Graduate Texts in Mathematics

Serge Lang

Fundamentals of Differential Geometry

With 22 Illustrations

Springer

Serge Lang
Department of Mathematics
Yale University
New Haven, CT 06520
USA

Mathematics Subject Classification (1991): 58-01

Library of Congress Cataloging-in-Publication Data
Lang, Serge, 1927–
Fundamentals of differential geometry / Serge Lang.
p. cm. — (Graduate texts in mathematics : 191)
Includes bibliographical references and index.
ISBN 0-387-98593-X (alk. paper)
1. Geometry, Differential. I. Title. II. Series.
QA641.L33 1999
516.3′6—dc21 98-29993

Printed on acid-free paper.

Production coordinated by Brian Howe and managed by Terry Kornak; manufacturing supervised by Jeffrey Taub.
Typeset by Asco Trade Typesetting Ltd., Hong Kong.
Printed and bound by Edwards Brothers, Inc., Ann Arbor, MI.
Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-98593-X Springer-Verlag New York Berlin Heidelberg SPIN 10687464

Foreword

The present book aims to give a fairly comprehensive account of the fundamentals of differential manifolds and differential geometry. The size of the book influenced where to stop, and there would be enough material for a second volume (this is not a threat).

At the most basic level, the book gives an introduction to the basic concepts which are used in differential topology, differential geometry, and differential equations. In differential topology, one studies for instance homotopy classes of maps and the possibility of finding suitable differentiable maps in them (immersions, embeddings, isomorphisms, etc.). One may also use differentiable structures on topological manifolds to determine the topological structure of the manifold (for example, à la Smale [Sm 67]). In differential geometry, one puts an additional structure on the differentiable manifold (a vector field, a spray, a 2-form, a Riemannian metric, ad lib.) and studies properties connected especially with these objects. Formally, one may say that one studies properties invariant under the group of differentiable automorphisms which preserve the additional structure. In differential equations, one studies vector fields and their integral curves, singular points, stable and unstable manifolds, etc. A certain number of concepts are essential for all three, and are so basic and elementary that it is worthwhile to collect them together so that more advanced expositions can be given without having to start from the very beginnings.

Those interested in a brief introduction could run through Chapters II, III, IV, V, VII, and most of Part III on volume forms, Stokes' theorem, and integration. They may also assume all manifolds finite dimensional.

Charts and local coordinates. A chart on a manifold is classically a representation of an open set of the manifold in some euclidean space.

Using a chart does not necessarily imply using coordinates. Charts will be used systematically. It will be observed equally systematically that finite dimensionality is hereby not used.

It is possible to lay down *at no extra cost* the foundations (and much more beyond) for manifolds modeled on Banach or Hilbert spaces rather than finite dimensional spaces. In fact, it turns out that the exposition gains considerably from the systematic elimination of the indiscriminate use of local coordinates $x_1, \dots, x_n$ and $dx_1, \dots, dx_n$. These are replaced by what they stand for, namely isomorphisms of open subsets of the manifold on open subsets of Banach spaces (local charts), and a local analysis of the situation which is more powerful and equally easy to use formally. In most cases, the finite dimensional proof extends at once to an invariant infinite dimensional proof. Furthermore, in studying differential forms, one needs to know only the definition of multilinear continuous maps. An abuse of multilinear algebra in standard treatises arises from an unnecessary double dualization and an abusive use of the tensor product.

I don't propose, of course, to do away with local coordinates. They are useful for computations, and are also especially useful when integrating differential forms, because the $dx_1 \wedge \cdots \wedge dx_n$. corresponds to the $dx_1 \cdots dx_n$ of Lebesgue measure, in oriented charts. Thus we often give the local coordinate formulation for such applications. Much of the literature is still covered by local coordinates, and I therefore hope that the neophyte will thus be helped in getting acquainted with the literature. I also hope to convince the expert that nothing is lost, and much is gained, by expressing one's geometric thoughts without hiding them under an irrelevant formalism.

I am aware of a widespread apprehensive reaction the moment some geometers or students see the words "Banach space" or "Hilbert manifold". As a possible palliative, I suggest reading the material assuming from the start that Banach space means finite dimensional space over the reals, and Hilbert manifold or Riemannian manifold means a finite dimensional manifold with a metric, with the local constant model being ordinary euclidean space. These assumptions will not make any proof shorter.

One major function of finding proofs valid in the infinite dimensional case is to provide proofs which are especially natural and simple in the finite dimensional case. Even for those who want to deal only with finite dimensional manifolds, I urge them to consider the proofs given in this book. In many cases, proofs based on coordinate free local representations in charts are clearer than proofs which are replete with the claws of a rather unpleasant prying insect such as Γ^i_{jkl}. Indeed, the bilinear map associated with a spray (which is the quadratic map corresponding to a symmetric connection) satisfies quite a nice local formalism in charts. I think the local representation of the curvature tensor as in Proposition 1.2 of Chapter IX shows the efficiency of this formalism and its superiority over local coordinates. Readers may also find it instructive to compare the proof of Proposition 2.6 of Chapter IX concerning the rate of growth of Jacobi fields with more classical ones involving coordinates as in [He 78], pp. 71–73.

Applications in Infinite Dimension

It is profitable to deal with infinite dimensional manifolds, modeled on a Banach space in general, a self-dual Banach space for pseudo Riemannian geometry, and a Hilbert space for Riemannian geometry. In the standard pseudo Riemannian and Riemannian theory, readers will note that the differential theory works in these infinite dimensional cases, with the Hopf–Rinow theorem as the single exception, but not the Cartan–Hadamard theorem and its corollaries. Only when one comes to dealing with volumes and integration does finite dimensionality play a major role. Even if via the physicists with their Feynman integration one eventually develops a coherent analogous theory in the infinite dimensional case, there will still be something special about the finite dimensional case.

The failure of Hopf–Rinow in the infinite dimensional case is due to a phenomenon of positive curvature. The validity of Cartan–Hadamard in the case of negative curvature is a very significant fact, and it is only recently being realized as providing a setting for major applications. It is a general phenomenon that spaces parametrizing certain structures are actually infinite dimensional Cartan–Hadamard spaces, in many contexts, e.g. Teichmuller spaces, spaces of Riemannian metrics, spaces of Kähler metrics, spaces of connections, spaces associated with certain partial differential equations, ad lib. Cf. for instance the application to the KdV equation in [ScTZ 96], and the comments at the end of Chapter XI, §3 concerning other applications.

Actually, the use of infinite dimensional manifolds in connection with Teichmuller spaces dates back some time, because as shown by Bers, these spaces can be embedded as submanifolds of a complex Banach space. Cf. [Ga 87], [Vi 73]. Viewing these as Cartan–Hadamard manifolds comes from newer insights.

For further comments on some recent aspects of the use of infinite dimension, including references to Klingenberg's book [Kl 83/95], see the introduction to Chapter XIII.

Of course, there are other older applications of the infinite dimensional case. Some of them are to the calculus of variations and to physics, for instance as in Abraham–Marsden [AbM 78]. It may also happen that one does not need formally the infinite dimensional setting, but that it is useful to keep in mind to motivate the methods and approach taken in various directions. For instance, by the device of using curves, one can reduce what is a priori an infinite dimensional question to ordinary calculus in finite dimensional space, as in the standard variation formulas given in Chapter XI, §1.

Similarly, the proper domain for the geodesic part of Morse theory is the loop space (or the space of certain paths), viewed as an infinite dimensional manifold, but a substantial part of the theory can be developed without formally introducing this manifold. The reduction to the finite dimensional case is of course a very interesting aspect of the situation, from which one can deduce deep results concerning the finite dimensional manifold itself, but it stops short of a complete analysis of the loop space. (Cf. Boot [Bo 60], Milnor [Mi 63].) See also the papers of Palais [Pa 63] and Smale [Sm 64].

In addition, given two finite dimensional manifolds X, Y it is fruitful to give the set of differentiable maps from X to Y an infinite dimensional manifold structure, as was started by Eells [Ee 58], [Ee 59], [Ee 61], [EeS 64], and [Ee 66]. By so doing, one transcends the purely formal translation of finite dimensional results getting essentially new ones, which would in turn affect the finite dimensional case. For other connections with differential geometry, see [El 67].

Foundations for the geometry of manifolds of mappings are given in Abraham's notes of Smale's lectures [Ab 60] and Palais's monograph [Pa 68].

For more recent applications to critical point theory and submanifold geometry, see [PaT 88].

In the direction of differential equations, the extension of the stable and unstable manifold theorem to the Banach case, already mentioned as a possibility in earlier versions of *Differential Manifolds*, was proved quite elegantly by Irwin [Ir 70], following the idea of Pugh and Robbin for dealing with local flows using the implicit mapping theorem in Banach spaces. I have included the Pugh–Robbin proof, but refer to Irwin's paper for the stable manifold theorem which belongs at the very beginning of the theory of ordinary differential equations. The Pugh–Robbin proof can also be adjusted to hold for vector fields of class H^p (Sobolev spaces), of importance in partial differential equations, as shown by Ebin and Marsden [EbM 70].

It is a standard remark that the C^∞-functions on an open subset of a euclidean space do not form a Banach space. They form a Fréchet space (denumerably many norms instead of one). On the other hand, the implicit function theorem and the local existence theorem for differential equations are not true in the more general case. In order to recover similar results, a much more sophisticated theory is needed, which is only beginning to be developed. (Cf. Nash's paper on Riemannian metrics [Na 56], and subsequent contributions of Schwartz [Sc 60] and Moser [Mo 61].) In particular, some additional structure must be added (smoothing operators). Cf. also my Bourbaki seminar talk on the subject [La 61]. This goes beyond the scope of this book, and presents an active topic for research.

On the other hand, for some applications, one may complete the C^∞-space under a suitable Hilbert space norm, deal with the resulting Hilbert manifold, and then use an appropriate regularity theorem to show that solutions of the equation under study actually are C^∞.

I have emphasized differential aspects of differential manifolds rather than topological ones. I am especially interested in laying down basic material which may lead to various types of applications which have arisen since the sixties, vastly expanding the perspective on differential geometry and analysis. For instance, I expect the books [BGV 92] and [Gi 95] to be only the first of many to present the accumulated vision from the seventies and eighties, after the work of Atiyah, Bismut, Bott, Gilkey, McKean, Patodi, Singer, and many others.

Negative Curvature

Most texts emphasize positive curvature at the expense of negative curvature. I have tried to redress this imbalance. In algebraic geometry, it is well recognized that negative curvature amounts more or less to "general type". For instance, curves of genus 0 are special, curves of genus 1 are semispecial, and curves of genus $\geqq 2$ are of general type. Thus I have devoted an entire chapter to the fundamental example of a space of negative curvature. Actually, I prefer to work with the Riemann tensor. I use "curvature" simply as a code word which is easily recognizable by people in the field. Furthermore, I include a complete account of the equivalence between seminegative curvature, the metric increasing property of the exponential map, and the Bruhat–Tits semiparallelogram law. Third, I emphasize the Cartan–Hadamard further by giving a version for the normal bundle of a totally geodesic submanifold. I am indebted to Wu for valuable mathematical and historical comments on this topic.

There are several current directions whereby spaces of negative curvature are the fundamental building blocks of some theories. They are quotients of Cartan–Hadamard spaces. I myself got interested in differential geometry because of the joint work with Jorgenson, which naturally led us to such spaces for the construction and theory of certain zeta functions. Quite generally, we were led to consider spaces which admit a stratification such that each stratum is a quotient of a Cartan–Hadamard space (especially a symmetric space) by a discrete group. That such stratifications exist very widely is a fact not generally taken into account. For instance, it is a theorem of Griffiths that given an algebraic variety over the complex numbers, there exists a proper Zariski closed subset whose complement is a quotient of a complex bounded domain, so in this way, every algebraic variety admits a stratification as above, even with constant negative curvature. Thurston's approach to 3-manifolds could be viewed from our perspective also. The general problem then arises how zeta functions, spectral invariants, homotopy and homology invariants, ad lib. behave with respect to stratifications, whether additively or otherwise. In the Jorgenson–Lang program, we associate a zeta function to each stratum, and the zeta functions of lower strata are the principal fudge factors in the functional equation of the zeta function associated to the main stratum. The spectral expansion of the heat kernel amounts to a theta relation, and we get the zeta function by taking the Gauss transform of the theta relation.

From a quite different perspective, certain natural "moduli" spaces for structures on finite dimensional manifolds have a very strong tendency to be Cartan–Hadamard spaces, for instance the space of Riemannian metrics, spaces of Kahler metrics, spaces of connections, etc. which deserve to be incorporated in a general theory.

In any case, I find the exclusive historical emphasis at the foundational level on positive curvature, spheres, projective spaces, grassmanians, at the expense of quotients of Cartan–Hadamard spaces, to be misleading as to the way manifolds are built up. Time will tell, but I don't think we'll have to wait very long before a radical change of view point becomes prevalent.

New Haven, 1998 SERGE LANG

Acknowledgments

I have greatly profited from several sources in writing this book. These sources include some from the 1960s, and some more recent ones.

First, I originally profited from Dieudonné's *Foundations of Modern Analysis*, which started to emphasize the Banach point of view.

Second, I originally profited from Bourbaki's *Fascicule de résultats* [Bou 69] for the foundations of differentiable manifolds. This provides a good guide as to what should be included. I have not followed it entirely, as I have omitted some topics and added others, but on the whole, I found it quite useful. I have put the emphasis on the differentiable point of view, as distinguished from the analytic. However, to offset this a little, I included two analytic applications of Stokes' formula, the Cauchy theorem in several variables, and the residue theorem.

Third, Milnor's notes [Mi 58], [Mi 59], [Mi 61] proved invaluable. They were of course directed toward differential topology, but of necessity had to cover ad hoc the foundations of differentiable manifolds (or, at least, part of them). In particular, I have used his treatment of the operations on vector bundles (Chapter III, §4) and his elegant exposition of the uniqueness of tubular neighborhoods (Chapter IV, §6, and Chapter VII, §4).

Fourth, I am very much indebted to Palais for collaborating on Chapter IV, and giving me his exposition of sprays (Chapter IV, §3). As he showed me, these can be used to construct tubular neighborhoods. Palais also showed me how one can recover sprays and geodesics on a Riemannian manifold by making direct use of the canonical 2-form and the metric (Chapter VII, §7). This is a considerable improvement on past expositions.

In the direction of differential geometry, I found Berger–Gauduchon–Mazet [BGM 71] extremely valuable, especially in the way they lead to the study of the Laplacian and the heat equation. This book has been very influential, for instance for [GHL 87/93], which I have also found useful.

I also found useful Klingenberg's book [Kl 83/95], see especially chapter XIII. I am very thankful to Karcher and Wu for instructing me on several matters, including estimates for Jacobi fields. I have also benefited from Helgason's book [Hel 84], which contains some material of interest independently of Lie groups, concerning the Laplacian. I am especially indebted to Wu's invaluable guidance in dealing with the trace of the second fundamental form and its application to the Laplacian, giving rise to a new exposition of some theorems of Helgason in Chapters XIV and XV.

Contents

CHAPTER I

Differential Calculus

We shall recall briefly the notion of derivative and some of its useful properties. As mentioned in the foreword, Chapter VIII of Dieudonné's book or my books on analysis [La 83], [La 93] give a self-contained and complete treatment for Banach spaces. We summarize certain facts concerning their properties as topological vector spaces, and then we summarize differential calculus. *The reader can actually skip this chapter* and start immediately with Chapter II if the reader is accustomed to thinking about the derivative of a map as a linear transformation. (In the finite dimensional case, when bases have been selected, the entries in the matrix of this transformation are the partial derivatives of the map.) We have repeated the proofs for the more important theorems, for the ease of the reader.

It is convenient to use throughout the language of categories. The notion of category and morphism (whose definitions we recall in §1) is designed to abstract what is common to certain collections of objects and maps between them. For instance, topological vector spaces and continuous linear maps, open subsets of Banach spaces and differentiable maps, differentiable manifolds and differentiable maps, vector bundles and vector bundle maps, topological spaces and continuous maps, sets and just plain maps. In an arbitrary category, maps are called morphisms, and in fact the category of differentiable manifolds is of such importance in this book that from Chapter II on, we use the word morphism synonymously with differentiable map (or p-times differentiable map, to be precise). All other morphisms in other categories will be qualified by a prefix to indicate the category to which they belong.

I, §1. CATEGORIES

A **category** is a collection of objects $\{X, Y, \ldots\}$ such that for two objects X, Y we have a set $\mathrm{Mor}(X, Y)$ and for three objects X, Y, Z a mapping (composition law)

$$\mathrm{Mor}(X, Y) \times \mathrm{Mor}(Y, Z) \to \mathrm{Mor}(X, Z)$$

satisfying the following axioms:

CAT 1. *Two sets* $\mathrm{Mor}(X, Y)$ *and* $\mathrm{Mor}(X', Y')$ *are disjoint unless* $X = X'$ *and* $Y = Y'$, *in which case they are equal.*

CAT 2. *Each* $\mathrm{Mor}(X, X)$ *has an element* id_X *which acts as a left and right identity under the composition law.*

CAT 3. *The composition law is associative.*

The elements of $\mathrm{Mor}(X, Y)$ are called **morphisms**, and we write frequently $f: X \to Y$ for such a morphism. The composition of two morphisms f, g is written fg or $f \circ g$.

A **functor** $\lambda: \mathfrak{A} \to \mathfrak{A}'$ from a category $\mathfrak{A}$ into a category $\mathfrak{A}'$ is a map which associates with each object X in $\mathfrak{A}$ an object $\lambda(X)$ in $\mathfrak{A}'$, and with each morphism $f: X \to Y$ a morphism $\lambda(f): \lambda(X) \to \lambda(Y)$ in $\mathfrak{A}'$ such that, whenever f and g are morphisms in $\mathfrak{A}$ which can be composed, then $\lambda(fg) = \lambda(f)\lambda(g)$ and $\lambda(\mathrm{id}_X) = \mathrm{id}_{\lambda(X)}$ for all X. This is in fact a covariant functor, and a contravariant functor is defined by reversing the arrows (so that we have $\lambda(f): \lambda(Y) \to \lambda(X)$ and $\lambda(fg) = \lambda(g)\lambda(f)$).

In a similar way, one defines functors of many variables, which may be covariant in some variables and contravariant in others. We shall meet such functors when we discuss multilinear maps, differential forms, etc.

The functors of the same variance from one category $\mathfrak{A}$ to another $\mathfrak{A}'$ form themselves the objects of a category $\mathrm{Fun}(\mathfrak{A}, \mathfrak{A}')$. Its morphisms will sometimes be called **natural transformations** instead of functor morphisms. They are defined as follows. If λ, μ are two functors from $\mathfrak{A}$ to $\mathfrak{A}'$ (say covariant), then a natural transformation $t: \lambda \to \mu$ consists of a collection of morphisms

$$t_X: \lambda(X) \to \mu(X)$$

as X ranges over $\mathfrak{A}$, which makes the following diagram commutative for any morphism $f: X \to Y$ in $\mathfrak{A}$:

$$\begin{array}{ccc} \lambda(X) & \xrightarrow{t_X} & \mu(X) \\ {\scriptstyle \lambda(f)}\downarrow & & \downarrow{\scriptstyle \mu(f)} \\ \lambda(Y) & \xrightarrow[t_Y]{} & \mu(Y) \end{array}$$

In any category $\mathfrak{A}$, we say that a morphism $f: X \to Y$ is an **isomorphism** if there exists a morphism $g: Y \to X$ such that fg and gf are the identities. For instance, an isomorphism in the category of topological spaces is called a topological isomorphism, or a homeomorphism. In general, we describe the category to which an isomorphism belongs by means of a suitable prefix. In the category of sets, a set-isomorphism is also called a bijection.

If $f: X \to Y$ is a morphism, then a **section** of f is defined to be a morphism $g: Y \to X$ such that $f \circ g = \mathrm{id}_Y$.

I, §2. TOPOLOGICAL VECTOR SPACES

The proofs of all statements in this section, including the Hahn–Banach theorem and the closed graph theorem, can be found in [La 93].

A **topological vector space** $\mathbf{E}$ (over the reals $\mathbf{R}$) is a vector space with a topology such that the operations of addition and scalar multiplication are continuous. It will be convenient to assume also, as part of the definition, that the space is **Hausdorff**, and **locally convex**. By this we mean that every neighborhood of 0 contains an open neighborhood U of 0 such that, if x, y are in U and $0 \leqq t \leqq 1$, then $tx + (1-t)y$ also lies in U.

The topological vector spaces form a category, denoted by TVS, if we let the morphisms be the continuous linear maps (by linear we mean throughout $\mathbf{R}$-linear). The set of continuous linear maps of one topological vector space $\mathbf{E}$ into $\mathbf{F}$ is denoted by $L(\mathbf{E}, \mathbf{F})$. The continuous r-multilinear maps

$$\psi: \mathbf{E} \times \cdots \times \mathbf{E} \to \mathbf{F}$$

of $\mathbf{E}$ into $\mathbf{F}$ will be denoted by $L^r(\mathbf{E}, \mathbf{F})$. Those which are symmetric (resp. alternating) will be denoted by $L^r_s(\mathbf{E}, \mathbf{F})$ or $L^r_{\mathrm{sym}}(\mathbf{E}, \mathbf{F})$ (resp. $L^r_a(\mathbf{E}, \mathbf{F})$). The isomorphisms in the category TVS are called **toplinear** isomorphisms, and we write $\mathrm{Lis}(\mathbf{E}, \mathbf{F})$ and $\mathrm{Laut}(\mathbf{E})$ for the toplinear isomorphisms of $\mathbf{E}$ onto $\mathbf{F}$ and the toplinear automorphisms of $\mathbf{E}$.

We find it convenient to denote by $L(\mathbf{E})$, $L^r(\mathbf{E})$, $L^r_s(\mathbf{E})$, and $L^r_a(\mathbf{E})$ the continuous linear maps of $\mathbf{E}$ into $\mathbf{R}$ (resp. the continuous, r-multilinear, symmetric, alternating maps of $\mathbf{E}$ into $\mathbf{R}$). Following classical terminology, it is also convenient to call such maps into $\mathbf{R}$ **forms** (of the corresponding type). If $\mathbf{E}_1, \ldots, \mathbf{E}_r$ and $\mathbf{F}$ are topological vector spaces, then we denote by $L(\mathbf{E}_1, \ldots, \mathbf{E}_r; \mathbf{F})$ the continuous multilinear maps of the product $\mathbf{E}_1 \times \cdots \times \mathbf{E}_r$ into $\mathbf{F}$. We let:

$$\mathrm{End}(\mathbf{E}) = L(\mathbf{E}, \mathbf{E}),$$

$$\mathrm{Laut}(\mathbf{E}) = \text{elements of } \mathrm{End}(\mathbf{E}) \text{ which are invertible in } \mathrm{End}(\mathbf{E}).$$

The most important type of topological vector space for us is the

Banachable space (a TVS which is complete, and whose topology can be defined by a norm). We should say **Banach** space when we want to put the norm into the structure. There are of course many norms which can be used to make a Banachable space into a Banach space, but in practice, one allows the abuse of language which consists in saying Banach space for Banachable space (unless it is absolutely necessary to keep the distinction). Continuous linear maps of Banach spaces are called **operators**.

For this book, we assume from now on that all our topological vector spaces are Banach spaces. We shall occasionally make some comments to indicate where it might be possible to generalize certain results to more general spaces. We denote our Banach spaces by $\mathbf{E}$, $\mathbf{F}, \ldots$.

The next two propositions give two aspects of what is known as the **closed graph theorem.**

Proposition 2.1. *Every continuous bijective linear map of* $\mathbf{E}$ *onto* $\mathbf{F}$ *is a toplinear isomorphism.*

Proposition 2.2. *If* $\mathbf{E}$ *is a Banach space, and* $\mathbf{F}_1$, $\mathbf{F}_2$ *are two closed subspaces which are complementary (i.e.* $\mathbf{E} = \mathbf{F}_1 + \mathbf{F}_2$ *and* $\mathbf{F}_1 \cap \mathbf{F}_2 = 0$), *then the map of* $\mathbf{F}_1 \times \mathbf{F}_2$ *onto* $\mathbf{E}$ *given by the sum is a toplinear isomorphism.*

We shall frequently encounter a situation as in Proposition 2.2, and if $\mathbf{F}$ is a closed subspace of $\mathbf{E}$ such that there exists a closed complement $\mathbf{F}_1$ such that $\mathbf{E}$ is toplinearly isomorphic to the product of $\mathbf{F}$ and $\mathbf{F}_1$ under the natural mapping, then we shall say that $\mathbf{F}$ **splits** in $\mathbf{E}$.

Next, we state a weak form of the Hahn–Banach theorem.

Proposition 2.3. *Let* $\mathbf{E}$ *be a Banach space and* $x \neq 0$ *an element of* $\mathbf{E}$. *Then there exists a continuous linear map* λ *of* $\mathbf{E}$ *into* $\mathbf{R}$ *such that* $\lambda(x) \neq 0$.

One constructs λ by Zorn's lemma, supposing that λ is defined on some subspace, and having a bounded norm. One then extends λ to the subspace generated by one additional element, without increasing the norm.

In particular, every finite dimensional subspace of $\mathbf{E}$ splits if $\mathbf{E}$ is complete. More trivially, we observe that a finite codimensional closed subspace also splits.

We now come to the problem of putting a topology on $L(\mathbf{E}, \mathbf{F})$. Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces, and let

$$A\colon \mathbf{E} \to \mathbf{F}$$

be a continuous linear map (also called a bounded linear map). We can then define the **norm** of A to be the greatest lower bound of all numbers K

such that

$$|Ax| \leqq K|x|$$

for all $x \in \mathbf{E}$. This norm makes $L(\mathbf{E}, \mathbf{F})$ into a Banach space.

In a similar way, we define the topology of $L(\mathbf{E}_1, \ldots, \mathbf{E}_r; \mathbf{F})$, which is a Banach space if we define the norm of a multilinear continuous map

$$A\colon \mathbf{E}_1 \times \cdots \times \mathbf{E}_r \to \mathbf{F}$$

by the greatest lower bound of all numbers K such that

$$|A(x_1, \ldots, x_r)| \leqq K|x_1| \cdots |x_r|.$$

We have:

Proposition 2.4. *If* $\mathbf{E}_1, \ldots, \mathbf{E}_r$, $\mathbf{F}$ *are Banach spaces, then the canonical map*

$$L(\mathbf{E}_1, L(\mathbf{E}_2, \ldots, L(\mathbf{E}_r, \mathbf{F}), \ldots)) \to L^r(\mathbf{E}_1, \ldots, \mathbf{E}_r; \mathbf{F})$$

from the repeated continuous linear maps to the continuous multilinear maps is a toplinear isomorphism, which is norm-preserving, i.e. a Banach-isomorphism.

The preceding propositions could be generalized to a wider class of topological vector spaces. The following one exhibits a property peculiar to Banach spaces.

Proposition 2.5. *Let* $\mathbf{E}$, $\mathbf{F}$ *be two Banach spaces. Then the set of toplinear isomorphisms* $\mathrm{Lis}(\mathbf{E}, \mathbf{F})$ *is open in* $L(\mathbf{E}, \mathbf{F})$.

The proof is in fact quite simple. If $\mathrm{Lis}(\mathbf{E}, \mathbf{F})$ is not empty, one is immediately reduced to proving that $\mathrm{Laut}(\mathbf{E})$ is open in $L(\mathbf{E}, \mathbf{E})$. We then remark that if $u \in L(\mathbf{E}, \mathbf{E})$, and $|u| < 1$, then the series

$$1 + u + u^2 + \cdots$$

converges. Given any toplinear automorphism w of $\mathbf{E}$, we can find an open neighborhood by translating the open unit ball multiplicatively from 1 to w.

Again in Banach spaces, we have:

Proposition 2.6. *If* $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$ *are Banach spaces, then the bilinear maps*

$$L(\mathbf{E}, \mathbf{F}) \times L(\mathbf{F}, \mathbf{G}) \to L(\mathbf{E}, \mathbf{G}),$$

$$L(\mathbf{E}, \mathbf{F}) \times \mathbf{E} \to \mathbf{F},$$

obtained by composition of mappings are continuous, and similarly for multilinear maps.

Remark. The preceding proposition is false for more general spaces than Banach spaces, say Fréchet spaces. In that case, one might hope that the following may be true. Let U be open in a Fréchet space and let

$$f\colon U \to L(\mathbf{E}, \mathbf{F}),$$
$$g\colon U \to L(\mathbf{F}, \mathbf{G}),$$

be continuous. Let γ be the composition of maps. Then $\gamma(f, g)$ is continuous. The same type of question arises later, with differentiable maps instead, and it is of course essential to know the answer to deal with the composition of differentiable maps.

I, §3. DERIVATIVES AND COMPOSITION OF MAPS

A real valued function of a real variable, defined on some neighborhood of 0 is said to be $o(t)$ if

$$\lim_{t \to 0} o(t)/t = 0.$$

Let $\mathbf{E}$, $\mathbf{F}$ be two topological vector spaces, and φ a mapping of a neighborhood of 0 in $\mathbf{E}$ into $\mathbf{F}$. We say that φ is **tangent to** 0 if, given a neighborhood W of 0 in $\mathbf{F}$, there exists a neighborhood V of 0 in $\mathbf{E}$ such that

$$\varphi(tV) \subset o(t)W$$

for some function $o(t)$. If both $\mathbf{E}$, $\mathbf{F}$ are normed, then this amounts to the usual condition

$$|\varphi(x)| \leqq |x|\psi(x)$$

with $\lim \psi(x) = 0$ as $|x| \to 0$.

Let $\mathbf{E}$, $\mathbf{F}$ be two topological vector spaces and U open in $\mathbf{E}$. Let $f\colon U \to \mathbf{F}$ be a continuous map. We shall say that f is **differentiable** at a point $x_0 \in U$ if there exists a continuous linear map λ of $\mathbf{E}$ into $\mathbf{F}$ such that, if we let

$$f(x_0 + y) = f(x_0) + \lambda y + \varphi(y)$$

for small y, then φ is tangent to 0. It then follows trivially that λ is uniquely determined, and we say that it is the **derivative** of f at x_0. We denote the derivative by $Df(x_0)$ or $f'(x_0)$. It is an element of $L(\mathbf{E}, \mathbf{F})$. If f is differentiable at every point of U, then f' is a map

$$f'\colon U \to L(\mathbf{E}, \mathbf{F}).$$

It is easy to verify the chain rule.

Proposition 3.1. *If $f\colon U \to V$ is differentiable at x_0, if $g\colon V \to W$ is differentiable at $f(x_0)$, then $g \circ f$ is differentiable at x_0, and*

$$(g \circ f)'(x_0) = g'(f(x_0)) \circ f'(x_0).$$

Proof. We leave it as a simple (and classical) exercise.

The rest of this section is devoted to the statements of the differential calculus. All topological vector spaces are assumed to be Banach spaces (i.e. Banachable). Then $L(\mathbf{E}, \mathbf{F})$ is also a Banach space, if $\mathbf{E}$ and $\mathbf{F}$ are Banach spaces.

Let U be open in $\mathbf{E}$ and let $f\colon U \to \mathbf{F}$ be differentiable at each point of U. If f' is continuous, then we say that f is **of class** C^1. We define maps of class C^p ($p \geqq 1$) inductively. The p-th derivative $D^p f$ is defined as $D(D^{p-1}f)$ and is itself a map of U into

$$L(\mathbf{E}, L(\mathbf{E}, \ldots, L(\mathbf{E}, \mathbf{F}) \cdots))$$

which can be identified with $L^p(\mathbf{E}, \mathbf{F})$ by Proposition 2.4. A map f is said to be **of class** C^p if its kth derivative $D^k f$ exists for $1 \leqq k \leqq p$, and is continuous.

Remark. *Let f be of class C^p, on an open set U containing the origin. Suppose that f is locally homogeneous of degree p near 0, that is*

$$f(tx) = t^p f(x)$$

for all t and x sufficiently small. Then for all sufficiently small x we have

$$f(x) = \frac{1}{p!} D^p f(0) x^{(p)},$$

where $x^{(p)} = (x, x, \ldots, x)$, p times.

This is easily seen by differentiating p times the two expressions for $f(tx)$, and then setting $t = 0$. The differentiation is a trivial application of the chain rule.

Proposition 3.2. *Let U, V be open in Banach spaces. If $f\colon U \to V$ and $g\colon V \to \mathbf{F}$ are of class C^p, then so is $g \circ f$.*

From Proposition 3.2, we can view open subsets of Banach spaces as the objects of a category, whose morphisms are the continuous maps of class C^p. These will be called **C^p-morphisms**. We say that f is of class C^∞ if it is of class C^p for all integers $p \geqq 1$. From now on, p is an integer $\geqq 0$ or ∞ (C^0 maps being the continuous maps). In practice, we omit the prefix C^p if the p remains fixed. Thus by **morphism**, throughout the rest of this book, we mean C^p-morphism with $p \leqq \infty$. We shall use the word morphism also for C^p-morphisms of manifolds (to be defined in the next chapter), *but morphisms in any other category will always be prefixed so as to indicate the category to which they belong* (for instance bundle morphism, continuous linear morphism, etc.).

Proposition 3.3. *Let U be open in the Banach space $\mathbf{E}$, and let $f\colon U \to \mathbf{F}$ be a C^p-morphism. Then $D^p f$ (viewed as an element of $L^p(\mathbf{E}, \mathbf{F})$) is symmetric.*

Proposition 3.4. *Let U be open in $\mathbf{E}$, and let $f_i\colon U \to \mathbf{F}_i$ $(i = 1, \dots, n)$ be continuous maps into spaces $\mathbf{F}_i$. Let $f = (f_1, \dots, f_n)$ be the map of U into the product of the $\mathbf{F}_i$. Then f is of class C^p if and only if each f_i is of class C^p, and in that case*

$$D^p f = (D^p f_1, \dots, D^p f_n).$$

Let U, V be open in spaces $\mathbf{E}_1$, $\mathbf{E}_2$ and let

$$f\colon U \times V \to \mathbf{F}$$

be a continuous map into a Banach space. We can introduce the notion of partial derivative in the usual manner. If (x, y) is in $U \times V$ and we keep y fixed, then as a function of the first variable, we have the derivative as defined previously. This derivative will be denoted by $D_1 f(x, y)$. Thus

$$D_1 f\colon U \times V \to L(\mathbf{E}_1, \mathbf{F})$$

is a map of $U \times V$ into $L(\mathbf{E}_1, \mathbf{F})$. We call it the **partial derivative** with respect to the first variable. Similarly, we have $D_2 f$, and we could take n factors instead of 2. The total derivative and the partials are then related as follows.

Proposition 3.5. *Let $U_1, \dots, U_n$ be open in the spaces $\mathbf{E}_1, \dots, \mathbf{E}_n$ and let $f\colon U_1 \times \cdots \times U_n \to \mathbf{F}$ be a continuous map. Then f is of class C^p if and only if each partial derivative $D_i f\colon U_1 \times \cdots U_n \to L(\mathbf{E}_i, \mathbf{F})$ exists and is of class C^{p-1}. If that is the case, then for $x = (x_1, \dots, x_n)$ and*

$$v = (v_1, \dots, v_n) \in \mathbf{E}_1 \times \cdots \times \mathbf{E}_n,$$

we have

$$Df(x) \cdot (v_1, \dots, v_n) = \sum D_i f(x) \cdot v_i.$$

The next four propositions are concerned with continuous linear and multilinear maps.

Proposition 3.6. *Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces and $f\colon \mathbf{E} \to \mathbf{F}$ a continuous linear map. Then for each $x \in \mathbf{E}$ we have*

$$f'(x) = f.$$

Proposition 3.7. *Let $\mathbf{E}$, $\mathbf{F}$, $\mathbf{G}$ be Banach spaces, and U open in $\mathbf{E}$. Let $f\colon U \to \mathbf{F}$ be of class C^p and $g\colon \mathbf{F} \to \mathbf{G}$ continuous and linear. Then $g \circ f$ is of class C^p and*

$$D^p(g \circ f) = g \circ D^p f.$$

Proposition 3.8. *If $\mathbf{E}_1, \dots, \mathbf{E}_r$ and $\mathbf{F}$ are Banach spaces and*

$$f\colon \mathbf{E}_1 \times \cdots \times \mathbf{E}_r \to \mathbf{F}$$

a continuous multilinear map, then f is of class C^∞, and its $(r + 1)$-st derivative is 0. If $r = 2$, then Df is computed according to the usual rule for derivative of a product (first times the derivative of the second plus derivative of the first times the second).

Proposition 3.9. *Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces which are toplinearly isomorphic. If $u\colon \mathbf{E} \to \mathbf{F}$ is a toplinear isomorphism, we denote its inverse by u^{-1}. Then the map*

$$u \mapsto u^{-1}$$

from $\mathrm{Lis}(\mathbf{E}, \mathbf{F})$ to $\mathrm{Lis}(\mathbf{F}, \mathbf{E})$ is a C^∞-isomorphism. Its derivative at a point u_0 is the linear map of $L(\mathbf{E}, \mathbf{F})$ into $L(\mathbf{F}, \mathbf{E})$ given by the formula

$$v \mapsto u_0^{-1} v u_0^{-1}.$$

Finally, we come to some statements which are of use in the theory of vector bundles.

Proposition 3.10. *Let U be open in the Banach space $\mathbf{E}$ and let $\mathbf{F}$, $\mathbf{G}$ be Banach spaces.*

(i) *If $f\colon U \to L(\mathbf{E}, \mathbf{F})$ is a C^p-morphism, then the map of $U \times \mathbf{E}$ into $\mathbf{F}$ given by*

$$(x, v) \mapsto f(x)v$$

is a morphism.

(ii) *If $f\colon U \to L(\mathbf{E}, \mathbf{F})$ and $g\colon U \to L(\mathbf{F}, \mathbf{G})$ are morphisms, then so is $\gamma(f, g)$ (γ being the composition).*

(iii) *If* $f: U \to \mathbf{R}$ *and* $g: U \to L(\mathbf{E}, \mathbf{F})$ *are morphisms, so is* fg (*the value of* fg *at* x *is* $f(x)g(x)$, *ordinary multiplication by scalars*).
(iv) *If* $f, g: U \to L(\mathbf{E}, \mathbf{F})$ *are morphisms, so is* $f + g$.

This proposition concludes our summary of results assumed without proof.

I, §4. INTEGRATION AND TAYLOR'S FORMULA

Let $\mathbf{E}$ be a Banach space. Let I denote a real, closed interval, say $a \leqq t \leqq b$. A **step mapping**

$$f: I \to \mathbf{E}$$

is a mapping such that there exists a finite number of disjoint sub-intervals $I_1, \ldots, I_n$ covering I such that on each interval I_j, the mapping has constant value, say v_j. We do not require the intervals I_j to be closed. They may be open, closed, or half-closed.

Given a sequence of mappings f_n from I into $\mathbf{E}$, we say that it converges uniformly if, given a neighborhood W of 0 into $\mathbf{E}$, there exists an integer n_0 such that, for all $n, m > n_0$ and all $t \in I$, the difference $f_n(t) - f_m(t)$ lies in W. The sequence f_n then converges to a mapping f of I into $\mathbf{E}$.

A **ruled** mapping is a uniform limit of step mappings. We leave to the reader the proof that every continuous mapping is ruled.

If f is a step mapping as above, we define its integral

$$\int_a^b f = \int_a^b f(t)\,dt = \sum \mu(I_j) v_j,$$

where $\mu(I_j)$ is the length of the interval I_j (its measure in the standard Lebesgue measure). This integral is independent of the choice of intervals I_j on which f is constant.

If f is ruled and $f = \lim f_n$ (lim being the uniform limit), then the sequence

$$\int_a^b f_n$$

converges in $\mathbf{E}$ to an element of $\mathbf{E}$ independent of the particular sequence f_n used to approach f uniformly. We denote this limit by

$$\int_a^b f = \int_a^b f(t)\,dt$$

and call it the **integral** of f. The integral is linear in f, and satisfies the

usual rules concerning changes of intervals. (If $b < a$ then we define $\int_a^b$ to be minus the integral from b to a.)

As an immediate consequence of the definition, we get:

Proposition 4.1. *Let* $\lambda: \mathbf{E} \to \mathbf{R}$ *be a continuous linear map and let* $f: I \to \mathbf{E}$ *be ruled. Then* $\lambda f = \lambda \circ f$ *is ruled, and*

$$\lambda \int_a^b f(t)\,dt = \int_a^b \lambda f(t)\,dt.$$

Proof. If f_n is a sequence of step functions converging uniformly to f, then λf_n is ruled and converges uniformly to λf. Our formula follows at once.

Taylor's Formula. *Let* $\mathbf{E}, \mathbf{F}$ *be Banach spaces. Let* U *be open in* $\mathbf{E}$. *Let* x, y *be two points of* U *such that the segment* $x + ty$ *lies in* U *for* $0 \leqq t \leqq 1$. *Let*

$$f: U \to \mathbf{F}$$

be a C^p*-morphism, and denote by* $y^{(p)}$ *the "vector"* $(y, \ldots, y)$ p *times. Then the function* $D^p f(x + ty) \cdot y^{(p)}$ *is continuous in* t, *and we have*

$$f(x+y) = f(x) + \frac{Df(x)y}{1!} + \cdots + \frac{D^{p-1}f(x)y^{(p-1)}}{(p-1)!} + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} D^p f(x+ty) y^{(p)}\,dt.$$

Proof. By the Hahn–Banach theorem, it suffices to show that both sides give the same thing when we apply a functional λ (continuous linear map into $\mathbf{R}$). This follows at once from Proposition 3.7 and 4.1, together with the known result when $\mathbf{F} = \mathbf{R}$. In this case, the proof proceeds by induction on p, and integration by parts, starting from

$$f(x+y) - f(x) = \int_0^1 Df(x+ty)y\,dt.$$

The next two corollaries are known as the **mean value theorem**.

Corollary 4.2. *Let* $\mathbf{E}, \mathbf{F}$ *be two Banach spaces,* U *open in* $\mathbf{E}$, *and* x, z *two distinct points of* U *such that the segment* $x + t(z - x)$ $(0 \leqq t \leqq 1)$ *lies in* U. *Let* $f: U \to \mathbf{F}$ *be continuous and of class* C^1. *Then*

$$|f(z) - f(x)| \leqq |z - x| \sup |f'(\xi)|,$$

the sup being taken over ξ *in the segment.*

Proof. This comes from the usual estimations of the integral. Indeed, for any continuous map $g\colon I \to \mathbf{F}$ we have the estimate

$$\left|\int_a^b g(t)\,dt\right| \leqq K(b-a)$$

if K is a bound for g on I, and $a \leqq b$. This estimate is obvious for step functions, and therefore follows at once for continuous functions.

Another version of the mean value theorem is frequently used.

Corollary 4.3. *Let the hypotheses be as in Corollary 4.2. Let x_0 be a point on the segment between x and z. Then*

$$|f(z) - f(x) - f'(x_0)(z-x)| \leqq |z-x| \sup |f'(\xi) - f'(x_0)|,$$

the sup taken over all ξ on the segment.

Proof. We apply Corollary 4.2 to the map

$$g(x) = f(x) - f'(x_0)x.$$

Finally, let us make some comments on the estimate of the remainder term in Taylor's formula. We have assumed that $D^p f$ is continuous. Therefore, $D^p f(x+ty)$ can be written

$$D^p f(x+ty) = D^p f(x) + \psi(y, t),$$

where ψ depends on y, t (and x of course), and for fixed x, we have

$$\lim |\psi(y, t)| = 0$$

as $|y| \to 0$. Thus we obtain:

Corollary 4.4. *Let $\mathbf{E}$, $\mathbf{F}$ be two Banach spaces, U open in $\mathbf{E}$, and x a point of U. Let $f\colon U \to \mathbf{F}$ be of class C^p, $p \geqq 1$. Then for all y such that the segment $x+ty$ lies in U $(0 \leqq t \leqq 1)$, we have*

$$f(x+y) = f(x) + \frac{Df(x)y}{1!} + \cdots + \frac{D^p f(x)y^{(p)}}{p!} + \theta(y)$$

with an error term $\theta(y)$ satisfying

$$\lim_{y \to 0} \theta(y)/|y|^p = 0.$$

I, §5. THE INVERSE MAPPING THEOREM

The inverse function theorem and the existence theorem for differential equations (of Chapter IV) are based on the next result.

Lemma 5.1 (Contraction Lemma or Shrinking Lemma). *Let M be a complete metric space, with distance function d, and let $f\colon M \to M$ be a mapping of M into itself. Assume that there is a constant K, $0 < K < 1$, such that, for any two points x, y in M, we have*

$$d\big(f(x), f(y)\big) \leqq K\, d(x, y).$$

Then f has a unique fixed point (a point such that $f(x) = x$). Given any point x_0 in M, the fixed point is equal to the limit of $f^n(x_0)$ (iteration of f repeated n times) as n tends to infinity.

Proof. This is a trivial exercise in the convergence of the geometric series, which we leave to the reader.

Theorem 5.2. *Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces, U an open subset of $\mathbf{E}$, and let $f\colon U \to \mathbf{F}$ a C^p-morphism with $p \geqq 1$. Assume that for some point $x_0 \in U$, the derivative $f'(x_0)\colon \mathbf{E} \to \mathbf{F}$ is a toplinear isomorphism. Then f is a local C^p-isomorphism at x_0.*

(By a **local C^p-isomorphism** at x_0, we mean that there exists an open neighborhood V of x_0 such that the restriction of f to V establishes a C^p-isomorphism between V and an open subset of $\mathbf{E}$.)

Proof. Since a toplinear isomorphism is a C^∞-isomorphism, we may assume without loss of generality that $\mathbf{E} = \mathbf{F}$ and $f'(x_0)$ is the identity (simply by considering $f'(x_0)^{-1} \circ f$ instead of f). After translations, we may also assume that $x_0 = 0$ and $f(x_0) = 0$.

We let $g(x) = x - f(x)$. Then $g'(x_0) = 0$ and by continuity there exists $r > 0$ such that, if $|x| < 2r$, we have

$$|g'(x)| < \tfrac{1}{2}.$$

From the mean value theorem, we see that $|g(x)| \leqq \frac{1}{2}|x|$ and hence g maps the closed ball of radius r, $\bar{B}_r(0)$ into $\bar{B}_{r/2}(0)$.

We contend: Given $y \in \bar{B}_{r/2}(0)$, there exists a unique element $x \in \bar{B}_r(0)$ such that $f(x) = y$. We prove this by considering the map

$$g_y(x) = y + x - f(x).$$

If $|y| \leqq r/2$ and $|x| \leqq r$, then $|g_y(x)| \leqq r$ and hence g_y may be viewed as a mapping of the complete metric space $\bar{B}_r(0)$ into itself. The bound of $\frac{1}{2}$ on the derivative together with the mean value theorem shows that g_y is a contracting map, i.e. that

$$|g_y(x_1) - g_y(x_2)| = |g(x_1) - g(x_2)| \leqq \tfrac{1}{2}|x_1 - x_2|$$

for x_1, $x_2 \in \bar{B}_r(0)$. By the contraction lemma, it follows that g_y has a unique fixed point. But the fixed point of g_y is precisely the solution of the equation $f(x) = y$. This proves our contention.

We obtain a local inverse $\varphi = f^{-1}$. This inverse is continuous, because

$$|x_1 - x_2| \leqq |f(x_1) - f(x_2)| + |g(x_1) - g(x_2)|$$

and hence

$$|x_1 - x_2| \leqq 2|f(x_1) - f(x_2)|.$$

Furthermore φ is differentiable in $B_{r/2}(0)$. Indeed, let $y_1 = f(x_1)$ and $y_2 = f(x_2)$ with y_1, $y_2 \in B_{r/2}(0)$ and x_1, $x_2 \in \bar{B}_r(0)$. Then

$$|\varphi(y_1) - \varphi(y_2) - f'(x_2)^{-1}(y_1 - y_2)| = \left|x_1 - x_2 - f'(x_2)^{-1}\big(f(x_1) - f(x_2)\big)\right|.$$

We operate on the expression inside the norm sign with the identity

$$\mathrm{id} = f'(x_2)^{-1}f'(x_2).$$

Estimating and using the continuity of f', we see that for some constant A, the preceding expression is bounded by

$$A|f'(x_2)(x_1 - x_2) - f(x_1) + f(x_2)|.$$

From the differentiability of f, we conclude that this expression is $o(x_1 - x_2)$ which is also $o(y_1 - y_2)$ in view of the continuity of φ proved above. This proves that φ is differentiable and also that its derivative is what it should be, namely

$$\varphi'(y) = f'(\varphi(y))^{-1},$$

for $y \in B_{r/2}(0)$. Since the mappings φ, f', "inverse" are continuous, it follows that φ' is continuous and thus that φ is of class C^1. Since taking inverses is C^∞ and f' is C^{p-1}, it follows inductively that φ is C^p, as was to be shown.

Note that this last argument also proves:

Proposition 5.3. *If f: $U \to V$ is a homeomorphism and is of class C^p with $p \geqq 1$, and if f is a C^1-isomorphism, then f is a C^p-isomorphism.*

In some applications it is necessary to know that if the derivative of a map is close to the identity, then the image of a ball contains a ball of only slightly smaller radius. The precise statement follows. In this book, it will be used only in the proof of the change of variables formula, and therefore may be omitted until the reader needs it.

Lemma 5.4. *Let U be open in* **E**, *and let f: $U \to \mathbf{E}$ be of class C^1. Assume that $f(0) = 0$, $f'(0) = I$. Let $r > 0$ and assume that $\bar{B}_r(0) \subset U$. Let $0 < s < 1$, and assume that*

$$|f'(z) - f'(x)| \leqq s$$

for all x, $z \in \bar{B}_r(0)$. If $y \in \mathbf{E}$ and $|y| \leqq (1 - s)r$, then there exists a unique $x \in \bar{B}_r(0)$ such that $f(x) = y$.

Proof. The map g_y given by $g_y(x) = x - f(x) + y$ is defined for $|x| \leqq r$ and $|y| \leqq (1 - s)r$, and maps $\bar{B}_r(0)$ into itself because, from the estimate

$$|f(x) - x| = |f(x) - f(0) - f'(0)x| \leqq |x| \sup|f'(z) - f'(0)| \leqq sr,$$

we obtain

$$|g_y(x)| \leqq sr + (1 - s)r = r.$$

Furthermore, g_y is a shrinking map because, from the mean value theorem, we get

$$\begin{aligned} |g_y(x_1) - g_y(x_2)| &= \left|x_1 - x_2 - \big(f(x_1) - f(x_2)\big)\right| \\ &= |x_1 - x_2 - f'(0)(x_1 - x_2) + \delta(x_1, x_2)| \\ &= |\delta(x_1, x_2)|, \end{aligned}$$

where

$$|\delta(x_1, x_2)| \leqq |x_1 - x_2| \sup|f'(z) - f'(0)| \leqq s|x_1 - x_2|.$$

Hence g_y has a unique fixed point $x \in \bar{B}_r(0)$ which is such that $f(x) = y$. This proves the lemma.

We shall now prove some useful corollaries, which will be used in dealing with immersions and submersions later. *We assume that morphism means C^p-morphism with $p \geqq 1$.*

Corollary 5.5. *Let U be an open subset of* **E**, *and f: $U \to \mathbf{F}_1 \times \mathbf{F}_2$ a morphism of U into a product of Banach spaces. Let $x_0 \in U$, suppose that $f(x_0) = (0, 0)$ and that $f'(x_0)$ induces a toplinear isomorphism of* **E** *and $\mathbf{F}_1 = \mathbf{F}_1 \times 0$. Then there exists a local isomorphism g of $\mathbf{F}_1 \times \mathbf{F}_2$ at $(0, 0)$ such that*

$$g \circ f\colon\ U \to \mathbf{F}_1 \times \mathbf{F}_2$$

maps an open subset U_1 of U into $\mathbf{F}_1 \times 0$ and induces a local isomorphism of U_1 at x_0 on an open neighborhood of 0 in $\mathbf{F}_1$.

Proof. We may assume without loss of generality that $\mathbf{F}_1 = \mathbf{E}$ (identify by means of $f'(x_0)$) and $x_0 = 0$. We define

$$\varphi\colon U \times \mathbf{F}_2 \to \mathbf{F}_1 \times \mathbf{F}_2$$

by the formula

$$\varphi(x, y_2) = f(x) + (0, y_2)$$

for $x \in U$ and $y_2 \in \mathbf{F}_2$. Then $\varphi(x, 0) = f(x)$, and

$$\varphi'(0, 0) = f'(0) + (0, \mathrm{id}_2).$$

Since $f'(0)$ is assumed to be a toplinear isomorphism onto $\mathbf{F}_1 \times 0$, it follows that $\varphi'(0, 0)$ is also a toplinear isomorphism. Hence by the theorem, it has a local inverse, say g, which obviously satisfies our requirements.

Corollary 5.6. *Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces, U open in $\mathbf{E}$, and $f\colon U \to \mathbf{F}$ a C^p-morphism with $p \geqq 1$. Let $x_0 \in U$. Suppose that $f(x_0) = 0$ and $f'(x_0)$ gives a toplinear isomorphism of $\mathbf{E}$ on a closed subspace of $\mathbf{F}$ which splits. Then there exists a local isomorphism $g\colon \mathbf{F} \to \mathbf{F}_1 \times \mathbf{F}_2$ at 0 and an open subset U_1 of U containing x_0 such that the composite map $g \circ f$ induces an isomorphism of U_1 onto an open subset of $\mathbf{F}_1$.*

Considering the splitting assumption, this is a reformulation of Corollary 5.5.

It is convenient to define the notion of splitting for injections. If $\mathbf{E}$, $\mathbf{F}$ are topological vector spaces, and $\lambda\colon \mathbf{E} \to \mathbf{F}$ is a continuous linear map, which is injective, then we shall say that λ **splits** if there exists a toplinear isomorphism $\alpha\colon \mathbf{F} \to \mathbf{F}_1 \times \mathbf{F}_2$ such that $\alpha \circ \lambda$ induces a toplinear isomorphism of $\mathbf{E}$ onto $\mathbf{F}_1 = \mathbf{F}_1 \times 0$. In our corollary, we could have rephrased our assumption by saying that $f'(x_0)$ is a splitting injection.

For the next corollary, dual to the preceding one, we introduce the notion of a **local projection**. Given a product of two open sets of Banach spaces $V_1 \times V_2$ and a morphism $f\colon V_1 \times V_2 \to \mathbf{F}$, we say that f is a **projection** (on the first factor) if f can be factored

$$V_1 \times V_2 \to V_1 \to \mathbf{F}$$

into an ordinary projection and an isomorphism of V_1 onto an open subset of $\mathbf{F}$. We say that f is a local projection at (a_1, a_2) if there exists an open neighborhood $U_1 \times U_2$ of (a_1, a_2) such that the restriction of f to this neighborhood is a projection.

Corollary 5.7. *Let U be an open subset of a product of Banach spaces $\mathbf{E}_1 \times \mathbf{E}_2$ and (a_1, a_2) a point of U. Let $f\colon U \to \mathbf{F}$ be a morphism into a Banach space, say $f(a_1, a_2) = 0$, and assume that the partial derivative*

$$D_2 f(a_1, a_2)\colon \mathbf{E}_2 \to \mathbf{F}$$

is a toplinear isomorphism. Then there exists a local isomorphism h of a product $V_1 \times V_2$ onto an open neighborhood of (a_1, a_2) contained in U such that the composite map

$$V_1 \times V_2 \xrightarrow{h} U \xrightarrow{f} \mathbf{F}$$

is a projection (on the second factor).

Proof. We may assume $(a_1, a_2) = (0, 0)$ and $\mathbf{E}_2 = \mathbf{F}$. We define

$$\varphi\colon \mathbf{E}_1 \times \mathbf{E}_2 \to \mathbf{E}_1 \times \mathbf{E}_2$$

by

$$\varphi(x_1, x_2) = (x_1, f(x_1, x_2))$$

locally at (a_1, a_2). Then φ' is represented by the matrix

$$\begin{pmatrix} \mathrm{id}_1 & O \\ D_1 f & D_2 f \end{pmatrix}$$

and is therefore a toplinear isomorphism at (a_1, a_2). By the theorem, it has a local inverse h which clearly satisfies our requirements.

Corollary 5.8. *Let U be an open subset of a Banach space $\mathbf{E}$ and $f\colon U \to \mathbf{F}$ a morphism into a Banach space $\mathbf{F}$. Let $x_0 \in U$ and assume that $f'(x_0)$ is surjective, and that its kernel splits. Then there exists an open subset U' of U containing x_0 and an isomorphism*

$$h\colon V_1 \times V_2 \to U'$$

such that the composite map $f \circ h$ is a projection

$$V_1 \times V_2 \to V_1 \to \mathbf{F}.$$

Proof. Again this is essentially a reformulation of the corollary, taking into account the splitting assumption.

Theorem 5.9 (The Implicit Mapping Theorem). *Let U, V be open sets in Banach spaces $\mathbf{E}$, $\mathbf{F}$ respectively, and let*

$$f\colon U \times V \to G$$

be a C^p mapping. Let $(a, b) \in U \times V$, and assume that

$$D_2 f(a, b)\colon \mathbf{F} \to G$$

is a toplinear isomorphism. Let $f(a, b) = 0$. Then there exists a continuous map $g\colon U_0 \to V$ defined on an open neighborhood U_0 of a such that $g(a) = b$ and such that

$$f(x, g(x)) = 0$$

for all $x \in U_0$. If U_0 is taken to be a sufficiently small ball, then g is uniquely determined, and is also of class C^p.

Proof. Let $\lambda = D_2 f(a, b)$. Replacing f by $\lambda^{-1} \circ f$ we may assume without loss of generality that $D_2 f(a, b)$ is the identity. Consider the map

$$\varphi\colon U \times V \to \mathbf{E} \times \mathbf{F}$$

given by

$$\varphi(x, y) = (x, f(x, y)).$$

Then the derivative of φ at (a, b) is immediately computed to be represented by the matrix

$$D\varphi(a, b) = \begin{pmatrix} \mathrm{id}_{\mathbf{E}} & O \\ D_1 f(a, b) & D_2 f(a, b) \end{pmatrix} = \begin{pmatrix} \mathrm{id}_{\mathbf{E}} & O \\ D_1 f(a, b) & \mathrm{id}_{\mathbf{F}} \end{pmatrix}$$

whence φ is locally invertible at (a, b) since the inverse of $D\varphi(a, b)$ exists and is the matrix

$$\begin{pmatrix} \mathrm{id}_{\mathbf{E}} & O \\ -D_1 f(a, b) & \mathrm{id}_{\mathbf{F}} \end{pmatrix}.$$

We denote the local inverse of φ by ψ. We can write

$$\psi(x, z) = (x, h(x, z))$$

where h is some mapping of class C^p. We define

$$g(x) = h(x, 0).$$

Then certainly g is of class C^p and

$$(x, f(x, g(x))) = \varphi(x, g(x)) = \varphi(x, h(x, 0)) = \varphi(\psi(x, 0)) = (x, 0).$$

This proves the existence of a C^p map g satisfying our requirements.

Now for the uniqueness, suppose that g_0 is a continuous map defined near a such that $g_0(a) = b$ and $f(x, g_0(x)) = c$ for all x near a. Then $g_0(x)$ is near b for such x, and hence

$$\varphi(x, g_0(x)) = (x, 0).$$

Since φ is invertible near (a, b) it follows that there is a unique point (x, y) near (a, b) such that $\varphi(x, y) = (x, 0)$. Let U_0 be a small ball on which g is defined. If g_0 is also defined on U_0, then the above argument shows that g and g_0 coincide on some smaller neighborhood of a. Let $x \in U_0$ and let $v = x - a$. Consider the set of those numbers t with $0 \leqq t \leqq 1$ such that $g(a + tv) = g_0(a + tv)$. This set is not empty. Let s be its least upper bound. By continuity, we have $g(a + sv) = g_0(a + sv)$. If $s < 1$, we can apply the existence and that part of the uniqueness just proved to show that g and g_0 are in fact equal in a neighborhood of $a + sv$. Hence $s = 1$, and our uniqueness statement is proved, as well as the theorem.

Note. The particular value $f(a, b) = 0$ in the preceding theorem is irrelevant. If $f(a, b) = c$ for some $c \neq 0$, then the above proof goes through replacing 0 by c everywhere.

CHAPTER II

Manifolds

Starting with open subsets of Banach spaces, one can glue them together with C^p-isomorphisms. The result is called a manifold. We begin by giving the formal definition. We then make manifolds into a category, and discuss special types of morphisms. We define the tangent space at each point, and apply the criteria following the inverse function theorem to get a local splitting of a manifold when the tangent space splits at a point.

We shall wait until the next chapter to give a manifold structure to the union of all the tangent spaces.

II, §1. ATLASES, CHARTS, MORPHISMS

Let X be a set. An **atlas of class** C^p ($p \geqq 0$) on X is a collection of pairs (U_i, φ_i) (i ranging in some indexing set), satisfying the following conditions:

AT 1. *Each U_i is a subset of X and the U_i cover X.*

AT 2. *Each φ_i is a bijection of U_i onto an open subset $\varphi_i U_i$ of some Banach space $\mathbf{E}_i$ and for any i, j, $\varphi_i(U_i \cap U_j)$ is open in $\mathbf{E}_i$.*

AT 3. *The map*

$$\varphi_j \varphi_i^{-1}\colon\ \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$$

is a C^p-isomorphism for each pair of indices i, j.

It is a trivial exercise in point set topology to prove that one can give X a topology in a unique way such that each U_i is open, and the φ_i are

topological isomorphisms. We see no reason to assume that X is Hausdorff. If we wanted X to be Hausdorff, we would have to place a separation condition on the covering. This plays no role in the formal development in Chapters II and III. It is to be understood, however, that any construction which we perform (like products, tangent bundles, etc.) would yield Hausdorff spaces if we start with Hausdorff spaces.

Each pair (U_i, φ_i) will be called a **chart** of the atlas. If a point x of X lies in U_i, then we say that (U_i, φ_i) is a **chart at** x.

In condition **AT 2**, we did not require that the vector spaces be the same for all indices i, or even that they be toplinearly isomorphic. If they are all equal to the same space $\mathbf{E}$, then we say that the atlas is an $\mathbf{E}$-atlas. If two charts (U_i, φ_i) and (U_j, φ_j) are such that U_i and U_j have a non-empty intersection, and if $p \geqq 1$, then taking the derivative of $\varphi_j\varphi_i^{-1}$ we see that $\mathbf{E}_i$ and $\mathbf{E}_j$ are toplinearly isomorphic. Furthermore, the set of points $x \in X$ for which there exists a chart (U_i, φ_i) at x such that $\mathbf{E}_i$ is toplinearly isomorphic to a given space $\mathbf{E}$ is both open and closed. Consequently, on each connected component of X, we could assume that we have an $\mathbf{E}$-atlas for some fixed $\mathbf{E}$.

Suppose that we are given an open subset U of X and a topological isomorphism $\varphi\colon U \to U'$ onto an open subset of some Banach space $\mathbf{E}$. We shall say that (U, φ) is **compatible** with the atlas $\{(U_i, \varphi_i)\}$ if each map $\varphi_i\varphi^{-1}$ (defined on a suitable intersection as in **AT 3**) is a C^p-isomorphism. Two atlases are said to be **compatible** if each chart of one is compatible with the other atlas. One verifies immediately that the relation of compatibility between atlases is an equivalence relation. An equivalence class of atlases of class C^p on X is said to define a structure of **C^p-manifold** on X. If all the vector spaces $\mathbf{E}_i$ in some atlas are toplinearly isomorphic, then we can always find an equivalent atlas for which they are all equal, say to the vector space $\mathbf{E}$. We then say that X is an **E-manifold** or that X is **modeled** on $\mathbf{E}$.

If $\mathbf{E} = \mathbf{R}^n$ for some fixed n, then we say that the manifold is **n-dimensional**. In this case, a chart

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

is given by n coordinate functions $\varphi_1, \ldots, \varphi_n$. If P denotes a point of U, these functions are often written

$$x_1(P), \ldots, x_n(P),$$

or simply $x_1, \ldots, x_n$. They are called **local coordinates** on the manifold.

If the integer p (which may also be ∞) is fixed throughout a discussion, we also say that X is a manifold.

The collection of C^p-manifolds will be denoted by Man^p. If we look only at those modeled on spaces in a category $\mathfrak{A}$ then we write $\mathrm{Man}^p(\mathfrak{A})$.

Those modeled on a fixed **E** will be denoted by $\text{Man}^p(\mathbf{E})$. We shall make these into categories by defining morphisms below.

Let X be a manifold, and U an open subset of X. Then it is possible, in the obvious way, to induce a manifold structure on U, by taking as charts the intersections

$$(U_i \cap U, \varphi_i|(U_i \cap U)).$$

If X is a topological space, covered by open subsets V_j, and if we are given on each V_j a manifold structure such that for each pair j, j' the induced structure on $V_j \cap V_{j'}$ coincides, then it is clear that we can give to X a unique manifold structure inducing the given ones on each V_j.

Example. Let X be the real line, and for each open interval U_i, let φ_i be the function $\varphi_i(t) = t^3$. Then the $\varphi_j\varphi_i^{-1}$ are all equal to the identity, and thus we have defined a C^∞-manifold structure on $\mathbf{R}$!

If X, Y are two manifolds, then one can give the product $X \times Y$ a manifold structure in the obvious way. If $\{(U_i, \varphi_i)\}$ and $\{(V_j, \psi_j)\}$ are atlases for X, Y respectively, then

$$\{(U_i \times V_j, \varphi_i \times \psi_j)\}$$

is an atlas for the product, and the product of compatible atlases gives rise to compatible atlases, so that we do get a well-defined product structure.

Let X, Y be two manifolds. Let $f\colon X \to Y$ be a map. We shall say that f is a **C^p-morphism** if, given $x \in X$, there exists a chart (U, φ) at x and a chart (V, ψ) at $f(x)$ such that $f(U) \subset V$, and the map

$$\psi \circ f \circ \varphi^{-1}\colon \varphi U \to \psi V$$

is a C^p-morphism in the sense of Chapter I, §3. One sees then immediately that this same condition holds for any choice of charts (U, φ) at x and (V, ψ) at $f(x)$ such that $f(U) \subset V$.

It is clear that the composite of two C^p-morphisms is itself a C^p-morphism (because it is true for open subsets of vector spaces). The C^p-manifolds and C^p-morphisms form a category. The notion of isomorphism is therefore defined, and we observe that in our example of the real line, the map $t \mapsto t^3$ gives an isomorphism between the funny differentiable structure and the usual one.

If $f\colon X \to Y$ is a morphism, and (U, φ) is a chart at a point $x \in X$, while (V, ψ) is a chart at $f(x)$, then we shall also denote by

$$f_{V,U}\colon \varphi U \to \psi V$$

the map $\psi f \varphi^{-1}$.

It is also convenient to have a local terminology. Let U be an open set (of a manifold or a Banach space) containing a point x_0. By a **local isomorphism** at x_0 we mean an isomorphism

$$f\colon U_1 \to V$$

from some open set U_1 containing x_0 (and contained in U) to an open set V (in some manifold or some Banach space). Thus a local isomorphism is essentially a change of chart, locally near a given point.

Manifolds of maps. Even starting with a finite dimensional manifold, the set of maps satisfying various smoothness conditions forms an infinite dimensional manifold. This story started with Eells [Ee 58], [Ee 59], [Ee 61]. Palais and Smale used such manifolds of maps in their Morse theory [Pa 63], [Ab 62], [Sm 64]. For a brief discussion of subsequent developments, see [Mar 74], p. 67, referring to [Eb 70], [Ee 66], [El 67], [Kr 72], [Le 67], [Om 70], and [Pa 68]. Two kinds of maps have played a role: the C^p maps of course, with various values of p, but also maps satisfying Sobolev conditions, and usually denoted by H^s. The latter form Hilbert manifolds (definition to be given later).

II, §2. SUBMANIFOLDS, IMMERSIONS, SUBMERSIONS

Let X be a topological space, and Y a subset of X. We say that Y is **locally closed** in X if every point $y \in Y$ has an open neighborhood U in X such that $Y \cap U$ is closed in U. One verifies easily that a locally closed subset is the intersection of an open set and a closed set. For instance, any open subset of X is locally closed, and any open interval is locally closed in the plane.

Let X be a manifold (of class C^p with $p \geqq 0$). Let Y be a subset of X and assume that for each point $y \in Y$ there exists a chart (V, ψ) at y such that ψ gives an isomorphism of V with a product $V_1 \times V_2$ where V_1 is open in some space $\mathbf{E}_1$ and V_2 is open in some space $\mathbf{E}_2$, and such that

$$\psi(Y \cap V) = V_1 \times a_2$$

for some point $a_2 \in V_2$ (which we could take to be 0). Then it is clear that Y is locally closed in X. Furthermore, the map ψ induces a bijection

$$\psi_1\colon Y \cap V \to V_1.$$

The collection of pairs $(Y \cap V, \psi_1)$ obtained in the above manner constitutes an atlas for Y, of class C^p. The verification of this assertion, whose formal details we leave to the reader, depends on the following obvious fact.

Lemma 2.1. *Let U_1, U_2, V_1, V_2 be open subsets of Banach spaces, and $g\colon U_1 \times U_2 \to V_1 \times V_2$ a C^p-morphism. Let $a_2 \in U_2$ and $b_2 \in V_2$ and assume that g maps $U_1 \times a_2$ into $V_1 \times b_2$. Then the induced map*

$$g_1\colon U_1 \to V_1$$

is also a morphism.

Indeed, it is obtained as a composite map

$$U_1 \to U_1 \times U_2 \to V_1 \times V_2 \to V_1,$$

the first map being an inclusion and the third a projection.

We have therefore defined a C^p-structure on Y which will be called a **submanifold** of X. This structure satisfies a universal mapping property, which characterizes it, namely:

Given any map $f\colon Z \to X$ from a manifold Z into X such that $f(Z)$ is contained in Y. Let $f_Y : Z \to Y$ be the induced map. Then f is a morphism if and only if f_Y is a morphism.

The proof of this assertion depends on Lemma 2.1, and is trivial.

Finally, we note that the inclusion of Y into X is a morphism.

If Y is also a closed subspace of X, then we say that it is a **closed submanifold**.

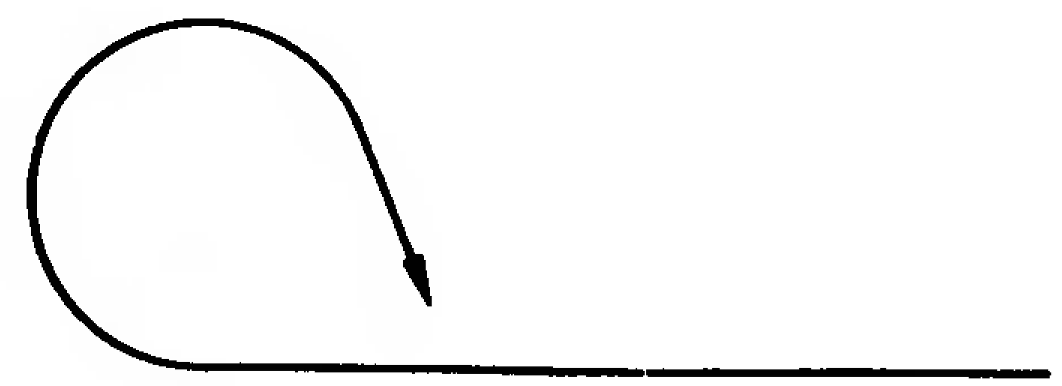
Suppose that X is finite dimensional of dimension n, and that Y is a submanifold of dimension r. Then from the definition we see that the local product structure in a neighborhood of a point of Y can be expressed in terms of local coordinates as follows. Each point P of Y has an open neighborhood U in X with local coordinates $(x_1, \dots, x_n)$ such that the points of Y in U are precisely those whose last $n - r$ coordinates are 0, that is, those points having coordinates of type

$$(x_1, \dots, x_r, 0, \dots, 0).$$

Let $f\colon Z \to X$ be a morphism, and let $z \in Z$. We shall say that f is an **immersion** at z if there exists an open neighborhood Z_1 of z in Z such that the restriction of f to Z_1 induces an isomorphism of Z_1 onto a submanifold of X. We say that f is an **immersion** if it is an immersion at every point.

Note that there exist injective immersions which are not isomorphisms

onto submanifolds, as given by the following example:



(The arrow means that the line approaches itself without touching.) An immersion which does give an isomorphism onto a submanifold is called an **embedding**, and it is called a **closed embedding** if this submanifold is closed.

A morphism $f\colon X \to Y$ will be called a **submersion** at a point $x \in X$ if there exists a chart (U, φ) at x and a chart (V, ψ) at $f(x)$ such that φ gives an isomorphism of U on a products $U_1 \times U_2$ (U_1 and U_2 open in some Banach spaces), and such that the map

$$\psi f \varphi^{-1} = f_{V,U}\colon U_1 \times U_2 \to V$$

is a projection. One sees then that the image of a submersion is an open subset (a submersion is in fact an open mapping). We say that f is a **submersion** if it is a submersion at every point.

For manifolds modelled on Banach spaces, we have the usual criterion for immersions and submersions in terms of the derivative.

Proposition 2.2. *Let X, Y be manifolds of class C^p ($p \geqq 1$) modeled on Banach spaces. Let $f\colon X \to Y$ be a C^p-morphism. Let $x \in X$. Then:*

(i) *f is an immersion at x if and only if there exists a chart (U, φ) at x and (V, ψ) at $f(x)$ such that $f'_{V,U}(\varphi x)$ is injective and splits.*

(ii) *f is a submersion at x if and only if there exists a chart (U, φ) at x and (V, ψ) at $f(x)$ such that $f'_{V,U}(\varphi x)$ is surjective and its kernel splits.*

Proof. This is an immediate consequence of Corollaries 5.4 and 5.6 of the inverse mapping theorem.

The conditions expressed in (i) and (ii) depend only on the derivative, and if they hold for one choice of charts (U, φ) and (V, ψ) respectively, then they hold for every choice of such charts. It is therefore convenient to introduce a terminology in order to deal with such properties.

Let X be a manifold of class C^p ($p \geqq 1$). Let x be a point of X. We consider triples (U, φ, v) where (U, φ) is a chart at x and v is an element of the vector space in which φU lies. We say that two such triples (U, φ, v) and (V, ψ, w) are **equivalent** if the derivative of $\psi\varphi^{-1}$ at φx maps

v on w. The formula reads:

$$(\psi\varphi^{-1})'(\varphi x)v = w$$

(obviously an equivalence relation by the chain rule). An equivalence class of such triples is called a **tangent vector** of X at x. The set of such tangent vectors is called the **tangent space** of X at x and is denoted by $T_x(X)$. Each chart (U, φ) determines a bijection of $T_x(X)$ on a Banach space, namely the equivalence class of (U, φ, v) corresponds to the vector v. By means of such a bijection it is possible to transport to $T_x(X)$ the structure of topological vector space given by the chart, and it is immediate that this structure is independent of the chart selected.

If U, V are open in Banach spaces, then to every morphism of class C^p $(p \geqq 1)$ we can associate its derivative $Df(x)$. If now $f\colon X \to Y$ is a morphism of one manifold into another, and x a point of X, then by means of charts we can interpret the derivative of f on each chart at x as a mapping

$$df(x) = T_xf\colon\ T_x(X) \to T_{f(x)}(Y).$$

Indeed, this map T_xf is the unique linear map having the following property. If (U, φ) is a chart at x and (V, ψ) is a chart at $f(x)$ such that $f(U) \subset V$ and $\bar{v}$ is a tangent vector at x represented by v in the chart (U, φ), then

$$T_xf(\bar{v})$$

is the tangent vector at $f(x)$ represented by $Df_{V,U}(x)v$. The representation of T_xf on the spaces of charts can be given in the form of a diagram

$$\begin{array}{ccc} T_x(X) & \longrightarrow & \mathbf{E} \\ {\scriptstyle T_xf}\downarrow & & \downarrow{\scriptstyle f'_{V,U}(x)} \\ T_{f(x)}(Y) & \longrightarrow & \mathbf{F} \end{array}$$

The map T_xf is obviously continuous and linear for the structure of topological vector space which we have placed on $T_x(X)$ and $T_{f(x)}(Y)$.

As a matter of notation, we shall sometimes write $f_{*,x}$ instead of T_xf.

The operation T satisfies an obvious functorial property, namely, if $f\colon X \to Y$ and $g\colon Y \to Z$ are morphisms, then

$$T_x(g \circ f) = T_{f(x)}(g) \circ T_x(f),$$

$$T_x(\mathrm{id}) = \mathrm{id}.$$

We may reformulate Proposition 2.2:

Proposition 2.3. *Let X, Y be manifolds of class C^p $(p \geqq 1)$ modelled on Banach spaces. Let $f\colon X \to Y$ be a C^p-morphism. Let $x \in X$. Then:*

(i) *f is an immersion at x if and only if the map T_xf is injective and splits.*
(ii) *f is a submersion at x if and only if the map T_xf is surjective and its kernel splits.*

Note. If X, Y are finite dimensional, then the condition that T_xf splits is superfluous. Every subspace of a finite dimensional vector space splits.

Example. Let $\mathbf{E}$ be a (real) Hilbert space, and let $\langle x, y\rangle \in \mathbf{R}$ be its inner product. Then the square of the norm $f(x) = \langle x, x\rangle$ is obviously of class C^∞. The derivative $f'(x)$ is given by the formula

$$f'(x)y = 2\langle x, y\rangle$$

and for any given $x \neq 0$, it follows that the derivative $f'(x)$ is surjective. Furthermore, its kernel is the orthogonal complement of the subspace generated by x, and hence splits. Consequently the unit sphere in Hilbert space is a submanifold.

If W is a submanifold of a manifold Y of class C^p $(p \geqq 1)$, then the inclusion

$$i\colon\ W \to Y$$

induces a map

$$T_wi\colon\ T_w(W) \to T_w(Y)$$

which is in fact an injection. From the definition of a submanifold, one sees immediately that the image of T_wi splits. It will be convenient to identify $T_w(W)$ in $T_w(Y)$ if no confusion can result.

A morphism $f\colon X \to Y$ will be said to be **transversal** over the submanifold W of Y if the following condition is satisfied.

Let $x \in X$ be such that $f(x) \in W$. Let (V, ψ) be a chart at $f(x)$ such that $\psi\colon V \to V_1 \times V_2$ is an isomorphism on a product, with

$$\psi(f(x)) = (0, 0) \quad \text{and} \quad \psi(W \cap V) = V_1 \times 0.$$

Then there exists an open neighborhood U of x such that the composite map

$$U \xrightarrow{f} V \xrightarrow{\psi} V_1 \times V_2 \xrightarrow{\mathrm{pr}} V_2$$

is a submersion.

In particular, if f is transversal over W, then $f^{-1}(W)$ is a submanifold of X, because the inverse image of 0 by our local composite map

$$\text{pr} \circ \psi \circ f$$

is equal to the inverse image of $W \cap V$ by ψ.

As with immersions and submersions, we have a characterization of transversal maps in terms of tangent spaces.

Proposition 2.4. *Let X, Y be manifolds of class C^p $(p \geqq 1)$ modeled on Banach spaces. Let $f\colon X \to Y$ be a C^p-morphism, and W a submanifold of Y. The map f is transversal over W if and only if for each $x \in X$ such that $f(x)$ lies in W, the composite map*

$$T_x(X) \xrightarrow{T_x f} T_w(Y) \to T_w(Y)/T_w(W)$$

with $w = f(x)$ is surjective and its kernel splits.

Proof. If f is transversal over W, then for each point $x \in X$ such that $f(x)$ lies in W, we choose charts as in the definition, and reduce the question to one of maps of open subsets of Banach spaces. In that case, the conclusion concerning the tangent spaces follows at once from the assumed direct product decompositions. Conversely, assume our condition on the tangent map. The question being local, we can assume that $Y = V_1 \times V_2$ is a product of open sets in Banach spaces such that $W = V_1 \times 0$, and we can also assume that $X = U$ is open in some Banach space, $x = 0$. Then we let $g\colon U \to V_2$ be the map $\pi \circ f$ where π is the projection, and note that our assumption means that $g'(0)$ is surjective and its kernel splits. Furthermore, $g^{-1}(0) = f^{-1}(W)$. We can then use Corollary 5.7 of the inverse mapping theorem to conclude the proof.

Remark. In the statement of our proposition, we observe that the surjectivity of the composite map is equivalent to the fact that $T_w(Y)$ is equal to the sum of the image of $T_x f$ and $T_w(W)$, that is

$$T_w(Y) = \text{Im}(T_x f) + \text{Im}(T_x i),$$

where $i\colon W \to Y$ is the inclusion. In the finite dimensional case, the other condition is therefore redundant.

If $\mathbf{E}$ is a Banach space, then the diagonal Δ in $\mathbf{E} \times \mathbf{E}$ is a closed subspace and splits: Either factor $\mathbf{E} \times 0$ or $0 \times \mathbf{E}$ is a closed complement. Consequently, the diagonal is a closed submanifold of $\mathbf{E} \times \mathbf{E}$. If X is any manifold of class C^p, $p \geqq 1$, then the diagonal is therefore also a submanifold. (It is closed of course if and only if X is Hausdorff.)

Let $f\colon X \to Z$ and $g\colon Y \to Z$ be two C^p-morphisms, $p \geqq 1$. We say that they are **transversal** if the morphism

$$f \times g\colon X \times Y \to Z \times Z$$

is transversal over the diagonal. We remark right away that the surjectivity of the map in Proposition 2.4 can be expressed in two ways. Given two points $x \in X$ and $y \in Y$ such that $f(x) = g(y) = z$, the condition

$$\text{Im}(T_x f) + \text{Im}(T_y g) = T_z(Z)$$

is equivalent to the condition

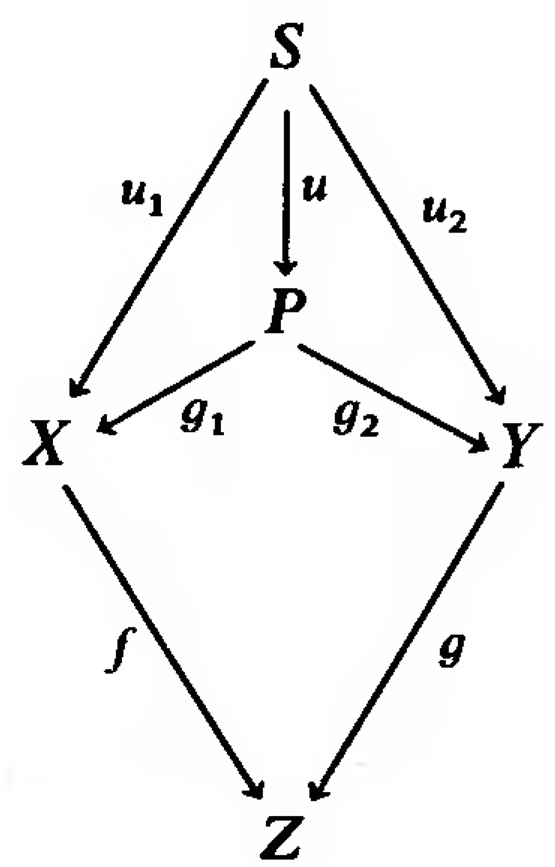
$$\text{Im}\big(T_{(x,y)}(f \times g)\big) + T_{(z,z)}(\Delta) = T_{(z,z)}(Z \times Z).$$

Thus in the finite dimensional case, we could take it as definition of transversality.

We use transversality as a sufficient condition under which the fiber product of two morphisms exists. We recall that in any category, the **fiber product** of two morphisms $f\colon X \to Z$ and $g\colon Y \to Z$ over Z consists of an object P and two morphisms

$$g_1\colon P \to X \quad \text{and} \quad g_2\colon P \to Y$$

such that $f \circ g_1 = g \circ g_2$, and satisfying the universal mapping property: Given an object S and two morphisms $u_1\colon S \to X$ and $u_2\colon S \to Y$ such that $fu_1 = gu_2$, there exists a unique morphism $u\colon S \to P$ making the following diagram commutative:



The triple (P, g_1, g_2) is uniquely determined, up to a unique isomorphism (in the obvious sense), and P is also denoted by $X \times_Z Y$.

One can view the fiber product unsymmetrically. Given two morphisms f, g as in the following diagram:

$$\begin{array}{ccc} & & Y \\ & & \downarrow{\scriptstyle g} \\ X & \xrightarrow[f]{} & Z \end{array}$$

assume that their fiber product exists, so that we can fill in the diagram:

$$\begin{array}{ccc} X \times_Z Y & \longrightarrow & Y \\ {\scriptstyle g_1}\downarrow & & \downarrow{\scriptstyle g} \\ X & \longrightarrow & Z \end{array}$$

We say that g_1 is the **pull back** of g by f, and also write it as $f^*(g)$. Similarly, we write $X \times_Z Y$ as $f^*(Y)$.

In our category of manifolds, we shall deal only with cases when the fiber product can be taken to be the set-theoretic fiber product on which a manifold structure has been defined. (The set-theoretic fiber product is the set of pairs of points projecting on the same point.) This determines the fiber product uniquely, and not only up to a unique isomorphism.

Proposition 2.5. *Let $f: X \to Z$ and $g: Y \to Z$ be two C^p-morphisms with $p \geqq 1$. If they are transversal, then*

$$(f \times g)^{-1}(\Delta_Z),$$

together with the natural morphisms into X and Y (obtained from the projections), is a fiber product of f and g over Z.

Proof. Obvious.

To construct a fiber product, it suffices to do it locally. Indeed, let $f: X \to Z$ and $g: Y \to Z$ be two morphisms. Let $\{V_i\}$ be an open covering of Z, and let

$$f_i: f^{-1}(V_i) \to V_i \quad \text{and} \quad g_i: g^{-1}(V_i) \to V_i$$

be the restrictions of f and g to the respective inverse images of V_i. Let $P = (f \times g)^{-1}(\Delta_Z)$. Then P consists of the points (x, y) with $x \in X$ and $y \in Y$ such that $f(x) = g(y)$. We view P as a subspace of $X \times Y$ (i.e. with the topology induced by that of $X \times Y$). Similarly, we construct P_i with f_i and g_i. Then P_i is open in P. The projections on the first and

second factors give natural maps of P_i into $f^{-1}(V_i)$ and $g^{-1}(V_i)$ and of P into X and Y.

Proposition 2.6. *Assume that each P_i admits a manifold structure (compatible with its topology) such that these maps are morphisms, making P_i into a fiber product of f_i and g_i. Then P, with its natural projections, is a fiber product of f and g.*

To prove the above assertion, we observe that the P_i form a covering of P. Furthermore, the manifold structure on $P_i \cap P_j$ induced by that of P_i or P_j must be the same, because it is the unique fiber product structure over $V_i \cap V_j$, for the maps f_{ij} and g_{ij} (defined on $f^{-1}(V_i \cap V_j)$ and $g^{-1}(V_i \cap V_j)$ respectively). Thus we can give P a manifold structure, in such a way that the two projections into X and Y are morphisms, and make P into a fiber product of f and g.

We shall apply the preceding discussion to vector bundles in the next chapter, and the following local criterion will be useful.

Proposition 2.7. *Let $f: X \to Z$ be a morphism, and $g: Z \times W \to Z$ be the projection on the first factor. Then f, g have a fiber product, namely the product $X \times W$ together with the morphisms of the following diagram:*

$$\begin{array}{ccc} X \times W & \xrightarrow{f \times \mathrm{id}} & Z \times W \\ {\scriptstyle \mathrm{pr}_1}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}_1} \\ X & \xrightarrow[f]{} & Z \end{array}$$

II, §3. PARTITIONS OF UNITY

Let X be a manifold of class C^p. A **function** on X will be a morphism of X into $\mathbf{R}$, of class C^p, unless otherwise specified. The C^p functions form a ring denoted by $\mathfrak{F}^p(X)$ or $\mathrm{Fu}^p(X)$. The **support** of a function f is the closure of the set of points x such that $f(x) \neq 0$.

Let X be a topological space. A covering of X is **locally finite** if every point has a neighborhood which intersects only finitely many elements of the covering. A **refinement** of a covering of X is a second covering, each element of which is contained in an element of the first covering. A topological space is **paracompact** if it is Hausdorff, and every open covering has a locally finite open refinement.

Proposition 3.1. *If X is a paracompact space, and if $\{U_i\}$ is an open covering, then there exists a locally finite open covering $\{V_i\}$ such that $V_i \subset U_i$ for each i.*

Proof. Let $\{V_k\}$ be a locally finite open refinement of $\{U_i\}$. For each k there is an index $i(k)$ such that $V_k \subset U_{i(k)}$. We let W_i be the union of those V_k such that $i(k) = i$. Then the W_i form a locally finite open covering, because any neighborhood of a point which meets infinitely many W_i must also meet infinitely many V_k.

Proposition 3.2. *If X is paracompact, then X is normal. If, furthermore, $\{U_i\}$ is a locally finite open covering of X, then there exists a locally finite open covering $\{V_i\}$ such that $\bar{V}_i \subset U_i$.*

Proof. We refer the reader to Bourbaki [Bou 68].

Observe that Proposition 3.1 shows that the insistence that the indexing set of a refinement be a given one can easily be achieved.

A **partition of unity** (of class C^p) on a manifold X consists of an open covering $\{U_i\}$ of X and a family of functions

$$\psi_i\colon X \to \mathbf{R}$$

satisfying the following conditions:

PU 1. *For all $x \in X$ we have $\psi_i(x) \geqq 0$.*

PU 2. *The support of ψ_i is contained in U_i.*

PU 3. *The covering is locally finite.*

PU 4. *For each point $x \in X$ we have*

$$\sum \psi_i(x) = 1.$$

(The sum is taken over all i, but is in fact finite for any given point x in view of **PU 3**.)

We sometimes say that $\{(U_i, \psi_i)\}$ is a partition of unity.

A manifold X will be said to **admit partitions of unity** if it is paracompact, and if, given a locally finite open covering $\{U_i\}$, there exists a partition of unity $\{\psi_i\}$ such that the support of ψ_i is contained in U_i.

If $\{U_i\}$ is a covering of X, then we say that a covering $\{V_k\}$ is subordinated to $\{U_i\}$ if each V_k is contained in some U_i.

It is desirable to give sufficient conditions on a manifold in order to insure the existence of partitions of unity. There is no difficulty with the topological aspects of this problem. It is known that a metric space is

paracompact (cf. Bourbaki [Bou 68], [Ke 55]), and on a paracompact space, one knows how to construct continuous partitions of unity (loc. cit.). However, in the case of infinite dimensional manifolds, certain difficulties arise to construct differentiable ones, and it is known that a Banach space itself may not admit partitions of unity (say of class C^∞). The construction of differentiable partitions of unity depends on the construction of a differentiable norm. Readers will find examples, theorems, and counterexamples in [BoF 65], [BoF 66], and [Re 64]. In the finite dimensional case, the existence will follow from the next theorem.

If $\mathbf{E}$ is a Banach space, we denote by $B_r(a)$ the open ball of radius r and center a, and by $\bar{B}_r(a)$ the closed ball of radius r and center a. If $a = 0$, then we write B_r and $\bar{B}_r$ respectively. Two open balls (of finite radius) are obviously C^∞-isomorphic. If X is a manifold and (V, φ) is a chart at a point $x \in X$, then we say that (V, φ) (or simply V) is a ball of radius r if φV is a ball of radius r in the Banach space.

Theorem 3.3. *Let X be a manifold which is locally compact, Hausdorff, and whose topology has a countable base. Given an open covering of X, then there exists an atlas $\{(V_k, \varphi_k)\}$ such that the covering $\{V_k\}$ is locally finite and subordinated to the given covering, such that $\varphi_k V_k$ is the open ball B_3, and such that the open sets $W_k = \varphi_k^{-1}(B_1)$ cover X.*

Proof. Let $U_1, U_2, \ldots$ be a basis for the open sets of X such that each $\bar{U}_i$ is compact. We construct inductively a sequence $A_1, A_2, \ldots$ of compact sets whose union is X, such that A_i is contained in the interior of A_{i+1}. We let $A_1 = \bar{U}_1$. Suppose we have constructed A_i. We let j be the smallest integer such that A_i is contained in $U_1 \cup \cdots \cup U_j$. We let A_{i+1} be the closed and compact set

$$\bar{U}_1 \cup \cdots \cup \bar{U}_j \cup \bar{U}_{i+1}.$$

For each point $x \in X$ we can find an arbitrarily small chart (V_x, φ_x) at x such that $\varphi_x V_x$ is the ball of radius 3 (so that each V_x is contained in some element of U). We let $W_x = \varphi_x^{-1}(B_1)$ be the ball of radius 1 in this chart. We can cover the set

$$A_{i+1} - \operatorname{Int}(A_i)$$

(intuitively the closed annulus) by a finite number of these balls of radius 1, say $W_1, \ldots, W_n$, such that, at the same time, each one of $V_1, \ldots, V_n$ is contained in the open set $\operatorname{Int}(A_{i+2}) - A_{i-1}$ (intuitively, the open annulus of the next bigger size). We let $\mathfrak{B}_i$ denote the collection $V_1, \ldots, V_n$ and let $\mathfrak{B}$ be composed of the union of the $\mathfrak{B}_i$. Then $\mathfrak{B}$ is locally finite, and we are done.

Corollary 3.4. *Let X be a manifold which is locally compact Hausdorff, and whose topology has a countable base. Then X admits partitions of unity.*

Proof. Let $\{(V_k, \varphi_k)\}$ be as in the theorem, and $W_k = \varphi_k^{-1}(B_1)$. We can find a function ψ_k of class C^p such that $0 \leqq \psi_k \leqq 1$, such that $\psi_k(x) = 1$ for $x \in W_k$ and $\psi_k(x) = 0$ for $x \notin V_k$. (The proof is recalled below.) We now let

$$\psi = \sum \psi_k$$

(a sum which is finite at each point), and we let $\gamma_k = \psi_k/\psi$. Then $\{(V_k, \gamma_k)\}$ is the desired partition of unity.

We now recall the argument giving the function ψ_k. First, given two real numbers r, s with $0 \leqq r < s$, the function defined by

$$\exp\left(\frac{-1}{(t-r)(s-t)}\right)$$

in the open interval $r < t < s$ and 0 outside the interval determines a bell-shaped C^∞-function from **R** into **R**. Its integral from minus infinity to t, divided by the area under the bell yields a function which lies strictly between 0 and 1 in the interval $r < t < s$, is equal to 0 for $t \leqq r$ and is equal to 1 for $t \geqq s$. (The function is even monotone increasing.)

We can therefore find a real valued function of a real variable, say $\eta(t)$, such that $\eta(t) = 1$ for $|t| < 1$ and $\eta(t) = 0$ for $|t| \geqq 1 + \delta$ with small δ, and such that $0 \leqq \eta \leqq 1$. If **E** is a Hilbert space, then $\eta(|x|^2) = \psi(x)$ gives us a function which is equal to 1 on the ball of radius 1 and 0 outside the ball of radius $1 + \delta$. This function can then be transported to the manifold by any given chart whose image is the ball of radius 3.

In a similar way, one would construct a function which is >0 on a given ball and $=0$ outside this ball.

Partitions of unity constitute the only known means of gluing together local mappings (into objects having an addition, namely vector bundles, discussed in the next chapter). It is therefore important, in both the Banach and Hilbert cases, to determine conditions under which they exist. In the Banach case, there is the added difficulty that the argument just given to get a local function which is 1 on B_1 and 0 outside B_2 fails if one cannot find a differentiable function of the norm, or of an equivalent norm used to define the Banachable structure.

Even though it is not known whether Theorem 3.3 extends to Hilbert manifolds, it is still possible to construct partitions of unity in that case. As Eells pointed out to me, Dieudonné's method of proof showing that separable metric space is paracompact can be applied for that purpose (this is Lemma 3.5 below), and I am indebted to him for the following exposition.

We need some lemmas. We use the notation ${}^cA$ for the complement of a set A.

Let M be a metric space with distance function d. We can then speak of open and closed balls. For instance $\bar{B}_a(x)$ denotes the closed ball of radius a with center x. It consists of all points y with $d(y, x) \leqq a$. An open subset V of M will be said to be **scalloped** if there exist open balls $U, U_1, \ldots, U_m$ in M such that

$$V = U \cap {}^c\bar{U}_1 \cap \cdots \cap {}^c\bar{U}_m.$$

A covering $\{V_i\}$ of a subset W of M is said to be locally finite (with respect to W) if every point $x \in W$ has a neighborhood which meets only a finite number of elements of the covering.

Lemma 3.5. *Let M be a metric space and $\{U_i\}$ $(i = 1, 2, \ldots)$ a countable covering of a subset W by open balls. Then there exists a locally finite open covering $\{V_i\}$ $(i = 1, 2, \ldots)$ of W such that $V_i \subset U_i$ for all i, and such that V_i is scalloped for all i.*

Proof. We define V_i inductively as follows. Each U_i is a ball, say $B_{a_i}(x_i)$. Let $V_1 = U_1$. Having defined V_{i-1}, let

$$r_{1i} = a_1 - \frac{1}{i}, \quad \ldots, \quad r_{i-1,i} = a_{i-1} - \frac{1}{i}$$

and let

$$V_i = U_i \cap {}^c\bar{B}_{r_{1i}}(x_1) \cap \cdots \cap {}^c\bar{B}_{r_{i-1,i}}(x_{i-1}),$$

it being understood that a ball of negative radius is empty. Then each V_i is scalloped, and is contained in U_i. We contend that the V_i cover W. Indeed, let x be an element of W. Let j be the smallest index such that $x \in U_j$. Then $x \in V_j$, for otherwise, x would be in the complement of V_j which is equal to the union of ${}^cU_j$ and the balls

$$\bar{B}_{r_{1j}}(x_1) \cup \cdots \cup \bar{B}_{r_{j-1,j}}(x_{j-1}).$$

Hence x would lie in some U_i with $i < j$, contradiction.

There remains to be shown that our covering $\{V_i\}$ is locally finite. Let $x \in W$. Then x lies in some U_n. Let s be such a small number >0 that the ball $B_s(x)$ is contained in U_n. Let $t = s/2$. For all i sufficiently large, the ball $B_t(x)$ is contained in $\bar{B}_{a_n - 1/i}(x_n) = \bar{B}_{r_{ni}}(x_n)$ and therefore this ball does not meet V_i. We have found a neighborhood of x which meets only a finite number of members of our covering, which is consequently locally finite (with respect to W).

Lemma 3.6. *Let U be an open ball in Hibert space* $\mathbf{E}$ *and let*

$$V = U \cap {}^c\bar{U}_1 \cap \cdots \cap {}^c\bar{U}_m$$

be a scalloped open subset. Then there exists a C^∞-function ω: $\mathbf{E} \to \mathbf{R}$ such that $\omega(x) > 0$ if $x \in V$ and $\omega(x) = 0$ otherwise.

Proof. For each U_i let φ_i: $\mathbf{E} \to \mathbf{R}$ be a function such that

$$0 \leqq \varphi_i(x) < 1 \quad \text{if} \quad x \in {}^c\bar{U}_i,$$

$$\varphi_i(x) = 1 \quad \text{if} \quad x \in \bar{U}_i.$$

Let $\varphi(x)$ be a function such that $\varphi(x) > 0$ on U and $\varphi(x) = 0$ outside U. Let

$$\omega(x) = \varphi(x) \prod (1 - \varphi_i(x)).$$

Then $\omega(x)$ satisfies our requirements.

Theorem 3.7. *Let A_1, A_2 be non-void, closed, disjoint subsets of a separable Hilbert space* $\mathbf{E}$. *Then there exists a C^∞-function ψ: $\mathbf{E} \to \mathbf{R}$ such that $\psi(x) = 0$ if $x \in A_1$ and $\psi(x) = 1$ if $x \in A_2$, and $0 \leqq \psi(x) \leqq 1$ for all x.*

Proof. By Lindelöf's theorem, we can find a countable collection of open balls $\{U_i\}$ $(i = 1, 2, \ldots)$ covering A_2 and such that each U_i is contained in the complement of A_1. Let W be the union of the U_i. We find a locally finite refinement $\{V_i\}$ as in Lemma 3.5. Using Lemma 3.6, we find a function ω_i which is >0 on V_i and 0 outside V_i. Let $\omega = \sum \omega_i$ (the sum is finite at each point of W). Then $\omega(x) > 0$ if $x \in A_2$, and $\omega(x) = 0$ if $x \in A_1$.

Let U be the open neighborhood of A_2 on which ω is >0. Then A_2 and ${}^cU$ are disjoint closed sets, and we can apply the above construction to obtain a function σ: $\mathbf{E} \to \mathbf{R}$ which is >0 on ${}^cU$ and $= 0$ on A_2. We let $\psi = \omega/(\sigma + \omega)$. Then ψ satisfies our requirements.

Corollary 3.8. *Let X be a paracompact manifold of class C^p, modeled on a separable Hilbert space* $\mathbf{E}$. *Then X admits partitions of unity (of class C^p).*

Proof. It is trivially verified that an open ball of finite radius in $\mathbf{E}$ is C^∞-isomorphic to $\mathbf{E}$. (We reproduce the formula in Chapter VII.) Given any point $x \in X$, and a neighborhood N of x, we can therefore always find a chart (G, γ) at x such that $\gamma G = \mathbf{E}$, and $G \subset N$. Hence, given an open covering of X, we can find an atlas $\{(G_\alpha, \gamma_\alpha)\}$ subordinated to the given covering, such that $\gamma_\alpha G_\alpha = \mathbf{E}$. By paracompactness, we can find a refinement $\{U_i\}$ of the covering $\{G_\alpha\}$ which is locally finite. Each U_i is contained in some $G_{\alpha(i)}$ and we let φ_i be the restriction of $\gamma_{\alpha(i)}$ to U_i. We now find open refinements $\{V_i\}$ and then $\{W_i\}$ such that

$$\bar{W}_i \subset V_i \subset \bar{V}_i \subset U_i,$$

the bar denoting closure in X. Each $\bar{V}_i$ being closed in X, it follows from our construction that $\varphi_i \bar{V}_i$ is closed in $\mathbf{E}$, and so is $\varphi_i \bar{W}_i$. Using the theorem, and transporting functions on $\mathbf{E}$ to functions on X by means of the φ_i, we can find for each i a C^p-function ψ_i: $X \to \mathbf{R}$ with is 1 on $\bar{W}_i$ and 0 on $X - V_i$. We let $\psi = \sum \psi_i$ and $\theta_i = \psi_i/\psi$. Then the collection $\{\theta_i\}$ is the desired partition of unity.

II, §4. MANIFOLDS WITH BOUNDARY

Let $\mathbf{E}$ be a Banach space, and λ: $\mathbf{E} \to \mathbf{R}$ a continuous linear map into $\mathbf{R}$. (This will also be called a **functional** on $\mathbf{E}$.) We denote by $\mathbf{E}_\lambda^0$ the kernel of λ, and by $\mathbf{E}_\lambda^+$ (resp. $\mathbf{E}_\lambda^-$) the set of points $x \in \mathbf{E}$ such that $\lambda(x) \geqq 0$ (resp. $\lambda(x) \leqq 0$). We call $\mathbf{E}_\lambda^0$ a **hyperplane** and $\mathbf{E}_\lambda^+$ or $\mathbf{E}_\lambda^-$ a **half plane**.

If μ is another functional and $\mathbf{E}_\lambda^+ = \mathbf{E}_\mu^+$, then there exists a number $c > 0$ such that $\lambda = c\mu$. This is easily proved. Indeed, we see at once that the kernels of λ and μ must be equal. Suppose $\lambda \neq 0$. Let x_0 be such that $\lambda(x_0) > 0$. Then $\mu(x_0) > 0$ also. The functional

$$\lambda - (\lambda(x_0)/\mu(x_0))\mu$$

vanishes on the kernel of λ (or μ) and also on x_0. Therefore it is the 0 functional, and $c = \lambda(x_0)/\mu(x_0)$.

Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces, and let $\mathbf{E}_\lambda^+$ and $\mathbf{F}_\mu^+$ be two half planes in $\mathbf{E}$ and $\mathbf{F}$ respectively. Let U, V be two open subsets of these half planes respectively. We shall say that a mapping

$$f\colon U \to V$$

is a morphism of class C^p if the following condition is satisfied. Given a point $x \in U$, there exists an open neighborhood U_1 of x in $\mathbf{E}$, an open neighborhood V_1 of $f(x)$ in $\mathbf{F}$, and a morphism f_1: $U_1 \to V_1$ (in the sense of Chapter I) such that the restriction of f_1 to $U_1 \cap U$ is equal to f. (We assume that all morphisms are of class C^p with $p \geqq 1$.)

If our half planes are full planes (i.e. equal to the vector spaces themselves), then our present definition is the same as the one used previously.

If we take as objects the open subsets of half planes in Banach spaces, and as morphisms the C^p-morphisms, then we obtain a category. The notion of isomorphism is therefore defined, and the definition of manifold by means of atlases and charts can be used as before. The manifolds of §1 should have been called **manifolds without boundary**, reserving the name of manifold for our new globalized objects. However, in most of this book, we shall deal exclusively with manifolds without boundary for simplicity. The following remarks will give readers the means of extending any result they wish (provided it is true) for the case of manifolds without boundaries to the case manifolds with.

First, concerning the notion of derivative, we have:

Proposition 4.1. *Let $f\colon U \to \mathbf{F}$ and $g\colon U \to \mathbf{F}$ be two morphisms of class C^p ($p \geqq 1$) defined on an open subset U of $\mathbf{E}$. Assume that f and g have the same restriction to $U \cap \mathbf{E}_\lambda^+$ for some half plane $\mathbf{E}_\lambda^+$, and let*

$$x \in U \cap \mathbf{E}_\lambda^+.$$

Then $f'(x) = g'(x)$.

Proof. After considering the difference of f and g, we may assume without loss of generality that the restriction of f to $U \cap \mathbf{E}_\lambda^+$ is 0. It is then obvious that $f'(x) = 0$.

Proposition 4.2. *Let U be open in $\mathbf{E}$. Let μ be a non-zero functional on $\mathbf{F}$ and let $f\colon U \to \mathbf{F}_\mu^+$ be a morphism of class C^p with $p \geqq 1$. If x is a point of U such that $f(x)$ lies in $\mathbf{F}_\mu^0$ then $f'(x)$ maps $\mathbf{E}$ into $\mathbf{F}_\mu^0$.*

Proof. Without loss of generality, we may assume that $x = 0$ and $f(x) = 0$. Let W be a given neighborhood of 0 in $\mathbf{F}$. Suppose that we can find a small element $v \in \mathbf{E}$ such that $\mu f'(0)v \neq 0$. We can write (for small t):

$$f(tv) = tf'(0)v + o(t)w_t$$

with some element $w_t \in W$. By assumption, $f(tv)$ lies in $\mathbf{F}_\mu^+$. Applying μ we get

$$t\mu f'(0)v + o(t)\mu(w_t) \geqq 0.$$

Dividing by t, this yields

$$\mu f'(0)v \geqq \frac{o(t)}{t}\mu(w_t).$$

Replacing t by $-t$, we get a similar inequality on the other side. Letting t tend to 0 shows that $\mu f'(0)v = 0$, a contradiction.

Let U be open in some half plane $\mathbf{E}_\lambda^+$. We define the **boundary** of U (written ∂U) to be the intersection of U with $\mathbf{E}_\lambda^0$, and the **interior** of U (written $\mathrm{Int}(U)$) to be the complement of ∂U in U. Then $\mathrm{Int}(U)$ is open in $\mathbf{E}$.

It follows at once from our definition of differentiability that a half plane is C^∞-isomorphic with a product

$$\mathbf{E}_\lambda^+ \approx \mathbf{E}_\lambda^0 \times \mathbf{R}^+$$

where $\mathbf{R}^+$ is the set of real numbers $\geqq 0$, whenever $\lambda \neq 0$. The boundary of $\mathbf{E}_\lambda^+$ in that case is $\mathbf{E}_\lambda^0 \times 0$.

Proposition 4.3. *Let λ be a functional on $\mathbf{E}$ and μ a functional on $\mathbf{F}$. Let U be open in $\mathbf{E}_\lambda^+$ and V open in $\mathbf{F}_\mu^+$ and assume $U \cap \mathbf{E}_\lambda^0$, $V \cap \mathbf{F}_\mu^0$ are not empty. Let $f\colon U \to V$ be an isomorphism of class C^p ($p \geqq 1$). Then $\lambda \neq 0$ if and only if $\mu \neq 0$. If $\lambda \neq 0$, then f induces a C^p-isomorphism of $\mathrm{Int}(U)$ on $\mathrm{Int}(V)$ and of ∂U on ∂V.*

Proof. By the functoriality of the derivative, we know that $f'(x)$ is a toplinear isomorphism for each $x \in U$. Our first assertion follows from the preceding proposition. We also see that no interior point of U maps on a boundary point of V and conversely. Thus f induces a bijection of ∂U on ∂V and a bijection of $\mathrm{Int}(U)$ on $\mathrm{Int}(V)$. Since these interiors are open in their respective spaces, our definition of derivative shows that f induces an isomorphism between them. As for the boundary, it is a submanifold of the full space, and locally, our definition of derivative, together with the product structure, shows that the restriction of f to ∂U must be an isomorphism on ∂V.

This last proposition shows that the boundary is a differentiable invariant, and thus that we can speak of the boundary of a manifold.

We give just two words of warning concerning manifolds with boundary. First, products do not exist in their category. Indeed, to get products, we are forced to define manifolds with **corners**, which would take us too far afield.

Second, in defining immersions or submanifolds, there is a difference in kind when we consider a manifold embedded in a manifold without boundary, or a manifold embedded in another manifold with boundary. Think of a closed interval embedded in an ordinary half plane. Two cases arise. The case where the interval lies inside the interior of the half plane is essentially distinct from the case where the interval has one end point touching the hyperplane forming the boundary of the half plane. (For instance, given two embeddings of the first type, there exists an automorphism of the half plane carrying one into the other, but there cannot exist an automorphism of the half plane carrying an embedding of the first type into one of the second type.)

We leave it to the reader to go systematically through the notions of tangent space, immersion, embedding (and later, tangent bundle, vector field, etc.) for arbitrary manifolds (with boundary). For instance, Proposition 2.2 shows at once how to get the tangent space functorially.

CHAPTER III

Vector Bundles

The collection of tangent spaces can be glued together to give a manifold with a natural projection, thus giving rise to the tangent bundle. The general glueing procedure can be used to construct more general objects known as vector bundles, which give powerful invariants of a given manifold. (For an interesting theorem see Mazur [Maz 61].) In this chapter, we develop purely formally certain functorial constructions having to do with vector bundles. In the chapters on differential forms and Riemannian metrics, we shall discuss in greater details the constructions associated with multilinear alternating forms, and symmetric positive definite forms.

Partitions of unity are an essential tool when considering vector bundles. They can be used to combine together a random collection of morphisms into vector bundles, and we shall give a few examples showing how this can be done (concerning exact sequences of bundles).

III, §1. DEFINITION, PULL BACKS

Let X be a manifold (of class C^p with $p \geqq 0$) and let $\pi\colon E \to X$ be a morphism. Let $\mathbf{E}$ be a Banach space.

Let $\{U_i\}$ be an open covering of X, and for each i, suppose that we are given a mapping

$$\tau_i\colon \pi^{-1}(U_i) \to U_i \times \mathbf{E}$$

satisfying the following conditions:

VB 1. *The map* τ_i *is a* C^p *isomorphism commuting with the projection on* U_i, *that is, such that the following diagram is commutative*:

$$\begin{array}{ccc} \pi^{-1}(U_i) & \xrightarrow{\tau_i} & U_i \times \mathbf{E} \\ & \searrow \quad \swarrow & \\ & U_i & \end{array}$$

In particular, we obtain an isomorphism on each fiber (*written* $\tau_i(x)$ *or* τ_{ix})

$$\tau_{ix}\colon \pi^{-1}(x) \to \{x\} \times \mathbf{E}$$

VB 2. *For each pair of open sets* U_i, U_j *the map*

$$\tau_{jx} \circ \tau_{ix}^{-1}\colon \mathbf{E} \to \mathbf{E}$$

is a toplinear isomorphism.

VB 3. *If* U_i *and* U_j *are two members of the covering, then the map of* $U_i \cap U_j$ *into* $L(\mathbf{E},\mathbf{E})$ (*actually* $\mathrm{Laut}(\mathbf{E})$) *given by*

$$x \mapsto (\tau_j\tau_i^{-1})_x$$

is a morphism.

Then we shall say that $\{(U_i, \tau_i)\}$ is a **trivializing covering** for π (or for E by abuse of language), and that $\{\tau_i\}$ are its **trivalizing** maps. If $x \in U_i$, we say that τ_i (or U_i) trivializes at x. Two trivializing coverings for π are said to be **VB-equivalent** if taken together they also satisfy conditions **VB 2**, **VB 3**. An equivalence class of such trivializing coverings is said to determine a structure of **vector bundle** on π (or on E by abuse of language). We say that E is the **total space** of the bundle, and that X is its **base space**. If we wish to be very functorial, we shall write E_π and X_π for these spaces respectively. The fiber $\pi^{-1}(x)$ is also denoted by E_x or π_x. We also say that the vector bundle has **fiber** $\mathbf{E}$, or is **modeled on** $\mathbf{E}$. Note that from **VB 2**, the fiber $\pi^{-1}(x)$ above each point $x \in X$ can be given a structure of Banachable space, simply by transporting the Banach space structure of $\mathbf{E}$ to $\pi^{-1}(x)$ via τ_{ix}. Condition **VB 2** insures that using two different trivializing maps τ_{ix} or τ_{jx} will give the same structure of Banachable space (with equivalent norms, of course not the same norms).

Conversely, we could replace **VB 2** by a similar condition as follows.

VB 2′. *On each fiber* $\pi^{-1}(x)$ *we are given a structure of Banachable space, and for* $x \in U_i$, *the trivializing map*

$$\tau_{ix}\colon \pi^{-1}(x) = E_x \to \mathbf{E}$$

is a toplinear isomorphism.

Then it follows that $\tau_{jx} \circ \tau_{ix}^{-1}\colon \mathbf{E} \to \mathbf{E}$ is a toplinear isomorphism for each pair of open sets U_i, U_j and $x \in U_i \cap U_j$.

In the finite dimensional case, condition **VB 3** is implied by **VB 2**.

Proposition 1.1. *Let* $\mathbf{E}$, $\mathbf{F}$ *be finite dimensional vector spaces. Let* U *be open in some Banach space. Let*

$$f\colon U \times \mathbf{E} \to \mathbf{F}$$

be a morphism such that for each $x \in U$, *the map*

$$f_x\colon \mathbf{E} \to \mathbf{F}$$

given by $f_x(v) = f(x, v)$ *is a linear map. Then the map of* U *into* $L(\mathbf{E}, \mathbf{F})$ *given by* $x \mapsto f_x$ *is a morphism.*

Proof. We can write $\mathbf{F} = \mathbf{R}_1 \times \cdots \times \mathbf{R}_n$ (n copies of $\mathbf{R}$). Using the fact that $L(\mathbf{E}, \mathbf{F}) = L(\mathbf{E}, \mathbf{R}_1) \times \cdots \times L(\mathbf{E}, \mathbf{R}_n)$, it will suffice to prove our assertion when $\mathbf{F} = \mathbf{R}$. Similarly, we can assume that $\mathbf{E} = \mathbf{R}$ also. But in that case, the function $f(x, v)$ can be written $g(x)v$ for some map $g\colon U \to \mathbf{R}$. Since f is a morphism, it follows that as a function of each argument x, v it is also a morphism. Putting $v = 1$ shows that g is a morphism and concludes the proof.

Returning to the general definition of a vector bundle, we call the maps

$$\tau_{jix} = \tau_{jx} \circ \tau_{ix}^{-1}$$

the **transition** maps associated with the covering. They satisfy what we call the **cocycle condition**

$$\tau_{kjx} \circ \tau_{jix} = \tau_{kix}.$$

In particular, $\tau_{iix} = \mathrm{id}$ and $\tau_{jix} = \tau_{ijx}^{-1}$.

As with manifolds, we can recover a vector bundle from a trivializing covering.

Proposition 1.2. *Let* X *be a manifold, and* $\pi\colon E \to X$ *a mapping from some set* E *into* X. *Let* $\{U_i\}$ *be an open covering of* X, *and for each* i *suppose that we are given a Banach space* $\mathbf{E}$ *and a bijection* (*commuting with the projection on* U_i),

$$\tau_i\colon \pi^{-1}(U_i) \to U_i \times \mathbf{E},$$

such that for each pair i, j *and* $x \in U_i \cap U_j$, *the map* $(\tau_j\tau_i^{-1})_x$ *is a* *toplinear isomorphism, and condition* **VB 3** *is satisfied as well as the cocycle condition. Then there exists a unique structure of manifold on E such that π is a morphism, such that τ_i is an isomorphism making π into a vector bundle, and $\{(U_i, \tau_i)\}$ into a trivialising covering.*

Proof. By Proposition 3.10 of Chapter I and our condition **VB 3**, we conclude that the map

$$\tau_j\tau_i^{-1}\colon (U_i \cap U_j) \times \mathbf{E} \to (U_i \cap U_j) \times \mathbf{E}$$

is a morphism, and in fact an isomorphism since it has an inverse. From the definition of atlases, we conclude that E has a unique manifold structure such that the τ_i are isomorphisms. Since π is obtained locally as a composite of morphisms (namely τ_i and the projections of $U_i \times \mathbf{E}$ on the first factor), it becomes a morphism. On each fiber $\pi^{-1}(x)$, we can transport the topological vector space structure of any $\mathbf{E}$ such that x lies in U_i, by means of τ_{ix}. The result is independent of the choice of U_i since $(\tau_j\tau_i^{-1})_x$ is a toplinear isomorphism. Our proposition is proved.

Remark. It is relatively rare that a vector bundle is **trivial**, i.e. VB-isomorphic to a product $X \times \mathbf{E}$. By definition, it is always trivial locally. In the finite dimensional case, say when E has dimension n, a trivialization is equivalent to the existence of sections $\xi_1, \dots, \xi_n$ such that for each x, the vectors $\xi_1(x), \dots, \xi_n(x)$ form a basis of E_x. Such a choice of sections is called a **frame** of the bundle, and is used especially with the tangent bundle, to be defined below. In this book where we give proofs valid in the infinite dimensional case, frames will therefore not occur until we get to strictly finite dimensional phenomenon.

The local representation of a vector bundle and the vector component of a morphism

For arbitrary vector bundles (and especially the tangent bundle to be defined below), we have a local representation of the bundle as a product in a chart. For many purposes, and especially the case of a morphism

$$f\colon Y \to E$$

of a manifold into the vector bundle, it is more convenient to use U to denote an open subset of a Banach space, and to let $\varphi\colon U \to X$ be an isomorphism of U with an open subset of X over which E has a trivialization $\tau\colon \pi^{-1}(\varphi U) \to U \times \mathbf{E}$ called a **VB-chart**. Suppose V is an

open subset of Y such that $f(V) \subset \pi^{-1}(\varphi U)$. We then have the commutative diagram:

$$\begin{array}{ccccc} V & \xrightarrow{f} & \pi^{-1}(\varphi U) & \xrightarrow{\tau} & U \times \mathbf{E} \\ & & \downarrow & & \downarrow \\ & & \varphi U & \xrightarrow{\varphi^{-1}} & U \end{array}$$

The composite $\tau \circ f$ is a morphism of V into $U \times \mathbf{E}$, which has two components

$$\tau \circ f = (f_{U1}, f_{U2})$$

such that $f_{U1}\colon V \to U$ and $f_{U2}\colon V \to \mathbf{E}$. We call f_{U2} the **vector component of f in the vector bundle chart** $U \times \mathbf{E}$ over U. Sometimes to simplify the notation, we omit the subscript, and merely agree that $f_U = f_{U2}$ denotes this vector component; or to simplify the notation further, we may simply state that f itself denotes this vector component if a discussion takes place entirely in a chart. In this case, we say that $f = f_U$ **represents the morphism** in the vector bundle chart, or in the chart.

Vector bundle morphisms and pull backs

We now make the set of vector bundles into a category.

Let $\pi\colon E \to X$ and $\pi'\colon E' \to X'$ be two vector bundles. A **VB-morphism** $\pi \to \pi'$ consists of a pair of morphisms

$$f_0\colon X \to X' \quad \text{and} \quad f\colon E \to E'$$

satisfying the following conditions.

VB Mor 1. *The diagram*

$$\begin{array}{ccc} E & \xrightarrow{f} & E' \\ \pi \downarrow & & \downarrow \pi' \\ X & \xrightarrow[f_0]{} & X' \end{array}$$

is commutative, and the induced map for each $x \in X$

$$f_x\colon E_x \to E'_{f(x)}$$

is a continuous linear map.

VB Mor 2. *For each $x_0 \in X$ there exist trivializing maps*

$$\tau\colon \pi^{-1}(U) \to U \times \mathbf{E}$$

and

$$\tau'\colon \pi'^{-1}(U') \to U' \times \mathbf{E}'$$

at x_0 and $f(x_0)$ respectively, such that $f_0(U)$ is contained in U', and such that the map of U into $L(\mathbf{E}, \mathbf{E}')$ given by

$$x \mapsto \tau'_{f_0(x)} \circ f_x \circ \tau^{-1}$$

is a morphism.

As a matter of notation, we shall also use f to denote the VB-morphism, and thus write $f\colon \pi \to \pi'$. In most applications, f_0 is the identity. By Proposition 1.1, we observe that **VB Mor 2** is redundant in the finite dimensional case.

The next proposition is the analogue of Proposition 1.2 for VB-morphisms.

Proposition 1.3. *Let π, π' be two vector bundles over manifolds X, X' respectively. Let $f_0\colon X \to X'$ be a morphism, and suppose that we are given for each $x \in X$ a continuous linear map*

$$f_x\colon \pi_x \to \pi'_{f_0(x)}$$

such that, for each x_0, condition **VB Mor 2** *is satisfied. Then the map f from π to π' defined by f_x on each fiber is a* VB*-morphism.*

Proof. One must first check that f is a morphism. This can be done under the assumption that π, π' are trivial, say equal to $U \times \mathbf{E}$ and $U' \times \mathbf{E}'$ (following the notation of **VB Mor 2**), with trivialising maps equal to the identity. Our map f is then given by

$$(x, v) \mapsto (f_0 x, f_x v).$$

Using Proposition 3.10 of Chapter I, we conclude that f is a morphism, and hence that (f_0, f) is a VB-morphism.

It is clear how to compose two VB-morphisms set theoretically. In fact, the composite of two VB-morphisms is a VB-morphism. There is no problem verifying condition **VB Mor 1**, and for **VB Mor 2**, we look at the situation locally. We encounter a commutative diagram of the following

type:

$$\begin{array}{ccccc}
\pi^{-1}(U) & \xrightarrow{f} & \pi'^{-1}(U') & \xrightarrow{g} & \pi''^{-1}(U'') \\
\downarrow{\scriptstyle \tau} & & \downarrow{\scriptstyle \tau'} & & \downarrow{\scriptstyle \tau''} \\
U \times \mathbf{E} & \longrightarrow & U' \times \mathbf{E}' & \longrightarrow & U'' \times \mathbf{E}''
\end{array}$$

and use Proposition 3.10 of Chapter I, to show that $g \circ f$ is a VB-morphism.

We therefore have a category, denoted by VB or VB^p, if we need to specify explicitly the order of differentiability.

The vector bundles over X from a subcategory $\mathrm{VB}(X) = \mathrm{VB}^p(X)$ (taking those VB-morphisms for which the map f_0 is the identity). If $\mathfrak{A}$ is a category of Banach spaces (for instance finite dimensional spaces), then we denote by $\mathrm{VB}(X, \mathfrak{A})$ those vector bundles over X whose fibers lie in $\mathfrak{A}$.

A morphism from one vector bundle into another can be given locally. More precisely, suppose that U is an open subset of X and $\pi\colon E \to X$ a vector bundle over X. Let $E_U = \pi^{-1}(U)$ and

$$\pi_U = \pi \mid E_U$$

be the restriction of π to E_U. Then π_U is a vector bundle over U. Let $\{U_i\}$ be an open covering of the manifold X and let π, π' be two vector bundles over X. Suppose, given a VB-morphism

$$f_i\colon \pi_{U_i} \to \pi'_{U_i}$$

for each i, such that f_i and f_j agree over $U_i \cap U_j$ for each pair of indices i, j. Then there exists a unique VB-morphism $f\colon \pi \to \pi'$ which agrees with f_i on each U_i. The proof is trivial, but the remark will be used frequently in the sequel.

Using the discussion at the end of Chapter II, §2 and Proposition 2.7 of that chapter, we get immediately:

Proposition 1.4. *Let $\pi\colon E \to Y$ be a vector bundle, and $f\colon X \to Y$ a morphism. Then*

$$f^*(\pi)\colon f^*(E) \to X$$

is a vector bundle called the **pull-back**, *and the pair $(f, \pi^*(f))$ is a* VB-*morphism*

$$\begin{array}{ccc}
f^*(E) & \xrightarrow{\pi^*(f)} & E \\
{\scriptstyle f^*(\pi)}\downarrow & & \downarrow{\scriptstyle \pi} \\
X & \xrightarrow[f]{} & Y
\end{array}$$

In Proposition 1.4, we could take f to be the inclusion of a submanifold. In that case, the pull-back is merely the restriction. As with open sets, we can then use the usual notation:

$$E_X = \pi^{-1}(X) \qquad \text{and} \qquad \pi_X = \pi \,|\, E_X.$$

Thus $\pi_X = f^*(\pi)$ in that case.

If X happens to be a point y of Y, then we have the constant map

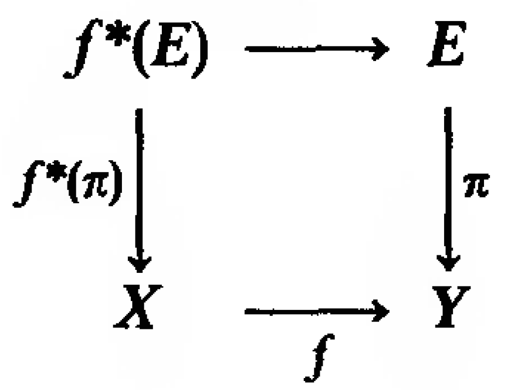
$$\pi_y\colon E_y \to y$$

which will sometimes be identified with E_y.

If we identify each fiber $(f^*E)_x$ with $E_{f(x)}$ itself (a harmless identification since an element of the fiber at x is simply a pair (x, e) with e in $E_{f(x)}$), then we can describe the pull-back f^* of a vector bundle $\pi\colon E \to Y$ as follows. It is a vector bundle $f^*\pi\colon f^*E \to X$ satisfying the following properties:

PB 1. *For each $x \in X$, we have $(f^*E)_x = E_{f(x)}$.*

PB 2. *We have a commutative diagram*



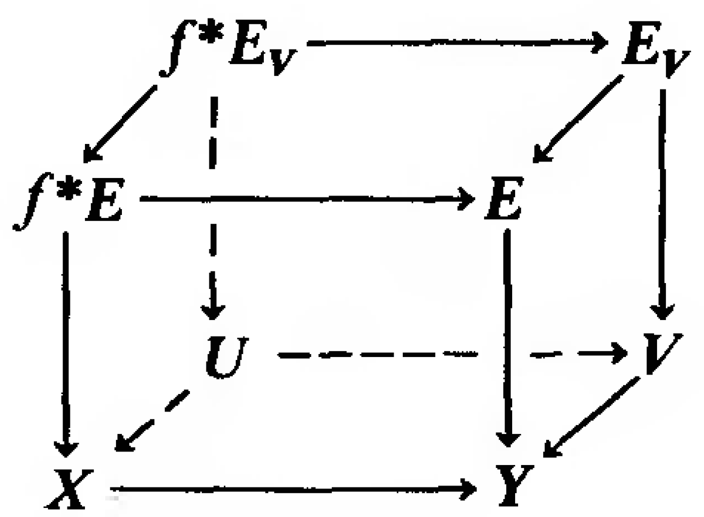
the top horizontal map being the identity on each fiber.

PB 3. *If E is trivial, equal to $Y \times \mathbf{E}$, then $f^*E = X \times \mathbf{E}$ and $f^*\pi$ is the projection.*

PB 4. *If V is an open subset of Y and $U = f^{-1}(V)$, then*

$$f^*(E_V) = (f^*E)_U,$$

and we have a commutative diagram:



III, §2. THE TANGENT BUNDLE

Let X be a manifold of class C^p with $p \geqq 1$. We shall define a functor T from the category of such manifolds into the category of vector bundles of class C^{p-1}.

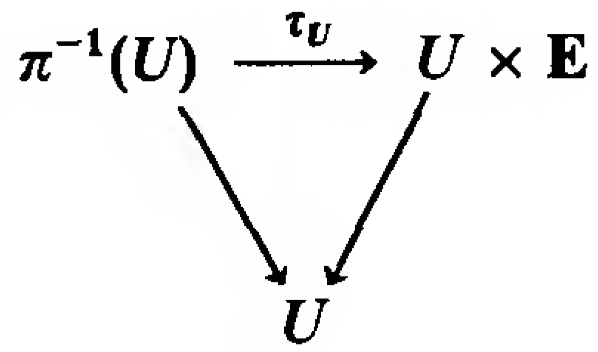
For each manifold X we let $T(X)$ be the disjoint union of the tangent spaces $T_x(X)$. We have a natural projection

$$\pi\colon T(X) \to X$$

mapping $T_x(X)$ on x. We must make this into a vector bundle. If (U, φ) is a chart of X such that φU is open in the Banach space $\mathbf{E}$, then from the definition of the tangent vectors as equivalence classes of triples (U, φ, v) we get immediately a bijection

$$\tau_U\colon \pi^{-1}(U) = T(U) \to U \times \mathbf{E}$$

which commutes with the projection on U, that is such that



is commutative. Furthermore, if (U_i, φ_i) and (U_j, φ_j) are two charts, and if we denote by φ_{ji} the map $\varphi_j\varphi_i^{-1}$ (defined on $\varphi_i(U_i \cap U_j)$), then we obtain a transition mapping

$$\tau_{ji} = (\tau_j\tau_i^{-1})\colon \varphi_i(U_i \cap U_j) \times \mathbf{E} \to \varphi_j(U_i \cap U_j) \times \mathbf{E}$$

by the formula

$$\tau_{ji}(x, v) = (\varphi_{ji}x, D\varphi_{ji}(x) \cdot v)$$

for $x \in U_i \cap U_j$ and $v \in \mathbf{E}$. Since the derivative $D\varphi_{ji} = \varphi'_{ji}$ is of class C^{p-1} and is an isomorphism at x, we see immediately that all the conditions of Proposition 1.2 are verified (using Proposition 3.10 of Chapter I), thereby making $T(X)$ into a vector bundle of class C^{p-1}.

We see that the above construction can also be expressed as follows. If the manifold X is glued together from open sets $\{U_i\}$ in Banach spaces by means of transition mappings $\{\varphi_{ij}\}$, then we can glue together products $U_i \times \mathbf{E}$ by means of transition mappings $(\varphi_{ij}, D\varphi_{ij})$ where the derivative $D\varphi_{ij}$ can be viewed as a function of two variables (x, v). Thus locally, for open subsets U of Banach spaces, the tangent bundle can be identified with the product $U \times \mathbf{E}$. The reader will note that our definition coincides with the oldest definition employed by geometers, our tangent vectors being vectors which transform according to a certain rule (namely the derivative).

If $f\colon X \to X'$ is a C^p-morphism, we can define

$$Tf\colon T(X) \to T(X')$$

to be simply $T_x f$ on each fiber $T_x(X)$. In order to verify that Tf is a VB-morphism (of class C^{p-1}), it suffices to look at the situation locally, i.e. we may assume that X and X' are open in vector spaces $\mathbf{E}$, $\mathbf{E}'$, and that $T_x f = f'(x)$ is simply the derivative. Then the map Tf is given by

$$Tf(x, v) = \big(f(x), f'(x)v\big)$$

for $x \in X$ and $v \in \mathbf{E}$. Since f' is of class C^{p-1} by definition, we can apply Proposition 3.10 of Chapter I to conclude that Tf is also of class C^{p-1}. The functoriality property is trivially satisfied, and we have therefore defined the functor T as promised.

It will sometimes be notationally convenient to write f_* instead of Tf for the induced map, which is also called the **tangent** map. The bundle $T(X)$ is called the **tangent bundle** of X.

Remark. The above definition of the tangent bundle fits with Steenrod's point of view [Ste 51]. I don't understand why many differential geometers have systematically rejected this point of view, when they take the definition of a tangent vector as a differential operator.

III, §3. EXACT SEQUENCES OF BUNDLES

Let X be a manifold. Let $\pi'\colon E' \to X$ and $\pi\colon E \to X$ be two vector bundles over X. Let $f\colon \pi' \to \pi$ be a VB-morphism. We shall say that the sequence

$$0 \to \pi' \xrightarrow{f} \pi$$

is **exact** if there exists a covering of X by open sets and for each open set U in this covering there exist trivializations

$$\tau'\colon E'_U \to U \times \mathbf{E}' \qquad \text{and} \qquad \tau\colon E_U \to U \times \mathbf{E}$$

such that $\mathbf{E}$ can be written as a product $\mathbf{E} = \mathbf{E}' \times \mathbf{F}$, making the following

diagram commutative:

$$\begin{array}{ccc} E'_U & \xrightarrow{\ f\ } & E_U \\ {\scriptstyle \tau'}\downarrow & & \downarrow{\scriptstyle \tau} \\ U \times \mathbf{E}' & \longrightarrow & U \times \mathbf{E}' \times \mathbf{F} \end{array}$$

(The bottom map is the natural one: Identity on U and the injection of $\mathbf{E}'$ on $\mathbf{E}' \times 0$.)

Let $\pi_1\colon E_1 \to X$ be another vector bundle, and let $g\colon \pi_1 \to \pi$ be a VB-morphism such that $g(E_1)$ is contained in $f(E')$. Since f establishes a bijection between E' and its image $f(E')$ in E, it follows that there exists a unique map $g_1\colon E_1 \to E'$ such that $g = f \circ g_1$. We contend that g_1 is a VB-morphism. Indeed, to prove this we can work locally, and in view of the definition, over an open set U as above, we can write

$$g_1 = \tau'^{-1} \circ \mathrm{pr} \circ \tau \circ g$$

where pr is the projection of $U \times \mathbf{E}' \times \mathbf{F}$ on $U \times \mathbf{E}'$. All the maps on the right-hand side of our equality are VB-morphisms; this proves our contention.

Let $\pi\colon E \to X$ be a vector bundle. A subset S of E will be called a **subbundle** if there exists an exact sequence $0 \to \pi' \to \pi$, also written

$$0 \to E' \xrightarrow{f} E,$$

such that $f(E') = S$. This gives S the structure of a vector bundle, and the previous remarks show that it is unique. In fact, given another exact sequence

$$0 \to E_1 \xrightarrow{g} E$$

such that $g(E_1) = S$, the natural map $f^{-1}g$ from E_1 to E' is a VB-isomorphism.

Let us denote by E/E' the union of all factor spaces E_x/E'_x. If we are dealing with an exact sequence as above, then we can give E/E' the structure of a vector bundle. We proceed as follows. Let $\{U_i\}$ be our covering, with trivialising maps τ'_i and τ_i. We can define for each i a bijection

$$\tau''_i\colon E_{U_i}/E'_{U_i} \to U_i \times \mathbf{F}$$

obtained in a natural way from the above commutative diagram. (Without loss of generality, we can assume that the vector spaces $\mathbf{E}'$, $\mathbf{F}$ are constant for all i.) We have to prove that these bijections satisfy the conditions of Proposition 1.2.

Without loss of generality, we may assume that f is an inclusion (of the total space E' into E). For each pair i, j and $x \in U_i \cap U_j$, the toplinear automorphism $(\tau_j \tau_i^{-1})_x$ is represented by a matrix

$$\begin{pmatrix} h_{11}(x) & h_{12}(x) \\ h_{21}(x) & h_{22}(x) \end{pmatrix}$$

operating on the right on a vector $(v, w) \in \mathbf{E}' \times \mathbf{F}$. The map $(\tau_j'' \tau_i''^{-1})_x$ on $\mathbf{F}$ is induced by this matrix. Since $\mathbf{E}' = \mathbf{E}' \times 0$ has to be carried into itself by the matrix, we have $h_{12}(x) = 0$. Furthermore, since $(\tau_j \tau_i^{-1})_x$ has an inverse, equal to $(\tau_i \tau_j^{-1})_x$, it follows that $h_{22}(x)$ is a toplinear automorphism of $\mathbf{F}$, and represents $(\tau_j'' \tau_i''^{-1})_x$. Therefore condition **VB 3** is satisfied, and E/E' is a vector bundle.

The canonical map

$$E_U \to E_U/E'_U$$

is a morphism since it can be expressed in terms of τ, the projection, and τ''^{-1}. Consequently, we obtain a VB-morphism

$$g\colon \pi \to \pi''$$

in the canonical way (on the total spaces, it is the quotient mapping of E on E/E'). We shall call π'' the **factor bundle**.

Our map g satisfies the usual universal mapping property of a cokernel. Indeed, suppose that

$$\psi\colon E \to G$$

is a VB-morphism such that $\psi \circ f = 0$ (i.e. $\psi_x \circ f_x = 0$ on each fiber E'_x). We can then define set theoretically a canonical map

$$\psi_*\colon E/E' \to G,$$

and we must prove that it is a VB-morphism. This can be done locally. Using the above notation, we may assume that $E = U \times \mathbf{E}' \times \mathbf{F}$ and that g is the projection. In that case, ψ_* is simply the canonical injection of $U \times \mathbf{F}$ in $U \times \mathbf{E}' \times \mathbf{F}$ followed by ψ, and is therefore a VB-morphism.

We shall therefore call g the **cokernel** of f.

Dually, let $g\colon \pi \to \pi''$ be a given VB-morphism. We shall say that the sequence

$$\pi \xrightarrow{g} \pi'' \to 0$$

is **exact** if g is surjective, and if there exists a covering of X by open sets, and for each open set U in this covering there exist spaces $\mathbf{E}'$, $\mathbf{F}$ and trivializations

$$\tau\colon E_U \to U \times \mathbf{E}' \times \mathbf{F} \qquad \text{and} \qquad \tau''\colon \mathbf{E}''_U \to \mathbf{F}$$

making the following diagram commutative:

$$\begin{array}{ccc} E_U & \xrightarrow{g} & E''_U \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau''} \\ U \times \mathbf{E}' \times \mathbf{F} & \longrightarrow & U \times \mathbf{F} \end{array}$$

(The bottom map is the natural one: Identity on U and the projection of $\mathbf{E}' \times \mathbf{F}$ on $\mathbf{F}$.)

In the same way as before, one sees that the "kernel" of g, that is, the union of the kernels E'_x of each g_x, can be given a structure of vector bundle. This union E' will be called the **kernel** of g, and satisfies the usual universal mapping property.

Proposition 3.1. *Let X be a manifold and let*

$$f\colon \pi' \to \pi$$

be a VB-morphism of vector bundles over X. Assume that, for each $x \in X$, the continuous linear map

$$f_x\colon E'_x \to E_x$$

is injective and splits. Then the sequence

$$0 \to \pi' \xrightarrow{f} \pi$$

is exact.

Proof. We can assume that X is connected and that the fibers of E' and E are constant, say equal to the Banach spaces $\mathbf{E}'$ and $\mathbf{E}$. Let $a \in X$. Corresponding to the splitting of f_a we know that we have a product decomposition $\mathbf{E} = \mathbf{E}' \times \mathbf{F}$ and that there exists an open set U of X containing a, together with trivializing maps

$$\tau\colon \pi^{-1}(U) \to U \times \mathbf{E} \qquad \text{and} \qquad \tau'\colon \pi'^{-1}(U) \to U \times \mathbf{E}'$$

such that the composite map

$$\mathbf{E}' \xrightarrow{\tau_a'^{-1}} E'_a \xrightarrow{f_a} E_a \xrightarrow{\tau_a} \mathbf{E}' \times \mathbf{F}$$

maps $\mathbf{E}'$ on $\mathbf{E}' \times 0$.

For any point x in U, we have a map

$$(\tau f \tau'^{-1})_x\colon \mathbf{E}' \to \mathbf{E}' \times \mathbf{F},$$

which can be represented by a pair of continuous linear maps

$$(h_{11}(x), h_{21}(x)).$$

We define

$$h(x)\colon \mathbf{E}' \times \mathbf{F} \to \mathbf{E}' \times \mathbf{F}$$

by the matrix

$$\begin{pmatrix} h_{11}(x) & 0 \\ h_{21}(x) & \mathrm{id} \end{pmatrix},$$

operating on the right on a vector $(v, w) \in \mathbf{E}' \times \mathbf{F}$. Then $h(x)$ restricted to $\mathbf{E}' \times 0$ has the same action as $(\tau f \tau'^{-1})_x$.

The map $x \mapsto h(x)$ is a morphism of U into $L(\mathbf{E}, \mathbf{E})$ and since it is continuous, it follows that for U small enough around our fixed point a, it maps U into the group of toplinear automorphisms of $\mathbf{E}$. This proves our proposition.

Dually to Proposition 3.1, we have:

Proposition 3.2. *Let X be a manifold and let*

$$g\colon \pi \to \pi''$$

be a VB-*morphism of vector bundles over X. Assume that for each $x \in X$, the continuous linear map*

$$g_x\colon E_x \to E''_x$$

is surjective and has a kernel that splits. Then the sequence

$$\pi \xrightarrow{g} \pi'' \to 0$$

is exact.

Proof. It is dual to the preceding one and we leave it to the reader.

In general, a sequence of VB-morphisms

$$0 \to \pi' \xrightarrow{f} \pi \xrightarrow{g} \pi'' \to 0$$

is said to be **exact** if both ends are exact, and if the image of f is equal to the kernel of g.

There is an important example of exact sequence. Let $f\colon X \to Y$ be an immersion. By the universal mapping property of pull backs, we have a canonical VB-morphism

$$T^*f\colon T(X) \to f^*T(Y)$$

of $T(X)$ into the pull back over X of the tangent bundle of Y. Furthermore, from the manner in which the pull back is obtained locally by taking products, and the definition of an immersion, one sees that the sequence

$$0 \to T(X) \xrightarrow{T^*f} f^*T(Y)$$

is exact. The factor bundle

$$f^*T(Y)/\mathrm{Im}(T^*f)$$

is called the **normal** bundle of f. It is denoted by $N(f)$, and its total space by $N_f(X)$ if we wish to distinguish between the two. We sometimes identify $T(X)$ with its image under T^*f and write

$$N(f) = f^*T(Y)/T(X).$$

Dually, let $f\colon X \to Y$ be a submersion. Then we have an exact sequence

$$T(X) \xrightarrow{T^*f} f^*T(Y) \to 0$$

whose kernel could be called the **subbundle** of f, or the **bundle along the fiber.**

There is an interesting case where we can describe the kernel more precisely. Let

$$\pi\colon E \to X$$

be a vector bundle. Then we can form the pull back of E over itself, that is, π^*E, and we contend that we have an exact sequence

$$0 \to \pi^*E \to T(E) \to \pi^*T(X) \to 0.$$

To define the map on the left, we look at the subbundle of π more closely. For each $x \in X$ we have an inclusion

$$E_x \to E,$$

whence a natural injection

$$T(E_x) \to T(E).$$

The local product structure of a bundle shows that the union of the $T(E_x)$ as x ranges over X gives the subbundle set theoretically. On the other hand, the total space of π^*E consists of pairs of vectors (v, w) lying over the same base point x, that is, the fiber at x of π^*E is simply $E_x \times E_x$. Since $T(E_x)$ has a natural identification with $E_x \times E_x$, we get for each x a bijection

$$(\pi^*E)_x \to T(E_x)$$

which defines our map from π^*E to $T(E)$. Considering the map locally in terms of the local product structure shows at once that it gives a VB-isomorphism between π^*E and the subbundle of π, as desired.

III, §4. OPERATIONS ON VECTOR BUNDLES

We consider subcategories of Banach spaces $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{C}$ and let

$$\lambda\colon \mathfrak{A} \times \mathfrak{B} \to \mathfrak{C}$$

be a functor in, say, two variables, which is, say, contravariant in the first and covariant in the second. (Everything we shall do extends in the obvious manner to functors of several variables, letting $\mathfrak{A}$, $\mathfrak{B}$ stand for n-tuples.)

Example. We took a functor in two variables for definiteness, and to illustrate both variances. However, we could consider a functor in one or more than two variables. For instance, let us consider the functor

$$\mathbf{E} \mapsto L(\mathbf{E}, \mathbf{R}) = L(\mathbf{E}) = \mathbf{E}^\vee,$$

which we call the **dual**. It is a contravariant functor in one variable. On the other hand, the functor

$$\mathbf{E} \mapsto L_a^r(\mathbf{E}, \mathbf{F})$$

of continuous multilinear maps of $\mathbf{E} \times \cdots \times \mathbf{E}$ into a Banach space $\mathbf{F}$ is contravariant in $\mathbf{E}$ and covariant in $\mathbf{F}$. The functor $\mathbf{E} \mapsto L_a^r(\mathbf{E}, \mathbf{R})$ gives rise later to what we call differential forms. We shall treat such forms systematically in Chapter V, §3.

If $f\colon \mathbf{E}' \to \mathbf{E}$ and $g\colon \mathbf{F} \to \mathbf{F}'$ are two continuous linear maps, with f a morphism of $\mathfrak{A}$ and g a morphism of $\mathfrak{B}$, then by definition, we have a map

$$L(\mathbf{E}', \mathbf{E}) \times L(\mathbf{F}, \mathbf{F}') \to L(\lambda(\mathbf{E}, \mathbf{F}), \lambda(\mathbf{E}', \mathbf{F}')),$$

assigning $\lambda(f, g)$ to (f, g).

We shall say that λ is of **class** C^p if the following condition is satisfied. Give a manifold U, and two morphisms

$$\varphi\colon U \to L(\mathbf{E}', \mathbf{E}) \quad \text{and} \quad \psi\colon U \to L(\mathbf{F}, \mathbf{F}'),$$

then the composite

$$U \to L(\mathbf{E}', \mathbf{E}) \times L(\mathbf{F}, \mathbf{F}') \to L(\lambda(\mathbf{E}, \mathbf{F}), \lambda(\mathbf{E}', \mathbf{F}'))$$

is also a morphism. (One could also say that λ is **differentiable**.)

Theorem 4.1. *Let λ be a functor as above, of class C^p, $p \geqq 0$. Then for each manifold X, there exists a functor λ_X, on vector bundles (of class C^p)*

$$\lambda_X\colon \mathrm{VB}(X, \mathfrak{A}) \times \mathrm{VB}(X, \mathfrak{B}) \to \mathrm{VB}(X, \mathfrak{C})$$

satisfying the following properties. For any bundles α, β in VB$(X, \mathfrak{A})$ *and* VB$(X, \mathfrak{B})$ *respectively, and* VB*-morphisms*

$$f\colon \alpha' \to \alpha \quad \textit{and} \quad g\colon \beta \to \beta'$$

in the respective categories, and for each $x \in X$, we have:

OP 1. $\lambda_X(\alpha, \beta)_x = \lambda(\alpha_x, \beta_x)$.

OP 2. $\lambda_X(f, g)_x = \lambda(f_x, g_x)$.

OP 3. *If α is the trivial bundle $X \times \mathbf{E}$ and β the trivial bundle $X \times \mathbf{F}$, then $\lambda_X(\alpha, \beta)$ is the trivial bundle $X \times \lambda(\mathbf{E}, \mathbf{F})$.*

OP 4. *If $h\colon Y \to X$ is a C^p-morphism, then*

$$\lambda_Y^*(h^*\alpha, h^*\beta) = h^*\lambda_X(\alpha, \beta).$$

Proof. We may assume that X is connected, so that all the fibers are toplinearly isomorphic to a fixed space. For each open subset U of X we let the total space $\lambda_U(E_\alpha, E_\beta)$ of $\lambda_U(\alpha, \beta)$ be the union of the sets

$$\{x\} \times \lambda(\alpha_x, \beta_x)$$

(identified harmlessly throughout with $\lambda(\alpha_x, \beta_x)$), as x ranges over U. We can find a covering $\{U_i\}$ of X with trivializing maps $\{\tau_i\}$ for α, and $\{\sigma_i\}$ for β,

$$\tau_i\colon \alpha^{-1}(U_i) \to U_i \times \mathbf{E},$$

$$\sigma_i\colon \beta^{-1}(U_i) \to U_i \times \mathbf{F}.$$

We have a bijection

$$\lambda(\tau_i^{-1}, \sigma_i)\colon\ \lambda_{U_i}(E_\alpha, E_\beta) \to U_i \times \lambda(\mathbf{E}, \mathbf{F})$$

obtained by taking on each fiber the map

$$\lambda(\tau_{ix}^{-1}, \sigma_{ix})\colon\ \lambda(\alpha_x, \beta_x) \to \lambda(\mathbf{E}, \mathbf{F}).$$

We must verify that **VB 3** is satisfied. This means looking at the map

$$x \to \lambda(\tau_{jx}^{-1}, \sigma_{jx}) \circ \lambda(\tau_{ix}^{-1}, \sigma_{ix})^{-1}.$$

The expression on the right is equal to

$$\lambda(\tau_{ix}\tau_{jx}^{-1}, \sigma_{jx}\sigma_{ix}^{-1}).$$

Since λ is a functor of class $\mathbf{C}^p$, we see that we get a map

$$U_i \cap U_j \to L\big(\lambda(\mathbf{E}, \mathbf{F}), \lambda(\mathbf{E}, \mathbf{F})\big)$$

which is a C^p-morphism. Furthermore, since λ is a functor, the transition mappings are in fact toplinear isomorphism, and **VB 2**, **VB 3** are proved.

