the terms $-(l_{11}+l_{22})(e_1^* \otimes e_1^* + e_2^* \otimes e_2^*)$.*

(iii) *If $n = 1$ then v has no influence on $Ricci(K_{\mathfrak{s}})$.*

Proof. We begin by calculating the contributions made by v and ν to $Ricci(K_{\mathfrak{s}})$. From Definition 4.8(d) we see that the terms in the $\mathfrak{s}$ block of the curvature matrix arising from v and ν are

$$
\begin{aligned}
\begin{pmatrix} \theta \wedge v + v^t \wedge \theta^t & \theta \wedge \nu \\ \nu^t \wedge \theta^t & 0 \end{pmatrix} &= \begin{pmatrix} \theta \\ 0 \end{pmatrix} \wedge (v\ \nu) + (v\ \nu)^t \wedge \begin{pmatrix} \theta \\ 0 \end{pmatrix}^t \\
&= \begin{pmatrix} \theta \\ 0 \end{pmatrix} \wedge \theta^t l + l^t \theta \wedge \begin{pmatrix} \theta \\ 0 \end{pmatrix}^t \\
&= \sum_{\substack{1\le p,k\le n \\ 1\le q\le n+r}} \theta_p \wedge \theta_k l_{kq} E_{pq} + \sum_{\substack{1\le p,k\le n \\ 1\le q\le n+r}} l_{kq} \theta_k \wedge \theta_p E_{qp} \\
&= \sum_{\substack{1\le p,k\le n \\ 1\le q\le n+r}} l_{kq} \theta_p \wedge \theta_k (E_{pq} - E_{qp}) \\
&= \sum_{\substack{1\le p,k\le n \\ 1\le q\le n+r}} l_{kq} \theta_p \wedge \theta_k e_{pq}.
\end{aligned}
$$

It follows that the terms in $K_{\mathfrak{s}}$ arising from v and ν are

$$
- \sum_{\substack{1\le p,k\le n \\ 1\le q\le n+r}} l_{kq} e_k^* \wedge e_p^* \otimes e_{pq}.
$$

Hence, by Exercise 1.22(i), in terms of $Ricci(K_{\mathfrak{s}})$ arising from v and ν are

$$
\begin{aligned}
&- \sum_{\substack{1\le p,k\le n \\ 1\le q\le n+r}} l_{kq} Ricci(e_k^* \wedge e_p^* \otimes e_{pq}) \\
&= - \sum_{\substack{1\le p,k\le n, 1\le q\le n+r \\ k\ne p\ne q}} l_{kq}(e_k^* \otimes e_q^* + \delta_{kq} e_p^* \otimes e_p^*) \\
&= - \sum_{\substack{1\le p,k\le n, 1\le q\le n+r \\ k\ne p\ne q, k\ne q}} l_{k1} e_k^* \otimes e_q^* - \sum_{\substack{1\le p,k\le n, 1\le q\le n+r \\ k\ne p\ne q, k=q}} l_{kq} e_k^* \otimes e_q^* \\
&\quad - \sum_{\substack{1\le p,k\le n, 1\le q\le n+r \\ k\ne p\ne q}} l_{kq} \delta_{kq} e_p^* \otimes e_p^*
\end{aligned}
$$

$$
\begin{aligned}
&= -(n-2) \sum_{\substack{1\le k\le n, 1\le q\le n+r \\ k\ne q}} l_{kq} e_k^* \otimes e_q^* - (n-1) \sum_{\substack{1\le k\le n, 1\le q\le n+r \\ k=q}} l_{kq} e_k^* \otimes e_q^* \\
&\quad - \sum_{\substack{1\le p,k\le n \\ k\ne p}} l_{kk} e_p^* \otimes e_p^* \\
&= -(n-2) \sum_{\substack{1\le k\le n \\ 1\le q\le n+r}} l_{kq} e_k^* \otimes e_q^* - \sum_{1\le k\le n} l_{kk} e_k^* \otimes e_k^* \\
&\quad - \sum_{1\le p,k\le n} l_{kk} e_p^* \otimes e_p^* + \sum_{1\le k\le n} l_{kk} e_k^* \otimes e_k^* \\
&= -(n-2) \sum_{\substack{1\le k\le n \\ 1\le q\le n+r}} l_{kq} e_k^* \otimes e_q^* - \sum_{1\le k\le n} l_{kk} \sum_{1\le p\le n} e_p^* \otimes e_p^* \qquad (*)
\end{aligned}
$$

The terms in the formula $(*)$ involving $e_k^* \otimes e_q^*$ for $1 \le k \le n$, $n+1 \le q \le n+r$, which is the same as the terms depending on ν, are

$$
-(n-2) \sum_{\substack{1\le k\le n \\ n+1\le q\le n+r}} l_{kq} e_k^* \otimes e_q^* \qquad (**)
$$

Since this vanishes for $n = 2$, we verify (a)(ii). When $n \ne 2$ it is clear from the expression $(**)$ that any variation of ν will be reflected in a variation of the coefficients of $e_k^* \otimes e_q^*$ for $1 \le k \le n$, $n+1 \le q \le n+r$ and, moreover, any variation in these coefficients may be achieved by a suitable choice of ν. This proves (a)(i).

The terms in the formula $(*)$ involving $e_k^* \otimes e_q^*$ for $1 \le k, q \le n$, which is the same as the terms depending on v, are

$$
-(n-2) \sum_{\substack{1\le k\le n \\ 1\le q\le n}} l_{kq} e_k^* \otimes e_q^* - \sum_{1\le k\le n} l_{kk} \sum_{1\le p\le n} e_p^* \otimes e_p^*. \qquad (***)
$$

Note that the sum of the coefficients of the terms $e_k^* \otimes e_k^*$, $1 \le k \le n$ in this expression is

$$
-2(n-1) \sum_{1\le k\le n} l_{kk}.
$$

Thus for $n \ne 1$, the expression $(***)$ determines the expression

$$
-(n-2) \sum_{\substack{1\le k\le n \\ 1\le q\le n}} l_{kq} e_k^* \otimes e_q^*
$$

and for $n \ne 2$ the converse holds. These remarks show that, for $n > 2$, any variation of v will be reflected in a variation of the coefficients of $e_k^* \otimes e_q^*$ for $1 \le k \le n$, $1 \le q \le n$ and, moreover, any variation in these coefficients may be achieved by a suitable choice of v. This proves (b)(i). If $n = 2$, the expression $(***)$ reduces to

$$-(l_{11} + l_{22})(e_1^* \otimes e_1^* + e_2^* \otimes e_2^*)$$

which verifies (b)(ii). Finally, if $n = 1$, the expression $(***)$ vanishes and (b)(iii) is verified. ■

Exercise 4.10.* Prove the following analog of Proposition 2.6. Let (P, ω) be a locally ambient Möbius geometry on M^n of type $\mathfrak{z} \oplus \mathfrak{q}$.

(i) If $\dim M \geq 3$, then the $\mathfrak{z}$ block of the curvature vanishes.

(ii) If $\dim M \geq 4$, then the curvature vanishes. ❑

The following result is an analog of Theorem **5**.3.15.

Proposition 4.11. *Let (P, ω) be a locally ambient Möbius geometry on M of codimension r and let $x \in M$. Then for each $p \in P$ over x, there is a unique linear isomorphism $\varphi_p: T_{\pi(p)}M \to \mathfrak{a}/\mathfrak{b}$ such that the following diagram commutes.*

$$\begin{array}{ccc}
T_p(P) & \xrightarrow[\approx]{\omega_p} & \mathfrak{h} + \mathfrak{a} \\
\downarrow & & \downarrow \\
T_x(P) & \overset{\varphi_p}{\underset{\approx}{\dashrightarrow}} & \mathfrak{a}/\mathfrak{b} \subset (\mathfrak{h} + \mathfrak{a})/\mathfrak{h}
\end{array}$$

Moreover, if $h \in Mob_m(\mathbf{R}) \times O_r(\mathbf{R})$, then $\varphi_{ph} = Ad(h^{-1})\varphi_p$.

Proof. The existence, linearity, uniqueness, and injectivity of φ_p follow from the fact that

$$\omega_p(v) = 0 \bmod \mathfrak{h} \Leftrightarrow v \text{ is tangent to the fiber of } P \to M.$$

Since $\dim M = \dim \mathfrak{a}/\mathfrak{b}$, it follows that φ_p is an isomorphism. The final fact comes from applying $\mathrm{Ad}(h)^{-1}$ to the equation $\varphi_p \circ \pi_{*p} = \omega_p \bmod \mathfrak{h}$ to get

$$\begin{aligned}
\mathrm{Ad}(h)^{-1}\varphi_p \circ \pi_{*p} &= \mathrm{Ad}(h)^{-1}\omega_p \bmod \mathfrak{h} \\
&= \omega_{ph} \circ R_{h*} \bmod \mathfrak{h} \\
&= \varphi_{ph} \circ \pi_{*ph} \circ R_{h*} \\
&= \varphi_{ph} \circ \pi_{*p} \quad (\text{since } \pi \circ R_h = \pi).
\end{aligned}$$

Since π_{*p} is surjective, it follows that $\varphi_{ph} = \mathrm{Ad}(h^{-1})\varphi_p$. ■

Corollary 4.12. *In the case of a locally ambient geometry arising from an immersion $f: M \to N$ into a torsion free Möbius geometry, the map $\varphi_p: T_x(M) \to \mathfrak{a}/\mathfrak{b}$ given by the theorem is the pullback, via f^*, of the corresponding map $\varphi_{\tilde{f}(p)}: T_{f(x)}(N) \to \mathfrak{g}/\mathfrak{h}$ for the Möbius geometry on N.*

Proof. This is simply the fact that, in terms of the form ω, in both cases the map φ is given by the block θ. ■

Definition 4.13. Let (P, ω) be a locally ambient Möbius geometry on M of codimension r. The *normal bundle* associated to (P, ω) is the r-dimensional vector bundle $\nu_{\text{nor}} = P \times_{H_\lambda} \mathbf{R}^r$, where the action of $H_\lambda = \text{Möb}_m(\mathbf{R})_0 \times O_r(\mathbf{R})$ on $\mathbf{R}^r$ is the standard second factor action. $q_{\text{nor}}: \nu_{\text{nor}} \to \mathbf{R}$ is the canonical (up to constant scale) metric on ν_{nor}. ❀

The point of Definitions 4.8 and 4.13 is that they describe the data associated to the immersion $f: M \to N$ we have been discussing.

Lemma 4.14. *The pair $(P_\lambda, \omega_\lambda)$ described on page 300 is a normal, locally ambient Möbius geometry on M of codimension r.*

Proof. Condition (i) of Definition 4.8 is obvious, as is the fact that ω_λ is an $(\mathfrak{h} + \mathfrak{a})$-valued form on P_λ. Condition (ii)(b) is an automatic consequence of the formula on P that $R_h^*\omega = \mathrm{Ad}(h^{-1})\omega$ for all $h \in H$. Condition (ii)(c) comes from the definition of a Cartan geometry and the fact that the inclusion $P_\lambda \subset f^*(P)$ is equivariant with respect to the right actions of H_λ; hence for $X \in \mathfrak{h}_\lambda$, the vector field $X^\dagger$ on $f^*(P)$ restricts to the vector field $X^\dagger$ on P_λ. The injectivity of (ii)(a) comes from the fact that $\omega_\lambda = \tilde{f}^*(\omega) \mid P_\lambda, f$ is an immersion, and ω is injective. The rest of (ii)(a) comes from combining (ii)(c) with the fact (Lemma 4.1) that $\omega_\tau(T_p(P_\tau)) = \mathfrak{a} \bmod \mathfrak{h}$ for all $p \in P_\lambda$. Condition (ii)(d) follows from the nature of the first and second reductions. ■

We emphasize this example of a locally ambient geometry by giving the following definition.

Definition 4.15. Let N be a Möbius manifold and let $f: M \to N$ be an immersion. *The locally ambient Möbius geometry of n along f* (or *along M* if f is an inclusion) is the pair $(P_\lambda, \omega_\lambda)$. ❀

Exercise 4.16. Let $\phi: N \to N$ be a geometrical automorphism of a Möbius geometry on N. If $f: M \to N$ is an immersion, show that the locally ambient Möbius geometry of N along f is equivalent to the locally ambient conformal geometry of N along ϕf. ❑

Now we can justify the terminology "ambient Möbius geometry localized along f" by showing that, at least in the case when the ambient geometry is the model Möbius sphere, it does indeed encapsulate all the geometry of f.

Proposition 4.17. (i) *Let $f: M^n \to S^{n+r}$ (standard Möbius sphere) be an immersion. If $(P_\lambda, \omega_\lambda)$ is the ambient geometry localized along f, then $d\omega_\tau + \omega_\tau \wedge \omega_\tau = 0$ and the monodromy representation Φ_{ω_τ} is trivial.*

(ii) *Conversely, let M^n be a connected n-dimensional manifold and let (P, ω) be a locally ambient Möbius geometry on M of codimension r satisfying $d\omega + \omega \wedge \omega = 0$ with trivial monodromy. Then there is an immersion $f: M^n \to S^{n+r}$, determined up to left multiplication with a Möbius motion of $\mathbf{R}^{n+r}$, such that the Möbius geometry of S^{n+r} localized along f is equivalent to the geometry (P, ω).*

Proof. This is just an application of the fundamental theorem of calculus, Theorem **3**.7.14. ■

Decomposing a Locally Ambient Geometry

Let (P, ω) be a torsion free locally ambient Möbius geometry on M of codimension r. As in the Riemannian case studied in the last chapter, we wish to decompose (P, ω) into three pieces. The first two of these pieces are the induced Möbius geometry on M and the Ehresmann connection on the normal bundle. These are described ahead in Propositions 4.18 and 4.20. The final piece is the second fundamental form described in Proposition 4.21 and Definition 4.22. As we mentioned before, although these three pieces exist in general, the locally ambient geometry may not in general be reconstructed from these pieces alone without the presence of some extra hypothesis such as that of normality (cf. Theorem 4.29).

First we show that the bundle P itself decomposes as the fiber product of principal bundles determined by the tangent bundle of M and the associated normal vector bundle ν_{nor}.

Proposition 4.18. *Let (P, ω) be a locally ambient Möbius geometry on M of codimension r and let $P_{tan} = P/O(r)$ and $P_{nor} = P/Mob_b(\mathbf{R})_0$.*

(i) *P_{tan} is a principal $Mob_n(\mathbf{R})_0$ bundle and P_{nor} is a principal $O_r(\mathbf{R})$ bundle.*

(ii) *$P = P_{tan} \times_M P_{nor}$ (fiber product).*

(iii) *There is a canonical bundle equivalent $P_{tan} \times_{Mob_n(R)_0} \mathfrak{a}/\mathfrak{b}$, where the action of $Mob_n(\mathbf{R})_0$ on $\mathfrak{a}/\mathfrak{b}$ is induced from the adjoint action.*

(iv) *Let ν_{nor} be the normal bundle over M associated to (P, ω). Then P_{nor} is the principal bundle associated to ν_{nor}. Moreover, if (P, ω) is the locally ambient geometry of an immersion f, then ν_{nor} is the normal bundle of f.*

Proof. (i) Since the right actions of $\text{Möb}_n(\mathbf{R})_0$ and of $O_r(\mathbf{R})$ on P commute, we see that $\text{Möb}_n(\mathbf{R})_0$ acts smoothly on $P_{\tan}$. Since the action of $\text{Möb}_n(\mathbf{R})_0 \times O_r(\mathbf{R})$ is proper on P and transitive and effective on the fibers of $P \to M$, it follows that the action of $\text{Möb}_n(\mathbf{R})_0$ is proper on $P_{\tan}$ (cf. the proof of Lemma **4**.3.12) and transitive and effective on the fibers of $P_{\tan} \to M$. Hence $P_{\tan}$ is a right principal $\text{Möb}_n(\mathbf{R})_0$ bundle over M. The case of $P_{\text{nor}} \to M$ is similar.

(ii) The two smooth bundle maps $O_r(\mathbf{R}) \to P \to P_{\tan}$ and $\text{Möb}_n(\mathbf{R})_0 \to P \to P_{\text{nor}}$ (defined by projection) induce a smooth map into the fiber product $P \to P_{\tan} \times_M P_{\text{nor}}$. This map covers the identity on M and is obviously an $\text{Möb}_n(\mathbf{R})_0 \times O_r(\mathbf{R})$ bundle isomorphism.

(iii) Define $P \times \mathfrak{a}/\mathfrak{b} \to T(M)$ by $(p, v) \to \varphi_p^{-1}(v)$ (cf. Proposition 4.11). Since every $h \in O_r(\mathbf{R})$ acts trivially on $\mathfrak{a}/\mathfrak{b}$, we have $\varphi_{ph}^{-1}(v) = \varphi_p^{-1}(\text{Ad}(h)v) = \varphi_p^{-1}(v)$. Thus, the map $P \times \mathfrak{a}/\mathfrak{b} \to T(M)$ induces a smooth bundle map $P_{\tan} \times \mathfrak{a}/\mathfrak{b} = (P/O_n(\mathbf{R})) \times \mathfrak{a}/\mathfrak{b} \to T(M)$. But we also have, for every $k \in \text{Möb}_n(\mathbf{R})_0$, $\varphi_{pk}^{-1}(v) = \varphi_p^{-1}(\text{Ad}(k)v)$. It follows that the points (p, v), $(pk, \text{Ad}(k^{-1})v) \in P_{\tan} \times \mathfrak{a}/\mathfrak{b}$ have the same image in $T(M)$. Thus, we get a further induced map $P_{\tan} \times_{\text{Möb}_n(R)_0} \mathfrak{a}/\mathfrak{b} \to T(M)$. This map is a vector bundle map covering the identity on M, and the fibers have the same dimensions. Thus, it is a bundle equivalence.

(iv) This is similar to the proof in the Riemannian case (cf. **6**.5.9(iii)) and we leave it to the reader. ■

Now we are in a position to study how the form ω of a locally ambient Möbius geometry (P, ω) decomposes. Let us write

$$\omega = \begin{pmatrix} \varepsilon & v & \nu & 0 \\ \theta & \alpha & -\beta^t & v^t \\ 0 & \beta & \gamma & \nu^t \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}, \quad \omega_{\tan} = \begin{pmatrix} \varepsilon & v & 0 & 0 \\ \theta & \alpha & 0 & v^t \\ 0 & 0 & 0 & 0 \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}. \tag{4.19}$$

The next step is to see that the parts $\omega_{\tan}$ and γ of the form ω determine, and are determined by, certain induced forms (to which we give the same names) on $P_{\tan}$ and P_{nor}.

Proposition 4.20. *Let (P, ω) be a locally ambient Möbius geometry on M of codimension r.*

(i) *The form ω_{tan} (respectively, γ) is basic for the canonical projection $P \to P_{tan}$ (respectively, P_{nor}). We use the same name ω_{tan} (respectively, ω_{nor}) to denote the corresponding form on P_{tan} (respectively, P_{nor}).*

(ii) *Let γ be the form on P_{nor} guaranteed by* (i). *Then γ is an Ehresmann connection on the principal bundle $O_r(\mathbf{R}) \to P_{nor} \to M$.*

(iii) *Let ω_{tan} be the form on P_{tan} guaranteed by* (i). *Then (P_{tan}, ω_{tan}) is a Möbius geometry on M. Moreover, if (P, ω) arises from an immersion*

$f: M \to N$, then the conformal metric on M associated to (P_{tan}, ω_{tan}) is the conformal metric induced on M from N.

(iv) *(P, ω) is torsion free $\Leftrightarrow$ (P_{tan}, ω_{tan}) is torsion free.*

Proof. (i) We deal with γ only, the cases of $\omega_{\tan}$ being similar. The transformation law in Definition 4.8(ii) (b) implies that, for $h \in \text{Möb}_n(\mathbf{R})_0$, $R_h^*\gamma = \gamma$, and moreover, 4.8(ii) (c) says that $\omega(X^\dagger) = X$ for $X \in \mathfrak{h}_\lambda$, so in particular $\gamma(X^\dagger) = 0$ for $X \in \mathfrak{a}$. Thus, by Lemma **1**.5.25, γ is basic for the principal bundle $P \to P/\text{Möb}_n(\mathbf{R})_0 = P_{\text{nor}}$.

(ii) The transformation law 4.8(ii) (b) implies that, for $h \in O_r(\mathbf{R})$, $R_h^*\gamma = \text{Ad}(h^{-1})\gamma$, and moreover 4.8(ii) (c) says that $\omega(X^\dagger) = X$ for $X \in \mathfrak{h}_\lambda$, so in particular, $\gamma(X^\dagger) = X$ for $X \in \mathfrak{o}(r)$. Since the projection $P \to P/\text{Möb}_n(\mathbf{R})_0$ is $O_r(\mathbf{R})$ equivalent and $\pi^*: A^1(P/\text{Möb}_n(\mathbf{R})_0, \mathfrak{o}(r)) \to A^1(P, \mathfrak{o}(r))$ is injective, it follows that the formulas $R_h^*\gamma = \text{Ad}(h^{-1})\gamma$ for $h \in O_r(\mathbf{R})$ and $\gamma(X^\dagger) = X$ for $X \in \mathfrak{o}(r)$ also hold for the version of γ on P_{nor}. Thus, γ satisfies the conditions for an Ehresmann connection (cf. Definition **6**.2.4 of Appendix A) on P_{nor}.

(iii) To see that $\omega_{\tan}$ is a trivialization of the tangent bundle of $P_{\tan}$, since $\dim P_{\tan} = \dim \mathfrak{o}$, it suffices to show that $\omega_{\tan}: T(P_{\tan}) \to \mathfrak{o}$ is surjective, or, equivalently, that the corresponding form $\omega_{\tan}: T(P) \to \mathfrak{o}$ is surjective. But since $\omega_{\tan}$ is just the canonical projection of ω on $\mathfrak{o}$, by condition (ii)(a) of Definition 4.8, it is surjective. The proof that $\omega_{\tan}$ satisfies the conditions

(a) $R_h^*\omega_{\tan} = \text{Ad}(h^{-1})\omega_{\tan}$ for $h \in O_n(\mathbf{R})$,

(b) $\omega_{\tan}(X^\dagger) = X$ for $X \in \mathfrak{b}$

is similar to the proof of (ii). Thus, $\omega_{\tan}$ is a Möbius geometry on M.

If (P, ω) arises from an immersion f, then the induced conformal metric $q_M: T(M) \to \mathbf{R}$ on M arising from the Möbius geometry $(P_{\tan}, \omega_{\tan})$ is, using Corollary 4.12, given as follows:

$$\begin{aligned} q_M(v) &= \|\varphi_p(v)\|, \quad \text{where } \varphi_p \text{ arises from the geometry of } M \\ &= \|\varphi_{\tilde{f}(p)}(f_*(v))\|, \quad \text{where } \varphi_{\tilde{f}(p)} \text{ arises from the} \\ &\qquad \text{locally ambient geometry of } f \\ &= q_N(f_*(v)). \end{aligned}$$

Part (iv) is obvious. ∎

The Second Fundamental Form

The second fundamental form is the last of the three pieces of a locally ambient geometry. From one point of view, the second fundamental form on M is just the restriction to P_λ of the trace zero matrices $h(j)$ defined on P_τ. This may be rephrased by saying that the second fundamental form is a differential form on M with values in a certain vector bundle $P_\lambda \times_\rho (S^2(\mathfrak{t}) \otimes \mathfrak{u})$. Equivalently, we may also say that it is a bilinear symmetric form B on $T(M)$ of trace zero and with values in the normal bundle ν_{nor}. All of this is described in the following result.

Proposition 4.21. *Let (P, ω) be a locally ambient geometry on M, and let v_{nor} be the associated normal bundle. The block β of the form ω on P given in Definition 4.8 determines, and is determined by, either of the following:*

(i) *the function $b \in A^0(P, \rho)$, where*

$$\rho = S^2(Ad^*) \otimes Ad: H_\lambda \to Gl(S^2(\mathfrak{t}^*) \otimes \mathfrak{u})),$$

given by

$$b(p)(v_1, v_2) = \beta(\omega_p^{-1}(v_1))v_2, \quad \text{where } v_i \in \mathfrak{t} \text{ for } i = 1, 2;$$

(ii) *The symmetric bilinear form $B \in Hom(S^2(T(M)), \nu_{nor})$ given by*

$$\varphi_p(B_x(v_1, v_2)) = b(p)(\varphi_p(v_1), \varphi_p(v_2)),$$

where $v_i \in T_x(M)$ for $i = 1, 2$.

Proof. (i) β is a 1-form on P taking values in $\text{Hom}(\mathfrak{t}, \mathfrak{u})$ $(= \mathfrak{t}^* \otimes \mathfrak{u})$. From the discussion on page 298, $b(p)(v_1, v_2) = \beta(\omega_p^{-1}(v_1))v_2$ is symmetric in v_1 and v_2, so that b may be regarded as a function on P with values in $S^2(\mathfrak{t}^*) \otimes \mathfrak{u}$. Now let us see how b transforms. Write $g \in \text{Möb}_n(\mathbf{R})_0 \times O_r(\mathbf{R})$ as in Eq. (4.7) so that, in particular, $\text{Ad}(g)v_1 = av_1$. Now β transforms according to $R_g^*\beta = c^{-1}\beta a$, and so b transforms according to

$$\begin{aligned} b(pg)(v_1, v_2) &= \beta(\omega_{pg}^{-1}(v_1))v_2 \\ &= \beta(R_{g*}\omega_p^{-1}(\text{Ad}(g)v_1))v_2 \\ &= (R_g^*\beta)(\omega_p^{-1}(av_1))v_2 \\ &= (c^{-1}\beta a)(\omega_p^{-1}(av_1))v_2 \\ &= c^{-1}\beta(\omega_p^{-1}(av_1))av_2 \\ &= c^{-1}b(p)(av_1, av_2) \\ &= ((\rho(g^{-1})b)(p))(v_1, v_2). \end{aligned}$$

Thus, $b \in A^0(P, \rho)$. Clearly, β determines and is determined by b.

(ii) Note that $b(p)(\varphi_p(v_1), \varphi_p(v_2)) \in \mathfrak{u}$. Since $\varphi_p^{-1}(\mathfrak{u}) = \nu_x =$ the normal space at $x \in M$, it follows that the formula defining $B_x(v_1, v_2)$ makes sense (i.e., it takes values in ν_{nor}); neverthless, we must show that it is independent of the choice of p lying over x. But this follows immediately from the transformation laws for b and for φ. Finally, we note that since φ_p is an isomorphism for each p, it follows that B and b determine each other. ■

Definition 4.22. Let (P, ω) be a locally ambient geometry on M of codimension r. The *second fundamental form* associated to (P, ω) is the bilinear symmetric form $B \in \mathrm{Hom}(S^2(T(M)), \nu_{\mathrm{nor}})$ given in Proposition 4.21(ii). Given a normal vector $X \in v_x$ $(x \in M)$, the corresponding *Weingarten map* $L_X \in \mathrm{End}(T_x(M))$ is defined by $\langle X, B(u,v)\rangle_v = \langle u, L_X(v)\rangle_M$. ❀

Definition 4.23. A bilnear symmetric form $B \in \mathrm{Hom}(S^2(T(m)), \nu_{\mathrm{nor}})$ has *trace zero* if, for any orthogonal equinormal basis $e_1, e_2, \dots, e_n$ of $T(M)$, $\sum_i B(e_i, e_i) = 0$. An element $b \in A^0(P, \rho)$, where $\rho = S^2(\mathrm{Ad}^*) \otimes \mathrm{Ad}\colon H_\lambda \to Gl(S^2(\mathfrak{t}^*) \otimes \mathfrak{u}))$, has *trace zero* if, for any othogonal equinormal basis $e_1, e_2, \dots, e_n$ of $\mathfrak{t}$, $\sum_i b(p)(e_i, e_i) = 0$ (cf. Exercise **1**.2.29). ❀

Exercise 4.24. (a) Show that the two notions of "trace zero" given in Definition 4.23 agree if B and b are related as in Proposition 4.21(ii).

(b) Show that an element $B \in \mathrm{Hom}(S^2(T(M)), \nu_{\mathrm{nor}})$ has trace zero if and only if the corresponding Weingarten maps L_X have trace zero for all $X \in \nu_x$.

(c) Show that the second fundamental form given in Definition 4.22 has trace zero. ❏

Lemma 4.25. *L_X is a self-adjoint map depending linearly on X.*

Proof. $\langle u, L_X(v)\rangle = \langle X, B(u,v)\rangle = \langle X, B(v,u)\rangle = \langle v, L_X(u)\rangle = \langle L_X(u), v\rangle$. Also,

$$\begin{aligned}\langle u, L_{aX+bY}(v)\rangle &= \langle aX + bY, B(u,v)\rangle = a\langle X, B(u,v)\rangle + b\langle Y, B(u,v)\rangle \\ &= a\langle u, L_X(v)\rangle + b\langle u, L_Y(v)\rangle = \langle u, aL_X(v) + bL_Y(v)\rangle,\end{aligned}$$

and so $L_{aX+bY} = aL_X + bL_Y$. ■

Definition 4.26. Let N be a torsion free Möbius geometry and let $f\colon M \to N$ be an immersion. A point $x \in M$ is called *umbilic* for this immersion if the second fundamental form vanishes at x. ❀

Note that when M is a curve, every point is umbilic, so this notion is not important for curves. (See Definition 5.1 for the appropriate notion for curves.)

Exercise 4.27. Let $M \subset N$ be a Riemannian inclusion. Then, by Proposition 3.1, we may regard N as a Möbius geometry. If B_{Rie} and $B_{\mathrm{M\ddot{o}b}}$ are the second fundamental forms in the corresponding geometries, show that $B_{\mathrm{M\ddot{o}b}} = B_{\mathrm{Rie}} - \frac{1}{n}\mathrm{M\ddot{o}b}\, B_{\mathrm{Rie}}$. Deduce that $x \in M$ is umbilic in the Riemannian sense if and only if it is umbilic in the Möbius sense. ❏

Reconstructing a Locally Ambient Geometry from Its Parts

Consider the following collection of data:

$$(4.28)\quad \begin{cases} \text{(a)} & \text{a Möbius geometry on } M \text{ of type } \mathfrak{s} \oplus \mathfrak{q} \\ & \text{(the "intrinsic geometry" } (P_{\tan}, \omega_{\tan})); \\[1ex] \text{(b)} & \text{a vector bundle on } M \text{ equipped with an Ehresmann} \\ & \text{connection } (\nu_{\mathrm{nor}} \text{ and } \gamma); \\[1ex] \text{(c)} & \text{a trace zero element of } \mathrm{Hom}(S^2(T(M)), \nu_{\mathrm{nor}}) \\ & \text{(the second fundamental form } B). \end{cases}$$

We have seen (Propositions 4.20 and 4.21(ii)) how a locally ambient geometry (P, ω) decomposes to determine data of this form. Now we study the converse.

Theorem 4.29. *Let M be a manifold of dimension $n \neq 2$. If we are given the data of the form in* (4.28)*, then there is a unique locally ambient geometry (P, ω) on M giving rise to it such that $Ricci(K_{\mathfrak{s}})$ takes values in the span of $e_i^* \otimes e_j^*$, $1 \le i, j \le n$.*

Proof. Note first that by Proposition 4.18(ii) we must have $P = P_{\tan} \times_M P_{\mathrm{nor}}$, and since P_{nor} is the principal bundle associated to ν_{nor} (Proposition 4.18(iv)), the data (a) and (b) determine P. We must now see that every block in the form

$$\omega = \begin{pmatrix} \varepsilon & v & \nu & 0 \\ \theta & \alpha & -\beta^t & v^t \\ 0 & \beta & \gamma & \nu^t \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}$$

is also determined. Pulling up the form

$$\omega_{\tan} = \begin{pmatrix} \varepsilon & v & 0 & 0 \\ \theta & \alpha & 0 & v^t \\ 0 & 0 & 0 & 0 \\ 0 & \theta^t & 0 & -\varepsilon \end{pmatrix}$$

of (a) to P determines every block of ω except for γ, β, and ν. But γ and β are determined by (b) and (c), respectively. Proposition 4.9(a) shows that ν is determined by the condition on $Ricci(K_{\mathfrak{s}})$. ■

Corollary 4.30. *Let M be a manifold of dimension $n \neq 2$. Then a normal locally ambient geometry (P, ω) on M is determined by the data of the form* (4.28) *it gives rise to.*

Proof. The final condition of Theorem 4.29 is automatic in the presence of normality. ■

Submanifolds of Möbius Spheres

Submanifolds M^m of the flat space S^{n+r} constitute an important special example of the general theory discussed above. We have already noted in Proposition 4.17 that the ambient geometry localized along f is a complete invariant for an immersion $f\colon M^n \to S^{n+r}$ up to a Möbius transformation of the ambient space S^{n+r}. Moreover, by Proposition 4.17, an arbitrary local ambient geometry arises from an immersion into the sphere if and only if its curvature and monodromy vanish. It is interesting to try to restate this result in terms of data of the form of (4.28). The immediate difficulty is that the monodromy condition doesn't seem to break up in this way. We avoid this difficulty by assuming that M is simply connected, so that the monodromy condition is automatic.

Theorem 4.31. *Let M be a simply connected manifold of dimension $n \geq 4$, and assume that*

(a) *(P_{tan}, ω_{tan}) is an arbitrary Möbius geometry on M of type $\mathfrak{s} \oplus \mathfrak{q}$ (cf. Exercise 1.13(ii)),*

(b) *ν_{nor} is an arbitrary r-dimensional vector bundle on M equipped with an Ehresmann connection γ,*

(c) *$B \in Hom(S^2(T(M)), \nu_{nor})$ is an arbitrary element of trace zero.*

Let β denote the $Hom(\mathfrak{t}, \mathfrak{u})$-valued form on P corresponding to B via the formula of Proposition 4.21. Let ν be the form guaranteed by Proposition 4.9(a). Then (a), (b), and (c) arise from an immersion $M \to S^{n+r}$ if and only if the following equations hold:

$$\begin{cases} d\alpha + \alpha \wedge \alpha + \theta \wedge \upsilon + \upsilon^t \wedge \theta^t = \beta^t \wedge \beta, \\ d\beta + \beta \wedge \alpha + \gamma \wedge \beta + \nu^t \wedge \theta^t = 0, \\ d\gamma + \gamma \wedge \gamma^t = \beta \wedge \beta^t \end{cases} \tag{4.32}$$

Proof. By Theorem 4.29, parts (a), (b), and (c) will always fit together to give a unique locally ambient geometry of codimension r of type $\mathfrak{s} \oplus \mathfrak{q}$ such that $Ricci(K_{\mathfrak{s}})$ takes values in the span of $e_i^* \otimes e_j^*$, $1 \leq i, j \leq n$. Referring to the curvature given in Definition 4.8(d), we see that the vanishing of the curvature implies Eq. (4.32). Conversely, in the presence of Eq. (4.32), the locally ambient geometry is normal and indeed of type $\mathfrak{q}$. Since $n \geq 4$, the curvature vanishes by Exercise 4.10(ii). The monodromy also vanishes since M is simply connected. Thus (a), (b) and (c) arise from an immersion $M \to S^{n+r}$. ∎

Comparison of Riemannian and Conformal Surfaces in Space

Suppose we are given a two-dimensional submanifold $M^2 \subset \mathbf{R}^{n+2}$ in Euclidean space. We may equip $\mathbf{R}^{n+2}$ with the canonical normal Möbius geometry associated with its Riemannian structure and then take the induced Möbius geometry on M. We may also take the induced Riemannian structure on M. This yields two geometries on M with different models and different curvatures. Here we study the question of the relation between them. It is convenient to work with the gauge version of the geometries. Because of the existence of the tangent reduction in the Riemannian case, we may choose a gauge on $\mathbf{R}^{n+2}$ such that the restriction to M of it and its curvature have the form

$$\omega_{\mathrm{Rie}} = \begin{pmatrix} 0 & 0 & 0 \\ \theta & \alpha & -\beta^t \\ 0 & \beta & \gamma \end{pmatrix},$$

$$\Omega_{\mathrm{Rie}} = \begin{pmatrix} 0 & 0 & 0 \\ d\theta + \alpha \wedge \theta & d\alpha + \alpha \wedge \alpha - \beta^t \wedge \beta & \star \\ 0 & d\beta + \beta \wedge \alpha + \gamma \wedge \beta & d\gamma + \gamma \wedge \gamma - \beta \wedge \beta^t \end{pmatrix} = 0.$$

The Riemannian geometry induced on the surface has gauge and curvature given by

$$\varpi_{\mathrm{Rie}} = \begin{pmatrix} 0 & 0 \\ \theta & \alpha \end{pmatrix},$$

$$\Xi_{\mathrm{Rie}} = \begin{pmatrix} 0 & 0 \\ d\theta + \alpha \wedge \theta & d\alpha + \alpha \wedge \alpha \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & \beta^t \wedge \beta \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & K\theta_1 \wedge \theta_2 \\ 0 & -K\theta_1 \wedge \theta_2 & 0 \end{pmatrix},$$

respectively, where K is the Gauss curvature.

Now we pass to the Möbius case. Once again the existence of the tangent reduction (which is, of course, the same as in the Riemannian case) allows us to choose a gauge on $\mathbf{R}^{n+2}$ such that the restriction to M of the gauge and its curvature are

$$\omega_{\mathrm{M\ddot{o}b}} = \begin{pmatrix} 0 & \upsilon & \nu & 0 \\ \theta & \alpha & -\beta^t & \upsilon^t \\ 0 & \beta & \gamma & \nu^t \\ 0 & \theta^t & 0 & 0 \end{pmatrix},$$

$$\Omega_{\mathrm{M\ddot{o}b}} = \begin{pmatrix} 0 & d\upsilon+\upsilon\wedge\alpha+\nu\wedge\beta & d\nu-\upsilon\wedge\beta^t+\nu\wedge\gamma & 0 \\ d\theta+\alpha\wedge\theta & d\alpha+\theta\wedge\upsilon+\alpha\wedge\alpha-\beta^t\wedge\beta+\upsilon^t\wedge\theta^t & \star & \star \\ 0 & d\beta+\beta\wedge\alpha+\gamma\wedge\beta & d\gamma-\beta\wedge\beta^t+\gamma\wedge\gamma & \star \\ 0 & \star & 0 & 0 \end{pmatrix} = 0,$$

where the θ, α, β, γ are the same as in the Riemannian case. But this is not yet the final reduction for a submanifold in conformal geometry (see the second reduction on page 302). The final reduction has the effect of making β trace free, replacing it with $\beta_0 = \beta - \frac{1}{2}$ Trace β, and hence replacing $h(i)$ by $h_0(i) = h(i) - \frac{1}{2}$ Trace $h(i)I$, while leaving the entries θ, α, and γ unaltered. The entries v and ν do change, to $\bar{v}$ and $\bar{\nu}$ say. Thus, the Möbius geometry induced on the surface itself has gauge and curvature given by

$$\varpi_{\text{Möb}} = \begin{pmatrix} 0 & \bar{v} & 0 \\ \theta & \alpha & \bar{v}^t \\ 0 & \theta^t & 0 \end{pmatrix},$$

$$\Xi_{\text{Möb}} = \begin{pmatrix} 0 & d\bar{v} + \bar{v}\wedge\alpha & 0 \\ d\theta + \alpha\wedge\theta & d\alpha + \theta\wedge\bar{v} + \alpha\wedge\alpha + \bar{v}^t\wedge\theta^t & \star \\ 0 & \star & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -\bar{v}\wedge\beta_0 & 0 \\ 0 & \beta_0^t\wedge\beta_0 & \star \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & L_1\theta_1\wedge\theta_2 & L_2\theta_1\wedge\theta_2 & 0 \\ 0 & 0 & -K_0\theta_1\wedge\theta_2 & \star \\ 0 & K_0\theta_1\wedge\theta_2 & 0 & \star \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Definition 4.33. K_0 is called the *Willmore curvature*[6] of the conformal surface M. ❀

Exercise 4.34. (a) Show that $K_0\colon M \to \mathbf{R}$ is independent of the choice of conformal gauge.

(b) Show that at points where K_0 doesn't vanish, there is no information contained in the functions L_1, L_2. [*Hint:* show that there is a choice of conformal gauge for which $L_1 = L_2 = 0$.]

(c) Show that at a point $x \in M$ where K_0 does vanish, the functions L_1, L_2 determine a canonical covector $\eta_x \in T_x(M)^*$. ❑

Writing $h(i) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, it follows that

$$h(i)\theta\wedge\theta^t h(i) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} 0 & \theta_1\wedge\theta_2 \\ -\theta_1\wedge\theta_2 & 0 \end{pmatrix}\begin{pmatrix} a & b \\ b & c \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\det(h(i))\theta_1\wedge\theta_2.$$

We also have

$$h_0(i) = \begin{pmatrix} a & b \\ b & c \end{pmatrix} - \begin{pmatrix} \frac{1}{2}(a+c) & 0 \\ 0 & \frac{1}{2}(a+c) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}(a-c) & b \\ b & -\frac{1}{2}(a-c) \end{pmatrix},$$

[6]The analogy with the Gauss curvature suggests the opposite sign for K_0. However, the present convention ensures that $K_0 \geq 0$.

so that

$$\det(h_0(i)) = -\frac{1}{4}(a-c)^2 - b^2 \quad (\leq 0)$$
$$= -\frac{1}{4}(a+c)^2 + ac - b^2$$
$$= -\left(\frac{1}{2}\text{Trace } h(i)\right)^2 + \det h(i).$$

Since

$$\beta^t\wedge\beta = (h(1)\theta\cdots h(n)\theta)\wedge\begin{pmatrix} \theta^t h(1) \\ \vdots \\ \theta^t h(n) \end{pmatrix} = \sum_{1\leq i\leq n} h(i)\theta\wedge\theta^t h(i),$$

it follows that

$$\beta^t\wedge\beta = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\theta_1\wedge\theta_2\sum_{1\leq i\leq n}\det h(i),$$

and similarly,

$$\beta_0^t\wedge\beta_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\theta_1\wedge\theta_2\sum_{1\leq i\leq n}\det h_0(i).$$

Thus, the Gauss and Willmore curvatures of the surface may be expressed as

$$K = \sum_{1\leq i\leq n}\det(h(i)),$$
$$K_0 = -\sum_{1\leq i\leq n}\det(h_0(i)) \quad (\geq 0)$$
$$= \sum_{1\leq i\leq n}\left\{\left(\frac{1}{2}\text{trace } h(i)\right)^2 - \det(h(i))\right\} = H^2 - K.$$

This is conceptually interesting as it shows that the conformal curvature, which is "intrinsic" to the conformal structure, when looked at from the Riemannian point of view involves the mean curvature, which is "extrinsic" in that view.

Willmore [T.J. Willmore, 1966] conjectures that for a torus T^2 embedded in $\mathbf{R}^3$, the inequality $\int_{T^2}(H^2-K)dA \geq 2\pi^2$ holds. This conjecture is still open. (See [T.J. Willmore, 1993] for a report of recent work.)

Fialkow's Relative Conformal Invariant Λ

Aaron Fialkow gave a classification of submanifolds of a conformal manifold provided the submanifold is free of umbilic points ([A. Fialkow, 1944]).

Since a conformal manifold supports a unique normal Möbius geometry (Proposition 3.1), the present classification is certainly more general than his. The assumption that the submanifold is free of umbilics allows Fialkow to associate to each representative metric q on M a canonical positive smooth function Λ on M that scales the same way the representative metric scales. Then q/Λ is a conformally invariant Riemannian metric on M. The same remarks apply to the metric on the normal bundle, and one may consider a conformally invariant second fundamental form b/Λ^2. In this way Fialkow converts conformal notions to Riemannian ones, making available the whole machinery of Riemannian geometry for the study of the conformal geometry of submanifolds. He gives a local classification in terms of symmetric tensors consisting of

(i) the metric q/Λ,

(ii) r additional tensors (some of which may vanish) of orders $4, 6, \ldots, 2r+2$,

(iii) the "deviation tensor" of order 2 (necessary only for surfaces).

The case of a surface in space is somewhat special in that (ii) consists of a single tensor, which is in fact the square of a tensor of order 2 so that three symmetric tensors of order 2 suffice for the local description of a hypersurface. For a comparison of Fialkow's classification and the present one in this case, see Exercises 6.6 and 6.7. Here we limit our discussion to Fialkow's function Λ.

Let $e_1, e_2, \ldots, e_r$ be an equinormal[7] basis for the normal space ν_x and let $L_1, L_2, \ldots, L_r$ be the corresponding Weingarten maps.

Lemma 4.35. *Let λ_{ik}, $i = 1, \ldots, n$, be the eigenvalues of L_k, $1 \le k \le r$. Then $\Lambda = \sum_{ijk}(\lambda_{ik} - \lambda_{jk})^2$ is invariant under an orthonormal change of the basis $e_1, e_2, \ldots, e_r$ of the normal space.*

Proof. Set $V = T_x(M)$ so that $L_k\colon V \to V$ for each k. Then $L_k^2\colon V \to V$ and $L_k \otimes L_k\colon V \otimes V \to V \otimes V$. Since each of the L_k is diagonalizable, it is easily verified that

$$\Lambda = 2n \sum_{1\le k\le n} \operatorname{Trace}(L_k^2) - 2 \sum_{1\le k\le n} \operatorname{Trace}(L_k \otimes L_k).$$

Replacing the basis $\{e_i\}$ by the basis $\{\bar{e}_i = \sum_j a_{ij} e_j\}$, where (a_{ij}) is an orthogonal matrix, will replace the corresponding Weingarten maps $\{L_k\}$ by $\{\bar{L}_i = \sum_j a_{ij} \bar{L}_j\}$. Then

[7] I.e., a basis whose elements are orthogonal and of equal lengths.

$$\begin{aligned}\sum_{1\le k\le n} \operatorname{Trace}(\bar{L}_k^2) &= \sum_{1\le i,j,k\le n} a_{ij}a_{ij} \operatorname{Trace}(L_j L_k)\\ &= \sum_{1\le i,j,k\le n} \delta_{jk} \operatorname{Trace}(L_j L_k) = \sum_{1\le k\le n} \operatorname{Trace}(L_k^2),\end{aligned}$$

and similarly,

$$\begin{aligned}\sum_{1\le k\le n} \operatorname{Trace}(\bar{L}_k \otimes \bar{L}_k) &= \sum_{1\le k,p,q\le n} a_{kp}a_{kq} \operatorname{Trace}(L_p \otimes L_q)\\ &= \sum_{1\le p,q\le n} \delta_{pq} \operatorname{Trace}(L_p \otimes L_q)\\ &= \sum_{1\le k\le n} \operatorname{Trace}(L_k \otimes L_k). \qquad \blacksquare\end{aligned}$$

Corollary 4.36. *$x \in M$ is umbilic $\Leftrightarrow \Lambda(x) = 0$.*

Proof. First note that if L_X is a multiple of the identity for each X in a basis of the normal space ν_x, then by Lemma 4.25, L_X is a multiple of the identity for every $X \in \nu_x$. Thus, we have

$$\begin{aligned}\Lambda(x) = 0 &\Leftrightarrow \lambda_{ik} = \lambda_{jk} \quad \text{for } 1 \le i,j,k \le r\\ &\Leftrightarrow L_k \text{ is a multiple of the identity for } 1 \le k \le r\\ &\Leftrightarrow x \text{ is umbilic.} \qquad \blacksquare\end{aligned}$$

Now let us reinterpret Λ as a function on P_λ. For this we fix, once and for all, an equinormal basis $\{\bar{e}_s\}$ for u. Then for each $p \in P_\lambda$, $\{\varphi_p^{-1}(\bar{e}_s)\}$ is an equinormal basis for ν_x, and we may write $\Lambda(p)$ for Λ with respect to this basis. The effect of replacing p by ph, where

$$h = \begin{pmatrix} \mu & 0 & 0 & 0\\ 0 & a & 0 & 0\\ 0 & 0 & d & 0\\ 0 & 0 & 0 & \mu^{-1}\end{pmatrix}\begin{pmatrix} 1 & p & 0 & s\\ 0 & 1 & 0 & p^t\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \in H_\lambda, \quad \text{where } s = \frac{1}{2} p \cdot p^t,$$

is to replace $\varphi_p^{-1}(\bar{e}_s)$ by

$$\begin{aligned}\varphi_{ph}^{-1}(\bar{e}_s) &= (\operatorname{Ad}(h^{-1})\varphi_p)^{-1}(\bar{e}_s) = \varphi_p^{-1}(\operatorname{Ad}(h)\bar{e}_s)\\ &= \varphi_p^{-1}\left(\mu^{-1}\sum_t d_{st}\bar{e}_t\right) = \mu^{-1}\sum_t d_{st}\varphi_p^{-1}(\bar{e}_t).\end{aligned}$$

The orthonormal change d of the basis $\{\varphi_p(e_s)\}$ has no effect on Λ, while the scaling μ^{-1} multiplies Λ by μ^{-1}. Thus, up to a universal choice (the choice of the equinormal basis $\{\bar{e}_s\}$ for u), Λ determines a well-defined function on P transforming according to $\Lambda(ph) = \mu^{-1}\Lambda(p)$.

In the case where M has no umbilics, so that Λ is never zero, we obtain a further reduction of the bundle P_λ by setting

$$P_\mu = \{p \in P_\lambda \mid \Lambda(p) = 1\}.$$

Exercise 4.37. Suppose that $M \subset N \subset Q$ are three conformal manifolds with M and N receiving their geometries from Q. Assume that M is totally umbilic in N and N is totally umbilic in Q. Show that M is totally umbilic in Q. ❑

§5. Immersed Curves

Let us assume we are given an abstract locally ambient geometry on a curve. We study this geometry in the case of a curve of codimension one or two that is *without vertices*. These hypotheses allow a reduction of the locally ambient geometry to a principal bundle with discrete fiber. In this case the locally ambient geometry on M may be fully described by certain canonical functions and forms on M itself.[8]

Definition 5.1. Let M be a connected manifold of dimension one, and let (P, ω) be a locally ambient Möbius geometry of codimension r on M. A point $x \in M$ is a *vertex* for this geometry if the block ν of ω vanishes (see Definition 4.8). ❀

Recall that by Lemma 4.14 an immersion of a submanifold M into a Möbius geometry determines a locally ambient geometry on M. It follows that Definition 5.1 defines the notion of a vertex for any immersion of a one-dimensional manifold into a Möbius geometry.

Exercise 5.2. Show that for curves in the plane (regarded as a Möbius geometry) a vertex is just an inflection point. ❑

Proposition 5.3. *Let M be a connected manifold of dimension one, and let (P, ω) be a vertex-free locally ambient Möbius geometry of codimension one on M. Then (P, ω) is determined (up to equivalence) by a canonical function $\kappa_1 \colon M \to \mathbf{R}$ and a canonical (up to sign) 1-form $\pm\theta \in A^1(M)$.*

Proof. We refer to Definition 4.8, where the locally ambient geometry has the connection form

[8] Actually, it would be more proper to say that these forms and functions take their values in certain flat line bundles over M according to the principles described in footnote 12 of Chapter 1. However, the situation is so elementary here that we confine ourselves to the brief Remark 5.4.

$$\omega = \begin{pmatrix} \varepsilon & v & \nu & 0 \\ \theta & 0 & 0 & v \\ 0 & 0 & 0 & \nu \\ 0 & \theta & 0 & -\varepsilon \end{pmatrix},$$

in which the blocks α and γ vanish because they are 1×1 skew-symmetric matrices and the block β vanishes because its ith row $\beta(i)$ is $\theta h(i)$ and $h(i)$ is a 1×1 matrix of trace zero (cf. 4.8(d)). Since ν is semibasic, we have $\nu = f\theta$ for some function f on P. The vertex-free condition means that ν never vanishes and so neither does f.

We are going to find two successive reductions of the bundle P. These reductions will be defined by $\nu = \theta$ and $\{\nu = \theta,\ \varepsilon = 0\}$, respectively. The initial group (i.e., of P) is

$$H_\lambda = \mathrm{M\ddot{o}b}_1(\mathbf{R})_0 \times \{\pm 1\}.$$

The First Reduction. ω transforms according to $R_g^*\omega = \mathrm{Ad}(g^{-1})\omega$. Let us write $g \in H_\lambda$ as

$$g = \begin{pmatrix} \mu & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 \\ 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & \mu^{-1} \end{pmatrix} \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$s = \frac{1}{2}p^2, \quad a_1 = \pm 1, \quad a_2 = \pm 1, \quad \mu > 0.$$

Then ν transforms according to $R_g^*\nu = a_2\mu^{-1}\nu$ and θ transforms according to $R_g^*\theta = a_1\mu\theta$. Thus,

$$\begin{aligned} \nu = f\theta &\Rightarrow R_g^*\nu = R_g^* f R_g^*\theta \\ &\Rightarrow a_2\mu^{-1}\nu = (R_g^* f)a_1\mu\theta \\ &\Rightarrow a_2\mu^{-1} f\theta = (R_g^* f)a_1\mu\theta \\ &\Rightarrow R_g^* f = a_1 a_2 \mu^{-2} f. \end{aligned}$$

Hence f assumes the value 1 on every fiber. It follows that we may reduce the bundle to a subbundle $P_1 \subset P$ defined by $f = 1$ (or $\nu = \theta$) which has group

$$\left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \;\middle|\; s = \frac{1}{2}p^2, a = \pm 1 \right\}.$$

On this reduction the form ε is semibasic, so we may write $\varepsilon = f\theta$ for some (new) function on the reduced bundle.

The Second Reduction. The final reduction depends how ε transforms under the element

$$g = \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p^t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad s = \frac{1}{2}p^2.$$

A simple calculation shows that $R_g^*\varepsilon = \varepsilon - p\theta$. By choosing $p = f$, we see that there is a point on each fiber where $\varepsilon = 0$. Thus, there is a final reduction $P_2 \subset P_1 \subset P$ via the equation $\varepsilon = 0$ to the discrete group $\mathbf{Z}_2$ generated by $w = \operatorname{diag}(1, -1, -1, 1)$. On this reduction, the form restricts to

$$\omega = \begin{pmatrix} 0 & v & \theta & 0 \\ \theta & 0 & 0 & v \\ 0 & 0 & 0 & \theta \\ 0 & \theta & 0 & 0 \end{pmatrix},$$

and this time v is semibasic. Thus, we may write $v = \kappa_1\theta$ for some function κ_1 on P_2. Now an easy calculation shows that v and θ change sign under the action of w. Thus, κ_1 is constant on the fibers of P_2 and hence is basic. On the other hand, θ is defined up to sign on M, so we may express the induced form on M as $\pm\theta$. Clearly, this function κ_1 and form $\pm\theta$ constitute a complete set of invariants for the immersion. ■

Remark 5.4. Note that prescribing an orientation for M would allow a canonical choice among the forms $\pm\theta$ (if $e_1 \in T_x(M)$ is positively oriented, choose the sign so that $\theta(e_1) > 0$). We remark that $\theta \in A^1(P_2, \mathbf{R})$ may also be regarded as an element $\theta \in A^1(M, \zeta)$, where ζ is the (trivial) line bundle associated to P_2 by the canonical representation of $\mathbf{Z}_2 \to O_1(\mathbf{R})$ sending $w \mapsto -1$, where $w = \operatorname{diag}(1, -1, -1, 1, 1)$ as above. Similar remarks apply to the θ and κ_2 appearing in Theorem 5.6.

Definition 5.5. $\pm\theta$ is called the *conformal arclength* of M, and κ_1 is called the *conformal curvature* of M. ✿

It may be shown (cf., e.g., [G. Cairns and R. Sharpe, 1990]) that for a curve in the (x, y)-plane with Taylor expansion at the origin of the form $y = \frac{1}{2}x^2 + \frac{1}{24}\kappa_1 x^4 + O(x^5)$, then, at the origin, $\theta = dx$ and κ_1 is the conformal curvature.

Now we turn to the case of a vertex-free curve in an ambient geometry of codimension two.

Theorem 5.6. *Let M be a connected manifold of dimension one, and let (P, ω) be a vertex-free, locally ambient Möbius geometry of codimension two on M. Then (P, ω) is determined (up to equivalence) by a canonical function $\kappa_1 \colon M \to \mathbf{R}$, a canonical (up to sign) function $\pm\kappa_2 \colon M \to \mathbf{R}$, and a canonical (up to sign) 1-form $\pm\theta \in A^1(M)$.*

Proof. Referring to Definition 4.8, the locally ambient geometry has the form

$$\omega = \begin{pmatrix} \varepsilon & v & \nu & 0 \\ \theta & 0 & 0 & v \\ 0 & 0 & \gamma & \nu \\ 0 & \theta & 0 & -\varepsilon \end{pmatrix},$$

where the block α vanishes because it is a 1×1 skew-symmetric matrix, and the block β vanishes because its ith row $\beta(i)$ is $\theta h(i)$ and $h(i)$ is a 1×1 matrix of trace zero. Since ν is semibasic, we may write $\nu = f\theta$ for some function $f \colon P \to \mathbf{R}^2$. Since ν never vanishes (cf. Definition 5.1), neither does the function f.

We are going to find two successive reductions of the bundle P defined by $\{\nu = \theta e_1\}$ and $\{\nu = \theta e_1,\ \varepsilon = 0\}$, respectively. The initial group of P is $H_\lambda = \text{Möb}_1(\mathbf{R})_0 \times O_2(\mathbf{R})$.

The First Reduction. ω transforms according to $R_g^*\omega = \operatorname{Ad}(g^{-1})\omega$. Let us write

$$g = \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix} \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $z > 0$, $a_1 \in \{\pm 1\}$, $c \in O_2(\mathbf{R})$, $s = \frac{1}{2}p^2$. Then ν and θ transform according to $R_g^*\nu = z^{-1}\nu c$ and $R_g^*\theta = a_1 z\theta$. Thus,

$$\begin{aligned} \nu = f\theta \Rightarrow R_g^*\nu &= R_g^* f R_g^*\theta \\ \Rightarrow z^{-1}\nu c &= a_1(R_g^* f)z\theta \\ \Rightarrow z^{-1} f c\theta &= a_1(R_g^* f)z\theta \\ \Rightarrow R_g^* f &= a_1 z^{-2} f c. \end{aligned}$$

Since f never vanishes, it follows that it assumes the value $e_i \in \mathbf{R}^2$ on every fiber. Thus, we may reduce the bundle to a subbundle $P_1 \subset P$ defined by $f = e_1$ (or $\nu = \theta e_1$), which has group

$$\left\{ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & a_1 & 0 & 0 \\ 0 & 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & p & 0 & 0 & s \\ 0 & 1 & 0 & 0 & p \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \;\middle|\; a_1, a_2 \in \{\pm 1\},\ s = \frac{1}{2}p^2 \right\}.$$

On this reduction, the form ω restricts to

$$\omega = \begin{pmatrix} \varepsilon & v & \theta & 0 & 0 \\ \theta & 0 & 0 & 0 & v \\ 0 & 0 & 0 & -\gamma & \theta \\ 0 & 0 & \gamma & 0 & 0 \\ 0 & \theta & 0 & 0 & -\varepsilon \end{pmatrix},$$

in which, except for v, all the forms are semibasic. In particular, we may write $\varepsilon = f\theta$ for some (new) function f on the reduced bundle.

The Second Reduction. The final reduction depends on how ε transforms under the element

$$g = \begin{pmatrix} 1 & p & 0 & s \\ 0 & 1 & 0 & p \\ 0 & 0 & I_2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad s = \frac{1}{2}p^2.$$

A simple calculation shows that $R_g^*\varepsilon = \varepsilon - p\theta$. By choosing $p = f$, we see that there is a point on each fiber where $\varepsilon = 0$. Thus, there is a final reduction $P_2 \subset P_1 \subset P$ via the equation $\varepsilon = 0$ to the group

$$\{\pm 1\} \times \{\pm 1\} = \left\{ g = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & a_1 & 0 & 0 \\ 0 & 0 & 0 & a_2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mid a_1, a_2 \in \{\pm 1\} \right\}. \tag{5.7}$$

On this reduction, the form restricts to

$$\omega = \begin{pmatrix} 0 & v & \theta & 0 & 0 \\ \theta & 0 & 0 & 0 & v \\ 0 & 0 & 0 & -\gamma & \theta \\ 0 & 0 & \gamma & 0 & 0 \\ 0 & \theta & 0 & 0 & 0 \end{pmatrix},$$

and this time every entry is semibasic. Thus, we may write $v = \kappa_1\theta$ and $\gamma = \kappa_2\theta$ for some functions $\kappa_1, \kappa_2 \colon P_2 \to \mathbf{R}$. An easy calculation shows that for g as in Eq. (5.7) we have

$$R_g^*\theta = a_1\theta, \quad R_g^*v = a_1 v, \quad R_g^*\gamma = a_1 a_2 \gamma,$$

so that $R_g^*\kappa_1 = \kappa_1$, $R_g^*\kappa_2 = a_2\kappa_2$. Thus, we may regard κ_1 as a function on M, κ_2 as a function on M defined up to sign, and θ as a 1-form on M defined up to sign. The functions κ_1, κ_2 and the form θ obviously constitute a complete set of invariants for (P, ω). ∎

It may be shown (cf., e.g., [G. Cairns, R. Sharpe, and L. Webb, 1994]) that, up to conformal transformation, every vertex-free curve in 3 space has a Taylor expansion at the origin of the form

$$\begin{aligned} y &= \frac{1}{3!}x^3 + \frac{1}{5!}(2\kappa_1 - \kappa_2^2)x^5 + O(x^6), \\ z &= \frac{1}{4!}\kappa_2 x^4 + O(x^5) \end{aligned}$$

and that, in this coordinate system, $\theta = dx$ at the origin and κ_1, κ_2 are the values at the origin of the corresponding functions described in Theorem 5.6.

Exercise 5.8. Generalize Theorems 5.3 and 5.6 to the case of arbitrary codimension. ❑

§6. Immersed Surfaces

For a surface M, the classification of immersions given in §4 fails in that a normal locally ambient geometry on M cannot be reconstructed from the intrinsic geometry on M, the Ehresmann connection on the normal bundle, and the second fundamental form (cf. Theorem 4.29).

In this section we provide, for surfaces in $\mathbf{R}^3$, a partial alternative of the theory of immersions given in §4, which is analogous to our study for curves in §5. We study in some detail the case of a flat locally ambient geometry on a surface under the restriction that the geometry is without umbilic points and both the ambient and induced geometries are oriented. The absence of umbilic points allows us to reduce the principal bundle of the local ambient geometry to a bundle with discrete fiber. This principal bundle may be described as the bundle of "double orientations" of M, where a *double orientation* is an orientation for each of the two principal subspaces which are compatible with the orientation of the surface at that point.[9] The local ambient geometry on M may then be fully described by certain canonical functions and forms on M itself. All but one of these forms and functions have values in a certain real flat line bundle L over the surface. L is the line bundle described in footnote 12 of Chapter 1.

Proposition 6.1. *Let M be a connected, topologically oriented manifold of dimension two, and let (P, ω) be an umbilic-free, topologically oriented, flat, locally ambient Möbius geometry of codimension one on M (e.g., one arising from an immersion $M \to S^3$). Let us write the form as*

$$\omega = \begin{pmatrix} \varepsilon & v_1 & v_2 & \nu & 0 \\ \theta_1 & 0 & -\alpha & \beta_1 & v_1 \\ \theta_2 & \alpha & 0 & \beta_2 & v_2 \\ 0 & \beta_1 & \beta_2 & 0 & \nu \\ 0 & \theta_1 & \theta_2 & 0 & -\varepsilon \end{pmatrix}.$$

Then, canonically associated to (P, ω) are

(i) *a two-fold cover $\tilde{M} \to M$ with associated flat line bundle L over M,*

(ii) *a function $\Psi \colon M \to \mathbf{R}$,*

(iii) *sections $p_1, p_2 \in A^0(M, L)$ $(= A^0_{\mathbf{Z}_2}(\tilde{M}, (\mathbf{R}, -)))$, where $(\mathbf{R}, -)$ denotes the nontrivial representation of $\mathbf{Z}_2$,*

(iv) *L-valued 1-forms $\theta_1, \theta_2 \in A^1(M, L)(= A^1_{\mathbf{Z}_2}(\tilde{M}, (\mathbf{R}, -)))$.*

Let ξ_1, ξ_2 be the vector fields dual to θ_1, θ_2. Then these data satisfy the relation

[9] Since the principal directions have distinct eigenvalues, they may be canonically ordered, so "compatible" makes sense here.

(v) $d\tilde{\omega} + \tilde{\omega} \wedge \tilde{\omega} = 0$, *where*

$$\tilde{\omega} = \begin{pmatrix} \frac{1}{2}p_1\theta_1 - \frac{1}{2}p_2\theta_2 & \star & \star & \star & 0 \\ \theta_1 & 0 & 0 & \star & -\frac{1}{2}((1+\Psi)\theta_1 + (\xi_1 p_2 + p_1 p_2)\theta_2) \\ \theta_2 & 0 & 0 & \star & \frac{1}{2}((\xi_2 p_1 - p_1 p_2)\theta_1 - (1-\Psi)\theta_2) \\ 0 & \theta_1 & -\theta_2 & 0 & -\frac{1}{2}(p_1\theta_1 + p_2\theta_2) \\ 0 & \star & \star & 0 & \star \end{pmatrix}$$

$$\in A^1_{\mathbf{Z}_2}(\tilde{M}, \mathfrak{g}),$$

where $\mathbf{Z}_2$ *acts by the covering transformation on* $\tilde{M}$, *and by* $Ad(\tau)$, $\tau = \operatorname{diag}(1,-1,-1,1,1)$ *on* $\mathfrak{g}$. *(As usual, the entries denoted by stars are determined by symmetry.)*

Moreover, given data as in (i), (ii), (iii) *and* (iv) *satisfying relation* (v), *there is, up to equivalence, a unique locally ambient geometry giving rise to it.*

Proof. As usual, we let H_λ be the group of the locally ambient Möbius geometry. We claim the absence of umbilic points means that it is possible to obtain a reduction of the principle bundle P. This reduction is given by

$$P_\rho = \left\{ p \in P_\lambda \mid h(p) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \alpha = 0 \right\}, \quad \text{where } \beta = \theta^t h.$$

Let us see that this reduction exists. According to Definition 4.8(b), we have $R_k^*\omega = \operatorname{Ad}(k^{-1})\omega$ for all $k \in H_\lambda$, so calculating as in Lemma 4.6, we get $R_k^* h = (z^{-1}cI)a^{-1}ha$, where

$$k = \begin{pmatrix} 1 & u & 0 & r \\ 0 & I_2 & 0 & u^t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & z^{-1} \end{pmatrix},$$

where $r = \frac{1}{2}uu^t$, $z > 0$, $a \in O_2(\mathbf{R})$, $c = \pm 1$. The absence of umbilic points means that h is a nonsingular matrix. Of course, h is also symmetric of trace zero. Thus, we may choose a rotation a and a scaling $z^{-1}c$ such that $(z^{-1}cI)a^{-1}ha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, proving the existence of the reduction.

Upon restriction to P_ρ, the form ω becomes

$$\omega = \begin{pmatrix} \varepsilon & v_1 & v_2 & \nu & 0 \\ \theta_1 & 0 & -\alpha & -\theta_1 & v_1 \\ \theta_2 & \alpha & 0 & \theta_2 & v_2 \\ 0 & \theta_1 & -\theta_2 & 0 & \nu \\ 0 & \theta_1 & \theta_2 & 0 & -\varepsilon \end{pmatrix}.$$

A simple calculation shows that the stabilizer of $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ in H_λ is the subgroup of H_ρ generated by elements of the form

$$g_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad g_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$u = \begin{pmatrix} 1 & u_1 & u_2 & 0 & \frac{1}{2}(u_1^2 + u_2^2) \\ 0 & 1 & 0 & 0 & u_1 \\ 0 & 0 & 1 & 0 & u_2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{6.2}$$

so that P_ρ is a principal H_ρ bundle. The restrictions of the forms ε and α to P_ρ vanish on the fibers, and thus they are semibasic. Another calculation shows that

$$\operatorname{Ad}(u)\omega = \begin{pmatrix} \star & \star & \star & \star & \star \\ \star & \star & \star & \star & \star \\ \star & \alpha - u_2\theta_1 + u_1\theta_2 & \star & \star & \star \\ \star & \star & \star & \star & \star \\ \star & \star & \star & \star & \star \end{pmatrix}.$$

It follows that on each fiber of P_ρ there is a point where $\alpha = 0$, so we get a further reduction to the finite group $\mathbf{Z}_2 \oplus \mathbf{Z}_4$, generated by g_1 and g_2, by setting

$$\hat{M} = \left\{ p \in P_\rho \mid h(p) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \alpha = 0 \right\}.$$

It follows that $\hat{M} \to M$ is an eight-sheeted principal covering space. Upon restriction to P_ρ, the form ω becomes

$$\omega = \begin{pmatrix} \varepsilon & v_1 & v_2 & \nu & 0 \\ \theta_1 & 0 & 0 & -\theta_1 & v_1 \\ \theta_2 & 0 & 0 & \theta_2 & v_2 \\ 0 & \theta_1 & -\theta_2 & 0 & \nu \\ 0 & \theta_1 & \theta_2 & 0 & -\varepsilon \end{pmatrix}.$$

Because $\hat{M} \to M$ has discrete fiber, all of the entries in this matrix restrict to zero on the fiber and hence are semibasic. Moreover, the structural equation implies certain relations among these forms, which we now determine.

The $(2,3)$ component of the structural equation reads

$$\theta_1 \wedge v_2 + \theta_1 \wedge \theta_2 + v_1 \wedge \theta_2 = 0,$$

so that if we express v_1 and v_2 in terms of the basis θ_1 and θ_2, they must have the form

$$v_1 = -\frac{1}{2}(1+\Psi)\theta_1 - \frac{1}{2}r_1\theta_2,$$
$$v_2 = \frac{1}{2}q_2\theta_1 - \frac{1}{2}(1-\Psi)\theta_2$$

for some functions r_1, q_2, and Ψ on $\tilde{M}$.

On the other hand, the structural equation may be used to compute $d\theta_i$ in two different ways (from the first column and from the fourth row). We get

$$d\theta_1 = -\theta_1 \wedge \varepsilon = \theta_1 \wedge \nu,$$
$$d\theta_2 = -\theta_2 \wedge \varepsilon = -\theta_2 \wedge \nu,$$

so that if we write $\varepsilon = \frac{1}{2}p_1\theta_1 - \frac{1}{2}p_2\theta_2$ for some functions p_1, p_2 on $\hat{M}$, then we must have

$$\nu = \frac{1}{2}p_1\theta_1 + \frac{1}{2}p_2\theta_2, \quad d\theta_1 = \frac{1}{2}p_2\theta_1 \wedge \theta_2, \quad \text{and} \quad d\theta_2 = \frac{1}{2}p_1\theta_1 \wedge \theta_2.$$

Putting these expressions in the $(1,1)$ block of the structural equation, we calculate

$$d\varepsilon = -v_1 \wedge \theta_1 - v_2 \wedge \theta_2$$

or

$$\begin{aligned} d\left(\frac{1}{2}p_1\theta_1 - \frac{1}{2}p_2\theta_2\right) &= \frac{1}{2}((1+\Psi)\theta_1 + r_1\theta_2) \wedge \theta_1 - \frac{1}{2}(q_2\theta_1 - (1-\Psi)\theta_2) \wedge \theta_2 \\ &= -\frac{1}{2}r_1\theta_1 \wedge \theta_2 - \frac{1}{2}q_2\theta_1 \wedge \theta_2. \end{aligned} \tag{6.3}$$

Substitution in the $(1,4)$ entry of the structural equation, which is

$$d\nu = -\varepsilon \wedge \nu + v_1 \wedge \theta_1 - v_2 \wedge \theta_2,$$

yields

$$\begin{aligned} d\left(\frac{1}{2}p_1\theta_1 + \frac{1}{2}p_2\theta_2\right) &= -\left(\frac{1}{2}p_1\theta_1 - \frac{1}{2}p_2\theta_2\right) \wedge \left(\frac{1}{2}p_1\theta_1 + \frac{1}{2}p_2\theta_2\right) \\ &\quad -\frac{1}{2}((1+\Psi)\theta_1 + r_1\theta_2) \wedge \theta_1 \\ &\quad -\frac{1}{2}(q_2\theta_1 - (1-\Psi)\theta_2) \wedge \theta_2 \\ &= -\frac{1}{2}p_1p_2\theta_1 \wedge \theta_2 + \frac{1}{2}r_1\theta_1 \wedge \theta_2 - \frac{1}{2}q_2\theta_1 \wedge \theta_2. \end{aligned} \tag{6.4}$$

Adding and subtracting Eqs. (6.3) and (6.4) yields

$$d(p_1 \wedge \theta_1) = -\left(q_2 + \frac{1}{2}p_1p_2\right)\theta_1 \wedge \theta_2$$

and

$$d(p_2 \wedge \theta_2) = \left(r_1 - \frac{1}{2}p_1p_2\right)\theta_1 \wedge \theta_2,$$

which in turn imply that

$$dp_1 \wedge \theta_1 + p_1 d\theta_1 = -\left(q_2 + \frac{1}{2}p_1p_2\right)\theta_1 \wedge \theta_2$$

and

$$dp_2 \wedge \theta_2 + p_2 d\theta_2 = \left(r_1 - \frac{1}{2}p_1p_2\right)\theta_1 \wedge \theta_2,$$

or

$$dp_1 \wedge q_1\theta_1 + (q_2 + p_1p_2)\theta_2 \quad \text{and} \quad dp_2 \wedge \theta_2 = (r_1 - p_1p_2)\theta_1 \wedge \theta_2.$$

Thus, we may write

$$dp_1 = q_1\theta_1 + (q_2 + p_1p_2)\theta_2 \quad \text{and} \quad dp_2 = (r_1 - p_1p_2)\theta_1 + r_2\theta_2$$

for some functions q_1 and r_2. Note that from this equation it follows that, if ξ_1 and ξ_2 are the vector fields dual to θ_1 and θ_2, then

$$\xi_1 p_1 = q_1, \quad \xi_2 p_1 = q_2 + p_1p_2, \quad \xi_1 p_2 = r_1 - p_1p_2 \quad \text{and} \quad \xi_2 p_2 = r_2.$$

From the equation $R_g^*\omega = \mathrm{Ad}(g^{-1})\omega$, we obtain the following table showing how the various items transform.

η	θ_1	θ_2	p_1	p_2	Ψ
$R_{g_1}\eta$	$-\theta_1$	θ_2	$-p_1$	p_2	Ψ
$R_{g_2}\eta$	$-\theta_2$	θ_1	p_2	$-p_1$	$-\Psi$

It follows that the 2-form $\theta_1 \wedge \theta_2$ is invariant under g_2 but changes sign under g_1. If there were a path on $\hat{M}$ joining $p \in \hat{M}$ to pg_1, this path would cover an orientation-reversing loop on M. By hypothesis, no such loop exists. Thus, $\hat{M} = M' \cup M'g_1$ for some submanifold (a union of components) of $M' \subset \hat{M}$, and $M' \to M$ is a four-fold cover (a principal $\mathbf{Z}_4$ bundle with group generated by g_2). If there were a path on M' joining $p \in M'$ to pg_2, this path would cover a loop on M that would be orientation reversing for the locally ambient geometry. Again by hypothesis, no such loop exists, so $M' = \tilde{M} \cup \tilde{M} g_2$ for some submanifold (a union of components) $\tilde{M} \subset M'$ and $\tilde{M} \to M$ is a two-fold cover (a principal $\mathbf{Z}_2$ bundle with group generated by $\tau = (g_2)^2$). From the table above we have the following:

η	θ_1	θ_2	p_1	p_2	Ψ
$R_\tau\eta$	$-\theta_1$	$-\theta_2$	$-p_1$	$-p_2$	Ψ

Let L be the real line bundle over M associated to $\tilde{M}$ via the representation $\mathbf{Z}_2 = \{\pm 1\} \subset Gl_1(\mathbf{R})$. It follows that $p_1, p_2 \in A^0(M, L)$, $\theta_1, \theta_2 \in A^1(M, L)$, $\Psi \in A^0(M, \mathbf{R})$ and that these items fit together to yield the form $\tilde{\omega}$ as described in (v). Conversely, given the data (i), (ii), (iii), and (iv) satisfying (v), we can construct the principal bundle $\mathbf{Z}_2 \to \tilde{M} \to M$ associated to L

and interpret Ψ, p_1, p_2 as functions on it and θ_1, θ_2 as forms on it. These data then determine an element $\tilde{\omega} \in A^1_{\mathbf{Z}_2}(\tilde{M}, \mathfrak{g})$ by the formula in (v), which in turn determines an umbilic-free, oriented, flat, locally ambient Möbius geometry of codimension one on M. We leave the details of this to the reader. ■

The three functions p_1, p_2, and Ψ given in this result are related to the Monge form of the equation for the surface. In [G. Cairns, R. Sharpe, and L. Webb, 1994] it is shown that for a surface in S^3 with invariants p_1, p_2, Ψ at some given point, there is an orientation-preserving Möbius transformation $g \in \text{Möb}_3(\mathbf{R})$ throwing the point to the south pole such that standard stereographic projection of the 3-sphere into space sends the image surface to a surface in space given locally by an equation of the form

$$\begin{aligned} z = &\frac{1}{2}(x^2 - y^2) + \frac{1}{6}(p_1 x^3 + p_2 y^3) \\ &+ \frac{1}{24}\{(3 + p_1^2 + \xi_1 p_1)x^4 + 4(\xi_2 p_1 - p_1 p_2)x^3 y + 6\Psi x^2 y^2 \\ &+ 4(\xi_1 p_2 + p_1 p_2)xy^3 + (-3 - p_2^2 + \xi_2 p_2)y^4\} + O(5). \end{aligned} \tag{6.5}$$

Moreover, the forms θ_1 and θ_2 agree with dx and dy at the origin. There are exactly two such elements $g \in \text{Möb}_3(\mathbf{R})$ that will do this job, and they differ by left multiplication with $\tau = \text{diag}(1, -1, -1, 1, 1)$. The effect of τ on the equation is to change the signs of x and y, and this accounts for the indeterminacy of the signs of p_1, p_2, θ_1, θ_2. We may interpret the existence of g as a map $\gamma\colon M \to \{e, \tau\} \setminus G$, in which case the canonical map

$$\begin{aligned} G &\to S^3 \\ g &\mapsto g \cdot N \end{aligned}$$

is a left inverse for γ.