The proof of the analogous statement for $\lambda_X(f, g)$, to the effect that it is a VB-morphism, proceeds in an analogous way, again using the hypothesis that λ is of class C^p. Condition **OP 3** is obviously satisfied, and **OP 4** follows by localizing. This proves our theorem.

The next theorem gives us the uniqueness of the operation λ_X.

Theorem 4.2. *If μ is another functor of class C^p with the same variance as λ, and if we have a natural transformation of functors $t\colon\ \lambda \to \mu$, then for each X, the mapping*

$$t_X\colon\ \lambda_X \to \mu_X,$$

defined on each fiber by the map

$$t(\alpha_x, \beta_x)\colon\ \lambda(\alpha_x, \beta_x) \to \mu(\alpha_x, \beta_x),$$

is a natural transformation of functors (in the VB-category).

Proof. For simplicity of notation, assume that λ and μ are both functors of one variable, and both covariant. For each open set $U = U_i$ of a trivializing covering for β, we have a commutative diagram:

$$\begin{array}{ccc} U \times \lambda(\mathbf{E}) & \xrightarrow{\ \mathrm{id} \times t(\mathbf{E})\ } & U \times \mu(\mathbf{E}) \\ {\scriptstyle \lambda_U(\sigma)}\uparrow & & \uparrow{\scriptstyle \mu_U(\sigma)} \\ \lambda_U(\beta) & \xrightarrow[\ t_U\ ]{} & \mu_U(\beta) \end{array}$$

The vertical maps are trivializing VB-isomorphism, and the top horizontal map is a VB-morphism. Hence t_U is a VB-morphism, and our assertion is proved.

In particular, for $\lambda = \mu$ and $t = \mathrm{id}$ we get the uniqueness of our functor λ_X.

(In the proof of Theorem 4.2, we do not use again explicitly the hypotheses that λ, μ are differentiable.)

In practice, we omit the subscript X on λ, and write λ for the functor on vector bundles.

Examples. Let $\pi\colon E \to X$ be a vector bundle. We take λ to be the dual, that is $\mathbf{E} \mapsto \mathbf{E}^\vee = L(\mathbf{E}, \mathbf{R})$. Then $\lambda(E)$ is denoted by $E^\vee$, and is called the **dual bundle**. The fiber at each point $x \in X$ is the dual space $E_x^\vee$. The dual bundle of the tangent bundle is called the **cotangent bundle** $T^\vee X$.

Similarly, instead of taking $L(E)$, we could take $L_a^r(E)$ to be the bundle of alternating multilinear forms on E. The fiber at each point is the space $L_a^r(E_x)$ consisting of all r-multilinear alternating continuous functions on E_x. When $E = TX$ is the tangent bundle, the sections of $L_a^r(TX)$ are called **differential forms** of degree r. Thus a 1-form is a section of $E^\vee$. Differential forms will be treated later in detail.

Recall that $\mathrm{End}(\mathbf{E}) = L(\mathbf{E}, \mathbf{E})$. In the theory of curvature, we shall deal with both functors

$$E \mapsto L^4(\mathbf{E}) = L^4(\mathbf{E}, \mathbf{R}) \quad \text{and} \quad \mathbf{E}^3 \mapsto L^2\big(\mathbf{E}, \mathrm{End}(\mathbf{E})\big) = L^2\big(\mathbf{E}, L(\mathbf{E}, \mathbf{E})\big).$$

In fact, if $R \in L^2\big(\mathbf{E}, L(\mathbf{E}, \mathbf{E})\big)$, then for each pair of elements $v, w \in \mathbf{E}$ and $z \in \mathbf{E}$, we see that $R(v, w) \in L(\mathbf{E}, \mathbf{E})$ and $R(v, w)z \in \mathbf{E}$, so we get a 3-linear map

$$(v, w, z) \mapsto R(v, w)z.$$

We shall apply both functors to the tangent bundle in Chapter IX.

For another type of operation, we have the **direct sum** (also called the **Whitney sum**) of two bundles α, β over X. It is denoted by $\alpha \oplus \beta$, and the fiber at a point x is

$$(\alpha \oplus \beta)_x = \alpha_x \oplus \beta_x.$$

Of course, the finite direct sum of vector spaces can be identified with their finite direct products, but we write the above operation as a direct sum in order not to confuse it with the following direct product.

Let $\alpha\colon E_\alpha \to X$ and $\beta\colon E_\beta \to Y$ be two vector bundles in VB(X) and VB(Y) respectively. Then the map

$$\alpha \times \beta\colon E_\alpha \times E_\beta \to X \times Y$$

is a vector bundle, and it is this operation which we call the **direct product** of α and β.

Let X be a manifold, and λ a functor of class C^p with $p \geqq 1$. The **tensor bundle** of type λ over X is defined to be $\lambda_X(T(X))$, also denoted by $\lambda T(X)$ or $T_\lambda(X)$. The sections of this bundle are called **tensor fields** of type λ, and the set of such sections is denoted by $\Gamma_\lambda(X)$. Suppose that we have a trivialization of $T(X)$, say

$$T(X) = X \times \mathbf{E}.$$

Then $T_\lambda(X) = X \times \lambda(\mathbf{E})$. A section of $T_\lambda(X)$ in this representation is completely described by the projection on the second factor, which is a morphism

$$f\colon X \to \lambda(\mathbf{E}).$$

We shall call it the **local representation** of the tensor field (in the given trivialization). If ξ is the tensor field having f as its local representation, then

$$\xi(x) = (x, f(x)).$$

Let $f\colon X \to Y$ be a morphism of class C^p ($p \geqq 1$). Let ω be a tensor field of type L^r over Y, which could also be called a **multilinear tensor field**. For each $y \in Y$, $\omega(y)$ (also written ω_y) is a continuous multilinear function on $T_y(Y)$:

$$\omega_y\colon T_y \times \cdots \times T_y \to \mathbf{R}.$$

For each $x \in X$, we can define a continuous multilinear map

$$f_x^*(\omega)\colon T_x \times \cdots \times T_x \to \mathbf{R}$$

by the composition of maps $(T_x f)^r$ and $\omega_{f(x)}$:

$$T_x \times \cdots \times T_x \to T_{f(x)} \times \cdots \times T_{f(x)} \to \mathbf{R}.$$

We contend that the map $x \mapsto f_x^*(\omega)$ is a tensor field over X, of the same type as ω. To prove this, we may work with local representation. Thus we can assume that we work with a morphism

$$f\colon U \to V$$

of one open set in a Banach space into another, and that

$$\omega\colon V \to L^r(\mathbf{F})$$

is a morphism, V being open in $\mathbf{F}$. If U is open in $\mathbf{E}$, then $f^*(\omega)$ (now denoting a local representation) becomes a mapping of U into $L^r(\mathbf{E})$, given by the formula

$$f_x^*(\omega) = L^r(f'(x)) \cdot \omega(f(x)).$$

Since $L^r\colon L(\mathbf{E}, \mathbf{F}) \to L(L^r(\mathbf{F}), L^r(\mathbf{E}))$ is of class C^∞, it follows that $f^*(\omega)$ is a morphism of the same class as ω. This proves what we want.

Of course, the same argument is valid for the other functors L_s^r and L_a^r (symmetric and alternating continuous multilinear maps). Special cases will be considered in later chapters. If λ denotes any one of our three functors, then we see that we have obtained a mapping (which is in fact linear)

$$f^*\colon \Gamma_\lambda(Y) \to \Gamma_\lambda(X)$$

which is clearly functorial in f. We use the notation f^* instead of the more correct (but clumsy) notation f_λ or $\Gamma_\lambda(f)$. No confusion will arise from this.

III, §5. SPLITTING OF VECTOR BUNDLES

The next proposition expresses the fact that the VB-morphisms of one bundle into another (over a fixed morhism) form a module over the ring of functions.

Proposition 5.1. *Let X, Y be manifolds and $f_0\colon X \to Y$ a morphism. Let α, β be vector bundles over X, Y respectively, and let $f, g\colon \alpha \to \beta$ be two* VB*-morphisms over f_0. Then the map $f + g$ defined by the formula*

$$(f + g)_x = f_x + g_x$$

is also a VB*-morphism. Furthermore, if $\psi\colon Y \to \mathbf{R}$ is a function on Y, then the map ψf defined by*

$$(\psi f)_x = \psi(f_0(x)) f_x$$

is also a VB*-morphism.*

Proof. Both assertions are immediate consequences of Proposition 3.10 of Chapter I.

We shall consider mostly the situation where $X = Y$ and f_0 is the identity, and will use it, together with partitions of unity, to glue VB-morphisms together.

Let α, β be vector bundles over X and let $\{(U_i, \psi_i)\}$ be a partition of unity on X. Suppose given for each U_i a VB-morphism

$$f_i\colon \alpha|U_i \to \beta|U_i.$$

Each one of the maps $\psi_i f_i$ (defined as in Proposition 5.1) is a VB-morphism. Furthermore, we can extend $\psi_i f_i$ to a VB-morphism of α into β simply by putting

$$(\psi_i f_i)_x = 0$$

for all $x \notin U_i$. If we now define

$$f\colon \alpha \to \beta$$

by the formula

$$f_x(v) = \sum \psi_i(x) f_{ix}(v)$$

for all pairs (x, v) with $v \in \alpha_x$, then the sum is actually finite, at each ponit x, and again by Proposition 5.1, we see that f is a VB-morphism. We observe that if each f_i is the identity, then $f = \sum \psi_i f_i$ is also the identity.

Proposition 5.2. *Let X be a manifold admitting partitions of unity. Let $0 \to \alpha \xrightarrow{f} \beta$ be an exact sequence of vector bundles over X. Then there exists a surjective* VB-*morphism $g\colon \beta \to \alpha$ whose kernel splits at each point, such that $g \circ f = \mathrm{id}$.*

Proof. By the definition of exact sequence, there exists a partition of unity $\{(U_i, \psi_i)\}$ on X such that for each i, we can split the sequence over U_i. In other words, there exists for each i a VB-morphism

$$g_i\colon \beta|U_i \to \alpha|U_i$$

which is surjective, whose kernel splits, and such that $g_i \circ f_i = \mathrm{id}_i$. We let $g = \sum \psi_i g_i$. Then g is a VB-morphism of β into α by what we have just seen, and

$$g \circ f = \sum \psi_i g_i f_i = \mathrm{id}.$$

It is trivial that g is surjective because $g \circ f = \mathrm{id}$. The kernel of g_x splits at each point x because it has a closed complement, namely $f_x \alpha_x$. This concludes the proof.

If γ is the kernel of β, then we have $\beta \approx \alpha \oplus \gamma$.

A vector bundle π over X will be said to be of **finite type** if there exists a finite trivialization for π (i.e. a trivialization $\{(U_i, \tau_i)\}$ such that i ranges over a finite set).

If k is an integer $\geqq 1$ and $\mathbf{E}$ a topological vector space, then we denote by $\mathbf{E}^k$ the direct product of $\mathbf{E}$ with itself k times.

Proposition 5.3. *Let X be a manifold admitting partitions of unity. Let π be a vector bundle of finite type in* VB$(X, \mathbf{E})$, *where $\mathbf{E}$ is a Banach space. Then there exists an integer $k > 0$ and a vector bundle α in* VB$(X, \mathbf{E}^k)$ *such that $\pi \oplus \alpha$ is trivializable.*

Proof. We shall prove that there exists an exact sequence

$$0 \to \pi \xrightarrow{f} \beta$$

with $E_\beta = X \times \mathbf{E}^k$. Our theorem will follow from the preceding proposition.

Let $\{U_i, \tau_i)\}$ be a finite trivialization of π with $i = 1, \ldots, k$. Let $\{(U_i, \psi_i)\}$ be a partition of unity. We define

$$f\colon E_\pi \to X \times \mathbf{E}^k$$

as follows. If $x \in X$ and v is in the fiber of E_π at x, then

$$f_x(v) = \big(x, \psi_1(x)\tau_1(v), \ldots, \psi_k(x)\tau_k(v)\big).$$

The expression on the right makes sense, because in case x does not lie in U_i then $\psi_i(x) = 0$ and we do not have to worry about the expression $\tau_i(v)$. If x lies in U_i, then $\tau_i(v)$ means $\tau_{ix}(v)$.

Given any point x, there exists some index i such that $\psi_i(x) > 0$ and hence f is injective. Furthermore, for this x and this index i, f_x maps E_x onto a closed subspace of $\mathbf{E}^k$, which admits a closed complement, namely

$$\mathbf{E} \times \cdots \times 0 \times \cdots \times \mathbf{E}$$

with 0 in the i-th place. This proves our proposition.

CHAPTER IV

Vector Fields and Differential Equations

In this chapter, we collect a number of results all of which make use of the notion of differential equation and solutions of differential equations.

Let X be a manifold. A vector field on X assigns to each point x of X a tangent vector, differentiably. (For the precise definition, see §2.) Given x_0 in X, it is then possible to construct a unique curve $\alpha(t)$ starting at x_0 (i.e. such that $\alpha(0) = x_0$) whose derivative at each point is the given vector. It is not always possible to make the curve depend on time t from $-\infty$ to $+\infty$, although it is possible if X is compact.

The structure of these curves presents a fruitful domain of investigation, from a number of points of view. For instance, one may ask for topological properties of the curves, that is those which are invariant under topological automorphisms of the manifold. (Is the curve a closed curve, is it a spiral, is it dense, etc.?) More generally, following standard procedures, one may ask for properties which are invariant under any given interesting group of automorphisms of X (discrete groups, Lie groups, algebraic groups, Riemannian automorphisms, ad lib.).

We do not go into these theories, each of which proceeds according to its own flavor. We give merely the elementary facts and definitions associated with vector fields, and some simple applications of the existence theorem for their curves.

Throughout this chapter, we assume all manifolds to be Hausdorff, of class C^p with $p \geqq 2$ from §2 on, and $p \geqq 3$ from §3 on. This latter condition insures that the tangent bundle is of class C^{p-1} with $p - 1 \geqq 1$ (or 2).

We shall deal with mappings of several variables, say $f(t, x, y)$, the first of which will be a real variable. We identify $D_1 f(t, x, y)$ with

$$\lim_{h \to 0} \frac{f(t+h, x, y) - f(t, x, y)}{h}.$$

IV, §1. EXISTENCE THEOREM FOR DIFFERENTIAL EQUATIONS

Let $\mathbf{E}$ be a Banach space and U an open subset of $\mathbf{E}$. In this section we consider vector fields locally. The notion will be globalized later, and thus for the moment, we define (the local representation of) a **time-dependent vector field** on U to be a C^p-morphism ($p \geqq 0$)

$$f\colon J \times U \to \mathbf{E},$$

where J is an open interval containing 0 in $\mathbf{R}$. We think of f as assigning to each point x in U a vector $f(t, x)$ in $\mathbf{E}$, depending on time t.

Let x_0 be a point of U. An **integral curve** for f with **initial condition** x_0 is a mapping of class C^r ($r \geqq 1$)

$$\alpha\colon J_0 \to U$$

of an open subinterval of J containing 0, into U, such that $\alpha(0) = x_0$ and such that

$$\alpha'(t) = f(t, \alpha(t)).$$

Remark. Let $\alpha\colon J_0 \to U$ be a continuous map satisfying the condition

$$\alpha(t) = x_0 + \int_0^t f(u, \alpha(u))\, du.$$

Then α is differentiable, and its derivative is $f(t, \alpha(t))$. Hence α is of class C^1. Furthermore, we can argue recursively, and conclude that if f is of class C^p, then so is α. Conversely, if α is an integral curve for f with initial condition x_0, then it obviously satisfies out integral relation.

Let

$$f\colon J \times U \to \mathbf{E}$$

be as above, and let x_0 be a point of U. By a **local flow** for f at x_0 we mean a mapping

$$\alpha\colon J_0 \times U_0 \to U$$

where J_0 is an open subinterval of J containing 0, and U_0 is an open subset of U containing x_0, such that for each x in U_0 the map

$$\alpha_x(t) = \alpha(t, x)$$

is an integral curve for f with initial condition x (i.e. such that $\alpha(0, x) = x$).

As a matter of notation, when we have a mapping with two arguments, say $\varphi(t, x)$, then we denote the separate mappings in each argument when the other is kept fixed by $\varphi_x(t)$ and $\varphi_t(x)$. The choice of letters will always prevent ambiguity.

We shall say that f satisfies a **Lipschitz condition** on U **uniformly with respect to** J if there exists a number $K > 0$ such that

$$|f(t, x) - f(t, y)| \leqq K|x - y|$$

for all x, y in U and t in J. We call K a **Lipschitz constant**. If f is of class C^1, it follows at once from the mean value theorem that f is Lipschitz on some open neighborhood $J_0 \times U_0$ of a given point $(0, x_0)$ of U, and that it is bounded on some such neighborhood.

We shall now prove that under a Lipschitz condition, local flows exist and are unique locally. In fact, we prove more, giving a uniformity property for such flows. If b is real > 0, then we denote by J_b the open interval $-b < t < b$.

Proposition 1.1. *Let J be an open interval of* $\mathbf{R}$ *containing* 0, *and U open in the Banach space* $\mathbf{E}$. *Let x_0 be a point of U, and $a > 0$, $a < 1$ a real number such that the closed ball $\bar{B}_{3a}(x_0)$ lies in U. Assume that we have a continuous map*

$$f\colon J \times U \to \mathbf{E}$$

which is bounded by a constant $L \geqq 1$ on $J \times U$, and satisfies a Lipschitz condition on U uniformly with respect to J, with constant $K \geqq 1$. If $b < a/LK$, then for each x in $\bar{B}_a(x_0)$ there exists a unique flow

$$\alpha\colon J_b \times B_a(x_0) \to U.$$

If f is of class C^p $(p \geqq 1)$, then so is each integral curve α_x.

Proof. Let I_b be the closed interval $-b \leqq t \leqq b$, *and let* x be a fixed point in $\bar{B}_a(x_0)$. Let M be the set of continuous maps

$$\alpha\colon I_b \to \bar{B}_{2a}(x_0)$$

of the closed interval into the closed ball of center x_0 and radius $2a$, such that $\alpha(0) = x$. Then M is a complete metric space if we define as usual the distance between maps α, β to be

$$\sup_{t \in I_b} |\alpha(t) - \beta(t)|.$$

We shall now define a mapping

$$S\colon M \to M$$

of M into itself. For each α in M, we let $S\alpha$ be defined by

$$(S\alpha)(t) = x + \int_0^t f(u, \alpha(u))\, du.$$

Then $S\alpha$ is certainly continuous, we have $S\alpha(0) = x$, and the distance of any point on $S\alpha$ from x is bounded by the norm of the integral, which is bounded by

$$b \sup|f(u, y)| \leqq bL < a.$$

Thus $S\alpha$ lies in M.

We contend that our map S is a shrinking map. Indeed,

$$\begin{aligned} |S\alpha - S\beta| &\leqq b \sup|f(u, \alpha(u)) - f(u, \beta(u))| \\ &\leqq bK|\alpha - \beta|, \end{aligned}$$

thereby proving our contention.

By the shrinking lemma (Chapter I, Lemma 5.1) our map has a unique fixed point α, and by definition, $\alpha(t)$ satisfies the desired integral relation. Our remark above concludes the proof.

Corollary 1.2. *The local flow α in Proposition* 1.1 *is continuous. Furthermore, the map $x \mapsto \alpha_x$ of $\bar{B}_a(x_0)$ into the space of curves is continuous, and in fact satisfies a Lipschitz condition.*

Proof. The second statement obviously implies the first. So fix x in $\bar{B}_a(x_0)$ and take y close to x in $\bar{B}_a(x_0)$. We let S_x be the shrinking map of the theorem, corresponding to the initial condition x. Then

$$\|\alpha_x - S_y\alpha_x\| = \|S_x\alpha_x - S_y\alpha_x\| \leqq |x - y|.$$

Let $C = bK$ so $0 < C < 1$. Then

$$\begin{aligned} \|\alpha_x - S_y^n\alpha_x\| &\leqq \|\alpha_x - S_y\alpha_x\| + \|S_y\alpha_x - S_y^2\alpha_x\| + \cdots + \|S_y^{n-1}\alpha_x - S_y^n\alpha_x\| \\ &\leqq (1 + C + \cdots + C^{n-1})|x - y|. \end{aligned}$$

Since the limit of $S_y^n\alpha_x$ is equal to α_y as n goes to infinity, the continuity of the map $x \mapsto \alpha_x$ follows at once. In fact, the map satisfies a Lipschitz condition as stated.

It is easy to formulate a uniqueness theorem for integral curves over their whole domain of definition.

Theorem 1.3 (Uniqueness Theorem). *Let U be open in $\mathbf{E}$ and let $f\colon U \to E$ be a vector field of class C^p, $p \geqq 1$. Let*

$$\alpha_1\colon J_1 \to U \quad \text{and} \quad \alpha_2\colon J_2 \to U$$

be two integral curves for f with the same initial condition x_0. Then α_1 and α_2 are equal on $J_1 \cap J_2$.

Proof. Let Q be the set of numbers b such that $\alpha_1(t) = \alpha_2(t)$ for

$$0 \leqq t < b.$$

Then Q contains some number $b > 0$ by the local uniqueness theorem. If Q is not bounded from above, the equality of $\alpha_1(t)$ and $\alpha_2(t)$ for all $t > 0$ follows at once. If Q is bounded from above, let b be its least upper bound. We must show that b is the right end point of $J_1 \cap J_2$. Suppose that this is not the case. Define curves β_1 and β_2 near 0 by

$$\beta_1(t) = \alpha_1(b + t) \quad \text{and} \quad \beta_2(t) = \alpha_2(b + t).$$

Then β_1 and β_2 are integral curves of f with the initial conditions $\alpha_1(b)$ and $\alpha_2(b)$ respectively. The values $\beta_1(t)$ and $\beta_2(t)$ are equal for small negative t because b is the least upper bound of Q. By continuity it follows that $\alpha_1(b) = \alpha_2(b)$, and finally we see from the local uniqueness theorem that

$$\beta_1(t) = \beta_2(t)$$

for all t in some neighborhood of 0, whence α_1 and α_2 are equal in a neighborhood of b, contradicting the fact that b is a least upper bound of Q. We can argue the same way towards the left end points, and thus prove our statement.

For each $x \in U$, let $J(x)$ be the union of all open intervals containing 0 on which integral curves for f are defined, with initial condition equal to x. The uniqueness statement allows us to define the integral curve uniquely on all of $J(x)$.

Remark. The choice of 0 as the initial time value is made for convenience. From the uniqueness statement one obtains at once (making a time translation) the analogous statement for an integral curve defined on any open interval; in other words, if J_1, J_2 do not necessarily contain 0, and t_0 is a point in $J_1 \cap J_2$ such that $\alpha_1(t_0) = \alpha_2(t_0)$, and also we have the differential equations

$$\alpha_1'(t) = f(\alpha_1(t)) \quad \text{and} \quad \alpha_2'(t) = f(\alpha_2(t)),$$

then α_1 and α_2 are equal on $J_1 \cap J_2$.

In practice, one meets vector fields which may be time dependent, and also depend on parameters. We discuss these to show that their study reduces to the study of the standard case.

Time-dependent vector fields

Let J be an open interval, U open in a Banach space $\mathbf{E}$, and

$$f\colon J \times U \to \mathbf{E}$$

a C^p map, which we view as depending on time $t \in J$. Thus for each t, the map $x \mapsto f(t, x)$ is a vector field on U. Define

$$\bar{f}\colon J \times U \to \mathbf{R} \times \mathbf{E}$$

by

$$\bar{f}(t, x) = (1, f(t, x)),$$

and view $\bar{f}$ as a time-independent vector field on $J \times U$. Let $\bar{\alpha}$ be its flow, so that

$$\bar{\alpha}'(t, s, x) = \bar{f}(\bar{\alpha}(t, s, x)), \quad \bar{\alpha}(0, s, x) = (s, x).$$

We note that $\bar{\alpha}$ has its values in $J \times U$ and thus can be expressed in terms of two components. In fact, it follows at once that we can write $\bar{\alpha}$ in the form

$$\bar{\alpha}(t, s, x) = (t + s, \bar{\alpha}_2(t, s, x)).$$

Then $\bar{\alpha}_2$ satisfies the differential equation

$$D_1\bar{\alpha}_2(t, s, x) = f(t + s, \bar{\alpha}_2(t, s, x))$$

as we see from the definition of $\bar{f}$. Let

$$\beta(t, x) = \bar{\alpha}_2(t, 0, x).$$

Then β is a flow for f, that is β satisfies the differential equation

$$D_1\beta(t, x) = f(t, \beta(t, x)), \quad \beta(0, x) = x.$$

Given $x \in U$, any value of t such that α is defined at (t, x) is also such that $\bar{\alpha}$ is defined at $(t, 0, x)$ because α_x and β_x are integral curves of the same vector field, with the same initial condition, hence are equal. Thus the study of time-dependent vector fields is reduced to the study of time-independent ones.

Dependence on parameters

Let V be open in some space $\mathbf{F}$ and let

$$g : J \times V \times U \to \mathbf{E}$$

be a map which we view as a time-dependent vector field on U, also depending on parameters in V. We define

$$G : J \times V \times U \to \mathbf{F} \times \mathbf{E}$$

by

$$G(t, z, y) = (0, g(t, z, y))$$

for $t \in J$, $z \in V$, and $y \in U$. This is now a time-dependent vector field on $V \times U$. A local flow for G depends on three variables, say $\beta(t, z, y)$, with initial condition $\beta(0, z, y) = (z, y)$. The map β has two components, and it is immediately clear that we can write

$$\beta(t, z, y) = (z, \alpha(t, z, y))$$

for some map α depending on three variables. Consequently α satisfies the differential equation

$$D_1\alpha(t, z, y) = g(t, z, \alpha(t, z, y)), \quad \alpha(0, z, y) = y,$$

which gives the flow of our original vector field g depending on the parameters $z \in V$. This procedure reduces the study of differential equations depending on parameters to those which are independent of parameters.

We shall now investigate the behavior of the flow with respect to its second argument, i.e. with respect to the points of U. We shall give two methods for this. The first depends on approximation estimates, and the second on the implicit mapping theorem in function spaces.

Let J_0 be an open subinterval of J containing 0, and let

$$\varphi : J_0 \to U$$

be of class C^1. We shall say that φ is an **ϵ-approximate solution** of f on J_0 if

$$|\varphi'(t) - f(t, \varphi(t))| \leqq \epsilon$$

for all t in J_0.

Proposition 1.4. *Let φ_1 and φ_2 be two ϵ_1- and ϵ_2-approximate solutions of f on J_0 respectively, and let $\epsilon = \epsilon_1 + \epsilon_2$. Assume that f is Lipschitz with constant K on U uniformly in J_0, or that $D_2 f$ exists and is bounded by K on $J \times U$. Let t_0 be a point of J_0. Then for any t in J_0, we have*

$$|\varphi_1(t) - \varphi_2(t)| \leqq |\varphi_1(t_0) - \varphi_2(t_0)|e^{K|t-t_0|} + \frac{\epsilon}{K} e^{K|t-t_0|}.$$

Proof. By assumption, we have

$$|\varphi_1'(t) - f(t, \varphi_1(t))| \leqq \epsilon_1,$$
$$|\varphi_2'(t) - f(t, \varphi_2(t))| \leqq \epsilon_2.$$

From this we get

$$|\varphi_1'(t) - \varphi_2'(t) + f(t, \varphi_2(t)) - f(t, \varphi_1(t))| \leqq \epsilon.$$

Say $t \geqq t_0$ to avoid putting bars around $t - t_0$. Let

$$\psi(t) = |\varphi_1(t) - \varphi_2(t)|,$$
$$\omega(t) = |f(t, \varphi_1(t)) - f(t, \varphi_2(t))|.$$

Then, after integrating from t_0 to t, and using triangle inequalities we obtain

$$\begin{aligned} |\psi(t) - \psi(t_0)| &\leqq \epsilon(t - t_0) + \int_{t_0}^{t} \omega(u)\, du \\ &\leqq \epsilon(t - t_0) + K \int_{t_0}^{t} \psi(u)\, du \\ &\leqq K \int_{t_0}^{t} [\psi(u) + \epsilon/K]\, du, \end{aligned}$$

and finally the recurrence relation

$$\psi(t) \leqq \psi(t_0) + K \int_{t_0}^{t} [\psi(u) + \epsilon/K]\, du.$$

On any closed subinterval of J_0, our map ψ is bounded. If we add ϵ/K to both sides of this last relation, then we see that our proposition will follow from the next lemma.

Lemma 1.5. *Let g be a positive real valued function on an interval, bounded by a number L. Let t_0 be in the interval, say $t_0 \leqq t$, and assume that there are numbers $A, K \geqq 0$ such that*

$$g(t) \leqq A + K \int_{t_0}^{t} g(u)\, du.$$

Then for all integers $n \geqq 1$ we have

$$g(t) \leqq A\left[1 + \frac{K(t-t_0)}{1!} + \cdots + \frac{K^{n-1}(t-t_0)^{n-1}}{(n-1)!}\right] + \frac{LK^n(t-t_0)^n}{n!}.$$

Proof. The statement is an assumption for $n = 1$. We proceed by induction. We integrate from t_0 to t, multiply by K, and use the recurrence relation. The statement with $n+1$ then drops out of the statement with n.

Corollary 1.6. *Let $f\colon J \times U \to \mathbf{E}$ be continuous, and satisfy a Lipschitz condition on U uniformly with respect to J. Let x_0 be a point of U. Then there exists an open subinterval J_0 of J containing 0, and an open subset of U containing x_0 such that f has a unique flow*

$$\alpha\colon J_0 \times U_0 \to U.$$

We can select J_0 and U_0 such that α is continuous and satisfies a Lipschitz condition on $J_0 \times U_0$.

Proof. Given x, y in U_0 we let $\varphi_1(t) = \alpha(t, x)$ and $\varphi_2(t) = \alpha(t, y)$, using Proposition 1.6 to get J_0 and U_0. Then $\epsilon_1 = \epsilon_2 = 0$. For s, t in J_0 we obtain

$$\begin{aligned}|\alpha(t, x) - \alpha(s, y)| &\leqq |\alpha(t, x) - \alpha(t, y)| + |\alpha(t, y) - \alpha(s, y)| \\ &\leqq |x - y|e^K + |t - s|L,\end{aligned}$$

if we take J_0 of small length, and L is a bound for f. Indeed, the term containing $|x-y|$ comes from Proposition 1.4, and the term containing $|t-s|$ comes from the definition of the integral curve by means of an integral and the bound L for f. This proves our corollary.

Corollary 1.7. *Let J be an open interval of $\mathbf{R}$ containing 0 and let U be open in $\mathbf{E}$. Let $f\colon J \times U \to \mathbf{E}$ be a continuous map, which is Lipschitz on U uniformly for every compact subinterval of J. Let $t_0 \in J$ and let φ_1, φ_2 be two morphisms of class C^1 such that $\varphi_1(t_0) = \varphi_2(t_0)$ and satisfying the relation*

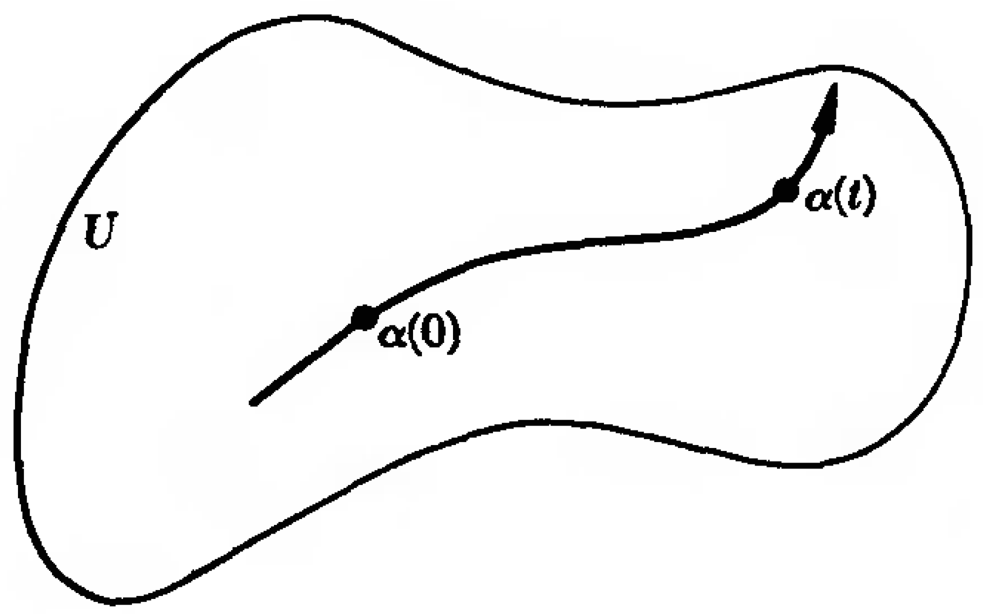
$$\varphi'(t) = f(t, \varphi(t))$$

for all t in J. Then $\varphi_1(t) = \varphi_2(t)$.

Proof. We can take $\epsilon = 0$ in the proposition.

The above corollary gives us another proof for the uniqueness of integral curves. Given $f\colon J \times U \to \mathbf{E}$ as in this corollary, we can define an integral curve α for f on a maximal open subinterval of J having a given value $\alpha(t_0)$ for a fixed t_0 in J. Let J be the open interval (a, b) and let (a_0, b_0) be the interval on which α is defined. We want to know when $b_0 = b$ (or $a_0 = a$), that is when the integral curve of f can be continued to the entire interval over which f itself is defined.

There are essentially two reasons why it is possible that the integral curve cannot be extended to the whole domain of definition J, or cannot be extended to infinity in case f is independent of time. One possibility is that the integral curve tends to get out of the open set U, as on the following picture:



This means that as t approaches b_0, say, the curve $\alpha(t)$ approaches a point which does not lie in U. Such an example can actually be constructed artificially. If we are in a situation when a curve can be extended to infinity, just remove a point from the open set lying on the curve. Then the integral curve on the resulting open set cannot be continued to infinity. The second possibility is that the vector field is unbounded. The next corollary shows that these possibilities are the only ones. In other words, if an integral curve does not tend to get out of the open set, and if the vector field is bounded, then the curve can be continued as far as the original data will allow a priori.

Corollary 1.8. *Let J be the open interval (a, b) and let U be open in $\mathbf{E}$. Let $f\colon J \times U \to \mathbf{E}$ be a continuous map, which is Lipschitz on U,* *uniformly for every compact subset of J. Let α be an integral curve of f, defined on a maximal open subinterval (a_0, b_0) of J. Assume:*

(i) *There exists $\epsilon > 0$ such that $\overline{\alpha((b_0 - \epsilon, b_0))}$ is contained in U.*
(ii) *There exists a number $B > 0$ such that $|f(t, \alpha(t))| \leqq B$ for all t in $(b_0 - \epsilon, b_0)$.*

Then $b_0 = b$.

Proof. From the integral expression for α, namely

$$\alpha(t) = \alpha(t_0) + \int_{t_0}^{t} f(u, \alpha(u))\, du,$$

we see that for t_1, t_2 in $(b_0 - \epsilon, b_0)$ we have

$$|\alpha(t_1) - \alpha(t_2)| \leqq B|t_1 - t_2|.$$

From this it follows that the limit

$$\lim_{t \to b_0} \alpha(t)$$

exists, and is equal to an element x_0 of U (by hypothesis (i)). Assume that $b_0 \neq b$. By the local existence theorem, there exists an integral curve β of f defined on an open interval containing b_0 such that $\beta(b_0) = x_0$ and $\beta'(t) = f(t, \beta(t))$. Then $\beta' = \alpha'$ on an open interval to the left of b_0, and hence α, β differ by a constant on this interval. Since their limit as $t \to b_0$ are equal, this constant is 0. Thus we have extended the domain of definition of α to a larger interval, as was to be shown.

The next proposition describes the solutions of **linear differential equations** depending on parameters.

Proposition 1.9. *Let J be an open interval of* **R** *containing* 0, *and let V be an open set in a Banach space. Let* **E** *be a Banach space. Let*

$$g\colon J \times V \to L(\mathbf{E}, \mathbf{E})$$

be a continuous map. Then there exists a unique map

$$\lambda\colon J \times V \to L(\mathbf{E}, \mathbf{E})$$

which, for each $x \in V$, is a solution of the differential equation

$$D_1\lambda(t, x) = g(t, x)\lambda(t, x), \quad \lambda(0, x) = \mathrm{id}.$$

This map λ is continuous.

Remark. In the present case of a linear differential equation, it is not necessary to shrink the domain of definition of its flow. Note that the differential equation is on the space of continuous linear maps. The corresponding linear equation on **E** itself will come out as a corollary.

Proof of Proposition 1.9. Let us first fix $x \in V$. Consider the differential equation

$$D_1\lambda(t, x) = g(t, x)\lambda(t, x),$$

with initial condition $\lambda(0, x) = \mathrm{id}$. This is a differential equation on $L(\mathbf{E}, \mathbf{E})$, where $f(t, z) = g_x(t)z$ for $z \in L(\mathbf{E}, \mathbf{E})$, and we write $g_x(t)$ instead of $g(t, x)$. Let the notation be as in Corollary 1.8. Then hypothesis (i) is automatically satisfied since the open set U is all of $L(\mathbf{E}, \mathbf{E})$. On every compact subinterval of J, g_x is bounded, being continuous. Omitting the index x for simplicity, we have

$$\lambda(t) = \mathrm{id} + \int_0^t g(u)\lambda(u)\, du,$$

whence for $t \geqq 0$, say

$$|\lambda(t)| \leqq 1 + B\int_0^t |\lambda(u)|\, du.$$

Using Lemma 1.5, we see that hypothesis (ii) of Corollary 1.8 is also satisfied. Hence the integral curve is defined on all of J.

We shall now prove the continuity of λ. Let $(t_0, x_0) \in J \times V$. Let I be a compact interval contained in J, and containing t_0 and 0. As a function of t, $\lambda(t, x_0)$ is continuous (even differentiable). Let $C > 0$ be such that $|\lambda(t, x_0)| \leqq C$ for all $t \in I$. Let V_1 be an open neighborhood of x_0 in V such that g is bounded by a constant $K > 0$ on $I \times V_1$.

For $(t, x) \in I \times V_1$ we have

$$|\lambda(t, x) - \lambda(t_0, x_0)| \leqq |\lambda(t, x) - \lambda(t, x_0)| + |\lambda(t, x_0) - \lambda(t_0, x_0)|.$$

The second term on the right is small when t is close to t_0. We investigate the first term on the right, and shall estimate it by viewing $\lambda(t, x)$ and $\lambda(t, x_0)$ as approximate solutions of the differential equation satisfied by $\lambda(t, x)$. We find

$$\begin{aligned} &|D_1\lambda(t, x_0) - g(t, x)\lambda(t, x_0)| \\ &\quad = |D_1\lambda(t, x_0) - g(t, x)\lambda(t, x_0) + g(t, x_0)\lambda(t, x_0) - g(t, x_0)\lambda(t, x_0)| \\ &\quad \leqq |g(t, x_0) - g(t, x)|\,|\lambda(t, x_0)| \leqq |g(t, x_0) - g(t, x)|C. \end{aligned}$$

By the usual proof of uniform continuity applied to the compact set $I \times \{x_0\}$, given $\epsilon > 0$, there exists an open neighborhood V_0 of x_0 contained in V_1, such that for all $(t, x) \in I \times V_0$ we have

$$|g(t, x) - g(t, x_0)| < \epsilon/C.$$

This implies that $\lambda(t, x_0)$ is an ϵ-approximate solution of the differential equation satisfied by $\lambda(t, x)$. We apply Proposition 1.4 to the two curves

$$\varphi_0(t) = \lambda(t, x_0) \quad \text{and} \quad \varphi_x(t) = \lambda(t, x)$$

for each $x \in V_0$. We use the fact that $\lambda(0, x) = \lambda(0, x_0) = \text{id}$. We then find

$$|\lambda(t, x) - \lambda(t, x_0)| < \epsilon K_1$$

for some constant $K_1 > 0$, thereby proving the continuity of λ at (t_0, x_0).

Corollary 1.10. *Let the notation be as in Proposition* 1.9. *For each* $x \in V$ *and* $z \in E$ *the curve*

$$\beta(t, x, z) = \lambda(t, x)z$$

with initial condition $\beta(0, x, z) = z$ *is a solution of the differential equation*

$$D_1\beta(t, x, z) = g(t, x)\beta(t, x, z).$$

Furthermore, β *is continuous in its three variables.*

Proof. Obvious.

Theorem 1.11 (Local Smoothness Theorem). *Let J be an open interval in* **R** *containing* 0 *and U open in the Banach space* **E**. *Let*

$$f\colon J \times U \to \mathbf{E}$$

be a C^p-morphism with $p \geqq 1$, and let $x_0 \in U$. There exists a unique local flow for f at x_0. We can select an open subinterval J_0 of J containing 0 *and an open subset U_0 of U containing x_0 such that the unique local flow*

$$\alpha\colon J_0 \times U_0 \to U$$

is of class C^p, and such that $D_2\alpha$ satisfies the differential equation

$$\boxed{D_1D_2\alpha(t, x) = D_2f\big(t, \alpha(t, x)\big)D_2\alpha(t, x)}$$

on $J_0 \times U_0$ with initial condition $D_2\alpha(0, x) = \text{id}$.

Proof. Let

$$g\colon J \times U \to L(\mathbf{E}, \mathbf{E})$$

be given by $g(t, x) = D_2f\big(t, \alpha(t, x)\big)$. Select J_1 and U_0 such that α is bounded and Lipschitz on $J_1 \times U_0$ (by Corollary 1.6), and such that g is continuous and bounded on $J_1 \times U_0$. Let J_0 be an open subinterval of J_1 containing 0 such that its closure $\bar{J}_0$ is contained in J_1.

Let $\lambda(t, x)$ be the solution of the differential equation on $L(\mathbf{E}, \mathbf{E})$ given by

$$D_1\lambda(t, x) = g(t, x)\lambda(t, x), \qquad \lambda(0, x) = \text{id},$$

as in Proposition 1.9. We contend that $D_2\alpha$ exists and is equal to λ on $J_0 \times U_0$. This will prove that $D_2\alpha$ is continuous, on $J_0 \times U_0$.

Fix $x \in U_0$. Let

$$\theta(t, h) = \alpha(t, x + h) - \alpha(t, x).$$

Then

$$\begin{aligned} D_1\theta(t, h) &= D_1\alpha(t, x + h) - D_1\alpha(t, x) \\ &= f\big(t, \alpha(t, x + h)\big) - f\big(t, \alpha(t, x)\big). \end{aligned}$$

By the mean value theorem, we obtain

$$\begin{aligned} &|D_1\theta(t, h) - g(t, x)\theta(t, h)| \\ &\quad = \big|f\big(t, \alpha(t, x + h)\big) - f\big(t, \alpha(t, x)\big) - D_2f\big(t, \alpha(t, x)\big)\theta(t, h)\big| \\ &\quad \leqq |h| \sup |D_2f(t, y) - D_2f\big(t, \alpha(t, x)\big)|, \end{aligned}$$

where y ranges over the segment between $\alpha(t, x)$ and $\alpha(t, x + h)$. By the compactness of $\bar{J}_0$ it follows that our last expression is bounded by $|h|\psi(h)$ where $\psi(h)$ tends to 0 with h, uniformly for t in $\bar{J}_0$. Hence we obtain

$$|\theta'(t, h) - g(t, x)\theta(t, h)| \leqq |h|\psi(h),$$

for all t in $\bar{J}_0$. This shows that $\theta(t, h)$ is an $|h|\psi(h)$ approximate solution for the differential equation satisfied by $\lambda(t, x)h$, namely

$$D_1\lambda(t, x)h - g(t, x)\lambda(t, x)h = 0,$$

with the initial condition $\lambda(0, x)h = h$. We note that $\theta(t, h)$ has the same initial condition, $\theta(0, h) = h$. Taking $t_0 = 0$ in Proposition 1.4, we obtain the estimate

$$|\theta(t, h) - \lambda(t, x)h| \leqq C_1|h|\psi(h)$$

for all t in $\bar{J}_0$. This proves that $D_2\alpha$ is equal to λ on $J_0 \times U_0$, and is therefore continuous on $J_0 \times U_0$.

We have now proved that $D_1\alpha$ and $D_2\alpha$ exist and are continuous on $J_0 \times U_0$, and hence that α is of class C^1 on $J_0 \times U_0$.

Furthermore, $D_2\alpha$ satisfies the differential equation given in the statement of our theorem on $J_0 \times U_0$. Thus our theorem is proved when $p = 1$.

A flow which satisfies the properties stated in the theorem will be called **locally of class** C^p.

Consider now again the linear equation of Proposition 1.9. We reformulate it to eliminate formally the parameters, namely we define a vector field

$$G\colon J \times V \times L(\mathbf{E}, \mathbf{E}) \to F \times L(\mathbf{E}, \mathbf{E})$$

to be the map such that

$$G(t, x, \omega) = \big(0, g(t, x)\omega\big)$$

for $\omega \in L(\mathbf{E}, \mathbf{E})$. The flow for this vector field is then given by the map Λ such that

$$\Lambda(t, x, \omega) = \big(x, \lambda(t, x)\omega\big).$$

If g is of class C^1 we can now conclude that the flow Λ is locally of class C^1, and hence putting $\omega = \mathrm{id}$, that λ is locally of class C^1.

We apply this to the case when $g(t, x) = D_2 f\big(t, \alpha(t, x)\big)$, and to the solution $D_2\alpha$ of the differential equation

$$D_1(D_2\alpha)(t, x) = g(t, x) D_2\alpha(t, x)$$

locally at each point $(0, x)$, $x \in U$. Let $p \geqq 2$ be an integer and assume out theorem proved up to $p - 1$, so that we can assume α locally of class C^{p-1}, and f of class C^p. Then g is locally of class C^{p-1}, whence $D_2\alpha$ is locally C^{p-1}. From the expression

$$D_1\alpha(t, x) = f\big(t, \alpha(t, x)\big)$$

we conclude that $D_1\alpha$ is C^{p-1}, whence α is locally C^p.

If f is C^∞, and if we knew that α is of class C^p for every integer p **on its domain of definition**, then we could conclude that α is C^∞; in other words, there is no shrinkage in the inductive application of the local theorem. We shall do this at the end of the section.

We shall now give another proof for the local smoothness of the flow, which depends on a simple application of the implicit mapping theorem in Banach spaces, and was found independently by Pugh and Robbin [Ro 68]. One advantage of this proof is that it extends to H^p vector fields, as noted by Ebin and Marsden [EbM 70].

Let U be open in $\mathbf{E}$ and let $f\colon U \to \mathbf{E}$ be a C^p map. Let $b > 0$ and let I_b be the closed interval of radius b centered at 0. Let

$$F = C^0(I_b, \mathbf{E})$$

be the Banach space of continuous maps of I_b into $\mathbf{E}$. We let V be the subset of $\mathbf{F}$ consisting of all continuous curves

$$\sigma\colon I_b \to U$$

mapping I_b into our open set U. Then it is clear that V is open in $\mathbf{F}$ because for each curve σ the image $\sigma(I_b)$ is compact, hence at a finite distance from the complement of U, so that any curve close to it is also contained in U.

We define a map

$$T\colon U \times V \to \mathbf{F}$$

by

$$T(x, \sigma) = x + \int_0 f \circ \sigma - \sigma.$$

Here we omit the dummy variable of integration, and x stands for the constant curve with value x. If we evaluate the curve $T(x, \sigma)$ at t, then by definition we have

$$T(x, \sigma)(t) = x + \int_0^t f\big(\sigma(u)\big)\, du - \sigma(t).$$

Lemma 1.12. *The map T is of class C^p, and its second partial derivative is given by the formula*

$$D_2 T(x, \sigma) = \int_0 Df \circ \sigma - I$$

where I is the identity. In terms of t, this reads

$$D_2 T(x, \sigma)h(t) = \int_0^t Df\big(\sigma(u)\big)h(u)\, du - h(t).$$

Proof. It is clear that the first partial derivative $D_1 T$ exists and is continuous, in fact C^∞, being linear in x up to a translation. To determine the second partial, we apply the definition of the derivative. The derivative of the map $\sigma \mapsto \sigma$ is of course the identity. We have to get the derivative with respect to σ of the integral expression. We have for small h

$$\left\| \int_0 f \circ (\sigma + h) - \int_0 f \circ \sigma - \int_0 (Df \circ \sigma)h \right\|$$
$$\leqq \int_0 |f \circ (\sigma + h) - f \circ \sigma - (Df \circ \sigma)h|.$$

We estimate the expression inside the integral at each point u, with u between 0 and the upper variable of integration. From the mean value theorem, we get

$$|f(\sigma(u) + h(u)) - f(\sigma(u)) - Df(\sigma(u))h(u)| \leqq \|h\| \sup |Df(z_u) - Df(\sigma(u))|$$

where the sup is taken over all points z_u on the segment between $\sigma(u)$ and $\sigma(u) + h(u)$. Since Df is continuous, and using the fact that the image of the curve $\sigma(I_b)$ is compact, we conclude (as in the case of uniform continuity) that as $\|h\| \to 0$, the expression

$$\sup |Df(z_u) - Df(\sigma(u))|$$

also goes to 0. (Put the ϵ and δ in yourself.) By definition, this gives us the derivative of the integral expression in σ. The derivative of the final term is obviously the identity, so this proves that $D_2 T$ is given by the formula which we wrote down.

This derivative does not depend on x. It is continuous in σ. Namely, we have

$$D_2 T(x, \tau) - D_2 T(x, \sigma) = \int_0 [Df \circ \tau - Df \circ \sigma].$$

If σ is fixed and τ is close to σ, then $Df \circ \tau - Df \circ \sigma$ is small, as one proves easily from the compactness of $\sigma(I_b)$, as in the proof of uniform continuity. Thus $D_2 T$ is continuous. By Proposition 3.5 of Chapter I, we now conclude that T is of class C^1.

The derivative of $D_2 T$ with respect to σ can again be computed as before if Df is itself of class C^1, and thus by induction, if f is of class C^p we conclude that $D_2 T$ is of class C^{p-1} so that by the same reference, we conclude that T itself is of class C^p. This proves our lemma.

We observe that a solution of the equation

$$T(x, \sigma) = 0$$

is precisely an integral curve for the vector field, with initial condition equal to x. Thus we are in a situation where we want to apply the implicit mapping theorem.

Lemma 1.13. *Let $x_0 \in U$. Let $a > 0$ be such that Df is bounded, say by a number $C_1 > 0$, on the ball $B_a(x_0)$ (we can always find such a since Df is continuous at x_0). Let $b < 1/C_1$. Then $D_2 T(x, \sigma)$ is invertible for all (x, σ) in $B_a(x_0) \times V$.*

Proof. We have an estimate

$$\left| \int_0^t Df(\sigma(u))h(u)\, du \right| \leqq bC_1 \|h\|.$$

This means that

$$|D_2 T(x, \sigma) + I| < 1,$$

and hence that $D_2 T(x, \sigma)$ is invertible, as a continuous linear map, thus proving Lemma 1.13.

We are ready to reprove the local smoothness theorem by the present means, when p is an integer, namely:

Theorem 1.14. *Let p be a positive integer, and let $f\colon U \to \mathbf{E}$ be a C^p vector field. Let $x_0 \in U$. Then there exist numbers $a, b > 0$ such that the local flow*

$$\alpha\colon J_b \times B_a(x_0) \to U$$

is of class C^p.

Proof. We take a so small and then b so small that the local flow exists and is uniquely determined by Proposition 1.1. We then take b smaller and a smaller so as to satisfy the hypotheses of Lemma 1.13. We can then apply the implicit mapping theorem to conclude that the map $x \mapsto \alpha_x$ is of class C^p. Of course, we have to consider the flow α and still must show that α itself is of class C^p. It will suffice to prove that $D_1\alpha$ and $D_2\alpha$ are of class C^{p-1}, by Proposition 3.5 of Chapter I. We first consider the case $p = 1$.

We could derive the continuity of α from Corollary 1.2 but we can also get it as an immediate consequence of the continuity of the map $x \mapsto \alpha_x$. Indeed, fixing (s, y) we have

$$|\alpha(t, x) - \alpha(s, y)| \leqq |\alpha(t, x) - \alpha(t, y)| + |\alpha(t, y) - \alpha(s, y)|$$
$$\leqq \|\alpha_x - \alpha_y\| + |\alpha_y(t) - \alpha_y(s)|.$$

Since α_y is continuous (being differentiable), we get the continuity of α.

Since

$$D_1\alpha(t, x) = f(\alpha(t, x)),$$

we conclude that $D_1\alpha$ is a composite of continuous maps, whence continuous.

Let φ be the derivative of the map $x \mapsto \alpha_x$, so that

$$\varphi\colon B_a(x_0) \to L(\mathbf{E}, C^0(I_b, \mathbf{E})) = L(\mathbf{E}, \mathbf{F})$$

is of class C^{p-1}. Then

$$\alpha_{x+w} - \alpha_x = \varphi(x)w + |w|\psi(w),$$

where $\psi(w) \to 0$ as $w \to 0$. Evaluating at t, we find

$$\alpha(t, x+w) - \alpha(t, x) = (\varphi(x)w)(t) + |w|\psi(w)(t),$$

and from this we see that

$$D_2\alpha(t, x)w = (\varphi(x)w)(t).$$

Then

$$\begin{aligned}&|D_2\alpha(t, x)w - D_2\alpha(s, y)w|\\ &\qquad \leqq |(\varphi(x)w)(t) - (\varphi(y)w)(t)| + |(\varphi(y)w)(t) - (\varphi(y)w)(s)|.\end{aligned}$$

The first term on the right is bounded by

$$|\varphi(x) - \varphi(y)|\,|w|$$

so that

$$|D_2\alpha(t, x) - D_2\alpha(t, y)| \leqq |\varphi(x) - \varphi(y)|.$$

We shall prove below that

$$|(\varphi(y)w)(t) - (\varphi(y)w)(s)|$$

is uniformly small with respect to w when s is close to t. This proves the continuity of $D_2\alpha$, and concludes the proof that α is of class C^1.

The following proof that $|(\varphi(y)w)(t) - (\varphi(y)w)(s)|$ is uniformly small was shown to me by Professor Yamanaka. We have

$$(1) \qquad \alpha(t, x) = x + \int_0^t f(\alpha(u, x))\, du.$$

Replacing x with $x + \lambda w$ ($w \in E$, $\lambda \neq 0$), we obtain

$$(2) \qquad \alpha(t, x+\lambda w) = x + \lambda w + \int_0^t f(\alpha(u, x+\lambda w))\, du.$$

Therefore

$$(3) \qquad \frac{\alpha(t, x+\lambda w) - \alpha(t, x)}{\lambda} = w + \int_0^t \frac{1}{\lambda}[f(\alpha(u, x+\lambda w)) - f(\alpha(u, x))]\, du.$$

On the other hand, we have already seen in the proof of Theorem 1.14 that

$$(4) \qquad \alpha(t, x+\lambda w) - \alpha(t, x) = \lambda(\varphi(x)w)(t) + |\lambda|\,|w|\psi(\lambda w)(t).$$

Substituting (4) in (3), we obtain:

$$\begin{aligned}(\varphi(t)w)(t) + \frac{|\lambda|}{\lambda}|w|\psi(\lambda w)(t) &= w + \int_0^t \frac{1}{\lambda}[f(\alpha(u, x+\lambda w)) - f(\alpha(u, x))]\, du\\ &= w + \int_0^t \int_0^1 G(u, \lambda, v)\, dv\, du,\end{aligned}$$

where

$$G(u, \lambda, v) = Df(\alpha(u, x) + v\epsilon_1(\lambda))((\varphi(x)w)(u) + \epsilon_2(\lambda))$$

with

$$\epsilon_1(\lambda) = \lambda(\varphi(x)w)(u) + |\lambda|\,|w|\psi(\lambda w)(u), \quad \epsilon_2(\lambda) = \frac{|\lambda|}{\lambda}\psi(\lambda w)(u).$$

Letting $\lambda \to 0$, we have

$$(5) \qquad (\varphi(x)w)(t) = w + \int_0^t Df(\alpha(u, x))(\varphi(x)w)(u)\, du.$$

By (5) we have

$$\begin{aligned}|(\varphi(x)w)(t) - (\varphi(x)w)(s)| &\leqq \left|\int_s^t Df(\alpha(u, x))(\varphi(x)w)(u)\, du\right|\\ &\leqq bC_1|\varphi(x)|\cdot|w|\cdot|t-s|,\end{aligned}$$

from which we immediately obtain the desired uniformity.

Returning to our main concern, the flow, we have

$$\alpha(t, x) = x + \int_0^t f(\alpha(u, x))\, du.$$

We can differentiate under the integral sign with respect to the parameter x and thus obtain

$$D_2\alpha(t, x) = I + \int_0^t Df(\alpha(u, x))D_2\alpha(u, x)\,du,$$

where I is a constant linear map (the identity). Differentiating with respect to t yields the linear differential equation satisfied by $D_2\alpha$, namely

$$\boxed{D_1D_2\alpha(t, x) = Df(\alpha(t, x))D_2\alpha(t, x)}$$

and this differential equation depends on time and parameters. We have seen earlier how such equations can be reduced to the ordinary case. We now conclude that locally, by induction, $D_2\alpha$ is of class C^{p-1} since Df is of class C^{p-1}. Since

$$D_1\alpha(t, x) = f(\alpha(t, x)),$$

we conclude by induction that $D_1\alpha$ is C^{p-1}. Hence α is of class C^p by Proposition 3.5 of Chapter I. Note that each time we use induction, the domain of the flow may shrink. We have proved Theorem 1.14, when p is an integer.

We now give the arguments needed to globalize the smoothness. We may limit ourselves to the time-independent case. We have seen that the time-dependent case reduces to the other.

Let U be open in a Banach space $\mathbf{E}$, and let $f\colon U \to \mathbf{E}$ be a C^p vector field. We let $J(x)$ be the domain of the integral curve with initial condition equal to w.

Let $\mathfrak{D}(f)$ be the set of all points (t, x) in $\mathbf{R} \times U$ such that t lies in $J(x)$. Then we have a map

$$\alpha\colon \mathfrak{D}(f) \to U$$

defined on all of $\mathfrak{D}(f)$, letting $\alpha(t, x) = \alpha_x(t)$ be the integral curve on $J(x)$ having x as initial condition. We call this the **flow** determined by f, and we call $\mathfrak{D}(f)$ its **domain of definition.**

Lemma 1.15. *Let $f\colon U \to E$ be a C^p vector field on the open set U of E, and let α be its flow. Abbreviate $\alpha(t, x)$ by tx, if (t, x) is in the domain of definition of the flow. Let $x \in U$. If t_0 lies in $J(x)$, then*

$$J(t_0x) = J(x) - t_0$$

(translation of $J(x)$ by $-t_0$), and we have for all t in $J(x) - t_0$:

$$t(t_0x) = (t + t_0)x.$$

Proof. The two curves defined by

$$t \mapsto \alpha(t, \alpha(t_0, x)) \qquad \text{and} \qquad t \mapsto \alpha(t + t_0, x)$$

are integral curves of the same vector field, with the same initial condition t_0x at $t = 0$. Hence they have the same domain of definition $J(t_0x)$. Hence t_1 lies in $J(t_0x)$ if and only if $t_1 + t_0$ lies in $J(x)$. This proves the first assertion. The second assertion comes from the uniqueness of the integral curve having given initial condition, whence the theorem follows.

Theorem 1.16 (Global Smoothness of the Flow). *If f is of class C^p (with $p \leqq \infty$), then its flow is of class C^p on its domain of definition.*

Proof. First let p be an integer $\geqq 1$. We know that the flow is locally of class C^p at each point $(0, x)$, by the local theorem. Let $x_0 \in U$ and let $J(x_0)$ be the maximal interval of definition of the integral curve having x_0 as initial condition. Let $\mathfrak{D}(f)$ be the domain of definition of the flow, and let α be the flow. Let Q be the set of numbers $b > 0$ such that for each t with $0 \leqq t < b$ there exists an open interval J containing t and an open set V containing x_0 such that $J \times V$ is contained in $\mathfrak{D}(f)$ and such that α is of class C^p on $J \times V$. Then Q is not empty by the local theorem. If Q is not bounded from above, then we are done looking toward the right end point of $J(x_0)$. If Q is bounded from above, we let b be its least upper bound. We must prove that b is the right end point of $J(x_0)$. Suppose that this is not the case. Then $\alpha(b, x_0)$ is defined. Let $x_1 = \alpha(b, x_0)$. By the local theorem, we have a unique local flow at x_1, which we denote by β:

$$\beta\colon J_a \times B_a(x_1) \to U, \qquad \beta(0, x) = x,$$

defined for some open interval $J_a = (-a, a)$ and open ball $B_a(x_1)$ of radius a centered at x_1. Let δ be so small that whenever $b - \delta < t < b$ we have

$$\alpha(t, x_0) \in B_{a/4}(x_1).$$

We can find such δ because

$$\lim_{t \to b} \alpha(t, x_0) = x_1$$

by continuity. Select a point t_1 such that $b - \delta < t_1 < b$. By the hypothesis on b, we can select an open interval J_1 containing t_1 and an open set U_1 containing x_0 so that

$$\alpha\colon J_1 \times U_1 \to B_{a/2}(x_1)$$

maps $J_1 \times U_1$ into $B_{a/2}(x_1)$. We can do this because α is continuous at

(t_1, x_0), being in fact C^p at this point. If $|t - t_1| < a$ and $x \in U_1$, we define

$$\varphi(t, x) = \beta(t - t_1, \alpha(t_1, x)).$$

Then

$$\varphi(t_1, x) = \beta(0, \alpha(t_1, x)) = \alpha(t_1, x)$$

and

$$\begin{aligned} D_1\varphi(t, x) &= D_1\beta(t - t_1, \alpha(t_1, x)) \\ &= f(\beta(t - t_1, \alpha(t_1, x))) \\ &= f(\varphi(t, x)). \end{aligned}$$

Hence both φ_x and α_x are integral curves for f with the same value at t_1. They coincide on any interval on which they are defined by the uniqueness theorem. If we take δ very small compared to a, say $\delta < a/4$, we see that φ is an extension of α to an open set containing (t_1, α_0), and also containing (b, x_0). Furthermore, φ is of class C^p, thus contradicting the fact that b is strictly smaller than the end point of $J(x_0)$. Similarly, one proves the analogous statement on the other side, and we therefore see that $\mathfrak{D}(f)$ is open in $\mathbf{R} \times U$ and that α is of class C^p on $\mathfrak{D}(f)$, as was to be shown.

The idea of the above proof is very simple geometrically. We go as far to the right as possible in such a way that the given flow α is of class C^p locally at (t, x_0). At the point $\alpha(b, x_0)$ we then use the flow β to extend differentiably the flow α in case b is not the right-hand point of $J(x_0)$. The flow β at $\alpha(b, x_0)$ has a fixed local domain of definition, and we simply take t close enough to b so that β gives an extension of α, as described in the above proof.

Of course, if f is of class C^∞, then we have shown that α is of class C^p for each positive integer p, and therefore the flow is also of class C^∞.

In the next section, we shall see how these arguments globalize even more to manifolds.

IV, §2. VECTOR FIELDS, CURVES, AND FLOWS

Let X be a manifold of class C^p with $p \geqq 2$. We recall that X is assumed to be Hausdorff. Let $\pi\colon T(X) \to X$ be its tangent bundle. Then $T(X)$ is of class C^{p-1}, $p - 1 \geqq 1$.

By a (time-independent) **vector field** on X we mean a cross section of the tangent bundle, i.e. a morphism (of class C^{p-1})

$$\xi\colon X \to T(X)$$

such that $\xi(x)$ lies in the tangent space $T_x(X)$ for each $x \in X$, or in other words, such that $\pi\xi = \mathrm{id}$.

If $T(X)$ is trivial, and say X is an $\mathbf{E}$-manifold, so that we have a VB-isomorphism of $T(X)$ with $X \times \mathbf{E}$, then the morphism ξ is completely determined by its projection on the second factor, and we are essentially in the situaiton of the preceding paragraph, except for the fact that our vector field is independent of time. In such a product representation, the projection of ξ on the second factor will be called the **local representation** of ξ. It is a C^{p-1}-morphism

$$f\colon X \to \mathbf{E}$$

and $\xi(x) = (x, f(x))$. We shall also say that ξ is **represented by** f **locally** if we work over an open subset U of X over which the tangent bundle admits a trivialisation. We then frequently use ξ itself to denote this local representation.

Let J be an open interval of $\mathbf{R}$. The tangent bundle of J is then $J \times \mathbf{R}$ and we have a canonical section $\imath$ such that $\imath(t) = 1$ for all $t \in J$. We sometimes write $\imath_t$ instead of $\imath(t)$.

By a **curve** in X we mean a morphism (always of class $\geqq 1$ unless otherwise specified)

$$\alpha\colon J \to X$$

from an open interval in $\mathbf{R}$ into X. If $g\colon X \to Y$ is a morphism, then $g \circ \alpha$ is a curve in Y. From a given curve α, we get an induced map on the tangent bundles:

$$\begin{array}{ccc} J \times \mathbf{R} & \xrightarrow{\alpha_*} & T(X) \\ \downarrow & & \downarrow \pi \\ J & \xrightarrow[\alpha]{} & X \end{array}$$

and $\alpha_* \circ \imath$ will be denoted by α' or by $d\alpha/dt$ if we take its value at a point t in J. Thus α' is a curve in $T(X)$, of class C^{p-1} if α is of class C^p. Unless otherwise specified, it is always understood in the sequel that we start with enough differentiability to begin with so that we never end up with maps of class < 1. Thus to be able to take derivatives freely we have to take X and α of class C^p with $p \geqq 2$.

If $g\colon X \to Y$ is a morphism, then

$$(g \circ \alpha)'(t) = g_*\alpha'(t).$$

This follows at once from the functoriality of the tangent bundle and the definitions.

Suppose that J contains 0, and let us consider curves defined on J and such that $\alpha(0)$ is equal to a fixed point x_0. We could say that two such curves α_1, α_2 are **tangent** at 0 if $\alpha_1'(0) = \alpha_2'(0)$. The reader will verify immediately that there is a natural bijection between tangency classes of curves with $\alpha(0) = x_0$ and the tangent space $T_{x_0}(X)$ of X at x_0. The tangent space could therefore have been defined alternatively by taking equivalence classes of curves through the point.

Let ξ be a vector field on X and x_0 a point of X. An **integral curve** for the vector field ξ with **initial condition** x_0, or starting at x_0, is a curve (of class C^{p-1})

$$\alpha: J \to X$$

mapping an open interval J of $\mathbf{R}$ containing 0 into X, such that $\alpha(0) = x_0$ and such that

$$\alpha'(t) = \xi(\alpha(t))$$

for all $t \in J$. Using a local representation of the vector field, we know from the preceding section that integral curves exist locally. The next theorem gives us their global existence and uniqueness.

Theorem 2.1. *Let α_1: $J_1 \to X$ and α_2: $J_2 \to X$ be two integral curves of the vector field ξ on X, with the same initial condition x_0. Then α_1 and α_2 are equal on $J_1 \cap J_2$.*

Proof. Let J^* be the set of points t such that $\alpha_1(t) = \alpha_2(t)$. Then J^* certainly contains a neighborhood of 0 by the local uniqueness theorem. Furthermore, since X is Hausdorff, we see that J^* is closed. We must show that it is open. Let t^* be in J^* and define β_1, β_2 near 0 by

$$\beta_1(t) = \alpha_1(t^* + t),$$
$$\beta_2(t) = \alpha_2(t^* + t).$$

Then β_1 and β_2 are integral curves of ξ with initial condition $\alpha_1(t^*)$ and $\alpha_2(t^*)$ respectively, so by the local uniqueness theorem, β_1 and β_2 agree in a neighborhood of 0 and thus α_1, α_2 agree in a neighborhood of t^*, thereby proving our theorem.

It follows from Theorem 2.1 that the union of the domains of all integral curves of ξ with a given initial condition x_0 is an open interval which we denote by $J(x_0)$. Its end points are denoted by $t^+(x_0)$ and $t^-(x_0)$ respectively. (We do not exclude $+\infty$ and $-\infty$.)

Let $\mathfrak{D}(\xi)$ be the subset of $\mathbf{R} \times X$ consisting of all points (t, x) such that

$$t^-(x) < t < t^+(x).$$

A (global) **flow** for ξ is a mapping

$$\alpha: \mathfrak{D}(\xi) \to X,$$

such that for each $x \in X$, the map α_x: $J(x) \to X$ given by

$$\alpha_x(t) = \alpha(t, x)$$

defined on the open interval $J(x)$ is a morphism and is an integral curve for ξ with initial condition x. When we select a chart at a point x_0 of X, then one sees at once that this definition of flow coincides with the definition we gave locally in the previous section, for the local representation of our vector field.

Given a point $x \in X$ and a number t, we say that tx is **defined** if (t, x) is in the domain of α, and we denote $\alpha(t, x)$ by tx in that case.

Theorem 2.2. *Let ξ be a vector field on X, and α its flows. Let x be a point of X. If t_0 lies in $J(x)$, then*

$$J(t_0x) = J(x) - t_0$$

(translation of $J(x)$ by $-t_0$), and we have for all t in $J(x) - t_0$:

$$t(t_0x) = (t + t_0)x.$$

Proof. Our first assertion follows immediately from the maximality assumption concerning the domains of the integral curves. The second is equivalent to saying that the two curves given by the left-hand side and right-hand side of the last equality are equal. They are both integral curves for the vector field, with initial condition t_0x and must therefore be equal.

In particular, if t_1, t_2 are two numbers such that t_1x is defined and $t_2(t_1x)$ is also defined, then so is $(t_1 + t_2)x$ and they are equal.

Theorem 2.3. *Let ξ be a vector field on X, and x a point of X. Assume that $t^+(x) < \infty$. Given a compact set $A \subset X$, there exists $\epsilon > 0$ such that for all $t > t^+(x) - \epsilon$, the point tx does not lie in A, and similarly for t^-.*

Proof. Suppose such ϵ does not exist. Then we can find a sequence t_n of real numbers approaching $t^+(x)$ from below, such that t_nx lies in A. Since A is compact, taking a subsequence if necessary, we may assume that t_nx converges to a point in A. By the local existence theorem, there exists a neighborhood U of this point y and a number $\delta > 0$ such that $t^+(z) > \delta$ for all $z \in U$. Taking n large, we have

$$t^+(x) < \delta + t_n$$

and $t_n x$ is in U. Then by Theorem 2.2,

$$t^+(x) = t^+(t_n x) + t_n > \delta + t_n > t^+(x)$$

contradiction.

Corollary 2.4. *If X is compact, and ξ is a vector field on X, then*

$$\mathfrak{D}(\xi) = \mathbf{R} \times X.$$

It is also useful to give one other criterion when $\mathfrak{D}(\xi) = \mathbf{R} \times X$, even when X is not compact. Such a criterion must involve some structure stronger than the differentiable structure (essentially a metric of some sort), because we can always dig holes in a compact manifold by taking away a point.

Proposition 2.5. *Let $\mathbf{E}$ be a Banach space, and X an $\mathbf{E}$-manifold. Let ξ be a vector field on X. Assume that there exist numbers $a > 0$ and $K > 0$ such that every point x of X admits a chart (U, φ) at x such that the local representation f of the vector field on this chart is bounded by K, and so is its derivative f'. Assume also that φU contains a ball of radius a around φx. Then $\mathfrak{D}(\xi) = \mathbf{R} \times X$.*

Proof. This follows at once from the global continuation theorem, and the uniformity of Proposition 1.1.

We shall prove finally that $\mathfrak{D}(\xi)$ is open and that α is a morphism.

Theorem 2.6. *Let ξ be a vector field of class C^{p-1} on the C^p-manifold X $(2 \leqq p \leqq \infty)$. Then $\mathfrak{D}(\xi)$ is open in $\mathbf{R} \times X$, and the flow α for ξ is a C^{p-1}-morphism.*

Proof. Let first p be an integer $\geqq 2$. Let $x_0 \in X$. Let J^* be the set of points in $J(x_0)$ for which there exists a number $b > 0$ and an open neighborhood U of x_0 such that $(t - b, t + b)$ U is contained in $\mathfrak{D}(\xi)$, and such that the restriction of the flow α to this product is a C^{p-1}-morphism. Then J^* is open in $J(x_0)$, and certainly contains 0 by the local theorem. We must therefore show that J^* is closed in $J(x_0)$.

Let s be in its closure. By the local theorem, we can select a neighborhood V of $sx_0 = \alpha(s, x_0)$ so that we have a unique local flow

$$\beta\colon J_a \times V \to X$$

for some number $a > 0$, with initial condition $\beta(0, x) = x$ for all $x \in V$, and such that this local flow β is C^{p-1}.

The integral curve with initial condition x_0 is certainly continuous on $J(x_0)$. Thus tx_0 approaches sx_0 as t approaches s. Let V_1 be a given small neighborhood of sx_0 contained in V. By the definition of J^*, we can find an element t_1 in J^* very close to s, and a small number b (compared to a) and a small neighborhood U of x_0 such that α maps the product

$$(t_1 - b, t_1 + b) \times U$$

into V_1, and is C^{p-1} on this product. For $t \in J_a + t_1$ and $x \in U$, we define

$$\varphi(t, x) = \beta(t - t_1, \alpha(t_1, x)).$$

Then $\varphi(t_1, x) = \beta(0, \alpha(t_1, x)) = \alpha(t_1, x)$, and

$$\begin{aligned} D_1\varphi(t, x) D_1\beta(t - t_1, \alpha(t_1, x)) \\ &= \xi(\beta(t - t_1, \alpha(t_1, x)) \\ &= \xi(\varphi(t, x)). \end{aligned}$$

Hence both φ_x, α_x are integral curves for ξ, with the same value at t_1. They coincide on any interval on which they are defined, so that φ_x is a continuation of α_x to a bigger interval containing s. Since α is C^{p-1} on the product $(t_1 - b, t_1 + b) \times U$, we conclude that φ is also C^{p-1} on $(J_a + t_1) \times U$. From this we see that $\mathfrak{D}(\xi)$ is open in $\mathbf{R} \times X$, and that α is of class C^{p-1} on its full domain $\mathfrak{D}(\xi)$. If $p = \infty$, then we can now conclude that α is of class C^r for each positive integer r on $\mathfrak{D}(\xi)$, and hence is C^∞, as desired.

Corollary 2.7. *For each $t \in \mathbf{R}$, the set of $x \in X$ such that (t, x) is contained in the domain $\mathfrak{D}(\xi)$ is open in X.*

Corollary 2.8. *The functions $t^+(x)$ and $t^-(x)$ are upper and lower semicontinuous respectively.*

Theorem 2.9. *Let ξ be a vector field on X and α its flow. Let $\mathfrak{D}_t(\xi)$ be the set of points x of X such that (t, x) lies in $\mathfrak{D}(\xi)$. Then $\mathfrak{D}_t(\xi)$ is open for each $t \in \mathbf{R}$, and α_t is an isomorphism of $\mathfrak{D}_t(\xi)$ onto an open subset of X. In fact, $\alpha_t(\mathfrak{D}_t) = \mathfrak{D}_{-t}$ and $\alpha_t^{-1} = \alpha_{-t}$.*

Proof. Immediate from the preceding theorem.

Corollary 2.10. *If x_0 is a point of X and t is in $J(x_0)$, then there exists an open neighborhood U of x_0 such that t lies in $J(x)$ for all $x \in U$, and* *the map*

$$x \mapsto tx$$

is an isomorphism of U onto an open neighborhood of tx_0.

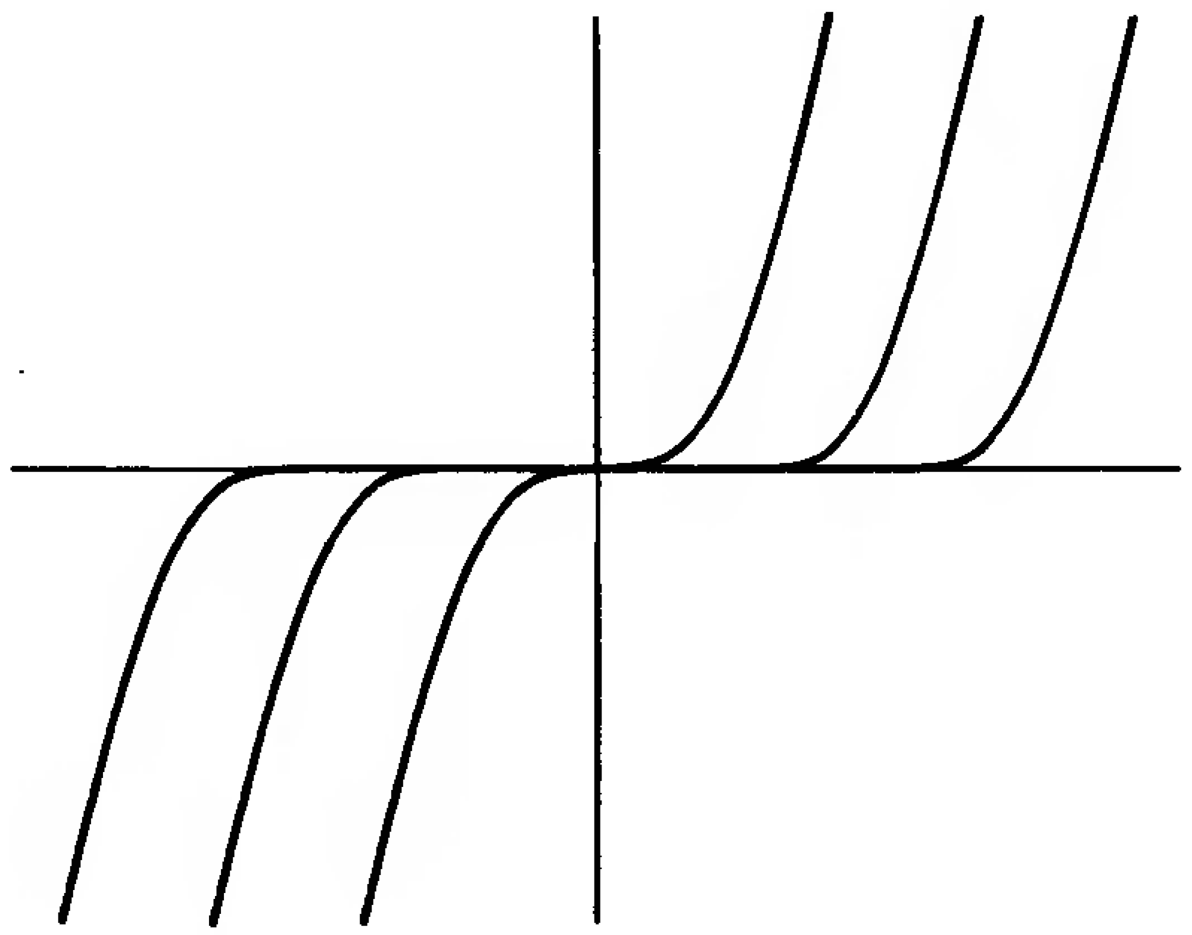
Critical points

Let ξ be a vector field. A **critical point** of ξ is a point x_0 such that $\xi(x_0) = 0$. Critical points play a significant role in the study of vector fields, notably in the Morse theory. We don't go into this here, but just make a few remarks to show at the basic level how they affect the behavior of integral curves.

Proposition 2.11. *If α is an integral curve of a C^1 vector field, ξ, and α passes through a critical point, then α is constant, that is $\alpha(t) = x_0$ for all t.*

Proof. The constant curve through x_0 is an integral curve for the vector field, and the uniqueness theorem shows that it is the only one.

Some smoothness of the vector field in addition to continuity must be assumed for the uniqueness. For instance, the following picture illustrates a situation where the integral curves are not unique. They consist in translations of the curve $y = x^3$ in the plane. The vector field is continuous but not locally Lipschitz.



Proposition 2.12. *Let ξ be a vector field and α an integral curve for ξ. Assume that all $t \geqq 0$ are in the domain of α, and that*

$$\lim_{t\to 0} \alpha(t) = x_1$$

exists. Then x_1 is a critical point for ξ, that is $\xi(x_1) = 0$.

Proof. Selecting t large, we may assume that we are dealing with the local representation f of the vector field near x_1. Then for $t' > t$ large, we have

$$\alpha(t') - \alpha(t) = \int_t^{t'} f(\alpha(u))\, du.$$

Write $f(\alpha(u)) = f(x_1) + g(u)$, where $\lim g(u) = 0$. Then

$$|f(x_1)|\,|t' - t| \leqq |\alpha(t') - \alpha(t)| + |t' - t| \sup|g(u)|,$$

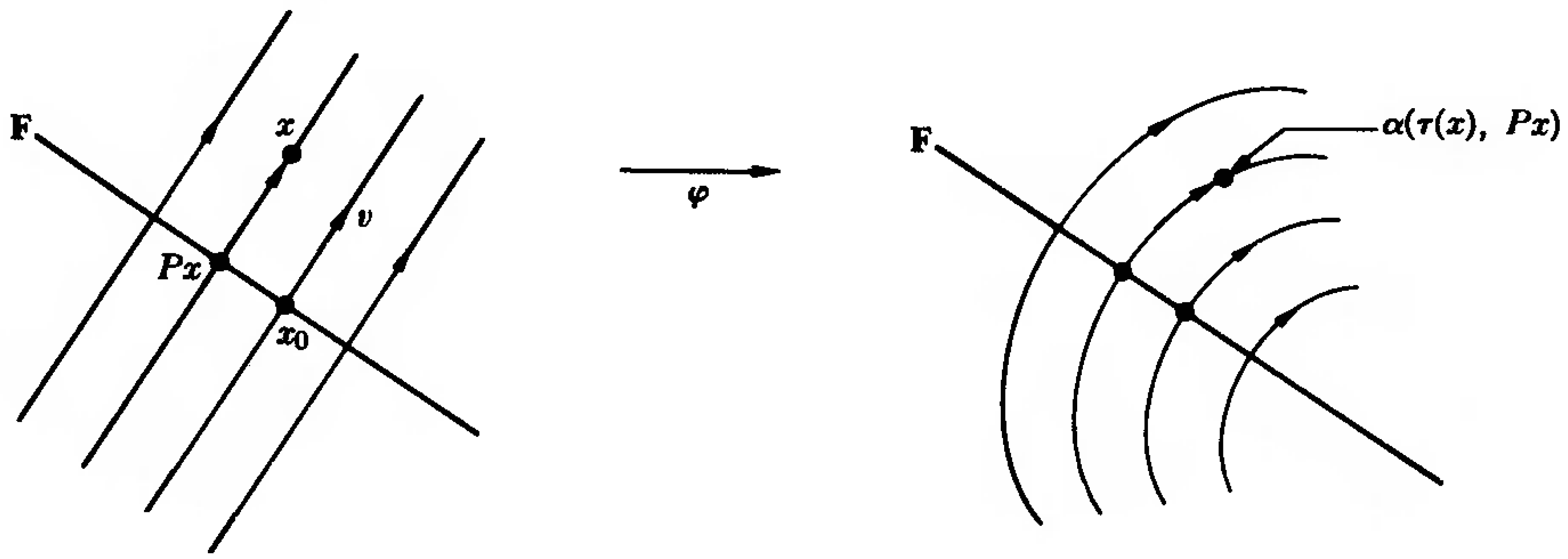
where the sup is taken for u large, and hence for small values of $g(u)$. Dividing by $|t' - t|$ shows that $f(x_1)$ is arbitrarily small, hence equal to 0, as was to be shown.

Proposition 2.13. *Suppose on the other hand that x_0 is not a critical point of the vector field ξ. Then there exists a chart at x_0 such that the local representation of the vector field on this chart is constant.*

Proof. In an arbitrary chart the vector field has a representation as a morphism

$$\xi\colon U \to E$$

near x_0. Let α be its flow. We wish to "straighten out" the integral curves of the vector field according to the next figure.



In other words, let $v = \xi(x_0)$. We want to find a local isomorphism φ at x_0 such that

$$\varphi'(x)v = \xi(\varphi(x)).$$

We inspire ourselves from the picture. Without loss of generality, we may assume that $x_0 = 0$. Let λ be a functional such that $\lambda(v) \neq 0$. We decompose $\mathbf{E}$ as a direct sum

$$\mathbf{E} = \mathbf{F} \oplus \mathbf{R}v,$$

where **F** is the kernel of λ. Let P be the projection on **F**. We can write any x near 0 in the form

$$x = Px + \tau(x)v,$$

where

$$\tau(x) = \frac{\lambda(x)}{\lambda(v)}.$$

We then bend the picture on the left to give the picture on the right using the flow α of ξ, namely we define

$$\varphi(x) = \alpha(\tau(x), Px).$$

This means that starting at Px, instead of going linearly in the direction of v for a time $\tau(x)$, we follow the flow (integral curve) for this amount of time. We find that

$$\varphi'(x) = D_1\alpha(\tau(x), Px)\frac{\lambda}{\lambda(v)} + D_2\alpha(\tau(x), Px)\,P.$$

Hence $\varphi'(0) = \text{id}$, so by the inverse mapping theorem, φ is a local isomorphism at 0. Furthermore, since $Pv = 0$ by definition, we have

$$\varphi'(x)v = D_1\alpha(\tau(x), Px) = \xi(\varphi(x)),$$

thus proving Proposition 2.13.

IV, §3. SPRAYS

Second-order vector fields and differential equations

Let X be a manifold of class C^p with $p \geqq 3$. Then its tangent bundle $T(X)$ is of class C^{p-1}, and the tangent bundle of the tangent bundle $T(T(x))$ is of class C^{p-2}, with $p-2 \geqq 1$.

Let α: $J \to X$ be a curve of class C^q ($q \leqq p$). A **lifting** of α into $T(X)$ is a curve β: $J \to T(X)$ such that $\pi\beta = \alpha$. We shall always deal with $q \geqq 2$ so that a lift will be assumed of class $q - 1 \geqq 1$. Such lifts always exist, for instance the curve α' discussed in the previous section, called the **canonical lifting** of α.

A **second-order** vector field over X is a vector field F on the tangent bundle $T(X)$ (of class C^{p-1}) such that, if π: $TX \to X$ denotes the canonical projection of $T(X)$ on X, then

$$\pi_* \circ F = \text{id.}, \qquad \text{that is} \quad \pi_* F(v) = v \quad \text{for all } v \text{ in } T(X).$$

Observe that the succession of symbols makes sense, because

$$\pi_*\colon TT(X) \to T(X)$$

maps the double tangent bundle into $T(X)$ itself.

A vector field F on TX is a second-order vector field on X if and only if it satisfies the following condition: Each integral curve β of F is equal to the canonical lifting of $\pi\beta$, in other words

$$(\pi\beta)' = \beta.$$

Here, $\pi\beta$ is the canonical projection of β on X, and if we put the argument t, then our formula reads

$$(\pi\beta)'(t) = \beta(t)$$

for all t in the domain of β. The proof is immediate from the definitions, because

$$(\pi\beta)' = \pi_*\beta' = \pi_* \circ F \circ \beta$$

We then use the fact that given a vector $v \in TX$, there is an integral curve $\beta = \beta_v$ with $\beta_v(0) = v$ (initial condition v).

Let α: $J \to X$ be a curve in X, defined on an interval J. We define α to be a **geodesic with respect to** F if the curve

$$\alpha'\colon J \to TX$$

is an integral curve of F. Since $\pi\alpha' = \alpha$, that is α' lies above α in TX, we can express the geodesic condition equivalently by stating that α satisfies the relation

$$\alpha'' = F(\alpha').$$

This relation for curves α in X is called the **second-order differential equation** for the curve α, determined by F. Observe that by definition, if β is an integral curve of F in TX, then $\pi\beta$ is a geodesic for the second order vector field F.

Next we shall give the representation of the second order vector field and of the integral curves in a chart.

Representation in charts

Let U be open in the Banach space $\mathbf{E}$, so that $T(U) = U \times \mathbf{E}$, and $T(T(U)) = (U \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E})$. Then π: $U \times \mathbf{E} \to U$ is simply the projection, and we have a commutative diagram:

$$\begin{array}{ccc} (U \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E}) & \xrightarrow{\pi_*} & U \times \mathbf{E} \\ \downarrow & & \downarrow \\ U \times \mathbf{E} & \xrightarrow[\pi]{} & U \end{array}$$

The map π_* on each fiber $\mathbf{E} \times \mathbf{E}$ is constant, and is simply the projection of $\mathbf{E} \times \mathbf{E}$ on the first factor $\mathbf{E}$, that is

$$\pi_*(x, v, u, w) = (x, u).$$

Any vector field on $U \times \mathbf{E}$ has a local representation

$$f\colon\ U \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}$$

which has therefore two components, $f = (f_1, f_2)$, each f_i mapping $U \times \mathbf{E}$ into $\mathbf{E}$. The next statement describes second order vector fields locally in the chart.

Let U be open in the Banach space $\mathbf{E}$, and let $T(U) = U \times \mathbf{E}$ be the tangent bundle. A C^{p-2}-morphism

$$f\colon\ U \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}$$

is the local representation of a second order vector field on U if and only if

$$f(x, v) = (v, f_2(x, v)).$$

The above statement is merely making explicit the relation $\pi_* F = \mathrm{id}$, in the chart. If we write $f = (f_1, f_2)$, then we see that

$$f_1(x, v) = v.$$

We express the above relations in terms of integral curves as follows. Let $\beta = \beta(t)$ be an integral curve for the vector field F on TX. In the chart, the curve has two components

$$\beta(t) = (x(t), v(t)) \in U \times \mathbf{E}.$$

By definition, if f is the local representation of F, we must have

$$\frac{d\beta}{dt} = \left(\frac{dx}{dt}, \frac{dv}{dt}\right) = f(x, v) = (v, f_2(x, v)).$$

Consequently, our differential equation can be rewritten in the following manner:

$$\frac{dx}{dt} = v(t),$$

(1) $$\frac{d^2x}{dt^2} = \frac{dv}{dt} = f_2\left(x, \frac{dx}{dt}\right),$$

which is of course familiar.

Sprays

We shall be interested in special kinds of second-order differential equations. Before we discuss these, we make a few technical remarks.

Let s be a real number, and π: $E \to X$ be a vector bundle. If v is in E, so in E_x for some x in X, then sv is again in E_x since E_x is a vector space. We write s_E for the mapping of E into itself given by this scalar multiplication. This maping is in fact a VB-morphism, and even a VB-isomorphism if $s \neq 0$. Then

$$T(s_E) = (s_E)_*\colon\ T(E) \to T(E)$$

is the usual induced map on the tangent bundle of E.

Now let $E = TX$ be the tangent bundle itself. Then our map s_{TX} satisfies the property

$$(s_{TX})_* \circ s_{TTX} = s_{TTX} \circ (s_{TX})_*,$$

which follows from the linearity of s_{TX} on each fiber, and can also be seen directly from the representation on charts given below.

We define a **spray** to be a second-order vector field which satisfies the homogeneous quadratic condition:

SPR 1. *For all $s \in \mathbf{R}$ and $v \in T(X)$, we have*

$$F(sv) = (s_{TX}) * sF(v).$$

It is immediate from the conditions defining sprays (second-order vector field satisfying **SPR 1**) that **sprays form a convex set**! Hence if we can exhibit sprays over open subsets of Banach spaces, then we can glue them together by means of partitions of unity, and we obtain at once the following global existence theorem.

Theorem 3.1. *Let X be a manifold of class C^p $(p \geqq 3)$. If X admits partitions of unity, then there exists a spray over X.*

Representations in a chart

Let U be open in $\mathbf{E}$, so that $TU = U \times \mathbf{E}$. Then

$$TTU = (U \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E}),$$

and the representations of s_{TU} and $(s_{TU})_*$ in the chart are given by the maps

$$s_{TU}\colon (x, v) \mapsto (x, sv) \quad \text{and} \quad (s_{TU})_*\colon (x, v, u, w) \mapsto (x, sv, u, sw).$$

Thus

$$s_{TTU} \circ (s_{TU})_*\colon (x, v, u, w) \mapsto (x, sv, su, s^2 w).$$

We may now give the local condition for a second-order vector field F to be a spray.

Proposition 3.2. *In a chart $U \times \mathbf{E}$ for TX, let $f\colon U \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}$ represent F, with $f = (f_1, f_2)$. Then f represents a spray if and only if, for all $s \in \mathbf{R}$ we have*

$$f_2(x, sv) = s^2 f_2(x, v).$$

Proof. The proof follows at once from the definitions and the formula giving the chart representation of $s(s_{TX})_*$.

Thus we see that the condition **SPR 1** (in addition to being a second-order vector field), simply means that f_2 is homogeneous of degree 2 in the variable v. By the remark in Chapter I, §3, it follows that f_2 is a quadratic map in its second variable, and specifically, this quadratic map is given by

$$f_2(x, v) = \tfrac{1}{2} D_2^2 f_2(x, 0)(v, v).$$

Thus the spray is induced by a symmetric bilinear map given at each point x in a chart by

$$(2) \qquad B(x) = \tfrac{1}{2} D_2^2 f_2(x, 0).$$

Conversely, suppose given a morphism

$$U \to L^2_{\text{sym}}(\mathbf{E}, \mathbf{E}) \qquad \text{given by} \qquad x \mapsto B(x)$$

from U into the space of symmetric bilinear maps $\mathbf{E} \times \mathbf{E} \to \mathbf{E}$. Thus for each $v, w \in \mathbf{E}$ the value of $B(x)$ at (v, w) is denoted by $B(x; v, w)$ or $B(x)(v, w)$. Define $f_2(x, v) = B(x; v, v)$. Then f_2 is quadratic in its second variable, and the map f defined by

$$f(x, v) = \big(v, B(x; v, v)\big) = \big(v, f_2(x, v)\big)$$

represents a spray over U. We call B the **symmetric bilinear map associated with the spray**. From the local representations in (1) and (2), we conclude that *a curve α is a geodesic if and only if α satisfies the differential equation*

$$(3) \qquad \alpha''(t) = B_{\alpha(t)}\big(\alpha'(t), \alpha'(t)\big) \qquad \text{for all } t.$$

We recall the trivial fact from linear algebra that the bilinear map B is determined purely algebraically from the quadratic map, by the formula

$$B(v, w) = \tfrac{1}{2}[f_2(v + w) - f_2(v) - f_2(w)].$$

We have suppressed the x from the notation to focus on the relevant second variable v. Thus the quadratic map and the symmetric bilinear map determine each other uniquely.

The above discussion has been local, over an open set U in a Banach space. In Proposition 3.4 and the subsequent discussion of connections, we show how to globalize the bilinear map B intrinsically on the manifold.

Examples. As a trivial special case, we can always take $f_2(x, v) = (v, 0)$ to represent the second component of a spray in the chart.

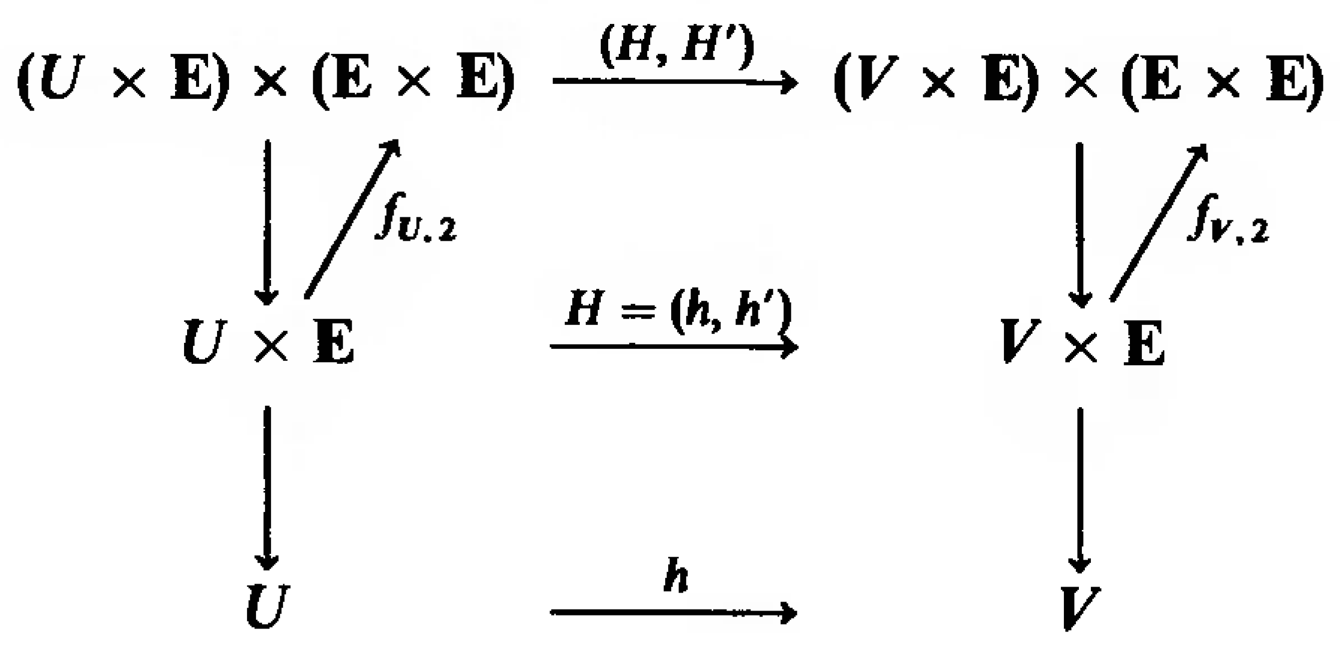
In the chapter on Riemannian metrics, we shall see how to construct a spray in a natural fashion, depending on the metric.

In the chapter on covariant derivatives we show how a spray gives rise to such derivatives.

Next, let us give the transformation rule for a spray under a change of charts, i.e. an isomorphism $h\colon U \to V$. On TU, the map Th is represented by a morphism (its vector component)

$$H\colon U \times \mathbf{E} \to \mathbf{E} \times \mathbf{E} \qquad \text{given by} \qquad H(x, v) = \big(h(x), h'(x)v\big).$$

We then have one further lift to the double tangent bundle TTU, and we may represent the diagram of maps symbolically as follows:



Then the derivative $H'(x, v)$ is given by the Jacobian matrix operating on column vectors ${}^t(u, w)$ with $u, w \in \mathbf{E}$, namely

$$H'(x, v) = \begin{pmatrix} h'(x) & 0 \\ h''(x)v & h'(x) \end{pmatrix} \quad \text{so} \quad H'(x, v)\begin{pmatrix} u \\ w \end{pmatrix} = \begin{pmatrix} h'(x) & 0 \\ h''(x)v & h'(x) \end{pmatrix}\begin{pmatrix} u \\ w \end{pmatrix}.$$

Thus the top map on elements in the diagram is given by

$$(H, H'): (x, v, u, w) \mapsto (h(x), h'(x)v, h'(x)u, h''(x)(u, v) + h'(x)w).$$

For the application, we put $u = v$ because $f_1(x, v) = v$, and $w = f_{U,2}(x, v)$, where f_U and f_V denote the representations of the spray over U and V respectively. It follows that f_U and f_V are related by the formula

$$f_V(h(x), h'(x)v) = (h'(x)v, h''(x)(v, v) + h'(x)f_{U,2}(x, v)).$$

Therefore we obtain:

Proposition 3.3. Change of variable formula for the quadratic part of a spray:

$$f_{V,2}(h(x), h'(x)v) = h''(x)(v, v) + h'(x)f_{U,2}(x, v),$$
$$B_V(h(x);\ h'(x)v, h'(x)w) = h''(x)(v, w) + h'(x)B_U(x;\ v, w).$$

Proposition 3.3 admits a converse:

Proposition 3.4. *Suppose we are given a covering of the manifold X by open sets corresponding to charts $U, V, \ldots$, and for each U we are given a morphism*

$$B_U: U \to L^2_{\text{sym}}(\mathbf{E}, \mathbf{E})$$

which transforms according to the formula of Proposition 3.3 under an isomorphism $h: U \to V$. Then there exists a unique spray whose associated bilinear map in the chart U is given by B_U.

Proof. We leave the verification to the reader.

Remarks. Note that $B_U(x; v, w)$ does not transform like a tensor of type $L^2_{\text{sym}}(\mathbf{E}, \mathbf{E})$, i.e. a section of the bundle $L^2_{\text{sym}}(TX, TX)$. There are several ways of defining the bilinear map B intrinsically. One of them is via second order bundles, or bundles of second order jets, and to extend the terminology we have established previously to such bundles, and even higher order jet bundles involving higher derivatives, as in [Po 62]. Another way will be done below, via connections. For our immediate purposes, it suffices to have the above discussion on second-order differential equations together with Proposition 3.3 and 3.4. Sprays were introduced by Ambrose, Palais, and Singer [APS 60], and I used them (as recommended by Palais) in the earliest version [La 62]. In [Lo 69] the bilinear map B_U is expressed in terms of second order jets. The basics of differential topology and geometry were being established in the early sixties. Cf. the bibliographical notes from [Lo 69] at the end of his first chapter.

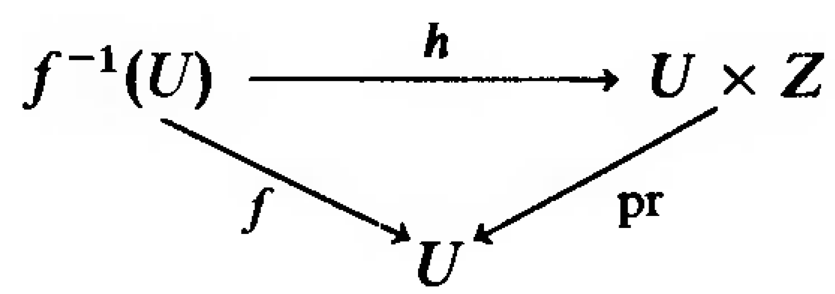
Connections

We now show how to define the bilinear map B intrinsically and directly.

Matters will be clearer if we start with an arbitrary vector bundle

$$p: E \to X$$

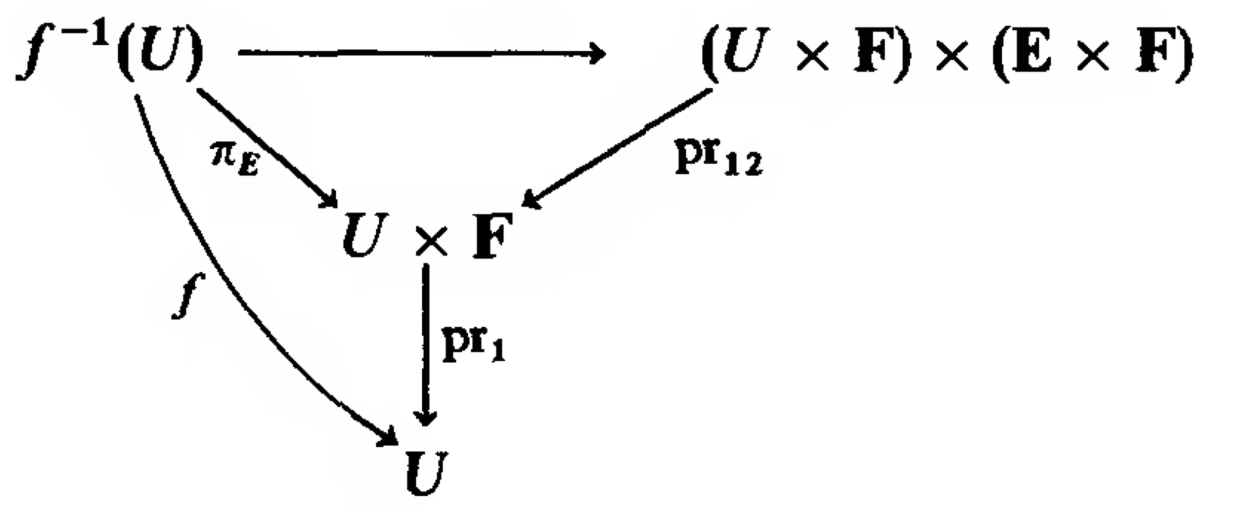
over a manifold X. As it happens we also need the notion of a fiber bundle when the fibers are not necessarily vector spaces, so don't have a linear structure. Let $f: Y \to X$ be a morphism. We say that f (or Y over X) is a **fiber bundle** if f is surjective, and if each point x of X has an open neighborhood U, and there is some manifold Z and an isomorphism $h: f^{-1}(U) \to U \times Z$ such that the following diagram is commutative:



Thus locally, $f: Y \to X$ looks like the projection from a product space. The reason why we need a fiber bundle is that the tangent bundle

$$\pi_E: TE \to E$$

is a vector bundle over E, but the composite $f = p \circ \pi_E: TE \to X$ is only a fiber bundle over X, a fact which is obvious by picking trivializations in charts. Indeed, if U is a chart in X, and if $U \times \mathbf{F} \to U$ is a vector bundle chart for E, with fiber $\mathbf{F}$, and $Y = TE$, then we have a natural isomorphism of fiber bundles over U:



Note that U being a chart in X implies that $U \times \mathbf{E} \to U$ is a vector bundle chart for the tangent bundle TU over U.

The tangent bundle TE has two natural maps making it a vector bundle:

$$\pi_E\colon TE \to E \text{ is a vector bundle over } E;$$

$$T(p)\colon TE \to TX \text{ is a vector bundle over } TX.$$

Therefore we have a natural morphism of fiber bundle (not vector bundle) over X:

$$(\pi_E, T(p))\colon TE \to E \oplus TX \qquad \text{given by} \qquad W \mapsto (\pi_E W, T(p)W)$$

for $W \in TE$. If $W \in T_e E$ with $e \in E_x$, then $\pi_E W \in E_x$ and $T(p)W \in T_x X$.

After these preliminaries, we define a **connection** to be a morphism of fiber bundles over X, from the direct sum $E \oplus TX$ into TE:

$$H\colon E \oplus TX \to TE$$

such that

$$(\pi_E, T(p)) \circ H = \mathrm{id}_{E \oplus TX},$$

and such that H is bilinear, in other words $H_x\colon E_x \oplus T_x X \to TE$ is bilinear.

Consider a chart U as in the above diagram, so

$$TU = U \times \mathbf{E} \qquad \text{and} \qquad T(U \times \mathbf{F}) = (U \times \mathbf{F}) \times (\mathbf{E} \times \mathbf{F}).$$

Then our map H has a coordinate representation

$$H(x, e, v) = (x, e, H_1(x, e, v), H_2(x, e, v)) \qquad \text{for } e \in \mathbf{F} \text{ and } v \in \mathbf{E}.$$

The fact that $(\pi_E, T(p)) \circ H = \mathrm{id}_{E \oplus TX}$ implies at once that $H_1(x, e, v) = v$.

The bilinearity condition implies that for fixed x, the map

$$(e, v) \mapsto H_2(x, e, v)$$

is bilinear as a map $\mathbf{F} \times \mathbf{E} \to \mathbf{E}$. We shall therefore denote this map by $B(x)$, and we write in the chart

$$H(x, e, v) = (x, e, v, B(x)(e, v)) \qquad \text{or also} \qquad (x, e, v, B(x, e, v)).$$

Now take the special case when $E = TX$. We say that the connection is **symmetric** if the bilinear map B is symmetric. Suppose this is the case. We may define the corresponding quadratic map $TX \to TTX$ by letting $f_2(x, v) = B(x, v, v)$. Globally, this amounts to defining a morphism

$$F\colon TX \to TTX \qquad \text{such that} \qquad F = H \circ \text{diagonal}$$

where the diagonal is taken in $TX \oplus TX$, in each fiber. Thus

$$F(v) = H(v, v) \qquad \text{for } v \in T_x X.$$

Then F is a vector field on TX, and the condition $(\pi_*, \pi_*) \circ H = \mathrm{id}$ on $TX \oplus TX$ implies that F is a second-order vector field on X, in other words, F defines a spray. It is obvious that all sprays can be obtained in this fashion. Thus we have shown how to describe geometrically the bilinear map associated with a spray.

Going back to the general case of a vector bundle E unrelated to TX, we note that the image of a connection H is a vector subbundle over E. Let V denote the kernel of the map $T(p)\colon TE \to TX$. We leave it to the reader to verify in charts that V is a vector subbundle of TE over E, and that the image of H is a complementary subbundle. One calls V the **vertical subbundle**, canonically defined, and one calls H the **horizontal subbundle** determined by the connection. See Chapter X, §4. Also note that Eliasson [El 67] introduced connections in Banach manifolds, with a view to applications to manifolds of mappings. Cf. Kobayashi [Ko 57], Dombrowski [Do 68], and Besse [Be 78] for more basic material on connections.

IV, §4. THE FLOW OF A SPRAY AND THE EXPONENTIAL MAP

The condition we have taken to define a spray is equivalent to other conditions concerning the integral curves of the second-order vector field F. We shall list these conditions systematically. We shall use the following relation. If $\alpha\colon J \to X$ is a curve, and α_1 is the curve defined by

$\alpha_1(t) = \alpha(st)$, then

$$\alpha_1'(t) = s\alpha'(st),$$

this being the chain rule for differentiation.

If v is a vector in TX, let β_v be the unique integral curve of F with initial condition v (i.e. such that $\beta_v(0) = v$). In the next three conditions, the sentence should begin with "for each v in TX".

SPR 2. *A number t is in the domain of β_{sv} if and only if st is in the domain of β_v and then*

$$\beta_{sv}(t) = s\beta_v(st).$$

SPR 3. *If s, t are numbers, st is in the domain of β_v if and only if s is in the domain of β_{tv}, and then*

$$\pi\beta_{tv}(s) = \pi\beta_v(st).$$

SPR 4. *A number t is in the domain of β_v if and only if 1 is in the domain of β_{tv}, and then*

$$\pi\beta_v(t) = \pi\beta_{tv}(1).$$

We shall now prove the equivalence between all four conditions.

Assume **SPR 1**, and let s be fixed. For all t such that st is in the domain of β_v, the curve $\beta_v(st)$ is defined and we have

$$\frac{d}{dt}(s\beta_v(st)) = s_* s\beta_v'(st) = s_* sF(\beta_v(st)) = F(s\beta_v(st)).$$

Hence the curve $s\beta_v(st)$ is an integral curve for F, with initial condition $s\beta_v(0) = sv$. By uniqueness we must have

$$s\beta_v(st) = \beta_{sv}(t).$$

This proves **SPR 2**.

Assume **SPR 2**. Since β_v is an integral curve of F for each v, with initial condition v, we have by definition

$$\beta_{sv}'(0) = F(sv).$$

Using our assumption, we also have

$$\beta_{sv}'(t) = \frac{d}{dt}(s\beta_v(st)) = s_* s\beta_v'(st).$$

Put $t = 0$. Then **SPR 1** follows because β_{sv} and β_v are integral curves of F with initial conditions sv and v respectively.

It is obvious that **SPR 2** implies **SPR 3**. Conversely, assume **SPR 3**. To prove **SPR 2**, we have

$$\beta_{sv}(t) = (\pi\beta_{sv})'(t) = \frac{d}{dt}\pi\beta_v(st) = s(\pi\beta_v)'(st) = s\beta_v(st),$$

which proves **SPR 2**.

Assume **SPR 4**. Then st is in the domain of β_r if and only if 1 is in the domain of β_{stv}, and s is in the domain of β_{tv} if and only if 1 is in the domain of β_{stv}. This proves the first assertion of **SPR 3**, and again by **SPR 4**, assuming these relations, we get **SPR 3**.

It is similarly clear that **SPR 3** implies **SPR 4**.

Next we consider further properties of the integral curves of a spray. Let F be a spray on X. As above, we let β_v be the integral curve with initial condition v. Let $\mathfrak{D}$ be the set of vectors v in $T(X)$ such that β_v is defined at least on the interval [0, 1]. We know from Corollary 2.7 that $\mathfrak{D}$ is an open set in $T(X)$, and by Theorem 2.6 the map

$$v \mapsto \beta_v(1)$$

is a morphism of $\mathfrak{D}$ into $T(X)$. We now define the **exponential** map

$$\exp\colon \mathfrak{D} \to X$$

to be

$$\exp(v) = \pi\beta_v(1).$$

Then exp is a C^{p-2}-morphism. We also call $\mathfrak{D}$ the **domain of the exponential map** (**associated with** F).

If $x \in X$ and 0_x denotes the zero vector in T_x, then from **SPR 1**, taking $s = 0$, we see that $F(0_x) = 0$. Hence

$$\exp(0_x) = x.$$

Thus our exponential map coincides with π on the zero cross section, and so induces an isomorphism of the cross section onto X. It will be convenient to denote the zero cross section of a vector bundle E over X by $\zeta_E(X)$ or simply ζX if the reference to E is clear. Here, E is the tangent bundle.

We denote by $\exp_x$ the restriction of exp to the tangent space T_x. Thus

$$\exp_x\colon T_x \to X.$$

Theorem 4.1. *Let X be a manifold and F a spray on X. Then*

$$\exp_x\colon T_x \to X$$

induces a local isomorphism at 0_x, and in fact $(\exp_x)_*$ is the identity at 0_x.

Proof. We prove the second assertion first because the main assertion follows from it by the inverse mapping theorem. Furthermore, since T_x is a vector space, it suffices to determine the derivative of $\exp_x$ on rays, in other words, to determine the derivative with respect to t of a curve $\exp_x(tv)$. This is done by using **SPR 3**, and we find

$$\frac{d}{dt}\pi\beta_{tv} = \beta_{tv}.$$

Evaluating this at $t = 0$ and taking into account that β_w has w as initial condition for any w gives us

$$(\exp_x)_*(0_x) = \mathrm{id}.$$

This concludes the proof of Theorem 4.1.

Helgason gave a general formula for the differential of the exponential map on analytic manifolds [He 61], reproduced in [He 78], Chapter I, Theorem 6.5. We shall study the differential of the exponential map in connection with Jacobi fields, in Chapter IX, §2.

Next we describe all geodesics.

Proposition 4.2. *The images of straight segments through the origin in T_x, under the exponential map $\exp_x$, are geodesics. In other words, if $v \in T_x$ and we let*

$$\alpha(v, t) = \alpha_v(t) = \exp_x(tv),$$

then α_v is a geodesic. Conversely, let $\alpha\colon J \to X$ be a C^2 geodesic defined on an interval J containing 0, and such that $\alpha(0) = x$. Let $\alpha'(0) = v$. Then $\alpha(t) = \exp_x(tv)$.

Proof. The first statement by definition means that α_v' is an integral curve of the spray F. Indeed, by the **SPR** conditions, we know that

$$\alpha(v, t) = \alpha_v(t) = \pi\beta_{tv}(1) = \pi\beta_v(t),$$

and $(\pi\beta_v)' = \beta_v$ is indeed an integral curve of the spray. Thus our assertion that the curves $t \mapsto \exp(tv)$ are geodesics is obvious from the definition of the exponential map and the **SPR** conditions.

Conversely, given a geodesic $\alpha\colon J \to X$, by definition α' satisfies the differential equation

$$\alpha''(t) = F(\alpha'(t)).$$

The two curves $t \mapsto \alpha(t)$ and $t \mapsto \exp_x(tv)$ satisfy the same differential equation and have the same initial conditions, so the two curves are equal. This proves the second statement and concludes the proof of the proposition.

Remark. From the theorem, we note that a C^1 curve in X is a geodesic if and only if, after a linear reparametrization of its interval of definition, it is simply $t \mapsto \exp_x(tv)$ for some x and some v.

We call the map $(v, t) \mapsto \alpha(v, t)$ the **geodesic flow** on X. It is defined on an open subset of $TX \times \mathbf{R}$, with $\alpha(v, 0) = x$ if $v \in T_xX$. Note that since $\pi(s\beta_v(t)) = \pi\beta_v(t)$ for $s \in \mathbf{R}$, we obtain from **SPR 2** the property

$$\alpha(sv, t) = \alpha(v, st)$$

for the geodesic flow. Precisely, t is in the domain of α_{sv} if and only if st is in the domain of α_v, and in that case the formula holds. As a slightly more precise version of Theorem 4.1 in this light, we obtain:

Corollary 4.3. *Let F be a spray on X, and let $x_0 \in X$. There exists an open neighborhood U of x_0, and an open neighborhood V of 0_{x_0} in TX satisfying the following condition. For every $x \in U$ and $v \in V \cap T_xX$, there exists a unique geodesic*

$$\alpha_v\colon (-2, 2) \to X$$

such that

$$\alpha_v(0) = x \quad \textit{and} \quad \alpha_v'(0) = v.$$

Observe that in a chart, we may pick V as a product

$$V = U \times V_2(0) \subset U \times \mathbf{E}$$

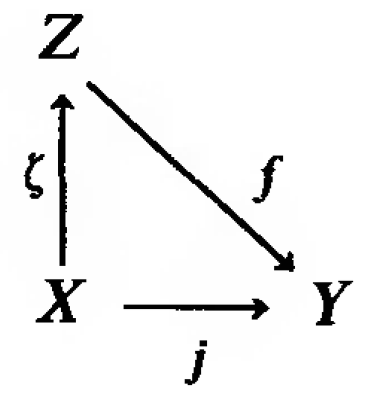
where $V_2(0)$ is a neighborhood of 0 in $\mathbf{E}$. Then the geodesic flow is defined on $U \times V_2(0) \times J$, where $J = (-2, 2)$. We picked $(-2, 2)$ for concreteness. What we really want is that 0 and 1 lie in the interval. Any bounded interval J containing 0 and 1 could have been selected in the statement of the corollary. Then of course, U and V (or $V_2(0)$) depend on J.

IV, §5. EXISTENCE OF TUBULAR NEIGHBORHOODS

Let X be a submanifold of a manifold Y. A **tubular neighborhood** of X in Y consists of a vector bundle $\pi\colon E \to X$ over X, an open neighborhood Z of the zero section $\zeta_E X$ in E, and an isomorphism

$$f\colon Z \to U$$

of Z onto an open set in Y containing X, which commutes with ζ:



We shall call f the **tubular map** and Z or its image $f(Z)$ the corresponding **tube** (in E or Y respectively). The bottom map j is simply the inclusion. We could obviously assume that it is an embedding and define tubular neighborhoods for embeddings in the same way. We shall say that our tubular neighborhood is **total** if $Z = E$. In this section, we investigate conditions under which such neighborhoods exist. We shall consider the uniqueness problem in the next section.

Theorem 5.1. *Let Y be of class C^p ($p \geqq 3$) and admit partitions of unity. Let X be a closed submanifold. Then there exists a tubular neighborhood of X in Y, of class C^{p-2}.*

Proof. Consider the exact sequence of tangent bundles:

$$0 \to T(X) \to T(Y)|X \to N(X) \to 0.$$

We know that this sequence splits, and thus there exists some splitting

$$T(Y)|X = T(X) \oplus N(X)$$

where $N(X)$ may be identified with a subbundle of $T(Y)|X$. Following Palais, we construct a spray ξ on $T(Y)$ using Theorem 3.1 and obtain the corresponding exponential map. We shall use its restriction to $N(X)$, denoted by $\exp|N$. Thus

$$\exp|N\colon \mathfrak{D} \cap N(X) \to Y.$$

We contend that this map is a local isomorphism. To prove this, we may work locally. Corresponding to the submanifold, we have a product

decomposition $U = U_1 \times U_2$, with $X = U_1 \times 0$. If U is open in $\mathbf{E}$, then we may take U_1, U_2 open in $\mathbf{F}_1$, $\mathbf{F}_2$ respectively. Then the injection of $N(X)$ in $T(Y)|X$ may be represented locally by an exact sequence

$$0 \to U_1 \times \mathbf{F}_2 \xrightarrow{\varphi} U_1 \times \mathbf{F}_1 \times \mathbf{F}_2,$$

and the inclusion of $T(Y)|X$ in $T(Y)$ is simply the inclusion

$$U_1 \times \mathbf{F}_1 \times \mathbf{F}_2 \to U_1 \times U_2 \times \mathbf{F}_1 \times \mathbf{F}_2.$$

We work at the point $(x_1, 0)$ in $U_1 \times \mathbf{F}_2$. We must compute the derivative of the composite map

$$U_1 \times \mathbf{F}_2 \xrightarrow{\varphi} U_1 \times U_2 \times \mathbf{F}_1 \times \mathbf{F}_2 \xrightarrow{\exp} Y$$

at $(x_1, 0)$. We can do this by the formula for the partial derivatives. Since the exponential map coincides with the projection on the zero cross section, its "horizontal" partial derivative is the identity. By Theorem 4.1 we know that its "vertical" derivative is also the identity. Let

$$\psi = (\exp) \circ \bar{\varphi}$$

(where $\bar{\varphi}$ is simply φ followed by the inclusion). Then for any vector (w_1, w_2) in $\mathbf{F}_1 \times \mathbf{F}_2$ we get

$$D\psi(x_1, 0) \cdot (w_1, w_2) = (w_1, 0) + \varphi_{x_1}(w_2),$$

where φ_{x_1} is the linear map given by φ on the fiber over x_1. By hypothesis, we know that $\mathbf{F}_1 \times \mathbf{F}_2$ is the direct sum of $\mathbf{F}_1 \times 0$ and of the image of φ_{x_1}. This proves that $D\psi(x_1, 0)$ is a toplinear isomorphism, and in fact proves that **the exponential map restricted to a normal bundle is a local isomorphism** on the zero cross section.

We have thus shown that there exists a vector bundle $E \to X$, an open neighborhood Z of the zero section in E, and a mapping $f\colon Z \to Y$ which, for each x in ζ_E, is a local isomorphism at x. We must show that Z can be shrunk so that f restricts to an isomorphism. To do this we follow Godement ([God 58], p. 150). We can find a locally finite open covering of X by open sets U_i in Y such that, for each i we have inverse isomorphisms

$$f_i\colon Z_i \to U_i \qquad \text{and} \qquad g_i\colon U_i \to Z_i$$

between U_i and open sets Z_i in Z, such that each Z_i contains a point x of X, such that f_i, g_i are the identity on X (viewed as a subset of both Z and Y) and such that f_i is the restriction of f to Z_i. We now find a locally finite covering $\{V_i\}$ of X by open sets of Y such that $\overline{V}_i \subset U_i$, and let $V = \bigcup V_i$. We let W be the subset of elements $y \in V$ such that, if y lies in an intersection $\overline{V}_i \cap \overline{V}_j$, then $g_i(y) = g_j(y)$. Then W certainly contains X. We contend that W contains an open subset containing X.

Let $x \in X$. There exists an open neighborhood G_x of x in Y *which meets only a finite number of* $\overline{V}_i$, *say* $\overline{V}_{i_1}, \dots, \overline{V}_{i_r}$. *Taking* G_x *small enough, we can assume that* x lies in each one of these, and that G_x is contained in each one of the sets $\overline{U}_{i_1}, \dots, \overline{U}_{i_r}$. Since x lies in each $\overline{V}_{i_1}, \dots, \overline{V}_{i_r}$, it is contained in $U_{i_1}, \dots, U_{i_r}$ and our maps $g_{i_1}, \dots, g_{i_r}$ take the same value at x, namely x itself. Using the fact that $f_{i_1}, \dots, f_{i_r}$ are restrictions of f, we see at once that our finite number of maps $g_{i_1}, \dots, g_{i_r}$ must agree on G_x if we take G_x small enough.

Let G be the union of the G_x. Then G is open, and we can define a map

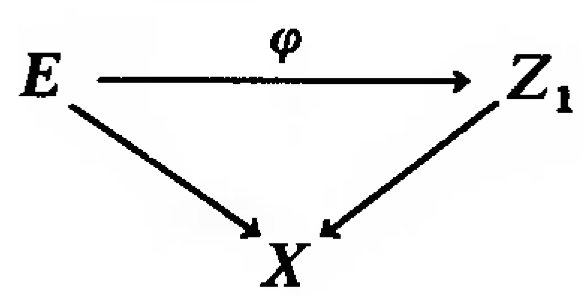
$$g\colon G \to g(G) \subset Z$$

by taking g equal to g_i on $G \cap V_i$. Then $g(G)$ is open in Z, and the restriction of f to $g(G)$ is an inverse for g. This proves that f, g are inverse isomorphisms on G and $g(G)$, and concludes the proof of the theorem.

A vector bundle $E \to X$ will be said to be **compressible** if, given an open neighborhood Z of the zero section, there exists an isomorphism

$$\varphi\colon E \to Z_1$$

of E with an open subset Z_1 of Z containing the zero section, which commutes with the projection on X:



It is clear that if a bundle is compressible, and if we have a tubular neighborhood defined on Z, then we can get a total tubular neighborhood defined on E. We shall see in the chapter on Riemannian metrics that certain types of vector bundles are compressible (Hilbert bundles, assuming that the base manifold admits partitions of unity).

IV, §6. UNIQUENESS OF TUBULAR NEIGHBORHOODS

Let X, Y be two manifolds, and $F\colon \mathbf{R} \times X \to Y$ a morphism. We shall say that F is an **isotopy** (of embeddings) if it satisfies the following conditions. First, for each $t \in \mathbf{R}$, the map F_t given by $F_t(x) = F(t, x)$ is an embedding. Second, there exist numbers $t_0 < t_1$ such that $F_t = F_{t_0}$ for all

$t \leqq t_0$ and $F_{t_1} = F_t$ for all $t \geqq t_1$. We then say that the interval $[t_0, t_1]$ is a **proper domain** for the isotopy, and the constant embeddings on the left and right will also be denoted by $F_{-\infty}$ and $F_{+\infty}$ respectively. We say that two embeddings $f\colon X \to Y$ and $g\colon X \to Y$ are **isotopic** if there exists an isotopy F_t as above such that $f = F_{t_0}$ and $g = F_{t_1}$ (notation as above). We write $f \approx g$ for f isotopic to g.

Using translations of intervals, and multiplication by scalars, we can always transform an isotopy to a new one whose proper domain is contained in the interval $(0, 1)$. Furthermore, the relation of isotopy between embeddings is an equivalence relation. It is obviously symmetric and reflexive, and for transitivity, suppose $f \approx g$ and $g \approx h$. We can choose the ranges of these isotopies so that the first one ends and stays constant at g before the second starts moving. Thus it is clear how to compose isotopies in this case.

If $s_0 < s_1$ are two numbers, and $\sigma\colon \mathbf{R} \to \mathbf{R}$ is a function (morphism) such that $\sigma(s) = t_0$ for $s \leqq s_0$ and $\sigma(s) = t_1$ for $s \geqq s_1$, and σ is monotone increasing, then from a given isotopy F_t we obtain another one, $G_t = F_{\sigma(t)}$. Such a function σ can be used to smooth out a piece of isotopy given only on a closed interval.

Remark. We shall frequently use the following trivial fact: If $f_t\colon X \to Y$ is an isotopy, and if $g\colon X_1 \to X$ and $h\colon Y \to Y_1$ are two embeddings, then the composite map

$$hf_tg\colon X_1 \to Y_1$$

is also an isotopy.

Let Y be a manifold and X a submanifold. Let $\pi\colon E \to X$ be a vector bundle, and Z an open neighborhood of the zero section. An isotopy $f_t\colon Z \to Y$ of open embeddings such that each f_t is a tubular neighborhood of X will be called an **isotopy of tubular neighborhoods**. In what follows, the domain will usually be all of E.

Proposition 6.1. *Let X be a manifold. Let $\pi\colon E \to X$ and $\pi_1\colon E_1 \to X$ be two vector bundles over X. Let*

$$f\colon E \to E_1$$

be a tubular neighborhood of X in E_1 (identifying X with its zero section in E_1). Then there exists an isotopy

$$f_t\colon E \to E_1$$

with proper domain $[0, 1]$ *such that $f_1 = f$ and f_0 is a* VB-*isomorphism. (If f, π, π_1 are of class C^p then f_t can be chosen of class C^{p-1}.)*

Proof. We define F by the formula

$$F_t(e) = t^{-1}f(te)$$

for $t \neq 0$ and $e \in E$. Then F_t is an embedding since it is composed of embeddings (the scalar multiplications by t, t^{-1} are in fact VB-isomorphism).

We must investigate what happens at $t = 0$.

Given $e \in E$, we find an open neighborhood U_1 of πe over which E_1 admits a trivialization $U_1 \times \mathbf{E}_1$. We then find a still smaller open neighborhood U of πe and an open ball B around 0 in the typical fiber $\mathbf{E}$ of E such that E admits a trivialization $U \times \mathbf{E}$ over U, and such that the representation $\bar{f}$ of f on $U \times B$ (contained in $U \times \mathbf{E}$) maps $U \times B$ into $U_1 \times \mathbf{E}_1$. This is possible by continuity. On $U \times B$ we can represent $\bar{f}$ by two morphisms,

$$\bar{f}(x, v) = \big(\varphi(x, v), \psi(x, v)\big)$$

and $\varphi(x, 0) = x$ while $\psi(x, 0) = 0$. Observe that for all t sufficiently small, te is contained in $U \times B$ (in the local representation).

We can represent F_t locally on $U \times B$ as the mapping

$$\bar{F}_t(x, v) = \big(\varphi(x, tv), t^{-1}\psi(x, tv)\big).$$

The map φ is then a morphism in the three variables x, v, and t even at $t = 0$. The second component of $\bar{F}_t$ can be written

$$t^{-1}\psi(x, tv) = t^{-1}\int_0^1 D_2\psi(x, stv)\cdot(tv)\,ds$$

and thus t^{-1} cancels t to yield simply

$$\int_0^1 D_2\psi(x, stv)\cdot v\,ds.$$

This is a morphism in t, even at $t = 0$. Furthermore, for $t = 0$, we obtain

$$\bar{F}_0(x, v) = \big(x, D_2\psi(x, 0)v\big).$$

Since f was originally assumed to be an embedding, it follows that $D_2\psi(x, 0)$ is a toplinear isomorphism, and therefore F_0 is a VB-isomorphism. To get our isotopy in standard form, we can use a function $\sigma\colon \mathbf{R} \to \mathbf{R}$ such that $\sigma(t) = 0$ for $t \leqq 0$ and $\sigma(t) = 1$ for $t \geqq 1$, and σ is monotone increasing. This proves our proposition.

Theorem 6.2. *Let X be a submanifold of Y. Let*

$$\pi\colon E \to X \quad \text{and} \quad \pi_1\colon E_1 \to X$$

be two vector bundles, and assume that E is compressible. Let $f\colon E \to Y$ and $g\colon E_1 \to Y$ be two tubular neighborhoods of X in Y. Then there exists a C^{p-1}-isotopy

$$f_t\colon E \to Y$$

of tubular neighborhoods with proper domain $[0, 1]$ *and a* VB-*isomorphism $\lambda\colon E \to E_1$ such that $f_1 = f$ and $f_0 = g\lambda$.*

Proof. We observe that $f(E)$ and $g(E_1)$ are open neighborhoods of X in Y. Let $U = f^{-1}\big(f(E) \cap g(E_1)\big)$ and let $\varphi\colon E \to U$ be a compression. Let ψ be the composite map

$$E \xrightarrow{\varphi} U \xrightarrow{f|U} Y$$

$\psi = (f|U) \circ \varphi$. Then ψ is a tubular neighborhood, and $\psi(E)$ is contained in $g(E_1)$. Therefore $g^{-1}\psi\colon E \to E_1$ is a tubular neighborhood of the same type considered in the previous proposition. There exists an isotopy of tubular neighborhoods of X:

$$G_t\colon E \to E_1$$

such that $G_1 = g^{-1}\psi$ and G_0 is a VB-isomorphism. Considering the isotopy gG_t, we find an isotopy of tubular neighborhoods

$$\psi_t\colon E \to Y$$

such that $\psi_1 = \psi$ and $\psi_0 = g\omega$ where $\omega\colon E \to E_1$ is a VB-isomorphism. We have thus shown that ψ and $g\omega$ are isotopic (by an isotopy of tubular neighborhoods). Similarly, we see that ψ and $f\mu$ are isotopic for some VB-isomorphism

$$\mu\colon E \to E.$$

Consequently, adjusting the proper domains of our isotopies suitably, we get an isotopy of tubular neighborhoods going from $g\omega$ to $f\mu$, say F_t. Then $F_t\mu^{-1}$ will give us the desired isotopy from $g\omega\mu^{-1}$ to f, and we can put $\lambda = \omega\mu^{-1}$ to conclude the proof.

(By the way, the uniqueness proof did not use the existence theorem for differential equations.)

CHAPTER V

Operations on Vector Fields and Differential Forms

If $E \to X$ is a vector bundle, then it is of considerable interest to investigate the special operation derived from the functor "multilinear alternating forms." Applying it to the tangent bundle, we call the sections of our new bundle differential forms. One can define formally certain relations between functions, vector fields, and differential forms which lie at the foundations of differential and Riemannian geometry. We shall give the basic system surrounding such forms. In order to have at least one application, we discuss the fundamental 2-form, and in the next chapter connect it with Riemannian metrics in order to construct canonically the spray associated with such a metric.

We assume throughout that our manifolds are Hausdorff, and sufficiently differentiable so that all of our statements make sense.

V, §1. VECTOR FIELDS, DIFFERENTIAL OPERATORS, BRACKETS

Let X be a manifold of class C^p and φ a function defined on an open set U, that is a morphism

$$\varphi\colon U \to \mathbf{R}.$$

Let ξ be a vector field of class C^{p-1}. Recall that

$$T_x\varphi\colon T_x(U) \to T_x(\mathbf{R}) = \mathbf{R}$$

is a continuous linear map. With it, we shall define a new function to be denoted by $\xi\varphi$ or $\xi \cdot \varphi$, or $\xi(\varphi)$. (There will be no confusion with this notation and composition of mappings.)

Proposition 1.1. *There exists a unique function $\xi\varphi$ on U of class C^{p-1} such that*

$$(\xi\varphi)(x) = (T_x\varphi)\xi(x).$$

If U is open in the Banach space $\mathbf{E}$ *and ξ denotes the local representation of the vector field on U, then*

$$(\xi\varphi)(x) = \varphi'(x)\xi(x).$$

Proof. The first formula certainly defines a mapping of U into $\mathbf{R}$. The local formula defines a C^{p-1}-morphism on U. It follows at once from the definitions that the first formula expresses invariantly in terms of the tangent bundle the same mapping as the second. Thus it allows us to define $\xi\varphi$ as a morphism globally, as desired.

Let Fu^p denote the ring of functions (of class C^p). Then our operation $\varphi \mapsto \xi\varphi$ gives rise to a linear map

$$\partial_\xi\colon \mathrm{Fu}^p(U) \to \mathrm{Fu}^{p-1}(U), \quad \text{defined by} \quad \partial_\xi\varphi = \xi\varphi.$$

A mapping

$$\partial\colon R \to S$$

from a ring R into an R-algebra S is called a **derivation** if it satisfies the usual formalism: Linearity, and $\partial(ab) = a\partial(b) + \partial(a)b$.

Proposition 1.2. *Let X be a manifold and U open in X. Let ξ be a vector field over X. If $\partial_\xi = 0$, then $\xi(x) = 0$ for all $x \in U$. Each ∂_ξ is a derivation of* $\mathrm{Fu}^p(U)$ *into* $\mathrm{Fu}^{p-1}(U)$.

Proof. Suppose $\xi(x) \neq 0$ for some x. We work with the local representations, and take φ to be a continuous linear map of $\mathbf{E}$ into $\mathbf{R}$ such that $\varphi(\xi(x)) \neq 0$, by Hahn–Banach. Then $\varphi'(y) = \varphi$ for all $y \in U$, and we see that $\varphi'(x)\xi(x) \neq 0$, thus proving the first assertion. The second is obvious from the local formula.

From Proposition 1.2 we deduce that if two vector fields induce the same differential operator on the functions, then they are equal.

Given two vector fields ξ, η on X, we shall now define a new vector field $[\xi, \eta]$, called their **bracket product**.

Proposition 1.3. *Let ξ, η be two vector fields of class C^{p-1} on X. Then there exists a unique vector field $[\xi, \eta]$ of class C^{p-2} such that for each open set U and function φ on U we have*

$$[\xi, \eta]\varphi = \xi(\eta(\varphi)) - \eta(\xi(\varphi)).$$

If U is open in $\mathbf{E}$ and ξ, η are the local representations of the vector fields, then $[\xi, \eta]$ is given by the local formula

$$[\xi, \eta]\varphi(x) = \varphi'(x)\big(\eta'(x)\xi(x) - \xi'(x)\eta(x)\big).$$

Thus the local representation of $[\xi, \eta]$ is given by

$$[\xi, \eta](x) = \eta'(x)\xi(x) - \xi'(x)\eta(x).$$

Proof. By Proposition 1.2, any vector field having the desired effect on functions is uniquely determined. We check that the local formula gives us this effect locally. Differentiating formally, we have (using the law for the derivative of a product):

$$\begin{aligned}(\eta\varphi)'\xi - (\xi\varphi)'\eta &= (\varphi'\eta)'\xi - (\varphi'\xi)\eta \\ &= \varphi'\eta'\xi + \varphi''\eta\xi - \varphi'\xi'\eta - \varphi''\xi\eta.\end{aligned}$$

The terms involving φ'' must be understood correctly. For instance, the first such term at a point x is simply $\varphi''(x)\big(\eta(x), \xi(x)\big)$ remembering that $\varphi''(x)$ is a bilinear map, and can thus be evaluated at the two vectors $\eta(x)$ and $\xi(x)$. However, we know that $\varphi''(x)$ is symmetric. Hence the two terms involving the second derivative of φ cancel, and give us our formula.

Corollary 1.4. *The bracket $[\xi, \eta]$ is bilinear in both arguments, we have $[\xi, \eta] = -[\eta, \xi]$, and Jacobi's identity*

$$[\xi, [\eta, \zeta]] = [[\xi, \eta], \zeta] + [\eta, [\xi, \zeta]].$$

In other words, for each ξ the map $\eta \mapsto [\xi, \eta]$ is a derivation with respect to the Lie product $(\eta, \zeta) \mapsto [\eta, \zeta]$.

If φ is a function, then

$$[\xi, \varphi\eta] = (\xi\varphi)\eta + \varphi[\xi, \eta], \quad \textit{and} \quad [\varphi\xi, \eta] = \varphi[\xi, \eta] - (\eta\varphi)\xi.$$

Proof. The first two assertions are obvious. The third comes from the definition of the bracket. We apply the vector field on the left of the equality to a function φ. All the terms cancel out (the reader will write it out as well or better than the author). The last two formulas are immediate.

We make some comments concerning the functoriality of vector fields. Let

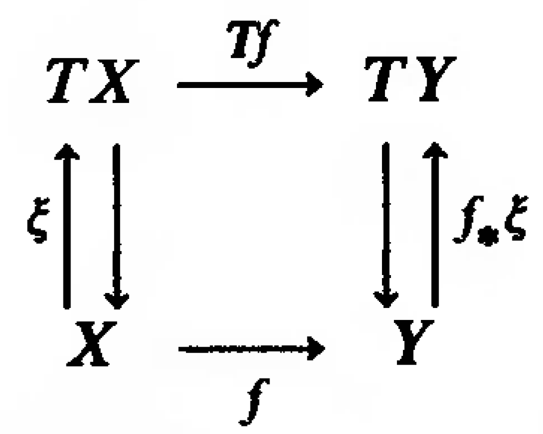
$$f\colon X \to Y$$

be an isomorphism. Let ξ be a vector field over X. Then we obtain an

induced vector field $f_*\xi$ over Y, defined by the formula

$$(f_*\xi)(f(x)) = Tf(\xi(x)).$$

It is the vector field making the following diagram commutative:



We shall also write f^* for $(f^{-1})_*$ when applied to a vector field. Thus we have the formulas

$$\boxed{f_*\xi = Tf \circ \xi \circ f^{-1}} \quad \text{and} \quad \boxed{f^*\xi = Tf^{-1} \circ \xi \circ f.}$$

If f is not an isomorphism, then one cannot in general define the direct or inverse image of a vector field as done above. However, let ξ be a vector field over X, and let η be a vector field over Y. If for each $x \in X$ we have

$$Tf(\xi(x)) = \eta(f(x)),$$

then we shall say that f maps ξ into η, or that ξ and η are ***f*-related**. If this is the case, then we may denote by $f_*\xi$ the map from $f(X)$ into TY defined by the above formula.

Let ξ_1, ξ_2 be vector fields over X, and let η_1, η_2 be vector fields over Y. If ξ_i is f-related to η_i for $i = 1, 2$ then as maps on $f(X)$ we have

$$\boxed{f_*[\xi_1, \xi_2] = [\eta_1, \eta_2].}$$

We may write suggestively the formula in the form

$$\boxed{f_*[\xi_1, \xi_2] = [f_*\xi_1, f_*\xi_2].}$$

Of course, this is meaningless in general, since $f_*\xi_1$ may not be a vector field on Y. When f is an isomorphism, then it is a correct formulation of the other formula. In any case, it suggests the correct formula.

To prove the formula, we work with the local representations, when

$X = U$ is open in $\mathbf{E}$, and $Y = V$ is open in $\mathbf{F}$. Then ξ_i, η_i are maps of U, V into the spaces $\mathbf{E}$, $\mathbf{F}$ respectively. For $x \in X$ we have

$$(f_*[\xi_1, \xi_2])(x) = f'(x)(\xi_2'(x)\xi_1(x) - \xi_1'(x)\xi_2(x)).$$

On the other hand, by assumption, we have

$$\eta_i(f(x)) = f'(x)\xi_i(x),$$

so that

$$\begin{aligned}[\eta_1, \eta_2](f(x)) &= \eta_2'(f(x))\eta_1(f(x)) - \eta_1'(f(x))\eta_2(f(x)) \\ &= \eta_2'(f(x))f'(x)\xi_1(x) - \eta_1'(f(x))f'(x)\xi_2(x) \\ &= (\eta_2 \circ f)'(x)\xi_1(x) - (\eta_1 \circ f)'(x)\xi_2(x) \\ &= f''(x) \cdot \xi_2(x) \cdot \xi_1(x) + f'(x)\xi_2'(x)\xi_1(x) \\ &\quad - f''(x) \cdot \xi_1(x) \cdot \xi_2(x) - f'(x)\xi_1'(x)\xi_2(x).\end{aligned}$$

Since $f''(x)$ is symmetric, two terms cancel, and the remaining two terms give the same value as $(f_*[\xi_1, \xi_2])(x)$, as was to be shown.

The bracket between vector fields gives an infinitesimal criterion for commutativity in various contexts. We give here one theorem of a general nature as an example of this phenomenon.

Theorem 1.5. *Let ξ, η be vector fields on X, and assume that $[\xi, \eta] = 0$. Let α and β be the flows for ξ and η respectively. Then for real values t, s we have*

$$\alpha_t \circ \beta_s = \beta_s \circ \alpha_t.$$

Or in other words, for any $x \in X$ we have

$$\alpha(t, \beta(s, x)) = \beta(s, \alpha(t, x)),$$

in the sense that if for some value of t a value of s is in the domain of one of these expressions, then it is in the domain of the other and the two expressions are equal.

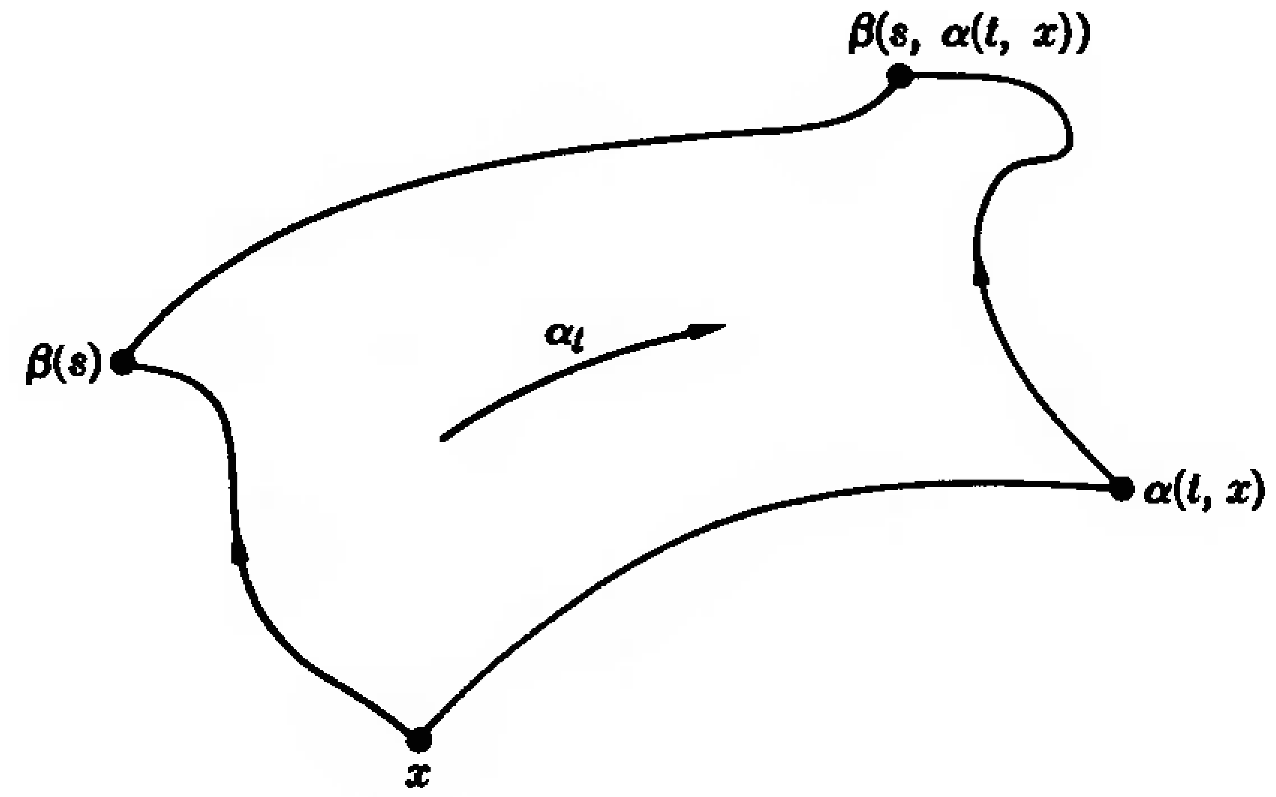
Proof. For a fixed value of t, the two curves in s given by the right- and left-hand side of the last formula have the same initial condition, namely $\alpha_t(x)$. The curve on the right

$$s \mapsto \beta(s, \alpha(t, x))$$

is by definition the integral curve of η. The curve on the left

$$s \mapsto \alpha(t, \beta(s, x))$$

is the image under α_t of the integral curve for η having initial condition x. Since x is fixed, let us denote $\beta(s, x)$ simply by $\beta(s)$. What we must show is that the two curves on the right and on the left satisfy the same differential equation.



In the above figure, we see that the flow α_t shoves the curve on the left to the curve on the right. We must compute the tangent vectors to the curve on the right. We have

$$\begin{aligned}\frac{d}{ds}(\alpha_t(\beta(s))) &= D_2\alpha(t, \beta(s))\beta'(s) \\ &= D_2\alpha(t, \beta(s))\eta(\beta(s)).\end{aligned}$$

Now fix s, and denote this last expression by $F(t)$. We must show that if

$$G(t) = \eta(\alpha(t, \beta(s))),$$

then

$$F(t) = G(t).$$

We have trivially $F(0) = G(0)$, in other words the curves F and G have the same initial condition. On the other hand,

$$F'(t) = \xi'(\alpha(t, \beta(s)))D_2\alpha(t, \beta(s))\eta(\beta(s))$$

and

$$\begin{aligned}G'(t) &= \eta'(\alpha(t, \beta(s)))\xi(\alpha(t, \beta(s))) \\ &= \xi'(\alpha(t, \beta(s)))\eta(\alpha(t, \beta(s))) \quad (\text{because } [\xi, \eta] = 0).\end{aligned}$$

Hence we see that our two curves F and G satisfy the same differential equation, whence they are equal. This proves our theorem.

Vector fields ξ, η such that $[\xi, \eta] = 0$ are said to **commute**. One can generalize the process of straightening out vector fields to a finite number of commuting vector fields, using the same method of proof, using Theorem 1.5. As another application, one can prove that if the Lie algebra of a connected Lie group is commutative, then the group is commutative. Cf. the section on Lie groups.

V, §2. LIE DERIVATIVE

Let λ be a differentiable functor on Banach spaces. For convenience, take λ to be covariant and in one variable. What we shall say in the rest of this section would hold in the same way (with slightly more involved notation) if λ had several variables and were covariant in some and contravariant in others.

Given a manifold X, we can take $\lambda(T(X))$. It is a vector bundle over X, which we denote by $T_\lambda(X)$ as in Chapter III. Its sections $\Gamma_\lambda(X)$ are the tensor fields of type λ.

Let ξ be a vector field on X, and U open in X. It is then possible to associate with ξ a map

$$\mathscr{L}_\xi \colon \Gamma_\lambda(U) \to \Gamma_\lambda(U)$$

(with a loss of two derivatives). This is done as follows.