Exercise 6.6. For a torus of revolution in space with radii r and R, show that the invariants p_1 and p_2 vanish identically while $\Psi = \frac{2-(R/r)^2}{4}$. ❑

Exercise 6.7. For a closed oriented surface M with generic umbilic points, show that the monodromy of L along any closed path counts the number, mod 2, of umbilics enclosed in it. ❑

Exercise 6.8. Read enough of Fialkow's paper [A. Fialkow, 1944] to show that, for the case of surfaces in space, the tensors described in Fialkow's Theorem 29.2 may be described in terms of our invariants by

(i) $G_{ij}dx_i dx_j = \theta_1^2 + \theta_2^2$,

(ii) $B_{ij}dx_i dx_j = \theta_1^2 - \theta_2^2$,

(iii) $E_{ij}dx_i dx_j = (2 + \frac{1}{8}(3p_1^2 - p_2^2) - \frac{1}{2}((\xi_1 p_1) - \Psi))\theta_1^2 + ((\xi_1 p_2) - (\xi_2 p_1) + p_1 p_2)\theta_1\theta_2 + (2 - \frac{1}{8}(p_1^2 - 3p_2^2) - \frac{1}{2}(\Psi - (\xi_2 p_2)))\theta_2^2$.

In particular, show that in the case of a torus of revolution, the deviation form is

$$E_{ij}dx_i dx_j = \left(2 + \frac{1}{2}\Psi\right)\theta_1^2 + \left(2 - \frac{1}{2}\Psi\right)\theta_2^2.$$ ❑

Exercise 6.9. Verify that Fialkow's invariants do indeed determine our invariants in the case of a surface in space. ❑

8
Projective Geometry

In the elementary sense, to solve a differential equation means to find a change of coordinates in which one recognizes the equation as one of known type. To truly be able to carry this out means that one is somehow able to deal with the differential equation in all possible coordinate systems. One of Elie Cartan's ideas was to develop a theory of differential equations that was independent of the coordinate system used to express it. This idea of Cartan is very ambitious and includes both ordinary and partial differential equations within its scope. It can be especially useful in the study of "overdetermined" partial differential equations, which often arise in differential geometric settings.

In this context, the following kind of question arises. Given a system of differential equations on a manifold, can one "geometrize" the system in the sense of associating to it, independently of any coordinate system, a canonical Cartan geometry (i.e., a connection) from which alone one can recover the original system of differential equations? Generally speaking, the answer is no. There do exist, however, some remarkable cases of both ordinary and partial differential equations in which the answer is yes. The first example of this is Cartan's five-variables paper referred to in the preface. There he introduced the *method of equivalence* for solving this question. Many authors have continued to study Cartan's method of equivalence. For example, Tanaka (see [N. Tanaka, 1978] and the references therein) took up this line of work and addressed the general problem of attaching a Cartan geometry to a certain kind of k-dimensional distribution on an n-dimensional manifold. In favorable cases, this yields a Cartan geometry.

We will not study Cartan's general program here. For that we refer the reader to [R. Bryant, S.-S. Chern, R. Gardner, and H. Goldschmidt, 1991]. In this chapter, the theme we have been discussing is taken up in the particular context of the "geometrization" of a certain class of systems of second-order ordinary differential equations on a manifold. These systems include the equations for geodesics in a Riemannian manifold as examples. The Cartan geometry that solves the equivalence problem here is modeled on projective space, and the solutions of the equations are recovered as the geodesics of the projective geometry.

In §1 we establish the basic notation for the n-dimensional real projective model space $\mathbf{P}^n$ with fundamental group G, the group of (almost) all projective transformations (Definition 1.3). In §2 we introduce the basic terminology of projective Cartan geometries and define the special geometries, including the normal geometries. We also show the existence of a canonical projective gauge associated to a fixed coordinate system (Proposition 2.2). In §3 we define geodesics in a projective Cartan geometry (Definition 3.4). We show that the set $\mathcal{C}$ of all curves on a manifold M that satisfy an ordinary second-order differential equation of a certain type determines a unique normal projective connection on M with $\mathcal{C}$ as its set of geodesics (Theorem 3.8(iii)). This is Cartan's version of the *geometry of paths* of O. Veblen and T.Y. Thomas. In Exercise 3.3 we point out that this theory applies to the second-order ODE

$$y'' = A(x,y) + B(x,y)y' + C(x,y)(y')^2 + D(x,y)(y')^3,$$

an example of which Cartan was particularly proud. In §4 we show that the set of geodesics of a Riemannian geometry on M is also of this type, and we determine the corresponding normal projective geometry (Theorem 4.1). We then apply this to prove a result of Beltrami stating that a Riemannian manifold M whose geodesics are all straight lines in some coordinate system must be of constant curvature (Theorem 4.2). The material in the first four sections is loosely modeled on [E. Cartan, 1924]. Finally, in §5 we take a brief tour of some of the literature associated with projective connections. We mention here the recent publication of [M.A. Akivis and V.V. Goldberg, 1993] in which the reader will find further details about projective Cartan geometries.

§1. The Projective Model

Projective space $\mathbf{P}^n = \mathbf{P}^n(\mathbf{R})$ was defined in Definition **1**.1.3. Three points of $\mathbf{P}^n$ are said to be collinear if they lie on a single line (cf. Definition 1.11). One version of the fundamental theorem of projective geometry (cf. [P. Samuel, 1988], p. 18) says that, for $n > 2$, the group of diffeomorphisms of $\mathbf{P}^n$ that preserve collinearity is $PGl_{n+1}(\mathbf{R})$ with its standard action on

$\mathbf{P}^n$ (induced from the standard action of $Gl_{n+1}(\mathbf{R})$ on $\mathbf{R}^{n+1}$). From that point, of view it would make sense to take $PGl_{n+1}(\mathbf{R})$ as the fundamental group of the geometry. However, it is more convenient to work with the subgroup $G = PSl_{n+1}(\mathbf{R})$, which is not much different from $PGl_{n+1}(\mathbf{R})$, as the following exercise shows.

Exercise 1.1.*

(i) Show that there is a canonical inclusion $PSl_{n+1}(\mathbf{R}) \subset PGl_{n+1}(\mathbf{R})$.

(ii) Show that $PSl_{n+1}(\mathbf{R})$ is the identity component of $PGl_{n+1}(\mathbf{R})$.

(iii) Show that $PGl_{n+1}(\mathbf{R})$ is

$$\begin{cases} \text{connected} & \text{if } n+1 \text{ is odd} \\ \text{has two components} & \text{if } n+1 \text{ is even.} \end{cases}$$ ❑

Lemma 1.2. *Let $Z = \{\lambda I_{n+1} \in Sl_{n+1}(\mathbf{R}) \mid \lambda \in \mathbf{R}^*\}$ so that $PSl_{n+1}(\mathbf{R}) = Sl_{n+1}(\mathbf{R})/Z$. Then Z is the center of $Sl_{n+1}(\mathbf{R})$ and*

$$Z = \begin{cases} I & \text{if } n+1 \text{ is odd,} \\ \pm I & \text{if } n+1 \text{ is even.} \end{cases}$$

Proof. To calculate Z, we have

$$\lambda I_{n+1} \in Sl_{n+1}(\mathbf{R}) \Leftrightarrow \lambda^{n+1} = 1 \Leftrightarrow \lambda = \begin{cases} 1 & \text{if } n+1 \text{ is odd,} \\ \pm 1 & \text{if } n+1 \text{ is even.} \end{cases}$$

It is clear that Z lies in the center. Let us show that if g is central, it must lie in Z. If $g \in Sl_{n+1}(\mathbf{R})$ is central, it must commute with diag($\pm1, \pm1, \ldots, \pm1) \in Sl_{n+1}(\mathbf{R})$ for all possible choices of the sign; when $n \geq 2$, this forces g to be diagonal. When $n = 1$, the fact that g must commute with $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ forces it to be diagonal. Since the diagonal matrix g must also commute with all matrices of the form

$$\begin{pmatrix} I_s & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & I_{n-s-1} \end{pmatrix},$$

it follows that $g = \lambda I_{n+1}$ for some $\lambda \in \mathbf{R}$. ■

In working with elements of $G = PSl_{n+1}(\mathbf{R})$, it will often be convenient in this chapter to represent these elements as matrices. Such an expression is only defined mod Z, but we shall not always include this proviso even though it must be understood to be present.

Let $e_0, e_1, \ldots, e_n$ be the standard basis of $\mathbf{R}^{n+1}$. It is easily seen that the stabilizer in $G = PSl_{n+1}(\mathbf{R})$ of the point $[e_0] \in \mathbf{P}^n$ is given by

$$H = \left\{ h \in PSl_{n+1}(\mathbf{R}) \mid h = \begin{pmatrix} z & p \\ 0 & a \end{pmatrix} \bmod Z \right\},$$

where the diagonal blocks of the matrix have sizes 1×1 and $n \times n$, so it follows that the map $G \to \mathbf{P}^n$ sending $g \mapsto g[e_0]$ induces a diffeomorphism $G/H \approx \mathbf{P}^n$. As usual, we will generally work with G/H instead of $\mathbf{P}^n$.

Definition 1.3. The group $PSl_{n+1}(\mathbf{R}) = G$ is called the *projective group* in n dimensions. The pair (G, H), where H is described above, is called the *projective model* in n dimensions. ❀

Exercise 1.4. Show that the Lie algebras of G and H are given by

$$\mathfrak{g} = \{g \in M_{n+1}(\mathbf{R}) \mid \operatorname{Trace}(g) = 0\}, \quad \mathfrak{h} = \left\{ g \in \mathfrak{g} \mid h = \begin{pmatrix} 1 & n \\ z & p \\ 0 & a \end{pmatrix} \right\}.$$ ❑

Exercise 1.5.* Show that the following three subgroups of H are connected, are normal in H, and have the following Lie algebras:

$$\begin{aligned}
H_e &= \left\{ h \in PSl_{n+1}(\mathbf{R}) \mid h \in \begin{pmatrix} z & p \\ 0 & a \end{pmatrix} \bmod Z, z > 0 \right\}, \\
\mathfrak{h} &= \left\{ h \in \mathfrak{g} \mid h = \begin{pmatrix} z & p \\ 0 & a \end{pmatrix} \right\}; \\
H_1 &= \left\{ h \in PSl_{n+1}(\mathbf{R}) \mid h \in \begin{pmatrix} 1 & p \\ 0 & a \end{pmatrix} \bmod Z \right\}, \\
\mathfrak{h}_1 &= \left\{ h \in M_{n+1}(\mathbf{R}) \mid h = \begin{pmatrix} 0 & p \\ 0 & a \end{pmatrix}, \operatorname{Trace} a = 0 \right\}; \\
H_2 &= \left\{ h \in PSl_{n+1}(\mathbf{R}) \mid h \in \begin{pmatrix} 1 & p \\ 0 & I \end{pmatrix} \bmod Z \right\}, \\
\mathfrak{h}_2 &= \left\{ h \in M_{n+1}(\mathbf{R}) \mid h = \begin{pmatrix} 0 & p \\ 0 & 0 \end{pmatrix} \right\}.
\end{aligned}$$ ❑

To describe all the normal subgroups of H, we shall use the homomorphism

$$\phi \colon H \to \begin{cases} \mathbf{R}^* & \text{for } n+1 \text{ odd}, \\ \mathbf{R}^+ & \text{for } n+1 \text{ even}, \end{cases}$$

sending

$$\begin{pmatrix} z & p \\ 0 & a \end{pmatrix} \mapsto \begin{cases} 1 & \text{for } n+1 \text{ odd}, \\ |z| & \text{for } n+1 \text{ even}. \end{cases}$$

Lemma 1.6.

(i) *H_e, H_1, H_2, and $\{e\}$ are the only connected normal subgroups of H.*

(ii) *The other normal subgroups of H are given by $\phi^{-1}(A)$, where A is a nontrivial subgroup of*

$$\begin{cases} \mathbf{R}^* & \text{for } n+1 \text{ odd}, \\ \mathbf{R}^+ & \text{for } n+1 \text{ even}. \end{cases}$$

Proof. Let $N \subset H$ be a normal subgroup, $N \neq \{e\}$. The commutator relation

$$\left[\begin{pmatrix} z & p \\ 0 & a \end{pmatrix}, \begin{pmatrix} 1 & q \\ 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 1 & zqa^{-1} - q \\ 0 & 1 \end{pmatrix}$$

shows that N contains an element of the form $\begin{pmatrix} z & p \\ 0 & a \end{pmatrix}$ (mod Z) with $p \neq 0$. The commutator relation

$$\left[\begin{pmatrix} z & p \\ 0 & a \end{pmatrix}, \begin{pmatrix} y & 0 \\ 0 & I \end{pmatrix} \right] = \begin{pmatrix} 1 & (I - y)pa^{-1} \\ 0 & 1 \end{pmatrix}$$

shows that N contains an element of the form $\begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix}$ (mod Z) with $p \neq 0$. The commutator relation

$$\left[\begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \right] = \begin{pmatrix} 1 & p(I - b^{-1}) \\ 0 & 1 \end{pmatrix}$$

shows that $N \supset H_2$. If $N = H_2$ we are done, so let us assume that N contains an element $\begin{pmatrix} z & p \\ 0 & a \end{pmatrix}$ (mod Z) with $z \neq 1$ or $a \neq 1$. Since $N \subset Sl_{n+1}(\mathbf{R})/Z$, it follows that if $a = I$ then $z = 1$. Thus, we may assume that $a \neq I$. Since $N \supset H_2$, it follows that $\begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix}$ (mod Z) $\in N$. The commutator relation

$$\left[\begin{pmatrix} z & 0 \\ 0 & a \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 \\ 0 & [a, b] \end{pmatrix}$$

shows that $\begin{pmatrix} 1 & 0 \\ 0 & [a, b] \end{pmatrix}$ (mod Z) $\in N$ for all $b \in Sl_n(\mathbf{R})$. Since $PSl_n(\mathbf{R})$ is simple, it follows that $N \supset H_1$. Since H_1 is the kernel of ϕ, it follows that $N = \phi^{-1}(A)$, where

$$A = \phi(N) \subset \begin{cases} \mathbf{R}^* & \text{for } n+1 \text{ odd}, \\ \mathbf{R}^+ & \text{for } n+1 \text{ even}. \end{cases}$$

Now apply Exercise 1.5. ∎

Now we work toward the introduction of the normal submodule of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$.

Lemma 1.7. *There is a canonical isomorphism ad: $\mathfrak{h}/\mathfrak{h}_2 \approx End(\mathfrak{g}/\mathfrak{h})$.*

Proof. Consider the homomorphism ad: $\mathfrak{h} \to \mathrm{End}(\mathfrak{g}/\mathfrak{h})$. This is given explicitly by

$$\mathrm{ad}\begin{pmatrix} z & p \\ 0 & a \end{pmatrix}\begin{pmatrix} \star & \star \\ v & \star \end{pmatrix} = \begin{pmatrix} z & p \\ 0 & a \end{pmatrix}\begin{pmatrix} \star & \star \\ v & \star \end{pmatrix} - \begin{pmatrix} \star & \star \\ v & \star \end{pmatrix}\begin{pmatrix} z & p \\ 0 & a \end{pmatrix}$$
$$= \begin{pmatrix} \star & \star \\ (a - zI)v & \star \end{pmatrix},$$

so it induces a map ad: $\mathfrak{h}/\mathfrak{h}_2 \to \mathrm{End}(\mathfrak{g}/\mathfrak{h})$, which is easily seen to be an isomorphism. ■

For the next definitions and for later purposes, we will need the following composite H module map, which is the projective analog of the Ricci homomorphism (cf. Definitions **6**.1.3 and **7**.1.19) and which we again denote by *Ricci*.

$$\begin{aligned}
\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h}) &\xrightarrow{\text{projection}} \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h}/\mathfrak{h}_2) \\
&\approx \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathfrak{h}/\mathfrak{h}_2 \overset{\mathrm{id}\otimes\mathrm{ad}}{\approx} \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathrm{End}(\mathfrak{g}/\mathfrak{h}) \\
&\approx \lambda^2(\mathfrak{g}/\mathfrak{h})^* \otimes \mathfrak{g}/\mathfrak{h} \otimes (\mathfrak{g}/\mathfrak{h})^* \xrightarrow{\text{contraction}} (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^* \\
&\qquad t^* \wedge u^* \otimes v \otimes w^* \longmapsto (u^*(v)\, t^* - t^*(v)\, u^*) \otimes w^*
\end{aligned} \tag{1.8}$$

Definition 1.9. In projective geometry, the *normal* submodule of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is the kernel of the H module map $Ricci$: $\mathrm{Hom}(\lambda^2(\mathfrak{g}, \mathfrak{h}), \mathfrak{h}) \to (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$ given by (1.8). Similarly, the *symmetric* submodule of $\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ is the kernel of the composite H module map

$$\mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \xrightarrow{Ricci} (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^* \to \lambda^2(\mathfrak{g}/\mathfrak{h})^*.$$ ■

Exercise 1.10.* For $0 \le p, q \le n$, define

$$e_{pq} = \begin{cases} E_{pq} & \text{if } p \ne q, \\ E_{pp} - \frac{1}{n+1}\sum_{0\le r\le n} E_{rr} & \text{if } p = q, \end{cases}$$

where the $E_{pq} \in M_{n+1}(\mathbf{R})$ are the standard elementary matrices. Let $e_i \in \mathfrak{g}/\mathfrak{h}$, $1 \le i \le n$, be the standard basis (i.e., e_i is the class of E_{i0}).

(i) Verify that the e_{pq} (for $1 \le p, q \le n$, and $0 = p < q \le n$) constitute a basis for $\mathfrak{h}$.

(ii) Show that the element $e_{pq} \in \mathfrak{h}$, $1 \le p, q \le n$, corresponds to $e_p \otimes e_q^*$ under the composite homomorphism $\mathfrak{h} \to \mathfrak{h}/\mathfrak{h}_2 \to \mathrm{End}(\mathfrak{g}/\mathfrak{h}) \approx (\mathfrak{g}/\mathfrak{h}) \otimes (\mathfrak{g}/\mathfrak{h})^*$.

(iii) Show that

$$Ricci(e_i^* \wedge e_j^* \otimes e_{pq}) = \begin{cases} (e_i^*\delta_{jp} - e_j^*\delta_{ip}) \otimes e_q^* & \text{if } p > 0, \\ 0 & \text{if } p = 0. \end{cases}$$ ❑

Lines

Definition 1.11. A *line* in $\mathbf{P}^n$ is the set of all one-dimensional subspaces of a two-dimensional subspace of $\mathbf{R}^{n+1}$. Three points of $\mathbf{P}^n$ are said to be *collinear* if they lie on a single line. ❀

Note that the projective group acts transitively on the set of lines of $\mathbf{P}^n$. A line on $\mathbf{P}^n$ meets any affine coordinate chart of $\mathbf{P}^n$ in a line with respect to the affine coordinate system. In such a coordinate system a line may be parametrized in the usual way; such a parameter is called an affine parameter. These lines, unparametrized, will be the models for geodesics in the next section.

Let us study the generalized circles (cf. Definition **5**.4.16) on $\mathbf{P}^n$ with a view toward determining which generalized circles are straight lines.

Lemma 1.12. *Let $X \in \mathfrak{g}$ have block decomposition of the form $X = \begin{pmatrix} x & p \\ q & a \end{pmatrix}$. The generalized circle corresponding to the vector X is a straight line if and only if $q \ne 0$ and is an eigenvector for a.*

Proof. Recall that a generalized circle is the projection to $\mathbf{P}^n$ of the integral curve of an ω_G constant vector field V on G. We shall exclude the generalized circles that degenerate to points. Let $X = \omega_G(V) \in \mathfrak{g}$. Then the integral curves of V are the left translates of the one-parameter subgroup generated by X. This subgroup $\varphi: \mathbf{R} \to G$ is given by

$$\varphi(t) = \sum_{n=0}^{\infty} \frac{1}{n!} t^n X^n.$$

Since left translation by g on $\mathbf{P}^n$ moves generalized circles to generalized circles and straight lines to straight lines, the character of any generalized circle corresponding to X will be determined by that of the projection of $\varphi(t)$ to $\mathbf{P}^n$, which is the curve $\varphi(t)[e_0]$. If this curve on $\mathbf{P}^n$ lies on a line, then $\varphi(t)e_0$ lies on a plane, which implies that the derivatives $\varphi^{(m)}(0)e_0 = X^m e_0$, $m = 0, 1, \ldots$, span a space of dimension at most two. Consider

$$e_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad Xe_0 = \begin{pmatrix} z \\ q \end{pmatrix}, \quad X^2 e_0 = \begin{pmatrix} z^2 + pq \\ (zI + a)q \end{pmatrix}.$$

If the span of the first two of these vectors is one-dimensional, it follows that $q = 0$ and hence that $X^k e_0 = z^k e_0$, so that $\varphi(t)[e_0] = [\varphi(t)e_0] = [e^{tz}e_0] = [e_0]$. This is the degenerate case that we are excluding. On the other hand, if the span of the first three of these vectors is two-dimensional, then $q \neq 0$ and the vectors $(zI + a)q$ and q are linearly dependent. This means that q is an eigenvector for a and it implies, by an easy calculation, that the span of *all* the $X^m e_0$, $m = 0, 1, \ldots$, is also two-dimensional and hence that $\varphi(t)[e_0]$ lies on a line. ■

§2. Projective Cartan Geometries

In this section we study some elementary aspects of projective geometry. We briefly study the canonical gauge associated to a coordinate system and the special geometries including the normal geometries.

Definition 2.1. Let M be a smooth manifold. A *projective geometry on* M is a Cartan geometry on M modeled on the projective model. ❀

Let us write the connection and curvature forms as

$$\pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix},$$

$$\Pi = \begin{pmatrix} E & Y \\ \Theta & A \end{pmatrix} = d\pi + \pi \wedge \pi = \begin{pmatrix} d\varepsilon + v \wedge \theta & dv + \varepsilon \wedge v + v \wedge \alpha \\ d\theta + \theta \wedge \varepsilon + \alpha \wedge \theta & d\alpha + \theta \wedge v + \alpha \wedge \alpha \end{pmatrix}.$$

As usual, we will also denote by π and Π the corresponding connection and curvature forms in their various gauge incarnations.

Proposition 2.2. *Suppose that (U, x) is a coordinate system on a manifold M equipped with a projective Cartan geometry. Then this geometry has a unique gauge $(U, \pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix})$ such that $\varepsilon = 0$ and $\theta_i = dx_i$ for $i = 1, \ldots, n$.*[1]

Proof. It is enough to show that this is true on a sufficiently small neighborhood V of each point $x \in U$ since, by the uniqueness, all these gauges will fit together to give a gauge on U.

Since, by Lemma 1.7, the adjoint representation $\mathrm{ad}\colon \mathfrak{h} \to \mathrm{End}(\mathfrak{g}/\mathfrak{h})$ is a surjection in projective geometry, then by Exercise **5**.1.5, there is a neighborhood V of each point $x \in U$ and a gauge (V, π) such that $\theta_i = dx_i$ for

[1] This result explains the frequent occurrence of the condition $\varepsilon = 0$ for projective gauges in the literature, e.g. [E. Cartan, 1937] and [S.-S. Chern, 1937].

$i = 1, \ldots, n$. Now we make a change of gauge via $h\colon V \to H$ of the form $h = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$ so that

$$\psi \Rightarrow_h \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} + h^* \omega_H = \begin{pmatrix} \varepsilon - b\theta & \star \\ \theta & \star \end{pmatrix}.$$

Since the components of θ span the space of 1-forms on V, we may choose the function $b\colon V \to \mathbf{R}^n$ in a unique manner so as to kill the first component of the new gauge. This yields the existence of the required gauge on V.

Now we study the uniqueness. Note than an *arbitrary* change of gauge $h\colon V \to H$ will alter a gauge with first column $\begin{pmatrix} 0 \\ dx \end{pmatrix}$ as follows. Writing $h = \begin{pmatrix} a & b \\ 0 & A \end{pmatrix}$, we have

$$\begin{aligned} \mathrm{Ad}(h^{-1})\pi + h^* \omega_H &= \begin{pmatrix} a^{-1} & -a^{-1} b A^{-1} \\ 0 & A \end{pmatrix} \begin{pmatrix} 0 & \star \\ dx & \star \end{pmatrix} \begin{pmatrix} a & b \\ 0 & A \end{pmatrix} \\ &\quad + \begin{pmatrix} a^* w_H & \star \\ 0 & \star \end{pmatrix} \\ &= \begin{pmatrix} -a^{-1} b A^{-1} dx + a^* \omega_H & \star \\ a A^{-1} dx & \star \end{pmatrix}. \end{aligned}$$

For the new gauge to have the same form as the old one, we must have $A = aI$ and $-a^{-1} b A^{-1} dx + a^* \omega_H = 0$. The first equation implies that $1 = \det h = a \det A = a^{n+1}$. Thus $a = \pm 1$, and so the second equation implies that $b = 0$. Hence $h = \pm I$, and so h is the identity in $PSl_{n+1}(\mathbf{R})$. ■

Torsion and Torsion Free Geometries

Recall that a torsion free geometry is one for which the curvature takes values in the H submodule $\mathfrak{h} \subset \mathfrak{g}$. Here is a criterion for a projective torsion free geometry.

Lemma 2.3. *Let M be equipped with a projective geometry. Fix a coordinate system on $U \subset M$, and let $(U, \pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix})$ be the gauge provided by Proposition 2.2. If we write $\alpha_{ik} = \sum_{1 \le l \le n} \Gamma_{ikl} \theta_l$, then the geometry is torsion free if and only if Γ_{ijk} is symmetric in the last two indices.*

Proof. Torsion free $\Leftrightarrow \Theta = 0 \Leftrightarrow d\theta + \theta \wedge \varepsilon + \alpha \wedge \theta = 0$. But in the gauge of Proposition 2.2, $\theta = dx$ so $d\theta = 0$. Also $\varepsilon = 0$, so

$$\text{torsion free} \Leftrightarrow \alpha \wedge \theta = 0$$

$$\Leftrightarrow \sum_{1\le k\le n} \alpha_{ik}\wedge\theta_k = 0 \quad \text{for } 1\le i\le n$$
$$\Leftrightarrow \sum_{1\le k\le n} \Gamma_{ikl}\theta_l\wedge\theta_k = 0 \quad \text{for } 1\le i\le n$$
$$\Leftrightarrow \Gamma_{ikl} = \Gamma_{ilk} \quad \text{for } 1\le i,k,l\le n. \qquad\blacksquare$$

Exercise 2.4.* Show that in a torsion free projective Cartan geometry the curvature satisfies the equation $A\wedge\theta - \theta\wedge\varepsilon = 0$. [*Hint:* use the Bianchi identity.] ❏

The Curvature Function

For later work it is useful to calculate the curvature function with respect to the gauge given in Proposition 2.2.

Proposition 2.5. *Let* $\pi = \begin{pmatrix} 0 & v \\ \theta & \alpha\end{pmatrix}$ *be a gauge of the form guaranteed by Proposition 2.2 and let* $\partial_i = \frac{\partial}{\partial x_i}$. *Assume the geometry is torsion free so that we may write the corresponding curvature as* $\Pi = \begin{pmatrix} \star & \star \\ 0 & A\end{pmatrix}$. *Then the curvature function satisfies*

(i) $K = \dfrac{1}{2}\displaystyle\sum_{1\le i,j,p,q\le n} A_{pq}(\partial_i,\partial_j)e_i^*\wedge e_j^*\otimes e_{pq} \bmod \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h}_2)$,

(ii) $Ricci(K) = \displaystyle\sum_{1\le i,j\le n}\left(\sum_{1\le q\le n} A_{qj}(\partial_i,\partial_q)\right) e_i^*\otimes e_j^*$.

Proof. (i) Since we can write the curvature as

$$\Pi = (\Pi_{pq}) = \sum_{0\le p,q\le n}\Pi_{pq}E_{pq} = \frac{1}{2}\sum_{\substack{0\le p,q\le n\\ 1\le i,j\le n}}\Pi_{pq}(\partial_i,\partial_j)\theta_i\wedge\theta_j E_{pq}$$

(where as usual $E_{pq}\in M_{n+1}(\mathbf{R})$ is the standard elementary matrix), it follows that we can write the curvature function (using the notation of Exercise 1.10) as

$$\begin{aligned}
K &= \frac{1}{2}\sum_{\substack{0\le p,q\le n\\ 1\le i,j\le n}}\Pi_{pq}(\partial_i,\partial_j)e_i^*\wedge e_j^*\otimes E_{pq}\\
&= \frac{1}{2}\sum_{\substack{0\le p,q\le n\\ 1\le i,j\le n}}\Pi_{pq}(\partial_i,\partial_j)e_i^*\wedge e_j^*\otimes e_{pq}\\
&\quad + \frac{1}{2(n+1)}\sum_{\substack{0\le p\le n\\ 1\le i,j\le n}}\Pi_{pp}(\partial_i,\partial_j)e_i^*\wedge e_j^*\otimes\sum_{0\le r\le n}E_{rr}\\
&= \frac{1}{2}\sum_{\substack{0\le p,q\le n\\ 1\le i,j\le n}}\Pi_{pq}(\partial_i,\partial_j)e_i^*\wedge e_j^*\otimes e_{pq}\\
&\qquad\left(\text{since } \sum_{0\le p\le n}\Pi_{pp}(\partial_i,\partial_j) = 0 \quad \text{for all } i,j\right)\\
&= \frac{1}{2}\sum_{1\le i,j,p,q\le n} A_{pq}(\partial_i,\partial_j)e_i^*\wedge e_j^*\otimes e_{pq} \bmod \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}),\mathfrak{h}_2).
\end{aligned}$$

(ii) From (i) and Exercise 1.10(iii), we have

$$\begin{aligned}
Ricci(K) &= \frac{1}{2}\sum_{1\le i,j,p,q\le n} A_{pq}(\partial_i,\partial_j)\,Ricci(e_i^*\wedge e_j^*\otimes e_{pq})\\
&= \frac{1}{2}\sum_{q\le i,j,p,q\le n} A_{pq}(\partial_i,\partial_j)(e_i^*\delta_{jp} - e_j^*\delta_{ip})\otimes e_q^*\\
&= \frac{1}{2}\sum_{1\le i,j,q\le n} A_{jq}(\partial_i,\partial_j)e_i^*\otimes e_q^* - \frac{1}{2}\sum_{1\le i,j,q\le n} A_{iq}(\partial_i,\partial_j)e_j^*\otimes e_q^*\\
&= \sum_{1\le i,j,q\le n} A_{jq}(\partial_i,\partial_j)e_i^*\otimes e_q^*\\
&= \sum_{1\le i,q\le n}\left(\sum_{1\le j\le n} A_{jq}(\partial_i,\partial_j)\right)e_i^*\otimes e_q^*. \qquad\blacksquare
\end{aligned}$$

Special Geometries

As we saw in Lemma 1.6, $\mathfrak{h}$ contains two ideals, $\mathfrak{h}\supset\mathfrak{h}_1\supset\mathfrak{h}_2$. We study the two subclasses of torsion free geometries for which the curvature takes values in $\mathfrak{h}_i$, $i = 1,2$, respectively.

Curvature type $\mathfrak{h}_1$. This means that the curvature has the form

$$\Pi = \begin{pmatrix} 0 & Y \\ 0 & A\end{pmatrix}.$$

Exercise 2.6. Show that for a geometry of curvature type $\mathfrak{h}_1$, the Bianchi identity implies, in addition to the relations of Exercise 2.4, the relation $Y\wedge\theta = 0$. ❏

Curvature type $\mathfrak{h}_2$. This means that the curvature has the form

$$\Pi = \begin{pmatrix} 0 & Y \\ 0 & 0\end{pmatrix}.$$

Proposition 2.7. *A projective Cartan geometry of dimension $n > 2$ with curvature in $\mathfrak{h}_2$ is flat.*

Proof. The Bianchi identity is

$$d\begin{pmatrix} E & Y \\ \Theta & A \end{pmatrix} = \left[\begin{pmatrix} E & Y \\ \Theta & A \end{pmatrix}, \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix}\right].$$

In the present case, where $E = \Theta = A = 0$, the $(2,2)$ component of this identity is $0 = \theta \wedge Y$. It follows that $Y_i = \theta_i \wedge \phi_i$ for some 1-forms ϕ_i. Thus, for $n > 2$, Y_i has all three of θ_1, θ_2, θ_3 as factors, which is impossible for a nonzero 2-form. Hence $Y_i = 0$ for $i \geq 1$, and so $Y = 0$. ∎

Normal and Symmetric Projective Cartan Geometries

Somewhat more subtle in their definition than the subclasses of geometries considered above are the *normal geometries* and the *symmetric geometries.*

Definition 2.8. A projective Cartan geometry is called *normal* (respectively, *symmetric*) if its curvature function $K\colon P \to \operatorname{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h})$ takes values in the normal (respectively, *symmetric*) submodule described in Definition 1.9. ❀

Note that a normal geometry is symmetric and that both geometries are of type $\mathfrak{h}_1$.

Proposition 2.9. *A type $\mathfrak{h}_1$ projective geometry is symmetric.*

Proof. It suffices to show that this is true for a single gauge, so we shall use the gauge of Proposition 2.2.

Let us consider first the case $n = 2$. In this case we have

$$\mathit{Ricci}(K) = \sum_{1 \leq i,j \leq 2} \left(\sum_{1 \leq q \leq 2} A_{qj}(\partial_i, \partial_q) \right) e_i^* \otimes e_j^*,$$

so the symmetry follows from the calculation

$$\begin{aligned}\sum_{1 \leq q \leq 2} A_{q2}(\partial_1, \partial_q) - \sum_{1 \leq q \leq 2} A_{q1}(\partial_2, \partial_q) &= A_{22}(\partial_1, \partial_2) - A_{11}(\partial_2, \partial_1) \\ &= A_{22}(\partial_1, \partial_2) + A_{11}(\partial_1, \partial_2)\end{aligned}$$

together with the fact that $E = 0$ and $E + \text{Trace } A = 0$.

For the case $n \geq 3$, we use the part of the Bianchi identity $d\Pi = [\Pi, \pi]$ coming from the $(2,1)$ block, which reads (from Exercise 2.4 and the fact that $\varepsilon = 0$) $A \wedge \theta = 0$. Evaluating this equation on the triple $\partial_q \wedge \partial_i \wedge \partial_j$, we get

$$\begin{aligned} 0 &= \sum_{1 \leq k \leq n} A_{qk} \wedge \theta_k(\partial_q \wedge \partial_i \wedge \partial_j) \\ &= \sum_{1 \leq k \leq n} \{A_{qk}(\partial_q \wedge \partial_i)\delta_{kj} + A_{qk}(\partial_i \wedge \partial_j)\delta_{kq} + A_{qk}(\partial_j \wedge \partial_q)\delta_{ki}\} \\ &= A_{qj}(\partial_q \wedge \partial_i) + A_{qq}(\partial_i \wedge \partial_j) + A_{qi}(\partial_j \wedge \partial_q). \end{aligned}$$

Now Trace A = Trace $\Pi - E = 0$, so summing on q we get

$$0 = \sum_{1 \leq q \leq n} A_{qj}(\partial_q \wedge \partial_i) + \sum_{1 \leq q \leq n} A_{qi}(\partial_j \wedge \partial_q),$$

or

$$\sum_{1 \leq q \leq n} A_{qj}(\partial_i \wedge \partial_q) = \sum_{1 \leq q \leq n} A_{qi}(\partial_j \wedge \partial_q).$$

It follows that *Ricci* K takes its values in the symmetric part of $(\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$. ∎

§3. The Geometry of Geodesics

In this section we are going to study collections of regular curves $c\colon I \to U \subset \mathbf{R}^n$ satisfying a system of ODEs of the form

$$E(P)\colon \frac{\ddot{c}_1 + P_1(\dot{c})}{\dot{c}_1} = \frac{\ddot{c}_2 + P_2(\dot{c})}{\dot{c}_2} = \cdots = \frac{\ddot{c}_n + P_n(\dot{c})}{\dot{c}_n}, \qquad (3.1)$$

where $P = (P_1, \ldots, P_n)$ with $P_i(v) = \sum a_{ij} v_i v_j$ having coefficients $a_{ij}\colon U \to \mathbf{R}$.

Lemma 3.2. (i) *For each point $x \in U$ and for each vector $v \in T_x(U)$, a system of the form $E(P)$ has a unique solution $c(t)$, defined on a neighborhood of 0, satisfying $c(0) = x$ and $\dot{c}(0) = v$.*

(ii) *Let $E(P)$ and $E(Q)$ be two systems of ODEs of the above type. They have the same solutions if and only if there is a 1-form $\phi \in A^1(U)$ such that $P_i(v) = Q_i(v) + v_i\phi(v)$ for all v.*

Proof. (i) This is a consequence of the usual existence result for solutions of ODEs. Of course, it does not depend on the Ps being quadratic functions.

(ii) $\Leftarrow$ Since

$$\frac{\ddot{c}_i + P_i(\dot{c})}{\dot{c}_i} = \frac{\ddot{c}_i + Q_i(\dot{c}) + \dot{c}_i\phi(\dot{c})}{\dot{c}_i} = \frac{\ddot{c}_i + Q_i(\dot{c})}{\dot{c}_i} + \phi(\dot{c}),$$

it follows that $E(P)$ is satisfied $\Leftrightarrow$ $E(Q)$ is satisfied.

$\Rightarrow$ Suppose that a curve c satisfies $E(P) \Leftrightarrow c$ satisfies $E(Q)$. Taking the difference of the equations $E(P)$ and $E(Q)$ yields

$$\frac{P_1(\dot{c}) - Q_1(\dot{c})}{\dot{c}_1} = \cdots = \frac{P_n(\dot{c}) - Q_n(\dot{c})}{\dot{c}_n} = \phi(\dot{c}) \quad \text{(say)}.$$

By (i) there are solutions c of $E(P)$ satisfying $c(0) = x$ and $\dot{c}(0) = v$ for x and v arbitrary so that ϕ is defined for every $x \in U$ and for every $v \in T_x(U)$. Since P and Q are quadratic, it follows that $\phi(v)$ is linear in v. Thus ϕ is a 1-form and $P_i(v) = Q_i(v) + v_i\phi(v)$ for all v. ■

Exercise 3.3. Show that for $n = 2$, the system of ordinary differential equations $E(P)$ reduces to a single equation that may be put in the form

$$y'' = A + By' + C(y')^2 + D(y')^3$$

(where A, B, C, D are functions of $(x, y) \in \mathbf{R}^2$ and $'$ denotes derivative with respect to x). ❑

Geodesics

Definition 3.4. Let M be a manifold equipped with a projective geometry. A *geodesic* on M is a curve that develops to a straight line in the model geometry. ❀

Note that by Proposition **5**.4.13, development is the same no matter what gauge is used. Thus, the notion of geodesic is well defined. Note also that we are concerned with *unparametrized* geodesics.[2] We are going to show that these geodesics of M satisfy a second-order ordinary differential equation of type $E(P)$ for some P. Then we shall show a converse, that a second-order differential equation of this type determines a unique normal projective geometry on M whose geodesics are the solutions of the differential equations.

Proposition 3.5. *Let* $(U, \pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix})$ *be a projective Cartan gauge and let* $c\colon (\alpha, \beta) \to U$ *be a smooth curve. Then c is a geodesic if and only if for all* $i = 1, 2, \ldots, n$, *the following expressions are equal:*

$$\frac{\theta_i(\dot{c})^{\cdot} + P_i(\dot{c})}{\theta_i(\dot{c})},$$

where "$\cdot$" $= d/dt$ and P_i $(1 \le i \le n)$ are the homogeneous quadratic polynomials

$$P_i(v) = \sum_{1 \le j \le n} \alpha_{ij}(v)\theta_j(v).$$

In particular, the geodesics are completely independent of the choice of v.

Proof. We must show that the given differential equation is a necessary and sufficient condition for the curve $c(t)$ to develop (cf. Definition **5**.4.15) to a straight line in the model space $\mathbf{P}^n$. We can develop the associated function $\pi(\dot{c})\colon (\alpha, \beta) \to \mathfrak{g}$, to obtain a curve $\tilde{c}\colon ((\alpha, \beta), 0) \to (G, e)$ satisfying $\omega_g(\dot{\tilde{c}}(t)) = \pi(\dot{c}(t))$. The projection of this curve to the model space $\mathbf{P}^n$ is $\tilde{c}(t)[e_0]$, where $[e_0] \in \mathbf{P}^n$, and this curve in $\mathbf{P}^n$ is covered by the curve $\tilde{c}(t)e_0$ in $\mathbf{R}^{n+1}$. Now $\tilde{c}(t)[e_0]$ lies in a straight-line segment in $\mathbf{P}^n$ if and only if $\tilde{c}(t)e_0$ lies in a plane of $\mathbf{R}^{n+1}$; the latter condition is equivalent to the condition that $\tilde{c}(t)e_0$, $\dot{\tilde{c}}(t)e_0$, and $\ddot{\tilde{c}}(t)e_0$ are linearly dependent. Thus, $c(t)$ is a geodesic if and only if there are functions $\mu(t)$ and $\lambda(t)$ such that $\ddot{\tilde{c}}e_0 = \mu\tilde{c}e_0 + \lambda\dot{\tilde{c}}e_0$ or, equivalently, $\tilde{c}^{-1}\ddot{\tilde{c}}e_0 = \mu e_0 + \lambda\tilde{c}^{-1}\dot{\tilde{c}}e_0$. But we have $\tilde{c}^{-1}\dot{\tilde{c}} = \omega_G(\dot{\tilde{c}}) = \pi(\dot{c})$. Also, since $(\tilde{c}^{-1})^{\cdot} = -\tilde{c}^{-1}\dot{\tilde{c}}\tilde{c}^{-1}$, we have

$$\pi(\dot{c})^{\cdot} = -\tilde{c}^{-1}\dot{\tilde{c}}\tilde{c}^{-1}\dot{\tilde{c}} + \tilde{c}^{-1}\ddot{\tilde{c}} = -\pi(\dot{c})^2 + \tilde{c}^{-1}\ddot{\tilde{c}}.$$

Using these identities, the equation that c be a geodesic becomes

$$\pi(\dot{c})^{\cdot}e_0 + \pi(\dot{c})^2 e_0 = \mu e_0 + \lambda\pi(\dot{c})e_0.$$

The e_0 component of this equation is the only component involving μ, and it may be satisfied by taking it as a definition of μ. The other components of this equation may be expressed in terms of the constituent forms of π to yield

$$\frac{\theta_i(\dot{c})^{\cdot} + \sum_{0 \le j \le n} \alpha_{ij}(\dot{c})\theta_j(\dot{c})}{\theta_i(\dot{c})} = \lambda, \quad \text{for } i = 1, \ldots, n. \qquad ■$$

Corollary 3.6. *If* $(U, \pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix})$ *is a gauge for which* $dx_i = \theta_i$, *then the equation for the geodesics is the system* $E(P)$, *where* $P = (P_1, \ldots, P_n)$ *and* $P_i(v) = \sum_{0 \le j \le n} \alpha_{ij}(v)\theta_j(v)$.

Proof. In this case we simply have $\pi_i(\dot{c}) = \dot{c}_i$, so $\frac{d}{dt}\pi_i(\dot{c}) = \frac{d}{dt}\dot{c}_i = \ddot{c}_i$. ■

Exercise 3.7. Verify that the condition of the equality for all $i = 1, 2, \ldots, n$ of the expressions

$$\frac{\theta_i(\dot{c})^{\cdot} + P_i(\dot{c})}{\theta_i(\dot{c})}$$

is a condition that is independent of the way in which the curve c is parametrized. ❑

[2] We could of course also discuss *parametrized* geodesics, that is, those geodesics for which the parameter of the development on $\mathbf{P}^n$ is a projective parameter, that is, a parameter that, in any affine coordinate system, is a linear fractional transformation of an affine parameter.

Theorem 3.8. *Suppose we are given, on an open set $U \subset \mathbf{R}^n$, a collection of curves $\mathcal{C}$ constituting the solutions of a system of ordinary differential equations of the form $E(P)$* (cf. Eq. (3.1)).

(i) *There are many projective geometries on U having the curves $\mathcal{C}$ as their geodesics. We can represent the geometrical equivalence class of any such geometry by a unique gauge $\pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix}$ with $\theta_i = dx_i$, $1 \le i \le n$, and $\varepsilon = 0$. Among these geometries,*

(a) *the block $\alpha = (\alpha_{ij})$ is determined by $\mathcal{C}$ up to addition by a matrix*

$$S = (S_{ij}) = (\sum \Lambda_{ijk}\theta_k)$$

where Λ_{ijk} is skew symmetric in the last two indices but is otherwise arbitrary,

(b) *the block v may be chosen arbitrarily, and*

(c) *the variation by the matrix S together with the form v parametrizes in a one-to-one fashion the geometric equivalence classes of geometries whose geodesics are $\mathcal{C}$.*

(ii) *Among the geometries described in* (i), *there is exactly one choice of the α yielding a torsion free geometry. This choice is independent of the block v, so this block parametrizes the torsion free geometries in a one-to-one fashion.*

(iii) *Among the geometries described in* (i), *there is a unique normal projective Cartan geometry.*

Proof. (i) Let us find all the projective Cartan geometries with $\mathcal{C}$ as geodesics.

First we note that although the curves $\mathcal{C}$ satisfy the system $E(P)$, this doesn't mean that P is determined by $\mathcal{C}$. In fact, by Lemma 3.2(ii) we know that $\mathcal{C}$ determines P only up to the addition by a term of the form $v\phi(v)$, where ϕ is an arbitrary 1-form on U.

By Proposition 2.2 we know that any projective geometry on $U \subset \mathbf{R}^n$, with standard coordinates $x = (x_1, \ldots, x_n)$, has a unique gauge $\pi = \begin{pmatrix} \varepsilon & v \\ \theta & \alpha \end{pmatrix}$ on $\mathbf{R}^n$ satisfying $\varepsilon = 0$ and $\theta = dx$. Thus, we merely need to determine all the ways of choosing the remaining blocks α and v so that the curves $\mathcal{C}$ are geodesics. The only a priori condition on the block α is that Trace $\alpha) = 0$ (recall that $\varepsilon = 0$).

Now by Proposition 3.5 the geodesics of a geometry are the smooth curves $c\colon (\alpha, \beta) \to U$ that are solutions of the equations stating the equality, for all $i = 1, 2, \ldots, n$, of the expressions

$$\frac{\ddot{c}_i + \sum_{1 \le i \le n} \alpha_{ij}(\dot{c})\dot{c}_j}{\dot{c}_i}.$$

It follows that the geometries in which the curves $\mathcal{C}$ are geodesics are those for which

$$\sum_{1 \le j \le n} \alpha_{ij}(v)v_j = P_i(v) + v_i\phi(v), \quad i = 1, \ldots, n, \tag{3.9}$$

where $\phi \in A^1(U)$ is arbitrary. Thus, the collection of geometric isomorphism classes of geometries with the required geodesics is in one-to-one correspondence with the collection of solutions α of (3.9) satisfying $\mathrm{Trace}(\alpha) = 0$.

Let us solve Eq. (3.9) for the forms α_{ij}. Write the homogeneous quadratic polynomials P_i as

$$P_i(v_1, \ldots, v_n) = \sum R_{ij}(v)v_j,$$

where $R_{ij}(v)$ is some linear combination of the v_is. The R_{ij} are determined by this equation up to addition with $S_{ij} = \sum \Lambda_{ijk}v_k$, where Λ is skew symmetric in the last two indices but is otherwise arbitrary.

Now set $\alpha_{ij} = R_{ij}$ for $i \ne j$, $1 \le i, j \le n$, and $\alpha_{ii} = R_{ii} - \frac{1}{n}\sum_{1 \le k \le n} R_{kk}$ for $1 \le i \le n$. Then $\mathrm{Trace}(\alpha) = 0$, so the $\pi = \begin{pmatrix} 0 & v \\ dx & \alpha \end{pmatrix}$ constitutes a projective gauge. Moreover, for $1 \le i \le n$, we have[3]

$$\sum_{1 \le j \le n} \alpha_{ij}dx_j = P_i(dx_1, \ldots, dx_n) - dx_i\left(\frac{1}{n}\sum_{1 \le k \le n} R_{kk}\right).$$

Thus, Eq. (3.9) is satisfied by this choice of π no matter how the block v is chosen and no matter how the α is varied by the matrix S. Moreover, by Proposition 2.2 each of these variations leads to a distinct geometric isomorphism class of projective Cartan geometry. This finishes the proof of (i).

(ii) By Lemma 2.3, a geometry is torsion free if and only if, for a gauge of the type given in Proposition 2.2 (as is the case in the proof of (i)), when we write $\alpha_{ik} = \sum \Gamma_{ikl}\theta_l$, then Γ_{ijk} is symmetric in the last two indices. For a geometry with geodesics $\mathcal{C}$, the choice of the α_{ik} is determined up to the replacement

$$\alpha_{ik} = \sum \Gamma_{ikl}\theta_l \mapsto \sum(\Gamma_{ikl} + \Lambda_{ijk})\theta_l,$$

where Λ_{ijk} is skew symmetric in the last two variables. Clearly, we may take $\Lambda_{ijk} = -\frac{1}{2}(\Gamma_{ijk} - \Gamma_{ikj})$ to be minus the skew-symmetric (in the last two indices) part of Γ_{ikl} and so replace π by a torsion free connection. It is

---

[3] The product of 1-forms is the symmetric product here.

clear from this that there is a unique block α such that $\pi = \begin{pmatrix} 0 & v \\ dx & \alpha \end{pmatrix}$ has geodesics C for which this is true. The only remaining freedom in choosing the connection comes from the choice of the block v, which may be chosen arbitrarily and which therefore parametrizes the geometric isomorphism classes of torsion free geometries with geodesics C.

(iii) Now we show that among the geometries constructed in (i) is a unique normal geometry. Since a normal geometry is torsion free, by (ii) only the freedom to choose the block v remains. Since a normal geometry is of type $\mathfrak{h}_1$, the $(1,1)$ block of the curvature vanishes. Let us write $v = \theta^t a$ for some function $a\colon U \to M_n(\mathbf{R})$, so that the curvature has the form

$$\begin{aligned}\Pi = \begin{pmatrix} 0 & Y \\ 0 & A \end{pmatrix} = d\pi + \pi\wedge\pi &= \begin{pmatrix} v\wedge\theta & \star \\ 0 & d\alpha + \theta\wedge v + \alpha\wedge\alpha \end{pmatrix} \\ &= \begin{pmatrix} \theta^t a\wedge\theta & \star \\ 0 & d\alpha + \theta\wedge\theta^t a + \alpha\wedge\alpha \end{pmatrix}.\end{aligned}$$

In particular, the vanishing of the $(1,1)$ block is equivalent to the matrix a's being symmetric since

$$0 = \theta^t a\wedge\theta = \sum_{1\le i,j\le n} \theta_i a_{ij}\wedge\theta_j = \sum_{1\le i<j\le n} (a_{ij} - a_{ji})\theta_i\wedge\theta_j.$$

Finally, we study the effect on the curvature function of varying the block v. We see that the change $v \mapsto v + \theta^t b$ (where b is a symmetric $n\times n$ matrix) has the following effect on A:

$$A = d\alpha + \theta\wedge v + \alpha\wedge\alpha \mapsto d\alpha + \theta\wedge(v + \theta^t b) + \alpha\wedge\alpha = A + \theta\wedge\theta^t b$$

so that

$$\begin{aligned}A_{pq}(\partial_i,\partial_j) \mapsto\; & A_{pq}(\partial_i,\partial_j) + (\theta\wedge\theta^t b)_{pq}(\partial_i,\partial_j) \\ &= A_{pq}(\partial_i,\partial_j) + \sum_{1\le k\le n} \theta_p\wedge\theta_k b_{kq}(\partial_i,\partial_j) \\ &= A_{pq}(\partial_i,\partial_j) + \sum_{1\le k\le n} (\delta_{pi}\delta_{kj} - \delta_{pj}\delta_{ki})b_{kq} \\ &= A_{pq}(\partial_i,\partial_j) + \delta_{pi}b_{jq} - \delta_{pj}b_{iq}.\end{aligned}$$

Since by Proposition 2.5(ii) we have

$$Ricci(K) = \sum_{1\le i,j\le n}\left(\sum_{1\le q\le n} A_{qj}(\partial_i,\partial_q)\right) e_i^*\otimes e_j^*,$$

it follows that

$$\begin{aligned}Ricci(K) \mapsto\; & Ricci(K) + \sum_{1\le i,j\le n}\left(\sum_{1\le q\le n} (\delta_{qi}b_{qj} - \delta_{qq}b_{ij})\right) e_i^*\otimes e_j^* \\ &= Ricci(K) - (n-1)\sum_{1\le i,j\le n} b_{ij}e_i^*\otimes e_j^*.\end{aligned}$$

By Proposition 2.9, $Ricci(K)$ takes values in the symmetric part of $(\mathfrak{g}/\mathfrak{h})^*\otimes(\mathfrak{g}/\mathfrak{h})^*$, so it follows that for $n > 1$ there is a unique choice of the symmetric matrix (b_{lk}) such that $Ricci(K) = 0$. Thus, among the projective geometries whose geodesics are C, there is a unique normal geometry. ∎

Corollary 3.10. *Suppose we are given a manifold M^n on which there is a system E of second-order ODEs such that M is covered by coordinate systems in each of which the system E assumes the form of equation $E(P)$ for some P (cf. Eq. (3.1)). Then there is a unique normal projective Cartan geometry on M whose geodesics are the solutions of E.*

Proof. If we choose a coordinate system in which the system E has the form of equation $E(P)$, then on each coordinate patch this corollary is just 3.8(iii). The uniqueness allows us to fit together the geometries on the various coordinate patches. ∎

§4. The Projective Connection in a Riemannian Geometry

In this study we study the geodesics in a torsion free Riemannian geometry on M and construct from them a normal projective Cartan geometry on M. Thus we have two model pairs, the Riemannian pair $(\mathfrak{g}_{\text{Rie}}, \mathfrak{h}_{\text{Rie}})$ with group H_{Rie} and the projective pair $(\mathfrak{g},\mathfrak{h})$ with group H. We shall regard $\mathfrak{g}_{\text{Rie}}$ as a subalgebra of $\mathfrak{g}$ and H_{Rie} as a subgroup of H according to

$$\begin{aligned}\mathfrak{g}_{\text{Rie}} &= \left\{\begin{pmatrix} 0 & 0 \\ c & A \end{pmatrix} \mid c\in\mathbf{R}^n, A^t = -A \in M_n(\mathbf{R})\right\} \\ &\subset \mathfrak{g} = \left\{\begin{pmatrix} a & b \\ c & A \end{pmatrix} \mid a + trace\ A = 0\right\}, \\ H_{\text{Rie}} &= \left\{\begin{pmatrix} 1 & 0 \\ c & A \end{pmatrix} \mid c\in\mathbf{R}^n, A\in O_n(\mathbf{R})\right\} \\ &\subset H = PSl_{n+1}(\mathbf{R}) = \left\{\begin{pmatrix} a & b \\ c & A \end{pmatrix} \in Sl_{n+1}(\mathbf{R})\right\}/Z.\end{aligned}$$

The proof of the following theorem is really an application of the ideas in the proof of Theorem 3.8. However, for the sake of clarity we include all the details.

Theorem 4.1. *The equation for the geodesics in a Euclidean geometry is of the type $E(P)$, so there is a canonical normal projective connection with these geodesics. If the geometry is Reimannian (i.e., a torsion free Euclidean geometry) with gauge and curvature given by*

$$\omega = \begin{pmatrix} 0 & 0 \\ \theta & \alpha \end{pmatrix}, \quad \Omega = \begin{pmatrix} 0 & 0 \\ 0 & A \end{pmatrix},$$

then the associated normal projective geometry has gauge and curvature given by

$$\pi = \begin{pmatrix} 0 & v \\ \theta & \alpha \end{pmatrix},$$

where

$$v = \frac{1}{n-1} \sum_{1 \le k \le n} \left\{ \sum_{1 \le l \le n} A_{lj}(e_k, e_l) \right\} \theta_k$$

and

$$\Pi = \begin{pmatrix} 0 & dv + v \wedge \theta \\ 0 & A + \theta \wedge v \end{pmatrix}.$$

Proof. By Proposition 6.2.7 the equation for the geodesics of a Euclidean geometry with gauge π is

$$\frac{\theta_1(\dot{c})^{\cdot} + P_1(\dot{c})}{\theta_1(\dot{c})} = \cdots = \frac{\theta_n(\dot{c})^{\cdot} + P_n(\dot{c})}{\theta_n(\dot{c})},$$

where

$$P_i(v) = \sum_{1 \le j \le n} \alpha_{ij}(v)\theta_j(v).$$

Since this has the form of equation $E(P)$ (see Eq. (3.1)), it follows that there is a unique normal projective Cartan geometry on M whose geodesics are the Euclidean geodesics.

Let us determine this projective geometry explicitly in the case when π is torsion free. We begin by studying a projective gauge of the form $\pi = \begin{pmatrix} 0 & v \\ \theta & \alpha \end{pmatrix}$ (i.e., θ and α come from the Riemannian gauge, $\varepsilon = 0$, and v is to be determined) so that the corresponding curvature is

$$\Pi = \begin{pmatrix} v \wedge \theta & dv + v \wedge \alpha \\ d\theta + \alpha \wedge \theta & d\alpha + \theta \wedge v + \alpha \wedge \alpha \end{pmatrix}.$$

Since ω is torsion free, it follows immediately that π is also torsion free. The equation for the projective geodesics is (by Proposition 3.5) the equality, for $1 \le i \le n$, of the expressions

$$\frac{\theta_i(\dot{c})^{\cdot} + \sum_{1 \le j \le n} \theta_{ij}(\dot{c})\theta_j(\dot{c})}{\theta_i(\dot{c})}.$$

Note that this is identical to the equation for the Riemannian geodesics and is independent of the choices for the block v.

It remains to show how to choose the block v to obtain a normal geometry. The projective curvature mod $\mathfrak{h}_2$ is

$$\Pi = \begin{pmatrix} v \wedge \theta & \star \\ 0 & d\alpha + \theta \wedge v + \alpha \wedge \alpha \end{pmatrix} = \Omega + \begin{pmatrix} v \wedge \theta & \star \\ 0 & \theta \wedge v \end{pmatrix}.$$

Since a normal geometry is of type $\mathfrak{h}_1$, it follows that we must choose the block v so that the $(1,1)$ block of the curvature vanishes. If we write $v = \theta^t b$ for some function $b \colon U \to M_n(\mathbf{R})$, the $(1,1)$ block of the curvature is

$$v \wedge \theta = \theta^t b \wedge \theta = \sum_{1 \le i,j \le n} \theta_i b_{ij} \wedge \theta_j = \sum_{1 \le i < j \le n} (b_{ij} - b_{ji})\theta_i \wedge \theta_j.$$

In particular, the vanishing of the $(1,1)$ block is equivalent to the matrix b's being symmetric. Let us assume this from now on. We have

$$\begin{aligned} \Pi = (\Pi_{pq}) &= \sum_{0 \le p,q \le n} \Pi_{pq} E_{pq} = \frac{1}{2} \sum_{\substack{0 \le p,q \le n \\ 1 \le i,j \le n}} \Pi_{pq}(\partial_i, \partial_j)\theta_i \wedge \theta_j E_{pq} \\ &= \frac{1}{2} \sum_{\substack{0 \le p,q \le n \\ 1 \le i,j \le n}} (A_{pq}(\partial_i, \partial_j) + \theta_p \wedge v_q(\partial_i, \partial_j))\theta_i \wedge \theta_j E_{pq} \end{aligned}$$

(where as usual $E_{pq} \in M_{n+1}(\mathbf{R})$ is the standard elementary matrix). It follows that we can write the curvature function (using the notation of Exercise 1.10 and the calculation of Theorem 3.8) as

$$K = \frac{1}{2} \sum_{1 \le i,j,p,q \le n} (A_{pq}(\partial_i, \partial_j) + \theta_p \wedge v_q(\partial_i, \partial_j)) e_i^* \wedge e_j^* \otimes e_{pq}$$

$$\operatorname{mod} Hom(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}_2).$$

Applying the Ricci homomorphism $Ricci : \mathrm{Hom}(\lambda^2(\mathfrak{g}/\mathfrak{h}), \mathfrak{h}) \to (\mathfrak{g}/\mathfrak{h})^* \otimes (\mathfrak{g}/\mathfrak{h})^*$ (Definition 1.9), the curvature function is mapped (Exercise 1.10(iii)) to

$$\begin{aligned} Ricci(K) &= \frac{1}{2} \sum_{1 \le i,j \le n} \left(\sum_{1 \le q \le n} (A_{qj}(\partial_i, \partial_q) + \theta_q \wedge v_j(\partial_i, \partial_q)) \right) e_i^* \otimes e_j^* \\ &= \frac{1}{2} \sum_{1 \le i,j \le n} \left(\sum_{1 \le q \le n} A_{qj}(\partial_i, \partial_q) \right) e_i^* \otimes e_j^* \end{aligned}$$

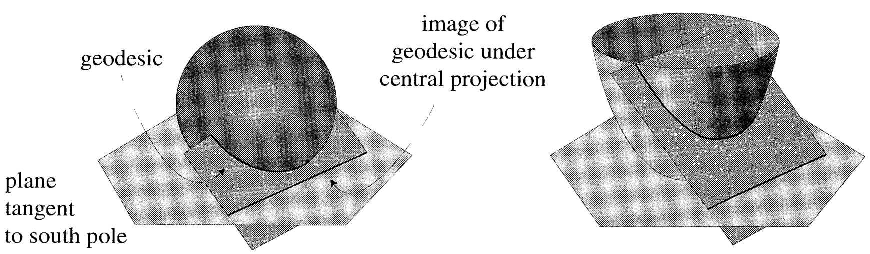
$$
\begin{aligned}
&+\frac{1}{2}\sum_{1\le i,j\le n}\left(\sum_{1\le q,r\le n} b_{rj}\theta_q\wedge\theta_r(\partial_i,\partial_q)\right)e_i^*\otimes e_j^*\\
&=\frac{1}{2}\sum_{1\le i,j\le n}\left(\sum_{1\le q\le n} A_{qj}(\partial_i,\partial_q)\right)e_i^*\otimes e_j^*\\
&+\frac{1}{2}\sum_{1\le i,j\le n}\left(\sum_{1\le q,r\le n} b_{rj}(\delta_{qi}\delta_{rq}-\delta_{qq}\delta_{ri})\right)e_i^*\otimes e_j^*\\
&=\frac{1}{2}\sum_{1\le i,j\le n}\left(\sum_{1\le q\le n} A_{qj}(\partial_i,\partial_q)\right)e_i^*\otimes e_j^*\\
&-\frac{1}{2}(n-1)\sum_{1\le i,j\le n} b_{ij}e_i^*\otimes e_j^*.
\end{aligned}
$$

Since $Ricci(K)$ takes values in the symmetric submodule of $(\mathfrak{g}/\mathfrak{h})^*\otimes(\mathfrak{g}/\mathfrak{h})^*$ (Proposition 2.9), it follows that $Ricci(K)=0$ if and only if we take

$$b_{ij}=\frac{1}{n-1}\sum_{1\le q\le n}A_{qj}(\partial_i,\partial_q). \qquad \blacksquare$$

The Configuration of Geodesics

The following figures serve to remind the reader that central projections provide, locally, a geodesic-preserving diffeomorphism between the n-sphere and Euclidean space and also between n-dimensional hyperbolic space and Euclidean space.



Now we wish to show the converse of this, due to Beltrami, that a Riemannian manifold whose geodesic configuration is locally the same as Euclidean space must have constant curvature.

Theorem 4.2. *Let M be a Riemannian manifold such that each point has a coordinate system in which every geodesic is a straight line. Then M has constant curvature.*

Proof. By Proposition **6**.4.5, it suffices to show that $A=c\theta\wedge\theta^t$ for some constant c. The hypotheses imply that the normal projective geometry on M with the same geodesics as the Riemannian structure on M must be flat.

Projective flatness means $\Pi=0$, so in particular, by Theorem 4.1, $A+\theta\wedge v=0$. Since A is skew symmetric, it follows that the diagonal terms of $\theta\wedge v$ must vanish. Thus, $v_i=\lambda_i\theta_i$ for some functions λ_i, $1\le i\le n$. Thus, $A_{ij}=-\lambda_j\theta_i\wedge\theta_j$, and then the skew symmetry of A implies that λ_i is independent of i, $1\le i\le n$, say $\lambda_i=-c$. Thus, $v=-c\theta^t$ and $A=c\theta\wedge\theta^t$, but we need to show that c is constant.

For $n>2$, the constancy of c is contained in Lemma **6**.4.1 (it is a consequence of the Riemannian Bianchi identity). For $n=2$, the projective gauge and curvature are

$$\pi=\begin{pmatrix}0&-c\theta_1&-c\theta_2\\ \theta_1&0&\theta_{12}\\ \theta_2&-\theta_{12}&0\end{pmatrix},$$

$$0=\Pi=\begin{pmatrix}0&-\{d(c\theta_1)-c\theta_2\wedge\theta_{12}\}&-\{d(c\theta_2)+c\theta_1\wedge\theta_{12}\}\\ d\theta_1+\theta_{12}\wedge\theta_2&0&d\theta_{12}-c\theta_1\wedge\theta_2\\ d\theta_2-\theta_{12}\wedge\theta_1&-d\theta_{12}+c\theta_1\wedge\theta_2)&0\end{pmatrix},$$

respectively. Thus,

$$
\begin{aligned}
0&=d(c\theta_1)-c\theta_2\wedge\theta_{12}=dc\wedge\theta_1+cd\theta_1-c\theta_2\wedge\theta_{12}\\
&=dc\wedge\theta_1-c\theta_{12}\wedge\theta_2-c\theta_2\wedge\theta_{12}=dc\wedge\theta_1,\\
0&=d(c\theta_2)+c\theta_1\wedge\theta_{12}=dc\wedge\theta_2+cd\theta_2+c\theta_1\wedge\theta_{12}\\
&=dc\wedge\theta_2+c\theta_{12}\wedge\theta_1+c\theta_1\wedge\theta_{12}=dc\wedge\theta_2.
\end{aligned}
$$

Hence, $dc=0$ and c is constant. $\blacksquare$

§5. A Brief Tour of Projective Geometry

In this section we describe very briefly a small sampling of the literature related to projective Cartan geometries.

Cartan's Notion of a Subgeometry

Cartan studied a general (and perhaps peculiar by the usual standard) notion for the induced geometry on a submanifold. In Cartan's version we have a submanifold $M^m\subset\mathbf{P}^n$. We select for each point $x\in M$ a "normal" space for M consisting of a projective subspace $\mathbf{P}^{n-m}(x)$ transverse to $T_x(M)$ and use it to obtain a reduction of the principal bundle. (A more usual way is to try to define a projective normal vector to a sufficiently nondegenerate surface. Both methods are discussed in [M.A. Akivis and V.V.

Goldberg, 1993], pp. 173ff). The result of this is that we obtain an induced projective Cartan geometry on M, but it will depend on the choice of "normal" spaces. Chern proved [S.-S. Chern, 1937] that in the analytic category every projective Cartan geometry on M arises in this way, provided

$$n \geq \begin{cases} \frac{1}{2}(m^2 + 2m - 1) & \text{if } m \text{ is odd,} \\ \frac{1}{2}(m^2 + 2m + 1) & \text{if } m \text{ is even.} \end{cases}$$

Families of Submanifolds

In [S.-S. Chern, 1943], Chern generalized both Cartan's result on geodesics in §3 and the analogous results of Hatchtroudi on hypersurfaces by showing that to certain $(k+1)(n-k)$ parameter families of k-dimensional submanifolds of M^n, there may be uniquely attached a normal projective connection. Yen [C.T. Yen, 1953] gave a geometric version of this. Chern and Griffiths [S.-S. Chern and P. Griffiths, 1978] showed that flat projective connections are characterized by having a two-parameter family of totally geodesic submanifolds, suitably distributed (cf. Chern's remark on page xxiv of [S.-S. Chern, 1978]).

Differential Equations

If $\mathbf{P}^2 = G/H$, then the bundle of *contact elements* (i.e., pairs (x, L) where $L \subset T_x(\mathbf{P}^2)$ is a one-dimensional subspace) is $B = G/K$ (where $K \subset H$ is the subgroup of H, of codimension one, fixing a given line in $\mathbf{P}^2$). Cartan showed that a *general* ordinary second-order differential equation on a two dimensional manifold M^2 determines a canonical Cartan geometry, modeled on $B = G/K$, on the bundle $B(M)$ of contact elements of M. There is a differential invariant arising from the curvature of this geometry which vanishes if and only if the equation is of the type considered in §3. In this case the geometry is the "pull up" of the projective geometry on M.

Chern [S.-S. Chern, 1940] generalized this to a third-order ODE on a surface M^2. He also found a different invariant I that can be calculated from the equation. When $I = 0$, he associates to the equation a Cartan geometry on the 3-manifold of solution of the ODE modeled on G_{10}/H_7, where G_{10} is the group of contact transformations in the plane (i.e., the Lie sphere group). On the other hand, if $I \neq 0$, then he associates to the equation a Cartan geometry with fundamental group the five-dimensional conformal group.

For some purposes (e.g., in control theory) the equivalence under change of variables of second-order ordinary differential equations needs to be studied with respect to a smaller group than the group of all diffeomorphisms of the space of variables. Taking the case of an equation of the form $y'' = F(x, y, y')$, $(x, y) \in \mathbf{R}^2$, we may wish to consider changes of variables of the form $x \mapsto \phi(x)$, $y \mapsto \psi(x, y)$. The advantage of this kind of change of variables is that it preserves the role of the independent variable. It is studied in [N. Kamran, K.G. Lamb, and W.F. Shadwick, 1985].

CR Geometry

CR Geometry is the study of the invariants of real hypersurfaces $M^{2n-1} \subset \mathbf{C}^n$ that are unchanged under complex analytic change of variables of $\mathbf{C}^n$. For a recent treatment of CR geometry, see [H. Jacobowitz, 1990]. Chern [S.-S. Chern, 1975] showed that when M is real analytic, there is a canonical projective Cartan geometry induced on M which is such an invariant.

Appendix A
Ehresmann Connections

Charles Ehresmann introduced his definition of an infinitesimal connection on a principal bundle in his paper [C. Ehresmann, 1950]. It put the notion of a Cartan connection on a rigorous footing and generalized it at the same time (cf. Proposition 3.1). Ehresmann's definition was accepted as the definitive definition of a connection. The reader can find a full exposition of this point of view in [S. Kobayashi and K. Nomizu, 1963].

And yet there has been some dissatisfaction with this state of affairs, leading one well-known mathematician to say that the theory of connections was founded on the wrong definition. This is a consequence, it seems, of two factors. First, Ehresmann connections are extremely general and so include perhaps much more than is actually interesting.[1] Secondly, Ehresmann's formulation hides the similarity that connections can have to the Maurer–Cartan form on a Lie group and so provides an unnecessary obstacle to one's understanding.

On the other hand Ehresmann connections appear as important components of the Cartan connections (cf. **6**.5.12 ii) and **7**.4.21) as well as being the foundation for the notion of covariant differentiation, so they do have a job to do.

Perhaps the relation between Cartan connections and Ehresmann connections is something like the relation between rings of algebraic integers

[1] In a different context, Chern [S.-S. Chern, 1966, p. 167] comments on what is perhaps the inevitability of this situation.

and arbitrary commutative rings in that each degree of generality illuminates the other.

We pass now to the definition given in Chapter 6 of an Ehresmann connection on a principal bundle. We repeat it here for convenience.

Definition 6.2.4. Let $H \to P \to M$ be an arbitrary principal bundle over M and let $\mathfrak{h}$ be the Lie algebra of H. An *Ehresmann connection* on P is an $\mathfrak{h}$-valued form γ satisfying the conditions

(i) $R_h^*\gamma = \mathrm{Ad}(h^{-1})\gamma$,

(ii) $\gamma(X^\dagger) = X$ for every $X \in \mathfrak{h}$.

The *curvature* of the Ehresmann connection γ is the $\mathfrak{h}$-valued 2-form on P given by $d\gamma + \frac{1}{2}[\gamma, \gamma]$. ❀

In view of the apparent formal similarity between this definition and the definition of a Cartan connection, it is appropriate to remark on an important difference between them. The form γ takes values in the Lie algebra of the *fiber* of the principal bundle, so its restriction to any tangent space of P, $\gamma_p: T_p(P) \to \mathfrak{h}$, must have a kernel whose dimension is at least the dimension of M. In fact, condition (ii) implies that γ is onto, so it follows that dim ker γ_p = dim M for all $p \in P$. This circumstance is at variance with that of a Cartan connection, which has no kernel at all.

What is the relation between Ehresmann connections and Cartan connections? There are two ways of considering this issue.

For a first-order reductive Cartan geometry, we show in §2 that the notion of a Cartan geometry is equivalent to the notion of a bundle of frames together with an Ehresmann connection on it. An example of this correspondence is given in Exercise **6**.3.6 in the case of Riemannian geometry.

On the other hand, we shall see in §3 how an arbitrary Cartan geometry (P, ω) modeled on the Klein pair (G, H) determines, and is determined by, a special kind of Ehresmann connection on the principal G bundle $P \times_H G$. This correspondence is different from—and perhaps less geometric than—the one described in §2.

§1. The Geometric Origin of Ehresmann Connections

Perhaps the motivating reason for Levi–Civita's search for what is now called the Levi–Civita connection was the need to find an analog of the vector gradient of Euclidean geometry (cf. Exercise **5**.3.50) in the context of Riemannian geometry. In fact, Levi–Civita called his discovery the *absolute derivative*. The idea was generalized in 1917 by H. Weyl to arbitrary vector bundles and was later developed by a number of people, including Schouten, T.Y. Thomas, E. Cartan, and finally C. Ehresmann. While this discovery had primarily analytical roots, it could also have resulted from certain simple geometric considerations, which I will sketch now.

Given a vector field Y on a manifold M and a vector $X_x \in T_x(M)$, what should $X_x(Y) \in T_x(M)$ mean? We may of course formally write

$$X_x(Y) = \lim_{t \to 0} \frac{Y_{x+tX_x} - Y_x}{t}.$$

There is a small difficulty with giving meaning to the expression Y_{x+tX_x}, but this may be solved using a Taylor expansion in t. More serious is the fact that the difference involves two vectors at different points of the manifold (i.e., in two different vector spaces), and so it has no meaning. In Euclidean space we can canonically identify any two tangent spaces by parallel translation, so the difference makes sense there, and in that case $X_x(Y)$ is just the usual vector gradient given by

$$X_x\left(f_1\frac{\partial}{\partial x_1} + \cdots + f_n\frac{\partial}{\partial x_n}\right) = X_x(f_1)\frac{\partial}{\partial x_1} + \cdots + X_x(f_n)\frac{\partial}{\partial x_n}.$$

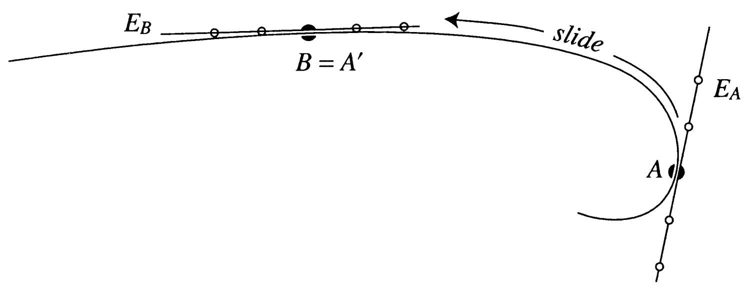
In a general Riemannian manifold, however, we would need to have an "infinitesimal connection," that is, an affine isometry[2] between the tangent spaces of the two "infinitesimally close" points x and "$x + tX_x$." This will lead to affine isometries $T_x(M) \approx T_y(M)$ whenever x and y can be joined by a path, with the isometry being obtained by integration along the curve, that is, by composing the infinitesimal isomorphisms step by step along the curve. Of course, the isometry may well depend on the curve so chosen. This leads us to ask, "is the notion of an infinitesimal connection a natural one in differential geometry?" We are going to give some examples to show that the answer is yes.

Involutes of Curves

The first examples of this refer to the case of a curve M in the Euclidean plane. This is a trivial case, but it serves to help us orient ourselves. Up to translation, there is a unique orientation-preserving isometry between the tangent spaces of M. The geometry suggests two ways of identifying the tangent spaces corresponding to the Ehresmann and Cartan connections, respectively. The first is to slide the tangent line E_A at A along the curve to B so that the point of contact on the moving line is fixed on that line.

[2] We do not wish to assume at the outset that the isometry preserves the origin. Of course, the Levi–Civita connection does preserve it. Even more generally, we can consider an infinitesimal connection on a manifold without any Riemannian structure, which would be a diffeomorphism between the tangent spaces of two "infinitesimally close" points.

We might imagine the tangent line to be fastened rigidly to a small bead fitted around the curve but free to slide along it as in the figure. In this scheme, parallel translation preserves the origin.



The second way is to allow the tangent line to roll without slipping along the curve. If a thread is wound along the curve and the part not in contact with the curve is pulled tight, then this part is tangent to the curve, and its successive positions as the thread is unwound identify the various tangent lines. In this scenario the points of the rolling tangent line trace out the involutes of the curve in the plane as in Figure A.1. The origin is not preserved under this version of parallel translation.