Given a point x of U and a local flow α for ξ at x, we have for each t sufficiently small a local isomorphism α_t in a neighborhood of our point x. Recall that locally, $\alpha_t^{-1} = \alpha_{-t}$. If η is a tensor field of type λ, then the composite mapping $\eta \circ \alpha_t$ has its range in $T_\lambda(X)$. Finally, we can take the tangent map $T(\alpha_{-t}) = (\alpha_{-t})_*$ to return to $T_\lambda(X)$ in the fiber above x. We thus obtain a composite map

$$F(t, x) = (\alpha_{-t})_* \circ \eta \circ \alpha_t(x) = (\alpha_t^* \eta)(x),$$

which is a morphism, locally at x. We take its derivative with respect to t and evaluate it at 0. After looking at the situation locally in a trivialization of $T(X)$ and $T_\lambda(X)$ at x, one sees that the map one obtains gives a section of $T_\lambda(U)$, that is a tensor field of type λ over U. This is our map $\mathscr{L}_\xi$. To summarize,

$$\mathscr{L}_\xi \eta = \frac{d}{dt}\bigg|_{t=0} (\alpha_{-t})_* \circ \eta \circ \alpha_t.$$

This map $\mathscr{L}_\xi$ is called the **Lie derivative**. We shall determine the Lie derivative on functions and on vector fields in terms of notions already discussed.

First let φ be a function. Then by the general definition, the Lie derivative of this function with respect to the vector field ξ with flow α is defined to be

$$\mathscr{L}_\xi \varphi(x) = \lim_{t \to 0} \frac{1}{t} \left[\varphi(\alpha(t, x)) - \varphi(x)\right],$$

or in other words,

$$\mathscr{L}_\xi \varphi = \frac{d}{dt}(\alpha_t^* \varphi)\bigg|_{t=0}.$$

Our assertion is then that

$$\boxed{\mathscr{L}_\xi \varphi = \xi \varphi.}$$

To prove this, let

$$F(t) = \varphi(\alpha(t, x)).$$

Then

$$\begin{aligned} F'(t) &= \varphi'(\alpha(t, x)) D_1 \alpha(t, x) \\ &= \varphi'(\alpha(t, x)) \xi(\alpha(t, x)), \end{aligned}$$

because α is a flow for ξ. Using the initial condition at $t = 0$, we find that

$$F'(0) = \varphi'(x) \xi(x),$$

which is precisely the value of $\xi\varphi$ at x, thus proving our assertion.

If ξ, η are vector fields, then

$$\boxed{\mathscr{L}_\xi \eta = [\xi, \eta].}$$

As before, let α be a flow for ξ. The Lie derivative is given by

$$\mathscr{L}_\xi \eta = \frac{d}{dt}(\alpha_t^* \eta)\bigg|_{t=0}.$$

Letting ξ and η denote the local representations of the vector fields, we note that the local representation of $(\alpha_t^* \eta)(x)$ is given by

$$(\alpha_t^* \eta)(x) = F(t) = D_2 \alpha(-t, x) \eta(\alpha(t, x)).$$

We must therefore compute $F'(t)$, and then $F'(0)$. Using the chain rule, the formula for the derivative of a product, and the differential equation satisfied by $D_2\alpha$, we obtain

$$\begin{aligned} F'(t) &= -D_1D_2\alpha(-t, x)\eta(\alpha(t, x)) + D_2\alpha(-t, x)\eta'(\alpha(t, x))D_1\alpha(t, x) \\ &= -\xi'(\alpha(-t, x))D_2\alpha(-t, x)\eta(\alpha(t, x)) + D_2\alpha(-t, x)\eta'(\alpha(t, x)). \end{aligned}$$

Putting $t = 0$ proves our formula, taking into account the initial conditions

$$\alpha(0, x) = x \quad \text{and} \quad D_2\alpha(0, x) = \text{id}.$$

V, §3. EXTERIOR DERIVATIVE

Let X be a manifold. The functor L_a^r (r-multilinear continuous alternating forms) extends to arbitrary vector bundles, and in particular, to the tangent bundle of X. A **differential form** of degree r, or simply an **r-form** on X, is a section of $L_a^r(T(X))$, that is a tensor field of type L_a^r. If X is of class C^p, forms will be assumed to be of a suitable class C^s with $1 \leqq s \leqq p - 1$. The set of differential forms of degree r will be denoted by $\mathscr{A}^r(X)$ ($\mathscr{A}$ for alternating). It is not only a vector space over $\mathbf{R}$ but a module over the ring of functions on X (of the appropriate order of differentiability). If ω is an r-form, then $\omega(x)$ is an element of $L_a^r(T_x(X))$, and is thus an r-multilinear alternating form of $T_x(X)$ into $\mathbf{R}$. We sometimes denote $\omega(x)$ by ω_x.

Suppose U is open in the Banach space $\mathbf{E}$. Then $L_a^r(T(U))$ is equal to $U \times L_a^r(\mathbf{E})$ and a differential form is entirely described by the projection on the second factor, which we call its **local representation**, following our general system (Chapter III, §4). Such a local representation is therefore a morphism

$$\omega\colon\ U \to L_a^r(\mathbf{E}).$$

Let ω be in $L_a^r(\mathbf{E})$ and $v_1, \ldots, v_r$ elements of $\mathbf{E}$. We denote the value $\omega(v_1, \ldots, v_r)$ also by

$$\langle \omega,\ v_1 \times \cdots \times v_r \rangle.$$

Similarly, let $\xi_1, \ldots, \xi_r$ be vector fields on an open set U, and let ω be an r-form on X. We denote by

$$\langle \omega,\ \xi_1 \times \cdots \times \xi_r \rangle$$

the mapping from U into $\mathbf{R}$ whose value at a point x in U is

$$\langle \omega(x),\ \xi_1(x) \times \cdots \times \xi_r(x) \rangle.$$

Looking at the situation locally on an open set U such that $T(U)$ is trivial, we see at once that this mapping is a morphism (i.e. a function on U) of the same degree of differentiability as ω and the ξ_i.

Proposition 3.1. *Let x_0 be a point of X and ω an r-form on X. If*

$$\langle \omega,\ \xi_1 \times \cdots \times \xi_r \rangle(x_0)$$

is equal to 0 for all vector fields $\xi_1, \ldots, \xi_r$ at x_0 (i.e. defined on some neighborhood of x_0), then $\omega(x_0) = 0$.

Proof. Considering things locally in terms of their local representations, we see that if $\omega(x_0)$ is not 0, then it does not vanish at some r-tuple of vectors $(v_1, \ldots, v_r)$. We can take vector fields at x_0 which take on these values at x_0 and from this our assertion is obvious.

It is convenient to agree that a differential form of degree 0 is a function. In the next proposition, we describe the exterior derivative of an r-form, and it is convenient to describe this situation separately in the case of functions.

Therefore let $f\colon X \to \mathbf{R}$ be a function. For each $x \in X$, the tangent map

$$T_xf\colon\ T_x(X) \to T_{f(x)}(\mathbf{R}) = \mathbf{R}$$

is a continuous linear map, and looking at local representations shows at once that the collection of such maps defines a 1-form which will be denoted by df. Furthermore, from the definition of the operation of vector fields on functions, it is clear that df is the unique 1-form such that for every vector field ξ we have

$$\langle df, \xi \rangle = \xi f.$$

To extend the definition of d to forms of higher degree, we recall that if

$$\omega\colon\ U \to L_a^r(\mathbf{E})$$

is the local representation of an r-form over an open set U of $\mathbf{E}$, then for each x in U,

$$\omega'(x)\colon\ \mathbf{E} \to L_a^r(\mathbf{E})$$

is a continuous linear map. Applied to a vector v in $\mathbf{E}$, it therefore gives rise to an r-form on $\mathbf{E}$.

Proposition 3.2. *Let ω be an r-form of class C^{p-1} on X. Then there exists a unique $(r+1)$-form $d\omega$ on X of class C^{p-2} such that, for any* *open set U of X and vector fields $\xi_0,\dots,\xi_r$ on U we have*

$$\langle d\omega, \xi_0 \times \cdots \times \xi_r\rangle$$
$$= \sum_{i=0}^{r}(-1)^i\xi_i\langle\omega, \xi_0 \times \cdots \times \hat{\xi}_i \times \cdots \times \xi_r\rangle$$
$$+ \sum_{i<j}(-1)^{i+j}\langle\omega, [\xi_i, \xi_j] \times \xi_0 \times \cdots \times \hat{\xi}_i \times \cdots \times \hat{\xi}_j \times \cdots \times \xi_r\rangle.$$

If furthermore U is open in **E** *and $\omega, \xi_0,\dots,\xi_r$ are the local representations of the form and the vector fields respectively, then at a point x the value of the expression above is equal to*

$$\sum_{i=0}^{r}(-1)^i\langle\omega'(x)\xi_i(x), \xi_0(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \times \xi_r(x)\rangle.$$

Proof. As before, we observe that the local formula defines a differential form. If we can prove that it gives the same thing as the first formulas, which is expressed invariantly, then we can globalize it, and we are done. Let us denote by S_1 and S_2 the two sums occurring in the invariant expression, and let L be the local expression. We must show that $S_1 + S_2 = L$. We consider S_1, and apply the definition of ξ_i operating on a function locally, as in Proposition 1.1, at a point x. We obtain

$$S_1 = \sum_{i=0}^{r}(-1)^i\langle\omega, \xi_0 \times \cdots \times \hat{\xi}_i \times \cdots \times \xi_r\rangle'(x)\xi_i(x).$$

The derivative is perhaps best computed by going back to the definition. Applying this definition directly, and discarding second order terms, we find that S_1 is equal to

$$\sum(-1)^i\langle\omega'(x)\xi_i(x), \xi_0(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \times \xi_r(x)\rangle$$
$$+\sum_i\sum_{i<j}(-1)^i\langle\omega(x), \xi_0(x)\times\cdots\times\xi_j'(x)\xi_i(x)\times\cdots\times\widehat{\xi_i(x)}\times\cdots\times\xi_r(x)\rangle$$
$$+\sum_i\sum_{j<i}\langle\omega(x), \xi_0(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \times \xi_j'(x)\xi_i(x) \times \cdots \times \xi_r(x)\rangle.$$

Of these there sums, the first one is the local formula L. As for the other two, permuting j and i in the first, and moving the term $\xi_j'(x)\xi_i(x)$ to the first position, we see that they combine to give (symbolically)

$$-\sum_i\sum_{i<j}(-1)^{i+j}\langle\omega, (\xi_j'\xi_i - \xi_i'\xi_j) \times \xi_0 \times \cdots \times \hat{\xi}_i \times \cdots \times \hat{\xi}_j \times \cdots \times \xi_r\rangle$$

(evaluated at x). Using Proposition 1.3, we see that this combination is equal to $-S_2$. This proves that $S_1 + S_2 = L$, as desired.

We call $d\omega$ the **exterior derivative** of ω. Leaving out the order of differentiability for simplicity, we see that d is an **R**-linear map

$$d\colon \mathscr{A}^r(X) \to \mathscr{A}^{r+1}(X).$$

We now look into the multiplicative properties of d with respect to the wedge product.

Let ω, ψ be continuous multilinear alternating forms of degree r and s respectively on the Banach space **E**. In multilinear algebra, one defines their **wedge product** as an $(r+s)$-continuous multilinear alternating form, by the formula

$$(\omega \wedge \psi)(v_1,\dots,v_{r+s}) = \frac{1}{r!\,s!}\sum\epsilon(\sigma)\omega(v_{\sigma1},\dots,v_{\sigma r})\psi(v_{\sigma(r+1)},\dots,v_{\sigma(r+s)})$$

the sum being taken over all permutations σ of $(1,\dots,r+s)$. This definition extends at once to differential forms on a manifold, if we view it as giving the value for $\omega \wedge \psi$ at a point x. The v_i are then elements of the tangent space T_x, and considering the local representations shows at once that the wedge product so defined gives a morphism of the manifold X into $L_a^{r+s}(T(X))$, and is therefore a differential form.

Remark. The coefficient $1/r!\,s!$ is not universally taken to define the wedge product. Some people, e.g. [He 78] and [KoN 63], take $1/(r+s)!$, which causes constants to appear later. I have taken the same factor as [AbM 78] and [GHL 87/93]. I recommend that the reader check out the case with $r = s = 1$ so $r+s=2$ to see how a factor $\frac{1}{2}$ comes in. With either convention, the wedge product between forms is associative, so with some care, one can carry out a consistent theory with either convention. I leave the proof of associativity to the reader. It follows by induction that if $\omega_1,\dots,\omega_m$ are forms of degrees $r_1,\dots,r_m$ respectively, and $r = r_1 + \cdots + r_m$, then

$$(\omega_1 \wedge \cdots \wedge \omega_m)(v_1,\dots,v_r) = \frac{1}{r_1!\cdots r_m!}\sum_\sigma\epsilon(\sigma)\Omega_\sigma,$$

where

$$\Omega_\sigma = \omega_1(v_{\sigma1},\dots,v_{\sigma r_1})\omega_2(v_{\sigma(r_1+1)},\dots,v_{\sigma(r_1+r_2)})\dots\omega_m(v_{\sigma(r-r_m+1)},\dots,v_{\sigma r}),$$

and where the sum is taken over all permutations of $(1,\dots,r)$.

If we regard functions on X as differential forms of degree 0, then the ordinary product of a function by a differential form can be viewed as the wedge product. Thus if f is a function and ω a differential form, then

$$f\omega = f \wedge \omega.$$

(The form on the left has the value $f(x)\omega(x)$ at x.)

The next proposition gives us more formulas concerning differential forms.

Proposition 3.3. *Let ω, ψ be differential forms on X. Then*

EXD 1. $d(\omega \wedge \psi) = d\omega \wedge \psi + (-1)^{\deg(\omega)}\omega \wedge d\psi$.

EXD 2. $dd\omega = 0$ *(with enough differentiability, say $p \geqq 4$).*

Proof. This is a simple formal exercise in the use of the local formula for the local representation of the exterior derivative. We leave it to the reader.

When the manifold is finite dimensional, then one can give a local representation for differential forms and the exterior derivative in terms of local coordinates, which are especially useful in integration which fits the notation better. We shall therefore carry out this local formulation in full. It dates back to Cartan [Ca 28]. There is in addition a theoretical point which needs clarifying. We shall use at first the wedge $\bigwedge$ in two senses. One sense is defined as above, giving rise to Proposition 3.3. Another sense will come from Theorem A. We shall comment on their relation after Theorem B.

We recall first two simple results from linear (or rather multilinear) algebra. We use the notation $\mathbf{E}^{(r)} = \mathbf{E} \times \mathbf{E} \times \cdots \times \mathbf{E}$, r times.

Theorem A. *Let $\mathbf{E}$ be a finite dimensional vector space over the reals of dimension n. For each positive integer r with $1 \leqq r \leqq n$ there exists a vector space $\bigwedge^r \mathbf{E}$ and a multilinear alternating map*

$$\mathbf{E}^{(r)} \to \bigwedge^r \mathbf{E}$$

denoted by $(u_1, \ldots, u_r) \mapsto u_1 \wedge \cdots \wedge u_r$, having the following property: If $\{v_1, \ldots, v_n\}$ is a basis of $\mathbf{E}$, then the elements

$$\{v_{i_1} \wedge \cdots \wedge v_{i_r}\}, \qquad i_1 < i_2 < \cdots < i_r,$$

form a basis of $\bigwedge^r \mathbf{E}$.

We recall that **alternating** means that $u_1 \wedge \cdots \wedge u_r = 0$ if $u_i = u_j$ for some $i \neq j$. We call $\bigwedge^r \mathbf{E}$ the r-th **alternating** product (or **exterior** product) on $\mathbf{E}$. If $r = 0$, we define $\bigwedge^0 \mathbf{E} = \mathbf{R}$. Elements of $\bigwedge^r \mathbf{E}$ which can be written in the form $u_1 \wedge \cdots \wedge u_r$ are called **decomposable**. Such elements generate $\bigwedge^r \mathbf{E}$. If $r > \dim E$, we define $\bigwedge^r \mathbf{E} = \{0\}$.

Theorem B. *For each pair of positive integers (r, s), there exists a unique product (bilinear map)*

$$\bigwedge^r \mathbf{E} \times \bigwedge^s \mathbf{E} \to \bigwedge^{r+s} \mathbf{E}$$

such that if $u_1, \ldots, u_r, w_1, \ldots, w_s \in \mathbf{E}$ then

$$(u_1 \wedge \cdots \wedge u_r) \times (w_1 \wedge \cdots \wedge w_s) \mapsto u_1 \wedge \cdots \wedge u_r \wedge w_1 \wedge \cdots \wedge w_s.$$

This product is associative.

The proofs for these two statements can be found, for instance, in my *Linear Algebra.*

Let $\mathbf{E}^\vee$ be the dual space, $\mathbf{E}^\vee = L(\mathbf{E}, \mathbf{R})$. If $\mathbf{E} = \mathbf{R}^n$ and $\lambda_1, \ldots, \lambda_n$ are the coordinate functions, then each λ_i is an element of the dual space, and in fact $\{\lambda_1, \ldots, \lambda_n\}$ is a basis of this dual space. Let $\mathbf{E} = \mathbf{R}^n$. There is an isomorphism

$$\boxed{\bigwedge^r \mathbf{E}^\vee \xrightarrow{\approx} L_a^r(\mathbf{E}, \mathbf{R})}$$

given in the following manner. If $g_1, \ldots, g_r \in \mathbf{E}^\vee$ and $v_1, \ldots, v_r \in \mathbf{E}$, then the value

$$\det(g_i(v_j))$$

is multilinear alternating both as a function of $(g_1, \ldots, g_r)$ and $(v_1, \ldots, v_r)$. Thus it induces a pairing

$$\bigwedge^r \mathbf{E}^\vee \times \mathbf{E}^r \to \mathbf{R}$$

and a map

$$\bigwedge^r \mathbf{E}^\vee \to L_a^r(\mathbf{E}, \mathbf{R}).$$

This map is the isomorphism mentioned above. Using bases, it is easy to verify that it is an isomorphism (at the level of elementary algebra).

Thus in the finite dimensional case, we may identify $L_a^r(\mathbf{E}, \mathbf{R})$ with the alternating product $\bigwedge^r \mathbf{E}^\vee$, and consequently we may view the local representation of a differential form of degree r to be a map

$$\omega\colon U \to \bigwedge^r \mathbf{E}^\vee$$

from U into the rth alternating product of $\mathbf{E}^\vee$. We say that the form is of class C^p if the map is of class C^p. (We view $\bigwedge^r \mathbf{E}^\vee$ as a normed vector space, using any norm. It does not matter which, since all norms on a finite dimensional vector space are equivalent.) The wedge product as we gave it, valid in the infinite dimensional case, is compatible with the wedge product and the isomorphism of $\bigwedge^r \mathbf{E}$ with $L^r_a(\mathbf{E}, \mathbf{R})$ given above. If we had taken a different convention for the wedge product of alternating forms, then a constant would have appeared in front of the above determinant to establish the above identification (e.g. the constant $\frac{1}{2}$ in the 2×2 case).

Since $\{\lambda_1, \dots, \lambda_n\}$ is a basis of $\mathbf{E}^\vee$, we can express each differential form in terms of its coordinate functions with respect to the basis

$$\{\lambda_{i_1} \wedge \cdots \wedge \lambda_{i_r}\}, \qquad (i_1 < \cdots < i_r),$$

namely for each $x \in U$ we have

$$\omega(x) = \sum_{(i)} f_{i_1 \cdots i_r}(x) \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_r},$$

where $f_{(i)} = f_{i_1 \cdots i_r}$ is a function on U. Each such function has the same order of differentiability as ω. We call the preceding expression the **standard form** of ω. We say that a form is **decomposable** if it can be written as just one term $f(x)\lambda_{i_1} \wedge \cdots \wedge \lambda_{i_r}$. Every differential form is a sum of decomposable ones.

We agree to the convention that functions are differential forms of degree 0.

As before, the differential forms on U of given degree r form a vector space, denoted by $\mathscr{A}^r(U)$.

Let $\mathbf{E} = \mathbf{R}^n$. Let f be a function on U. For each $x \in U$ the derivative

$$f'(x)\colon \mathbf{R}^n \to \mathbf{R}$$

is a linear map, and thus an element of the dual space. Thus

$$f'\colon U \to \mathbf{E}^\vee$$

represents a differential form of degree 1, which is usually denoted by df. If f is of class C^p, then df is class C^{p-1}.

Let λ_i be the i-th coordinate function. Then we know that

$$d\lambda_i(x) = \lambda_i'(x) = \lambda_i$$

for each $x \in U$ because $\lambda'(x) = \lambda$ for any continuous linear map λ. Whenever $\{x_1, \dots, x_n\}$ are used systematically for the coordinates of a point in $\mathbf{R}^n$, it is customary in the literature to use the notation

$$d\lambda_i(x) = dx_i.$$

This is slightly incorrect, but is useful in formal computations. We shall also use it in this book on occasions. Similarly, we also write (incorrectly)

$$\omega = \sum_{(i)} f_{(i)}\, dx_{i_1} \wedge \cdots \wedge dx_{i_r}$$

instead of the correct

$$\omega(x) = \sum_{(i)} f_{(i)}(x) \lambda_{i_1} \wedge \cdots \wedge \lambda_{i_r}.$$

In terms of coordinates, the map df (or f') is given by

$$df(x) = f'(x) = D_1 f(x)\lambda_1 + \cdots + D_n f(x)\lambda_n,$$

where $D_i f(x) = \partial f/\partial x_i$ is the i-th partial derivative. This is simply a restatement of the fact that if $h = (h_1, \dots, h_n)$ is a vector, then

$$f'(x)h = \frac{\partial f}{\partial x_1} h_1 + \cdots + \frac{\partial f}{\partial x_n} h_n.$$

Thus in old notation, we have

$$df(x) = \frac{\partial f}{\partial x_1}\, dx_1 + \cdots + \frac{\partial f}{\partial x_n}\, dx_n.$$

We shall develop the theory of the alternating product and the exterior derivative directly without assuming Propositions 3.2 or 3.3 in the finite dimensional case.

Let ω and ψ be forms of degrees r and s respectively, on the open set U. For each $x \in U$ we can then take the alternating product $\omega(x) \wedge \psi(x)$ and we define the **alternating product** $\omega \wedge \psi$ by

$$(\omega \wedge \psi)(x) = \omega(x) \wedge \psi(x).$$

(It is an exercise to verify that this product corresponds to the product defined previously before Proposition 3.3 under the isomorphism between $L^r_a(\mathbf{E}, \mathbf{R})$ and the r-th alternating product in the finite dimensional case.) If f is a differential form of degree 0, that is a function, then we have again

$$f \wedge \omega = f\omega,$$

where $(f\omega)(x) = f(x)\omega(x)$. By definition, we then have

$$\omega \wedge f\psi = f\omega \wedge \psi.$$

We shall now define the **exterior derivative** $d\omega$ for any differential form ω. We have already done it for functions. We shall do it in general first in terms of coordinates, and then show that there is a characterization independent of these coordinates. If

$$\omega = \sum_{(i)} f_{(i)}\, d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r},$$

we define

$$d\omega = \sum_{(i)} df_{(i)} \wedge d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r}.$$

Example. Suppose $n = 2$ and ω is a 1-form, given in terms of the two coordinates (x, y) by

$$\omega(x, y) = f(x, y)\, dx + g(x, y)\, dy.$$

Then

$$\begin{aligned} d\omega(x, y) &= df(x, y) \wedge dx + dg(x, y) \wedge dy \\ &= \left(\frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy\right) \wedge dx + \left(\frac{\partial g}{\partial x}\, dx + \frac{\partial g}{\partial y}\, dy\right) \wedge dy \\ &= \frac{\partial f}{\partial y}\, dy \wedge dx + \frac{\partial g}{\partial x}\, dx \wedge dy \\ &= \left(\frac{\partial f}{\partial y} - \frac{\partial g}{\partial x}\right) dy \wedge dx \end{aligned}$$

because the terms involving $dx \wedge dx$ and $dy \wedge dy$ are equal to 0.

Proposition 3.4. *The map d is linear, and satisfies*

$$d(\omega \wedge \psi) = d\omega \wedge \psi + (-1)^r \omega \wedge d\psi$$

if $r = \deg \omega$. The map d is uniquely determined by these properties, and by the fact that for a function f, we have $df = f'$.

Proof. The linearity of d is obvious. Hence it suffices to prove the formula for decomposable forms. We note that for any function f we have

$$d(f\omega) = df \wedge \omega + f\, d\omega.$$

Indeed, if ω is a function g, then from the derivative of a product we get

$d(fg) = f\, dg + g\, df$. If

$$\omega = g\, d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r},$$

where g is a function, then

$$\begin{aligned} d(f\omega) = d(fg\, d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r}) &= d(fg) \wedge d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r} \\ &= (f\, dg + g\, df) \wedge d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r} \\ &= f\, d\omega + df \wedge \omega, \end{aligned}$$

as desired. Now suppose that

$$\begin{aligned} \omega &= f\, d\lambda_{i_1} \wedge \cdots \wedge d\lambda_{i_r} \quad \text{and} \quad & \psi &= g\, d\lambda_{j_1} \wedge \cdots \wedge d\lambda_{j_s} \\ &= f\tilde{\omega}, & &= g\tilde{\psi}, \end{aligned}$$

with $i_1 < \cdots < i_r$ and $j_1 < \cdots < j_s$ as usual. If some $i_\nu = j_\mu$, then from the definitions we see that the expressions on both sides of the equality in the theorem are equal to 0. Hence we may assume that the sets of indices $i_i, \ldots, i_r$ and $j_1, \ldots, j_s$ have no element in common. Then $d(\tilde{\omega} \wedge \tilde{\psi}) = 0$ by definition, and

$$\begin{aligned} d(\omega \wedge \psi) = d(fg\tilde{\omega} \wedge \tilde{\psi}) &= d(fg) \wedge \tilde{\omega} \wedge \tilde{\psi} \\ &= (g\, df + f\, dg) \wedge \tilde{\omega} \wedge \tilde{\psi} \\ &= d\omega \wedge \psi + f\, dg \wedge \tilde{\omega} \wedge \tilde{\psi} \\ &= d\omega \wedge \psi + (-1)^r f\tilde{\omega} \wedge dg \wedge \tilde{\psi} \\ &= d\omega \wedge \psi + (-1)^r \omega \wedge d\psi, \end{aligned}$$

thus proving the desired formula, in the present case. (We used the fact that $dg \wedge \tilde{\omega} = (-1)^r \tilde{\omega} \wedge dg$ whose proof is left to the reader.) The formula in the general case follows because any differential form can be expressed as a sum of forms of the type just considered, and one can then use the bilinearity of the product. Finally, d is uniquely determined by the formula, and its effect on functions, because any differential form is a sum of forms of type $f\, d\lambda_i \wedge \cdots \wedge d\lambda_{i_r}$ and the formula gives an expression of d in terms of its effect on forms of lower degree. By induction, if the value of d on functions is known, its value can then be determined on forms of degree $\geqq 1$. This proves our assertion.

Proposition 3.5. *Let ω be a form of class C^2. Then $dd\omega = 0$.*

Proof. If f is a function, then

$$df(x) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\, dx_j$$

and

$$ddf(x) = \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_j}\, dx_k \wedge dx_j.$$

Using the fact that the partials commute, and the fact that for any two positive integers r, s we have $dx_r \wedge dx_s = -dx_s \wedge dx_r$, we see that the preceding double sum is equal to 0. A similar argument shows that the theorem is true for 1-forms, of type $g(x)\, dx_i$ where g is a function, and thus for all 1-forms by linearity. We proceed by induction. It suffices to prove the formula in general for decomposable forms. Let ω be decomposable of degree r, and write

$$\omega = \eta \wedge \psi,$$

where $\deg \psi = 1$. Using the formula for the derivative of an alternating product twice, and the fact that $dd\psi = 0$ and $dd\eta = 0$ by induction, we see at once that $dd\omega = 0$, as was to be shown.

We conclude this section by giving some properties of the pull-back of forms. As we saw at the end of Chapter III, §4, if $f\colon X \to Y$ is a morphism and if ω is a differential form on Y, then we get a differential form $f^*(\omega)$ on X, which is given at a point $x \in X$ by the formula

$$f^*(\omega)_x = \omega_{f(x)} \circ (T_x f)^r,$$

if ω is of degree r. This holds for $r \geqq 1$. The corresponding local representation formula reads

$$\langle f^*\omega(x), \xi_1(x) \times \cdots \times \xi_r(x)\rangle = \langle \omega(f(x)), f'(x)\xi_1(x) \times \cdots \times f'(x)\xi_r(x)\rangle$$

if $\xi_1, \ldots, \xi_r$ are vector fields.

In the case of a 0-form, that is a function, its pull-back is simply the composite function. In other words, if φ is a function on Y, viewed as a form of degree 0, then

$$\boxed{f^*(\varphi) = \varphi \circ f.}$$

It is clear that the pull-back is linear, and satisfies the following properties.

Property 1. *If ω, ψ are two differential forms on Y, then*

$$f^*(\omega \wedge \psi) = f^*(\omega) \wedge f^*(\psi).$$

Property 2. *If ω is a differential form on Y, then*

$$df^*(\omega) = f^*(d\omega).$$

Property 3. *If $f\colon X \to Y$ and $g\colon Y \to Z$ are two morphisms, and ω is a differential form on Z, then*

$$f^*(g^*(\omega)) = (g \circ f)^*(\omega).$$

Finally, in the case of forms of degree 0:

Property 4. *If $f\colon X \to Y$ is a morphism, and g is a function on Y, then*

$$d(g \circ f) = f^*(dg)$$

and at a point $x \in X$, the value of this 1-form is given by

$$T_{f(x)}g \circ T_x f = (dg)_x \circ T_x f.$$

The verifications are all easy, and even trivial, except possibly for **Property 2**. We shall give the proof of **Property 2** in the finite dimensional case and leave the general case to the reader.

For a form of degree 1, say

$$\omega(y) = g(y)\, dy_1,$$

with $y_1 = f_1(x)$, we find

$$(f^* d\omega)(x) = (g'(f(x)) \circ f'(x)) \wedge df_1(x).$$

Using the fact that $ddf_1 = 0$, together with Proposition 3.4 we get

$$(df^*\omega)(x) = (d(g \circ f))(x) \wedge df_1(x),$$

which is equal to the preceding expression. Any 1-form can be expressed as a linear combination of form $g_i\, dy_i$, so that our assertion is proved for forms of degree 1.

The general formula can now be proved by induction. Using the linearity of f^*, we may assume that ω is expressed as $\omega = \psi \wedge \eta$ where ψ,

η have lower degree. We apply Proposition 3.3 and **Property 1** to

$$f^* d\omega = f^*(d\psi \wedge \eta) + (-1)^r f^*(\psi \wedge d\eta)$$

and we see at once that this is equal to $df^*\omega$, because by induction, $f^* d\psi = df^*\psi$ and $f^* d\eta = df^*\eta$. This proves **Property 2**.

Example 1. Let $y_1, \ldots, y_m$ be the coordinates on V, and let μ_j be the jth coordinate function, $j = 1, \ldots, m$, so that $y_j = \mu_j(y_1, \ldots, y_m)$. Let

$$f\colon U \to V$$

be the map with coordinate functions

$$y_j = f_j(x) = \mu_j \circ f(x).$$

If

$$\omega(y) = g(y)\, dy_{j_1} \wedge \cdots \wedge dy_{j_s}$$

is a differential form on V, then

$$\boxed{f^*\omega = (g \circ f)\, df_{j_1} \wedge \cdots \wedge df_{j_s}.}$$

Indeed, we have for $x \in U$:

$$(f^*\omega)(x) = g(f(x))(\mu_{j_1} \circ f'(x)) \wedge \cdots \wedge (\mu_{j_s} \circ f'(x))$$

and

$$f_j'(x) = (\mu_j \circ f)'(x) = \mu_j \circ f'(x) = df_j(x).$$

Example 2. Let $f\colon [a, b] \to \mathbf{R}^2$ be a map from an interval into the plane, and let x, y be the coordinates of the plane. Let t be the coordinate in $[a, b]$. A differential form in the plane can be written in the form

$$\omega(x, y) = g(x, y)\, dx + h(x, y)\, dy,$$

where g, h are functions. Then by definition,

$$f^*\omega(t) = g(x(t), y(t)) \frac{dx}{dt}\, dt + h(x(t), y(t)) \frac{dy}{dt}\, dt,$$

if we write $f(t) = (x(t), y(t))$. Let $G = (g, h)$ be the vector field whose components are g and h. Then we can write

$$f^*\omega(t) = G(f(t)) \cdot f'(t)\, dt,$$

which is essentially the expression which is integrated when defining the integral of a vector field along a curve.

Example 3. Let U, V be both open sets in n-space, and let $f\colon U \to V$ be a C^p map. If

$$\omega(y) = g(y)\, dy_1 \wedge \cdots \wedge dy_n,$$

where $y_j = f_j(x)$ is the j-th coordinate of y, then

$$dy_j = D_1 f_j(x)\, dx_1 + \cdots + D_n f_j(x)\, dx_n$$
$$= \frac{\partial y_j}{\partial x_1}\, dx_1 + \cdots + \frac{\partial y_j}{\partial x_n}\, dx_n,$$

and consequently, expanding out the alternating product according to the usual multilinear and alternating rules, we find that

$$f^*\omega(x) = g(f(x))\Delta_f(x)\, dx_1 \wedge \cdots \wedge dx_n,$$

where Δ_f is the determinant of the Jacobian matrix of f.

V, §4. THE POINCARÉ LEMMA

If ω is a differential form on a manifold and is such that $d\omega = 0$, then it is customary to say that ω is **closed**. If there exists a form ψ such that $\omega = d\psi$, then one says that ω is **exact**. We shall now prove that locally, every closed form is exact.

Theorem 4.1 (Poincaré Lemma). *Let U be an open ball in $\mathbf{E}$ and let ω be a differential form of degree $\geqq 1$ on U such that $d\omega = 0$. Then there exists a differential form ψ on U such that $d\psi = \omega$.*

Proof. We shall construct a linear map k from the r-forms to the $(r-1)$-forms $(r \geqq 1)$ such that

$$dk + kd = \mathrm{id}.$$

From this relation, it will follow that whenever $d\omega = 0$, then

$$dk\omega = \omega,$$

thereby proving our proposition. We may assume that the center of the ball is the origin. If ω is an r-form, then we define $k\omega$ by the formula

$$\langle (k\omega)_x, v_1 \times \cdots \times v_{r-1} \rangle = \int_0^1 t^{r-1} \langle \omega(tx), x \times v_1 \times \cdots \times v_{r-1} \rangle \, dt.$$

We can assume that we deal with local representations and that $v_i \in \mathbf{E}$. We have

$$\begin{aligned}\langle (dk\omega)_x, v_1 \times \cdots \times v_r \rangle &= \sum_{i=1}^{r} (-1)^{i+1} \langle (k\omega)'(x) v_i, v_1 \times \cdots \times \hat{v}_i \times \cdots \times v_r \rangle \\ &= \sum (-1)^{i+1} \int_0^1 t^{r-1} \langle \omega(tx), v_i \times v_1 \times \cdots \times \hat{v}_i \times \cdots \times v_r \rangle \, dt \\ &\quad + \sum (-1)^{i+1} \int_0^1 t^r \langle \omega'(tx) v_i, x \times v_1 \times \cdots \times \hat{v}_i \times \cdots \times v_r \rangle \, dt.\end{aligned}$$

On the other hand, we also have

$$\begin{aligned}\langle (kd\omega)(x), v_1 \times \cdots \times v_r \rangle &= \int_0^1 t^r \langle d\omega(x), x \times v_1 \times \cdots \times v_r \rangle \, dt \\ &= \int_0^1 t^r \langle \omega'(tx) x, v_1 \times \cdots \times v_r \rangle \, dt \\ &\quad + \sum (-1)^i \int_0^1 t^r \langle \omega'(tx) v_i, x \times v_1 \times \cdots \times \hat{v}_i \times \cdots \times v_r \rangle \, dt.\end{aligned}$$

We observe that the second terms in the expressions for $kd\omega$ and $dk\omega$ occur with opposite signs and cancel when we take the sum. As to the first terms, if we shift v_i to the i-th place in the expression for $dk\omega$, then we get an extra coefficient of $(-1)^{i+1}$. Thus

$$\begin{aligned}dk\omega + kd\omega &= \int_0^1 r t^{r-1} \langle \omega(tx), v_1 \times \cdots \times v_r \rangle \, dt \\ &\quad + \int_0^1 t^r \langle \omega'(tx) x, v_1 \times \cdots \times v_r \rangle \, dt.\end{aligned}$$

This last integral is simply the integral of the derivative with respect to t of

$$\langle t^r \omega(tx), v_1 \times \cdots \times v_r \rangle.$$

Evaluating this expression between $t = 0$ and $t = 1$ yields

$$\langle \omega(x), v_1 \times \cdots \times v_r \rangle$$

which proves the theorem.

We observe that we could have taken our open set U to be star-shaped instead of an open ball. For more information on the relationship between closed and exact forms, see Chapter XIII, §1.

V, §5. CONTRACTIONS AND LIE DERIVATIVE

Let ξ be a vector field and let ω be an r-form on a manifold X, $r \geqq 1$. Then we can define an $(r-1)$-form $C_\xi \omega$ by the formula

$$(C_\xi \omega)(x)(v_2, \ldots, v_r) = \omega(\xi(x), v_2, \ldots, v_r),$$

for $v_2, \ldots, v_r \in T_x$. Using local representations shows at once that $C_\xi \omega$ has the appropriate order of differentiability (the minimum of ω and ξ). We call $C_\xi \omega$ the **contraction** of ω by ξ, and also denote $C_\xi \omega$ by

$$\omega \circ \xi.$$

If f is a function, we define $C_\xi f = 0$. Leaving out the order of differentiability, we see that contraction gives an **R**-linear map

$$C_\xi \colon \mathscr{A}^r(X) \to \mathscr{A}^{r-1}(X).$$

This operation of contraction satisfies the following properties.

CON 1. $C_\xi \circ C_\xi = 0$.

CON 2. *The association $(\xi, \omega) \mapsto C_\xi \omega = \omega \circ \xi$ is bilinear. It is in fact bilinear with respect to functions, that is if φ is a function, then*

$$C_{\varphi\xi} = \varphi C_\xi \quad \text{and} \quad C_\xi(\varphi\omega) = \varphi C_\xi \omega.$$

CON 3. *If ω, ψ are differential forms and $r = \deg \omega$, then*

$$C_\xi(\omega \wedge \psi) = (C_\xi \omega) \wedge \psi + (-1)^r \omega \wedge C_\xi \psi.$$

These three properties follow at once from the definitions.

Example. Let $X = \mathbf{R}^n$, and let

$$\omega(x) = dx_1 \wedge \cdots \wedge dx_n.$$

If ξ is a vector field on $\mathbf{R}^n$, then we have the local representation

$$(\omega \circ \xi)(x) = \sum_{i=1}^{n} (-1)^{i+1} \xi_i(x)\, dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n.$$

We also have immediately from the definition of the exterior derivative,

$$d(\omega \circ \xi) = \sum_{i=1}^{n} \frac{\partial \xi_i(x)}{\partial x_i}\, dx_1 \wedge \cdots \wedge dx_n,$$

letting $\xi = (\xi_1, \ldots, \xi_n)$ in terms of its components ξ_i.

We can define the **Lie derivative** of an r-form as we did before for vector fields. Namely, we shall evaluate the following limit:

$$(\mathscr{L}_\xi \omega)(x) = \lim_{t \to 0} \frac{1}{t} [(\alpha_t^* \omega)(x) - \omega(x)],$$

or in other words,

$$\mathscr{L}_\xi \omega = \frac{d}{dt} (\alpha_t^* \omega) \Big|_{t=0}$$

where α is the flow of the vector field ξ, and we call $\mathscr{L}_\xi$ the **Lie derivative** again, applied to the differential form ω. We may rewrite this definition in terms of the value on vector fields $\xi_1, \ldots, \xi_r$ as follows:

$$(\mathscr{L}_\xi \omega)(\xi_1, \ldots, \xi_r) = \frac{d}{dt} \langle \omega \circ \alpha_t,\ \alpha_{t*}\xi_1 \times \cdots \times \alpha_{t*}\xi_r \rangle \Big|_{t=0}$$

Proposition 5.1. *Let ξ be a vector field and ω a differential form of degree $r \geqq 1$. The Lie derivative $\mathscr{L}_\xi$ is a derivation, in the sense that*

$$\mathscr{L}_\xi(\omega(\xi_1, \ldots, \xi_r)) = (\mathscr{L}_\xi \omega)(\xi_1, \ldots, \xi_r) + \sum_{i=1}^{r} \omega(\xi_1, \ldots, \mathscr{L}_\xi \xi_i, \ldots, \xi_r)$$

where of course $\mathscr{L}_\xi \xi_i = [\xi, \xi_i]$.

If ξ, ξ_i, ω denote the local representations of the vector fields and the form respectively, then the Lie derivative $\mathscr{L}_\xi \omega$ has the local representation

$$\begin{aligned}\langle (\mathscr{L}_\xi \omega)(x), \xi_1(x) \times \cdots \times \xi_r(x) \rangle &= \langle \omega'(x)\xi(x), \xi_1(x) \times \cdots \times \xi_r(x) \rangle \\ &\quad + \sum_{i=1}^{r} \langle \omega(x), \xi_1(x) \times \cdots \times \xi'(x)\xi_i(x) \times \cdots \times \xi_r(x) \rangle.\end{aligned}$$

Proof. The proof is routine using the definitions. The first assertion is obvious by the definition of the pull back of a form. For the local expression we actually derive more, namely we derive a local expression for $\alpha_t^*\omega$ and $\frac{d}{dt}\alpha_t^*\omega$ which are characterized by their values at $(\xi_1, \ldots, \xi_r)$. So we let

$$\begin{aligned}(1) \qquad F(t) &= \langle (\alpha_t^*\omega)(x), \xi_1(x) \times \cdots \times \xi_r(x) \rangle \\ &= \langle \omega(\alpha(t, x)), D_2\alpha(t, x)\xi_1(x) \times \cdots \times D_2\alpha(t, x)\xi_r(x) \rangle.\end{aligned}$$

Then the Lie derivative $(\mathscr{L}_\xi \omega)(x)$ is precisely $F'(0)$, but we obtain also the local representation for $\frac{d}{dt}\alpha_t^*\omega$:

$$(2) \quad F'(t) = \left\langle \frac{d}{dt}\alpha_t^*\omega(x), \xi_1(x) \times \cdots \times \xi_r(x) \right\rangle =$$

$$(3) \qquad \langle \omega'(\alpha(t, x))D_1\alpha(t, x), D_2\alpha(t, x)\xi_1(x) \times \cdots \times D_2\alpha(t, x)\xi_r(x) \rangle$$

$$+ \sum_{i=1}^{r} \langle \omega(\alpha(t, x)), D_2\alpha(t, x)\xi_1(x) \times \cdots \times D_1 D_2\alpha(t, x)\xi_i(x) \times \cdots \times D_2\alpha(t, x)\xi_r(x) \rangle$$

by the rule for the derivative of a product. Putting $t = 0$ and using the differential equation satisfied by $D_2\alpha(t, x)$, we get precisely the local expression as stated in the proposition. Remember the initial condition $D_2\alpha(0, x) = \mathrm{id}$.

From Proposition 5.1, we conclude that the Lie derivative gives an $\mathbf{R}$-linear map

$$\mathscr{L}_\xi \colon \mathscr{A}^r(X) \to \mathscr{A}^r(X).$$

We may use expressions (1) and (3) in the above proof to derive a formula which holds even more generally for time-dependent vector fields.

Proposition 5.2. *Let ξ_t be a time-dependent vector field, α its flow, and let ω be a differential form. Then*

$$\frac{d}{dt}(\alpha_t^*\omega) = \alpha_t^*(\mathcal{L}_{\xi_t}\omega) \quad \text{or} \quad \frac{d}{dt}(\alpha_t^*\omega) = \alpha_t^*(\mathcal{L}_{\xi}\omega)$$

for a time-independent vector field.

Proof. Proposition 5.1 gives us a local expression for $(\mathcal{L}_{\xi_t}\omega)(y)$, replacing x by y because we shall now put $y = \alpha(t, x)$. On the other hand, from (1) in the proof of Proposition 5.1, we obtain

$$\alpha_t^*(\mathcal{L}_{\xi_t}\omega)(x) = \langle(\mathcal{L}_{\xi_t}\omega)(y),\ D_2\alpha(t, x)\xi_1(x) \times \cdots \times D_2\alpha(t, x)\xi_r(x)\rangle.$$

Substituting the local expression for $(\mathcal{L}_{\xi_t}\omega)(y)$, we get expression (3) from the proof of Proposition 5.1, thereby proving Proposition 5.2.

Proposition 5.3. *As a map on differential forms, the Lie derivative satisfies the following properties.*

LIE 1. $\mathcal{L}_\xi = d \circ C_\xi + C_\xi \circ d$, *so* $\mathcal{L}_\xi = C_\xi \circ d$ *on functions.*

LIE 2. $\mathcal{L}_\xi(\omega \wedge \psi) = \mathcal{L}_\xi\omega \wedge \psi + \omega \wedge \mathcal{L}_\xi\psi$.

LIE 3. $\mathcal{L}_\xi$ *commutes with* d *and* C_ξ.

LIE 4. $\mathcal{L}_{[\xi,\eta]} = \mathcal{L}_\xi \circ \mathcal{L}_\eta - \mathcal{L}_\eta \circ \mathcal{L}_\xi$.

LIE 5. $C_{[\xi,\eta]} = \mathcal{L}_\xi \circ C_\eta - C_\eta \circ \mathcal{L}_\xi$.

LIE 6. $\mathcal{L}_{f\xi}\omega = f\mathcal{L}_\xi\omega + df \wedge C_\xi\omega$ *for all forms* ω *and functions* f.

Proof. Let $\xi_1, \ldots, \xi_r$ be vector fields, and ω an r-form. Using the definition of the contraction and the local formula of Proposition 5.1, we find that $C_\xi\, d\omega$ is given locally by

$$\begin{aligned}
&\langle C_\xi\, d\omega(x),\ \xi_1(x) \times \cdots \times \xi_r(x)\rangle \\
&\quad = \langle \omega'(x)\xi(x),\ \xi_1(x) \times \cdots \times \xi_r(x)\rangle \\
&\qquad + \sum_{i=1}^{r}(-1)^i\langle\omega'(x)\xi_i(x),\ \xi(x) \times \xi_1(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \xi_r(x)\rangle.
\end{aligned}$$

On the other hand, $dC_\xi\omega$ is given by

$$\begin{aligned}
&\langle dC_\xi\omega(x),\ \xi_1(x) \times \cdots \times \xi_r(x)\rangle \\
&\quad = \sum_{i=1}^{r}(-1)^{i+1}\langle(C_\xi\omega)'(x)\xi_i(x),\ \xi_1(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \times \xi_r(x)\rangle.
\end{aligned}$$

To compute $(C_\xi\omega)'(x)$ is easy, going back to the definition of the derivative. At vectors $v_1, \ldots, v_{r-1}$, the form $C_\xi\omega(x)$ has the value

$$\langle\omega(x),\ \xi(x) \times v_1 \times \cdots \times v_{r-1}\rangle.$$

Differentiating this last expression with respect to x and evaluating at a vector h we get

$$\langle\omega'(x)h,\ \xi(x) \times v_1 \times \cdots \times v_{r-1}\rangle + \langle\omega(x),\ \xi'(x)h \times v_1 \times \cdots \times v_{r-1}\rangle.$$

Hence $\langle dC_\xi\omega(x),\ \xi_1(x) \times \cdots \times \xi_r(x)\rangle$ is equal to

$$\begin{aligned}
&\sum_{i=1}^{r}(-1)^{i+1}\langle\omega'(x)\xi_i(x),\ \xi(x) \times \xi_1(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \times \xi_r(x)\rangle \\
&\qquad + \sum_{i=1}^{r}(-1)^{i+1}\langle\omega(x),\ \xi'(x)\xi_i(x) \times \xi_1(x) \times \cdots \times \widehat{\xi_i(x)} \times \cdots \times \xi_r(x)\rangle.
\end{aligned}$$

Shifting $\xi'(x)\xi_i(x)$ to the i-th place in the second sum contributes a sign of $(-1)^{i-1}$ which gives 1 when multiplied by $(-1)^{i+1}$. Adding the two local representations for $dC_\xi\omega$ and $C_\xi\, d\omega$, we find precisely the expression of Proposition 5.1, thus proving **LIE 1**.

As for **LIE 2**, it consists in using the derivation rule for d and C_ξ in Proposition 3.3, **EXD 1**, and **CON 3**. The corresponding rule for $\mathcal{L}_\xi$ follows at once. (Terms will cancel just the right way.) The other properties are then clear.

V, §6. VECTOR FIELDS AND 1-FORMS UNDER SELF DUALITY

Let $\mathbf{E}$ be a Banach space and let

$$(v, w) \mapsto \langle v, w\rangle$$

be a continuous bilinear function of $\mathbf{E} \times \mathbf{E} \to \mathbf{R}$. We call such a function a **bilinear form**. This form induced a linear map

$$\lambda\colon \mathbf{E} \to \mathbf{E}^\vee$$

which to each $v \in \mathbf{E}$ associates the functional λ_v such that

$$\lambda_v(w) = \langle v, w\rangle.$$

We have a similar map on the other side. If both these mappings are toplinear isomorphisms of $\mathbf{E}$ and $\mathbf{E}^\vee$ then we say that the bilinear form is **non-singular**. If such a non-singular form exists, then we say that $\mathbf{E}$ is **self-dual**. For instance, a Hilbert space is self-dual.

If $\mathbf{E}$ is finite dimensional, it suffices for a bilinear form to be non-singular that its kernels on the right and on the left be 0. (The kernels are the kernels of the associated maps λ as above.) However, in the infinite dimensional case, this condition on the kernels is not sufficient any more.

Let $\mathbf{E}$ be a self dual Banach space with respect to the non-singular form $(v, w) \mapsto \langle v, w\rangle$, and let

$$\Omega: \mathbf{E} \times \mathbf{E} \to \mathbf{R}$$

be a continuous bilinear map. There exists a unique operator A such that

$$\Omega(v, w) = \langle Av, w\rangle$$

for all $v, w \in \mathbf{E}$. (An **operator** is a continuous linear map by definition.)

Remarks. *Suppose that the form* $(v, w) \mapsto \langle v, w\rangle$ *is symmetric,* i.e.

$$\langle v, w\rangle = \langle w, v\rangle$$

for all $v, w \in \mathbf{E}$. Then Ω is **symmetric** (resp. **alternating**) if and only if A is symmetric (resp. skew-symmetric). Recall that A symmetric (with respect to $\langle\ ,\ \rangle$) means that

$$\langle Av, w\rangle = \langle v, Aw\rangle \qquad \text{for all} \quad v, w \in \mathbf{E}.$$

That A is skew-symmetric means that $\langle Av, w\rangle = -\langle Aw, w\rangle$ for all $v, w \in \mathbf{E}$. For any operator A: $\mathbf{E} \to \mathbf{E}$ there is another operator ${}^tA$ (the transpose of A with respect to the non-singular form $\langle\ ,\ \rangle$) such that for all $v, w \in \mathbf{E}$ we have

$$\langle Av, w\rangle = \langle v, {}^tAw\rangle.$$

Thus A is symmetric (resp. skew-symmetric) if and only if ${}^tA = A$ (resp. ${}^tA = -A$).

The above remarks apply to any continuous bilinear form Ω. For invertibility, we have the criterion:

The form Ω is non-singular if and only if the operator A representing the form with respect to $\langle\ ,\ \rangle$ is invertible.

The easy verification is left to the reader. Of course, in the finite dimensional case, invertibility or non-singularity can be checked by verifying

that the matrix representing the linear map with respect to bases has non-zero determinant. Similarly, the form is also represented by a matrix with respect to a choice of bases, and its being non-singular is equivalent to the matrix representing the form being invertible.

We recall that the set of invertible operators in Laut(E) is an open subset. Alternatively, the set of non-singular bilinear forms on $\mathbf{E}$ is an open subset of $L^2(\mathbf{E})$.

We may now globalize these notions to a vector bundle (and eventually especially to the tangent bundle) as follows.

Let X be a manifold, and π: $E \to X$ a vector bundle over X with fibers which are toplinearly isomorphic to $\mathbf{E}$, or as we shall also say, **modeled** on $\mathbf{E}$. Let Ω be a tensor field of type L^2 on E, that is to say, a section of the bundle $L^2(E)$ (or $L^2(\pi)$), or as we shall also say, a **bilinear tensor field** on E. Then for each $x \in X$, we have a continuous bilinear form Ω_x on E_x.

If Ω_x is non-singular for each $x \in X$ then we say that Ω is **non-singular**. If π is trivial, and we have a trivialisation $X \times E$, then the local representation of Ω can be described by a morphism of X into the Banach space of operators. If Ω is non-singular, then the image of this morphism is contained in the open set of invertible operators. (If Ω is a 2-form, this image is contained in the submanifold of skew-symmetric operators.) For example, in a chart U, we can represent Ω over U by a morphism

$$A\colon U \to L(\mathbf{E}, \mathbf{E}) \qquad \text{such that} \qquad \Omega_x(v, w) = \langle A_x v, w\rangle$$

for all $v, w \in \mathbf{E}$. Here we wrote A_x instead of $A(x)$ to simplify the typography.

A non-singular Ω as above can be used to establish a linear isomorphism

$$\Gamma(E) \to \Gamma L^1(E), \qquad \text{also denoted by} \quad \Gamma L(E) \text{ or } \Gamma E^\vee,$$

between the $\mathbf{R}$-vector spaces of sections $\Gamma(E)$ of E and the 1-forms on E in the following manner. Let ξ be a section of E. For each $x \in X$ we define a continuous linear map

$$(\Omega \circ \xi)_x\colon E_x \to \mathbf{R}$$

by the formula

$$(\Omega \circ \xi)_x(w) = \Omega_x(\xi(x), w).$$

Looking at local trivialisations of π, we see at once that $\Omega \circ \xi$ is a 1-form on E.

Conversely, let ω be a given 1-form on E. For each $x \in X$, ω_x is therefore a 1-form on E_x and since Ω is non-singular, there exists a unique element $\xi(x)$ of E_x such that

$$\Omega_x(\xi(x), w) = \omega_x(w)$$

for all $w \in E_x$. In this fashion, we obtain a mapping ξ of X into E and we contend that ξ is a morphism (and therefore a section).

To prove our contention we can look at the local representations. We use Ω and ω to denote these. They are represented over a suitable open set U by two morphisms

$$A\colon U \to \mathrm{Aut}(\mathbf{E}) \quad \text{and} \quad \eta\colon U \to \mathbf{E}$$

such that

$$\Omega_x(v, w) = \langle A_x v, w\rangle \quad \text{and} \quad \omega_x(w) = \langle \eta(x), w\rangle.$$

From this we see that

$$\xi(x) = A_x^{-1}\eta(x),$$

from which it is clear that ξ is a morphism. We may summarize our discussion as follows.

Proposition 6.1. *Let X be a manifold and $\pi\colon E \to X$ a vector bundle over X modeled on $\mathbf{E}$. Let Ω be a non-singular bilinear tensor field on E. Then Ω induces an isomorphism of* $\mathrm{Fu}(X)$*-modules*

$$\Gamma E \to \Gamma E^\vee.$$

A section ξ corresponds to a 1-form ω if and only if $\Omega \circ \xi = \omega$.

In many applications, one takes the differential form to be df for some function f. The vector field corresponding to df is then called the **gradient of f with respect to Ω.**

Remark. There is no universally accepted notation to denote the correspondence between a 1-form and a vector field under Ω as above. Some authors use sharps and flats, which have two disadvantages. First, they do not provide a symbols for the mapping, and second they do not contain the Ω in the notation. I would propose the check sign $\bigvee_\Omega$ to denote either isomorphism

$$\bigvee_\Omega\colon \Gamma L(E) \to \Gamma E \qquad \text{denoted on elements by} \qquad \omega \mapsto \bigvee_\Omega \omega = \omega^\vee = \xi_\omega$$

and also

$$\bigvee_\Omega\colon \Gamma E \to \Gamma L(E) \qquad \text{denoted on elements by} \qquad \xi \mapsto \bigvee_\Omega \xi = \xi^\vee = \omega_\xi.$$

If Ω is fixed throughout a discussion and need not be referred to, then it is useful to write $\xi^\vee$ or $\lambda^\vee$ in some formulas. We have $\bigvee_\Omega \circ \bigvee_\Omega = \mathrm{id}$. Instead of the sharp and flat superscript, I prefer the single $^\vee$ sign.

Many important applications of the above duality occur when Ω is a non-singular symmetric bilinear tensor field on the tangent bundle TX. Such a tensor field is then usually denoted by g. If ξ, η are vector fields, we may then define their scalar product to be the function

$$\langle \xi, \eta\rangle_g = g(\xi, \eta).$$

On the other hand, by the duality of Proposition 6.1, if i.e. ω, λ are 1-forms, i.e. sections of the dual bundle $T^\vee X$, then $\omega^\vee$ and $\lambda^\vee$ are vector fields, and we define the scalar product of the 1-forms to be

$$\langle \omega, \lambda\rangle_g = \langle \omega^\vee, \lambda^\vee\rangle_g.$$

This duality is especially important for Riemannian metrics, as in Chapter X.

The rest of this section will not be used in the book.

In Proposition 6.1, we dealt with a quite general non-singular bilinear tensor field on E. We now specialize to the case when $E = TX$ is the tangent bundle of X, and Ω is a 2-form, i.e. Ω is alternating. A pair (X, Ω) consisting of a manifold and a non-singular closed 2-form is called a **symplectic manifold.** (Recall that **closed** means $d\Omega = 0$.)

We denote by ξ, η vector fields over X, and by f, h functions on X, so that df, dh are 1-forms. We let ξ_{df} be the vector field on X which corresponds to df under the 2-form Ω, according to Proposition 6.1. Vector fields on X which are of type ξ_{df} are called **Hamiltonian** (with respect to the 2-form). More generally, we denote by ξ_ω the vector field corresponding to a 1-form ω. By definition we have the formula

Ω 1. $\qquad \Omega \circ \xi_\omega = \omega \quad$ so in particular $\quad \Omega \circ \xi_{df} = df.$

In Chapter VII, §6 we shall consider a particularly important example, when the base manifold is the cotangent bundle; the function is the **kinetic energy**

$$K(v) = \tfrac{1}{2}\langle v, v\rangle_g$$

with respect to the scalar product g of a Riemannian or pseudo Riemannian metric, and the 2-form Ω arises canonically from the pseudo Riemannian metric.

In general, by **LIE 1** of Proposition 5.3 formula **Ω 1**, and the fact that $d\Omega = 0$, we find for any 1-form ω that:

Ω 2. $\qquad \mathscr{L}_{\xi_\omega}\Omega = d\omega.$

The next proposition reinterprets this formula in terms of the flow when $d\omega = 0$.

Proposition 6.2. *Let ω be such that $d\omega = 0$. Let α be the flow of ξ_ω. Then $\alpha_t^*\Omega = \Omega$ for all t (in the domain of the flow).*

Proof. By Proposition 5.2,

$$\frac{d}{dt}\alpha_t^*\Omega = \alpha_t^*\mathscr{L}_{\xi_\omega}\Omega = 0 \quad \text{by } \Omega\, \mathbf{2}.$$

Hence $\alpha_t^*\Omega$ is constant, equal to $\alpha_0^*\Omega = \Omega$, as was to be shown.

A special case of Proposition 6.2 in Hamiltonian mechanics is when $\omega = dh$ for some function h. Next by **LIE 5**, we obtain for any vector fields ξ, η:

$$\mathscr{L}_\xi(\Omega \circ \eta) = (\mathscr{L}_\xi\Omega) \circ \eta + \Omega \circ [\xi, \eta].$$

In particular, since $ddf = 0$, we get

Ω 3. $$\mathscr{L}_{\xi_{df}}(\Omega \circ \xi_{dh}) = \Omega \circ [\xi_{df}, \xi_{dh}].$$

One defines the **Poisson bracket** between two functions f, h to be

$$\{f, h\} = \xi_{df} \cdot h.$$

Then the preceding formula may be rewritten in the form

Ω 4. $$[\xi_{df}, \xi_{dh}] = \xi_{d\{f,h\}}.$$

It follows immediately from the definitions and the antisymmetry of the ordinary bracket between vector fields that the Poisson bracket is also antisymmetric, namely

$$\{f, h\} = -\{h, f\}.$$

In particular, we find that

$$\xi_{df} \cdot f = 0.$$

In the case of the cotangent bundle with a symplectic 2-form as in the next section, physicists think of f as an energy function, and interpret this formula as a law of conservation of energy. The formula expresses the property that f is constant on the integral curves of the vector field ξ_{df}. This property follows at once from the definition of the Lie derivative of a function. Furthermore:

Proposition 6.3. *If $\xi_{df} \cdot h = 0$ then $\xi_{dh} \cdot f = 0$.*

This is immediate from the antisymmetry of the Poisson bracket. It is interpreted as conservation of momentum in the physical theory of

Hamiltonian mechanics, when one deals with the canonical 2-form on the cotangent bundle, to be defined in the next section.

V, §7. THE CANONICAL 2-FORM

Consider the functor $\mathbf{E} \mapsto L(\mathbf{E})$ (continuous linear forms). If $E \to X$ is a vector bundle, then $L(E)$ will be called the **dual bundle**, and will be denoted by $\mathbf{E}^\vee$. For each $x \in X$, the fiber of the dual bundle is simply $L(E_x)$.

If $E = T(X)$ is the tangent bundle, then its dual is denoted by $T^\vee(X)$ and is called the **cotangent bundle**. Its elements are called **cotangent vectors**. The fiber of $T^\vee(X)$ over a point x of X is denoted by $T_x^\vee(X)$. For each $x \in X$ we have a pairing

$$T_x^\vee \times T_x \to \mathbf{R}$$

given by

$$\langle \lambda, u \rangle = \lambda(u)$$

for $\lambda \in T_x^\vee$ and $u \in T_x$ (it is the value of the linear form λ at u).

We shall now describe how to construct a canonical 1-form on the cotangent bundle $T^\vee(X)$. For each $\lambda \in T^\vee(X)$ we must define a 1-form on $T_\lambda(T^\vee(X))$.

Let π: $T^\vee(X) \to X$ be the canonical projection. Then the induced tangent map

$$T\pi = \pi_*: T(T^\vee(X)) \to T(X)$$

can be applied to an element z of $T_v(T^\vee(X))$ and one sees at once that $\pi_* z$ lies in $T_x(X)$ if λ lies in $T_x^\vee(X)$. Thus we can take the pairing

$$\langle \lambda, \pi_* z \rangle = \theta_\lambda(z)$$

to define a map (which is obviously continuous linear):

$$\theta_\lambda: T_\lambda(T^\vee(X)) \to \mathbf{R}.$$

Proposition 7.1. *This map defines a 1-form on $T^\vee(X)$. Let $X = U$ be open in $\mathbf{E}$ and*

$$T^\vee(U) = U \times \mathbf{E}^\vee, \quad T(T^\vee(U)) = (U \times \mathbf{E}^\vee) \times (\mathbf{E} \times \mathbf{E}^\vee).$$

If $(x, \lambda) \in U \times \mathbf{E}^\vee$ and $(u, \omega) \in \mathbf{E} \times \mathbf{E}^\vee$, then the local representation $\theta_{(x,\lambda)}$ is given by

$$\langle \theta_{(x,\lambda)}, (u, \omega) \rangle = \lambda(u).$$

Proof. We observe that the projection $\pi\colon U \times \mathbf{E}^\vee \to U$ is linear, and hence that its derivative at each point is constant, equal to the projection on the first factor. Our formula is then an immediate consequence of the definition. The local formula shows that θ is in fact a 1-form locally, and therefore globally since it has an invariant description.

Our 1-form is called the **canonical 1-form on the cotangent bundle**. We define the **canonical 2-form** Ω on the cotangent bundle $T^\vee X$ to be

$$\Omega = -d\theta.$$

The next proposition gives a local description of Ω.

Proposition 7.2. *Let U be open in* $\mathbf{E}$, *and let Ω be the local representation of the canonical 2-form on $T^\vee U = U \times \mathbf{E}^\vee$. Let $(x, \lambda) \in U \times \mathbf{E}^\vee$. Let (u_1, ω_1) and (u_2, ω_2) be elements of $\mathbf{E} \times \mathbf{E}^\vee$. Then*

$$\begin{aligned}\langle \Omega_{(x,\lambda)}, (u_1, \omega_1) \times (u_2, \omega_2)\rangle &= \langle u_1, \omega_2\rangle - \langle u_2, \omega_1\rangle \\ &= \omega_2(u_1) - \omega_1(u_2).\end{aligned}$$

Proof. We observe that θ is linear, and thus that θ' is constant. We then apply the local formula for the exterior derivative, given in Proposition 3.2. Our assertion becomes obvious.

The canonical 2-form plays a fundamental role in Lagrangian and Hamiltonian mechanics, cf. [AbM 78], Chapter 3, §3. I have taken the sign of the canonical 2-form both so that its value is a 2×2 determinant, and so that it fits with, for instance, [LoS 68] and [AbM 78]. We observe that Ω is closed, that is $d\Omega = 0$, because $\Omega = -d\theta$. Thus $(T^\vee X, \Omega)$ is a symplectic manifold, to which the properties listed at the end of the last section apply.

In particular, let ξ be a vector field on X. Then to ξ is **associated a function** called the **momentum function**

$$f_\xi\colon T^\vee X \to \mathbf{R} \qquad \text{such that} \qquad f_\xi(\lambda_x) = \lambda_x(\xi(x))$$

for $\lambda_x \in T_x^\vee X$. Then df_ξ is a 1-form on $T^\vee X$. Classical Hamiltonian mechanics then applies Propositions 6.2 and 6.3 to this situation. We refer the interested reader to [LoS 68] and [AbM 78] for further information on this topic. For an important theorem of Marsden–Weinstein [MaW 74] and applications to vector bundles, see [Ko 87].

V, §8. DARBOUX'S THEOREM

If $\mathbf{E} = \mathbf{R}^n$ then the usual scalar product establishes the self-duality of $\mathbf{R}^n$. This self-duality arises from other forms, and in this section we are especially interested in the self-duality arising from alternating forms. If $\mathbf{E}$ is finite dimensional and ω is an element of $L_a^2(\mathbf{E})$, that is an alternating 2-form, which is non-singular, then one sees easily that the dimension of $\mathbf{E}$ is even.

Example. An example of such a form on $\mathbf{R}^{2n}$ is the following. Let

$$v = (v_1, \dots, v_n, v_1', \dots, v_n'),$$
$$w = (w_1, \dots, w_n, w_1', \dots, w_n'),$$

be elements of $\mathbf{R}^{2n}$, with components v_i, v_i', w_i, w_i'. Letting

$$\omega(v, w) = \sum_{i=1}^n (v_i w_i' - v_i' w_i)$$

defines a non-singular 2-form ω on $\mathbf{R}^{2n}$. It is an exercise of linear algebra to prove that any non-singular 2-form on $\mathbf{R}^{2n}$ is linearly isomorphic to this particular one in the following sense. If

$$f\colon E \to F$$

is a linear isomorphism between two finite dimensional spaces, then it induces an isomorphism

$$f^*\colon L_a^2(F) \to L_a^2(E).$$

We call forms ω on E and ψ on F **linearly isomorphic** if there exists a linear isomorphism f such that $f^*\psi = \omega$. Thus up to a linear isomorphism, there is only one non-singular 2-form on $\mathbf{R}^{2n}$. (For a proof, cf. for instance my book *Algebra*.)

We are interested in the same question on a manifold locally. Let U be open in the Banach space $\mathbf{E}$ and let $x_0 \in U$. A 2-form

$$\omega\colon U \to L_a^2(\mathbf{E})$$

is said to be **non-singular** if each form $\omega(x)$ is non-singular. If ξ is a vector field on U, then $\omega \circ \xi$ is a 1-form, whose value at (x, ω) is given

$$(\omega \circ \xi)(x)(w) = \omega(x)(\xi(x), w).$$

As a special case of Proposition 6.1, we have:

Let ω be a non-singular 2-form on an open set U in **E**. *The association*

$$\xi \mapsto \omega \circ \xi$$

is a linear isomorphism between the space of vector fields on U and the space of 1-forms on U.

Let

$$\omega\colon U \to L_a^2(U)$$

be a 2-form on an open set U in **E**. If there exists a local isomorphism f at a point $x_0 \in U$, say

$$f\colon U_1 \to V_1,$$

and a 2-form ψ on V_1 such that $f^*\psi = \omega$ (or more accurately, ω restricted to U_1), then we say that ω is **locally isomorphic** to ψ at x_0. Observe that in the case of an isomorphism we can take a direct image of forms, and we shall also write

$$f_*\omega = \psi$$

instead of $\omega = f^*\psi$. In other words, $f_* = (f^{-1})^*$.

Example. On $\mathbf{R}^{2n}$ we have the constant form of the previous example. In terms of local coordinates $(x_1, \dots, x_n, y_1, \dots, y_n)$, this form has the local expression

$$\omega(x, y) = \sum_{i=1}^{n} dx_i \wedge dy_i.$$

This 2-form will be called the **standard 2-form** on $\mathbf{R}^{2n}$.

The Darboux theorem states that any non-singular closed 2-form in $\mathbf{R}^{2n}$ is locally isomorphic to the standard form, that is that in a suitable chart at a point, it has the standard expression of the above example. A technique to show that certain forms are isomorphic was used by Moser [Mo 65], who pointed out that his arguments also prove the classical Darboux theorem. Moser's theorem will be given in Chapter XVIII, §2.

Alan Weinstein observed that Moser's proof applies to the infinite dimensional case, whose statement is as follows.

Theorem 8.1 (Darboux Theorem). *Let* **E** *be a self-dual Banach space. Let*

$$\omega\colon U \to L_a^2(\mathbf{E})$$

be a non-singular closed 2-form on an open set of **E**, *and let $x_0 \in U$. Then ω is locally isomorphic at x_0 to the constant form $\omega(x_0)$.*

Proof. Let $\omega_0 = \omega(x_0)$, and let

$$\omega_t = \omega_0 + t(\omega - \omega_0), \qquad 0 \leqq t \leqq 1.$$

We wish to find a time-dependent vector field ξ_t locally at 0 such that if α denotes its flow, then

$$\alpha_t^*\omega_t = \omega_0.$$

Then the local isomorphism α_1 satisfies the requirements of the theorem. By the Poincaré lemma, there exists a 1-form θ locally at 0 such that

$$\omega - \omega_0 = d\theta,$$

and without loss of generality, we may assume that $\theta(x_0) = 0$. We contend that the time-dependent vector field ξ_t, such that

$$\omega_t \circ \xi_t = -\theta,$$

has the desired property. Let α be its flow. If we shrink the domain of the vector field near x_0 sufficiently, and use the fact that $\theta(x_0) = 0$, then we can use the local existence theorem (Proposition 1.1 of Chapter IV) to see that the flow can be integrated at least to $t = 1$ for all points x in this small domain. We shall now verify that

$$\frac{d}{dt}(\alpha_t^*\omega_t) = 0.$$

This will prove that $\alpha_t^*\omega_t$ is constant. Since we have $\alpha_0^*\omega_0 = \omega_0$ because

$$\alpha(0, x) = x \quad \text{and} \quad D_2\alpha(0, x) = \text{id},$$

it will conclude the proof of the theorem.

We compute locally. We use the local formula of Proposition 5.2, and formula **LIE 1**, which reduces to

$$\mathscr{L}_{\xi_t}\omega_t = d(\omega_t \circ \xi_t),$$

because $d\omega_t = 0$. We find

$$\begin{aligned}\frac{d}{dt}(\alpha_t^*\omega_t) &= \alpha_t^*\left(\frac{d}{dt}\omega_t\right) + \alpha_t^*(\mathscr{L}_{\xi_t}\omega_t)\\ &= \alpha_t^*\left(\frac{d}{dt}\omega_t + d(\omega_t \circ \xi_t)\right)\\ &= \alpha_t^*(\omega - \omega_0 - d\theta)\\ &= 0.\end{aligned}$$

This proves Darboux's theorem.

Remark 1. For the analogous uniqueness statement in the case of a non-singular symmetric form, see the Morse–Palais lemma of Chapter VII, §5. Compare also with Theorem 2.2 of Chapter XVIII.

Remark 2. The proof of the Poincaré lemma can also be cast in the above style. For instance, let $\phi_t(x) = tx$ be a retraction of a star shaped open set around 0. Let ξ_t be the vector field whose flow is ϕ_t, and let ω be a closed form. Then

$$\frac{d}{dt}\phi_t^*\omega = \phi_t^*\mathscr{L}_{\xi_t}\omega = \phi_t^*\, dC_{\xi_t}\omega = d\phi_t^* C_{\xi_t}\omega.$$

Since $\phi_0^*\omega = 0$ and ϕ_1 is the identity, we see that

$$\omega = \phi_1^*\omega - \phi_0^*\omega = \int_0^1 \frac{d}{dt}\phi_t^*\omega\, dt = d\int_0^1 \phi_t^* C_{\xi_t}\omega\, dt$$

is exact, thus concluding a proof of Poincaré's theorem.

CHAPTER VI

The Theorem of Frobenius

Having acquired the language of vector fields, we return to differential equations and give a generalization of the local existence theorem known as the Frobenius theorem, whose proof will be reduced to the standard case discussed in Chapter IV. We state the theorem in §1. Readers should note that one needs only to know the definition of the bracket of two vector fields in order to understand the proof. It is convenient to insert also a formulation in terms of differential forms, for which the reader needs to know the local definition of the exterior derivative. However, the condition involving differential forms is proved to be equivalent to the vector field condition at the very beginning, and does not reappear explicitly afterwards.

We shall follow essentially the proof given by Dieudonné in his *Foundations of Modern Analysis*, allowing for the fact that we use freely the geometric language of vector bundles, which is easier to grasp.

It is convenient to recall in §2 the statements concerning the existence theorems for differential equations depending on parameters. The proof of the Frobenius theorem proper is given in §3. An important application to Lie groups is given in §5, after formulating the theorem globally.

The present chapter will not be used in the rest of this book.

VI, §1. STATEMENT OF THE THEOREM

Let X be a manifold of class C^p $(p \geqq 2)$. A subbundle E of its tangent bundle will also be called a **tangent subbundle** over X. We contend that the following two conditions concerning such a subbundle are equivalent.

FR 1. *For each point $z \in X$ and vector fields ξ, η at z (i.e. defined on an open neighborhood of z) which lie in E (i.e. such that the image of each point of X under ξ, η lies in E), the bracket $[\xi, \eta]$ also lies in E.*

FR 2. *For each point $z \in X$ and differential form ω of degree 1 at z which vanishes on E, the form $d\omega$ vanishes on $\xi \times \eta$ whenever ξ, η are two vector fields at z which lie in E.*

The equivalence is essentially a triviality. Indeed, assume **FR 1**. Let ω vanish to E. Then

$$\langle d\omega, \xi \times \eta\rangle = -\langle \omega, [\xi, \eta]\rangle - \eta\langle \omega, \xi\rangle + \xi\langle \omega, \eta\rangle.$$

By assumption the right-hand side is 0 when evaluated at z. Conversely, assume **FR 2**. Let ξ, η be two vector fields at z lying in E. If $[\xi, \eta](z)$ is not in E, then we see immediately from a local product representation and the Hahn-Banach theorem that there exists a differential form ω of degree 1 defined on a neighborhood of z which is 0 on E_z and non-zero on $[\xi, \eta](z)$, thereby contradicting the above formula.

We shall now give a third condition equivalent to the above two, and actually, we shall not refer to **FR 2** any more. We remark merely that in the finite dimensional case, it is easy to prove that when a differential form ω satisfies condition **FR 2**, then $d\omega$ can be expressed locally in a neighborhood of each point as a finite sum

$$d\omega = \sum \gamma_i \wedge \omega_i$$

where γ_i and ω_i are of degree 1 and each ω_i vanishes on E. We leave this as an exercise to the reader.

Let E be a tangent subbundle over X. We shall say that E is **integrable** at a point x_0 if there exists a submanifold Y of X containing x_0 such that the tangent map of the inclusion

$$j\colon\ Y \to X$$

induces a VB-isomorphism of TY with the subbundle E restricted to Y. Equivalently, we could say that for each point $y \in Y$, the tangent map

$$T_y j\colon\ T_y Y \to T_y X$$

induces a toplinear isomorphism of $T_y Y$ on E_y. Note that our condition defining integrability is local at x_0. We say that E is **integrable** if it is integrable at every point.

Using the functoriality of vector fields, and their relations under tangent maps and the bracket product, we see at once that if E is integrable, then it satisfies **FR 1**. Indeed, locally vector fields having their values in E are related to vector fields over Y under the inclusion mapping.

Frobenius' theorem asserts the converse.

Theorem 1.1. *Let X be a manifold of class C^p ($p \geqq 2$) and let E be a tangent subbundle over X. Then E is integrable if and only if E satisfies condition* **FR 1**.

The proof of Frobenius' theorem will be carried out by analyzing the situation locally and reducing it to the standard theorem for ordinary differential equations. Thus we now analyze the condition **FR 1** in terms of its local representation.

Suppose that we work locally, over a product $U \times V$ of open subsets of Banach spaces $\mathbf{E}$ and $\mathbf{F}$. Then the tangent bundle $T(U \times V)$ can be written in a natural way as a direct sum. Indeed, for each point (x, y) in $U \times V$ we have

$$T_{(x,y)}(U \times V) = T_x(U) \times T_y(V).$$

One sees at once that the collection of fibers $T_x(U) \times 0$ (contained in $T_x(U) \times T_y(V)$) forms a subbundle which will be denoted by $T_1(U \times V)$ and will be called the **first factor** of the tangent bundle. One could define $T_2(U \times V)$ similarly, and

$$T(U \times V) = T_1(U \times V) \oplus T_2(U \times V).$$

A subbundle E of $T(X)$ is integrable at a point $z \in X$ if and only if there exists an open neighborhood W of z and an isomorphism

$$\varphi\colon\ U \times V \to W$$

of a product onto W such that the composition of maps

$$T_1(U \times V) \xrightarrow{\text{inc.}} T(U \times V) \xrightarrow{T\varphi} T(W)$$

induces a VB-isomorphism of $T_1(U \times V)$ onto $E|W$ (over φ). Denoting by φ_y the map of U into W given by $\varphi_y(x) = \varphi(x, y)$, we can also express the integrability condition by saying that $T_x\varphi_y$ should induce a toplinear isomorphism of $\mathbf{E}$ onto $E_{\varphi(x,y)}$ for all (x, y) in $U \times V$. We note that in terms of our local product structure, $T_x\varphi_y$ is nothing but the partial derivative $D_1\varphi(x, y)$.

Given a subbundle of $T(X)$, and a point in the base space X, we know from the definition of a subbundle in terms of a local product decomposition that we can find a product decomposition of an open neighborhood of this point, say $U \times V$, such that the point has coordinates (x_0, y_0) and such that the subbundle can be written in the form of an exact sequence

$$0 \to U \times V \times \mathbf{E} \xrightarrow{\tilde{f}} U \times V \times \mathbf{E} \times \mathbf{F}$$

with the map

$$f(x_0, y_0)\colon \mathbf{E} \to \mathbf{E} \times \mathbf{F}$$

equal to the canonical embedding of $\mathbf{E}$ on $\mathbf{E} \times 0$. For a point (x, y) in $U \times V$ the map $f(x, y)$ has two components $f_1(x, y)$ and $f_2(x, y)$ into $\mathbf{E}$ and $\mathbf{F}$ respectively. Taking a suitable VB-automorphism of $U \times V \times \mathbf{E}$ if necessary, we may assume without loss of generality that $f_1(x, y)$ is the identity. We now write $f(x, y) = f_2(x, y)$. Then

$$f\colon U \times V \to L(\mathbf{E}, \mathbf{F})$$

is a morphism (of class C^{p-1}) which describes our subbundle completely.

We shall interpret condition **FR 1** in terms of the present situation. If

$$\xi\colon U \times V \to \mathbf{E} \times \mathbf{F}$$

is the local representation of a vector field over $U \times V$, we let ξ_1 and ξ_2 be its projections on $\mathbf{E}$ and $\mathbf{F}$ respectively. Then ξ lies in the image of $\tilde{f}$ if and only if

$$\xi_2(x, y) = f(x, y)\xi_1(x, y)$$

for all (x, y) in $U \times V$, or in other words, if and only if ξ is of the form

$$\xi(x, y) = (\xi_1(x, y), f(x, y)\xi_1(x, y))$$

for some morphism (of class C^{p-1})

$$\xi_1\colon U \times V \to \mathbf{E}.$$

We shall also write the above condition symbolically, namely

$$\xi = (\xi_1, f \cdot \xi_1). \tag{1}$$

If ξ, η are the local representations of vector fields over $U \times V$, then the reader will verify at once from the local definition of the bracket (Proposition 1.3 of Chapter V) that $[\xi, \eta]$ lies in the image of $\tilde{f}$ if and only if

$$Df(x, y) \cdot \xi(x, y) \cdot \eta_1(x, y) = Df(x, y) \cdot \eta(x, y) \cdot \xi_1(x, y)$$

or symbolically,

$$Df \cdot \xi \cdot \eta_1 = Df \cdot \eta \cdot \xi_1. \tag{2}$$

We have now expressed all the hypotheses of Theorem 1.1 in terms of local data, and the heart of the proof will consist in proving the following result.

Theorem 1.2. *Let U, V be open subsets of Banach spaces $\mathbf{E}$, $\mathbf{F}$ respectively. Let*

$$f\colon U \times V \to L(\mathbf{E}, \mathbf{F})$$

be a C^r-morphism ($r \geqq 1$). Assume that if

$$\xi_1, \eta_1\colon U \times V \to \mathbf{E}$$

are two morphisms, and if we let

$$\xi = (\xi_1, f \cdot \xi_1) \quad \text{and} \quad \eta = (\eta_1, f \cdot \eta_1)$$

then relation (2) above is satisfied. Let (x_0, y_0) be a point of $U \times V$. Then there exists open neighborhoods U_0, V_0 of x_0, y_0 respectively, contained in U, V, and a unique morphism $\alpha\colon U_0 \times V_0 \to V$ such that

$$D_1\alpha(x, y) = f(x, \alpha(x, y))$$

and $\alpha(x_0, y) = y$ for all (x, y) in $U_0 \times V_0$.

We shall prove Theorem 1.2 in §3. We now indicate how Theorem 1.1 follows from it. We denote by α_y the map $\alpha_y(x) = \alpha(x, y)$, viewed as a map of U_0 into V. Then our differential equation can be written

$$D\alpha_y(x) = f(x, \alpha_y(x)).$$

We let

$$\varphi\colon U_0 \times V_0 \to U \times V$$

be the map $\varphi(x, y) = (x, \alpha_y(x))$. It is obvious that $D\varphi(x_0, y_0)$ is a toplinear isomorphism, so that φ is a local isomorphism at (x_0, y_0). Furthermore, for $(u, v) \in \mathbf{E} \times \mathbf{F}$ we have

$$D_1\varphi(x, y) \cdot (u, v) = (u, D\alpha_y(x) \cdot u) = (u, f(x, \alpha_y(x)) \cdot u)$$

which shows that our subbundle is integrable.

VI, §2. DIFFERENTIAL EQUATIONS DEPENDING ON A PARAMETER

Proposition 2.1. *Let U, V be open sets in Banach spaces* $\mathbf{E}$, $\mathbf{F}$ *respectively. Let J be an open interval of* $\mathbf{R}$ *containing* 0, *and let*

$$g\colon J \times U \times V \to \mathbf{F}$$

be a morphism of class C^r ($r \geqq 1$). Let (x_0, y_0) be a point in $U \times V$. Then there exists open balls J_0, U_0, V_0 centered at 0, x_0, y_0 and contained J, U, V respectively, and a unique morphism of class C^r

$$\beta\colon J_0 \times U_0 \times V_0 \to V$$

such that $\beta(0, x, y) = y$ and

$$D_1\beta(t, x, y) = g\big(t, x, \beta(t, x, y)\big)$$

for all $(t, x, y) \in J_0 \times U_0 \times V_0$.

Proof. This follows from the existence and uniqueness of local flows, by considering the ordinary vector field on $U \times V$

$$G\colon J \times U \times V \to \mathbf{E} \times \mathbf{F}$$

given by $G(t, x, y) = \big(0, g(t, x, y)\big)$. If $B(t, x, y)$ is the local flow for G, then we let $\beta(t, x, y)$ be the projection on the second factor of $B(t, x, y)$. The reader will verify at once that β satisfies the desired conditions. The uniqueness is clear.

Let us keep the initial condition y fixed, and write

$$\beta(t, x) = \beta(t, x, y).$$

From Chapter IV, §1, we obtain also the differential equation satisfied by β in its second variable:

Proposition 2.2. *Let notation be as in Proposition* 2.1, *and with y fixed, let $\beta(t, x) = \beta(t, x, y)$. Then $D_2\beta(t, x)$ satisfies the differential equation*

$$D_1D_2\beta(t, x) \cdot v = D_2g\big(t, x, \beta(t, x)\big) \cdot v + D_3g\big(t, x, \beta(t, x)\big) \cdot D_2\beta(t, x) \cdot v,$$

for every $v \in \mathbf{E}$.

Proof. Here again, we consider the vector field as in the proof of Proposition 2.1, and apply the formula for the differential equation satisfied by $D_2\beta$ as in Chapter IV, §1.

VI, §3. PROOF OF THE THEOREM

In the application of Proposition 2.1 to the proof of Theorem 1.2, we take our morphism g to be

$$g(t, z, y) = f(x_0 + tz, y) \cdot z$$

with z in a small ball $\mathbf{E}_0$ around the origin in $\mathbf{E}$, and y in V. It is convenient to make a translation, and without loss of generality we can assume that $x_0 = 0$ and $y_0 = 0$. From Proposition 2.1 we then obtain

$$\beta\colon J_0 \times \mathbf{E}_0 \times V_0 \to V$$

with initial condition $\beta(0, z, y) = y$ for all $z \in \mathbf{E}_0$, satisfying the differential equation

$$D_1\beta(t, z, y) = f\big(tz, \beta(t, z, y)\big) \cdot z.$$

Making a change of variables of type $t = as$ and $z = a^{-1}x$ for a small positive number a, we see at once that we may assume that J_0 contains 1, provided we take $\mathbf{E}_0$ sufficiently small. As we shall keep y fixed from now on, we omit it from the notation, and write $\beta(t, z)$ instead of $\beta(t, z, y)$. Then our differential equation is

$$D_1\beta(t, z) = f\big(tz, \beta(t, z)\big) \cdot z. \tag{3}$$

We observe that if we knew the existence of α in the statement of Theorem 1.2, then letting $\beta(t, z) = \alpha(x_0 + tz)$ would yield a solution of our differential equation. Thus the uniqueness of α follows. To prove its existence, we start with β and contend that the map

$$\alpha(x) = \beta(1, x)$$

has the required properties for small $|x|$. To prove our contention it will suffice to prove that

$$D_2\beta(t, z) = tf\big(tz, \beta(t, z)\big) \tag{4}$$

because if that relation holds, then

$$D\alpha(x) = D_2\beta(1, x) = f\big(x, \beta(1, x)\big) = f\big(x, \alpha(x)\big)$$

which is precisely what we want.

From Proposition 2.2, we obtain for any vector $v \in \mathbf{E}$,

$$D_1 D_2 \beta(t, z) \cdot v = tD_1 f(tz, \beta(t, z)) \cdot v \cdot z + D_2 f(tz, \beta(t, z)) \cdot D_2\beta(t, z) \cdot v \cdot z + f(tz, \beta(t, z)) \cdot v.$$

We now let $k(t) = D_2\beta(t, z) \cdot v - tf(tz, \beta(t, z)) \cdot v$. Then one sees at once that $k(0) = 0$ and we contend that

$$Dk(t) = D_2 f(tz, \beta(t, z)) \cdot k(t) \cdot z. \tag{5}$$

We use the main hypothesis of our theorem, namely relation (2), in which we take ξ_1 and η_1 to be the fields v and z respectively. We compute Df using the formula for the partial derivatives, and apply it to this special case. Then (5) follows immediately. It is a linear differential equation satisfied by $k(t)$, and by Corollary 1.7 of Chapter IV, we know that the solution 0 is the unique solution. Thus $k(t) = 0$ and relation (4) is proved. The theorem also.

VI, §4. THE GLOBAL FORMULATION

Let X be a manifold. Let F be a tangent subbundle. By an **integral manifold** for F, we shall mean an injective immersion

$$f\colon Y \to X$$

such that at every point $y \in Y$, the tangent map

$$T_y f\colon T_y Y \to T_{f(y)} X$$

induces a toplinear isomorphism of $T_y Y$ on the subspace $F_{f(y)}$ of $T_{f(y)}X$. Thus Tf induces locally an isomorphism of the tangent bundle of Y with the bundle F over $f(Y)$.

Observe that the image $f(Y)$ itself may not be a submanifold of X. For instance, if F has dimension 1 (i.e. the fibers of F have dimension 1), an integral manifold for F is nothing but an integral curve from the theory of differential equations, and this curve may wind around X in such a way that its image is dense. A special case of this occurs if we consider the torus as the quotient of the plane by the subgroup generated by the two unit vectors. A straight line with irrational slope in the plane gets mapped on a dense integral curve on the torus.

If Y is a submanifold of X, then of course the inclusion $j\colon Y \to X$ is an injective immersion, and in this case, the condition that it be an integral manifold for F simply means that $T(Y) = F|Y$ (F restricted to Y).

We now have the local uniqueness of integral manifolds, corresponding to the local uniqueness of integral curves.

Theorem 4.1. *Let Y, Z be integral submanifolds of X for the subbundle F of TX, passing through a point x_0. Then there exists an open neighborhood U of x_0 in X, such that*

$$Y \cap U = Z \cap U.$$

Proof. Let U be an open neighborhood of x_0 in X such that we have a chart

$$U \to V \times W$$

with

$$x_0 \mapsto (y_0, w_0),$$

and Y corresponds to all points (y, w_0), $y \in V$. In other words, Y corresponds to a factor in the product in the chart. If V is open in $\mathbf{F}_1$ and W open in $\mathbf{F}_2$, with $\mathbf{F}_1 \times \mathbf{F}_2 = \mathbf{E}$, then the subbundle $\mathbf{F}$ is represented by the projection

$$\begin{array}{c} V \times W \times \mathbf{F}_1 \\ \downarrow \\ V \times W \end{array}$$

Shrinking Z, we may assume that $Z \subset U$. Let $h\colon Z \to V \times W$ be the restriction of the chart to Z, and let $h = (h_1, h_2)$ be represented by its two components. By assumption, $h'(x)$ maps $\mathbf{E}$ into $\mathbf{F}_1$ for every $x \in Z$. Hence h_2 is constant, so that $h(Z)$ is contained in the factor $V \times \{w_0\}$. It follows at once that $h(Z) = V_1 \times \{w_0\}$ for some open V_1 in V, and we can shrink U to a product $V_1 \times W_1$ (where W_1 is a small open set in W containing w_0) to conclude the proof.

We wish to get a maximal connected integral manifold for an integrable subbundle F of TX passing through a given point, just as we obtained a maximal integral curve. For this, it is just as easy to deal with the nonconnected case, following Chevalley's treatment in his book on *Lie Groups*. (Note the historical curiosity that vector bundles were invented about a year after Chevalley published his book, so that the language of vector bundles, or the tangent bundle, is absent from Chevalley's presentation. In fact, Chevalley used a terminology which now appears terribly confusing for the notion of a tangent subbundle, and it will not be repeated here!)

We give a new manifold structure to X, depending on the integrable tangent subbundle F, and the manifold thus obtained will be denoted by

X_F. This manifold has the same set of points as X. Let $x \in X$. We know from the local uniqueness theorem that a submanifold Y of X which is at the same time an integral manifold for F is locally uniquely determined. A chart for this submanifold locally at x is taken to be a chart for X_F. It is immediately verified that the collection of such charts is an atlas, which defines our manifold X_F. (We lose one order of differentiability.) The identity mapping

$$j\colon X_F \to X$$

is then obviously an injective immersion, satisfying the following universal properties.

Theorem 4.2. *Let F be an integrable tangent subbundle over X. If*

$$f\colon Y \to X$$

is a morphism such that $Tf\colon TY \to TX$ maps TY into F, then the induced map

$$f_F\colon Y \to X_F$$

(same values as f but viewed as a map into the new manifold X_F) is also a morphism. Furthermore, if f is an injective immersion, then f_F induces an isomorphism of Y onto an open subset of X_F.

Proof. Using the local product structure as in the proof of the local uniqueness Theorem 4.1, we see at once that f_F is a morphism. In other words, locally, f maps a neighborhood of each point of Y into a submanifold of X which is tangent to F. If in addition f is an injective immersion, then from the definition of the charts on X_F, we see that f_F maps Y bijectively onto an open subset of X_F, and is a local isomorphism at each point. Hence f_F induces an isomorphism of Y with an open subset of X_F, as was to be shown.

Corollary 4.3. *Let $X_F(x_0)$ be the connected component of X_F containing a point x_0. If $f\colon Y \to X$ is an integral manifold for F passing through x_0, and Y is connected, then there exists a unique morphism*

$$h\colon Y \to X_F(x_0)$$

making the following diagram commutative:

$$\begin{array}{ccc} Y & \xrightarrow{h} & X_F(x_0) \\ & {\scriptstyle f}\searrow \quad \swarrow{\scriptstyle j} & \\ & X & \end{array}$$

and h induces an isomorphism of Y onto an open subset of $X_F(x_0)$.

Proof. Clear from the preceding discussion.

Note the general functorial behavior of the integral manifold. If

$$g\colon X \to X'$$

is an isomorphism, and F is an integrable tangent subbundle over X, then $F' = (Tg)(F) = g_*F$ is an integrable bundle over X'. Then the following diagram is commutative:

$$\begin{array}{ccc} X_F & \xrightarrow{g_F} & X'_{F'} \\ {\scriptstyle j}\downarrow & & \downarrow{\scriptstyle j'} \\ X & \xrightarrow[g]{} & X' \end{array}$$

The map g_F is, of course, the map having the same values as g, but viewed as a map on the manifold X_F.

VI, §5. LIE GROUPS AND SUBGROUPS

It is not our purpose here to delve extensively into Lie groups, but to lay the groundwork for their theory. For more results, we refer the reader to texts on Lie groups, differential geometry, and also to the paper by W. Graeub [Gr 61]. Although seemingly written to apply only to the finite dimensional case, this paper holds essentially in its entirety for the Banach case (and Hilbert case when dealing with Riemannian metrics), and is written on foundations corresponding to those of the present book.

By a **group manifold**, or a **Lie group** G, we mean a manifold with a group structure, that is a law of composition and inverse,

$$\tau\colon G \times G \to G \qquad \text{and} \qquad G \to G$$

which are morphisms. Thus each $x \in G$ gives rise to a left translation

$$\tau^x\colon G \to G$$

such that $\tau^x(y) = xy$.

When dealing with groups, we shall have to distinguish between isomorphisms in the category of manifolds, and isomorphisms in the category of group manifolds, which are also group homomorphisms. Thus we shall use prefixes, and speak of group manifold isomorphism, or manifold isomorphism as the case may be. We abbreviate these by GM-isomorphism or M-isomorphism. We see that left translation is an M-isomorphism, but not a GM-isomorphism.

Let e denote the origin (unit element) of G. If $v \in T_eG$ is a tangent vector at the origin, then we can translate it, and we obtain a map

$$(x, v) \mapsto \tau_*^x v = \xi_v(x)$$

which is easily verified to be a VB-isomorphism

$$G \times T_eG \to TG$$

from the product bundle to the tangent bundle of G. This is done at once using charts. Recall that T_eG can be viewed as a Banachable space, using any local trivialization of G at e to get a toplinear isomorphism of T_eG with the standard Banachable space on which G is modeled. Thus we see that the tangent bundle of a Lie group is trivializable.

A vector field ξ over G is called **left invariant** if $\tau_*^x \xi = \xi$ for all $x \in G$. Note that the map

$$x \mapsto \xi_v(x)$$

described above is a left invariant vector field, and that the association

$$v \mapsto \xi_v$$

obviously establishes a linear isomorphism between T_eG and the vector space of left invariant vector fields on G. The space of such vector fields will be denoted by $\mathfrak{g}$ or $\mathfrak{l}(G)$, and will be called the **Lie algebra** of G, because of the following results.

Proposition 5.1. *Let ξ, η be left invariant vector fields on G. Then $[\xi, \eta]$ is also left invariant.*

Proof. This follows from the general functorial formula

$$\tau_*^x[\xi, \eta] = [\tau_*^x \xi, \tau_*^x \eta] = [\xi, \eta].$$

Under the linear isomorphism of T_eG with $\mathfrak{l}(G)$, we can view $\mathfrak{l}(G)$ as a Banachable space. By a **Lie subalgebra** of $\mathfrak{l}(G)$ we shall mean a closed subspace $\mathfrak{h}$ which splits, and having the property that if $\xi, \eta \in \mathfrak{h}$, then $[\xi, \eta] \in \mathfrak{h}$ also.

Note. In the finite dimensional case, every subspace is closed and splits, so that only this last condition about the bracket product need be mentioned explicitly.

Let G, H be Lie groups. A map

$$f\colon H \to G$$

will be called a **homomorphism** if it is a group homomorphism and a morphism in the category of manifolds. Such a homomorphism induces a continuous linear map

$$T_ef = f_*\colon T_eH \to T_eG,$$

and it is clear that it also induces a corresponding linear map

$$\mathfrak{l}(H) \to \mathfrak{l}(G),$$

also denoted by f_*. Namely, if $v \in T_eH$ and ξ_v is the left invariant vector field on H induced by v, then

$$f_*\xi_v = \xi_{f_*v}.$$

The general functorial property of related vector fields applies to this case, and shows that the induced map

$$f_*\colon \mathfrak{l}(H) \to \mathfrak{l}(G)$$

is also a Lie algebra homomorphism, namely for $\xi, \eta \in \mathfrak{l}(H)$ we have

$$f_*[\xi, \eta] = [f_*\xi, f_*\eta].$$

Now suppose that the homomorphism $f\colon H \to G$ is also an immersion at the origin of H. Then by translation, one sees that it is an immersion at every point. If in addition it is an injective immersion, then we shall say that f is a **Lie subgroup** of G. We see that in this case, f induces a splitting injection

$$f_*\colon \mathfrak{l}(H) \to \mathfrak{l}(G).$$

The image of $\mathfrak{l}(H)$ in $\mathfrak{l}(G)$ is a Lie subalgebra of $\mathfrak{l}(G)$.

In general, let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{l}(G)$ and let F_e be the corresponding subspace of T_eG. For each $x \in G$, let

$$F_x = \tau_*^x F_e.$$

Then F_x is a split subspace of T_xG, and using local charts, it is clear that the collection $F = \{F_x\}$ is a subbundle of TG, which is left invariant. Furthermore, if

$$f\colon H \to G$$

is a homomorphism which is an injective immersion, and if $\mathfrak{h}$ is the image of $\mathfrak{l}(H)$, then we also see that f is an integral manifold for the subbundle F. We shall now see that the converse holds, using Frobenius' theorem.

Theorem 5.2. *Let G be a Lie group, $\mathfrak{h}$ a Lie subalgebra of $\mathfrak{l}(G)$, and let F be the corresponding left invariant subbundle of TG. Then F is integrable.*

Proof. I owe the proof to Alan Weinstein. It is based on the following lemma.

Lemma 5.3. *Let X be a manifold, let ξ, η be vector fields at a point x_0, and let F be a subbundle of TX. If $\xi(x_0) = 0$ and ξ is contained in F, then $[\xi, \eta](x_0) \in F$.*

Proof. We can deal with the local representations, such that $X = U$ is open in $\mathbf{E}$, and F corresponds to a factor, that is

$$TX = U \times \mathbf{F}_1 \times \mathbf{F}_2 \quad \text{and} \quad F = U \times \mathbf{F}_1.$$

We may also assume without loss of generality that $x_0 = 0$. Then $\xi(0) = 0$, and $\xi\colon U \to \mathbf{F}_1$ may be viewed as a map into $\mathbf{F}_1$. We may write

$$\xi(x) = A(x)x,$$

with a morphism $A\colon U \to L(\mathbf{E}, \mathbf{F}_1)$. Indeed,

$$\xi(x) = \int_0^1 \xi'(tx)\,dt \cdot x,$$

and $A(x) = \mathrm{pr}_1 \circ \int_0^1 \xi'(tx)\,dt$, where pr_1 is the projection on $\mathbf{F}_1$. Then

$$\begin{aligned}[\xi, \eta](x) &= \eta'(x)\xi(x) - \xi'(x)\eta(x)\\ &= \eta'(x)A(x)x - A'(x)\cdot x \cdot \eta(x) - A(x)\cdot\eta(x),\end{aligned}$$

whence

$$[\xi, \eta](0) = A(0)\eta(0).$$

Since $A(0)$ maps $\mathbf{E}$ into $\mathbf{F}_1$, we have proved our lemma.

Back to the proof of the proposition. Let ξ, η be vector fields at a point x_0 in G, both contained in the invariant subbundle F. There exist invariant vector fields ξ_0 and η_0 and x_0 such that

$$\xi(x_0) = \xi_0(x_0) \quad \text{and} \quad \eta(x_0) = \eta_0(x_0).$$

Let

$$\xi_1 = \xi - \xi_0 \quad \text{and} \quad \eta_1 = \eta - \eta_0.$$

Then ξ_1, η_1 vanish at x_0 and lie in F. We get:

$$[\xi, \eta] = \sum_{i,j} [\xi_i, \eta_j].$$

The proposition now follows at once from the lemma.

Theorem 5.4. *Let G be a Lie group, let $\mathfrak{h}$ be a Lie subalgebra of $\mathfrak{l}(G)$, and let F be its associated invariant subbundle. Let*

$$j\colon H \to G$$

be the maximal connected integral manifold of F passing through e. Then H is a subgroup of G, and $j\colon H \to G$ is a Lie subgroup of G. The association between $\mathfrak{h}$ and $j\colon H \to G$ establishes a bijection between Lie subalgebras of $\mathfrak{l}(G)$ and Lie subgroups of G.

Proof. Let $x \in H$. The M-isomorphism τ^x induces a VB-isomorphism of F onto itself, in other words, F is invariant under τ^x_*. Furthermore, since H passes through e, and xe lies in H, it follows that $j\colon H \to G$ is also the maximal connected integral manifold of F passing through x. Hence x maps H onto itself. From this we conclude that if $y \in H$, then $xy \in H$, and there exists some $y \in H$ such that $xy = e$, whence $x^{-1} \in H$. Hence H is a subgroup. The other assertions are then clear.

If H is a Lie subgroup of G, belonging to the Lie algebra $\mathfrak{h}$, and F is the associated integrable left invariant tangent subbundle, then the integral manifold for F passing through a given point x is simply the translation xH, as one sees from first functorial principles.

When $\mathfrak{h}$ is 1-dimensional, then it is easy to see that the Lie subgroup is in fact a homomorphic image of an integral curve

$$\alpha\colon \mathbf{R} \to G$$

which is a homomorphism, and such that $\alpha'(0) = v$ is any vector in T_eG which is the value at e of a non-zero element of $\mathfrak{h}$. Changing this vector merely reparametrizes the curve. The integral curve may coincide with the subgroup, or it comes back on itself, and then the subgroup is essentially a circle. Thus the integral curve need not be equal to the subgroup. However, locally near $t = 0$, they do coincide. Such an integral curve is called a **one-parameter subgroup** of G.

Using Theorem 1.5 of Chapter V, it is then easy to see that if the Lie algebra of a connected Lie group G is commutative, then G itself is commutative. One first proves this for elements in a neighborhood of the origin, using 1-parameter subgroups, and then one gets the statement globally by expressing G as a union of products

$$UU \cdots U,$$

where U is a symmetric connected open neighborhood of the unit element. All of these statements are easy to prove, and belong to the first chapter of a book on Lie groups. Our purpose here is merely to lay the general foundations essentially belonging to general manifold theory.

Warning. The group of differential automorphisms of a finite dimensional manifold is "infinite dimensional" but usually not a Lie group, because multiplication is usually continuous only in each variable separately. For an analysis of this, also in the context of H^p (Sobolev) spaces, cf. Ebin and Marsden [EbM 70].

PART II

Metrics, Covariant Derivatives, and Riemannian Geometry

CHAPTER VII

Metrics

In our discussion of vector bundles, we put no greater structure on the fibers than that of topological vector space (of the same category as those used to build up manifolds). One can strengthen the notion so as to include the metric structure, and we are thus led to consider Hilbert bundles, whose fibers are Hilbert spaces.

Aside from the definitions, and basic properties, we deal with two special topics. On the one hand, we complete our uniqueness theorem on tubular neighborhoods by showing that when a Riemannian metric is given, a tubular neighborhood can be straightened out to a metric one. Secondly, we show how a Riemannian metric gives rise in a natural way to a spray, and thus how one recovers geodesics. The fundamental 2-form is used to identify the vector fields and 1-forms on the tangent bundle, identified with the cotangent bundle by the Riemannian metric.

We assume throughout that our manifolds are Hausdorff and are sufficiently differentiable so that all our statements make sense. (For instance, when dealing with sprays, we take $p \geqq 3$.)

Of necessity, we shall use the standard spectral theorem for (bounded) symmetric operators. A self-contained treatment will be given in the appendix.

VII, §1. DEFINITION AND FUNCTORIALITY

For Riemannian geometry, we shall deal with a Hilbertable vector space, that is a topological vector space which is complete, and whose topology can be defined by the norm associated with a bilinear form, which is symmetric and positive definite. All facts needed in the sequel concerning Hilbert spaces can be found in the Appendix.

It turns out that some basic properties have only to do with a weaker property of the space **E** on which a manifold is modeled, namely that the Banach space **E** is self dual, via a symmetric non-singular bilinear form. Thus we only assume this property until more is needed. We recall that such a form is a continuous bilinear map

$$(v, w) \mapsto \langle v, w \rangle \qquad \text{of} \quad \mathbf{E} \times \mathbf{E} \to \mathbf{R}$$

such that $\langle v, w \rangle = \langle w, v \rangle$ for all $v, w \in \mathbf{E}$, and the corresponding map of **E** into the dual space $L(\mathbf{E})$ is a topological isomorphism.

Examples. Of course, the standard positive definite scalar product on Euclidean space provides the easiest (in some sense) example of a self dual vector space. But the physicists are interested in $\mathbf{R}^4$ with the scalar product such that the square of a vector (x, y, z, t) is $x^2 + y^2 + z^2 - t^2$. This scalar product is non-singular. For one among many nice applications of the indefinite case, cf. for instance [He 84] and [Gu 91], dealing with Huygens' principle.

We consider $L^2_{\text{sym}}(\mathbf{E})$, the vector space of continuous bilinear forms

$$\lambda\colon \mathbf{E} \times \mathbf{E} \to \mathbf{R}$$

which are symmetric. If x is fixed in **E**, then the continuous linear form $\lambda_x(y) = \lambda(x, y)$ is represented by an element of **E** which we denote by Ax, where A is a continuous linear map of **E** into itself. The symmetry of λ implies that A is symmetric, that is we have

$$\lambda(x, y) = \langle Ax, y \rangle = \langle x, Ay \rangle$$

for all $x, y \in \mathbf{E}$. Conversely, given a symmetric continuous linear map $A\colon \mathbf{E} \to \mathbf{E}$ we can define a continuous bilinear form on **E** by this formula. Thus $L^2_{\text{sym}}(\mathbf{E})$ is in bijection with the set of such operators, and is itself a Banach space, the norm being the usual operator norm. Suppose **E** is a Hilbert space, and in particular, **E** is self dual.

The subset of $L^2_{\text{sym}}(\mathbf{E})$ consisting of those forms corresponding to symmetric **positive definite operators** (by definition such that $A \geqq \epsilon I$ for some $\epsilon > 0$) will be called the **Riemannian** of **E** and be denoted by $\text{Ri}(\mathbf{E})$. Forms λ in $\text{Ri}(\mathbf{E})$ are called positive definite. The associated operator A of such a form is invertible, because its spectrum does not contain 0 and the continuous function $1/t$ is invertible on the spectrum.

In general, suppose only that **E** is self dual. The space $L^2_{\text{sym}}(\mathbf{E})$ contains as an open subset the set of non-singular symmetric bilinear forms, which

we denote by $\text{Met}(\mathbf{E})$, and which we call the set of **metrics or pseudo Riemannian metrics.** In view of the operations on vector bundles (Chapter III, §4) we can apply the functor L^2_{sym} to any bundle whose fibers are self dual. Thus if $\pi\colon E \to X$ is such a bundle, then we can form $L^2_{\text{sym}}(\pi)$. A section of $L^2_{\text{sym}}(\pi)$ will be called by definition a **symmetric bilinear form** on π. A **(pseudo Riemannian) metric** on π (or on E) is defined to be a symmetric bilinear form on π, whose image lies in the open set of metrics at each point. We let $\text{Met}(\pi)$ be the set of metrics on π , which we also call the set of metrics on E, and may denote by $\text{Met}(E)$.

If **E** is a Hilbert space and the image of the section of $L^2_{\text{sym}}(\pi)$ lies in the Riemannian space $\text{Ri}(\pi_x)$ at each point x, in order words, if on the fiber at each point the non-singular symmetric bilinear form is actually positive definite, then we call the metric **Riemannian**. Let us denote a metric by g, so that $g(x) \in \text{Met}(E_x)$ for each $x \in X$, and lies in $\text{Ri}(E_x)$ if the metric is Riemannian. Then $g(x)$ is a non-singular symmetric bilinear form in general, and in the Riemannian case, it is positive definite in addition.

A pair (X, g) consisting of a manifold X and a (pseudo Riemannian) metric g will be called a **pseudo Riemannian manifold**. It will be called a **Riemannian manifold** if the manifold is modeled on a Hilbert space, and the metric is Riemannian.

Observe that the sections of $L^2_{\text{sym}}(\pi)$ form a vector space (abstract) but that the Riemannian metrics do not. They form a convex cone. Indeed, if $a, b > 0$ and g_1, g_2 are two Riemannian metrics, then $ag_1 + bg_2$ is also a Riemannian metric.

Suppose we are given a VB-trivialization of π over an open subset U of X, say

$$\tau\colon \pi^{-1}(U) \to U \times \mathbf{E}.$$

We can transport a given pseudo Riemannian metric g (or rather its restriction to $\pi^{-1}(U)$) to $U \times \mathbf{E}$. In the local representation, this means that for each $x \in U$ we can identify $g(x)$ with a symmetric invertible operator A_x giving rise to the metric. The operator A_x is positive definite in the Riemannian case. Furthermore, the map

$$x \mapsto A_x$$

from U into the Banach space $L(\mathbf{E}, \mathbf{E})$ is a morphism.

As a matter of notation, we sometimes write g_x instead of $g(x)$. Thus if v, w are two vectors in E_x, then $g_x(v, w)$ is a number, and is more convenient to write than $g(x)(v, w)$. We shall also write $\langle v, w \rangle_x$ if the metric g is fixed once for all.

Proposition 1.1. *Let X be a manifold admitting partitions of unity. Let $\pi\colon E \to X$ be a vector bundle whose fibers are Hilbertable vector spaces. Then π admits a Riemannian metric.*

Proof. Find a partition of unity $\{U_i, \varphi_i\}$ such that $\pi|U_i$ is trivial, that is such that we have a trivialization

$$\pi_i\colon \pi^{-1}(U_i) \to U_i \times \mathbf{E}$$

(working over a connected component of X, so that we may assume the fibers toplinearly isomorphic to a fixed Hilbert space $\mathbf{E}$). We can then find a Riemannian metric on $U_i \times \mathbf{E}$ in a trivial way. By transport of structure, there exists a Riemannian metric g_i on $\pi|U_i$ and we let

$$g = \sum \varphi_i g_i.$$

Then g is a Riemannian metric on x.

Let us investigate the functorial behavior of metrics.

Consider a VB-morphism

$$\begin{array}{ccc} E' & \xrightarrow{f} & E \\ {\scriptstyle \pi'}\downarrow & & \downarrow{\scriptstyle \pi} \\ X & \xrightarrow[f_0]{} & Y \end{array}$$

with vector bundles E' and E over X and Y respectively, whose fibers are self dual spaces. Let g be a symmetric bilinear form on π, so that for each $y \in Y$ we have a continuous, bilinear, symmetric map

$$g(y)\colon E_y \times E_y \to \mathbf{R}.$$

Then the composite map

$$E'_x \times E'_x \to E_y \times E_y \to \mathbf{R}$$

with $y = f(x)$ is a symmetric bilinear form on E'_x and one verifies immediately that it gives rise to such a form, on the vector bundle π', which will be denoted by $f^*(g)$. Then f induces a map

$$L^2_{\text{sym}}(f) = f^*\colon L^2_{\text{sym}}(\pi) \to L^2_{\text{sym}}(\pi').$$

Furthermore, if f_x is injective and splits for each $x \in X$, and g is a metric (resp. g is a Riemannian metric in the Hilbert case), then obviously so is $f^*(g)$, and we can view f^* as mapping $\text{Met}(\pi)$ into $\text{Met}(\pi')$ (resp. $\text{Ri}(\pi)$ into $\text{Ri}(\pi')$ in the Riemannian case).

Let X be a manifold modeled on a Hilbertable space and let $T(X)$ be

its tangent bundle. By abuse of language, we call a metric on $T(X)$ also a metric on X and write $\text{Met}(X)$ instead of $\text{Met}(T(X))$. Similarly, we write $\text{Ri}(X)$ instead of $\text{Ri}(T(X))$.

Let $f\colon X \to Y$ be an immersion. Then for each $x \in X$, then linear map

$$T_x f\colon T_x(X) \to T_{f(x)}(Y)$$

is injective, and splits, and thus we obtain a contravariant map

$$f^*\colon \text{Ri}(Y) \to \text{Ri}(X),$$

each Riemannian metric on Y inducing a Riemannian metric on X.

A similar result applies in the pseudo Riemannian case. If (Y, g) is Riemannian, and f is merely of class C^1 but not necessarily an immersion, then the pull back $f^*(g)$ is not necessarily positive definite, but is merely what we call **semipositive**. In general, if (X, h) is pseudo Riemannian and $h(v, v) \geqq 0$ for all $v \in T_x X$, all x, then (X, h) is called **semi Riemannian**. Thus the pull back of a semi Riemannian metric is semi Riemannian.

For a major result concerning Riemannian embeddings of manifolds in Euclidean space, see Nash [Na 56], followed by Moser [Mo 61], as well as the exposition I gave in [La 61]. Even though dealing a priori with finite dimensional manifolds, the imbedding problem is essentially concerned with the infinite dimensional manifold of Riemannian metrics. The problem partly amounts to obtaining an inverse mapping theorem in a context more complicated than that of Banach spaces, namely Frechet spaces, when all C^p norms intervene, for $p = 1, 2, \ldots$. Newton approximation is used instead of the shrinking lemma to solve the local isomorphism problem in this case.

The next five sections will be devoted to considerations which apply specifically to the Riemannian case, where positivity plays a central role.

VII, §2. THE HILBERT GROUP

Let $\mathbf{E}$ be a Hilbert space. The group of toplinear automorphisms $\text{Laut}(\mathbf{E})$ contains the group $\text{Hilb}(\mathbf{E})$ of Hilbert automorphisms, that is those toplinear automorphisms which preserve the inner product:

$$\langle Av, Aw \rangle = \langle v, w \rangle$$

for all $v, w \in \mathbf{E}$. We note that A is Hilbertian if and only if $A^*A = I$.

As usual, we say that a linear continuous map $A\colon \mathbf{E} \to \mathbf{E}$ is **symmetric** if $A^* = A$ and that it is **skew-symmetric** if $A^* = -A$. We have a direct sum decomposition of the Banach space $L(\mathbf{E}, \mathbf{E})$ in terms of the two closed subspaces of symmetric and skew-symmetric operators:

$$A = \tfrac{1}{2}(A + A^*) + \tfrac{1}{2}(A - A^*).$$

We denote by Sym(**E**) and Sk(**E**) the Banach spaces of symmetric and skew-symmetric maps respectively. The word **operator** will always mean continuous linear map of **E** into itself.

Proposition 2.1. *For all operators A, the series*

$$\exp(A) = I + A + \frac{A^2}{2!} + \cdots$$

converges. If A commutes with B, then

$$\exp(A + B) = \exp(A)\exp(B).$$

For all operators sufficiently close to the identity I, the series

$$\log(A) = \frac{(A - I)}{1} + \frac{(A - I)^2}{2} + \cdots$$

converges, and if A commutes with B, then

$$\log(AB) = \log(A) + \log(B).$$

Proof. Standard.

We leave it as an exercise to the reader to show that the exponential function gives a C^∞-morphism of $L(\mathbf{E}, \mathbf{E})$ into itself. Similarly, a function admitting a development in power series say around 0 can be applied to the set of operators whose bound is smaller than the radius of convergence of the series, and gives a C^∞-morphism.

Proposition 2.2. *If A is symmetric (resp. skew-symmetric), then* $\exp(A)$ *is symmetric positive definite (resp. Hilbertian). If A is toplinear automorphism sufficiently close to I and is positive definite symmetric (resp. Hilbertian), then* $\log(A)$ *is symmetric (resp. skew-symmetric).*

Proof. The proofs are straightforward. As an example, let us carry out the proof of the last statement. Suppose A is Hilbertian and sufficiently close to I. Then $A^*A = I$ and $A^* = A^{-1}$. Then

$$\log(A)^* = \frac{(A^* - I)}{1} + \cdots$$
$$= \log(A^{-1}).$$

If A is close to I, so is A^{-1}, so that these statements make sense. We now conclude by noting that $\log(A^{-1}) = -\log(A)$. All the other proofs are carried out in a similar fashion, taking a star operator in series term by term, under conditions which insure convergence.

The exponential and logarithm functions give inverse C^∞ mappings between neighborhoods of 0 in $L(\mathbf{E}, \mathbf{E})$ and neighborhoods of I in Laut(**E**). Furthermore, the direct sum decomposition of $L(\mathbf{E}, \mathbf{E})$ into symmetric and skew-symmetric subspaces is reflected locally in a neighborhood of I by a C^∞ direct product decomposition into positive definite and Hilbertian automorphisms. This direct product decomposition can be translated multiplicatively to any toplinear automorphism, because if $A \in$ Laut(**E**) and B is close to A, then

$$B = AA^{-1}B = A\big(I - (I - A^{-1}B)\big)$$

and $(I - A^{-1}B)$ is small. This proves:

Proposition 2.3. *The Hilbert group of automorphisms of* **E** *is a closed submanifold of* Laut(**E**).

In addition to this local result, we get a global one also:

Proposition 2.4. *The exponential map gives a C^∞-isomorphism from the space* Sym(**E**) *of symmetric endomorphisms of* **E** *and the space* Pos(**E**) *of symmetric positive definite automorphisms of* **E**.

Proof. We must construct its inverse, and for this we use the spectral theorem. Given A, symmetric positive definite, the analytic function $\log t$ is defined on the spectrum of A, and thus $\log A$ is symmetric. One verifies immediately that it is the inverse of the exponential function (which can be viewed in the same way). We can expand $\log t$ around a large positive number c, in a power series uniformly and absolutely convergent in an interval $0 < \epsilon \leqq t \leqq 2c - \epsilon$, to achieve our purposes.

Proposition 2.5. *The manifold of toplinear automorphisms of the Hilbert space* **E** *is C^∞-isomorphic to the product of the Hilbert automorphisms and the positive definite symmetric automorphisms, under the mapping*

$$\text{Hilb}(\mathbf{E}) \times \text{Pos}(\mathbf{E}) \to \text{Laut}(\mathbf{E})$$

given by

$$(H, P) \to HP.$$

Proof. Our map is induced by a continuous bilinear map of

$$L(\mathbf{E}, \mathbf{E}) \times L(\mathbf{E}, \mathbf{E})$$

into $L(\mathbf{E}, \mathbf{E})$ and so is C^∞. We must construct an inverse, or in other words express any given toplinear automorphism A in a unique way as a product $A = HP$ where H is Hilbertian, P is symmetric positive definite, and both H, P depend C^∞ on A. This is done as follows. First we note that A^*A is symmetric positive definite (because $\langle A^*Av, v\rangle = \langle Av, Av\rangle$, and furthermore, A^*A is a toplinear automorphism, so that 0 cannot be in its spectrum, and hence $A^*A \geqq \epsilon I > O$ since the spectrum is closed). We let

$$P = (A^*A)^{1/2}$$

and let $H = AP^{-1}$. Then H is Hilbertian, because

$$H^*H = (P^{-1})^*A^*AP^{-1} = I.$$

Both P and H depend differentiably on A since all constructions involved are differentiable.

There remains to be shown that the expression as a product is unique. If $A = H_1P_1$ where H_1, P_1 are Hilbertian and symmetric positive definite respectively, then

$$H^{-1}H_1 = PP_1^{-1},$$

and we get $H_2 = PP_1^{-1}$ for some Hilbertian automorphism H_2. By definition,

$$I = H_2^*H_2 = (PP_1^{-1})^*PP_1^{-1}$$

and from the fact that $P^* = P$ and $P_1^* = P_1$, we find

$$P^2 = P_1^2.$$

Taking the log, we find $2 \log P = 2 \log P_1$. We now divide by 2 and take the exponential, thus giving $P = P_1$ and finally $H = H_1$. This proves our proposition.

VII, §3. REDUCTION TO THE HILBERT GROUP

We define a new category of bundles, namely the **Hilbert bundles** over X, denoted by $\mathrm{HB}(X)$. As before, we would denote by $\mathrm{HB}(X, \mathbf{E})$ or $\mathrm{HB}(X, \mathfrak{A})$ those Hilbert bundles whose fiber is a Hilbert space $\mathbf{E}$ or lies in a category $\mathfrak{A}$.

Let $\pi: E \to X$ be a vector bundle over X, and assume that it has a trivialization $\{(U_i, \tau_i)\}$ with trivializing maps

$$\tau_i: \pi^{-1}(U_i) \to U_i \times \mathbf{E}$$

where $\mathbf{E}$ is a Hilbert space, such that each toplinear automorphism $(\tau_j\tau_i^{-1})_x$ is a Hilbert automorphism. Equivalently, we could also say that τ_{ix} is a Hilbert isomorphism. Such a trivialization will be called a **Hilbert trivialization**. Two such trivializations are called **Hilbert-compatible** if their union is again a Hilbert trivialization. An equivalence class of such compatible trivializations constitutes what we call a **Hilbert bundle** over X. Any such Hilbert bundle determines a unique vector bundle, simply by taking the VB-equivalence class determined by the trivialization.

Given a Hilbert trivialization $\{(U_i, \tau_i)\}$ of a vector bundle π over X, we can define on each fiber π_x a Hilbert space structure. Indeed, for each x we select an open set U_i in which x lies, and then transport to π_x the scalar product in $\mathbf{E}$ by means of τ_{ix}. By assumption, this is independent of the choice of U_i in which x lies. Thus in a Hilbert bundle, we can assume that the fibers are Hilbert spaces, not only Hilbertable.

It is perfectly possible that several distinct Hilbert bundles determine the same vector bundle.

Any Hilbert bundle determining a given vector bundle π will be said to be a **reduction of π to the Hilbert group**.

We can make Hilbert bundles into a category, if we take for the **HB-morphisms** the VB-morphisms which are injective and split at each point, and which preserve the metric, again at each point.

Each reduction of a vector bundle to the Hilbert group determines a Riemannian metric on the bundle. Indeed, defining for each $z \in X$ and $v, w \in \pi_x$ the scalar product

$$g_x(v, w) = \langle \tau_{ix}v, \tau_{ix}w\rangle$$

with any Hilbert-trivializing map τ_{ix} such that $x \in U_i$, we get a morphism

$$x \mapsto g_x$$

of X into the sections of $L^2_{\mathrm{sym}}(\pi)$ which are positive definite. We also have the converse.

Theorem 3.1. *Let π be a vector bundle over a manifold X, and assume that the fibers of π are all toplinearly isomorphic to a Hilbert space* $\mathbf{E}$. *Then the above map, from reductions of π to the Hilbert group, into the Riemannian metrics, is a bijection.*

Proof. Suppose that we are given an ordinary VB-trivialization $\{(U_i, \tau_i)\}$ of π. We must construct an HB-trivialization. For each i, let g_i be the Riemannian metric on $U_i \times \mathbf{E}$ transported from $\pi^{-1}(U_i)$ by means of τ_i. Then for each $x \in U_i$, we have a positive definite symmetric operator A_{ix} such that

$$g_{ix}(v, w) = \langle A_{ix}v, w\rangle$$

for all $v, w \in \mathbf{E}$. Let B_{ix} be the square root of A_{ix}. We define the trivialization σ_i by the formula

$$\sigma_{ix} = B_{ix}\tau_{ix}$$

and contend that $\{(U_i, \sigma_i)\}$ is a Hilbert trivialization. Indeed, from the definition of g_{ix}, it suffices to verify that the VB-isomorphism

$$B_i\colon\ U_i \times \mathbf{E} \to U_i \times \mathbf{E}$$

given by B_{ix} on each fiber, carries g_i on the usual metric. But we have, for $v, w \in E$:

$$\langle B_{ix}v, B_{ix}w\rangle = \langle A_{ix}v, w\rangle$$

since B_{ix} is symmetric, and equal to the square root of A_{ix}. This proves what we want.

At this point, it is convenient to make an additional comment on normal bundles.

Let α, β be two Hilbert bundles over the manifold X, and let $f\colon \alpha \to \beta$ be an HB-morphism. Assume that

$$0 \to \alpha \xrightarrow{f} \beta$$

is exact. Then by using the Riemannian metric, there is a natural way of constructing a splitting for this sequence (cf. Chapter III, §5).

Using Theorem 1.2 of the Appendix, we see at once that if $\mathbf{F}$ is a (closed) subspace of a Hilbert space, then $\mathbf{E}$ is the direct sum

$$\mathbf{E} = \mathbf{F} \oplus \mathbf{F}^{\perp}$$

of $\mathbf{F}$ and its orthogonal complement, consisting of all vectors perpendicular to $\mathbf{F}$.

In our exact sequence, we may view f as an injection. For each x we let $\alpha_x^{\perp}$ be the orthogonal complement of α_x in β_x. Then we shall find an exact sequence of VB-morphisms

$$\beta \xrightarrow{h} \alpha \to 0$$

whose kernel is $\alpha^{\perp}$ (set theoretically). In this manner, the collection of orthogonal complements $\alpha_x^{\perp}$ can be given the structure of a Hilbert bundle.

For each x we can write $\beta_x = \alpha_x \oplus \alpha_x^{\perp}$ and we define h_x to be the

projection in this direct sum decomposition. This gives us a mapping $h\colon \beta \to \alpha$, and it will suffice to prove that h is a VB-morphism. In order to do this, we may work locally. In that case, after taking suitable VB-automorphisms over a small open set U of X, we can assume that we deal with the following situation.

Our vector bundle β is equal to $U \times \mathbf{E}$ and α is equal to $U \times \mathbf{F}$ for some subspace $\mathbf{F}$ of $\mathbf{E}$, so that we can write $\mathbf{E} = \mathbf{F} \times \mathbf{F}^{\perp}$. Our HB-morphism is then represented for each x by an injection $f_x\colon \mathbf{F} \to \mathbf{E}$:

$$U \times \mathbf{F} \xrightarrow{f} U \times \mathbf{E}.$$

By the definition of exact sequences, we can find two VB-isomorphisms τ and σ such that the following diagram is commutative:

$$\begin{array}{ccc} U \times \mathbf{F} & \xrightarrow{f} & U \times \mathbf{E} \\ \sigma \downarrow & & \downarrow \tau \\ U \times \mathbf{F} & \longrightarrow & U \times \mathbf{E} \end{array}$$

and such that the bottom map is simply given by the ordinary inclusion of $\mathbf{F}$ in $\mathbf{E}$. We can transport the Riemannian structure of the bundles on top to the bundles on the bottom by means of σ^{-1} and τ^{-1} respectively. We are therefore reduced to the situation where f is given by the simple inclusion, and the Riemannian metric on $U \times \mathbf{E}$ is given by a family A_x of symmetric positive definite operators on $\mathbf{E}$ ($x \in U$). At each point x, we have $\langle v, w\rangle_x = \langle A_x v, w\rangle$. We observe that the map

$$A\colon\ U \times \mathbf{E} \to U \times \mathbf{E}$$

given by A_x on each fiber is a VB-automorphism of $U \times \mathbf{E}$. Let $\mathrm{pr}_{\mathbf{F}}$ be the projection of $U \times \mathbf{E}$ on $U \times \mathbf{F}$. It is a VB-morphism. Then the composite

$$h = \mathrm{pr}_{\mathbf{F}} \circ A$$

gives us a VB-morphism of $U \times \mathbf{E}$ on $U \times \mathbf{F}$, and the sequence

$$U \times \mathbf{E} \xrightarrow{h} U \times \mathbf{F} \to 0$$

is exact. Finally, we note that the kernel of h consists precisely of the orthogonal complement of $U \times \mathbf{F}$ in each fiber. This proves what we wanted.

VII, §4. HILBERTIAN TUBULAR NEIGHBORHOODS

Let **E** be a Hilbert space. Then the open ball of radius 1 is isomorphic to **E** itself under the mapping

$$v \mapsto \frac{v}{(1-|v|^2)^{1/2}},$$

the inverse mapping being

$$w \mapsto \frac{w}{(1+|w|^2)^{1/2}}.$$

If $a > 0$, then any ball of radius a is isomorphic to the unit ball under multiplication by the scalar a (or a^{-1}).

Let X be a manifold, and $\sigma\colon X \to \mathbf{R}$ a function (morphism) such that $\sigma(x) > 0$ for all $x \in X$. Let $\pi\colon E \to X$ be a Hilbert bundle over X. We denote by $E(\sigma)$ the subset of E consisting of those vectors v such that, if v lies in E_x, then

$$|v|_x < \sigma(x).$$

Then $E(\sigma)$ is an open neighborhood of the zero section.

Proposition 4.1. *Let X be a manifold and $\pi\colon E \to X$ a Hilbert bundle. Let $\sigma\colon X \to \mathbf{R}$ be a morphism such that $\sigma(x) > 0$ for all x. Then the mapping*

$$w \to \frac{\sigma(\pi w)w}{(1+|w|^2)^{1/2}}$$

gives an isomorphism of E onto $E(\sigma)$.

Proof. Obvious. The inverse mapping is constructed in the obvious way.

Corollary 4.2. *Let X be a manifold admitting partitions of unity, and let $\pi\colon E \to X$ be a Hilbert bundle over X. Then E is compressible.*

Proof. Let Z be an open neighborhood of the zero section. For each $x \in X$, there exists an open neighborhood V_x and a number $a_x > 0$ such that the vectors in $\pi^{-1}(V_x)$ which are of length $< a_x$ lie in Z. We can find a partition of unity $\{(U_i, \varphi_i)\}$ on X such that each U_i is contained in some $V_{x(i)}$. We let σ be the function

$$\sum a_{x(i)}\varphi_i.$$

Then $E(\sigma)$ is contained in Z, and our assertion follows from the proposition.

Proposition 4.3. *Let X be a manifold. Let $\pi\colon E \to X$ and $\pi_1\colon E_1 \to X$ be two Hilbert bundles over X. Let*

$$\lambda\colon E \to E_1$$

be a VB-*isomorphism. Then there exists an isotopy of* VB-*isomorphisms*

$$\lambda_t\colon E \to E_1$$

with proper domain [0, 1] *such that $\lambda_1 = \lambda$ and λ_0 is an* HB-*isomorphism.*

Proof. We find reductions of E and E_1 to the Hilbert group, with Hilbert trivializations $\{(U_i, \tau_i)\}$ for E and $\{(U_i, \rho_i)\}$ for E_1. We can then factor $\rho_i\lambda\tau_i^{-1}$ as in Proposition 2.5, applied to each fiber map:

$$\begin{array}{ccccc}
U_i \times \mathbf{E} & \longrightarrow & U_i \times \mathbf{E} & \longrightarrow & U_i \times \mathbf{E} \\
\uparrow{\scriptstyle \tau_i} & & \uparrow{\scriptstyle \tau_i} & & \uparrow{\scriptstyle \rho_i} \\
\pi^{-1}(U_i) & \xrightarrow[\lambda_P]{} & \pi(U_i^{-1}) & \xrightarrow[\lambda_H]{} & \pi_1^{-1}(U_i)
\end{array}$$

and obtain a factorization of λ into $\lambda = \lambda_H\lambda_P$ where λ_H is a HB-isomorphism and λ_P is a positive definite symmetric VB-automorphism. The latter form a convex set, and our isotopy is simply

$$\lambda_t = \lambda_H \circ \big(tI + (1+t)\lambda_P\big).$$

(Smooth out the end points if you wish.)

Theorem 4.4. *Let X be a submanifold of Y. Let $\pi\colon E \to X$ and $\pi_1\colon E_1 \to X$ be two Hilbert bundles. Assume that E is compressible. Let $f\colon E \to Y$ and $g\colon E_1 \to Y$ be two tubular neighborhoods of X in Y. Then there exists an isotopy*

$$f_t\colon E \to Y$$

of tubular neighborhoods with proper domain [0, 1] *and there exists an* HB-*isomorphism $\mu\colon E \to E_1$ such that $f_1 = f$ and $f_0 = g\mu$.*

Proof. From Theorem 6.2 of Chapter IV, we know already that there exists a VB-isomorphism λ such that $f \approx g\lambda$. Using the preceding proposition, we know that $\lambda \approx \mu$ where μ is a HB-isomorphism. Thus $g\lambda \approx g\mu$ and by transitivity, $f \approx \mu$, as was to be shown.

Remark. In view of Proposition 4.1, we could of course replace the condition that E be compressible by the more useful condition (in practice) that X admit partitions of unity.

VII, §5. THE MORSE–PALAIS LEMMA

Let U be an open set in some (real) Hilbert space $\mathbf{E}$, and let f be a C^{p+2} function on U, with $p \geqq 1$. We say that x_0 is a **critical point** for f if $Df(x_0) = 0$. We wish to investigate the behavior of f at a critical point. After translations, we can assume that $x_0 = 0$ and that $f(x_0) = 0$. We observe that the second derivative $D^2f(0)$ is a continuous bilinear form on E. Let $\lambda = D^2f(0)$, and for each $x \in \mathbf{E}$ let λ_x be the functional such that $y \mapsto \lambda(x, y)$. If the map $x \mapsto \lambda_x$ is a toplinear isomorphism of $\mathbf{E}$ with its dual space $\mathbf{E}^\vee$, then we say that λ is **non-singular**, and we say that the critical point is **non-degenerate**.

We recall that a local C^p-isomorphism φ at 0 is a C^p-invertible map defined on an open set containing 0.

Theorem 5.1. *Let f be a C^{p+2} function defined on an open neighborhood of 0 in the Hilbert space $\mathbf{E}$, with $p \geqq 1$. Assume that $f(0) = 0$, and that 0 is a non-degenerate critical point of f. Then there exists a local C^p-isomorphism at 0, say φ, and an invertible symmetric operator A such that*

$$f(x) = \langle A\varphi(x), \varphi(x)\rangle.$$

Proof. We may assume that U is a ball around 0. We have

$$f(x) = f(x) - f(0) = \int_0^1 Df(tx)x\,dt,$$

and applying the same formula to Df instead of f, we get

$$f(x) = \int_0^1\int_0^1 D^2f(stx)tx \cdot x\,ds\,dt = g(x)(x, x)$$

where

$$g(x) = \int_0^1\int_0^1 D^2f(stx)t\,ds\,dt.$$

Then g is a C^p map into the Banach space of continuous bilinear maps on $\mathbf{E}$, and even the space of symmetric such maps. We know that this

Banach space is toplinearly isomorphic to the space of symmetric operators on $\mathbf{E}$, and thus we can write

$$f(x) = \langle A(x)x, x\rangle$$

where $A\colon U \to \mathrm{Sym}(\mathbf{E})$ is a C^p map of U into the space of symmetric operators on E. A straightforward computation shows that

$$D^2f(0)(v, w) = \langle A(0)v, w\rangle.$$

Since we assumed that $D^2f(0)$ is non-singular, this means that $A(0)$ is invertible, and hence $A(x)$ is invertible for all x sufficiently near 0.

Theorem 5.1 is then a consequence of the following result, which expresses locally the uniqueness of a non-singular symmetric form.

Theorem 5.2. *Let $A\colon U \to \mathrm{Sym}(\mathbf{E})$ be a C^p map of U into the open set of invertible symmetric operators on $\mathbf{E}$. Then there exists a C^p isomorphism of an open subset U_1 containing 0, of the form*

$$\varphi(x) = C(x)x, \qquad \text{with a } C^p \text{ map} \quad C\colon U_1 \to \mathrm{Laut}(\mathbf{E})$$

such that

$$\langle A(x)x, x\rangle = \langle A(0)\varphi(x), \varphi(x)\rangle = \langle A(0)C(x)x, C(x)x\rangle.$$

Proof. We seek a map C such that

$$C(x)^*A(0)C(x) = A(x).$$

If we let $B(x) = A(0)^{-1}A(x)$, then $B(x)$ is close to the identity I for small x. The square root function has a power series expansion near 1, which is a uniform limit of polynomials, and is C^∞ on a neighborhood of I, and we can therefore take the square root of $B(x)$, so that we let

$$C(x) = B(x)^{1/2}.$$

We contend that this $C(x)$ does what we want. Indeed, since both $A(0)$ and $A(x)$ $\left(\text{or } A(x)^{-1}\right)$ are self-adjoint, we find that

$$B(x)^* = A(x)A(0)^{-1},$$

whence

$$B(x)^*A(0) = A(0)B(x).$$

But $C(x)$ is a power series in $I - B(x)$, and $C(x)^*$ is the same power series in $I - B(x)^*$. The preceding relation holds if we replace $B(x)$ by any power of $B(x)$ (by induction), hence it holds if we replace $B(x)$ by any polynomial in $I - B(x)$, and hence finally, it holds if we replace $B(x)$ by $C(x)$, and thus

$$C(x)^*A(0)C(x) = A(0)C(x)C(x) = A(0)B(x) = A(x).$$

which is the desired relation.

All that remains to be shown is that φ is a local C^p-isomorphism at 0. But one verifies that in fact, $D\varphi(0) = C(0)$, so that what we need follows from the inverse mapping theorem. This concludes the proof of Theorems 5.1 and 5.2.

Corollary 5.3. *Let f be a C^{p+2} function near 0 on the Hilbert space* $\mathbf{E}$, *such that 0 is a non-degenerate critical point. Then there exists a local C^p-isomorphism ψ at 0, and an orthogonal decomposition $\mathbf{E} = \mathbf{F} + \mathbf{F}^\perp$, such that if we write $\psi(x) = y + z$ with $y \in \mathbf{F}$ and $z \in \mathbf{F}^\perp$, then*

$$f(\psi(x)) = \langle y, y\rangle - \langle z, z\rangle.$$

Proof. On a space where A is positive definite, we can always make the toplinear isomorphism $x \mapsto A^{1/2}x$ to get the quadratic form to become the given hermitian product $\langle\,,\rangle$, and similarly on a space where A is negative definite. In general, we use the spectral theorem to decompose $\mathbf{E}$ into a direct orthogonal sum such that the restriction of A to the factors is positive definite and negative definite respectively.

Note. The Morse–Palais lemma was proved originally by Morse in the finite dimensional case, using the Gram–Schmidt orthogonalization process. The elegant generalization and its proof in the Hilbert space case is due to Palais [Pa 69]. It shows (in the language of coordinate systems) that a function near a critical point can be expressed as a quadratic form after a suitable change of coordinate system (satisfying requirements of differentiability). It comes up naturally in the calculus of variations. For instance, one considers a space of paths (of various smoothness) $\sigma\colon [a, b] \to \mathbf{E}$ where $\mathbf{E}$ is a Hilbert space. One then defines a length function (see next section) or the **energy function**

$$f(\sigma) = \int_a^b \langle\sigma'(t), \sigma'(t)\rangle\, dt,$$

and one investigates the critical points of this function, especially its minimum values. These turn out to be the solutions of the variational problem, by definition of what one means by a variational problem. Even if $\mathbf{E}$ is finite dimensional, so a Euclidean space, the space of paths is infinite dimensional. Cf. [Mi 63] and [Pa 63].

VII, §6. THE RIEMANNIAN DISTANCE

Let (X, g) be a Riemannian manifold. For each C^1 curve

$$\gamma\colon [a, b] \to X$$

we define its **length**

$$L_g(\gamma) = L(\gamma) = \int_a^b \langle\gamma'(t), \gamma'(t)\rangle_g^{1/2}\, dt = \int_a^b \|\gamma'(t)\|_g\, dt.$$

The norm is the one associated with the positive definite scalar product, i.e. the Hilbert space norm at each point. We can extend the length to piecewise C^1 paths by taking the sum over the C^1 curves constituting the path. *We assume that X is connected, which is equivalent to the property that any two points can be joined by a piecewise C^1 path.* (If X is connected, then the set of points which can be joined to a given point x_0 by a piecewise C^1 path is immediately verified to be open and closed, so equal to X. The converse, that pathwise connectedness implies connectedness, is even more obvious.)

We define the **g-distance** on X for any two points $x, y \in X$ by:

$\mathrm{dist}_g(x, y)$ = greatest lower bound of $L(\gamma)$ for paths γ in X joining x and y.

When g is fixed throughout, we may omit g from the notation and write simply $\mathrm{dist}(x, y)$. It is clear that dist_g is a semidistance, namely it is symmetric in (x, y) and satisfies the triangle inequality. To prove that it is a distance, we have to show that if $x \neq y$ then $\mathrm{dist}_g(x, y) > 0$. In a chart, there is a neighborhood U of x which contains a closed ball $\bar{\mathbf{B}}(x, r)$ with $r > 0$, and such that y lies outside this closed ball. Then any path between x and y has to cross the sphere $\mathbf{S}(x, r)$. *Here we are using the Hilbert space norm in the chart.* We can also take r so small that the norm in the chart is given by

$$\langle v, w\rangle_{g(x)} = \langle v, A(x)w\rangle,$$

for $v, w \in \mathbf{E}$, and $x \mapsto A(x)$ is a morphism from U into the set of invertible symmetric positive definite operators, such that there exist a number $C_1 > 0$ for which

$$A(x) \geqq C_1 I \qquad \text{for all } x \in \bar{\mathbf{B}}(x, r).$$

We then claim that there exists a constant $C > 0$ depending only on r, such that for any piecewise C^1 path γ between x and a point on the sphere $\mathbf{S}(x, r)$ we have

$$L(\gamma) \geqq Cr.$$

This will prove that $\mathrm{dist}_g(x, y) \geqq Cr > 0$, and will conclude the proof that dist_g is a distance.

By breaking up the path into a sum of C^1 curves, we may assume without loss of generality that our path is such a curve. Furthermore, we may take the interval $[a, b]$ on which γ is defined to be such that $\gamma(b)$ is the first point such that $\gamma(t)$ lies on $\mathbf{S}(x, r)$, and otherwise $\gamma(t) \in \bar{\mathbf{B}}(x, r)$ for $t \in [a, b]$. Let $\gamma(b) = ru$, where u is a unit vector. Write $\mathbf{E}$ as an orthogonal direct sum

$$\mathbf{E} = \mathbf{R}u \perp \mathbf{F},$$

where $\mathbf{F}$ is a closed subspace. Then $\gamma(t) = s(t)u = w(t)$ with $|s(t)| \leqq r$, $s(a) = 0$, $s(b) = r$ and $w(t) \in \mathbf{F}$. Then

$$\begin{aligned} L(\gamma) = \int_a^b \|\gamma'(t)\|_g \, dt &= \int_a^b \langle \gamma'(t), A(\gamma(t))\gamma'(t)\rangle^{1/2} \, dt \\ &\geqq C_1^{1/2} \int_a^b \langle \gamma'(t), \gamma'(t)\rangle^{1/2} \, dt \\ &\geqq C_1^{1/2} \int_a^b |s'(t)| \, dt \quad \text{by Pythagoras} \\ &\geqq C_1^{1/2} r \end{aligned}$$

as was to be shown.

In addition, the above local argument also proves:

Proposition 6.1. *The distance* dist_g *defines the given topology on* X. *Equivalently, a sequence* $\{x_n\}$ *in* X *converges to a point* x *in the given topology if and only if* $\mathrm{dist}_g(x_n, x)$ *converges to* 0.

We conclude this section with some remarks on reparametrization. Let

$$\gamma\colon [a, b] \to X$$

be a piecewise C^1 path in X. To reparametrize γ, we may do so on each subinterval where γ is actually C^1, so assume γ is C^1. Let

$$\varphi\colon [c, d] \to [a, b]$$

be a C^1 map such that $\varphi(c) = a$ and $\varphi(d) = b$. Then $\gamma \circ \varphi$ is C^1, and is called a **reparametrization** of γ. The chain rule shows that

$$L(\gamma \circ \varphi) = L(\gamma).$$

Define the function $s\colon [a, b] \to \mathbf{R}$ by

$$s(t) = \int_a^t \|\gamma(t)\|_g \, dt, \qquad \text{so} \quad s(b) = L = L(\gamma).$$

Then s is monotone and $s(a) = 0$, while $s(b) = L(\gamma)$. Suppose that there is only a finite number of values $t \in [a, b]$ such that $\gamma'(t) = 0$. We may then break up $[a, b]$ into subintervals where $\gamma'(t) \neq 0$ except at the end points of the subintervals. Consider each subinterval separately, and say

$$a < a_1 < b_1 < b$$

with $\gamma'(t) \neq 0$ for $t \in (a_1, b_1)$. Let $s(a_1)$ be the length of the curve over the interval $[a, a_1]$. Define

$$s(t) = s(a_1) + \int_{a_1}^t \|\gamma'(t)\|_g \, dt \qquad \text{for} \quad a_1 \leqq t \leqq b_1.$$

Then s is strictly increasing, and therefore the inverse function $t = \varphi(s)$ is defined over the interval. Thus we can reparametrize the curve by the variable s over the interval $a_1 \leqq t \leqq b_1$, with the variable s satisfying

$$s(a_1) \leqq s \leqq s(b_1).$$

Thus the whole path γ on $[a, b]$ is reparametrized by another path

$$\gamma \circ \varphi\colon [0, L] \to X$$

via a piecewise map $f\colon [0, L] \to [a, b]$, such that

$$\|(\gamma \circ \varphi)'(s)\|_g = 1 \qquad \text{and} \qquad L_0^s(\gamma \circ \varphi) = s.$$

We now define a path $\gamma\colon [a, b] \to X$ to be **parametrized by arc length** if $\|\gamma'(t)\|_g = 1$ for all $t \in [a, b]$. We see that starting with any path γ, with the condition that there is only a finite number of points where $\gamma'(t) = 0$ for convenience, there is a reparametrization of the path by arc length.

Let $f\colon Y \to X$ be a C^p map with $p \geqq 1$. We shall deal with several notions of isomorphisms in different categories, so in the C^p category, we may call f a differential morphism. Suppose (X, g) and (Y, h) are Riemannian manifolds. We say that f is an **isometry**, or a **differential metric isomorphism** if f is a differential isomorphism and $f^*(g) = h$. If f is an isometry, then it is immediate that f preserves distances, i.e. that

$$\mathrm{dist}_g(f(y_1), f(y_2)) = \mathrm{dist}_h(y_1, y_2) \qquad \text{for all} \quad y_1, y_2 \in Y.$$

Note that there is another circumstance of interest with somewhat weaker conditions when $f\colon Y \to X$ is an immersion, so induces an injection $Tf(y)\colon T_yY \to T_{f(y)}X$ for every $y \in Y$, and we can speak of f being a metric immersion if $f^*(g) = h$. It may even happen that f is a local differential isomorphism at each point of y, as for instance if f is covering map. In such a case, f may be a local isometry, but not a global one, whereby f may not preserve distances on all of Y, possibly because two points $y_1 \neq y_2$ may have the same image $f(y_1) = f(y_2)$.

VII, §7. THE CANONICAL SPRAY

We now come back to the pseudo Riemannian case.

Let X be a pseudo Riemannian manifold, modeled on the self dual space $\mathbf{E}$. The scalar product $\langle\, , \rangle$ in $\mathbf{E}$ identifies $\mathbf{E}$ with its dual $\mathbf{E}^\vee$. The metric on X gives a toplinear isomorphism of each tangent space $T_x(X)$ with $T_x^\vee(X)$. If we work locally with $X = U$ open in $\mathbf{E}$ and we make the identification

$$T(U) = U \times \mathbf{E} \qquad \text{and} \qquad T^\vee(U) = U \times E^\vee \approx T(U)$$

then the metric gives a VB-isomorphism

$$h\colon T(U) \to T(U)$$

by means of a morphism

$$g\colon U \to L(\mathbf{E}, \mathbf{E})$$

such that $h(x, v) = (x, g(x)v)$. (In the finite dimensional case, with respect to an orthonormal basis, $g(x)$ is represented by a symmetric matrix $(g_{ij}(x))$, so the notation here fits what's in other books with their g_{ij}.) The scalar product of the metric at each point x is then given by the formula

$$\langle v, w\rangle_x = \langle v, g(x)w\rangle = \langle g(x)v, w\rangle \qquad \text{for} \quad v, w \in \mathbf{E}.$$

For each $x \in U$ we note that $g'(x)$ maps $\mathbf{E}$ into $L(\mathbf{E}, \mathbf{E})$. For $x \in U$ and $u, v \in E$ we write

$$(g'(x)u)(v) = g'(x)u \cdot v = g'(x)(u, v).$$

From the symmetry of g, differentiating the symmetry relation of the scalar product, we find that for all $u, v, w \in \mathbf{E}$,

$$\langle g'(x)u \cdot w, v\rangle = \langle g'(x)u \cdot v, w\rangle.$$

So we can interchange the last two arguments in the scalar product without changing the value.