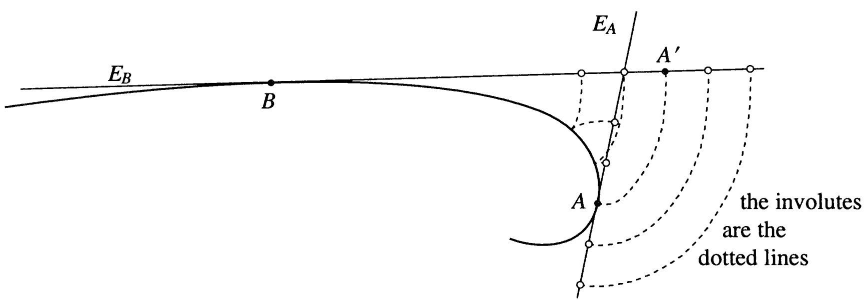


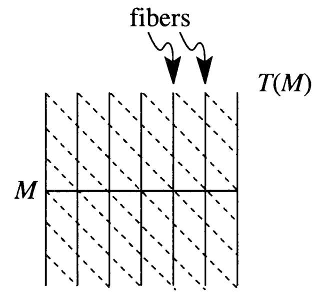
FIGURE A.1

In each of these cases the identification between tangent spaces at nearby points on the curve may be regarded as a one-dimensional distribution on the tangent bundle $T(M)$ of the curve. This is perhaps clearest in the case of involutes, which, although they lie in the plane, may also be regarded as lying on the image of the canonical map

$$\begin{array}{ccc} T(M) & \to & \mathbf{R}^2. \\ (x,v) & \mapsto & x+v \end{array}$$

As curves in $T(M)$, the involutes are integral curves of the distribution on $T(M)$. Although the involutes have cusps, as seen in Figure A.1, this is

merely an artifact of the mapping from the tangent bundle into the plane. There are no singularities in the corresponding distribution in the tangent bundle itself (as in the following picture).



This distribution gives an identification between the fibers of $T(M)$ and indicates how, for an arbitrary manifold M of dimension n, a distribution $\mathcal{H}$ of dimension n on $T(M)$ that is transverse to the fibers of $T(M)$ describes an infinitesimal connection.

In none of these examples have we been dealing with an Ehresmann connection on a principal bundle, so the reader has the right to ask, "what, if any, is the relationship between these notions and the definition of an Ehresmann connection?" We provide the answer to this question in the following two exercises.

Exercise 1.2. (i) Let γ be an Ehresmann connection on the principal bundle $H \to P \to M$. Show that the distribution $\mathcal{D}$ on P given by $\mathcal{D}_p = \ker \gamma_p$ satisfies the conditions

(a) $T_p(P) = T_p(pH) \otimes \mathcal{D}_p$ for all $p \in P$,

(b) $R_{h*}\mathcal{D}_p = \mathcal{D}_{ph}$ for all $p \in P$, $h \in H$.

(ii) Let $H \to P \to M$ be a principal bundle and let $\mathcal{D}$ be a distribution on P satisfying conditions (a) and (b). Show that there is a unique Ehresmann connection γ on P giving rise to $\mathcal{D}$ as in (i) (cf. [S. Kobayashi and K. Nomizu, 1963], pp. 63–64). ❏

Exercise 1.3. Let $H \to P \to M$ be a principal H bundle and let $\mathcal{D}$ be a distribution on P satisfying conditions (a) and (b) of Exercise 1.2(i). Let $V \to E \to M$ be the vector bundle associated to the principal bundle P by the representation (V, ρ) (i.e., $E = P \times_H V$; cf. p. 39). Show that there is a unique distribution $\mathcal{D}_E$ on E such that the derivatives of the canonical maps $f_v: P \to E = P \times_H V$ sending $p \mapsto (p, v)$ yield isomorphisms in the following diagram.

$$\begin{tikzcd} T_p(P) \arrow[d, "f_{v*}"'] & \supset & \mathcal{D}_p \arrow[d, "\approx"] \\ T_{(p,v)}(E) & \supset & (\mathcal{D}_E)_{(p,v)} \end{tikzcd}$$

Show also that the distribution $\mathcal{D}_E$ is transverse to the fiber (i.e., no nonzero vector in $\mathcal{D}_E$ is tangent to the fiber). ❑

Rolling Without Slipping or Twisting Along a Surface in 3-Space

A procedure generalizing the notion of an involute can be carried out for a surface M in 3-space. Given two points $A, B \in M$ and a curve $\sigma: ([a,b], a, b) \to (M, A, B)$, we may allow the tangent plane to "roll without slipping or twisting" along the curve in order to identify the tangent planes at A and B. If we then translate the final position of the tangent plane so that its origin coincides with the point B, this procedure describes the Levi–Civita parallelism (in the induced metric) on the surface.[3] The new ingredient here is that the distributions on the tangent bundle that arise in this way are generally no longer integrable, so the parallel translation definitely depends on the choice of the path.

In Appendix B the notion of one manifold rolling without slipping or twisting on another in Euclidean space and its relationship to Levi–Civita connection and the Ehresmann connection on the normal bundle are studied in detail.

§2. The Reductive Case

In this section our aim is to show that in the case of a first-order reductive Cartan geometry, we have the following equivalence:

$$\left\{\begin{array}{c}\text{first-order Cartan geometries on } M \text{ with}\\ \text{reductive model } (\mathfrak{h}\oplus\mathfrak{p}, \mathfrak{h}), \text{ group } H,\\ \text{and Cartan connection } \omega = \omega_{\mathfrak{p}} + \omega_{\mathfrak{h}}\end{array}\right\}$$

$$\overset{\text{bijection}}{\longleftrightarrow}\left\{\begin{array}{c}\text{principal } H \text{ frame bundles } P \text{ over } M\\ \text{with fundamental form } \omega_{\mathfrak{p}} \text{ on } P \text{ and}\\ \text{Ehresmann connection } \omega_{\mathfrak{h}} \text{ on } P\end{array}\right\}.$$

Lemma 2.1. *Let (P, ω) be a reductive Cartan geometry on M modeled on $(\mathfrak{g}, \mathfrak{h})$ with group H so that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ is an H module decomposition. Let $\omega_{\mathfrak{h}}$ be the $\mathfrak{h}$ projection of ω. Then $\omega_{\mathfrak{h}}$ is an Ehresmann connection on P and P has a canonical (up to a choice of basis for $\mathfrak{p}$) interpretation as a frame bundle with group H.*

[3]If we take the raw, untranslated identification, we get the Cartan connection on the surface. See also the remarks at the end of §3.

Proof. Let us decompose the Cartan connection as $\omega = \omega_{\mathfrak{h}} + \omega_{\mathfrak{p}}$.

Proof of Property **6.2.4**(i) *of an Ehresmann connection:* By Definition **5.3.1**(c)(ii), for a Cartan connection we have $R_h^*\omega = \mathrm{Ad}(h^{-1})\omega$. Hence

$$\begin{aligned} R_h^*\omega_{\mathfrak{h}} + R_h^*\omega_{\mathfrak{p}} &= \mathrm{Ad}(h^{-1})\omega_{\mathfrak{h}} + \mathrm{Ad}(h^{-1})\omega_{\mathfrak{p}}, \quad \text{or}\\ R_h^*\omega_{\mathfrak{h}} - \mathrm{Ad}(h^{-1})\omega_{\mathfrak{h}} &= -R_h^*\omega_{\mathfrak{p}} + \mathrm{Ad}(h^{-1})\omega_{\mathfrak{p}}. \end{aligned}$$

Since the left side takes values in $\mathfrak{h}$ and the right side takes values in $\mathfrak{p}$, both sides vanish, and so

$$R_h^*\omega_{\mathfrak{h}} = \mathrm{Ad}(h^{-1})\omega_{\mathfrak{h}}.$$

Proof of Property (ii) of an Ehresmann connection. By Definition **5.3.1**(c)(iii), for a Cartan connection we have $\omega(X^\dagger) = X$ for every $X \in \mathfrak{h}$. Hence

$$\begin{aligned} X = \omega(X^\dagger) &= \omega_{\mathfrak{h}}(X^\dagger) + \omega_{\mathfrak{p}}(X^\dagger), \quad \text{or}\\ X - \omega_{\mathfrak{h}}(X^\dagger) &= \omega_{\mathfrak{p}}(X^\dagger). \end{aligned}$$

Since the left side takes values in $\mathfrak{h}$ and the right side takes values in $\mathfrak{p}$, both sides vanish so

$$\omega_{\mathfrak{h}}(X^\dagger) = X.$$

To interpret P as a frame bundle, we refer to the procedure outlined in Exercise **5.3.21**. ■

There is a kind of converse to this result that depends on the following notion of the fundamental forms on the principal frame bundle of a manifold.

Definition 2.2. Let $Gl_n(\mathbf{R}) \to P \to M^n$ be the principal $Gl_n(\mathbf{R})$ bundle of frames on M. The *fundamental*[4] forms $\theta_1, \theta_2, \ldots, \theta_n$ are the forms on P defined as follows. A point $p \in P$ consists of a pair (x, e), where $e = (e_1, e_2, \ldots, e_n)$ is a frame at x (i.e., a basis of $T_x(M)$). Thus, for $v \in T_p(P)$, we may write $\pi_{*p}(\nu) = \theta_1(\nu)e_1 + \theta_2(\nu)e_2 + \cdots + \theta_n(\nu)e_n$, and this defines

[4]Such forms have also been called "canonical" or "solder" forms. The terminology depends on the context. When we are given the frame bundle of a manifold, the correct terminology is "fundamental" or "canonical" forms since, as described in the definition, they are canonically determined on the frame bundle. On the other hand, if we are given a principal bundle, then the "solder forms" are not canonical, but once they are given they allow us to uniquely identify this principal bundle with a bundle of frames so that the solder forms correspond to the restriction of the canonical forms. Thus, the solder forms allow us to "solder" (in the sense of "identify") the two bundles.

the fundamental forms. Moreover, given any reduction $H \to Q \to M$, the restrictions of the fundamental forms to Q are called the *fundamental forms* of the frame bundle Q. ❀

Proposition 2.3. *Let $H \to P \to M^n$ be any reduction of the principal $Gl_n(\mathbf{R})$ frame bundle associated to the tangent bundle of M. Let $\mathfrak{h}$ be the Lie algebra of H and let γ be an Ehresmann connection on P. Let $\mathfrak{p} = \mathbf{R}^n$ be the standard reprsentation of $Gl_n(\mathbf{R})$ restricted to H, and let $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{p}$ be the Lie algebra for which $\mathfrak{h}$ is a subalgebra, the action of $\mathfrak{h}$ on $\mathfrak{p}$ is the standard adjoint action, and $[\mathfrak{p}, \mathfrak{p}] = 0$. Let $\theta = (\theta_1, \theta_2, \ldots, \theta_n)^t$ denote the $\mathfrak{p}$-valued 1-form on P determined by the fundamental forms. Then*

$$\omega = \begin{pmatrix} 0 & 0 \\ \theta & \gamma \end{pmatrix}$$

is a Cartan connection on P modeled on the pair $(\mathfrak{g}, \mathfrak{h})$ with group H.

Proof. It suffices to verify condition **5**.3.1(c).

First we verify the condition **5**.3.1(c)(iii). By Definition **6**.2.4(ii), we have $\gamma(X^\dagger) = X$ for every $X \in \mathfrak{h}$. Since $X^\dagger$ is tangent to the fiber, by Definition 2.2, $\theta_p(X^\dagger) = 0$. Thus, $\omega_p(X^\dagger) = X$ for all $X \in \mathfrak{h}$.

Next we verify the condition **5**.3.1(c)(i), which says that $\omega_p: T_p(P) \to \mathfrak{g}$ is an isomorphism for each $p \in P$. Since the dimensions are the same, it suffices to show surjectivity. By the verification of condition **5**.3.1(c)(iii) above, the algebra $\mathfrak{h}$ lies in the image of ω_p. On the other hand, ker γ_p has dimension $n = \dim M$ and is complementary to the tangent space of the fiber at p. Thus, by Definition 2.2, $\theta_p \mid \ker \gamma_p: \ker \gamma_p \to \mathfrak{p}$ is an isomorphism and hence $\omega_p: T_p(P) \to \mathfrak{g}$ is surjective.

Finally we verify the condition **5**.3.1(c)(ii), which says that $(R_h)^*\omega = \mathrm{Ad}(h^{-1})\omega$ for all $h \in H$. The right action of H on P is given by the formula

$$(e_1, e_2, \ldots, e_n) \cdot h = \left(\sum_{1 \le i \le n} h_{i1} e_i, \sum_{1 \le i \le n} h_{i2} e_1, \ldots, \sum_{1 \le i \le n} h_{in} e_i) \right),$$

where $h \in H \subset Gl_n(\mathbf{R})$. Now, writing $p = (e_1, e_2, \ldots, e_n) \in P$, the defining equation $\pi_{*p}(\nu) = \theta_1(\nu)e_1 + \cdots + \theta_n(\nu)e_n$ for the fundamental form yields

$$\begin{aligned} &\theta_1(\nu)e_1 + \cdots + \theta_n(\nu)e_n \\ &= \pi_{*p}\nu \\ &= (\pi \circ R_h)_{*p}\nu \quad (\text{since } \pi \circ R_h = \pi) \\ &= \pi_{*ph}(R_{h*}\nu) \\ &= \theta_1(R_{h*}\nu)(e_1 \cdot h) + \cdots + \theta_n(R_{h*}\nu)(e_n \cdot h) \\ &= \sum_{1 \le i,j \le n} \theta_j(R_{h*}\nu) h_{ij} e_i \end{aligned}$$

$$= \sum_{1 \le j \le n} \theta_j(R_{h*}\nu) h_{1j} e_1 + \cdots + \sum_{1 \le j \le n} \theta_j(R_{h*}\nu) h_{nj} e_n,$$

from which it follows that $\theta_i(\nu) = \sum_{1 \le j \le n} \theta_j(R_{h*}\nu) h_{ij}$ and hence that $R_h^*\theta = h^{-1}\theta$. Thus,

$$\begin{aligned} R_h^*\omega &= R_h^*\gamma + R_h^*\theta \\ &= \mathrm{Ad}(h^{-1})\gamma + h^{-1}\theta \quad (\text{the first term comes from Definition } \mathbf{6}.2.4(\mathrm{i})) \\ &= \mathrm{Ad}(h^{-1})\omega. \end{aligned}$$ ■

Exercise 2.4. Verify that the correspondences given in Propositions 2.1 and 2.3 are inverse to each other and so complete the verification of the equivalence given at the beginning of §2. ❑

§3. Ehresmann Connections Generalize Cartan Connections

Suppose we are given a Cartan geometry (P, ω) on M modeled on the Klein geometry (G, H) with Lie algebras $(\mathfrak{g}, \mathfrak{h})$. From the right principal H bundle $H \to P \to M$ and the action of H on G by left multiplication, we obtain the associated right principal G bundle $G \to Q \to M$ and we have the canonical inclusion $P \subset Q = P \times_H G$ sending $p \mapsto (p, e)$. We are going to show that certain Ehresmann connections on Q restrict to give all Cartan connections on P.

Proposition 3.1. *Let (G, H) be a Klein geometry and let P and Q be principal H and G bundles, respectively, over a manifold M. Assume that $\dim G = \dim P$ and that $\varphi: P \to Q$ is an H bundle map. Then the correspondence*

$$\left\{ \begin{array}{c} \text{Ehresmann connections} \\ \text{on } Q \text{ whose kernels do} \\ \text{not meet } \varphi_*(T(P)) \end{array} \right\} \xrightarrow{\varphi^*} \left\{ \begin{array}{c} \text{Cartan} \\ \text{connections} \\ \text{on } P \end{array} \right\}$$

is a bijection of sets.

Proof. *The correspondence* $\varpi \mapsto \omega = \varphi^*\varpi$. We must show that ω is a Cartan connection.

Since $\varphi_*(T(P)) \cap \ker \varpi = 0$, it follows that $\omega = \varphi^*\varpi$ is a $\mathfrak{g}$-valued 1-form on P with no kernel.

First we verify condition **5**.3.1(c)(i). Since $\dim P = \dim \mathfrak{g}$ and $\omega_p: T_p(P) \to \mathfrak{g}$ is injective, it follows that it is an isomorphism.

Next we verify condition **5**.3.1(c)(iii). Since $\varphi: P \to Q$ is an H bundle map, it follows from Exercise **3**.2.15(b) that for all $\nu \in \mathfrak{h}$, the vector fields

$X = \nu^\dagger$ on P (calculated via the right H action on P) and $Y = \nu^\dagger$ on Q (calculated via the right G action on Q) are φ related (i.e., $\varphi_{*p}(X_p) = Y_{\varphi(p)}$ for all $p \in P$). Thus $\omega(X_P) = \varphi^*\varpi(X_p) = \varpi(\varphi_*(X_p)) = \varpi(Y_{\varphi(p)}) = \nu$ for all $\nu \in \mathfrak{g}$.

Finally, we verify condition **5**.3.1(c)(ii). We have

$$\begin{aligned} R_h^*\omega &= R_h^*\varphi^*\varpi = (\varphi R_h)^*\varpi = (R_h\varphi)^*\varpi = \varphi^* R_h^*\varpi \\ &= \varphi^*\mathrm{Ad}(h^{-1})\varpi = \mathrm{Ad}(h^{-1})\varphi^*\omega. \end{aligned}$$

Thus, $\omega = \varphi^*\varpi$ is indeed a Cartan connection.

The correspondence $\omega \mapsto \varpi = j(\omega)$. Given a Cartan connection ω on P, we may extend it to a form $\varpi = j(\omega)$ on $P \times G$ by means of the formula

$$\varpi_{(p,g)} = \mathrm{Ad}(g^{-1})\pi_p^*\omega + \pi_G^*\omega_G,$$

where $\pi_p\colon P \times G \to P$ and $\pi_G\colon P \times G \to G$ are the canonical projections.

We are going to show first that ϖ is the pull-up of an Ehresmann connection on $P \times_H G$ whose kernel does not meet $\phi_*(T(P))$. This will show that j is a map

$$\left\{\begin{matrix}\text{Cartan} \\ \text{connections} \\ \text{on } P\end{matrix}\right\} \xrightarrow{j} \left\{\begin{matrix}\text{Ehresmann connections} \\ \text{on } Q \text{ whose kernels do} \\ \text{not meet } \varphi_*(T(P))\end{matrix}\right\},$$

which means that it is a possible inverse for the map φ^*.

Note that the form ϖ satisfies the condition $\varpi(0 \times \nu^\dagger) = \nu$ for all $\nu \in \mathfrak{g}$. This verifies condition **6**.2.4(ii) for (the pull-up to $P \times G$ of) an Ehresmann connection on $P \times_H G$. We also see that ϖ restricts to ω on $P \times e$. In particular, ϖ does not vanish on $T(P \times e)$.

Now we study the effect on ϖ of right multiplication by $\gamma \in G$ on the second factor of $P \times G$, which we denote by $\mathrm{id} \times R_\gamma$. Under this map, ϖ pulls back by

$$\begin{aligned} (\mathrm{id} \times R_\gamma)^*\varpi_{(p,g\gamma)} &= \varpi_{(p,g\gamma)} \circ (\mathrm{id} \times R_\gamma)_* \\ &= (\mathrm{Ad}(g\gamma)^{-1}\pi_p^*\omega + \pi_G^*\omega_G) \circ (\mathrm{id} \times R_\gamma)^* \\ &= \mathrm{Ad}(g\gamma)^{-1}\omega(\pi_{p*} \circ (\mathrm{id} \times R_\gamma)_*) + \omega_G(\pi_{G*} \circ (\mathrm{id} \times R_\gamma)_*) \\ &= \mathrm{Ad}(g\gamma)^{-1})\omega(\pi_{p*}) + \omega_G(R_{\gamma*} \circ \pi_{G*}) \\ &= \mathrm{Ad}(g\gamma)^{-1}\pi_p^*\omega + \mathrm{Ad}(\gamma^{-1})\omega_G(\pi_{G*}) \\ &= \mathrm{Ad}(\gamma)^{-1}(\mathrm{Ad}(g)^{-1}\pi_p^*\omega + \pi_G^*\omega_G) \\ &= \mathrm{Ad}(\gamma)^{-1}\varpi. \end{aligned}$$

This verifies condition **6**.2.4(i) for (the pull-up to $P \times G$ of) an Ehresmann connection on $P \times_H G$.

Finally, let us verify that ϖ is basic for the H bundle map $P \times G \to P \times_H G$ (i.e., that it is the pull-up of a form on $P \times_H G$). By Lemma **1**.5.21, it suffices to show that ϖ is invariant under the maps

$$\begin{aligned} \alpha_h\colon P \times G &\to P \times G \\ (p,g) &\mapsto (ph, h^{-1}g) \end{aligned}$$

and vanishes in the fiber directions of $P \times G \to P \times_H G$. For the former condition, we calculate

$$\begin{aligned} (\alpha_h^*\varpi)_{(p,g)} &= \varpi_{(ph,h^{-1}g)} \circ \alpha_{h*} \\ &= \mathrm{Ad}(h^{-1}g)^{-1}\pi_P^*\omega \circ \alpha_{h*} + \pi_G^*\omega_G \circ \alpha_{h*} \\ &= \mathrm{Ad}(h^{-1}g)^{-1}\omega \circ \pi_{P*} \circ \alpha_{h*} + \omega_G \circ \pi_{G*} \circ \alpha_{h*} \\ &= \mathrm{Ad}(h^{-1}g)^{-1}\omega \circ R_{h*} \circ \pi_{P*} + \omega_G \circ L_{h^{-1}*} \circ \pi_{G*} \\ &= \mathrm{Ad}(h^{-1}g)^{-1}\mathrm{Ad}(h^{-1})\omega \circ \pi_{P*} + \omega_G \circ \pi_{G*} \\ &= \mathrm{Ad}(g)^{-1}\pi_P^*\omega + \pi_G^*\omega_G = \varpi_{(p,g)}. \end{aligned}$$

Now consider the latter condition. Fix $\nu \in \mathfrak{h} \subset \mathfrak{g}$. Let $\nu^\dagger$ be the vector on $P \times G$ corresponding to ν with respect to the right H action on $P \times G$ given by

$$\begin{aligned} (\nu^\dagger)_{(p,g)} &= ((\mu_P \times \mu_G) \circ (\mathrm{id} \times \mathrm{id} \times \iota \times \mathrm{id}) \circ (\mathrm{id} \times \Delta \times \mathrm{id}) \circ \rho)_{*(p,g,e)}(0,0,\nu) \\ &= (\mu_P \times \mu_G)_* \circ (\mathrm{id} \times \mathrm{id} \times \iota \times \mathrm{id})_* \circ (\mathrm{id} \times \Delta \times \mathrm{id})_* \circ \rho_{*(p,g,e)}(0,0,\nu) \\ &= (\mu_P \times \mu_G)_* \circ (\mathrm{id} \times \mathrm{id} \times \iota \times \mathrm{id})_{*(p,e,e,g)}(0,\nu,\nu,0) \\ &= (\mu_P \times \mu_G)_{*(p,e,e,g)}(0,\nu,-\nu,0) \\ &= (\mu_{P*(p,e)}(0,\nu), \mu_{G*(e,g)}(-\nu,0)) \\ &= (\omega^{-1}(\nu)_p, -\omega^{-1}(\mathrm{Ad}(g)\nu)_g), \end{aligned}$$

so that

$$\begin{aligned} \varpi_{(p,g)}(\nu^\dagger) &= \varpi_{(p,g)}(\omega^{-1}(\nu), -\omega_G^{-1}(\mathrm{Ad}(g^{-1})\nu)) \\ &= \mathrm{Ad}(g^{-1})\pi_P^*\omega(\omega^{-1}(\nu), -\omega_G^{-1}(\mathrm{Ad}(g^{-1})\nu)) \\ &\quad + \pi_G^*\omega_G(\omega^{-1}(\nu), -\omega_G^{-1}(\mathrm{Ad}(g^{-1})\nu)) \\ &= \mathrm{Ad}(g^{-1})\nu - (\mathrm{Ad}(g^{-1})\nu) \\ &= 0. \end{aligned}$$

Since the vector fields $\nu^\dagger$ span the space of fiber directions, this finishes the proof of the second condition. Thus, ϖ does pass down to yield a $\mathfrak{g}$-valued form on $P \times_H G$. The resulting form obviously satisfies conditions (i) and (ii) for an Ehresmann connection as well as the condition that $\ker \varpi$ meets $\varphi_*(T(P))$ only in zero.

The correspondences φ^* *and* j *are inverse to each other.* To see this, we first calculate

$$\begin{aligned}\varphi^*(j(\omega_p)) &= \varphi^*\varpi_{(p,e)}\\ &= \operatorname{Ad}(e^{-1})\varphi^*\pi_P^*\omega_p + \varphi^*\pi_G^*\omega_{Ge}\\ &= (\pi_P\circ\varphi)^*\omega_p + 0\\ &= \omega_p,\end{aligned}$$

showing that the composite $\varphi^* j$ is the identity.

On the other hand, φ^* is injective, for if $\varphi^*(\varpi_1) = \varphi^*(\varpi_2)$, then ϖ_1 and ϖ_2 agree on the image $\varphi_*(T(P))$ and hence on all the right translates $R_{g*}\varphi_*(T(P))$. But they also agree on the vector fields $\nu^\dagger$ for all $\nu \in \mathfrak{g}$, and these two kinds of vectors span the whole of $T(Q)$.

Thus, φ^* and j are inverse to each other. ■

Two Notions of Development

Suppose the manifold M is equipped with a Cartan geometry modeled on the Klein pair (G, H). Let $\sigma: I \to M$ be a path on M. In §4 of Chapter 5, we showed how to develop the path σ to a path $\tilde{\sigma}: I \to G/H$ on the model space G/H. The same methods showed how to associate to a loop σ on M the (Cartan) holonomy of σ which is an element of H. For an Ehresmann connection on a principal bundle over M, there is, in general, no analog of the development of a path on M. However, there is always an analog of the holonomy of a loop. It is this concept that we study here.

Our study depends on the following generalization of the notion of horizontal vectors given in Definition **5**.3.44.

Definition 3.2. Let $H \to P \to M$ be a principal H bundle over M, let $\mathfrak{h}$ be the Lie algebra of H, and let γ be an Ehresmann connection on P. A vector $\nu \in T_p(P)$ is *horizontal* if $\gamma(\nu) = 0$. A path $\sigma: I \to P$ is *horizontal* if $\sigma^*\gamma = 0$ (i.e., the tangent vectors of σ are all horizontal). ❀

Exercise 3.3.* Let M be equipped with a principal bundle Q and an Ehresmann connection on it.

(i) Show that the map $\pi_{*p}: T_p(P) \to T_x(M)$ induced by the projection $\pi: P \to M$ restricts to give an isomorphism between the horizontal vectors at $p \in P$ and $T_x(M)$ (cf. Lemma **5**.3.45).

(ii) Fix $x_0 \in M$ and fix a point $q_0 \in Q$ lying over it. Show that every path on M starting at x_0 is covered by a unique horizontal path on Q starting at q_0 (cf. the proof of Proposition B.2.3). ❏

Definition 3.4. Let $Q \to M$ be a principal G bundle equipped with an Ehresmann connection and let $c: (I, \partial I) \to (M, x_0)$ be a closed path on M. Fix $q_0 \in Q$ lying over x_0, and let $\tilde{c}: (I, 0, 1) \to (Q, q_0, q_1)$ be the unique horizontal lift of c starting at q_0. Since q_0 and q_1 lie in the same fiber of Q, we may write $q_0 = q_1 g$ for some $g \in G$. The (*Ehresmann*) *holonomy* of c with respect to q_0 is g. ❀

The behavior of this notion of holonomy is somewhat different than that of the notion of Cartan geometries given in Definition **5**.4.18. Here is a table comparing these behaviors of holonomies of loops on M.

Comparison of Holonomy of Loops

Cartan Holonomy	*Ehresmann Holonomy*
• Has values in G	• Has values in H
• Varies by conjugacy according to the starting point of the lift.	• Varies by conjugacy according to the starting point of the lift.
• Only applies to loops on M that lift to loops on P.	• Applies to all loops on M.

Proposition 3.5. *Assume the context of Definition* 3.4. *If the Ehresmann holonomy of c with respect to q_0 is g, then the Ehresmann holonomy of c with respect to q_0h is $h^{-1}gh$.*

Proof. Let $\tilde{c}: (I, 0, 1) \to (Q, q_0, q_1)$ be the unique horizontal lift of c starting at q_0. We claim that the unique horizontal lift of c starting at q_0h is $R_h \circ \tilde{c}: (I, 0, 1) \to (Q, q_0h, q_1h)$. Clearly, this is a lift of c, but it is also horizontal since

$$(R_h \circ \tilde{c})^*\gamma = \tilde{c}^* R_h^*\gamma = \tilde{c}^*\operatorname{Ad}(h^{-1})\gamma = \operatorname{Ad}(h^{-1})\tilde{c}^*\gamma = 0$$

(since $\tilde{c}$ is horizontal). But $q_0h = q_1gh = q_1h(h^{-1}gh)$, so $h^{-1}gh$ is the holonomy of c starting at q_0h. ■

Since the Ehresmann connection may be regarded as a generalization of the Cartan connection in two distinct ways, we may compare the notions of holonomy in each of these two ways.

We study first the relation between the holonomy of a Cartan connection and the Ehresmann connection derived from it via the procedure of Proposition 3.1.

Proposition 3.6. *Let (P, ω) be a Cartan geometry on M modeled on (G, H) and let (Q, ϖ) be the corresponding Ehresmann connection as given in Proposition* 3.1. *Fix a point $x_0 \in M$ and a loop $c: (I, \partial I) \to (M, x_0)$ that lifts to a loop based at $q_0 \in P_0$. Then the Cartan and Ehresmann holonomies of c are the same.*

Proof. Let $\hat{c}: (I, 0, 1) \to (P \times_H G, (p_0, e), (p_0, g))$ be any horizontal lift of c and let $\hat{\hat{c}} = (c_1, c_2): (I, 0, 1) \to (P \times G, (p_0, e), (p_0, g))$ be a lift of $\hat{c}$ to $P \times G$. By Definition 3.4, the Ehresmann holonomy of c is g^{-1}.

On the other hand, according to the proof of Proposition 3.1,

$$0 = \hat{c}^*\varpi_{(p,g)} \Rightarrow 0 = \mathrm{Ad}(g^{-1})\hat{\hat{c}}^*\pi_P^*\omega + \hat{\hat{c}}^*\pi_G^* = \mathrm{Ad}(c_2^{-1})c_1^*\omega + c_2^*\omega_G,$$

so that by Proposition **3**.4.10(ii)

$$c_1^*\omega = -\mathrm{Ad}(c_2)c_2^*\omega_G = (c_2^{-1})^*\omega_G.$$

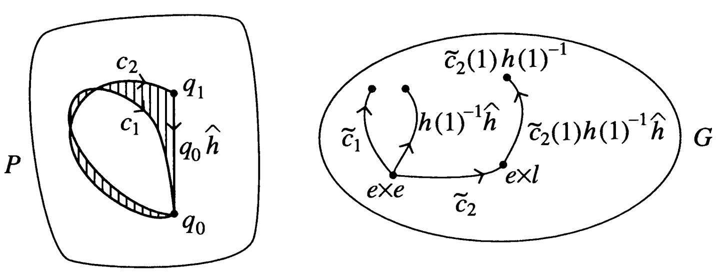
Since c_1 is a lift to P of c, it follows (cf. Definition **3**.7.2) that c_2^{-1} is the development of c in the Cartan sense, so that the Cartan holonomy of c is $c_2^{-1}(1) = g^{-1}$. Thus, the two holonomies are the same. ■

Finally, we study a simple case of the relation between the holonomy of a Cartan connection and an Ehresmann connection when the two connections are related as in §2.

Proposition 3.7. *Let (P,ω) be such a Cartan geometry on M modeled on $(G = H \times_\rho L, H)$, where $H \times_\rho L$ is a semidirect product (with L the normal subgroup). Fix a point $x_0 \in M$ and a loop $c\colon (I,\partial I) \to (M,x_0)$ that lifts to a loop based at $q_0 \in P_0$ (lying over x_0). Let $\mathfrak{g}$, $\mathfrak{h}$, and $\mathfrak{p}$ be the Lie algebras of G, H, and L, respectively. Decompose the Cartan connections as $\omega = \omega_{\mathfrak{h}} + \omega_{\mathfrak{p}}$, so that (by Lemma 2.1) $\omega_{\mathfrak{h}}$ is an Ehresmann connection on P. Then the Ehresmann holonomy of c is the H component of the Cartan holonomy of c.*

Proof. On the one hand, we may fix a lift $c\colon (I,0,1) \to (P,q_0,q_0)$ of c to a closed loop on P. On the other hand, by Exercise 3.3(ii), the loop c also has a unique horizontal lift $c_2\colon (I,0,1) \to (Q,q_0,q_1)$ starting at q_0.

Because c_1 and c_2 are both lifts of the same loop c, there is a unique map $h\colon (I,0) \to (H,e)$ such that $c_2(t) = c_1(t)h(t)$. Let $\hat{h}(t) = h(1-t)$. We have $q_1 = c_2(1) = c_1(1)h(1) = q_0h(1)$. Thus, $h(1)^{-1}$ is the Ehresmann holonomy of c based at q_0. On the other hand, the Cartan holonomy of c is obtained as the endpoint of the development, starting at the identity, of any loop on P based at q_0 and covering the loop c; for instance, the loop $(q_0\hat{h}) \star c_2$. The development of this loop is obtained (cf. Exercise **3**.7.5(b)) by path multiplication of the development of c_2 with the development of $q_0\hat{h}$.



Now $c_2^*\omega = c_2^*\omega_{\mathfrak{p}}$ (here we use the fact that $c_2^*\omega_{\mathfrak{h}} = 0$, which holds since c_2 is a horizontal curve). Thus, $c_2^*\omega$ is a form with values in $\mathfrak{p} \subset \mathfrak{g}$, so its development on G, starting at $e \times e \in H \times_\rho L$, lies entirely on the subgroup L and hence ends at an element of the form $e \times l \in H \times_\rho L$. On the other hand, the curve $q_0\hat{h}$ on P develops (cf. Lemma **5**.4.12) to the curve, starting at $e \times e \in H \times_\rho L$, given by $h(1)^{-1}\hat{h}$ and hence it also develops to the curve, starting at $e \times l \in H \times_\rho L$, given by $(e \times l)h(1)^{-1}\hat{h}$. Thus, the development of $(q_0\hat{h}) \star c_2$, starting at $e \times e \in H \times_\rho L$, ends at

$$(e \times l)h(1)^{-1}\hat{h}(1) = h(1)^{-1} \times \rho(h(1))(l).$$

This is the Cartan holonomy and its H component is the Ehresmann holonomy. ■

It is interesting to compare this result to the case of rolling without slipping or twisting described in Appendix B. For a 2-sphere rolling on a plane along a closed loop on the sphere, the initial and final points of contact are the same on the sphere but differ by a translation on the plane. The total effect on the sphere is to undergo a two-dimensional rotation and translation. The rotation part is the Ehresmann holonomy of the loop on the sphere, whereas the rotation and translation are the Cartan holonomy of the loop on the sphere.

§4. Covariant Derivative

In Chapter 5 we discussed the covariant derivative associated to a reductive Cartan geometry (cf. Definition **5**.3.46 through Exercise **5**.3.52). In this section we show this notion generalizes to arbitrary Ehresmann connections (Proposition 4.4). As in Definition **5**.3.46, the definition of covariant derivative depends on the notion of a horizontal vector given, in the present context, in Definition 3.2.

Definition 4.1. Let

(i) $G \to Q \to M$ be a principal bundle over M,

(ii) γ be an Ehresmann connection on Q,

(iii) E be the vector bundle associated to a representation (V,ρ) of G (i.e., $E = Q \times_G V$), and

(iv) Y be a tangent vector field on M and $\tilde{Y}$ its horizontal lift to Q.

The *covariant derivative* $D_Y\colon \Gamma(E) \to \Gamma(E)$ associated to these data is defined by the following diagram.

$$
\begin{tikzcd}
A^0(Q,\rho) \arrow[r, "Y"] \arrow[d, "\psi"', "\approx"] & A^0(Q,\rho) \arrow[d, "\approx"', "\psi"] \\
\Gamma(E) \arrow[r, "D_Y X"] & \Gamma(E)
\end{tikzcd}
$$

(Cf. Exercise **1.3.28** for a description of the isomorphism ψ.) ❀

In preparation for the calculation given in Proposition B.3.7, it is useful to reformulate the definition of covariant derivative for the case where Q can be interpreted as a frame bundle of E.

Proposition 4.2. *When the representation (V, ρ) in Definition* 4.1 *is faithful, then*

(i) *given a fixed basis $\bar{e}_1, \dots, \bar{e}_\nu$ for V, there is a canonical interpretation of Q as the bundle of frames of E given by*

$$q \in Q \mapsto (e_1(q), \dots, e_\nu(q)), \quad \text{where } e_i(q) = (q, \bar{e}_i) \in Q \times_G V = E,$$

(ii) *the covariant derivative $D_Y X$, where $X \in \Gamma(E)$, may be calculated as follows. Express the vector field X in terms of the basis $e_i(q)$ as $X = a_1(q)e_1(q) + \cdots + a_\nu(q)e_\nu(q)$. Let $\tilde{Y}_q = T_q(Q)$ denote the horizontal lift of Y_x. Then $(D_Y X)_x = \tilde{Y}_q(a_1)e_1(q) + \cdots + \tilde{Y}_q(a_\nu)e_\nu(q)$.*

Proof. (i) For this we refer the reader to the discussion on page 37.

(ii) Define $f: Q \to V$ by $f(q) = \sum_{1 \le j \le \nu} a_j(q)\bar{e}_j$. We claim that $f \in A^0(Q, \rho)$. To see this, first note that

$$
\begin{aligned}
e_j(qh) &= (qh, \bar{e}_j) \in Q \times_G V \\
&= (q, \rho(h)\bar{e}_j) \\
&= \sum_k \rho(h)_{jk}(q, \bar{e}_k).
\end{aligned}
$$

Thus,

$$\sum_{1 \le j \le \nu} a_j(q)e_j(q) = X = \sum_{1 \le j \le \nu} a_j(qh)e_j(qh) = \sum_{1 \le j,k \le \nu} a_j(qh)\rho(h)_{jk}e_k(q)$$

and hence

$$a_k(q) = \sum_{1 \le j \le \nu} a_j(qh)\rho(h)_{jk},$$

or

$$a_j(qh) = \sum_{1 \le k \le \nu} a_k(q)\rho(h^{-1})_{kj}.$$

It follows that

$$
\begin{aligned}
f(qh) &= \sum_{1 \le j \le \nu} a_j(qh)\bar{e}_j = \sum_{1 \le j,k \le \nu} a_k(q)\rho(h^{-1})_{kj}\bar{e}_j \\
&= \rho(h^{-1}) \sum_{1 \le j,k \le \nu} a_k(q)\bar{e}_k = \rho(h^{-1})f(q).
\end{aligned}
$$

Next we claim that $\psi(f) = X$, which follows from the fact that $\psi(f) \in \Gamma(E) = \Gamma(Q \times_G V)$ is induced by the map $Q \to Q \times V$ sending

$$
\begin{aligned}
q \mapsto (q, f(q)) &= \left(q, \sum_{1 \le j \le \nu} a_j(q)\bar{e}_j\right) = \sum_{1 \le j \le \nu} a_j(q)(q, \bar{e}_j) \\
&= \sum_{1 \le j \le \nu} a_j(q)\bar{e}_j(q) = X.
\end{aligned}
$$

Now, according to Definition 4.1 we have $(D_Y X)_x = \psi(\tilde{Y}_q(f))$, and the latter is induced by the map $Q \to Q \times V$ sending

$$
\begin{aligned}
q \mapsto (q, \tilde{Y}_q(f)) &= \left(q, \tilde{Y}_q\left(\sum_{1 \le j \le \nu} a_j\bar{e}_j\right)\right) = \left(q, \sum_{1 \le j \le \nu} \tilde{Y}_q(a_j)\bar{e}_j\right) \\
&= \sum_{1 \le j \le \nu} \tilde{Y}_q(a_j)(q, \bar{e}_j).
\end{aligned}
$$
■

We note that in the case described in Proposition 4.2, it is clear that the covariant derivative $(D_Y X)_x$ depends only on the value of Y at x. This is also true in general.

Exercise 4.3. (i) Verify that the properties of Proposition **5.3.48** remain true for the covariant derivative of Definition 4.1.

(ii) Verify that, for each $x \in M$, the vector $(D_Y X)_x \in E_x$ (= the fiber of E over x) depends only on the value of Y at x and the values of X along any curve in M tangent to Y at X.

(iii) Strengthen the result of (ii) by showing that the dependence on the curve is actually only a dependence on the first two terms of the Taylor series[5] for the curve at x. ❏

Proposition 4.4. *The covariant derivative associated to a reductive Cartan geometry is the same as the covariant derivative associated to the Ehresmann geometry obtained from it by the procedure of* §2.

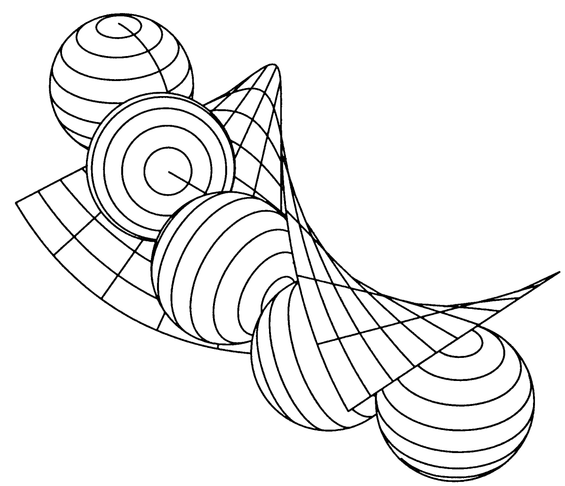
Proof. Let us decompose the Cartan connection as $\omega = \omega_{\mathfrak{h}} + \omega_{\mathfrak{p}}$ as in Lemma 2.1. Then $\omega_{\mathfrak{h}}$ is an Ehresmann connection, and the procedure for associating a covariant derivative to it described in Definitions **5.3.46** and **5.3.47** is identical to that of Definitions 3.2 and 4.1. ■

[5]In any coordinate system.

Appendix B

Rolling Without Slipping or Twisting

In this appendix we study the most classical of all nonholonomic differential systems, the system describing how one manifold may roll without slipping or twisting on another in Euclidean space.



A sphere rolling along the central line of a helicoid without slipping or twisting.

Such a system plays a role, for instance, in robotics, where one wishes to use robotic fingers to manipulate some object. Our interest, however, is to show the relationship of this concrete physical notion to the idea of Levi–Civita connection of a Riemannian manifold and the canonical Ehresmann connection on its normal bundle in Euclidean space.[1]

In §1 we give a definition of the notion of rolling without slipping or twisting in an "elementary" form which appeals directly to the intuition in Euclidean space. In §2 we reformulate this definition in terms of a differential sysem (i.e., a distribution) on a state space. In this formulation we easily obtain the existence and uniqueness of rolling without slipping or twisting. In particular, this yields isometries between the ambient tangent spaces at each point of the rolling. In §3 we show how this notion gives rise to both the Levi–Civita connection and the normal connection for a submanifold of Euclidean space. The section ends by relating the notion of rolling a manifold M^n on $\mathbf{R}^n$ to the notion of development described in Chapter 5. In particular, we show that a curve on M rolls without slipping on a straight line in $\mathbf{R}^n$ if and only if the curve is a geodesic. We also show that a vector field along a curve in M is parallel if and only if its image in $\mathbf{R}^n$ is parallel. This result (Proposition 3.5) makes the notion of the Levi–Civita connection and the normal connection for a submanifold of Euclidean space quite concrete. Finally, in §4, we study what happens when we roll one manifold on a second while the second is itself rolling on a third.

§1. Rolling Maps

In order to study rolling maps we need to have a way to represent elements of $\mathrm{Euc}_N(\mathbf{R})$, the group of isometries of Euclidean N space. If $g \in \mathrm{Euc}_N(\mathbf{R})$, then we may write

$$g\colon \mathbf{R}^n \to \mathbf{R}^n, \quad \nu \mapsto A\nu + p$$

where $A \in O_n(\mathbf{R})$ and $p \in \mathbf{R}^N$. It is easy to verify that the map

$$\mathrm{Euc}_N(\mathbf{R}) \to \left\{ \begin{pmatrix} 1 & 0 \\ p & A \end{pmatrix} \in Gl_{N+1}(\mathbf{R}) \mid A \in O_n(\mathbf{R}), p \in \mathbf{R}^n \right\}$$

is an isomorphism.

A rolling map will be a certain kind of one-parameter family $g\colon I \to \mathrm{Euc}_N(\mathbf{R})$. We write $g(t)v = A(t)v + p(t)$. For fixed $t \in I$, we have $g(t)\colon \mathbf{R}^N \to \mathbf{R}^N$ with derivative $g(t)_* = A(t)\colon \mathbf{R}^N \to \mathbf{R}^N$; here "$*$" refers to the partial derivative with respect to the space coordinates. We write the partial derivative of $g(t)$ with respect to the parameter t as $\dot{g}(t)\colon \mathbf{R}^n \to \mathbf{R}^n$. In particular, the composite mapping $\dot{g}(t)g(t)^{-1}\colon \mathbf{R}^N \to \mathbf{R}^N$ makes sense.

[1]Both of these connections are described in Chapter 6. See also [C. Dodson and T. Poston, 1991], pp. 207ff., for a treatment of the tangential part of this result. We also refer the reader to [R.L. Bryant and L. Hsu, 1993], p. 456, for a more advanced study of aspects of this kind of differential system.

The following definition formalizes the intuitive notion of one manifold rolling on another in Euclidean space without slipping or twisting. The subsequent interpretation is meant to explain the sense in which it does so.

Definition 1.1. Let M_0^n, $M_1^n \subset \mathbf{R}^N$ (where $N = n + r$) be submanifolds. Then a map $g\colon I \to \mathrm{Euc}_N(\mathbf{R})$ satisfying the following properties for each $t \in I$ is called a *rolling of* M_1 *on* M_0 *without slipping or twisting.* More briefly, we shall call g a *rolling map.*

The "rolling" condition:

(1) There is a piecewise smooth "rolling curve on M_1" $\sigma_1\colon I \to M_1$ such that

  (a) $g(t)\sigma_1(t) \in M_0$, and

  (b) $T_{g(t)\sigma_1(t)}(g(t)M_1) = T_{g(t)\sigma_1(t)}(M_0)$ for all t.

The curve $\sigma_0\colon I \to M_0$ defined by $\sigma_0(t) = g(t)\sigma_1(t)$ is called the *development*[2] *of* σ_1 *on* M_0.

The "no-slip" condition:

(2) $\dot{g}(t)g(t)^{-1}\sigma_0(t) = 0$ for all t.

The tangential part of the "no-twist" condition:

(3) $(\dot{g}(t)g(t)^{-1})_* T_{\sigma_0(t)}(M_0) \subset T_{\sigma_0(t)}(M_0)^\perp$ for all t.

The normal part of the "no-twist" condition:

(4) $(\dot{g}(t)g(t)^{-1})_* T_{\sigma_0(t)}(M_0)^\perp \subset T_{\sigma_0(t)}(M_0)$ for all t. ❀

Interpretation

Condition (1) says that M_1 moves under the action of $g(t)$ so as to be tangent to M_0 at the point $\sigma_0(t) = g(t)\sigma_1(t)$ at time t. Condition (2) says that the infinitesimal isometry $\dot{g}(t)g(t)^{-1}$ fixes $\sigma_0(t)$, namely, that it is an infinitesimal rotation about the point $\sigma_0(t)$. This is the "no-slip" condition. Next we consider conditions (3) and (4). Each tangent vector $w \in T_q(M_1)$ (respectively, normal vector $w \in T_q(M_1)^\perp$), as it is carried along by the motion $g(t)$, determines a one-parameter family of vectors $w_t = g(t)_* w \in T_{g(t)q}(g(t)M_1)$ (respectively, $T_{g(t)q}(g(t)M_1)^\perp$) with velocity $\dot{w}_t$. Now suppose $v \in T_{g(t_0)\sigma_1(t_0)}(g(t_0)M_1)$ $(= T_{\sigma_0(t_0)}(M_0))$. Setting $w = g(t_0)_*^{-1}v$, we have $v = w_{t_0} = g(t_0)_* w$ and

[2]See §5 for the relationship of this to the definition of development given in §4 of Chapter 5.

$$\dot{w}_t\big|_{t=t_0} = \dot{g}(t_0)_* w = \dot{g}(t_0)_* g(t_0)_*^{-1} v = (\dot{g}(t_0) g(t_0)^{-1})_* v.$$

Thus, condition (3) says that the motion of a vector $v \in T_{g(t_0)\sigma_1(t_0)}(g(t_0)M_1)$ $(= T_{\sigma_0(t_0)}(M_0))$ that is "stuck to $g(t)M_1$" has no component of velocity in the tangential direction. This is the tangential part of the "no-twist" condition. Condition (4) says that a vector $v \in T_{\sigma(t)}(g(t)M_1)^\perp$, as it is carried along by $g(t)$, has no component of velocity in the normal direction. This is the normal part of the "no twist" condition.

§2. The Existence and Uniqueness of Rolling Maps

The interpretation given in §1 makes it clear that the conditions of Definition 1.1 should be required of a rolling map. What is not clear is the existence and uniqueness, for each curve $\sigma(t)$, of a rolling map with $\sigma(t)$ as its rolling curve.

The most convenient[3] way to deal with the existence and uniqueness question is to reformulate the equations of Definition 1.1 as a differential system, that is, as an n-dimensional distribution on a certain configuration space Σ. Roughly speaking, the configuration space Σ is the space of all positions of M_1 in which it is tangent to M_0. More precisely, we define the configuration space to be

$$\Sigma = \{(p, A, q) \in M_0 \times O_N(\mathbf{R}) \times M_1 \mid AT_q(M_1) = T_p(M_0)\}$$

(where we are identifying $T_p(M_0)$ and $T_q(M_1)$ with subspaces of $\mathbf{R}^N$ in the usual way).

Why should Σ be regarded as the configuration space? First note that there is a canonical smooth map

$$\begin{aligned} \phi\colon \Sigma &\to \mathrm{Euc}_n(\mathbf{R}), \\ (p,A,q) &\mapsto g \end{aligned}$$

where $g = \phi(p, A, q)$ is defined by

$$\begin{aligned} g\colon \mathbf{R}^n &\to \mathbf{R}^n. \\ x &\mapsto A(x-q)+p \end{aligned}$$

Since $gq = p$ and $g_* = A$, it follows that $g_* T_q(M_1) = T_p(M_0)$. Thus we see that a point in Σ determines a point $p \in M_0$ and an isometry $g \in \mathrm{Euc}_N(\mathbf{R})$ such that gM_1 is tangent to M_0 at $p = gq$. This explains the sense in which Σ may be called the configuration space of all positions of M_1 in which it is tangent to M_0.

---

[3] This is instructive as well as convenient, as it introduces the reader to a beautiful application of differential systems.

The first order of business is to determine the tangent bundle of Σ. It is not difficult to verify that the canonical projection $\Sigma \to M_0 \times M_1$ is a fiber bundle with fibers $O_n(\mathbf{R}) \times O_r(\mathbf{R})$. It follows that Σ is a manifold of dimension $2n + \frac{1}{2}n(n-1) + \frac{1}{2}r(r-1)$. We shall also need to use the second fundamental form as it is described in Exercise **1.2.24**. In particular, we shall interpret the second fundamental form as having values in the normal bundle of the manifold on which it is defined.

Proposition 2.1. *Let B_0 and B_1 denote the second fundamental forms for M_0 and M_1, respectively. Let $V \subset T_p(M_0) \times T_A(O_N(\mathbf{R})) \times T_p(M_0)$ be given by*

$$V = \{(\dot{p}, \dot{A}, \dot{q}) \mid \dot{A}A^{-1}u = AB_1(\dot{q}, A\dot{u}) - B_0(\dot{p}, u) \bmod T_p(M_0), \\ \forall u \in T_p(M_0)\}.$$

Then $T_{(p,A,q)}(\Sigma) = V$.

Proof. We know that $(\dot{p}, \dot{A}, \dot{q}) \in T_{(p,A,q)}(\Sigma)$ if and only if it is the derivative at (p, A, q) of a curve $(p(t), A(t), q(t))$ on Σ. Let $u_1(t)$ be a vector field tangent to M_1 defined along $q(t)$. Then, by the definition of Σ, the vector field $u_0(t) = A(t)u_1(t)$ is tangent to M_0 along $p(t)$. Differentiating produces

$$\dot{u}_0(t) = \dot{A}(t)u_1(t) + A(t)\dot{u}_1(t).$$

By Exercise **1.2.24**, we have

$$\dot{u}_0(t) = -B_0(\dot{p}, u_0) \bmod T_p(M_0) \quad \text{and} \quad \dot{u}_1(t) = -B_1(\dot{q}, u_1) \bmod T_q(M_1).$$

Thus, at (p, A, q) we have

$$-B_0(\dot{p}, u_0) = \dot{A}u_1 - AB_1(\dot{q}, u_1) \bmod T_p(M_0), \quad \text{or}$$

$$-B_0(\dot{p}, u) = \dot{A}A^{-1}u - AB_1(\dot{q}, A^{-1}u) \bmod T_p(M_0) \quad \text{for all } u \in T_p(M). \tag{2.2}$$

Thus, $T_{(p,A,q)}(\Sigma) \subset V$.

On the other hand, we claim that $\dim V \le 2n + \frac{1}{2}n(n-1) + \frac{1}{2}r(r-1)$, from which the proposition follows.

To see this, we note that $\dot{A}A^{-1}$ is skew symmetric (since $AA^t = I \Rightarrow \dot{A}A^t + A\dot{A}^t = 0 \Rightarrow \dot{A}A^{-1} + (\dot{A}A^{-1})^t = 0$). If we express the skew-symmetric matrix $\dot{A}A^{-1}$ with respect to an orthonormal basis of $\mathbf{R}^N$ that is adapted to $T_q(M_1)$ (i.e., the first n basis elements span $T_q(M_1)$, which implies that the last r basis elements span $T_q(M_1)^\perp$), then Eq. (2.2) determines the (2,1) block of $\dot{A}A^{-1}$ in terms of $\dot{p}$ and $\dot{q}$, so we have (in the given basis)

$$\dot{A}A^{-1} = \left\{ \begin{matrix} \begin{pmatrix} \text{unknown } n \times n \\ \text{skew-symmetric matrix} \end{pmatrix} & \begin{pmatrix} \text{known} \\ \text{by symmetry} \end{pmatrix} \\ (\text{known}) & \begin{pmatrix} \text{unknown } r \times r \\ \text{skew-symmetric matrix} \end{pmatrix} \end{matrix} \right\}.$$

Thus, we see that specifying the variables $\dot{p} \in T_p(M_0)$ and $\dot{q} \in T_q(M_1)$, each with at most n degrees of freedom, forces $\dot{A}A^{-1}$ to lie in a subspace of dimension $\frac{1}{2}n(n-1) + \frac{1}{2}r(r-1)$. It follows that we have the inequality

$$\dim V \leq 2n + \frac{1}{2}n(n-1) + \frac{1}{2}r(r-1). \qquad \blacksquare$$

A smooth curve $c: I \to \Sigma$ determines a smooth curve $\phi \circ c: I \to \mathrm{Euc}_N(\mathbf{R})$. The next order of business is to determine the conditions on c that will ensure that $\phi \circ c$ is a rolling map.

Lemma 2.3. *$\phi \circ c$ is a rolling map if and only if c is tangent to the n-dimensional distribution $\mathcal{R}$ on Σ given by*

(i) $\dot{p} = A\dot{q}$,

(ii) $\dot{A}A^{-1}u = AB_1(A^{-1}\dot{p}, A^{-1}u) - B_0(\dot{p}, u)$ *for all* $u \in T_p(M_0)$,

(iii) $\dot{A}A^{-1}\nu = -AB_1^t(A^{-1}\dot{p}, A^{-1}\nu) + B_0^t(\dot{p}, \nu)$ *for all* $\nu \in T_p(M_0)^\perp$. *(See Exercise* **1**.2.24(iii) *for the definition of* B^t.)

Moreover, the derivative of the canonical map $\Sigma \to M$ induces the isomorphism shown in the following diagram.

$$\begin{tikzcd}[column sep=tiny] T_{(p,A,q)}(\Sigma) \arrow[dr] & \supset & \mathcal{R}_{(p,A,q)} \arrow[dl, "\approx"] \\ & T_p(M) & \end{tikzcd}$$

Proof. Let $c(t)$ be any curve on Σ. We write $c = (p, A, q)$ and $\dot{c} = (\dot{p}, \dot{A}, \dot{q})$ and we try to express conditions (2), (3), and (4) of Definition 1.1 in terms of $\dot{c}$. (Condition (1) is, of course, built into the very definiton of Σ.) We have

$$gv = A(v-q) + p \quad \text{so} \quad \dot{g}v = \dot{A}(v-q) + \dot{p} - A\dot{q}$$

and

$$g^{-1}v = A^{-1}(v-p) + q,$$

so that

$$\dot{g}g^{-1}v = \dot{A}(A^{-1}(v-p)) + \dot{p} - A\dot{q}$$

and

$$(\dot{g}g^{-1})_* v = \dot{A}A^{-1}v.$$

Since $\dot{g}g^{-1}\sigma = \dot{g}g^{-1}p = \dot{p} - A\dot{q}$, it follows that condition 1.1(2), namely, $\dot{g}g^{-1}\sigma = 0$, is equivalent to $\dot{p} = A\dot{q}$.

Now $(\dot{g}g^{-1})_* = \dot{A}A^{-1}$ implies that condition 1.1(3), namely,

$$(\dot{g}g^{-1})_* T_{\sigma_0(t)}(M_0) \subset T_{\sigma_0(t)}(M_0)^\perp$$

(and, respectively, condition 1.1(4), i.e., $(\dot{g}g^{-1})_* T_p(M_0)^\perp \subset T_p(M_0)$), is equivalent to $\dot{A}A^{-1}T_p(M_0) \subset T_p(M_0)^\perp$ (respectively, $\dot{A}A^{-1}T_p(M_0)^\perp \subset T_p(M_0)$), which by Proposition 2.1, is equivalent to the equation $\dot{A}A^{-1}u = AB_1(A^{-1}\dot{p}, A^{-1}u) - B_0(\dot{p}, u)$ for all $u \in T_p(M_0)$ (respectively, to the equation $\dot{A}A^{-1}v = -AB_1^t(A^{-1}\dot{p}, A^{-1}v) + B_0^t(\dot{p}, v)$ for all $v \in T_p(M_0)^\perp$). Of course, by Exercise **1**.2.24, this latter equation is always true mod $T_p(M_0)$ (respectively, mod $T_p(M_0)^\perp$), and the right-hand side always takes values in the normal bundle (respectively, the tangent bundle) of M_0. But here the left side also lies in $T_p(M_0)^\perp$ (respectively, mod $T_p(M_0)$), so we get equality.

Thus, $\phi \circ c$ is a rolling map if and only if c is tangent to the distribution $\mathcal{R}$.

To calculate the dimension of $\mathcal{R}$, we again fix, as in the proof of Proposition 2.1, an orthonormal basis of $\mathbf{R}^{n+r}$ that is adapted to $T_q(M_1)$ and consider the matrix $\dot{A}A^{-1}$ which is skew symmetric in this basis. By Definition 1.1(3) and 1.1(4), $\dot{A}A^{-1}$ must, in this basis, have the form

$$\begin{pmatrix} 0 & -S^t \\ S & 0 \end{pmatrix}.$$

Thus $\dot{A}A^{-1}$ is completely determined by its value on $T_p(M_0)$, which is known by equation 2.3(ii) at the point (p, A, q) once either $\dot{p}$ or $\dot{q}$ is known. By equation 2.3(i), $\dot{q}$ determines $\dot{p}$, so it follows that $\dim \mathcal{R} \leq n$. On the other hand, for any choice of $\dot{q} \in T_q(M_1)$ equations (i), (ii), and (iii) can be solved for $\dot{p}$ and $\dot{A}$ so that $\dim \mathcal{R} = n$. Finally, this same argument shows that the canonical projection $\pi: \Sigma \to M$ inducing the map $\pi_{*(p,A,q)}: T_{(p,A,q)}(\Sigma) \to T_p(M)$ has the property that $\pi_{*(p,A,q)} \mid \mathcal{R}$ is an isomorphism. $\blacksquare$

Now we are in a position to obtain the existence and uniqueness of rolling maps.

Proposition 2.4. *Let $M_0^n, M_1^n \subset \mathbf{R}^{n+r}$ be submanifolds and let $(p_0, A_0, q_0) \in \Sigma$. Assume we are given a piecewise smooth curve $\sigma_1: (I, 0) \to (M_1, q_0)$ (respectively, $\sigma_0: (I, 0) \to (M_0, p_0)$). Then there is a unique rolling map $g: (I, 0) \to (Euc_n(\mathbf{R}), id)$ with rolling curve σ_1 (respectively, development σ_0).*

Proof. We may as well assume that the curve is smooth since we can patch the smooth pieces together to obtain the general case. We will deal only with the case of $\sigma_1: (I, 0) \to (M_1, q_0)$, as the other case is similar in the configuration space setting we are using. Pulling back the bundle

$$M_0 \times O_n(\mathbf{R}) \times O_r(\mathbf{R}) \to \Sigma \to M_1$$

over I along σ_1 yields a bundle

$$M_0 \times SO_n(\mathbf{R}) \times SO_r(\mathbf{R}) \to \Sigma_{\sigma_1} \to I$$

on which the restriction $\mathcal{R}_{\sigma_1}$ of the distribution $\mathcal{R}$ has dimension 1. Since a one-dimensional distribution is always integrable, there exists a unique curve in Σ_{σ_1} through $(p_0, A_0, q_0) \in \Sigma_{\sigma_1}$ which is tangent to $\mathcal{R}_{\sigma_1}$. The image of this curve in Σ is tangent to $\mathcal{R}$ and covers σ_1. By Lemma 2.3, this proves the existence and uniqueness of the rolling map given in Definition 1.1.

§3. Relation to Levi–Civita and Normal Connections

In this section we relate our discussion of "rolling without slipping or twisting" to the Levi–Civita connection and to the Ehresmann connection on the normal bundle. Some of this material necessarily covers ground similar to that of §6.5.

We consider the case of the standard n-plane $\mathbf{R}^n$ ($= M_1$ in the notation above) $\subset \mathbf{R}^N$ rolling on a manifold M^n ($= M_0$ in the notation above) $\subset \mathbf{R}^N$. Then

$$\Sigma = \{(p, A, q) \in M \times O_N(\mathbf{R}) \times \mathbf{R}^n \mid A\mathbf{R}^n = T_p(M)\}$$

is the configuration space for this pair of manifolds, and we again have the distribution $\mathcal{R}$ on it described in Lemma 2.3. We are going to study the following diagram of fiber bundles.

$$\begin{tikzcd}[row sep=small] & & P_{\tan} \\ \Sigma \arrow[r] & P_\tau \arrow[ur] \arrow[dr] & \\ & & P_{\mathrm{nor}} \end{tikzcd}$$

The three spaces $P_{\tan}$, P_{nor} and P_τ are the bundles over M that we have already considered in §6.5. We are going to describe the images of $\mathcal{R}$ under these maps and show in particular that its images in $T(P_{\tan})$ and $T(P_{\mathrm{nor}})$ are (the kernels of) the Levi–Civita and the normal connections. Since, by Exercise A.1.3, these kernels determine the connections, we see that for a manifold in Euclidean space the notion of rolling without slipping or twisting includes the notions of the Levi–Civita connection on the tangent bundle and the Ehresmann connection on the normal bundle.

We define

$$\mathrm{Iso}(\mathbf{R}^k, \mathbf{R}^N) = \{\phi \in \mathrm{Hom}(\mathbf{R}^k, \mathbf{R}^N) \mid \langle\varphi(u), \varphi(v)\rangle = \langle u, v\rangle \forall u, v \in \mathbf{R}^k\}$$

and set

$$\begin{aligned} P_\tau &= \{(p, A) \in M \times O_N(\mathbf{R}) \mid A(\mathbf{R}^n) = T_p(M)\}, \\ P_{\tan} &= \{(p, \varphi) \in M \times \mathrm{Iso}(\mathbf{R}^n, \mathbf{R}^N) \mid \varphi(\mathbf{R}^n) = T_p(M)\}, \\ P_{\mathrm{nor}} &= \{(p, \varphi) \in M \times \mathrm{Iso}(\mathbf{R}^r, \mathbf{R}^N) \mid \varphi(\mathbf{R}^r) = T_p(M)^\perp\}. \end{aligned}$$

Lemma 3.1. *The map $P_\tau \to M$ (respectively, $P_{tan} \to M$, $P_{nor} \to M$) induced by canonical projection is a principal $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ (respectively, $O_n(\mathbf{R})$, $O_r(\mathbf{R})$) bundle map.*

Proof. We deal with the map $P_\tau \to M$ only, the others being similar. Consider the right action

$$P_\tau \times O_n(\mathbf{R}) \times O_r(\mathbf{R}) \to P_\tau. \atop ((p,A),a,b) \mapsto \left(p, A\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\right)$$

This action is free and proper (note that $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ is compact) and acts transitively on the fibers of the projection $P_\tau \to M$. It follows from Theorem 4.2.4 that $P_\tau \to M$ is a principal bundle with group $O_n(\mathbf{R}) \times O_r(\mathbf{R})$. ■

Exercise 3.2. Verify that $P_{\tan}$ and P_{nor} are the tangent and normal orthonormal frame bundles associated to the inclusion $M \subset \mathbf{R}^N$ as in §6.5. Show also that P_τ corresponds to the bundle of the same name in §6.5. ◇

Lemma 3.3. *For $p \in M$ let $V \subset T_p(\mathbf{R}^N) \times T_A(O_N(\mathbf{R}))$ be defined by*

$$(\dot p, \dot A) \in V \Leftrightarrow \dot p \textit{ and } \dot A \textit{ satisfy the following conditions}$$

(i) $\dot A A^{-1} u = -B(\dot p, u) \bmod T_p(M)$ *for all* $u \in T_p(M)$.

(ii) $\dot A A^{-1} v = B^t(\dot p, v) \bmod T_p(M)^\perp$ *for all* $v \in T_p(M)^\perp$.

Then $T_{(p,A)}(P_\tau) \subset V$.

Proof. Let $(p, A) \in P_\tau$, and consider the vector space V. In an orthonormal basis of $T_p(\mathbf{R}^N)$ adapted to $T_p(M)$ (cf. the proof of Proposition 2.1), conditions (i) and (ii) imply that the matrix $\dot A A^{-1}$ has the block form

$$\dot A A^{-1} = \begin{pmatrix} \text{arbitrary} & \text{determined by } \dot p \\ \text{determined by } \dot p & \text{arbitrary} \end{pmatrix}.$$

In particular, the dimension of V is

$$\dim M + \dim O_n(\mathbf{R}) + \dim O_r(\mathbf{R}) = \dim P_\tau.$$

Thus, it suffices to verify the inclusion $T_{(p,A)}(P_\tau) \subset V$. Suppose that $(\dot p, \dot A) \in T_{(p,A)}(P_\tau)$. Let $(p(t), A(t))$ be a curve on P_τ with tangent $(\dot p, \dot A)$ at

(p, A). Let $u(t)$ be a tangent field (respectively, normal field) to M along the curve $p(t)$. Then $A^{-1}(t)u(t) \in \mathbf{R}^n \times 0$ (respectively, $0 \times \mathbf{R}^r$). Differentiating, we get

$$-A^{-1}(t)\dot{A}(t)A^{-1}(t)u(t) + A^{-1}(t)\dot{u}(t) \in \mathbf{R}^n \times 0 \text{ respectively, } 0 \times \mathbf{R}^r)$$

or

$$-\dot{A}(t)A^{-1}(t)u(t) + \dot{u}(t) \in T_p(M) \text{ (respectively, } T_p(M)^{\perp})).$$

But we know from Exercise **1**.2.24(i) (respectively, (iii)) that

$$\dot{u}(t) = -B(\dot{p}(t), u(t)) \bmod T_p(M)$$

respectively,

$$\dot{u}(t) = B^t(\dot{p}(t), u(t)) \bmod T_p(M)^{\perp}),$$

so

$$\dot{A}(t)A^{-1}(t)u(t) + B(\dot{p}(t), u(t)) \in T_p(M)$$

respectively,

$$\dot{A}(t)A^{-1}(t)u(t) - B^t(\dot{p}(t), u(t)) \in T_p(M)^{\perp}).$$

This verifies the inclusion $\subset$. ■

Exercise 3.4. Show that

$$\begin{aligned} T_{(p,\varphi)}(P_{tan}) &= \{(\dot{p}, \dot{\varphi}) \in T_p(M) \times \mathrm{Hom}(\mathbf{R}^n, \mathbf{R}^N) \mid \dot{\varphi} = -B(\dot{p}, \varphi(\text{—})) \\ &\qquad \bmod T_p(M)\}, \\ T_{(p,\varphi)}(P_{\mathrm{nor}}) &= \{(\dot{p}, \dot{\varphi}) \in T_p(M) \times \mathrm{Hom}(\mathbf{R}^n, \mathbf{R}^N) \mid \dot{\varphi} = -B^t(\dot{p}, \varphi(\text{—})) \\ &\qquad \bmod T_p(M)^{\perp}\}. \end{aligned}$$ ■

It is convenient now to regard the group of Euclidean isometries $\mathrm{Euc}_N(\mathbf{R})$ as the matrix group (with 1×1 and $N \times N$ blocks down the diagonal in the block form)

$$\mathrm{Euc}_N(\mathbf{R}) = \left\{ \begin{pmatrix} 1 & 0 \\ p & A \end{pmatrix} \in Gl_{N+1}(\mathbf{R}) \mid p \in \mathbf{R}^N,\ A \in O_N(\mathbf{R}) \right\}$$

with Lie algebra

$$\mathfrak{euc}_N(\mathbf{R}) = \left\{ \begin{pmatrix} 0 & 0 \\ \dot{p} & \dot{A} \end{pmatrix} \in M_{N+1}(\mathbf{R}) \mid \dot{p} \in \mathbf{R}^N,\ \dot{A} \in \mathfrak{o}_N(\mathbf{R}) \right\}.$$

In this picture, the isometry given by $gv = Av + p$ corresponds to the matrix $\begin{pmatrix} 1 & 0 \\ p & A \end{pmatrix}$, and the (left invariant) Maurer–Cartan form on $\mathrm{Euc}_N(\mathbf{R})$ may be expressed as

$$\begin{aligned} \omega_{\mathrm{Euc}_N(\mathbf{R})} &= \begin{pmatrix} 1 & 0 \\ p & A \end{pmatrix}^{-1} \begin{pmatrix} 0 & 0 \\ \dot{p} & \dot{A} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -A^{-1}p & A^{-1} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \dot{p} & \dot{A} \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ A^{-1}\dot{p} & A^{-1}\dot{A} \end{pmatrix}. \end{aligned}$$

Let us express the Maurer–Cartan form on $\mathrm{Euc}_N(\mathbf{R})$ (with 1×1, $n \times n$, and $r \times r$ blocks down the diagonal in the block form) as

$$\omega_{\mathrm{Euc}_{n+r}(\mathbf{R})} = \begin{pmatrix} 0 & 0 & 0 \\ \star & \alpha & -\beta^t \\ \star & \beta & \gamma \end{pmatrix}.$$

Then we have

Proposition 3.5. *The inclusion $M \subset \mathbf{R}^{n+r}$ is covered by the right $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ equivariant map*

$$\begin{aligned} \iota\colon P_\tau &\subset \mathrm{Euc}_{n+r}(\mathbf{R}). \\ (p, A) &\mapsto \begin{pmatrix} 1 & 0 \\ p & A \end{pmatrix} \end{aligned}$$

Let us define $\mathcal{H} = \ker(\alpha \mid P_\tau) \cap \ker(\delta \mid P_\tau)$. Then, for each $(p, A, q) \in \Sigma$, the derivative of the canonical map $\Sigma \to P_\tau$ induces the isomorphism in the following diagram.

$$\begin{array}{ccc} T_{(p,A,q)}(\Sigma) & \supset & \mathcal{R}_{(p,A,q)} \\ \downarrow & & \downarrow{\scriptstyle\approx} \\ T_{(p,A)}(P_\tau) & \supset & \mathcal{H}_{(p,A)} \end{array}$$

Proof. Since

$$\begin{aligned} &(\iota^*\omega_{\mathrm{Euc}_N(\mathbf{R})})_{(p,A)}(\dot{p}, \dot{A}) \\ &\quad = (\omega_{\mathrm{Euc}_N(\mathbf{R})})_{\begin{pmatrix} 1 & 0 \\ p & A \end{pmatrix}}(\iota_*(\dot{p}, \dot{A})) = \begin{pmatrix} 0 & 0 \\ A^{-1}\dot{p} & A^{-1}\dot{A} \end{pmatrix}, \end{aligned}$$

using the block form of the Maurer-Cartan form, we have

$$A^{-1}\dot{A} = \begin{pmatrix} \alpha(\dot{p}, \dot{A}) & -\beta^*(\dot{p}, \dot{A}) \\ \beta(\dot{p}, \dot{A}) & \gamma(\dot{p}, \dot{A}) \end{pmatrix}.$$

The proof of Lemma 3.1 shows that $P_\tau \subset \mathrm{Euc}_{n+r}(\mathbf{R})$ is a right $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ equivariant map. Now the derivative of the map $\Sigma \to P_\tau$ is given by $(\dot{p}, \dot{A}, \dot{q}) \mapsto (\dot{p}, \dot{A})$, and so, according to the description of $\mathcal{R}$ in Lemma 2.3, (and in the present case, $B_0 = B$ and $B_1 = 0$) the image of $\mathcal{R}_{(p,A,q)}$ under this map is $W_{(p,A)} \subset T_p(M) \times T_A(O_n(\mathbf{R}))$ where

$$(\dot{p}, \dot{A}) \in W_{(p,A)} \Leftrightarrow \text{the following two conditions hold:}$$

(i) $\dot{A}A^{-1}u = -B(\dot{p}, u)$, for all $u \in T_p(M)$

(ii) $\dot{A}A^{-1}\nu = -B^t(\dot{p}, \nu)$, for all $\nu \in T_p(M)^\perp$.

Thus,

$$
\begin{aligned}
(\dot{p}, \dot{A}) \in W_{(p,A)} &\Rightarrow \begin{cases} \dot{A}A^{-1}T_p(M) \subset T_p(M)^\perp, \text{ and} \\ \dot{A}A^{-1}T_p(M)^\perp \subset T_p(M) \end{cases} \\
&\Rightarrow \begin{cases} A^{-1}\dot{A}(\mathbf{R}^n \times 0) \subset 0 \times \mathbf{R}^r, \text{ and} \\ A^{-1}\dot{A}(0 \times \mathbf{R}^r) \subset \mathbf{R}^n \times 0 \end{cases} \\
&\Rightarrow \begin{cases} \alpha(\dot{p}, \dot{A}) = 0, \text{ and} \\ \delta(\dot{p}, \dot{A}) = 0 \end{cases} \\
&\Rightarrow (\dot{p}, \dot{A}) \in \mathcal{H}_{(p,A)}.
\end{aligned}
$$

Thus $W_{(p,A)} \subset \mathcal{H}_{(p,A)}$. On the other hand, by Lemma 2.3, the composite $\mathcal{R}_{(p,A,q)} \to T_{(p,A)}(P_\tau) \to T_p(M)$ is an isomorphism, so that $\mathcal{R}_{(p,A,q)} \to T_{(p,A)}(P_\tau)$ is injective. Since P_τ is a principal $O_n(\mathbf{R}) \times O_r(\mathbf{R})$ bundle over M and $\alpha \oplus \gamma$ restricts to the fiber of P_τ to yield the Maurer–Cartan form on $O_n(\mathbf{R}) \times O_r(\mathbf{R})$, it follows that dim $\mathcal{H} = n =$ dim $\mathcal{R}$, and so the image of $\mathcal{R}_{(p,A,q)}$ is $\mathcal{H}_{(p,A)}$. ∎

Corollary 3.6. *For each $(p, A) \in P_\tau$, the canonical projections $P_\tau \to P_{tan}$ and $P_\tau \to P_{nor}$ induces the isomorphisms in the following diagram.*

$$
\begin{array}{ccc}
T_{(p,A)}(P_\tau) \supset \mathcal{H}_{(p,A)} & \qquad & T_{(p,A)}(P_\tau) \supset \mathcal{H}_{(p,A)} \\
\downarrow \qquad\quad \downarrow \approx & & \downarrow \qquad\quad \downarrow \approx \\
T_{(p,A|\mathbf{R}^n)}(P_{tan}) \supset ker\,(\alpha)_{(p,A|\mathbf{R}^n)} & & T_{(p,A|\mathbf{R}^r)}(P_{nor}) \supset ker\,(\gamma)_{(p,A|\mathbf{R}^r)}
\end{array}
$$

Proof. The proofs are similar, so we consider only the projection $P_\tau \to P_{\text{tan}}$. Now the form α on P_τ is basic for this projection (Proposition **6**.5.12) so the distribution ker α makes sense as a distribution on P_{tan}. Since $\mathcal{H} = \ker(\alpha \mid P_\tau) \cap \ker(\delta \mid P_\tau)$, the canonical projection maps $\mathcal{H}$ to ker α. Finally, since the derivative of the composite map $P_\tau \to P_{\text{tan}} \to M$ is injective on $\mathcal{H}$, it follows that $\mathcal{H} \to$ ker α is injective, and since the dimensions are the same, it is an isomorphism. ∎

Proposition 3.5 establishes the link between "rolling without slipping or twisting" and the usual connection forms on a submanifold of $\mathbf{R}^N$ since we know from Proposition **6**.5.12 that $\ker(\alpha \mid P_\tau) \cap \ker(\delta \mid P_\tau)$ projects to the Levi–Civita connection on P_{tan} and to the canonical normal connection P_{nor}.

Proposition 3.7. *Let $t \mapsto (\sigma_0(t), A(t), \sigma_1(t)) \in \Sigma$ be an integral curve for the distribution $\mathcal{R}$.*

(i) *$t \mapsto \sigma_1(t) \in \mathbf{R}^n$ is the development (in the sense of Definition* **5**.4.15*) of $t \mapsto \sigma_0(t) \in M$.*

(ii) *If $v(t)$ is a vector field along $t \mapsto \sigma_0(t) \in M$ which is tangent (respectively, normal) to M, then*

$$D_{\dot{\sigma}_0(t)}v(t) = A(t)(A(t)^{-1}v(t)),$$

where $D_{\dot{\sigma}_0(t)}v(t)$ is the Levi–Civita (respectively, the canonical normal) covariant derivative.