Observe that locally, the tangent linear map

$$T(h)\colon T(T(U)) \to T(T(U))$$

is then given by

$$T(h)\colon (x, v, u_1, u_2) \mapsto (x, g(x)v, u_1, g'(x)u_1 \cdot v + g(x)u_2).$$

If we pull back the canonical 2-form described in Proposition 7.2 of Chapter V from $T^\vee(U) \approx T(U)$ to $T(U)$ by means of h then its description locally can be written on $U \times \mathbf{E}$ in the following manner.

$$(1) \quad \langle \Omega_{(x,v)}, (u_1, u_2) \times (w_1, w_2)\rangle = \langle u_1, g(x)w_2\rangle - \langle u_2, g(x)w_1\rangle \\ - \langle g'(x)u_1 \cdot v, w_1\rangle + \langle g'(x)w_1 \cdot v, u_1\rangle.$$

From the simple formula giving our canonical 2-form on the cotangent bundle in Chapter V, we see at once that it is nonsingular on $T(U)$. Since h is a VB-isomorphism, it follows that the pull-back of this 2-form to the tangent bundle is also non-singular.

We shall now apply the results of the preceding section. To do so, we construct a 1-form on $T(X)$. Indeed, we have a function **(kinetic energy!)**

$$K\colon T(X) \to \mathbf{R}$$

given by $K(v) = \frac{1}{2}\langle v, v\rangle_x$ if v is in T_x. Then dK is a 1-form. By Proposition 6.1 of Chapter V, it corresponds to a vector field on $T(X)$, and we contend:

Theorem 7.1. *The vector field F on $T(X)$ corresponding to $-dK$ under the canonical 2-form is a spray over X, called the* **canonical spray**.

Proof. We work locally. We take U open in $\mathbf{E}$ and have the double tangent bundle

$$\begin{array}{c} (U \times \mathbf{E}) \times (\mathbf{E} \times \mathbf{E}) \\ \downarrow \\ U \times \mathbf{E} \\ \downarrow \\ U. \end{array}$$

Our function K can be written

$$K(x, v) = \tfrac{1}{2}\langle v, v\rangle_x = \tfrac{1}{2}\langle v, g(x)v\rangle,$$

and dK at a point (x, v) is simply the ordinary derivative

$$DK(x, v)\colon \mathbf{E} \times \mathbf{E} \to \mathbf{R}.$$

The derivative DK is completely described by the two partial derivatives, and we have

$$DK(x, v) \cdot (w_1, w_2) = D_1K(x, v) \cdot w_1 + D_2K(x, v) \cdot w_2.$$

From the definition of derivative, we find

$$D_1K(x, v) \cdot w_1 = \tfrac{1}{2}\langle v, g'(x)w_1 \cdot v\rangle$$

$$D_2K(x, v) \cdot w_2 = \langle w_2, g(x)v\rangle = \langle v, g(x)w_2\rangle.$$

We use the notation of Proposition 3.2 of Chapter IV. We can represent the vector field F corresponding to dK under the canonical 2-form Ω by a morphism $f\colon U \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}$, which we write in terms of its two components:

$$f(x, v) = \big(f_1(x, v), f_2(x, v)\big) = (u_1, u_2).$$

Then by definition:

$$(2)\quad \big\langle \Omega_{(x,v)}, \big(f_1(x, v), f_2(x, v)\big) \times (w_1, w_2)\big\rangle = \langle DK(x, v), (w_1, w_2)\rangle$$
$$= D_1K(x, v) \cdot w_1 + \langle v, g(x)w_2\rangle.$$

Comparing expressions (1) to (2), we find that as functions of w_2 they have only one term on the right side depending on w_2. From the equality of the two expressions, we conclude that

$$\langle f_1(x, v), g(x)w_2\rangle = \langle v, g(x)w_2\rangle$$

for all w_2, and hence that $f_1(x, v) = v$, whence our vector field F is a second order vector field on X.

Again we compare expression (1) and (2), using the fact just proved that $u_1 = f_1(x, v) = v$. Setting the right sides of the two expressions equal to each other, and using $u_2 = f_2(u, v)$, we obtain:

Proposition 7.2. *In the chart U, let $f = (f_1, f_2)\colon U \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}$ represent F. Then $f_2(x, v)$ is the unique vector such that for all $w_1 \in \mathbf{E}$ we have:*

$$\langle f_2(x, v), g(x)w_1\rangle = \tfrac{1}{2}\langle g'(x)w_1 \cdot v, v\rangle - \langle g'(x) \cdot v \cdot v, w_1\rangle.$$

From this one sees that f_2 is homogeneous of degree 2 in the second variable v, in other words that it represents a spray. This concludes the proof of Theorem 7.1.

Remark. Having represented $f_{U,2}(x, v)$ in the chart, we could also represent he associated bilinear map B_U. We shall give the formula for B_U in the context of Theorem 4.2 of Chapter VIII.

CHAPTER VIII

Covariant Derivatives and Geodesics

Throughout this chapter, by a manifold, we shall mean a C^∞ manifold, for simplicity of language. Vector fields, forms and other objects will also be assumed to be C^∞ unless otherwise specified. We let X be a manifold. We denote the **R**-*vector space of vector fields by $\Gamma T(X)$. Observe that $\Gamma T(X)$ is also a module over the ring of functions $\mathfrak{F} = \mathfrak{F}^\infty(X) = \mathrm{Fu}(X)$. We let*

$$\pi\colon TX \to X$$

be the natural map of the tangent bundle onto X.

VIII, §1. BASIC PROPERTIES

By a **covariant derivative** D we mean an **R**-bilinear map

$$D\colon \Gamma T(X) \times \Gamma T(X) \to \Gamma T(X),$$

denoted by $(\xi, \eta) \mapsto D_\xi \eta$, satisfying the two conditions:

COVD 1. (a) In the first variable ξ, $D_\xi \eta$ is Fu-linear.
(b) For a function φ, define $D_\xi \varphi = \xi\varphi = \mathscr{L}_\xi \varphi$ to be the Lie derivative of the function. Then in the second variable η, $D_\xi \eta$ is a derivation. Thus (a) and (b) can be written in the form:

$$D_{\varphi\xi}\eta = \varphi D_\xi \eta \qquad \text{and} \qquad D_\xi(\varphi\eta) = (D_\xi\varphi)\eta + \varphi D_\xi \eta.$$

COVD 2. $D_\xi \eta - D_\eta \xi = [\xi, \eta]$.

Remark. This second condition can be eliminated to give rise to a more general notion, following the ideas of a connection as described at the end of Chapter IV, §3. However, we concentrate here on what we need for some basic results, rather than develop systematically the general theory of connections.

Having defined D_ξ on functions and vector fields, we may extend the definition to all differential forms, or even to multilinear tensor fields. Let ω be in $\Gamma L^r(T(X))$, i.e. ω is a multilinear tensor field on X, *not necessarily alternating.* We define $D_\xi \omega$ by giving its value on vector fields $\eta_1, \ldots, \eta_r$, namely

$$(D_\xi\omega)(\eta_1, \ldots, \eta_r) = \mathscr{L}_\xi(\omega(\eta_1, \ldots, \eta_r)) - \sum_{j=1}^{r} \omega(\eta_1, \ldots, D_\xi\eta_j, \ldots, \eta_r).$$

The definition of D_ξ is such that D_ξ satisfies the derivation property with respect to the $r+1$ variables ω, $\eta_1, \ldots, \eta_r$, that is

$$D_\xi(\omega(\eta_1, \ldots, \eta_r)) = (D_\xi\omega)(\eta_1, \ldots, \eta_r) + \sum_{j=1}^{r} \omega(\eta_1, \ldots, D_\xi\eta_j, \ldots, \eta_r).$$

Recall that $D_\xi = \mathscr{L}_\xi$ on functions, as on the left side of this equation. Looking in a local chart shows that $D_\xi\omega$ is again a multilinear tensor field. It is immediate from the definition that if ω is alternating, then so is $D_\xi\omega$. In particular, D_ξ is a derivation with respect to contractions and it is also a derivation with respect to the wedge product, that is:

COVD 3. $D_\xi(\omega \circ \eta_1) = (D_\xi\omega) \circ \eta_1 + \omega \circ D_\xi\eta_1$.

COVD 4. On the algebra of alternating forms, the covariant derivative D_ξ is a derivation, in the sense that for two forms ω and γ, we have

$$D_\xi(\omega \wedge \gamma) = D_\xi\omega \wedge \gamma + \omega \wedge D_\xi\gamma.$$

The proof comes directly from the definition of the wedge product in Chapter V, §3. In the finite dimensional case, when a form is a sum of decomposable forms, i.e. wedge products of forms of degree 0 and 1, it follows that the above definition is the unique extension of D_ξ to the algebra of differential forms. Furthermore, similarly to the formula of Proposition 5.1 of Chapter V, for the Lie derivative of a form, one has:

COVD 5. $(\mathscr{L}_\xi \omega)(\eta_1, \dots, \eta_r)$

$$= (D_\xi \omega)(\eta_1, \dots, \eta_r) + \sum_{i=1}^{r} \omega(\eta_1, \dots, D_{\eta_i}\xi, \dots, \eta_r),$$

which is an alternative to

$$\mathscr{L}_\xi(\omega(\eta_1, \dots, \eta_r)) = (D_\xi \omega)(\eta_1, \dots, \eta_r) + \sum_{i=1}^{r} \omega(\eta_1, \dots, D_\xi \eta_i, \dots, \eta_r).$$

COVD 6. $(d\omega)(\xi_0, \xi_1, \dots, \xi_r)$

$$= \sum_{i=0}^{r} (-1)^i (D_{\xi_i}\omega)(\xi_1, \dots, \xi_{i-1}, \xi_0, \xi_{i+1}, \dots, \xi_r).$$

Proof. One uses the formulas given in propositions of Chapter V, Proposition 3.2, for $d\omega$, and Proposition 5.1 for the Lie derivative. One replaces brackets $[\beta, \gamma]$ by $D_\beta \gamma - D_\gamma \beta$. The desired formulas drop out. Note that only **COVD 2** has been used in the proof.

Next we give a finite dimensional formula. Recall that a frame of vector fields $\xi_1, \dots, \xi_n$ is such that for each x, $\{\xi_1(x), \dots, \xi_n(x)\}$ is a basis of $T_x X$.

Proposition 1.1. *Let $\{\xi_1, \dots, \xi_n\}$ be a frame of vector fields. Let $\{\lambda_1, \dots, \lambda_n\}$ be the dual frame of 1-forms (so $\lambda_i(\xi_j) = \delta_{ij}$). For any form $\omega \in \mathscr{A}^r(X)$ we have*

$$d\omega = \sum_{i=1}^{n} \lambda_i \wedge D_{\xi_i}\omega.$$

Proof. Let $d'\omega = \sum \lambda_i \wedge D_{\xi_i}\omega$. Then d' defines an anti-derivation of the alternating algebra of forms, that is if $\psi \in \mathscr{A}^q(x)$ for any q, then

$$d'(\omega \wedge \psi) = (d'\omega) \wedge \psi + (-1)^r \omega \wedge d'\psi.$$

Furthermore, $d' = d$ on functions (as is immediately verified), and we verify that $d' = d$ on $\mathscr{A}^1(X)$ as follows:

$$\begin{aligned}(d'\omega)(\xi, \eta) &= \sum (\lambda_i \wedge D_{\xi_i}\omega)(\xi, \eta) \\ &= \sum [\lambda_i(\xi)\langle D_{\xi_i}\omega, \eta\rangle - \lambda_i(\eta)\langle D_{\xi_i}\omega, \xi\rangle] \\ &= \sum [\langle D_{\lambda_i(\xi)}\omega, \eta\rangle - \langle D_{\lambda_i(\eta)}\omega, \xi\rangle] \\ &= \langle D_\xi \omega, \eta\rangle - \langle D_\eta \omega, \xi\rangle \\ &= (d\omega)(\xi, \eta) \qquad \text{by } \textbf{COVD 6},\end{aligned}$$

which concludes the proof for 1-forms. Since 1-forms generate the algebra of forms in the finite dimensional case, the proposition is proved in general.

The above finite dimensional formula won't be used until we meet strictly finite dimensional results, in connection with volume forms and integration. We included it here for completeness of the general formalism. We now return to the general case which may be infinite dimensional.

We can extend the covariant derivative to TX-valued forms i.e. sections of the bundle $L^r(TX, TX)$. If ω is such a section, we define $D_\xi \omega$ by its values on vector fields $\eta_1, \dots, \eta_r$ to be

$$(D_\xi \omega)(\eta_1, \dots, \eta_r) = D_\xi(\omega(\eta_1, \dots, \eta_r)) - \sum_{j=1}^{r} \omega(\eta_1, \dots, D_\xi \eta_j, \dots, \eta_r),$$

so D_ξ satisfies the derivation property with respect to the $r+1$ variables ω, $\eta_1, \dots, \eta_r$. We note that $\omega(\eta_1, \dots, \eta_r) \in \Gamma TX$ is a vector field, so we know how to apply the covariant derivative $D_\xi(\omega(\eta_1, \dots, \eta_r))$ instead of $\mathscr{L}_\xi(\omega(\eta_1, \dots, \eta_r))$ for ordinary **R**-valued forms, in which case $\omega(\eta_1, \dots, \eta_r)$ is a function on X. When ω is TX-valued, we have on the other hand

$$\mathscr{L}_\xi(\omega(\eta_1, \dots, \eta_r)) = [\xi, \omega(\eta_1, \dots, \eta_r)].$$

A local formula will be given in Proposition 2.2.

VIII, §2. SPRAYS AND COVARIANT DERIVATIVES

Let F be a spray over a manifold X. In a chart U, we index geometric objects by U to indicate their representatives in the chart. Thus the representative ξ_U of a vector field over U is a morphism

$$\xi_U\colon U \to \mathbf{E}.$$

Similarly, we have the symmetric bilinear map associated with the spray, and its representative

$$B_U(x) = \tfrac{1}{2} D_2^2 f_{U,2}(x, 0),$$

where $f_{U,2}$ is the second component of the representative for the spray, as described in Chapter IV, §3.

Theorem 2.1. *Given a spray F over X, there exists a unique covariant derivative D such that in a chart U, the derivative is given by the local* *formula*

$$(D_\xi\eta)_U(x) = \eta'_U(x)\xi_U(x) - B_U(x;\xi_U(x),\eta_U(x)).$$

Or, suppressing the index U for simplicity, and thus using ξ, η to denote the local representatives of the vector fields in the chart, we have

$$(D_\xi\eta)(x) = \eta'(x)\xi(x) - B(x;\xi(x),\eta(x))$$

or simply

$$D_\xi\eta = \eta'\cdot\xi - B(\xi,\eta).$$

Proof. Let us define $D_\xi\eta$ over U by the formula of the theorem. It is immediately verified that $D_\xi\eta$ is a vector field over U, and that the association $(\xi,\eta)\mapsto D_\xi\eta$ is a covariant derivative over U: It is Fu(U)-linear in the variable ξ, it is a derivation in the variable η with respect to multiplication by functions, and we have

$$D_\xi\eta - D_\eta\xi = [\xi,\eta].$$

This last property follows from the representation of the bracket in a chart given by Proposition 1.3 of Chapter V. Thus a spray gives rise to a covariant derivative in a chart, in a natural fashion.

We now claim that when the spray is given globally, there exists a unique covariant derivative on the manifold X which has the above representation in a chart. For this we must verify how the local representation changes under a change of chart. Let

$$h\colon U\to V$$

be a C^∞-isomorphism, i.e. a change of chart. Then we claim that the natural image of $D_{\xi_U}\eta_U$ under the change of chart is $D_{\xi_V}\eta_V$, so that we may define $D_\xi\eta$ for any two vector fields on the manifold via the local representations.

In other words, we have to verify that

$$(D_{\xi_V}\eta_V)(h(x)) = h'(x)(D_{\xi_U}\eta_U)(x).$$

But we have

$$\eta_V(h(x)) = h'(x)\eta_U(x),$$

whence by the rule for the derivative of a product, we obtain

$$(\eta_V\circ h)'(x) = h''(x)\eta_U(x) + h'(x)\eta'_U(x).$$

Hence putting $v=\xi_U(x)$, $w=\eta_U(x)$, we get by using the change of variable formula for a spray in a chart, Proposition 3.3 of Chapter IV, together with the fact that $h''(x)$ is a symmetric bilinear map:

$$\begin{aligned}(D_{\xi_V}\eta_V)(h(x)) &= \eta'_V(h(x))h'(x)\xi_U(x) - B_V(h(x);h'(x)v,h'(x)w)\\ &= (\eta_V\circ h)'(x)\xi_U(x) - h''(x)(v,w) - h'(x)B_U(x;v,w)\\ &= h''(x)(w,v) + h'(x)\eta'_U(x)\xi_U(x)\\ &\quad - h''(x)(v,w) - h'(x)B_U(x;v,w)\\ &= h'(x)(\eta'_U(x)\xi_U(x) - B_U(x;v,w))\\ &\quad\text{(appreciate the cancellation!)}\\ &= h'(x)(D_\xi\eta)_U(x),\end{aligned}$$

which proves the change of variable formula, and therefore concludes the proof of Theorem 2.1.

The covariant derivative defined in Theorem 2.1 will be called the **covariant derivative determined by the spray**, or **associated with the spray**. As mentioned previously, one could give a similar definition of a covariant derivative associated to any connection (even without the symmetry condition on the bilinear map).

There is of course an analogous local representation for differential forms as follows.

Proposition 2.2. *Let $\omega\in\Gamma L^r(TX,\mathbf{R})$ or $\Gamma L^r(TX,TX)$. Let ξ, $\eta_1,\dots,\eta_r$ be vector fields over X. If $\omega\in\Gamma L^r(TX,\mathbf{R})$, then in a chart U we have the formula*

$$\begin{aligned}&(D_\xi\omega)_U(\eta_{1U},\dots,\eta_{rU})\\ &\quad= \omega'_U(\xi_U)(\eta_{1U},\dots,\eta_{rU}) + \sum_{j=1}^r \omega_U(\eta_{1U},\dots,B_U(\xi_U,\eta_{jU}),\dots,\eta_{rU}).\end{aligned}$$

If $\omega\in\Gamma L^r(TX,TX)$, then

$$(D_\xi\omega)_U(\eta_{1U},\dots,\eta_{rU}) = \textit{same expression} - B_U(\xi_U,\omega_U(\eta_{1U},\dots,\eta_{rU})).$$

Proof. This comes directly from the definitions in §1. Observe that in applying the definitions, the sum

$$\sum_{j=1}^r \omega_U(\eta_{1U},\dots,\eta'_{1U}\cdot\xi,\dots,\eta_{rU})$$

occurs twice, once with a + sign and once with a − sign, so cancels in the end.

For the limited purposes of this book, we will not need the proposition. It has an analogue for lifts of curves, which we shall discuss briefly at the end of §3.

Converse, from covariant derivatives to sprays

We now wish to discuss the converse of Theorem 2.1, and for this purpose, we have to make general remarks on localization. Let **E** be a Banach space. We say that **E admits cut off functions** if given two positive real numbers $0 < r < s$, there exists a C^∞-function (simply called function) φ such that $\varphi = 1$ on the ball $B_r(0)$ and $\varphi = 0$ on the complement of $B_s(0)$. Given any point $x_0 \in \mathbf{E}$, we may then find similarly a function which is 1 in the ball $B_r(x_0)$ and 0 outside $B_s(x_0)$. If X is a manifold modeled on **E**, then one can then find such cut off functions equal to 1 in a given neighborhood of a point, and 0 outside a slightly larger neighborhood. Manifolds modelled on a Hilbert space, and especially finite dimensional manifolds, admit cut off functions.

Assume that X admits cut off functions. Let E be a vector bundle over X, and let ξ be a section of E. Let $x_0 \in X$. Let φ a cut off function near x_0. Then $\varphi\xi$ is a section of E, having the same values as ξ in a neighborhood of x_0. Suppose that $E = TX$ and that D is a covariant derivative. Then

$$(D_\xi \eta)(x) = (D_{\varphi\xi}\eta)(x)$$

for all x in a sufficiently small neighborhood of x_0, because D is Fu-linear in the first variable. Since φ is constant near x_0, it follows that

$$(\mathscr{L}_\xi \varphi)(x) = 0 \qquad \text{for } x \text{ near } x_0,$$

and it therefore follows also that

$$(D_\xi(\varphi\eta))(x) = (D_\xi \eta)(x)$$

for all x sufficiently close to x_0.

Now given an open neighborhood U_0 of x_0 corresponding to a chart, we pick out off functions φ, ψ near x_0 such that the supports of φ, ψ are contained in U_0, and φ, $\psi = 1$ on an open neighborhood U of x whose closure is contained in U_0. Then U also corresponds to a chart, and we may compute

$$(D_\xi \eta)(x) = \big(D_{\varphi\xi}(\psi\eta)\big)(x) \qquad \text{for} \quad x \in U.$$

Thus the determination of the values of a covariant derivative can be carried out locally in a chart. We still need a criterion when the value of the covariant derivative at a given point depends only on the value of ξ at the given point.

Lemma 2.3. *Let E, F be vector bundles over X, with E finite dimensional and X admitting cut off functions. Let*

$$H\colon \Gamma E \to \Gamma F$$

be a linear map which is Fu(X)*-linear, that is $H(\varphi\xi) = \varphi H(\xi)$ for $\varphi \in$* Fu. *Given a point $x \in X$, the value $H(\xi)(x)$ depends only on the value $\xi(x)$.*

Proof. It suffices to prove that if $\xi(x_0) = 0$ then $H(\xi)(x_0) = 0$. There exists a cut off function φ near x_0 by assumption, so we may give the proof locally. By assumption, there exists a finite number of sections $e_1, \dots, e_r$ of E which form a basis for the sections locally, so there exist functions $\varphi_1, \dots, \varphi_r$ such that

$$\xi = \varphi_1 e_1 + \cdots + \varphi_r e_r$$

locally. Then

$$H(\xi) = \varphi_1 H(e_1) + \cdots + \varphi_r H(e_r).$$

The condition $\xi(x_0) = 0$ is equivalent with the conditions $\varphi_i(x_0) = 0$ for all i. Hence $H(\xi)(x_0) = 0$, thus proving the lemma.

Observe that when we obtain a covariant derivative from a spray, the value of the covariant derivative at a point x depends only on the value of the vector field $\xi(x)$ (a derivative of η however enters). This was clear from the local formula in Theorem 2.1, because for instance $B_U(x; u, w)$ is defined for arbitrary vectors u, w which can then be taken to be the values $\xi_U(x)$ and $\eta_U(x)$ respectively.

Conversely, we are now interested in reversing the procedure. Specifically, let D be a covariant derivative. We assume the existence of cut off functions throughout. In a chart over an open set U in **E**, define

$$(\mathbf{B_U}) \qquad B_U(x; \xi, \eta) = \eta'(x)\xi(x) - (D_{\xi_U}\eta_U)(x).$$

It is immediately verified from the two properties of a covariant derivative that $B_U(x)$ is symmetric in ξ_U, η_U by **COVD 2**, and then $B_U(x)$ is Fu(U)-bilinear in ξ_U, η_U. Given vectors u, $w \in \mathbf{E}$ one wants to define

$$B_U(x)(u, w) = B_U\big(x;\ \xi(x), \eta(x)\big)$$

for any vector fields ξ, η such that $\xi(x) = u$ and $\eta(x) = w$. At this point, we need to know that the value on the right of $(\mathbf{B_U})$ is independent of the vector fields ξ, η chosen so that $\xi(x) = v$ and $\eta(x) = w$. By Lemma 2.3 we can certainly achieve this in the finite dimensional case, and in that case we obtain:

Theorem 2.4. *Assume X finite dimensional. Then the association of a covariant derivative to a spray establishes a bijection between sprays over X and covariant derivatives.*

In practice, Theorem 2.4 is not that useful (and it will NOT be used in this book) because one either starts from a spray to get a covariant derivative, or if one starts from some natural covariant derivative, and one needs the spray, the situation provides the tools to show that a spray can indeed be defined in a natural manner to give the covariant derivative. We shall see an example of this in §4, when we discuss the Riemannian covariant derivative. Furthermore, the finite dimensional device used in Lemma 2.3 has had historically the unfortunate effect of obscuring the natural bilinear map B, thus obscuring a fundamental structure in expositions of differential geometry. Quite generally, connections on any vector bundle give rise to covariant derivatives. These are applicable to many contexts of topology and analysis, see for example [BGV 92], Chapter I, and also for instance [MokSY 93] for an entirely different direction.

VIII, §3. DERIVATIVE ALONG A CURVE AND PARALLELISM

Instead of using vector fields ξ, η we may carry out a similar construction of a differentiation dealing only with curves, as follows. (For arbitrary maps instead of curves, see Eliasson [El 67].) We continue to denote by F a spray over X. Let $\pi\colon TX \to X$ be the tangent bundle, and let

$$\alpha\colon J \to X$$

be a C^1 curve. By a **lift** γ of α to TX we mean a C^1 curve $\gamma\colon J \to TX$ such that $\pi\gamma = \alpha$. We then also say that γ **lies above** α. We denote the set of lifts of α by Lift(α). It is clear that Lift(α) is a vector space over $\mathbf{R}$, and a module over the ring of functions on J. We wish to define $D_{\alpha'}\gamma$ in a way analogous to the way we defined $D_\xi\eta$ for vector fields ξ, η. This is done by the next theorem. As in §2, we let B_U denote the bilinear map associated to the spray in a chart U.

Theorem 3.1. *There exists a unique linear map*

$$D_{\alpha'}\colon \text{Lift}(\alpha) \to \text{Lift}(\alpha)$$

which in a chart U has the expression

$$(D_{\alpha'}\gamma)_U(t) = \gamma_U'(t) - B_U(\alpha(t); \alpha_U'(t), \gamma_U(t)).$$

The map $D_{\alpha'}$ satisfies the derivation property for a C^1 function φ on J:

$$(D_{\alpha'}(\varphi\gamma))(t) = \varphi'(t)(D_{\alpha'}\gamma)(t) + \varphi(t)(D_{\alpha'}\gamma)(t).$$

Remark. In the present context, the local representation γ_U of a curve in $TU = U \times \mathbf{E}$ is taken to be the map on the second component, i.e.

$$\gamma_U\colon J \to \mathbf{E}.$$

Thus $\gamma_U'(t)$ is the ordinary derivative, with values $\gamma_U'(t) \in \mathbf{E}$. Note that in the case of the representation $\alpha_U\colon J \to U$, we have $\alpha_U'(t) \in \mathbf{E}$ also. Thus $\alpha_U'(t)$, $\gamma_U(t)$ and $\gamma_U'(t)$ are "vectors."

Proof of Theorem 3.1. The proof is entirely analogous to the proof for Theorem 2.1, using the local representation of the bilinear map B_U associated with a spray in charts. We have to verify that the formula of Theorem 3.1 transforms in the proper way under a change of charts, i.e. under an isomorphism $h\colon U \to V$. Note that the local representation γ_V of the curve by definition is given by

$$\gamma_V(t) = h'(\alpha_U(t))\gamma_U(t).$$

Therefore by the rule for the derivative of a product, we find:

$$\gamma_V'(t) = h''(\alpha_U(t))(\alpha_U'(t), \gamma_U(t)) + h'(\alpha_U(t), \gamma_U'(t)).$$

Hence using the transformation rule from B_U to B_V, Proposition 3.3 of Chapter IV, we get

$$\begin{aligned}(D_{\alpha'}\gamma)_V(t) &= \gamma_V'(t) - B_V(\alpha(t); \alpha_V'(t), \gamma_V(t))\\ &= h''(\alpha_U(t))(\alpha_U'(t), \gamma_U(t)) + h'(\alpha_U(t))\gamma_U'(t)\\ &\quad - h''(\alpha_U(t))(\alpha_U'(t), \gamma_U(t))\\ &\quad - h'(\alpha_U(t))B_U(\alpha(t), \alpha_U'(t), \gamma_U(t))\\ &= h'(\alpha_U(t))(D_{\alpha'}\gamma)_U(t) \quad \text{(because the } h'' \text{ term cancels!)},\end{aligned}$$

which proves the desired transformation formula for $(D_{\alpha'}\gamma)_U$ in charts. Thus we have proved the existence of $D_{\alpha'}\gamma$ as asserted. Its being a derivation is immediate from the local representation in charts. This concludes the proof of Theorem 3.1.

Corollary 3.2. *Let η be a vector field and suppose $\gamma(t) = \eta(\alpha(t))$, $t \in J$. Let ξ be a vector field on X such that $\alpha'(t_0) = \xi(\alpha(t_0))$ for some $t_0 \in J$. Then*

$$(D_{\alpha'}\gamma)(t_0) = (D_\xi \eta)(\alpha(t_0)).$$

Proof. Immediate from the chain rule and the local representation of Theorem 3.1.

Let $\alpha\colon J \to X$ be a C^2-morphism. We say that a lift $\gamma\colon J \to TX$ of α is **α-parallel** if $D_{\alpha'}\gamma = 0$. In the chart U, this is equivalent to the condition that

$$\gamma_U'(t) = B_U(\alpha_U(t)\,;\,\alpha_U'(t),\,\gamma_U(t)),$$

which defines a first-order linear differential equation for γ_U. From Chapter IV, §3, (3), we conclude:

A curve α is a geodesic for the spray if and only if $D_{\alpha'}\alpha' = 0$, that is, if and only if α' is α-parallel.

Theorem 3.3. *Let $\alpha\colon J \to X$ be a C^2 curve in X. Let $t_0 \in J$. Given $v \in T_{\alpha(t_0)}X$, there exists a unique lift $\gamma_v\colon J \to TX$ which is α-paralled and such that $\gamma_v(t_0) = v$. Let $\mathrm{Par}(\alpha)$ denote the set of α-parallel lifts of α. The map $v \mapsto \gamma_v$ is a linear isomorphism of $T_{\alpha(t_0)}X$ with $\mathrm{Par}(\alpha)$.*

Proof. The existence and uniqueness simply comes from the existence and uniqueness of solutions of differential equations. Note that from the linearity of the equation, the integral curve γ is defined on the whole interval of definition J by Proposition 1.9 of Chapter IV.

Of course, the notion of parallelism is *with respect to the given spray*, which has been left out of the notation. We express the linearity of Theorem 3.3 another way in the next theorem.

Theorem 3.4. *Fix $t_0 \in J$. For $t \in J$ define the map*

$$P^t_{t_0,\alpha} = P^t\colon T_{\alpha(t_0)}X \to T_{\alpha(t)}X \qquad \text{by} \qquad P^t(v) = \gamma(t, v),$$

where $t \mapsto \gamma(t, v)$ is the unique curve in TX which is α-parallel and $\gamma(t_0, v) = v$. Then P^t is a linear isomorphism.

Proof. We must verify that

$$P^t(sv) = sP^t(v) \quad\text{and}\quad P^t(v + w) = P^t(v) + P^t(w) \quad \text{for } s \in \mathbf{R} \text{ and } v, w \in T_xX.$$

But these properties follow at once from the linearity of the differential equation satisfied by γ, and the uniqueness theorem for its solutions with given initial conditions.

The map P_t is called **parallel translation** along α.

Multilinear tensor fields

Instead of dealing with vector fields, we may deal with TX-valued multilinear tensor fields, or $\mathbf{R}$-valued multilinear tensor fields at essentially no extra cost. *Let E denote either TX or $\mathbf{R}$.* We extend $D_{\alpha'}$ to a linear map

$$D_{\alpha'}\colon \mathrm{Lift}(\alpha, L^r(TX, E)) \to \mathrm{Lift}(\alpha, L^r(TX, E))$$

as follows. Let $\omega\colon J \to L^r(TX, E)$ be a lift of $\alpha\colon J \to X$. Let $\eta_1, \dots, \eta_r$ be lifts of α in TX (sometimes called **vector fields along the curve** α). We **define** $D_{\alpha'}\omega$ by its values on $(\eta_1, \dots, \eta_r)$ to be

$$(D_{\alpha'}\omega)(\eta_1, \dots, \eta_r) = D_{\alpha'}(\omega(\eta_1, \dots, \eta_r)) - \sum_{j=1}^r \omega(\eta_1, \dots, D_{\alpha'}\eta_j, \dots, \eta_r).$$

Thus $D_{\alpha'}$ satisfies the Leibniz rule for the derivative of a multifold product with the $r+1$ variables $\omega, \eta_1, \dots, \eta_r$. Note that *if $\eta_1, \dots, \eta_r$ are* α-parallel, so $D_{\alpha'}\eta_j = 0$, then the formula simplifies to

$$(D_{\alpha'}\omega)(\eta_1, \dots, \eta_r) = D_{\alpha'}(\omega(\eta_1, \dots, \eta_r)).$$

We shall obtain a local formula as usual. Given an index j, we define a linear operator $C_{j,B,\alpha}$ of $\Gamma L^r(TX, E)$ into itself by

$$(C_{j,\beta,\alpha}\omega)(\eta_1, \dots, \eta_r) = \omega(\eta_1, \dots, B(\alpha\,;\,\alpha', \eta_j), \dots, \eta_r).$$

Proposition 3.5 (Local Expression). *Let $\omega = \omega_U$, $\eta_j = \eta_{jU}$ etc. represent the respective objects in a chart U, omitting the subscript U to simplify the notation. Then*

$$(D_{\alpha'}\omega)(\eta_1, \dots, \eta_r) = \omega'(\eta_1, \dots, \eta_r) - B(\alpha\,;\,\alpha', \omega(\eta_1, \dots, \eta_r))\delta_{E,TX} + \sum_{j=1}^r \omega(\eta_1, \dots, B(\alpha\,;\,\alpha', \eta_j), \dots, \eta_r)$$

or also

$$D_{\alpha'}\omega = \omega' - B(\alpha\,;\,\alpha', \omega)\delta_{E,TX} + \sum_{j=1}^r C_{j,B,\alpha}\omega,$$

where $\delta_{E,TX} = 1$ if $E = TX$ and 0 if $E = \mathbf{R}$.

This comes from the definition at the end of §1, and the fact that the ordinary derivative

$$(\omega_U(\eta_{1U}, \dots, \eta_{rU}))'$$

in the chart is obtained by the Leibniz rule (suppressing the index U)

$$(\omega(\eta_1,\dots,\eta_r))' = \omega'(\eta_1,\dots,\eta_r) + \sum \omega(\eta_1,\dots,\eta_j',\dots,\eta_r).$$

Corollary 3.6. *Let $E = TX$ or $\mathbf{R}$ as above. Let $\Omega\colon X \to L^r(TX, E)$ be a section (so a tensor field), and let $\omega(t) = \Omega(\alpha(t))$, $t \in J$. Let $t_0 \in J$. Let ξ be a vector field such that $\alpha'(t_0) = \xi(\alpha(t_0))$. Then*

$$(D_{\alpha'}\omega)(t_0) = (D_\xi\Omega)(\alpha(t_0)).$$

Proof. Immediate from the chain rule and the local representation formula.

A lift $\gamma\colon J \to L^r(TX, E)$ is called **α-parallel** if $D_{\alpha'}\gamma = 0$. The local expression in a chart U shows that the condition $D_{\alpha'}\gamma = 0$ is locally equivalent to the condition

$$\gamma' = B(\alpha; \alpha', \gamma) - \sum_{j=1}^{r} C_{j,B,\alpha}\gamma.$$

Of course, we have suppressed the subscript U from the notation. Thus the condition of being α-parallel defines locally an ordinary linear differential equation, and we obtain from the standard existence and uniqueness theorems:

Theorem 3.7. *Let $t_0 \in J$ and $\omega_0 \in \Gamma L^r(T_{\alpha(t_0)}X, E_{\alpha(t_0)})$. There exists a unique curve $\gamma\colon J \to L^r(TX, E)$ which is α-parallel and such that $\gamma(t_0) = \omega_0$. Denote this curve by γ_{ω_0}. The map*

$$\omega_0 \mapsto \gamma_{\omega_0}$$

establishes a linear isomorphism between the Banach space $L^r(T_{\alpha(t_0)}X, E_{\alpha(t_0)})$ and the space of lifts $\mathrm{Lift}(\alpha, L^r(TX, E))$.

We have now reached a point where we have the parallelism analogous to the simplest case of the tangent bundle as in Theorem 3.4.

Theorem 3.8. *Let the notation be as in Theorem 3.7. For $t \in J$ define the map*

$$P^t_{t_0,\alpha} = P^t_\alpha\colon L^r(T_{\alpha(t_0)}X, E_{\alpha(t_0)}) \to L^r(T_{\alpha(t)}X, E_{\alpha(t)})$$

by

$$P^t_\alpha(\omega_0) = \gamma(t, \omega_0),$$

where $t \mapsto \gamma(t, \omega_0)$ is the unique α-parallel lift of α with $\gamma(0, \omega_0) = \omega_0$. Then P^t_α is a linear isomorphism.

Proof. This follows at once from the linearity of the differential equation satisfied by γ, and the uniqueness theorem for its solutions with given initial conditions.

Example. The metric g itself is a symmetric bilinear $\mathbf{R}$-valued tensor to which the above results can be applied.

VIII, §4. THE METRIC DERIVATIVE

Let (X, g) be a pseudo Riemannian manifold. Let

$$\langle v, w\rangle_g = \langle v, w\rangle_{g(x)}$$

denote the scalar product on the tangent bundle, with $v, w \in T_x$ for some x. If ξ, η are vector fields, then $\langle \xi, \eta\rangle_g$ is a function on X, whose value at a point x is

$$\langle \xi(x), \eta(x)\rangle_g = \langle \xi(x), \eta(x)\rangle_{g(x)}.$$

If ζ is a vector field, we denote

$$\zeta\langle \xi, \eta\rangle_g = D_\zeta\langle \xi, \eta\rangle_g = \mathscr{L}_\zeta\langle \xi, \eta\rangle_g.$$

Theorem 4.1. *Let (X, g) be a pseudo Riemannian manifold. There exists a unique covariant derivative D such that for all vector fields ξ, η, ζ we have*

MD 1. $$D_\xi\langle \eta, \zeta\rangle_g = \langle D_\xi\eta, \zeta\rangle_g + \langle \eta, D_\xi\zeta\rangle_g.$$

This covariant derivative is called the **pseudo Riemannian derivative**, *or* **metric derivative**, *or* **Levi-Civita derivative**.

Proof. For the uniqueness, we shall express $\langle D_\xi\eta, \zeta\rangle_g$ entirely in terms of operations which do not involve the derivative D. To do this, we write down the first defining property of a connection for a cyclic permutation of the three variables:

$$\xi\langle \eta, \zeta\rangle_g = \langle D_\xi\eta, \zeta\rangle_g + \langle \eta, D_\xi\zeta\rangle_g,$$
$$\eta\langle \zeta, \xi\rangle_g = \langle D_\eta\zeta, \xi\rangle_g + \langle \zeta, D_\eta\xi\rangle_g,$$
$$\zeta\langle \xi, \eta\rangle_g = \langle D_\zeta\xi, \eta\rangle_g + \langle \xi, D_\zeta\eta\rangle_g.$$

We add the first two relations and subtract the third. Using the second defining property of a covariant derivative, the following property drops out:

MD 2. $$2\langle D_\xi \eta, \zeta\rangle_g = \xi\langle \eta, \zeta\rangle_g + \eta\langle \zeta, \xi\rangle_g - \zeta\langle \xi, \eta\rangle_g + \langle [\xi, \eta], \zeta\rangle_g - \langle [\xi, \zeta], \eta\rangle_g - \langle [\eta, \zeta], \xi\rangle_g.$$

This proves the uniqueness.

As to existence, define $\langle D_\xi \eta, \zeta\rangle_g$ to be $\frac{1}{2}$ of the right side of **MD 2**. If we view ξ, η as fixed, and ζ as variable, then this right side can be checked in a chart to give a continuous linear functional on vector fields. By Proposition 6.1 of Chapter V, such a functional can be represented by a vector, and this vector defines $D_\xi \eta$ at each point of the manifold. Thus $D_\xi \eta$ is itself a vector field. Using the basic property of the bracket product with a function φ:

$$[\xi, \varphi\eta] = \varphi[\xi, \eta] + (\xi\varphi)\eta \qquad \text{and} \qquad [\varphi\xi, \eta] = \varphi[\xi, \eta] - (\eta\varphi)\xi$$

it is routinely verified that $\langle D_\xi \eta, \zeta\rangle_g$ is Fu-linear in its first variable ξ, and also Fu-linear in the third variable ζ. One also verifies routinely that **COVD 2** is also satisfied, whence existence follows and the theorem is proved.

Recall that we defined $D_\xi \omega$ for any multilinear tensor ω. In particular, let $\omega = g$ be the metric. Then the defining property of the metric connection can now be phrased by stating that for all vector fields ξ,

$$D_\xi g = 0.$$

For each vector field η let $\bigvee_g \eta$ or $\eta^\vee$ be the 1-form corresponding to η under the metric, i.e. for all vector fields ζ, $(\bigvee_g \eta)(\zeta) = \langle \eta, \zeta\rangle_g$.

Corollary. *For the metric derivative D and all vector fields ξ, we have the commutation rule*

$$D_\xi \circ \bigvee_g = \bigvee_g \circ D_\xi \qquad \text{or} \qquad D_\xi(\eta^\vee) = (D_\xi \eta)^\vee.$$

Proof. One line:

$$\left(\bigvee_g (D_\xi \eta)\right)(\zeta) = \langle D_\xi \eta, \zeta\rangle_g = D_\xi\langle \eta, \zeta\rangle_g - \langle \eta, D_\xi\rangle_g = D_\xi\left(\bigvee_g \eta\right)(\zeta).$$

Local representation of the metric derivative

From **MD 2**, we derive a local formula in a chart U. *In the next formula, we write ξ, η, ζ: $U \to \mathbf{E}$ for the representatives of vector fields in the chart, instead of the correct ξ_U, η_U, ζ_U.* Omitting the index U simplifies the

notation when U is fixed throughout the discussion. Here

$$g\colon U \to L(\mathbf{E}, \mathbf{E})$$

denotes the operator defining the metric relative to the given non-singular form on E, so that

$$\langle \xi, \eta\rangle_g = \langle g\xi, \eta\rangle = \langle \xi, g\eta\rangle.$$

Observe that in **COVD 2** and **MD 2**, we took the scalar product in the tangent space, but in the next formula, the scalar product $\langle\ ,\ \rangle$ without an index is the one given by our original non-singular symmetric bilinear form on **E**.

MD 3. *Locally in a chart U, the metric derivative is determined by the formula*:

$$2\langle D_\xi \eta, g\zeta\rangle = 2\langle g\zeta, \eta' \cdot \xi\rangle + \langle \eta, g' \cdot \xi \cdot \zeta\rangle + \langle \xi, g' \cdot \eta \cdot \zeta\rangle - \langle \xi, g' \cdot \zeta \cdot \eta\rangle.$$

Proof. We apply **MD 2**. We express a g-scalar product in terms of the standard scalar product, and we use the local representations of the Lie derivative and the bracket from Chapter V, Proposition 1.1 and Proposition 1.3. For instance, we have the local representation

$$\begin{aligned} \xi\langle \eta, \zeta\rangle_g &= \langle \eta, g\zeta\rangle'\xi \\ &= \langle \eta' \cdot \xi, g\zeta\rangle + \langle \eta, g' \cdot \xi \cdot \zeta\rangle + \langle \eta, g\zeta' \cdot \xi\rangle \end{aligned}$$

by using the rule for the derivative of a product. This formula is meant to be evaluated at each point x. Note that $g'(x)$: $\mathbf{E} \times \mathbf{E} \to \mathbf{E}$ is a bilinear map, which is such that, for instance:

$$g'(x) \cdot \xi(x) \cdot \zeta(x) = g'(x)(\xi(x), \zeta(x)).$$

One can work formally without putting the (x) in the notation. Similarly,

$$\begin{aligned} \langle [\xi, \eta], \zeta\rangle_g &= \langle \eta' \cdot \xi - \xi' \cdot \eta, g\zeta\rangle \\ &= \langle g\zeta, \eta' \cdot \xi - \xi' \cdot \eta\rangle \\ &= \langle g\zeta, \eta' \cdot \xi\rangle - \langle g\zeta, \xi' \cdot \eta\rangle. \end{aligned}$$

Thus we can transform each term appearing on the right of **MD 2**. Then all the terms involving g (rather than g') will cancel except two of them which are equal, and add to yield $2\langle g\zeta, \eta' \cdot \xi\rangle$. The remaining terms are those which are shown on the right side of **MD 3**. This concludes the proof.

Remark. Let us denote by D^g the covariant derivative associated with the metric g. Let $c \in \mathbf{R}^+$. Then cg is also a metric, called a **scaling** of g, and it follows immediately from **MD 3** that

$$D^{cg} = D^g,$$

i.e. the covariant derivative is invariant under a scaling of the metric.

Observe that the definition of the metric derivative in Theorem 4.1 is given by a formula, namely **MD 2**, with its local representation **MD 3**. We want to see that the metric derivative is the one associated with a spray. We recall that quadratic maps and symmetric bilinear maps correspond to each other via the formulas

$$Q(v) = B(v, v) \quad \text{and} \quad B(v, w) = \tfrac{1}{2}[Q(v+w) - Q(v) - Q(w)].$$

The next theorem summarizes the situation.

Theorem 4.2. *Let (X, g) be a pseudo Riemannian manifold. There exists a unique spray on X satisfying the following two equivalent conditions.*

MS 1. *In a chart U, the associated bilinear map B_U satisfies the following formula for all $v, w, z \in \mathbf{E}$:*

$$-2\langle B_U(x; v, w), g(x)z\rangle$$
$$= \langle g'(x)\cdot v\cdot z, w\rangle + \langle g'(x)\cdot w\cdot z, v\rangle - \langle g'(x)\cdot z\cdot w, v\rangle.$$

Thus if we let

$$f_{U,2}(x, v) = B_U(x; v, v) \quad \textit{and} \quad f_U(x, v) = (v, f_{U,2}(x, v)),$$

then f_U represents the spray on $TU = U \times \mathbf{E}$.

MS 2. *The covariant derivative associated to the spray is the metric derivative satisfying Theorem 4.1.*

This spray is the same as the canonical spray of Chapter VII, *Theorem* 7.1.

Proof. First observe that B_U as defined by the formula is symmetric in (v, w). The symmetry is built in the sum of the first two terms, and to see that the third term is symmetric, one differentiates with respect to x the formula

$$\langle g(x)z, v\rangle = \langle g(x)v, z\rangle,$$

which merely expresses the symmetry of $g(x)$ itself. Thus we may form the quadratic map $f_{U,2}(x, v) = B_U(x; v, v)$ from the symmetric bilinear map $B_U(x; v, w)$. It follows that f_U as defined represents a spray F_U over TU. At this point, one may argue in two ways to globalize.

Comparing **MD 3** with **MS 1** we see that the covariant derivative on U determined by the spray F_U is precisely the metric derivative. Theorem 2.1 shows that if two sprays determine the same covariant derivative on U then they are equal. If U, V are two charts, then f_U and f_V are the local representatives of sprays F_U and F_V on U and V respectively, which must therefore coincide on $U \cap V$. Hence the family $\{F_U\}$ defines a spray F on X. Once again, Theorem 2.1 and **MD 3** show that covariant derivative determined by F is the metric derivative.

Furthermore, if we substitute $v = w$ (and $z = w_1$) in the chart formula of **MS 1**, thus giving the quadratic expression $f_{U,2}(x, v)$, then one sees that this expression coincides with the chart expression of Proposition 7.2 of Chapter VII, and hence that the spray obtained in a natural way from the metric derivative is equal to the canonical spray of Chapter VII, Theorem 7.1.

Another possibility is to admit Theorems 7.1 and 7.2 of Chapter VII, which already proved the existence of a spray whose quadratic map $f_{U,2}$ is obtained from the symmetric bilinear map B_U as defined in **MS 1**. This gives immediately the existence of a unique spray on X having the representation of **MS 1** in a chart U, and this spray is the canonical spray. That **MS 2** is equivalent to **MS 1** then follows from **MD 3**. This concludes the proof.

The spray of Theorem 4.2 will be called the **metric spray**. Since it is equal to the canonical spray, we really don't need two names for it.

Remark. To connect with other texts, note that in terms of local coordinates, the metric spray is given by a map f_2 satisfying the second order differential equation

$$\frac{d^2x_i}{dt^2} = f_2(x, v) \quad \text{and} \quad v_i = \frac{dx_i}{dt}.$$

As a function of the variable v, the map f is quadratic, and minus its coefficients are functions of x, called the **Christoffel symbols**, Γ^i_{jk}. Thus by definition, the above differential equation is of type

$$\frac{d^2x_i}{dt^2} = -\sum_{j,k} \Gamma^i_{jk}(x)\frac{dx_k}{dt}\frac{dx_j}{dt}.$$

In terms of the standard basis for $\mathbf{R}^n$, the metric is represented by a matrix

$$(g_{ij}(x)),$$

and we let (g^{ij}) be the inverse matrix. Then the formula of Theorem 4.2 can be written in terms of the local coordinates in terms of the Christoffel symbols, namely

$$\Gamma^j_{ik} = \frac{1}{2}\sum_\nu g^{j\nu}\left(\frac{\partial g_{\nu k}}{\partial x_i} + \frac{\partial g_{i\nu}}{\partial x_k} - \frac{\partial g_{ik}}{\partial x_\nu}\right)$$

If I gave priority to fit classical notation, I would have written $-\Gamma_U$ instead of B_U for the bilinear map associated with the spray. However, using the letter B suggests bilinearity, whereas using the letter Γ would suggest the above mess. Besides, using B is more natural for the bilinear map associated to the quadratic map of the second order differential equation, and eliminates a minus sign from that equation.

Theorem 4.3. *Let $\alpha : J \to X$ be a C^2 curve in a Riemannian manifold (X, g). For the metric derivative, and curves $\gamma, \zeta \in \mathrm{Lift}(\alpha, TX)$, we have the formula*

$$\langle \gamma, \zeta\rangle'_g = \langle D_{\alpha'}\gamma, \zeta\rangle_g + \langle \gamma, D_{\alpha'}\zeta\rangle_g.$$

Furthermore, parallel translation is a metric isomorphism. In particular, let $t_0 \in J$. If γ_v, γ_w are the unique α-parallel lifts of α with $\gamma_v(t_0) = v$ and $\gamma_w(t_0) = w$, then for all t,

$$\langle \gamma_v(t), \gamma_w(t)\rangle_g = \langle v, w\rangle_g.$$

Proof. The formula is proved in the same way that the computation proving Theorem 3.1 was parallel to the computation proving Theorem 2.1 (giving the behavior under changes of charts). From the formula, if $D_{\alpha'}\gamma = D_{\alpha'}\zeta = 0$, it follows that $\langle \gamma, \zeta\rangle_g$ is constant, whence the second assertion follows.

Corollary 4.4. *Let φ be a C^2 function on X. Let α be a geodesic for the metric spray. Then*

$$(\varphi \circ \alpha)'' = \langle D_{\alpha'}(\mathrm{grad}\ \varphi) \circ \alpha, \alpha'\rangle_g.$$

Proof. Taking the first derivative of $\varphi \circ \alpha$ yields

$$(\varphi \circ \alpha)'(t) = (d\varphi)(\alpha(t))\alpha'(t) = \langle (\mathrm{grad}\ \varphi)(\alpha(t)), \alpha'(t)\rangle_g.$$

Now take the next derivative using Theorem 4.3 and the fact that $D_{\alpha'}\alpha' = 0$. The desired formula drops out.

VIII, §5. MORE LOCAL RESULTS ON THE EXPONENTIAL MAP

In this section, we give further results on the exponential map obtained from a spray. We follow the same notation as in Chapter IV, §4, and at first we just deal with a spray. We do not need to know whether it comes from a metric or not.

Throughout the section, we let X be a manifold with a spray F.

Instead of looking at the exponential map restricted to the tangent space at a given point, we may consider this map in the neighborhood of a point in the whole tangent bundle. Let $\pi : TX \to X$ be the projection as always. Let $x_0 \in X$, with zero element $0_{x_0} \in T_{x_0}X$. There exists an open neighborhood V of 0_{x_0} in TX on which we can define the map

$$G: V \to X \times X \quad \text{such that} \quad G(v) = (\pi v, \exp_{\pi v}(v)).$$

It is sometimes useful to express this map in a different notation. Specifically, if we denote a point in the tangent bundle by a pair (x, v) if $v \in T_xX$, then

$$G(x, v) = (x, \exp_x(v)).$$

Using a pair (x, v) is certainly the way we would write a point in the tangent bundle as represented in a chart $U \times \mathbf{E}$, with $x \in U$ and $v \in \mathbf{E}$.

Proposition 5.1. *The map G is a local isomorphism at $(x_0, 0)$.*

Proof. The Jacobian matrix of G in a chart is given immediately from Chapter IV, Theorem 4.1 by

$$\begin{pmatrix} \mathrm{id} & \mathrm{id} \\ 0 & \mathrm{id} \end{pmatrix}$$

which is invertible. The inverse mapping theorem concludes the proof.

For the next local results, it is convenient to express certain uniformities in a chart, where we can measure distances uniformly in the model Banach space $\mathbf{E}$, with a given norm. It is irrelevant to know whether this norm has any smoothness properties or not. It will be used just to describe neighborhoods of a vector 0 in the tangent bundle. I found [Mi 63] useful.

Let $x_0 \in X$. For $\epsilon > 0$, we let $\mathbf{E}(\epsilon)$ denote the open ball of elements $v \in \mathbf{E}$ with $|v| < \epsilon$. Arbitrarily small open neighborhoods of $(x_0, 0)$ in a chart for TX are of the form

$$U_0 \times \mathbf{E}(\epsilon),$$

where U_0 is an open neighborhood of x_0 in X, and ϵ is arbitrarily small.

Corollary 5.2. *Given $x_0 \in X$. Let V be an open neighborhood of $(x_0, 0)$ in TX such that G induces an isomorphism of V with its image, and in a chart, for some $\epsilon > 0$,*

$$V = U_0 \times \mathbf{E}(\epsilon).$$

Let W be a neighborhood of x_0 in X such that $G(V) \supset W \times W$. Then:

(1) *Any two points $x, y \in W$ are joined by a unique geodesic in X lying in U_0, and this geodesic depends C^∞ on the pair (x, y). In other words, if $t \mapsto \exp_x(tv)$ $(0 \leqq t \leqq 1)$ is the geodesic joining x and y, with $y = \exp_x(v)$, then the correspondence*

$$(x, v) \leftrightarrow (x, y)$$

is C^∞.

(2) *For each $x \in W$ the exponential $\exp_x$ maps the open set in T_xX represented by $(x, \mathbf{E}(\epsilon))$ isomorphically onto an open set $U(x)$ containing W.*

Proof. The properties are merely an application of the definitions and Proposition 5.1.

The pair (V, W) will be said to constitute a **normal neighborhood** of x_0 in X. Dealing with the pair rather than a single neighborhood is slightly inelegant, but to eliminate one of the neighborhoods requires a little more work, which most of the time is not necessary. It has to do with "convexity" properties, and a theorem of Whitehead [Wh 32]. We shall do the work at the end of this section for the Riemannian case.

In the Riemannian case, given $x \in X$, by a **normal chart** at x we mean an open ball $B_g(x, c)$ such that the exponential map

$$\exp_x\colon \mathbf{B}_g(0_x, c) \to B_g(x, c)$$

is an isomorphism. We call $B_g(x, c)$ a **normal ball**.

We shall need a lemma which gives us the analogue of the commutation rule of partial derivatives in the context of covariant derivatives. Let J_1, J_2 be open intervals, and let

$$\sigma\colon J_1 \times J_2 \to X$$

be a C^2 map. For each fixed $t \in J_2$ we obtain a curve $\sigma\colon J_1 \to X$ such that $\sigma_t(r) = \sigma(r, t)$. We can then take the ordinary partial derivative

$$\partial_1\sigma(r, t) = \sigma_t'(r) = \frac{\partial\sigma}{\partial r}.$$

Similarly, we can define $\partial_2(r, t) = \partial\sigma/\partial t$. Observe that for each t, the curves $r \mapsto \partial_1\sigma(r, t)$ and $r \mapsto \partial_2\sigma(r, t)$ are lifts of $r \mapsto \sigma(r, t)$ in TX.

More generally, let Q be a lift of σ in TX. Then one may apply the covariant derivative with respect to functions of the first variable r, with the various notation

$$(D_{\partial_1\sigma, t}Q_t)(r) = (D_1Q)(r, t) = \frac{DQ}{\partial r}.$$

Similarly, we have $D_2Q(r, t)$.

Lemma 5.3. *We have the rules on lifts of σ to TX:*

(a) *$D_1\partial_2 = D_2\partial_1$; and*

(b) *$\partial_2\langle\partial_1\sigma, \partial_1\sigma\rangle_g = 2\langle D_1\partial_2\sigma, \partial_1\sigma\rangle_g$.*

Proof. Let σ_U represent σ in a chart. Then from Theorem 3.1,

$$D_1\partial_2\sigma_U = \partial_1\partial_2\sigma_U - B_U(\sigma_U; \partial_1\sigma_U, \partial_2\sigma_U).$$

Since B_U is symmetric in the last two arguments, this proves (a). As to (b), we use the metric derivative to yield

$$\partial_2\langle\partial_1\sigma, \partial_1\sigma\rangle_g = 2\langle D_2\partial_1\sigma, \partial_1\sigma\rangle_g,$$

and we use (a) to permute the partials variables on the right, to conclude the proof of (b), and therefore the proof of the lemma.

Let now (X, g) be a pseudo Riemannian manifold. For each $x \in X$ we have the scalar product $\langle v, w\rangle_g = \langle v, w\rangle_{g(x)}$ for $v, w \in T_xX$. Let $c > 0$. The equation

$$\langle v, v\rangle_g = c^2$$

defines a submanifold in T_xX, which may be empty. If the metric is Riemannian, the equation defines what we call a **sphere**. In the case when the metric is pseudo Riemannian, say indefinite in the finite dimensional case, then one thinks of the equation as defining something like a hyperboloid in the vector space T_xX. We can still define the **level "hypersurface"** $\mathbf{S}_g(c)$ to be the set of solutions of the above equation. Even in infinite dimension, we can say that the codimension of this hypersurface is 1. Note that

$$\mathbf{S}_g(rc) = r\mathbf{S}_g(c) \qquad \text{for} \quad r > 0.$$

In a neighborhood of the origin 0_x in T_xX, the exponential map is defined, and gives an isomorphism which may be restricted to $\mathbf{S}_g(c)$ intersected with this neighborhood. The image of this intersection is then a submanifold of a neighborhood of x in X. We look at the geodesics starting at x.

Theorem 5.4. *Let $t \mapsto u(t)$ be a curve in $\mathbf{S}_g(1)$. Let $0 \leqq r \leqq b$ where b is such that the points $ru(t)$ are in the domain of the exponential $\exp_x$. Define*

$$\sigma(r, t) = \exp_x(ru(t)) \qquad \textit{for} \quad 0 \leqq r \leqq b.$$

Then

$$\langle \partial_1\sigma, \partial_1\sigma \rangle_g = \langle u, u \rangle_g = 1.$$

Proof. This is immediate since parallel translation is an isometry by Theorem 4.3.

Corollary 5.5. *Assume (X, g) Riemannian. Let $v \in T_xX$. Suppose $\|v\|_g = r$, with $r > 0$. Also suppose the segment $\{tv\}$ $(0 \leqq t \leqq 1)$ is contained in the domain of the exponential. Let $\alpha(t) = \exp_x(tv)$. Then $L(\alpha) = r$.*

Proof. Special case of the length formula in Theorem 5.4, followed by an integration to get the length.

Remark. The corollary is also valid in the pseudo Riemannian case, if one assume that $v^2 = r^2 > 0$, so the notion of length makes sense for the curve $t \mapsto \exp_x(tv)$.

Lemma 5.6. *Let X be pseudo Riemannian. Let $\sigma\colon J_1 \times J_2 \to X$ be a C^2 map. For each $t \in J_2$ let $\alpha_t(s) = \sigma(s, t)$. Assume that each α_t is a geodesic, and that $\alpha_t'^2$ is independent of t. Then for each $t \in J_2$, the map $s \mapsto \langle \partial_1\sigma, \partial_2\sigma \rangle_g(s, t)$ is constant.*

Proof. Let D be the metric derivative. Then $D_1\partial_1\sigma = 0$ because for a geodesic α, we know that the metric derivative has the property that $D_{\alpha'}\alpha' = 0$. Thus we get

$$\begin{aligned}\partial_1\langle \partial_1\sigma, \partial_2\sigma \rangle_g &= \langle D_1\partial_1\sigma, \partial_2\sigma \rangle_g + \langle \partial_1\sigma, D_1\partial_2\sigma \rangle_g \\ &= \tfrac{1}{2}\partial_2\langle \partial_1\sigma, \partial_1\sigma \rangle_g \quad \text{by the above and Lemma 5.3} \\ &= 0 \qquad\qquad\qquad\quad \text{by hypothesis.}\end{aligned}$$

This concludes the proof.

The next theorem expresses the fact that locally near x, the geodesics are orthogonal to the images of the level sets $\mathbf{S}_g(c)$ under the exponential map.

Theorem 5.7. *Let (X, g) be pseudo Riemannian. Let $x_0 \in X$ and let W be a small open neighborhood of x_0, selected as in Corollary 5.2, with ϵ sufficiently small. Let $x \in W$. Then the geodesics through x are orthogonal to the image of $\mathbf{S}_g(c)$ under $\exp_x$, for c sufficiently small positive.*

Proof. For ϵ sufficiently small positive, the exponential map is defined on $\mathbf{S}_g(r)$ for $0 < r \leqq \epsilon$, and as we have seen, the level sets $\mathbf{S}_g(r)$ are submanifolds of X. Then our assertion amounts to proving that for every curve $u\colon J \to \mathbf{S}_g(1)$ and $0 < r < c$, if we define

$$\sigma(r, t) = \exp(ru(t)),$$

then the two curves

$$t \mapsto \exp_x(r_0(u(t)) \qquad \text{and} \qquad r \mapsto \exp_x(ru(t_0))$$

are orthogonal for any given value (r_0, t_0), which amounts to proving that

$$\langle \partial_1\sigma, \partial_2\sigma \rangle_g = 0.$$

But for $r = 0$, we have $\sigma(0, t) = \exp_x(0) = x$, independent of t. Hence $\partial_2\sigma(0, t) = 0$. We can apply Lemma 5.6 to conclude the proof.

In the Riemannian case, the theorem is known as **Gauss' lemma.** Helgason [He 61] showed in the analytic case that it is valid in the pseudo Riemannian case as well. I followed the proof given in [Mi 63], which I found applicable to the present context without coordinates, and without assuming analyticity.

Convexity

We conclude this section with the more systematic study of convexity, which was bypassed in Corollary 5.2. We shall treat the Riemannian case, which is slightly simpler. So we assume that (X, g) is Riemannian.

We need to know:

Given $x \in X$, there exists $c > 0$ such that if $0 < r < c$, then the geodesic α such that $\alpha(t) = \exp_x(tv)$, with $0 \leqq t \leqq 1$, and $\|v\|_g = r$, is the shortest piecewise C^1 path between x and $\exp_x(v)$.

This will be proved in Theorems 6.2 and 6.4 of the next section. In particular, $\operatorname{dist}_g(x, \exp_x(v)) = r$ for r sufficiently small. As usual, we let:

$$\mathbf{B}_g(0_x, r) = \text{open ball in } T_xX \text{ centered at } 0_x, \text{ of radius } r;$$

$$B_g(x, r) = \text{open ball in } X \text{ centered at } x, \text{ of radius } r;$$

$$\mathbf{S}_g(0_x, r) = \text{sphere of radius } r \text{ in } T_xX, \text{ centered at } 0_x; \text{and}$$

$$S_g(x, r) = \text{sphere of radius } r \text{ in } X, \text{ centered at } x.$$

Here we shall deal only with r sufficiently small.

We define an open set U of X to be **convex** if given $x, y \in U$ there exists a unique geodesic in U joining x to y, and such that the length of the geodesic is $\operatorname{dist}_g(x, y)$. We shall prove Whitehead's theorem [Wh 32] in the form:

Theorem 5.8. *Let (X, g) be a Riemannian manifold. Given $x \in X$, there exists $c > 0$ such that for all r with $0 < r < c$ the open neighborhood $B_g(x, r) = \exp_x \mathbf{B}_g(0_x, r)$ is convex.*

Proof. We need a lemma.

Lemma 5.9. *Given $x \in X$, there exists $c > 0$ such that if $r < c$, and if α is a geodesic in X, tangent to $S_g(x, r)$ at $y = \alpha(t_0)$, then $\alpha(t)$ lies outside $S_g(x, r)$ for $t \neq t_0$ in some neighborhood of t_0.*

Proof. We pick c such that the exponential map $\exp_x$ is a differential isomorphism on $\mathbf{B}_g(0_x, r)$ for all $r < c$ and preserves distances on rays from 0_x to $v \in T_xX$ with $\|v\|_g = r$. Without loss of generality, we can suppose $t_0 = 0$, so $\alpha(0) = y$. We shall view y as variable, so we index α by y. Also we have to look at the other initial condition $\alpha'(0) = u \in T_yY$, so we write $\alpha_{y,u}$ for the geodesic. Now let

$$\eta_{y,u}(t) = \exp_x^{-1} \alpha_{y,u}(t) \qquad \text{and} \qquad f_{y,u}(t) = \eta_{y,u}(t)^2.$$

Then $\eta_{y,u}$ is a curve in the fixed Hilbert space T_xX, so

$$f'_{y,u}(t) = 2\langle \eta'_{y,u}(t), \eta_{y,u}(t)\rangle_{g(x)},$$
$$f''_{y,u}(t) = 2\eta'_{y,u}(t)^2 + 2\langle \eta''_{y,u}(t), \eta_{y,u}(t)\rangle_{g(x)}.$$

Let $h(y, u) = f''_{y,u}(0)$. Then $h(x, u) = 2u^2$, so h_x as a function on T_xX is positive definite. Therefore there exists $c > 0$ such that for $0 < r < c$ and $\|y\|_g = r$ the function h_y is positive definite on T_yY, and in particular $h(y, u) > 0$ for $u^2 \neq 0$. Under the assumption that $\alpha_{y,u}$ is tangent to $S_g(x, r)$ at y, we must have

$$f'_{y,u}(0) = 0 \qquad \text{and} \qquad f''_{y,u}(0) = h(y, u) > 0,$$

whence for sufficiently small $|t|$, we get

$$f_{y,u}(t) > f_{y,u}(0) = \left(\exp_x^{-1} \alpha_{y,x}(0)\right)^2 = \left(\exp_x^{-1}(u)\right)^2 = r^2,$$

which proves the lemma.

We can now conclude the proof of Theorem 5.8. Using Corollary 5.2, we can find $c_1 > 0$ such that putting $W = B_g(x, c_1)$ satisfies the condition of Corollary 5.2. Let $c < c_1$. We show that $r \leqq c$ implies $B_g(x, r)$ is convex. Let $y, z \in B_g(x, r)$. Then by that corollary, there exists a unique geodesic α in the neighborhood V of x joining y and z. As in the lemma, let

$$f(t) = \left(\exp_x^{-1} \alpha(t)\right)^2, \qquad \text{with} \quad a \leqq t \leqq b.$$

It now suffices to prove that $f(t) < r^2$. Suppose $f(t) \geqq r^2$ for some t, and let $t_0 \in [a, b]$ be the maximum of f on this interval, so $f(t_0) \geqq r^2$. Then $t_0 \neq a, b$ so $f'(t_0) = 0$, whence α is tangent to the sphere $S_g(x, r_0)$ where $r_0 = f(t_0)^{1/2}$. The lemma now gives a contradiction, which concludes the proof of Theorem 5.8.

Remark. In the pseudo Riemannian case, with metric g, one has to use an auxiliary Riemannian metric h to apply a similar argument, which makes the proof slightly longer.

VIII, §6. RIEMANNIAN GEODESIC LENGTH AND COMPLETENESS

Throughout this section, we let (X, g) be a Riemannian manifold.

We return to the Riemannian case, where we use the positive definiteness of the metric. In Chapter VII, §6 we defined the length of a piecewise C^1 path. We want to compare the length locally with the length of straight lines in the tangent space at a point, under the exponential map. In the process, we shall see that locally, a geodesic is the shortest path between two points.

Thus let $x_0 \in X$ and let (V, W) be a normal neighborhood as in Corollary 5.2. Let $x \in W$. For each piecewise C^1 path

$$\gamma\colon [a, b] \to U(x) - \{x\},$$

with $U(x)$ being as in Corollary 5.2(2), we can use the fact that the exponential map is invertible and so there exists a unique curve $t \mapsto u(t)$ in T_xM such that $\|u(t)\|_g = 1$ and

$$\gamma(t) = \exp_x(r(t)u(t)) \quad \text{with} \quad 0 < r(t) < \epsilon.$$

In a chart, the vector $r(t)u(t)$ is obtained by the inverse of the exponential map followed by a projection, so in particular, the functions $t \mapsto r(t)$ and $t \mapsto u(t)$ on $[a, b]$ are piecewise C^1. We call these functions the local **polar coordinates** for γ.

Lemma 6.1. *For a piecewise C^1 curve γ: $[a, b] \to U(x) - \{x\}$ as above, we have the inequality*

$$L(\gamma) \geqq |r(b) - r(a)|.$$

Equality holds only if the function $t \mapsto r(t)$ is monotone and the map $t \mapsto u(t)$ is constant.

Proof. Let $\sigma(r, t) = \exp_x(ru(t))$. Then $\gamma(t) = \sigma(r(t), t)$. We have

$$\gamma'(t) = \frac{d\gamma}{dt} = \frac{\partial\sigma}{\partial r} r'(t) + \frac{\partial\sigma}{\partial t}.$$

By the Gauss Lemma Theorem 5.7, we know that $\partial\sigma/\partial r$ and $\partial\sigma/\partial t$ are orthogonal. Since $\|\partial\sigma/\partial r\|_g = 1$ by Lemma 5.4, it follows that

$$\|\gamma'(t)\|_g^2 = |r'(t)|^2 + \left\|\frac{\partial\sigma}{\partial t}\right\|_g^2 \geqq |r'(t)|^2,$$

with equality holding only if $\partial\sigma/\partial t = 0$, or equivalently, $du/dt = 0$. Hence

$$L(\gamma) = \int_a^b \|\gamma'(t)\|_g \, dt \geqq \int_a^b |r'(t)| \, dt \geqq |r(b) - r(a)|;$$

and equality holds only if $t \mapsto r(t)$ is monotone and $t \mapsto u(t)$ is constant. This completes the proof.

Theorem 6.2. *Let (V, W) constitute a normal neighborhood of a point $x_0 \in X$. Let α: $[0, 1] \to V$ be the geodesic (up to reparametrization) in V joining two points of W (namely $\alpha(0)$ and $\alpha(1)$). Let γ: $[0, 1] \to X$ be any other piecewise C^1 path in X joining these two points. Then*

$$L(\alpha) \leqq L(\gamma).$$

If equality holds, then the polar component $t \mapsto v(t)$ for γ is constant, the function $t \mapsto r(t)$ is monotone, and a reparametrization of γ is equal to α.

Proof. Let x, $y \in W$ and let $y = \exp_x(ru)$ with $0 < r < \epsilon$, and $\|u\|_g = 1$. Then for $\delta > 0$ and $0 < \delta < r$ the path γ contains a segment joining the shell $\mathrm{Sh}_g(x, \delta)$ with the shell $\mathrm{Sh}_g(x, r)$ and lying between the two shells. By Lemma 6.1, the length of this segment is $\geqq r - \delta$. Letting δ tend to 0 shows that $L(\gamma) \geqq r$. The same lemma proves the conditions on the polar functions as asserted.

Corollary 6.3. *Let α: $[0, 1] \to X$ be a piecewise C^1 path, parametrized by are length. If $L(\alpha) \leqq L(\gamma)$ for all paths from $\alpha(0)$ to $\alpha(1)$ in X, then α is a geodesic.*

Proof. We can find a partition of $[0, 1]$ such that the image under α of each small interval in the partition is contained in some neighborhood W as in the theorem, and its length is small so the image of the segment is contained in a normal neighborhood. By Theorem 6.2, the path restricted to this segment must be a geodesic. Hence the entire path is a geodesic, as was to be shown.

Let α: $[a, b] \to X$ be a geodesic. We say that α is a **minimal geodesic** if $L(\alpha) \leqq L(\gamma)$ for every path γ joining $\alpha(a)$ and $\alpha(b)$ in X. Theorem 6.2 gives us the existence of minimal geodesics locally. We can then formulate another application. Let $x \in X$. Let dist_g be the Riemannian distance. Let:

$$\mathbf{B}_g(0_x, r),\ \mathbf{S}_g(0_x, r),\ B_g(x, r),\ S_g(x, r)$$

be the open balls and spheres of radius r, centered at 0_x in T_xX and at x in X, respectively. We now know enough to show that $S_g(x, r)$ is the image of $\mathbf{S}_g(0_x, r)$ under the exponential map, and similarly for the open ball, for sufficiently small r.

Theorem 6.4. *Let (X, g) be a Riemannian manifold and let $x \in X$. There exists $c > 0$ such that for all $r < c$ the map $\exp_x$ is defined on $\mathbf{B}_g(0_x, c)$, gives a differential isomorphism*

$$\exp_x\colon \mathbf{B}_g(0_x, r) \to B_g(x, r) \quad \textit{for all } r \textit{ with } \quad 0 < r < c,$$

and also a differential isomorphism

$$\exp_x\colon \mathbf{S}_g(0_x, r) \to S_g(x, r) \quad \textit{for } 0 < r < c.$$

Proof. Immediate from Corollary 5.5 and Theorem 6.2.

Next we consider completeness. Since X is a metric space (in the ordinary sense), with respect to the distance dist_g, the notion of X being complete is standard: every Cauchy sequence for dist_g converges. On the other hand, we can now define another notion of completeness.

We say that (X, g) is **geodesically complete** if and only if the maximal interval of definition of every geodesic in X is all of $\mathbf{R}$. Alternatively, we could say that for each point $x \in X$, the exponential map $\exp_x$ is defined on all of T_x, because under one normalization of the parametrization of a geodesic, it is simply the curve $t \mapsto \exp_x(tv)$ for some $v \in T_xX$. To be systematic, let us consider the following conditions:

COM 1. *As a metric space under* dist_g, *X is complete.*

COM 2. *All geodesics in X are defined on* $\mathbf{R}$.

COM 3. *For every $x \in X$, the exponential $\exp_x$ is defined on all of T_xX.*

COM 4. *For some $x \in X$, the exponential $\exp_x$ is defined on all of T_xX.*

Proposition 6.5. *Each condition implies the next, i.e.*

$$\textbf{COM 1} \Rightarrow \textbf{COM 2} \Rightarrow \textbf{COM 3} \Rightarrow \textbf{COM 4}.$$

Proof. Assume **COM 1**. Let $\alpha\colon J \to X$ be a geodesic parametrized by arc length on some interval, and take J to be maximal in $\mathbf{R}$. By the existence and uniqueness theorem for differential equations, J is open in $\mathbf{R}$, and it will suffice to prove that J is closed, or in other words, that J contains its end points. For $t_1, t_2 \in J$ we have

$$\text{dist}(\alpha(t_1), \alpha(t_2)) \leqq |t_2 - t_1|.$$

Suppose for instance that J is bounded above, and let $\{t_n\}$ be a sequence in J converging to the right end point of J. Then the sequence $\{\alpha(t_n)\}$ is Cauchy by the above inequality, so $\{\alpha(t_n)\}$ converges to a point x_0 by **COM 1**. Then for all n sufficiently large, $\alpha(t_n)$ lies in a small normal neighborhood of x_0, and there is some $\epsilon > 0$, independent of n, such that the geodesic can be extended to an interval of length at least ϵ beyond t_n, thus contradicting the maximality of J, and proving **COM 2**. The subsequent implications are trivial, so the proposition is proved.

We are now interested when geodesic completeness implies completeness. We shall give two criteria for this. One of them is that the manifold has finite dimension, and the other one will be important for its application to conditions on curvature in Chapter IX. The finite dimensional case depends on the next result.

Theorem 6.6 (Hopf–Rinow). *Assume that (X, g) is finite dimensional connected geodesically complete at a point p, that is, $\exp_p$ is defined on T_pX. Then any point in X can be joined to p by a minimal geodesic.*

Proof. I follow here the variation of the proof given in [Mi 63]. Let y be a point with $p \neq y$. Let W be a normal neighborhood of p containing the image of a small ball under the exponential map $\exp_p$. Let $r = \text{dist}(p, y)$, and let δ be small $< r$. Then the shell $\text{Sh}_g(p, \delta) = \text{Sh}(p, \delta)$ is contained in W. Since $\text{Sh}(p, \delta)$ is the image of the sphere of radius δ in T_pX, it follows that $\text{Sh}(p, \delta)$ is compact. Hence there exists a point x_0 on $\text{Sh}(p, \delta)$ which is at minimal g-distance from y, that is

$$\text{dist}(x_0, y) \leqq \text{dist}(x, y) \qquad \text{for all} \quad x \in \text{Sh}(p, \delta).$$

We can write $x_0 = \exp_p(\delta u)$ for some $u \in T_x$ with $\|u\|_g = 1$. Let $\alpha(t) = \exp_p(tu)$. We shall prove that $\exp_p(ru) = y$. We prove this by "continuous induction" on t, as it were. More precisely, we shall prove:

$$(\textbf{dist}_t) \qquad \text{We have } \text{dist}(\alpha(t), y) = r - t \qquad \text{for} \quad \delta \leqq t \leqq r.$$

Taking $t = r$ will prove the theorem. First we note that $(\textbf{dist}_\delta)$ is true. Indeed, every path from p to y intersects the shell $\text{Sh}(p, \delta)$, so

$$(1) \qquad \begin{aligned} \text{dist}(p, y) &= \min_x (\text{dist}(p, x) + \text{dist}(x, y)) \qquad \text{for} \quad x \in \text{Sh}(p, \delta) \\ &= \delta + \min_x \text{dist}(x, y) \\ &= \delta + \text{dist}(x_0, y), \end{aligned}$$

so $(\textbf{dist}_\delta)$ is true. Now "inductively", assume that (dist_t) is true for all $t < r'$, with $\delta \leqq r' \leqq r$. Let r_1 be the least upper bound of such r'. Since the distance dist_g is continuous, it follows that $(\textbf{dist}_{r_1})$ is true, and it suffices to prove that $r_1 = r$. Suppose $r_1 < r$. Pick δ_1 small so we get as usual a spherical shell $\text{Sh}(\alpha(r_1), \delta_1)$ around $\alpha(r_1)$, contained in a normal neighborhood of $\alpha(r_1)$. As in (1), there is a point x_1 on $\text{Sh}(\alpha(r_1), \delta_1)$ at minimal distance from y, and we have the relation as in (1), namely

$$\text{dist}(\alpha(r_1), y) = \delta_1 + \text{dist}(x_1, y).$$

Since $(\textbf{dist}_{r_1})$ is true, we find

$$(2) \qquad \text{dist}(x_1, y) = r - r_1 - \delta_1.$$

We claim that $x_1 = \alpha(r_1 + \delta_1)$. To see this, first observe that

$$\mathrm{dist}(p, x_1) \geqq \mathrm{dist}(p, y) - \mathrm{dist}(x_1, y) = r_1 + \delta_1.$$

But the path consisting of the two minimal geodesics from $p = \alpha(0)$ to $\alpha(r_1)$ and from $\alpha(r_1)$ to x_1 has length $r_1 + \delta_1$, so this path (which is viewed as a broken geodesic) has minimal length, so it is an unbroken geodesic by Corollary 6.3. Hence the path is actually equal to α, so $\alpha(r_1 + \delta_1) = x_1$, and therefore $\mathrm{dist}(x_1, y) = r - (r_1 + \delta_1)$, so $(\mathbf{dist}_{r_1+\delta_1})$ is true, thus concluding the proof of the continuous induction, and also concluding the proof the Hopf–Rinow theorem.

Corollary 6.7. *In the finite dimensional case the four completeness conditions* **COM 1** *through* **COM 4** *are equivalent to a fifth:*

COM 5. *A closed* dist_g*-bounded subset of* X *is compact.*

Proof. Assume **COM 4** with $\exp_{x_0}$ defined on $T_{x_0}X$. Let S be closed and bounded in X. Without loss of generality, we may assume $x_0 \in S$. Let b be a bound for the diameter of S. Then by Theorem 6.6 (Hopf–Rinow), every point of S can be joined to x_0 by a geodesic of length $\leqq b$, so S is contained in the image under $\exp_{x_0}$ of the closed ball of radius b in $T_{x_0}X$, so S is compact, thus proving **COM 5**.

Assume **COM 5**. Let $\{x_n\}$ be a Cauchy sequence in X. Then $\{x_n\}$ lies in a bounded set, whose closure is compact by assumption, so $\{x_n\}$ has a point of accumulation which is actually a limit in X. This proves **COM 1**, and concludes the proof of the corollary.

Remark. In his thesis [McA 65], McAlpin gave the following example which shows a divergence of behavior in the case of infinite dimensional Hilbert manifolds. Let $\mathbf{E}$ be a Hilbert space with orthonormal basis $\{e_n\}$ $(n \geqq 0)$. Let $T\colon \mathbf{E} \to \mathbf{E}$ *be the linear map such that for a vector* $v = \sum x_n e_n \in \mathbf{E}$

$$T\left(\sum x_n e_n\right) = \sum a_n x_n e_n$$

where $a_0 = 1$ and $a_n = 1 + 1/2^n$ for $n \geqq 1$. Then

$$\|v\| \leqq \|Tv\| \leqq \tfrac{3}{2}\|v\|,$$

and therefore T is invertible in $\mathrm{Laut}(\mathbf{E})$. Let $\mathbf{S}$ be the unit sphere in $\mathbf{E}$ and let $X = T(\mathbf{S})$, so X is a submanifold of $\mathbf{E}$, to which we give the induced metric. Let α be a path joining e_0 to $-e_0$ in $\mathbf{S}$. Then $T\alpha$ is a path joining e_0 to $-e_0$ in X, and T is length increasing, that is

$$L(\alpha) \leqq L(T\alpha).$$

Hence the length of any path in X joining e_0 to $-e_0$ is $\geqq \pi$, which is the minimal length of paths between the two points in $\mathbf{E}$. However, let α_n be the half great circle joining the two points in the upper (e_0, e_n)-half plane. Then

$$L(T\alpha_n) < (1 + 1/2^n)\pi \to \pi = \mathrm{dist}_X(e_0, -e_0).$$

Hence there is no minimal path joining the two points in X. Note that each $T\alpha_n$ is a geodesic in X joining the two points, because

$$\mathrm{Im}(T\alpha_n) \cup -\mathrm{Im}(T\alpha_n)$$

is the fixed point set of the isometry F_n defined by

$$F_n\left(\sum x_k e_k\right) = x_0 e_0 + x_n e_n - \sum_{k \neq 0, n} x_k e_k.$$

McAlpin refers to [Gros 64] for results on the distribution of degenerate points of the exponential map in similar examples.

Next we give another criterion for (X, g) to be complete. We start with a lemma.

Lemma 6.8. *Let* $f\colon Y \to X$ *be a* C^1 *map between Riemannian manifolds* (Y, h) *and* (X, g). *Assume that there is a constant* $C > 0$ *such that for all* $y \in Y$ *and* $w \in T_yY$ *we have*

$$\|Tf(y)w\|_g \geqq C\|w\|_h.$$

If $\gamma\colon [a, b] \to \gamma$ *is a piecewise* C^1 *path in* Y, *then*

$$L(f \circ \gamma) \geqq CL(\gamma).$$

Proof. We have

$$\begin{aligned} L_g(f \circ \gamma) = \int_a^b \|(f \circ \gamma)'(t)\|_g \, dt &= \int_a^b \|Tf(\gamma(t))\gamma'(t)\|_g \, dt \\ &\geqq \int_a^b C\|\gamma'(t)\|_h \, dt \\ &= CL_h(\gamma), \end{aligned}$$

as was to be shown.

Let $f\colon Y \to X$ be a C^1 map of manifolds. We say that f has the **unique path lifting property** if given a point $x \in X$, a piecewise C^1 path α in X starting from x, and a point $y \in Y$ such that $f(y) = x$, then there exists a unique piecewise C^1 path γ in Y such that $f \circ \gamma = \alpha$ and γ starts at y.

Theorem 6.9. *Let $f\colon Y \to X$ be a local C^1 isomorphism of a Riemannian manifold (Y, h) into a Riemannian manifold (X, g). Assume that (Y, h) is complete, and X is connected. Also assume that there is a constant $C > 0$ such that for all $y \in Y$ and $w \in T_yY$ we have*

$$\|Tf(y)w\|_g \geqq C\|w\|_h.$$

Then f is surjective, f is a covering and has the unique path lifting property, and (X, g) is complete.

Proof. The proof is in three steps. *The first step is to prove that f is surjective and has the unique path lifting property.* Let $x \in X$, $x = f(y)$. Every point in X can be joined to x by a piecewise C^1 path. Let $\alpha\colon [a, b] \to X$ be such a path, joining $\alpha(a) = x$ with $\alpha(b)$. We shall prove that α can be lifted uniquely to a path in Y starting from y. This will prove the first step. Let S be the set of elements $t \in [a, b]$ such that the path α restricted to $[0, t]$ can be lifted uniquely to a path γ starting at y. Without loss of generality, we may assume that $a < b$. The set S is not empty because $a \in S$, and it is open because f is a local isomorphism. So it remains to show that S is closed. Let $\{t_n\}$ be a sequence in S increasing to the least upper bound b_0 of S. Then $\{\alpha(t_n)\}$ converges to $\alpha(b_0)$, and by Lemma 6.8 the lengths of the lifted path between $\gamma(t_n)$ and $\gamma(t_m)$ tend to 0 as m, n tend to infinity, so the sequence $\{\gamma(t_n)\}$ is Cauchy in Y, converging to some element y_0 since Y is assumed complete. Then $f(y_0) = \alpha(b_0)$, so S is closed, whence $S = X$ by assumption. Therefore f is surjective, and we have also proved the existence and uniqueness of path liftings.

The next step in the proof is to reduce the theorem to the case when the map f is a local isometry. We do this as follows. Let $g^* = f^*(g)$ be the pull-back of the metric g from X to Y by f. Then for all $y \in Y$ and $w \in T_yY$ we have

$$\|w\|_{g^*} = \|Tf(y)w\|_g \geqq C\|w\|_h.$$

Hence on Y we find that $\mathrm{dist}_{g^*} \geqq C\,\mathrm{dist}_h$. We now claim that Y is complete for the distance dist_{g^*}. To see this, first observe that if $\{y_n\}$ is g^*-Cauchy, then $\{y_n\}$ is also h-Cauchy, so $\{y_n\}$ is h-convergent to an element $y_0 \in Y$. Then $\{f(y_n)\}$ converges to $f(y_0)$. But f induces an isomorphism from some neighborhood V of y_0 to an open neighborhood of $f(y_0)$, and hence for all but a finite number of n, the points $f(y_n)$ lie in $f(V)$, so $\{y_n\}$ is also g^*-convergent to y_0 since $g^* = f^*(g)$. This proves that Y is g^*-complete. Furthermore, we have the inequality

$$\mathrm{dist}_{g^*}(y_1, y_2) \geqq \mathrm{dist}_g\big(f(y_1), f(y_2)\big) \qquad \text{for all} \quad y_1, y_2 \in Y.$$

In this final step, we prove that f is a covering. Since Y is g^*-complete, this will also prove that (X, g) is complete, and will conclude the proof

of the theorem. By the second step, we may assume without loss of generality that f is a local isometry, and that

$$(*) \qquad \mathrm{dist}_h(y_1, y_2) \geqq \mathrm{dist}_g\big(f(y_1), f(y_2)\big) \qquad \text{for all} \quad y_1, y_2 \in Y.$$

Let $x \in X$. From Theorem 6.4 we know that

$$\exp_x\colon \mathbf{B}_g(0_x, r) \to B_g(x, r)$$

is an isomorphism for all r sufficiently small, say $r < c$ with $c > 0$. Let $y \in f^{-1}(x)$. Since f is a local isometry, the following diagram is commutative (using $(*)$):

$$\begin{array}{ccc} \mathbf{B}_h(0_y, r) & \xrightarrow{df(y)} & \mathbf{B}_g(0_x, r) \\ {\scriptstyle \exp_y}\big\downarrow & & \big\downarrow{\scriptstyle \exp_x} \\ B_h(y, r) & \xrightarrow[f]{} & B_g(x, r) \end{array}$$

Note that the right vertical arrow is a differential isomorphism because we have picked r small enough, but so far we have made no such assertion for the left vertical arrow. For the proof of the theorem, it will suffice to show that $f^{-1}B_g(x, r)$ is the disjoint union of the balls $B_h(y, r)$ for $y \in f^{-1}(x)$, if r is taken small enough. We take r so small that given $x' \in B_g(x, r)$ there is a unique geodesic in $B_g(x, r)$ joining x to x' (namely $\exp_x(tv)$ for some v). Then, first, we have $f\big(B_h(y, r)\big) \subset B_g(x, r)$, so the union is contained in $f^{-1}B_g(x, r)$. Conversely, given a point $z \in f^{-1}B_g(x, r)$, we can join $f(z)$ to x by a geodesic of length $< r$ in $B_g(x, r)$, and by the path lifting property already proved in step 1, we can join z to a point y in $f^{-1}(x)$ by a geodesic of the same length, so

$$f^{-1}\big(B_g(x, r)\big) = \bigcup_y B_h(y, r),$$

where the union is taken over $y \in f^{-1}(x)$. Finally, let y_1, $y_2 \in f^{-1}(x)$ and suppose $y_1 \neq y_2$. We claim that $B_h(y_1, r)$ is disjoint from $B_h(y_2, r)$. Suppose there is some point z in the intersection. Then z can be joined to y_1 by a geodesic α_1 in $B_h(y_1, r)$, and z can also be joined to y_2 by a geodesic α_2 in $B_h(y_2, r)$, and these geodesics are distinct. Their images under f are geodesics in $B_g(x, r)$ joining x with $f(z)$. By the uniqueness of path lifting, this would mean we have two distinct geodesics in $B_g(x, r)$ joining x and z, and that these geodesics have length $< r$. This contradicts the local uniqueness statement, and proves that the balls $B_h(y_1, r)$ and $B_h(y_2, r)$ are disjoint. This concludes the proof of the theorem.

Remark. In the next chapter, *under a condition of seminegative curvature* (to be defined), we shall take $Y = T_x X$, and we shall prove that

$$f = \exp_x: \; T_x X \to X$$

satisfies the hypotheses of Theorem 6.9, and therefore in particular that geodesic completeness implies completeness. In this manner, we shall be able to replace the local compactness condition by a curvature condition to insure the equivalence between the two notions of completeness. The whole technique goes back to Hadamard [Ha 1898] in the case of surfaces with seminegative curvature, and Cartan [Ca 28] in the general case, still in this context of seminegative curvature. The notion of a "covering space" was not so clear during this early period. Except for a minor variation, the theorem is apparently due to Ambrose [Am 56], and occurs in the standard treatments of differential geometry as in [He 62] later replaced by [He 78], Chapter I, Lemma 13.4; [KoN 63], Chapter IV, Theorem 4.6 and Chapter VIII, §8, Theorem 8.1 and especially Lemma 1. The theorem is at the base of the Cartan–Hadamard theorem, to be proved later.

CHAPTER IX

Curvature

This chapter is a continuation of the preceding one, and is concerned with the iteration of covariant derivatives, from a formal point of view, and also from the point of view of their effect on the geometry of the manifold.

IX, §1. THE RIEMANN TENSOR

Let X be a manifold with a spray, and the covariant derivative D associated with the spray. If ξ, η, ζ are vector fields on X, we are concerned with the operator

$$D_\xi D_\eta - D_\eta D_\xi - D_{[\xi,\eta]}: \; \Gamma TX \to \Gamma TX,$$

which is a linear map of ΓTX into itself.

Proposition 1.1. *There exists a unique tensor field R, section of $L^3(TX, TX)$, i.e. arising from the functor $\mathbf{E} \mapsto L^3(\mathbf{E}, \mathbf{E})$ (continuous trilinear maps of $\mathbf{E}$ into itself) such that for all vector fields ξ, η, ζ we have*

$$R(\xi, \eta, \zeta) = D_\xi D_\eta \zeta - D_\eta D_\xi \zeta - D_{[\xi,\eta]}\zeta.$$

Proof. The expression on the right-hand side gives a well-defined vector field on X. To show that this association comes from a tensor field, we can compute in a chart. To do this, we use the local expression for the covariant derivative given in Theorem 2.1 of Chapter VIII. So for the rest of the argument, ξ, η, ζ stand for ξ_U, η_U, ζ_U in a chart U. Then, for example, we have

(1) $$D_\eta \zeta = \zeta' \cdot \eta - B(\eta, \zeta).$$

We determine $D_\xi(D_\eta\zeta)$ by substitution in this formula. As a first step, we have to write down the derivative

$$(D_\eta\zeta)'\cdot\xi = \zeta''\cdot\xi\cdot\eta + \zeta'\cdot\eta'\cdot\xi - B(\eta'\xi, \zeta) - B(\eta, \zeta'\cdot\xi) - (B'\cdot\xi)(\eta, \zeta).$$

Then it follows that

$$\begin{aligned} D_\xi(D_\eta\zeta) = \zeta''\cdot\xi\cdot\eta + \zeta'\cdot\eta'\cdot\xi - B(\eta'\cdot\xi, \zeta) - B(\eta, \zeta'\cdot\xi) - B(\zeta'\cdot\eta, \xi) \\ - (B'\cdot\xi)(\eta, \zeta) + B(B(\eta, \zeta), \xi). \end{aligned}$$

Permuting ξ and η gives us the second term. Using the local expression for the bracket

$$[\xi, \eta] = \eta'\cdot\xi - \xi'\cdot\eta$$

as well as (1) will give us the third term. The reader will then verify that all the expressions containing a derivative cancel, leaving only trilinear expressions involving ξ, η, and ζ. This proves Proposition 1.1.

In addition, after the cancellation of the terms with derivatives, we obtain a local expression for R, namely:

Proposition 1.2. *Letting ξ, η, ζ represent vector fields in a chart:*

$$\begin{aligned} R(\xi, \eta, \zeta) = B(B(\eta, \zeta), \xi) - B(B(\xi, \zeta), \eta) \\ + (B'\cdot\eta)(\xi, \zeta) - (B'\cdot\xi)(\eta, \zeta). \end{aligned}$$

Remark. There is no universal convention as to the sign of R. I use the same sign as [KoN 63], [ChE 75], [He 78], and [BGV 92], but the opposite sign to [BGM 71], [HGL 87/93], and [Mi 63]. For further comments, see the discussion after the definition of sectional curvature.

Let $v, w, z \in T_xX$. It is customary to write

$$\begin{aligned} R(v, w, z) = R(v, w)z = R_x(\xi(x), \eta(x), \zeta(x)) \\ = R(\xi, \eta, \zeta)(x), \end{aligned}$$

if ξ, η, ζ are any vector fields such that $\xi(x) = v$, $\eta(x) = w$, $\zeta(x) = z$. One writes

$$R(\xi, \eta)\colon \Gamma TX \to \Gamma TX$$

for the linear map of ΓTX into itself, given by

$$R(\xi, \eta) = D_\xi D_\eta - D_\eta D_\xi - D_{[\xi,\eta]}.$$

As a function of two variables, according to this definition, one may view R as a section of the bundle $L^2(TX, L(TX, TX))$, which is formed by applying the functor $\mathbf{E}^3 \mapsto L^2(\mathbf{E}, L(\mathbf{E}, \mathbf{E}))$ to the tangent bundle.

Next we list some identities.

Proposition 1.3.

$R(v, w) = -R(w, v)$ (*skew-symmetry*).

$R(v, w, z) + R(w, z, v) + R(z, v, w) = 0$ (*cyclicity*, **Bianchi's identity**).

Proof. The first relation is obvious from the definition. The second one is immediate from the local representation of Proposition 1.2.

For the next two properties, we assume that the spray is the one associated with a metric, so the covariant derivative is the metric derivative. We let $\langle\, ,\, \rangle_g$ be the scalar product associated with the metric. Then we define a function of four variables

$$R(v, w, z, u) = \langle R(v, w)z, u\rangle_g \qquad \text{for} \quad v, w, z, u \in T_xX.$$

Then R is a tensor of type L^4, that is a section of $L^4(TX) = L^4(TX, \mathbf{R})$. We shall call R the **Riemann 4-tensor** (canonical with respect to g). We call $-R$ the **curvature tensor**. The properties of Proposition 1.3 may be formulated for this 4-tensor, and we shall see in a moment that it also satisfies two other important properties. Thus it is useful to make a general definition. A tensor R of type L^4 is called a tensor of **Riemann type** if it satisfies the following four properties:

RIEM 1. $R(v, w, z, u) = -R(w, v, z, u)$

RIEM 2. $R(v, w, z, u) = -R(v, w, u, z)$

RIEM 3. $R(v, w, z, u) + R(w, z, v, u) + R(z, v, w, u) = 0$

RIEM 4. $R(v, w, z, u) = R(z, u, v, w)$.

The first two conditions express the property of being alternating in the first two variables, and also in the last two variables. The third condition is called the **Bianchi identity**, and expresses the property that the cyclic symmetrization of the tensor is 0. The fourth property states that the tensor is symmetric in the pairs of variables (v, w) and (z, u). In particular, we note right away that from **RIEM 4**, we obtain:

$R(v, w, v, u)$ is symmetric in (w, u), that is $R(v, w, v, u) = R(v, u, v, w)$.

We shall make more comments on these properties after the next proposition, which justifies the terminology.

Proposition 1.4. *On a pseudo Riemannian manifold, the Riemann tensor satisfies all the above four properties. Furthermore,* **RIEM 4** *follows from* **RIEM 1, 2, 3.**

Proof. Properties **RIEM 1** and **RIEM 3** have been proved in Proposition 1.3. Property **RIEM 2** amounts to proving that $R(v, w, z, z) = 0$ for all v, w, z; or in terms of vector fields, $R(\xi, \eta, \zeta, \zeta) = 0$. We will need to differentiate. Since all the terms with derivatives vanish in the local formula of Proposition 1.2, we may assume without loss of generality that $[\xi, \eta] = 0$. Then

$$\langle R(\xi, \eta)\zeta, \zeta\rangle_g = \langle D_\xi D_\eta \zeta - D_\eta D_\xi \zeta, \zeta\rangle_g,$$

and we must show that the right side is symmetric in ξ, η. But $[\xi, \eta] = 0$ implies that

$$\mathscr{L}_\xi \mathscr{L}_\eta \langle \zeta, \zeta\rangle_g$$

is symmetric in ξ, η. Since we are dealing with the metric covariant derivative, it follows that

$$\mathscr{L}_\eta \langle \zeta, \zeta\rangle_g = 2\langle D_\eta \zeta, \zeta\rangle_g$$

and therefore

$$\mathscr{L}_\xi \mathscr{L}_\eta \langle \zeta, \zeta\rangle_g = 2\langle D_\xi D_\eta \zeta, \zeta\rangle_g + 2\langle D_\xi \zeta, D_\eta \zeta\rangle_g,$$

from which it follows at once that $\langle D_\xi D_\eta \zeta, \zeta\rangle_g$ is symmetric in ξ, η, thus proving **RIEM 2.**

The formula **RIEM 4** is a formal consequence of the preceding three formulas. It is basically an exercise in algebra, which we carry out. In the cyclic identity **RIEM 3**, interchange u with z, v, w successively, and add the resulting three relations. One gets, using **RIEM 1** and **RIEM 3**:

$$(*) \qquad R(u, v, w, z) + R(u, w, z, v) + R(u, z, v, w) = 0.$$

From cyclicity and **RIEM 1**, one gets

$$R(z, v, u, w) = R(u, v, z, w) - R(u, z, v, w) \quad \text{or}$$
$$R(u, z, v, w) = R(u, v, z, w) - R(z, v, u, w).$$

We substitute the value on the left in $(*)$, and use **RIEM 1** to conclude the proof of **RIEM 4.**

We shall be dealing with a contraction of the canonical 4-tensor. We defined the **canonical 2-tensor** R_2 by

$$R_2(v, w) = R(v, w, v, w).$$

Proposition 1.5. *The canonical 2-tensor determines the Riemann tensor. Or similarly, if the canonical tensor R satisfies*

$$R(v, w, v, w) = 0 \qquad \text{for all} \quad v, w,$$

then $R = 0$.

Proof. Say we prove the second assertion first. From **RIEM 4**, which implies that $R(v, w, v, z)$ is symmetric in (w, z), if $R(v, w, v, w) = 0$ for all v, w then $R(v, w, v, z) = 0$ for all v, w, z. From the alternating properties of **RIEM 1** and **RIEM 2**, it follows that $R = 0$ identically.

To show that the canonical 2-tensor determines the Riemann tensor, we note that the problem is essentially equivalent to the other statement, but one may argue directly as when one recovers a symmetric bilinear form from a quadratic form, namely

$$\left[\frac{\partial^2}{\partial t \partial s} R(v + tz, w + su, v + tz, w + su) - \frac{\partial^2}{\partial t \partial s} R(v + tu, w + sz, v + tu, w + sz)\right]_{s=t=0} = 6R(v, w, z, u).$$

This proves the proposition.

An important case arises when $R_2 \geqq 0$. We define (X, g) to have **seminegative curvature** if $R_2 \geqq 0$. The following discussion explains this terminology in terms of its historical development.

Curvature discussion

A large part of the theory we are developing is fundamentally a theory of commutative rings with certain types of derivation, and possibly scalar products, in which positivity or negativity plays no role. This theory contains a number of formulas with precise equality between various terms. There would be some value in redoing this chapter and the preceding one completely in such a context of commutative differential algebra. At some point, for certain applications, the positivity or negativity properties of the real numbers are used, as in the second statement of Proposition 2.6 below. For such applications, the question arises as to what is the natural sign to be used, if indeed there is a natural sign.

Historically, the theory arose in a geometric context, based on geometric intuition. To each pair of vectors (v, w) in a tangent space T_xX, we define the **area square** of the parallelogram spanned by these vectors to be

$$\mathrm{Ar}_g(v, w)^2 = v^2w^2 - \langle v, w\rangle_g^2.$$

As usual, $v^2 = \langle v, v\rangle_g$. Then, when $\mathrm{Ar}_g(v, w)^2 \neq 0$, we define the **sectional curvature** to be

$$\mathrm{Sec}_g(v, w) = -\frac{R(v, w, v, w)}{\mathrm{Ar}_g(v, w)^2}.$$

In the Riemannian case, $\mathrm{Ar}_g(v, w) \neq 0$ if and only if v, w are linearly independent. If v^2 and $w^2 > 0$, then the value on the right depends only on the unit vectors in the direction of v, w respectively; and if v, w are orthogonal unit vectors, then

$$\mathrm{Sec}_g(v, w) = -R(v, w, v, w).$$

In the Riemannian case, it is immediate that the value of the sectional curvature on (v, w) depends only on the plane generated by v and w, because of the skew-symmetry of **RIEM 1** and **RIEM 2**. For the complex analogue, see [La 87], Chapter V, §3.

Let $c \in \mathbf{R}^+$ be a positive number. The multiple cg is called a **scaling** of the metric g. Since the covariant derivative D^{cg} is the same as D^g, it follows from the definitions that under scaling, the curvature changes as

$$\mathrm{Sec}_{cg} = c^{-1}\,\mathrm{Sec}_g.$$

Directly from the definition, we then see in the Riemannian case that:

The sectional curvature has constant value -1 *if and only if*

$$R_2(v, w) = v^2w^2 - \langle v, w\rangle_g^2 \qquad \textit{for all} \quad v, w \in T_xX.$$

Viewing v as fixed, the above expression is quadratic in w, and the corresponding symmetric bilinear form is

$$R(v, w, v, z) = v^2\langle w, z\rangle_g - \langle v, w\rangle_g\langle v, z\rangle_g.$$

Thus $R(v, w)v$ is given by

$$R(v, w)v = v^2w - \langle v, w\rangle_g v \qquad \text{for all} \quad v, w \in T_xX.$$

Similarly, the sectional curvature has constant value $+1$ *if and only if the analogous formula holds with a minus sign inserted on one side, so that for instance*

$$-R(v, w)v = v^2w - \langle v, w\rangle_g v.$$

In the applications of this book (the rest of this chapter, the Cartan–Hadamard theorem, the variation formulas of Chapter XI, §1, etc.) what matters is not the "curvature" as defined above, but the canonical tensor R itself. Furthermore, formulas in these applications come out much neater with R than with "curvature" for two reasons:

First, for such formulas, dividing to normalize as in the curvature quotient is unnatural, partly because the term by which one divides, algebraically, may be equal to 0 unless extra conditions are imposed.

Second, even for inequalities as distinguished from equalities, the natural condition which arises is $R(v, w, v, w) \geqq 0$ rather than curvature $\leqq 0$. If one takes R with the sign as we have defined it, then only plus signs occur in all the formulas (cf. Lemma 2.5 and the variation formula, Theorem 1.3 of Chapter XI, for instance). This universal occurrence of plus signs is obscured if one introduces minus signs artificially. I regard this universal occurrence of plus signs as structurally important.

The naturality of R in the real case is similar to the naturality of its counterpart in the complex case, where formulas involving positivity come out neatly by using the analogue of R rather than its negative (as already noted by Griffiths). Cf. [La 87], the comments pp. 136–137 about holomorphic sectional curvature. The lesson is that the "curvature" in classical terminology is minus the natural object R (aside from questions of normalizing the dilation to the unit sphere).

Classically, starting with surface theory, people wanted some formulas such as Gauss-Bonnet or formulas relating "curvature" and Betti numbers, using $\pm R$, to come out so that on the sphere, one gets a value of certain integral to be 4π and not -4π. So they picked the minus sign, and gave the notion $-R$ (normalized) the name of curvature, which makes the sphere have positive curvature. The bottom line is that depending on what applications one makes, both R and $-R$ are "natural." However, from the point of view of universal algebraic manipulations, R is the clearest functorial notion.

One can define two other curvatures, at least. Actually, all we need is a tensor of curvature type. From such a tensor R, we obtain two other tensors. First observe that to each pair of vectors $v, z \in \mathbf{E}$ we can associate an endomorphism of $\mathbf{E}$, denoted by $\mathrm{Ric}(v, z)$, and defined by

$$\mathrm{Ric}_R(v, z)w = R(v, w)z.$$

Thus Ric gives a bilinear map

$$\mathrm{Ric}_R\colon \mathbf{E} \times \mathbf{E} \to L(\mathbf{E}, \mathbf{E}).$$

Applied to the tangent bundle, and the Riemann tensor R itself, Ric is called the **Ricci tensor.**

Furthermore, in the finite dimensional case, the trace

$$\mathrm{tr}: L(\mathbf{E}, \mathbf{E}) \to \mathbf{R}$$

is a continuous linear map. Then the composite

$$\mathrm{Sc}_R = \mathrm{tr} \circ \mathrm{Ric}_R$$

is a function of pairs of vectors, which when applied to the tangent bundle defines what is called the **scalar curvature**. In the infinite dimensional case, one has to give an additional structure, assuming that the Ricci tensor is of "trace class", or defining the sectional curvature with respect to a given "trace," i.e. a continuous functional on $L(\mathbf{E}, \mathbf{E})$ which is equal on products AB and BA. But this now leads far afield.

Suppose we are in the Riemannian case. We can then give an explicit formula for the scalar curvature. In the neighborhood of a point, we can find vector fields $\xi_1, \ldots, \xi_n$ (with $n = \dim X$) which are orthonormal, by the usual orthogonalization process. Such a sequence of vector fields is called an **orthonormal frame** at the point.

Proposition 1.6. *Let $\{\xi_1, \ldots, \xi_n\}$ be an orthonormal frame on an open set. Then for vector fields ξ, η we have*

$$\mathrm{Sc}_R(\xi, \eta) = \sum_{i=1}^{n} R(\xi, \xi_i, \eta, \xi_i).$$

Proof. This is immediate from the definition of the trace of an endomorphism of a finite dimensional vector space.

The conditions **RIEM 1** and **RIEM 2** express the property of depending only on the wedge product of each pair of variables $v \wedge w$ and $z \wedge u$. Property **RIEM 4** is a symmetric property in these pairs of variables. Thus we may say that the four-variable tensor R defines a symmetric bilinear form on $\bigwedge^2 TX$, which we denote by

$$R^\wedge : \bigwedge^2 TX \wedge \bigwedge^2 TX \to \mathbf{R}, \quad \text{such that } R^\wedge(v \wedge w, z \wedge u) = R(v, w, z, u).$$

On the other hand, we also have the pseudo Riemannian metric, which induces a non-singular scalar product on $\bigwedge^2 TX$ by the formula

$$\langle v \wedge w, z \wedge u \rangle_g = \det \begin{vmatrix} \langle v, z \rangle_g & \langle v, u \rangle_g \\ \langle w, z \rangle_g & \langle w, u \rangle_g \end{vmatrix}.$$

These scalar products are of course evaluated at each point $x \in X$, v, w, z, $u \in T_x X$. The scalar product on $\bigwedge^2 TX$ with respect to the non-singular symmetric form g then corresponds to a symmetric operator which is called the **curvature operator**.

In the infinite dimensional case, from the self duality, each tangent space can be interpreted as the dual of its dual, and the wedge product is defined as in Chapter V, §3 so the above notions still make sense.

Readers wanting to pursue the topic of curvature are now referred to other books on differential geometry, including [BGM 71], [ChE 75], [doC 92], and [GHL 87/93].

IX, §2. JACOBI LIFTS

Let (X, g) be pseudo Riemannian. We write w^2 for $\langle w, w \rangle_g$, and $v \perp w$ for $\langle v, w \rangle_g = 0$. We let α: $[a, b] \to X$ be a geodesic. Unless otherwise specified, (X, g) is not necessarily Riemannian.

A lift $\eta \in \mathrm{Lift}(\alpha)$ to the tangent bundle will be called a **Jacobi lift**, or more classically a **Jacobi field**, if it satisfies the **Jacobi differential equation**

$$D^2_{\alpha'} \eta = R(\alpha', \eta)\alpha'.$$

Theorem 3.1 of Chapter VIII and Proposition 1.2 in the preceding section of the present chapter show that locally, the above equation is a linear differential equation. Therefore, by the existence and uniqueness theorem for linear differential equations, we get:

Theorem 2.1. *Let (X, g) be pseudo Riemannian, let α: $[a, b] \to X$ be a geodesic. Given vectors z, $w \in T_{\alpha(a)}X$, there exists a unique Jacobi lift $\eta = \eta_{z,w}$ of α to TX such that*

$$\eta(a) = z \quad \text{and} \quad D_{\alpha'}\eta(a) = w.$$

In particular, the set of Jacobi lifts of α is a vector space linearly isomorphic to $T_{\alpha(a)} \times T_{\alpha(a)}$ under the map $(z, w) \mapsto \eta_{z,w}$.

We denote the space of Jacobi lifts of α by $\mathrm{Jac}(\alpha)$.

Let $v \in T_x$ and consider the unique geodesic

$$\alpha(t) = \exp_x(tv)$$

such that $\alpha(0) = x$ and $\alpha'(0) = v$, with α defined on an open interval. Let $w \in T_x X$ and let η_w be the unique Jacobi lift of α such that

$$\eta_w(0) = 0 \quad \text{and} \quad D_{\alpha'}\eta_w(0) = w.$$

Example. Let $w = v$. Then

$$\eta_v(t) = t\alpha'(t).$$

Proof. One verifies at once that $\eta_v(0) = 0$, and since $D_{\alpha'}\alpha' = 0$, we also have

$$D_{\alpha'}\eta_v(t) = \alpha'(t) \quad \text{and} \quad D^2_{\alpha'}\eta_v = 0 = R(\alpha', \alpha')\alpha'.$$

Remark 1. Defining a Jacobi lift implicitly has the geodesic α in its definition. If, say, $a = 0$, this geodesic is uniquely determined by its initial condition $\alpha'(0) = v$, so the Jacobi lift is also determined by v. Thus one could write $\eta_w^{(v)}$ for the Jacobi lift. In the present discussions, this won't be necessary since we deal systematically with a fixed α.

Remark 2. In a chart, the derivative $\eta'(0)$ can be computed in the naive way since $\eta: J \to TX$ is defined on an interval. The naive derivative and the covariant derivative $D_{\alpha'}\eta$ differ locally in a chart by a term linear in η, which therefore vanishes at 0 if $\eta(0) = 0$. Hence the naive derivative and the covariant derivative have the same value at 0, that is

$$\eta'(0) = D_{\alpha'}\eta(0).$$

We note that, α being fixed, the association $w \mapsto \eta_w$ is linear. We now have the possibility of orthogonalization.

Proposition 2.2. *Let (X, g) be pseudo Riemannian. Let $\alpha: [a, b] \to X$ be a geodesic, and let η be a Jacobi lift of α. Then there are numbers c, d such that*

$$\langle \eta, \alpha' \rangle_g(t) = c(t-a) + d.$$

In fact, $d = \langle \eta, \alpha' \rangle_g(a)$ and $c = \langle D_{\alpha'}\eta, \alpha' \rangle_g(a)$. If $\eta(a)$ and $D_{\alpha'}\eta(a)$ are orthogonal to $\alpha'(a)$, then $\eta(t)$ is orthogonal to $\alpha'(t)$ for all t.

Proof. Using the metric derivative, and $D_{\alpha'}\alpha' = 0$ since α is a geodesic, we find that $\partial\langle \eta, \alpha' \rangle_g = \langle D_{\alpha'}\eta, \alpha' \rangle_g$, and then

$$\partial^2\langle \eta, \alpha' \rangle_g = \langle D^2_{\alpha'}\eta, \alpha' \rangle_g = R(\alpha', \eta, \alpha', \alpha') = 0.$$

Hence $\langle \eta, \alpha' \rangle_g$ is a linear function, whose coefficients are immediately determined to be those written down in the proposition.

Proposition 2.3. *As above, let $\alpha'(0) = v$. Write $w = cv + w_1$ with $\langle w_1, v \rangle_g = 0$. Then η_w has the decomposition*

$$\eta_w = c\eta_v + \eta_{w_1}, \quad \textit{also written} \quad \eta_w(t) = ct\alpha'(t) + \eta_{w_1}(t).$$

Furthermore η_{w_1} is orthogonal to α', that is $\langle \eta_{w_1}, \alpha' \rangle_g = 0$.

Proof. Immediate from Proposition 2.2.

Next we get the similar orthogonalization of $D_{\alpha'}\eta_w$.

Proposition 2.4. *Notation as in Proposition 2.3, we have an orthogonal decomposition*

$$D_{\alpha'}\eta_w = cD_{\alpha'}\eta_v + D_{\alpha'}\eta_{w_1} \quad \textit{also written} \quad D_{\alpha'}\eta_w(t) = c\alpha'(t) + D_{\alpha'}\eta_{w_1}(t).$$

In other words, if $w_1 \perp \alpha'(0)$, then $D_{\alpha'}\eta_{w_1} \perp \alpha'$. Furthermore $(D_{\alpha'}\eta_v)^2$ is constant.

Proof. For the first assertion, we take the derivative and use Proposition 2.3 to get

$$0 = \partial\langle \eta_{w_1}, \alpha' \rangle_g = \langle D_{\alpha'}\eta_{w_1}, \alpha' \rangle_g.$$

For the second, we then obtain for $\eta = \eta_w$:

$$\begin{aligned} \partial\langle D_{\alpha'}\eta, D_{\alpha'}\eta \rangle_g &= 2\langle D^2_{\alpha'}\eta, D_{\alpha'}\eta \rangle_g \\ &= 2\langle R(\alpha', \eta)\alpha', D_{\alpha'}\eta \rangle_g. \end{aligned}$$

If $\eta = \eta_v$ so $\eta_v(t) = t\alpha'(t)$, then the right side is 0 because R_4 is alternating in its last two variables. This concludes the proof.

Remark. If $\eta = \eta_w$ with $w \perp \alpha'(0)$, then we cannot necessarily conclude that the derivative is 0.

The next lemma will give us information on the rate of growth of a Jacobi lift, and the convexity of its square.

Lemma 2.5. *Assume (X, g) Riemannian. Let η be a Jacobi lift of α. Let $f(t) = \|\eta(t)\|$. Then at those values of $t > 0$ such that $\eta(t) \neq 0$, we have*

$$f'' = \frac{1}{\|\eta\|^3}\left((D_{\alpha'}\eta)^2\eta^2 - \langle D_{\alpha'}\eta, \eta \rangle_g^2\right) + \frac{1}{\|\eta\|}R_2(\alpha', \eta).$$

Proof. Straightforward calculus, using the covariant derivative. The first derivative f' is given by

$$f' = (\eta^2)^{-1/2}\langle \eta, D_{\alpha'}\eta \rangle_g = \frac{1}{\|\eta\|}\langle \eta, D_{\alpha'}\eta \rangle_g.$$

Then f'' is computed by using the rule for the derivative of a product. In the term containing $\langle D^2_{\alpha'}\eta, \eta\rangle_g$, we replace $D^2_{\alpha'}\eta$ by $R(\alpha', \eta)\alpha'$ (using the definition of a Jacobi lift) to conclude the proof of the lemma.

In the above lemma, we note that on the right side, the first term is $\geqq 0$, and the second term is $\geqq 0$ if $R_2 \geqq 0$.

Proposition 2.6. *Let* α: $[0, b] \to X$ *be a geodesic. Let* $w \in T_{\alpha(0)}X$, $w \neq 0$. *Let* $\eta_w = \eta_{0,w} = \eta$ *be the unique Jacobi lift satisfying*

$$\eta_w(0) = 0 \quad \textit{and} \quad D_{\alpha'}\eta_w(0) = w.$$

If (X, g) *is Riemannian and* $R_2 \geqq 0$ (*so* (X, g) *has seminegative curvature*), *then for* $t \in [0, b]$ *we have*

$$\|\eta(t)\| \geqq \|w\|t \quad \textit{and in particular} \quad \|\eta(1)\| \geqq \|w\| \textit{ if } b = 1.$$

Proof. Let $h(t) = \|\eta(t)\| - \|w\|t$ for $0 \leqq t \leqq b$. Then h is continuous, $h(0) = 0$, and by Lemma 2.5, $h'' = f'' \geqq 0$ whenever $\eta(t) \neq 0$. One cannot have $\eta(t) = 0$ for arbitrarily small values of $t \neq 0$, otherwise $D_{\alpha'}\eta(0)$ would be 0 (because in a chart U, $\eta_U'(0) = D_{\alpha'}\eta(0)$). In fact, we shall prove that there is no value of $t \neq 0$ such that $\eta(t) = 0$. Suppose there is such a value, and let t_0 be the smallest value > 0. In the interval $(0, t_0)$ we have $h'' \geqq 0$ by Lemma 2.5, so h' in increasing. But the beginning of the Taylor expansion of η in a chart is

$$\eta_U(t) = wt + O(t^2), \qquad \text{so} \qquad \lim_{t\to 0} f'(t) = \|w\|.$$

Furthermore, $h'(0)$ exists and is equal to 0, so $h' \geqq 0$ on $[0, t_0)$, so h is increasing, and there cannot be a value $t_0 > 0$ with $\eta(t_0) = 0$. Then the above argument applies on the whole interval $[0, b]$ to prove the desired inequality on the whole interval. This concludes the proof of Proposition 2.6.

Remark. These results essentially stem from Cartan [Ca 28]. The above version without coordinates, which extends to the infinite dimensional case, comes from [BiC 64]. Readers may find it instructive to compare this version with the one involving coordinates given in [He 78], pp. 71–73.

Proposition 2.6 is used for the subsequent application to the Cartan–Hadamard theorem (Theorem 3.7), based on Theorem 6.9 of Chapter VIII, whose origin is in Hadamard for surfaces [Ha 1898] and Cartan in general. (Here and at several other places, I rely on Helgason's very useful bibliographical comments.)

Variations of geodesics

By a **variation** of a curve α one means a C^2 map

$$\sigma\colon [a, b] \times J \to X$$

where J is some interval containing 0, such that $\sigma(s, 0) = \alpha(s)$ for all s. One then writes

$$\sigma(s, t) = \alpha_t(s),$$

and one views $\{\alpha_t\}$ as a family of curves defined on $[a, b]$. If all curves α_t are geodesics for $t \in J$ then one says that σ is a **variation through geodesics.**

Lemma 2.7. *Let* σ: $J_1 \times J_2 \to X$ *be a* C^2 *map. Then on lifts of* σ *to the tangent bundle, we have the equality of operators*

$$D_1D_2 - D_2D_1 = R(\partial_1\sigma, \partial_2\sigma).$$

Proof. The formula can be verified in a chart. It follows directly from the definitions, especially using the local expression of Proposition 1.2.

Proposition 2.8. *Let* σ: $[a, b] \times J \to X$ *be a variation of a geodesic* α *through geodesics. Let*

$$\eta(s) = \partial_2\sigma(s, 0).$$

Then η *is a Jacobi lift of* α, *said to* **come from** σ *or* **associated with** σ.

Proof. Given σ, we have

$$\begin{aligned} D_1^2\partial_2\sigma = D_1D_1\partial_2\sigma &= D_1D_2\partial_1\sigma \quad \text{by Lemma 5.3 of Chapter VIII} \\ &= D_2D_1\partial_1\sigma + R(\partial_1\sigma, \partial_2\sigma)\partial_1\sigma \quad \text{by Lemma 2.7.} \end{aligned}$$

But $D_1\partial_1\sigma(s, t) = 0$ because α_t is a geodesic, whence $D^2_{\alpha'}\eta = R(\alpha', \eta)\alpha'$, so η is a Jacobi lift of α, as was to be shown.

Theorem 2.9 (Variation at the Beginning Point). *Let* α *be a geodesic in* X *with initial value* $\alpha(0) = x$. *Let* $z, w \in T_xX$. *Let* β *be a curve such that*

$$\beta(0) = \alpha(0) \quad \textit{and} \quad \beta'(0) = z.$$

Let

$$\zeta(t) = P^t_{0,\beta}(\alpha'(0) + tw) = P^t_{0,\beta}(\alpha'(0)) + tP^t_{0,\beta}(w),$$
$$\sigma(s, t) = \exp_{\beta(t)} s\zeta(t).$$

Let $\alpha_t(s) = \sigma(s, t)$. Then $\alpha_0 = \alpha$, σ is a variation of α by geodesics $\{\alpha_t\}$, and α_t is the unique geodesic such that

$$\alpha_t(0) = \beta(t) \quad \text{and} \quad \alpha_t'(0) = \zeta(t).$$

In particular, if $w = 0$, then $\alpha_t'(0) = P^t_{0,\beta}(\alpha'(0))$. Furthermore, let

$$\eta(s) = \partial_2\sigma(s, 0).$$

Then $\eta = \eta_{z,w}$ is the unique Jacobi lift of α with initial conditions

$$\eta(0) = z \quad \text{and} \quad D_{\alpha'}\eta(0) = w.$$

Proof. The stated values for $\alpha_t(0)$ and $\alpha_t'(0)$ are immediate. Then from the definition of parallel translation,

$$(*) \qquad \zeta(0) = \alpha'(0) \quad \text{and} \quad D_{\beta'}\zeta(0) = w,$$

because if $\gamma_v(t) = P^t_{0,\beta}(v)$, then $D_{\beta'}\gamma_v = 0$ and we can use the standard rule for the derivative of the product $tP^t_{0,\beta}(w)$.

Then $\sigma(0, t) = \beta(t)$, so we obtain the initial conditions:

$$\eta(0) = \partial_2\sigma(0, 0) = \beta'(0) = z;$$

$$\begin{aligned} D_{\alpha'}\eta(0) &= D_1\partial_2\sigma(0, 0) = D_2\partial_1\sigma(0, 0) \quad \text{by Chapter VIII, Lemma 5.3} \\ &= (D_{\beta'}T\exp_\beta(0)\zeta)(0) \\ &= D_{\beta'}\zeta(0) = w \quad \text{by } (*), \end{aligned}$$

thus concluding the proof.

Example. Constant curvature. Let (X, g) be Riemannian. As an example, we shall now determine more explicitly the Jacobi lifts when (X, g) has constant curvature. Since the covariant derivative is invariant under a scaling of the metric, we may as well assume that the curvature is 0 or ± 1. In the next three proposition, we let $x \in X$ and we let $v \in T_xX$ be a *unit vector*. As usual, we let $\alpha = \alpha_v$ be the geodesic

$$\alpha(t) = \exp_x(tv).$$

For $w \in T_xX$ we let $\eta_w = \eta_w^{(v)}$ be the *Jacobi lift* of α_v satisfying the usual initial conditions

$$\eta_w(0) = 0 \quad \text{and} \quad D_{\alpha'}\eta_w(0) = w.$$

Finally, we let $\gamma_w = \gamma_w^{(v)}$ be *parallel translation* of w along α_v, so

$$\gamma_w(0) = w \quad \text{and} \quad D_{\alpha'}\gamma_w = 0.$$

Proposition 2.10. *Assume that the curvature is 0, or equivalently that the Riemann tensor R is identically 0. Then for all $w \in T_xX$ we have*

$$\eta_w(t) = t\gamma_w(t).$$

Proof. The two curves $t \mapsto \eta_w(t)$ and $t \mapsto t\gamma_w(t)$ have the same initial conditions. Also they satisfy the same differential equation, namely

$$D^2_{\alpha'}\eta_w = 0 \quad \text{and} \quad D^2_{\alpha'}(t\gamma_w(t)) = 0.$$

Hence they are equal, thereby proving the proposition.

The next two propositions deal with constant curvature ± 1. We recall that we wrote down the Riemann tensor explicitly in those cases in §1. We may therefore write down the differential equation for a Jacobi lift more explicitly in those cases, as follows.

Proposition 2.11. *Assume that (X, g) has constant curvature -1. Then the Jacobi differential equation has the form*

$$(1) \qquad D^2_{\alpha'}\eta_w = \eta_w - \langle \eta_w, \alpha' \rangle_g \alpha'.$$

Furthermore, if we orthogonalize w with respect to v, so write

$$w = c_0v + c_1u \quad \text{with } c_0, c_1 \in \mathbf{R} \text{ and a unit vector } u \perp v,$$

then

$$(2) \qquad \eta_w(t) = c_0t\alpha'(t) + (\sinh t)c_1\gamma_u(t).$$

Proof. The orthogonalization of Jacobi lifts comes from Proposition 2.3, so we want to identify the orthogonal components of the Jacobi lift of α_v with scalar multiples of parallel translation. It suffices to do so when $w = v$ and $w = u \perp v$ separately. The example following Theorem 2.1 already gives us the v-component, so we may assume $w = u$. In this case, the reader will verify that the two curves

$$t \mapsto \eta_u(t) \quad \text{and} \quad t \mapsto (\sinh t)\gamma_u(t)$$

have the same initial conditions at 0 (for their value, and the value of their first covariant derivative). They also satisfy the same differential equation, namely

$$D_{\alpha'}^2 \eta_u = \eta_u$$

and similarly for the other curve, since $D_{\alpha'}\gamma_u = 0$. Hence the two curves are equal, as was to be shown.

Thirdly we deal with constant positive curvature.

Proposition 2.12. *Assume X has constant curvature $+1$. Let $x \in X$. Then the same formulas hold as in Proposition* 2.11, *except for a minus sign on one side in formula* (1), *and with* sinh t *replaced by* sin t *in formula* (2).

Proof. The arguments are the same. Using sin t instead of sinh t just guarantees that the differential equation

$$D_{\alpha'}^2 \eta_u = -\eta_u$$

is satisfied, with the minus sign.

This concludes our analysis of the Jacobi lifts in the cases of constant curvature.

The Jacobi differential equation has at least two main aspects. One of them will be applied to a study of the differential of the exponential map in the next section. The other will be applied to variational questions in Chapter XI, §1.

IX, §3. APPLICATION OF JACOBI LIFTS TO $T\exp_x$

We continue to assume that (X, g) is pseudo Riemannian, unless otherwise specified.

We are interested in Jacobi lifts because they give precise information concerning the differential of the exponential map, for instance as in the following result. In the statement, if $v \in T_x$ then we identify $T_v T_x$ with T_x, as we usually do for a Banach space.

Theorem 3.1. *Let $x \in X$ and $v \in T_x$. Let α (defined on an open interval containing 0) be the geodesic such that $\alpha(0) = x$ and $\alpha'(0) = v$. Let $w \in T_x$ and let $\eta_w = \eta_{0,w}$ be the Jacobi lift of α such that*

$$\eta_w(0) = 0 \quad \text{and} \quad D_{\alpha'}\eta_w(0) = w.$$

Then for $r > 0$, in the interval of definition of α, we have the formula

$$T\exp_x(rv)w = \frac{1}{r}\eta_w(r).$$

In particular, w lies in the kernel of $T\exp_x(rv)$ if and only if $\eta_w(r) = 0$. Furthermore, if we let

$$\sigma(s, t) = \exp_x(s(v + tw)),$$

then $\eta_w(s) = \partial_2\sigma(s, 0)$.

Proof. The curve σ_t is a geodesic for each t, and

$$\sigma_0(s) = \exp_x(sv) = \alpha(s),$$

so σ is a variation of α through geodesics. Let $\eta(s) = \partial_2\sigma(s, 0)$. Then η is a Jacobi lift of α by Proposition 2.8. Let $f(s, t) = s(v + tw)$. Then

$$\partial_2\sigma(s, t) = (T\exp_x)(f(s, t))(\partial f/\partial t) = (T\exp_x)(f(s, t))(sw).$$

Hence $\eta(0) = 0$. Furthermore this same expression yields the formula of the theorem,

$$\eta(r) = (T\exp_x)(f(r, 0))rw = (T\exp_x)(rv)rw.$$

Taking the limit as $r \to 0$ in the formula, noting that in a chart $D_{\alpha'}\eta(0) = \eta'(0)$, and using $T\exp_x(0) = \text{id}$ proves that $D_{\alpha'}\eta(0) = w$ and concludes the proof of Proposition 3.1.

The Jacobi lifts also allow us to give a more global version of the Gauss lemma of Chapter VIII, Theorem 5.6.

Proposition 3.2 (Gauss Lemma, Global). *Let (X, g) be pseudo Riemannian. Let $x \in X$ and $v \in T_x X$. Let the exponential map $r \mapsto \exp_x(rv)$ be defined on an open interval J. Then for all $w \in T_x X$ we have*

$$\langle T\exp_x(rv)v, T\exp_x(rv)w\rangle_g = \langle v, w\rangle_g.$$

Proof. Immediate from Proposition 3.1 and the orthogonalization of Proposition 2.3.

Variation of a geodesic at its end point

Next we shall give another way of constructing Jacobi lifts, which will not be used until Chapter XV, Proposition 2.5. Readers interested in seeing at once the application of Jacobi lifts to the Cartan–Hadamard theorem, say, may omit the following construction.

Let $x, y \in X$ with $x \neq y$ be points such that y lies in the exponential image of a ball centered at 0_x in $T_x X$ and such that the exponential map $\exp_x$ is an isomorphism on this ball.

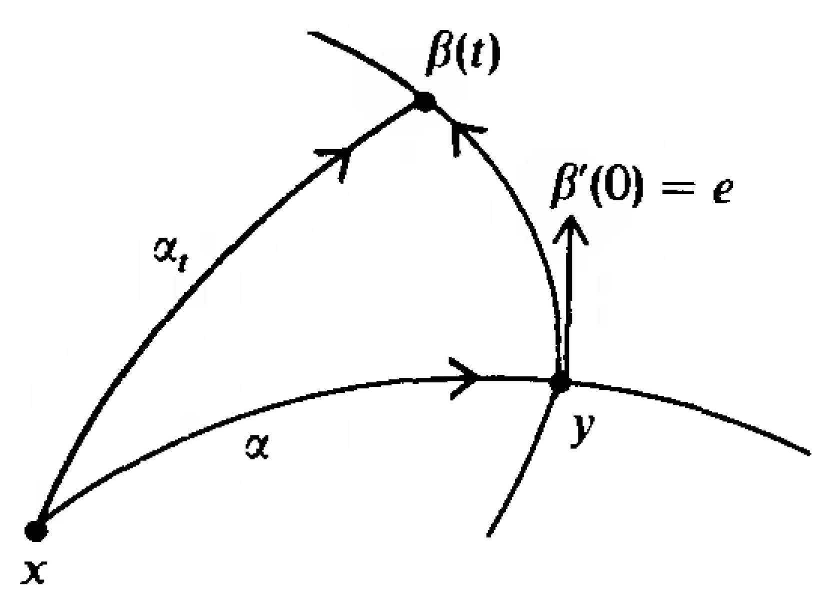
Thus this ball provides a normal chart at x. Let α be the geodesic parametrized by arc length joining x to y, so there is a unit vector $u \in T_x$ such that

$$\alpha(s) = \exp_x(su) \quad \text{and} \quad y = \exp_x(ru) \quad \text{for some } r > 0.$$

Thus $\text{dist}_g(x, y) = r$. Let e be a unit vector in T_yX, and let β be the geodesic such that

$$\beta(0) = y \quad (\text{so } \beta \text{ starts at } y) \quad \text{and} \quad \beta'(0) = e.$$

We consider an interval of the variable t such that $\beta(t)$ is contained in the image of the previous ball around x. For each t we let α_t be the unique geodesic from x to $\beta(t)$, parametrized by arc length. Then $\{\alpha_t\}$ is a variation of α, namely $\alpha_0 = \alpha$, and it is a variation through geodesics, illustrated on the next figure, drawn when e is perpendicular to $\alpha'(r)$ to illustrate Proposition 3.3.



The above variation will be called the **variation of α at its end point, in the direction of e.**

Proposition 3.3. *Let $y = \exp_x(ru)$ be in a normal chart at x as above, with the unit vector u. Let $\alpha(s) = \exp_x(su)$, and let $\{\alpha_t\}$ be the variation of α at its end point y in the direction of the unit vector $e \in T_yX$. Also denote this variation by σ, and let $\eta(s) = \partial_2\sigma(s, 0)$. Assume that e is orthogonal to $\alpha'(r)$. Then $D_{\alpha'}\eta$ is orthogonal to α', and η is the unique Jacobi lift of α such that*

$$\eta(0) = 0 \quad \text{and} \quad \eta(r) = e.$$

Proof. First note the uniqueness. If there is another Jacobi lift having the last stated property, then the difference vanishes at 0 and r, and by Theorem 3.1 this difference must be 0 since the exponential map is assumed to be an isomorphism from a ball to its image, which contains $y = \exp(ru)$.

Next, the variation σ is given by the formula

$$\sigma(s, t) = \alpha_t(s) = \exp_x(su(t)) \quad \text{such that} \quad \exp_x(s(t)u(t)) = \beta(t),$$

where $u(t)$ is a unit vector, and $s(t)u(t)$ is the vector whose exponential is $\beta(t)$. The polar coordinates $s(t)$ and $u(t)$ depend as smoothly on t as the exponential map, or its inverse. Then

$$\partial_2\sigma(s, t) = T\exp_x(su(t))su'(t),$$

so that (since $u = u(0)$),

$$\begin{aligned} \eta(s) &= T\exp_x(su)su'(0) \\ &= \eta_{u'(0)}(s), \end{aligned}$$

because from Theorem 3.1, we see that $D_{\alpha'}\eta(0) = u'(0)$. Since $u(t)^2 = 1$, it follows that $u'(0)$ is perpendicular to $\alpha'(0) = u$, so $D_{\alpha'}\eta$ is orthogonal to α'. Furthermore

$$\beta'(t) = T\exp_x(s(t)u(t))\big(s(t)u'(t) + s'(t)u(t)\big),$$

and since $s(0) = r$, we find

$$\begin{aligned} e = \beta'(0) &= T\exp_x(ru)(ru'(0) + s'(0)u) \\ &= T\exp_x(ru)ru'(0) + T\exp_x(ru)s'(0)u. \end{aligned}$$

Since e is assumed orthogonal to $\alpha'(r) = T\exp_x(ru)u$, and $u'(0)$ is also orthogonal to u, we must have $s'(0) = 0$, whence the relation

$$e = T\exp_x(ru)ru'(0) \quad \text{or} \quad \eta_{u'(0)}(r) = e.$$

This proves the proposition.

Transpose of $T\exp_x$

In the next results we are concerned with the differential of the exponential map at arbitrary points, namely for $v \in T_x$ such that $\exp_x$ is defined on the segment $[0, v]$, we are concerned with

$$T\exp_x(v)\colon T_x \to T_y, \quad \text{where} \quad y = \exp_x(v),$$

especially whether this map is an isomorphism, or what is its kernel.

Theorem 3.1 describes a condition for an element w to be in the kernel in terms of a zero for a suitable Jacobi lift. We shall exploit this condition to see that under some circumstances, there cannot be a non-trivial zero. We first give a lemma from Ambrose [Am 60].

Lemma 3.4. *Let (X, g) be pseudo Riemannian. Let η, ζ be Jacobi lifts of a geodesic α. Then*

$$\langle D_{\alpha'}\eta, \zeta\rangle_g - \langle \eta, D_{\alpha'}\zeta\rangle_g \quad \textit{is constant.}$$

Proof. We differentiate the above expression and expect to get 0. From the defining property of the covariant derivative, the derivative of the above expression is equal to

$$\begin{aligned}\langle D^2_{\alpha'}\eta, \zeta\rangle + \langle D_{\alpha'}\eta, D_{\alpha'}\zeta\rangle - \langle D_{\alpha'}\eta, D_{\alpha'}\zeta\rangle - \langle \eta, D^2_{\alpha'}\zeta\rangle \\ = \langle D^2_{\alpha'}\eta, \zeta\rangle - \langle D^2_{\alpha'}\zeta, \eta\rangle. \\ = R(\alpha', \eta, \alpha', \zeta) - R(\alpha', \zeta, \alpha', \eta) \\ = 0\end{aligned}$$

by the symmetry property of R. This proves the lemma.

The next lemma, from McAlpin's thesis [McA 65], describes the adjoint of the differential of the exponential map.

Lemma 3.5. *Let (X, g) be pseudo Riemannian. Let α (defined at least on $[0, 1]$) be the geodesic such that $\alpha(0) = x$ and $\alpha'(0) = v$. Let*

$$z \in T_{\alpha(1)}, \qquad w \in T_{\alpha(0)},$$

and let

$$v^* = -\alpha'(1) = -Pv,$$

where P is the parallel translation along α'. Then

$$\langle T\exp_{\alpha(0)}(v)w, z\rangle_{\alpha(1)} = \langle w, T\exp_{\alpha(1)}(v^*)z\rangle_{\alpha(0)}.$$

Proof. Let ζ be the Jacobi lift of α such that $\zeta(1) = 0$ and $D_{\alpha'}\zeta(1) = z$. Let η be the Jacobi lift as in Theorem 3.1. Then

$$\langle T\exp_x(v)w, z\rangle = \langle \eta(1), D_{\alpha'}\zeta(1)\rangle = \langle D_{\alpha'}\eta(1), \zeta(1)\rangle + C = C,$$

where C is the constant of Lemma 3.4. We compute C to be

$$C = -\langle D_{\alpha'}\eta(0), \zeta(0)\rangle = -\langle w, \zeta(0)\rangle.$$

Let $\mathrm{rev}(\alpha)$ be the reverse curve, so that $\mathrm{rev}(\alpha)(t) = \alpha(1-t)$, and let ξ be the unique Jacobi lift of $\mathrm{rev}(\alpha)$ such that

$$\xi(0) = 0 \qquad \text{and} \qquad D_{\mathrm{rev}(\alpha)'}\xi(0) = z.$$

Then in fact $\xi(t) = \zeta(1-t)$, and applying Theorem 3.1 concludes the proof.

Seminegative curvature

We apply the above results to the case of seminegative curvature. The next proposition gives us a criterion for the kernel of the differential of the exponential to be trivial, and we use Jacobi lifts in the proof.

Theorem 3.6. *Let (X, g) be Riemannian. Assume (X, g) has seminegative curvature. Then for all $x \in X$ and $v \in T_x$, $v \neq 0$, such that $\exp_x$ is defined on the segment $[0, v]$ in T_x, we have*

$$\|T\exp_x(v)w\|_g \geqq \|w\|_g \qquad \textit{for all } w \in T_xX.$$

In particular,

$$\operatorname{Ker} T\exp_x(v) = 0.$$

Proof. Let η_w be the Jacobi lift as in Proposition 3.1, so that

$$T\exp_x(v)w = \eta_w(1).$$

The asserted inequality is then a special case of the inequality found in Proposition 2.6. This inequality implies that $\operatorname{Ker} T\exp_x(v) = 0$, which concludes the proof.

Observe that the estimate on the differential of the exponential states that the inverse $T\exp_x(v)^{-1}$ is bounded by 1, as a continuous linear map. Of course, so far, this inverse is defined only on the image of $T\exp_x(v)$. In the finite dimensional case, invertibility is immediate. In the infinite dimensional case, it is in McAlpin's thesis [McA 65], as follows.

Theorem 3.7 (McAlpin [McA 65]). *Let (X, g) be a Riemannian–Hilbertian manifold with seminegative curvature, and let $x \in X$. Assume that $\exp_x$ is defined on all of T_x (what we called geodesically complete at x). Then for all $v \in T_x$ the map $T\exp_x(v)$ is a topological linear isomorphism, and in particular, $\exp_x$ is a local isomorphism.*

Proof. We have already proved that $T\exp_x(v)$ is injective and has a continuous inverse on its image. Lemma 3.5 shows that we can apply the same reasoning to the adjoint $(T\exp_x(v))^* = T\exp_y(v^*)$ for $y = \exp_x(v)$, so this adjoint also has kernel 0. Hence $T\exp_x(v)$ is surjective, thereby concluding the proof of the theorem. (See also Chapter X, §2.)

The next theorem was proved by Hadamard for surfaces [Ha 1898], by Cartan for finite dimensional Riemannian manifolds [Ca 28], and by McAlpin in the Hilbertian case [McA 65].

Theorem 3.8 (Cartan–Hadamard). *Let (X, g) be a Riemannian manifold, connected, and such that $\exp_x$ is defined on all of T_x for some $x \in X$ (so geodesically complete). If $R_2 \geqq 0$ (i.e. X has seminegative curvature), then the exponential map $\exp_x$: $T_xX \to X$ is a covering. In particular, if X is simply connected, then $\exp_x$ is an isomorphism.*

Proof. We have already proved that $\exp_x$ is a local isomorphism. There remains to prove that $\exp_x$ is surjective, and that it is a covering. But all the work has been done, because we simply apply Theorem 6.9 of Chapter VIII with $Y = T_x$ having the given metric $h = g(x)$, for which Y is certainly complete. Theorem 3.6 guarantees that the essential estimate hypothesis is satisfied, so that proof is complete.

Corollary 3.9. *Let (X, g) be a connected Riemannian manifold with seminegative curvature. Then (X, g) is complete if and only if the exponential map $\exp_x$ is defined on all of T_x for some $x \in X$, and therefore for every $x \in X$.*

Proof. That (X, g) complete implies $\exp_x$ defined on all of T_x was proved under all circumstances in Proposition 6.5 of Chapter VIII. The converse is now immediate from Theorem 2.10 and Theorem 6.9 of Chapter VIII.

We define a **Cartan–Hadamard manifold** to be a Riemannian manifold (X, g) which is complete, simply connected and with $R_2 \geqq 0$, i.e. seminegative curvature.

Corollary 3.10. *Let (X, g) be a Cartan–Hadamard manifold. Let $x \in X$. Then for all v, $w \in T_xX$ we have the inequality*

$$\mathrm{dist}_g(\exp_x(v), \exp_x(w)) \geqq \|v - w\|_g.$$

Proof. By Theorem 3.8 the exponential map has an inverse

$$\varphi\colon X \to T_xX$$

and by Theorem 3.6 this inverse satisfies

$$\|T\varphi(z)\|_g \leqq 1$$

for all $z \in X$, where the norm is that of a continuous linear map from T_zX to $T_{\varphi(z)}X$, with their structures of Hilbert spaces due to g. The inequality of the corollary is then immediate from the definition of the length of curves.

Corollary 3.11. *Suppose that (X, g) is a Cartan–Hadamard manifold. Then any two points can be joined by a unique geodesic whose length is the g-distance between the two points.*

Proof. Immediate from Corollary 3.10, because if x, y are the two points, then $y = \exp_x(v)$ for some $v \in T_xX$, and the geodesic α such that $\alpha(t) = \exp_x(tv)$ joins the two points, is unique by the Hadamard–Cartan theorem, and has length $\|v\|_g$.

Remark 1. The above corollary is of course mostly subsumed in the finite dimensional case by the Hopf–Rinow theorem, but it was noticed in the Hilbert case in McAlpin's thesis [McA 65]. Furthermore, McAlpin observed that one can define on the ball $\mathbf{B}(2/c)$ with $c > 0$ of a Hilbert space $\mathbf{E}$ a bounded seminegative metric, as in the finite dimensional case, namely for $x \in \mathbf{B}(2/c)$ and v, $w \in \mathbf{E}$ we let

$$\langle v, w\rangle_x = \frac{4\langle v, w\rangle}{4 - c^2x^2}.$$

Then the ball has curvature $-c^2$. Note that for constant curvature K one has

$$R(v, w)z = K(\langle z, w\rangle v - \langle z, v\rangle w).$$

Similarly one shows that the sphere has constant positive curvature. Standard proofs that the only simply connected manifolds with constant curvature are all of $\mathbf{E}$, a sphere of finite radius for positive curvature, and the above example for negative curvature, work in the Hilbert case, and will be given below.

Remark 2. Proposition 3.2 can be interpreted as implying that the geodesics which come from rays starting at the origin in the tangent space are orthogonal to the g-spheres in X. Of course it may happen that the exponential map is not an injective map of T_x into X (as on the circle or 2-sphere), so the orthogonality interpretation holds only when it makes sense. In the particular case of seminegative curvature and completeness of the Cartan–Hadamard theorem, the interpretation is valid everywhere. Note that Proposition 3.2 in the case of seminegative curvature is also a special case of the "local" result on orthogonality, Theorem 5.6 of Chapter VIII, because we have a global chart coming from the Cartan–Hadamard theorem, and the previous arguments are valid for this chart.

We now work out as examples the cases of constant curvature.

Theorem 3.12. *Let X be Riemannian complete, simply connected. Let $x_0 \in X$.*

(a) *If $R = 0$, i.e. if X has 0 curvature, then the exponential map*

$$\exp_{x_0}: T_{x_0}X \to X$$

is an isometry.

(b) *Suppose the curvature is constant, equal to -1. Let Y also be Riemannian complete, simply connected, and let $y_0 \in Y$. Let $L: T_{x_0}X \to T_{y_0}Y$ be a linear isometry, and let $f: X \to Y$ be defined by*

$$f = \exp_{y_0} \circ L \circ \exp_{x_0}^{-1}$$

so f is a differential isomorphism according to Theorem 3.8. Then f is an isometry. In other words, up to an isometry, there is only one complete Riemannian manifold with given constant negative curvature modeled on a given Hilbert space (finite dimensional or not).

Proof. For (a), we use Theorem 3.1 and Proposition 2.10 which shows that the exponential map amounts to parallel translation, so is an isometry. For (b), we argue in a similar way, but a bit more complicated. We have to show that for each $x \in T_{x_0}X$ the map

$$Tf(x): T_xX \to T_{f(x)}Y$$

is a linear isometry. Since $T\exp_{x_0}(0) = \text{id}$, it follows that $Tf(x_0) = L$, so $Tf(x_0)$ is a linear isometry. Assume $x \neq x_0$. Let $x = \exp_{x_0}(rv)$ with some unit vector $v \in T_{x_0}X$ and $r > 0$. Let $\eta^{(v)}$ denote the map which to each $w \in T_{x_0}X$ associates the Jacobi lift η_w of Theorem 3.1. Then

$$\begin{aligned} Tf(x) &= T\exp_{y_0}(L(rv)) \circ L \circ T\exp_{x_0}(rv)^{-1} \\ &= T\exp_{y_0}(rL(v)) \circ L \circ T\exp_{x_0}(rv)^{-1} \\ &= \frac{1}{r}\eta^{L(v)}(r) \circ L \circ \left(\frac{1}{r}\eta^{(v)}(r)\right)^{-1}. \end{aligned}$$

The map $T\exp_{x_0}(rv): T_{x_0}X \to T_xX$ is a linear isomorphism. To show that $Tf(x)$ preserves norms is equivalent to showing that

$$\|Tf(x) \circ T\exp_{x_0}(rv)w\| = \|T\exp_{x_0}(rv)w\| \qquad \text{for all } w \in T_xX.$$

But we have

$$\|Tf(x) \circ T\exp_{x_0}(rv)w\| = \|T\exp_{y_0}(rL(v))L(w)\| \qquad \text{for all } w \in T_{x_0}X.$$

We may now use Proposition 2.11 and Theorem 3.1 which describe $T\exp$ in terms of its components in the v-direction and a direction orthogonal to v, and parallel translation. Since in Proposition 2.11 the respective coefficients 1 and $(\sinh r)/r$ are the same whether we take $T\exp_{x_0}(rv)$ or $T\exp_{y_0}(rL(v))$ because L is an isometry, preserving orthogonality and changing unit vectors to unit vectors, it follows that in the notation of Proposition 2.11,

$$\|T\exp_{x_0}(rv)w\|^2 = c_0^2 + c_1^2\left(\frac{\sinh r}{r}\right)^2 = \|T\exp_{y_0}(rL(v))L(w)\|,$$

thus proving (b), and concluding the proof of the theorem.

We also have the following variation in the case of positive curvature.

Theorem 3.13. *Let X be Riemannian, complete, simply connected, with sectional curvature $+1$. Then X is isometric to the ordinary sphere of the same dimension in Hilbert space.*

Proof. The proof is similar, except that one cannot deal with the exponential defined on the whole tangent space $T_{x_0}X$. For convenience, we let X be the unit sphere in Hilbert space of a given dimension, and we let Y be Riemannian, complete simply connected with sectional curvature $+1$. We can then define the map f on the open ball of radius π. The same argument as before, replacing $\sinh r$ by $\sin r$, shows that f is a local isometry. We then pick another point $x_1 \neq \pm x_0$. We let

$$Tf(x_1) = L_1: T_{x_1}X \to T_{f(x_1)}Y.$$

Just as we defined $f = f_{x_0}$ from x_0, we can define $f_1 = f_{x_1}$ from x_1. Then f and f_1 coincide on the intersection of their domain, and thus define a local isometry $X \to Y$. By Theorem 6.9 of Chapter VIII, this local isometry is a covering map, and since Y is assumed simply connected it follows that f is a differential isomorphism, and hence a global isometry, thus proving the theorem.

Remark. The above theorems may be viewed as fitting a special case of a theorem of Cartan, cf. [BGM 71], Proposition E.III.2.

IX, §4. CONVEXITY THEOREMS

We begin with a formula for the variation of geodesics, and apply it to get a convexity theorem.

Let $\sigma = \sigma(s, t)$ be a variation of geodesics $\alpha_t\colon [a, b] \to X$ in a Riemannian manifold, so $\alpha_t(s) = \sigma(s, t)$. The geodesics α_t are not necessarily parametrized by arc length. We let

$$f(t) = L(\alpha_t)$$

be the length, and we put

$$h(s, t) = \|\partial_1\sigma(s, t)\|_g \qquad \text{so that} \qquad f(t) = \int_a^b h(s, t)\, ds.$$

For simplicity, we omit the subscript g and write $\|\partial_1\sigma(s, t)\|$. We can differentiate under the integral sign, so that

$$f'(t) = \frac{d}{dt}L(\alpha_t) = \int_a^b \partial_2 h(s, t)\, ds,$$

$$f''(t) = \frac{d^2}{dt^2}L(\alpha_t) = \int_a^b \partial_2^2 h(s, t)\, ds.$$

Hence to determine $f''(t)$ it suffices to determine $\partial_2 h$ and $\partial_2^2 h$. Having assumed that every α_t is a geodesic simplifies the computation. We note that

$$h = \langle \partial_1\sigma, \partial_1\sigma\rangle_g^{1/2}.$$

Theorem 4.1. *We have*

(1) $$\partial_2 h = \frac{1}{\|\partial_1\sigma\|}\langle D_2\partial_1\sigma, \partial_1\sigma\rangle_g,$$

(2) $$\partial_2^2 h = \frac{1}{\|\partial_1\sigma\|^3}\left((D_2\partial_1\sigma)^2(\partial_1\sigma)^2 - \langle D_2\partial_1\sigma, \partial_1\sigma\rangle_g^2\right) + \frac{1}{\|\partial_1\sigma\|}R_2(\partial_2\sigma, \partial_1\sigma).$$

Proof. The first formula comes directly from the definition of the metric (Levi-Civita) derivative. The second is obtained at once by using the rule for the derivative of a product, and setting

$$D_2^2\partial_1\sigma = R_2(\partial_2\sigma, \partial_1\sigma)\partial_2\sigma,$$

which is the Jacobi equation satisfied by the variation of geodesics. Then we take the scalar product with $\partial_1\sigma$ to obtain the term on the far right, with the Riemann tensor

$$R_2(\partial_2\sigma, \partial_1\sigma) = \langle R(\partial_2\sigma, \partial_1\sigma)\partial_2\sigma, \partial_1\sigma\rangle_g.$$

This concludes the proof. It is essentially the same as Lemma 2.5.

Theorem 4.2. *Let X be a Riemannian manifold, and let $\sigma = \sigma(s, t)$ be a variation of geodesics $\{\alpha_t\}$. Let u be the (varying) unit vector tangent to these geodesics, namely*

$$u = \partial_1\sigma / \|\partial_1\sigma\|_g.$$

Let $\tilde{v}$ be the orthogonalization

$$\tilde{v} = D_2\partial_1\sigma - \langle D_2\partial_1\sigma, u\rangle_g u.$$

Then

$$\tilde{v}^2 = (D_2\partial_1\sigma)^2 - \langle D_2\partial_1\sigma, u\rangle_g^2 \geqq 0,$$

and for the length $\ell(t) = L(\alpha_t)$, we have

$$\ell''(t) = \int_a^b \frac{1}{\|\partial_1\sigma\|}\left(\tilde{v}^2 + R_2(\partial_1\sigma, \partial_2\sigma)\right)(s, t)\, ds.$$

Proof. Immediate from Lemma 4.1 and the definitions.

Remark. From the expression for ℓ'', we see that usually one has the strict convexity $\ell'' > 0$. This occurs for instance if R_2 is strictly positive, or if $\tilde{v}^2$ is strictly positive. If there is some value of t such that $\tilde{v}(s, t)^2 = 0$ for all s, then $D_2\partial_1\sigma$ is proportional to u at this value of t.

In Chapter X, §1 we won't assume that each α_t is a geodesic, but we will be interested in another aspect, namely the special value at $t = 0$, that is $\ell''(0)$, so we shall carry out the computation in that context.

We define a **Hadamard**, or **Cartan–Hadamard** manifold to be a complete Riemannian manifold, simply connected, with seminegative curvature. We formulate the next two theorems locally on a convex set in a manifold with seminegative curvature. They apply globally as a special case to Cartan–Hadamard manifolds, where we can use Corollary 3.11.

Theorem 4.3. *Let X be a Riemannian manifold with seminegative curvature ($R_2 \geqq 0$), and U a convex open set. Let β_1, β_2 be disjoint geodesics in U, defined on the same interval. Let $\alpha_t\colon [a, b] \to U$ be the geodesic joining $\beta_1(t)$ with $\beta_2(t)$, and let $\ell(t) = L(\alpha_t)$. Then $\ell''(t) \geqq 0$ for all t.*

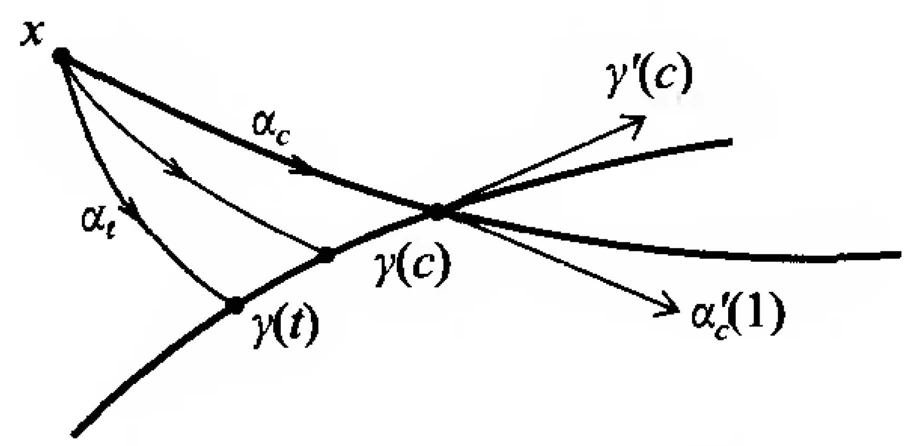
Proof. Immediate from Theorem 4.2 and the hypothesis that $R_2 \geqq 0$.

For estimates of $\ell(t)$, see [BuK 81], 6.6.

Next we consider the special case when β_1 is constant, i.e. we measure the distance between a given point and the points on a geodesic which does not contain the given point. We obtain a strict convexity result as follows.

Theorem 4.4. *Let X have seminegative curvature. Let U be a convex open subset. Let γ be a geodesic in U not containing a point $x \in U$. For each t at which γ is defined, let α_t: $[0, 1] \to U$ be the geodesic joining x with $\gamma(t)$. Let $\ell(t) = L(\alpha_t)$. Then $\ell''(t) > 0$ for all t. In particular, on an interval $[t_1, t_2]$ where γ is defined, the maximum of $L(\alpha_t)$ for $t \in [t_1, t_2]$ occurs only at the end points, with $t = t_1$ or $t = t_2$.*

Proof. The picture is as follows. We suppose there is a point c such that $\ell''(c) = 0$.



As in Theorem 4.2, let $\sigma(s, t) = \alpha_t(s)$, put $\alpha(s) = \sigma(s, c)$ and let

$$\eta(s) = \partial_2\sigma(s, c),$$

so η is a Jacobi lift of α. From the integral expression for $\ell''(c)$, using the variation formula (2), we conclude from the Schwarz inequality that $D_2\partial_1\sigma$ is proportional to $\partial_1\sigma$ at $t = c$. Using the standard fact $D_2\partial_1 = D_1\partial_2$ (Chapter VIII, Lemma 5.3), we conclude that $D_{\alpha'}\eta$ is proportional to α', i.e. there exists a function φ such that

$$D_1\partial_2\sigma(s, c) = \varphi(s)\partial_1\sigma(s, c), \qquad \text{that is} \quad D_{\alpha'}\eta = \varphi\alpha'.$$

We finish the proof using an argument shown to me by Quian. By Proposition 2.3, we can orthogonalize

$$\eta = \psi\alpha' + \xi,$$

where ξ is a lift of α orthogonal to α', and ψ is some function. By Proposition 2.4, we have also an orthogonal decomposition after applying $D_{\alpha'}$, that is

$$D_{\alpha'}\eta = \psi'\alpha' + D_{\alpha'}\xi.$$

Since $D_{\alpha'}\eta$ has been shown to be proportional to α', we conclude that $D_{\alpha'}\xi = 0$. Since $\eta(0) = 0$ it follows that $\xi(0) = 0$, and ξ being a Jacobi lift, it follows that $\xi = 0$, because a Jacobi lift is determined by initial conditions at a given point. Thus finally we obtain

$$\eta(1) = \psi(1)\alpha'(1), \qquad \text{that is} \qquad \gamma'(c) = \partial_2\sigma(1, c) = \psi(1)\alpha'(1).$$

This means that the geodesic γ is tangent to the geodesic α at the point $\gamma(c)$, and hence these two geodesics coincide since a geodesic is determined by its initial conditions at a given point. However, we assumed that x does not lie on γ, so we get a contradiction which concludes the proof.

Corollary 4.5. *Let X be a Cartan–Hadamard manifold. Then every ball in X is convex.*

Proof. Let x be the center of the ball, and let x_0, x_1 be points in the ball. If x lies on the geodesic between x_0 and x_1 then the Cartan–Hadamard theorem shows that this geodesic is the ray passing through the origin of the ball, so lies in the ball. If not, then we can apply Theorem 4.4.

Note. Concerning Theorem 4.4, readers may find it instructive to compare the version here with Lemma 5.15 in [ChE 75].

We shall complement the convexity theorem by a geometric description of the first derivative of the length. The statement is quite general.

Theorem 4.6. *Let X be a Riemannian manifold and let $x \in X$. Let U be a convex open set in X such that*

$$\exp_x\colon V \to U$$

is an isomorphism of some convex open set V in T_x containing 0_x, with U. Let γ be a curve in U not containing x, and let α_t be the geodesic segment from x to $\gamma(t)$. Let $\theta(t)$ be the angle between γ and α_t. Let the length of α_t be

$$\ell(t) = L(\alpha_t).$$

Then $\ell'(t) = \|\gamma'(t)\| \cos\theta(t)$.

In particular, if t_0 is such that $\ell(t_0)$ is a local minimum and $\gamma'(t_0) \neq 0$, then α_{t_0} is perpendicular to γ at $\gamma(t_0)$.

Proof. Let us first prove the result in euclidean space. Let $t \to v(t)$ be a curve in a euclidean space, and let $F(t) = \|v(t)\|$, with the euclidean norm denoted by the double bar. Then

$$\lim_{h\to 0+} \frac{\|v(t+h)\| - \|v(t)\|}{h} = \lim_{h\to 0+} \frac{1}{2\|v(t)\|h}\left(\|v(t+h)\|^2 - \|v(t)\|^2\right)$$

$$= \lim_{h\to 0+} \frac{1}{2\|v(t)\|h}\left(\|v(t+h) - v(t)\|^2 + 2\|v(t)\|\,\|v(t+h)-v(t)\|\cos\Theta_t\right)$$

where Θ_t is the euclidean angle from the law of cosines; namely for a euclidean triangle with sides a, b, c one has

$$c^2 = a^2 + b^2 + 2ab\cos\Theta$$

where Θ is the angle opposite the side c. But

$$\|v(t+h) - v(t)\|^2 = O(h^2) \qquad \text{for} \quad h \to 0,$$

so

$$F'(t) = \lim_{h\to 0+} \frac{\|v(t+h)\| - \|v(t)\|}{h} = \|v'(t)\|\cos\Theta_t.$$

This proves the formula in the euclidean case.

For the general case, let $t \mapsto v(t)$ be a curve in V such that $\exp_x v(t) = \gamma(t)$. Let

$$\alpha_t(s) = \exp_x(sv(t)), \qquad 0 \leqq s \leqq 1,$$

so that α_t is the geodesic between x and $\gamma(t) = \alpha_t(1)$. Then

$$\alpha_t'(1) = T\exp_x(v(t))v(t) \qquad \text{and} \qquad \gamma'(t) = T\exp_x(v(t))v'(t).$$

By the global Gauss lemma, Proposition 3.2, we have

$$\langle \alpha_t'(1), \gamma'(t)\rangle_g = \langle v(t), v'(t)\rangle_{g(x)},$$

where the scalar product on the left is taken in the tangent space at $\gamma(t)$, and the scalar product on the right is taken in the tangent space at x. By definition of the usual formula for scalar products, we obtain

$$\|\alpha_t'(1)\|_g \|\gamma'(t)\|_g \cos\theta(t) = \|v(t)\|_g \|v'(t)\|_g \cos\Theta_t.$$

We have $\|\alpha_t'(1)\|_g = \|v(t)\|_g$ because $\exp_x$ preserves distance along rays. Thus we obtain the relation

$$\text{(3)} \qquad \|\gamma'(t)\|_g \cos\theta(t) = \|v'(t)\|_g \cos\Theta_t.$$

We apply this to show that $\ell' = F'$, namely

$$\ell'(t) = \lim_{h\to 0+} \frac{\ell(t+h) - \ell(t)}{h} = \lim_{h\to 0+} \frac{1}{2\ell(t)h}\left(\ell(t+h)^2 - \ell(t)^2\right)$$

$$= F'(t) = \|v'(t)\|_{g(x)} \cos\Theta_t,$$

because $\ell(t) = L(\alpha_t) = \|v(t)\|_{g(x)}$, and we can apply the euclidean result. This proves the theorem. The proof more or less follows that of Helgason, [He 78], Chapter I, Lemma 13.6, but the theorem is valid in greater generality than stated by Helgason.

Remarks. If the curve γ is parametrized by arc length, then the factor $\|\gamma'(t)\|$ disappears from the formula, and we simply get

$$\ell'(t) = \cos\theta(t).$$

Note that the theorem applies globally to a Cartan–Hadamard manifold, but curvature considerations were not necessary for the formula to be true. However, the next theorem provides seminegative conditions under which Theorem 4.6 can be applied.

Corollary 4.7. *Let X be a Cartan–Hadamard manifold. Let $x \in X$ and let γ be a geodesic which does not contain x. Then the distance $d(x, \gamma(t))$ has a unique minimum for some value t_0. The geodesic from x to $\gamma(t_0)$ is perpendicular to γ at $\gamma(t_0)$.*

Proof. That the distance has a minimum comes from the fact that the geodesic distance goes to infinity as $t \to \pm\infty$. Because the line is locally compact, there is some minimum, and the convexity Theorem 4.4 shows that this is the only minimum, with the distance being strictly decreasing for $t \leqq t_0$ and strictly increasing for $t \geqq t_0$. Theorem 4.6 concludes the proof.

Since two distinct points in a Cartan–Hadamard space are joined by a unique geodesic, it follows that two distinct geodesics can intersect in only one point.

Next we give an application of the metric increasing property as in Helgason [He 78], Chapter I, Corollary 13.2.

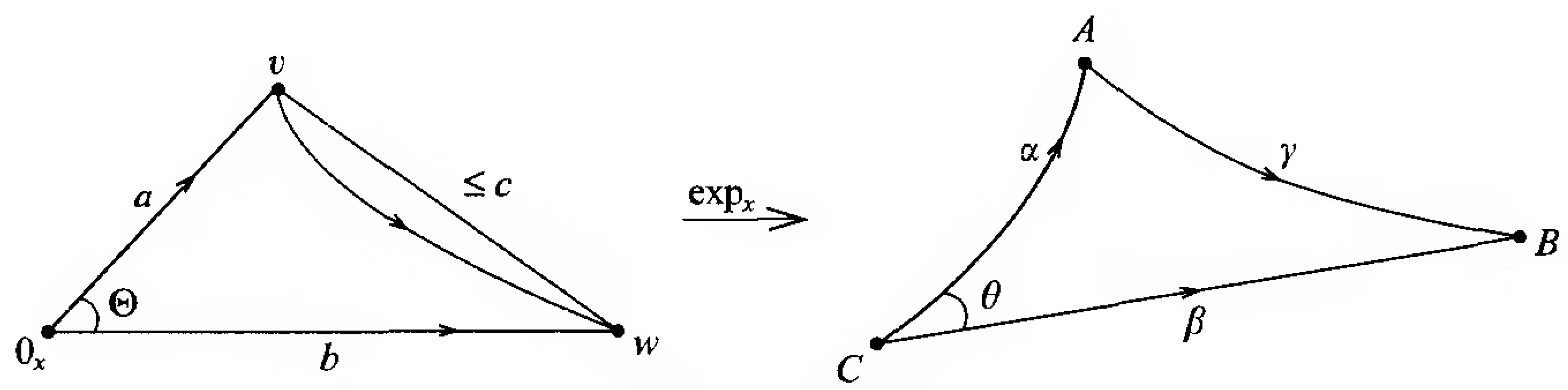
Theorem 4.8. *Let X be a Cartan–Hadamard manifold. Let ABC be a geodesic triangle whose angles are A, B, C and whose sides are geodesics of lengths a, b, and c. Then:*

(i) $a^2 + b^2 \leqq c^2 + 2ab\cos C$;
(ii) $A + B + C \leqq \pi$.

Proof. Let x be the vertex of angle C. Let $\exp_x(v)$ and $\exp_x(w)$ with v, $w \in T_x$ be the vertices with angles A, B respectively. Then the geodesic sides of angle C are α, β respectively, with

$$\alpha(s) = \exp_x(sv) \quad \text{and} \quad \beta(s) = \exp_x(sw),$$

and $0 \leqq s \leqq 1$. The picture is as follows.



We let Θ and θ be the angles as shown, and $\cos \theta = \cos C$ by definition. Actually, we have

(4) $$\cos \theta = \cos \Theta.$$

Indeed,

$$\langle \alpha'(0), \beta'(0) \rangle_g = \langle T\exp_x(0)v, T\exp_x(0)w \rangle_{g(x)} = \langle v, w \rangle_{g(x)}.$$

The left side is equal to $\|\alpha'(0)\|_g \|\beta'(0)\|_g \cos \theta$, and the right side is equal to $\|v\|_x \|w\|_x \cos \Theta$. Trivially $\alpha'(0) = v$ and $\beta'(0) = w$, so (4) follows. So far, we have not used seminegative curvature. It comes next.

We have $a^2 + b^2 = \text{dist}(v, w)^2 + 2ab \cos \Theta$. By the distance increasing property of the exponential map, the inequality (i) follows.

As for (ii), since each geodesic side of the geodesic triangle has length at most equal to the sum of the other two sides, it follows that there exists a euclidean triangle with sides of lengths a, b, c. Let Θ_C be the angle of this euclidean triangle corresponding to C. Then

$$a^2 + b^2 = c^2 + 2ab \cos \Theta_C.$$

By (i) it follows that $\cos C \geqq \cos \Theta_C$, and hence $\Theta_C \geqq C$. Similarly, $\Theta_A \geqq A$ and $\Theta_B \geqq B$. But

$$\Theta_A + \Theta_B + \Theta_C = \pi.$$

This proves (ii) and concludes the proof of the theorem.

The moral lesson of the above results is that Cartan–Hadamard manifolds behave like, or better than, ordinary euclidean space.

IX, §5. TAYLOR EXPANSIONS

We shall deal systematically with the Taylor expansion of various curves. We consider a curve in X, say of class C^2, not necessarily a geodesic,

$$\alpha \colon J \to X,$$

and we assume $0 \in J$, so 0 is an origin. We suppose given a spray on TX, giving rise to the covariant derivative. For $w \in T_{\alpha(0)}X$ we let

$$\gamma \colon t \mapsto \gamma(t, w)$$

be the unique α-parallel curve with initial condition $\gamma(0, w) = w$. Recall that α-parallel means $D_{\alpha'}\gamma = 0$. We denote parallel translation by

$$P^t = P^t_\alpha = P^t_{0,\alpha} \colon T_{\alpha(0)}X \to T_{\alpha(t)}X$$

Then P^t is topological linear isomorphism, as we saw in Chapter VIII, §3.

Proposition 5.1. *Let $\eta \colon J \to TX$ be a lift of α in TX. Then*

$$\eta(t) = P^t \sum_{k=0}^{m} D^k_{\alpha'}\eta(0) \frac{t^k}{k!} + O(t^{m+1}) \qquad \textit{for } t \to 0;$$

or alternatively,

$$\eta(t) = \sum_{k=0}^{m} \gamma(t, D^k_{\alpha'}\eta(0)) \frac{t^k}{k!} + O(t^{m+1}) \qquad \textit{for } t \to 0.$$

Proof. The second expression is merely a reformulation of the first, taking into account the definition of parallel translation. Since $t \to 0$, the formula is local, and we may prove it in a chart, so we use η, γ to denote the vector components η_U, γ_U in a chart U, suppressing the index U. Let

$$\beta(t) = \eta(t) - \sum_{k=0}^{m} \gamma(t, D^k_{\alpha'}\eta(0)) \frac{t^k}{k!}.$$

From the existence and uniqueness of the ordinary Taylor formula, it will suffice to prove that for the ordinary derivatives of β, we have

$$\partial^k \beta(0) = \beta^{(k)}(0) = 0 \qquad \text{for } k = 0, \ldots, m.$$

By definition, note that $\beta(0) = 0$. Let $w_k = D^k_{\alpha'}\beta(0)$. Since $D_{\alpha'}\gamma = 0$, we have

$$D^j_{\alpha'}\beta(t) = D^j_{\alpha'}\eta(t) - \sum_{k \geqq j} \gamma(t, w_k)\frac{t^{k-j}}{(k-j)!}.$$

Therefore

$$\begin{aligned} D^j_{\alpha'}\beta(0) &= D^j_{\alpha'}\eta(0) - \gamma(0, w_j) \\ &= w_j - w_j \\ &= 0. \end{aligned}$$

We now need a lemma. We let $\mathbf{E}$ be the Banach space on which X is modeled.

Lemma 5.2. *Let $\beta\colon J \to \mathbf{E}$ be the vector component of a lift of α. If $D^j_{\alpha'}\beta(0) = 0$ for $0 \leqq j \leqq m$ then $\partial^j\beta(0) = 0$ for $0 \leqq j \leqq m$.*

Proof. By definition,

$$D_{\alpha'}\beta = \beta' - B(\alpha; \alpha', \beta).$$

Hence $D_{\alpha'}\beta(0) = \beta'(0)$. We can proceed by induction. Let us carry out the case of the second derivative so the reader sees what's going on. Hence suppose in addition that $D^2_{\alpha'}\beta(0) = 0$. From the definitions, we get

$$\begin{aligned} D^2_{\alpha'}\beta = \beta'' &- [\partial_1 B(\alpha; \alpha', \beta)\alpha' + B(\alpha; \alpha'', \beta) + B(\alpha; \alpha', \beta')] \\ &- B(\alpha; \alpha', \beta' - B(\alpha; \alpha', \beta)). \end{aligned}$$

Since $\beta(0) = \beta'(0) = D_{\alpha'}\beta(0) = 0$ we find that

$$0 = D^2_{\alpha'}\beta(0) = \beta''(0),$$

thus proving the assertion for $m = 2$. The inductive proof is the same in general.

We apply the above considerations to Jacobi lifts.

Proposition 5.3. *Suppose that α is a geodesic. Let $w \in T_{\alpha(0)}X$ and let η_w be the Jacobi lift of α such that $\eta_w(0) = 0$ and $D_{\alpha'}\eta_w(0) = w$. Then*

$$\eta_w(t) = P^t\left[wt + R(\alpha'(0), w, \alpha'(0))\frac{t^3}{3!}\right] + O(t^4).$$

Proof. We plug in Proposition 5.1. Since $D^2_{\alpha'}\eta_w = R(\alpha', \eta_w, \alpha')$ contains η_w linearly, the evaluation of the second term of the Taylor expansion at 0 is 0. As for the third term, we have to use the chain rule. To be sure we don't forget anything, we should write more precisely

$$R(\alpha', \eta_w, \alpha') = R(\alpha; \alpha', \eta_w, \alpha')$$

to make explicit the dependence on the extra position variable. But it turns out that it does not matter in the end, because no matter what, the chain rule gives

$$D^3_{\alpha'}\eta_w = R(\alpha', D_{\alpha'}\eta_w, \alpha') + \text{terms containing } \eta_w \text{ linearly},$$

so $D^3_{\alpha'}\eta_w(0) = R(\alpha'(0), w, \alpha'(0))$, which proves the proposition.

From Proposition 5.3, we get information on the pull back of the metric g of a pseudo Riemannian manifold, to the tangent space at a given point.

Proposition 5.4. *Let (X, g) be a pseudo Riemannian manifold, and let $x \in X$, Fix v, $w \in T_xX$. Then*

$$\exp_x^*(tv)(g)(w, w) = w^2 + \tfrac{1}{3}R_2(v, w)t^2 + O(t^3) \qquad \textit{for } t \to 0.$$

where we recall that $R_2(v, w) = R(v, w, v, w)$.

Proof. From the theory of Jacobi lifts, applied to $\alpha(t) = \exp_x(tv)$, we have the formula

$$\frac{1}{t}\eta_w(t) = T\exp_x(tv)w.$$

Therefore modulo functions which are $O(t^3)$ for $t \to 0$, we get from Proposition 5.3

$$\begin{aligned} \exp_x^*(g)(tv)(w, w) &\equiv \left\langle \frac{1}{t}\eta_w(t), \frac{1}{t}\eta_w(t)\right\rangle_{g(\alpha(t))} \\ &\equiv \left\langle P^t\left[w + R(v, w, v)\frac{t^2}{3!}\right], P^t\left[w + R(v, w, v)\frac{t^2}{3!}\right]\right\rangle_{g(\alpha(t))} \\ &\equiv \left\langle w + R(v, w, v)\frac{t^2}{3!}, w + R(v, w, v)\frac{t^2}{3!}\right\rangle_{g(x)} \\ &\equiv w^2 + 2R_2(v, w)\frac{t^2}{3!}, \end{aligned}$$

which proves the proposition.

The preceding proposition gives us the Taylor expansion of the mapping

$$f(v) = \exp_x^*(g)(v) \quad \text{for} \quad v \in T_xX$$

along rays through the origin. Observe that

$$f\colon T_xX \to L^2_{\text{sym}}(T_xX)$$

is a map of the self-dual Banach space T_xX to the space of symmetric bilinear forms on T_xX (actually the open subset of non-singular forms). The map f has a Taylor expansion

$$(*) \qquad f(v) = f(0) + f_1(v) + f_2(v) + O(|v|^3) \quad \text{for} \quad |v| \to 0,$$

where $v \mapsto |v|$ is a Banach norm on T_x, and where f_1 and f_2 are homogeneous of degree 1 and 2 respectively. Since the homogeneous terms in the Taylor expansion are uniquely determined by f, and since we computed their restrictions to rays through the origin in Proposition 5.4, we now obtain:

Theorem 5.5. *Let (X, g) be a pseudo Riemannian manifold. Let $x \in X$. For $v \in T_xX$ let $q(v) \in L^2_{\text{sym}}(T_xX)$ be the symmetric bilinear function such that*

$$q(v)(w_1, w_2) = \frac{1}{3} R_x(v, w_1, v, w_2).$$

Let the metric g be viewed as a tensor in $L^2_{\text{sym}}(TX)$, and let f be the pull back of the metric g in a star shaped neighborhood of 0_x in T_xX where the exponential map is defined. Then

$$f(v) = g(x) + q(v) + O(|v|^3) \quad \textit{for} \quad |v| \to 0.$$

CHAPTER X

Jacobi Lifts and Tensorial Splitting of the Double Tangent Bundle

Chapter IX dealt with two related main topics, involving estimates for the exponential maps via Jacobi lifts, and the Cartan–Hadamard theorem concerning seminegative curvature. The present chapter goes somewhat deeper into both topics. In addition, it is instructive to treat systematically the splitting of the double tangent bundle. A special case is treated ad hoc in the extension of the Cartan–Hadamard theorem to the normal bundle of a totally geodesic submanifold in §2, but it is worth while understanding the fuller structure in general.

Logically, the rest of the book does not depend on this chapter, which may therefore be omitted by readers interested in the other subsequent topics. In Chapter XI, we return to manifolds with seminegative curvature in another context. Chapter XII gives a prototype example for Cartan–Hadamard manifolds. Thus Chapters IX through XII provide a much more systematic account of seminegative curvature than is usually found in differential geometry texts.

X, §1. CONVEXITY OF JACOBI LIFTS

This section complements §2 of Chapter IX. We investigate the convexity of Jacobi lifts. We compute the second derivative of the square, which turns out to be semipositive or positive under the condition $R_2 \geqq 0$ (seminegative curvature). This section and the next are based on notes of Wu, leading up to an extension of the Cartan–Hadamard theorem. I am much indebted to Wu for his notes and explanations. Cf. also the historical note in §2.

Lemma 1.1. *Let (X, g) be a pseudo-Riemannian manifold and let α be a geodesic. Let η be a Jacobi lift, and*

$$f(s) = \eta(s)^2 = \langle \eta(s), \eta(s)\rangle_g.$$

Then

$$f' = 2\langle D_{\alpha'}\eta, \eta\rangle_g \quad \text{and} \quad f'' = 2R_2(\alpha', \eta) + 2(D_{\alpha'}\eta)^2.$$

If X is Riemannian and $R_2 \geqq 0$ (seminegative curvature), then $f'' \geqq 0$.

Proof. The first derivative comes from the defining property of the Levi-Civita (metric) derivative along curves, as in Chapter VIII, Theorem 4.3. This same reference then also yields the second derivative

$$\begin{aligned} f'' &= 2\langle D^2_{\alpha'}\eta, \eta\rangle_g + 2\langle D_{\alpha'}\eta, D_{\alpha'}\eta\rangle_g \\ &= 2\langle R(\alpha', \eta)\alpha', \eta\rangle_g + 2(D_{\alpha'}\eta)^2 \end{aligned}$$

by the Jacobi differential equation. This proves the formulas. The final statement is then immediate, thus concluding the proof.

Theorem 1.2. *Let X be a Riemannian manifold with $R_2 \geqq 0$ (seminegative curvature). Let α be a geodesic and η a Jacobi lift with $\eta(0) = 0$ but $D_{\alpha'}\eta(0) \neq 0$. Let*

$$f(s) = \eta(s)^2.$$

Then $f(0) = f'(0) = 0$. Furthermore, we have convexity:

$$f''(s) \geqq 0 \quad \text{for all } s.$$

Thus $f'(s) \leqq 0$ for $s < 0$ and $f'(s) \geqq 0$ for $s > 0$, with the corresponding semi-decreasing and semi-increasing properties of f for $s \leqq 0$ and $s \geqq 0$ respectively.

Proof. Immediate from the definitions and assumption on R_2, taking Lemma 1.1 into account.

Remark. The quantity $R_2(\alpha', \xi) + (D_{\alpha'}\xi)^2$ with various lifts ξ of α will come up repeatedly in Chapter XI, in several variational contexts. Lemma 1.1 and Theorem 1.2 perhaps give the simplest manifestation of the phenomenon involved.

Next we consider the situation of Jacobi lifts which are not 0 at the initial point, but whose covariant derivative may or may not be 0.

We let α be a geodesic with initial value

$$\alpha(0) = y \in X, \quad \text{so that} \quad \alpha(s) = \exp_y(s\alpha'(0)).$$

We let β be a geodesic with initial conditions

$$\beta(0) = \alpha(0) \quad \text{and} \quad \beta'(0) = z \in T_yX,$$

so α and β start at the same point.

We let ζ be a lift of β in TX, with the same initial value as α', that is

$$\zeta(0) = \alpha'(0).$$

With these data, we define the **ζ-variation of α along β**, or also the **(β, ζ)-variation of α** (at the beginning point) to be

$$\alpha_t(s) = \sigma(s, t) = \exp_{\beta(t)} s\zeta(t).$$

This is trivially a variation through geodesics, and we have the initial conditions

(1) $$\alpha_t(0) = \beta(t), \quad \alpha_t'(0) = \zeta(t) = \partial_1\sigma(0, t).$$

We let η be the **Jacobi lift of α coming from this variation**, that is

(2) $$\eta(s) = \partial_2\sigma(s, 0), \quad \text{so that} \quad \eta(0) = z.$$

The variation and Jacobi lift are designed to investigate the growth of η as in Proposition 2.6 of Chapter IX, rather than $D_{\alpha'}\eta(0)$. As in Lemma 1.1, we let

$$f(s) = \eta(s)^2 = \langle \eta(s), \eta(s)\rangle_g.$$

We shall find an expression for f' different from the one in Lemma 1.1.

Lemma 1.3. *Let η be the Jacobi lift of α coming from its (β, ζ)-variation at the beginning point. Let $f = \eta^2$. Then*

$$f'(0) = 2\langle D_{\eta(0)}\zeta, \eta(0)\rangle_g.$$

Proof. Starting with the expression in Lemma 1.1, we get

$$\begin{aligned} f'(0) &= 2\langle D_{\alpha'(0)}\eta, \eta\rangle_g(0) \\ &= 2\langle D_{\zeta(0)}\eta, \eta(0)\rangle_g \\ &= 2\langle D_{\eta(0)}\zeta, \eta(0)\rangle_g. \end{aligned}$$

In this step, we need for the covariant derivatives of curves the analogue of the formula for the covariant derivative of vector fields, with the difference being formally equal to $[\eta, \zeta](0)$. Furthermore, from (1) and (2), η and ζ are obtained as the images under σ of the commuting vertical and horizontal unit vector fields in the (s, t)-plane, so the bracket is equal to 0. We let the reader fill in the details of the above arguments, to conclude the proof.

We end this section by specializing the situation somewhat with conditions which arise for the application we have in mind in §2.

Let Y be a submanifold of a Riemannian manifold X. We define Y to be **totally geodesic** if Y is closed, and if every geodesic in X with initial conditions in (Y, TY) is contained in Y. It is a simple matter of foundations, which will be taken care of systematically in Chapter XIV, §1, especially Theorem 1.4, that a geodesic in Y is also a geodesic in X. The next proposition provides a tool corresponding to Chapter IX, Proposition 2.6.

Proposition 1.4. *Let X be a Riemannian manifold and let Y be a totally geodesic submanifold. Let α be a geodesic in X, $\alpha(0) = y \in Y$. Let σ be the (β, ζ)-variation of α defined above. We suppose that β is a geodesic in Y, so in particular, $\beta'(0) = z \in T_Y Y$. Let η be the corresponding Jacobi lift of α, and let $f = \eta^2$.*

(i) *If ζ is orthogonal to Y, i.e. its values are in NY, then $f'(0) = 0$.*
(ii) *If in addition X has $R_2 \geqq 0$ (seminegative curvature), then $f(s)$ is weakly decreasing for $s \leqq 0$, weakly increasing for $s \geqq 0$, and*

$$f(s) \geqq f(0) \qquad \textit{for all } s,$$

so $\|\eta(s)\| \geqq \|\eta(0)\|$ *for all* s.

Proof. Since Y is totally geodesic, the second fundamental form $h_{12}(\eta, \zeta)(0) = 0$ by Theorem 1.4 of Chapter XIV. Then combining Theorem 1.5 of Chapter XIV and Lemma 1.3 which was just proved, we obtain $f'(0) = 0$. The other assertions are immediate from the convexity $f''(0) \geqq 0$ of Lemma 1.1. This concludes the proof.

Remarks. Proposition 1.4 will be used only in the next section. The foundational material of Chapter XIV occurs in parallel to the considerations of the present chapter, with intersection just at this single point. Thus I decided in favor of the present organization, with a localized appeal to Chapter XIV, which will not interfere with the general logical development. Similarly, in the next section, we shall also appeal to Chapter XIV for the fact that in a totally geodesic submanifold, parallel translation is the same as in the ambient manifold.

More convexity results will be given in §3, but I found it worthwhile to give immediately an application of Proposition 1.4 in the next section.

X, §2. GLOBAL TUBULAR NEIGHBORHOOD OF A TOTALLY GEODESIC SUBMANIFOLD

In Chapter IX, we dealt mostly with the exponential map defined on a fixed tangent space $T_x X$. We shall now consider systematically the exponential map on the tangent bundle, and some applications extending the Cartan–Hadamard theorem. Like the preceding section, the exposition is based on notes of Wu.

Let X be a Riemannian manifold. We recall that an open subset U of X is said to be **convex** if between two points of X there is a unique geodesic of X lying in U joining the two points, and the length of this geodesic is the distance between the points. In particular, X itself is an open set, in which case we may speak of X being **convex**. For example, a Cartan–Hadamard manifold is convex by Corollary 3.11 of Chapter IX. However, some of the formalism to be used is local, and it is appropriate to formulate it as such, but this involves additional notation to describe small neighborhoods of points.

We shall again deal with a totally geodesic submanifold Y, and from Chapter XIV, Theorem 1.4, we shall invoke that along geodesics in Y, parallel translation with respect to Y is the same as with respect to X. In particular, parallel translation in X between two points y_0, y in Y maps $T_{y_0} Y$ on $T_y Y$. Since parallel translation preserves the scalar product, we conclude that it induces a metric toplinear isomorphism

$$P_{y_0}^y : N_{y_0} Y \to N_y Y.$$

Let X be convex (so in general, we may be dealing with an open subset of an arbitrary Riemannian manifold). *Let Y be a totally geodesic submanifold.* If X is complete then we may compose the exponential map with parallel translation, and for given y_0, we may define the mapping

$$E = E_{y_0} : Y \times N_{y_0} Y \to X$$

by the formula

$$E(y, v) = \exp_y P_{y_0}^y(v) \qquad \text{for } v \in N_{y_0} Y.$$

If X is not complete, then we may fix v, and let y lie in some open set where $\exp_y$ is defined at $P_{y_0}^y(v)$. Then E is defined on a product

$$E : U \times V \to X,$$

where U is a neighborhood of some point y, and $v \in V$.

Lemma 2.1. *Let β be a geodesic in Y with $\beta(0) = y$ and $\beta'(0) = z$. For $v \in N_{y_0}Y$, let $\beta_v(t) = (\beta(t), v)$. Let*

$$\varphi_v(t) = E(\beta_v(t)) = \exp_{\beta(t)} P_{y_0}^{\beta(t)}(v).$$

Then

$$\varphi_v'(0) = TE(y, v)(z, 0).$$

Proof. This is just the chain rule.

We let y, y_0 be two points of X, and $v \in N_{y_0}Y$. We let α be the geodesic with initial conditions

$$\alpha(0) = y \quad \text{and} \quad \alpha'(0) = P_{y_0}^y(v);$$

so that

$$\alpha(s) = \exp_y sP_{y_0}^y(v).$$

We shall now associate a (β, ζ)-variation of α at its beginning point. Let β be a geodesic in Y, with initial conditions

(1) $$\beta(0) = \alpha(0) \quad \text{and} \quad \beta'(0) = z,$$

so α and β start at the same point. Let

(2) $$\zeta(t) = P_{y_0}^{\beta(t)}(v) \quad \text{so that} \quad \zeta(0) = \alpha'(0),$$

and

(3) $$\sigma(s, t) = \exp_{\beta(t)} s\zeta(t).$$

This (β, ζ)-variation of α will be called the **parallel variation of α along β** depending on (y_0, y).

Proposition 2.2. *Let X be convex and let Y be a totally geodesic submanifold. Let $y_0, y \in Y$. Let $v \in N_{y_0}Y$. Let β, ζ be the curves defined in* (1) *and* (2) *above, and let η be the Jacobi lift associated with the variation σ defined in* (3). *Then*

$$\eta(0) = z \quad \textit{and} \quad \eta(1) = TE(y, v)(z, 0).$$

Proof. Putting $s = 0$ in the definition of σ, we obtain

$$\sigma(0, t) = \exp_{\beta(t)}(0) = \beta(t),$$

so the value $\eta(0) = z$ drops out. For $\eta(1)$, we just apply Lemma 2.1 to conclude the proof.

Next we start global considerations.

Theorem 2.3. *Let X be a convex complete Riemannian manifold, and let Y be a totally geodesic submanifold. Then Y is also convex complete. Let $y_0 \in Y$ and let*

$$P_{y_0}\colon Y \times N_{y_0}Y \to NY$$

be the map such that for each $y \in Y$ and $v \in N_{y_0}Y$ we have

$$P_{y_0}(y, v) = P_{y_0}^y(v).$$

Then P_{y_0} is a vector bundle isomorphism, trivializing the normal bundle.

Proof. This simply amount to the fact that flows of differential equations depend smoothly on parameters, and that parallel translation is invertible by parallel translation along the reverse geodesic.

Given a chart U of Y at y_0, it follows that $U \times T_{y_0}$ is a chart at the corresponding point in NY. Of course, Y itself admits a global chart, given for instance by its own exponential mapping at y_0. So once the point y_0 is selected, there is a canonical way of constructing a global chart for the normal bundle. The next application will be global.

We shall always take $Y \times N_{y_0}Y$ with its Riemannian product structure. Thus Y has the Riemann metric restricted from X, and N_{y_0} has its positive definite scalar product restricted from $T_{y_0}X$, so the "constant" Riemann metric is on the fiber. At each point, the product has the Hilbert space product metric satisfying the Pythagoras theorem.

Theorem 2.4 (Wu). *Let X be a Cartan–Hadamard manifold. Let Y be a totally geodesic submanifold. Fix a point $y_0 \in Y$. Let*

$$E\colon Y \times N_{y_0}Y \to X$$

be defined by $E(y, v) = \exp_y P_{y_0}^y(v)$ for $v \in N_{y_0}Y$. Then E is metric semi-increasing.

Proof. For $z \in T_yY$ and v, $w \in N_{y_0}Y$ we have to show that

$$\|TE(y, v)(z, w)\| \geqq \|(z, w)\|.$$

The product Hilbert space metric by definition gives

$$\|(z, w)\|^2 = \|(z, 0)\|^2 + \|(0, w)\|^2 = \|z\|^2 + \|w\|^2.$$

The Gauss Lemma 5.6 of Chapter VIII, §5 implies that

$$\|TE(y, v)(z, w)\|^2 = \|TE(y, v)(z, 0)\|^2 + \|TE(y, v)(0, w)\|^2.$$

Hence we need only prove separately that

$$\|TE(y, v)(z, 0)\| \geqq \|(z, 0)\| = \|z\|$$

and

$$\|TE(y, v)(0, w)\| \geqq \|(0, w)\| = \|w\|.$$

The second inequality is simply the metric semi-increasing property of Chapter IX, Theorem 3.6. As to the first inequality, we may now quote Proposition 1.4 (ii) and Lemma 2.1 to conclude the proof.

We arrive at the extension of the Cartan–Hadamard theorem, what we shall call the **global tubular neighborhood property**.

Theorem 2.5. *Let X be a Cartan–Hadamard manifold and let Y be a totally geodesic submanifold. Let $NY = N_X Y$ be the normal bundle over Y. Let*

$$\exp_{NY} : NY \to X$$

be the restriction of the exponential map to the normal bundle, so we call $\exp_{NY}$ *the* **tubular neighborhood map**. *Then* $\exp_{NY}$ *is a differential isomorphism, so that Y admits a global tubular neighborhood.*

Proof. Fix a point $y_0 \in Y$. Parallel translation P_{y_0} gives a differential isomorphism from $Y \times N_{y_0}$ to NY be Theorem 2.3, and we transport the product Riemannian metric to NY via this isomorphism. Then $\exp_{NY}$ is metric semi-increasing. By Theorem 6.9 of Chapter VIII, it suffices to prove that $\exp_{NY}$ is a local C^1-isomorphism. In the finite dimensional case, we are done, just as for the ordinary Cartan–Hadamard theorem. In the infinite dimensional case, we have to argue a bit longer.

We now let v denote an element of NY, say $v \in N_y Y$, and it suffices to prove that $T\exp_{NY}(v)$ is invertible, by the inverse mapping theorem. Suppose $T\exp_{NY}(v)$ is not invertible. Let $r > 0$ be the smallest value such that $T\exp_{NY}(rv)$ is not invertible. Such a value exists because $\exp_{NY}$ is locally invertible at each point of the zero section, this being the tubular neighborhood theorem. Let

$$L(s) = T\exp_{NY}(sv) \qquad \text{for} \qquad 0 \leqq s \leqq r.$$

Then $\{L(s)\}$ is a family of bounded operators

$$L(s) : \mathbf{E}(s) \to \mathbf{F}(s),$$

where $\{\mathbf{E}(s)\}$, $\{\mathbf{F}(s)\}$ are the families of tangent Banach spaces varying continuously. One has the following trivial lemma.

Lemma 2.6. *Let $\mathbf{E}$, $\mathbf{F}$ be Banach spaces. Let $\{A(s)\}$ $(0 \leqq s \leqq r)$ be a continuous family of bounded operators, such that $A(s) : \mathbf{E} \to \mathbf{F}$ is invertible for $0 \leqq s < r$, and there is a uniform lower bound $c > 0$ such that*

$$|A(s)| \geqq c \qquad \text{for} \qquad 0 \leqq s < r.$$

Then $\lim_{s \to r} A(s)^{-1}$ exists and is a bounded operator inverse of $A(r)$.

Proof. We write

$$A(s)^{-1} - A(s')^{-1} = A(s')^{-1}\big(A(s') - A(s)\big)A(s)^{-1}.$$

Taking the norm, we see that the family $\{A(s)^{-1}\}$ is Cauchy, so has a limit, which is the desired inverse by continuity.

Now using charts, there are fixed Banach spaces $\mathbf{E}$ and $\mathbf{F}$ such that $\mathbf{E}(s)$ and $\mathbf{F}(s)$ are isomorphic to $\mathbf{E}$ and $\mathbf{F}$ respectively, by invertible operators varying continuously. Theorem 2.5 shows that there is a lower bound $|L(s)| \geqq c > 0$ independently of s for $0 \leqq s < r$. The operator $L(s)$ corresponds to an operator $A(s) : \mathbf{E} \to \mathbf{F}$ in the charts, and we can apply the lemma to the family $\{A(s)\}$, to yield an invertible limit operator $A(r)^{-1}$. Going back to the family $L(s)$ concludes the proof.

Example. See Theorem 3.7 of Chapter XII and subsequent comments.

Historical note. A statement equivalent to the map $\exp_{NY}$ being a bijection is given by Helgason [He 78], Chapter I, Theorem 14.6, in the finite dimensional case. Helgason does not mention the normal bundle, and does not deal with the further item that the map is a local isomorphism. Helgason saw his theorem "as a generalization of a decomposition theorem due to Mostow for a semisimple Lie group", see p. 96 and his Theorem 1.4 of Chapter VI, as well as the historical comment, p. 279, referring to Mostow [Mo 53]. The differential structure is missing in his and Mostow's statements, as when they make the decomposition only "topologically". This can be traced back to the way Theorem 14.6 of Chapter I was stated and proved. A result in this direction, in the context of semisimple Lie groups and symmetric spaces is given by Loos [Lo 69], pp. 160–161, with the differentiability property.

We see the fact that $\exp_{NY}$ is a differential isomorphism as a generalization of the tubular neighborhood theorem to the global context of Cartan–Hadamard manifolds. When I wrote to Wu to suggest that Theorem 2.5 should be valid, he wrote back: "A very good observation. It will add fuel to your frustration with geometers, however, if I tell you that it is well known, but nobody bothers to write it up. A student here in the seventies asked me and I wrote it up for her, but of course I didn't dare to publish it." Wu and others also dealt only with the finite dimensional case, where it suffices to show that $T\exp_{NY}$ has trivial kernel at each point to prove that $\exp_{NY}$ is a local differential isomorphism. Wu's argument went through with slight difference in wording in the infinite dimensional case. However, as for the ordinary Cartan–Hadamard theorem, an additional step has to be carried out. The step I have given with the family of operators showing that the limit is invertible, replaces Lemmas 3.4 and 3.5 of Chapter IX. When the submanifold Y is not just a point, the situation of Chapter IX, Lemma 3.5 is not symmetric, and it is not clear to me how to formulate a description of the adjoint of the normal exponential map TE or $T\exp_{NY}$, i.e. how to formulate the analogue of McAlpin's lemma.

For further comments, see the end of Chapter XI, §4.

The next three sections expand the fundamental material on Jacobi lifts and their convexity properties, as well as returning to the splitting of the double tangent bundle alluded to in Chapter IV, §3. Readers interested in seeing at once further considerations about Cartan–Hadamard spaces may skip the rest of this chapter, and go directly to §1, §2, §3, and §4 of the next chapter.

X, §3. MORE CONVEXITY AND COMPARISON RESULTS

I am much indebted to Karcher for explaining to me the material of the present section. Cf. [Ka 89] for more material, especially p. 182.

We continue to let X be a Riemannian manifold, and η a Jacobi lift of a curve in X. We don't need a symbol for the projection $\pi\eta$ in X, and we shall use the notation $D_*\eta$ to denote the covariant derivative taken over this projection, that is $D_{(\pi\eta)'}\eta$ in full notation.

We consider the function

$$\varphi(s) = \frac{\eta^2(s)}{s^2} = \frac{f(s)}{s^2}$$

instead of the function η^2 as in the previous section. *We let $J = (0, b)$ be an interval on which η is defined and such that $\eta(s) \neq 0$ for $s \in J$.* For simplicity we omit the subscript g from the scalar product in the tangent bundle.

We let

(1) $$h_0(s) = \frac{1}{s} \quad \text{and} \quad h(s) = \frac{\langle D_*\eta, \eta\rangle}{\langle \eta, \eta\rangle}(s).$$

We are interested in the growth properties of η, and so in the derivative φ'.

Lemma 3.1. *Let the notation be as above, with $\varphi(s) = f(s)/s^2$. Then*

$$\varphi'(s) = \frac{2}{s^2}\eta^2(s)\left(h(s) - \frac{1}{s}\right) \quad \textit{so} \quad \varphi'/\varphi(s) = 2\left(h(s) - \frac{1}{s}\right)$$

and

$$\varphi''(s) = \frac{f''(s)}{s^2} - \frac{4f'(s)}{s^3} + \frac{6f(s)}{s^4}.$$

Proof. Ordinary differentiation.

Observe that the middle term occurs with a minus sign, and that a priori, we don't know where it is positive or negative. Of course, the third term is $\geqq 0$, and Proposition 1.4 gave some information on the first term.

One of our goals is Theorem 3.4. We reach it by some calculus. We want growth properties of h, so we shall compute h'. The orthogonalization of $D_*\eta$ with respect to η will occur naturally, so we label it, and define

$$\mu_\eta = \mu = D_*\eta - \langle D_*\eta, \eta\rangle\frac{\eta}{\eta^2}.$$

Then we define the **orthogonal term**

$$\frac{\mu^2}{\eta^2} = \frac{(D_*\eta)^2}{\eta^2} - h^2 = \frac{(D_*\eta)^2}{\eta^2} - \frac{\langle D_*\eta, \eta\rangle}{(\eta^2)^2} \geqq 0.$$

Lemma 3.2. *Let $\alpha = \pi\eta$. Then*

(2) $$h' = \frac{(D_*\eta)^2}{\eta^2} + \frac{\langle D_*^2\eta, \eta\rangle}{\eta^2} - 2\frac{\langle D_*\eta, \eta\rangle^2}{(\eta^2)^2} = \frac{\mu^2}{\eta^2} + \frac{R_2(\alpha', \eta)}{\eta^2} - h^2.$$

Proof. The first equation for h' is immediate from the definition of the Levi-Civita metric derivative. The second comes from the definition of R_2 and the Jacobi equation for η, as well as the definition of the orthogonal term. This concludes the proof.

Since the orthogonal term is $\geqq 0$, we see that we get inequalities relating h' and h, depending on the sign of the term R_2, i.e. depending on the curvature. In particular:

(3) *If $R_2 \geqq 0$ then $h' \geqq -h^2$.*

In [Ka 89], Karcher views Lemma 3.2 as establishing a Ricatti equation for h. From such an equation, one obtains an inequality as follows.

Lemma 3.3. *Let h_1, h be a pair of functions on some interval, satisfying*

$$h_1' \leqq -h_1^2 \quad \textit{and} \quad h' \geqq -h^2.$$

Then

$$((h_1 - h)e^{\int (h_1+h)})' \leqq 0.$$

So if $h_1(s_1) \geqq h(s_1)$ for some s_1 in the interval, then

$$h_1(s) \geqq h(s) \quad \textit{for} \quad s \leqq s_1.$$

Proof. First note that a constant of integration added to the indefinite integral in the inequality would not affect the truth of the inequality. Next, routine differentiation yields

$$((h_1 - h)e^{\int (h_1+h)})' = (h_1' - h' + h_1^2 - h^2)e^{\int (h_1+h)}.$$

The exponential term on the right is > 0, and its coefficient is $\leqq 0$ by hypothesis, thus concluding the proof of the first inequality. It follows that the function $(h_1 - h)\exp(\int(h_1 + h))$ is semi-decreasing. If $h_1(s_1) \geqq h(s_1)$ at some point s_1, then this function is $\geqq 0$ for $s \leqq s_1$, thus concluding the proof.

Let h be the function defined in (1), and suppose $\eta(0) = 0$. Then

$$h(s) \to \infty \quad \text{as } s \to 0,$$

as one sees immediately from the Taylor expansion of the covariant derivative. We are now ready for the main result of this section.

Theorem 3.4. *Let X be a Riemannian manifold and η the Jacobi lift of a curve in X. Assume $\eta(0) = 0$ but $D_*\eta(0) \neq 0$. Suppose $R_2 \geqq 0$ (seminegative curvature). Let h be as in (1), defined on an interval $J = (0, b)$ such that $\eta(s) \neq 0$ for $s \in J$. Then*

$$\frac{1}{s} \leqq h(s) \quad \textit{for } s \in J.$$

In other words, the function $\varphi(s) = \eta^2(s)/s^2$ is semi-increasing on J.

Proof. Suppose $h_0(s_1) > h(s_1)$ for some $s_1 \in J$. Then for some $\delta > 0$,

$$h_0(s_1 + \delta) \geqq h(s_1).$$

Let $h_1(s) = 1/(s + \delta)$. Then $h_1' = -h_1^2$ and $h_1(s_1) \geqq h(s_1)$. We apply Lemma 3.3 and let $s \to 0$ (so $s \leqq s_1$). Then $h_1(s)$ is bounded, but $h(s) \to \infty$, a contradiction which proves the inequality $1/s \leqq h(s)$. By Lemma 3.1 we conclude that $\varphi(s)$ is semi-increasing. This proves the theorem.

Remark. The semi-increasing property of $\eta^2(s)/s^2$ gives a refinement of Proposition 2.6 in Chapter IX. Furthermore, Karcher mentions in [Ka 89] that the case of $R_2 \leqq 0$ (positive curvature) can also be reduced to a Ricatti equation, but unfortunately he did not provide the details, which involve a formula with the second fundamental form. This item, together with many others, would form a continuation of the present chapter in another book. (This is not a threat.) Keeping the comparison estimate to positive or negative inequalities, the result is:

If $R_2 \leqq 0$, then the function $\varphi(s) = \eta_w^2(s)/s^2$ is semi-decreasing for $s \in J$, and in particular, $\varphi(s) \leqq w^2$ (because φ can be extended to the value $\varphi(0) = w^2$ by continuity).

X, §4. SPLITTING OF THE DOUBLE TANGENT BUNDLE

We return to basics concerning covariant derivatives and sprays. In Chapter IV, Propositions 3.3 and 3.4, we gave the transformation formula in charts for the bilinear map associated with a spray, and we pointed out that a local object satisfying this transformation formula comes from a unique spray. Given such a local object, one can also define a covariant derivative directly without going through the spray, by means of the transformation formula, that is:

Proposition 4.1. *Let X be a differential manifold modeled on a Banach space $\mathbf{E}$. Suppose that we are given a covering of X by open sets corresponding to charts $U, V, \ldots$, and for each U we are given a morphism*

$$B_U \colon U \to L^2_{\mathrm{sym}}(\mathbf{E}, \mathbf{E})$$

satisfying the transformation rule of Chapter IV, Proposition 3.3. In other words, for each change of chart by a differential isomorphism

$$h \colon U \to V,$$

we have for $v, w \in \mathbf{E}$ representing tangent vectors:

$$B_V(h(x); h'(x)v, h'(x)w) = h''(x)(v, w) + h'(x)B_U(x; v, w).$$

Then there exists a unique covariant derivative D such that in a chart U for vector fields η, ξ we have

$$(D_\xi \eta)_U(x) = \eta'_U(x)\xi_U(x) - B_U(x; \xi_U(x), \eta_U(x)).$$

The proof is routine, just like Proposition 3.4 of Chapter IV.

If one takes a stronger definition of a covariant derivative to incorporate the existence of the bilinear map B_U in each chart U, then there is a bijection between sprays and covariant derivatives. In the finite dimensional case, Chapter IV, Theorem 2.4, we went through the argument going backward without using the stronger definition, to connect with the practice of many differential geometers. Others, like Klingenberg [Kl 83/95], do indeed include the additional structure of the bilinear map in the definition of a covariant derivative.

From now on, we shall always assume that a **covariant derivative** *is associated to a spray, or equivalently, that in each chart U there is a morphism B_U satisfying the stated transformation law, and such that the covariant derivative has the expression repeated in Proposition* 4.1.

We are ready to describe at greater length the double tangent bundle. At the end of Chapter IV, §3 we mentioned the possibility of splitting TTX. We now deal systematically with this splitting, which arose shortly after Ambrose, Palais, and Singer introduced sprays [APS 60]. A splitting was given by Dombrowski [Do 61], together with several other results which we mention below. See also [Wu 65]. On the other hand, Karcher explained to me another formulation of the splitting in terms of the pull back of the tangent bundle, and we shall start with this, expressed in Theorem 4.3. We go into the Dombrowski formulation afterward. Some important applications are given in Eliasson [El 67].

We start with complements to the basic discussion of Chapter III, §1 concerning the pull back of a vector bundle, and we shall apply it to the tangent bundle $\pi = \pi_X: TX \to X$. Quite generally, given a morphism

$$f: X' \to X$$

and a vector bundle $p: E \to X$ over X, the pull back f^*E (or $f^*(p)$) satisfies the universal mapping property for vector bundles over X', so that a VB morphism $E' \to E$ over f can be factored uniquely through f^*E. This is immediate from Chapter III, §1. In particular, let us take $f = \pi$. We may take the pull back:

(1)
$$\begin{array}{ccc} \pi^*TX & \xrightarrow{\pi^*(\pi)} & TX \\ \big\downarrow & & \big\downarrow \pi \\ TX & \xrightarrow[\pi]{} & X \end{array}$$

In a chart U of X, a point of TU consists of a pair (x, v) with a vector $v \in \mathbf{E}$. Then as we said in Chapter III, §1, we identify the fiber $(\pi^*TX)_{(x,u)}$ with $(TX)_x = T_xX$. In any case, the chart U on X, and the vector bundle chart $U \times \mathbf{E}$ for TX over U can be complemented with a vector bundle chart for the pull back

(1_U)
$$\begin{array}{ccc} (\pi^*TX)_U & \xleftrightarrow{\approx} & (U \times \mathbf{E}) \times \mathbf{E} \\ \big\downarrow & & \big\downarrow \\ (TX)_U & \longleftrightarrow & U \times \mathbf{E} \end{array}$$

so that a point of π^*TX in the chart is a triple

$$(x, v, z) \in (U \times \mathbf{E}) \times \mathbf{E}.$$

In addition, the double tangent bundle TTX has the vector bundle chart

(2_U)
$$\begin{array}{ccc} (TTX)_U & \longleftrightarrow & (U \times \mathbf{E}) \times \mathbf{E} \times \mathbf{E} \\ \big\downarrow & & \big\downarrow \\ (TX)_U & \longleftrightarrow & U \times \mathbf{E} \end{array}$$

We have a vector bundle morphism $S_1 = T\pi$ as in the following diagram:

(3)
$$\begin{array}{ccc} TTX & \xrightarrow{S_1 = T\pi} & TX \\ \pi_{TX} \big\downarrow & & \big\downarrow \pi \\ TX & \xrightarrow[\pi]{} & X \end{array}$$

which in a chart gives

(3_U)
$$\begin{array}{ccc} (U \times \mathbf{E}) \times \mathbf{E} \times \mathbf{E} & \xrightarrow{(T\pi)_U} & U \times E \\ \big\downarrow & & \big\downarrow \\ U \times E & \longrightarrow & U \end{array} \qquad \text{with} \quad (T\pi)_U(x, v, z, w) = (x, z).$$

Indeed, the projection $\mathrm{pr}_1: U \times \mathbf{E} \to U$ is linear, so its derivative at every point is equal to pr_1 itself. The pair (z, w) represents a tangent vector at (x, v) in the tangent vector space $\mathbf{E} \times \mathbf{E}$.

The above diagram then gives rise to the factoring map

$$\kappa_1 \colon TTX \to \pi^* TX$$

of the double tangent bundle into the pull back, so we get the diagram of VB morphisms

$$\text{(4)}\qquad \begin{array}{ccccc} TTX & \xrightarrow{\kappa_1} & \pi^*TX & \xrightarrow{\pi^*(\pi)} & TX \\ {\scriptstyle \pi_{TX}}\downarrow & & \downarrow & & \downarrow{\scriptstyle \pi} \\ TX & \xrightarrow[\mathrm{id}]{} & TX & \xrightarrow[\pi]{} & X \end{array} \qquad \text{with} \quad T\pi = \pi^*(\pi) \circ \kappa_1.$$

In the chart, these two squares become

$$(4_U)\qquad \begin{array}{ccccc} (U \times \mathbf{E}) \times \mathbf{E} \times \mathbf{E} & \xrightarrow{\kappa_{1,U}} & (U \times \mathbf{E}) \times \mathbf{E} & \xrightarrow{\pi^*(\pi)_U} & U \times \mathbf{E} \\ {\scriptstyle \mathrm{pr}_{12}}\downarrow & & \downarrow{\scriptstyle \mathrm{pr}_{12}} & & \downarrow{\scriptstyle \mathrm{pr}_1} \\ U \times \mathbf{E} & \xrightarrow[\mathrm{id}]{} & U \times \mathbf{E} & \xrightarrow[\mathrm{pr}_1]{} & U \end{array}$$

with

$$\kappa_{1,U}(x, v, z, w) = (x, v, z) \qquad \text{and} \qquad \pi^*(\pi)_U(x, v, z) = (x, z).$$

Thus (4_U) gives the factorization of (3_U) in the vector bundle charts.

So far, these diagrams concern the tangent and double tangent bundle without any further structure. We now suppose given a spray or covariant derivative, so that we have the bilinear map B_U in a chart U.

Lemma 4.2. *Given a spray or covariant derivative on X, there is a unique vector bundle morphism over TX,*

$$\kappa_2 \colon TTX \to \pi^* TX$$

such that over a chart U, we have

$$(5_U)\qquad \kappa_{2,U}(x, v, z, w) = \big(x, v, w - B_U(x; v, z)\big).$$

Proof. Let $h \colon U \to V$ be a change of charts, i.e. a differential isomorphism. In Chapter IV, §3 we gave the change of chart (2_U) of TTX. Let $H = (h, h')$. Then the change of chart for $(TTX)_U$ is given by the map

$$(U \times \mathbf{E}) \times \mathbf{E} \times \mathbf{E} \xrightarrow{(H,H')} (V \times \mathbf{E}) \times \mathbf{E} \times \mathbf{E}$$

such that

$$(H, H')(x, v, z, w) = \big(h(x), h'(x)z, h''(x)(v, z) + h'(x)w\big).$$

Then

$$\kappa_{2,V} \circ (H, H')(x, v, z, w) = \big(h(x), h'(x)v, h'(x)w\big)$$

because the term $h''(x)(v, z)$ cancels in the last coordinate on the right. This proves the lemma.

The next theorem puts together both maps κ_1 and κ_2.

Theorem 4.3 (Tensorial Splitting Theorem). *Given a spray, or covariant derivative on a differential manifold X, the map*

$$\kappa = (\kappa_1, \kappa_2) \colon TTX \to \pi^*TX \oplus_{TX} \pi^*TX$$

is a vector bundle isomorphism over TX. In the chart

$$(TTX)_U = (U \times \mathbf{E}) \times \mathbf{E} \times \mathbf{E}$$

this map is given by

$$(6_U)\qquad \kappa_U(x, v, z, w) = \big(x, v, z, w - B_U(x; v, z)\big).$$

Proof. With the notation h, H, (H, H') as in Lemma 4.2, we conclude that

$$\kappa_V \circ (H, H')(x, v, z, w) = \big(h(x), h'(x)v, h'(x)z, h'(x)w\big),$$

so the family $\{\kappa_U\}$ defines a *VB* morphism over *TX*. The expression of the map in a chart shows that over U it is a *VB* isomorphism, which concludes the proof. Note that the map κ_U is represented by a 2×2 matrix acting on the last two coordinates, and having the identity on the diagonal.

Of course, one may phrase a variation of Theorem 4.3 by using the mappings going all the way to TX instead of the pull back π^*TX. More precisely, let us define

$$S_i \colon TTX \to TX \quad (i = 1, 2) \qquad \text{by} \qquad S_i = \pi^*(\pi) \circ \kappa_i.$$

Then: $S_1 = T\pi$, so in a chart U, $S_{1,U}(x, v, z, w) = (x, z)$;
S_2 is the unique VB morphism such that in the chart,

$$S_{2,U}(x, v, z, w) = \big(x, w - B_U(x; v, z)\big).$$

Thus we obtain a morphism $S = (S_1, S_2)$ of vector bundles

(7)
$$\begin{array}{ccc} TTX & \xrightarrow{S = (S_1, S_2)} & TX \oplus TX \\ {\scriptstyle \pi_{TX}} \downarrow & & \downarrow \\ TX & \xrightarrow[\pi]{} & X \end{array}$$

whose local representation is actually given by a similar formula as (6_U), namely

(7_U)
$$S_U(x, v, z, w) = \big(x, z, w - B_U(x; v, z)\big),$$

which just drops the v coordinate in the term on the right side.

One calls κ_1 or S_1 the **horizontal component**, and κ_2 or S_2 the **vertical component**. The maps $S_1 = T\pi$ and S_2 are in fact the maps used by Dombrowski [Do 61]. We now go into his formulation of the splitting.

We need to make some remarks about the covariant derivative D acting on vector fields. Let ζ, ξ denote vector fields over X. We have defined $D_\xi \zeta$, but it is also convenient to use D without a subscript. Let $VF(X)$ denote the $\mathbf{R}$-vector space of vector fields over X. Then we let

$$D\colon VF(X) \to \operatorname{Hom}\big(VF(X), VF(X)\big)$$

be the linear map such that $D\zeta \in \operatorname{Hom}\big(VF(X), VF(X)\big)$ and $(D\zeta)\xi = D_\xi \zeta$. The next lemma gives Dombrowski's direct description of S_2 in terms of the covariant derivative.

Lemma 4.4. *Let X be a manifold with a spray or covariant derivative D. There exists a unique vector bundle morphism (over π)*

$$K\colon TTX \to TX$$

such that for all vector fields ξ, ζ on X, we have

(8)
$$D_\xi \zeta = K \circ T\zeta \circ \xi, \quad \textit{in other words,} \quad D = K \circ T$$

as operators on vector fields, so the following diagram is commutative:

$$\begin{array}{ccc} TX & \xrightarrow{T\xi} & TTX \\ {\scriptstyle \xi} \uparrow & & \downarrow {\scriptstyle K} \\ X & \xrightarrow[D_\xi \xi]{} & TX \end{array}$$

In fact, $K = S_2$.

Proof. In a chart U, we let the local representation

$$K_{U,(x,v)}\colon \mathbf{E} \times \mathbf{E} \to \mathbf{E}$$

be given by

(8_U)
$$K_{U,(x,v)}(z, w) = w - B_U(x; v, z),$$

so $K = S_2$ satisfies the requirements of the lemma.

Remark. Dombrowski gives invariant definitions, but goes finite dimensional in his formulas, p. 76, and the footnote: "Here and in the sequel we apply a non-orthodox usage of summation signs, in order to make formulas more concise." As far as making formulas more concise, the use of the bilinear map B_U and the chart notation rather than local coordinates are more effective. And they happen to make the statements valid in infinite dimensions.

We have accumulated three maps

$$TTX \to TX,$$

namely the maps $S_2 = K$, $S_1 = T\pi$, and π_{TX}. So far, we have put two of them together. We shall now put all three together. As observed in Chapter IV, §3, we view TTX as a fiber bundle over X, in addition to being a vector bundle over TX.

Theorem 4.5 (Dombrowski Splitting Theorem). *Let X be a manifold with a spray or a covariant derivative. Then the map*

$$(\pi_{TX}, S_1, S_2)\colon TTX \to TX \oplus TX \oplus TX$$

is an isomorphism of fiber bundles over X.

Proof. The map is well defined, and the previous chart formulas show that it is both a bijection and a local differential isomorphism. We let readers check this out in the charts to conclude the proof.

As Dombrowski remarks, if X is a Riemannian manifold, then one can use the splitting theorem to define a natural Riemannian metric on TTX. Indeed, let g as usual denote the Riemannian metric on X. Let $v \in TX$, so $v \in T_x X$ with $x = \pi v$. Let $Z, W \in T_v TX$. We define the **splitting metric** $\tilde{g} = g_{TTX}$ by the formula

$$\langle Z, W \rangle_{\tilde{g}} = \tilde{g}(Z, W) = \langle S_1 Z, S_1 W \rangle_g + \langle S_2 Z, S_2 W \rangle_g.$$

This metric was first defined in terms of local coordinates by Sasaki [Sas 58], but the above formula was given by Dombrowski, who also twists the metric by $\sqrt{-1}$, and defines an almost complex structure, thus obtaining further results for which we refer to his paper. The Dombrowski splitting was used subsequently by Eliasson [El 67] to define connections in Banach manifolds, with applications to the manifold of mappings between two manifolds.

Ambrose–Palais–Singer [APS 60] showed that there is one and only one torsionless covariant derivative (connection) whose geodesics are the Jacobi lifts, and that any other connection with the same property differs from this one by a torsion tensor. Eliasson used the splitting to define such connections [El 67], Theorem 3.2, p. 178. I regard such matters as topics for another book.

X, §5. TENSORIAL DERIVATIVE OF A CURVE IN *TX* AND OF THE EXPONENTIAL MAP

I am further indebted to Karcher for this section, partly based on his paper [Ka 77], p. 536. The paper contains more interesting material, especially Appendix C in connection with present considerations.

We continue to consider a manifold X with a spray, or equivalently with a covariant derivative. When we first introduced the covariant derivative, we used vector fields as in §4, and then discussed the analogous notion for curves. We follow the same pattern here, and we deal with curves in the present section.

So let ζ be a curve in TX. In terms of the vector bundle morphism S defined in §3 (7), we may give the splitting formula for the derived curve ζ' in TX, that is

(1)
$$S\zeta' = \big(T(\pi)\zeta', D_{(\pi\zeta)'}\zeta\big) = \big((\pi\zeta)', D_{(\pi\zeta)'}\zeta\big),$$
$$\text{so}\quad S\zeta' = (\beta', D_{\beta'}\zeta),$$

letting $\beta = \pi\zeta$, or equivalently, ζ is a lift of β. We call $S\zeta'$ the **tensorial derivative of** ζ. It has values in $TX \oplus TX$ (as vector bundle over X).

We recall from Chapter VIII that in a chart U,

(2)
$$(D_{\beta'}\zeta)_U = \zeta'_U - B_U(\beta_U; \beta'_U, \zeta_U).$$

As remarked following Theorem 3.1 of Chapter VIII, the local representation ζ_U of a curve in $TU = U \times \mathbf{E}$ is taken to be the map on the second component, i.e. $\zeta_U\colon J \to \mathbf{E}$, and $\zeta'_U(t)$ is the **ordinary derivative** with values $\zeta'_U(t) \in \mathbf{E}$ also. Thus $\beta'_U(t)$, $\zeta_U(t)$, $\zeta'_U(t)$ are "vectors", giving the local representation of these curves.

Let U be a chart in X, so that $U \times \mathbf{E}$ is a chart for TX over U. Let $\mathscr{D}_U \subset U \times \mathbf{E}$ be the domain of the exponential, which thus has the local representation
$$\exp_U\colon \mathscr{D}_U \to U.$$

We think in terms of pairs, so by abuse of notation, we sometimes write
$$\exp_U\colon U \times \mathbf{E} \to U,$$

with the understanding the $\exp_U$ is only defined on $\mathscr{D}_U$. The tangent map
$$T\exp\colon TTX \to TX$$

then has a local representation at a point $(x, v) \in U \times \mathbf{E}$ given by the linear map
$$\exp'_{U\times\mathbf{E}}(x, v)\colon \mathbf{E} \times \mathbf{E} \to \mathbf{E} \times \mathbf{E}.$$

The following remark is merely a translation of Theorem 4.3 in terms of curves.

Lemma 5.1. *Let X be a manifold with a spray, or equivalently a covariant derivative. Let β be a curve in X, and let ζ be a lift of β in TX. Let*
$$\varphi(t) = \exp_{\beta(t)}\zeta(t),$$

so φ is a curve in X. Then in a chart U, $\varphi'(t)$ has the representation

(3_U)
$$\varphi'_U(t) = \exp'_U(\beta_U(t), \zeta_U(t))\big(\beta'_U(t), (D_{\beta'}\zeta)_U(t)\big),$$

or suppressing t,

(4_U)
$$\varphi'_U = \exp'_U(\beta_U, \zeta_U)\big(\beta'_U, (D_{\beta'}\zeta)_U\big).$$

Proof. This is immediate from Theorem 4.3, the local expression (2) for the covariant derivative, and formula (1).

By abuse of notation, one sometimes omits the subscript U, and one writes

(4?)
$$\varphi' = T\exp(\zeta)(\beta', D_{\beta'}\zeta).$$

This way of writing exhibits an identification of TTX with $\pi^*TX \oplus \pi^*TX$ as in Theorem 3.3, and a further identification of the fibers of π^*TX with the fibers of TX itself. These identifications are not as dangerous as one might think, for a variety of reasons. First, they are made routinely in the calculus of several variables in euclidean space or in Banach spaces, for that matter. Second, one can give an invariant formulation which is strictly correct at the cost of additional notation, as follows.

From our spray or covariant derivative, we had the vector bundle isomorphism

$$TTX \xrightarrow{\kappa} \pi^*TX \oplus \pi^*TX \qquad \text{over } TX.$$

The tangent map $T\exp$ is a VB morphism

$$T\exp\colon TTX \to TX \qquad \text{over} \qquad \exp\colon TX \to X.$$

We now define the **tensorial tangent map** or **tensorial derivative**

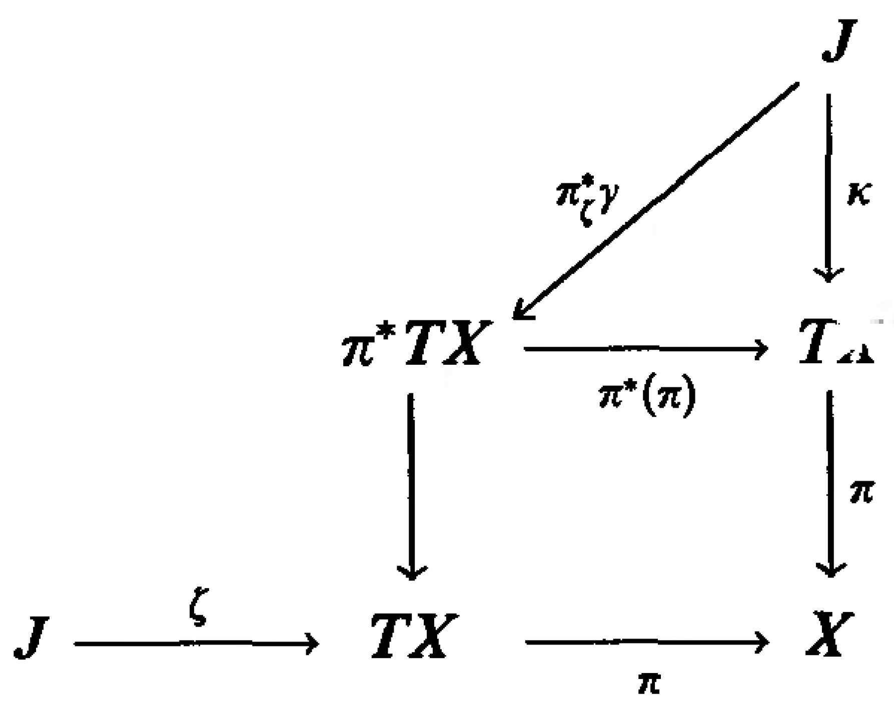
$$\mathbf{T}\exp\colon \pi^*TX \oplus \pi^*TX \to \pi^*TX$$

to be the unique *VB* morphism such that

$$\pi^*(\pi) \circ \mathbf{T}\exp \circ \kappa = T\exp.$$

We shall complement this map by another one going all the way to TX in Lemma 5.2.

Let $\zeta\colon J \to TX$ be a curve in TX, lying above its projection $\beta = \pi\zeta$ in X. Let $\gamma\colon J \to TX$ be another curve in TX, defined on the same interval, such that $\pi\gamma = \pi\zeta$. Then γ has a pull back $\pi_\zeta^*\gamma$ to π^*TX, depending on ζ, and making the following diagram commutative:



Then we have the valid formula

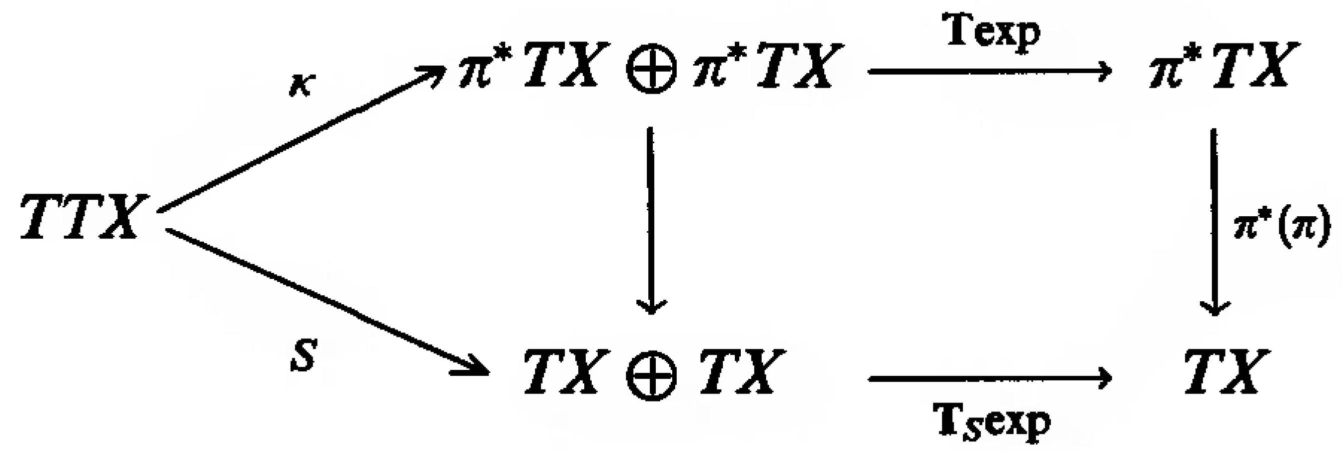
(5) $$\pi_\zeta^*\varphi' = \mathbf{T}\exp(\zeta)(\pi_\zeta^*\beta', \pi_\zeta^* D_{\beta'}\zeta).$$

On the other hand, instead of pulling back to π^*TX and using the map $\kappa = (\kappa_1, \kappa_2)$ one may stay on TX and use the map $S = (S_1, S_2)$. Then we may define the ***S*-tensorial derivative** $\mathbf{T}_S\exp$ by means of the following lemma:

Lemma 5.2. *There exists a unique vector bundle morphism over X,*

$$\mathbf{T}_S\exp\colon TX \oplus TX \to TX$$

such that the following diagram commutes:



The two vertical maps are vector bundle morphisms, the top vector bundle being over TX and the bottom one over X. The composite map is

$$T\exp\colon TTX \to TX,$$

so both $\mathbf{T}\exp$ *and* $\mathbf{T}_S\exp$ *represent* $T\exp$ *under the splitting maps.*

Proof. Routine verification that everything makes sense.

With these definitions, we are in a position to formulate (4?) without any identifications. Hypotheses being as in Lemma 5.1, with ζ a curve in TX and $\beta = \pi\zeta$, we have

(6) $$\varphi' = \mathbf{T}_S\exp(\zeta)(\beta', D_{\beta'}\zeta).$$

Thus the only thing needed to make (4?) meaningful (and valid) was to replace the ordinary tangent map by the tensorial map, using the splitting map S. Things could be worse.

Variations and Jacobi lifts

We now stick an extra parameter s to make up a variation. We let

$$\alpha(s) = \exp_{\beta(0)}(s\zeta(0)).$$

We define the **(β, ζ)-variation** of α by letting

(7) $$\alpha_t(s) = \sigma(s, t) = \exp(s\zeta(t)) = \exp_{\beta(t)}(s\zeta(t)).$$

Note that for each real s (such that $s\zeta(t)$ is in the domain of the exponential), the tangent vector $s\zeta(t)$ lies in $T_{\beta(t)}X$, so $s\zeta = \zeta_s$ is also a lift of β. For fixed t, the curve α_t is a usual geodesic, and we have

(8) $$\alpha_t'(s) = T\exp_{\beta(t)}(s\zeta s(t))\zeta(t),$$

with initial conditions

(9) $$\alpha_t(0) = \beta(t) \quad \text{and} \quad \alpha_t'(0) = \zeta(t).$$

Indeed, let us put

$$Z(s, t) = s\zeta(t) = Z_t(s).$$

Then Z_t is a curve in the tangent space $T_{\pi\zeta(t)}X = T_{\beta(t)}X$. Then

$$\partial_1 Z(s, t) = Z_t'(s)$$

is now taken to be the ordinary derivative of a curve in a Banach space, so $Z_t'(s)$ is also a vector, element of $T_{\beta(t)}X$.

Next we define the usual Jacobi lift

$$\eta_t(s) = \partial_2\sigma(s, t).$$

We let further

$$\varphi_s(t) = \sigma(s, t) = \alpha_t(s),$$

so φ_s is a curve in X, and φ_s' is a curve in TX. Then by definition,

$$\varphi_s'(t) = \eta_t(s).$$

We have the initial conditions

(10a) $$\eta_t(0) = \beta'(t) = \varphi_0'(t).$$

(10b) $$(D_{\alpha_t'}\eta_t)(0) = D_1\partial_2\sigma(0, t) = D_2\partial_1\sigma(0, t) = (D_{\beta'}\zeta)(t).$$

In general,

(11a) $$\eta_t(s) = \mathbf{T}_S \exp(s\zeta(t))(S\zeta'(t)) = \mathbf{T}_S \exp(s\zeta(t))(\beta'(t), D_{\beta'(t)}(s\zeta)),$$

(11b) $$(D_{\alpha_t'}\eta_t)(s) = D_1\partial_2\sigma(s, t) = D_2\partial_1\sigma(s, t).$$

X, §6. THE FLOW AND THE TENSORIAL DERIVATIVE

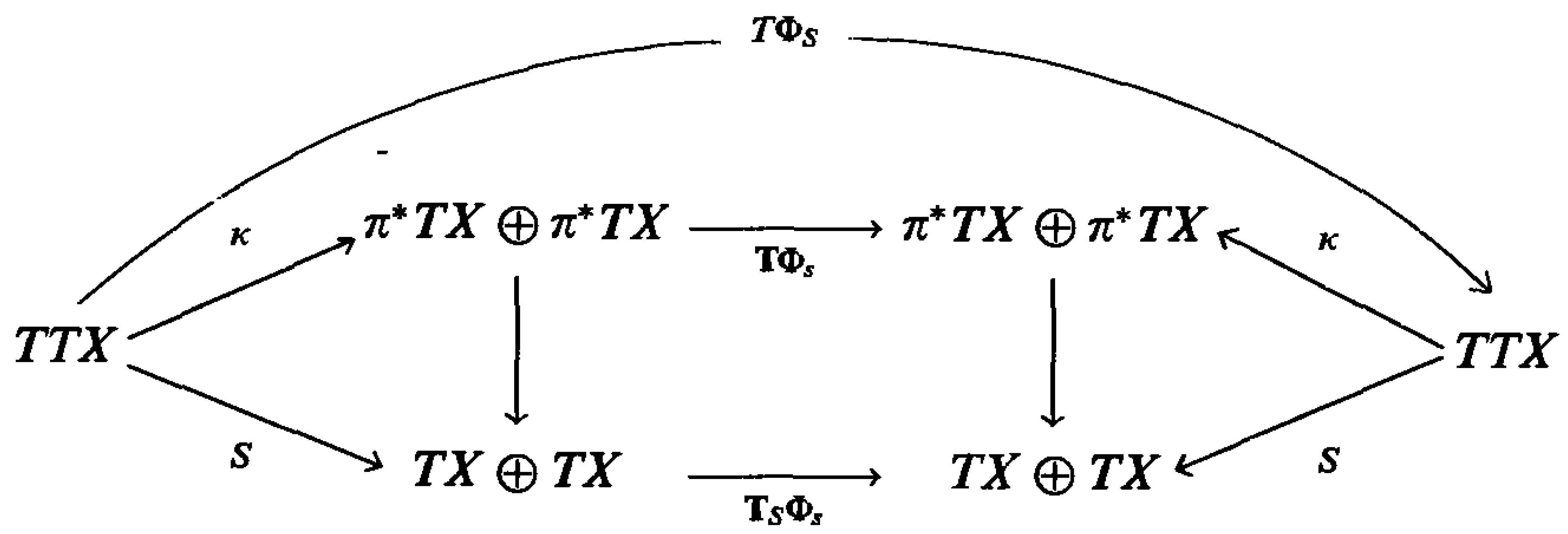
Dombrowski actually defines his map K in terms of the flow, but for the exposition of this section I again benefited from Karcher's instructions. We may complement (10) and (11) in §5 from another point of view. Let Φ be the flow of the spray on TX. Thus Φ is the global geodesic flow, and is a function of two variables. For an element $v \in TX$, so $v \in T_xX$ for some x, we have by definition

$$\Phi(s, v) = \Phi_s(v) = \frac{d}{ds}\exp(sv) = T\exp(sv)v.$$

Then Φ_s is a differential isomorphism

$$\Phi_s\colon TX \to TX,$$

The tangent map of the flow then fits in a commutative diagram which we call the **tensorial flow diagram**, with maps $\mathbf{T}\Phi_s$ and $\mathbf{T}_S\Phi_s$ as follows.



With the notation of §5, $t \mapsto \Phi_s(\zeta(t))$ is a curve in TX. Directly from the definitions, we have

(1) $$\alpha_t'(s) = \Phi_s(\zeta(t)).$$

We may summarize some of the tabulations of §5 in terms of the flow.

Theorem 6.1. *Let ζ be a curve in TX, let*

$$\sigma(s, t) = \exp(s\zeta(t)) = \alpha_t(s) \quad \textit{and} \quad \eta_t(s) = \partial_2\sigma(s, t).$$

Then the tensorial derivative of $\Phi_s \circ \zeta$ is given by

$$S(\Phi_s \circ \zeta)'(t) = (\eta_t, D_{\alpha_t'}\eta_t)(s).$$

Remark on notation. If ζ is a lift in TX of a curve β in X, we have used systematically the notation $D_{\beta'}\zeta$ for the covariant derivative. We could also have written $D_{(\pi\zeta)'}\zeta$. We used this notation to accompany the notation when one takes the covariant derivative of vector fields. However, in the context of curves, the index of D is completely determined by the curve ζ in the tangent bundle. Therefore we also write $D_*\zeta$, the $*$ being forced, namely $(\pi\zeta)'$. Then the formula of Theorem 6.1 reads

$$S(\Phi_s \circ \zeta)'(t) = (\eta_t, D_*\eta_t)(s),$$

which saves some double indices.

At Karcher's suggestion, I include an application to a classical Liouville theorem, formulated in contemporary language. We return to the canonical 2-form on the tangent bundle of the Riemannian manifold X, as defined in Chapter V, Proposition 7.2. Of course, without the Riemannian structure, the canonical 2-form is on the cotangent bundle, but with the Riemannian structure, it is transferred to the tangent bundle as in Chapter VII, §7, formula (1). The next result gives the representation of the canonical 2-form in terms of the splitting coordinates. For this purpose, if (A_1, B_1) and (A_2, B_2) are pairs of vectors in T_xX, the formula

$$\Omega_S((A_1, B_1), (A_2, B_2)) = \langle A_1, B_2\rangle_g - \langle A_2, B_1\rangle_g$$

defines a 2-form Ω_S on $TX \oplus TX$.

Proposition 6.2. *Let X be a Riemannian manifold, and let Ω be the canonical 2-form on the tangent bundle. Let $v \in TX$, $Z, W \in T_vTX$. Write*

$$SZ = (A_1, B_1) \quad \textit{and} \quad SW = (A_2, B_2).$$

Then the canonical 2-form can be expressed in the form

$$\Omega(Z, W) = \Omega_S(SZ, SW) = \langle A_1, B_2\rangle_g - \langle A_2, B_1\rangle_g.$$

Proof. This is a routine verification, which nevertheless has to be taken seriously. We use a chart. Write $Z = (z_1, z_2)$ and $W = (w_1, w_2)$ in the chart, i.e. in $\mathbf{E} \times \mathbf{E}$. Put together Chapter VII, §7, formula (1) for the canonical 2-form on the tangent bundle, and Theorem 4.2 of Chapter VIII, formula **MS 1**, giving the chart expression for the bilinear map B_U, depending on the metric. Keep cool, calm, and collected; there will be cancellations, due to the symmetry

$$\langle g'(x)u \cdot w, v\rangle = \langle g'(x)u \cdot v, w\rangle$$

noted in Chapter VII, §7; you will use $\frac{1}{2} + \frac{1}{2} = 1$; and the formula of Proposition 6.2 will drop out to conclude the proof.

In light of Proposition 6.2, we call Ω_S the **splitting of the canonical 2-form.**

Theorem 6.3. *Let X be a Riemannian manifold. Let Ω be the canonical 2-form on TX. Let $v \in TX$ (so $v \in T_xX$ for some x), and let $Z, W \in T_vTX$. Let Φ be the flow of the spray on TX. Let*

$$\psi(s) = \Omega(T\Phi_s(v)Z, T\Phi_s(v)W).$$

Then ψ is constant. In other words, Ω is invariant under the flow.

Proof. We use Proposition 6.2. Let η_1, η_2 be the Jacobi lifts of the curve $s \mapsto \alpha_v(s) = \exp(sv)$ with initial conditions

$$\eta_i(0) = A_i \quad \text{and} \quad D_v\eta_i(0) = B_i.$$

Then using Theorem 6.1 and Proposition 6.2, we get

$$\psi = \langle \eta_1, D_*\eta_2\rangle - \langle \eta_2, D_*\eta_1\rangle.$$

Hence using the basic property of the Levi-Civita metric derivative,

$$\begin{aligned}\psi' &= \langle \eta_1, D_*^2\eta_2\rangle - \langle \eta_2, D_*^2\eta_1\rangle \\ &= \langle \eta_1, R(v, \eta_2)v\rangle - \langle \eta_2, R(v, \eta_1)v\rangle \quad \text{by the Jacobi equation} \\ &= 0\end{aligned}$$

by one of the fundamental identities of the Riemann tensor, Chapter IX, §1, **RIEM 4**. This concludes the proof.

Remark. In Chapter XIII we shall investigate Killing fields, whose flow preserves the metric in the Riemannian case. The situation is similar here with the canonical 2-form, although the spray is usually not called a Killing field.

CHAPTER XI

Curvature and the Variation Formula

This chapter is a direct continuation of Chapter IX, but in a new context, the variation formula. Given a family of curve $\{\alpha_t\}$, their lengths $L(\alpha_t)$ defines a function, and we are interested in the singular points of this function on the space of curves especially the relative minima and the second derivative test. We do not formalize the infinite dimensional space of curves but work simply with families. We shall see that the Riemann tensor plays an essential role in the expression for the second derivative, which allows us to go futher than we did in Chapter IX, and especially in proving the converse of Theorem 3.6, for which we have to deal with positive curvature. The variation formula will allow us to estimate growths of Jacobi lifts more generally than in Chapter IX.

XI, §1. THE INDEX FORM, VARIATIONS, AND THE SECOND VARIATION FORMULA

We let (X, g) be a pseudo Riemannian manifold, with the corresponding covariant derivative D. As a matter of notation, if w is vector in a tangent space, then we write $w^2 = \langle w, w\rangle_g$. If $w^2 \geqq 0$, then we define

$$\|w\| = \langle w, w\rangle_g^{1/2}.$$

We begin by a general discussion concerning the Jacobi expression defining Jacobi lifts. Let $\alpha\colon [a, b] \to X$ be a geodesic. Let $\eta \in \mathrm{Lift}(\alpha)$. We are interested in the expression

$$D_{\alpha'}^2\eta - R(\alpha', \eta)\alpha'$$

and its square. By definition, η is a Jacobi lift if and only if this expression is equal to 0, and in the Riemannian case, it is equal to 0 if and only if its square is equal to 0. We shall also deal with a subspace of $\mathrm{Lift}(\alpha)$, namely:

$\mathrm{Lift}_0(\alpha)$ = vector space of lifts η of α such that

$$\eta(a) = 0 \quad \text{and} \quad \eta(b) = 0.$$

For $\eta, \gamma \in \mathrm{Lift}(\alpha)$ we define the **index**

$$I(\eta, \gamma) = \int_a^b [\langle D_{\alpha'}\eta, D_{\alpha'}\gamma\rangle_g + R(\alpha', \eta, \alpha', \gamma)](s)\, ds.$$

Then I is a symmetric bilinear form on $\mathrm{Lift}(\alpha)$, whose corresponding quadratic form is

$$I(\eta, \eta) = \int_a^b [(D_{\alpha'}\eta)^2 + R_2(\alpha', \eta)](s)\, ds.$$

Similarly we define

$\mathrm{Jac}_0(\alpha)$ = subspace of Jacobi lifts of α lying in $\mathrm{Lift}_0(\alpha)$, that is, vanishing at the end points.

Proposition 1.1. *Let $\alpha\colon [a, b] \to X$ be a geodesic. The index form I on* $\mathrm{Lift}(\alpha)$ *also has the expression*

$$I(\eta, \gamma) = -\int_a^b [\langle D_{\alpha'}^2\eta, \gamma\rangle_g - R(\alpha', \eta, \alpha', \gamma)](s)\, ds$$
$$+ \langle D_{\alpha'}\eta, \gamma\rangle_g(b) - \langle D_{\alpha'}\eta, \gamma\rangle_g(a).$$

In particular, if η is a Jacobi lift, then

$$I(\eta, \gamma) = \langle D_{\alpha'}\eta, \gamma\rangle_g(b) - \langle D_{\alpha'}\eta, \gamma\rangle_g(a);$$

and if in addition $\gamma \in \mathrm{Lift}_0(\gamma)$, then $I(\eta, \gamma) = 0$.

Proof. From the defining property of the metric derivative, we know that

$$\partial\langle D_{\alpha'}\eta, \gamma\rangle_g = \langle D_{\alpha'}^2\eta, \gamma\rangle_g + \langle D_{\alpha'}\eta, D_{\alpha'}\gamma\rangle_g.$$

Then the first formula is clear. If in addition η is a Jacobi lift, then the expression under the integral is 0 by definition, so the second formula follows; and if $\gamma \in \mathrm{Lift}_0(\alpha)$ then the expressions belonging to the end points are equal to 0, so the proposition is proved.

Theorem 1.2. *Let $\eta \in \mathrm{Lift}(\alpha)$. Then $I(\eta, \gamma) = 0$ for all $\gamma \in \mathrm{Lift}_0(\alpha)$ if and only if*

$$(D^2_{\alpha'}\eta - R(\alpha', \eta)\alpha')^2 = 0.$$

In the Riemannian case, this happens if and only if η is a Jacobi lift.

Proof. If η is a Jacobi lift, then by definition

$$D^2_{\alpha'}\eta = R(\alpha', \eta)\alpha',$$

so $I(\eta, \gamma) = 0$ for all $\gamma \in \mathrm{Lift}_0(\alpha)$. Conversely, assume this is the case. Let φ be a C^∞ function on $[a, b]$ such that $\varphi(a) = \varphi(b) = 0$. Let

$$\gamma_1 = D^2_{\alpha'}\eta - R(\alpha', \eta)\alpha' \quad \text{and} \quad \gamma = \varphi\gamma_1.$$

Then $\gamma \in \mathrm{Lift}_0(\alpha)$ and by Proposition 1.1,

$$0 = I(\eta, \gamma) = \int_a^b \varphi(s)\gamma_1(s)^2\, ds.$$

This being true for all φ as above, it follows that $\gamma_1^2 = 0$, whence the theorem follows.

The previous discussion belongs to the general theory of the Jacobi differential equation. Previously, we developed this theory to get information about the differential of the exponential map. The differential equation has another side to it, to which we now turn. We shall be interested in two functions of paths $\alpha: [a, b] \to X$:

The length function

$$L_a^b(\alpha) = L(\alpha) = \int_a^b \|\alpha'(s)\|_g\, ds \quad \text{whenever } \alpha'(s)^2 \geqq 2.$$

The energy function

$$E_a^b(\alpha) = E(\alpha) = \int_a^b \alpha'(s)^2\, ds.$$

Note that the length does not depend on the parametrization, but the energy does. We are interested in minimizing those functions.

In calculus, one applies the second derivative test at a critical point of a function, that is a point where the first derivative is 0. The second derivative then has geometric meaning. One wants to do a similar thing on function spaces, or the space of paths. Ultimately, one can define a manifold structure on this space, but there is a simple device which at first avoids defining such a structure for some specific computations. We are specifically interested here in the example of the second derivatives

$$\frac{d^2}{dt^2}L(\alpha_t) \quad \text{and} \quad \frac{d^2}{dt^2}E(\alpha_t)$$

at $t = 0$. To compute these derivatives, we don't need to give a differential structure to the path space, we need only be able to differentiate under the integral sign in the usual way. The computation of these derivatives is called the **second variation formula**, and the end result is as follows, for the variation of a geodesic. Observe how the index form enters into the result.

Theorem 1.3. *Let $\alpha: [a, b] \to X$ be a geodesic parametrized by arc length, that is $\alpha'(s)^2 = 1$ for all s. Let σ be a variation of α, so that $\alpha_t(s) = \sigma(s, t)$. Define*

$$\eta(s) = \partial_2\sigma(s, 0) \quad \textit{and} \quad \nu(s) = D_{\alpha'}\eta(s) - \langle D_{\alpha'}\eta(s), \alpha'(s)\rangle_g \alpha'(s),$$

so $\nu(s)$ is the normal projection of $D_{\alpha'}\eta(s)$ with respect to the unit vector $\alpha'(s)$. Also define a second component along $\alpha'(s)$, namely

$$\gamma_2(s) = \langle D_2\partial_2\sigma, \partial_1\sigma\rangle_g(s, 0) = \langle D_2\partial_2\sigma(s, 0), \alpha'(s)\rangle_g.$$

Let $R_2(v, w) = R(v, w, v, w)$ be the canonical 2-tensor. Then

$$\begin{aligned}\frac{d^2}{dt^2}E(\alpha_t)\Big|_{t=0} &= I(\eta, \eta) + \gamma_2(b) - \gamma_2(a)\\ &= \int_a^b [(D_{\alpha'}\eta)^2 + R_2(\alpha', \eta)](s)\, ds + \gamma_2(b) - \gamma_2(a).\end{aligned}$$

As for the length, assuming the variation satisfies $\alpha_t'(s)^2 \geqq 0$ for all t, s:

$$\frac{d^2}{dt^2}L(\alpha_t)\Big|_{t=0} = \int_a^b [\nu^2 + R_2(\alpha', \eta)](s)\, ds + \gamma_2(b) - \gamma_2(a),$$

so this is the same expression as for the energy, except that $D_{\alpha'}\eta(s)$ is replaced by the normal projection $\nu(s)$.

If the curves $t \mapsto \sigma(a, t)$ and $t \mapsto \sigma(b, t)$ are geodesics, then

$$\gamma_2(b) = \gamma_2(a) = 0,$$

so the terms involving the end points are equal to 0.

Remark 1. The last assertion is immediate, since for any geodesic γ, we have $D_{\gamma'}\gamma' = 0$.

Remark 2. An example comes from Theorem 4.6 of Chapter IX, where α_t is the geodesic between points $\beta_1(t)$, $\beta_2(t)$, and β_1, β_2 are also geodesics. Then indeed, the constant of integration is equal to 0, and

$$\frac{d^2}{dt^2}L(\alpha_t)\bigg|_{t=0} = \int_a^b \left(\nu^2 + R_2(\alpha', \eta)\right).$$

For the proof of Theorem 1.3, we ned a lemma giving some expressions for the square of the derivative of a family of geodesics. This lemma is independent of the integrals which have just been considered, and we state it in full. It is similar but more elaborate than the lemmas of Chapter IX, §4.

Lemma 1.4. *Let (X, g) be a pseudo-Riemannian manifold. Let α be a geodesic (not necessarily parametrized by arc length), and let $\sigma = \sigma(s, t)$ be a variation of α (not necessarily by geodesics), so $\alpha = \alpha_0$, and $\alpha_t(s) = \sigma(s, t)$. Put*

$$e(s, t) = \langle\partial_1\sigma, \partial_1\sigma\rangle_g(s, t) = \alpha_t'(s)^2.$$

Define $\eta(s) = \partial_2\sigma(s, 0)$ and

$$\gamma_2(s) = \langle D_2\partial_2\sigma, \partial_1\sigma\rangle_g(s, 0) = \langle D_2\partial_2\sigma(s, 0), \alpha'(s)\rangle_g.$$

Then

(1) $$\partial_2 e(s, 0) = 2\langle D_{\alpha'}\eta(s), \alpha'(s)\rangle_g,$$

(2) $$\partial_2^2 e(s, 0) = 2\gamma_2'(s) + 2R_2(\alpha'(s), \eta(s)) + 2\left(D_{\alpha'}\eta(s)\right)^2.$$

Proof. We shall keep in mind that from the definitions,

$$D_{\alpha'}\eta(s) = D_1\partial_2\sigma(s, 0).$$

For the first derivative, we have

$$\begin{aligned}\partial_2 e &= \partial_2\langle\partial_1\sigma, \partial_1\sigma\rangle_g \\ &= 2\langle D_2\partial_1\sigma, \partial_1\sigma\rangle_g \quad \text{because } D \text{ is the metric derivative} \\ &= 2\langle D_1\partial_2\sigma, \partial_1\sigma\rangle_g \quad \text{by Lemma 5.3 of Chapter VIII.}\end{aligned}$$

This proves the first formula. For the second, we continue to differentiate, and obtain

$$\partial_2^2 e = \partial_2\langle D_1\partial_2\sigma, \partial_1\sigma\rangle_g$$

(3) $$= 2\langle D_2D_1\partial_2\sigma, \partial_1\sigma\rangle_g + 2\langle D_1\partial_2\sigma, D_2\partial_1\sigma\rangle_g.$$

In the first term on the right, we use Lemma 2.7 of Chapter IX to write

$$D_2D_1 = D_1D_2 - R(\partial_1\sigma, \partial_1\sigma).$$

In the second term on the right, we use Lemma 5.3 of Chapter VIII to write $D_2\partial_1 = D_1\partial_2$. Then we find

$$\partial_2^2 e = 2\langle D_1D_2\sigma, \partial_1\sigma\rangle_g - 2\langle R(\partial_1\sigma, \partial_2\sigma)\partial_2\sigma, \partial_1\sigma\rangle_g + 2(D_1\partial_2\sigma)^2$$

(4) $$= 2\langle D_1D_2\partial_2\sigma, \partial_1\sigma\rangle_g + 2R_2(\partial_1\sigma, \partial_2\sigma) + 2(D_1\partial_2\sigma)^2.$$

Finally, we use the metric derivative again to compute:

(5) $$\partial_1\langle D_2\partial_2\sigma, \partial_1\sigma\rangle_g = \langle D_1D_2\partial_2\sigma, \partial_1\sigma\rangle_g + \langle D_2\partial_2\sigma, D_1\partial_1\sigma\rangle_g.$$

However, $D_1\partial_1\sigma(s, 0) = D_1\partial_1\alpha(s) = D_{\alpha'}\alpha'(s) = 0$, because α is assumed to be a geodesic. Hence from (4) and (5) we find

(6) $$\partial_2^2 e(s, 0) = 2\gamma_2'(s) + 2R_2(\partial_1\sigma, \partial_2\sigma)(s, 0) + 2\left(D_1\partial_2\sigma(s, 0)\right)^2.$$

This proves (2), and concludes the proof of the lemma.

We are now ready to prove Theorem 1.3. We shall compute the second derivative of the length, which if anything is harder than that of the energy because of the square root sign. The computation for the energy follows exactly the same pattern. We begin with the first derivative, also called the **first variation**,

$$\frac{d}{dt}L(\alpha_t) = \frac{d}{dt}\int_a^b e(s, t)^{1/2}\, dt$$

L 1. $$= \int_a^b \frac{1}{2}e(s, t)^{-1/2}\partial_2 e(s, t)\, ds.$$

Then we take the second derivative

$$\frac{d^2}{dt^2}L(\alpha_t) = \int_a^b \frac{1}{2}e(s, t)^{-1/2}\partial_2^2 e(s, t)\, ds - \int_a^b \frac{1}{4}e(s, t)^{-3/2}\left(\partial_2 e(s, t)\right)^2 ds$$

whence using the hypothesis that α is parametrized by arc length,

L 2. $$\left.\frac{d^2}{dt^2}L(\alpha_t)\right|_{t=0} = \int_a^b \frac{1}{2}\partial_2^2 e(s, 0)\, ds - \int_a^b \frac{1}{4}\left(\partial_2 e(s, 0)\right)^2 ds.$$

We plug in the value from Lemma 1.4 and use a projection. If w, u are vectors and u is a unit vector, and $v = w - (w \cdot u)u$ is the orthogonalization of w with respect to u, then we have trivially

$$v^2 = w^2 - (w \cdot u)^2.$$

We apply this to $w = D_{\alpha'}\eta$ and $u = \alpha'$. Integrating as in **L 2** yields the asserted answer, and proves the formula for the second derivative of $L(\alpha_t)$ at $t = 0$. As already mentioned, the formula for the energy is easier and will be left to the reader. This concludes the proof of Theorem 1.3.

Remark 3. For simplicity we limited ourselves to curves rather than piecewise C^2 maps. Milnor [Mi 63] gives a thorough discussion of paths where end-point terms will appear where the path is broken. See for instance his Theorem 12.2 and Theorem 13.1 of Chapter III.

Remark 4. Observe how R_2 comes naturally in the formula. At a minimum one wants the second derivative to be semipositive, so having all plus signs in the variation formula is desirable.

Corollary 1.5. *Let η be a Jacobi lift of α, and σ a variation of α such that $\eta(s) = \partial_2\sigma(s, 0)$. Assume that $t \mapsto \sigma(a, t)$ and $t \mapsto \sigma(b, t)$ are geodesics. Then*

$$\left.\frac{d^2}{dt^2}E(\alpha_t)\right|_{t=0} = \langle D_{\alpha'}\eta(b), \eta(b)\rangle_g - \langle D_{\alpha'}\eta(a), \eta(a)\rangle_g.$$

In particular, if $D_{\alpha'}\eta$ is perpendicular to α' then this equality also holds if E is replaced by the length L.

Proof. Immediate from Theorem 1.3 and the alternative expressions of Proposition 1.1.

Concerning the orthogonality assumption which will recur, we recall that if a Jacobi lift η of α is such that $D_{\alpha'}\eta$ is orthogonal to α' at some

point, then $D_{\alpha'}\eta$ is orthogonal to α' on the whole interval of definition. See Proposition 2.4 of Chapter IX.

Proposition 1.6. *Assumptions being as in Theorem 1.3, suppose that $D_{\alpha'}\eta$ is orthogonal to α'. Then*

$$\left.\frac{d}{dt}L(\alpha_t)\right|_{t=0} = 0.$$

Proof. This is immediate from Lemma 1.4 (1) and **I 1**.

The next application of Theorem 1.3 gives the semipositivity of the index on the subspace of $\mathrm{Lift}_0(\alpha)$ orthogonal to α', under a natural condition.

Theorem 1.7. *Suppose that α is a geodesic whose length is the distance between its end points. Let $\zeta \in \mathrm{Lift}_0(\alpha)$ be orthogonal to α'. Then*

$$I(\zeta, \zeta) \geqq 0.$$

Proof. I owe the proof to Wu. Define

$$\sigma(s, t) = \exp_{\alpha(s)}\left(t\zeta(s)\right)$$

with $0 \leqq s \leqq b$ and $0 \leqq t \leqq \epsilon$. For each t, σ_t is a curve, not necessarily a geodesic, joining the endpoints of α, that is

$$\sigma_t(a) = \alpha(a) \quad \text{and} \quad \sigma_t(b) = \alpha(b),$$

because of the assumption $\zeta \in \mathrm{Lift}_0(\alpha)$. Furthermore, $\sigma(s, 0) = \alpha(s)$, so $\{\sigma_t\} = \{\alpha_t\}$ is a variation of α, leaving the end points fixed. Note that

$$\partial_2\sigma(s, 0) = \zeta(s).$$

Finally, the curves $t \mapsto \sigma(a, t)$ and $t \mapsto \sigma(b, t)$ are geodesics, and $D_{\alpha'}\zeta \perp \alpha'$ (differentiating $\langle\zeta, \alpha'\rangle = 0$). Therefore, if we define the function

$$\ell(t) = L(\alpha_t) = L(\sigma_t),$$

then by Theorem 1.3 we get

$$\ell''(0) = I(\zeta, \zeta).$$

Since by assumption $L(\alpha_0) \leqq L(\alpha_t)$ (because $L(\alpha)$ is the distance between the end points), it follows that the function ℓ has a minimum at $t = 0$, which proves the theorem.

Remark. The above theorem is a special case of the Morse index theorem, but will suffice for our applications. For full index theorems, cf. Milnor [Mi 63]; Kobayashi–Nomizu [KoN 69], Chapter VIII, §6; Cheeger–Ebin [ChE 75], Chapter 4; do Carmo [doC 92]; and Klingenberg [Kl 83/95], Chapter 2, Sections 2.4 and 2.5, where he deals with the energy functional on the loop space as an infinite dimensional manifold, rather than the length. However, in that chapter, the original manifolds are finite dimensional. I hope to have convinced the reader further about the irrelevance whether the manifold is finite dimensional or not.

Corollary 1.8. *Let η be a Jacobi lift of α, and let ξ be any lift of α, with the same end points as η, that is*

$$\eta(0) = \xi(0) \quad \text{and} \quad \eta(b) = \xi(b).$$

Suppose that $\eta - \xi$ is orthogonal to α'. Then

$$I(\eta, \eta) \leqq I(\xi, \xi).$$

Proof. Let $\zeta = \xi - \eta$. By Theorem 1.7 we have $I(\zeta, \zeta) \geqq 0$, so by the bilinearity of the index,

$$I(\xi, \xi) - 2I(\eta, \xi) + I(\eta, \eta) \geqq 0. \tag{7}$$

But

$$\begin{aligned} I(\eta, \xi) &= \langle D_\alpha, \eta, \xi \rangle |_a^b - \int_a^b \langle D_{\alpha'}^2 \eta, \eta \rangle - \langle R(\alpha', \eta)\alpha', \eta \rangle \\ &= \langle D_\alpha, \eta, \xi \rangle |_a^b \quad \text{because } \eta \text{ is a Jacobi lift (Proposition 1.1)} \\ &= \langle D_\alpha, \eta, \eta \rangle |_a^b \quad \text{by assumption} \\ &= I(\eta, \eta) \quad \text{because } \eta \text{ is a Jacobi lift.} \end{aligned}$$

Hence inequality (1) becomes the inequality asserted in the corollary.

For some applications, one wants to compute the second derivative of a composite function $f(L(\alpha_t))$, where f is a function of a real variable, for instance when we determine the Laplacian in polar coordinates later. So we give here the relevant formula, since it is essentially a corollary of the above considerations.

Proposition 1.9. *Let f be a C^2 function of a real variable. As in Theorem 1.3, let σ be a variation of α, and let $\eta(s) = \partial_2\sigma(s, 0)$. Assume $D_{\alpha'}\eta$ orthogonal to α'. Then*

$$\frac{d^2}{dt} f(L(\alpha_t)) \bigg|_{t=0} = f'(L(\alpha_0)) \left[\langle D_{\alpha'}\eta(b), \eta(b) \rangle_g - \langle D_{\alpha'}\eta(a), \eta(a) \rangle_g \right].$$

Proof. Let $F(t) = f(L(\alpha_t))$. Then

$$F'(t) = f'(L(\alpha_t)) \frac{d}{dt} L(\alpha_t)$$

and

$$F''(t) = f''(L(\alpha_t)) \frac{d}{dt} L(\alpha_t) + f'(L(\alpha_t)) \left(\frac{d}{dt}\right)^2 L(\alpha_t).$$

Then at $t = 0$ the first term on the right is 0 because of Proposition 1.6. The second term at $t = 0$ is the asserted one by Corollary 1.5 and the orthogonality assumption. This concludes the proof.

Example. Proposition 3.3 of Chapter IX provides an example of the situation in Proposition 1.9. Both will be used in Chapter XV, §2.

Theorem 1.3, i.e. the second variation formula, also has some topological applications which we don't prove in this book, but which we just mention. If R_2 is negative, so the curvature is positive, then one has a theorem of Synge [Sy 36]:

Let X be a compact even dimensional orientable Riemannian manifold with strictly positive sectional curvature. Then X is simply connected.

The idea is that in each homotopy class one can find a geodesic of minimal length. By the second derivative test, the expression for the second derivative of the length is 0 for such a geodesic. One has to prove that one can choose the variation such that the "orthogonal" term containing the integral of $\nu(s)^2$ is 0. The boundary term will vanish if one works with a variation to which we can apply Remark 1. Finally, having strictly positive curvature will yield a negative term, which gives a contradiction. Details can be found in other texts on Riemannian geometry.

The same ideas and the theorem of Synge lead to a theorem of Weinstein [We 67]:

Let X be a compact oriented Riemannian manifold of positive sectional curvature. Let f be an isometry of X preserving the orientation if dim X *is even, and reversing orientation if* dim X *is odd. Then f has a fixed point, i.e. a point x such that $f(x) = x$.*

Proofs of both the Synge and the Weinstein theorems are given in [doC 92]. The Synge theorem is given in [BGM 71] and [GHL 87/93].

This as far as we go in the direction of the calculus of variations. These are treated more completely in Morse theory, for instance in [Mi 63], [Pa 63], [Sm 64], and in differential geometry texts such as [KoN 69], [BGM 71], [ChE 75], [doC 92], [GHL 87/93].

Klingenberg's book [Kl 83/95] also contains topological applications, see especially Chapter 2, where Klingenberg uses the energy function rather than the length function.

INTRODUCTION TO §2

In Chapter IX, §3 we showed that when the Riemann tensor R_2 is semipositive (seminegative curvature), then the exponential map is metric semi-increasing. We now want to prove the converse. In part the argument is similar, using the Jacobi lift which gives an explicit formula for the differential of the exponential map. However, at a crucial point the argument gets somewhat more involved because instead of a straightforward convexity computation as in Chapter IX, Lemma 2.5 and Proposition 2.6, we now have to appeal to the second variation formula, especially Corollary 1.8. The basic result we are after is an immediate consequence of the Rauch comparison theorem, which is proved in standard texts on Riemannian geometry. Essentially they all use the same proof, which is a simplification by Ambrose of Rauch's original argument. See for instance [KoN 69], Vol. II, Chapter VIII, Theorem 4.1; [Kl 83/95], Chapter II, Lemma 2.7.2 and Corollary 2.7.3; [doC 92], Chapter X, §2. They all formulate the theorem in finite dimension, unnecessarily. For our purposes, we need only a special case, describing the effect of the exponential map on the metric under the curvature conditions, positive or negative. A presentation of the proof can be given more simply in this special case, as was shown to me by Wu, to whom I owe the exposition in the next section. A proof of the full Rauch theorem will be reproduced in §4.

For an alternative approach to Jacobi lift inequalities, cf. [Ka 89].

XI, §2. GROWTH OF A JACOBI LIFT

Basic Assumptions. *Throughout, we let (X, g) be a Riemannian manifold. Let $x \in X$ and let $u \in T_X$ be a unit vector. Let $\alpha: [0, b] \to X$ be the geodesic segment defined by $\alpha(s) = \exp_x(su)$. Thus α is parametrized by arclength, and the segment $\{su\}$, $0 \leqq s \leqq b$ is assumed to be in the domain of the exponential.*

We also let $w \in T_X$, $w \neq 0$ and we let $\eta = \eta_w$ be the Jacobi lift of α such that

$$\eta(0) = 0 \quad \text{and} \quad D_{\alpha'}\eta(0) = w.$$

As shown in Chapter IX, Theorem 3.1 and its proof, we have for $0 < r \leqq b$:

$$T\exp_x(ru)w = \frac{1}{r}\eta(r). \tag{1}$$

Furthermore, let

$$\sigma(s, t) = \exp_x(s(u + tu)) \qquad \text{and} \qquad \sigma_t(s) = \alpha_t(s).$$

Then $\{\sigma_t\} = \{\alpha_t\}$ will be called the **standard variation of $\alpha = \alpha_0$ in the direction of w.** We have

$$\eta(s) = \partial_2\sigma(s, 0). \tag{2}$$

Thirdly, by Chapter IX, Proposition 3.2 we have the global **Gauss lemma**

$$\langle T\exp_x(ru)u, T\exp_x(ru)w\rangle_g = \langle u, w\rangle_g. \tag{3}$$

For simplicity, we shall usually omit the subscript g.

If w is a scalar multiple of u, then

$$\|T\exp_x(ru)w\|^2 = \|w\|^2$$

by the Gauss lemma. This is another way of seeing what is also in Chapter VIII, Corollary 5.5, namely:

Proposition 2.1. *The exponential map is metric preserving on rays from the origin.*

Whereas in Chapter IX we considered the norm, we now consider the square of the norm of the Jacobi lift η, so we let

$$f(s) = \eta(s)^2 = \|\eta(s)\|^2.$$

We want to estimate the growth of $\|\eta\|$, in other words, the growth of f. We fix a value r with $0 < r \leqq b$, and we let

$$\zeta(s) = \frac{1}{\|\eta(r)\|}\eta(s) \qquad \text{for} \quad 0 \leqq s \leqq r.$$

Since $D_{\alpha'}\xi(0) = w \neq 0$, it follows that for all $r > 0$ sufficiently small, we have $\eta(r) \neq 0$. Cf. Proposition 2.6 of Chapter IX.

Lemma 2.2. *Assume that $w \perp u$ and that α is contained in a convex open set. Given r as above, there exists a lift ξ of α such that on $[0, r]$, $\xi \neq 0$, $\xi \perp \alpha'$, and*

$$\frac{1}{2}\frac{f'}{f}(r) = I_0^r(\zeta, \zeta) \leqq \frac{1}{r} + \int_0^r R_2(\alpha', \zeta)(s)\, ds.$$

Proof. We have directly from the definitions

$$\frac{1}{2}\frac{f'}{f}(r) = \langle D_{\alpha'}\zeta(r), \zeta(r)\rangle = \langle D_{\alpha'}\zeta, \zeta\rangle\Big|_0^r$$
$$= I_0^r(\zeta, \zeta)$$

because ζ is a Jacobi lift of α, and we use Proposition 1.1.

For the second inequality, let $P_0^s = P_{0,\alpha}^s$ be parallel translation along α, with $P_0^0 = \mathrm{id}$. Let v be the vector such that

$$P_0^r(v) = \zeta(r).$$

Define the lift ξ by

$$\xi(s) = P_0^s\left(\frac{s}{r}v\right). \tag{4}$$

Note that:

(5) $\xi(0) = \zeta(0)$, $\xi(r) = \zeta(r)$, $D_{\alpha'}\xi(s) = P_0^s\left(\dfrac{1}{r}v\right)$ (see Lemma 2.3).

Thus $(D_{\alpha'}\xi)^2 = 1/r^2$. By Corollary 1.8, we obtain

$$I_0^r(\zeta, \zeta) \leqq I_0^r(\xi, \xi)$$
$$= \int_0^r [(D_{\alpha'}\xi)^2 + R_2(\alpha', \xi)]$$
$$= \frac{1}{r} + \int_0^r R_2(\alpha', \xi),$$

thereby proving the lemma.

In determining $D_{\alpha'}\xi$ we used the following general lemma, which really belongs to Chapter VIII, §3, and which we state in a self-contained way.

Lemma 2.3. *Let X be any manifold with a spray, and let $\alpha: [a, b] \to X$ be a curve in X. Let $\beta: [a, b] \to T_{\alpha(a)}$ be a curve in $T_{\alpha(a)}$, let P be parallel translation along α, and let $\xi(t) = P_a^t(\beta(t))$. Then*

$$D_{\alpha'}\xi(t) = P_a^t(\beta'(t)).$$

Proof. We prove the relation in a chart, where we have the formula

$$D_{\alpha'}\xi = \xi' - B(\alpha; \alpha', \xi).$$

Let $\gamma(t, v)$ be parallel translation of $v \in T_{\alpha(a)}$. Then

$$\xi'(t) = \partial_1\sigma(t, \beta(t)) + \partial_2\gamma(t, \beta(t))\beta'(t)$$
$$= \partial_1\gamma(t, \beta(t)) + \gamma(t, \beta'(t)),$$

because $v \mapsto \gamma_t(v)$ is linear, and the derivative of a linear map is equal to the linear map. The lemma follows from the local definition of the covariant derivative, and the definition of parallel translation (Theorem 3.3 of Chapter VIII).

Lemma 2.4. *Let $h(s) = s^2w^2$. Then*

$$\lim_{s\to 0} f(s)/h(s) = 1.$$

Proof. This is immediate from the first term of the Taylor expansion given in Chapter IX, Proposition 5.1.

Theorem 2.5. *Under the basic assumptions, assume that $w \perp u$. Let U_x be an open convex neighborhood of x, and V_x an open neighborhood of 0_x such that $\exp_x: V_x \to U_x$ is an isomorphism. We suppose α is contained in U_x. If the curvature is $\geqq 0$ (resp. > 0) on U_x then*

$$\|\eta_w(r)\| \leqq r\|w\| \quad (\text{resp.} < r\|w\|) \qquad \textit{for} \quad 0 < r \leqq b.$$

Proof. By lemma 2.2, for $\epsilon > 0$ we find

$$\int_\epsilon^r f'/f \leqq \int_\epsilon^r h'/h + \text{the Riemann tensor integral.}$$

Since by hypothesis, the Riemann tensor integrand is $\leqq 0$, we obtain

$$\log f(r)/h(r) \leqq \log f(\epsilon)/h(\epsilon),$$

and therefore

$$\frac{f(r)}{h(r)} \leqq \frac{f(\epsilon)}{h(\epsilon)} \to 1 \quad \text{as } \epsilon \to 0 \text{ by Lemma 2.4.}$$

Then $f(r) \leqq h(r)$, which proves the theorem with the weak inequality sign. For the strict inequality case, one takes into account the Riemann tensor integral, and the fact that the integrand is < 0, so all inequalities are strict. This concludes the proof.

Theorem 2.6. *Let (X, g) be a Riemannian manifold. Let $x \in X$ and let $\exp_X: V_x \to U_x$ be an isomorphism of a neighborhood of 0_x with an open convex neighborhood of x. Suppose g has curvature $\geqq 0$ on U_x. Then $\exp_x$ is metric semidecreasing from V_x to U_x. If the curvature is > 0 on U_x, then for $v \in V_x$, $v \neq 0$ and $w \in T_x$, w unequal to a scalar multiple of v, we have*

$$\|T\exp_x(v)w\| < \|w\|.$$

Thus $\exp_x$ is metric strictly decreasing on V_x, except in the direction of rays from the origin.

Proof. We let u be the unit vector in the direction of v, $v = bu$. If w is orthogonal to u, then the inequality of Theorem 2.5 together with (1) shows that

$$\|T\exp_x(ru)w\|^2 < \|w\|^2.$$

For arbitrary w, we write $w = w_0 + w_1$ with $w_0 = cu$ (some $c \in \mathbf{R}$), and $w_1 \perp u$. Then by the Gauss lemma, $T\exp_x(ru)w_0 \perp T\exp_x(ru)w_1$, so

$$\|T\exp_x(ru)w\|^2 = \|T\exp_x(ru)w_0\|^2 + \|T\exp_x(ru)w_1\|^2,$$

which proves the theorem, in light of Proposition 2.1 and the inequality in Theorem 2.5.

For estimates concerning Jacobi lifts and geodesic constructions, see Buser and Karcher [BuK 81], especially 6.3 and 6.5.

XI, §3. THE SEMI PARALLELOGRAM LAW AND NEGATIVE CURVATURE

The usual parallelogram law will be semified in two ways: first, we stop at midpoint of one of the diagonals; and second, we write an inequality instead of the equality. We can then formulate things as follows.

Let X be a metric space. We say that the **semi parallelogram law** holds in X if for any two points $x_1, x_2 \in X$ there is a point z which satisfies for all $x \in X$:

$$d(x_1, x_2)^2 + 4d(x, z)^2 \leqq 2d(x, x_1)^2 + 2d(x, x_2)^2.$$

It follows that

$$d(z, x_1) = d(z, x_2) = \tfrac{1}{2}d(x_1, x_2).$$

This is obtained by setting $x = x_1$ and $x = x_2$ in the semi parallelogram law to get the inequalities $2d(x_1, z) \leqq d(x_1, x_2)$ and $2d(x_2, z) \leqq d(x_1, x_2)$. The opposite inequalities follow from the triangle inequality

$$d(x_1, x_2) \leqq d(x_1, z) + d(z, x_2).$$

Note that the point z is uniquely determined by x_1, x_2 because if z' is another such point, we put $x = z'$ in the law to see that $z = z'$. Thus we call z the **midpoint** between x_1 and x_2.

A **Bruhat–Tits space** is defined to be a complete metric space which satisfies the semi parallelogram law.

Theorem 3.1 (Serre). *Let X be a Bruhat–Tits space. Let S be a bounded subset of X. Then there exists a unique closed ball $\bar{\mathbf{B}}_r(x_1)$ in X of minimal radius containing S.*

Proof. We first prove uniqueness. Suppose there are two balls $\bar{\mathbf{B}}_r(x_1)$ and $\bar{\mathbf{B}}_r(x_2)$ of minimal radius containing S, but $x_2 \neq x_1$. Let x be any point of S, so $d(x, x_2) \leqq r$ and $d(x, x_1) \leqq r$. Let z be the midpoint between x_1 and x_2. By the semi parallelogram law, we have

$$d(x_1, x_2)^2 \leqq 4r^2 - 4d(x, z)^2.$$

By definition, for each $\epsilon > 0$ there is a point $x \in S$ such that $d(x, z) \geq r - \epsilon$. It follows that $d(x_1, x_2) = 0$, that is $x_1 = x_2$.

As to existence, let $\{x_n\}$ be a sequence of points which are centers of balls of radius r_n approaching the inf of all such radii such that $\bar{\mathbf{B}}_{r_n}(x_n)$ contains S. Let r be this inf. If the sequence $\{x_n\}$ is a Cauchy sequence, then it converges to some point which is the center of a closed ball of the minimal radius containing S, and we are done. We show this must always happen. Let z_{mn} be the midpoint between x_n and x_m. By the minimality of r, given ϵ there exists a point $x \in s$ such that

$$d(x, z_{mn})^2 \geqq r^2 - \epsilon.$$

We apply the semi parallelogram law with $z = z_{mn}$. Then

$$\begin{aligned} d(x_m, x_n)^2 &\leqq 2d(x, x_m)^2 + 2d(x, x_n)^2 - 4d(x, z_{mn})^2 \\ &\leqq \epsilon(m, n) + 4\epsilon, \end{aligned}$$

where $\epsilon(m, n) \to 0$, thus proving that $\{x_n\}$ is Cauchy, and concluding the proof of the theorem.

The center of the ball in Theorem 3.1 is called the **circumcenter** of the set S.

Theorem 3.2 (Bruhat–Tits). *Let X be a Bruhat-Tits metric space. Let G be a group of isometries of X, with the action of G denoted by $(g, x) \mapsto g \cdot x$. Suppose G has a bounded orbit (this occurs if, for instance, G is compact). Then G has a fixed point, for instance the circumcenter of the orbit.*

Proof. Let $p \in X$ and let $G \cdot p$ be the orbit. Let $\bar{\mathbf{B}}_r(x_1)$ be the unique closed ball of minimal radius containing this orbit. For any $g \in G$, the image $g \cdot \bar{\mathbf{B}}_r(x_1) = \bar{\mathbf{B}}_r(x_2)$ is a closed ball of the same radius containing the orbit, and $x_2 = g \cdot x_1$, so by the uniqueness of Theorem 3.1, it follows that x_1 is a fixed point, thus concluding the proof.

Corollary 3.3. *Let G be a topological group, H a closed subgroup. Let K be a subgroup of G, so that K acts by translation on the coset space G/H. Suppose G/H has a metric (distance function) such that translation by elements of K are isometries, G/H is a Bruhat–Tits space, and one orbit is bounded. Then a conjugate of K is contained in H.*

Proof. By Corollary 3.2, the action of K has a fixed point, i.e. there exists a coset xH such that $kxH = xH$ for all $k \in K$. Then $x^{-1}KxH \subset H$, whence $x^{-1}Kx \subset H$, as was to be shown.

We shall now discuss one of the roles of the above theorems in differential geometry. Unless otherwise specified, manifolds can be infinite dimensional, and Riemannian manifolds may therefore be Hilbertian.

One question arises: which kinds of spaces have metrics as discussed above? First we give a sufficient global condition, perhaps the most useful in practice, the **exponential metric increasing property**.

EMI. The space X is a complete Riemannian manifold, and for every $z \in X$, the exponential map

$$\exp_z : T_z \to X$$

is a differential isomorphism, is metric preserving along rays from the origin, and in general is metric semi-increasing.

A complete Riemannian manifold, simply connected, with $R_2 \geqq 0$ (seminegative curvature) is called a **Hadamard**, or **Cartan–Hadamard manifold**.

Proposition 3.4. *A complete Riemannian manifold satisfying* **EMI** *is a Bruhat–Tits space. A Cartan–Hadamard manifold is a Bruhat–Tits space.*

Proof. On a Hilbert space, we have equality in the parallelogram law. Using the hypothesis in **EMI** with z as the midpoint, we see that the left side in the parallelogram law remains the same under the exponential map, the right side only increases, and hence the semi parallelogram law falls out.

Next we give equivalent local conditions. Let X be a Riemannian manifold. We say that the **semi parallelogram law** holds **locally** on X if every point x has an open convex neighborhood U_x with an isomorphism $\exp_x : V_x \to U_x$ of a neighborhood of 0_x in T_x, such that the semi parallelogram law holds in U_x.

Theorem 3.5. *Let X be a Riemannian manifold. The following three conditions are equivalent:*

(a) *The curvature is seminegative.*
(b) *The exponential map is locally metric semi-increasing at every point.*
(c) *The semi parallelogram law holds locally on X.*

Proof. This is merely putting together results which have been proved individually. Theorem 3.6 of Chapter IX shows that (a) implies (b). That (b) implies (c) is a local version of Proposition 3.4. Indeed, the parallelogram law holds in the tangent space T_z, and if the exponential map at z is metric semi-increasing, then the semi parallelogram law holds locally by applying the exponential map. Specifically, given x_1, x_2 in some convex open set, we let z be the midpoint on the geodesic joining x_2 and x_2, so that there is some $v_1 \in T_z$ such that, putting $v_2 = -v_1$,

$$x_1 = \exp_z(v_1), \qquad x_2 = \exp_z(v_2), \qquad z = \exp_z(0_z).$$

Given $x = \exp_z(v)$ with $v \in T_z$ the parallelogram law in T_z reads

$$d(v_1, v_2)^2 + 4d(v, 0)^2 = 2d(v, v_1)^2 + 2d(v, v_2)^2,$$

where $d(v, w) = |v - w|$ for $v, w \in T_z$. Under the exponential map, the distances on the left side are preserved, and the distances on the right side are expanded if condition (b) is satisfied, so under the exponential map, we get

$$d(x_1, x_1)^2 + 4d(x, z)^2 \leqq 2d(x, x_1)^2 + 2d(x, x_2)^2,$$

which is the semi parallelogram law. Finally, to show that (c) implies (a), we merely follow the same argument with the reverse inequality. So assume (c). Suppose the curvature is positive at some point, and hence is positive on a convex open neighborhood of the point, which we denote by z. We pick a vector $v \in T_z$, and let $v_1 = -v$, $v_2 = v$. We let $w \perp v$, $w \neq 0$. Then

$$d(v_1, v_2)^2 + 4d(v + w, 0)^2 = 2d(v + w, v_1)^2 + 2d(v + w, v_2)^2,$$

because this relation is one with the norm in the Hilbert space T_z. Now we apply the exponential map, that is, we let

$$x_1 = \exp_z(v_1), \qquad x_2 = \exp_z(v_2), \qquad x = \exp_z(v + w), \qquad z = \exp_z(0_z).$$

The distances on the left side of the equation are preserved under the exponential map (taking the norms of v, w sufficiently small). By Theorem 2.6, the distances on the right are strictly decreased, contradicting the semi parallelogram law (actually giving an anti semi parallelogram inequality). This concludes the proof.

The equivalence of the semi parallelogram law and seminegative curvature is due to Bruhat–Tits [BrT 72].

In the next chapter we shall give the most important classical example of a Cartan–Hadamard manifold. We note that of the three equivalent conditions in Theorem 3.5, the most subtle is the curvature condition, and the simplest is the semi parallelogram law, which can be formulated independently of the theory of manifolds. In the example, we show that the conditions are satisfied by actually proving the metric increasing property of the exponential map, which is an intermediate condition establishing a link between the other two. It may be useful to formulate here a weak variation of Chapter VIII, Theorem 6.9, because it summarizes in an easy way some consequences of the metric increasing property which we shall prove in a concrete case in the next section.

Theorem 3.6. *Let* $\mathbf{E}$ *be a Hilbert space and* X *a Riemannian manifold. Let* $h\colon \mathbf{E} \to X$ *be a differential isomorphism which is metric semi-increasing, that is*

$$|Th(v)w|^2_{h(v)} \geqq |w|^2_{\mathbf{E}} \qquad \text{for all} \quad w \in \mathbf{E},$$

and also such that h *is metric preserving on rays from the origin. Then* X *is complete. Let* $v \in \mathbf{E}$, $v \neq 0$. *Then*

$$t \mapsto h(tv)$$

is a geodesic passing through $h(0)$ *and* $h(v)$, *and is the unique such geodesic. If the group of isometries of* X *operates transitively, then there is a unique geodesic through two distinct points of* X.

Proof. The map $h^{-1}\colon X \to \mathbf{E}$ is distance semi-decreasing. If $\{x_n\}$ is Cauchy in X, then $\{h^{-1}(x_n)\}$ is Cauchy in $\mathbf{E}$, converging to some point v, and by continuity of h, it follows that $\{x_n\}$ converges to $h(v)$, so X is complete. If α is a geodesic in X between two points x and y, then $h^{-1} \circ \alpha$ is a curve in $\mathbf{E}$ between $h^{-1}(x)$ and $h^{-1}(y)$. In $\mathbf{E}$, the geodesics with respect to the Hilbert space norm are just the lines, which minimize distance. From the property that h preserves distances on rays from the origin, it follows at once that if $x = h(v)$ and ξ is the line segment from 0 (in $\mathbf{E}$) to v, then $h \circ \xi$ minimizes the distance between $h(0)$ and $h(v)$, and so $h \circ \xi$ is the unique geodesic between $h(0)$ and $h(v)$. If the group of isometries of X operates transitively, then the last statement is clear, thus concluding the proof.

Historical note. The presentation of the above material essentially follows a path which is the reverse of the historical path. It took almost a century before certain ideas were given their full generality and simplicity.

Historically, things start at the end of the nineteenth century. Klingenberg [Kl 83/95] asserts that von Mangoldt essentially proved what is called today the Cartan–Hadamard theorem for surfaces [vM 1881], 15 years before Hadamard did so [Ha 1896]. Actually, von Mangoldt refers to previous papers by others before him, Hadamard refers to von Mangoldt, and Cartan [Ca 28] refers to Hadamard (Cartan dealt with arbitrary Riemannian manifolds). I am unable to read the original articles.

Helgason [He 62] gave a proof of Cartan's fixed point theorem following Cartan's ideas, see the revised version [He 78], Chapter I, Theorem 13.5, namely: On a Riemannian manifold of seminegative curvature, a compact group of isometries has a fixed point. Cartan's immediate application was to show that all maximal compact subgroups of a semisimple Lie group are conjugate. Mostow [Mo 53] gave a similar exposition, but in a more limited context. See also Kobayashi–Nomizu [KoN 69], Vol. II, Chapter VIII, Theorems 9.1 and 9.2. They use a center of mass rather than the circumcenter.

Then Bruhat–Tits [BrT 72] formulated their fixed point theorem, here stated as Theorem 1.2, setting up the parallelogram law condition prominently. Serre used a variation of their proof and the formulation of Theorem 3.1 to reach the currently ultimate result with the very simple proof we have given. I don't know that Serre published this, but it is referred to exactly as we have stated it in Brown [Bro 89], Chapter VI, Theorem 2 of §5. Thus a line of thoughts which started a century before, abuts to a basic elementary theorem about metric spaces. The condition of compactness is replaced by the condition of boundedness, and the more complicated notion of curvature is replaced by the semi parallelogram law.

In addition, the center of mass which occurred in Cartan's treatment (and others following him), is replaced by the circumcenter, following Bruhat–Tits.

In [Kl 83/95], Theorem 1.10.17 (also 1.10.18), Klingenberg formulated a version of Cartan's theorem under convexity hypotheses, in the context of manifolds rather than merely metric spaces. He attributed the idea of his proof to Eberlein. What Klingenberg proves is actually Theorem 3.1 for compact sets in a differential geometric context, although one has to analyze the proof to see this in 1.10.17. We note that Theorem 3.1 for compact sets occurs in [BGS 85], Lemma 1, p. 10.

Bruhat–Tits actually *characterized* Cartan–Hadamard spaces by the semi parallelogram law [BrT 72]. From that point on, a theory of curvature for metric spaces rather than manifolds developed separately, with an extensive exposition in Ballman [Ba 95], containing Theorem 3.1. Ballman refers to Brown for Theorem 3.1, cf. [Ba 95], Theorem 5.1 and Proposition 5.10 of Chapter I.

Note that in the metric theory which has been developed in parallel, the geodesic between two points has been obtained as the curve arising by taking successive midpoints ad infinitum. Helgason has pointed out to me that this limiting procedure was already used by Cartan [Ca 46], pp. 360–363!

There remains to say a few more words here about the infinite dimensional case. The importance of infinite dimensional manifolds was recognized in the sixties, e.g. for function spaces, for Morse theory, and for the Nash–Moser theorem on Riemannian metrics. We note that Klingenberg [Kl 83/95] has a nice chapter on the H^1-loop space. However, differential properties of curvature are not fully carried out in these works. For example, Klingenberg does not do the Cartan–Hadamard theorem in the infinite dimensional case, and he also defines symmetric spaces only in the finite dimensional case.

I like Marsden's book [Ma 74], especially §7 and §9, where he already suggests infinite dimensional contexts for various notions of differential geometry, although his alternation between finite dimensional manifolds and infinite dimensional manifolds does not present a clear account of theorems and proofs valid in the infinite dimensional case, as it applies to Hamiltonian mechanics, quantum mechanics, and relativity, including the infinite dimensional space of Riemannian metrics, and its differential geometry.

Garland's study of loop groups [Ga 80] is another candidate to be placed in the setting of infinite dimensional Bruhat–Tits spaces. A similar possibility exists for the article by Atiyah and Pressley [AtP 83].

Other possibilities are spaces having to do with the KdV equation, e.g. [ScTZ 96], and "moduli" spaces in a broad sense, e.g. Teichmuller spaces, spaces of Riemannian metrics as in Ebin [Eb 70] and Fried–Groisser [FrG 89], spaces of Kahler metrics, spaces of connections, ad lib. Anders Karlsson has told me that the metric in [Eb 70] has seminegative curvature. Karlsson has also pointed out that once it is proved that some Teichmuller space has seminegative curvature and some completeness properties, then the Bruhat–Tits fixed point theorem can be applied routinely to establish the so-called Nielsen Realization problem for the corresponding mapping class group. In the Riemann surface case, this is due to Kerckhoff [Ker 83], see also Wolpert's paper [Wo 87].

Once one becomes aware of the possibility of applying the Cartan–Hadamard theory in infinite dimensions, one realizes that examples abound. A systematic account of the general theory of symmetric spaces and their applications in the infinite dimensional case remains to be worked out.

XI, §4. TOTALLY GEODESIC SUBMANIFOLDS

The main concrete point of this section is to consider certain submanifolds of Cartan–Hadamard manifolds which are themselves Cartan–Hadamard.

Let X be a Riemannian manifold and let Y be a closed submanifold. We define Y to be **totally geodesic** if every geodesic in X with initial conditions in (Y, TY) is contained in Y. There is an alternative condition which we discuss.

Theorem 4.1. *Let X be a Cartan–Hadamard manifold. Let Y be a totally geodesic submanifold. Then:*

(i) *Y is a Cartan–Hadamard manifold.*
(ii) *Given two distinct points of Y, the unique geodesic in X passing through these points actually lies in Y.*

Proof. Note that from the definition of a totally geodesic submanifold, it follows that the exponential map on X, restricted to TY, is equal to the exponential map on Y, or in a formula, for $y \in Y$,

$$\exp_{y,Y} = \exp_{y,X} \quad \text{restricted to } T_y Y.$$

By hypothesis and the definitions, it follows that $\exp_{y,Y}$ is metric semi-increasing, so Y has seminegative curvature by Theorem 3.5. By hypothesis, Y is geodesically complete, and hence complete by Corollary 3.9 of Chapter IX. By Theorem 3.8 of Chapter IX, given $y \in Y$, the exponential

$$\exp_y : T_y Y \to Y$$

is a covering, and since it is injective because $\exp_y : T_y X \to X$ is injective, it follows that $\exp_y : T_y Y \to Y$ is an isomorphism, so Y is simply connected. Thus we have shown that Y is Cartan–Hadamard. Then (ii) is trivial from (i), because the unique geodesic in Y passing through two distinct points is the same as the unique geodesic in X passing through these points. This concludes the proof.

We complement the situation by a general statement, converse of (ii) in the theorem. It is included for completeness, but will not be used.

Proposition 4.2. *Let X be a complete Riemannian manifold, such that given two distinct points of X, there is a unique geodesic passing through these two points. Let Y be a closed submanifold. Suppose that locally, given two distinct points in Y, the unique geodesic segment in X joining these points actually lies in Y. Then Y is totally geodesic.*

Proof. I owe the following simple argument to Wu. One has mostly to prove that a Y-geodesic is an X-geodesic. Let $\alpha : [0, c) \to X$ be a geodesic in X having initial conditions in Y, that is

$$\alpha(0) = y \in Y \quad \text{and} \quad \alpha'(0) \in T_y Y.$$

Suppose α does not lie in Y. Then there is a largest number b such that $\alpha([0, b]) \subset Y$ but $\alpha(b + \epsilon) \notin Y$ for all small $\epsilon > 0$. Note that b could be 0. Since $\alpha([0, b]) \subset Y$, it follows that $\alpha'(b) \in T_{\alpha(b)} Y$. This is true even if $b = 0$, by assumption. Let

$$\beta : [b, b + \epsilon] \to Y$$

be the geodesic in Y such that $\beta'(b) = \alpha'(b)$, with sufficiently small ϵ, so that $\beta(b + \epsilon)$ lies in a convex X-ball centered at $\alpha(b)$, and also in a convex Y-ball centered at this same point $\alpha(b)$. Let

$$\gamma : [b, b + \epsilon] \to X$$

be the geodesic segment in X joining $\beta(b)$ and $\beta(b + \epsilon)$. By hypothesis, we have $\gamma([b, b + \epsilon]) \subset Y$. But a geodesic of X lying in Y is necessarily a geodesic of Y, say by the minimizing characterization of geodesics. By uniqueness, we have $\gamma = \beta$ on $[b, b + \epsilon]$. But then

$$\gamma'(b) = \beta'(b) = \alpha'(b),$$

and so γ is in fact the continuation of the restriction of α to $[0, b]$. Hence $\alpha([b, b + \epsilon])$ is contained in Y, contradiction concluding the proof.

In Chapter X, §2 we proved an extension of the Cartan–Hadamard theorem, to the normal bundle. We shall now show how part of the proof can be replaced by another argument. Specifically, we prove:

Lemma 4.3. *Let X be a Cartan–Hadamard manifold. Let Y be a totally geodesic submanifold. Then the map*

$$\exp_{NY} : NY \to X$$

is a bijection.

Proof. The argument will follow the same pattern that is used routinely to show that given a point not in a closed subspace of a Hilbert space, there is a line through the point perpendicular to the subspace. We first prove that given $x \in X$ but $x \notin Y$, there exists a point $y_0 \in Y$ such that

$$d(x, y_0) = d(x, Y) = \inf_{y \in Y} d(x, y).$$

Let $\{y_n\}$ be a sequence in Y such that $d(x, y_n)$ approaches $r = d(x, y)$ as n goes to infinity. We can apply the semi parallelogram law in X exactly as in the proof of Theorem 3.1. The midpoint in X is on the geodesic between the two points, and lies in Y because of the assumption that Y is totally geodesic. Then the semi parallelogram law shows at once that $\{y_n\}$ is Cauchy, and therefore converges to the desired point y_0. The unique geodesic through x and y_0 is perpendicular to Y at y_0 by Corollary 4.7 of Chapter IX. Furthermore, this geodesic cannot intersect Y in another point y_1, otherwise the existence of this geodesic and the geodesic in Y between y_0 and y_1 would contradict Corollary 3.11 of Chapter IX. Thus we conclude that the map $E : NY \to X$ is bijective.

The above lemma provides a variation for part of the proof of Chapter X, Theorem 2.5, avoiding further appeal to the Ambrose Theorem 6.9 of Chapter VIII. However, an important additional step is still required to prove the local C^1-isomorphism property, unavoidably using some estimates for Jacobi lifts as in proposition 2.6 of Chapter IX. Of course, one may use at this point Theorem 2.4 and Lemma 2.6 of Chapter X.

XI, §5. RAUCH COMPARISON THEOREM

Because it does not take very long, we shall give the proof of Rauch's comparison theorem. I follow the exposition given in [KoN 69], Vol. II, Chapter VIII, Theorem 4.1; or Cheeger–Ebin [ChE 75], Chapter I, §10. As already mentioned, it is derived from Ambrose. We first make some preliminary remarks.

Let (X, g_X) and (Y, g_Y) be Riemannian manifolds. Let

$$\alpha_X: [a, b] \to X \qquad \text{and} \qquad \alpha_Y: [a, b] \to Y$$

be geodesics, defined on the same interval, and parametrized by arc length. We have the Riemann tensors $R_{2,X}$ and $R_{2,Y}$ on X and Y, respectively, and thus we have their values

$$R_{2,X}(\alpha_X(s)) \qquad \text{and} \qquad R_{2,Y}(\alpha_Y(s)) \qquad \text{for all} \quad s \in [a, b].$$

We define $R_{2,X} \leqq R_{2,Y}$ **along** (α_X, α_Y) if for each s and every pairs of orthogonal vectors $v, w \in T_{\alpha_X(s)}X$ and $v', w' \in T_{\alpha_Y(s)}Y$ such that v, v' have the same length and w, w' have the same length, we have

$$R_{2,X}(v, w) \leqq R_{2,Y}(v', w').$$

If K denotes the curvature, this means that

$$K_X(P) \geqq K_Y(Q)$$

for every plane P contained in $T_{\alpha_X(s)}X$ and every plane Q contained in $T_{\alpha_Y(s)}Y$. The Rauch comparison theorem will compare Jacobi lifts of the two geodesics in terms of the Riemann tensor (curvature, with an opposite sign).