(iii) *A vector field $v(t)$ along $t \mapsto \sigma_0(t) \in M$ which is tangent (respectively, normal) to M is parallel if and only if $A(t)^{-1}v(t)$ is constant.*

Proof. (i) Exercise **5**.4.14 tells us how to develop a curve $t \mapsto \sigma_0(t) \in M$ to a curve on the model space of a reductive geometry. It says we must first take the horizontal lift $\hat{\sigma}: (I, a) \to (P_{\text{tan}}, p)$, which in the present case is the projection of $t \mapsto (\sigma_0(t), A(t), \sigma_1(t)) \in \Sigma$, to P_{tan}. Then we must solve the equation $\hat{\sigma}^*\omega_{\text{tan}} = d\tilde{\sigma}$ (i.e., $\omega_{\text{tan}}(\dot{\hat{\sigma}}) = \dot{\tilde{\sigma}}$) for the development $\tilde{\sigma}: (I, a) \to (\mathbf{R}^n, 0)$. But this equation is, in the present circumstance, the equation $A(t)^{-1}\dot{\sigma}_0(t) = \dot{\tilde{\sigma}}(t)$. But since, by the definition of $\mathcal{R}$, $A(t)^{-1}\dot{\sigma}_0(t) = \dot{\sigma}_1(t)$, the uniqueness of development implies $\tilde{\sigma}(t) = \sigma_1(t)$ for all t.

(ii) Case (1): $v(t)$ is a tangent vector field. According to Definition A.4.1, the covariant derivative D_XY, for $X \in T_x(M)$ and Y a tangent vector field on M, is calculated as follows: express Y as a function $f_Y: P \to \mathfrak{g}/\mathfrak{h}$; lift X to a horizontal vector $\tilde{X}_p = T_p(P)$; compute $\tilde{X}_p(f_Y) \in \mathfrak{g}/\mathfrak{h}$; and then set $D_XY = \varphi_p^{-1}(\tilde{X}_p(f_Y)) \in T_x(M)$.

First consider the tangent vector field $v(t)$ on M. It determines the function $t \mapsto \varphi_{(\sigma(t),A(t))}(v(t)) \in \mathfrak{g}/\mathfrak{h}$.

Next, since the curve $t \mapsto \sigma_0(t)$ has the lift $t \mapsto (\sigma_0(t), A(t), \sigma_1(t)) \in \Sigma$, which is an integral curve for the distribution $\mathcal{R}$, it follows by Corollary 3.6 that the curve $t \mapsto (\sigma_0(t), A(t) \mid \mathbf{R}^n \times 0)$ is a horizontal lift to P_{tan}, that is, $(\dot{\sigma}_0(t), \dot{A}(t) \mid \mathbf{R}^n \times 0))$ is a horizontal vector covering $\dot{\sigma}_0(t)$.

Putting this all together, we have

$$
\begin{aligned}
\varphi_{(\sigma(t),A(t))}(D_{\dot{\sigma}_0(t)}v(t)) &= (\varphi_{(\sigma(t),A(t))}(v(t)))^{\cdot} \\
&= (A(t)^{-1}(v(t)))^{\cdot} \\
&= \varphi_{(\sigma(t),A(t))}(A(t)(A(t)^{-1}(v(t)))^{\cdot}).
\end{aligned}
$$

Since $\varphi_{(\sigma(t),A(t))}$ is injective, this verifies case (1).

Case (2): $v(t)$ is a normal vector field. This time the covariant derivative is taken with respect to the Ehresmann connection on the principal $O_r(\mathbf{R})$ bundle $O_r(\mathbf{R}) \to Q \to M$ of normal frames on M.

According to Definition A.4.1, the covariant derivative D_XY, for $X \in T_x(M)$ and Y a normal vector field on M, may be calculated as follows: express Y as a function $f_Y: Q \to \mathbf{R}^r$ (i.e., write Y in terms of the basis given by the frame $q \in Q$); lift X to a horizontal vector $\tilde{X}_q \in T_q(Q)$; compute $\tilde{X}_q(f_Y) \in \mathbf{R}^r$. Then $D_X Y$ is the normal vector to M which, when expressed in the basis $q \in Q$, is $\tilde{X}_q(f_Y)$.

First consider the normal vector field $v(t)$ on M. It determines the function $t \mapsto A(t)^{-1}v(t) \in \mathbf{R}^r$ with derivative $t \mapsto (A(t)^{-1}v(t))^{\cdot} \in \mathbf{R}^r$.

Since the curve $t \mapsto \sigma_0(t)$ has the lift $t \mapsto (\sigma_0(t), A(t), \sigma_1(t)) \in \Sigma$, which is an integral curve for the distribution $\mathcal{R}$, it follows by Corollary 3.6 that the curve $t \mapsto (\sigma_0(t), A(t) \mid 0 \times \mathbf{R}^r)$ is a horizontal lift to P_{nor}, that is, for each t, $(\dot{\sigma}_0(t), \dot{A}(t) \mid 0 \times \mathbf{R}^r)$ is a horizontal vector covering $\dot{\sigma}_0(t)$.

Putting all this together, we have

$$D_{\dot{\sigma}_0(t)} v(t) = A(t)(A(t)^{-1}(v(t)))^{\cdot}.$$

(iii) This is an immediate consequence of (ii). ■

§4. Transitivity of Rolling Without Slipping or Twisting

In this section we will study the "composite" of two "rollings without slipping or twisting."

Theorem 4.1. *Let three submanifolds $M_0, M_1, M_2 \subset \mathbf{R}^N$ be given which are tangent to each other at some point $p \in M_0 \cap M_1 \cap M_2$, and let a path $\sigma_0 \colon (I, 0) \to (M_0, p)$ be given. Suppose that*

(a) *M_1 rolls on M_0 along $\sigma_1 \colon (I, 0) \to (M_1, p)$ without slipping or twisting via $g_1 \colon (I, 0) \to (Euc_N(\mathbf{R}), I)$, with development $\sigma_0 \colon (I, 0) \to (M_0, p)$,*

(b) *M_2 rolls on M_1 along σ_2 without slipping or twisting via $g_2 \colon (I, 0) \to (Euc_N(\mathbf{R}), I)$, with development $\sigma_1 \colon (I, 0) \to (M_1, p)$.*

Then it follows that

(c) *M_2 rolls on M_0 along σ_2 without slipping or twisting via $g_1 g_2 \colon (I, 0) \to (Euc_N(\mathbf{R}), I)$, with development $\sigma_0 \colon (I, 0) \to (M_0, p)$.*

Proof. It suffices to show that $g_1 g_2$ verifies properties (1), (2), (3), and (4) of Definition 1.1.

(1) By (a), $\sigma_0(t) = g_1(t)\sigma_1(t)$ and $T_{\sigma_0(t)}(g_1(t)M_1) = T_{\sigma_0(t)}(M_0)$.
By (b), $\sigma_1(t) = g_2(t)\sigma_2(t)$ and $T_{\sigma_1(t)}(g_2(t)M_2) = T_{\sigma_1(t)}(M_1)$.
Thus, $\sigma_1(t) = g_1(t)g_2(t)\sigma_3(t)$ and

$$\begin{aligned} T_{\sigma_1(t)}(g_1(t)g_2(t)M_3) &= g_1(t)_* T_{\sigma_2(t)}(g_2(t)M_3) \\ &= g_1(t)_* T_{\sigma_2(t)}(M_2) \\ &= T_{\sigma_1(t)}(g_1(t)M_2) \\ &= T_{\sigma_1(t)}(M_1). \end{aligned}$$

In the proof of the remaining parts, we shall need the (easily proved) formula

$$(g_1(t)g_2(t))^{\cdot}(g_1(t)g_2(t))^{-1} = \dot{g}_1(t)g_1(t)^{-1} + g_1(t)\dot{g}_2(t)g_2(t)^{-1}g_1(t)^{-1}.$$

(2)

$$\begin{aligned} &(g_1(t)g_2(t))^{\cdot}(g_1(t)g_2(t))^{-1}\sigma_1(t) \\ &\quad = \dot{g}_1(t)g_1(t)^{-1}\sigma_1(t) + g_1(t)\dot{g}_2(t)g_2(t)^{-1}g_1(t)^{-1}\sigma_1(t) \\ &\quad = 0 + g_1(t)\dot{g}_2(t)\dot{g}_2(t)^{-1}\sigma_2(t) = 0. \end{aligned}$$

(3)

$$\begin{aligned} &((g_1(t)g_2(t))^{\cdot}(g_1(t)g_2(t))^{-1})_* T_{\sigma_1(t)}(M_1) \\ &\quad = (\dot{g}_1(t)g_1(t)^{-1})_* T_{\sigma_1(t)}(M_1) + (g_1(t)\dot{g}_2(t)g_2(t)^{-1}g_1(t)^{-1})_* T_{\sigma_1(t)}(M_1) \\ &\quad \subset T_{\sigma_1(t)}(M_1)^{\perp} + g_1(t)_*(\dot{g}_2(t)g_2(t)^{-1})_* g_1(t)_*^{-1} T_{\sigma_1(t)}(M_1) \\ &\quad \subset T_{\sigma_1(t)}(M_1)^{\perp} + g_1(t)_*(\dot{g}_2(t)g_2(t)^{-1})_* T_{\sigma_2(t)}(M_2) \\ &\quad \subset T_{\sigma_1(t)}(M_1)^{\perp} + g_1(t)_* T_{\sigma_2(t)}(M_2)^{\perp} \subset T_{\sigma_1(t)}(M_1)^{\perp}. \end{aligned}$$

(4) is similar to (3). ■

Corollary 4.2. *Suppose that two manifolds $M_1, M_2 \subset \mathbf{R}^N$ are given which are tangent to each other at some point $p \in M_1 \cap M_2$, and let a path $\sigma_1 \colon (I, 0) \to (M_1, p)$ be given. Suppose that*

(a) *M_2 rolls on M_1 along σ_1 without slipping or twisting via $g_1 \colon (I, 0) \to (Euc_n(\mathbf{R}), I)$, with development $\sigma_2 \colon (I, 0) \to (M_2, p)$.*

Then it follows that

(b) *M_1 rolls on M_2 along σ_2 without slipping or twisting via $g_2 = g_1^{-1}$: $(I, 0) \to (Euc_n(\mathbf{R}, I)$, with development $\sigma_1 \colon (I, 0) \to (M_1, p)$.*

Proof. Apply the theorem to the case $M_3 = M_1$ to see that M_1 rolls on M_1 along σ_1 without slipping or twisting via $g_1 g_2 \colon (I, 0) \to (\mathrm{Euc}_n(\mathbf{R}), I)$, with development $\sigma_3 \colon (I, 0) \to (M_2, p)$. But obviously $g_1 g_2 = I$ and $\sigma_3 = \sigma_1$. ■

Corollary 4.3. *The "rolling data" consisting of a fixed manifold M and a curve σ on it depend only on the curve in space and the tangent spaces to M along the curve.*

Proof. Construct a manifold M_1 out of σ and the tangent spaces along the curve by letting V_t be the subspace of $T_{\sigma(t)}(M)$ orthogonal to $\sigma'(t)$ and setting $M_1 = \cup V_t$. It is easy to see that M_1 is a manifold in some neighborhood of σ which is a smooth curve on M_1. Moreover, the rolling map of M_1 on M is the identity. Thus, by multiplication, if we wish to find the rolling map of $\tilde{M}$ along (M, σ), it is enough to roll it along (M_1, σ) or indeed any other manifold tangent to M along σ. ■

Exercise 4.4. Verify that the picture given at the beginning of this appendix is correct. That is, show that a sphere rolling on the central line of a helicoid rolls along a line of longitude. [*Hint:* use the fact that the central line of a helicoid is a geodesic.] ❑

Appendix C

Classification of One-Dimensional Effective Klein Pairs

The classification of the one-dimensional effective Klein pairs was first given by Lie in the nineteenth century. In §1 we give a modern proof. In fact, the method we follow forms the basis for the "rough" classification of the primitive effective Klein pairs of any dimension over any field of characteristic zero (cf., e.g., [K. DePaepe, 1996]). More information on the classification of Klein pairs over the real and complex fields may be found in [K. Yamaguchi, 1993].

§1. Classification of One-Dimensional Effective Klein Pairs

Proposition 4.3.19. *Let* $(\mathfrak{g}, \mathfrak{h})$ *be an effective Klein pair with* $\dim \mathfrak{g}/\mathfrak{h} = 1$. *Then* $(\mathfrak{g}, \mathfrak{h})$ *is isomorphic to one of the three Klein pairs described in Definition* 4.3.18.

Proof. Assume first that the adjoint action $\mathrm{ad}\colon \mathfrak{h} \to \mathrm{End}(\mathfrak{g}/\mathfrak{h})$ is injective. Since $\dim \mathfrak{g}/\mathfrak{h} = 1$, we get $\mathrm{End}\,(\mathfrak{g}/\mathfrak{h}) = \mathbf{R}$. This means that either

(i) $\mathfrak{h} = 0$, in which case $(\mathfrak{g}, \mathfrak{h}) = (\mathbf{R}, 0)$, or

(ii) $\mathfrak{h} = \mathbf{R}$ and the adjoint map is an isomorphism.

In this latter case we may choose a unique $h \in \mathfrak{h}$ such that $\mathrm{ad}(h) = \mathrm{id}_{\mathfrak{g}/\mathfrak{h}}$. Thus, if $g_1 \in \mathfrak{g}$ represents a basis element of $\mathfrak{g}/\mathfrak{h}$, then we have $[h, g_1] = g_1 + h_1$ for some $h_1 \in \mathfrak{h}$. But then the element $g = g_1 + h_1$ also represents the same basis element of $\mathfrak{g}/\mathfrak{h}$, and moreover $[h,g]=[h,g_1+h_1]=[h,g_1]+0=g_1+h_1=g$. The map sending

$$g\mapsto\begin{pmatrix}0&0\\1&0\end{pmatrix}\quad\text{and}\quad h\mapsto\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

establishes the isomorphism with the Klein pair of the affine line.

This leaves us with the most complicated case where the adjoint action $\mathrm{ad}\colon\mathfrak{h}\to\mathrm{End}(\mathfrak{g}/\mathfrak{h})$ has a non zero kernel, which we assume for the rest of the proof. Define a sequence of subalgebras $U_{-1}\supset U_0\supset U_1\supset U_2\supset\dots$ by setting $U_{-1}=\mathfrak{g}$, $U_0=\mathfrak{h}$, and then, inductively, $U_{s+1}=\{x\in U_s\mid[\mathfrak{g},x]\subset U_s\}$ for $s=0,1,\dots$. In particular, we note that

$$U_1=\ker(\mathrm{ad}\colon\mathfrak{h}\to\mathrm{End}(\mathfrak{g}/\mathfrak{h})\neq 0.$$

Sublemma 1.1.

(i) $[U_i,U_j]\subset U_{i+j}$.

(ii) *For some* $\mu>0$, *we have* $\dim(U_s/U_{s+1})=1$ *for* $-1\le s\le\mu$ *and* $U_s=0$ *for* $s>\mu$.

(iii) $[U_{-1},U_s]\not\subset U_s$ *for* $0\le s\le\mu$.

Proof. (i) The result is obviously true if $i\le 0$ and $j\le 0$. Fix $s\ge 1$, and assume by induction that the result is true for $-1\le i$, $-1\le j$, and $i+j<s$. Suppose $i+j=s$. Then

$$\begin{aligned}\mathrm{ad}(\mathfrak{g})[U_i,U_j]&\subset[\mathrm{ad}(\mathfrak{g})U_i,U_j]+[U_i,\mathrm{ad}(\mathfrak{g})U_j]\\&\subset[U_{i-1},U_j]+[U_i,U_{j-1}]\subset U_{i+j-1}\quad\text{(by induction).}\end{aligned}$$

Thus, $[U_i,U_j]\subset U_{i+j}$.

(ii) Since $\dim\,\mathfrak{g}<\infty$ and $U_0\supset U_1\supset U_2\dots$, it follows that, for some minimal μ, $U_{\mu+1}=U_{\mu+2}$. But then $[\mathfrak{g},U_{\mu+1}]\subset U_{\mu+1}$. Thus, $U_{\mu+1}$ is an ideal in $\mathfrak{g}$ contained in $\mathfrak{h}$ and hence vanishes by the hypothesis of effectiveness. Since $U_1\neq 0,\mu>0$. Next consider the maps $\mathrm{ad}\colon U_s\to\mathrm{Hom}(\mathfrak{g},U_{s-1})$. These induce maps

$$U_s/U_{s+1}\to\mathrm{Hom}(\mathfrak{g},U_{s-1}/U_s),$$

which are injective because, by definition of the U_ss, if $x\in U_s$ and $[x,\mathfrak{g}]\subset U_s$, then $x\in U_{s+1}$. Since $[U_s,\mathfrak{h}]\subset U_s$, these maps induce injective maps

$$U_s/U_{s+1}\to\mathrm{Hom}(\mathfrak{g}/\mathfrak{h},U_{s-1}/U_s).$$

Since $\dim\,\mathfrak{g}/\mathfrak{h}=1$, it follows that $\dim(U_s/U_{s+1})\le\dim(U_{s-1}/U_s)$. But $\dim\,U_{-1}/U_0=\dim\,\mathfrak{g}/\mathfrak{h}=1$. Assuming inductively that $\dim(U_{s-1}/U_s)=1$, it follows that $\dim(U_s/U_{s+1})\le 1$. Either $\dim(U_s/U_{s+1})=1$, in which case $s\le\mu$ and the induction continues, or $\dim(U_s/U_{s+1})=0$, in which case $U_s=U_{s+1}$ so that $s=\mu+1$.

(iii) $[U_{-1},U_s]\subset U_s\Rightarrow U_s\subset U_{s+1}\Rightarrow U_s=U_{s+1}\Rightarrow s>\mu$. ■

Sublemma 1.2. *The associated graded Lie algebra* $U_{-1}/U_0\oplus U_0/U_1\oplus\cdots\oplus U_\mu/U_{\mu+1}$ *has a basis* $e_s\in U_s/U_{s+1}, s=-1,0\dots,\mu$, *such that*

(i) $[e_0,e_s]=-2se_s$ *for* $s=-1,0,1,\dots$ *and*

(ii) $[e_{-1},e_{s+1}]=e_s$ *for* $s=0,1,\dots$.

Proof. Part (ii) of the Sublemma 1.1 tells us that the associated graded Lie algebra $U_{-1}/U_0\oplus U_0/U_1\oplus\cdots\oplus U_\mu/U_{\mu+1}$ has every grade of dimension one, and part (iii) says that $[U_{-1}/U_0,\ U_s/U_{s+1}]\neq 0$ for $0\le s\le\mu$. Fix basis elements $e_{-1}\in U_{-1}/U_0$ and (temporarily) $e_\mu\in U_\mu/U_{\mu+1}$, and define $e_s=[e_{-1},e_{s+1}]$ for $s=\mu-1,\dots,0$ inductively. By Sublemma 1.1(iii) none of the e_ss vanish. Now define λ_j by $[e_0,e_s]=\lambda_se_s$ for $s=-1,0,1,\dots,\mu$, and note that

$$\begin{aligned}\lambda_se_s&=[e_0,e_s]\\&=[e_0,[e_{-1},e_{s+1}]]\\&=[[e_0,e_{-1}],e_{s+1}]+[e_{-1},[e_0,e_{s+1}]]\\&=[\lambda_{-1}e_{-1},e_{s+1}]+[e_{-1},\lambda_{s+1}e_{s+1}]\\&=\lambda_{-1}e_s+\lambda_{s+1}e_s.\end{aligned}$$

This yields $\lambda_{s+1}=\lambda_s-\lambda_{-1}$. Thus, $\lambda_s=-s\lambda_{-1}$. Since $[U_{-1}/U_0,\ U_0/U_1]$ has $-\lambda_{-1}e_{-1}=[e_{-1},e_0]$ as a basis, it follows that $\lambda_{-1}\neq 0$. If we replace our choice of e_μ by $2e_\mu/\lambda_{-1}$, this replaces e_s by $2e_s/\lambda_{-1}$ for $s=\mu-1,\dots,0$ and also replaces λ_{-1} by 2 so that $\lambda_{-s}=-2s$. However, the equations $e_s=[e_{-1},e_{s+1}]$ for $s=\mu-1,\dots,0$ continue to hold. ■

Sublemma 1.3. *The Lie algebra* $\mathfrak{g}$ *has a basis* $x_s,-1\le s\le\mu$, *with* $x_s\in U_s$ *a lift of* $e_s\in U_s/U_{s+1}$, *such that*

(i)	$[x_0,x_s]=-2sx_z$	*for* $s=-1,0,1,\dots,\mu$	*and*
(ii)	$[x_{-1},x_{s+1}]=x_z$	*for* $s=0,1,\dots,\mu-1$	*and*
(iii)	$[x_t,x_s]\in\langle x_{t+s}\rangle$	*(the span of* x_{t+s}*)*	

Proof. (i) Let $x_0\in\mathfrak{h}$ be any lift of e_0. We know from linear algebra that the characteristic polynomials of $\mathrm{ad}(x_0)$ and $\mathrm{ad}(e_0)$ are the same. Since, by Sublemma 1.2, $\mathrm{ad}(e_0)$ is diagonalizable, with all eigenspaces of dimension one and eigenvalues $2,0,-2,\dots,-2\mu$, the equality of the characteristic polynomials shows the same is true of $\mathrm{ad}(x_0)$. For $s\neq 0$, let x_s denote an eigenvector of $\mathrm{ad}(x_0)$ with eigenvalue $-2s$. Replacing each $x_s,-1\le s\le\mu,s\neq 0$, by some multiple, we may assume that $x_s\in U_s$ is a lift of $e_s \in U_s/U_{s+1}$ for $-1 \le s \le \mu$. It is clear that $x_s - 1 \le s \le \mu$ forms a basis for $\mathfrak{g}$. This proves (i).

Now note that

$$\begin{aligned}\mathrm{ad}(x_0)[x_t, x_s] &= [\mathrm{ad}(x_0)x_t, x_s] + [x_t, \mathrm{ad}(x_0)x_s]\\ &= -2t[x_t, x_s] - 2s[x_t, x_s]\\ &= -2(t+s)[x_t, x_s].\end{aligned}$$

(ii) Now $[x_{-1}, x_{s+1}] \in U_s$ covers $[e_{-1}, e_{s+1}] = e_s \in U_s/U_{s+1}$ for all s. Since, by the preceding formula,

$$\mathrm{ad}(x_0)[x_{-1}, x_{s+1}] = -2s[x_{-1}, x_{s+1}],$$

we see that $[x_{-1}, x_{s+1}]$ is the unique eigenvector of $\mathrm{ad}(x_0)$ lifting e_s. Thus $[x_{-1}, x_{s+1}] = x_s$ for $s = 0, 1, \ldots, \mu - 1$.

(iii) Since $\mathrm{ad}(x_0)[x_t, x_s] = -2(t+s)[x_t, x_s]$ we see that $[x_t, x_s]$ lies in the $-2(t+s)$-eigenspace of $\mathrm{ad}(x_0)$ which, by (i), is $\langle x_{t+s}\rangle$. ■

Next we use the Killing form B on g defined by

$$B(x, y) = \mathrm{Trace}(\mathrm{ad}(x)\mathrm{ad}(y)).$$

Note that

$$B(x_0, x_0) = \mathrm{Trace}(\mathrm{ad}(x_0)^2) = (-2)^2 + 0^2 + 2^2 + 4^2 + \cdots + 4\mu^2 > 0.$$

Sublemma 1.4. $B([x,y],z) = B(x,[y,z])$ *for all* $x, y, z \in \mathfrak{g}$. *Moreover, the radical of* B *is an ideal of* $\mathfrak{g}$.

Proof. We have $\mathrm{ad}([x,y]) = [\mathrm{ad}(x), \mathrm{ad}(y)] \in \mathrm{End}(\mathfrak{g})$, so, setting $X = \mathrm{ad}(x)$, $Y = \mathrm{ad}(y)$, $Z = \mathrm{ad}(z)$, we have

$$\begin{aligned}B([x,y],z) - B(x,[y,z]) &= \mathrm{Trace}(\mathrm{ad}[x,y]\mathrm{ad}(z)) - \mathrm{Trace}(\mathrm{ad}(x)\mathrm{ad}[y,z])\\ &= \mathrm{Trace}([X,Y]Z) - \mathrm{Trace}(X[Y,Z])\\ &= \mathrm{Trace}(XYZ - YXZ - (XYZ - XZY))\\ &= \mathrm{Trace}(-YXZ + XZY)\\ &= \mathrm{Trace}([XZ,Y]) = 0.\end{aligned}$$

By definition $\mathrm{Rad}(B) = \{x \in \mathfrak{g} \mid B(x, \mathfrak{g}) = 0\}$ is the radical of B. If $x \in \mathrm{Rad}(B)$ and $y \in \mathfrak{g}$, then $B([x,y],\mathfrak{g}) = B(x,[y,\mathfrak{g}]) = 0$ by the first part of the lemma. Thus, $\mathrm{Rad}(B)$ is an ideal. ■

Now we are in a position to complete the proof of the proposition using the basis $x_{-1}, x_0, \ldots$ for $\mathfrak{g}$ constructed in Sublemma 1.3. Sublemma 1.3(iii) shows us that $\mathrm{ad}(x_s)\mathrm{ad}(x_t)x_u \in \langle x_{s+t+u}\rangle$. It follows that for $s + t \neq 0$,

$\mathrm{ad}(x_s)\mathrm{ad}(x_t)$ is nilpotent. In particular, $B(x_s, x_t) = 0$ unless $s + t = 0$. Thus, $\mathrm{Rad}\, B \supset U_2$.

Suppose $U_2 \neq 0$. We derive a contradiction as follows.

Since $0 \neq U_2 \subset \mathrm{Rad}(B)$ and $\mathrm{Rad}(B) \not\subset \mathfrak{h}$ (since $(\mathfrak{g}, \mathfrak{h})$ is effective) it follows that there is an element of the form

$$\alpha x_{-1} + \beta x_0 + \gamma \alpha_1 \in \mathrm{Rad}(B) \quad \text{with } \alpha \neq 0.$$

Since $B(x_0, x_0) \neq 0$, we have

$$\begin{aligned}0 &= B(\alpha x_{-1} + \beta x_0 + \gamma x_1, x_0)\\ &= \beta B(x_0, x_0) \Rightarrow \beta = 0.\end{aligned}$$

Since $\mathrm{Rad}(B)$ is an ideal we also have

$$[x_{-1}, \alpha x_{-1} + \gamma x_1] = \gamma x_0 \in \mathrm{Rad}(B)$$

so $\gamma = 0$. Finally we have

$$[x_1, \alpha x_{-1}] = \alpha x_0 \in \mathrm{Rad}(B)$$

so $\alpha = 0$, which is a contradiction.

Thus, $U_2 = 0$ and so $\mu \le 1$. But by Sublemma 1.1(ii), $\mu \ge 1$. Thus, $\mu = 1$ and the map sending

$$x_{-1} \mapsto \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$x_0 \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$x_1 \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

establishes an isomorphism between $(\mathfrak{g}, \mathfrak{h})$ and the Klein pair of the projective line. ■

Appendix D

Differential Operators Obtained from Symmetry

In this appendix we apply the ideas about geometric differential operators described toward the end of §5 of Chapter 3 to show how curl, divergence, and the Cauchy–Riemann operator arise from symmetry considerations on a two-dimensional Riemannian manifold. This appendix may be regarded as an introduction to [E. Cartan, 1966]. The main aim of that book is to obtain the Dirac operator on a four-dimensional Lorentzian manifold. We present the simplest case of Cartan's method, with the aim of clarifying the principles involved in obtaining geometric differential operators from symmetry considerations in a Cartan geometry.[1]

Throughout this appendix we are considering an oriented Riemannian surface with model $(\mathfrak{g}, \mathfrak{h}) = (\mathfrak{euc}_2(\mathbf{R}), \mathfrak{so}_2(\mathbf{R}))$ and group $H = SO_2(\mathbf{R})$.

In §1 we give a very pedestrian introduction to the "plethysm" of representations of the circle group. In §2 we apply this to decompose the covariant derivative on Riemannian surfaces.

§1. Real Representations of $SO_2(\mathbf{R})$

Let $V_k = (\mathbf{R}^2, \rho_k)$, $k \in \mathbf{Z}$, denote the real two-dimensional representation of $SO_2(\mathbf{R})$ given by

[1] On p. 55 of [E. Cartan, 1966] one can find a very brief description of the analogous material in the case of Euclidean 3-space.

$$\rho_k \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos k\theta & -\sin k\theta \\ \sin k\theta & \cos k\theta \end{pmatrix} \in SO_2(\mathbf{R}).$$

The isomorphism

$$\begin{array}{ccc} SO_1(\mathbf{R}) & \to & U_1(\mathbf{C}) \\ \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} & \mapsto & e^{i\theta} \end{array}$$

allows us to make the reinterpretation $V_k = (\mathbf{C}, \rho_k)$, with $\rho_k(e^{i\theta}) = e^{ik\theta}$. Let e_k, f_k denote the standard real basis for V_k and let e_k^*, $f_k^* \in V_k^*$ be the dual basis.[2] Under the identification $\mathbf{R}^2 \approx \mathbf{C}$ we have the correspondence $x_k e_k + y_k f_k \leftrightarrow z_k = x_k + iy_k$.

Exercise 1.1. Show that the representations $V_{\pm k}$ are isomorphic (over $\mathbf{R}$). ❏

Note that $V_0 = U \oplus U$, where U is the trivial one-dimensional representation. In fact, the representations V_k, $k > 0$, together with U, constitute a complete set of isomorphism classes of all the irreducible representations. Every finite-dimensional representation is isomorphic to a direct sum of copies of these representations. However, we don't really need these facts. What is important for our study is the Clebsch–Gordan (or plethysm) problem of describing the decomposition of the tensor products of the representations V_k. This is given in Lemma 1.2. Further information on plethysm may be found in [W. Fulton and J. Harris, 1991].

Lemma 1.2. (i) *For $k, l \neq 0$, there is an isomorphism of representations $\phi\colon V_k \otimes V_l = V_{k-l} \oplus V_{k+l}$ defined by*

$$\begin{cases} \phi(e_k \otimes e_l) = \frac{1}{2} e_{k-l} + \frac{1}{2} e_{k+l}, \\ \phi(f_k \otimes f_l) = \frac{1}{2} e_{k-l} - \frac{1}{2} e_{k+l}, \\ \phi(e_k \otimes f_l) = -\frac{1}{2} f_{k-l} + \frac{1}{2} f_{k+l}, \\ \phi(f_k \otimes e_l) = \frac{1}{2} f_{k-l} + \frac{1}{2} f_{k+l} \end{cases}$$

or, equivalently,

$$\begin{cases} \phi^{-1}(e_{k+l}) = e_k \otimes e_l - f_k \otimes f_l, \\ \phi^{-1}(e_{k-l}) = e_k \otimes e_l + f_k \otimes f_l, \\ \phi^{-1}(f_{k+l}) = e_k \otimes f_l + f_k \otimes e_l, \\ \phi^{-1}(f_{k-1}) = f_k \otimes e_l - e_k \otimes f_l. \end{cases}$$

(ii) *There is an isomorphism of representations $\tau\colon V_k^* \to V_k$ defined by*

[2] Note that the left $SO_n(\mathbf{R})$ module structure on $\mathrm{Hom}_{\mathbf{R}}(\mathbf{R}^n, \mathbf{R})$ is always taken to be $(h \cdot \varphi)(v) = \varphi(h^{-1}v)$. While it would be possible to drop the inverse in the present abelian ($n = 2$) case, there would eventually be conflict with the other cases.

$$\begin{cases} \tau(e_k^*) = e_k, \\ \tau(f_k^*) = f_k. \end{cases}$$

Proof. (i) It is clear that ϕ is a linear isomorphism. Let $z_k \in V_k$, $z_l \in V_l$, regarding these as complex numbers. Then the formulas above imply that $\phi(z_k \otimes z_l) = (\frac{1}{2} z_k \bar{z}_l, \frac{1}{2} z_k z_l) \in V_{k-l} \oplus V_{k+l}$, which makes obvious the fact that ϕ is an $SO_2(\mathbf{R})$ module isomorphism.

(ii) Clearly, τ is a linear isomorphism. We must see that it satisfies

$$\tau(e^{i\theta} \cdot \varphi) = e^{i\theta} \cdot \tau(\varphi) \text{ for all } \varphi \in V_k^* = \mathrm{Hom}_{\mathbf{R}}(V_k, \mathbf{R}).$$

We may write

$$\varphi = \varphi(e_k) e_k^* + \varphi(f_k) f_k^*$$

so that we have

$$\begin{aligned} e^{i\theta} \cdot \varphi &= (e^{i\theta} \cdot \varphi)(e_k) e_k^* + (e^{i\theta} \cdot \varphi)(f_k) f_k^* \\ &= \varphi(e^{-ik\theta} e_k) e_k^* + \varphi(e^{-ik\theta} f_k) f_k^*. \end{aligned}$$

Thus,

$$\begin{aligned} \tau(e^{i\theta} \cdot \varphi) &= \tau(\varphi(e^{-ik\theta} e_k) e_k^* + \varphi(e^{-ik\theta} f_k) f_k^*) \\ &= \varphi(e^{-ik\theta} e_k) e_k + \varphi(e^{-ik\theta} f_k) f_k \\ &= \varphi(\cos(k\theta) e_k - \sin(k\theta) f_k) e_k + \varphi(\sin(k\theta) e_k + \cos(k\theta) f_k) f_k \\ &= \varphi(e_k)(\cos(k\theta) e_k + \sin(k\theta) f_k) + \varphi(f_k)(-\sin(k\theta) e_k + \cos(k\theta) f_k) \\ &= \varphi(e_k)(e^{i\theta} \cdot e_k) + \varphi(f_k)(e^{i\theta} \cdot f_k) \\ &= e^{i\theta} \cdot (\varphi(e_k) e_k + \varphi(f_k) f_k) \\ &= e^{i\theta} \cdot \tau(\varphi). \end{aligned}$$ ■

Let

$$E_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \in \mathfrak{g},$$

and let E_0^*, E_1^*, $E_2^* \in \mathfrak{g}^*$ be the dual basis.

Corollary 1.3. $V_k \otimes \mathfrak{g}^* \approx V_{k-1} \oplus V_k \oplus V_{k+1}$, *where the isomorphism sends*

$$\begin{array}{ll} e_k \otimes E_0^* \mapsto e_k, & f_k \otimes E_0^* \mapsto f_k, \\ e_k \otimes E_1^* \mapsto \frac{1}{2} e_{k-1} + \frac{1}{2} e_{k+1}, & f_k \otimes E_1^* \mapsto \frac{1}{2} f_{k-1} + \frac{1}{2} f_{k+1}, \\ e_k \otimes E_2^* \mapsto -\frac{1}{2} f_{k-1} + \frac{1}{2} f_{k+1}, & f_k \otimes E_2^* \mapsto \frac{1}{2} e_{k-1} - \frac{1}{2} e_{k+1}. \end{array}$$

Proof. Since $\mathfrak{h} \approx U$ and $\mathfrak{p} \approx V_1$, it follows from Lemma 1.2(ii) that

$$\begin{aligned}\mathfrak{g}^* = \mathfrak{h}^* \oplus \mathfrak{p}^* &\approx U \oplus V_1.\\ E_0^* &\mapsto u \text{ (a generator of } U)\\ E_1^* &\mapsto e_1\\ E_2^* &\mapsto f_1\end{aligned}$$

The rest is a direct application of Lemma 1.2(i). ■

Exercise 1.4. Find analogs of Exercise 1.1, Lemma 1.2, and Corollary 1.3 for the group $O_2(\mathbf{R})$. ❏

§2. Operators on Riemannian Surfaces

Let M be an oriented Riemannian surface and let P be the oriented orthonormal frame bundle over M with canonical projection $\pi: P \to M$. Namely,

$$P = \{(x, e_1, e_2) \mid x \in M, e_1, e_2 \in T_x(M)\}$$

where the pairs (e_1, e_2) are oriented orthonormal bases of $T_x(M)$ and $\pi(x, e_1, e_2) = x$. The right action of $SO_2(\mathbf{R})$ on P is given by

$$(x, e_1, e_2) \cdot R(s) = (x, e_1 \cos s + \sin s e_2, -e_1 \sin s + \cos s e_2),$$

where

$$R(s) = \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix}.$$

We are going to link up the representation theory of §1 with the derivative on P. We begin by expressing the Cartan connection form on P as

$$\omega = \begin{pmatrix} 0 & 0 & 0 \\ \theta_1 & 0 & -\alpha \\ \theta_2 & \alpha & 0 \end{pmatrix} \in A^1(P, \mathfrak{euc}_2(\mathbf{R})).$$

The forms α, θ_1, θ_2 are defined (cf. Chapter 6) by the equations

$$\begin{aligned}&\pi_*(\nu) = \theta_1(\nu)e_1 + \theta_2(\nu)e_2 \quad \text{for all } \nu \in T_{(x,u_1,u_2)}(P),\\ &d\theta_1 - \alpha \wedge \theta_2 = 0,\\ &d\theta_2 + \alpha \wedge \theta_1 = 0.\end{aligned}$$

Let X_0, X_1, X_2 denote the vector fields on P dual to α, θ_1, θ_2.

Lemma 2.1.

(i) $\pi_*(X_1) = e_1$, $\pi_*(X_2) = e_2$.

(ii) *If X is an arbitrary vector field on M, its horizontal lift (i.e., the unique lift on which α vanishes) at $(x, e_1, e_2) \in P$ is given by $\tilde{X} = (X \cdot e_1)X_1 + (X \cdot e_2)X_2$.*

Proof. (i) This follows from the equation $\pi_*(\nu) = \theta_1(\nu)e_1 + \theta_2(\nu)e_2$ and the fact that X_1, X_2 are dual to θ_1, θ_2.

(ii) Since $\alpha(X_1) = \alpha(X_2) = 0$, it follows that $(X \cdot e_1)X_1 + (X \cdot e_2)X_2$ is horizontal, and by (i) it is a lift of X. ■

Lemma 2.2.

(i) *Fix $p \in P$, and define $c: I \to P$ by $c(s) = p \cdot R(s)$. Then $c'(0) = X_{0p}$.*

(ii) *Let $F: P \to V_k$ transform according to $F(ph) = \rho_k(h^{-1})F(p)$ for $h \in SO_2(\mathbf{R})$. Then $(X_0F)(p) = k\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} F(p)$.*

(iii) *Let X be a vector field on M, and consider the functions $X \cdot e_i: P \to \mathbf{R}$, $i = 1, 2$. Then*

$$\begin{aligned}X_0\langle X, e_1\rangle &= \langle X, e_2\rangle,\\ X_0\langle X, e_2\rangle &= -\langle X, e_1\rangle.\end{aligned}$$

Proof. (i) Factor the curve c as the composite. Then

$$\begin{aligned}c^*\omega &= (\mu \circ (\mathrm{id} \times R) \circ \iota_2)^*\omega\\ &= \iota_2^* \circ (\mathrm{id}^* \times R^*) \circ \mu^*\omega\\ &= \iota_2^* \circ (\mathrm{id}^* \times R^*)(\mathrm{Ad}(R(s)^{-1})\pi_p^*\omega + \pi_H^*\omega_H)\\ &= \mathrm{Ad}(R(s)^{-1})(\pi_p \circ (\mathrm{id} \times R) \circ \iota_2)^*\omega + (\pi_H \circ (\mathrm{id} \times R) \circ \iota_2)^*\omega_H\\ &= \mathrm{Ad}(R(s)^{-1})(\pi_p \circ \iota_2)^*\omega + R^*\omega_H = R^*\omega_H,\end{aligned}$$

since $\pi_p \circ \iota_2$ is constant. Thus,

$$\begin{aligned}\omega(c'(0)) &= \omega\left(c_*\left(\frac{d}{ds}\right)\Big|_{s=0}\right) = c^*\omega\left(\frac{d}{ds}\right)\Big|_{s=0} = R^*\omega_H\left(\frac{d}{ds}\right)\Big|_{s=0}\\ &= \omega_H\left(R_*\frac{d}{ds}\right)\Big|_{s=0} = R_*\left(\frac{d}{ds}\right)\Big|_{s=0} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} = \omega(X_0).\end{aligned}$$

Thus, $c'(0) = X_0$.