Theorem 5.1 (Rauch Comparison Theorem). *Let (X, g_X) and (Y, g_Y) be Riemannian manifolds of the same dimension, which may be infinite. Let α_X (resp. α_Y) be geodesics in X (resp. Y), parametrized by arc length, and defined on the same interval $[a, b]$. Let η_X (resp. η_Y) be Jacobi lifts or these geodesics, orthogonal to α'_X (resp. α'_Y).*
Assume:

(i) *$\eta_X(a) = \eta_Y(a) = 0$, and $\eta_X(r)$, $\eta_Y(r) \neq 0$ for $0 < r \leqq b$.*
(ii) *$\|D_{\alpha'_X}\eta_X(a)\| = \|D_{\alpha'_Y}\eta_Y(a)\|$.*
(iii) *The length of α_X is the distance between its end points.*
(iv) *We have $R_{2,X} \leqq R_{2,Y}$ along (α_X, α_Y).*

Then

$$\|\eta_X(s)\|^2 \leqq \|\eta_Y(s)\|^2 \qquad \textit{for all} \qquad s \in [a, b].$$

Proof. We shall use the definition of the index and Proposition 1.1, that is, for a Jacobi lift η of α such that $\eta(a) = 0$ we have

$$(1) \qquad I_a^s(\eta, \eta) = \int_a^s (D_{\alpha'}\eta)^2 + R_2(\alpha', \eta) = \langle D_{\alpha'}\eta, \eta\rangle(s).$$

We may index η by X and Y as well. We define

$$f(s) = \|\eta(s)\|^2 = \eta(s)^2, \qquad \text{also written} \quad \eta^2(s),$$

and again we may index f and η by X, and also by Y. Define

$$h(s) = I_a^s(\eta, \eta)/\eta^2(s) \qquad \text{for} \qquad 0 < s \leqq b.$$

Thus we have h_X and h_Y. Note that by (1),

$$f'(s) = 2I_a^s(\eta, \eta) \qquad \text{and} \qquad f'/f = 2h.$$

For $a < c < b$, we get

$$\log \eta^2(s) = \log \eta^2(c) + 2\int_c^s h,$$

whence

$$\log(\eta_X^2(s)/\eta_Y^2(s)) = \log(\eta_X^2(c)/\eta_Y^2(c)) + 2\int_c^s (h_X - h_Y).$$

By assumptions (i) and (ii), and the first term of the Taylor expansion of a Jacobi lift (Chapter IX, Proposition 5.1), we get

$$\lim_{c \to a} \log \eta_X^2(c)/\eta_Y^2(c) = 0.$$

Hence

$$\log \eta_X^2(s)/\eta_Y^2(s) = \lim_{c \to a} 2\int_c^s (h_X - h_Y).$$

It will therefore suffice to prove that $h_X(s) \leqq h_Y(s)$ for $a < s \leqq b$. Fix r with $a < r < b$. It will suffice to prove $h_X(r) \leqq h_Y(r)$. Define

$$\zeta(s) = \frac{1}{\|\eta(r)\|}\eta(s),$$

so we may index ζ by X (resp. Y) to get ζ_X and ζ_Y. Let $W(s) = \alpha'(s)^\perp$ be the orthogonal complement of $\alpha'(s)$, so we have $W_X(s)$ and $W_Y(s)$ in the tangent spaces at $\alpha_X(s)$ and $\alpha_Y(s)$, respectively. Let

$$L_r\colon W_Y(r) \to W_X(r)$$

be a linear metric isomorphism such that

$$L_r(\zeta_Y(r)) = \zeta_X(r).$$

Such a metric isomorphism exists since X, Y are assumed to have the same (possibly infinite) dimension. Let P_X (resp. P_Y) be parallel translation along α_X (resp. α_Y). For each s we obtain a metric linear isomorphism

$$L_s\colon W_Y(s) \to W_X(s) \qquad \text{defined by} \qquad L_s = P^s_{r,X} \circ L_r \circ P^r_{s,Y}.$$

Define

$$\xi(s) = L_s(\zeta_Y(s)) = P^s_{r,X} \circ L_r \circ P^r_{s,Y}(\zeta_Y(s)).$$

Then ζ_X and ξ have the same end points at $s = a$ and $s = r$. Furthermore

(2) $$\xi(s)^2 = \zeta_Y(s)^2 \qquad \text{and} \qquad (D_{\alpha'_X}\xi)^2 = (D_{\alpha'_Y}\zeta_Y)^2.$$

The first equality follows from the fact that parallel translation is a metric linear isomorphism. The second follows at once from Lemma 2.3, by using the curve

$$\beta(s) = P^r_{s,X} \circ \xi(s) \qquad \text{in} \quad T_{\alpha_X(r)}X.$$

Now we find:

$$I^r_a(\zeta_{X'}, \zeta_X) \leqq I^r_a(\xi, \xi) \quad \text{by Corollary 1.8 and assumption (iii)}$$

$$= \int_a^r (D_{\alpha'_X}\xi)^2 + R_{2,X}(\alpha'_X, \xi) \quad \text{by definition (1)}$$

$$\leqq \int_a^r (D_{\alpha'_Y}\zeta_Y)^2 + R_{2,Y}(\alpha'_Y, \zeta_Y) \quad \text{by (2) and by assumption (iv)}$$

(3) $$= I^r_a(\zeta_Y, \zeta_Y) \quad \text{by definition.}$$

From the definition of ζ, inequality (3) can be rewritten

$$I^r_a(\eta_X, \eta_X)/\eta^2_X(r) \leqq I^r_a(\eta_Y, \eta_Y)/\eta^2_Y(r)$$

which means by definition that $h_X(r) \leqq h_Y(r)$, and concludes the proof.

Remark. Instead of inequality (3), one has the precise relation

(4) $$I^r_a(\zeta_X, \zeta_X) = I^r_a(\zeta_Y, \zeta_Y) + H^r_a$$

where

$$H^r_a = \int_a^r \big(R_{2,X}(\alpha'_X, \zeta) - R_{2,Y}(\alpha'_Y, \zeta_Y)\big).$$

The integrand on the right gives the precise contribution coming from assumption (iv), and shows that if there is a strict inequality in (iv), then there is a strict inequality in the conclusion $h_X(r) < h_Y(r)$.

CHAPTER XII

An Example of Seminegative Curvature

The present chapter gives a prototype for a Cartan–Hadamard manifold, so complete, with seminegative curvature, simply connected, namely the space of symmetric positive definite real matrices of a given dimension. The example is obtained by "bending" a flat euclidean space by an exponential map. The chapter is essentially elementary, and could be read after the reader is acquainted with Riemannian metrics.

XII, §1. $\mathrm{Pos}_n(\mathbf{R})$ AS A RIEMANNIAN MANIFOLD

Let:

$\mathrm{Mat}_n(\mathbf{R})$ = space of $n \times n$ real matrices;
$\mathrm{Pos}_n(\mathbf{R})$ = space of symmetric positive definite $n \times n$ matrices v (we write $v > 0$);
$\mathrm{Sym}_n(\mathbf{R})$ = vector space of symmetric $n \times n$ real matrices;
$\mathrm{GL}_n(\mathbf{R}) = G$ = group of invertible real $n \times n$ matrices.

We usually omit $\mathbf{R}$ for simplicity, and write simply Pos_n and Sym_n. We recall that a matrix p is called **positive** (or **positive definite**) if it is symmetric and

$$\langle p\xi, \xi\rangle > 0 \qquad \text{for all} \quad \xi \in \mathbf{R}^n, \quad \xi \neq 0.$$

We have the exponential map

$$\exp\colon \mathrm{Mat}_n \to \mathrm{GL}_n$$

which we shall actually consider on the symmetric matrices

$$\exp\colon \mathrm{Sym}_n \to \mathrm{Pos}_n,$$

given by the usual power series

$$\exp(v) = \sum_{n=0}^{\infty} \frac{v^n}{n!}.$$

The image lies in Pos_n because if v is symmetric then $p = q^2$ where $q = \exp(v/2)$ and q is symmetric commuting with p, so positive. From linear algebra, the exponential map is differential (i.e. C^∞) isomorphism, namely it has a C^∞ inverse, which can be called the **logarithm**. To see this, let p be a positive matrix. We can diagonalize p, that is there exists a basis $\xi_1, \dots, \xi_n$ of $\mathbf{R}^n$ and numbers $\lambda_1, \dots, \lambda_n > 0$ such that

$$p\xi_i = \lambda_i \xi_i \qquad \text{for} \quad i = 1, \dots, n.$$

Then one defines $\log p = v$ to be the linear map represented by the diagonal matrix

$$\begin{pmatrix} \log \lambda_1 & & \\ & \ddots & \\ & & \log \lambda_n \end{pmatrix}$$

with respect to the basis $\xi_1, \dots, \xi_n$. Similarly, one can define a square root of p to be the linear map represented by the matrix

$$\begin{pmatrix} \lambda_1^{1/2} & & \\ & \ddots & \\ & & \lambda_n^{1/2} \end{pmatrix}$$

with respect to the basis $\xi_1, \dots, \xi_n$. The appendix shows how to define the similar notions on Hilbert space, in a more invariant fashion.

The restriction of exp to lines through 0 is a group isomorphism from each line to its image, and is called a **one-parameter subgroup**.

On Sym_n we have a natural positive definite scalar product, defined by

$$\langle v, w\rangle_{\mathrm{tr}} = \langle v, w\rangle = \mathrm{tr}(vw),$$

where tr is the trace of a matrix. The tangent space at a point $p \in \mathrm{Pos}_n$ is a translation of Sym_n. Without using more sophisticated language, we may identify it with Sym_n (as we identify $\mathbf{R}^n$ with the tangent space at any point). We then define a positive definite scalar product depending on a point $p \in \mathrm{Pos}_n$ by letting

$$\langle v, w\rangle_p = \mathrm{tr}(p^{-1}vp^{-1}w) \qquad \text{so that} \qquad |v|_p^2 = |v|_{p,\mathrm{tr}}^2 = \mathrm{tr}\big((p^{-1}v)^2\big).$$

The positive definiteness comes from the fact that

$$\mathrm{tr}(p^{-1}vp^{-1}v) = \mathrm{tr}(p^{-1/2}vp^{-1/2}p^{-1/2}vp^{-1/2}) = \mathrm{tr}(w^2)$$
$$\text{with} \quad w = p^{-1/2}vp^{-1/2},$$

and $\mathrm{tr}(w^2) > 0$ if $v \neq 0$. If $t \mapsto p(t)$ is a curve in Pos_n, then the differential of the length is given by

$$(ds/dt)^2 = \mathrm{tr}\big(p(t)^{-1}p'(t)\big)^2 \qquad \text{abbreviated} \quad ds^2 = \mathrm{tr}(p^{-1}\,dp)^2.$$

We call the above metric the **trace metric**. It defines Pos_n as a Riemannian manifold X. It is equal to a constant times the Killing metric, which the reader will find in other books.

The group $G = \mathrm{GL}_n(\mathbf{R})$ acts on Pos_n. For each $g \in G$ we let $[g]$ be the map of Pos_n into itself defined by

$$p \mapsto gp{}^tg = [g]p, \qquad \text{for} \quad p \in \mathrm{Pos}_n.$$

Indeed, $[g]p$ is positive, because of all $\xi \in \mathbf{R}^n$, $\xi \neq 0$ we have

$$\langle [g]p\xi, \xi\rangle = \langle gp{}^tg\xi, \xi\rangle = \langle p{}^tg\xi, {}^tg\xi\rangle > 0.$$

Since, as we have remarked earlier, every positive definite matrix is the square of a positive matrix, it follows that G acts transitively on Pos_n. In particular, if $p = g^2$ with $g \in \mathrm{Pos}_n$ then $p = [g]e$, where e is the unit matrix. The appendix shows that all these statements remain valid in Hilbert space, except that the definition of the scalar product $\langle v, w\rangle_p$ by means of the trace is a finite dimensional phenomenon, so this is one point where it remains to be seen whether the theory has an extension to the infinite dimensional case. We shall list the properties we are using carefully, to make proof analyses easier.

Theorem 1.1. *The association $g \mapsto [g]$ is a representation of G in the group of isometries of Pos_n, that is each $[g]$ is an isometry.*

Proof. First we note that $[g]$ can also be viewed as a map on the whole vector space Sym_n, and this map is linear as a function of such matrices.

Hence its derivative is given by

$$[g]'(p)w = gw{}^tg \qquad \text{for all} \quad w \in \mathrm{Sym}_n.$$

Now we verify that $[g]$ preserves the scalar product, or the norm. We have:

$$\begin{aligned} |[g]'(p)w|_{[g]p}^2 &= \mathrm{tr}\big(([g]p)^{-1}gw{}^tg)^2\big) \\ &= \mathrm{tr}\big((gp{}^tg)^{-1}gw{}^tg)^2\big) \\ &= \mathrm{tr}({}^tg^{-1}p^{-1}g^{-1}gw{}^tg{}^tg^{-1}p^{-1}g^{-1}gw{}^tg) \\ &= \mathrm{tr}({}^tg^{-1}p^{-1}wp^{-1}w{}^tg) \\ &= \mathrm{tr}\big((p^{-1}w)^2\big) \\ &= |w|_p^2 \end{aligned}$$

which proves the theorem.

Let $K = \mathrm{Uni}_n(\mathbf{R}) = O(n)$ be the group of real unitary matrices. Then K is a compact subgroup of G, and a standard elementary fact of linear algebra asserts that the map

$$\mathrm{Pos}_n \times K \to G \qquad \text{given by} \quad (p, k) \mapsto pk$$

is a differential (even real analytic) isomorphism. However, we use an isomorphism of G-homogeneous spaces $\varphi\colon G/K \to \mathrm{Pos}_n$ given by

$$\varphi\colon gK \mapsto g{}^tg = [g]e.$$

The elements $k \in K$ are precisely the elements $k \in G$ such that ${}^tk = k^{-1}$. Left translation by an element $g \in G$ acting on G/K corresponders to $[g]$ under φ.

Theorem 1.2. *The map* $\exp\colon \mathrm{Sym}_n \to \mathrm{Pos}_n$ *is metric semi-increasing. On rays from the origin, it is metric preserving. It is equal to the exponential map associated to the metric, i.e. the curves $t \mapsto \exp(tw)$ with $w \in \mathrm{Sym}_n$ are geodesics. Thus* Pos_n *satisfies condition* **EMI**, *and is a Cartan–Hadamard and Bruhat–Tits space.*

The main part of the proof will be given in the next section. Here we shall make some remarks, taking care of the easier aspects of the theorem. First note that the two stated metric properties imply that our naive exponential series is actually the exponential map associated to the metric, by applying Theorem 3.6 of Chapter X.

Next, since the action $[g]$ of $g \in G$ preserves the metric, and hence preserves distances, and since every element of Pos_n can be written as $[g]e$, to prove condition **EMI**, it suffices to do it at the origin, namely at e.

Next, we give a simple self-contained proof that our naive exponential map preserves the metric along rays from the origin.

Theorem 1.3. *The exponential map* $\exp\colon \mathrm{Sym}_n \to \mathrm{Pos}_n$ *is metric preserving on a line through the origin.*

Proof. Such a line has the form $t \mapsto tv$ with some $v \in \mathrm{Sym}_n$, $v \neq 0$. We need to prove

$$|v|_{\mathrm{tr}}^2 = |\exp'(tv)v|_{\exp tv}^2.$$

Note that

$$\begin{aligned}\frac{d}{dt}\exp(tv) &= \exp'(tv)v\\ &= \frac{d}{dt}\sum \frac{t^n v^n}{n!}\\ &= \sum \frac{t^{n-1}}{(n-1)!}v^n\\ &= \exp(tv)v.\end{aligned}$$

Hence

$$\begin{aligned}|\exp'(tv)v|_{\exp tv}^2 &= \mathrm{tr}(((\exp tv)^{-1}(\exp tv)v)^2)\\ &= \mathrm{tr}(v^2)\\ &= |v|_{\mathrm{tr}}^2,\end{aligned}$$

which proves the theorem.

As an application, we can determine explicitly the distance between two points in Pos_n, as follows.

Theorem 1.4. *Let* $p, q \in \mathrm{Pos}_n$. *Let* $a_1, \ldots, a_n$ *be the roots of* $\det(tp - q)$. *Then*

$$\mathrm{dist}(p, q) = \sum (\log a_i)^2.$$

Proof. Suppose first $p = e$ and q is the diagonal matrix of $a_1, \ldots, a_n$. Let $v = \log q$, so v is diagonal with components $\log a_1, \ldots, \log a_n$. The theorem is then a consequence of Theorem 1.3, since v^2 has components $(\log a_i)^2$. We reduce the general case to the above special case. First we claim that there exists $g \in G$ such that $[g]p = e$ and $[g]q = d$ is diagonal. Indeed, we first translate p to e, so without loss of generality we may assume $p = e$. There exists an orthonormal basis of $\mathbf{R}^n$ diagonalizing q, so there exists a diagonal matrix d and $k \in K$ such that $q = k\, d\,{}^t k = k\, dk^{-1}$.

But $e = kk^{-1}$, so taking $[k]q$ proves our claim. Finally, from the equations $gp\,{}^tg = e$ and $gq\,{}^tg = d$ we get $p = g^{-1}\,{}^tg^{-1}$ and $q = g^{-1}d\,{}^tg^{-1}$, so

$$\begin{aligned}\det(tp - q) &= \det(tg^{-1}\,{}^tg^{-1} - g^{-1}d\,{}^tg^{-1})\\ &= \det(g)^{-2}\det(te - d).\end{aligned}$$

Since $\mathrm{dist}(p, q) = \mathrm{dist}(e, d)$, the theorem follows.

The next section contains the main part of the proof for the metric increasing property of the exponential map, and §3 contains further results about totally geodesic submanifolds of Pos. Except for a slight axiomatization, I follow Mostow's very elegant exposition of Cartan's work [Mo 53], in both sections.

XII, §2. THE METRIC INCREASING PROPERTY OF THE EXPONENTIAL MAP

We shall need only a few very specific properties of the exponential map, and the trace scalar product, so we axiomatize them to make the logic clearer.

We let $\mathscr{A}$ *be a finite dimensional algebra over* $\mathbf{R}$, *with an anti-involution, that is a linear automorphism* $v \mapsto {}^tv$ *of order* 2 *such that* ${}^t(vw) = {}^tw\,{}^tv$ *for all* $v, w \in \mathscr{A}$. *We let* **Sym** *be the subspace of* $\mathscr{A}$ *consisting of the symmetric elements, i.e.* v *such that* $v = {}^tv$. *We suppose given a* **trace**, *that is a functional*

$$\mathrm{tr}\colon \mathscr{A} \to \mathbf{R}$$

such that $\mathrm{tr}(vw) = \mathrm{tr}(wv)$ *for all* $v, w \in \mathscr{A}$, *and we asume that* $\mathrm{tr}(w^2) > 0$ *for all* $w \in \mathrm{Sym}$, $w \neq 0$. *Thus the functional gives rise to the* tr-*scalar product*

$$\langle v, w\rangle_{\mathrm{tr}} = \mathrm{tr}(vw), \qquad |v|_{\mathrm{tr}}^2 = \mathrm{tr}(v^2),$$

which is positive definite on Sym. *We shall also assume a Schwarzian property, see below.*

The standard example is when $\mathscr{A} = \mathrm{Mat}_n(\mathbf{R})$. Note that we have the exponential map

$$\exp\colon \mathscr{A} \to \mathscr{A} \qquad \text{given by} \quad \exp(v) = \sum \frac{v^n}{n!} = e^v.$$

We define

$$\mathrm{Pos} = \exp(\mathrm{Sym}).$$

Then every element of Pos has a square root in Pos, namely $p = \exp(v)$ implies $p^{1/2} = \exp(v/2)$.

Let $v, w \in \mathscr{A}$. We define

$$F_v(w) = \exp(-v/2)\exp'(v)w \cdot \exp(-v/2) = e^{-v/2}\exp'(v)we^{-v/2}.$$

Note that

$$\exp'(v)w = \frac{d}{dt}\exp(v+tw)\bigg|_{t=0}.$$

Directly from the definitions, we get

(1) $$\exp'(v)w = \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{r+s=n-1} v^r w v^s.$$

Since powers of an element commute with each other, we note that $\exp(-v/2)$ commutes with powers v^r, v^s.

Lemma 2.1. *The maps F_v and $\exp'(v)$ are hermitian with respect to the tr-scalar product on $\mathscr{A}$. If $v \in \mathrm{Sym}$, then F_v and $\exp'(v)$ map Sym into itself.*

Proof. A routine verification gives for $u, v, w \in \mathscr{A}$:

$$\begin{aligned}\operatorname{tr}(F_v(w)u) &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{r+s=n-1}\operatorname{tr}(\exp(-v/2)v^r w v^s \exp(-v/2)u)\\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\sum_{r+s=n-1}\operatorname{tr}(wv^s\exp(-v/2)u\exp(-v/2)v^r)\\ &= \operatorname{tr}(wF_v(v))\end{aligned}$$

because $\exp(-v/2)$ commutes with v^r and v^s. This concludes the proof that F_v is hermitian with respect to the tr-scalar product. If $v \in \mathrm{Sym}$, then formula (1) shows that F_v maps Sym into itself. The statements about $\exp'(v)$ follow the same pattern of proof.

We define $L_v\colon \mathscr{A} \to \mathscr{A}$ to be left multiplication, $L_v(w) = vw$, and R_v is right multiplication. We let $D_v = L_v - R_v$, so

$$D_v(w) = vw - wv = [v, w] \quad \text{for all} \quad v, w \in \mathscr{A}.$$

Lemma 2.2. *Let $v \in \mathrm{Sym}$. Then D_v^2 is hermitian on* Sym.

Proof. Again this is routine, namely:

$$\begin{aligned} D_v(w) &= vw - wv,\\ D_v^2(w) &= v^2w - 2vwv + wv^2,\\ (D_v^2w)u &= v^2wu - 2vwvu + wv^2u,\\ wD_v^2u &= wv^2u - 2wvuv + wuv^2.\end{aligned}$$

Applying tr to these last two expressions and using its basic property $\operatorname{tr}(xy) = \operatorname{tr}(yx)$ yields the proof of the lemma.

We recall that a hermitian operator B on Hilbert space is called **semipositive**, written $B \geqq 0$, if we have $\langle Bw, w\rangle \geqq 0$ for all $w \neq 0$ in the Hilbert space. Then one defines $B_1 \geqq B_2$ if $B_1 - B_2 \geqq 0$.

In the proofs that follow, we shall use two basic properties.

Spectral Property. *Let M be a symmetric linear map of a finite dimensional vector space over* **R**, *with a positive definite scalar product. Let $b \geqq 0$. Let $f_0(t)$ be a convergent power series such that $f_0(t) \geqq b$ for all t in an interval containing the eigenvalues of M. Then $f_0(M) \geqq bI$.*

Proof. Immediate by diagonalizing the linear map with respect to a basis. Of course, the Appendix proves the analogous property in Hilbert space.

We also **assume**:

Schwarzian Property. *For all $v, w \in \mathrm{Sym}$,*

$$\operatorname{tr}((vw)^2) \leqq \operatorname{tr}(v^2w^2).$$

For the convenience of the reader, we recall the proof in the cases of matrices. The matrices (linear maps) can be simultaneously diagonalized, if one of them is positive definite, and in that case the inequality amounts to the usual Schwartz inequality. If both matrices are singular, then one can consider a matrix $w + \epsilon e$ with the identity matrix e, and $\epsilon > 0$. Then $w + \epsilon e$ is non-singular for all sufficiently small $\epsilon \neq 0$, and one can then use the preceding non-singular case, followed by taking a limit as $\epsilon \to 0$. This concludes the proof.

We define a formal power series

$$f(t) = \sum_{k=0}^{\infty}\frac{(t/2)^{2k}}{(2k+1)!} = \frac{\sinh t/2}{t/2} = \frac{\exp(t/2) - \exp(-t/2)}{t}.$$

We note that L_v and R_v commute with each other, and so

$$\exp(D_v/2) = \exp(L_v/2)\exp(-R_v/2).$$

We may take $f(D_v)$. Since only even powers of D_v occur in the power series for f, it follows that if $v \in \mathrm{Sym}$, then $f(D_v)$ maps Sym into itself, and the operator

$$f(D_v)\colon \mathrm{Sym} \to \mathrm{Sym}$$

is hermitian for the tr-scalar product.

Lemma 2.3. *For any $v \in \mathscr{A}$, we have $D_v F_v = D_v f(D_v)$.*

Proof. Let $t \mapsto x(t)$ be a smooth curve in $\mathscr{A}$. Then

$$x(\exp x) = (\exp x)x.$$

Differentiating both sides gives

$$x' \exp x + x(\exp x)' = (\exp x)'x + (\exp x)x',$$

and therefore

$$x' \exp x - (\exp x)x' = (\exp x)'x - x(\exp x)'.$$

Multiplying on the left and right by $\exp(-x/2)$, and using the fact that x commutes with $\exp(-x/2)$ yields

(2)
$$\begin{aligned}\exp(-x/2)x'\exp(x/2) &- \exp(x/2)x'\exp(-x/2)\\ &= \exp(-x/2)(\exp x)'\exp(-x/2)x\\ &\quad - x\exp(-x/2)(\exp x)'\exp(-x/2).\end{aligned}$$

Since L_x and R_x commute, we have

$$\exp(D_x/2) = \exp(L_x/2)\exp(-R_x/2),$$

so (2) can be written in the form

(3)
$$(\exp(D_x/2) - \exp(-D_x/2))x' = D_x F_x x'.$$

We now take the curve $x(t) = v + tw$, and evaluate the preceding identity at $t = 0$, so $x'(0) = w$, to conclude the proof of the lemma.

Theorem 2.4. *Let $v \in \mathrm{Sym}$. Then $F_v = f(D_v)$ on* Sym. *Hence for $w \in \mathrm{Sym}$, we have*

$$\exp'(v)w = \exp(v/2)\cdot f(D_v)w\cdot\exp(v/2).$$

Proof. Let $h_v = F_v - f(D_v)$. Then $h_v\colon \mathrm{Sym} \to \mathrm{Sym}$ is hermitian, and its image is contained in the subspace $E = \mathrm{Ker}\, D_v \cap \mathrm{Sym}$. Since Sym is assumed finite dimensional, it is the direct sum of E and its orthogonal complement $E^\perp$ in Sym. Since h_v is hermitian, it maps $E^\perp$ into $E^\perp$, but h_v also maps $E^\perp$ into E, so $h_v = 0$ on $E^\perp$. In addition, E is the commutant of v in Sym, and hence $f(D_v) = \mathrm{id} = E_v$ on E, so $h_v = 0$ on E. Hence $h_v = 0$ on Sym, thus concluding the proof of the theorem.

Theorem 2.5. *Let $v \in \mathrm{Sym}$. Then D_v^2 is semipositive, and $f(D_v) \geqq I$.*

Proof. By Lemma 2.2, for $w \in \mathrm{Sym}$ we have

$$\begin{aligned}\langle D_v^2 w, w\rangle_{\mathrm{tr}} &= \mathrm{tr}(wv^2w - 2vwvw + v^2w^2)\\ &= 2\,\mathrm{tr}(v^2w^2 - (vw)^2).\end{aligned}$$

Thus the semipositivity of D_v^2 results from the Schwarzian property of tr. Now we can write

$$f(t) = f_0(t^2)$$

where $f_0(t)$ is the obvious power series, which has positive coefficients. Note that $f_0(t) \geqq 1$ for all $t \geqq 0$. Therefore by the spectral property of power series, it follows that

$$f(D_v) = f_0(D_v^2) \geqq I.$$

This concludes the proof.

Theorem 2.6. *The exponential map* exp *is* tr-*norm semi-increasing on* Sym, *that is for all v, $w \in \mathrm{Sym}$, putting $p = \exp(v)$, we have*

$$|w|_{\mathrm{tr}}^2 = \mathrm{tr}(w^2) \leqq \mathrm{tr}((p^{-1}\exp'(v)w)^2) = |\exp'(v)w|_{p,\mathrm{tr}}^2.$$

Proof. The right side of the above inequality is equal to

$$\begin{aligned}\mathrm{tr}((p^{-1}\exp'(v)w)^2) &= \mathrm{tr}\big((\exp(-v/2)\cdot\exp'(v)w\cdot\exp(-v/2)\big)^2\\ &= \mathrm{tr}\big(F_v(w)^2\big)\\ &= |f(D_v)w|_{\mathrm{tr}}^2 \quad \text{by Theorem 2.4.}\end{aligned}$$

Applying Theorem 2.5 now concludes the proof.

Corollary 2.7. *For each* $v \in \mathrm{Sym}$, *the maps*

$$F_v \quad \textit{and} \quad \exp'(v): \mathrm{Sym} \to \mathrm{Sym}$$

are linear automorphisms.

Proof. Theorem 2.6 shows that Ker $\exp'(v) = 0$, and $\exp'(v)$ is a linear isomorphism. The statement for F_v then follows because F_v is composed of $\exp'(v)$ and multiplicative translations by invertible elements in Sym. This concludes the proof.

We note that Theorem 2.6 concludes the proof of Theorem 1.2.

Remark. There is no trace on the full algebra of endomorphisms of an infinite dimensional Hilbert space, satisfying the conditions listed at the beginning of the section. There are such traces on some other types of infinite dimensional algebras with units. Lemmas 2.1, 2.2 and 2.3 don't depend on anything more. However, Theorem 2.4 depends on an additional **hilbertian property**, namely that there exists a constant $C > 0$ such that

$$\mathrm{tr}(v^2) \geqq C|v|^2 \qquad \text{for all} \quad v \in \mathrm{Sym},$$

where $|v|$ is the original Banach norm. Since all norms on a finite dimensional vector space are equivalent, this condition is valid in finite dimension. But as Rieffel has informed me, if this condition is satisfied, and so makes $\mathscr{A}$ into a Hilbert space, then necessarily $\mathscr{A}$ is finite dimensional. There remains to be seen whether there are natural infinite dimensional cases where a weaker condition is still true so that consequences of this property, similar to Theorem 2.5 and 2.6, are valid in some sense, without $\mathscr{A}$ being complete for the Hilbert trace norm.

XII, §3. TOTALLY GEODESIC AND SYMMETRIC SUBMANIFOLDS

We continue with the same notation as in the preceding section. We follow Mostow's exposition of Cartan's work, as before [Mo 53]. It can be shown that all finite dimensional symmetric spaces of non-compact type are symmetric submanifolds of Pos_n for some n, depending on the choice of a suitable representation. Thus Theorems 3.3, 3.5, 3.7 and 3.9 below apply quite generally.

We abbreviate some mappings which occur frequently as follows. For $v \in \mathrm{Sym}$:

$$J_v = \exp'(v) \quad \text{(the Jacobian of the exponential map)},$$
$$A_v = R_p + L_p \qquad \text{where} \quad p = \exp v = e^v,$$
$$\text{so} \quad A_v(w) = e^v w + w e^v.$$

At the moment, we do not yet know that A_v is a linear isomorphism. To prove this, we shall use another function besides f, namely we let

$$g(t) = 2f(t)^{-1}\cosh(t/2) = t\coth(t/2).$$

Since $g(t) \to \infty$ as $t \to \pm\infty$, $g(t) \to 1$ as $t \to 0$, and $g(t) > 0$ for all $t \in \mathbf{R}$, it follows that g is a continuous function on every closed interval $[0, c]$, and is bounded away from 0 and ∞ on this closed interval. Furthermore, $g(t) = g_0(t^2)$ where g_0 satisfies $g_0 \geqq \epsilon > 0$ on $[0, c]$, for some $\epsilon > 0$.

Theorem 3.1. *The Hermitian operator* A_v *is invertible on* Sym. *Furthermore, we have the formula*

$$A_v = \exp'(v) g_0(D_v^2) \quad \textit{on } \mathrm{Sym}.$$

Proof. From the definitions and Theorem 2.4, we know that

$$J_v = \exp'(v) = \exp(L_v/2)\exp(R_v/2) f(D_v) \quad \text{on Sym}.$$

Note that $\exp L_v = L_p$ and $\exp R_v = R_p$. Abbreviate $L = L_v$, $R = R_v$, $D = L - R$. By Corollary 2.7, we find

$$\begin{aligned} J_v^{-1} A_v = J_v^{-1}(e^R + e^L) &= f(D)^{-1} e^{-L/2} e^{-R/2}(e^L + e^R) \\ &= f(D)^{-1}(e^{(L-R)/2} + e^{(R-L)/2}) \\ &= f(D)^{-1}(e^{D/2} + e^{-D/2}) \\ &= 2f(D)^{-1}\cosh(D/2) = g(D), \end{aligned}$$

which proves the formula. Now from the fact that g_0 is bounded away from 0, strictly positive on an interval $[0, c]$ such that $0 \leqq D_v^2 \leqq cI$, we deduce the invertibility, and conclude the proof of the theorem.

The next considerations will depend on the existence of symmetries, so the present context may also be viewed as an example of symmetric spaces, which will be defined in general in Chapter XIII.

Let $p \in \text{Pos}$, and consider the mappings of Pos $\to$ Pos given by:

$$S(x) = x^{-1}, \qquad B_p(x) = pxp, \qquad S_p(x) = px^{-1}p, \qquad \text{so} \quad S_p^2 = \text{id}.$$

One calls S_p the **Cartan symmetry** and we have $S_p = B_p \circ S$. We know that B_p is an isometry, and we claim that S is also an isometry. Indeed, the differentials of S and B_p are given by

$$S'(p)w = -p^{-1}wp^{-1}, \qquad B_p'(v)w = pwp.$$

as one verifies directly from the definitions. This is a pure Banach algebra relation. Then the isometry relation

$$|S'(p)w|_{p^{-1}}^2 = |w|_p^2 \qquad \text{for} \quad p \in \text{Pos}, \quad w \in \text{Sym},$$

is immediate from the definitions of the trace and the scalar product. It follows that S_p is an isometry, being composed of isometries. In addition, we note that $S_p(p) = p$, that is, p is a fixed point of S_p.

We have $S_p'(p) = -\text{id}$.

Proof. Immediate from the chain rule $S_p'(p) = B_p'(S(p))S'(p)$.

The above properties show that S_p is a symmetry in the sense defined generally in the next chapter.

We shall study submanifolds, both in the Lie vector space Sym and in the symmetric space Pos = exp Sym. So let V be a vector subspace of Sym, and let $X = \exp(V)$. Note that if $y = \exp(w) \in X$, then $y^{-1} = \exp(-w)$ is also in X, so X is stable under the map $y \mapsto y^{-1}$. By a **symmetric submanifold** of Pos, we mean a submanifold of the form $X = \exp(V)$ such that X satisfies the condition

SYM 1. $\qquad x, y \in X$ implies $xyx \in X$.

In other words, X is stable under the operation $(x, y) \mapsto xyx$. Observe that this condition is equivalent with the condition that S_x leaves X stable for all $x \in X$, i.e. X is stable under all Cartan symmetries with $x \in X$.

Example. Let Sym_n be the standard space of symmetric $n \times n$ real matrices, and let V be the subspace of matrices with trace 0. Thus $X = \exp U$ consists of the positive definite symmetric matrices with determinant 1, which is symmetric. As we shall we below, it follows that X is a totally geodesic submanifold of Pos_n, usually denoted by SPos_n (the special positive elements).

Let $X = \exp(V)$ be a symmetric submanifold. Then for each $y \in X$ we have the operator

$$[y]: X \to X \qquad \text{defined by} \qquad [y]x = yxy.$$

Lemma 3.2. *Suppose $X = \exp(V)$ symmetric. Given $p, q \in X$ there exists $y \in X$ such that $ypy = q$. In other words, X acts transitively on itself.*

Proof. The condition $ypy = q$ is equivalent with

$$\begin{aligned} p^{1/2}yp^{1/2}p^{1/2}yp^{1/2} = p^{1/2}qp^{1/2} \quad &\Leftrightarrow \quad (p^{1/2}yp^{1/2})^2 = p^{1/2}qp^{1/2}, \\ &\Leftrightarrow \quad p^{1/2}yp^{1/2} = (p^{1/2}qp^{1/2})^{1/2}, \\ &\Leftrightarrow \quad y = p^{-1/2}(p^{1/2}qp^{1/2})^{1/2}p^{-1/2}, \end{aligned}$$

which concludes the proof.

Note. A similar proof shows that given $p, q \in X$ there exists $y \in X$ such that $ypy = q^{-1}$ and thus also $yqy = p^{-1}$. Written in terms of the operator $[y]$, these read

$$[y]p = q^{-1} \qquad \text{and} \qquad [y]q = p^{-1}.$$

We shall now describe equivalent conditions for a manifold to be symmetric. First we derive a formal relation about the exponential on Sym.

For all $w \in \mathcal{A}$ we have

$$\frac{d}{dt}\exp(tw) = \exp'(tw)w = e^{tw}w.$$

This follows at once from the definition of the differential, and the fact that all elements with which we operate commute with each other, so one can take the derivative of $\exp(tw)$ in the usual way from ordinary calculus. Now given $x = e^w$ and $p \in \text{Pos}$, we can define a curve $\xi(t)$ in Sym by the formula

$$\exp \xi(t) = e^{tw}pe^{tw}.$$

Note that $\xi(0) = \log p$ and $\exp(\xi(1)) = xpx$. Differentiating with respect to t, we obtain

$$\begin{aligned} \exp'(\xi(t))\xi'(t) &= e^{tw}wpe^{tw} + e^{tw}pt^{tw}w \\ &= A_{\xi(t)}(w) \end{aligned}$$

by the definition of A_v and the first observation in the proof. It follows from Theorem 3.1 that

$$\begin{aligned}\xi'(t) &= \exp'(\xi(t))^{-1}A_{\xi(t)}(w)\\ &= f(D^2_{\xi(t)}/2)^{-1}\cosh(D_{\xi(t)}/2)(w)\\ (1)\qquad &= g_0(D^2_{\xi(t)})(w).\end{aligned}$$

Theorem 3.3. *Let V be a vector subspace of* Sym, *and let $X = \exp(V)$. Then X is a symmetric submanifold of* Pos *if and only if*:

SYM 2. *The map D_v^2 maps V into itself for all $v \in V$.*

Proof. Suppose that D_u^2 maps V into itself for all $u \in V$. Note that g_0 is actually real analytic, and the above equation is an ordinary differential equation for $\xi(t)$ in V. It has a unique solution with initial condition $\xi(0) = \log p$, and of course, this solution lies in V, that is $\xi(t) \in V$ for all t. Taking $t = 1$ shows that $xpx \in \exp(V)$, thus proving one implication.

Conversely, assume that $x, p \in \exp(V)$ implies $xpx \in \exp(V)$. Let w be as before, and also $\xi(t)$ as before, with say $\xi(0) = v = \log p$. We have to show $D_v^2(w) \in V$. By assumption, ξ is a curve in V, and hence so is ξ', which we computed above, with the power series f of §2. Thus $\xi'(t)$ is a power series in t, whose coefficients lie in V. The coefficient of t^2 is directly computed to be

$$\tfrac{1}{12}D^2_{\xi(0)}(w) \in V,$$

thus completing the proof of the theorem.

Remark 3.4. The condition that V is stable under D_v^2 for all $v \in V$ is actually a Lie algebra condition, because in an arbitrary Lie algebra one may define D_v by

$$D_v(w) = [v, w].$$

One may then use the following purely Lie algebraic result to get an equivalent condition.

Lemma 3.5. *Let L be a Lie algebra and V a linear subspace. Then V is stable under D_v^2 for all $v \in V$ if and only if V is stable under all operators D_uD_v with $u, v \in V$.*

Proof. Applying the hypothesis that D_v^2 leaves V stable to $u + v$ (polarization) shows that $D_uD_v + D_vD_u$ leaves V stable, or in other words,

$$(*)\qquad [u, [v, w]] + [v, [u, w]] \in V \quad \text{for all} \quad u, v, w \in V.$$

From $D_{[u,v]} = D_uD_v - D_vD_u$, we see that $D_{[u,v]} + 2D_vD_u$ leaves V stable, that is

$$(**)\qquad [[u, v], w] + 2[v, [u, w]] \in V.$$

Interchanging u and w in $(*)$ shows that $[[u, v], w] + [v, [u, w]] \in V$. Combining this with $(**)$ proves the lemma. Note that proof is valid for a Lie algebra over any commutative ring.

A linear subspace of a Lie algebra which satisfies the property that for all u, v, w in the subspace the element $[u, [v, w]]$ lies in the subspace is called a **Lie triple system.** Thus the lemma implies

Theorem 3.6. *A submanifold $X = \exp(V)$ of* Pos *is symmetric if and only if V is a Lie triple system.*

The previous theorem established an equivalent between a Lie property and the symmetry property of the submanifold. The next theorem gives another equivalent condition in terms of geodesics.

We say that $X = \exp(V)$ is a **geodesic submanifold** if given two points $x, y \in X$, the geodesic between these two points lies in X.

Theorem 3.7. *Let $X = \exp(V)$. Then X is a geodesic submanifold if and only if X satisfies the* (*equivalent*) *conditions of Theorem* 3.3, *e.g. X is a symmetric submanifold.*

Proof. Assume X is symmetric. The image of the line through 0 and an element $v \in V$, $v \neq 0$ is a geodesic which is contained in X. Since the maps $x \mapsto yxy$ (for $y \in X$) leave X stable, and act transitively on X, it follows that X contains the geodesic between any two of its points. Conversely, assume X is a geodesic submanifold. Let $x \in X$, $v \in V$. Then S_x maps the geodesic $x^{1/2}\exp(tv)x^{1/2}$ to $x^{1/2}\exp(-tv)x^{1/2}$, and so this geodesic is stable under S_x (as a submanifold). Hence S_x maps X into itself, so X is symmetric, thus concluding the proof.

Examples. Let $\mathscr{A} = \text{Mat}_n(\mathbf{R})$ and $\text{Pos} = \text{Pos}_n$ the space of symmetric positive definite matrices. Let A be the submanifold of diagonal matrices with positive diagonal components. Then A is totally geodesic, as one sees by taking $V =$ vector space of all diagonal matrices. The bracket of two elements in V is 0, so V trivially satisfies the criterion of Theorem 3.6. One usually denotes V by $\mathfrak{a}$. The orthogonal complement of $\mathfrak{a}$ for the trace form is immediately determined to be $\text{Sym}_n^{(0)} = \mathfrak{a}^{\perp}$, consisting of the matrices with zero diagonal components. We now obtain an example of the global tubular neighborhood theorem for Cartan–Hadamard spaces, by applying Theorem 4.4 of Chapter X to the present case.

Theorem 3.8. *Let A be the group of diagonal $n \times n$ matrices with positive diagonal components, and let $\mathfrak{a} = \mathrm{Lie}(A)$ be the vector space of diagonal matrices. Let $\mathfrak{a}^{\perp} = \mathrm{Sym}_n^{(0)}$ be the space of matrices with zero diagonal components. Then the map*

$$A \times \mathfrak{a}^{\perp} \to \mathrm{Pos}_n \qquad \text{given by} \quad (a, v) \mapsto [a]\exp(v)$$

is a differential isomorphism.

Similarly, instead of A, one could consider other totally geodesic submanifolds obtained as follows. Given positive integers n_i $(i = 1, \ldots, m)$ such that $\sum n_i = n$, we let V be the subspace of Sym consisting of diagonal blocks of dimensions $n_1, \ldots, n_m$. Then V is a Lie triple system, and $\exp(V)$ is totally geodesic.

Finally, the Riemann tensor R can be described explicitly as follows.

Theorem 3.9. *Let R be the Riemann tensor. Then at the unit element e, with $u, v, w \in T_e(\mathrm{Pos}) = \mathrm{Sym}$, we have*

$$R(v, w)u = -[[v, w], u],$$

and $R_2(v, w) = \langle R(v, w)v, w\rangle_{\mathrm{tr}} \geqq 0.$

Proof. Assume the formula for the 4-tensor R. Substituting $u = v$ and taking the tr-scalar product immediately shows that

$$\langle R(v, w)v, w\rangle_{\mathrm{tr}} = -2\,\mathrm{tr}\big((vw)^2 - v^2w^2\big).$$

Hence the semipositivity of R_2 comes from the Schwarzian property. So there remains to prove the formula for R. But this is a special case of a formula which holds much more generally for Killing fields, since for symmetric spaces, we know that $\mathfrak{m}_e = T_e$, see Chapter XIII, Theorem 5.8 and Theorem 4.6.

CHAPTER XIII

Automorphisms and Symmetries

By a **covariant derivative** D on a manifold X we shall always mean a covariant derivative associated with a spray. Thus for each vector field ξ, the association $\eta \mapsto D_\eta \xi$ is a tensor, and D could therefore be called a **tensorial derivative**. As Wu informed me, Hermann Weyl was the first to point out the importance of this notion, independently of a metric [Wey 18]. A pair (X, D) consisting of a manifold and such a covariant derivative will be called a ***D*-manifold**. We also note that D is often called a **connection** in the literature, or an **affine connection**, following Hermann Weyl.

The curvature involved a second derivative, and we went immediately into it without stopping to consider the second derivative for its own sake. We now do so in a first section. The second derivative is even more important than in ordinary calculus, and we shall see several applications, both in this chapter and the next. Among other things, it is used to define the Laplace operator in §1. We also give the formula relating it to the Riemann tensor.

The second derivative is an operator discovered in certain contexts by Killing, and Karcher pointed out to me that this operator is tensorial in its arguments. The **Killing operator** is defined for two vector fields by

$$D_\eta D_\zeta - D_{D_\eta \zeta} = Q(\eta, \zeta).$$

We relate it to D^2 in §1.

The rest of this chapter deals with the context of D-manifolds and their automorphisms.

Let (X, D^X), (Y, D^Y) be D-manifolds, which we write more simply (X, D) and (Y, D). An isomorphism ρ: $X \to Y$ (differential) is said to be a **D-isomorphism** if $\rho^*D = D$. The pull back ρ^*D is defined by functoriality, so by the formula

$$(\rho^*D)_{\rho^*\zeta}(\rho^*\eta) = \rho^*(D_\zeta\eta)$$

for all vector fields η, ζ on Y. Recall that for any vector field ξ on Y,

$$(\rho^*\xi)(x) = T\rho(x)^{-1}\xi(\rho(x)), \qquad \text{i.e.} \quad \rho^*\xi = T\rho^{-1}\cdot(\xi\circ\rho).$$

A **D-automorphism** of X is just an isomorphism $(X, D) \to (X, D)$ (same D).

Since a spray is uniquely determined by its associated covariant derivative, it follows that a D-automorphism is also an automorphism for the spray.

Suppose X is pseudo Riemannian with metric g. Then there is a unique covariant derivative called the metric or Levi-Civita, or pseudo Riemannian derivative (connection) associated with g. In one important example, many metrics may have the same covariant derivative: any two positive definite scalar products on a finite dimensional vector space have the same covariant derivative, namely the ordinary one. A metric automorphism (i.e. an isometry) of X induces a D-automorphism, but the converse is not necessarily true. A number of properties of metric automorphisms actually depend only on their being D-automorphisms, and I think it clarifies matters considerably to obtain certain results as special cases of the more general results concerning D-automorphisms, and to lay the foundations in general, especially since some manifolds have a covariant derivative for which no Riemannian metric exists for which D is the metric derivative.

When such a Riemannian metric exists, then Kobayashi has brought to my attention some important facts. Let D be the metric derivative associated to g. Then the two notions of D-automorphism and g-automorphism coincide in the following cases of finite dimensional manifolds:

When the manifold is compact (due to Yano).
When the manifold is complete "irreducible" (due to Kobayashi).

Cf. [KoN 63], p. 242. Furthermore, as Kobayashi also informed me, de Rham's holonomy decomposition theorem states that the universal covering space of a Riemannian manifold is metrically isomorphic to a product of a euclidean space and irreducible manifolds. Thus in the finite dimensional case, when D is a metric derivative, D-automorphisms are fairly close to being affine transformations of a euclidean space, combined with g-automorphisms.

Quite generally, let ξ be a vector field and let $\{\rho_t\}$ be its flow. We define ξ to be **D-Killing** (resp. **metric** or **g-Killing** in the pseudo Riemannian case) if each $\rho_t\colon X \to X$ is a D-automorphism (resp. a metric automorphism). The definition is usually given only in the pseudo Riemannian case. A D-automorphism has been called an **affine transformation.** However, I find it more appropriate to unify and functorialize the terminology. We denote by $\mathrm{Kill}_D(x)$ and $\mathrm{Kill}_g(X)$ the set of D-Killing fields and metric Killing fields respectively. Each one will be seen to be a vector space over the reals, and we have the inclusion

$$\mathrm{Kill}_D(X) \supset \mathrm{Kill}_g(X)$$

if D is the metric derivative associated with g.

For the rest of this chapter, by a Killing field we shall mean a D-Killing field unless otherwise specified.

It turns out that Killing fields can be characterized by a second-order differential equation due to Killing [Ki 1891], namely

$$D_\eta D_\zeta\xi - D_{D_\eta\zeta}\xi = R(\eta, \xi)\zeta \qquad \text{for all} \quad \eta, \zeta.$$

Helgason gave me the above reference after looking into the literature, when I expressed interest in the history of this equation. Helgason also pointed out that the equation was subsequently referred to in Cartan [Ca 51], Section 5, especially p. 328; and Eisenhart [Ei 26], p. 247, formula 71.1. It has usually been discussed only in the Riemannian context, not only in the older literature, but more recently in Klingenberg [Kl 83/95] and Sakai [Ca 96]. However, a form of it is given in [KoN 63], Chapter VI, Proposition 2.6, as was pointed out to me by Kobayashi. I would not have recognized it otherwise.

In §2, we deal with the characterization of Killing fields by the Killing equation, and in §3 we discuss the pseudo Riemannian case. In §4, we list some Lie algebra properties of Killing fields. In §5 we introduce Cartan's symmetries, and describe some of their implications for Killing fields. In §6 we give further properties of symmetric spaces. I originally learned some of the material from Klingenberg [Kl 83/95], whose approach I liked very much. However, Klingenberg limits himself to the Riemannian case, whereas we work in the general situation of an arbitrary covariant derivative (connection), since a Riemannian or pseudo Riemannian hypothesis is unnecessary for the main results. Klingenberg does a very nice and beautiful job. He even states in Chapter I: "... we consider manifolds modelled on Hilbert spaces rather than on finite dimensional spaces. This will be useful in Chapter 2 and presents no conceptual difficulties anyway, as was demonstrated by Lang [1]." However, in his Chapter 2, he finks out, by considering symmetric spaces (for instance) only in the finite dimensional context. About this, he says in the Preface to the book: "In

Chapter 2, entitled *Curvature and Topology*, I restrict myself to finite dimensional manifolds, because the local compactness of the manifold is needed." In fact, the hypothesis of local compactness is needed only in some cases, notably involving positive curvature, but it is definitely not needed for others. Klingenberg directed his book to certain applications having to do with the existence of closed geodesics and pinching, and he gives priority to the Hopf–Rinow theorem, which is the only exception to the general principle that all basic results of differential geometry hold in the infinite dimensional case.

Thus Klingenberg makes unnecessary assumptions about finite dimension when they are not needed, for certain results concerning Killing fields. What is needed on some occasions is that the exponential map at a point (or at all points) is surjective. The Hopf–Rinow theorem guarantees this property in finite dimension, but as an assumption, the surjectivity of the exponential map is weaker than finite dimension since it includes the infinite dimensional case of seminegative curvature, when the Cartan–Hadamard theorem and its corollaries are valid. Thus I found the above assumption the most natural one to make.

In addition, some results of Klingenberg's Chapter 2 are given only in the context of symmetric spaces but they are valid for arbitrary D-manifolds, without any further assumption concerning the existence of symmetries. Furthermore, this validity gave rise to a question by Helgason: to what extent can the Lie algebra of Killing fields, or the Lie subalgebra associated with a certain subspace (denoted by $\mathfrak{h}_p + \mathfrak{m}_p$), be integrated inside an arbitrary manifold so that a manifold may contain in some sense a maximal symmetric submanifold (locally at a point p)? Thus a systematic analysis of proofs in the Riemannian case, and the elimination of superfluous hypotheses, actually suggested further topics of research.

In any case, a new exposition of the material in this chapter was in order on several counts, among which: to deal with arbitrary manifolds with a covariant derivative, not just Riemannian or pseudo Riemannian manifolds; to include the infinite dimensional case; and to free the general theory of Killing fields from the context of symmetric spaces.

XIII, §1. THE TENSORIAL SECOND DERIVATIVE

We begin with some remarks on the first derivative. Let D be a covariant derivative. We have used D essentially with a subscript, such as D_η, applied to various tensors (vector fields, forms, etc.). However, it will now be essential to apply D without a subscript. For instance if f is a function, then Df is a 1-form, defined on a vector field η by

$$(Df)(\eta) = D_\eta f = \eta \cdot f.$$

On a vector field ξ, we have

$$(D\xi)(\eta) = D_\eta \xi.$$

If ω is an r-multilinear tensor, **R** or TX-valued, we have

$$(D\omega)(\eta) = D_\eta \omega;$$

so $D\omega$ is an $(r+1)$-multilinear tensor. And so forth.

Let (X, D) be a D-manifold. For vector fields η, ζ we define the **Killing operator**, or **second tensorial derivative** $Q(\eta, \zeta)$ to be

$$Q(\eta, \zeta) = D_\eta D_\zeta - D_{D_\eta \zeta}.$$

We shall now see that this operator amounts to the second derivative, and we discuss it systematically on functions, vector fields, and multilinear tensors.

On Functions. We start with functions. In ordinary calculus on vector spaces, if f is a C^2 function on an open set in a Banach space, then the second derivative $f''(x)$ is a symmetric bilinear form called the Hessian of f at x. We shall now consider the Hessian, and more generally the second derivative D^2 on functions. In this case, we call $Q(\eta, \zeta)$ the **tensorial Hessian**, or ***D*-Hessian**, or simply the **Hessian**. Let f be a function. Then $Df = df$ is a 1-form. We claim that

$$(1) \qquad (D^2 f)(\eta, \zeta) = Q(\eta, \zeta) f.$$

To see this, note that if ω is a 1-form, then

$$(D\omega)(\eta, \zeta) = (D_\eta \omega)(\zeta) = D_\eta(\omega(\zeta)) - \omega(D_\eta \zeta)$$

so

$$(D^2 f)(\eta, \zeta) = D(Df)(\eta, \zeta) = D_\eta(D_\zeta f) - (Df)(D_\eta \zeta)$$
$$= D_\eta D_\zeta f - (D_\eta \zeta) \cdot f = Q(\eta, \zeta) f,$$

thus proving the claim.

Let η, ζ denote the representations of the vector fields in a chart, and let B be the symmetric bilinear map whose quadratic map represents the spray. Then it follows immediately from the above definition that we have the

Local Represention. We have

$$Q(\eta, \zeta) f(x) = f''(x)(\eta(x), \zeta(x)) + f'(x) B(x)(\eta(x), \zeta(x)),$$

or omitting the x,

$$Q(\eta, \zeta)f = f'' \cdot \eta \cdot \zeta + f' \cdot B(\eta, \zeta). \tag{2}$$

Thus $D^2 f$ is symmetric, or in other words, $Q(\eta, \zeta)$ is symmetric in (η, ζ) as an operator on functions. Warning: it is not necessarily symmetric on higher degree forms or vector fields, cf. Proposition 1.3.

Representation Along a Geodesic. *Let $x \in X$ and let $v \in T_x X$. Let α be the geodesic such that $\alpha(0) = x$ and $\alpha'(0) = v$. Then*

$$D^2 f(v, v) = \left(\frac{d}{dt}\right)^2 f(\alpha(t))\Big|_{t=0}. \tag{3}$$

This simply comes by looking at the Killing operator along the geodesic, so that

$$Q(\alpha', \alpha') = D_{\alpha'} D_{\alpha'} - D_{D_{\alpha'}\alpha'}.$$

Since α is a geodesic, it follows that the second term on the right is 0, and the desired formula comes out, since the covariant derivative on functions is just the Lie derivative.

In the pseudo Riemannian case, we can give additional information about the Hessian in terms of the scalar product, as follows.

Theorem 1.1. *Suppose X is pseudo Riemannian and D is the metric derivative. Let $\operatorname{gr}(\varphi)$ denote the gradient of a function φ. Then*

$$D^2\varphi(\eta, \zeta) = \langle D_\eta \operatorname{gr}(\varphi), \zeta\rangle.$$

In particular, $\langle D_\eta \operatorname{gr}(\varphi), \zeta\rangle$ is symmetric in (η, ζ).

Proof. Let $f = \langle \operatorname{gr}(\varphi), \zeta\rangle = D_\zeta \varphi$. By the definition of the metric derivative,

$$\begin{aligned} D_\eta f &= \langle D_\eta \operatorname{gr}(\varphi), \zeta\rangle + \langle \operatorname{gr}(\varphi), D_\eta \zeta\rangle \\ &= \langle D_\eta \operatorname{gr}(\varphi), \zeta\rangle + (D_\eta \zeta) \cdot \varphi. \end{aligned}$$

On the other hand, by (1),

$$D_\eta f = D_\eta(D_\zeta \varphi) = D^2\varphi(\eta, \zeta) + (D_\eta \zeta) \cdot \varphi,$$

which proves the theorem.

Example (The Laplacian). Let X be finite dimensional. Let φ be a function on X. We define the **Laplacian** on functions to be minus the trace of the Killing operator on functions. Thus if $\{\xi_1, \dots, \xi_n\}$ is an orthonormal frame on X, then

$$\Delta = -\sum_{i=1}^{n} Q(\xi_i, \xi_i).$$

Applied to φ, this reads

$$\begin{aligned} \Delta\varphi = -\operatorname{tr}(D^2\varphi) &= -\sum_{i=1}^{n} D^2\varphi(\xi_i, \xi_i) = -\sum_{i=1}^{n} Q(\xi_i, \xi_i)\varphi. \\ &= -\sum_{i=1}^{n} \langle D_{\xi_i} \operatorname{gr}(\varphi), \xi_i\rangle \end{aligned} \tag{4}$$

by Theorem 1.1.

In Chapter XV, §1 and §2 we shall give another definition of the Laplacian, and Corollary 2.4 of Chapter XV shows that it coincides with the definition we are now using.

In the next chapter, we shall give formulas comparing the Laplacian on submanifolds and in submersions.

On Vector Fields. Let ξ, η, ζ represent vector fields in a chart. The local representation of the Killing operator reads:

$$\begin{aligned} Q(\eta, \zeta)\xi = {} & \xi'' \cdot \zeta \cdot \eta + \xi' \cdot B(\eta, \zeta) - B(\eta, \xi' \cdot \zeta) - B(\xi' \cdot \eta, \zeta) \\ & - (B' \cdot \zeta)(\eta, \xi) - B(B(\eta, \zeta), \xi) + B(B(\eta, \xi), \zeta). \end{aligned} \tag{5}$$

This is analogous to the local expression of the Riemann tensor, cf. Chapter IX, Proposition 1.2. The proof is routine, following the same pattern as in that reference. We simply use the local expression for the iterated derivative $D_\eta D_\zeta$ found there.

Proposition 1.2. *For each vector field ξ, $Q(\eta, \zeta)\xi$ defines a bilinear tensor as a function of (η, ζ). Furthermore, just as with functions, we have*

$$(D^2\xi)(\eta, \zeta) = Q(\eta, \zeta)\xi.$$

Proof. The expression $Q(\eta, \zeta)$ is well defined at each point of X, and the local expression shows that it is a section of the vector bundle of bilinear maps of TX into TX. The formula relating it to D^2 is proved by exactly the same argument as (1). Note that by definition, $(D\xi)(\zeta) = D_\zeta\xi$, so:

$$\begin{aligned}(D^2\xi)(\eta, \zeta) &= D(D\xi)(\eta, \zeta)\\ &= (D_\eta(D\xi))(\zeta)\\ &= D_\eta D_\zeta\xi - (D\xi)(D_\eta\zeta),\end{aligned}$$

which proves the formula.

Proposition 1.3. *For all vector fields* η, ζ *we have*

$$Q(\eta, \zeta) - Q(\zeta, \eta) = R(\eta, \zeta).$$

Proof. This is a short computation, namely:

$$\begin{aligned}Q(\eta, \zeta) - Q(\zeta, \eta) &= D_\eta D_\zeta - D_\zeta D_\eta - D_{D_\eta\zeta} + D_{D_\zeta\eta}\\ &= D_\eta D_\zeta - D_\zeta D_\eta - D_{[\eta,\zeta]}\\ &= R(\eta, \zeta).\end{aligned}$$

This concludes the proof.

On Multilinear Tensors. Let ω be an r-multilinear tensor, i.e. a section of the vector bundle $L^r(TX, E)$, where $E = \mathbf{R}$ or $E = TX$. Then $D\omega$ is an $(r+1)$-multilinear tensor, defined by the contraction

$$(D\omega)(\xi) = D_\xi\omega.$$

If λ is an r-multilinear tensor, we recall the contraction with respect to a vector field η_1, given by

$$\lambda(\eta_1)(\eta_2, \ldots, \eta_r) = \lambda(\eta_1, \eta_2, \ldots, \eta_r).$$

Thus $\lambda(\eta_1)$ is an $(r-1)$-multilinear tensor. We have the analogue of Proposition 1.2 for an r-multilinear tensor, namely:

Proposition 1.4.

$$(D^2\omega)(\eta, \zeta) = Q(\eta, \zeta)\omega.$$

Proof. As before,

$$\begin{aligned}(D^2\omega)(\eta, \zeta) &= (D_\eta(D\omega))(\zeta)\\ &= D_\eta D_\zeta\omega - (D\omega)(D_\eta\zeta)\\ &= D_\eta D_\zeta\omega - D_{D_\eta\zeta}\omega,\end{aligned}$$

which proves the proposition.

I end this section with remarks due to Karcher, from whom I enjoyed learning some differential geometry. They will not be used in the sequel, but I thought they might prove useful to familiarize readers with the tensorial derivative.

Proposition 1.5. *Let A be a tensor field of endomorphisms of TX, i.e. a section of $L(TX, TX)$. As a function of its (η, ζ) variables, $Q(\eta, \zeta)A$ is tensorial. Furthermore, $R(\eta, \zeta)$ is a derivation in the sense that for all vector fields ξ,*

$$R(\eta, \zeta)(A\xi) = (R(\eta, \zeta)A)\xi + AR(\eta, \zeta)\xi.$$

Proof. This follows directly from Proposition 1.3 and the fact that $D_\zeta(A\eta) = (D_\zeta A)\eta + AD_\zeta\eta$, i.e. D_ζ is a derivation.

In addition, the tensorial derivative can be extended inductively to arbitrarily many tensor fields, by the formula

$$Q_n(\eta_n, \eta_{n-1}, \ldots, \eta_1) = D_{\eta_n} \circ Q_{n-1}(\eta_{n-1}, \ldots, \eta_1) - \sum_{j=1}^{n-1} Q_{n-1}(\eta_{n-1}, \ldots, D_{\eta_n}\eta_j, \ldots, \eta_1). \tag{6}$$

Applied with $n = 1$ or 2 to functions or vector fields, one recovers the operators mentioned at the beginning of this section. Furthermore, Q_3 can be used to give another proof of the Bianchi identity in the pseudo Riemannian case. Indeed, using the symmetry of the Hessian, one verifies that

$$Q_3(\xi, \eta, \zeta)f - Q_3(\eta, \zeta, \xi)f = -df \cdot R(\xi, \eta, \zeta)$$

for every function f. It follows that $df \cdot \text{Bianchi}(R) = 0$, whence $\text{Bianchi}(R) = 0$. Finally, observe that in the pseudo Riemannian case, one has the expression analogous to Theorem 1.1, namely

$$Q_3(\xi, \eta, \zeta)f = \langle Q_2(\xi, \eta)\,\text{grad}\, f, \zeta\rangle. \tag{7}$$

Thus the tensorial derivative plays the same role as iterated derivatives in local charts, with its own theory, Taylor's formula, etc.

XIII, §2. ALTERNATIVE DEFINITIONS OF KILLING FIELDS

Throughout, we let (X, D) be a D-manifold.

We shall give several equivalent conditions defining Killing fields. The first one states that the flow of the field consists of D-automorphisms.

Let ξ be a vector field with flow $\{\rho_t\}$. The following properties define ξ to be a **Killing field**, i.e. a ***D*-Killing field**.

Kill 1. The flow preserves D, that is for all t, $\rho_t^* D = D$ on the open set where ρ_t is defined. Equivalently, for all vector fields η, ζ, locally where defined we have

$$T\rho_t \cdot D_\zeta \eta = D_{T\rho_t \cdot \zeta}(d\rho_t \cdot \eta).$$

Kill 2. For all vector fields η, ζ:

$$[\xi, D_\zeta \eta] = D_{[\xi,\zeta]}\eta + D_\zeta[\xi, \eta].$$

Kill 3. The vector field ξ satisfies the Killing differential equation

$$Q(\eta, \zeta)\xi = R(\eta, \xi)\zeta \qquad \text{for all vector fields } \eta, \zeta.$$

Concerning **Kill 2**, we interpret the association $(\zeta, \eta) \mapsto D_\zeta \eta$ as a "product" from pairs of vector fields to vector fields. This product is bilinear with respect to scalar multiplication. The condition **Kill 2** asserts that *bracketing with* ξ (i.e. $\mathscr{L}_\xi$) *is a derivation with respect to this product.* See condition **Kill**$_g$ **2** in §3 for the analogous derivation property in the metric case.

Before proving the equivalence between the three Killing conditions, we formulate a general lemma showing how the bracket product is related to the Killing equation.

Lemma 2.1. *For all vector fields* ξ, η, ζ *we have*

$$[\xi, D_\zeta \eta] = Q(\zeta, \eta)\xi - R(\zeta, \xi)\eta + D_{[\xi,\zeta]}\eta + D_\zeta[\xi, \eta].$$

Proof. This is a short computation as follows:

$$\begin{aligned}
[\xi, D_\zeta \eta] &= D_\xi D_\zeta \eta - D_{D_\zeta \eta}\xi \\
&= D_\zeta D_\xi \eta + D_{[\xi,\zeta]}\eta + R(\xi, \zeta)\eta - D_{D_\zeta \eta}\xi \\
&= D_\zeta D_\eta \xi + D_\zeta[\xi, \eta] + D_{[\xi,\zeta]}\eta + R(\xi, \zeta)\eta - D_{D_\zeta \eta}\xi
\end{aligned}$$

the first step by the definition of the covariant derivative, the second step by the definition of the Riemann tensor, and the third step again by the definition of the covariant derivative. This prove the lemma.

We remark that the lemma gives a natural context for the Killing equation. It shows how bracketing with ξ decomposes into two pieces: one piece exhibits the derivation property, and the other, the Killing piece, exhibits the obstruction. As a result, to prove the equivalence of the conditions, first observe that **Kill 3** is equivalent to **Kill 2** in light of the lemma. For the other equivalence with **Kill 1**, one proceeds as in Sakai [Sa 96], taking the Lie derivative, and using the formulas $\rho_t^*(\xi) = \xi$ and

$$\frac{d}{dt}\rho_t^* \lambda = \rho_t^*(\mathscr{L}_\xi \lambda) \quad \text{for all vector fields } \lambda.$$

We have:

$$\frac{d}{dt}\rho_t^*(D_\zeta \eta) = \rho_t^*[\xi, D_\zeta \eta] = [\rho_t^* \xi, \rho_t^*(D_\zeta \eta)] = [\xi, \rho_t^*(D_\zeta \eta)],$$

$$\frac{d}{dt}\left(D_{\rho_t^*\zeta}(\rho_t^* \eta)\right) = D_{\rho_t^*[\xi,\zeta]}(\rho_t^* \eta) + D_{\rho_t^*\zeta}\left(\rho_t^*[\xi, \eta]\right).$$

Assuming **Kill 1**, we put $t = 0$ to obtain **Kill 2**. Conversely, assume **Kill 2**. Fix $x \in X$, η and ζ. Define a curve β in $T_x X$ by

$$\beta(t) = T\rho_{-t} \cdot D_{d\rho_t \cdot \zeta(x)}(T\rho_t \cdot \eta).$$

It suffices to prove that $\beta'(t) = 0$. But putting $\zeta_t = T\rho_t \cdot \zeta$ and $\eta_t = T\rho_t \cdot \eta$, we get:

$$\begin{aligned}
\beta'(t) &= \frac{d}{ds}\bigg|_{s=0} \beta(t+s) \\
&= \frac{d}{ds}\bigg|_{s=0} T\rho_{-t-s} \cdot D_{T\rho_{t+s}} \cdot \zeta(x)(T\rho_{t+s} \cdot \eta) \\
&= d\rho_{-t}\left\{\frac{d}{ds}\bigg|_{s=0} T\rho_{-s} \cdot D_{T\rho_s} \cdot \zeta_t(x)\left(T\rho_s(\eta_t)\right)\right\}.
\end{aligned}$$

The expression inside { } is 0 essentially because of the computation in the first part of the proof, which shows that **Kill 2** is the infinitesimal property corresponding to **Kill 1**. But now we have obtained

$$\beta'(t) = T\rho_{-t} \cdot 0 = 0,$$

so β is constant, equal to $D_{\zeta(x)}\eta$ which concludes the proof.

Remark 1. Note that conditions **Kill 2** and **Kill 3** are conditions of differential algebra over commutative rings. The formulations and proofs of most basic results in this and the next section depend only on such differential algebra, which means they can be transcribed to algebraic geometric contexts, freed of the real differential geometry.

Remark 2. In [KoN 63], Kobayashi–Nomizu define an operator $A_\xi = \mathscr{L}_\xi - D_\xi$ for every vector field ξ. Then as in their Chapter VI, Proposition 2.5, one has (when "torsion" is 0 as here) $A_\xi \eta = -D_\eta \xi$ for all vector fields ξ, η; and in the subsequent Proposition 2.6, one sees that ξ is Killing if and only if

$$D_\eta(A_\xi) = R(\xi, \eta) \qquad \text{for all vector fields } \eta.$$

Let α be a geodesic of (X, D). If ρ is a D-automorphism of X, then $\rho \circ \alpha$ is a geodesic by first principles, because geodesics like all the rest of the paraphernalia defined in terms of D behave functorially with respect to isomorphisms preserving the covariant derivative.

Proposition 2.2. *A vector field is a Killing field if and only if its restriction to every geodesic is a Jacobi lift of the geodesic.*

Proof. First let α be a geodesic and let ξ be Killing. We shall give two proofs that the restriction of ξ to α is a Jacobi lift of α. We take $\rho = \rho_s$ to be the flow of α, and use **Kill 1**. Put

$$\sigma(s, t) = \rho(s, \alpha(t)).$$

Then $\sigma(s, t)$ is a variation of α through geodesics, and

$$\xi(\alpha(t)) = \partial_1 \sigma(0, t).$$

Thus $\xi(\alpha(t))$ is a Jacobi lift of $\alpha(t)$ by Chapter IX, Proposition 2.8. This gives one proof. For the second proof, recall the Jacobi equation

$$D^2_{\alpha'}(\xi \circ \alpha) = R(\alpha', \xi \circ \alpha)\alpha'.$$

This equation comes out directly from condition **Kill 3** by setting $\eta = \zeta = \alpha'$ over α, and inducing ξ on α. Since $D_{\alpha'}\alpha' = 0$, the term not involving $D^2_{\alpha'}$ becomes 0, and the Jacobi equation drops out from the Killing equation. These proofs are essentially those in [KoN 69] (Vol. II), p. 66, Proposition 1.3.

Next we give Karcher's proof for the converse. Let ξ be a vector field whose restriction to every geodesic is a Jacobi lift, and let α be a geodesic. Then

$$D^2_{\alpha'}(\xi \circ \alpha) = R(\alpha', \xi \circ \alpha)\alpha' = Q(\alpha', \alpha')(\xi \circ \alpha),$$

because $D_{\alpha'}\alpha' = 0$. At a given point x, there is a geodesic α such that $\alpha(0) = x$ and $\alpha'(0)$ is a given tangent vector. By polarization on the values $\alpha'(0) = v$, replacing v by $v + w$, we find that for all vector fields η, ζ we have

$$(*) \qquad Q(\eta, \zeta)\xi + Q(\zeta, \eta)\xi = R(\eta, \xi)\zeta + R(\zeta, \xi)\eta.$$

On the other hand, by Proposition 1.3, we have

$$(**) \qquad Q(\eta, \zeta)\xi - Q(\zeta, \eta)\xi = R(\eta, \zeta)\xi.$$

We add equations (*) and (**). The term $Q(\zeta, \eta)\xi$ drops out, and we obtain

$$\begin{aligned} 2Q(\eta, \zeta)\xi &= R(\eta, \xi)\zeta - R(\xi, \zeta)\eta - R(\zeta, \eta)\xi \\ &= 2R(\eta, \xi)\zeta \quad \text{by the Bianchi identity,} \end{aligned}$$

which concludes the proof of the converse, and of the proposition.

XIII, §3. METRIC KILLING FIELDS

We now turn to properties having specially to do with the pseudo Riemannian case. The metric is denoted by g. We omit the subscript g in the scalar product $\langle \xi, \eta \rangle$ for simplicity.

Proposition 3.1. *Suppose X is pseudo Riemannian. The following conditions are equivalent to a vector field ξ being g-Killing.*

Kill$_g$ 1. $\mathscr{L}_\xi g = 0$.

Kill$_g$ 2. $\mathscr{L}_\xi$ is a derivation with respect to the metric product, that is, for all vector fields η, ζ we have

$$\mathscr{L}_\xi \langle \eta, \zeta \rangle = \langle [\xi, \eta], \zeta \rangle + \langle \eta, [\xi, \zeta] \rangle.$$

Kill$_g$ 3. The map $(\eta, \zeta) \mapsto \langle D_\eta \xi, \zeta \rangle$ is skew-symmetric, or in other words for all vector fields η,

$$\langle D_\eta \xi, \eta \rangle = 0.$$

Proof (Cf. [O'N 83]). Assume that ξ is g-Killing. The property **Kill$_g$ 1** then follows essentially directly from the definition of Lie derivative, because for all t, $\rho_t^*(g) = g$, so the Lie derivative of g is 0. The converse is also immediate, because in general

$$\frac{d}{ds}\rho_s^*(g) = \rho_s^*(\mathscr{L}_\xi g).$$

Hence assuming **Kill$_g$ 1**, we conclude that the left side is 0, whence $\rho_s^*(g)$ is constant, whence equal to g, thus proving that ξ is g-Killing.

The other two equivalences come from general formulas exhibiting the obstruction to being a derivation.

Lemma 3.2. *For all vector fields ξ, η, ζ we have*

$$\begin{aligned}\mathscr{L}_\xi\langle \eta, \zeta\rangle &= \langle \mathscr{L}_\xi \eta, \zeta\rangle + \langle \eta, \mathscr{L}_\xi\zeta\rangle + \mathscr{L}_\xi(g)\langle \eta, \zeta\rangle \\ &= \langle \mathscr{L}_\xi \eta, \zeta\rangle + \langle \eta, \mathscr{L}_\xi\zeta\rangle + \langle D_\eta\xi, \zeta\rangle + \langle \eta, D_\zeta\xi\rangle.\end{aligned}$$

Proof. The first identity exhibits the fact that $\mathscr{L}_\xi(g(\eta, \zeta))$ satisfies the Leibniz derivation product rule, relative to the triple (g, η, ζ), cf. Chapter V, Proposition 5.1, which applies to all multilinear forms, not just alternating forms. Thus **Kill$_g$ 2** is immediately equivalent to **Kill$_g$ 1**. The second formula follows from the metric derivative property

$$\mathscr{L}_\xi\langle \eta, \zeta\rangle = \langle D_\xi\eta, \zeta\rangle + \langle \eta, D_\xi\zeta\rangle,$$

after using $D_\xi\eta = D_\eta\xi + [\xi, \eta]$ and similarly for $D_\xi\zeta$.

Remark. Lemma 3.2 plays the same role as Lemma 2.1. The vanishing of $\mathscr{L}_\xi(g)$ (resp. $\langle D_\eta\xi, \zeta\rangle + \langle \eta, D_\zeta\xi\rangle$) is equivalent to $\mathscr{L}_\xi$ being a derivation. Both the D-Killing and g-Killing fields are thus characterized as being derivations with respect to a natural product, as in **Kill 2** and **Kill$_g$ 2.**

Next we give some properties of metric Killing fields. We begin by a property of all vector fields ξ, with flow $\{\rho_s\}$.

(1) $$\partial_s\rho(s, x) = \xi(\rho(s, x)) = T\rho_s(x)\xi(x).$$

Proof. We have

$$\begin{aligned}\partial_s\rho(s, x) &= \partial_t\rho(s+t, x)|_{t=0} \\ &= \partial_t\rho(s, \rho(t, x))|_{t=0} \\ &= \partial_t\rho_s(\rho(t, x))|_{t=0} \\ &= T\rho_s(x)\partial_t\rho(t, x)|_{t=0} = T\rho_s(x)\xi(x)\end{aligned}$$

as was to be shown.

Let α be a curve in X. Then by the chain rule,

(2) $$\partial_t\rho_s(\alpha(t)) = T\rho_s(\alpha(t))\alpha'(t).$$

We then obtain:

Proposition 3.3. *Let ξ be a g-Killing field.*

(i) *For any curve α, we have*

$$\langle \xi(\rho_s \circ \alpha), (\rho_s \circ \alpha)'\rangle = \langle \xi \circ \alpha, \alpha'\rangle.$$

Equivalently, the left side is independent of s.

(ii) *If α is a geodesic, then $\langle \xi \circ \alpha, \alpha'\rangle$ is constant.*

Proof. The proof of (i) is immediate from (1) and (2). As for (ii), we take the derivative of the function $\langle \xi \circ \alpha, \alpha'\rangle$, and find

$$\langle D_{\alpha'}(\xi \circ \alpha), \alpha'\rangle$$

because $D_{\alpha'}\alpha' = 0$ by definition of a geodesic. By **Kill$_g$ 3** it follows that the above expression is 0, thus proving the proposition.

I learned the following two results from Klingenberg [Kl 83/95].

Proposition 3.4. *Let ξ be a g-Killing field. As usual let $\xi^2 = \langle \xi, \xi\rangle$. Then*

$$\operatorname{grad}\xi^2 = -2D_\xi\xi.$$

Proof. Again let α be a curve with $\alpha(0) = x$, $\alpha'(0) = v = \xi(x)$. Consider the derivative

$$h(s, t) = \partial_t\langle \partial_s\rho(s, \alpha(t)), \partial_s\rho(s, \alpha(t))\rangle.$$

Since $\partial_s\rho(s, \alpha(t)) = \xi(\rho_s(\alpha(t)))$, putting $f = \xi^2$ we find

$$k(0, 0) = df(x)v = \langle \operatorname{grad}\xi^2, v\rangle.$$

On the other hand, from a basic property of the Riemannian covariant derivative, we also have

$$\begin{aligned}h(s, t) &= 2\langle \partial_s\rho(s, \alpha(t)), D_t\partial_s\rho(s, \alpha(t))\rangle \\ &= 2\langle \partial_s\rho(s, \alpha(t)), D_s\partial_t\rho(s, \alpha(t))\rangle\end{aligned}$$

by the usual commutation rule of Chapter VIII, Lemma 5.3,

$$\begin{aligned}&= 2\langle \xi(\rho_s \circ \alpha(t)), D_s(\rho_s \circ \alpha)'(t)\rangle \\ &= 2\partial_s\langle \xi(\rho_s \circ \alpha(t)), (\rho_s \circ \alpha)'(t)\rangle \\ &\qquad - 2\langle D_s\xi(\rho_s \circ \alpha(t)), (\rho_s \circ \alpha)'(t)\rangle \\ &= -2\langle D_s\xi(\rho_s \circ \alpha(t)), (\rho_s \circ \alpha)'(t)\rangle\end{aligned}$$

because $\rho_s \circ \alpha$ is a geodesic and we apply Proposition 3.3(ii). We then evaluate at $t = 0$, $s = 0$ to get

$$h(0, 0) = -2\langle D_\xi \xi(x), v\rangle.$$

The two expressions for $h(0, 0)$ are valid for all v, and hence the proposition follows.

Corollary 3.5. *Let ξ be a g-Killing field and $\rho = \rho(s, x)$ its flow. For fixed x, the curve $s \mapsto \rho(s, x)$ is a non-constant geodesic if and only if $\xi(x) \neq 0$ and $d\xi^2(x) = 0$.*

Proof. A curve $s \mapsto \beta(s)$ is a geodesic if and only if $D_{\beta'}\beta' = 0$. In our context, with $\beta(s) = \rho(s, x)$, this means $D_s \partial_s \rho(s, x) = 0$, and so the equivalence is clear from the proposition.

XIII, §4. LIE ALGEBRA PROPERTIES OF KILLING FIELDS

We continue to assume that (X, D) is a D-manifold.

Proposition 4.1. *Killing fields form a Lie subalgebra of all vector fields.*

Proof. It suffices to prove that if ξ, η satisfy **Kill 2**, then so does $[\xi, \eta]$. This is a special case of the following lemma, formulated in an abstract context because at this point I want to emphasize the extent to which the present arguments depend only on Lie algebras over rings.

Lemma 4.2. *Let V be a Lie algebra (over a commutative ring). Suppose given a bilinear map $V \times V \to V$, which we denote*

$$(y, z) \mapsto yz,$$

and call the bilinear product. Let W be the submodule of V consisting of all elements $w \in V$ such that the map

$$y \mapsto [w, y]$$

is a derivation for this bilinear product, namely

$$[w, yz] = [w, y]z + y[w, z].$$

Then W is a Lie subalgebra of V.

Proof. We carry out the short computation in full, but note that having formulated the result, the computation is forced, and no surprise occurs. For $v, w \in W$ we have to show that $[v, w]$ acts as a derivation with respect to the bilinear product. We shall use the defining property of the bracket product of a Lie algebra, which says that bracketing with an element is a derivation with respect to the bracket product. Let $y, z \in V$. Then

$$\begin{aligned}
[[v, w], yz] &= [[v, yz], w] + [v, [w, yz]] \\
&= [[v, y]z + y[v, z], w] + [v, [w, y]z + y[w, z]] \\
&= [[v, y], w]z + [v, y][z, w] + [y, w][v, z] + y[[v, z], w] \\
&\quad + [v, [w, y]]z + [w, y][v, z] + [v, y][w, z] + y[v, [w, z]].
\end{aligned}$$

The middle terms cancel, and using the bracket derivation property, what is left is

$$= [[v, w], y]z + y[[v, w], z],$$

which proves the lemma.

As noted at the beginning of §2, we apply the lemma to the bilinear map

$$(\zeta, \eta) \mapsto D_\zeta \eta.$$

We take the real numbers as the ring of coefficients. This concludes the proof of Proposition 3.1.

The above proposition does not avoid having to give a separate but similar argument for the analogous property of metric Killing fields.

Proposition 4.3. *Suppose D is the metric derivative in the pseudo Riemannian case. Then the metric Killing fields form a Lie subalgebra of the Killing fields.*

Proof. Property **Kill$_g$ 2** states that ξ is Killing if and only if the Lie derivative $\mathscr{L}_\xi$ is a derivation with respect to the metric product. As in Lemma 4.2, one proves that the set of vector fields which act as a derivation with respect to such a product is a Lie subalgebra. One uses the fact that on the space of functions, one has

$$\mathscr{L}_{[\xi,\eta]} = \mathscr{L}_\xi \circ \mathscr{L}_\eta - \mathscr{L}_\eta \circ \mathscr{L}_\xi.$$

The steps essentially follow those of Lemma 4.2 and will be left to the reader, as well as the possible formulation of an abstract lemma to cover the situation. On the other hand, one can also argue from **Kill$_g$ 1**, since one has the same formula for $\mathscr{L}_{[\xi,\eta]}$ acting on the metrics, showing at once that if $\mathscr{L}_\xi(g) = \mathscr{L}_\eta(g) = 0$ then also $\mathscr{L}_{[\xi,\eta]}(g) = 0$. Take your pick.

We now come to some very different considerations, fixing a point p on the manifold. Let Kill(X) be the vector space of Killing fields. We shall define certain subspaces $\mathfrak{h}_p$ and $\mathfrak{m}_p$. The following comes from analyzing the proofs in [Kl 83/95], 2.2.18, 2.2.21, and showing that they work in complete generality, quite independently of the context of symmetric spaces in which they occur in Klingenberg. Furthermore, they are valid in infinite dimensions.

We define two subspaces of Kill(X), depending on the choice of point p, as follows.

$$\mathfrak{h}_p = \text{vector space of } \eta \in \mathrm{Kill}(X) \text{ such that } \eta(p) = 0,$$

$$\mathfrak{m}_p = \text{vector space of } \xi \in \mathrm{Kill}(X) \text{ such that } D_\zeta\xi(p) = 0 \text{ for all vector fields } \zeta.$$

Remark. The above definitions apply in each case, $\mathrm{Kill}_D(X)$ and $\mathrm{Kill}_g(X)$, where g denotes the metric. Then we may denote the subspaces by

$$\mathfrak{h}_p(D),\quad \mathfrak{m}_p(D) \qquad \text{and} \qquad \mathfrak{h}_p(g),\quad \mathfrak{m}_p(g)$$

to distinguish the two types of Killing fields. Note that

$$(1) \qquad \mathfrak{h}_p(g) = \mathfrak{h}_p(D) \cap \mathrm{Kill}_g \qquad \text{and} \qquad \mathfrak{m}_p(g) = \mathfrak{m}_p(D) \cap \mathrm{Kill}_g,$$

whenever D is the metric derivative in the pseudo Riemannian case.

The following discussion and results apply to each case separately, so we formulate them by omitting the D and g from the notation. The results for $\mathfrak{h}_p(g)$ and $\mathfrak{m}_p(g)$ follow from those with D instead of g, in light of (1).

Observe that if $\xi \in \mathfrak{m}_p$ and η is any vector field, then

$$[\xi, \eta](p) = D_\xi\,\eta(p).$$

The following proposition gives commutation rules usually listed for symmetric spaces, but they hold in general.

Proposition 4.4.

(a) $[\mathfrak{m}_p, \mathfrak{m}_p] \subset \mathfrak{h}_p$.
(b) $[\mathfrak{h}_p, \mathfrak{h}_p] \subset \mathfrak{h}_p$.
(c) $[\mathfrak{h}_p, \mathfrak{m}_p] \subset \mathfrak{m}_p$.

In particular, $\mathfrak{h}_p + \mathfrak{m}_p$ *is a Lie subalgebra of* Kill(X).

Proof. For (a), we let $\xi, \eta \in \mathfrak{m}_p$ and we evaluate at p to get 0, by the definition of $\mathfrak{m}_p$. For (b), let $\eta, \zeta \in \mathfrak{h}_p$. Then

$$[\eta, \zeta](p) = D_\eta\zeta(p) - D_\zeta\eta(p) = 0$$

because $\eta(p) = 0$ and $\zeta(p) = 0$ in the indices, and $D_0 = 0$. For (c), let $\eta \in \mathfrak{h}_p$ and $\xi \in \mathfrak{m}_p$. We use the relation

$$[\eta, \xi] = D_\eta\xi - D_\xi\eta.$$

We have to show that $D_\zeta[\eta, \xi](p) = 0$ for all ζ. It suffices to show that

$$D_\zeta D_\eta\xi(p) = 0 \qquad \text{and} \qquad D_\zeta D_\xi\eta(p) = 0.$$

We use an elegant argument of Klingenberg. We have by **Kill 3**:

$$(D_\zeta D_\eta\xi)(p) = D_{D_\zeta\eta}\xi(p) + R(\zeta, \xi)\eta(p)$$
$$= 0,$$

the first term because $\xi \in \mathfrak{m}_p$, and the second because $\eta(p) = 0$. The second equation $D_\zeta D_\xi\eta(p) = 0$ follows the same way. This concludes the proof of Proposition 4.4.

Proposition 4.5. *Assume that the exponential map* $\exp_p\colon T_p \to X$ *is surjective. Then* $\mathfrak{h}_p \cap \mathfrak{m}_p = \{0\}$, *so* $\mathfrak{h}_p + \mathfrak{m}_p$ *is a direct sum. More generally, the map*

$$\mathrm{Kill}(X) \to T_p \times \mathrm{End}(T_p) \quad \textit{given by} \quad \xi \mapsto (\xi(p), D\xi(p))$$

is injective. (By definition, $D\xi(p)(v) = (D_v\xi)(p)$ *for* $v \in T_p$.)

Proof. The first assertion is a consequence of the second, so suppose that $\xi(p) = 0$ and $D_\zeta\xi(p) = 0$ for all vector fields ζ. We restrict ξ to a geodesic α with $\alpha(0) = p$. Then by Proposition 2.2, $\xi \circ \alpha$ is the unique Jacobi lift of α with (0, 0) initial conditions, so $\xi \circ \alpha = 0$. By the assumption that the exponential map is surjective, there exists a geodesic from p to any point of X, so $\xi = 0$, concluding the proof of the proposition.

Remark. For more comments on and use of the hypothesis about the exponential map, see the next section. A question also arises how large is $\mathfrak{m}_p$. In the next section, we give conditions under which $\mathfrak{m}_p$ is isomorphic to the whole tangent space T_p. Such conditions insure the existence of "enough" isometries. Similarly, $\mathfrak{h}_p + \mathfrak{m}_p$ can be smaller than Kill(X), but will be shown equal to equal Kill(X) in the symmetric case.

We conclude this section with a result usually stated on symmetric spaces. It gives the value of the Riemann tensor at a given point p for vector fields in $\mathfrak{m}_p$. Note that the proof is short and uses practically nothing of what precedes, basically only **Kill 3** and Proposition 4.4(a).

Theorem 4.6. *Fix a point $p \in X$. For all vector fields $\xi, \eta, \zeta \in \mathfrak{m}_p$, we have*

$$R(\xi, \eta)\zeta(p) = D_\zeta[\xi, \eta](p) = [\zeta, [\xi, \eta]](p).$$

Proof. By **Kill 3**, using $D_\eta\zeta(p) = 0$ and $D_\zeta\xi(p) = 0$, we get

$$D_\eta D_\zeta \xi(p) + R(\xi, \eta)\zeta(p) = 0,$$
$$D_\zeta D_\xi \eta(p) + R(\eta, \zeta)\xi(p) = 0.$$

But $R(\eta, \zeta)\xi = D_\eta D_\zeta \xi - D_\zeta D_\eta \xi - D_{[\eta,\zeta]}\xi$, and by definition, $D_{[\eta,\zeta]}\xi(p) = 0$. Using this, and subtracting the above two relations yields

$$\begin{aligned} R(\xi, \eta)\zeta(p) &= (D_\zeta D_\xi \eta - D_\zeta D_\eta \xi)(p) \\ &= D_\zeta(D_\xi \eta - D_\eta \xi)(p) \\ &= D_\zeta[\xi, \eta](p) \\ &= [\zeta, [\xi, \eta]](p), \end{aligned}$$

because putting $\lambda = [\xi, \eta]$ we know from Proposition 4.4(a) that $\lambda \in \mathfrak{h}_p$, and

$$[\zeta, \lambda](p) = D_\zeta\lambda(p) - D_\lambda\zeta(p) = D_\zeta\lambda(p),$$

thus concluding the proof of the theorem.

XIII, §5. SYMMETRIC SPACES

Throughout this section we let (X, D) be a D-manifold. After giving appropriate definitions, and more precisely after Proposition 5.2, we assume that X is a symmetric space, possibly infinite dimensional.

We begin with some remarks on isomorphisms in general. Let

$$\sigma\colon (X, D^X) \to (Y, D^Y)$$

be a D-isomorphism. Then σ carries all objects defined naturally in terms of the covariant derivative to similar objects. For instance, if α is a geodesic in X, then $\sigma \circ \alpha$ is a geodesic in Y. If γ is a lift of α in TX, and this lift is α-parallel, that is $D_{\alpha'}\gamma = 0$, then $(T\sigma) \circ \gamma$ is a $(\sigma \circ \alpha)$-parallel lift of $\sigma \circ \alpha$ in TY. In particular, let $x \in X$ and let $v \in T_xX$. Say $\gamma = \gamma(t, v)$ is α-parallel translation of v in X. Then

$$(T\sigma)\big(\gamma(t, v)\big) = (T_{\alpha(t)}\sigma)\big(\gamma(t, v)\big)$$

is $(\sigma \circ \alpha)$-parallel translation of $(T\sigma)(v)$ along $\sigma \circ \alpha$.

We shall need a lemma showing one role of the exponential map.

Lemma 5.1. *Let $x, y \in X$. Assume that $\exp_x\colon T_x \to X$ is surjective. Given a linear isomorphism $L\colon T_x \to T_y$, there is at most one D-automorphism $f\colon X \to X$ such that $f(x) = y$ and $T_xf = L$.*

Proof. A D-automorphism f maps geodesics to geodesics, and a geodesic is uniquely determined by its initial conditions, namely the value at 0 and the derivative at 0. Thus the condition that $\exp_x$ is surjective is just what is needed to determine f globally on X from its initial conditions at x.

Next we come to symmetries. By a ***D*-symmetry** (resp. ***g*-symmetry**, or **metric symmetry**), we mean a D-isomorphism (resp. metric isomorphism) $\sigma_x\colon X \to X$ such that σ_x leaves x fixed, i.e. $\sigma_x(x) = x$ and $T_x\sigma_x = -\mathrm{id}$.

Proposition 5.2. *Suppose X has a symmetry at every point $x \in X$. Then X is geodesically complete, that is $\exp_x$ is defined on T_x for all x.*

Proof. Let $\alpha\colon [0, c] \to X$ be a geodesic, defined on a finite interval. Let $x = \alpha(c)$. Then $T_x\sigma_x$ maps $-\alpha'(c)$ to $\alpha'(c)$. But σ_x being a D-isomorphism maps geodesics to geodesics, and by the uniqueness of geodesics satisfying initial conditions, it follows that σ_x maps $\alpha(t)$ with $t \in [0, c]$ to $\alpha(2c - t)$, in other words, α is defined on the interval $[0, 2c]$, whence on $\mathbf{R}$ by symmetry, thus concluding the proof.

A manifold will be called ***D*-symmetric** (resp. ***g*-symmetric**) if it has a D- (resp. g-) symmetry at every point, and if $\exp_x\colon T_x \to X$ is surjective for all $x \in X$.

Remark. If X is finite dimensional, then the surjectivity is implied by geodesic completeness because of the Hopf–Rinow theorem. This theorem may be false in infinite dimension, but it is the only basic theorem which has this remarkable property. In particular, the Cartan–Hadamard theorem is true in infinite dimension, and Hopf–Rinow is true in the case of seminegative curvature. Hence it is important not to exclude infinite dimensional symmetric spaces. Klingenberg assumes finite dimensionality at this point, unnecessarily so, as will be evident from the rest of this chapter.

The theory of Riemannian symmetric spaces is originally due to Cartan [Ca 27], [Ca 28/46]. Here we follow Klingenberg [Kl 83/95], 2.2., except for going infinite dimensional, and dealing with arbitrary D-manifolds, not just Riemannian or pseudo Riemannian manifolds.

For the key background on the surjectivity of exp, see Chapter VIII, Theorem 6.9; and Chapter IX, §3, especially Theorems 3.7 and 3.8 with its corollaries, which give conditions under which the exponential map is surjective, notably seminegative curvature.

A symmetric pair (X, D) will also be called a ***D*-symmetric space**. We often leave out the D, and simply speak of a symmetric space.

For the rest of this section, we let (X, D) be a symmetric space.

As a consequence of Lemma 5.1, we note that:

(a) *The symmetry σ_x is the unique D-automorphism of X such that $\sigma_x(x) = x$ and $T_x\sigma_x = -\text{id}$.*
(b) *We have $\sigma_x^2 = \text{id}$.*

In particular, let $\alpha\colon \mathbf{R} \to X$ be a geodesic, with $\alpha(0) = x$. Then

$$\sigma_x(\alpha(t)) = \alpha(-t).$$

This is just a special case of the more general formula which already occurred in Proposition 5.2, namely

$$(1) \qquad \sigma_{\alpha(c)}(\alpha(t)) = \alpha(2c - t) \quad \text{or} \quad \sigma_{\alpha(c)}(\alpha(2c - t)) = \alpha(t).$$

Thus symmetries are just the maps which reverse the geodesics.

For real numbers a, b we denote parallel translation along α by

$$P_{a,\alpha}^b\colon T_{\alpha(a)} \to T_{\alpha(b)}.$$

We may omit the subscript α from the notation when α is fixed throughout a discussion. We shall use the basic formalism of parallel translation, including the formulas:

PAR 1. $$P_{b,\alpha}^c \circ P_{a,\alpha}^b = P_{a,\alpha}^c.$$

PAR 2. *Let $\beta(t) = \alpha(L(t))$ be a linear reparametrization of α, with $L(t) = c_1 t + c_2$, $c_1 \neq 0$. Suppose $\alpha(c) = \beta(c)$ for some c.* Then

$$P_{a,\beta}^b = P_{L(a),\alpha}^{L(b)}.$$

Proof. We prove the second. Let $v \in T_{\alpha(c)}X$ and let $\gamma_\alpha(t) = \gamma_\alpha(t, v)$ be the α-parallel translation of v along α. Let $\eta(t) = \gamma_\alpha(L(t))$. We want to show η is β-parallel translation of v. We have $\gamma_\alpha(c) = \gamma_\beta(c) = v$, and also $\eta(c) = v$. In a chart, letting B be the bilinear map defining the covariant derivative, we have

$$\begin{aligned} D_{\beta'}\eta(t) &= \eta'(t) - B_{\beta(t)}(\eta(t), \beta'(t)) \\ &= c_1[\gamma_\alpha'(L(t)) - B_{\alpha(L(t))}(\gamma_\alpha(L(t)), \alpha'(L(t))] \\ &= 0 \end{aligned}$$

because $D_{\alpha'}\gamma_\alpha = 0$. Hence $\eta = \gamma_\beta$ is β-parallel translation of v, with the prescribed initial condition. Thus we have shown

$$P_{c,\beta}^t = P_{c,\alpha}^{L(t)}.$$

The general formula then follows from **PAR 1**.

Proposition 5.3. *Let $x, y \in X$. Let α be a non-constant geodesic such that $\alpha(c) = x$ and $\alpha(b) = y$. Then*

$$T_y\sigma_x = -P_{b,\alpha}^{2c-b} \quad \text{on} \quad T_yX.$$

Proof. Let $v \in T_{\alpha(c)}X$ be a tangent vector as above. By the remarks at the beginning of this section, $(T\sigma_{\alpha(c)})(\gamma(t, v))$ is parallel translation of $(T\sigma)(v)$ along $\sigma_{\alpha(c)} \circ \alpha$, and we may apply **PAR 2** with $\beta(t) = \alpha(2c - t)$. Note that $(T\sigma)(v) = -v$. Hence

$$(T_{\alpha(b)}\sigma_x)(P_{c,\alpha}^b(v)) = -P_{c,\alpha}^{L(b)}(v).$$

Putting $w = P_{c,\alpha}^b(v)$ so $v = P_{b,\alpha}^c(w) = -P_{c,\alpha}^{L(b)} \circ P_{b,\alpha}^c(w)$ yields the proposition.

One may get rid of the flipping and minus sign by defining **α-translation**, or **translation along** α, to be the map

$$\tau_{\alpha,s}\colon X \to X \quad \text{such that} \quad \tau_{\alpha,s} = \sigma_{\alpha(s/2)} \circ \sigma_{\alpha(0)}.$$

Such translations stem from Cartan [Ca 28]. Note that from (1), we get

$$(2) \qquad \tau_\alpha(s, \alpha(t)) = \tau_{\alpha,s}(\alpha(t)) = \alpha(t + s).$$

Proposition 5.4. *Let $P_{t,\alpha}^{t+s}\colon T_{\alpha(t)} \to T_{\alpha(t+s)}$ be parallel translation. Then*

$$T_{\alpha(t)}\tau_{\alpha,s} = P_{t,\alpha}^{t+s}.$$

In particular, for $v \in T_{\alpha(0)}$, *we have*

$$T_{\alpha(0)}\tau_{\alpha,s}(v) = P^s_{0,\alpha}(v).$$

Proof. This is immediate from Proposition 5.3, using the chain rule for the tangent map of a composite mapping, and **PAR 1**.

Proposition 5.5. *Let* α *be a non-constant geodesic.*

(i) *Then* $\{\tau_{\alpha,s}\}$ *is the flow of a Killing field, i.e. it is a one-parameter group of D-automorphisms. In other words, if for* $x \in X$ *we define*

$$\xi_\alpha(x) = \partial_1\tau_\alpha(0, x),$$

then ξ_α *is Killing, and* τ_α *is its flow.*

(ii) *The geodesic* α *is an integral curve of* ξ_α, *that is, for all* t,

$$\alpha'(t) = \xi_\alpha(\alpha(t)).$$

If the symmetries are metric symmetries, then $\tau_{\alpha,s}$ *is the flow of a metric Killing field, and* ξ_α *is metric Killing.*

Proof. We first show that $\tau_{\alpha,s+t} = \tau_{\alpha,s} \circ \tau_{\alpha,t}$ for all $s, t \in \mathbf{R}$. Both sides are D-automorphisms. By Lemma 5.1 it suffices to show that they coincide at one point and that their tangent maps coincide at this point. We can select the point to be, say, $\alpha(0)$, in which case the equality of both sides at $x = \alpha(0)$ is given by (2). Then the equality of the tangent maps at $\alpha(0)$ is given by Proposition 5.4, which concludes the proof that $\{\tau_{\alpha,s}\}$ is a one-parameter group of D-automorphisms. It is then a property of all one-parameter groups of differential automorphisms, that if one defines $\xi_\alpha(x)$ as in the formula given in (i), then $\{\tau_\alpha\}$ is the flow of ξ_α. The proof is in any case immediate by differentiating $\tau_\alpha(s + t, x)$.

For (ii), we differentiate the equation in (2) with respect to s, and then set $s = 0$ to obtain the fact that α is an integral curve of ξ_α.

The remark about metric symmetries is immediate, due to the fact that parallel translation in the metric case is an isometry. This concludes the proof of the proposition.

Proposition 5.6. *Let* α, β *be non-constant geodesics with*

$$\alpha(0) = \beta(0) = p.$$

Let $\alpha'(0) = w$. *Let* τ_α *be translation along* α *as above, and let*

$$\eta(t) = \partial_1\tau_\alpha(0, \beta(t)) = \xi_\alpha(\beta(t)).$$

Then

$$\eta(0) = w \quad \text{and} \quad D_{\beta'}\eta(0) = 0.$$

Thus η *is the unique Jacobi lift of* β *satisfying these initial conditions.*

Proof. That $\eta(0) = \alpha'(0) = w$ comes from Proposition 5.5(ii), so we next have to show that $D_{\beta'}\eta(0) = 0$. Let $v = \beta'(0)$. By Proposition 5.4, we know that $T_p\tau_{\alpha,s} = P^s_{0,\alpha}$. Essentially from the definition of parallel translation, it follows that $D_sT_p\tau_{\alpha,s}(v) = 0$. (Cf. Chapter VIII, Theorems 3.3 and 3.4.) Let $\varphi(s, t) = \tau_\alpha(s, \beta(t))$. Since $\partial_2\varphi(s, Q) = T_p\tau_{\alpha,s}(v)$, we get:

$$0 = D_1\partial_2\varphi(0, 0) = D_2\partial_1\varphi(0, 0) = D_{\beta'}\eta(0).$$

The assertion about Jacobi lifts is merely a reminder of standard properties of Jacobi lifts, cf. Chapter IX, Theorem 2.1 and Proposition 2.8. This concludes the proof of Proposition 5.6.

Corollary 5.7. *Let* α *be a non-constant geodesic, and put* $\alpha(0) = p$. *Then* $\xi_\alpha \in \mathfrak{m}_p$.

Proof. Special case of Proposition 5.6, because given $v \in T_pX$ we can find a geodesic β such that $\beta(0) = p$ and $\beta'(0) = v$.

We are now in a position to summarize a number of results into an exact sequence, which we call the **Killing sequence** at a point p on a symmetric space X:

$$0 \to \mathfrak{h}_p \to \mathrm{Kill}(X) \rightleftarrows T_p \to 0.$$

The arrow $\mathrm{Kill}(X) \to T_p$ is simply $\xi \mapsto \xi(p)$. By definition, $\mathfrak{h}_p$ is the kernel. Corollary 5.7 allows us to split this sequence as follows. A vector $v \in T_p$ determines a geodesic α uniquely such that $\alpha(0) = p$ and $\alpha'(0) = v$. This geodesic in turn determines ξ_α, which we may now denote by ξ_v.

Theorem 5.8. *The Killing sequence is exact, and is split by the map* $v \mapsto \xi_v$. *The map* $\xi \mapsto \xi(p)$ *thus induces an isomorphism*

$$\mathfrak{m}_p \xrightarrow{\approx} T_pX$$

of $\mathfrak{m}_p$ *with the tangent space at* p. *We have a direct sum decomposition*

$$\mathrm{Kill}(X) = \mathfrak{h}_p \oplus \mathfrak{m}_p.$$

If $\xi \in \mathfrak{m}_p$, $\xi \neq 0$ *then* $\xi = \xi_\alpha = \xi_v$, *where* α *is the geodesic such that* $\alpha(0) = p$ *and* $\alpha'(0) = \xi(p) = v$.

Proof. That $\mathfrak{h}_p$ is the kernel of $\xi \mapsto \xi(p)$ comes from the definition of $\mathfrak{h}_p$. The map is surjective, because at the given point p we can find a geodesic α such that $\alpha(0) = p$ and $\alpha'(0)$ is equal to a given tangent vector at p. We can then apply Proposition 5.5. The direct sum decomposition follows from Proposition 4.5. The last statement is merely a rephrasing of these results, in light of Proposition 5.5. This concludes the proof.

On a symmetric space, there is a complement to Proposition 2.2, namely every Jacobi field comes from a Killing field by a theorem of Bott–Samelson [BoS 58].

Examples of symmetric spaces

Let G be a Lie group, possibly infinite dimensional. Let $\sigma \neq \text{id}$ be an automorphism of G such that $\sigma^2 = \text{id}$. If we put ${}^t x = \sigma(x)^{-1}$ for $x \in G$, then $\sigma(x) = {}^t x^{-1}$ and the "transpose" is an anti automorphism, so one can work either with σ or the transpose, sometimes written as x^*. Let:

$K = G_{\text{fix}} =$ the fixed submanifold, which is a closed submanifold;

$G_\sigma =$ submanifold of elements of the form $x\sigma(x)^{-1}$, $x \in G$.

We have a differential isomorphism

$$\varphi\colon G/K \xrightarrow{\approx} G_\sigma \qquad \text{given by} \qquad xK \mapsto x\sigma(x)^{-1}.$$

There is a differential representation of G on G_σ defined by

$$\psi\colon G \to \text{Aut}(G_\sigma) \qquad \text{such that} \qquad \psi(x)y = xy{}^t x = xy\sigma(x)^{-1}.$$

On the other hand, there is a differential representation of G on G/K by translation

$$\tau\colon G \to \text{Aut}(G/K) \qquad \text{such that} \qquad \tau(x)yK = xyK.$$

Under the isomorphism φ, translation $\tau(x)$ corresponds to $\psi(x)$.

One also has the **Cartan symmetry** S_x on G_σ for $x \in G_\sigma$,

$$S_x\colon G_\sigma \to G_\sigma \qquad \text{given by} \qquad S_x(y) = xy^{-1}x.$$

This symmetry gives a morphism (viewed as a non-associative product)

$$G_\sigma \times G_\sigma \to G_\sigma \qquad \text{denoted by} \qquad S_x(y) = x \cdot y.$$

The corresponding morphism of $G/K \times G/K \to G/K$ is given by

$$xK \cdot yK = x\sigma(x)^{-1}\sigma(y)K.$$

All the above assertions are immediately verified.

Loos [Lo 69] was the first to note that the theory of symmetric spaces could be based on essentially algebraic properties satisfied by the product $(x, y) \mapsto x \cdot y$. So let us define a **Loos space** to be a connected manifold X with a morphism $X \times X \to X$ denoted $(x, y) \mapsto x \cdot y$, satisfying the following conditions for all $x, y, z \in X$.

LO 1. $x \cdot x = x$.
LO 2. $x \cdot (x \cdot y) = y$.
LO 3. $x \cdot (y \cdot z) = (x \cdot z) \cdot (y \cdot z)$.
LO 4. x is an isolated fixed point of the morphism $y \mapsto x \cdot y$.

The last property means that there is an open neighborhood U of x such that for all $y \in U$, if $x \cdot y = y$ then $y = x$. Loos spaces obviously form a category.

The spaces G_σ and G/K (which are differentially isomorphic under φ) are Loos spaces under the above defined products, and φ is a Loos isomorphism.

The verification is immediate from the definitions.

One may denote the morphism $y \mapsto x \cdot y$ by $\ell_x\colon X \to X$, and similarly for the right operation $r_x\colon y \to y \cdot x$. Then, for instance, **LO 2** means that $\ell_x^2 = \text{id}$.

Note that, instead of taking $K = G_{\text{fix}}$, one could take any subgroup contained in G_{fix} but containing the connected component of the identity.

In finite dimension (at least) symmetric spaces essentially all come from the above example. Expositions may start from Lie groups (as in Helgason) or from the Riemannian geometry point of view (as in Klingenberg). The present chapter gives an introduction to both points of view (see also Chapter XII).

XIII, §6. PARALLELISM AND THE RIEMANN TENSOR

We begin with some basic properties of the Riemann tensor R on an arbitrary D-manifold (X, D).

Let $x \in X$. For each $u \in T_x X$ we have a continuous linear operator

$$R_u\colon T_x \to T_x \qquad \text{given by} \qquad R_u(v) = R(u, v)u.$$

In fact, the trilinear map $(u, v, w) \mapsto R(u, v, w)$ is continuous on $T_x \times T_x \times T_x$. Note that the sign selected here for R_u is opposite to the sign of Klingenberg [Kl 83/95], 2.2.9, because the expression $R(u, v)u$ is the one which occurs in the way we wrote the Jacobi differential equation. Similarly, if Ω is a trilinear tensor field, one defines $\Omega_u(v) = \Omega(u, v, u)$.

If the spray is associated with a pseudo Riemannian metric $\langle\ ,\ \rangle$, then the standard properties of the Riemann tensor immediately show that R_u is self-adjoint, that is

$$\langle R_u(v), w\rangle = \langle v, R_u(w)\rangle.$$

In other words, it is equal to its transpose on the tangent space. But at the beginning, we work in greater generality without assuming that the spray comes from a metric. We let $P^b_{a,\alpha}$ be parallel translation along a geodesic α.

Proposition 6.1. *Let Ω: $X \to L^3(TX, TX)$ be a trilinear tensor field on a D-manifold X. Then $D_\xi\Omega = 0$ for all ξ if and only if parallel translation commutes with Ω, that is for every geodesic α,*

$$P^b_{a,\alpha} \circ \Omega_{\alpha'(a)} = \Omega_{\alpha'(b)} \circ P^b_{a,\alpha}.$$

Proof. If $D_\xi\Omega = 0$ for all vector fields ξ, then the commutation comes directly from the definition of $D_\xi\Omega = 0$, and, say, the local expression as in Chapter VIII, 3.5, 3.6, and 3.7. Conversely, for a trilinear tensor field Ω and a geodesic α, we have

$$(D_{\alpha'}\Omega)(\alpha(0)) = \lim_{t\to 0} \frac{P_t\Omega_{\alpha(t)} - \Omega_{\alpha(0)}}{t}.$$

The converse (actually the equivalence) follows immediately. The proposition could have been given in Chapter VIII.

As an example of Proposition 6.1, we have:

Proposition 6.2. *Let (X, D) be a symmetric space. Then for all vector fields ξ we have*

$$D_\xi R = 0.$$

In other words, the Riemann tensor is parallel.

Proof. At a given point x, we compute $T_x\sigma_x$ applied to $(D_uR)(v, w, z)$ in two ways, with vectors $u, v, w, z \in T_x$. First,

$$T_x\sigma_x \cdot (D_uR)(v, w, z) = -(D_uR)(v, w, z) \quad \text{because} \quad T_x\sigma_x = -\text{id}.$$

On the other hand, by functoriality, and the fact that σ_x is a D-automorphism,

$$\begin{aligned} T_x\sigma_x \cdot (D_uR)(v, w, z) &= (D_{T_x\sigma_x \cdot u}R)(T_x\sigma_x \cdot v, T_x\sigma_x \cdot w, T_x\sigma_x \cdot z) \\ &= (D_{-u}R)(-v, -w, -z) = (D_uR)(v, w, z). \end{aligned}$$

This proves the proposition.

In light of Propositions 6.2 and 6.1, on a symmetric space we have

$$P^b_{a,\alpha} \circ R_{\alpha'(a)} = R_{\alpha'(b)} \circ P^b_{a,\alpha}.$$

The next proposition will also apply to symmetric space, but depends only on the parallelism of the Riemann tensor. We again follow Klingenberg [Kl 83/95].

Proposition 6.3. *Let (X, D) be a D-manifold. Let α be a geodesic, $\alpha(0) = x$, $\alpha'(0) = u \neq 0$. Let η be the Jacobi lift of α with initial conditions*

$$\eta(0) = v_0 \quad \text{and} \quad D_{\alpha'}\eta(0) = v_1.$$

Let

$$A(t) = \sum_{K=0}^{\infty} \frac{t^{2k}}{(2k)!} R^k_u(v_0) + \sum_{k=0}^{\infty} \frac{t^{2k+1}}{(2k+1)!} R^k_u(v_1).$$

Let $P^t_{0,\alpha}$ be parallel translation along α, and assume that parallel translation commutes with the Riemann tensor. Then

$$\eta(t) = P^t_{0,\alpha}A(t).$$

Proof. Let $\eta_1(t) = P^t_{0,\alpha}A(t)$. Trivially, $\eta_1(0) = v_0$. By Chapter IX, Proposition 5.1, we also see that $D_{\alpha'}\eta_1(0) = v_1$ because $D_{\alpha'}P_{0,\alpha} = 0$. There remains to prove that η_1 satisfies the Jacobi differential equation. Because of the absolute convertence of the series, it suffices to check what happens to each term. Let γ denote parallel translation along α. Then for $v \in T_{\alpha(0)}X$, since $D_{\alpha'}\gamma = 0$, we find:

$$D^2_{\alpha'}\gamma\left(t, \frac{t^m}{m!} R^k_u(v)\right) = \frac{t^{m-2}}{(m-2)!}\gamma(t, R^k_u(v)) = \frac{t^{m-2}}{(m-2)!} P^t_0 \circ R^k_u(v).$$

By hypothesis,

$$P^t_0 \circ R_u \circ R^{k-1}_u = R_{\alpha'(t)} \circ P^t_{0,\alpha} \circ R^{k-1}_u.$$

Applying the definition $R_w(w_1) = R(w, w_1)w$, and the definition of the power series A, the assertion of the proposition drops out.

The assumption that parallel translation commutes with the Riemann tensor of course applies to a symmetric space, which is our main interest at this time. In this case, we obtain:

Corollary 6.4. *Let (X, D) be a symmetric space. Let η be the Jacobi lift of a geodesic α such that*

$$\eta(0) = -v \in T_{\alpha(0)} \quad \text{and} \quad \eta(t/2) = 0.$$

Then

$$\eta(t) = P^t_{0,\alpha}(v).$$

CHAPTER XIV

Immersions and Submersions

In this chapter, we investigate systematically the relationships of some of the differential geometric notions for submanifolds and for submersions. These involve the covariant derivative, Hessian, and curvature. The determination of the Hessian can be applied to compare the Laplacian in both contexts, because we can define the Laplacian as the trace of the Hessian in the finite dimensional case. The connection with the definition in terms of the divergence of the gradient will be given in Chapter XV.

The material of this chapter was systematized during the 1960s. Aside from Kobayashi–Nomizu, readers can consult O'Neill [O'N 66], and Dombrowski [Do 68].

XIV, §1. THE COVARIANT DERIVATIVE ON A SUBMANIFOLD

Let X be a Riemannian manifold (not necessarily finite dimensional), and let Y be a submanifold, with the induced Riemannian structure. We have an orthogonal decomposition of the tangent space at a point $x \in Y$ given by

$$T_xX = T_xY + N_xY$$

where $N_xY = (T_xY)^\perp$ is the orthogonal complement of T_xY in T_xX. Immediately from a chart, we seen that $\{N_x\}_{x\in Y}$ are the fibers of a vector bundle, called the **normal bundle** of Y in X, and denoted by N_XY. We let pr_{TY} and pr_{NY} be the orthogonal projections from TX to TY and NY respectively.

We have metric derivatives D^X on X and D^Y on Y. This section is devoted to comparing them. We can take projections on TY or NY inside TX, and thus we have two orthogonal components. We study each component separately. Note that we may view the restriction of TX to Y as a vector bundle over Y. This restriction splits as an orthogonal sum

$$(TX)_Y = TY \perp NY,$$

and a section over Y therefore has two components (ζ, ν), where ζ is a vector field on Y, and ν is a **normal field**, that is a section of the normal bundle. If η is a vector field over Y, then we can summarize the results of this section in the following theorem.

Theorem 1.1. *Let ζ_X, ν_X be extensions of ζ, ν to X. The covariant derivatives of ζ_X and ν_X on Y can be expressed in the form*

$$D^X_\eta \zeta_X = D^Y_\eta \zeta + h_{12}(\eta, \zeta),$$
$$D^X_\eta \nu_X = h_{21}(\eta, \nu) + \nabla_\eta \nu,$$

where:

h_{12} is a symmetric bilinear bundle map $TY \times TY \to N_X Y$.

h_{21} is a bilinear bundle map $TY \times NY \to TY$.

$\nabla_\eta \nu = \mathrm{pr}_{NY}\, D^X_\eta \nu_X$ is independent of the extension ν_X of ν, and ∇ is a metric derivative on NY (to be defined in Proposition 1.6).

We may then define an operator

(1) $H_\eta\colon \Gamma TY \to \Gamma NY$ *by the condition* $H_\eta(\zeta) = h_{12}(\eta, \zeta)$,

and then

(2) $$h_{21}(\eta, \nu) = -\,{}^t H_\eta(\nu).$$

As usual, the transpose is defined by the condition that for all vector fields ξ on Y, and normal fields μ on Y, we have

$$\langle H_\eta(\xi), \mu\rangle = \langle \xi, {}^t H_\eta(\mu)\rangle.$$

Formula (2) will be proved in Theorem 1.5. Thus we give precise information on the four components h_{ij} with $i, j = 1, 2$. In particular, we see from (1) and (2) that D^X_η is represented by the matrix

$$\begin{pmatrix} D^Y_\eta & -{}^t H_\eta \\ H_\eta & \nabla_\eta \end{pmatrix} \text{ acting on } \begin{pmatrix} \zeta \\ \nu \end{pmatrix}.$$

On the other hand, there is still another operator commonly used, see for instance [KoN 69], Chapter VII, §3, defined as follows. For each normal field μ on Y there is a linear map

$$S_\mu\colon \Gamma TY \to \Gamma TY$$

defined by the condition

(3) $$\langle S_\mu(\eta), \xi\rangle = \langle \mu, h_{12}(\eta, \xi)\rangle \quad \text{for all } \xi, \eta \in \Gamma TY.$$

Note that S_μ is Fu(Y)-linear. Furthermore, since h_{12} is symmetric, it follows that S_μ is symmetric as a linear operator, that is, for all vector fields ξ, η on Y we have

(4) $$\langle S_\mu(\xi), \eta\rangle = \langle \xi, S_\mu(\eta)\rangle.$$

Directly from Theorem 1.1 and the definition we obtain what is called the **Weingarten formula**

(5) $$h_{21}(\eta, \mu) = S_\mu(\eta) = -{}^t H_\eta(\mu).$$

Thus S_μ is a representation of the second fundamental form. From the definitions and Theorem 1.1, we may write

(6) $$D^X_\eta \nu = \nabla_\eta \nu + S_\nu \eta \quad \text{on } Y.$$

Finite dimensional case: the trace. Suppose that Y is finite dimensional. We may then define the **trace** of h_{12} as follows. Let $p = \dim Y$, and let $\{\xi_1, \dots, \xi_p\}$ be an orthonormal frame of vector fields on Y. On Y, we define

$$\mathrm{tr}\, h_{12} = \sum_{i=1}^{p} h_{12}(\xi_i, \xi_i).$$

For each normal field μ we can take the scalar product with μ, and then the trace is simply the trace of the linear automorphism S_μ, at each point of Y. It is independent of the choice of orthonormal frame on Y.

We now proceed systematically with propositions proving all these statements. The first proposition determines h_{11}.

Proposition 1.2. *Let η, ζ be vector fields on Y. Let ζ_X be a vector field on X extending ζ locally one some open set. Then on Y,*

$$\mathrm{pr}_{TY}\, D^X_\eta \zeta_X = D^Y_\eta \zeta.$$

Proof. Let $\nabla_\eta \zeta_X = \mathrm{pr}_{TY}\, D^X_\eta \zeta_X$. Let η_X be an extension of η to a vector field on an open set in X. Then for $x \in Y$ in this open set we have

$$[\eta_X, \zeta_X]_X(x) = [\eta, \zeta]_Y(x), \tag{7}$$

so

$$(D^X_\eta \zeta_X)(x) - D^X_\zeta \eta_X(x) = [\eta_X, \zeta_X](x) = [\eta, \zeta](x).$$

Hence

$$\nabla_\eta \zeta_X - \nabla_\zeta \eta_X = [\eta, \zeta].$$

Fix η and η_X. At x, $D^X_\eta \zeta_X$ depends only on $\eta(x)$. Also $D^X_\zeta \eta_X$ depends only on $\zeta(x)$. Hence this formula shows that $\nabla_\eta \zeta_X$ is independent of the choice of extension ζ_X of ζ. Thus we may omit the subscript X, and write simply $\nabla_\eta \zeta$. Furthermore, we have proved one of the defining properties of the covariant derivative.

By Theorem 4.1 of Chapter VII, it will suffice to show that ∇ is a covariant derivative. Note that ∇_η is Fu(Y)-linear in the variable η, and satisfies the product rule of a derivative because it is satisfied by D^X_η. Finally, we verify the metric property. Let ξ be another vector field on Y. Then on Y,

$$\begin{aligned}\xi \cdot \langle \eta, \zeta \rangle &= \xi_X \cdot \langle \eta_X, \zeta_X \rangle \\ &= \langle D^X_\xi \eta_X, \zeta_X \rangle + \langle \eta_X, D^X_\xi \zeta_X \rangle \\ &= \langle \mathrm{pr}_{TY}\, D^X_\xi \eta_X, \zeta \rangle + \langle \eta, \mathrm{pr}_{TY}\, D^X_\xi \zeta_X \rangle\end{aligned}$$

because for $x \in Y$, the vectors $\zeta(x)$ and $\eta(x)$ lie in T_xY, so the normal component is annihilated in the scalar product. This proves the metric property, and concludes the proof of the proposition.

Next we look at the normal component. We shall obtain a canonical symmetric bilinear bundle map

$$h^{Y,X}_{12} = h_{12}\colon TY \times TY \to NY = N_XY.$$

Proposition 1.3. *Let $x \in Y$. Let $v, w \in T_xY$. Let η, ζ be sections of TY on a neighborhood of x such that $\eta(x) = v$ and $\zeta(x) = w$. Let η_X and ζ_X be extensions of η, ζ to local vector fields on X near x. Then we have the symmetric relation*

$$\mathrm{pr}_{NY}\, D^X_\eta \zeta_X(x) = \mathrm{pr}_{NY}\, D^X_\zeta \eta_X(x).$$

In particular, $\mathrm{pr}_{NY}\, D^X_\eta \zeta_X(x)$ is independent of the choice of sections η, ζ having the given values v, w at x.

Proof. By definition of the covariant derivative,

$$D^X_\eta \zeta_X - D^X_\zeta \eta_X = [\eta_X, \zeta_X] = [\eta, \zeta] \quad \text{at } x.$$

But η, ζ being sections of TY, so is $[\eta, \zeta]$. Hence the normal bundle components of $(D^X_\eta \zeta_X)(x)$ and $(D^X_\zeta \eta_X)(x)$ are the same, thus proving the formula. We know from basic definitions that $(D^X_\eta \zeta)(x)$ is independent of the choice of η, and $(D^X_\zeta \eta)(x)$ is independent of the choice of ζ. Then the last assertion follows, thus proving the proposition.

As a matter of notation, we may define

$$h_{12}(\zeta, \eta)(x) = \mathrm{pr}_{NY}\, D^X_\eta \zeta(x) = D^X_\eta \zeta(x)_{NY}$$

to denote its normal component. By abuse of notation, we omit the subscript X on ζ in light of Proposition 1.3. We may define $h_{12}(v, w)$ for $v, w \in T_xY$ by letting

SFF 1. $$h_{12}(v, w) = \mathrm{pr}_{NY}\, D^X_\eta \zeta(x) = \mathrm{pr}_{NY}\, D^X_\eta \zeta_X(x).$$

Proposition 1.3 shows that this is well defined, and symmetric, that is

$$h_{12}(v, w) = h_{12}(w, v).$$

Thus h_{12} is a symmetric bilinear vector bundle map.

In light of Propositions 1.2 and 1.3, for every point $x \in Y$, sections η, ζ of TY near x, and any extension ζ_X of ζ near x, we obtain the **Gauss formula**:

SFF 2. $$D^X_\eta \zeta_X(x) = D^Y_\eta \zeta(x) + h_{12}(\eta(x), \zeta(x)).$$

Before going to a discussion of h_{21}, we mention the significance of the condition $h_{12} = 0$. One defines Y to be **totally geodesic** if every geodesic in X with initial conditions in (Y, TY) is contained in Y.

Corollary 1.4. *The submanifold Y is totally geodesic if and only if its second fundamental form is 0 at every point. Let Y be totally geodesic. Let α be a geodesic in Y. Then α is also a geodesic in X. Let $P_{\alpha,X}$ and $P_{\alpha,Y}$ be the corresponding parallel translations. Then*

$$D^X_{\alpha'} = D^Y_{\alpha'} \quad \textit{and} \quad P_{\alpha,X} = P_{\alpha,Y}.$$

Proof. The condition that a curve α is a geodesic is that $D_{\alpha'}\alpha' = 0$. Suppose Y is totally geodesic. Let α be a geodesic in Y with $\alpha(0) = x$ and $\alpha'(0) = v \in T_xY$. Then by assumption, α is also a geodesic in X, so taking the covariant derivatives along α, we get

$$D^X_{\alpha'}\alpha' = D^Y_{\alpha'}\alpha' \quad \text{at } 0,$$

whence $h_{12}(v, v) = 0$ for all $v \in T_xY$. Since h_{12} is symmetric, it follows that $h_{12} = 0$. Conversely, suppose $h_{12} = 0$. Let α be a geodesic in X, with, say, $\alpha(a) = y \in Y$ and $\alpha'(a) \in T_yY$. Let β be the geodesic in Y with the same initial condition at y. By **SFF 2**, for any x on β in a small neighborhood of y, we have

$$D^X_{\beta'}\beta'(x) = D^Y_{\beta'}\beta'(x) = 0.$$

Hence β is also a geodesic in X. Since α and β have the same initial conditions, they are equal, thus concluding the proof of the first statement. The fact that the covariant derivatives and parallel translations are equal then follows at once from the defininition of h_{12} in Theorem 1.1. This concludes the proof of Theorem 1.4.

We have now concluded the discussion of h_{12}, and we pass to the discussion of h_{21}, and to the proof that it is minus the transpose of h_{12}.

Let η as before be a vector field on Y but now let μ be a normal field on Y. We define

$$h_{21}(\eta, \mu) = \mathrm{pr}_{TY}\, D^X_\eta \mu_X \quad \text{on } Y,$$

where as before μ_X denotes an extension of μ locally on X. The formula immediately shows that the value of h_{21} depends only on the value $\eta(x)$ at a point $x \in Y$, but we have the similar question arising about the dependence on μ and its extension μ_X. This time, the matter is settled by the next result.

Theorem 1.5. *Let η, ξ be vector fields on Y and let μ be a normal field on Y. Then on Y,*

$$\langle D^X_\eta \mu_X, \xi\rangle = \langle h_{21}(\eta, \mu), \xi\rangle = \langle \mu, -h_{12}(\eta, \xi)\rangle.$$

Proof. We take D^X_η (Lie derivative) of $\langle \xi_X, \mu_X\rangle$, evaluated at a point of Y. The scalar product is taken in TX, of course. To find the derivative at a point $x \in Y$, one may differentiate along any curve passing through that point, such that the derivative of the curve is the $\eta(x)$, and such a curve may be taken in Y. Therefore at such a point $x \in Y$, we have

$$\begin{aligned} 0 = D^X_\eta\langle \mu_X, \xi_X\rangle &= \langle D^X_\eta \mu_X, \xi_X\rangle + \langle \mu_X, D^X_\eta \xi_X\rangle \\ &= \langle \mathrm{pr}_{TY}\, D^X_\eta \mu_X, \xi_X\rangle + \langle \mu_X, \mathrm{pr}_{NY}\, D^X_\eta \xi_X\rangle \\ &= \langle h_{21}(\eta_X, \mu_X), \xi\rangle + \langle \mu, h_{12}(\eta, \xi)\rangle. \end{aligned}$$

We omit the subscript X in h_{12} because we know the independence from the extension to X by Proposition 1.3. This proves the formula of the theorem, at the same time that it also shows that $h_{21}(\eta_X, \mu_X)$ is independent of the extension of η, μ to X. Similarly, the relation shows that at the given point x, $h_{21}(\eta, \mu)(x)$ depends only on $\eta(x)$ and $\mu(x)$ respectively. This concludes the proof.

From Theorem 1.5, we see that h_{12} and h_{21} determine each other uniquely, and one is minus the transpose of the other under the metric product on X. They are both called the **second fundamental form**, but a distinction should be made between them.

Theorem 1.5 allows us to write the formula analogous to **SFF 1**, namely for $x \in Y$, $v \in T_xY$ and $w \in N_xY$, $\eta(x) = v$, $\mu(x) = w$, we have

SFF 3. $$h_{21}(v, w) = \mathrm{pr}_{TY}\, D^X_\eta \mu(x).$$

Note that Theorem 1.5 also concludes the proof of formula (2) in Theorem 1.1.

There remains but to deal with the fourth component $h_{22}(\eta, \mu)$, where η is a vector field on Y and μ is a normal field. For the first time, we have to deal with the more general notion of a covariant derivative on a vector bundle. Quite generally, let E be a vector bundle over Y. A **derivative** on E **relative to** TY is a mapping

$$\nabla\colon\ \Gamma TY \times \Gamma E \to \Gamma E$$

denoted by $(\eta, \mu) \mapsto \nabla_\eta\mu$, which is $\mathrm{Fu}(Y)$-linear in η, and is a derivation in μ, that is for any function φ on Y,

$$\nabla_\eta(\varphi\mu) = (\eta\cdot\varphi)\mu + \varphi\nabla_\eta\mu.$$

Suppose E has a metric. We say that ∇ is **metric** if it satisfies the condition

$$\eta\cdot\langle \mu, \nu\rangle = \langle \nabla_\eta\mu, \nu\rangle + \langle \mu, \nabla_\eta\nu\rangle$$

for all vector fields η on Y, and sections μ, ν of E. We shall apply this notion to the normal bundle $E = NY$.

Proposition 1.6. *Let μ_X be an extension of a normal field to X. Then $\mathrm{pr}_{NY}\, D^X_\eta \mu_X$ is independent of this extension, so we may denote*

$$\nabla_\eta\mu = \mathrm{pr}_{NY}\, D^X_\eta \mu_X.$$

Furthermore, ∇ is a metric derivative on NY.

Proof. We prove the metric formula first. By definition of the covariant derivative on X, we know that on Y, for any normal field ν,

$$\eta \cdot \langle \mu_X, \nu_X \rangle = \langle D_\eta^X \mu_X, \nu_X \rangle + \langle \mu_X, D_\eta^X \nu_X \rangle.$$

For $x \in Y$, the values $\mu_X(x)$ and $\nu_X(x)$ lie in $N_x Y$, so the covariant derivatives in the above relation can be replaced by their projections on the normal bundle NY. The Lie derivative on the left can be computed at x along curve whose derivative at x is $\eta(x)$, and this curve can be taken to lie entirely in Y. Therefore the left side is independent of the extensions μ_X, ν_X of μ, ν locally near x, so we may write it as $\eta \cdot \langle \mu, \nu \rangle$. Then we write

$$\langle (D_\eta^X \mu_X)(x), \nu(x) \rangle = (\eta \cdot \langle \mu, \nu \rangle)(x) - \langle \mu(x), (D_\eta^X \nu_X)(x) \rangle.$$

The right side is independent of the extension μ_X of μ, and therefore so is the left side. Similarly for ν_X. Thus we have proved simultaneously the metric formula and the independence which allows us to define $\nabla_\eta \mu$. Note that the Fu(Y)-linearity in η is then immediate from the metric formula. The derivation property in μ follows from that of D_η^X. This concludes the proof.

XIV, §2. THE HESSIAN AND LAPLACIAN ON A SUBMANIFOLD

We continue with a submanifold Y of a Riemannian manifold X. We remind the reader of the Hessian of a function f on Y. We need here only formula (1) of Chapter XIII, §1. For vector fields ξ, η on Y, the Hessian is

$$D_Y^2 f(\xi, \eta) = \xi \cdot \eta \cdot f - (D_\xi^Y \eta) \cdot f.$$

We put Y as a subscript on the left for typographical reasons, involving the square as a superscript.

Proposition 2.1. *Let f_X be an extension of f to X. Let ξ, η be vector fields on Y. Then on Y, we have*

$$D_X^2 f_X(\xi, \eta) = D_Y^2 f(\xi, \eta) - h_{12}(\xi, \eta) \cdot f_X,$$

where $h_{12}(\xi, \eta) = \mathrm{pr}_{NY}\, D_\xi^X \eta_X$ as in §1.

Proof. We have

$$D_X^2 f_X(\xi_X, \xi_X) = \xi \cdot \eta \cdot f - (D_\xi^X \eta_X) \cdot f_X.$$

By Theorem 1.1, at points of Y we have

$$D_\xi^X \eta_X = (D_\xi^Y \eta) + h_{12}(\xi, \eta),$$

which concludes the proof by definition of D_Y^2.

The tangential component

We can use the normal bundle to obtain a tubular neighborhood of Y. Locally, we can find a function $r > 0$ such that, if $N_r Y$ denotes the vectors w with norm $\|w\| < r(x)$ for $w \in N_x Y$, $x \in Y$, then the exponential map

$$\exp\colon NY \to X \qquad \text{given by} \qquad w \mapsto \exp_x(w) \qquad \text{for } w \in N_x Y$$

gives an isomorphism of $N_r Y$ with an open neighborhood of Y in X. Given a function f on Y, we may extend f to this tubular neighborhood by making f constant in the normal directions, that is we define

$$f_X(\exp_x(w)) = f(x).$$

This extension will be called the **normal extension** of f to a tubular neighborhood of Y.

Proposition 2.2. *Let f_X be the normal extension of f to a tubular neighborhood of Y. Then for vector fields ξ, η on Y, we have*

$$D_X^2 f_X(\xi, \eta) = D_Y^2 f(\xi, \eta).$$

Proof. This is immediate, because if ν is a normal vector field on Y, then $(\nu \cdot f_X)(x) = 0$ for $x \in Y$, immediately from the definitions. Indeed, the Lie derivative may be taken along a geodesic from x, along which f is constant, so its Lie derivative is 0. We can apply Proposition 2.1 with $\nu = h_{12}(\xi, \eta)$ to conclude the proof.

Next we look at normal fields on Y.

Proposition 2.3. *Let ν be a normal field on Y. Let f be a function on Y and f_X its normal extension to a tubular neighborhood of Y. Then on Y,*

$$D_X^2 f_X(\nu, \nu) = 0.$$

Proof. Let ν_X be any extension of ν to a neighborhood of a point x_0 in in Y. Then at x_0,

$$(1) \qquad D_X^2 f_X(\nu, \nu) = \nu_X \cdot \nu_X \cdot f_X - (D_{\nu(x_0)} \nu_X) \cdot f_X.$$

We select a suitable extension of v. For $y \in Y$ near x_0, $w \in N_y Y$ and $|w|$ sufficiently small, let $\alpha_{y,w}$ be the geodesic in X with $\alpha_{y,w}(0) = y$ and $\alpha'_{y,w}(0) = w$. Thus $\exp_y(w) = \alpha_y(1, w) = \alpha_{y,w}(1)$. Define the **normal extension** v_X be the formula

$$v_X(\exp_y(w)) = P^1_{0,\alpha_{y,w}}(v(y)).$$

where P is parallel translation as in Chapter VIII, Theorems 3.3 and 3.4. Then $v_X f_X = 0$, so the first term on the right of (1) is 0. As for the second term, letting $\alpha = \alpha_{x_0,w}$ with $w = v(x_0)$, α is a geodesic so $D_{\alpha'}\alpha' = 0$, and we get

$$(D_{v(x_0)}v_X)(x_0) = (D_{\alpha'}\alpha')(x_0) = 0.$$

So having chosen v_X suitably, we conclude that both terms are 0, which proves the proposition.

Theorem 2.4. *Let X be a finite dimensional Riemannian manifold, and let Y be a submanifold. Let f be a function on Y and let f_X be its normal extension to X. Then on Y,*

$$\Delta_Y f = \Delta_X f_X.$$

Proof. Let $\{\xi_1, \ldots, \xi_p\}$ be an orthonormal frame of vector fields locally on Y, and let $\{v_1, \ldots, v_q\}$ be an orthonormal frame of normal fields. Together they form an orthonormal frame of sections of TX restricted to Y. Then at a point $x \in Y$, we have

$$\Delta_X f_X(x) = -\sum D^2_X f_X(\xi_i, \xi_i) - \sum D^2_X f_X(v_j, v_j) \quad \text{at } x.$$

We may now apply Propositions 2.2 and 2.3 to conclude the proof.

Remark. Readers may compare the above proof with that of Helgason [He 84], Chapter II, Theorem 3.2. Of course, the theorem on the Laplacian depends on the manifolds being finite dimensional. However, the basic result behind it concerns the Hessian, and is independent of this restrictive condition.

Full decomposition of the Laplacian

We shall now return to the use of Proposition 2.1. The rest of this section was written with Wu's collaboration.

For our next purposes, we let f be a function on X and we let f_Y be its restriction to the submanifold Y. We let:

$\{\xi_1, \ldots, \xi_p\}$ be an orthonormal frame of vector fields on Y;
$\{v_1, \ldots, v_q\}$ be an orthonormal frame of normal vector fields.

Thus $\{v_1(x), \ldots, v_q(x)\}$ is an orthonormal basis of $N_x Y$ for $x \in Y$. Letting as usual Δ_X be the Laplacian on X, by the definition of Chapter XII, §1, and Proposition 2.1, we have

$$(2) \quad \Delta_X f = \sum_{i=1}^{p} -D^2_X f(\xi_i, \xi_i) + \sum_{j=1}^{q} -D^2_X f(v_j, v_j)$$

$$(3) \quad = \sum_{i=1}^{p} -D^2_X f_Y(\xi_i, \xi_i) + \sum_{i=1}^{p} h_{12}(\xi_i, \xi_i)\cdot f + \sum_{j=1}^{q} -D^2_X f(v_j, v_j).$$

By Proposition 2.2, the first term on the right is just $\Delta_Y f_Y$. Note that the second and third terms involve normal components, and thus it is natural to define the **Y-transversal Laplacian**

$$(4) \quad \Delta^T_{X,Y} f = \sum_{i=1}^{p} h_{12}(\xi_i, \xi_i)\cdot f - \sum_{j=1}^{q} D^2_X f(v_j, v_j).$$

Then using Proposition 2.2, we may reformulate (3) in the form:

Proposition 2.5. *Let X be a finite dimensional Riemannian manifold and Y a submanifold. Let f be a function on X and let f_Y be its restriction to Y. Then*

$$\Delta_X f = \Delta_Y f_Y + \Delta^T_{X,Y} f \quad \text{on } Y.$$

Thus the Laplacian has been decomposed into a tangential component, which is the Laplacian on Y, and a transversal component. Note that Theorem 2.4 describes the special case of the tangential component, for functions which are constant in the normal direction. For convenience, we also define the **normal trace**

$$(\mathrm{tr}_{N,Y}\, D^2_X) f = \sum_{j=1}^{q} D^2_X f(v_j, v_j).$$

Fix a point $x_0 \in Y$, and let W'_0 be an open ball centered at 0 in the normal space $N_{x_0} Y$. We suppose W'_0 sufficiently small that the exponential map

$$\exp_{x_0}\colon W'_0 \to W_0$$

is an isomorphism of W'_0 onto its image W_0, which is a submanifold of X

called the **normal submanifold** to Y at x_0. Let f be a function on X. Instead of considering its restriction to Y, we now consider its restriction f_{W_0} to the normal submanifold W_0. This restriction depends on the choice of point x_0 in Y. For the moment, we don't use finite dimensionality.

Proposition 2.6. *Let X be a Riemannian manifold, Y a submanifold, and $x_0 \in Y$. Let W_0 be the normal submanifold to Y at x_0. Let f be a function on X, and let $w \in N_{x_0}Y$. Then $D_X^2 f(w, w)$ depends only on the restriction f_{W_0}. More precisely, if α is the geodesic defined by $\alpha(t) = \exp_{x_0}(tw)$, then*

$$D_X^2 f(w, w) = (D_{\alpha'}^2 f_{W_0})(x_0).$$

Furthermore, for every $v \in T_{x_0}Y$, the value $h_{12}(v, v) \cdot f$ at x_0 also depends only on the restriction f_{W_0}.

Proof. By the Killing definition of the second tensorial derivative (Chapter XIII, §1) and Corollary 3.2 of Chapter VIII, §3 we may compute this derivative along the geodesic, that is

$$D_X^2 f(w, w) = (D_{\alpha'} D_{\alpha'} f)(x_0) - (D_{\alpha'}\alpha') \cdot f(x_0).$$

Since α is a geodesic, the second term on the right vanishes, and the first term depends only on f along the geodesic $t \mapsto \exp_{x_0}(tw)$, and so depends only on the transversal part f_{W_0}. This concludes the proof of the first part. The second statement is even simpler, because the derivative $h_{12}(v, v) \cdot f$ at x_0 may be computed by using the same geodesic

$$\alpha(t) = \exp_{x_0}(tw), \qquad \text{with} \qquad w = h_{12}(v, v).$$

This concludes the proof.

Proposition 2.7. *Suppose X finite dimensional. Let Y be a submanifold, and $x_0 \in Y$. Let W_0 be the normal submanifold of Y at x_0. Let f be a function on X. Then*

$$((\mathrm{tr}\ h_{12}) \cdot f)(x_0) = (\mathrm{tr}\ h_{12})(x_0) \cdot f_{W_0}$$

and thus finally

$$\Delta_X f(x_0) = \Delta_Y f_Y(x_0) + (\mathrm{tr}\ h_{12}) \cdot f(x_0) - \mathrm{tr}_{N,Y}\ D_X^2 f(x_0).$$

Proof. Immediate from (4) and Proposition 2.6, using $w = w_j = v_j(x_0)$ and $v = v_i = \xi_i(x_0)$.

We note the symmetry between the submanifold Y and its normal submanifold W_0 at the given point x_0. In Proposition 2.3, when we take the normal extension of a function on Y, the normal term vanishes. Similarly, if f_Y is constant on a neighborhood of x_0 in Y, then the term $\Delta_Y f_Y(x_0)$ vanishes.

We may now extend the above more globally as follows. Let

$$\pi\colon X \to Z$$

be a submersion. For each $z \in Z$ we let $Y_z = \pi^{-1}(z)$ be the fiber above z. Then Y_z is a submanifold, to which we can apply Propositions 2.5, 2.6 and 2.7. Furthermore, we can use (4) to define the normal part of the Laplacian depending on π. We had already defined the trace of the second fundamental form h_{12}. We now define the **normal trace** $\mathrm{tr}_{N,\pi}\ D_X^2$ in the similar way, namely for any function f,

$$(\mathrm{tr}_{N,\pi}\ D_X^2)(f) = \sum_{j=1}^{q} D_X^2 f(v_j, v_j).$$

Then we define the **transversal** part of the Laplacian by the formula

$$\Delta_{T,\pi} = [\mathrm{tr}\ h_{12}] - \mathrm{tr}_{N,\pi}\ D_X^2,$$

where $[\mathrm{tr}\ h_{12}]$ denotes the Lie derivative $\mathscr{L}_{\mathrm{tr}\ h_{12}}$ to simplify the notation. Having fixed the submersion, we may omit π from the notation, but the definition of the normal part depends on the choice of submersion, because the traces depends on the submersion (the submanifolds Y_z and their normal submanifolds).

Of course, we may give a similar definition for the **vertical** or **tangential** part of the Laplacian $\Delta_{V,\pi}$, namely for $x \in Y_{\pi(x)} = Y$,

$$(\Delta_{V,\pi} f)(x) = \Delta_Y f_Y(x).$$

Proposition 2.8. *Suppose X finite dimensional. Let $\pi\colon X \to Z$ be a submersion. Then*

$$\Delta_X = \Delta_{V,\pi} + \Delta_{T,\pi}.$$

Proof. This is just a reformulation of Proposition 2.7, taking the previous definitions into account.

Example 1. Let Y be a submanifold of X and fix a point $x_0 \in Y$. Let V_0 be an open ball centered at x_0 in Y. Let W_0' be a neighborhood of 0 in the normal space $N_{x_0}Y$. For V_0 sufficiently small, there exists a unique geodesic in X from x_0 to x. For W_0' sufficiently small and $w \in W_0'$, we define the map

$$\varphi\colon V_0 \times W_0' \to X \quad \text{by the formula} \quad \varphi(x, w) = \exp_x P_{x_0}^x(w),$$

where $P_{x_0}^x$ is parallel translation along the geodesic from x_0 to x.

Lemma 2.9. *The above map φ is a local isomorphism at $(x_0, 0)$. Its differential at this point is in fact the identity.*

Proof. This is a routine verification left to the reader. Note that the tangent space of $V_0 \times W_0'$ at the point is precisely $T_{x_0}Y \times N_{x_0}Y$, which we identify with $T_{x_0}X$. The second statement about the differential implies the first about the local isomorphism by the inverse mapping theorem.

We note that the lemma provides a local product decomposition. Let $U_0 = \varphi(V_0 \times W_0')$, so U_0 is an open neighborhood of x_0 in X. The projection

$$\pi\colon \varphi(V_0 \times W_0') = U_0 \to W_0'$$

is a submersion to which we can apply Proposition 2.8.

Example 2. Let H be a Lie group acting smoothly on X as a group of metric automorphisms. We say that H **acts regularly**, or that the action is **regular**, if there exists a submersion

$$\pi\colon X \to Z$$

such that the fibers are the orbits of H. For instance, the orthogonal group $O(n) = \mathrm{Uni}_n(\mathbf{R})$ acts regularly on $\mathbf{R}^n$ from which the origin is deleted. Under a regular action, for each $x \in X$ the map $H \to Hx$ (the orbit) given by $h \mapsto hx$ gives an embedding of H/H_x in X, so gives an isomorphism with H/H_x and the orbit Hx. Fix a point $x_0 \in X$. The map

$$H \to H/H_{x_0}$$

being a submersion, there exists a local section $\sigma\colon V_0' \to H$ defined on an open neighborhood of the identity coset eH_{x_0}, and passing through e, so $\sigma(eH_{x_0}) = e$. We put $V_0 = \sigma(V_0')$. We let Y_0 be the orbit Hx_0. We note that we have a natural linear isomorphism of tangent spaces

$$T_{x_0}V_0 \xleftrightarrow{\approx} T_{x_0}Y_0.$$

Let W_0' be an open neighborhood of 0 in the normal space $N_{x_0}Y_0$, equal to the orthogonal complement of $T_{x_0}Y_0$ in $T_{x_0}X$, such that the exponential

map

$$\exp\colon W_0' \to \exp(W_0')$$

is an isomorphism. Put $W_0 = \exp(W_0')$. We call (e, x_0) the **origin** of $V_0 \times W_0$.

Lemma 2.10. *Under a regular action by H, the map*

$$\varphi\colon V_0 \times W_0 \to X \quad \text{given by} \quad (h, x) \mapsto hx$$

is a local isomorphism at the origin (e, x_0).

Proof. This is a simple exercise in computing the differential of the map at the origin, and showing that it is the identity.

As for Example 1, we may then apply Proposition 2.8 to the submersion

$$\varphi(V_0 \times W_0) = U_0 \to W_0.$$

Example 2 is essentially the one used by Helgason to construct his transversal part of the Laplacian [He 84], Chapter II, §3, especially Theorems 3.4 and 3.5. He does not use the second fundamental form, but uses a construction applicable to all differential operators. This generality requires some general results, notably his Theorem 1.4 characterizing differential operators. Such considerations are completely bypassed by the direct local differential geometric approach used in the present section.

Note that the submersion used in Proposition 2.8 is just that. No other requirement is made. In the next section, we shall consider a stronger version, with an additional metric condition.

Example 3. In Chapter XV, Theorem 3.8, we shall describe the polar decomposition of the Laplacian, in a normal chart, namely let $\exp_x\colon \mathbf{B}_c(0_x) \to B_c(x)$ be a differential isomorphism for some $c > 0$. After deleting the origin, the ball is isomorphic to a product $\mathrm{S}_1(x) \times (0, c)$, projecting on the open interval $(0, c)$. The submanifolds are the spheres $S_r(x)$, $0 < r < c$. The transversal part is called the **radial part** in this case.

XIV, §3. THE COVARIANT DERIVATIVE ON A RIEMANNIAN SUBMERSION

Let X, Z be Riemannian manifolds, and let

$$\pi\colon X \to Z$$

be a submersion. We assume that π is **Riemannian**, *meaning that for each*

$x \in X$, *the differential*

$$T\pi(x): T_xX \to T_{\pi(x)}Z$$

is an orthogonal projection. For each $z \in Z$ we let $Y_z = \pi^{-1}(z)$ be the fiber. Then Y_z is a submanifold of X, and the kernel of $T\pi(x)$ is $T_xY_{\pi(x)}$. We also have the normal bundle $N_xY_{\pi(x)}$, and normal sections. By definition, $T\pi(x)$ induces a linear metric isomorphism

$$\pi_* = T\pi(x): N_xY_{\pi(x)} \to T_{\pi(x)}Z.$$

I am indebted to Wu for his explanation of the behavior of the Laplacian in submersions, which led to the exposition of this section.

Lemma 3.1. *Let $x \in Y_{\pi(x)}$ be a point in a fiber. Let f be a function on Z. Then for $w \in N_xY_{\pi(x)}$ we have*

$$(D_w\pi^*f)(x) = (D\pi_{*w}f)(\pi(x)),$$

or in different notation, if v is a normal field at x, then

$$(v \cdot \pi^*f)(x) = (\pi_*v(x) \cdot f)(\pi(x)).$$

On the other hand, if $v \in T_xY_{\pi(x)}$, then

$$(D_v\pi^*f)(x) = 0.$$

Proof. One may prove the formulas in a chart, in which case both merely come from the chain rule

$$(f \circ \pi)'(x) = f'(\pi(x))T\pi(x),$$

applied to any vector in $T_xX = T_xY_{\pi(x)} + N_xY_{\pi(x)}$. So the lemma is clear.

The tangent bundle TX has an orthogonal sum decomposition into two subbundles

$$TX = F \perp E,$$

where at a point x, $F_x = T_xY_{\pi(x)}$ is the tangent space to the fiber, and $E_x = N_xY_{\pi(x)}$ is the space normal to the fiber. One also calls F the **vertical subbundle** and E the **horizontal subbundle**. The differential $T\pi$ gives a metric isomorphism at each point

$$T\pi(x): E_x \to T_xZ.$$

A vector field μ on Z lifts uniquely to a horizontal field μ_X, i.e. a vector field such that

$$\mu_X(x) \in E_x = N_xY_{\pi(x)} \quad \text{and} \quad T\pi(x)\mu_X(x) = \mu(\pi(x)),$$

at each point $x \in X$. We call μ_X the **horizontal lifting** of μ. On the other hand, a vector field on X having values in F is called a **vertical field**. Both notions are of course relative to the submersion π, and one could write $F(\pi)$ and $E(\pi)$ to bring π into the notation. But π is now fixed, so we omit it from the notation.

Next we have some formulas for the lifting to normal fields. First,

$$(1) \qquad \pi_*[\mu_X, \nu_X] = [\mu, \nu].$$

The proof is immediate no matter what, and can be verified in a chart.
We also have the tangential component, i.e. for any vertical field ξ,

$$(2) \qquad \langle[\mu_X, \nu_X], \xi\rangle = \langle\mu_X, D_{\nu_X}\xi\rangle - \langle\nu_X, D_{\mu_X}\xi\rangle.$$

In particular, the value of the vertical component of $[\mu_X, \nu_X]$ at a point x depends only on $\mu_X(x)$, $\nu_X(x)$. To prove (2), we first write the defining formula

$$\langle[\mu_X, \nu_X], \xi\rangle = \langle D_{\mu_X}\nu_X - D_{\nu_X}\mu_X, \xi\rangle.$$

We use the fact that $\langle\nu_X, \xi\rangle = \langle\mu_X, \xi\rangle = 0$. We apply D_{μ_X} and D_{ν_X}, respectively, to these equalities, and use the defining property of the metric derivative. Then (2) falls out.

We shall use the formula giving the metric derivative explicitly, namely **MD 2** of Chapter VIII, §4. For any vector field ξ on X, we have

$$(3) \qquad 2\langle D_{\mu_X}\nu_X, \xi\rangle = \mu_X \cdot \langle\nu_X, \xi\rangle + \nu_X \cdot \langle\mu_X, \xi\rangle - \xi \cdot \langle\mu_X, \nu_X\rangle$$
$$+ \langle[\mu_X, \nu_X], \xi\rangle - \langle[\mu_X, \xi], \nu_X\rangle - \langle[\nu_X, \xi], \mu_X\rangle.$$

Proposition 3.2. *Let μ, ν be vector fields on Z, and μ_X, ν_X their horizontal liftings to X. Then*

$$\mathrm{pr}_E(D_{\mu_X}\nu_X) = (D_\mu\nu)_X,$$

or equivalently, for every horizontal field λ_X,

$$\langle D_{\mu_X}\nu_X, \lambda_X\rangle = \langle D_\mu\nu, \lambda\rangle.$$

Proof. The expression $\langle D_{\mu_X} v_X, \lambda_X \rangle$ coming from (3) involves only the Lie derivative, scalar product of vector fields and brackets. The scalar product is preserved under lifting, by definition of a Riemannian submersion. Formula (1) gives the preservation of the bracket. The Lie derivative is also preserved under lifting by Lemma 3.1. This concludes the proof.

The rest of this section will not be used until §6.

Proposition 3.3. *Let* μ, v, λ, ζ *be vector fields on* Z. *Then*

$$\mu_X \cdot \langle D_{v_X} \lambda_X, \zeta_X \rangle = \pi^*(\mu \cdot \langle D_v \lambda, \zeta \rangle).$$

Proof. Again, direct consequence of (3) and Proposition 3.2.

Next we determine the vertical component. If η is a vector field on X, we define its **vertical component** be

$$\eta^V = \mathrm{pr}_F\, \eta \quad \text{where } F \text{ is the vertical subbundle of } TX.$$

Proposition 3.4. *Let* μ, v *be vector fields on* Z. *Then*

$$D_{\mu_X} v_X = \tfrac{1}{2}[\mu_X, v_X]^V + (D_\mu v)_X.$$

Proof. The horizontal component was already determined in Proposition 3.2, which gives the second term on the right of the equation. As for the vertical component, we use (3) with a vertical field ξ. Since $\langle \mu_X, v_X \rangle = \langle \mu, v \rangle$, if ξ is vertical, we have $\xi \cdot \langle \mu_X, v_X \rangle = 0$. The first two terms and the last two terms of (3) on the right vanish by (1). The value for the vertical component then drops out, thus proving the proposition.

Proposition 3.5. *Let* $\alpha\colon [a, b] \to Z$ *be a curve such that* $\alpha'(t) \neq 0$ *for all* t.

(i) *Let* $y \in Y_{\alpha(a)}$. *There exists a unique lifting* $A = A_y$ *of* α *in* X *which is horizontal, i.e. such that* $A'(t)$ *lies in the horizontal subbundle for all* t, *and with the given initial condition* $A(a) = y$.

(ii) *The curve* α *is a geodesic if and only if* A *is a geodesic.*

(iii) *For each* y, *define* $F(y, t) = A_y(t)$, *and let* $F_t(y) = A_y(t)$. *Then* $F_t\colon Y_{\alpha(a)} \to Y_{\alpha(t)}$ *is a differential isomorphism.*

Proof. The existence and uniqueness of the lifting are elementary, at the level of the existence and uniqueness of solutions of a differential equation. We give the details. The global assertion is a consequence of local existence and uniqueness, so we may suppose that there is a vector field v locally on Z such that $v(\alpha(t)) = \alpha'(t)$ for all t, i.e. v extends α'. For simplicity of notation, shrinking Z if necessary to some open subset, we

suppose v is defined on all of Z. Let $y \in X$. By the fundamental theorem on differential equations, there exists a unique curve $A\colon [a, b] \to X$ such that $A'(t) = v_X(A(t))$ for all t. We claim that A lifts α, that is $A(t) \in Y_{\alpha(t)}$ (the fiber above $\alpha(t)$). Indeed,

$$(\pi \circ A)'(t) = T\pi(A(t))A'(t) = T\pi(A(t))v_X(A(t)),$$

and $v_X(A(t)) \in E_{A(t)}$. Let $\beta = \pi \circ A$. Then β satisfies the differential equation $\beta'(t) = v(\beta(t))$, with the same initial conditions as α, so $\beta = \alpha$, and thus A lifts α. As for uniqueness, suppose v_1, v_2 are two extensions of α' to local vector fields on Z. Let A_1, A_2 be the liftings of α corresponding to these two vector fields. Then they satisfy $A_1'(t) = A_2'(t)$ for all t, and so they are equal, thus proving the first part of the proposition. For the geodesic property, we put $\mu = v$ in Proposition 3.4. Then the bracket term on the right is 0. We evaluate along α'. Then (ii) follows from the characterization of a geodesic by the condition $D_{\alpha'}\alpha' = 0$.

Finally (iii) is now essentially formal. Say for $t = b$, we consider the reverse curve of α, and its lift from $A(b)$ which is necessarily the reverse curve of A by uniqueness. Hence F_b has an inverse mapping. This concludes the proof of Proposition 3.5.

In Proposition 3.2 we considered horizontal fields. The next proposition gives a similar result for differentiation with respect to a vertical field.

Proposition 3.6. *Let* ξ *be a vertical field. Then*

$$\langle D_\xi \mu_X, v_X \rangle = -\tfrac{1}{2}\langle [\mu_X, v_X]^V, \xi \rangle.$$

Proof. By the metric derivative formula (3) and Proposition 3.4, we obtain

$$\begin{aligned}\langle D_\xi \mu_X, v_X \rangle &= \langle D_{\mu_X} \xi, v_X \rangle + \langle [\xi, \mu_X], v_X \rangle \\ &= -\langle D_{\mu_X} v_X, \xi \rangle \\ &= -\tfrac{1}{2}\langle [\mu_X, v_X], \xi \rangle \\ &= -\tfrac{1}{2}\langle [\mu_X, v_X]^V, \xi \rangle,\end{aligned}$$

thereby proving the proposition.

XIV, §4. THE HESSIAN AND LAPLACIAN ON A RIEMANNIAN SUBMERSION

We continue with a Riemannian submersion

$$\pi\colon X \to Z$$

as in §3, but we shall use only Lemma 3.1 and Proposition 3.2. We shall deal with the Hessian. As in §3, we let F and E be the vertical and horizontal subbundles respectively, giving rise to the two orthogonal projections pr_F and pr_E. We may apply these in the way we did in the previous sections, along each fiber $Y_{\pi(x)}$.

Proposition 4.1. *Let ξ, η be vertical fields on X. Then for every function f on Z, we have*

$$\begin{aligned}(D_X^2\pi^*f)(\xi, \eta) &= -h_{12}(\xi, \eta)\cdot\pi^*f \\ &= -\pi_*h_{12}(\xi, \eta)\cdot f.\end{aligned}$$

Proof. We have

$$\begin{aligned}D^2\pi^*f(\xi, \eta) &= \xi\cdot(\eta\cdot\pi^*f) - (D_\xi\eta)\cdot\pi^*f \\ &= -(D_\xi\eta)\cdot\pi^*f,\end{aligned}$$

because $\eta\cdot\pi^*f = 0$ since π^*f is constant on the fibers and $\eta\cdot\pi^*f$ can be computed along a curve contained in the fiber $Y_{\pi(x)}$. Furthermore, the constancy of f on a fiber also yields

$$(D_\xi\eta)\cdot\pi^*f = \text{pr}_E(D_\xi\eta)\cdot\pi^*f.$$

Then Lemma 3.1 and Proposition 2.1 conclude the proof.

Suppose that the fibers are finite dimensional, of dimension p. Let $\xi_1, \ldots, \xi_p$ be an orthonormal frame of vertical fields locally on X. Recall that in §1, we defined the trace of the second fundamental form to be

$$\text{tr } h_{12} = \sum_{i=1}^{p} h_{12}(\xi_i, \xi_i).$$

Proposition 4.2. *With a vertical orthonormal frame $\xi_1, \ldots, \xi_p$, and a function f on Z, we have*

$$\sum_{i=1}^{p}(D_X^2\pi^*f)(\xi_i, \xi_i) = -(\text{tr } h_{12})\cdot\pi^*f = -\pi_*\text{tr } h_{12}\cdot f.$$

Next we go to horizontal fields.

Proposition 4.3. *Let μ, ν be vector fields on Z, with horizontal liftings μ_X, ν_X. Then*

$$D_X^2\pi^*f(\mu_X, \nu_X) = D_Z^2 f(\mu, \nu).$$

Proof. We have

$$D_Z^2 f(\mu, \nu) = \mu\cdot\nu\cdot f - (D_\mu\nu)\cdot f,$$

and the similar expression on X with subscript X. The vertical component of $D_{\mu_X}\nu_X$ annihilates π^*f because π^*f is constant on fibers. For the horizontal component, Proposition 3.2 shows that the last terms on the right on X and on Z give the same value. As to the first term on the right, Lemma 3.1 shows that

$$\nu_X\cdot\pi^*f = \pi^*(\nu\cdot f),$$

so doing the same thing with μ_X shows that the first terms on the right of the equation on X and Z give the same value. This concludes the proof.

We shall give the relation between the Laplacians as an application. I am indebted to Wu for the next theorems.

Theorem 4.4. *Assume that X, and hence Z, are finite dimensional. Then for all functions f on Z we have*

$$\Delta_X\pi^*f = \pi^*\Delta_Z f + (\text{tr } h_{12})\cdot\pi^*f.$$

Proof. Let $\{\xi_1, \ldots, \xi_p\}$ be an orthonormal frame of local sections of the vertical bundle F, and let $\{\mu_1, \ldots, \mu_q\}$ be an orthonormal frame of sections on Z. Let $\{\mu_{1X}, \ldots, \mu_{qX}\}$ be their lifts to the horizontal bundle. Then

$$\{\xi_1, \ldots, \xi_p, \mu_{1X}, \ldots, \mu_{qX}\}$$

is a local orthonormal frame on X. We get:

$$\begin{aligned}\Delta_X\pi^*f &= -\sum_i D_X^2\pi^*f(\xi_i, \xi_i) - \sum_j D_X^2\pi^*f(\mu_{jX}, \mu_{jX}) \\ &= (\text{tr } h_{12})\cdot\pi^*f - \sum_j D_Z^2 f(\mu_j, \mu_j)\end{aligned}$$

by Propositions 4.2 and 4.3 respectively. This proves the theorem.

The trace of the second fundamental form is defined on X. Under some homogeneity condition that we shall now describe, we can descend it to Z.

Let σ be a metric automorphism of X, preserving the fibers, that is σ induces a differential metric automorphism of each fiber $Y_{\pi(x)}$ for all x. Then

$$\pi\circ\sigma = \pi \quad \text{so} \quad \pi_*\circ\sigma_* = \pi_*.$$

Let $x_1, x_2 \in Y_{\pi(x)}$ be points in the same fiber, and suppose $\sigma x_1 = x_2$. Let ξ, η be vertical fields on X. Then

$$\sigma_*(h_{12}(\xi, \eta)(x_1)) = h_{12}(\sigma_*\xi, \sigma_*\eta)(x_2).$$

With a vertical frame $\{\xi_1, \ldots, \xi_p\}$ as above, we can then define the **trace of** h_{12} **on** Z, by the formula

$$\mathrm{tr}_Z\, h_{12} = \pi_*\, \mathrm{tr}\, h_{12} \quad \text{so for} \quad x \in X, \quad (\mathrm{tr}_Z\, h_{12})(x) = \sum_{i=1}^{p} \pi_*(h_{12}(\xi_i, \xi_i)(x)).$$

Suppose that there is a group of metric automorphisms of X, preserving the fibers, and acting transitively on the fibers. Then given any two points x_1, x_2 in the same fiber $Y_{\pi(x)}$, it follows that

$$\mathrm{tr}_Z\, h_{12}(x_1) = \mathrm{tr}_Z\, h_{12}(x_2),$$

and therefore we may view $\mathrm{tr}_Z\, h_{12}$ as a vector field on Z, which we call the **trace of the second fundamental form on Z**. Then Theorem 4.4 can be formulated as follows.

Theorem 4.5. *Suppose that X, and hence Z, are finite dimensional. Suppose also that there is a group of isometries of X, preserving the fibers and acting transitively on each fiber. Let $\mathrm{tr}_Z\, h_{12}$ be the trace of the second fundamental form on Z. Then for all functions f on Z, we have*

$$\Delta_X \pi^* f = \pi^*(\Delta_Z f + (\mathrm{tr}_Z\, h_{12}) \cdot f).$$

Remark. Readers may compare the above version with Helgason [He 84], Chapter II, Theorem 3.7. To obtain the version in Helgason, there remains to identify the trace of the second fundamental form with the gradient of the appropriate function, which we shall do in Chapter XV, §6 and §8.

XIV, §5. THE RIEMANN TENSOR ON SUBMANIFOLDS

Let Y be a submanifold of a Riemannian manifold X. Then we have the two Riemann tensors R_Y and R_X, which we wish to compare on Y.

Theorem 5.1 (Gauss Equation). *For v_i $(i = 1, 2, 3, 4)$ in T_xY, we have*

$$\begin{aligned} R_X(v_1, v_2, v_3, v_4) = R_Y(v_1, v_2, v_3, v_4) \\ + \langle h_{12}(v_2, v_3), h_{12}(v_1 \cdot v_4)\rangle - \langle h_{12}(v_2 \cdot v_4), h_{12}(v_1, v_3)\rangle. \end{aligned}$$

Or, if ξ, η, ζ, τ are vector fields on Y,

$$\begin{aligned} R_X(\xi, \eta, \zeta, \tau) = R_Y(\xi, \eta, \zeta, \tau) \\ + \langle h_{12}(\eta, \zeta), h_{12}(\xi, \tau)\rangle - \langle h_{12}(\eta, \tau), h_{12}(\xi, \zeta)\rangle. \end{aligned}$$

Proof. The proof is routine, and forced. We have by Theorem 1.1, or **SFF 2** in §1, on Y:

$$D_\eta^Y \zeta = D_\eta^X \zeta_X + h_{12}(\eta, \zeta);$$

so iterating,

$$D_\xi^Y D_\eta^Y \zeta = \mathrm{pr}_{TY}\left(D_\xi^X D_\eta^X \zeta_X + D_\xi^X (h_{12}(\eta, \zeta)_X)\right).$$

We interchange ξ and η and subtract. We also note that

$$[\xi, \eta] \cdot \zeta = \mathrm{pr}_{TY}[\xi_X, \eta_X] \cdot \zeta_X \quad \text{on } Y.$$

Hence by the definition of the Riemann tensor, for all vector fields τ on Y,

$$\begin{aligned} \langle R_Y(\xi, \eta)\zeta, \tau\rangle = \langle R_X(\xi, \eta)\zeta, \tau\rangle \\ - \langle D_\eta^X (h_{12}(\xi, \zeta)_X), \tau\rangle + \langle D_\xi^X (h_{12}(\eta, \zeta)_X), \tau\rangle. \end{aligned}$$

Applying Theorem 1.4 concludes the proof.

For the next theorem, we define $\nabla_\xi h_{12}$ following the general principle in defining covariant derivatives of tensors to be derivatives in all variables. So it is defined on Y by the equation

$$\nabla_\xi(h_{12}(\eta, \zeta)) = (\nabla_\xi h_{12})(\eta, \zeta) + h_{12}(D_\xi \eta, \zeta) + h_{12}(\eta, D_\xi \zeta).$$

Theorem 5.2 (Codazzi Equation). *For vector fields ξ, η, ζ on Y,*

$$\mathrm{pr}_{NY}\, R_X(\xi, \eta, \zeta) = (\nabla_\xi h_{12})(\eta, \zeta) - (\nabla_\eta h_{12})(\xi, \zeta).$$

Proof. We start again with

$$D_\eta^X \zeta_X = D_\eta^Y \zeta + h_{12}(\eta, \zeta)$$

so

$$\begin{aligned} D_\xi^X D_\eta^X \zeta_X &= D_\xi^X((D_\eta^Y \zeta)_X + D^X(h_{12}(\eta, \zeta))_X \\ &= D_\xi^Y D_\eta^Y \zeta + h_{12}(\xi, D_\eta^Y \zeta) - {}^tH_\xi(h_{12}(\eta, \zeta)) + \nabla_\xi(h_{12}(\eta, \zeta)). \end{aligned}$$

Since ${}^tH_\xi$ is TY-valued, it is killed by pr_{NY}, and we obtain

$$\mathrm{pr}_{NY}\, D^X_\xi D^X_\eta \zeta_X = h_{12}(\xi, D^Y_\eta \zeta) + \nabla_\xi(h_{12}(\eta, \zeta)).$$

We interchange ξ and η and subtract. We use the definition of R_X to get:

$$\begin{aligned}\mathrm{pr}_{NY}\, R(\xi, \eta)\zeta &= \nabla_\xi(h_{12}(\eta, \zeta)) - \nabla_\eta(h_{12}(\xi, \zeta)) \\ &\quad + h_{12}(\xi, D^Y_\eta \zeta) - h_{12}(\eta, D^Y_\xi \zeta) - \mathrm{pr}_{NY}\, D^X_{[\xi,\eta]}\zeta_X.\end{aligned}$$

But $\mathrm{pr}_{NY}\, D^X_{[\xi,\eta]}\zeta_X = h_{12}([\xi, \eta], \zeta)$. We use the defining equation of $\nabla_\xi h_{12}$ and similarly with ξ, η interchanged, which we subtract. Note that

$$h_{12}(D_\xi \eta, \zeta) - h_{12}(D_\eta \xi, \zeta) = h_{12}([\xi, \eta], \zeta).$$

Then we get cancellations, from which the Codazzi equation follows, thus proving the theorem.

The formalism can go on. We define the **normal Riemann tensor** on Y by

$$R_{NY}(\eta, \zeta) = \nabla_\eta \nabla_\zeta - \nabla_\zeta \nabla_\eta - \nabla_{[\eta,\zeta]},$$

so for vector fields η, ζ on Y

$$R_{NY}(\eta, \zeta)\colon \Gamma NY \to \Gamma NY$$

is an operator on normal fields. As with R_Y we may form the tensors in three and four variables with normal fields μ, ν:

$$\begin{aligned}R_{NY}(\eta, \zeta)\mu &= R_{NY}(\eta, \zeta, \mu), \\ R_{NY}(\eta, \zeta, \mu, \nu) &= \langle R_{NY}(\eta, \zeta)\mu, \nu\rangle.\end{aligned}$$

We recall the operator S_μ for a normal field μ, giving a representation of the second fundamental form in §1, (3), (4), (5). As usual, we may form the bracket

$$[S_\mu, S_\nu] = S_\mu \circ S_\nu - S_\nu \circ S_\mu.$$

Theorem 5.3 (Ricci Equation). *We have*

$$R_X(\xi, \eta, \mu, \nu) = R_{NY}(\xi, \eta, \mu, \nu) - \langle [S_\mu, S_\nu]\xi, \eta\rangle.$$

Proof. More of the same type of computation. We use (6) in §1 twice to get

$$\begin{aligned}R_X(\xi, \eta)\mu &= D^X_\xi D^X_\eta \mu - D^X_\eta D^X_\xi \mu - D^X_{[\xi,\eta]}\mu \\ &= R_{NY}(\xi, \eta) + S_{\nabla_\xi \mu}\eta + D_\eta(S_\mu \xi) + h_{12}(A_\mu \xi, \eta) \\ &\quad - S_{\nabla_\eta \mu}\xi - D_\xi(S_\mu \eta) - h_{12}(\xi, S_\mu \eta) + S_\mu[\xi, \eta].\end{aligned}$$

We take the scalar product with ν, and use formula (3) to find:

$$\begin{aligned}\langle R_X(\xi, \eta)\mu, \nu\rangle &= \langle R_{NY}(\xi, \eta)\mu, \nu\rangle + \langle h_{12}(S_\mu \xi, \eta), \nu\rangle - \langle h_{12}(\xi, S_\mu \eta), \nu\rangle \\ &= \langle R_{NY}(\xi, \eta)\mu, \nu\rangle - \langle (S_\mu S_\nu - S_\nu S_\mu)\xi, \eta\rangle \\ &= \langle R_{NY}(\xi, \eta)\mu, \nu\rangle - \langle [S_\mu, S_\nu]\xi, \eta\rangle,\end{aligned}$$

which concludes the proof.

XIV, §6. THE RIEMANN TENSOR ON A RIEMANNIAN SUBMERSION

We return to a Riemannian submersion

$$\pi\colon X \to Z$$

as in §3 *and* §4, *and use the same notation.* This section is due to O'Neill [O'N 66], some of whose results have been reproduced in various differential geometry texts, e.g. [ChE 75] and [Kl 83/95]. We let R_X and R_Z denote the Riemann tensors on X and Z respectively. If μ is a vector field on Z, we let μ_X (as in §3 and §4) be its horizontal lifting to X.

Theorem 6.1. *Let μ, ν, λ, ζ be vector fields on Z. Then*

$$R_X(\mu_X, \nu_X, \lambda_X, \zeta_X) = R_Z(\mu, \nu, \lambda, \zeta) + V_R(\mu_X, \nu_X, \lambda_X, \zeta_X)$$

where V_R denotes the vertical component,

$$\begin{aligned}V_R(\mu_X, \nu_X, \lambda_X, \zeta_X) &= \tfrac{1}{4}\langle [\mu_X, \lambda_X]^V, [\nu_X, \zeta_X]^V\rangle - \tfrac{1}{4}\langle [\nu_X, \lambda_X]^V, [\mu_X, \zeta_X]^V\rangle \\ &\quad + \tfrac{1}{2}\langle [\lambda_X, \zeta_X]^V, [\mu_X, \nu_X]^V\rangle.\end{aligned}$$

Proof. The Riemann tensor involves second derivatives, but all the formulas needed to perform the iteration easily have been proved in §3.

So we forge ahead. First, by Propositions 3.3, 3.5, and 3.6, we find

$$(1) \quad \langle D_{\mu_X} D_{\nu_X} \lambda_X, \zeta_X \rangle = \mu_X \cdot \langle D_{\nu_X} \lambda_X, \zeta_X \rangle - \langle D_{\nu_X} \lambda_X, D_{\mu_X} \zeta_X \rangle$$
$$= \mu \cdot \langle D_\nu \lambda, \zeta \rangle - \langle D_\nu \lambda, D_\mu \zeta \rangle - \tfrac{1}{4} \langle [\nu_X, \lambda_X], [\mu_X, \zeta_X] \rangle$$
$$= \langle D_\mu D_\nu \lambda, \zeta \rangle - \tfrac{1}{4} \langle [\nu_X, \lambda_X]^V, [\mu_X, \zeta_X]^V \rangle.$$

Decomposing $[\mu_X, \nu_X]$ into horizontal and vertical component and using Proposition 3.6, we get

$$(2) \qquad \langle D_{[\mu_X, \nu_X]} \lambda_X, \zeta_X \rangle = \langle D_{[\mu, \nu]} \lambda, \zeta \rangle - \tfrac{1}{2} \langle [\lambda_X, \zeta_X]^V, [\mu_X, \nu_X]^V \rangle.$$

By (1) and (2), and the definition of the Riemann tensor

$$R(\mu, \nu) = D_\mu D_\nu - D_\nu D_\mu - D_{[\mu, \nu]}$$

and similarly with the subscript X, the formula of Theorem 6.1 falls out, and the proof is concluded.

Corollary 6.2. *For the tensor R_2 such that $R_2(v, w) = R(v, w, v, w)$, we get*

$$R_{2X}(\mu_X, \nu_X) = R_{2Z}(\mu, \nu) + \tfrac{3}{4} \| [\mu_X, \nu_X] \|^2.$$

In particular, the tensor R_2 decreases under submersions.

Proof. This is immediate from the definition and Theorem 6.1.

For the curvature, which is minus R_2, Corollary 6.2 means that curvature increases under submersions.

Remark. In O'Neill [O'N 66], he defines two operators, and formulates his results in terms of these operators. The first result amounts to Theorem 6.1, and is the analogue of the Gauss formula for submersions. The other is the analogue of the Codazzi formula, which I omit. Note that an expression

$$\tfrac{1}{2} [\mu_X, \nu_X]$$

should probably receive a name, as a single item, to make the coefficients 1/2, 1/4, or 3/4 more structural. It remains to be seen what is the best convention to adopt about these expressions.

PART III

Volume Forms and Integration

CHAPTER XV

Volume Forms

For the first time we meet a strictly finite dimensional phenomenon: If X is of finite dimension n, then the n-forms $\mathscr{A}^n(X)$ play a distinguished role whose extension to the infinite dimensional case is not evident. So this chapter is devoted to these forms of maximal degree. In the next chapter, we shall study how to integrate them, so the present chapter also provides a transition from the differential theory to the integration theory.

Although for organization and reference purposes it is convenient to place together here a number of results on volume forms, only the first section giving a basic definition will be used in the next three chapters, so the other sections may be skipped by a reader wanting to get immediately into integration.

XV, §1. VOLUME FORMS AND THE DIVERGENCE

Let V be a finite dimensional vector space over $\mathbf{R}$, of dimension n. We assume given a positive definite symmetric scalar product g, denoted by

$$(v, w) \mapsto \langle v, w\rangle_g = g(v, w) \qquad \text{for } v, w \in V.$$

The space $\bigwedge^n V$ has dimension 1. If $\{e_1, \dots, e_n\}$ and $\{u_1, \dots, u_n\}$ are orthonormal bases of V, then

$$e_1 \wedge \cdots \wedge e_n = \pm u_1 \wedge \cdots \wedge u_n.$$

Two such orthonormal bases are said to have the **same orientation**, or to be **orientation equivalent**, if the plus sign occurs in the above relation. A choice of an equivalence class of orthonormal bases having the same orientation is defined to be an **orientation** of V. Thus an orientation determines a basis for the one-dimensional space $\bigwedge^n V$ over $\mathbf{R}$. Such a basis will be called a **volume**. There exists a unique n-form Ω on V (alternating), also denoted by vol_g, such that for every oriented orthonormal basis $\{e_1, \dots, e_n\}$ we have

$$\Omega(e_1, \dots, e_n) = 1.$$

Conversely, given a non-zero n-form Ω on V, all orthonormal bases $\{e_1, \dots, e_n\}$ such that $\Omega(e_1, \dots, e_n) > 0$ are orientation equivalent, and on such bases Ω has a constant value.

Let (X, g) be a Riemannian manifold. By an **orientation** of (X, g) we mean a choice of a volume form Ω, and an orientation of each tangent space T_xX $(x \in X)$ such that for any oriented orthonormal basis $\{e_1, \dots, e_n\}$ of T_xX we have

$$\Omega_x(e_1, \dots, e_n) = 1.$$

The form gives a coherent way of making the orientations at different points compatible. It is an exercise to show that if (X, g) has such an orientation, and X is connected, then (X, g) has exactly two orientations. In Chapter XVI, we shall give a variation of this definition. By an **oriented chart**, with coordinates $x_1, \dots, x_n$ in $\mathbf{R}^n$, we mean a chart such that with respect to these coordinates, the form has the representation

$$\Omega(x) = \varphi(x)\, dx_1 \wedge \cdots \wedge dx_n$$

with a function φ which is positive at every point of the chart. We call Ω the **Riemannian volume form**, and also denote it by vol_g, so

$$\mathrm{vol}_g(x) = \Omega(x) = \Omega_x.$$

We return to our vector space V, with positive definite metric g, and oriented.

Proposition 1.1. *Let $\Omega = \mathrm{vol}_g$. Then for all n-tuples of vectors $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$ in V, we have*

$$\Omega(v_1, \dots, v_n)\Omega(w_1, \dots, w_n) = \det\langle v_i, w_j\rangle_g.$$

In particular,

$$\Omega(v_1, \dots, v_n)^2 = \det\langle v_i, v_j\rangle_g.$$

Proof. The determinant on the right side of the first formula is multilinear and alternating in each n-tuple $\{v_1, \dots, v_n\}$ and $\{w_1, \dots, w_n\}$.

Hence there exists a number $c \in \mathbf{R}$ such that

$$\det\langle v_i, w_j\rangle_g = c\Omega(v_1, \dots, v_n)\Omega(w_1, \dots, w_n)$$

for all such n-tuples. Evaluating on an oriented orthonormal basis shows that $c = 1$, thus proving the proposition.

Applying Proposition 1.1 to an oriented Riemannian manifold yields:

Proposition 1.2. *Let (X, g) be an oriented Riemannian manifold. Let $\Omega = \mathrm{vol}_g$. For all vector fields $\{\xi_1, \dots, \xi_n\}$ and $\{\eta_1, \dots, \eta_n\}$ on X, we have*

$$\Omega(\xi_1, \dots, \xi_n)\Omega(\eta_1, \dots, \eta_n) = \det\langle \xi_i, \eta_j\rangle_g.$$

In particular,

$$\Omega(\xi_1, \dots, \xi_n)^2 = \det\langle \xi_i, \xi_j\rangle_g.$$

Furthermore, if $\xi^\vee$ denotes the one-form dual to ξ (characterized by $\xi^\vee(\eta) = \langle \xi, \eta\rangle_g$ for all vector fields η), then

$$\Omega(\xi_1, \dots, \xi_n)\Omega = \xi_1^\vee \wedge \cdots \wedge \xi_n^\vee.$$

This last formula is merely an application of the definition of the wedge product of forms, taking into account the preceding formulas concerning the determinant.

At a point, the space of n-forms is 1-dimensional. Hence any n-form on a Riemannian manifold can be written as a product $\varphi\Omega$ where φ is a function and Ω is the Riemannian volume form.

If ξ is a vector field, then $\Omega \circ \xi$ is an $(n-1)$-form, and so there exists a function φ such that

$$d(\Omega \circ \xi) = \varphi\Omega.$$

We call φ the **divergence** of ξ with respect to Ω, or with respect to the Riemannian metric. We denote it by $\mathrm{div}_\Omega\, \xi$ or simply $\mathrm{div}\, \xi$. Thus by definition,

$$d(\Omega \circ \xi) = (\mathrm{div}\, \xi)\Omega.$$

Example. Looking back at Chapter V, §3 we see that if

$$\Omega(x) = dx_1 \wedge \cdots \wedge dx_n$$

is the canonical form on $\mathbf{R}^n$ and ξ is a vector field, $\xi = \sum \varphi_i u_i$ where $\{u_1, \dots, u_n\}$ are the standard unit vectors, and φ_i are the coordinate functions, then

$$\operatorname{div}_\Omega \xi = \sum_{i=1}^{n} \frac{\partial \varphi_i}{\partial x_i}.$$

For the formula with a general metric, see Proposition 1.5.

We shall study the divergence from a differential point of view in the next section, and from the point of view of Stokes' theorem in Chapter XVIII.

On 1-forms, we define the operator

$$d^*: \mathscr{A}^1(X) \to \mathscr{A}^0(X)$$

by duality, that is if $\lambda^\vee$ denotes the vector field corresponding to λ under the Riemannian metric, then we define

$$d^*\lambda = -\operatorname{div} \lambda^\vee.$$

Let us define the **Laplacian** or **Laplace operator** *on functions* by the formula

$$\Delta = d^*d = -\operatorname{div} \circ \operatorname{grad}.$$

In Corollary 2.4, we shall prove the equivalence of this definition with that of Chapter XII, §1.

Proposition 1.3. *For functions φ, ψ we have*

$$\Delta(\varphi\psi) = \varphi\Delta\psi + \psi\Delta\varphi - 2\langle d\varphi, d\psi\rangle_g.$$

Proof. The routine gives:

$$\begin{aligned}\Delta(\varphi\psi) = d^*d(\varphi\psi) &= d^*(\psi\, d\varphi + \varphi\, d\psi)\\ &= -\operatorname{div}(\psi\xi_{d\varphi}) - \operatorname{div}(\varphi\xi_{d\psi})\\ &= -\psi \operatorname{div} \xi_{d\varphi} - (d\psi)\xi_{d\varphi} - \varphi \operatorname{div} \xi_{d\psi} - (d\varphi)\xi_{d\psi}\\ &= \psi\Delta\varphi + \varphi\Delta\psi - 2\langle d\varphi, d\psi\rangle_g\end{aligned}$$

as was to be shown.

Recall that

$$\langle d\varphi, d\psi\rangle_g = \langle \operatorname{grad} \varphi, \operatorname{grad} \psi\rangle_g,$$

so there is an alternative expression for the last term in the formula.

We shall tabulate some formulas concerning the gradient. For simplicity of notation, we shall omit the subscript g in the scalar product, because we now fix g. We shall also write simply gr ψ instead of $\operatorname{grad}_g \varphi$.

gr 1. For functions φ, ψ we have

$$\operatorname{gr}(\varphi\psi) = \varphi \operatorname{gr}(\psi) + \psi \operatorname{gr}(\varphi).$$

gr 2. The map $\varphi \mapsto (\operatorname{gr}(\varphi))/\varphi = \varphi^{-1} \operatorname{gr}(\varphi)$ is a homomorphism, from the multiplicative group of functions never 0, to the additive group of functions. In particular, for a positive function φ,

$$2\varphi^{-1/2} \operatorname{gr}(\varphi^{1/2}) = \varphi^{-1} \operatorname{gr}(\varphi) = \operatorname{gr} \log \varphi$$

because $d \log \varphi = \varphi^{-1}\, d\varphi$.

gr 3. $$\operatorname{gr}(\varphi) \cdot \psi = \operatorname{gr}(\psi) \cdot \varphi = \langle \operatorname{gr}(\varphi), \operatorname{gr}(\psi)\rangle_g.$$

We use these formulas to give two versions of certain operators which arise in practice. For any function φ, we write for the Lie derivative

$$[\operatorname{gr} \varphi] = \mathscr{L}_{\operatorname{gr} \varphi}.$$

Corollary 1.4. *Let δ be a positive function. Then*

$$\Delta - [\operatorname{gr} \log \delta] = \delta^{-1/2}\Delta \circ \delta^{1/2} - \delta^{-1/2}\Delta(\delta^{1/2}).$$

Proof. For a function ψ, by Proposition 1.3,

$$\begin{aligned}(\Delta \circ \delta^{1/2})\psi &= \Delta(\delta^{1/2}\psi)\\ &= \delta^{1/2}\Delta\psi + \psi\Delta(\delta^{1/2}) - 2(\operatorname{gr} \delta^{1/2}) \cdot \psi.\end{aligned}$$

We apply the right side of the equality to be proved to a function ψ. We use the formula just derived, mutliplied by $\delta^{-1/2}$. The term $\delta^{-1/2}\Delta(\delta^{1/2})\psi$ cancels, and we obtain

$$(\text{right side})(\psi) = \Delta\psi - 2\delta^{-1/2}(\operatorname{gr} \delta^{1/2}) \cdot \psi.$$

We use **gr 2** to conclude the proof.

Remark. In Helgason [He 84a], Chapter II, Theorem 3.7, he uses the identity of Corollary 1.4. The difference in sign comes from the fact that we take the Laplacian to be the positive one, and his Laplacian is negative, as an operator.

More formulas concerning the Laplacian will be given in the next section, using the covariant derivative and the variation formula. For applications of such formulas and theory to the heat kernel, cf. [Cha 84], especially Chapters II and III, in addition to [BGM 71].

The remainder of this section will go more systematically into the general theory of the divergence. It will not be used in the sequel except in §6 *and* §8. *Hence readers may proceed immediately to* §2.

General definition of the divergence

Although the most important case of the divergence is on a Riemannian manifold, some properties are most clearly expressed in a more general case which we now describe. Let T be a vector space of finite dimension n over $\mathbf{R}$. Then $\bigwedge^n T$ is of dimension 1, and will be called the **determinant** of T, so by definition,

$$\det T = \bigwedge^{\max} T = \bigwedge^n T.$$

Observe that we also have $\det T^\vee$. A non-zero element of $\det T^\vee$ will be called a **volume form** on T.

The vector space of sections of $\bigwedge^n T^\vee X$ on a manifold X of dimension n is also a module over the ring of functions. By a **volume form on** X we mean section which is nowhere 0, so a volume form is a basis for this space over the ring of functions. Instead of saying that Ω is a volume form, one may also say that Ω is **non-singular**. If Ψ is any n-form on X, then there exists a function f such that $\Psi = f\Omega$. So let Ω be a volume form. Let ξ be a vector field on X. We define the **divergence of** ξ **with respect to** Ω just as we did for the Riemannian volume form, namely $\operatorname{div}_\Omega(\xi)$ is defined by the property

DIV 1. $$d(\Omega \circ \xi) = (\operatorname{div}_\Omega(\xi))\Omega.$$

From Chapter V, Proposition 5.3, **LIE 1**, we also have the equivalent defining property

DIV 2. $$\mathscr{L}_\xi \Omega = (\operatorname{div}_\Omega(\xi))\Omega.$$

Directly from **DIV 2** and **LIE 2**, we get for any functions φ, f:

DIV 3. $$\operatorname{div}_\Omega(\varphi\xi) = \varphi \operatorname{div}_\Omega(\xi) + \xi \cdot \varphi.$$

DIV 4. $$df \wedge (\Omega \circ \xi) = (\xi \cdot f)\Omega.$$

Proof. First we have $\mathscr{L}_\xi(f\Omega) = (\xi \cdot f)\Omega + f \operatorname{div}_\Omega(\xi)\Omega$, and second,

$$\begin{aligned}\mathscr{L}_\xi(f\Omega) = d(f\Omega \circ \xi) &= df \wedge (\Omega \circ \xi) + fd(\Omega \circ \xi)\\ &= df \wedge (\Omega \circ \xi) + f \operatorname{div}_\Omega(\xi)\Omega.\end{aligned}$$

Then **DIV 4** follows from these two expressions.

One can define an orientation on the general vector space T depending on the non-singular form Ω. Of course in general, we don't have the notion of orthogonality. But we say that a basis $\{v_1, \dots, v_n\}$ of T is **positively oriented**, or simply **oriented**, with respect to Ω if $\Omega(v_1, \dots, v_n) > 0$. Let Ω, Ψ be volume forms. We say that they have the **same orientation**, or that they are **positive with respect to each other**, if there exists a positive function h such that $\Omega = h\Psi$. Forms with the same orientation define the same orientation on bases. A manifold which admits a volume form is said to be **orientable**, and the class of volume forms having the same orientation is said to define the **orientation.**

Let δ be a positive function on X, and let Ψ be a volume form. Then:

DIV 5. $$\operatorname{div}_{\delta\Psi}(\xi) = (\xi \cdot \log \delta) + \operatorname{div}_\Psi(\xi).$$

Proof. By Proposition 5.3 of Chapter V, **LIE 1**, we have

$$\begin{aligned}d(\delta\Psi \circ \xi) = \mathscr{L}_\xi(\delta\Psi) &= (\xi \cdot \delta)\delta^{-1}\delta\Psi + \delta\mathscr{L}_\xi(\Psi)\\ &= (\xi \cdot \log \delta)(\delta\Psi) + \delta \operatorname{div}_\Psi(\xi)\Psi,\end{aligned}$$

which proves the formula.

The divergence in a chart

Next we obtain an expression for the divergence in a chart. Let U be an open set of a chart for X in $\mathbf{R}^n$ with its standard unit vectors $u_1, \dots, u_n$. There exists a function δ never 0 on U such that in this chart,

$$\Omega = \delta\, dx_1 \wedge \cdots \wedge dx_n.$$

Suppose U is connected. Then we have $\delta > 0$ on U or $\delta < 0$ on U since Ω is assumed non-singular. For simplicity, assume $\delta > 0$.

Example. If $\Omega = \Omega_g$ is the Riemannian volume form, then

$$\delta = (\det g)^{1/2}.$$

In other words,

$$\Omega_g(x) = (\det g(x))^{1/2}\, dx_1 \wedge \cdots \wedge dx_n.$$

Here $g(x)$ denotes the matrix representing g with respect to the standard basis of $\mathbf{R}^n$.

We write ξ in the chart U as a linear combination

$$\xi = \sum \varphi_i u_i$$

with coordinate functions $\varphi_1, \ldots, \varphi_n$. We let ∂_i be the i-th partial derivative. We write the coordinate vector of ξ vertically, that is

$$\Phi = \Phi_\xi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix}.$$

We let ${}^t\mathbf{D}_\Omega$ be the *row* vector of operators

$${}^t\mathbf{D}_\Omega = (\partial_1 + \partial_1 \log \delta, \ldots, \partial_n + \partial_n \log \delta).$$

Proposition 1.5.

$$\begin{aligned} \operatorname{div}_\Omega \xi &= \delta^{-1} \sum \partial_i(\delta\varphi_i) \\ &= \sum \partial_i\varphi_i + \sum (\partial_i \log \delta)\varphi_i. \end{aligned}$$

In matrix form,

$$\operatorname{div}_\Omega \xi = {}^t\mathbf{D}_\Omega \Phi_\xi \qquad \textit{or also} \qquad \operatorname{div}_\Omega = \delta^{-1}\, {}^tD \circ \delta.$$

Proof. We have

$$\begin{aligned} (\Omega \circ \xi)(u_1, \ldots, \hat{u}_i, \ldots, u_n) &= \Omega(\xi, u_1, \ldots, \hat{u}_i, \ldots, u_n) \\ &= (-1)^{i-1}\Omega(u_1, \ldots, \xi, \ldots, u_n) \\ &= (-1)^{i-1}\delta\varphi_i. \end{aligned}$$

Hence

$$(\Omega \circ \xi) = \sum (-1)^{i-1}\delta_{\varphi_i}\, dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n,$$

and since $ddx_j = 0$ for all j, we obtain

$$\begin{aligned} d(\Omega \circ \xi) &= \sum (-1)^{i-1}\partial_i(\delta\varphi_i)\, dx_i \wedge dx_1 \wedge \cdots \wedge \widehat{dx_i} \wedge \cdots \wedge dx_n \\ &= \sum \partial_i(\delta\varphi_i)\, dx_1 \wedge \cdots \wedge dx_n \\ &= \delta^{-1} \sum \partial_i(\delta\varphi_i)\Omega. \end{aligned}$$

This proves the proposition.

We return to the gradient, for which we give an expression in local coordinates, with an application to the Laplacian.

Proposition 1.6. *Let* $\operatorname{gr}(\psi) = \sum \varphi_i u_i$. *Let* $g(x)$ *be the* $n \times n$ *matrix representing the metric at a point* x. *Then the coordinate vector of* $\operatorname{gr}(\psi)$ *is*

$$\Phi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} = g(x)^{-1} \begin{pmatrix} \partial_1\psi \\ \vdots \\ \partial_n\psi \end{pmatrix}.$$

In other words,

$$\Phi = g^{-1}\partial\psi,$$

where ∂ *is the vector differential operator such that* ${}^t\partial = (\partial_1, \ldots, \partial_n)$.

Proof. By definition,

$$\langle \operatorname{gr}(\psi), u_j \rangle_g = (d\psi)(u_j) = \partial_j\psi.$$

The left side is equal to $\langle \operatorname{gr}(\psi), g(x)u_j \rangle$ at a point x. Note that here the scalar product is the usual dot product on $\mathbf{R}^n$, without the subscript g. The formula of the proposition then follows at once.

Proposition 1.7. *Let* f *and* ψ *be function, and let* $\operatorname{gr}(\psi) = \sum \varphi_j u_j$ *as in Proposition* 1.6. *Then*

$$\operatorname{gr}(\psi) \cdot f = \sum_{j=1}^{n} (\partial_j f)\varphi_j.$$

Proof. Since $u_j \cdot f = \partial_j f$, the formula is clear.

From Propositions 1.5 and 1.6, we obtain the coordinate representation of the Laplacian via a matrix:

Proposition 1.8. *On an open set of* $\mathbf{R}^n$, *with metric matrix* g, $\delta = (\det g)^{1/2}$, *and Laplacian* Δ_g, *we have*

$$\begin{aligned} -\Delta_g &= \operatorname{div}_g \operatorname{gr}_g = {}^t\mathbf{D}_g g^{-1}\partial \\ &= \delta^{-1}\, {}^t\partial\, \delta_g^{-1}\partial. \end{aligned}$$

Here, $\mathbf{D}_g$ *abbreviates* $\mathbf{D}_{\Omega_g}$, *and* div_g *abbreviates* $\operatorname{div}_{\Omega_g}$.

Putting all the indices in, we get

(1) $$-\Delta_g f = \delta^{-1} \sum_i \partial_i \left(\delta \sum_j g^{ij}\partial_j f \right)$$

where in classical notation, $g^{-1}(x)$ is the matrix $(g^{ij}(x))$ for $x \in \mathbf{R}^n$. Using the rule for the derivative of a product, we write (1) in the form

$$(2) \qquad -\Delta_g f = \sum_{i,j=1}^{n} g^{ij}\,\partial_i \partial_j f + L_g f,$$

where L_g is a first-order differential operator, that is a linear combination of the partials $\partial_1, \ldots, \partial_n$ with coefficients which are functions, depending on g. From this expression, we see that the matrix $g^{-1} = (g^{ij})$ is the matrix of the second-order term, quadratic in the partials ∂_i, ∂_j. Hence we obtain:

Theorem 1.9. *Let X be a Riemannian manifold. Then the Laplacian determines the metric,* i.e. *if two Riemannian metrics have the same Laplacian, they are equal. If $F\colon X \to Y$ is a differential isomorphism of Riemannian manifolds, and F maps Δ_X on Δ_Y, that is F commutes with the Laplacians, then F is an isometry.*

Note that the second statement about the differential isomorphism is just a piece of functorial abstract nonsense, in light of the first statement. Indeed, F maps the metric g_X to a metric $F_* g_X$ on Y, and similarly for the Laplacian. By assumption, $F_*\Delta_X = \Delta_Y$. Hence Δ_Y is the Laplacian of g_Y and of $F_* g_X$, so $g_Y = F_* g_X$ by the first statement in the theorem.

Example. Let $A = \mathbf{R}^+ \times \cdots \times \mathbf{R}^+$ be the product of positive multiplicative groups, taken n times, so we view A as an open subset of $\mathbf{R}^n$. We let a denote the variable in A, so $a = {}^t(a_1, \ldots, a_n)$ with $a_i > 0$. We identify the tangent space $T_a A = T_a$ with $\mathbf{R}^n$, so a vector $v \in T_a$ is an ordinary n-tuple,

$$v = {}^t(c_1, \ldots, c_n) \qquad \text{with} \qquad c_i \in \mathbf{R}.$$

Let g be the metric on A defined by the formula

$$\langle v, v\rangle_a = \sum_{i=1}^{n} c_i^2/a_i^2.$$

Then g is represented by the diagonal matrix $g(a) = \operatorname{diag}(a_1^{-2}, \ldots, a_n^{-2})$, that is

$$\langle v, v\rangle_a = \langle v, g(a)v\rangle,$$

where the scalar product without indices denotes the standard scalar product on $\mathbf{R}^n$. Then

$$\delta(a) = \det g(a)^{1/2} = \prod_{i=1}^{n} a_i^{-1} = \mathbf{d}(a)^{-1}$$

where $\mathbf{d}(a) = a_1 \cdots a_n$ is the product of the coordinates. Thus for a function ψ on A, we have the explicit determination for the gradient,

$$(1) \qquad (\operatorname{gr}_A \psi)(a) = g(a)^{-1}\partial\psi = {}^t(a_1^2\partial_1\psi, \ldots, a_n^2\partial_n\psi)(a).$$

The Laplacian Δ_A from Proposition 1.8 is seen to be

$$(2) \qquad -\Delta_A = \sum_{i=1}^{n} a_i\partial_i + \sum_{i=1}^{n} a_i^2\partial_i^2.$$

This comes from matrix multiplication,

$$\mathbf{d}(a)(\partial_1, \ldots, \partial_n)\begin{pmatrix} \mathbf{d}(a)^{-1}a_1^2\partial_1 \\ \vdots \\ \mathbf{d}(a)^{-1}a_n^2\partial_n \end{pmatrix}.$$

XV, §2. COVARIANT DERIVATIVES

In this section, we gather together a number of results on the covariant derivative in connection with volume forms on the oriented Riemannian manifold (X, g) of dimension n.

We begin by some remarks extending the formalism of the covariant derivative to volume forms. First, we recall from multilinear algebra that the metric g induces a natural metric on the dual space, i.e. the cotangent space, identified with the tangent space via g. In other words, for two vector fields ξ, η we have

$$\langle \xi^\vee, \eta^\vee\rangle_g = \langle \xi, \eta\rangle_g.$$

Then we get a scalar product on differential forms of all degree. This is just a matter of punctual multilinear algebra. On p-forms which are decomposable, the scalar product is defined by the determinant,

$$\langle \xi_1^\vee \wedge \cdots \wedge \xi_p^\vee, \eta_1^\vee \wedge \cdots \wedge \eta_p^\vee\rangle_g = \det\langle \xi_i^\vee, \eta_j^\vee\rangle_g = \det\langle \xi_i, \eta_j\rangle_g.$$

Let D be the metric covariant derivative. Its characterizing property for the scalar product of two vector fields extends at once to forms, and specifically to 1-forms, and then for p-forms, ω, ψ of any degree, and any vector field ξ, we have

$$(1) \qquad \xi \cdot \langle \omega, \psi\rangle_g = \langle D_\xi\omega, \psi\rangle_g + \langle \omega, D_\xi\psi\rangle_g.$$

This applies in particular to volume forms Ω and Ψ.

The proof for 1-forms comes directly from the metric property of D. For forms of higher degree, it comes also at once from the multilinearity of the determinant as a function of columns and rows, and from the fact that the derivative of a product satisfies the standard rule. One applies this rule both to the determinant viewed as a product of p column variables, and the scalar products $\langle \xi_i, \eta_j \rangle_g$. The reader can write all this down faster than I could.

We recall that a sequence $\{\xi_1, \ldots, \xi_n\}$ of vector fields is called an **orthonormal frame** (on some open subset of X) if they are orthonormal for the metric g, that is

$$\langle \xi_i, \xi_j \rangle_g = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

Given a point $x \in X$, such an orthonormal frame always exists in a neighborhood of x.

Theorem 2.1. *Let D be the metric covariant derivative. Then*

$$D_\xi \operatorname{vol}_g = 0$$

for all vector fields ξ.

Proof. Let $\Omega = \operatorname{vol}_g$ be the Riemannian volume form. If $\{\xi_1, \ldots, \xi_n\}$ is an orthonormal frame, then $\Omega = \pm \xi_1^\vee \wedge \cdots \wedge \xi_n^\vee$ and $\langle \Omega, \Omega \rangle_g = 1$. Taking the Lie derivative with ξ yields 0, and also yields

$$0 = 2\langle D_\xi \Omega, \Omega \rangle_g.$$

But $D_\xi \Omega = \varphi \Omega$ for some function φ, so $0 = 2\varphi \langle \Omega, \Omega \rangle_g$, whence $\varphi = 0$, which proves the proposition.

Remark. The above result remains true suitably formulated in the non-oriented case, because the theorem is local, and locally, the absolute value of the form differs by ± 1 from the itself.

The next theorem will give an application of Theorem 2.1.

The metric derivative D operates on vector fields and also on r-forms for all r, especially $r = 1$ and $r = n$. For any vector field ξ we let $D\xi$ be the endomorphism of ΓTX such that

$$(D\xi)\eta = D_\eta \xi.$$

At each point $x \in X$ we have the operator

$$(D\xi)_x\colon T_x X \to T_x X \qquad \text{such that} \qquad (D\xi)_x(v) = (D_v \xi)(x),$$

on the finite dimensional vector space $T_x X$. This allows us to take the trace $\operatorname{tr}(D\xi)$ of this operator at each point, so to take $\operatorname{tr}(D\xi)_x$. The trace can be computed as usual by using an orthonormal basis.

Similarly, we can define $D\lambda$ for a 1-form $\lambda \in \mathscr{A}^1(X)$, whereby

$$D\lambda\colon \Gamma TX \to \Gamma T^\vee X \qquad \text{is such that} \qquad (D\lambda)(\xi) = D_\xi \lambda.$$

Thus for each $x \in X$, $(D\lambda)_x$ may be viewed as a linear map

$$(D\lambda)_x\colon T_x X \to T_x^\vee X,$$

whose trace can be computed by using duality, namely

$$\operatorname{tr}(D\lambda) = \sum_i \langle D_{\xi_i} \lambda, \xi_i \rangle.$$

On the right side, we use the convenient notation $\langle \lambda, \xi \rangle = \lambda(\xi)$ for a 1-form λ and a vector field ξ. In such a case, there is no subscript g on the scalar bilinear pairing between functionals and vectors.

Theorem 2.2. *Let $\xi_1, \ldots, \xi_n$ be an orthonormal frame of vector fields, and let ξ be a vector field. Then*

$$\operatorname{div} \xi = \sum_{i=1}^n \langle D_{\xi_i} \xi, \xi_i \rangle_g = \operatorname{tr}(D\xi).$$

In particular, for $\lambda \in \mathscr{A}^1(X)$ we have

$$\operatorname{div} \lambda^\vee = \operatorname{tr}(D\lambda).$$

Proof. Let $\Omega = \operatorname{vol}_g$ be the volume form. By **COVD 6** of Chapter VIII, §1, and Proposition 2.1, we get

$$\begin{aligned} d(\Omega \circ \xi)(\xi_1, \ldots, \xi_n) &= \sum_{i=1}^n (-1)^{i-1} D_{\xi_i}(\Omega \circ \xi)(\xi_1, \ldots, \hat{\xi}_i, \ldots, \xi_n) \\ &= \sum_{i=1}^n (-1)^{i-1} (\Omega \circ D_{\xi_i}\xi)(\xi_1, \ldots, \hat{\xi}_i, \ldots, \xi_n) \\ &= \sum_{i=1}^n \Omega(\xi_1, \ldots, D_{\xi_i}\xi, \ldots, \xi_n) \end{aligned}$$

and since $D_{\xi_i}\xi$ has the Fourier expression $D_{\xi_i}\xi = \sum_j \langle D_{\xi_i}\xi, \xi_j \rangle_g \xi_j$,

$$= \sum_{i=1}^n \langle D_{\xi_i}\xi, \xi_i \rangle_g \Omega(\xi_1, \ldots, \xi_n).$$

But also $d(\Omega \circ \xi)(\xi_1, \ldots, \xi_n) = (\mathrm{div}\, \xi)\mathrm{vol}_g(\xi_1, \ldots, \xi_n)$. Hence

$$\mathrm{div}\, \xi = \sum_{i=1}^{n} \langle D_{\xi_i}\xi, \xi_i\rangle_g,$$

which proves the first formula. The second is a mere rephrasing, applied to the vector field $\lambda^\vee$.

Directly from the definition of the operator d^* in the preceding section, we now obtain:

Corollary 2.3. *On a 1-form λ, we have $d^*\lambda = -\mathrm{tr}(D\lambda)$.*

We can then apply this to the Laplacian, to get:

Corollary 2.4. *Let $\xi_1, \ldots, \xi_n$ be an orthonormal frame as in Theorem 2.2. Let φ be a function. Then*

$$\Delta\varphi = -\mathrm{tr}(D\, d\varphi) = -\sum_{i=1}^{n} \langle D_{\xi_i} d\varphi, \xi_i\rangle = -\sum_{i=1}^{n} \langle D_{\xi_i}(\mathrm{grad}\, \varphi), \xi_i\rangle_g.$$

If $\{u_1, \ldots, u_n\}$ is an orthonormal basis of the tangent space T_xX at some point $x \in X$, and α_i is the geodesic with $\alpha_i(0) = x$ and $\alpha_i'(0) = u_i$, then

$$\Delta\varphi(x) = -\sum_{i=1}^{n} (\varphi \circ \alpha_i)''(0).$$

Proof. The first assertion comes from applying Theorem 2.2 to $\lambda = d\varphi$. The second assertion then follows by using Corollary 4.4 of Chapter VIII.

From the preceding corollary, we can obtain an expression for the Laplacian in polar coordinates. I follow [BGM 71]. *We pick a point $x \in X$ as an origin, with its tangent space T_xX. We let U_x be an open ball centered at 0_x on which $\exp_x$ induces an isomorphism to its image, and we let $y \in U_x$. We want to determine $\Delta\varphi(y)$ for a function φ which depends only on the Riemannian distance from x, say*

$$\varphi(y) = f(r(y)) \quad \text{where} \quad r(y) = \mathrm{dist}_g(x, y),$$

and f is a C^2 function of a real variable.

Proposition 2.5. *Let $\alpha = \alpha_1$ be the unique geodesic from x to $y \neq x$, parametrized by arc length, and let $e_1 = \alpha'(r) \in T_yX$. Let $e_2, \ldots, e_n$ be*

unit vectors in T_yX such that $\{e_1, \ldots, e_n\}$ is an orthonormal basis of T_yX. Let η_i $(i = 2, \ldots, n)$ be the Jacobi lift of α such that

$$\eta_i(0) = 0 \quad \text{and} \quad \eta_i(r) = e_i.$$

Then

$$\Delta\varphi(y) = -f''(r) - f'(r)\sum_{i=2}^{r} \langle D_{\alpha'}\eta_i(r), \eta_i(r)\rangle_g.$$

Proof. Let β_i $(i = 1, \ldots, n)$ be the geodesic from y such that

$$\beta_i(0) = y \quad \text{and} \quad \beta_i'(0) = e_i.$$

Observe that $\beta_1(t) = \alpha_1(r + t)$ for small t, by the uniqueness of the integral curve of the corresponding differential equation. We apply Corollary 2.4 to the Laplacian at y, and the geodesics β_i $(i = 1, \ldots, n)$ to get

$$\Delta\varphi(y) = -\sum_{i=1}^{n} (\varphi \circ \beta_i)''(0).$$

Since $\beta_1(t) = \alpha_1(r + t)$, we can split off the first term, to obtain

$$\Delta\varphi(y) = -f''(r) - \sum_{i=2}^{n} (\varphi \circ \beta_i)''(0).$$

Let $\alpha_{i,t}$ be the unique geodesic from x to $\beta_i(t)$ (for small t), parametrized by arc length. Thus $\alpha_{i,t}$ is what we called the variation of α at its end point, in the direction of e_i, for $i = 2, \ldots, n$. Then by Propositions 3.3 of Chapter IX, Proposition 1.9 of Chapter XI, and the fact that

$$(\varphi \circ \beta_i)(t) = f(L(\alpha_{i,t})),$$

we obtain

$$(\varphi \circ \beta_i)''(0) = f'(r)\langle D_{\alpha'}\eta_i(r), \eta_i(r)\rangle_g,$$

which proves our proposition.

The trace $\mathrm{tr}(D\xi)$ in a chart

Just as we ended the last section with formulas in a chart U, we end the present section with the corresponding formula. Again we let Φ be the coordinate vector of ξ, so ${}^t\Phi = (\varphi_1, \ldots, \varphi_n)$ is the coordinate vector of ξ with respect to the basis $\{u_1, \ldots, u_n\}$. We let B_U be the bilinear map occurring in the definition of the covariant derivative, so the chart representation of $D_\eta \xi$ is

$$D_\eta \xi = \xi'_U \eta - B_U(\eta, \xi).$$

Then:

Proposition 2.6. *Let $\delta = (\det g)^{1/2}$. For each j we have*

$$\partial_j \log \delta = -\sum_k B_U(u_j, u_k),$$

and

$$\begin{aligned} \operatorname{div} \xi = \operatorname{tr}(D\xi) &= \sum_i \partial_i \varphi_i - \sum_{i,k} \varphi_i \langle B_U(u_i, u_k), u_k \rangle \\ &= \sum_i \partial_i \varphi_i - \sum_k B_U(\xi, u_k). \end{aligned}$$

Proof. The second formula for the trace comes from the definition of the trace and the definition of $D\xi$. The first formula then follows componentwise from Proposition 1.4. This concludes the proof.

XV, §3. THE JACOBIAN DETERMINANT OF THE EXPONENTIAL MAP

We continue to consider a Riemannian manifold (X, g). We let $x \in X$, and we let $\mathbf{B}_x$ be an open ball in T_xX centered at 0_x, such that $\exp_x$ gives an isomorphism of $\mathbf{B}_x$ with its image in X. Thus without loss of generality, we may assume X oriented, and we let vol_g be the volume form on X. We call $\mathbf{B}_x$ a normal chart at x. For $y \in \exp_x(\mathbf{B}_x)$. We write $y = \exp_x(v_y)$, so $v_y = \log_x(y)$, as it were.

We note that the differential

$$T\exp_x(v_y)\colon T_x \to T_y$$

is a linear isomorphism, and both T_x and T_y have the positive definite scalar products of the Riemannian metric, so we may define the absolute value of the determinant of $(T\exp_x)(v_y)$. One simply picks orthonormal bases in each one of these vector spaces, and the determinant of the matrix representing $(T\exp_x)(v_y)$ with respect to these bases. Picking oriented bases actually makes the determinant positive, so we don't need to take an absolute value. We let J denote the Jacobian determinant, so

$$\exp_x^* \mathrm{vol}_g = J\, \mathrm{vol}_{\mathrm{euc}} \quad \text{or also} \quad \exp_x^* \mathrm{vol}_g(v) = J(v)\, \mathrm{vol}_{\mathrm{euc}}(v),$$

where $\mathrm{vol}_{\mathrm{euc}}$ is the euclidean volume on T_xX, determined by the positive

definite metric $g(x)$, and v is the vector variable in T_xX. We shall express J in polar coordinates.

Let $\mathbf{S}(1)$ be the unit sphere in T_xX. Any vector $v \in T_xX$, $v \neq 0$, can be written uniquely in the form

$$v = ru,$$

where u is a unit vector in the direction of v, and $r > 0$. We call (r, u) the **polar coordinates** of v. Then the euclidean volume has the usual decomposition

$$\mathrm{vol}_{\mathrm{euc}}(r, u) = r^{n-1}\, dr\, d\mu(u),$$

where $d\mu(u)$ is the usual spherical measure ($d\theta$ in dimension 2). Then

$$(\exp_x^* \mathrm{vol}_g)(ru) = J(r, u) r^{n-1}\, dr\, d\mu(u).$$

Proposition 3.1. *Let u be a unit vector in T_xX and let α be the geodesic parametrized by arc length such that $\alpha(0) = x$ and $\alpha'(0) = u$. Put $u = w_1$ and let $\{u, w_2, \ldots, w_n\}$ be a basis of T_xX such that $w_i \perp u$ for $i = 2, \ldots, n$. Let η_i $(i = 2, \ldots, n)$ be the Jacobi lift of α such that*

$$\eta_i(0) = 0 \quad \text{and} \quad D_{\alpha'}\eta_i(0) = w_i.$$

Then

$$r^{n-1} J(r, u) = \frac{\det(\eta_2(r), \ldots, \eta(r))}{\det(w_2, \ldots, w_n)} = \frac{\det^{1/2}\langle \eta_i(r), \eta_j(r) \rangle_g}{\det(w_2, \ldots, w_n)}.$$

The determinant on the right is taken for $i, j = 2, \ldots, n$.

Proof. Observe that we may also use η_1, which is such that $\eta_1(t) = t\alpha'(t)$. The equality between the two expressions on the right of the equality sign follows from Proposition 1.1. Let $f = \exp_x$. Then for any vectors $w_1, \ldots, w_n \in T_xX$ we have

$$\begin{aligned} (\exp_x^* \mathrm{vol}_g)(v)(w_1, \ldots, w_n) &= \mathrm{vol}_g(Tf(v)w_1, \ldots, Tf(v)w_n) \\ &= \det(Tf(v)w_1, \ldots, Tf(v)w_n) \\ &= J(v) \det(w_1, \ldots, w_n). \end{aligned}$$

We put $v = rw_1 = ru$. By Theorem 3.1 of Chapter IX we know that

$$T\exp_x(ru)w_i = \frac{1}{r}\eta_i(r).$$

Then for $i = 1$, $\eta_1(r)/r = \alpha'(r)$, which is a unit vector perpendicular to the others. Thus to compute the volume of the parallelotope in euclidean

n-space, we may disregard this vector, and simply compute the volume of the projection on $(n-1)$-space, and thus we may compute only the $(n-1)\times(n-1)$ determinant of the vectors

$$\det(\eta_2(r)/r, \dots, \eta_n(r/r) = \frac{1}{r^{n-1}} \det(\eta_2(r), \dots, \eta_n(r)),$$

from which the proposition falls out.

Proposition 3.1 is applied in several cases.

Corollary 3.2. *If in Proposition* 3.1 *all the vectors w_i are unit vectors u_i such that $\{u_1, \dots, u_n\}$ is an orthonormal basis of T_xX, and $u = u_1$, then we have simply*

$$r^{n-1}J(r, u) = \det{}^{1/2}\langle \eta_i(r), \eta_j(r)\rangle_g.$$

From this case and the asymptotic expansion for the Jacobi lifts, we obtain:

Corollary 3.3. *Again with an orthonormal basis $\{u_1, \dots, u_n\}$ of T_xX, let $u = u_1$ and*

$$\operatorname{Ric}(u, u) = \sum_{i=2}^{n} R_2(u, u_i).$$

Then

$$\exp_x^* \operatorname{vol}_g(ru) = \left[1 + \operatorname{Ric}(u, u)\frac{r^2}{3!} + O(r^3)\right] \operatorname{vol}_{\mathrm{euc}}(ru) \qquad \text{for } r \to 0.$$

Proof. By Corollary 3.2, $J(r, u)$ is $\det^{1/2}\langle \eta_i(r)/r, \eta_j(r)/r\rangle_g$ with the determinant taken for $i, j = 1, \dots, n$ or $i, j = 2, \dots, n$. Using the asymptotic expansion of Chapter IX, Proposition 5.4 and the orthonormality, one gets that

$$J(r, u) = \prod_{i=2}^{n}\left(1 + 2R_2(u, u_i)\frac{r^2}{3!}\right)^{1/2} + O(r^3) \qquad \text{for } r \to 0,$$

which is immediately expanded to yield the corollary.

Example. Suppose $\dim X = 2$. Then $\operatorname{Ric}(u, u) = R_2(u, u_2) = R_2(u_1, u_2)$. Putting $u_2 = u'$, we get

$$J(r, u) = 1 + R_2(u, u')\frac{r^2}{3!} + O(r^3) \qquad \text{for } r \to 0.$$

If we keep u fixed, and use Δ in polar coordinates, $\Delta = -\partial_r^2 - r^{-1}\partial_r$, then we see that

$$R_2(u, u') = -\tfrac{3}{2}\Delta J(0).$$

Compare with [He 78], Chapter I, Lemma 12.1 and Theorem 12.2.

For the further asymptotic expansion of the volume, see [Gray 73], as well as applications referred to in the bibliography of this paper.

On the other hand, we shall also meet a situation where $\{w_1, \dots, w_n\}$ is not an orthonormal basis as in the next corollary. Cf. Chapter IX, Proposition 3.3.

Corollary 3.4. *Let $\exp_x\colon \mathbf{B}_x \to X$ be the normal chart in X as at the beginning of the section, and $y = \exp_x(ru)$ with $ru \in \mathbf{B}_x$, and some unit vector u. Let $\alpha(s) = \exp_x(su)$ and let $e_1 = \alpha'(r)$. Complete e_1 to an orthonormal basis $\{e_1, \dots, e_n\}$ of T_yX, and let η_i be the Jacobi lift of α (depending on y, or r if u_1 is viewed as fixed), such that*

$$\eta_i(0) = 0 \quad \text{and} \quad \eta_i(r) = e_i \quad \text{for } i = 2, \dots, n.$$

Let $J'(s, u) = \partial_1 J(s, u)$. Then

$$J'/J(r, u) + \frac{n-1}{r} = \sum_{i=2}^{n} \langle D_{\alpha'}\eta_i(r), \eta_i(r)\rangle_g.$$

Proof. In the present case, $D_{\alpha'}\eta_i(0) = w_i$ is whatever it is, but we observe that the determinant $\det(w_2, \dots, w_n)$ is constant, so disappears in taking the logarithmic derivative of the expression in Proposition 3.1. We also observe that in the present case,

$$\langle \eta_i(r), \eta_j(r)\rangle_g = \delta_{ij},$$

so the matrix formed with these scalar products is the unit matrix. Taking the logarithmic derivative of one side, we obtain

$$J'/J(r, u) + (n-1)/r.$$

Let $h_{ij} = \langle \eta_i, \eta_j\rangle_g$, and let $H = (h_{ij})$. On the other side, we obtain the logarithmic derivative

$$\frac{1}{2}\frac{(\det H)'}{\det H}.$$

Let $H_2, \dots, H_n$ be the columns of H. By Leibniz's rule, we know that

$$(\det H)' = \sum_{i=2}^{n} \det(H_2, \dots, H_i', \dots, H_n).$$

Observe that

$$\langle \eta_i, \eta_j \rangle_g' = \langle D_{\alpha'}\eta_i, \eta_j \rangle_g + \langle \eta_i, D_{\alpha'}\eta_j \rangle_g.$$

and in particular,

$$\langle \eta_i, \eta_i \rangle_g' = 2\langle D_{\alpha'}\eta_i, \eta_i \rangle_g.$$

What we want follows from a purely algebraic property of determinants, namely:

Lemma 3.5. *Let $A = (A^1, \dots, A^m)$ be a non-singular $m \times m$ matrix over a field, where $A^1, \dots, A^m$ are the columns of A. Let $B = (B^1, \dots, B^m)$ be any $m \times m$ matrix over the field. Then*

$$\sum_i \det(A^1, \dots, B^i, \dots, A^m) = (\det A) \operatorname{tr}(A^{-1}B).$$

Proof. Let $X = (x_{ij})$ be the matrix such that

$$x_{1i}A^1 + \cdots + x_{mi}A^m = B^i \qquad \text{for} \quad i = 1, \dots, m.$$

By Cramer's rule,

$$x_{ii} \det(A) = \det(A^1, \dots, B^i, \dots, A^m).$$

But $AX = B$ so $X = A^{-1}B$, and the lemma follows.

We apply the lemma to the case when $A = H(r)$ is the unit matrix and $B^j = H_j'(r)$ to conclude the proof.

Corollary 3.6. *Let φ be a C^2 function on a normal ball centered at the point $x \in X$. Suppose that φ depends only on the g-distance r from x, say $\varphi(y) = f(r(y))$. Let $y = \exp(ru)$, with a unit vector u. Then*

$$\Delta\varphi(y) = -f''(r) - f'(r)\left(J'/J(r, u) + \frac{n-1}{r}\right).$$

Proof. Combine Corollary 3.4 with Proposition 2.5.

The formulas in Proposition 2.5 and Corollary 3.6 apply to a function which is constant on the spheres centered at the point x. However, it is only a formal matter to obtain the more general formula for any function. We rely on a general lemma about the exponential. Consider a normal ball B centered at a point $x \in X$ as in Chapter VIII, §5. Thus the exponential

$$\exp_x \colon \mathbf{B} \to B$$

gives an isomorphism of a ball $\mathbf{B}$ in T_xX centered at 0_x, with the Riemannian ball B in X, centered at x. For $y \neq x$ in the ball B, let $\mathbf{n}(y)$ be the unit normal vector to the sphere $S_r(x)$ with $r = \operatorname{dist}_g(x, y)$. Then $y \mapsto \mathbf{n}(y)$ is a vector field on the punctured ball, normal to each sphere, and called the **unit radial field** from x.

Lemma 3.7. *Let u be a unit vector in T_xX. Let φ be a C^2 function on a normal ball centered at x, and define the function f_u by*

$$f_u(r) = \varphi(\exp_x(ru)).$$

Then

$$f_u'(r) = (D_{\mathbf{n}}\varphi)(\exp_x(ru))$$

and

$$f_u''(r) = (D_{\mathbf{n}}^2\varphi)(\exp_x(ru)).$$

Proof. Let $y = \exp_x(ru)$ with some unit vector $u \in T_xX$. Let α be the geodesic defined by $\alpha(t) = \exp_x(tu)$. Then

$$f_u'(r) = (T\varphi)(y)\, T\exp_x(ru)u = (T\varphi)(y)\alpha'(r).$$

By the global Gauss lemma of Chapter IX, Proposition 3.2, $\alpha'(r)$ is precisely the unit normal vector $\mathbf{n}(y)$. Hence the right side of the above equation is the Lie derivative of φ in the direction of this unit normal vector, which is none other than $(D_{\mathbf{n}}\varphi)(y)$. This proves the first formula. The second comes by iterating the first, thereby completing the proof.

Theorem 3.8. *Let φ be a C^2 function on a normal ball centered at the point $x \in X$. Let $S_r(x)$ for $r > 0$ be the Riemannian sphere of radius r centered at x, and contained in the ball. Let Δ_S denoted the Laplacian on $S = S_r(x)$. Let $\mathbf{n}$ be the unit radial field from x, let u be a unit vector in T_xX. Then for $y = \exp_x(ru)$ we have*

$$\Delta_X\varphi(y) = (\Delta_S\varphi_S)(y) - (D_{\mathbf{n}}^2\varphi)(y) - \left(J^{-1}D_{\mathbf{n}}J(r, u) + \frac{n-1}{r}\right)(D_{\mathbf{n}}\varphi)(y).$$

Proof. We apply Proposition 2.5 of Chapter XIV, which decomposes the Laplacian into a tangential part relative to a submanifold, which we now take to be the sphere $Y = S$; and a transversal part. The tangential part gives precisely the term $\Delta_S \varphi_S$ at y. For the transversal part, we apply Proposition 2.6 of Chapter XIV, which tells us that the value depends only on the restriction of φ to the normal manifold. But then, we can apply Lemma 3.7 and the formula which we found in Corollary 3.6 to conclude the proof.

For further applications of Jacobi lifts to volumes, cf. for instance [GHL 87/93], Chapter 3H.

XV, §4. THE HODGE STAR ON FORMS

We already touched on the star operation on functions, and we defined d^* on 1-forms. We now deal systematically with the star operation on alternating forms. I shall follow **Koszul's formalism in formulas S 1 through S 8** [Ko 57], which is quite elegant. A direct very brief treatment of just what is needed to get the global duality and adjointness of d, d^* using Stokes' theorem, will be done in a self-contained way ad hoc in Chapter XVIII, so that the reader need not go through the systematic formalism just to understand that particular application of Stokes' theorem.

Until further notice, we don't differentiate, and the theory is punctual, so:

We let T be a finite dimensional vector space over $\mathbf{R}$, *of dimension n, with r-forms φ, ψ in $L^r_a(T)$, and with vectors $v \in T$. We suppose that T has a positive definite scalar product g, and is oriented so we have a volume form $\Omega_g = \Omega$. We let $v^\vee$ be the 1-form dual to v under g.*

S 1. *There exists a unique isomorphism* $*\colon L^r_a(T) \to L^{n-r}_a(T)$ *such that for all $v_1, \dots, v_{n-r} \in T$ and $\varphi \in L^r_a(T)$ we have*

$$(*\varphi)(v_1, \dots, v_{n-r})\Omega = \varphi \wedge v_1^\vee \wedge \cdots \wedge v_{n-r}^\vee.$$

Proof. Given φ, the right side of the above equation is a multilinear alternating function of $v_1, \dots, v_{n-r}$ into the 1-dimensional space of n-forms, so having chosen Ω as a basis for this space, we get a real-valued form, which constitutes the coefficient of Ω on the left side. The association

$$\varphi \mapsto *\varphi$$

is obviously linear.

S 2. *We have* $*\Omega = 1$ *and* $*1 = \Omega$, *and for a function f,* $*(f\Omega) = f$.

Proof. Immediate from the definition **S 1** and Proposition 1.1.

S 3. *For $\varphi \in L^r_a(T)$ and $v_1, \dots, v_{n-r} \in T$ we have*

$$(*\varphi)(v_1, \dots, v_{n-r}) = *(\varphi \wedge v_1^\vee \wedge \cdots \wedge v_{n-r}^\vee).$$

Proof. Using **S 2** and **S 1**, we find:

$$\begin{aligned} *(\varphi \wedge v_1^\vee \wedge \cdots \wedge v_{n-r}^\vee) &= *[(*\varphi)(v_1 \wedge \cdots \wedge v_{n-r})\Omega] \\ &= (*\varphi)(v_1, \dots, v_{n-r})(*\Omega) \\ &= (*\varphi)(v_1, \dots, v_{n-r}). \end{aligned}$$

S 4. *For $\varphi \in L^r_a(T)$ and $v \in T$ we have*

$$*(\varphi \wedge v^\vee) = (*\varphi) \circ v.$$

Proof. Indeed,

$$\begin{aligned} (*(\varphi \wedge v^\vee))(v_1, \dots, v_{n-r-1}) &= *(\varphi \wedge v^\vee \wedge v_1^\vee \wedge \cdots \wedge v_{n-r-1}) \\ &= (*\varphi)(v, v_1, \dots, v_{n-r-1}) \\ &= ((*\varphi) \circ v)(v_1, \dots, v_{n-r-1}). \end{aligned}$$

For the next property we need a lemma independently of the star operation.

Lemma 4.1. *For $\varphi \in L^r_a(T)$, and $v, v_1, \dots, v_{n-r+1} \in T$, we have*

$$\begin{aligned} &(\varphi \circ v) \wedge v_1^\vee \wedge \cdots \wedge v_{n-r+1}^\vee \\ &\qquad = \sum_{i=1}^{n-r+1} (-1)^{r+i} \langle v^\vee, v_i \rangle (\varphi \wedge v_1^\vee \wedge \cdots \wedge \widehat{v_i^\vee} \wedge \cdots \wedge v_{n-1+1}^\vee). \end{aligned}$$

Proof. The basic formalism of forms tells us that the contraction with a vector is an anti-derivation on the algebra of forms (Chapter V, §5, **CON 3**). Since $\varphi \wedge v_1^\vee \wedge \cdots \wedge v_{n-r+1}^\vee$ has degree $n + 1$ and so is equal to 0, we find

$$\begin{aligned} 0 &= (\varphi \wedge v_1^\vee \wedge \cdots \wedge v_{n-r+1}^\vee) \circ v \\ &= (\varphi \circ v) \wedge v_1^\vee \wedge \cdots \wedge v_{n-r+1}^\vee \\ &\quad + \sum_{i=1}^{n-r+1} (-1)^{r+i-1} \varphi \wedge v_1^\vee \wedge \cdots \wedge (v_i^\vee \circ v) \wedge \cdots \wedge v_{n-r+1}^\vee. \end{aligned}$$

We observe that

$$v_i^\vee \circ v = \langle v_i, v\rangle_g = \langle v, v_i\rangle_g = \langle v^\vee, v_i\rangle,$$

to conclude the proof of the lemma.

S 5. *For any form $\varphi \in L_a^r(T)$ and $v \in T$ we have*

$$*(\varphi \circ v) = (-1)^{n-1}(*\varphi) \wedge v^\vee.$$

Proof. First, for all $v_1, \ldots, v_{n-r} \in T$ we have

$$(1) \qquad (*(\varphi \circ v))(v_1, \ldots, v_{n-r+1}) = *((\varphi \circ v) \wedge v_1^\vee \wedge \cdots \wedge v_{n-r+1}^\vee).$$

On the other hand,

$$(-1)^{n-1}(*\varphi) \wedge v^\vee = (-1)^{r+1} v^\vee \wedge *\varphi.$$

Hence

$$\begin{aligned}((-1)^{n-1}(*\varphi) \wedge v^\vee)(v_1, \ldots, v_{n-1+1}) &= (-1)^{r+1}(v^\vee \wedge *\varphi)(v_1, \ldots, v_{n-r+1}) \\ &= \sum_{i=1}^{n-r+1} (-1)^{r+1}\langle v^\vee, v_i\rangle((*\varphi)(v_1, \ldots, \hat{v}_i, \ldots, v_{n-r+1}) \\ &= \sum_{i=1}^{n-r+1} *(-1)^{r+1}\langle v^\vee, v_i\rangle(\varphi \wedge v_1^\vee \wedge \cdots \wedge \widehat{v_i^\vee} \wedge \cdots \wedge v_{n-r+1}^\vee).\end{aligned}$$

Using Lemma 4.1 and (1) concludes the proof.

We can do an induction on **S 5**, and also get a corollary:

S 6. *For $\varphi \in L_a^r(T)$, $** = (-1)^{r(n-1)}$ and*

$$(*\varphi) \wedge v_1^\vee \wedge \cdots \wedge v_r^\vee = (-1)^{r(n-1)} * (\varphi \circ v_1 \circ \cdots \circ v_r).$$

Proof. We have

$$\begin{aligned}(**\varphi)(v_1, \ldots, v_r) &= *((*\varphi) \wedge v_1^\vee \wedge \cdots \wedge v_r^\vee), \\ \text{[applying S 5 repeatedly]} &= (-1)^{r(n-1)} * (*(\varphi \circ v_1, \circ \cdots \circ v_r)).\end{aligned}$$

Since for any function f we have $*f = f*1$ and $**f = *f\Omega = f$, property **S 6** follows.

S 7. *Let S denote the $*$ operation. Then $S\colon L_a^r(T) \to L_a^{n-r}(T)$ is an isomorphism.*

This is immediate, but is stated for the record.

S 8. *Let $\varphi, \psi \in L_a^r(T)$. Then*

$$\varphi \wedge *\psi = \psi \wedge *\varphi.$$

Proof. The pairings of φ, ψ given by the expressions on the left and right are bilinear, so it suffices to verify the equality when

$$\varphi = v_1^\vee \wedge \cdots \wedge v_r^\vee \qquad \text{and} \qquad \psi = w_1^\vee \wedge \cdots \wedge w_r^\vee.$$

In this case, we obtain

$$\begin{aligned}\varphi \wedge *\psi &= (-1)^{r(n-r)}(*\psi) \wedge \varphi \\ &= (-1)^{r(n-r)} *(w_1^\vee \wedge \cdots \wedge w_r^\vee) \wedge v_1^\vee \wedge \cdots \wedge v_r^\vee \\ \text{[by S 6]} \quad &= (-1)^{r(n-r)}(-1)^{r(n-1)} *[(w_1^\vee \wedge \cdots \wedge w_r^\vee) \circ v_1 \circ \cdots \circ v_r] \\ \text{[by S 2]} \quad &= \det\langle w_i, v_j\rangle_g \Omega.\end{aligned}$$

But $\det\langle w_i, v_j\rangle_g = \det\langle v_i, w_j\rangle_g$, from which **S 8** follows.

The next formula proves that the star operation is given in a simple-minded way on natural basis elements for the wedge products. We shall use this property in Chapter XVIII, §4, in a self-contained way to make the results on integration independent of the general star formalism, but the next formula won't be used in the rest of this section or the next.

Proposition 4.2. *Let $\{v_1, \ldots, v_n\}$ be an orthonormal basis of T. Let $\omega_1, \ldots, \omega_n$ be the dual basis of 1-forms. Let $I = (i_1, \ldots, i_r)$ with $i_1 < \cdots < i_r$ and let $J = (j_1, \ldots, j_{n-r})$ with $j_1 < \cdots j_{n-r}$ be the complementary set such that $\{1, \ldots, n\}$ is a permutation of (I, J). Let $\epsilon(I, J)$ be the sign of the permutation. Assume $v_1, \ldots, v_n$ oriented. Let*

$$\omega_I = \omega_{i_1} \wedge \cdots \wedge \omega_{i_r} \qquad \text{and} \qquad \omega_J = \omega_{j_1} \wedge \cdots \wedge \omega_{j_{n-r}}.$$

Then

$$*\omega_I = \epsilon(I, J)\omega_J.$$

Proof. Directly from the definition of $\Omega = \Omega_g$ we have that

$$\Omega_g = \omega_1 \wedge \cdots \wedge \omega_n = v_1^\vee \wedge \cdots \wedge v_n^\vee.$$

At first, let J be an arbitrary sequence of $n-r$ indices among $(1,\dots,n)$. Then by **S 3**,

$$(*\omega_I)(v_{j_1},\dots,v_{j_{n-r}}) = *(\omega_I \wedge \omega_J),$$

which is $\neq 0$ if and only if J is the complementary set, i.e. (I, J) is a permutation of $(1,\dots,n)$. In this case, the right side of the above expression is simply $\epsilon(I, J) * \Omega = \epsilon(I, J)$. Alternatively, one may write

$$*\omega_I = \epsilon(I,J)\omega_J,$$

if (I, J) is a permutation of $(1,\dots,n)$, from which Proposition 4.2 follows.

We are now through with the punctual theory, and we pass to a Riemannian manifold (X, g), where the vector space T is replaced by the tangent bundle TX, and vectors are replaced by vector fields. We let D be the metric covariant derivative as usual. Also

$$\mathscr{A}^r(X) = \Gamma L_a^r(TX).$$

Proposition 4.3. *The star operation commutes with every D_ξ, i.e. for any vector field ξ and $\varphi \in \mathscr{A}^r(X)$, we have*

$$*D_\xi\varphi = D_\xi *\varphi.$$

Proof. For 0-forms (functions) and n-forms (functions times the volume form) the assertion is immediate by using Proposition 2.1, to the effect that $D_\xi \operatorname{vol}_g = 0$. Now let $\varphi \in \Gamma L_a^r(TX)$. Then:

$$\begin{aligned}
&(D_\xi *\varphi)(\xi_1,\dots,\xi_{n-r}) + \sum_{i=1}^{n-r} (*\varphi)(\xi_1,\dots,D_\xi\xi_i,\dots,\xi_{n-r}) \\
&= D_\xi\big((*\varphi)(\xi_1,\dots,\xi_{n-r})\big) \qquad \text{[because } D_\xi \text{ is a derivation]} \\
&= D_\xi *(\varphi \wedge \xi_1^\vee \wedge \cdots \wedge \xi_n^\vee) \qquad \text{[by } \mathbf{S\ 3}\text{]} \\
&= (*D_\xi)(\varphi \wedge \xi_1^\vee \wedge \cdots \wedge \xi_{n-r}^\vee) \qquad \text{[by the proposition for } n\text{-forms]} \\
&= *(D_\xi\varphi \wedge \xi_1^\vee \wedge \cdots \wedge \xi_{n-r}^\vee) + \sum_{i=1}^{n-r} *(\varphi \wedge \xi_1^\vee \wedge \cdots \wedge D_\xi\xi_i^\vee \wedge \cdots \wedge \xi_{n-r}^\vee) \\
&= (*D_\xi\varphi)(\xi_1,\dots,\xi_{n-r}) + \sum_{i=1}^{n-r} (*\varphi)(\xi_1,\dots,D_\xi\xi_i,\dots,\xi_{n-r}),
\end{aligned}$$

which proves the proposition.

We now define d^* in general to be

$$d^* = (-1)^{nr+n+1} *d* \quad \text{on } \mathscr{A}^r(X).$$

In Chapter XIII, §3 we shall define a scalar product on forms with compact support for which d^* will be seen to be the adjoint of d. For the moment, we continue with an essentially differential algebraic theory.

Proposition 4.4. *For φ, $\psi \in \mathscr{A}^r(X)$ we have*

$$d\varphi \wedge *\psi = \varphi \wedge (*d^*\psi) + d(\varphi \wedge *\psi).$$

Proof. Immediate from the definition of d^*, **S 6**, and the basic formula for d of a wedge product (a graded derivation).

Proposition 4.5. *Let $\xi_1,\dots,\xi_n$ be a frame of vector fields, and let $\xi_1',\dots,\xi_n'$ be the dual frame, that is $\langle \xi_i', \xi_j\rangle_g = \delta_{ij}$. Then for any form $\varphi \in \mathscr{A}^r(X)$ we have*

$$d^*\varphi = \sum_{i=1}^{n} (D_{\xi_i}\varphi) \circ \xi_1'.$$

Proof. Proposition 1.1. of Chapter VIII gives us an expression for $d(*\varphi)$ in terms of the frame. The dual frame is such that $\lambda_i^\vee = \xi_i'$. Then the formula of Proposition 4.4 is an immediate consequence of **S 5**.

Remark. *If the frame $\xi_1,\dots,\xi_n$ is orthonormal, then of course $\xi_i' = \xi_i$.*

We define the **Laplacian** associated with the Riemannian manifold (X, g) to be

$$\Delta = dd^* + d^*d, \quad \text{operating on each } \mathscr{A}^r(X).$$

On Euclidean space $\mathbf{R}^n$ with its standard positive definite scalar product, the Laplacian on functions is the usual operator (with the minus sign)

$$\Delta = -\sum \left(\frac{\partial}{\partial x_i}\right)^2.$$

As a more general example illustrating the role of Ricci curvature, we give the one higher dimensional version of Corollary 2.4. Let $\lambda \in \mathscr{A}^1(X)$. With the Ricci curvature in mind, we define $\operatorname{Ric}(\lambda)$ to be the scalar valued form such that, with respect to an orthonormal frame $\xi_1,\dots,\xi_n$, and any vector field ξ we have

$$\operatorname{Ric}(\lambda)(\xi) = \sum_i \langle (D_\xi D_{\xi_i} - D_{\xi_i} D_\xi)\lambda, \xi_i\rangle,$$

where we denote by $\langle \lambda, \xi\rangle$ the value of a 1-form λ on a vector field ξ.

Proposition 4.6. *Let $\xi_1, \dots, \xi_n$ be an orthonormal frame. As an operator on 1-forms, $\Delta\colon \mathscr{A}^1(X) \to \mathscr{A}^1(X)$ is given by*

$$\Delta = -\sum D^2_{\xi_i} - \text{Ric}.$$

Written in terms of the variables, this means

$$\langle \Delta\lambda, \xi \rangle = -\sum_i \langle D_{\xi_i} D_{\xi_i} \lambda, \xi \rangle - \sum_i \langle (D_\xi D_{\xi_i} - D_{\xi_i} D_\xi)\lambda, \xi_i \rangle.$$

Proof. By Proposition 4.5, we have

$$d^*\lambda = -\sum (D_{\xi_i}\lambda)(\xi_i)$$

and so by a general formula on covariant derivatives we get a value for $dd^*\lambda$, namely

$$\langle dd^*\lambda, \xi \rangle = -\sum_i \langle D_\xi D_{\xi_i}\lambda, \xi_i \rangle.$$

On the other hand, to get $d^*d\lambda$, we first note that by **COVD 6** of Chapter VIII, §1,

$$(d\lambda)(\xi, \eta) = \langle D_\xi \lambda, \eta \rangle - \langle D_\eta \lambda, \xi \rangle.$$

Again by Proposition 4.5,

$$\langle d^*d\lambda, \xi \rangle = \sum \langle D_{\xi_i} D_\xi \lambda, \xi_i \rangle - \sum \langle D_{\xi_i} D_{\xi_i} \lambda, \xi \rangle.$$

Adding the two expressions yields the formula of the proposition.

XV, §5. HODGE DECOMPOSITION OF DIFFERENTIAL FORMS

In this section we carry out a bit of pure algebra, applicable to the situation of the previous section, and also applicable to other situations, especially in the complex case. See for instance [Wel 80], pp. 147–148 and [GriH 76], Chapter 0, §6. We work axiomatically. To prove the axioms **H 1** and **H 2** below requires more extensive analytical tools than we use in this book, and specifically it requires the basic theory of elliptic operators. What is needed is carried out in the above references, and the essential is done in a self-contained way in Appendix 4 of [La 75].

Since the algebraic set up which follows applies to other differential operators besides the d we have been using, I use a more neutral letter D, which in the complex theory is taken to be the so-called $\bar{\partial}$ operator.

None of this section will be used in the rest of the book. It is included here only for the convenience of a reader wanting to see how the theory further develops, and to isolate clearly what is purely algebraic from what demands more differential analysis.

Let A be a vector space of dimension n over $\mathbf{R}$, with a positive definite scalar product $\langle\ ,\ \rangle$ and corresponding norm $\|\ \|$; or alternatively, the vector space may be over $\mathbf{C}$, with a positive definite hermitian product. Let

$$D\colon A \to A$$

be a linear map which has an adjoint (algebraic) D^*, that is

$$\langle Du, v \rangle = \langle u, D^*u \rangle \qquad \text{all } u, v \in A.$$

and such that $DD = 0$. We define the **Laplacian** of D to be

$$\Delta_D = \Delta = DD^* + D^*D.$$

We define $\mathbf{H}_D = \mathbf{H} = \ker \Delta$ to be the ***D*-harmonic space**. We assume the **Hodge Conditions**:

H 1. *The kernel $\mathbf{H} = \ker \Delta$ is finite dimensional.*

H 2. *We have $\mathbf{H}^\perp = \Delta A$.*

We then prove further properties as follows.

Since $\mathbf{H}$ is assumed finite dimensional, there is an orthogonal projection of A on $\mathbf{H}$, which we denote also by $\mathbf{H}$ if necessary, that is $\mathbf{H}(u)$ is the orthogonal projection of u on $\mathbf{H}$.

Theorem 5.1. *Under the above two Hodge conditions, we have*

$$\mathbf{H}^\perp = DA + D^*A,$$

and an orthogonal decomposition

$$A = \mathbf{H} \perp \Delta A = \mathbf{H} \perp DA \perp D^*A.$$

The restriction of Δ to $\mathbf{H}^\perp$ is invertible, and

$$\text{Ker } D = \mathbf{H} + DA.$$

Proof. By orthogonalization and **H 2**, given $u \in A$ we have

$$u = \mathbf{H}u + \Delta v = \mathbf{H}u + DD^*v + D^*Dv$$

with some $v \in A$. Hence A is contained in $\mathbf{H} + DA + D^*A$, so we get equality. Furthermore

$$\langle \Delta u, u \rangle = \|Du\|^2 + \|D^*u\|^2.$$

Hence $\Delta u = 0$ if and only if $Du = D^*u = 0$. (Each implication is immediate.) The adjointness relation then shows that DA, D^*A are orthogonal to $\mathbf{H}$, and $D^2 = 0$ implies the DA is orthogonal to D^*A, so we get the orthogonal decomposition

$$A = \mathbf{H} \perp DA \perp D^*A,$$

and $\Delta A = DA + D^*A$ by **H 2**. Since $\Delta\mathbf{H} = 0$ it follows that

$$\Delta\colon DA + D^*A \to DA + D^*A$$

is surjective, and so is an isomorphism, and thus Δ is invertible on $\mathbf{H}^\perp$. Finally $\mathbf{H} + DA$ is contained in the kernel of D, and D is injective on D^*A because

$$DD^*u = 0 \Rightarrow \langle DD^*u, u\rangle = 0 \Rightarrow \|D^*u\|^2 = 0.$$

This proves the theorem.

Remark 1. As a special case of the last formula, suppose $u \in A$ and u is perpendicular to Ker Δ. If u is **D-closed**, that is $Du = 0$, then $u = Dv$ for some $v \in A$, that is u is **D-exact**.

Remark 2. *If we denote by $H(A)$ the homology* Ker D/Im D *then we get an isomorphism of the homology with the harmonic space*

$$\mathbf{H} \approx H(A).$$

We let

$$G\colon A \to \mathbf{H}^\perp = \Delta A$$

be equal to 0 on $\mathbf{H}$, and be the inverse of Δ on ΔA. Then by definition,

$$G\Delta = \Delta G \quad \text{and} \quad I = \Delta G + H.$$

Furthermore:

G and Δ commute with D and D^*.

Proof. We have

$$\Delta D = (DD^* + D^*D)D = DD^*D \quad \text{and} \quad D\Delta = D(DD^* + D^*D) = DD^*D$$

so D commutes with Δ. Similarly for D^*. The commutation of D and D^* with G then follows since $G = \Delta^{-1}$ on ΔA.

Graded structure

Suppose that in addition that A is graded,

$$A = \bigoplus_{p=0}^{n} A^p,$$

that A^p is orthogonal to A^q for $q \neq p$, and that

$$D^p = D\colon A^p \to A^{p+1}$$

raises degrees by 1, so $D^*\colon A^p \to A^{p-1}$ lowers degrees by 1.

Under the above assumptions, we can define the homology of D in degree p to be

$$H^p(A) = \text{Ker } D^p/\text{Im } D^{p-1},$$

where D^p is D viewed as map from A^p to A^{p+1}. Immediately from Theorem 5.1 we obtain:

Theorem 5.2. *Let $\mathbf{H}^p = \mathbf{H} \cap H^p(A)$. Then*

$$\mathbf{H} = \bigoplus_{p=0}^{n} \mathbf{H}^p$$

and $H^p(A) \approx \mathbf{H}^p$, that is every class in Ker D^p mod Im D^{p-1} *has a unique representative in the harmonic space $\mathbf{H}^p$.*

The star operator

We suppose given an automorphism $S\colon A \to A$ which is an isomorphism

$$S\colon A^p \to A^{n-p}.$$

We assume:

S 1. *On A^p we have $S^2 = (-1)^{p(n-1)}$.*

S 2. $D^* = (-1)^{np+n+1}$ *on A^p.*

Proposition 5.3. *Under these assumptions, $D = SD^*S$ and $\mathbf{H}$, Δ, G commute with S.*

Proof. We give the proof when n is even for simplicity. For $u \in A^p$, we have:

$$SD^*Su = -S^2DS^2u = -S^2D(-1)^pu$$
$$= -(-1)^p(-1)^{p+1}Du$$
$$= Du,$$

so $D = SD^*S$.

For the commutation of S with Δ, we write, using the above,

$$S\Delta = -SDSDS - SSDSD,$$
$$\Delta S = -DSDSS - SDSDS.$$

On A^p, $SS = (-1)^p$, so it is immediate that SS commutes with DSD, thus showing that S commutes with Δ.

Since S commutes with Δ, it follows that

$$S\colon \mathbf{H} \to \mathbf{H}$$

induces an automorphism of $\mathbf{H}$ with itself. For $u \in A$ we have:

$Su - \mathbf{H}Su \in \mathbf{H}$ by definition of the orthogonal projection; and

$Su - S\mathbf{H}u = S\Delta Gu = \Delta SGu$ since Δ commutes with S.

Then

$Su - S\mathbf{H}u \perp \mathbf{H}$ since it lies in ΔA.

Subtracting shows that $\mathbf{H}Su - S\mathbf{H}u$ is both orthogonal to $\mathbf{H}$, and also lies in $\mathbf{H}$, so must be 0, whence $\mathbf{H}$ commutes with S. Since $G = \Delta^{-1}$ on $\mathbf{H}^\perp$ it follows that G also commutes with S, thus proving the proposition.

XV, §6. VOLUME FORMS IN A SUBMERSION

In this section we return to volume forms in general, in a way which leads naturally into the considerations of the next chapter on integration.

We begin by recalling some simple facts of multilinear algebra. Consider an exact sequence of finite dimensional real vector spaces

(1) $$0 \to T_y \to T_x \to T_z \to 0,$$

with $\dim T_y = p$, $\dim T_x = n$, and $\dim T_z = q$, so $p + q = n$. Then we

have the dual sequence of dual spaces (homs into the scalars)

(2) $$0 \to T_z^\vee \to T_x^\vee \to T_y^\vee \to 0.$$

The surjection on the right gives rise to a surjective linear map by restriction:

(3) $$\bigwedge^p T_x^\vee \to \bigwedge^p T_y^\vee \to 0,$$

and the injection on the left gives rise to an injective linear map

(4) $$0 \to \bigwedge^q T_z^\vee \to \bigwedge^q T_x^\vee.$$

Lemma 6.1. *There is a canonical isomorphism*

$$\bigwedge^p T_y^\vee \otimes \bigwedge^q T_z^\vee \to \bigwedge^n T_x^\vee$$

defined as follows. For $\omega \in \bigwedge^q T_z^\vee$ and $\eta \in \bigwedge^p T_y^\vee$, let $\tilde{\eta} \in \bigwedge^p T_x^\vee$ map on η in sequence (3). *The map*

$$(\eta, \omega) \mapsto \tilde{\eta} \wedge \omega$$

is independent of the choice of $\tilde{\eta}$, and defines the isomorphism.

Proof. Routine algebraic verification. The above lemma is sometimes stated in the form

$$\det(T_x^\vee) = \det(T_y^\vee) \otimes \det(T_z^\vee).$$

By a **non-singular** or **volume form** in $\bigwedge^n T_x$ we simply mean a non-zero form, so a basis for $\bigwedge^n T_x$. Of course this is merely the definition we have given previously, in case the manifold is a point.

As a consequence of Lemma 6.1, given a volume form $\Omega \in \bigwedge^n T_x^\vee$ and a volume form $\omega \in \bigwedge^q T_z^\vee$, there is a unique $\eta \in \bigwedge^p T_y^\vee$ such that

$$\Omega = \eta \otimes \omega,$$

or in other words, for any pre-image $\tilde{\eta}$,

$$\Omega = \tilde{\eta} \wedge \omega.$$

The above discussion was punctual. It applies to the case when x, z are points in a submersion

$$\pi\colon X \to Z,$$

with $\pi(x) = z$. We let $Y = Y_{\pi(x)}$ be a fiber, with the natural injection $j\colon Y \to X$, and we let $y \in Y$, $x = j(y)$. Then (1) is the exact sequence of tangent spaces

$$0 \to T_y Y \to T_x X \to T_{\pi(x)} Z \to 0.$$

Let Ω, ω be volume forms on X, Z respectively. For each $y \in X$ there is a volume form η_y on $T_y Y$ such that

$$\Omega_y = \eta_y \otimes \omega_{\pi(y)}.$$

Lemma 6.2. *Let $\pi\colon X \to Z$ be a submersion. Suppose X is orientable. Then every fiber Y_z is orientable. If Ω and ω are volume forms on X, Z respectively, then there exists a p-form $\tilde{\eta}$ on X whose restriction to each fiber $Y_{\pi(y)}$ as above is the form η_y such that $\Omega_y = \eta_y \otimes \omega_{\pi(y)}$. For any such $\tilde{\eta}$, we have*

$$\Omega = \tilde{\eta} \wedge \omega.$$

Proof. The orientability comes from the existence of the family of forms $\{\eta_y\}$, which is verified to be C^∞ in terms of coordinates. The local existence of $\tilde{\eta}$ is immediate. The global existence follows by using a partition of unity.

A p-form on X whose restriction to all fibers is 0 will be called **fiber null**. Two p-forms Ψ_1, Ψ_2 are thus called **fiber equivalent** if their difference $\Psi_1 - \Psi_2$ is fiber null. Two forms $\tilde{\eta}_1$ and $\tilde{\eta}_2$ which restrict to the same forms on the fibers in Lemma 6.2 are thus fiber equivalent.

Riemannian submersions

Next we deal with the Riemannian case. We start with punctual remarks. Let T_y, T_x, T_z be finite dimensional real vector spaces, with the exact sequence

$$0 \to T_y \to T_x \to T_z \to 0.$$

We suppose in addition that T_x, T_z have positive definite scalar products, and that $T_x \to T_z$ is **metric**, in the sense that it induces a linear isometry $T_y^\perp \xrightarrow{\approx} T_z$.

Lemma 6.3. *Under the above assumptions, let Ω_x and Ω_z be metric volume forms on T_x and T_z (so they determine an orientation). Then one of the possible (up to sign) metric volume forms Ω_y on T_y satisfies the relation*

$$\Omega_x = \Omega_y \otimes \Omega_z.$$

Proof. Let $\{e_1, \ldots, e_p\}$ be an orthonormal basis for T_y, and $\{e_{p+1}, \ldots, e_{p+q}\}$ an orthonormal basis for $T_y^\perp$. Together they form an orthonormal basis for T_x. The metric dual bases $\{e_1^\vee, \ldots, e_p^\vee\}$ and $\{e_{p+1}^\vee, \ldots, e_{p+q}^\vee\}$ form an orthonormal basis of the dual space, and with the appropriate orientation of $\{e_1, \ldots, e_p\}$,

$$\Omega_x = e_1^\vee \wedge \cdots \wedge e_p^\vee \wedge e_{p+1}^\vee \wedge \cdots e_{p+q}^\vee.$$

Note that $e_{p+1}^\vee, \ldots, e_{p+q}^\vee$ are the images of an orthonormal basis of $T_z^\vee$ under the natural injection

$$0 \to T_z^\vee \to T_x^\vee.$$

Then the lemma is an immediate consequence of the definitions.

The lemma applies to the case of a Riemannian submersion

$$\pi\colon X \to Z.$$

If $y \in X$, we apply the lemma by putting $T_y = T_y Y$, $T_x = T_y X$, and $T_z = T_{\pi(y)} Z$. Then we get:

Proposition 6.4. *Let $\pi\colon X \to Z$ be a Riemannian submersion. Suppose X, Z oriented, so Y_z is oriented for each z. Let Ω_X, Ω_Z be the Riemannian volume forms on X, Z respectively. Then for each $z \in Z$, the Riemannian volume form Ω_{Y_z} (with the determined orientation of Y_z) satisfies*

$$\Omega_X(y) = \Omega_{Y_z}(y) \otimes \Omega_Z(z).$$

The relation of Proposition 6.4 is punctual. However, the individual volume forms on the fibers locally are the restriction of a form on an open set of X itself. Indeed, if $\{\xi_1, \ldots, \xi_p\}$ is an orthonormal frame of vertical vector fields on X, suitably oriented, then

$$\Omega_Y = \xi_1^\vee \wedge \cdots \wedge \xi_p^\vee.$$

Then Ω_Y restricted to each fiber Y_z is the Riemannian volume form on Y_z. We call Ω_Y the **vertical metric volume form,** which is independent of the choice of vertical orthonormal frame, with the orientation determined by that of X and Z. In general, by a **vertical volume form** we mean a form equal to a positive function times Ω_Y, or equivalently, a form which can be expressed locally as a wedge product $\xi_1' \wedge \cdots \wedge \xi_p'$, where $\{\xi_1, \ldots, \xi_p\}$ is a suitably oriented orthogonal frame of vertical vector fields, and $\{\xi_1', \ldots, \xi_p'\}$ is the dual frame (in the sense of dual basis of vector spaces) vanishing on horizontal fields. Any two such forms differ by a function nowhere 0. Note that if $\{\xi_1, \ldots, \xi_p\}$ is a vertical orthonormal frame, then $\xi_i^\vee = \xi_i'$ for $i = 1, \ldots, p$.

Let Ψ be a p-form, and let v be a horizontal vector field. We define Ψ to be **v-constant over the fibers** if the Lie derivative $\mathscr{L}_v\Psi$ is fiber null, i.e. restricted to every fiber is 0. An important example of this condition when $v = v_X$ is lifted from a vector field on Z will be given in §8. For now, we have a general result.

Proposition 6.5. *Let $\pi: X \to Z$ be a Riemannian submersion. Let Ω_X and Ω_Z be Riemannian volume forms on X, Z respectively. Let v be a vector field on Z, and v_X its horizontal lift to X. Abbreviate div_X for $\operatorname{div}_{\Omega_X}$, and similarly for Z. Let Ω_Y be the vertical metric volume form, and let φ be the function such that*

$$(\mathscr{L}_{v_X}\Omega_Y) \wedge \Omega_Z = \varphi\Omega_X.$$

Then

$$\operatorname{div}_X(v_X) = \pi^*\operatorname{div}_Z(v) + \varphi.$$

If one can write $\Omega_Y = (\pi^*\delta)\Psi$ with some positive function δ, and a form Ψ which is v_X-constant over the fibers, then

$$\operatorname{div}_X(v_X) = \pi^*\operatorname{div}_Z(v) + \pi^*(v \cdot \log \delta).$$

Proof. The first formula comes from definition **DIV 2** of the divergence, and the fact that the Lie derivative is a derivation for the wedge product, by Chapter V, Proposition 5.3, **LIE 2**, namely

$$\begin{aligned}\mathscr{L}_{v_X}(\Omega_X) &= \mathscr{L}_{v_X}\Omega_Y \wedge \Omega_Z + \Omega_Y \wedge \mathscr{L}_{v_X}\Omega_Z\\ &= \varphi\Omega_X + \Omega_Y \wedge \pi^*\operatorname{div}_Z(v)\Omega_Z.\end{aligned}$$

The second condition is then immediate, because $\mathscr{L}_{v_X}\Psi$ is a form $\tilde{0}$, in the notation of Lemma 6.1. This concludes the proof.

Remark. One doesn't really need to assume that the function δ is positive, but then one must put the absolute value sign in the formula, with $\log|\delta|$. In any case, if a function δ exists, positive or negative, one can change the orientation to make it positive.

Next we give a result of Wu, tying together the trace of the second fundamental form, and the volume forms on fibers of the Riemannian submersion. This result will not be used later, but is included for its intrinsic interest. It shows directly how the divergence is related to the second fundamental form.

Theorem 6.6 (Wu). *Let $\pi: X \to Z$ be a Riemannian submersion, with X, Z connected. Let Ω_Y be the vertical metric volume form. Let v be a horizontal vector field. Let h_{12} be the second fundamental form. Then:*

(a) *The restriction of $\mathscr{L}_v\Omega_Y + \langle \operatorname{tr} h_{12}, v\rangle_g\Omega_Y$ to each fiber is 0.*

(b) *Let Ψ be a vertical volume form on X, v-constant over the fibers, and δ the function on X defined by $\Omega_Y = \delta\Psi$. Then*

$$v \cdot (\log \delta) = -\langle \operatorname{tr} h_{12}, v\rangle_g.$$

In particular, if ξ_{hor} denotes the horizontal component of a vector field ξ, and Ψ is v-constant over the fibers for all v, then

$$(\operatorname{gr} \log \delta)_{\text{hor}} = -\operatorname{tr} h_{12}.$$

Before going into the proof proper, we make some remarks concerning an operator which has already come up, but which now intervenes in a more systematic way.

Let X be a Riemannian manifold, with metric covariant derivative D. For any vector field η, we define

A 1. $$A_\eta = \mathscr{L}_\eta - D_\eta.$$

Since $[\eta, \xi] = D_\eta\xi - D_\xi\eta$, it follows immediately that for any vector fields ξ, η we have

A 2. $$A_\eta\xi = -D_\xi\eta.$$

We can extend A_η to a derivation on tensor fields, especially multilinear forms, since both the Lie and covariant derivatives extend, as in Chapters V, Proposition 5.3, **LIE 2**; and Chapter VIII, §1, **COVD 4**.

Now we come to the proof proper. The theorem is local, so we argue locally. Let $\{\xi_1, \ldots, \xi_p\}$ be an orthonormal frame of vertical vector fields, and $\{v_1, \ldots, v_q\}$ an orthonormal frame of horizontal ones. We let

$$\{\xi_1', \ldots, \xi_p', v_1', \ldots, v_q'\}$$

be the dual frame (in the sense of dual bases of algebra). Let f_{ij} be the functions and μ_i the horizontal field such that

(1) $$D_v\xi_i = \sum_j f_{ij}\xi_j + \mu_i.$$

Thus μ_i is a linear combination of $v_1, \dots, v_q$. Note that $f_{ij} = -f_{ji}$, because

$$\begin{aligned} f_{ij} = \langle D_v\xi_i, \xi_j\rangle_g &= v \cdot \langle \xi_i, \xi_j\rangle_g - \langle \xi_i, D_v\xi_j\rangle_g \\ &= -\langle \xi_i, D_v\xi_j\rangle_g \qquad \text{because } \langle \xi_i, \xi_j\rangle_g = \delta_{ij} \\ &= -f_{ji}. \end{aligned}$$

In particular,

(2) $$f_{ii} = 0 \quad \text{and} \quad \sum_i f_{ii} = 0.$$

We now claim that

(3) $$D_v\xi'_i = \sum_j f_{ij}\xi'_j + \mu'_i$$

where μ'_i is a linear combination of $v'_1, \dots, v'_q$. To see this, write

$$D_v\xi'_i = \sum \varphi_{ij}\xi'_j + \mu'_i.$$

Then

$$\varphi_{ij} = (D_v\xi'_i)(\xi_j) = -\xi'_i(D_v\xi_j) = -f_{ji} = f_{ij},$$

thus proving the claim.

Third, we claim that

(4) $$A_v\xi'_i = \sum_j \psi_{ij}\xi'_j + \lambda'_i \quad \text{with} \quad \psi_{ij} = -\langle v, D_{\xi_j}\xi_i\rangle g,$$

where λ'_i is a linear combination of $v'_1, \dots, v'_q$. To see this, we have

$$\begin{aligned} \psi_{ij} = (A_v\xi'_i)(\xi_j) &= -\xi'_i(A_v\xi_j) = -\langle A_v\xi_j, \xi_i\rangle_g \\ &= \langle D_{\xi_j}v, \xi_i\rangle_g = -\langle v, D_{\xi_j}\xi_i\rangle_g, \end{aligned}$$

thus proving the claim. In particular,

(5) $$\sum_i \psi_{ii} = -\langle \operatorname{tr} h_{12}, v\rangle_g.$$

We now prove Theorem 6.6(a). We have, using (3), (4):

$$\begin{aligned} \mathscr{L}_v\Omega_Y = \mathscr{L}_v(\xi'_1 \wedge \cdots \wedge \xi'_p) &= \sum_i \xi'_1 \wedge \cdots \wedge (D_v + A_v)\xi'_i \wedge \cdots \wedge \xi'_p \\ &= \sum_i (f_{ii} + \psi_{ii})\xi'_1 \wedge \cdots \wedge \xi'_p + \sum_i \xi'_1 \wedge \cdots \wedge (\mu'_i + \lambda'_i) \wedge \cdots \wedge \xi'_p \end{aligned}$$

(6) $$= -\langle \operatorname{tr} h_{12}, v\rangle_g\Omega_Y + \text{second term, using (2), (5).}$$

The restriction of the second term to the fibers is 0 because each $v'_1, \dots, v'_q$ restricts to 0 on the fibers, and μ'_i, λ'_i are linear combinations of these horizontal dual fields. This concludes the proof of (a).

Next suppose that Ψ is vertical and v-constant over the fibers. Let $\Omega_Y = \delta\Psi$. Then

(7) $$\begin{aligned} \mathscr{L}_v\Omega_Y &= (v \cdot \delta)\Psi + \delta(\mathscr{L}_v\Psi) \\ &= v \cdot (\log\delta)\Omega_Y + \delta\mathscr{L}_v\Psi. \end{aligned}$$

From (6) and (7) we conclude that the restriction of the function $v \cdot (\log\delta) + \langle \operatorname{tr} h_{12}, v\rangle_g$ to each fiber is 0, whence it is the zero function, thus proving (b). The last assertion is immediate from the definition of the gradient, thus concluding the proof.

XV, §7. VOLUME FORMS ON LIE GROUPS AND HOMOGENEOUS SPACES

Let G be a finite dimensional Lie group of dimension n, with unit element e. We denote L_a, R_a left and right translations by an element $a \in G$, so

$$L_a(x) = ax \quad \text{and} \quad R_a(x) = xa.$$

For an element $x \in G$, we define **conjugation** $\mathbf{c}_x$ or $\mathbf{c}(x)$ by

$$\mathbf{c}_x(y) = xyx^{-1} \quad \text{so that} \quad \mathbf{c}_x\colon G \to G \text{ is a Lie group automorphism.}$$

Note that $\mathbf{c}_x = L_x \circ R_x^{-1} = R_x^{-1} \circ L_x$ (left and right translation commute). We **define** the **Lie conjugation** $\mathbf{c}_{\text{Lie}}(x)$ by an element $x \in G$ by the functorial effect, that is

$$\mathbf{c}_{\text{Lie}}(x) = T\mathbf{c}_x(e),$$

so $|\mathbf{c}_{\text{Lie}}(x)|$ is the Jacobian of conjugation at the origin.

Remark. Suppose G is given as a Lie subgroup of $\mathrm{GL}_N(\mathbf{R})$ for some N. Then $T_e\mathrm{GL}_N(\mathbf{R}) = \mathrm{Mat}_N(\mathbf{R})$ is the space of $N \times N$ matrices, and T_eG is a subspace of $\mathrm{Mat}_N(\mathbf{R})$. Then it is immediate that for $g \in G$,

$$\mathbf{c}_{\text{Lie}}(g)v = gvg^{-1}$$

is actually conjugation in the ordinary sense of the word. Hence it does no harm to think of $\mathbf{c}_{\text{Lie}}(x)$ as such a conjugation. In any case, the map

$$x \mapsto \mathbf{c}_{\text{Lie}}(x)$$

is a representation of G in the group of linear automorphisms of $T_eG = \mathrm{Lie}(G)$. This representation will be called the **conjugation representation** of G.

Proposition 7.1. *The exponential commutes with conjugation, namely for* $v \in T_eG$, *we have*

$$\exp \mathbf{c}_{\mathrm{Lie}}(x)v = \mathbf{c}_x(\exp v) = x \exp(v)x^{-1}.$$

Proof. This is actually a special case of the general fact that if $f\colon G \to G'$ is a Lie group homomorphism, and $v \in T_eG$, then

$$f(\exp v) = \exp\big(Tf(e)v\big).$$

We apply this formula to $f = \mathbf{c}_x$. As to the general formula, one notes that $\alpha(t) = f\big(\exp(tv)\big)$ defines a 1-parameter subgroup α of G', and that $\alpha'(0) = Tf(e)v$ by the chain rule, so $\alpha(t) = \exp\big(Tf(e)tv\big)$ for all t, concluding the proof.

Left and right translations induce maps on vector fields and forms. Note that on contravariant objects such as a function φ, we have

$$(L_a\varphi)(x) = \varphi(a^{-1}x) \qquad \text{and} \qquad (R_a\varphi)(x) = \varphi(xa^{-1}),$$

because we want $\varphi(x) = (L_a\varphi)(L_ax)$, and similarly with R_a.

Let $f\colon X \to Y$ be a differential morphism of manifolds. For each p-form ω on Y we can take its pull back $f^*\omega$ given by the formula

$$(f^*\omega)_x(v_1, \dots, v_p) = \omega_{f(x)}\big(Tf(x)v_1, \dots, Tf(x)v_p\big).$$

If f is a differential isomorphism, and Ω is a volume form on Y, then $f^*\Omega$ is a volume form on X, and we also have the direct image f_* such that $f_*^{-1} = f^*$. We apply these to the two translations L_a and R_a.

Suppose Ω is a volume form on G. For simplicity of notation we omit the star, and write the transformation formula as

$$(L_a\Omega)_{ax}(L_aV) = \Omega_x(V) \qquad \text{where} \quad V \in \textstyle\bigwedge^n T_x = \det T_x.$$

Of course by L_aV we mean $\big(\det TL_a(x)\big)(V)$, where

$$\det TL_a(x) = \textstyle\bigwedge^n TL_a(x)\colon \bigwedge^n T_x \to \bigwedge^n T_{ax}$$

is the induced linear map on $\bigwedge^n T_x$.

Suppose Ω is a volume form on G, invariant under left translation, that is $L_a\Omega = \Omega$ for all $a \in G$. Then Ω is uniquely determined by its value at the origin e, that is by Ω_e, and the form Ω_x at a point x is obtained by translating Ω_e to x via L_x. Conversely, given a volume form on T_eG, i.e. a non-singular form of maximal degree on the tangent space at the origin, we can translate it to obtain an invariant volume form on G. Hence the left invariant volume forms on G constitute the non-zero elements of a 1-dimensional vector space over $\mathbf{R}$.

Let Ω be a left invariant volume form on G. Then $R_a\Omega$ is also left invariant, and hence there exists a real number $\chi(a) \neq 0$ such that

$$R_a\Omega = \chi(a)\Omega.$$

The number $\chi(a)$ does not depend on Ω, and is immediately seen to be a continuous homomorphism $\chi\colon G \to \mathbf{R}^*$ (multiplicative group of non-zero elements). If G is connected, then $\chi(G) \subset \mathbf{R}^+$. We say that G is **strictly unimodular** if $\chi = 1$, that is χ is trivial, and **unimodular** if $|\chi|$ is trivial, this corresponding to the standard terminology. A compact group is unimodular. For a connected group, the two notions of strictly unimodular and unimodular coincide.

Proposition 7.2. *We have* $\chi(a) = \det \mathbf{c}_{\mathrm{Lie}}(a)$ *for* $a \in G$.

Proof. We use $\mathbf{c}_a = L_a \circ R_a^{-1}$, and abbreviate $\mathbf{c}_aV = \det \mathbf{c}_{\mathrm{Lie}}(a)V$. Then for $V \neq 0$,

$$\begin{aligned}\Omega(V) &= (\mathbf{c}_a\Omega)_e(\mathbf{c}_aV) = (R_{a^{-1}}\Omega)_e\big((\det \mathbf{c}_{\mathrm{Lie}}(a)V\big)\\ &= \det \mathbf{c}_{\mathrm{Lie}}(a)(R_{a^{-1}}\Omega)_e(V) = \det \mathbf{c}_{\mathrm{Lie}}(a)\chi(a)^{-1}\Omega(V).\end{aligned}$$

Cancelling $\Omega(V)$ concludes the proof of the proposition.

Proposition 7.3. *Let Ω be a left invariant volume form on G. Then $\chi\Omega$ is right invariant, i.e. is a right Haar form.*

Proof. We have

$$R_a(\chi\Omega) = R_a(\chi)R_a(\Omega) = \chi(a^{-1})\chi\chi(a)\Omega = \chi\Omega,$$

thus proving the proposition.

Let X be a homogeneous space for G. For each $x \in X$, the isotropy group G_x is the closed subgroup of elements $g \in G$ such that $gx = x$. Thus G_x is a Lie subgroup. We have a G-homogeneous space isomorphism

$$G/G_x \to X \qquad \text{given by} \qquad g \mapsto gx.$$

If x, y are two elements of X, and $a \in G$, $ax = y$, then

$$G_y = aG_xa^{-1}.$$

In other words, the isotropy groups of the two points are conjugate in a natural way.

Let us look more closely at the standard model $X = G/H$ for a homogeneous space, with a Lie subgroup H. We denote by e_G, e_H, and $e_{G/H}$ or $e(G/H)$ the unit element in G, H and G/H respectively. By definition, the unit element of G/H is the coset eH of H in the space of cosets. Conjugation by an element $g \in G$ induces a differential homogeneous space isomorphism

$$\mathbf{c}_g\colon\ G/H \to G/\mathbf{c}_g(H) = G/gHg^{-1}.$$

Thus we have the tangent map $T\mathbf{c}_g(e_{G/H})$, which is a linear isomorphism on the tangent space at $e_{G/H}$. If $h \in H$, then

$$T\mathbf{c}_h(e_{G/H})\colon\ T_e(G/H) \to T_e(G/H)$$

is a linear automorphism of the tangent space of G/H at its natural origin. Of course, we also have conjugation both on H and on G, that is

$$T\mathbf{c}_h(e_G)\colon\ T_eG \to T_eG \quad \text{and} \quad T\mathbf{c}_h(e_H)\colon\ T_eH \to T_eH.$$

We may then take the determinant of the previous three linear maps, namely $\det_{G/H}$, $\det_G$ and $\det_H$, although we shall omit the subscript from det, since the reference to the ambient space is made clear by the points at which the maps are evaluated, that is $e_{G/H}$, e_G and e_H respectively.

Proposition 7.4. *For $h \in H$, we have*

$$\det T\mathbf{c}_h(e_G) = \det T\mathbf{c}_h(e_{G/H}) \cdot \det T\mathbf{c}_h(e_H).$$

More generally, let $\pi\colon X \to Z$ be a submersion, with a differential automorphism $f\colon X \to X$ commuting with π. Let $y \in X$ be a fixed point of f, and Y the fiber containing y. Then $f = f_X$ induces differential automorphisms f_Y and f_Z of the fiber and of Z; and

$$\det Tf_X(y) = \big(\det Tf_Y(y)\big)\big(\det Tf_z(\pi(y))\big).$$

Proof. Let $T_y = T_yY$, $T_x = T_yX$ and $T_z = T_{\pi(y)}Z$, so we have the exact sequence

$$0 \to T_y \to T_x \to T_z \to 0.$$

The map f induces tangent linear maps on each of those spaces, and we denote these by L_y, L_x, L_z, so

$$L_x = Tf_X(y), \qquad L_y = Tf_Y(y) \qquad \text{and} \qquad L_z = Tf_Z(z).$$

If V is a finite dimensional vector space of dimension p, we let $\det V = \bigwedge^p V$ be its maximal exterior product with itself. Similarly to Lemma 6.1, we have a natural isomorphism

$$\det T_x = (\det T_y) \otimes (\det T_z).$$

Concretely, if $\{v_1, \ldots, v_p\}$ is a basis of T_y and $\{w_1, \ldots, w_q\}$ is a basis of T_z, with representatives $\{\tilde{w}_1, \ldots, \tilde{w}_q\}$ in T_x, then

$$v_1 \wedge \cdots \wedge v_p \wedge \tilde{w}_1 \wedge \cdots \wedge \tilde{w}_q$$

is a basis of $\det T_x = \bigwedge^{p+q} T_x$. The scaling effect of $\det L_x$ is then equal to the product of the scaling effect on each factor, $(\det L_y)(\det L_z)$, which proves the general formula. The special case first stated in Proposition 7.4 occurs with $f = \mathbf{c}_h$ $(h \in H)$. This concludes the proof.

We define G/H to be **strictly unimodular** if $\chi_G = \chi_H$ on H. If X is a homogeneous space for G, and H is one of the isotropy groups, so X is G-homogeneous space isomorphic to G/H, we say that X is **strictly unimodular** if G/H is strictly unimodular. We make the similar definition for G/H being unimodular, using $|\chi|$ instead of χ. The next result gives the first significant application of strict modularity.

Proposition 7.5. *Let X be a homogeneous space for G. If X is strictly unimodular, then there exists a left G-invariant volume form on X, unique up to a constant multiple.*

Proof. We want to define the invariant form on G/H by translating a given volume form ω_e on $T_e(G/H)$. On G/H, the left translation L_h is induced by conjugation $\mathbf{c}_h$ on G. By Proposition 7.4 and the hypothesis, we have

$$\det TL_h(e_{G/H}) = \det T\mathbf{c}_h(e_{G/H}) = 1.$$

Hence $L_h\omega_e(G/H) = \omega_{e(G/H)}$, that is $\omega_{e(G/H)}$ is invariant under translations by elements of H. Then for any $g \in G$ we define

$$\omega_{gH} = L_g\omega_e(G/H).$$

The value on the right is independent of the coset representative g, and it is then clear that translation yields the desired G-invariant volume form on G/H. The uniqueness up to a constant factor follows because the invariant forms are determined linearly from their values at the origin, and the forms at the origin constitute a 1-dimensional space. This concludes the proof.

Remark. If both G and H are unimodular, then so is G/H. If H is compact, then H is unimodular. If G is unimodular in addition, so is G/H. The same goes for strict unimodularity. When one applies the above considerations to Haar measures and integration, what matters is modularity, not the strict modularity. Cf. Chapter XVI, Theorem 4.3, and Chapter XVI, §5 for a derivation of some of these results in the context of Haar measure. Of course, the existence of an invariant volume form on Lie groups was known at the end of nineteenth century. At the time, into the twentieth century, it was a problem whether an invariant measure could be found on any locally compact group, and this problem was solved by Haar, whence the name Haar measure. In the next section, we shall accordingly define Haar forms, to fit into the psychology which has developed since Haar's result, even though invariant forms were known long before this result.

I found dealing with the Haar forms rather than Haar measure to provide additional flexibility. Then one has to make a distinction between modularity and strict modularity, but it isn't at all serious for local results. In all examples I know, the number of components is finite, and local results can be reduced to the case when the groups are connected, sometimes by passing to finite covering.

XV, §8. HOMOGENEOUSLY FIBERED SUBMERSIONS

In [He 72], Helgason obtained a formula for the Laplacian in a Riemannin submersion admitting horizontal metric sections. The result was reproduced in his book [He 84], Chapter II, Theorem 3.7, and concerns the case when there is a homogeneity condition on the fibers of the submersion. The present section developed from the attempt by Wu and myself to understand Helgason's situation better, from the point of view of local Riemannian geometry. The results of §6 were developed with this goal in mind, and will thus have their first application here, together with an important fact due to Wu.

We start without a Riemannian structure. For the first two basic properties, we don't need finite dimensionality. *So let X, Z be connected possibly infinite dimensional manifolds, and let*

$$\pi: X \to Z$$

be a submersion. We shall say that the submersion is **homogeneously fibered** if it satisfies the following condition.

HF Condition. *There is a possibly infinite dimensional Lie group H acting as a group of differential automorphisms on X, preserving the*

fibers, such that at each point x, we have a differential isomorphism

$$H/H_x \to Y_{\pi(x)} = \pi^{-1}(\pi(x))$$

of H/H_x-principal homogeneous space given by $h \mapsto hx$.

Note that a submersion always admits local differential sections, but in general these do not need to be metric. Furthermore, the submersion need not admit a global section. The next proposition applies to local sections when the need arises, but we shall not use it in this book.

Proposition 8.1. *Suppose there is a section $\sigma: Z \to X$ of a homogeneously fibered submersion. Define*

$$\gamma: H \times Z \to X \qquad by \qquad \gamma(h, z) = h\sigma(z).$$

Then γ is a submersion.

Proof. The tangent map $T\gamma(h, z)$ is a surjective homomorphism of tangent spaces at each point. In fact, if we let $\gamma_h(\sigma(z)) = h(\sigma z) = \gamma(h, z)$, then $T\gamma_h(\sigma(z))$ gives a linear isomorphism of the tangent spaces to the fiber. On the other hand, $T\sigma$ gives a linear isomorphism of the tangent space $T_z Z$ to a subspace of $T_{\sigma(z)}X$, and we have the direct sum decomposition at the point $x = \sigma(z)$,

$$T_{\sigma(z)}X = T_x(Hx) \oplus \sigma_* T_z Z.$$

This concludes the proof.

Suppose in addition that $\pi: X \to Z$ is a Riemannian submersion, and H acts isometrically. We then say that $\pi: X \to Z$ is a **metrically homogeneously fibered submersion.** *We suppose this is the case from now on.* An immediate question which arises about the isotropy groups H_x is the extent to which they can vary (up to conjugation). I owe the next key result to Wu.

Theorem 8.2 (Wu). *Let $\pi: X \to Z$ be a metrically homogeneously fibered submersion. For any two points $x, y \in X$, the isotropy groups H_x, H_y are conjugate in H. In fact, let x, y be points of X which can be joined by the horizontal lift of a curve in Z. Then $H_x = H_y$, and the flow of the horizontal lift induces an H-homogeneous space isomorphism between the fibers at x and at y.*

Proof. We recall that the horizontal lift was defined in Chapter XIV, §3. Suppose first that x, y can be joined by a horizontal lift A. Let

$h \in H_x$. Since H acts isometrically on X, $h \circ A$ is the unique horizontal lift from $hx = x$ to hy. But $h \circ A$ has the same initial conditions as A, and so coincides with A by the uniqueness of solutions of differential equations. Hence $hy = y$, and $h \in H_y$. The reverse inclusion $H_y \subset H_x$ follows by symmetry, so $H_x = H_y$. Next, for arbitrary points $x, y \in X$, consider any curve in Z between $\pi(x)$ and $\pi(y)$. Then the horizontal lift of this curve in X joins x to a point y' in the same fiber as y, and the isotropy groups of y and y' are conjugate. Finally, let F be the flow of horizontal lifts, that is $F_t(x) = A_x(t)$, where A_x is the horizontal lift of a curve $\alpha_{\pi(x)}$ with initial condition $\pi(x)$ on Z. Then $t \mapsto F_t(hx)$ and $t \mapsto hA_x(t)$ are horizontal lifts with the same initial conditions, and so are equal. This concludes the proof.

We now assume finite dimensionality, so we have volume forms. In addition, we assume that the fibers are strictly unimodular, i.e. H/H_x is strictly unimodular for all x, in which case we say that the homogeneous fibration is **strictly unimodular**.

We then select a fixed Haar form on one of the coset spaces H/H_o with one of the isotropy groups. Then conjugation transforms this Haar form to a Haar form Haar_{H/H_x} for all $x \in X$.

Let Y_z be a fiber of the submersion, with $z = \pi(x)$, so we obtain a homogeneous space isomorphism $H/H_x \xrightarrow{\approx} Y_z = Y_{\pi(x)}$. Selecting two different points in the fiber above z give rise to different isomorphisms, but the unimodularity condition implies that the Haar form on H/H_x corresponds to a Haar form on Y_z independent of the choice of point x in the fiber. We denote this Haar form by Haar_{Y_z}.

Let Ω_X, Ω_Z be the Riemannian volume forms on X (resp. Z). Then there is a function δ on Z such that for each z, and $y \in Y_z$ we have

$$\Omega_{Y_z}(y) = \delta(z)\ \text{Haar}_{Y_z}(y).$$

It is immediate that δ is C^∞ (say from a local coordinate representation). We call δ the **Riemannian Haar density**. The **Haar form** Ψ on X is defined to be the p-form ($p =$ fiber dimension) whose restriction to each fiber is the Haar form as above, and which is 0 on decomposable elements containing a horizontal field. Equivalently, let $\{\xi_1, \ldots, \xi_p\}$ be a frame of vertical fields on some open subset of X, and $\{\mu_1, \ldots, \mu_q\}$ a frame of horizontal fields. Then there exists a function φ such that with the dual frame $\{\xi'_1, \ldots, \xi'_p, \mu'_1, \ldots, \mu'_q\}$ we have

$$\Psi = \varphi \xi'_1 \wedge \cdots \wedge \xi'_p,$$

and the restriction of Ψ to each fiber is the Haar form on the fiber. Thus in terms of the natural basis for p-forms arising from a choice of vector field frames, the Haar form has only a vertical component. If we denote by η_z the Haar form on Y_z, then in the notation of §6, we see that Ψ is one of the possible choices for $\tilde{\eta}$. Any other choice when expressed as a linear combination of decomposable p-forms would contain some factor μ'_j in each term other than the above expression for Ψ. Such terms restrict to 0 on the fibers.

Recall that if v is a vector field on Z, we let v_X be its unique horizontal lift in X.

Theorem 8.3. *Let $\pi: X \to Z$ be a metrically homogeneously fibered strictly unimodular submersion. Let v be a vector field on Z. Then the Haar form Ψ is v_X-constant over the fibers. If δ is the Riemannian Haar density, then*

$$\text{div}_X(v_X) = \pi^* \text{div}_Z(v) + \pi^*(v \cdot \log \delta).$$

Proof. Let α be an integral curve of v in Z and let A be its horizontal lift, so v_X restricts to A' on the curve. By Theorem 8.2, the flow F_t gives a homogeneous space isomorphism $Y_{\alpha(0)} \to Y_{\alpha(t)}$ of the fibers. Let $\Psi_{\alpha(t)}$ be the Haar form restricted to the fiber. By the unimodularity condition, $F_t^* \Psi_{\alpha(t)} = \Psi_{\alpha(0)}$, which is constant. We now use frames as in the remarks preceding the theorem. In taking $F_t^*(\Psi)$, we note that each term $F_t^*(\xi'_i)$ may have a horizontal component, so that in a neighborhood (in X) of a point of the fiber $Y_{\alpha(0)}$,

$$F_t^*(\Psi) = \Psi + \Phi_t,$$

where Φ_t contains a horizontal factor. The restriction of Φ_t to the fiber $Y_{\alpha(0)}$ is 0, so the restriction of $\mathscr{L}_{v_X}\Psi$ to the fiber $Y_{\alpha(0)}$ is 0. Hence Ψ is v_X-constant over the fibers. We can then apply Proposition 6.5 to conclude the proof.

Theorem 8.4 (Helgason). *Let $\pi: X \to Z$ be a Riemannian submersion metrically homogeneously fibered, and unimodular. Let δ be the Riemannian Haar density. Let Δ_X, Δ_Z be the Laplacians. Then for a function ψ on Z, we have*

$$\Delta_X(\pi^*\psi) = \pi^*((\Delta_Z \psi) - (\text{gr}_Z \log \delta) \cdot \psi).$$

Proof. All the work has been done, and the statement merely puts together Proposition 6.5 via Theorem 8.3, and the definition of the Laplacian as minus the divergence of the gradient.

Remark. Actually, Theorem 8.4 as stated above somewhat refines Helgason's original statement. In the original paper [He 72] the isotropy groups are compact in Theorem 3.2. Helgason normalizes the Haar measures on them to have total measure 1, and from a fixed Haar measure on H, he can then normalize the measures on the homogeneous spaces H/H_x with varying x. Here, we are able to use another normalization which stems from Wu's theorem that all the isotropy groups are in fact conjugate. The compactness condition is relaxed in Theorem 3.3, but other conditions intervene.

In addition, Helgason assumed the existence of local horizontal sections. He gave a semiglobal proof for his theorem, using the symmetry of the Laplace operator vis-à-vis the scalar product defined by integration. Helgason's argument is very nice, but it is completely bypassed here with a direct analysis based on local differential geometric considerations.

Finally, the minus sign differs from Helgason because he uses the negative Laplacian and we use the positive Laplacian.

Appendix. Direct Image of Differential Operators

In the preceding chapter and the present chapter we have been principally concerned with the behavior of differential geometric invariants under immersions and submersions, especially the Laplacian which we analyzed directly. It may be instructive to the reader to see how a somewhat more general object behaves, namely an arbitrary differential operator, which we now discuss briefly.

Let X be a finite dimensional manifold. By a **differential operator** on X we mean a linear map on the space of C^∞ functions on X,

$$D: \mathrm{Fu}(X) \to \mathrm{Fu}(X)$$

such that given a point in X, there is a chart U at that point with coordinates $(x) = (x_1, \ldots, x_n)$ such that in terms of these coordinates, D can be written in the form

$$D_U = \sum \varphi_{(j)}(x)\left(\frac{\partial}{\partial x_1}\right)^{j_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{j_n},$$

with coefficient functions $\varphi_{(j)}$ which are C^∞, and the sum is taken over a finite number of n-tuples $(j) = (j_1, \ldots, j_n)$ of integers $\geqq 0$. If $f\colon U \to V$ is a differential isomorphism, so a change of charts, then it is immediately verified from the chain rule that D has a similar expression D_V on V. Furthermore, at a given point a in U, let

$$P_a(T_1, \ldots, T_n) = \sum \varphi_{(j)}(a) T_1^{j_1} \cdots T_n^{j_n}.$$

We call P the **polynomial representing** D **in the chart at** a. Then from the chain rule, it is immediate that the polynomial representing D in V at $f(a)$ has degree at most equal to the degree of the polynomial representing D in U at a. Hence these degrees are equal, and define the **degree of the differential operator** at a point (independently of the chart).

We denote the set of differential operators on X by $\mathrm{DO}(X)$. It is clear that $\mathrm{DO}(X)$ is an **R**-algebra.

Suppose given a submanifold Y of X and an open subset S of X containing Y, together with a submersion

$$\pi\colon S \to Y$$

which is the identity on Y. Thus locally in a chart, π is a projection and Y itself is a section of the submersion. Given these data, we define the **direct image**

$$\pi_*\colon \mathrm{DO}(X) \to \mathrm{DO}(Y)$$

as follows. Given a function f on Y, we consider the composite function $f \circ \pi$, apply D, and restrict the resulting function to Y, so by definition

$$(\pi_* D)f = D(f \circ \pi)_Y,$$

the subscript denoting restriction to Y. The operator $\pi_* D$ is a linear operator on functions. It is in fact a differential operator. One sees this by picking a chart such that in this chart, π is a projection

$$\pi\colon W \times V \to V,$$

with V a chart in Y. Let the coordinates be (w, y) with $y \in V$ and $w \in W$. Then D is represented in the chart $W \times V$ as a sum

$$D = \sum \varphi_{(j)}(w, y)\partial_1^{j_1} \cdots \partial_r^{j_r} + E(w, y),$$

where $r = \dim Y$, $\partial_i = \partial/\partial_{y_i}$, and $E(y, w)$ is a differential operator in the left ideal generated by $\partial/\partial_{w_1}, \ldots, \partial/\partial_{w_s}$ $(s = \dim W)$. For any function $f = f(y)$ on V, the function $f \circ \pi$ given by $(f \circ \pi)(w, y) = f(y)$ is annihilated by $E(w, y)$. If $b \in W$ and the charts are chosen such that the section Y is (b, Y) in the chart, then the above sum decomposition for D shows that

$$D(f \circ \pi)(y) = \sum \varphi_{(j)}(b, y)\, \partial_1^{j_1} \cdots \partial_r^{j_r},$$

so $\pi_* D$ is a differential operator.

We have two basic examples which arise in practice, for instance [He 84a], Chapter II, §5.

Example 1. Let X be a Riemannian manifold, and Y a submanifold with normal bundle NY. Let S be a tubular neighborhood of Y obtained from the exponential map on a neighborhood of the zero section, which we identify with Y. Let $\pi\colon S \to Y$ be the orthogonal projection, which projects a point to Y along the normal geodesic in S. We can always pick the neighborhood of the zero section in NY so that there is a unique such normal geodesic locally. Then we are in the situation discussed above, and the direct image $\pi_*\colon \mathrm{DO}(X) \to \mathrm{DO}(Y)$ is called the **normal projection** of differential operators on Y.

Example 2. Let $\pi\colon X \to Z$ be a homogeneously fibered submersion, as defined at the beginning of §8. By Proposition 8.1, we can always find locally a section in the neighborhood of a given point of Z. Then the map γ defined in Proposition 8.1, is a submersion, and we may identify $Y = \sigma(Z)$ with (e, Z) (letting e be the unit element of H). We cannot define

$$\pi_*\colon \mathrm{DO}(X) \to \mathrm{DO}(Z).$$

in general, but we can define π_* in a natural way on a subset of $\mathrm{DO}(X)$. Indeed, an element of the group H acting on X also acts on any object functorially associated with X, especially on $\mathrm{DO}(X)$. By definition, given $h \in H$, let $[h]D$ for $D \in \mathrm{DO}(X)$ be defined by

$$\big(([h]D)f\big) = \big(D(f \circ L_h)\big) \circ L_h^{-1}$$

where L_h is left translation by h, so that for $x \in X$,

$$\big(([h]D)f\big)(x) = D(f \circ L_h)(h^{-1}x).$$

We say that D is ***H*-invariant** if $[h]D = D$ for all $h \in H$. The set of H-invariant differential operators is a subalgebra of $\mathrm{DO}(X)$, which we denote by $\mathrm{DO}(X)^H$. We can then define

$$\pi_*\colon \mathrm{DO}(X)^H \to \mathrm{DO}(Z)$$

as follows. For a function, f on Z, we let

$$(\pi_* D)f = D(f \circ \pi)_Z.$$

This means that $D(f \circ \pi)$ is constant on the fibers of π, that is $D(f \circ \pi)$ is an H-invariant function, which therefore factors through a function on Z. We denote this function by inserting the subscript Z. To verify that $D(f \circ \pi)$ is constant on fibers, put $F = f \circ \pi$, so that F is a function on X, constant on fibers. For $h \in H$, let $[h]F = F \circ L_h^{-1}$. Then

$$[h](DF) = \big([h]D\big)\big([h]F\big) = \big([h]D\big)(F \circ L_h^{-1})$$
$$= DF,$$

because we assumed $D \in \mathrm{DO}(X)^H$. Thus $D(f \circ \pi)$ is constant on orbits of H. Hence $(\pi_* D)f = D(f \circ \pi)_Z$ defines a linear map $\mathrm{DO}(X)^H \to \mathrm{DO}(Z)$. This map is a differential operator. One can see this either as a special case of the general discussion, using the section of Proposition 8.1, or one can simply rewrite the local formula for the differential operator on the submersion, and use the H-invariance to see that the coefficient functions $\varphi_{(j)}(w, x)$ are H-invariant, that is $\varphi_{(j)}(hw, x) = \varphi_{(j)}(w, x)$ for $h \in H$ and $w \in W$ as before.

There is even a more jazzed up way of seeing that a linear operator on the space of functions is a differential operator, namely:

Theorem (Peetre–Carleson). *Let X be a manifold, and let*

$$L\colon \mathrm{Fu}(X) \to \mathrm{Fu}(X)$$

be a linear map which decreases supports, that is

$$\mathrm{supp}(Lf) \subset \mathrm{supp}(f)$$

for all functions $f \in \mathrm{Fu}(X)$. Then L is a differential operator.

The proof takes about two pages. Cf. for instance [Nar 68], whose proof is reproduced in Helgason [He 84], Chapter II, Theorem 1.4. See also [GHL 87/93], pp. 40 and 191, and further references therein.

CHAPTER XVI

Integration of Differential Forms

The material of this chapter is also contained in my book on real analysis [La 93], but it may be useful to the reader to have it also here in a rather self contained way, based only on standard properties of integration in Euclidean space.

Throughout this chapter, μ is Lebesgue measure on $\mathbf{R}^n$.
If A is a subset of $\mathbf{R}^n$, we write $\mathscr{L}^1(A)$ instead of $\mathscr{L}^1(A, \mu, \mathbf{C})$.
All manifolds are assumed finite dimensional.
They may have a boundary.

XVI, §1. SETS OF MEASURE 0

We recall that a set has **measure 0** in $\mathbf{R}^n$ if and only if, given ϵ, there exists a covering of the set by a sequence of rectangles $\{R_j\}$ such that $\sum \mu(R_j) < \epsilon$. We denote by R_j the closed rectangles, and we may always assume that the interiors R_j^0 cover the set, at the cost of increasing the lengths of the sides of our rectangles very slightly (an $\epsilon/2^n$ argument). We shall prove here some criteria for a set to have measure 0. We leave it to the reader to verify that instead of rectangles, we could have used cubes in our characterization of a set of a measure 0 (a cube being a rectangle all of whose sides have the same length).