(ii) Fix $p \in P$, and take c as in (i). Then by (i) we have $(X_0F)(p) = F(c(s))'\big|_{s=0}$. Now

$$F(c(s)) = F(p \cdot R(s)) = \rho_k(R(s))^{-1}F(p) = R(-ks)F(p),$$

so $(X_0F)(p) = R(-ks)'\big|_{s=0}F(p)$. But

$$\begin{aligned}R(-ks)'\big|_{s=0} &= \begin{pmatrix} \cos ks & \sin ks \\ -\sin ks & \cos ks \end{pmatrix}'\Bigg|_{s=0} = \begin{pmatrix} -k\sin ks & k\cos ks \\ -k\cos ks & -k\sin ks \end{pmatrix}\Bigg|_{s=0}\\ &= \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix}.\end{aligned}$$

(iii) Define $F\colon P \to \mathbf{R}^2$ by $F(p) = \begin{pmatrix} \langle X, e_1\rangle \\ \langle X, e_2\rangle \end{pmatrix}$, where $p = (x, e_1, e_2)$. Then

$$\begin{aligned} F(p\cdot R(s)) &= \begin{pmatrix} \langle X, \cos s e_1 + \sin s e_2\rangle \\ \langle X, -\sin s e_1 + \cos s e_2\rangle \end{pmatrix} \\ &= \begin{pmatrix} \cos s & \sin s \\ -\sin s & \cos s \end{pmatrix} \begin{pmatrix} \langle X, e_1\rangle \\ \langle X, e_2\rangle \end{pmatrix} = R(-s)F(p). \end{aligned}$$

Thus, F transforms as a function $F\colon P \to V_1$. Hence,

$$(X_0F)(p) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} F(p),$$

or

$$\begin{pmatrix} X_0\langle X, e_1\rangle \\ X_0\langle X, e_2\rangle \end{pmatrix} = \begin{pmatrix} \langle X, e_2\rangle \\ -\langle X, e_1\rangle \end{pmatrix}.$$ ❑

Proposition 2.3. *Let $\Gamma(k)$ denote the space of smooth functions $F\colon P \to V_k$ transforming according to $F(ph) = \rho_k(h^{-1})F(p)$ for $h \in H = SO_2(\mathbf{R})$.*[3] *Write $F \in \Gamma(k)$ as*

$$F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = F_1 e_k + F_2 f_k.$$

(i) *The universal covariant derivative $\tilde{D}F\colon P \to V_k \otimes \mathfrak{g}^*$ is given by*

$$\tilde{D}F = X_1(F) \otimes E_1^* + X_2(F) \otimes E_2^* + X_0(F) \otimes E_0^*.$$

(ii) *Decomposing $\tilde{D}F$ in terms of the components of $V_k \otimes \mathfrak{g}^* = V_{k-1} \oplus V_k \oplus V_{k+1}$ yields (when $k \neq 0$) the three operators*

$$\tilde{D}\colon \Gamma(k) \xrightarrow{(\partial, J_k, \bar\partial)} \Gamma(k-1) \oplus \Gamma(k) \oplus \Gamma(k+1),$$

where

$$\partial = \frac{1}{2}\begin{pmatrix} X_1 & X_2 \\ -X_2 & X_1 \end{pmatrix}, \quad J_k = k\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \bar\partial = \frac{1}{2}\begin{pmatrix} X_1 & -X_2 \\ X_2 & X_1 \end{pmatrix}.$$

Proof. For $\nu \in \mathfrak{g}$, we have $(\tilde{D}F)\nu = \omega^{-1}(\nu)F\colon P \to V_k$. In particular, $(\tilde{D}F)E_i = X_i(F)$, $0 \le i \le 2$, which verifies (i). Now calculate

[3] Of course, $\Gamma(k)$ may also be interpreted as the space of smooth sections of the associated 2-plane bundle $P \times_H V_k$.

$$\begin{aligned} \tilde{D}F &= X_0(F_1e_k + F_2f_k) \otimes E_0^* + X_1(F_1e_k + F_2f_k) \otimes E_1^* \\ &\quad + X_2(F_1e_k + F_2f_k) \otimes E_2^* \\ &= X_0(F_1)e_k \otimes E_0^* + X_0(F_2)f_k \otimes E_0^* \\ &\quad + X_1(F_1)e_k \otimes E_1^* + X_1(F_2)f_k \otimes E_1^* + X_2(F_1)e_k \otimes E_2^* \\ &\quad + X_2(F_2)f_k \otimes E_2^*. \end{aligned}$$

Breaking up this expression according to the decomposition of $V_k \otimes \mathfrak{g}^*$ indicated in Corollary 1.3 yields

$$\begin{aligned} \tilde{D}F &= X_0(F_1)e_k + X_0(F_2)f_k \\ &\quad + X_1(F_1)\left(\frac{1}{2}e_{k-1} + \frac{1}{2}e_{k+1}\right) + X_1(F_2)\left(\frac{1}{2}f_{k-1} + \frac{1}{2}f_{k+1}\right) \\ &\quad + X_2(F_1)\left(-\frac{1}{2}f_{k-1} + \frac{1}{2}f_{k+1}\right) + X_2(F_2)\left(\frac{1}{2}e_{k-1} - \frac{1}{2}e_{k+1}\right) \\ &= X_0(F_1)e_k + X_0(F_2)f_k \\ &\quad + \frac{1}{2}((X_1(F_1) + X_2(F_2))e_{k-1} + (X_1(F_2) - X_2(F_1))f_{k-1}) \\ &\quad + \frac{1}{2}((X_1(F_1) - X_2(F_2))e_{k+1} + (X_1(F_2) + X_2(F_1))f_{k+1}). \end{aligned}$$

This, together with Lemma 2.2, yields the result. ■

Corollary 2.4. *The covariant derivative $DF\colon P \to V_k \otimes \mathfrak{p}^*$ is given by*

$$DF = X_1(F) \otimes E_1^* + X_2(F) \otimes E_2^*$$

and decomposes as

$$\Gamma(k) \xrightarrow{(\partial, \bar\partial)} \Gamma(k-1) \oplus \Gamma(k+1).$$

Proof. This follows directly from the proposition. ■

Definition 2.5. The operator $\bar\partial$ is called the *Cauchy–Riemann operator.* ❀

The following lemma prepares the way for the calculation of ∂ and $\bar\partial$ in a gauge.

Lemma 2.6. *Fix a local section $\sigma(x) = (x, e_1, e_2)$ of P (so e_1, e_2 is an orthonormal frame on M).*

(i) *At points of P in the image of σ, we have $X_i = \sigma_*(e_i) - \alpha(\sigma_*(e_i))X_0$, $i = 1, 2$. Now let*

$$\begin{pmatrix} 0 & 0 & 0 \\ \theta_1 & 0 & -\alpha \\ \theta_2 & \alpha & 0 \end{pmatrix}$$

denote the infinitesimal gauge corresponding to σ.

(ii)

$$X_i(X\cdot e_j)=\begin{cases} e_i(X\cdot e_1)-\alpha(e_i)X\cdot e_2 & \text{if } j=1,\\ e_i(X\cdot e_2)-\alpha(e_i)X\cdot e_1 & \text{if } j=2.\end{cases}$$

Proof. (i) By Lemma 2.1(i), $\pi_*(X_i)=e_i$, $i=1,2$, so $\pi_*(X_i-\sigma_*(e_i))=0$. Thus, $X_i=\sigma_*(e_i)+\lambda X_0$ for some function λ. Applying α to this equation yields $0=\alpha(\sigma_*(e_i))+\lambda$.

(ii)

$$\begin{aligned} X_i(X\cdot e_j)&=(e_i-\alpha(e_i)X_0)(X\cdot e_j)\quad \text{by (i)}\\ &=e_i(X\cdot e_j)-\alpha(e_i)X_0(X\cdot e_j)\\ &=\begin{cases} e_i(X\cdot e_1)-\alpha(e_i)X\cdot e_2 & \text{if } j=1,\\ e_i(X\cdot e_2)-\alpha(e_i)X\cdot e_1 & \text{if } j=2,\end{cases}\end{aligned}$$

where the final equalities come from Corollary 2.2(iii). ■

Theorem 2.7. (i) *If we interpret the operator $\partial\colon\Gamma(1)\to\Gamma(0)$ as a map from vector fields on M to pairs of functions on M, it assumes the form, with respect to the gauge above,*

$$\partial F=\frac{1}{2}\begin{pmatrix} e_1(X\cdot e_1)+e_2(X\cdot e_2)-\alpha(e_2)X\cdot e_1-\alpha(e_1)X\cdot e_2\\ e_1(X\cdot e_2)-e_2(X\cdot e_1)+\alpha(e_2)X\cdot e_2-\alpha(e_1)X\cdot e_1\end{pmatrix}.$$

In particular, in the case when M is the Euclidean plane with the standard coordinates x_1, x_2, we may take $e_1=\partial/\partial x_1$, $e_2=\partial/\partial x_2$ (in which case, $\alpha=0$), and the last expression reduces to

$$\partial F=\frac{1}{2}\begin{pmatrix} \frac{\partial f_1}{\partial x_1}+\frac{\partial f_2}{\partial x_2}\\ \frac{\partial f_2}{\partial x_1}-\frac{\partial f_1}{\partial x_2}\end{pmatrix}$$

whose components are half of the usual expressions for divergence and curl.

(ii) *If we reinterpret $\bar\partial\colon\Gamma(0)\to\Gamma(1)$ as a map from $\mathbf{R}^2$-valued functions on M to vector fields on M, this operator takes the form*

$$\bar\partial f=\frac{1}{2}((e_1(f_1)-e_2(f_2))e_1+(e_2(f_1)+e_1(f_2))e_2).$$

The right side is the same no matter what orthonormal basis (e_1,e_2) is used. In particular, when M is the Euclidean (x_1,x_2) plane and $e_1=\partial/\partial x_1$, $e_2=\partial/\partial x_2$, the equation $\bar\partial f=0$ becomes the Cauchy–Riemann equations

$$\frac{\partial f_1}{\partial x_1}=\frac{\partial f_2}{\partial x_2},\quad \frac{\partial f_1}{\partial x_2}=-\frac{\partial f_2}{\partial x_1}.$$

Proof. (i) The vector field X on M corresponds (cf. Corollary 5.3.16) to the function $F\colon P\to V_1$ given by $F(x,e_1,e_2)=(X\cdot e_1,X\cdot e_2)$. Thus,

$$\begin{aligned}\partial F&=\frac{1}{2}\begin{pmatrix} X_1 & X_2\\ -X_2 & X_1\end{pmatrix}\begin{pmatrix} F_1\\ F_2\end{pmatrix}=\frac{1}{2}\begin{pmatrix} X_1(F_1)+X_2(F_2)\\ -X_2(F_1)+X_1(F_2)\end{pmatrix}\\ &=\frac{1}{2}\begin{pmatrix} X_1(X\cdot e_1)+X_2(X\cdot e_2)\\ -X_2(X\cdot e_1)+X_1(X\cdot e_2)\end{pmatrix}.\end{aligned}$$

By Lemma 2.6 we get

$$\partial F=\frac{1}{2}\begin{pmatrix} e_1(X\cdot e_1)-\alpha(e_1)X\cdot e_2+e_2(X\cdot e_2)-\alpha(e_2)X\cdot e_1\\ -e_2(X\cdot e_1)+\alpha(e_2)X\cdot e_2+e_1(X\cdot e_2)-\alpha(e_1)X\cdot e_1\end{pmatrix},$$

which yields the formula. Thus the final remark is clear.

(ii) Reinterpreting $f=\begin{pmatrix} f_1\\ f_2\end{pmatrix}\colon M\to\mathbf{R}^2$ as the element $F=\pi^*f=\begin{pmatrix}\pi^*f_1\\ \pi^*f_2\end{pmatrix}\in\Gamma(0)$, we calculate

$$\begin{aligned}\bar\partial F&=\frac{1}{2}\begin{pmatrix} X_1 & -X_2\\ X_2 & X_1\end{pmatrix}\begin{pmatrix}\pi^*f_1\\ \pi^*f_2\end{pmatrix}=\frac{1}{2}\begin{pmatrix} X_1(\pi^*f_1)-X_2(\pi^*f_2)\\ X_2(\pi^*f_1)+X_1(\pi^*f_2)\end{pmatrix}\\ &=\frac{1}{2}\begin{pmatrix}(\pi_*(X_1))f_1-(\pi_*(X_2))f_2\\ (\pi_*(X_2))f_1+(\pi_*(X_1))f_2\end{pmatrix}=\frac{1}{2}\begin{pmatrix} e_1(f_1)-e_2(f_2)\\ e_2(f_1)+e_1(f_2)\end{pmatrix}\\ &=\frac{1}{2}\varphi_{(x,e_1,e_2)}((e_1(f_1)-e_2(f_2))e_1+(e_2(f_1)+e_1(f_2))e_2).\end{aligned}$$

Interpreting this as a vector field on M gives the result. ■

We remark that the Cauchy–Riemann operator on a Riemannian surface is a special case of the Dirac operator on an even-dimensional manifold.

Appendix E

Characterization of Principal Bundles

We give a proof of the following result stated in Chapter 4.

Theorem 4.2.4. *Let P be a smooth manifold, H a Lie group, and $\mu: P \times H \to P$ a smooth, free, proper right action. Then*

(i) *P/H with the quotient topology is a topological manifold* ($\dim P/H = \dim P - \dim H$),

(ii) *P/H has a unique smooth structure for which the canonical projection $P \to P/H$ is a submersion,*

(iii) *$\xi = (P, \pi, M, H)$ is a smooth principal right H bundle.*

Proof. (i) *Step 1.* The orbits of H yield a foliation on P.

Since the composite $\mu \circ \lambda_h: P \to P \times H \to P$ sending $p \mapsto (p, h) \mapsto ph$ is a diffeomorphism for all $h \in H$, it follows that the action $\mu: P \times H \to P$ is a submersion and hence induces a foliation on $P \times H$ with leaves of the form

$$\mathcal{L}_p = \{(q, h) \in P \times H \mid qh = p\}.$$

The leaves of this foliation are permuted under right action of H on the second factor of $P \times H$. This foliation projects, via the first factor projection $P \times H \to P$, to a foliation on P whose leaves are the orbits of H.

Step 2. The leaves of the foliation on P are properly embedded submanifolds.

Define $\varphi_p: H \to P$ by $h \to ph$. Every leaf is the image of one of these maps. Since the map φ_p factors as $\varphi_p = \mu \circ \rho_p$, where $\rho_p: H \to P \times H$ sends $h \to (p, h)$, it is smooth; since the action is free, φ_p is injective.

If $K \subset P$ is a compact set, then $\varphi_p^{-1}(K) = \{h \in H \mid \{p\}h \cap K \neq \emptyset\}$, which is compact since $\{p\}$ and K are compact and the action is proper.

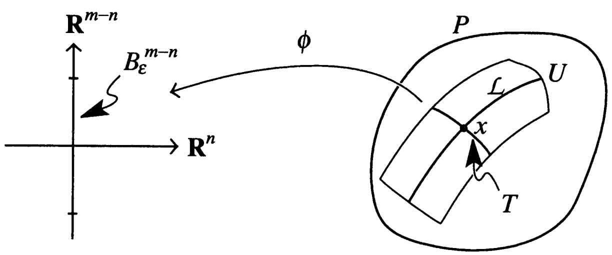
Finally, we show that $\varphi_p: H \to P$ is an immersion. If it is *not* an immersion, then for some $h \in H$ the map $\varphi_{p*h}: T_h(H) \to T_{ph}(P)$ is not injective. But then the commutativity of the diagrams (where $r_a: P \to P$ sends $q \mapsto qa$ for $a \in H$)

$$\begin{tikzcd} H \arrow[r, "\varphi_p"] \arrow[d, "R_a"] & P \arrow[d, "r_a"] \\ H \arrow[r, "\varphi_p"'] & P \end{tikzcd} \qquad \begin{tikzcd} T_h(H) \arrow[r, "\varphi_{p*h}"] \arrow[d, "R_{a*}", "\approx"'] & T_{ph}(P) \arrow[d, "r_{a*}", "\approx"'] \\ T_{ha}(H) \arrow[r, "\varphi_{p*ha}"'] & T_{pha}(P) \end{tikzcd}$$

shows that $\varphi_{p*e}: T_e(H) \to T_p(P)$ is not injective (take $a = h^{-1}$) and that if $\nu \in \ker \varphi_{p*e}$, then $R_{a*}\nu \in \operatorname{Ker} \varphi_{p*a}$ for all $a \in H$. Choose such a $\nu \neq 0$, and let $\psi: \mathbf{R} \to H$ be the one-parameter subgroup satisfying $\psi_{*e}(1) = \nu$. Let $s \in \mathbf{R}$, and set $h = \psi(s)$. Now the derivative of the composite $\varphi_p \circ \psi$ at $s \in \mathbf{R}$ is $\varphi_{p*}\psi_{*h}: T_s(\mathbf{R}) \to T_h(H) \to T_{ph}(P)$, which vanishes. Thus the composite $\varphi_p \circ \psi$ is constant, and in particular φ_p is not injective, which is a contradiction.

Step 3. For each point $p \in P$ there is a foliated chart (U, φ), with $\varphi: (U, p) \to (\mathbf{R}^n \times \mathbf{R}^{m-n}, 0)$ and a set $T = \varphi^{-1}(0 \times B_\varepsilon^{m-n})$, with $\varepsilon > 0$, such that the map $\alpha: T \times H \to P$ sending $(t, h) \mapsto th$ is a diffeomorphism onto its image.

Fix $p \in P$, and let $\mathcal{L}$ be the leaf through p. Since $\mathcal{L}$ is proper, we may choose a foliated chart (U, φ), with $\varphi: (U, U \cap \mathcal{L}, p) \to (\mathbf{R}^n \times \mathbf{R}^{m-n}, \mathbf{R}^n \times 0, 0 \times 0)$. Now we set $T = T_\varepsilon = \varphi^{-1}(0 \times B_\varepsilon^{m-n})$, where B_ε^{m-n} is a ball of radius ε about the origin in $\mathbf{R}^{m-n}$, so that T has compact closure.



We claim that there is a neighborhood of the identity $V \subset H$ such that $TV \subset U$. For if this fails, we can find a sequence of neighborhoods of the identity $V_n \subset H$, $n = 1, 2, \ldots$, which all have compact closures and satisfy

$$V_1 \supset \bar{V}_2 \supset V_2 \supset \bar{V}_3 \supset \cdots \supset \bigcap_{n=1}^{\infty} V_n = e$$

such that each set TV_n meets the complement of U. Now choose sequences $t_n \in T$ and $v_n \in V_n$ with $t_n v_n$ in the complement of U. Since T and V_n have compact closures, we may, by passing to subsequences, assume that t_n and v_n converge to $t \in \bar{T}$ and $e \in H$, respectively. Thus, $t_n v_n \to t$, which again lies in the complement of U. But this is impossible since $t \in \bar{T} \subset U$. This proves the existence of $V \subset H$ such that $TV \subset U$.

Replacing $V \subset H$ by its identity component, we may assume that it is an open path-connected neighborhood of the identity satisfying $TV \subset U$. Set $K_\varepsilon = K = \{h \in H \mid \bar{T}h \cap \bar{T} \neq \emptyset\}$. Now since $\bar{T}$ is compact and the action is proper, it follows that K is compact. We claim that $e \in K$ is an isolated point of K so that $K - \{e\}$ is again compact. It suffices to show that $V \cap K = \{e\}$. Let $h \in V \cap K$, and choose a path $\sigma: (I, 0, 1) \to (V, e, h)$. Since the path

$$p\sigma: (I, 0, 1) \to (U, p, ph)$$

lies in a single plaque and both $p, ph \in \bar{T}$, it follows that $ph = p$ so that $h = e$.

Next we claim that for some $\varepsilon > 0$, the map $\alpha = \alpha_\varepsilon: T_\varepsilon \times H \to P$ sending $(t, h) \to th$ is injective. Suppose not. Then we may choose, for each n, two unequal pairs of points $(t_n, h_n), (u_n, k_n) \in T_{1/n} \times H$ with $t_n h_n = u_n k_n$ for all n. Thus, $t_n = u_n \lambda_n$ (where $\lambda_n = k_n h_n^{-1}$) and hence $\lambda_n \in K_{1/n} \subset K_1$. Now if we were to have $t_n = u_n$, it would follow that $\lambda_n = e$, since the action is free, and then the pairs (t_n, h_n), (u_n, k_n) would be equal. Since we assume this fails, it follows that we must have $\lambda_n \in K_1 - \{e\}$ for all n. Since K_1 is compact, a subsequence of the λ_ns converges to some $\lambda_\infty \neq e$. But $t_n = u_n \lambda_n$ and u_n both converge to p; therefore, taking the limit, we get $p = p\lambda_\infty$, which implies that $\lambda_\infty = e$, since the action is free. This is a contradiction. Since for the choice of ε in the last paragraph the map α is a smooth injection, and the dimensions of its domain and range are the same, we need only show that it is an immersion to finish step 3. Because of the commutative diagram

$$\begin{tikzcd} T \times H \arrow[r, "\alpha"] \arrow[d, "\mathrm{id} \times R_a"'] & P \arrow[d, "r_a"] \\ T \times H \arrow[r, "\alpha"] & P \end{tikzcd}$$

it suffices to show that $\alpha_*: T_{(t,e)}(T \times H) \to T_t(P)$ is injective for all $t \in T$. But $T_{(t,e)}(T \times H) = T_t(T) \oplus T_e(H)$, and the restriction of α_* to the first factor is the inclusion-induced map $T_t(T) \to T_t(P)$, while the restriction to the second factor is a map $\varphi_t: T_e(H) \to T_t(P)$. These maps are both inclusions, and the images lie in subspaces having only 0 in common, so α_* is injective.

Step 4. The quotient space P/H is Hausdorff.

Let $p, q \in P$ lie in distinct H orbits. Choose a transversal T containing p as in step 3. We may assume that the oribit of q does not meet the closure of T, for it meets the coordinate system about p in at most one plaque, which we may avoid by shrinking the transversal T. Similarly, let U be a transversal containing q whose closure does not meet the orbit of p. Now the orbits $\bar{U}H$ form a closed set that meets T in a closed set T_0. Since $p \in T - T_0$ and the latter is open in T, we may further shrink T so that $\bar{T}H \cap \bar{U}H = \emptyset$. In particular, TH and UH are disjoint, so their images separate the images of p and q in P/H.

Step 5. The quotient space P/H is locally Euclidean of dimension equal to $\dim P - \dim H$. If T is one of the transversals through p guaranteed by step 3 and $\pi: P \to P/H$ is the canonical projection map, then $(\pi(T), (\pi \mid T)^{-1})$ is a coordinate chart for P/H.

It suffices to show that the induced map $\pi \mid T: T \to \pi(T)$ is a homeomorphism. It is certainly continuous, and by step 3 it is bijective. But it is also open, for if $U \subset T$ is open, then UH is open and hence $\pi(U)$ is open in $\pi(T)$. Finally, $\dim P/H = \dim T = \dim P - \dim H$.

(ii) The charts $(\pi(T), (\pi \mid T)^{-1})$ arising from the transversals of step 3 yield a smooth structure of G/H.

It suffices to show that two such charts, $(\pi(T), (\pi \mid T)^{-1})$ and $(\pi(U), (\pi \mid U)^{-1})$ say, are smoothly compatible. But $(\pi \mid T)^{-1}(\pi(T) \cap \pi(T)) = T \cap UH$ and $(\pi \mid U)^{-1}(\pi(T) \cap \pi(T)) = TH \cap U$. Moreover, the change of coordinate mapping is just the "sliding" map $T \cap UH \to TH \cap U$, which is known to be smooth from Chapter 2.

(iii) $\xi = (P, \pi, M, H)$ is a smooth principal right H bundle.

Let T be a transversal as in step 3. Then $\pi(T)$ is an open set in P/H. Define

$$\psi: \pi(T) \times H \to \pi^{-1}(\pi(T)) \quad \text{by } \psi(\pi(t), h) = th.$$

This was shown to be a diffeomorphism in step 3. Given two transversals T_1 and T_2, we get two maps ψ_1 and ψ_2 and the associated transition function

$$\psi_2^{-1}\psi_1: \pi(T_1) \cap \pi(T_2) \times H \to \pi(T_1) \cap \pi(T_2) \times H.$$

Define $f: T_1 \to H$ by setting $\psi_2^{-1}\psi_1(\pi(t), e) = (\pi(t), f(t))$. Then f is smooth and

$$\psi_2^{-1}\psi_1(\pi(t), h) = (\pi(t), f(t)^{-1}h).$$

Thus, the maps ψ constitute an H atlas for the bundle $P \to P/H$. ■

Bibliography

M.A. Akivis and V.V. Goldberg, *Conformal Differential Geometry and Its Generalizations*, Wiley Interscience, 1996.

M.A. Akivis and V.V. Goldberg, *Projective Differential Geometry of Submanifolds*, North-Holland, Amsterdam, 1993.

J. Beem and P. Ehrlich, *Global Lorentzian Geometry*, Dekker, New York, 1991.

M. Berger, *Geometry II*, Springer-Verlag, New York, 1987.

W. Boothby, *An Introduction to Differentiable Manifolds and Riemannian Geometry*, 2nd ed., Academic Press, San Diego, 1986.

R.L. Bryant and L. Hsu, Rigidity of integral curves of rank 2 distributions, *Invent. Math.* **114**, 435–461, 1993.

R. Bryant, S.-S. Chern, R. Gardner, H. Goldschmidt, and P.A. Griffiths, *Exterior Differential Systems*, Springer-Verlag, New York, 1991.

W.K. Bühler, *Gauss, A Biographical Study*, Springer-Verlag, New York, 1981.

G. Cairns and R. Sharpe, The inversive differential geometry of plane curves, *L'Enseignement Math.* **36**, 175–196, 1990.

G. Cairns, R. Sharpe, and L. Webb, Conformal invariants for curves and surfaces in three dimensional space, *Rocky Mountain Journal of Mathematics*, **24**, no. 3, 933–959, 1994.

E. Cartan, Les systèmes de Pfaff à cinq variables et les équations aux dérivées partielles du second ordre, *Ann. Ec. Norm.* **27**, 109–192, 1910.

E. Cartan, Les espaces à connexion conforme, *Ann. Soc. Pol. Math.* **2**, 171–221, 1923.

E. Cartan, Sur les variétés à connexion projective, *Bull. Soc. Math. France*, **52**, 205–241. See also the collected works, vol. IIIa, 825–861, 1924.

E. Cartan, Le calcul tensoriel en géométrie projective, *C.R. Acad. Sci.* **198**, 2033–2037, 1934.

E. Cartan, *La Methode du Repère Mobile, la Théorie des Groupes Continus et les Espaces Généralisés*, Exposés de Géométrie, V, Paris, Hermann, 1935.

E. Cartan, Le calcul tensoriel projectif, *Recueil Soc. Math.* Moscow, **42**, 131–148, 1935.

E. Cartan, L'extension du calcul tensoriel aux géométries non affines, *Ann. Math.* **38**, 1–13, 1937.

E. Cartan, *Leçons sur la Theorie des Espaces à Connexion Projective*, Paris, Gautier, 1937.

E. Cartan, Familles de surfaces isoparametriques dans les espaces à courbure constant, *Annali di Mat.* **17**, 177–191, 1938. Also in the collected works.

E. Cartan, *Leçons sur la Théorie des Spineurs I, II*, Exposés de Géometrie, XI, Hermann, Paris, 1938.

E. Cartan, *Selecta Jubilé Scientifique*, Paris, Gauthier, 1939.

E. Cartan, La notion d'orientation dans les différentes géométries, *Bull. Soc. Math. France*, **69**, 47–70, 1941.

E. Cartan, *The Theory of Spinors*, Hermann, Paris, 1966, translation of *Leçons sur la Théorie des Spineurs I, II*, Exposés de Geométrie, XI, Hermann, Paris, 1938.

E. Cartan and A. Einstein, *Letters on Absolute Parallelism 1929–1932*, Princeton University Press, Princeton, 1979.

T. Cecil, *Lie Sphere Geometry*, Springer-Verlag, New York, 1992.

T. Cecil and S.-S. Chern, Tautness and Lie sphere geometry, *Math. Ann.* **278**, 381–399, 1987.

S.-S. Chern, Sur la possibilité de plonger un espace à connexion projective donné dans un espace projectif, *Bull. Sci. Mat.* **61**, 234–243, 1937.

S.-S. Chern, The geometry of the differential equation $y''' = F(x, y, y', y'')$, *Science Reports Tsing Hua Univ.* **4**, 97–111, 1940.

S.-S. Chern, A generalization of the projective geometry of linear spaces, *Proc. Nat. Acad. of Sci. USA*, **29**, 33–43, 1943.

S.-S. Chern, The geometry of G-structures, *Bull. Amer. Math. Soc.* **72**, 167–219, 1966.

S.-S. Chern, On the projective structure of real hypersurfaces in $\mathbf{C}^{n+1}$, *Math. Scandinavia*, **36**, 74–82, 1975.

S.-S. Chern, Circle bundles, in *Geometry and Topology, III Latin Amer. School of Math.*, Lecture Notes in Math., Springer-Verlag, **597**, 114–131, 1977.

S.-S. Chern, *Selected Papers*, Springer-Verlag, New York, 1978.

S.-S. Chern, From triangles to manifolds, *Amer. Math. Monthly*, **86**, 339–349, 1979.

S.-S. Chern, Finsler geometry is just Riemannian geometry without the quadratic condition, *Notices of the AMS*, **43**, no. 9, 959–963, 1996.

S.-S. Chern and C. Chevalley, Elie Cartan and his mathematical work, *Bull. AMS*, **58**, 217–250, 1952.

S.-S. Chern and P.A. Griffiths, An inequality for the rank of a web and webs of maximal rank, *Annali Sc. Norm. Sup.-Pisa*, Serie IV, **5**, 539–557, 1978.

S.-S. Chern and J. Wolfson, A simple proof of Frobenius theorem, *Manifolds and Lie Groups, Papers in Honor of Y. Matsushima*, Birkhäuser, 347–349, 1981.

K. DePaepe, *Primitive Effective Pairs of Lie Algebras over a Field of Characteristic Zero*, Ph.D. Thesis, University of Toronto, 1996.

J. Dieudonné, *Treatise on Analysis*, **4**, translated by I.G. Macdonald, Pure and Applied Mathematics, Vol. 10-IV, Academic Press, New York-London, 1974.

C. Dodson and T. Poston, *Tensor Geometry*, 2nd ed., Springer-Verlag, New York, 1991.

S. Donaldson and P.B. Kronheimer, *The Geometry of Four-Manifolds*, Oxford Science Publications, 1990.

J. Dugundji, *Topology*, Allyn & Bacon Inc., Boston, 1966.

C. Ehresmann, Sur les espaces localement homogènes, *L'Enseignement Math.* **35**, 317–333, 1936.

C. Ehresmann, Les connexions infinitésimales dans un espace fibré differentiable, *Colloque de Topologie, Bruxelles*, 29–55, Thone, Liège, 1951.

C. Ehresmann, Structures feuilletées, *Proc. Vth Can. Congress*, Montreal, 109–172, 1961.

A. Einstein, Ether and Relativity, in *Sidelights on Relativity*, translated by G.B. Jeffrey and W. Perrett, Methuin, 1922.

L.P. Eisenhart, *Non-Riemannian Geometry*, AMS Colloquium Publications VIII, fifth printing, 1964.

D. Ferus, H. Karcher, and H. Munzner, Cliffordalgebren und neue isoparametrische Hyperflaschen, *Math. Z.* **177**, 479–502, 1981.

A. Fialkow, Conformal differential geometry of a subspace, *T.A.M.S.* **56**, 309–433, 1944.

B.L. Foster, Would Liebniz Lie to You? *Math. Intelligencer*, **8**, no. 3, 34–40, 1986.

D. Freed and K. Uhlenbeck, *Instantons and Four-Manifolds*, Springer-Verlag, Berlin, New York, 1984.

W. Fulton and J. Harris, *Representation Theory, A First Course*, Springer-Verlag Graduate Texts in Mathematics, New York, 1991.

R.B. Gardner, *The Method of Equivalence and Its Applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, **58**, Society for Industrial and Applied Mathematics, Philadelphia, 1989.

M. Golubitski, Primitive actions and maximal subgroups of Lie groups, *J. Diff. Geometry*, **7**, 175–191, 1972.

R. Guénon, *The Reign of Quantity and the Signs of the Times*, Penguin Metaphyscial Library. This is a reprint of the 1952 translation by Lord Northbourne of the French original published by Édition Gallimard, Paris, 1945.

D. Howard and J. Stachel, ed., *Einstein and the History of General Relativity*, Birkháuser, Boston, 1989.

J. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Springer-Verlag, New York, 1972.

D. Husemoller, *Fiber Bundles*, McGraw-Hill, New York, 1966.

H. Jacobowitz, *An Introduction to CR Structures*, Math. Surveys and Monographs, **32**, Amer. Math. Soc., Providence, RI, 1990.

N. Jacobson, *Lectures in Abstract Algebra*, **2**, Van Nostrand, London, 1953.

N. Kamran, K. Lamb, and W. Shadwick, The local equivalence problem for $y'' = F(x, y, y')$ and the Panlevé transcendants, *J. of Diff. Geometry*, **22**, 139–150, 1985.

B. Kamté, *Cercles généralisés et complétude dans la géométrie de Cartan*, Master's thesis, University of Quebec at Montreal, 1995.

R. Kirby, *The Topology of Four Manifolds*, Springer-Verlag, New York, 1989.

F. Klein, from Klein's Erlangen Program of 1872, English translation "A Comparative Review of Recent Researches in Geometry" translated by M.W. Haskell, *Bull. AMS* (formally *Bull. of the New York Math. Soc.*), July 1893.

S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, Vol. 1*, Interscience, New York, 1963.

B. Komrakov, A. Churyumov, and B. Doubrov, *Two Dimensional Homogeneous Spaces*, Preprint Series, Inst. of Math., Univ. of Oslo, 1993.

A. Kosinski, *Differential Manifolds*, Academic Press, Boston, 1993.

Ursula K. Le Guin, Excerpt from *The World is a Forest*, ©1972 by Ursula K. Le Guin; first appeared in *Again, Dangerous Visions*; reprinted by permission of the author and the author's agent, Virginia Kidd.

L. Loomis and S. Sternberg, *Advanced Calculus*, Addison-Wesley, 1968.

Y. Manin, Strings, *Math. Intelligencer*, **11**, no. 2, 59–65, 1989.

P. Meyer, Qu'est ce qu'une differentielle d'ordre n? *Expo. Math.* **7**, 249–264, 1989.

J. Milnor, *Topology from the Differentiable Viewpoint*, Univ. of Virginia Press, Charlottesville, Virginia, 1965.

P. Molino, *Riemannian Foliations*, Birkhäuser, Boston, 1988.

M. Monastyrski, *Riemann, Topology and Physics*, Birkhäuser, Boston, 1987.

H. Ozeki and M. Takeuchi, On some types of isoparametric hypersurfaces in spheres I, *Tôkohu Math. J.* **127**, no. 2, 515–559, 1975.

H. Ozeki and M. Takeuchi, On some types of isoparametric hypersurfaces in spheres II, *Tôkohu Math. J.* **128**, no. 2, 7–55, 1976.

H. Poincaré, *Science and Hypothesis*, Dover, 1952, page 50. This is a translation of the French original published in 1902.

W.A. Poor, *Differential Geometric Structures*, McGraw-Hill, New York, 1981.

M. Postnikov, *Lectures in Geometry, Semester V, Lie Groups and Lie Algebras*, Mir, Moscow, 1986.

E. Ruh, Cartan connections, *Proc. Sympos. Pure Math.* **54**, part 3, A.M.S., Providence, RI, 1993.

P. Samuel, *Projective Geometry*, Undergradute Texts in Mathematics, Springer-Verlag, New York, 1988.

I. Singer and J. Thorpe, The curvature of 4-dimensional Einstein spaces, 355–366, in *Global Analysis, Papers in Honor of Kunihiko Kodaira*, edited by D. Spencer and S. Iyanaga, Princeton Math. Series No. 29, 1969.

M. Spivak, *A Comprehensive Introduction to Differential Geometry, Vol. 1 & 2*, second edition, Publish or Perish Press, Houston, Texas, 1970 and 1979.

N. Tanaka, On the equivalence problem associated with simple graded Lie algebras, *Hokkaido Math. J.* **8**, 23–84, 1978.

C.-L. Terng, Recent progess in submanifold geometry, *Proc. of Symp. in Pure Math.* **54**, Part 1, 1993.

G. Thorbergson, Isoparametric foliations and their buildings, *Ann. Math.* **133**, 429–446, 1991.

H. Weyl, *Space, Time, Matter*, Dover Publications, New York, 1952 (a reprint of Weyl's book, the first edition of which was originally published in German in 1918).

J.E. White, *The Method of Iterated Tangents with Applications in Local Riemannian Geometry*, Pitmann, Boston, 1982.

J.H.C. Whitehead, Locally homogeneous spaces in differential geometry, *Ann. of Math.* **33**, no. 2, 681–687, 1932.

H. Whitney, Differentiable Manifolds, *Annals of Math.* **37**, no. 3, 645–680, 1936.

T. Willmore, Note on embedded surfaces, *An Stiint. Univ. "Al I. Cusa" Iasi s. I. a. Mat.* **11**, 493–496, 1965.

T. Willmore, *Riemannian Geometry*, Clarendon, Oxford, 1993.

C.N. Yang, Magnetic monopoles, fibre bundles and gauge fields, *An. New York Acad. Sci.* **294**, 86–97, 1977.

H. Yamabe, On an arcwise connected subgroup of a Lie group, *Osaka Math. J.* **2**, no. 1, 13–14, 1950.

K. Yamaguchi, Differential systems associated with simple graded Lie algebras, in *Progress in Differential Geometry*, 413–494, in the series Advanced Studies in Mathematics, **22**, ed. by K. Shiohama, Math. Soc. of Japan, 1993.

C.T. Yen, Sur la connexion projective normale associée à un feuilletage du deuxieme ordre, *Ann. Mat. Pura Appl.* **34**, no. 4, 55–94, 1953.

Index