We recall that a map f satisfies a **Lipschitz condition** on a set A if there exists a number C such that

$$|f(x) - f(y)| \leqq C|x - y|$$

for all $x, y \in A$. Any C^1 map f satisfies locally at each point a Lipschitz condition, because its derivative is bounded in a neighborhood of each point, and we can then use the mean value estimate,

$$|f(x) - f(y)| \leqq |x - y| \sup|f'(z)|,$$

the sup being taken for z on the segment between x and y. We can take the neighborhood of the point to be a ball, say, so that the segment between any two points is contained in the neighborhood.

Lemma 1.1. *Let A have measure 0 in $\mathbf{R}^n$ and let $f\colon A \to \mathbf{R}^n$ satisfy a Lipschitz condition. Then $f(A)$ has measure 0.*

Proof. Let C be a Lipschitz constant for f. Let $\{R_j\}$ be a sequence of cubes covering A such that $\sum \mu(R_j) < \epsilon$. Let r_j be the length of the side of R_j. Then for each j we see that $f(A \cap S_j)$ is contained in a cube R_j' whose sides have length $\leqq 2Cr_j$. Hence

$$\mu(R_j') \leqq 2^n C^n r_j^n = 2^n C^n \mu(R_j).$$

Our lemma follows.

Lemma 1.2. *Let U be open in $\mathbf{R}^n$ and let $f\colon U \to \mathbf{R}^n$ be a C^1 map. Let Z be a set of measure 0 in U. Then $f(Z)$ has measure 0.*

Proof. For each $x \in U$ there exists a rectangle R_x contained in U such that the family $\{R_x^0\}$ of interiors covers Z. Since U is separable, there exists a denumerable subfamily covering Z, say $\{R_j\}$. It suffices to prove that $f(Z \cap R_j)$ has measure 0 for each j. But f satisfies a Lipschitz condition on R_j since R_j is compact and f' is bounded on R_j, being continuous. Our lemma follows from Lemma 1.1.

Lemma 1.3. *Let A be a subset of $\mathbf{R}^m$. Assume that $m < n$. Let*

$$f\colon A \to \mathbf{R}^n$$

satisfy a Lipschitz condition. Then $f(A)$ has measure 0.

Proof. We view $\mathbf{R}^m$ as embedded in $\mathbf{R}^n$ on the space of the first m coordinates. Then $\mathbf{R}^m$ has measure 0 in $\mathbf{R}^n$, so that A has also n-dimensional measure 0. Lemma 1.3 is therefore a consequence of Lemma 1.1.

Note. All three lemmas may be viewed as stating that certain parametrized sets have measure 0. Lemma 1.3 shows that parametrizing a set by strictly lower dimensional spaces always yields an image having measure 0. The other two lemmas deal with a map from one space into another of the same dimension. Observe that Lemma 1.3 would be false if f is only assumed to be continuous (Peano curves).

The next theorem will be used later only in the proof of the residue theorem, but it is worthwhile inserting it at this point.

Let $f\colon X \to Y$ be a morphism of class C^p, with $p \geqq 1$, and assume throughout this section that X, Y are finite dimensional. A point $x \in X$ is called a **critical point** of f if f is not a submersion at x. This means that

$$T_x f\colon T_x X \to T_{f(x)} Y$$

is not surjective, according to our differential criterion for a submersion.

Assume that a manifold X has a countable base for its charts. Then we can say that a set has measure 0 in X if its intersection with each chart has measure 0.

Theorem 1.4 (Sard's Theorem). *Let $f\colon X \to Y$ be a C^∞ morphism of finite dimensional manifolds having a countable base. Let Z be the set of critical points of f in X. Then $f(Z)$ has measure 0 in Y.*

Proof. (Due to Dieudonné.) By induction on the dimension n of X. The assertion is trivial if $n = 0$. Assume $n \geqq 1$. It will suffice to prove the theorem locally in the neighborhood of a point in Z. We may assume that $X = U$ is open in $\mathbf{R}^n$ and

$$f\colon U \to \mathbf{R}^p$$

can be expressed in terms of coordinate functions,

$$f = (f_1, \ldots, f_p).$$

We let us usual

$$D^\alpha = D_1^{\alpha_1} \cdots D_n^{\alpha_n}$$

be a differential operator, and call $|\alpha| = \alpha_1 + \cdots + \alpha_n$ its **order**. We let $Z_0 = Z$ and for $m \geqq 1$ we let Z_m be the set of points $x \in Z$ such that

$$D^\alpha f_j(x) = 0$$

for all $j = 1, \ldots, p$ and all α with $1 \leqq |\alpha| \leqq m$. We shall prove:

(1) *For each $m \geqq 0$ the set $f(Z_m - Z_{m+1})$ has measure 0.*

(2) *If $m \geqq n/p$, then $f(Z_m)$ has measure 0.*

This will obviously prove Sard's theorem.

Proof of **(1)**. Let $a \in Z_m - Z_{m+1}$. Suppose first that $m = 0$. Then for some coordinate function, say $j = 1$, and after a renumbering of the variables if necessary, we have

$$D_1 f_1(a) \neq 0.$$

The map

$$g\colon x \mapsto (f_1(x), x_2, \ldots, x_p)$$

obviously has an invertible derivative at $x = a$, and hence is a local isomorphism at a. Considering $f \circ g^{-1}$ instead of f, we are reduced to the case where f is given by

$$f(x) = (x_1, f_2(x), \ldots, f_p(x)) = (x_1, h(x)),$$

where h is the projection of f on the last $p - 1$ coordinates and is therefore a morphism $h\colon V \to \mathbf{R}^{p-1}$ defined on some open V containing a. Then

$$Df(x) = \begin{pmatrix} 1 & 0 \\ * & Dh(x) \end{pmatrix}.$$

From this it is clear that x is a critical point for f if and only if x is a critical point for h, and it follows that $h(Z \cap V)$ has measure 0 in $\mathbf{R}^{p-1}$. Since $f(Z)$ is contained in $\mathbf{R}^1 \times h(Z)$, we conclude that $f(Z)$ has measure 0 in $\mathbf{R}^p$ as desired.

Next suppose that $m \geqq 1$. Then for some α with $|\alpha| = m + 1$, and say $j = 1$, we have

$$D^\alpha f_1(a) \neq 0.$$

Again after a renumbering of the indices, we may write

$$D^\alpha f_1 = D_1 g_1$$

for some function g_1, and we observe that $g_1(x) = 0$ for all $x \in Z_m$, in a neighborhood of a. The map

$$g\colon x \mapsto (g_1(x), x_2, \ldots, x_n)$$

is then a local isomorphism at a, say on an open set V containing a, and we see that

$$g(Z_m \cap V) \subset \{0\} \times R^{n-1}.$$

We view g as a change of charts, and considering $f \circ g^{-1}$ instead of f, together with the invariance of critical points under changes of charts, we may view f as defined on an open subset of $\mathbf{R}^{n-1}$. We can then apply induction again to conclude the proof of our first assertion.

Proof of **(2)**. Again we work locally, and we may view f as defined on the closed n-cube of radius r centered at some point a. We denote this cube by $C_r(a)$. For $m \geqq n/p$, it will suffice to prove that

$$f(Z_m \cap C_r(a))$$

has measure 0. For large N, we cut up each side of the cube into N equal segments, thus obtaining a decomposition of the cube into N^n small cubes. By Taylor's formula, if a small cube contains a critical point $x \in Z_m$, then for any point y of this small cube we have

$$|f(y) - f(x)| \leqq K|x - y|^{m+1} \leqq K(2r/N)^{m+1},$$

where K is a bound for the derivatives of f up to order $m+1$, and we use the sup norm. Hence the image of Z_m contained in small cube is itself contained in a cube whose radius is given by the right-hand side, and whose volume in $\mathbf{R}^p$ is therefore bounded by

$$K^p(2r/N)^{p(m+1)}.$$

We have at most N^n such images to consider and we therefore see that

$$f(Z_m \cap C_r(a))$$

is contained in a union of cubes in $\mathbf{R}^p$, the sum of whose volumes is bounded by

$$K^p N^n (2r/N)^{p(m+1)} \leqq K^p(2r)^{p(m+1)} N^{n-p(m+1)}.$$

Since $m \geqq n/p$, we see that the right-hand side of this estimate behaves like $1/N$ as N becomes large, and hence that the union of the cubes in $\mathbf{R}^p$ has arbitrarily small measure, thereby proving Sard's theorem.

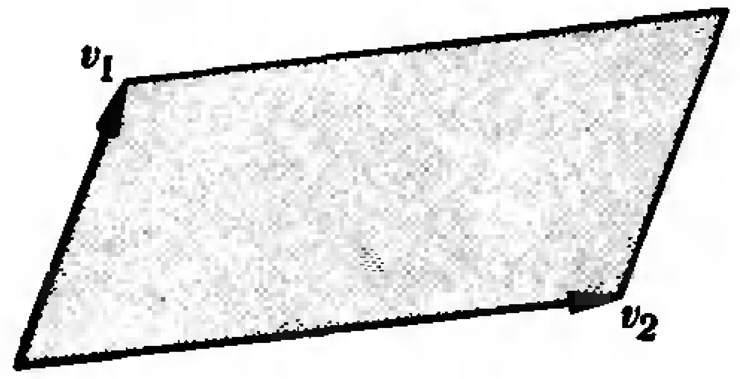
Sard's theorem is harder to prove in the case f is C^p with finite p [29], but $p = \infty$ already is quite useful.

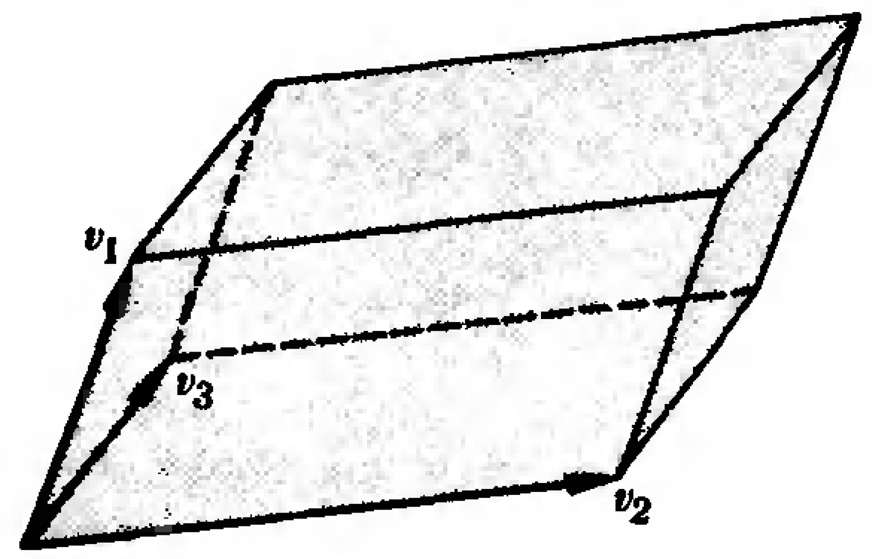
XVI, §2. CHANGE OF VARIABLES FORMULA

We first deal with the simplest of cases. We consider vectors $v_1, \ldots, v_n$ in $\mathbf{R}^n$ and we define the **block** B spanned by these vectors to be the set of points

$$t_1 v_1 + \cdots + t_n v_n$$

with $0 \leqq t_i \leqq 1$. We say that the block is **degenerate** (in $\mathbf{R}^n$) if the vectors $v_1, \ldots, v_n$ are linearly dependent. Otherwise, we say that the block is **non-degenerate**, or is a **proper block** in $\mathbf{R}^n$.





We see that a block in $\mathbf{R}^2$ is nothing but a parallelogram, and a block in $\mathbf{R}^3$ is nothing but a parallelepiped (when not degenerate).

We shall sometimes use the word volume instead of measure when applied to blocks or their images under maps, for the sake of geometry.

We denote by $\text{Vol}(v_1, \ldots, v_n)$ the volume of the block B spanned by $v_1, \ldots, v_n$. We define the **oriented volume**

$$\text{Vol}^0(v_1, \ldots, v_n) = \pm \text{Vol}(v_1, \ldots, v_n),$$

taking the $+$ sign if $\text{Det}(v_1, \ldots, v_n) > 0$ and the $-$ sign if

$$\text{Det}(v_1, \ldots, v_n) < 0.$$

The determinant is viewed as the determinant of the matrix whose column vectors are $v_1, \ldots, v_n$, in that order.

We recall the following characterization of determinants. Suppose that we have a product

$$(v_1, \ldots, v_n) \mapsto v_1 \wedge v_2 \wedge \cdots \wedge v_n$$

which to each n-tuple of vectors associates a number, such that the product is multilinear, alternating, and such that

$$e_1 \wedge \cdots \wedge e_n = 1$$

if $e_1, \ldots, e_n$ are the unit vectors. Then this product is necessarily the determinant, that is, it is uniquely determined. "Alternating" means that if $v_i = v_j$ for some $i \neq j$, then

$$v_1 \wedge \cdots \wedge v_n = 0.$$

The uniqueness is easily proved, and we recall this short proof. We can write

$$v_i = a_{i1}e_1 + \cdots + a_{in}e_n$$

for suitable numbers a_{ij}, and then

$$\begin{aligned} v_1 \wedge \cdots \wedge v_n &= (a_{11}e_1 + \cdots + a_{1n}e_n) \wedge \cdots \wedge (a_{n1}e_1 + \cdots + a_{nn}e_n) \\ &= \sum_\sigma a_{1,\sigma(1)}e_{\sigma(1)} \wedge \cdots \wedge a_{n,\sigma(n)}e_{\sigma(n)} \\ &= \sum_\sigma a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)}. \end{aligned}$$

The sum is taken over all maps $\sigma\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, but because of the alternating property, whenever σ is not a permutation the term corresponding to σ is equal to 0. Hence the sum may be taken only over all permutations. Since

$$e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = \epsilon(\sigma) e_1 \wedge \cdots \wedge e_n$$

where $\epsilon(\sigma) = 1$ or -1 is a sign depending only on σ, it follows that the alternating product is completely determined by its value $e_1 \wedge \cdots \wedge e_n$, and in particular is the determinant if this value is equal to 1.

Proposition 2.1. *We have*

$$\mathrm{Vol}^0(v_1, \ldots, v_n) = \mathrm{Det}(v_1, \ldots, v_n)$$

and

$$\mathrm{vol}(v_1, \ldots, v_n) = |\mathrm{Det}(v_1, \ldots, v_n)|.$$

Proof. If $v_1, \ldots, v_n$ are linearly dependent, then the determinant is equal to 0, and the volume is also equal to 0, for instance by Lemma 1.3. So our formula holds in the case. It is clear that

$$\mathrm{Vol}^0(e_1, \ldots, e_n) = 1.$$

To show that Vol^0 satisfies the characteristic properties of the determinant, all we have to do now is to show that it is linear in each variable, say the

first. In other words, we must prove

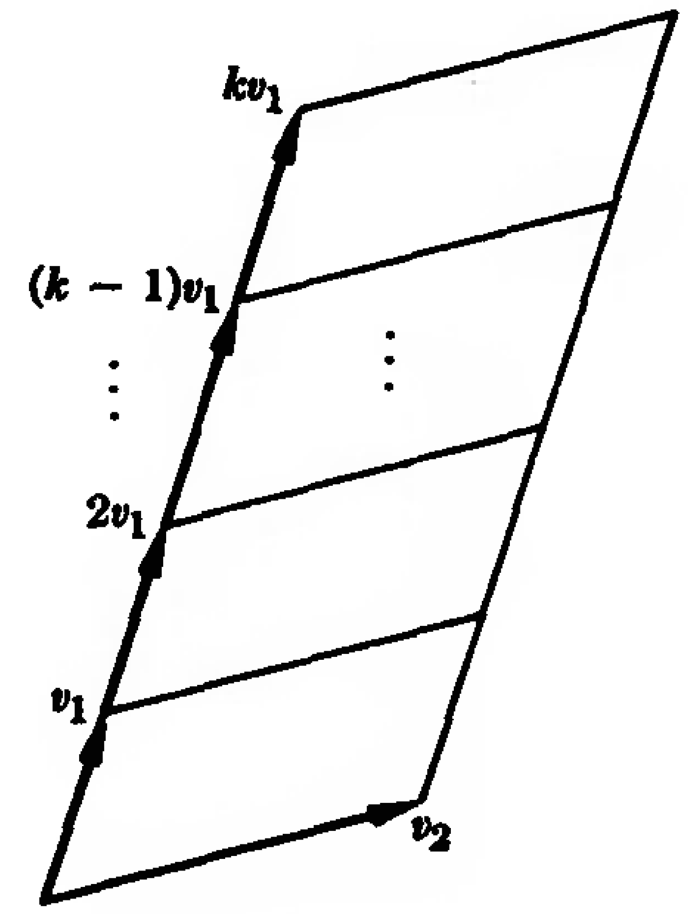
$$(*) \qquad \mathrm{Vol}^0(cv, v_2, \ldots, v_n) = c\,\mathrm{Vol}^0(v, v_2, \ldots, v_n) \qquad \text{for } c \in \mathbf{R},$$

$$(**) \qquad \mathrm{Vol}^0(v + w, v_2, \ldots, v_n) = \mathrm{Vol}^0(v, v_2, \ldots, v_n) + \mathrm{Vol}^0(w, v_2, \ldots, v_n).$$

As to the first assertion, suppose first that c is some positive integer k. Let B be the block spanned by $v, v_2, \ldots, v_n$. We may assume without loss of generality that $v, v_2, \ldots, v_n$ are linearly independent (otherwise, the relation is obviously true, both sides being equal to 0). We verify at once from the definition that if $B(v, v_2, \ldots, v_n)$ denotes the block spanned by $v, v_2, \ldots, v_n$ then $B(kv, v_2, \ldots, v_n)$ is the union of the two sets

$$B((k-1)v, v_2, \ldots, v_n) \qquad \text{and} \qquad B(v, v_2, \ldots, v_n) + (k-1)v$$

which have only a set of measure 0 in common, as one verifies at once from the definitions.



Therefore, we find that

$$\begin{aligned} \mathrm{Vol}(kv, v_2, \ldots, v_n) &= \mathrm{Vol}((k-1)v, v_2, \ldots, v_n) + \mathrm{Vol}(v, v_2, \ldots, v_n) \\ &= (k-1)\,\mathrm{Vol}(v, v_2, \ldots, v_n) + \mathrm{Vol}(v, v_2, \ldots, v_n) \\ &= k\,\mathrm{Vol}(v, v_2, \ldots, v_n), \end{aligned}$$

as was to be shown.

Now let

$$v = v_1/k$$

for a positive integer k. Then applying what we have just proved shows that

$$\mathrm{Vol}\left(\frac{1}{k}v_1, v_2, \ldots, v_n\right) = \frac{1}{k}\mathrm{Vol}(v_1, \ldots, v_n).$$

Writing a positive rational number in the form $m/k = m \cdot 1/k$, we conclude that the first relation holds when c is a positive rational number. If r is a positive real number, we find positive rational numbers c, c' such that $c \leqq r \leqq c'$. Since

$$B(cv, v_2, \ldots, v_n) \subset B(rv, v_2, \ldots, v_n) \subset B(c'v, v_2, \ldots, v_n),$$

we conclude that

$$c\,\mathrm{Vol}(v, v_2, \ldots, v_n) \leqq \mathrm{Vol}(rv, v_2, \ldots, v_n) \leqq c'\,\mathrm{Vol}(v, v_2, \ldots, v_n).$$

Letting c, c' approach r as a limit, we conclude that for any real number $r \geqq 0$ we have

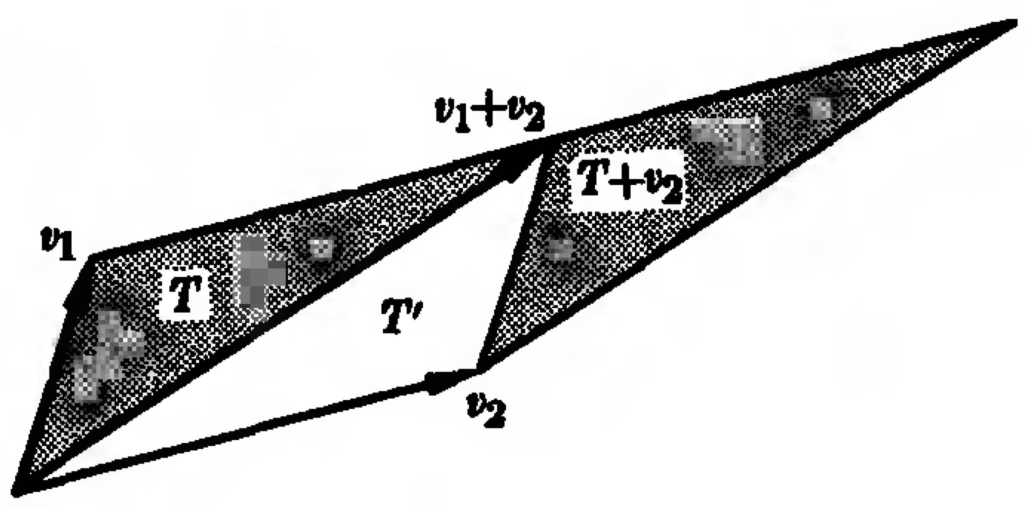
$$\mathrm{Vol}(rv, v_2, \ldots, v_n) = r\,\mathrm{Vol}(v, v_2, \ldots, v_n).$$

Finally, we note that $B(-v, v_2, \ldots, v_n)$ is the translation of

$$B(v, v_2, \ldots, v_n)$$

by $-v$ so that these two blocks have the same volume. This proves the first assertion.

As for the second, we look at the geometry of the situation, which is made clear by the following picture in case $v = v_1$, $w = v_2$.



The block spanned by $v_1, v_2, \ldots$ consists of two "triangles" T, T' having only a set of measure zero in common. The block spanned by $v_1 + v_2$ and v_2 consists of T' and the translation $T + v_2$. It follows that these two blocks have the same volume. We conclude that for any number c,

$$\mathrm{Vol}^0(v_1 + cv_2, v_2, \ldots, v_n) = \mathrm{Vol}^0(v_1, v_2, \ldots, v_n).$$

Indeed, if $c = 0$ this is obvious, and if $c \neq 0$ then

$$c\,\mathrm{Vol}^0(v_1 + cv_2, v_2) = \mathrm{Vol}^0(v_1 + cv_2, cv_2)$$
$$= \mathrm{Vol}^0(v_1 + cv_2) = c\,\mathrm{Vol}^0(v_1, v_2).$$

We can then cancel c to get our conclusion.

To prove the linearity of Vol^0 with respect to its first variable, we may assume that $v_2, \ldots, v_n$ are linearly independent, otherwise both sides of (**) are equal to 0. Let v_1 be so chosen that $\{v_1, \ldots, v_n\}$ is a basis of $\mathbf{R}^n$. Then by induction, and what has been proved above,

$$\begin{aligned}\mathrm{Vol}^0(c_1v_1 + \cdots + c_nv_n, v_2, \ldots, v_n) &= \mathrm{Vol}^0(c_1v_1 + \cdots + c_{n-1}v_{n-1}, v_2, \ldots, v_n)\\ &= \mathrm{Vol}^0(c_1v_1, v_2, \ldots, v_n)\\ &= c_1\,\mathrm{Vol}^0(v_1, \ldots, v_n).\end{aligned}$$

From this the linearity follows at once, and the theorem is proved.

Corollary 2.2. *Let S be the unit cube spanned by the unit vectors in $\mathbf{R}^n$. Let λ: $\mathbf{R}^n \to \mathbf{R}^n$ be a linear map. Then*

$$\mathrm{Vol}\,\lambda(S) = |\mathrm{Det}(\lambda)|.$$

Proof. If $v_1, \ldots, v_n$ are the images of $e_1, \ldots, e_n$ under λ, then $\lambda(S)$ is the block spanned by $v_1, \ldots, v_n$. If we represent λ by the matrix $A = (a_{ij})$, then

$$v_i = a_{1i}e_1 + \cdots + a_{ni}e_n,$$

and hence $\mathrm{Det}(v_1, \ldots, v_n) = \mathrm{Det}(A) = \mathrm{Det}(\lambda)$. This proves the corollary.

Corollary 2.3. *If R is any rectangle in $\mathbf{R}^n$ and λ: $\mathbf{R}^n \to \mathbf{R}^n$ is a linear map, then*

$$\mathrm{Vol}\,\lambda(R) = |\mathrm{Det}(\lambda)|\mathrm{Vol}(R).$$

Proof. After a translation, we can assume that the rectangle is a block. If $R = \lambda_1(S)$ where S is the unit cube, then

$$\lambda(R) = \lambda \circ \lambda_1(S),$$

whence by Corollary 2.2,

$$\mathrm{Vol}\,\lambda(R) = |\mathrm{Det}(\lambda \circ \lambda_1)| = |\mathrm{Det}(\lambda)\,\mathrm{Det}(\lambda_1)| = |\mathrm{Det}(\lambda)|\,\mathrm{Vol}(R).$$

The next theorem extends Corollary 2.3 to the more general case where the linear map λ is replaced by an arbitrary C^1-invertible map. The proof then consists of replacing the linear map by its derivative and estimating the error thus introduced. For this purpose, we have the **Jacobian determinant**

$$\Delta_f(x) = \operatorname{Det} J_f(x) = \operatorname{Det} f'(x),$$

where $J_f(x)$ is the Jacobian matrix, and $f'(x)$ is the derivative of the map $f\colon U \to \mathbf{R}^n$.

Proposition 2.4. *Let R be a rectangle in $\mathbf{R}^n$, contained in some open set U. Let $f\colon U \to \mathbf{R}^n$ be a C^1 map, which is C^1-invertible on U. Then*

$$\mu(f(R)) = \int_R |\Delta_f|\, d\mu.$$

Proof. When f is linear, this is nothing but Corollary 2.3 of the preceding theorem. We shall prove the general case by approximating f by its derivative. Let us first assume that R is a cube for simplicity. Given ϵ, let P be a partition of R, obtained by dividing each side of R into N equal segments for large N. Then R is partitioned into N^n subcubes which we denote by S_j $(j = 1, \ldots, N^n)$. We let a_j be the center of S_j.

We have

$$\operatorname{Vol} f(R) = \sum_j \operatorname{Vol} f(S_j)$$

because the images $f(S_j)$ have only sets of measure 0 in common. We investigate $f(S_j)$ for each j. The derivative f' is uniformly continuous on R. Given ϵ, we assume that N has been taken so large that for $x \in S_j$ we have

$$f(x) = f(a_j) + \lambda_j(x - a_j) + \varphi(x - a_j),$$

where $\lambda_j = f'(a_j)$ and

$$|\varphi(x - a_j)| \leqq |x - a_j|\epsilon.$$

To determine $\operatorname{Vol} f(S_j)$ we must therefore investigate $f(S)$ where S is a cube centered at the origin, and f has the form

$$f(x) = \lambda x + \varphi(x), \qquad |\varphi(x)| \leqq |x|\epsilon.$$

on the cube S. (We have made suitable translations which don't affect volumes.) We have

$$\lambda^{-1} \circ f(x) = x + \lambda^{-1} \circ \varphi(x),$$

so that $\lambda^{-1} \circ f$ is nearly the identity map. For some constant C, we have for $x \in S$

$$|\lambda^{-1} \circ \varphi(x)| \leqq C\epsilon.$$

From the lemma after the proof of the inverse mapping theorem, we conclude that $\lambda^{-1} \circ f(S)$ contains a cube of radius

$$(1 - C\epsilon)(\text{radius } S),$$

and trivial estimates show that $\lambda^{-1} \circ f(S)$ is contained in a cube of radius

$$(1 + C\epsilon)(\text{radius } S).$$

We apply λ to these cubes, and determine their volumes. Putting indices j on everything, we find that

$$\begin{aligned} &|\operatorname{Det} f'(a_j)| \operatorname{Vol}(S_j) - \epsilon C_1 \operatorname{Vol}(S_j) \\ &\qquad \leqq \operatorname{Vol} f(S_j) \leqq |\operatorname{Det} f'(a_j)| \operatorname{Vol}(S_j) + \epsilon C_1 \operatorname{Vol}(S_j) \end{aligned}$$

with some fixed constant C_1. Summing over j and estimating $|\Delta_f|$, we see that our theorem follows at once.

Remark. We assumed for simplicity that R was a cube. Actually, by changing the norm on each side, multiplying by a suitable constant, and taking the sup of the adjusting norms, we see that this involves no loss of generality. Alternatively, we can approximate a given rectangle by cubes.

Corollary 2.5. *If g is continuous on $f(R)$, then*

$$\int_{f(R)} g\, d\mu = \int_R (g \circ f)|\Delta_f|\, d\mu.$$

Proof. The functions g and $(g \circ f)|\Delta_f|$ are uniformly continuous on $f(R)$ and R respectively. Let us take a partition of R and let $\{S_j\}$ be the subrectangles of this partition. If δ is the maximum length of the sides of the subrectangles of the partition, then $f(S_j)$ is contained in a rectangle whose sides have length $\leqq C\delta$ for some constant C. We have

$$\int_{f(R)} g\, d\mu = \sum_j \int_{f(S_j)} g\, d\mu.$$

The sup and inf of g of $f(S_j)$ differ only by ϵ if δ is taken sufficiently small. Using the theorem, applied to each S_j, and replacing g by its minimum m_j and maximum M_j on S_j, we see that the corollary follows at once.

Theorem 2.6 (Change of Variables Formula). *Let U be open in $\mathbf{R}^n$ and let $f\colon U \to \mathbf{R}^n$ be a C^1 map, which is C^1 invertible on U. Let g be in $\mathscr{L}^1(f(U))$. Then $(g \circ f)|\Delta_f|$ is in $\mathscr{L}^1(U)$ and we have*

$$\int_{f(U)} g\, d\mu = \int_U (g \circ f)|\Delta_f|\, d\mu.$$

Proof. Let R be a closed rectangle contained in U. We shall first prove that the restriction of $(g \circ f)|\Delta_f|$ to R is in $\mathscr{L}^1(R)$, and that the formula holds when U is replaced by R. We know that $C_c(f(U))$ is L^1-dense in $\mathscr{L}^1(f(U))$, by [La 93], Theorem 3.1 of Chapter IX. Hence there exists a sequence $\{g_k\}$ in $C_c(f(U))$ which in L^1-convergent to g. Using [La 93], Theorem 5.2 of Chapter VI, we may assume that $\{g_k\}$ converges pointwise to g except on a set Z of measure 0 in $f(U)$. By Lemma 1.2, we know that $f^{-1}(Z)$ has measure 0.

Let $g_k^* = (g_k \circ f)|\Delta_f|$. Each function g_k^* is continuous on R. The sequence $\{g_k^*\}$ converges almost everywhere to $(g \circ f)|\Delta_f|$ restricted to R. It is in fact an L^1-Cauchy sequence in $\mathscr{L}^1(R)$. To see this, we have by the result for rectangles and continuous functions (corollary of the preceding theorem):

$$\int_R |g_k^* - g_m^*|\, d\mu = \int_{f(R)} |g_k - g_m|\, d\mu,$$

so the Cauchy nature of the sequence $\{g_k^*\}$ is clear from that of $\{g_k\}$. It follows that the restriction of $(g \circ f)|\Delta_f|$ to R is the L^1-limit of $\{g_k^*\}$, and is in $\mathscr{L}^1(R)$. It also follows that the formula of the theorem holds for R, that is

$$\int_{f(A)} g\, d\mu = \int_A (g \circ f)|\Delta_f|\, d\mu$$

when $A = R$.

The theorem is now seen to hold for any measurable subset A of R, since $f(A)$ is measurable, and since a function g in $\mathscr{L}^1(f(A))$ can be extended to a function in $\mathscr{L}^1(f(R))$ by giving it the value 0 outside $f(A)$. From this it follows that the theorem holds if A is a finite union of rectangles contained in U. We can find a sequence of rectangles $\{R_m\}$ contained in U whose union is equal to U, because U is separable. Taking the usual stepwise complementation, we can find a disjoint sequence of measurable sets

$$A_m = R_m - (R_1 \cup \cdots \cup R_{m-1})$$

whose union is U, and such that our theorem holds if $A = A_m$. Let

$$h_m = g_{f(A_m)} = g\chi_{f(A_m)} \quad \text{and} \quad h_m^* = (h_m \circ f)|\Delta_f|.$$

Then $\sum h_m$ converges to g and $\sum h_m^*$ converges to $(g \circ f)|\Delta_f|$. Our theorem follows from Corollary 5.13 of the dominated convergence theorem in [La 93].

Note. In dealing with polar coordinates or the like, one sometimes meets a map f which is invertible except on a set of measure 0, e.g. the polar coordinate map. It is now trivial to recover a result covering this type of situation.

Corollary 2.7. *Let U be open in $\mathbf{R}^n$ and let $f\colon U \to \mathbf{R}^n$ be a C^1 map. Let A be a measurable subset of U such that the boundary of A has measure 0, and such that f is C^1 invertible on the interior of A. Let g be in $\mathscr{L}^1(f(A))$. Then $(g \circ f)|\Delta_f|$ is in $\mathscr{L}^1(A)$ and*

$$\int_{f(A)} g\, d\mu = \int_A (g \circ f)|\Delta_f|\, d\mu.$$

Proof. Let U_0 be the interior of A. The sets $f(A)$ and $f(U_0)$ differ only by a set of measure 0, namely $f(\partial A)$. Also the sets A, U_0 differ only by a set of measure 0. Consequently we can replace the domains of integration $f(A)$ and A by $f(U_0)$ and U_0, respectively. The theorem applies to conclude the proof of the corollary.

XVI, §3. ORIENTATION

Let U, V be open sets in half spaces of $\mathbf{R}^n$ and let $\varphi\colon U \to V$ be a C^1 isomorphism. We shall say that φ is **orientation preserving** if the Jacobian determinant $\Delta_\varphi(x)$ is > 0, all $x \in U$. If the Jacobian determinant is negative, then we say that φ is orientation **reversing.**

Let X be a C^p manifold, $p \geqq 1$, and let $\{(U_i, \varphi_i)\}$ be an atlas. We say that this atlas is **oriented** if all transition maps $\varphi_j \circ \varphi_i^{-1}$ are orientation preserving. Two atlases $\{(U_i, \varphi_i)\}$ and $\{(V_\alpha, \psi_\alpha)\}$ are said to **define the same orientation**, or to be **orientation equivalent**, if their union is oriented. We can also define locally a chart (V, ψ) to be **orientation compatible** with the oriented atlas $\{(U_i, \varphi_i)\}$ if all transition maps $\varphi_i \circ \varphi^{-1}$ (defined whenever $U_i \cap V$ is not empty) are orientation preserving. An orientation equivalence class of oriented atlases is said to define an **oriented** manifold, or to be an **orientation** of the manifold. It is a simple exercise to verify that if a connected manifold has an orientation, then it has two distinct orientations.

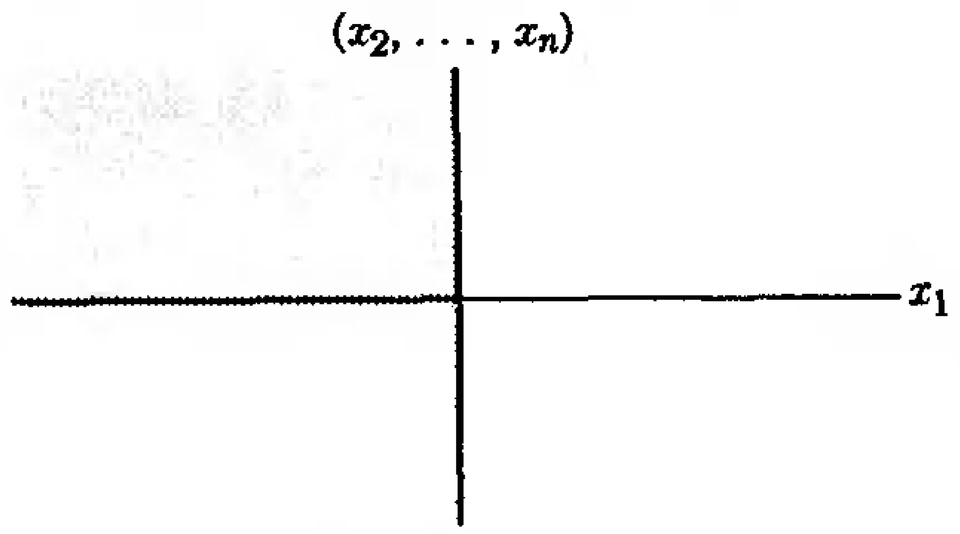
The standard examples of the Moebius strip or projective plane show that not all manifolds admit orientations. We shall now see that the boundary of an oriented manifold with boundary can be given a natural orientation.

Let φ: $U \to \mathbf{R}^n$ *be an oriented chart at a boundary point of X, such that:*

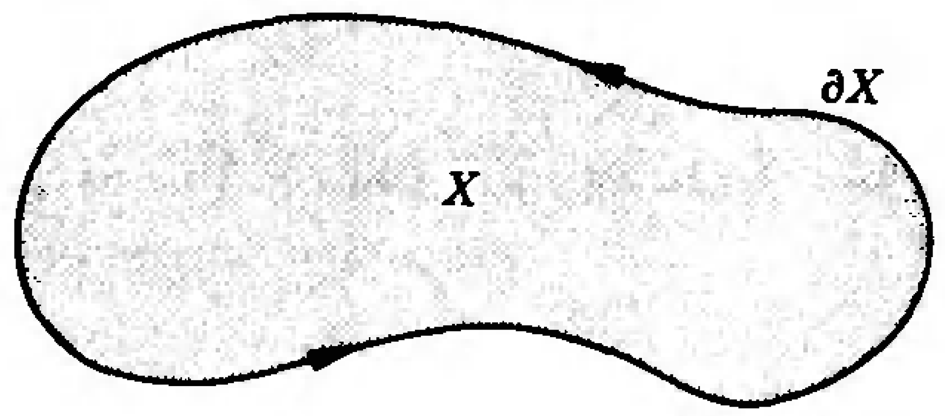
(1) *if* $(x_1, \dots, x_n)$ *are the local coordinates of the chart, then the boundary points correspond to those points in* $\mathbf{R}^n$ *satisfying* $x_1 = 0$; *and*

(2) *the points of U not in the boundary have coordinates satisfying* $x_1 < 0$.

Then $(x_2, \dots, x_n)$ are the local coordinates for a chart of the boundary, namely the restriction of φ to $\partial X \cap U$, and the picture is as follows.



We may say that we have considered a chart φ such that the manifold lies to the left of its boundary. If the reader thinks of a domain in $\mathbf{R}^2$, having a smooth curve for its boundary, as on the following picture, the reader will see that our choice of chart corresponds to what is usually visualized as "counterclockwise" orientation.



The collection of all pairs $(U \cap \partial X, \varphi|(U \cap \partial X))$, *chosen according to the criteria described above, is obviously an atlas for the boundary* ∂X, *and we contend that it is an oriented atlas.*

We prove this easily as follows. If

$$(x_1, \dots, x_n) = x \quad \text{and} \quad (y_1, \dots, y_n) = y$$

are coordinate systems at a boundary point corresponding to choices of charts made according to our specifications, then we can write $y = f(x)$ where $f = (f_1, \dots, f_n)$ is the transition mapping. Since we deal with

oriented charts for X, we know that $\Delta_f(x) > 0$ for all x. Since f maps boundary into boundary, we have

$$f_1(0, x_2, \dots, x_n) = 0$$

for all $x_2, \dots, x_n$. Consequently the Jacobian matrix of f at a point $(0, x_2, \dots, x_n)$ is equal to

$$\begin{pmatrix} D_1 f_1(0, x_2, \dots, x_n) & 0 \cdots\cdots 0 \\ * & \\ * & \Delta_g^{(n-1)} \\ * & \end{pmatrix},$$

where $\Delta_g^{(n-1)}$ is the Jacobian matrix of the transition map g induced by f on the boundary, and given by

$$\begin{aligned} y_2 &= f_2(0, x_2, \dots, x_n), \\ &\vdots \\ y_n &= f_n(0, x_2, \dots, x_n). \end{aligned}$$

However, we have

$$D_1 f_1(0, x_2, \dots, x_n) = \lim_{h \to 0} \frac{f_1(h, x_2, \dots, x_n)}{h},$$

taking the limit with $h < 0$ since by prescription, points of X have coordinates with $x_1 < 0$. Furthermore, for the same reason we have

$$f_1(h, x_2, \dots, x_n) < 0.$$

Consequently

$$D_1 f_1(0, x_2, \dots, x_n) > 0.$$

From this it follows that $\Delta_g^{(n-1)}(x_2, \dots, x_n) > 0$, thus proving our assertion that the atlas we have defined for ∂X is oriented.

From now on, when we deal with an oriented manifold, it is understood that its boundary is taken with orientation described above, and called the induced orientation.

XVI, §4. THE MEASURE ASSOCIATED WITH A DIFFERENTIAL FORM

Let X be a manifold of class C^p *with* $p \geqq 1$. *We assume from now on that X is Hausdorff and has a countable base. Then we know that X admits* C^p *partitions of unity, subordinated to any given open covering.*

(Actually, instead of the conditions we assumed, we could just as well have assumed the existence of C^p partitions of unity, which is the precise condition to be used in the sequel.)

We can define the **support** of a differential form as we defined the support of a function. It is the closure of the set of all $x \in X$ such that $\omega(x) \neq 0$. If ω is a form of class C^p and α is a C^q function on X, then we can form the product $\alpha\omega$, which is the form whose value at x is $\alpha(x)\omega(x)$. If α has compact support, then $\alpha\omega$ has compact support. Later, we shall study the integration of forms, and reduce this to a local problem by means of partitions of unity, in which we multiply a form by functions.

We assume that the reader is familiar with the correspondence between certain functionals on continuous functions with compact support and measures. Cf. [La 93] for this. We just recall some terminology.

We denote by $C_c(X)$ the vector space of continuous functions on X with **compact support** (i.e. vanishing outside a compact set). We write $C_c(X, \mathbf{R})$ or $C_c(X, \mathbf{C})$ if we wish to distinguish between the real or complex valued functions.

We denote by $C_K(X)$ the subspace of $C_c(X)$ consisting of those functions which vanish outside K. (Same notation $C_S(X)$ for those functions which are 0 outside any subset S of X. Most of the time, the useful subsets in this context are the compact subsets K.)

A linear map λ of $C_c(X)$ into the complex numbers (or into a normed vector space, for that matter) is said to be **bounded** if there exists some $C \geqq 0$ such that we have

$$|\lambda f| \leqq C\|f\|$$

for all $f \in C_c(X)$. Thus λ is bounded if and only if λ is continuous for the norm topology.

A linear map λ of $C_c(X)$ into the complex numbers is said to be **positive** if we have $\lambda f \geqq 0$ whenever f is real and $\geqq 0$.

Lemma 4.1. *Let λ: $C_c(X) \to \mathbf{C}$ be a positive linear map. Then λ is bounded on $C_K(X)$ for any compact K.*

Proof. By the corollary of Urysohn's lemma, there exists a continuous real function $g \geqq 0$ on X which is 1 on K has compact support. If $f \in C_K(X)$, let $b = \|f\|$. Say f is real. Then $bg \pm f \geqq 0$, whence

$$\lambda(bg) \pm \lambda f \geqq 0$$

and $|\lambda f| \leqq b\lambda(g)$. Thus λg is our desired bound.

A complex valued linear map on $C_c(X)$ which is bounded on each subspace $C_K(X)$ for every compact K will be called a **C_c-functional** on $C_c(X)$, or more simply, a **functional**. A functional on $C_c(X)$ which is also continuous for the sup norm will be called a **bounded** functional. It is clear that a bounded functional is also a C_c-functional.

Lemma 4.2. *Let $\{W_\alpha\}$ be an open covering of X. For each index α, let λ_α be a functional on $C_c(W_\alpha)$. Assume that for each pair of indices α, β the functionals λ_α and λ_β are equal on $C_c(W_\alpha \cap W_\beta)$. Then there exists a unique functional λ on X whose restriction to each $C_c(W_\alpha)$ is equal to λ_α. If each λ_α is positive, then so is λ.*

Proof. Let $f \in C_c(X)$ and let K be the support of f. Let $\{h_i\}$ be a partition of unity over K subordinated to a covering of K by a finite number of the open sets W_α. Then each $h_i f$ has support in some $W_{\alpha(i)}$ and we define

$$\lambda f = \sum_i \lambda_{\alpha(i)}(h_i f).$$

We contend that this sum is independent of the choice of $\alpha(i)$, and also of the choice of partition of unity. Once this is proved, it is then obvious that λ is a functional which satisfies our requirements. We now prove this independence. First note that if $W_{\alpha'(i)}$ is another one of the open sets W_α in which the support of $h_i f$ is contained, then $h_i f$ has support in the intersection $W_{\alpha(i)} \cap W_{\alpha'(i)}$, and our assumption concerning our functionals λ_α shows that the corresponding term in the sum does not depend on the choice of index $\alpha(i)$. Next, let $\{g_k\}$ be another partition of unity over K subordinated to some covering of K by a finite number of the open sets W_α. Then for each i,

$$h_i f = \sum_k g_k h_i f,$$

whence

$$\sum_i \lambda_{\alpha(i)}(h_i f) = \sum_i \sum_k \lambda_{\alpha(i)}(g_k h_i f).$$

If the support of $g_k h_i f$ is in some W_α, then the value $\lambda_\alpha(g_k h_i f)$ is independent of the choice of index α. The expression on the right is then symmetric with respect to our two partitions of unity, whence our theorem follows.

Theorem 4.3. *Let $\dim X = n$ and let ω be an n-form on X of class C^0, that is continuous. Then there exists a unique positive functional λ on $C_c(X)$ having the following property. If (U, φ) is a chart and*

$$\omega(x) = f(x)\, dx_1 \wedge \cdots \wedge dx_n$$

is the local representation of ω in this chart, then for any $g \in C_c(X)$ with *support in U, we have*

$$\lambda g = \int_{\varphi U} g_\varphi(x)|f(x)|\,dx, \tag{1}$$

where g_φ represents g in the chart [*i.e.* $g_\varphi(x) = g(\varphi^{-1}(x))$], *and dx is Lebesgue measure.*

Proof. The integral in (1) defines a positive functional on $C_c(U)$. The change of variables formula shows that if (U, φ) and (V, ψ) are two charts, and if g has support in $U \cap V$, then the value of the functional is independent of the choice of charts. Thus we get a positive functional by the general localization lemma for functionals.

The positive measure corresponding to the functional in Theorem 4.3 will be called the **measure associated with** $|\omega|$, and can be denoted by $\mu_{|\omega|}$.

Theorem 4.3 does not need any orientability assumption. With such an assumption, we have a similar theorem, obtained without taking the absolute value.

Theorem 4.4. *Let* $\dim X = n$ *and assume that X is oriented. Let ω be an n-form on X of class C^0. Then there exists a unique functional λ on $C_c(X)$ having the following property. If (U, φ) is an oriented chart and*

$$\omega(x) = f(x)\,dx_1, \wedge \cdots \wedge dx_n$$

is the local representation of ω in this chart, then for any $g \in C_c(X)$ with support in U, we have

$$\lambda g = \int_{\varphi U} g_\varphi(x) f(x)\,dx,$$

where g_φ represents g in the chart, and dx is Lebesgue measure.

Proof. Since the Jacobian determinant of transition maps belonging to oriented charts is positive, we see that Theorem 4.4 follows like Theorem 4.3 from the change of variables formula (in which the absolute value sign now becomes unnecessary) and the existence of partitions of unity.

If λ is the functional of Theorem 4.4, we shall call it the **functional associated with** ω. For any function $g \in C_c(X)$, we define

$$\int_X g\omega = \lambda g.$$

If in particular ω has compact support, we can also proceed directly as follows. Let $\{\alpha_i\}$ be a partition of unity over X such that each α_i has compact support. We define

$$\int_X \omega = \sum_i \int_X \alpha_i \omega,$$

all but a finite number of terms in this sum being equal to 0. As usual, it is immediately verified that this sum is in fact independent of the choice of partition of unity, and in fact, we could just as well use only a partition of unity over the support of ω. Alternatively, if α is a function in $C_c(X)$ which is equal to 1 on the support of ω, then we could also define

$$\int_X \omega = \int_X \alpha\omega.$$

It is clear that these two possible definitions are equivalent. In particular, we obtain the following variation of Theorem 4.4.

Theorem 4.5. *Let X be an oriented manifold of dimension n. Let $\mathscr{A}^n_c(X)$ be the* **R***-space of differential forms with compact support. There exists a unique linear map*

$$\omega \mapsto \int_X \omega \quad \text{of} \quad \mathscr{A}^n_c(X) \to \mathbf{R}$$

such that, if ω has support in an oriented chart U with coordinates $x_1, \dots, x_n$ and $\omega(x) = f(x)\,dx_1 \wedge \cdots \wedge dx_n$ in this chart, then

$$\int_X \omega = \int_U f(x)\,dx_1 \cdots dx_n.$$

Let X be an oriented manifold. By **a volume form** Ω we mean a form such that in every oriented chart, the form can be written as

$$\Omega(x) = f(x)\,dx_1 \wedge \cdots \wedge dx_n$$

with $f(x) > 0$ for all x. In the next section, we shall see how to get a volume form from a Riemannian metric. Here, we shall consider the non-oriented case to get the notion of density.

Even when a manifold is not orientable, one may often reduce certain questions to the orientable case, because of the following result. We assume that readers are acquainted with basic facts about coverings.

Proposition 4.6. *Let X be a connected C^1 manifold. If X is not orientable, then there exists a covering $X' \to X$ of degree 2 such that X' is orientable.*

Sketch of Proof. Suppose first that X is simply connected. Let $x \in X$. Fix a chart (U_0, φ_0) at x such that the image of the chart is an open ball in euclidean space. Let y be any point of X, and let $\alpha: [a, b] \to X$ be a piecewise C^1 path from x to y. We select a sufficiently fine partition

$$[a, b] = [t_0, t_1, \ldots, t_n],$$

and open sets U_i containing $\alpha([t_i, t_{i+1}])$, such that U_i has an isomorphism φ_i onto an open ball in euclidean space, and such that the charts φ_i and φ_{i+1} have the same orientation. It is easy to verify that if two paths are homotopic, then the charts which we obtain at y by "continuation" as above along the two paths are orientation equivalent. This is done first for paths which are "close together," and then extended to homotopic paths, according to the standard technique which already appears in analytic continuation. Thus fixing one orientation in the neighborhood of a give point determines an orientation on all of X when X is simply connected.

Now suppose X not simply connected, and let $\tilde{X}$ be its universal covering space. Let Γ be the fundamental group. Then the subgroup of elements $\gamma \in \Gamma$ which preserve an orientation of $\tilde{X}$ is of index 2, and the covering corresponding to this subgroups has degree 2 over X and can be given an orientation by using charts which lift to oriented charts in the universal covering space. This concludes the proof.

Densities

The rest of this section will not be used, especially not for Stokes' theorem in the next chapter. However, Theorem 4.3 for the non-orientable case is important for other applications, and we make further comments about this other context.

Let s be a real number. Let $\mathbf{E}$ be a finite dimensional vector space over $\mathbf{R}$, of dimension n. We denote by E^* the set of non-zero elements of E, and by $\bigwedge^n E^*$ the set of non-zero elements of $\bigwedge^n E$. By an s-**density** on E we mean a function

$$\delta: \bigwedge^n E^* \to \mathbf{R}_{\geqq 0} \quad \text{such that} \quad \delta(cw) = |c|^s \delta(w)$$

for all $c \neq 0$ in $\mathbf{R}$ and $w \in \bigwedge^n E^*$. Equivalently, we could say that there exists an n-form $\omega \in L^n_a(\mathbf{E}, \mathbf{R})$ such that for $v_i \in \mathbf{E}$ we have

$$\delta(v_1 \wedge \cdots \wedge v_n) = |\omega(v_1, \ldots, v_n)|^s.$$

We let $\text{den}^s(\mathbf{E})$ denote the set of densities of $\mathbf{E}$. An element of $\text{den}^s(\mathbf{E})$ amounts to picking a basis of $\bigwedge^n \mathbf{E}$, up to a factor ± 1, and assigning a number to this basis.

Let U be an open subset of $\mathbf{E}$. By a C^p **density** on U we mean a C^p morphism $\delta: U \to \text{den}^s(\mathbf{E})$. Note that $\text{den}^s(\mathbf{E})$ is an open half line, so a density on U amounts to selecting a differential form of class C^p on a neighborhood of each point of U, such that the absolute values of these forms coincide on intersections of these neighborhoods.

Let $f: U \to V$ be a C^p isomorphism. Then f induces a map on densities, by the change of variable formula on forms with the Jacobian determinant, and then taking absolute values to the s-power. Thus we may form the bundle (not vector bundle) of densities, with charts

$$U \times \text{den}^s(\mathbf{E})$$

over U, and density-bundle morphisms just as we did with differential forms. For example let $E = \mathbf{R}^n$, with coordinates $x_1, \ldots, x_n$. Then

$$|dx_1 \wedge \cdots \wedge dx_n| = dx_1 \cdots dx_n$$

defines a 1-density, and $|dx_1 \wedge \cdots \wedge dx_n|^s$ defines an s-density, denoted by $|dx|^s$.

Observe that s-densities form a cone, i.e. if δ_1, δ_2 are s-densities on a manifold X, and $a_1, a_2 \in \mathbf{R}^+$ (the set of positive real numbers), then $a_1\delta_1 + a_2\delta_2$ is also an s-density. In particular, continuing to assume that X admits continuous partitions of unity, we can reformulate and prove Theorem 4.3 for densities. Indeed, the differential form in Theorem 4.3 need not be globally defined, because one needs only its absolute value to define the integral. Thus with the language of densities, Theorem 4.3 reads as follows.

Theorem 4.7. *Let δ be a C^0 density on X, i.e. a continuous density. Then there is a positive functional λ on $C_c(X)$ having the following property. If U is a chart and δ is represented by the density $|f(x)\, dx_1 \wedge \cdots \wedge dx_n|$ on this chart, then for any function $\varphi \in C_c(U)$ we have*

$$\lambda(\varphi) = \int_U \varphi_U(x) |f(x)|\, dx_1 \cdots dx_n,$$

where $dx_1 \cdots dx_n$ is the usual symbol for ordinary integration on $\mathbf{R}^n$, and φ_U is the representation of φ in the chart.

Examples. We have already given the example of integration with respect to $|dx_1 \wedge \cdots \wedge dx_n|$ in euclidean space. Here is a less trivial example. Let X be a Riemannian manifold of finite dimension n, with

Riemannian metric g. Locally in a chart U, we view g as a morphism

$$g\colon\ U \to L(\mathbf{E}, \mathbf{E}),$$

with $\mathbf{E}$ having a fixed positive definite scalar product. With respect to an orthonormal basis, we have a linear metric isomorphism $\mathbf{E} \approx \mathbf{R}^n$, and $g(x)$ at each point x can be represented by a matrix $(g_{ij}(x))$. If we put

$$\delta^2(x) = |\det g_{ij}(x)|\, dx_1 \cdots dx_n$$

then δ^2 defines a density, called the **Riemannian square density**; and

$$\delta(x) = |\det g_{ij}(x)|^{1/2}\, dx_1 \cdots dx_n$$

defines the **Riemannian density**.

Remark. Locally, a manifold is always orientable. Hence a formula or result which is local, and is proved in the orientable case, also applies to densities, sometimes by inserting an absolute value sign. For example, Proposition 1.2 of Chapter XV applies after inserting absolute value signs, but Proposition 2.1 of Chapter XV applies as stated for the Riemannian density instead of the Riemannian volume.

Integration on a submersion

Let $\pi\colon X \to Z$ be a submersion. In Chapter XV, §6 we discussed the relationship between volume forms on X, Z and the fibers, and we use the same notation as before. We now derive the consequences of this relation for integration.

Theorem 4.8. *Let $\pi\colon X \to Z$ be a submersion. Let Ω be a volume form on X and ω a volume form on Z. Let $\Omega = \eta \otimes \omega$. Let $\tilde{\eta}$ be a form on X, of the same degree as η, restricting to η on the fibers. Then for all $f \in C_c(X)$, we have*

$$\int_X f\Omega = \int_Z \left(\int_{Y_z} f(y)\tilde{\eta}_z(y) \right) \omega(z).$$

Proof. The proposition is local, since by using a partition of unity, we are reduced to the case when the support of f is in a given neighborhood of a point. Then the submersion is represented in a chart as a projection $U \times W \to W$, where U, W are open in $\mathbf{R}^p$ and $\mathbf{R}^q$ respectively, U being a chart on X and W a chart on Z. On $U \times W$ we have the coordinate representation

$$(f\Omega)_{U\times W}(y, z) = f(y, z)\varphi(y, z)\, dy_1 \wedge \cdots \wedge dy_p \wedge \rho(z)\, dz_1 \wedge \cdots \wedge dz_q,$$

where $y_1, \ldots, y_p$ are the coordinate functions on U, $z_1, \ldots, z_q$ are the coordinate functions on W,

$$\tilde{\eta}_z(y) = \varphi(y, z)\, dy_1 \wedge \cdots \wedge dy_p \quad \text{and} \quad \omega(z) = \rho(z)\, dz_1 \wedge \cdots \wedge dz_q.$$

Then the proposition merely amounts to Fubini's theorem, which concludes the proof.

XVI, §5. HOMOGENEOUS SPACES

For the convenience of the reader, we reproduce some results on locally compact groups, corresponding to the results on volume forms in Chapter XV, §7. When dealing with manifolds, the results of §7 provide a more natural setting, but it is worthwhile to develop the results dealing just with Haar measure on locally compact groups, so here goes. See also [La 93], Chapter XII, Theorem 4.3.

Let G be a locally compact group. Let Δ_G be the so-called **modular function** on G, relating right and left Haar measure. Thus by definition, for $f \in C_c(G)$ and left Haar measure dx on G, we have

MOD 1. $$\int_G f(xy)\, dx = \Delta(y) \int_G f(x)\, dx.$$

Then $\Delta\colon G \to \mathbf{R}^+$ is a continuous homomorphism.

For $f \in C_c(G)$, we have

MOD 2. $$\int_G f(x^{-1})\Delta(x)\, dx = \int_G f(x)\, dx.$$

Proof. First we show that the functional on $C_c(G)$ defined by the left side of the equation is left G-invariant. Applying it to the left translate of f by an element $a \in G$, and putting $g(u) = f(u^{-1})\Delta(u)$, we get

$$\begin{aligned}\int_G (\tau_a f)(x^{-1})\Delta(x)\, dx &= \int f(a^{-1}x^{-1})\Delta(x)\, dx \\ &= \int f((xa)^{-1})\Delta(xa)\Delta(a)^{-1}\, dx \\ &= \int g(xa)\Delta(a)^{-1}\, dx \\ &= \int g(x)\, dx,\end{aligned}$$

thus proving the left G-invariance. Hence there exists a constant c such that for all f,

$$\int f(x^{-1})\Delta(x)\,dx = c\int f(x)\,dx.$$

To see that $c = 1$, let $f(x) = \varphi(x)\Delta(x)$, where φ is an even function $\geqq 0$, $\varphi(x) = \varphi(x^{-1})$ for all x, φ has support close to the unit element e, and

$$\int_G \varphi(x)\,dx = 1.$$

Since Δ is continuous, it follows that $\Delta(x)$ is close to 1 for x near e. We let the support of φ come closer and closer to e. From the formula

$$1 = \int \varphi(x)\,dx = c\int \varphi(x)\Delta(x)\,dx,$$

letting the support of φ tend to e, we conclude that $c = 1$, thus proving **MOD 2**.

The functional $\Delta(x)\,dx$ is right invariant, and in fact, for all $y \in G$, we have

MOD 3. $$\int f(xy)\Delta(x)\,dx = \int f(x^{-1})\,dx = \int f(x)\Delta(x)\,dx.$$

Proof. Let $g(u) = f(u^{-1})$ and $h(u) = g(y^{-1}u)$. Then using **MOD 2**, we get

$$\int f(xy)\Delta(x)\,dx = \int g(y^{-1}x^{-1})\Delta(x)\,dx = \int h(x^{-1})\Delta(x)\,dx$$
$$= \int h(x)\,dx = \int g(y^{-1}x)\,dx = \int g(x)\,dx = \int f(x^{-1})\,dx,$$

which proves the first equality. The second is only a special case with $y = e$.

Let H be a closed subgroup of G, with corresponding function Δ_H.

Theorem 5.1. *Suppose that $\Delta_G = \Delta_H$ on H. Then there exists a unique G-invariant positive functional on $C_c(G/H)$, so a unique G-invariant positive σ-regular measure on G/H. "Uniqueness" is up to a positive constant multiple.*

Proof. For $f \in C_c(G)$, we define

$$f^H(x) = \int_H f(uh)\,dh.$$

It is standard that $f \mapsto f^H$ maps $C_c(G)$ onto $C_c(G/H)$, cf. for instance [La 93], Chapter XII, Theorem 4.1. The map $f \mapsto f^H$ also preserves positivity. Given $f \in C_c(G/H)$, to define its invariant integral on G/H, we let $f^\sharp \in C_c(G)$ be such that $(f^\sharp)^H = f$. We want to define

(1) $$\int_{G/H} f(\dot{x})\,d\dot{x} = \int_G f^\sharp(x)\,dx.$$

The problem is to show that this definition is independent of the choice of $f^\sharp$. This is settled by the following lemma.

Lemma 5.2. *Let $f \in C_c(G)$. If $f^H = 0$, that is*

$$\int_H f(xh)\,dh = 0$$

for all $x \in G$, then

$$\int_G f(x)\,dx = 0.$$

Proof. For all $\varphi \in C_c(G)$, we have:

$$\begin{aligned}\int_G \varphi(x)\left(\int_H f(xh)\,dh\right)dx &= \int_G\left(\int_H \varphi(x)f(xh)\,dh\right)dx\\ &= \int_H\left(\int_G \varphi(x)f(xh)\,dx\right)dh\\ &= \int_H\left(\int_G \Delta_G(h)\varphi(xh^{-1})f(x)\,dx\right)dh\\ &= \int_G\left(\int_H \Delta_H(h)\varphi(xh^{-1})\,dh\right)f(x)\,dx\\ &= \int_G\left(\int_H \varphi(xh)\,dh\right)f(x)\,dx.\end{aligned}$$

By the surjectivity $C_c(G) \to C_c(G/H)$ we can find φ such that $\varphi^H = 1$ on the support of f. Since by assumption the left side of the equation is 0, this concludes the proof of the lemma.

Now knowing that (1) is well defined, it is immediate that the functional

$$f \mapsto \int_{G/H} f(\dot{x})\, d\dot{x}$$

is G-invariant, and positive, so we have proved the existence of the desired functional. Uniqueness is proved by reducing it to uniqueness for Haar measure, since we have the **repeated integral formula on** G, H, G/H, namely

$$\int_G f(x)\, dx = \int_{G/H} \left(\int_H f(xh)\, dh \right) d\dot{x}.$$

This concludes the proof of the theorem.

A G-invariant measure on G/H will be called a **Haar measure**, as for groups.

A group G is called **unimodular** if $\Delta_G = 1$, so right and left Haar measures are equal. Suppose this is the case. In particular, if $a \in G$ and $\mathbf{c}_a$ is conjugation,

$$\mathbf{c}_a(x) = axa^{-1},$$

then $\mathbf{c}_a$ preserves a given Haar measure, i.e. $\mathbf{c}_a$ is a measure preserving group isomorphism. Suppose K is a compact subgroup. Then $\Delta_G = \Delta_K$ on K, since both functions provide continuous homomorphisms of K into $\mathbf{R}^+$, so both functions are trivial on K. Thus we always have a G-invariant measure on the coset space G/K. For $a \in G$ we have an isomorphism of G-homogeneous spaces

$$\mathbf{c}_a\colon\ G/K \to \mathbf{c}_a(G)/\mathbf{c}_a(K) = G/\mathbf{c}_a(K).$$

Fix the Haar measure on G. Fix the Haar measures on K and $\mathbf{c}_a(K)$ to have total measure 1, which is possible since K is compact. Then these measures determine uniquely the Haar measure on G/K. Since $\mathbf{c}_a$ preserves the fixed Haar measure on G, it follows that it also preserves the Haar measure on the homogeneous space G/K, to satisfy the repeated integral formula on G, K, G/K.

CHAPTER XVII

Stokes' Theorem

Throughout the chapter, all manifolds are assumed finite dimensional. They may have a boundary.

XVII, §1. STOKES' THEOREM FOR A RECTANGULAR SIMPLEX

If X is a manifold and Y a submanifold, then any differential form on X induces a form on Y. We can view this as a very special case of the inverse image of a form, under the embedding (injection) map

$$\mathrm{id}\colon\ Y \to X.$$

In particular, if Y has dimension $n-1$, and if $(x_1, \dots, x_n)$ is a system of coordinates for X at some point of Y such that the points of Y correspond to those coordinates satisfying $x_j = c$ for some fixed number c, and index j, and if the form on X is given in terms of these coordinates by

$$\omega(x) = f(x_1, \dots, x_n)\, dx_1 \wedge \cdots \wedge dx_n,$$

then the restriction of ω to Y (or the form induced on Y) has the representation

$$f(x_1, \dots, c, \dots, x_n)\, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n.$$

We should denote this induced form by ω_Y, although occasionally we omit the subscript Y. We shall use such an induced form especially when Y is the boundary of a manifold X.

Let

$$R = [a_1, b_1] \times \cdots \times [a_n, b_n]$$

be a rectangle in n-space, that is a product of n closed intervals. The set theoretic boundary of R consists of the union over all $i = 1, \ldots, n$ of the pieces

$$R_i^0 = [a_1, b_1] \times \cdots \times \{a_i\} \times \cdots \times \{a_n, b_n\},$$

$$R_i^1 = [a_1, b_1] \times \cdots \times \{b_i\} \times \cdots \times [a_n, b_n].$$

If

$$\omega(x_1, \ldots, x_n) = f(x_1, \ldots, x_n)\, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n$$

is an $(n-1)$-form, and the roof over anything means that this thing is to be omitted, then we define

$$\int_{R^0} \omega = \int_{a_i}^{b_1} \cdots \widehat{\int_{a_i}^{b_i}} \cdots \int_{a_n}^{b_n} f(x_1, \ldots, a_i, \ldots, x_n)\, dx_1 \cdots \widehat{dx_j} \cdots dx_n,$$

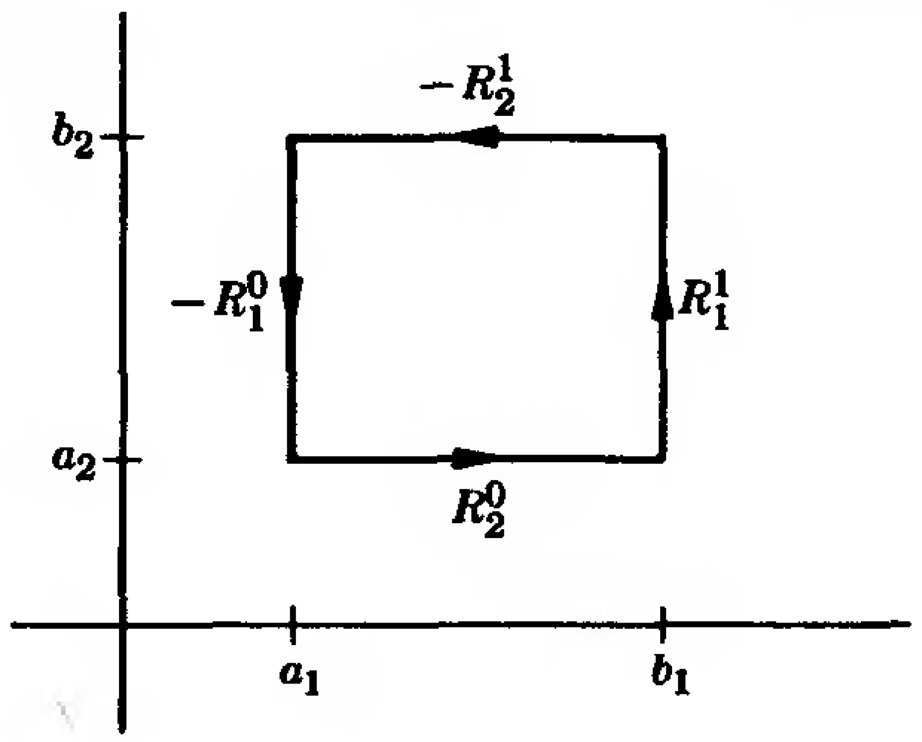
if $i = j$, and 0 otherwise. And similarly for the integral over R_i^1. We define the integral over the oriented boundary to be

$$\int_{\partial^0 R} = \sum_{i=1}^{n} (-1)^i \left[\int_{R_i^0} - \int_{R_i^1} \right].$$

Stokes' Theorem for Rectangles. *Let R be a rectangle in an open set U in n-space. Let ω be an $(n-1)$-form on U. Then*

$$\int_R d\omega = \int_{\partial^0 R} \omega.$$

Proof. In two dimensions, the picture looks like this:



It suffices to prove the assertion when ω is a decomposable form, say

$$\omega(x) = f(x_1, \ldots, x_n)\, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n.$$

We then evaluate the integral over the boundary of R. If $i \neq j$, then it is clear that

$$\int_{R_i^0} \omega = 0 = \int_{R_i^1} \omega,$$

so that

$$\int_{\partial^0 R} \omega =$$

$$(-1)^j \int_{a_1}^{b_1} \cdots \widehat{\int_{a_j}^{b_j}} \cdots \int_{a_n}^{b_n} [f(x_1, \ldots, a_j, \ldots, x_n) - f(x_1, \ldots, b_j, \ldots, x_n)]\, dx_1 \cdots \widehat{dx_j} \cdots dx_n.$$

On the other hand, from the definitions we find that

$$d\omega(x) = \left(\frac{\partial f}{dx_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n \right) \wedge dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n$$

$$= (-1)^{j-1} \frac{\partial f}{\partial x_j} dx_1 \wedge \cdots \wedge dx_n.$$

(The $(-1)^{j-1}$ comes from interchanging dx_j with $dx_1, \ldots, dx_{j-1}$. All other terms disappear by the alternating rule.)

Integrating $d\omega$ over R, we may use repeated integration and integrate $\partial f/\partial x_j$ with respect to x_j first. Then the fundamental theorem of calculus for one variable yields

$$\int_{a_j}^{b_j} \frac{\partial f}{\partial x_j} dx_j = f(x_1, \ldots, b_j, \ldots, x_n) - f(x_1, \ldots, a_j, \ldots, x_n).$$

We then integrate with respect to the other variables, and multiply by $(-1)^{j-1}$. This yields precisely the value found for the integral of ω over the oriented boundary $\partial^0 R$, and proves the theorem.

Remark. Stokes' theorem for a rectangle extends at once to a version in which we parametrize a subset of some space by a rectangle. Indeed, if $\sigma\colon R \to V$ is a C^1 map of a rectangle of dimension n into an open set V in $\mathbf{R}^N$, and if ω is an $(n-1)$-form in V, we may define

$$\int_\sigma d\omega = \int_R \sigma^* d\omega.$$

One can define

$$\int_{\partial\sigma} \omega = \int_{\partial^0 R} \sigma^*\omega,$$

and then we have a formula

$$\boxed{\int_{\sigma} d\omega = \int_{\partial\sigma} \omega,}$$

In the next section, we prove a somewhat less formal result.

XVII, §2. STOKES' THEOREM ON A MANIFOLD

Theorem 2.1. *Let X be an oriented manifold of class C^2, dimension n, and let ω be an $(n-1)$-form on X, of class C^1. Assume that ω has compact support. Then*

$$\int_X d\omega = \int_{\partial X} \omega.$$

Proof. Let $\{\alpha_i\}_{i\in I}$ be a partition of unity, of class C^2. Then

$$\sum_{i\in I} \alpha_i\,\omega = \omega,$$

and this sum has only a finite number of non-zero terms since the support of ω is compact. Using the additivity of the operation d, and that of the integral, we find

$$\int_X d\omega = \sum_{i\in I} \int_X d(\alpha_i\,\omega).$$

Suppose that α_i has compact support in some open set V_i of X and that we can prove

$$\int_{V_i} d(\alpha_i\,\omega) = \int_{V_i\cap\partial X} \alpha_i\,\omega,$$

in other words we can prove Stokes' theorem locally in V_i. We can write

$$\int_{V_i\cap\partial X} \alpha_i\,\omega = \int_{\partial X} \alpha_i\,\omega,$$

and similarly

$$\int_{V_i} d(\alpha_i\,\omega) = \int_X d(\alpha_i\,\omega).$$

Using the additivity of the integral once more, we get

$$\int_X d\omega = \sum_{i\in I} \int_X d(\alpha_i\,\omega) = \sum_{i\in I} \int_{\partial X} \alpha_i\,\omega = \int_{\partial X} \omega,$$

which yields Stokes' theorem on the whole manifold. Thus our argument with partitions of unity reduces Stokes' theorem to the local case, namely it suffices to prove that for each point of X these exists an open neighborhood V such that if ω has compact support in V, then Stokes' theorem holds with X replaced by V. We now do this.

If the point is not a boundary point, we take an oriented chart (U, φ) at the point, containing an open neighborhood V of the point, satisfying the following conditions: φU is an open ball, and φV is the interior of a rectangle, whose closure is contained in φU. If ω has compact support in V, then its local representation in φU has compact support in φV. Applying Stokes' theorem for rectangles as proved in the preceding section, we find that the two integrals occurring in Stokes' formula are equal to 0 in this case (the integral over an empty boundary being equal to 0 by convention).

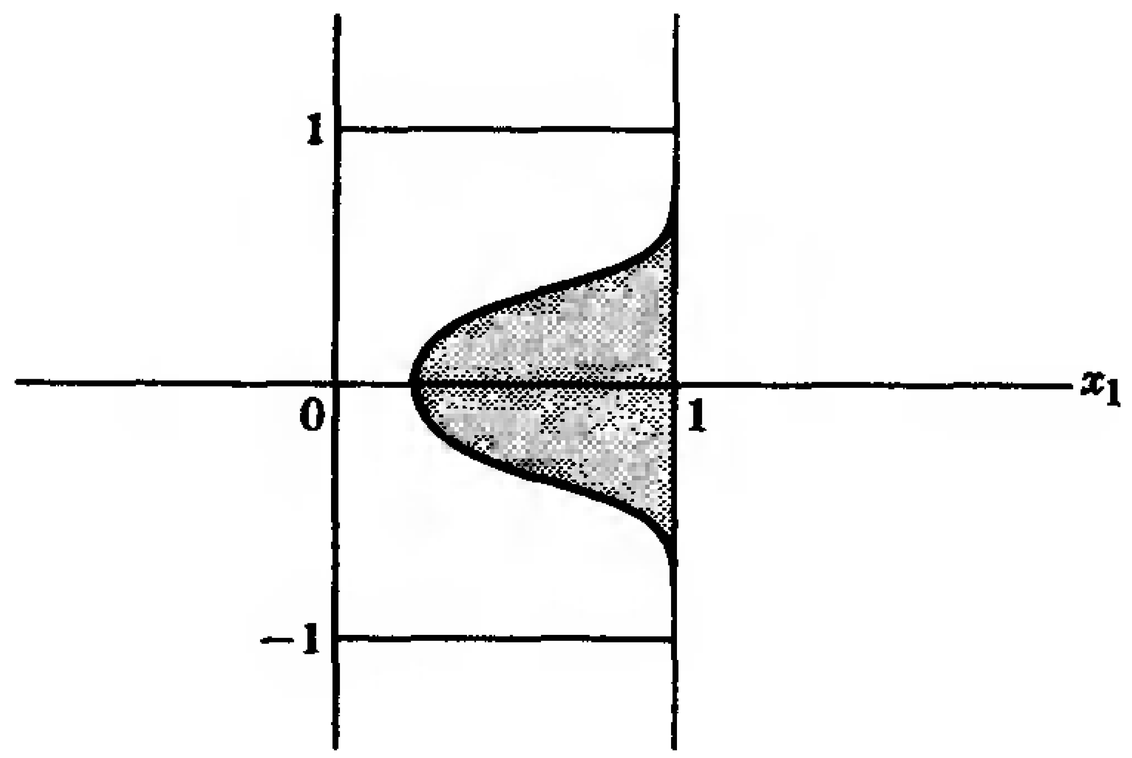
Now suppose that we deal with a boundary point. We take an oriented chart (U, φ) at the point, having the following properties. First, φU is described by the following inequalities in terms of local coordinates $(x_1, \dots, x_n)$:

$$-2 < x_1 \leqq 1 \quad \text{and} \quad -2 < x_j < 2 \quad \text{for } j = 2, \dots, n.$$

Next, the given point has coordinates $(1, 0, \dots, 0)$, and that part of U on the boundary of X, namely $U \cap \partial X$, is given in terms of these coordinates by the equation $x_1 = 1$. We then let V consist of those points whose local coordinates satisfy

$$0 < x_1 \leqq 1 \quad \text{and} \quad -1 < x_j < 1 \quad \text{for } j = 2, \dots, n.$$

If ω has compact support in V, then ω is equal to 0 on the boundary of the rectangle R equal to the closure of φV, except on the face given by $x_1 = 1$, which defines that part of the rectangle corresponding to $\partial X \cap V$. Thus the support of ω looks like the shaded portion of the following picture.



In the sum giving the integral over the boundary of a rectangle as in the previous section, only one term will give a non-zero contribution, corresponding to $i = 1$, which is

$$(-1)\left[\int_{R_1^0} \omega - \int_{R_1^1} \omega\right].$$

Furthermore, the integral over R_1^0 will also be 0, and in the contribution of the integral over R_1^1, the two minus signs will cancel, and yield the integral of ω over the part of the boundary lying in V, because our charts are so chosen that $(x_2, \ldots, x_n)$ is an oriented system of coordinates for the boundary. Thus we find

$$\int_V d\omega = \int_{V \cap \partial X} \omega,$$

which proves Stokes' theorem locally in this case, and concludes the proof of Theorem 2.7.

Corollary 2.2. *Suppose X is an oriented manifold without boundary, and ω has compact support. Then*

$$\int_X d\omega = 0.$$

For any number of reasons, some of which we consider in the next section, it is useful to formulate conditions under which Stokes' theorem holds even when the form ω does not have compact support. We shall say that ω has **almost compact support** if there exists a decreasing sequence of open sets $\{U_k\}$ in X such that the intersection

$$\bigcap_{k=1}^{\infty} U_k$$

is empty, and a sequence of C^1 functions $\{g_k\}$, having the following properties:

AC 1. *We have $0 \leqq g_k \leqq 1$, $g_k = 1$ outside U_k, and $g_k\omega$ has compact support.*

AC 2. *If μ_k is the measure associated with $|dg_k \wedge \omega|$ on X, then*

$$\lim_{k\to\infty} \mu_k(\bar{U}_k) = 0.$$

We then have the following application of Stokes' theorem.

Corollary 2.3. *Let X be a C^2 oriented manifold, of dimension n, and let ω be an $(n-1)$-form on X, of class C^1. Assume that ω has almost compact support, and that the measures associated with $|d\omega|$ on X and $|\omega|$ on ∂X are finite. Then*

$$\int_X d\omega = \int_{\partial X} \omega.$$

Proof. By our standard form of Stokes' theorem we have

$$\int_{\partial X} g_k\omega = \int_X d(g_k\omega) = \int_X dg_k \wedge \omega + \int_X g_k\, d\omega.$$

We estimate the left-hand side by

$$\left|\int_{\partial X} \omega - \int_{\partial X} g_k\,\omega\right| = \left|\int_{\partial X} (1 - g_k)\,\omega\right| \leqq \mu_{|\omega|}(U_k \cap \partial X).$$

Since the intersection of the sets U_k is empty, it follows for a purely measure-theoretic reason that

$$\lim_{k\to\infty} \int_{\partial X} g_k\,\omega = \int_{\partial X} \omega.$$

Similarly,

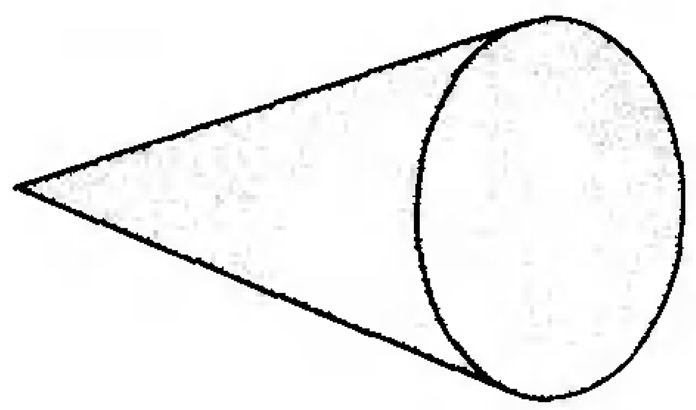
$$\lim_{k\to\infty} \int_X g_k\, d\omega = \int_X d\omega.$$

The integral of $dg_k \wedge \omega$ over X approaches 0 as $k \to \infty$ by assumption, and the fact that $dg_k \wedge \omega$ is equal to 0 on the complement of $\bar{U}_k$ since g_k is constant on this complement. This proves our corollary.

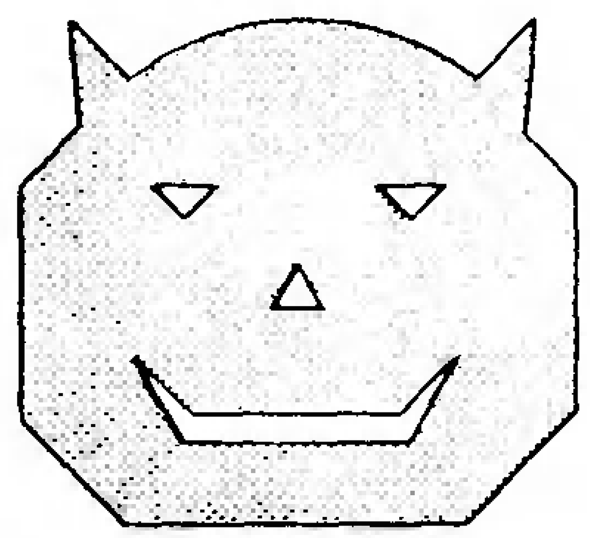
The above proof shows that the second condition **AC 2** is a very natural one to reduce the integral of an arbitrary form to that of a form with compact support. In the next section, we relate this condition to a question of singularities when the manifold is embedded in some bigger space.

XVII, §3. STOKES' THEOREM WITH SINGULARITIES

If X is a compact manifold, then of course every differential form on X has compact support. However, the version of Stokes' theorem which we have given is useful in contexts when we start with an object which is not a manifold, say as a subset of $\mathbf{R}^n$, but is such that when we remove a portion of it, what remains is a manifold. For instance, consider a cone (say the solid cone) as illustrated in the next picture.



The vertex and the circle surrounding the base disc prevent the cone from being a submanifold of $\mathbf{R}^3$. However, if we delete the vertex and this circle, what remains is a submanifold with boundary embedded in $\mathbf{R}^3$. The boundary consists of the conical shell, and of the base disc (without its surrounding circle). Another example is given by polyhedra, as on the following figure.



The idea is to approximate a given form by a form with compact support, to which we can apply Theorem 2.1, and then take the limit. We shall indicate one possible technique to do this.

The word "boundary" has been used in two senses: The sense of point set topology, and the sense of boundary of a manifold. Up to now, they were used in different contexts so no confusion could arise. We must now make a distinction, and therefore use the word boundary only in its manifold sense. If X is a subset of $\mathbf{R}^N$, we denote its closure by $\overline{X}$ as usual. We call the set-theoretic difference $\overline{X} - X$ the **frontier** of X in $\mathbf{R}^N$, and denote it by $\text{fr}(X)$.

Let X be a submanifold without boundary of $\mathbf{R}^N$, of dimension n. We know that this means that at each point of X there exists a chart for an open neighborhood of this point in $\mathbf{R}^N$ such that the points of X in this chart correspond to a factor in a product. A point P of $\overline{X} - X$ will be called a **regular** frontier point of X if there exists a chart at P in $\mathbf{R}^N$ with local coordinates $(x_1, \ldots, x_N)$ such that P has coordinates $(0, \ldots, 0)$; the points of X are those with coordinates

$$x_{n+1} = \cdots = x_N = 0 \qquad \text{and} \qquad x_n < 0;$$

and the points of the frontier of X which lie in the chart are those with coordinates satisfying

$$x_n = x_{n+1} = \cdots = x_N = 0.$$

The set of all regular frontier points of X will be denoted by ∂X, and will be called the **boundary** of X. We may say that $X \cup \partial X$ is a submanifold of $\mathbf{R}^N$, possibly with boundary.

A point of the frontier of X which is not regular will be called **singular**. It is clear that the set of singular points is closed in $\mathbf{R}^N$. We now formulate a version of Theorem 2.1 when ω does not necessarily have compact support in $X \cup \partial X$. Let S be a subset of $\mathbf{R}^N$. By a **fundamental sequence** of open neighborhoods of S we shall mean a sequence $\{U_k\}$ of open sets containing S such that, if W is an open set containing S, then $U_k \subset W$ for all sufficiently large k.

Let S be the set of singular frontier points of X and let ω be a form defined on an open neighborhood of $\overline{X}$, and having compact support. The intersection of supp ω with $(X \cup \partial X)$ need not be compact, so that we cannot apply Theorem 2.1 as it stands. The idea is to find a fundamental sequence of neighborhods $\{U_k\}$ of S, and a function g_k which is 0 on a neighborhood of S and 1 outside U_k so that $g_k\omega$ differs from ω only inside U_k. We can then apply Theorem 2.1 to $g_k\omega$ and we hope that taking the limit yields Stokes' theorem for ω itself. However, we have

$$\int_X d(g_k\,\omega) = \int_X dg_k \wedge \omega + \int_X g_k\, d\omega.$$

Thus we have an extra term on the right, which should go to 0 as $k \to \infty$ if we wish to apply this method. In view of this, we make the following definition.

Let S be a closed subset of $\mathbf{R}^N$. We shall say that S is **negligible for** X if there exists an open neighborhood U of S in $\mathbf{R}^N$, a fundamental sequence of open neighborhoods $\{U_k\}$ of S in U, with $\overline{U}_k \subset U$, and a sequence of C^1 functions $\{g_k\}$, having the following properties.

NEG 1. *We have* $0 \leqq g_k \leqq 1$. *Also,* $g_k(x) = 0$ *for* x *in some open neighborhood of* S, *and* $g_k(x) = 1$ *for* $x \notin U_k$.

NEG 2. *If* ω *is an* $(n-1)$*-form of class* C^1 *on* U, *and* μ_k *is the measure associated with* $|dg_k \wedge \omega|$ *on* $U \cap X$, *then* μ_k *is finite for large* k, *and*

$$\lim_{k\to\infty} \mu_k(U \cap X) = 0.$$

From our first condition, we see that $g_k\omega$ vanishes on an open neighborhood of S. Since $g_k = 1$ on the complement of $\bar{U}_k$, we have $dg_k = 0$ on this complement, and therefore our second condition implies that the measures induced on X near the singular frontier by $|dg_k \wedge \omega|$ (for $k = 1, 2, \ldots$), are concentrated on shrinking neighborhoods and tend to 0 as $k \to \infty$.

Theorem 3.1 (Stokes' Theorem with Singularities). *Let* X *be an oriented,* C^3 *submanifold without boundary of* $\mathbf{R}^N$. *Let* $\dim X = n$. *Let* ω *be an* $(n-1)$*-form of class* C^1 *on an open neighborhood of* $\bar{X}$ *in* $\mathbf{R}^N$, *and with compact support. Assume that:*

(i) *If* S *is the set of singular points in the frontier of* X, *then* $S \cap \operatorname{supp} \omega$ *is negligible for* X.

(ii) *The measures associated with* $|d\omega|$ *on* X, *and* $|\omega|$ *on* ∂X, *are finite.*

Then

$$\int_X d\omega = \int_{\partial X} \omega.$$

Proof. Let U, $\{U_k\}$, and $\{g_k\}$ satisfy conditions **NEG 1** and **NEG 2**. Then $g_k\omega$ is 0 on an open neighborhood of S, and since ω is assumed to have compact support, one verifies immediately that

$$(\operatorname{supp} g_k \omega) \cap (X \cup \partial X)$$

is compact. Thus Theorem 2.1 is applicable, and we get

$$\int_{\partial X} g_k \omega = \int_X d(g_k \omega) = \int_X dg_k \wedge \omega + \int_X g_k \, d\omega.$$

We have

$$\left|\int_{\partial X} \omega - \int_{\partial X} g_k \omega\right| \leqq \left|\int_{\partial X} (1 - g_k)\omega\right|$$

$$\leqq \int_{U_k \cap \partial X} 1 \, d\mu_{|\omega|} = \mu_{|\omega|}(U_k \cap \partial X).$$

Since the intersection of all sets $U_k \cap \partial X$ is empty, it follows from purely measure-theoretic reasons that the limit of the right-hand side is 0 as $k \to \infty$. Thus

$$\lim_{k\to\infty} \int_{\partial X} g_k \omega = \int_{\partial X} \omega.$$

For similar reasons, we have

$$\lim_{k\to\infty} \int_X g_k \, d\omega = \int_X d\omega.$$

Our second assumption **NEG 2** guarantees that the integral of $dg_k \wedge \omega$ over X approaches 0. This proves our theorem.

Criterion 1. *Let* S, T *be compact negligible sets for a submanifold* X *of* $\mathbf{R}^N$ *(assuming* X *without boundary). Then the union* $S \cup T$ *is negligible for* X.

Proof. Let U, $\{U_k\}$, $\{g_k\}$ and V, $\{V_k\}$, $\{h_k\}$ be triples associated with S and T respectively as in condition **NEG 1** and **NEG 2** (with V replacing U and h replacing g when T replaces S). Let

$$W = U \cup V, \qquad W_k = U_k \cup V_k, \qquad \text{and} \qquad f_k = g_k h_k.$$

Then the open sets $\{W_k\}$ form a fundamental sequence of open neighborhoods of $S \cup T$ in W, and **NEG 1** is trivially satisfied. As for **NEG 2**, we have

$$d(g_k h_k) \wedge \omega = h_k \, dg_k \wedge \omega + g_k \, dh_k \wedge \omega,$$

so that **NEG 2** is also trivially satisfied, thus proving our criterion.

Criterion 2. *Let* X *be an open set, and let* S *be a compact subset in* $\mathbf{R}^n$. *Assume that there exists a closed rectangle* R *of dimension* $m \leqq n-2$ *and a* C^1 *map* $\sigma\colon R \to \mathbf{R}^n$ *such that* $S = \sigma(R)$. *Then* S *is negligible for* X.

Before giving the proof, we make a couple of simple remarks. First, we could always take $m = n - 2$, since any parametrization by a rectangle of dimension $< n - 2$ can be extended to a parametrization by a rectangle of dimension $n - 2$ simply by projecting away coordinates. Second, by our first criterion, we see that a finite union of sets as described above, that is parametrized smoothly by rectangles of codimension $\geqq 2$, are negligible. Third, our Criterion 2, combined with the first criterion, shows that negligibility in this case is local, that is we can subdivide a rectangle into small pieces.

We now prove Criterion 2. Composing σ with a suitable linear map, we may assume that R is a unit cube. We cut up each side of the cube into k equal segments and thus get k^m small cubes. Since the derivative of σ is bounded on a compact set, the image of each small cube is contained in an n-cube in $\mathbf{R}^N$ of radius $\leqq C/k$ (by the mean value theorem), whose n-dimensional volume is $\leqq (2C)^n/k^n$. Thus we can cover the image by small cubes such that the sum of their n-dimensional volumes is

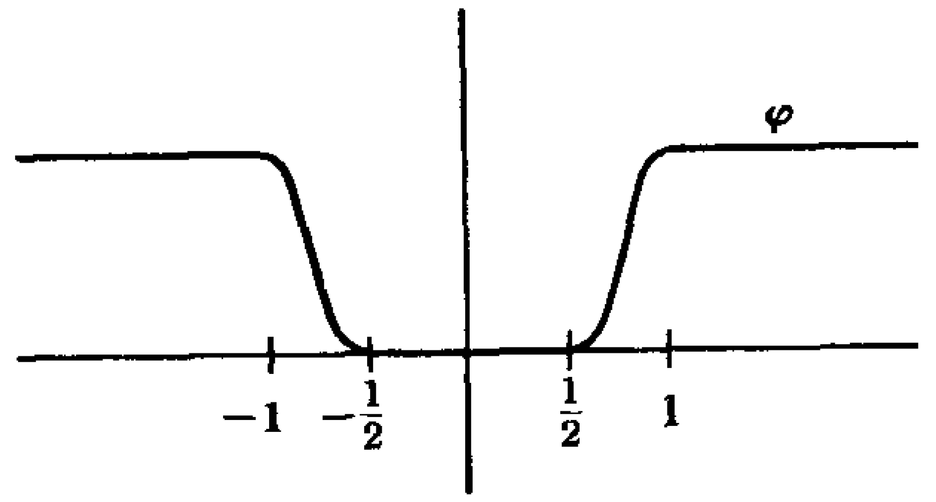
$$\leqq (2C)^n/k^{n-m} \leqq (2C)^n/k^2.$$

Lemma 3.2. *Let S be a compact subset of $\mathbf{R}^n$. Let U_k be the open set of points x such that $d(x, S) < 2/k$. There exists a C^∞ function g_k on $\mathbf{R}^N$ which is equal to 0 in some open neighborhood of S, equal to 1 outside U_k, $0 \leqq g_k \leqq 1$, and such that all partial derivatives of g_k are bounded by $C_1 k$, where C_1 is a constant depending only on n.*

Proof. Let φ be a C^∞ function such that $0 \leqq \varphi \leqq 1$, and

$$\varphi(x) = 0 \quad \text{if} \quad 0 \leqq \|x\| \leqq \tfrac{1}{2},$$
$$\varphi(x) = 1 \quad \text{if} \quad 1 \leqq \|x\|.$$

We use $\| \ \|$ for the sup norm in $\mathbf{R}^n$. The graph of φ looks like this:



For each positive integer k, let $\varphi_k(x) = \varphi(kx)$. Then each partial derivative $D_i\varphi_k$ satisfies the bound

$$\|D_i\varphi_k\| \leqq k\|D_i\varphi\|,$$

which is thus bounded by a constant times k. Let L denote the lattice of integral points in $\mathbf{R}^n$. For each $l \in L$, we consider the function

$$x \mapsto \varphi_k\left(x - \frac{l}{2k}\right).$$

This function has the same shape as φ_k but is translated to the point $l/2k$. Consider the product

$$g_k(x) = \prod \varphi_k\left(x - \frac{l}{2k}\right)$$

taken over all $l \in L$ such that $d(l/2k, S) \leqq 1/k$. If x is a point of $\mathbf{R}^n$ such that $d(x, S) < 1/4k$, then we pick an l such that

$$d(x, l/2k) \leqq 1/2k.$$

For this l we have $d(l/2, S) < 1/k$, so that this l occurs in the product, and

$$\varphi_k(x - l/2k) = 0.$$

Therefore g_k is equal to 0 in an open neighborhood of S. If, on the other hand, we have $d(x, S) > 2/k$ and if l occurs in the product, that is

$$d(l/2k, S) \leqq 1/k,$$

then

$$d(x, l/2k) > 1/k,$$

and hence $g_k(x) = 1$. The partial derivatives of g_k the bounded in the desired manner. This is easily seen, for if x_0 is a point where g_k is not identically 1 in a neighborhood of x_0, then $\|x_0 - l_0/2k\| \leqq 1/k$ for some l_0. All other factors $\varphi_k(x - 1/2k)$ will be identically 1 near x_0 unless $\|x_0 - l/2k\| \leqq 1/k$. But then $\|l - l_0\| \leqq 4$ whence the number of such l is bounded as a function of n (in fact by 9^n). Thus when we take the derivative, we get a sum of a most 9^n terms, each one having a derivative bounded by $C_1 k$ for some constant C_1. This proves our lemma.

We return to the proof of Criterion 2. We observe that when an $(n-1)$-form ω is expressed n terms of its coordinates,

$$\omega(x) = \sum f_j(x)\, dx_1 \wedge \cdots \wedge \widehat{dx_j} \wedge \cdots \wedge dx_n,$$

then the coefficients f_j are bounded on a compact neighborhood of S. We take U_k as in the lemma. Then for k large, each function

$$x \mapsto f_j(x) D_j g_k(x)$$

is bounded on U_k by a bound $C_2 k$, where C_2 depends on a bound for ω, and on the constant of the lemma. The Lebesgue measure of U_k is bounded by C_3/k^2, as we saw previously. Hence the measure of U_k associated with $|dg_k \wedge \omega|$ is bounded by C_4/k, and tends to 0 as $k \to \infty$. This proves our criterion.

As an example, we now state a simpler version of Stokes' theorem, applying our criteria.

Theorem 3.3. *Let X be an open subset of $\mathbf{R}^n$. Let S be the set of singular points in the closure of X, and assume that S is the finite union of C^1 images of m-rectangles with $m \leqq n-2$. Let ω be an $(n-1)$-form defined on an open neighborhood of $\bar{X}$. Assume that ω has compact support, and that the measure associated with $|\omega|$ on ∂X and with $|d\omega|$ on X are finite. Then*

$$\int_X d\omega = \int_{\partial X} \omega.$$

Proof. Immediate from our two criteria and Theorem 3.2.

We can apply Theorem 3.3 when, for instance, X is the interior of a polyhedron, whose interior is open in $\mathbf{R}^n$. When we deal with a submanifold X of dimension n, embedded in a higher dimensional space $\mathbf{R}^N$, then one can reduce the analysis of the singular set to Criterion 2 provided that there exists a finite number of charts for X near this singular set on which the given form ω is bounded. This would for instance be the case with the surface of our cone mentioned at the beginning of the section. Criterion 2 is also the natural one when dealing with manifolds defined by algebraic inequalities. By using Hironaka's resolution of singularities, one can parametrize a compact set of algebraic singularities as in Criterion 2.

Finally, we note that the condition that ω have compact support in an open neighborhood of $\bar{X}$ is a very mild condition. If for instance X is a bounded open subset of $\mathbf{R}^n$, then $\bar{X}$ is compact. If ω is any form on some open set containing $\bar{X}$, then we can find another form η which is equal to ω on some open neighborhood of $\bar{X}$ and which has compact support. The integrals of η entering into Stokes' formula will be the same as those of ω. To find η, we simply multiply ω with a suitable C^∞ function which is 1 in a neighborhood of $\bar{X}$ and vanishes a little further away. Thus Theorem 3.3 provides a reasonably useful version of Stokes' theorem which can be applied easily to all the cases likely to arise naturally.

CHAPTER XVIII

Applications of Stokes' Theorem

In this chapter we give a survey of applications of Stokes' theorem, concerning many situations. Some come just from the differential theory, such as the computation of the maximal de Rham cohomology (the space of all forms of maximal degree modulo the subspace of exact forms); some come from Riemannian geometry; and some come from complex manifolds, as in Cauchy's theorem and the Poincaré residue theorem. I hope that the selection of topics will give readers an outlook conducive for further expansion of perspectives. The sections of this chapter are logically independent of each other, so the reader can pick and choose according to taste or need.

XVIII, §1. THE MAXIMAL DE RHAM COHOMOLOGY

Let X be a manifold of dimension n without boundary. Let r be an integer $\geqq 0$. We let $\mathscr{A}^r(X)$ be the $\mathbf{R}$-vector space of differential forms on X of degree r. Thus $\mathscr{A}^r(X) = 0$ if $r > n$. If $\omega \in \mathscr{A}^r(X)$, we define the **support** of ω to be the closure of the set of points $x \in X$ such that $\omega(x) \neq 0$.

Examples. If $\omega(x) = f(x)\, dx_1 \wedge \cdots \wedge dx_n$ on some open subset of $\mathbf{R}^n$, then the support of ω is the closure of the set of x such that $f(x) \neq 0$.

We denote the support of a form ω by $\operatorname{supp}(\omega)$. By definition, the support is closed in X. We are interested in the space of maximal degree forms $\mathscr{A}^n(X)$. Every form $\omega \in \mathscr{A}^n(X)$ is such that $d\omega = 0$. On the other hand, $\mathscr{A}^n(X)$ contains the subspace of **exact** forms, which are defined to be those forms equal to $d\eta$ for some $\eta \in \mathscr{A}^{n-1}(X)$. The factor space is defined to be the **de Rham cohomology** $H^n(X) = H^n(X, \mathbf{R})$. The main theorem of this section can then be formulated.

Theorem 1.1. *Assume that X is compact, orientable, and connected. Then the map*

$$\omega \mapsto \int_X \omega$$

induces an isomorphism of $H^n(X)$ with $\mathbf{R}$ itself. In particular, if ω is in $\mathscr{A}^n(X)$ then there exists $\eta \in \mathscr{A}^{n-1}(X)$ such that $d\eta = \omega$ if and only if

$$\int_X \omega = 0.$$

Actually the hypothesis of compactness on X is not needed. What is needed is compactness on the support of the differential forms. Thus we are led to define $\mathscr{A}_c^r(X)$ to be the vector space of n-forms with compact support. We call a form **compactly exact** if it is equal to $d\eta$ for some $\eta \in \mathscr{A}_c^{r-1}(X)$. We let

$$H_c^n(X) = \text{factor space } \mathscr{A}_c^n(X)/d\mathscr{A}_c^{n-1}(X).$$

Then we have the more general version:

Theorem 1.2. *Let X be a manifold without boundary, of dimension n. Suppose that X is orientable and connected. Then the map*

$$\omega \mapsto \int_X \omega$$

induces an isomorphism of $H_c^n(X)$ with $\mathbf{R}$ itself.

Proof. By Stokes' theorem (Chapter XVII, Corollary 2.2) the integral vanishes on exact forms (with compact support), and hence induces an $\mathbf{R}$-linear map of $H_c^n(X)$ into $\mathbf{R}$. The theorem amounts to proving the converse statement: if

$$\int_X \omega = 0,$$

then there exists some $\eta \in \mathscr{A}_c^{n-1}(X)$ such that $\omega = d\eta$. For this, we first have to prove the result locally in $\mathbf{R}^n$, which we now do.

As a matter of notation, we let

$$I^n = (0, 1)^n$$

be the open n-cube in $\mathbf{R}^n$. What we want is:

Lemma 1.3. *Let ω be an n-form on I^n, with compact support, and such that*

$$\int_{I^n} \omega = 0.$$

Then there exists a form $\eta \in \mathscr{A}_c^{n-1}(I^{n-1})$ with compact support, such that

$$\omega = d\eta.$$

We will prove Lemma 1.3 by induction, but it is necessary to load to induction to carry it out. So we need to prove a stronger version of Lemma 1.3 as follows.

Lemma 1.4. *Let ω be an $(n-1)$-form on I^{n-1} whose coefficient is a function of n variables $(x_1, \ldots, x_n)$ so*

$$\omega(x) = f(x_1, \ldots, x_n)\, dx_1 \wedge \cdots \wedge dx_{n-1}.$$

(Of course, all functions, like forms, are assumed C^∞.) Suppose that ω has compact support in I^{n-1}. Assume that

$$\int_{I^{n-1}} \omega = 0.$$

Then there exists an $(n-1)$-form η, whose coefficients are C^∞ functions of $x_1, \ldots, x_n$ with compact support such that

$$\omega(x_1, \ldots, x_{n-1}; x_n) = d_{n-1}\, \eta(x_1, \ldots, x_{n-1}; x_n).$$

The symbol d_{n-1} here means the usual exterior derivative taken with respect to the first $n-1$ variables.

Proof. By induction. We first prove the theorem when $n - 1 = 1$. First we carry out the proof leaving out the extra variable, just to see what's going on. So let

$$\omega(x) = f(x)\, dx,$$

where f has compact support in the open interval $(0, 1)$. This means there exists $\epsilon > 0$ such that $f(x) = 0$ if $0 < x \leqq \epsilon$ and if $1 - \epsilon \leqq x \leqq 1$. We assume

$$\int_0^1 f(x)\, dx = 0.$$

Let

$$g(x) = \int_0^x f(t)\,dt.$$

Then $g(x) = 0$ if $0 < x \leqq \epsilon$, and also if $1 - \epsilon \leqq x \leqq 1$, because for instance if $1 - \epsilon \leqq x \leqq 1$, then

$$g(x) = \int_0^1 f(t)\,dt = 0.$$

Then $f(x)\,dx = dg(x)$, and the lemma is proved in this case. Note that we could have carried out the proof with the extra variable x_2, starting from

$$\omega(x) = f(x_1, x_2)\,dx_1,$$

so that

$$g(x_1, x_2) = \int_0^1 f(t, x_2)\,dt.$$

We can differentiate under the integral sign to verify that g is C^∞ in the pair of variables (x_1, x_2).

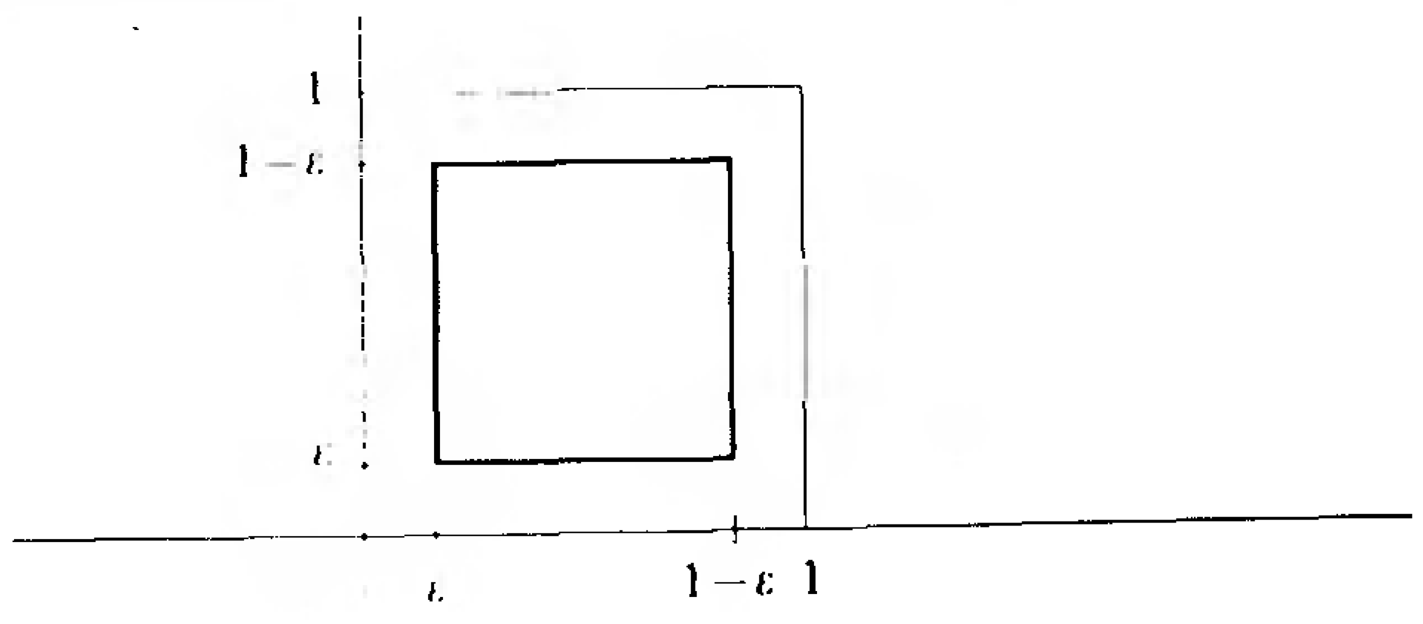
Now let $n \geqq 3$ and assume that theorem proved for $n-1$ by induction. To simplify the notation, let us omit the extra variable x_{n+1}, and write

$$\omega(x) = f(x_1, \dots, x_n)\,dx_1 \wedge \cdots \wedge dx_n,$$

with compact support in I^n. Then there exists $\epsilon > 0$ such that the support of f is contained in the closed cube

$$\bar{I}^n(\epsilon) = [\epsilon, 1-\epsilon]^n.$$

The following figure illustrates this support in dimension 2.



Let ψ be an $(n-1)$-form on I^{n-1}, $\psi(x) = \psi(x_1, \dots, x_{n-1})$ such that

$$\int_{I^{n-1}} \psi = 1,$$

and ψ has compact support. Let

$$\begin{aligned} g(x_n) &= \int_{I^{n-1}} f(x_1, \dots, x_{n-1}; x_n)\,dx_1 \wedge \cdots \wedge dx_{n-1} \\ &= \int_{\bar{I}^{n-1}(\epsilon)} f(x_1, \dots, x_{n-1}; x_n)\,dx_1 \wedge \cdots \wedge dx_{n-1}. \end{aligned}$$

Note here that we do have the parameter x_n coming in at the inductive step. Let

$$\mu(x) = f(x)\,dx_1 \wedge \cdots \wedge dx_{n-1} - g(x_n)\psi(x_1, \dots, x_{n-1}),$$

so

$$(*) \qquad \mu(x) \wedge dx_n = \omega(x) - g(x_n)\psi(x) \wedge dx_n.$$

Then

$$\int_{I^{n-1}} \mu = g(x_n) - g(x_n) = 0.$$

Furthermore, since f has compact support, so does g (look at the figure). By induction, there exists an $(n-1)$-form η, of the first $n-1$ variables, but depending on the parameter x_n, that is

$$\eta(x) = \eta(x_1, \dots, x_{n-1}; x_n)$$

such that

$$\mu(x_1, \dots, x_{n-1}; x_n) = d_{n-1}\eta(x_1, \dots, x_{n-1}; x_n).$$

Here d_{n-1} denotes the exterior derivative with respect to the first $n-1$ variables. Then trivially,

$$\begin{aligned} \mu(x_1, \dots, x_{n-1}; x_n) \wedge dx_n &= d_{n-1}\,\eta(x_1, \dots, x_{n-1}; x_n) \wedge dx_n \\ &= d\eta(x), \end{aligned}$$

where $d\eta$ is now the exterior derivative taken with respect to all n variables. Hence finally from equation $(*)$ we obtain

$$(**) \qquad \omega(x) = d\eta(x) + g(x_n)\psi(x_1, \dots, x_{n-1}) \wedge dx_n.$$

To conclude the proof of Lemma 1.3, it suffices to show that the second term on the right of $(**)$ is exact. We are back to a one-variable problem. Let

$$h(x_n) = \int_0^{x_n} g(t)\, dt.$$

Then $dh(x_n) = g(x_n)dx_n$, and h has compact support in the interval $(0, 1)$, just as in the start of the induction. Then

$$\begin{aligned} d\big(h(x_n)\psi(x_1,\dots,x_{n-1})\big) &= dh(x_n) \wedge \psi(x_1,\dots,x_{n-1}) \\ &= (-1)^{n-1} g(x_n)\psi(x_1,\dots,x_{n-1}) \wedge dx_n \end{aligned}$$

because $d\psi = 0$. Of course we could have carried along the extra parameter all the way through. This concludes the proof of Lemma 1.3.

We formulate an immediate consequence of Lemma 1.3 directly on the manifold.

Lemma 1.5. *Let U be an open subset of X, isomorphic to I^n. Let $\psi \in \mathscr{A}_c^n(U)$ be such that*

$$\int_U \psi \neq 0.$$

Let $\omega \in \mathscr{A}_c^n(U)$. Then there exists $c \in \mathbf{R}$ and $\eta \in \mathscr{A}_c^{n-1}(U)$ such that

$$\omega - c\psi = d\eta.$$

Proof. We take $c = \int_U \omega \Big/ \int_U \psi$ and apply Lemma 1.3 to $\omega - c\psi$.

Observe that the hypothesis of connectedness has not yet entered the picture. The preceding lemmas were purely local. We now globalize.

Lemma 1.6. *Assume that X is connected and oriented. Let U, ψ be as in Lemma 1.5. Let V be the set of points $x \in X$ having the following property. There exists a neighborhood $U(x)$ of x isomorphic to I^n such that for every $\omega \in \mathscr{A}_c^n(U(x))$ there exist $c \in \mathbf{R}$ and $\eta \in \mathscr{A}_c^{n-1}(X)$ such that*

$$\omega - c\psi = d\eta.$$

Then $V = X$.

Proof. Lemma 1.5 asserts that $V \supset U$. Since X is connected, it suffices to prove that V is both open and closed. It is immediate from the definition of V that V is open, so there remains to prove its closure. Let z be in the closure of V. Let W be a neighborhood of z isomorphic to I^n. There exists a point $x \in V \cap W$. There exists a neighborhood $U(x)$ as in

the definition of V such that $U(x) \subset W$. For instance, we may take

$$U(x) \approx (a_1, b_1) \times \cdots \times (a_n, b_n) \approx I^n$$

with a_i sufficiently close to 0 and b_i sufficiently close to 1, and of course $0 < a_i < b_i$ for $i = 1,\dots,n$. Let $\psi_1 \in \mathscr{A}_c^n(U(x))$ be such that

$$\int_{U(x)} \psi_1 = 1.$$

Let $\omega \in \mathscr{A}_c^n(W)$. By the definition of V, there exist $c_1 \in \mathbf{R}$ and $\eta_1 \in \mathscr{A}_c^n(X)$ such that

$$\psi_1 - c_1\psi = d\eta_1.$$

By Lemma 1.5, there exists $c_2 \in \mathbf{R}$ and $\eta_2 \in \mathscr{A}_c^n(X)$ such that

$$\omega - c_2\psi_1 = d\eta_2.$$

Then

$$\omega - c_2c_1\psi = d(\eta_2 + c_2\eta_1),$$

thus concluding the proof of Lemma 1.6.

We have now reached the final step in the proof of Theorem 5.2, namely we first fix a form $\psi \in \mathscr{A}_c^n(U)$ with $U \approx I^n$ and $\int_X \psi \neq 0$. Let $\omega \in \mathscr{A}_c^n(X)$. It suffices to prove that there exist $c \in \mathbf{R}$ and $\eta \in \mathscr{A}_c^{n-1}(X)$ such that

$$\omega - c\psi = d\eta.$$

Let K be the compact support of ω. Cover K by a finite number of open neighborhoods $U(x_1),\dots,U(x_m)$ satisfying the property of Lemma 1.6. Let $\{\varphi_i\}$ be a partition of unity subordinated to this covering, so that we can write

$$\omega = \sum \varphi_i\omega.$$

Then each form $\varphi_i\omega$ has support in some $U(x_j)$. Hence by Lemma 5.6, there exist $c_i \in \mathbf{R}$ and $\eta_i \in \mathscr{A}_c^{n-1}(X)$ such that

$$\varphi_i\omega - c_i\psi = d\eta_i,$$

whence $\omega - c\psi = d\eta$, with $c = \sum c_i$ and $\eta = \sum \eta_i$. This concludes the proof of Theorems 1.1 and 1.2.

XVIII, §2. MOSER'S THEOREM

We return here to the techniques of proof in Chapter V, as for Poincaré's lemma, Theorem 5.1 and Darboux's Theorem 7.3 of that chapter. However, we now have a similar theorem in the context of integration.

We first make the general comment, similar to the one we made previously, for general forms. Let $\mathbf{E}$ be a Banach space, and let ω be an r-multilinear alternating form on $\mathbf{E}$ (so $\mathbf{R}$-valued). We say that ω is non-singular if for each vector $v \in \mathbf{E}$, defining ω_v by

$$\omega_v\colon (v_1, \dots, v_{r-1}) \mapsto (v, v_1, \dots, v_{r-1}),$$

the map $v \mapsto \omega_v$ is a toplinear isomorphism between $\mathbf{E}$ and $L_a^{r-1}(\mathbf{E})$. We previously considered bilinear forms, in Chapter V, §6.

We can globalize the notion to a manifold, so a form $\omega \in \mathscr{A}^r(X)$ is called **non-singular** if $\omega(x)$ is non-singular for each x. It is clear that in the finite dimensional case, a volume form is non-singular. With this globalization, we obtain:

Proposition 2.1. *Let ω be a non-singular r-form on X. Given a form $\eta \in \mathscr{A}^{r-1}(X)$, there exists a unique vector field ξ such that*

$$\omega \circ \xi = \eta.$$

We could also write the relation with the contraction notation, i.e. $C_\xi \omega = \eta$.

We now come to Moser's theorem [Mo 65].

Theorem 2.2. *Let X be a compact, connected oriented manifold of dimension n. Let $\omega, \psi \in \mathscr{A}^n(X)$ $\left(= \mathscr{A}_c^n(X)\right)$ be volume forms such that*

$$\int_X \omega = \int_X \psi.$$

Then there exists an automorphism $f\colon X \to X$ of X such that $\omega = f^*\psi$.

Proof. Let

$$\omega_s = (1-s)\omega + s\psi \qquad \text{for} \quad 0 \leqq s \leqq 1.$$

Then ω_s is a volume form for each s, and in particular is non-singular. By Theorem 1.1, there exists $\eta \in \mathscr{A}^{n-1}(X)$ such that $\psi - \omega = d\eta$. Note also that $\psi - \omega = d\omega_s/ds$. Since ω_s is non-singular, there exists a unique vector field ξ_s such that

$$\omega_s \circ \xi_s = -\eta.$$

Let α_s be the flow of ξ_s. Then α_s is defined on $\mathbf{R} \times X$ by Corollary 2.4 of Chapter IV. Then we get:

$$\begin{aligned} \frac{d}{ds}(\alpha_s^*\omega_s) &= \frac{d}{du}(\alpha_u^*\omega_s)\Big|_{u=s} + \alpha_s^*\left(\frac{d\omega_s}{ds}\right) \\ &= \alpha_s^*\, d(\omega_s \circ \xi_s) + \alpha_s^*(\psi - \omega) \quad \text{by Proposition 5.2 of Chapter V} \\ &= -\alpha_s^* d\eta + \alpha_s^* d\eta \\ &= 0. \end{aligned}$$

Therefore $\alpha_s^*\omega_s$ is constant as a function of s, so we find

$$\omega = \alpha_0^*\omega_0 = \alpha_1^*\omega_1 = f^*\psi, \qquad \text{with } f = \alpha_1,$$

thereby proving the theorem.

XVIII, §3. THE DIVERGENCE THEOREM

Let X be an oriented manifold of dimension n possibly with boundary, and let Ω be an n-form on X. Let ξ be a vector field on X. Then $d\Omega = 0$, and hence the basic formula for the Lie derivative (Chapter V, Proposition 5.3) shows that

$$\mathscr{L}_\xi \Omega = d(\Omega \circ \xi).$$

Consequently in this case, Stokes' theorem yields:

Theorem 3.1 (Divergence Theorem).

$$\boxed{\int_X \mathscr{L}_\xi \Omega = \int_{\partial X} \Omega \circ \xi.}$$

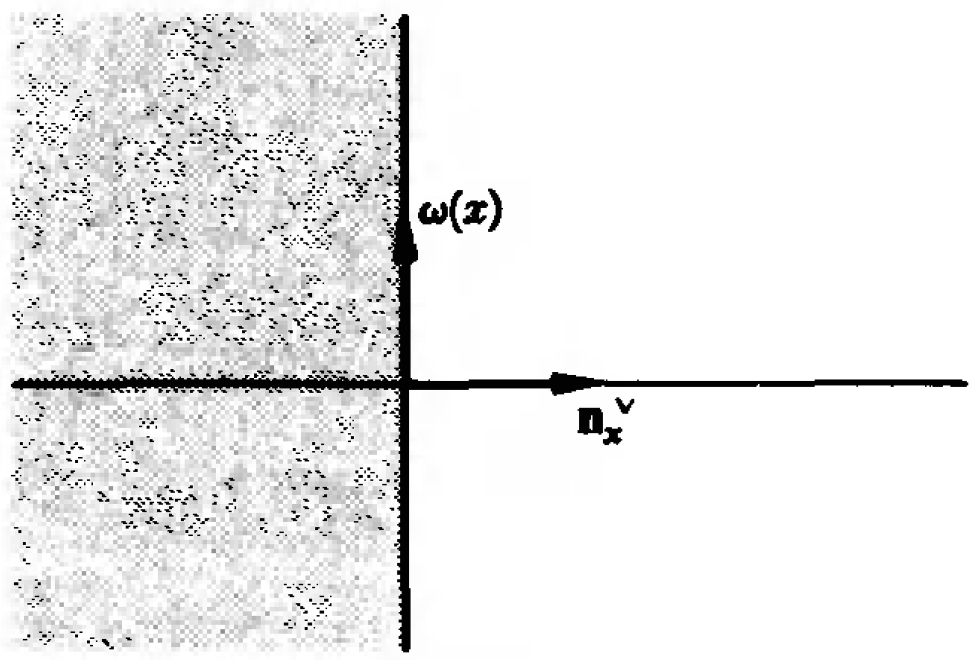
Remark. Even if the manifold is not orientable, it is possible to use the notion of density to formulate a Stokes theorem for densities. Cf. Loomis–Sternberg [Los 68] for the formulation, due to Rasala. However, this formulation reduces at once to a local question (using partitions of unity on densities). Since locally every manifold is orientable, and a density then amounts to a differential form, this more general formulation again reduces to the standard one on an orientable manifold.

Suppose that (X, g) is a Riemannian manifold, assumed oriented for simplicity. We let Ω or vol_g be the volume form defined in Chapter XV,

§1. Let ω be the canonical Riemannian volume form on ∂X for the metric induced by g on the boundary. Let $\mathbf{n}_x$ be a unit vector in the tangent space $T_x(X)$ such that u is perpendicular to $T_x(\partial X)$. Such a unit vector is determined up to sign. Denote by $\mathbf{n}_x^\vee$ its dual functional, i.e. the component on the projection along $\mathbf{n}_x$. *We select $\mathbf{n}_x$ with the sign such that*

$$\mathbf{n}_x^\vee \wedge \omega(x) = \Omega(x).$$

We then shall call $\mathbf{n}_x$ the **unit outward normal vector** to the boundary at x. In an oriented chart, it looks like this.



Then by formula **CON 3** of Chapter V, §5 we find

$$\Omega \circ \xi = \langle \mathbf{n}, \xi \rangle \omega - \mathbf{n}^\vee \wedge (\omega \circ \xi),$$

and the restriction of this form to ∂X is simply $\langle \mathbf{n}, \xi \rangle \omega$. Thus we get:

Theorem 3.2 (Gauss Theorem). *Let X be a Riemannian manifold. Let ω be the canonical Riemannian volume form on ∂X and let Ω be the canonical Riemannian volume form on X itself. Let $\mathbf{n}$ be the unit outward normal vector field to the boundary, and let ξ be a C^1 vector field on X, with compact support. Then*

$$\int_X (\operatorname{div}_\Omega \xi)\Omega = \int_{\partial X} \langle \mathbf{n}, \xi \rangle \omega.$$

The next thing is to show that the map d^* from Chapter XV, §1 is the adjoint for a scalar product defined by integration. First we expand slightly the formalism of d^* for this application. Recall that for any vector field ξ, the divergence of ξ is defined by the property

(1) $$d(\mathrm{vol}_g \circ \xi) = (\operatorname{div} \xi)\mathrm{vol}_g.$$

Note the trivial derivation formula for a function φ:

(2) $$\operatorname{div}(\varphi\xi) = \varphi \operatorname{div} \xi + (d\varphi)(\xi).$$

If λ is a 1-form, i.e. in $\Gamma L^1(TX) = \mathscr{A}^1(X)$, we have the corresponding vector field $\xi_\lambda = \lambda^\vee$ uniquely determined by the condition that

$$\langle \xi_\lambda, \eta \rangle_g = \lambda(\eta) \qquad \text{for all vector fields } \eta.$$

For a 1-form λ, we define the operator

$$d^*\colon \mathscr{A}^1(X) \to \mathscr{A}^0(X) = \mathrm{Fu}(X) \qquad \text{by} \qquad d^*\lambda = -\operatorname{div} \xi_\lambda,$$

so by (1),

(3) $$(d^*\lambda)\, \mathrm{vol}_g = d(\mathrm{vol}_g \circ \xi_\lambda).$$

We get a formula analogous to (2) for d^*, namely

(4) $$d^*(\varphi\lambda) = \varphi d^*\lambda - \langle d\varphi, \lambda \rangle.$$

Indeed, $d^*(\varphi\lambda) = -\operatorname{div} \xi_{\varphi\lambda} = -\operatorname{div}(\varphi\xi_\lambda) = -\varphi \operatorname{div} \xi_\lambda - (d\varphi)(\xi_\lambda)$ by (2), which proves the formula.

Let $\lambda,\ \omega \in \mathscr{A}^1(TX)$. We define the **scalar product** via duality

$$\langle \lambda, \omega \rangle_g = \langle \xi_\lambda, \xi_\omega \rangle_g.$$

Then for a function φ we have the formula

(5) $$\langle d\varphi, \lambda \rangle_g \mathrm{vol}_g = (\varphi d^*\lambda)\, \mathrm{vol}_g - d(\mathrm{vol}_g \circ \varphi\xi_\lambda).$$

Indeed,

$$\begin{aligned} \langle d\varphi, \lambda \rangle_g \mathrm{vol}_g &= [\varphi\, d^*\lambda - d^*(\varphi\lambda)]\, \mathrm{vol}_g && \text{by (4)} \\ &= (\varphi d^*\lambda)\, \mathrm{vol}_g - d(\mathrm{vol}_g \circ \varphi\xi_\lambda) && \text{by (3)} \end{aligned}$$

thus proving (5). Note that the congruence of the two forms $\langle d\varphi, \lambda \rangle_g \mathrm{vol}_g$ and $(\varphi d^*\lambda)\, \mathrm{vol}_g$ modulo exact forms is significant, and is designed for Proposition 3.3 below.

Observe that the scalar product between two forms above is a function, which when multiplied by the volume form vol_g may be integrated over X. Thus we define the **global scalar product** on 1-forms with compact support to be

$$\langle \lambda, \omega \rangle_{(X,g)} = \langle \lambda, \omega \rangle_X = \int_X \langle \lambda, \omega \rangle_g \mathrm{vol}_g.$$

Applying Stokes' theorem, we then find:

Proposition 3.3. *Let (X, g) be a Riemannian manifold, oriented and without boundary. Then d^* is the adjoint of d with respect to the global scalar product, i.e.*

$$\langle d\varphi, \lambda\rangle_X = \langle \varphi, d^*\lambda\rangle_X.$$

We define the **Laplacian** (operating on functions) to be the operator

$$\Delta = d^*d.$$

For the Laplacian operating on higher degree forms, we shall give the expression $d^*d + dd^*$ in the next section, but here for functions, the second term disappears.

For a manifold with boundary, we define the **normal derivative** of a function φ to be the function *on the boundary* given by

$$\partial_{\mathbf{n}}\varphi = \langle \mathbf{n}, \xi_{d\varphi}\rangle_g = \langle \mathbf{n}, \operatorname{grad}_g \varphi\rangle_g.$$

Theorem 3.4 (Green's Formula). *Let (X, g) be an oriented Riemannian manifold possibly with boundary, and let φ, ψ be functions on X with compact support. Let ω be the canonical volume form associated with the induced metric on the boundary. Then*

$$\int_X (\varphi\Delta\psi - \psi\Delta\varphi)\,\mathrm{vol}_g = -\int_{\partial X} (\varphi\partial_{\mathbf{n}}\psi - \psi\partial_{\mathbf{n}}\varphi)\omega.$$

Proof. From formula (4) we get

$$d^*(\varphi\, d\psi) = \varphi\Delta\psi - \langle d\varphi, d\psi\rangle_g,$$

whence

$$\begin{aligned}\varphi\Delta\psi - \psi\Delta\varphi &= d^*(\varphi\, d\psi) - d^*(\psi\, d\varphi)\\ &= -\operatorname{div}(\varphi\, d\psi) + \operatorname{div}(\psi\, d\varphi).\end{aligned}$$

We apply Theorem 3.2 to conclude the proof.

Remark. Of course, if X has no boundary in Theorem 3.7, then the integral on the left side is equal to 0.

Corollary 3.5 (E. Hopf). *Let X be a Riemannian manifold without boundary, and let f be a C^2 function on X with compact support, such that $\Delta f \geqq 0$. Then f is constant. In particular, every harmonic function with compact support is constant.*

Proof. We first give the proof assuming that X is oriented. By Green's formula we get

$$\int_X \Delta f\,\mathrm{vol}_g = 0.$$

Since $\Delta f \geqq 0$, it follows that in fact $\Delta f = 0$, so we are reduced to the harmonic case. We now apply Green's formula to f^2, and get

$$0 = \int_X \Delta f^2\,\mathrm{vol}_g = \int_X 2f\Delta f\,\mathrm{vol}_g - \int_X 2(\operatorname{grad} f)^2\,\mathrm{vol}_g.$$

Hence $(\operatorname{grad} f)^2 = 0$ because $\Delta f = 0$, and finally $\operatorname{grad} f = 0$, so $df = 0$ and f is constant, thus proving the corollary in the oriented case. For the non-oriented case, by Proposition 4.6 of Chapter XVI, there exists a covering of degree 2 of X which is oriented, and then one can pull back all the objects from X to this covering to conclude the proof in this case.

XVIII, §4. THE ADJOINT OF d FOR HIGHER DEGREE FORMS

We extend the results of the preceding section to arbitrary forms. Given the vector space V of dimension n over $\mathbf{R}$, with a positive definite scalar product g, we note that the exterior powers $\bigwedge^r V$ are self dual, with a positive definite scalar product such that

$$(v_1 \wedge \cdots \wedge v_r, w_1 \wedge \cdots \wedge w_r) \mapsto \det\langle v_i, w_j\rangle_g.$$

We defined the notion of orientation on V in Chapter XV, §1, and we now assume that V is oriented.

Proposition 4.1. *Given $1 \leqq r \leqq n$, there exists a unique isomorphism*

$$*\colon \bigwedge^r V \to \bigwedge^{n-r} V$$

such that for $\varphi, \psi \in \bigwedge^r V$ we have

$$\langle \varphi, \psi\rangle_g\,\mathrm{vol}_g = \varphi \wedge *\psi.$$

Proof. The proof will give an explicit determination of the isomorphism on the usual for $\bigwedge^r V$. Let $I = [i_1 < i_2 < \cdots < i_r]$ be an ordered set of r indices. We let

$$e_I = e_{i_1} \wedge \cdots \wedge e_{i_r}.$$

If I' is another such ordered set with $n-r$ elements, and $I \cup I' = \{1,\dots,n\}$ then we let ϵ_I be the sign of the permutation (I, J) of $(1,\dots,n)$. We then define

$$*e_I = \epsilon_I e_{I'},$$

and extend this operation by linearity to all of $\bigwedge^r V$. Then directly from the definition, we see that if J is an ordered set of r indices, then

$$\begin{aligned} e_I \wedge *e_J &= \langle e_I, e_J\rangle e_1 \wedge \cdots \wedge e_n \\ &= \delta_{IJ}\, e_1 \wedge \cdots \wedge e_n. \end{aligned}$$

Thus on the standard basis elements of $\bigwedge^r V$ the desired relation of the proposition is satisfied. The same relation is therefore satisfied for all elements of $\bigwedge^r V$, as desired.

We define the operator $\mathbf{w}$ on the direct sum $\bigoplus_r \bigwedge^r V$ to have the effect

$$\mathbf{w} = (-1)^{nr+r} \quad \text{on} \quad \bigwedge^r V.$$

Proposition 4.2. *We have* $*\mathbf{w} = \mathbf{w}*$. *If* n *is even, then* $\mathbf{w} = (-1)^r$ *on* $\bigwedge^r V$. *Furthermore,* $** = \mathbf{w}$.

Proof. Direct, simple computations.

We now apply the above to a Riemannian manifold X of dimension n, and to real differential forms. We let

$$\mathscr{A}_c^r(X)$$

be the space of C^∞ differential forms of degree r, with compact support on the manifold. At each point $x \in X$, we use the space $V = T_x^\vee$ (the dual space of the tangent space). The usual operator

$$d\colon \mathscr{A}_c^r(X) \to \mathscr{A}_c^{r+1}(X)$$

is $\mathbf{R}$-linear. By Stokes' theorem, if ω has compact support, then

$$\int_X d\omega = 0.$$

We shall give an application of this fact in a Riemannian context. We have the volume form vol_g (which does not necessarily have compact support) and we define a **scalar product** on $\mathscr{A}_c^r(X)$ by the formula

$$\langle \varphi, \psi\rangle_{X,g} = \langle \varphi, \psi\rangle_X = \int_X \varphi \wedge *\psi = \int_X \langle \varphi, \psi\rangle_g \,\mathrm{vol}_g,$$

where we usually omit the index g and merely write X as in $\langle \varphi, \psi\rangle_X$.

Proposition 4.3. *The exterior derivative* d *has an adjoint* d^* *with respect to the scalar product* $\langle\, ,\rangle_X$, *namely for* $\varphi \in \mathscr{A}_c^{r-1}(X)$ *and* $\psi \in \mathscr{A}_c^r(X)$ *we have*

$$\langle d\varphi, \psi\rangle_X = \langle \varphi, d^*\psi\rangle_X.$$

Furthermore, the adjoint is given by the explicit formula

$$\begin{aligned} d^* &= (-1)^{nr+n+1} * d * \quad &&\text{on } \mathscr{A}_c^r(X). \\ &= -*d* &&\text{if } n \text{ is even.} \end{aligned}$$

Proof. By Stokes' theorem, we have:

$$\begin{aligned} \int_X d\varphi \wedge *\psi &= \int_X d(\varphi \wedge *\psi) - (-1)^{r-1}\int_X \wedge d*\psi \\ &= (-1)^r \int_X \varphi \wedge d*\psi. \end{aligned}$$

Now

$$\begin{aligned} (-1)^r \varphi \wedge d*\psi &= (-1)^r \varphi \wedge **\mathbf{w}\, d*\psi \\ &= (-1)^r \varphi \wedge \mathbf{w}*(*d*)\psi \\ &= (-1)^{nr+n+1} \varphi \wedge *(*d*)\psi, \end{aligned}$$

which proves the proposition.

XVIII, §5. CAUCHY'S THEOREM

It is possible to define a complex analytic (analytic, for short) manifold, using open sets in $\mathbf{C}^n$ and charts such that the transition mappings are analytic. Since analytic amps are C^∞, we see that we get a C^∞ manifold, but with an additional structure, and we call such a manifold **complex analytic**. It is verified at once that the analytic charts of such a manifold define an orientation. Indeed, under a complex analytic change of charts, the Jacobian changes by a complex number times its complex conjugate, so changes by a positive real number.

If $z_1, \dots, z_n$ are the complex coordinates of $\mathbf{C}^n$, then

$$(z_1, \dots, z_n, \bar{z}_1, \dots, \bar{z}_n)$$

can be used a C^∞ local coordinates, viewing $\mathbf{C}^n$ as $\mathbf{R}^{2n}$. If $z_k = x_k + iy_k$, then

$$dz_k = dx_k + i\,dy_k \quad \text{and} \quad d\bar{z}_k = dx_k - i\,dy_k.$$

Differential forms can then be expressed in terms of wedge products of the dz_k and $d\bar{z}_k$. For instance

$$dz_k \wedge d\bar{z}_k = 2i\,dy_k \wedge dx_k.$$

The complex standard expression for a differential form is then

$$\omega(z) = \sum_{(i,j)} \varphi_{(i,j)}(z)\,dz_{i_1} \wedge \cdots \wedge dz_{i_r} \wedge d\bar{z}_{j_1} \wedge \cdots \wedge d\bar{z}_{j_s}.$$

Under an analytic change of coordinates, one sees that the numbers r and s remain unchanged, and that if $s = 0$ in one analytic chart, then $s = 0$ in any other analytic chart. Similarly for r. Thus we can speak of a form of type (r, s). A form is said to be **analytic** if $s = 0$, that is if it is of type $(r, 0)$.

We can decompose the exterior derivative d into two components. Namely, we note that if ω is of type (r, s), then $d\omega$ is a sum of forms of type $(r+1, s)$ and $(r, s+1)$, say

$$d\omega = (d\omega)_{(r+1,\,s)} + (d\omega)_{(r,\,s+1)}.$$

We define

$$\partial\omega = (d\omega)_{(r+1,\,s)} \quad \text{and} \quad \bar{\partial}\omega = (d\omega)_{(r,\,s+1)}.$$

In terms of local coordinates, it is then easy to verify that if ω is decomposable, and is expressed as

$$\omega(z) = \varphi(z)dz_{i_1} \wedge \cdots \wedge dz_{i_r} \wedge d\bar{z}_{j_1} \wedge \cdots \wedge d\bar{z}_{j_s} = \varphi\tilde{\omega},$$

then

$$\partial\omega = \sum \frac{\partial\varphi}{\partial z_k}\,dz_k \wedge \tilde{\omega}.$$

and

$$\bar{\partial}\omega = \sum \frac{\partial\varphi}{\partial \bar{z}_k}\,d\bar{z}_k \wedge \tilde{\omega}.$$

In particular, we have

$$\frac{\partial}{\partial z_k} = \frac{1}{2}\left(\frac{\partial}{\partial x_k} - i\frac{\partial}{\partial y_k}\right) \quad \text{and} \quad \frac{\partial}{\partial \bar{z}_k} = \frac{1}{2}\left(\frac{\partial}{\partial x_k} + i\frac{\partial}{\partial y_k}\right).$$

(*Warning*: Note the position of the plus and minus signs in these expressions.)

Thus we have

$$d = \partial + \bar{\partial},$$

and operating with ∂ or $\bar{\partial}$ follows rules similar to the rules for operating with d.

Note that f is analytic if and only if $\bar{\partial}f = 0$. Similarly, we say that a differential form is **analytic** if in its standard expression, the functions $\varphi_{(i,j)}$ are analytic and the form is of type $(r, 0)$, that is there are no $d\bar{z}_j$ present. Equivalently, this amounts to saying that $\bar{\partial}\omega = 0$. The following extension of Cauchy's theorem to several variables is due to Martinelli.

We let $|z|$ be the euclidean norm,

$$|z| = (z_1\bar{z}_1 + \cdots + z_n\bar{z}_n)^{1/2}.$$

Theorem 5.1 (Cauchy's Theorem). *Let f be analytic on an open set in $\mathbf{C}^n$ containing the closed ball of radius R centered at a point ζ. Let*

$$\omega_k(z) = dz_1 \wedge \cdots \wedge dz_n \wedge d\bar{z}_1 \wedge \cdots \wedge \widehat{d\bar{z}_k} \wedge \cdots \wedge d\bar{z}_n$$

and

$$\omega(z) = \sum_{k=1}^{n} (-1)^k \bar{z}_k \omega_k(z).$$

Let S_R be the sphere of radius R centered at ζ. Then

$$f(\zeta) = \epsilon(n)\frac{(n-1)!}{(2\pi i)^n}\int_{S_R} \frac{f(z)}{|z-\zeta|^{2n}}\,\omega(z-\zeta)$$

where $\epsilon(n) = (-1)^{n(n+1)/2}$.

Proof. We may assume $\zeta = 0$. First note that

$$\bar{\partial}\omega(z) = \sum_{k=1}^{n} (-1)^k d\bar{z}_k \wedge \omega_k(z) = (-1)^{n+1} n\,dz \wedge d\bar{z},$$

where $dz = dz_1 \wedge \cdots \wedge dz_n$ and similarly for $d\bar{z}$. Next, observe that if

$$\psi(z) = \frac{f(z)}{|z|^{2n}}\,\omega(z),$$

then

$$d\psi = 0.$$

This is easily seen. On the one hand, $\partial\psi = 0$ because ω already has $dz_1 \wedge \cdots \wedge dz_n$, and any further dz_i wedged with this gives 0. On the other hand, since f is analytic, we find that

$$\bar{\partial}\psi(z) = f(z)\,\bar{\partial}\left(\frac{\omega(z)}{|z|^{2n}}\right) = 0$$

by the rule for differentiating a product and a trivial computation.

Therefore, by Stokes' theorem, applied to the annulus between two spheres, for any r with $0 < r \leqq R$ we get

$$\int_{S_R} \psi - \int_{S_r} \psi = 0,$$

or in other words,

$$\int_{S_R} f(z)\frac{\omega(z)}{|z|^{2n}} = \int_{S_r} f(z)\frac{\omega(z)}{|z|^{2n}} = \frac{1}{r^{2n}}\int_{S_r} f(z)\omega(z).$$

Using Stokes' theorem once more, and the fact that $\partial\omega = 0$, we see that this is

$$= \frac{1}{r^{2n}}\int_{B_r} \bar{\partial}(f\omega) = \frac{1}{r^{2n}}\int_{B_r} f\bar{\partial}\omega.$$

We can write $f(z) = f(0) + g(z)$, where $g(z)$ tends to 0 as z tends to 0. Thus in taking the limit as $r \to 0$, we may replace f by $f(0)$. Hence our last expression has the same limit as

$$f(0)\frac{1}{r^{2n}}\int_{B_r} \bar{\partial}\omega = f(0)\frac{1}{r^{2n}}\int_{B_r} (-1)^{n+1} n\, dz \wedge d\bar{z}.$$

But

$$dz \wedge d\bar{z} = (-1)^{n(n-1)/2} i^n 2^n\, dy_1 \wedge dx_1 \wedge \cdots \wedge dy_n \wedge dx_n.$$

Interchanging dy_k and dx_k to get the proper orientation gives another contribution of $(-1)^n$, together with the form giving Lebesgue measure. Hence our expression is equal to

$$f(0)(-1)^{n(n+1)/2} n(2i)^n \frac{1}{r^{2n}} V(B_r),$$

where $V(B_r)$ is the Lebesgue volume of the ball of radius r in $\mathbf{R}^{2n}$, and is classically known to be equal to $\pi^n r^{2n}/n!$. Thus finally we see that our expression is equal to

$$f(0)(-1)^{n(n+1)/2}\frac{(2\pi i)^n}{(n-1)!}.$$

This proves Cauchy's theorem.

XVIII, §6. THE RESIDUE THEOREM

Let f be an analytic function in an open set U of $\mathbf{C}^n$. The set of zeros of f is called a **divisor**, which we denote by $V = V_f$. In the neighborhood of a regular point a, that is a point where $f(a) = 0$ but some complex partial derivative of f is not zero, the set V is a complex submanifold of U. In fact, if, say, $D_n f(a) \neq 0$, then the map

$$(z_1, \ldots, z_n) \mapsto (z_1, \ldots, z_{n-1}, f(z))$$

gives a local analytic chart (analytic isomorphism) in a neighborhood of a. Thus we may use f as the last coordinate, and locally V is simply obtained by the projection on the set $f = 0$. This is a special case of the complex analytic inverse function theorem.

It is always true that the function $\log|f|$ is locally in $\mathscr{L}^1$. We give the proof only in the neighborhood of a regular point a. In this case, we can change f by a chart (which is known as a change-of-variable formula), and we may therefore assume that $f(z) = z_n$. Then $\log|f| = \log|z_n|$, and the Lebesgue integral decomposes into a simple product integral, which reduces our problem to the case of one variable, that is to the fact that $\log|z|$ is locally integrable near 0 in the ordinary complex plane. Writing $z = re^{i\theta}$, our assertion is obvious since the function $r\log r$ is locally integrable near 0 on the real line.

Note. In a neighborhood of a singular point the fastest way and formally clearest, is to invoke Hironaka's resolution of singularities, which reduces the question to the non-singular case.

For the next theorem, it is convenient to let

$$d^c = \frac{1}{4\pi i}(\partial - \bar{\partial}).$$

Note that

$$dd^c = \frac{i}{2\pi}\partial\bar{\partial}.$$

The advantage of dealing with d and d^c is that they are real operators.

The next theorem, whose proof consists of repeated applications of Stokes' theorem, is due to Poincaré. It relates integration in V and U by a suitable kernel.

Theorem 6.1 (Residue Theorem). *Let f be analytic on an open set U of $\mathbf{C}^n$ and let V be its divisor of zeros in U. Let ψ be a C^∞ form with compact support in U, of degree $2n-2$ and type $(n-1, n-1)$. Then*

$$\int_V \psi = \int_U \log |f|^2 \, dd^c\psi.$$

(As usual, the integral on the left is the integral of the restriction of ψ to V, and by definition, it is taken over the regular points of V.)

Proof. Since ψ and $dd^c\psi$ have compact support, the theorem is local (using partitions of unity). We give the proof only in the neighborhood of a regular point. Therefore we may assume that U is selected sufficiently small so that every point of the divisor of f in U is regular, and such that, for small ϵ, the set of points

$$U_\epsilon = \{z \in U, |f(z)| \geqq \epsilon\}$$

is a submanifold with boundary in U. The boundary of U_ϵ is then the set of points z such that $|f(z)| = \epsilon$. (Actually to make this set a submanifold we only need to select ϵ to be a regular value, which can be done for arbitrarily small ϵ by Sard's theorem.) For convenience we let S_ϵ be the boundary of U_ϵ, that is the set of points z such that $|f(z)| = \epsilon$.

Since $\log |f|$ is locally in $\mathscr{L}^1$, it follows that

$$\int_{U_\epsilon} \log |f| \, dd^c\psi = \lim_{\epsilon \to 0} \int_{U_\epsilon} \log |f| \, dd^c\psi.$$

Using the trivial identity

$$d(\log |f| \, d^c\psi) = d \log |f| \wedge d^c\psi + \log |f| \, dd^c\psi,$$

we conclude by Stokes' theorem that this limit is equal to

$$\lim_{\epsilon \to 0} \left[\int_{S_\epsilon} \log |f| \, d^c\psi - \int_{U_\epsilon} d \log |f| \wedge d^c\psi \right].$$

The first integral under the limit sign approaches 0. Indeed, we may assume hat $f(z) = z_n = re^{i\theta}$. On S_ϵ we have $|f(z)| = \epsilon$, so $\log |f| = \log \epsilon$.

There exist forms ψ_1, ψ_2 in the first $n-1$ variables such that

$$d^c\psi = \psi_1 \wedge dz_n + \psi_2 \wedge d\bar{z}_n,$$

and the restriction of dz_n to S_ϵ is equal to

$$\epsilon i e^{i\theta} \, d\theta,$$

with a similar expression for $d\bar{z}_n$. Hence our boundary integral is of type

$$\epsilon \log \epsilon \int_{S_\epsilon} \omega,$$

where ω is a bounded form. From this it is clear that the limit is 0.

Now we compute the second integral. Since ψ is assumed to be of type $(n-1, n-1)$ it follows that for any function g,

$$\partial g \wedge \partial\psi = 0 \qquad \text{and} \qquad \bar{\partial} g \wedge \bar{\partial}\psi = 0.$$

Replacing d and d^c by their values in terms of ∂ and $\bar{\partial}$, it follows that

$$-\int_{U_\epsilon} d \log |f| \wedge d^c\psi = \int_{U_\epsilon} d^c \log |f| \wedge d\psi.$$

We have

$$d(d^c \log |f| \wedge \psi) = dd^c \log |f| \wedge \psi - d^c \log |f| \wedge d\psi.$$

Furthermore dd^c is a constant times $\partial\bar{\partial}$, and $dd^c \log |f|^2 = 0$ in any open set where $f \neq 0$, because

$$\partial\bar{\partial} \log |f|^2 = \partial\bar{\partial}(\log f + \log \bar{f}) = 0$$

since $\partial \log \bar{f} = 0$ and $\bar{\partial} \log f = 0$ by the local analyticity of $\log f$. Hence we obtain the following values for the second integral by Stokes:

$$\int_{U_\epsilon} d^c \log |f|^2 \wedge d\psi = \int_{S_\epsilon} d^c \log |f|^2 \wedge \psi.$$

Since

$$d^c \log |f|^2 = -\frac{i}{4\pi}(\partial - \bar{\partial})(\log f + \log \bar{f})$$

$$= -\frac{i}{4\pi}\left(\frac{dz_n}{z_n} - \frac{d\bar{z}_n}{\bar{z}_n}\right)$$

(always assuming $f(z) = z_n$), we conclude that if $z_n = re^{i\theta}$, then the restriction of $d^c \log |f|^2$ to S_ϵ is given by

$$\text{res}_{S_\epsilon} d^c \log f = \frac{d\theta}{2\pi}.$$

Now write ψ in the form

$$\psi = \psi_1 + \psi_2$$

where ψ_1 contains only dz_j, $d\bar{z}_j$ for $j = 1, \ldots, n-1$ and ψ_2 contains dz_n or $d\bar{z}_n$. Then the restriction of ψ_2 to S_ϵ contains $d\theta$, and consequently

$$\int_{S_\epsilon} d^c \log |f|^2 \wedge \psi = \int_{S_\epsilon} \frac{d\theta}{2\pi} \wedge (\psi_1 \mid S_\epsilon).$$

The integral over S_ϵ decomposes into a product integral, we respect to the first $n-1$ variables, and with respect to $d\theta$. Let

$$\int^{(n-1)} \psi_1(z) \mid S_\epsilon = g(z_n).$$

Then simply by the continuity of g we get

$$\lim_{\epsilon \to 0} \frac{1}{2\pi} \int_0^{2\pi} g(\epsilon e^{i\theta})\, d\theta = g(0).$$

Hence

$$\lim_{\epsilon \to 0} \int_{S_\epsilon} \frac{d\theta}{2\pi} \wedge (\psi_1 \mid S_\epsilon) = \int_{z_n = 0} \psi_1.$$

But the restriction of ψ_1 to the set $z_n = 0$ (which is precisely V) is the same as the restriction of ψ to V. This proves the residue theorem.

APPENDIX

The Spectral Theorem

The following is a set of notes from a seminar of Von Neumann around 1950.

APP., §1. HILBERT SPACE

Let **E** be a vector space over **C** (The real theory follows exactly the same pattern.) By an **inner product** on **E** we mean an **R**-bilinear pairing $\langle x, y \rangle \in \mathbf{C}$ of $\mathbf{E} \times \mathbf{E}$ into **C** such that, for all complex numbers α, we have:

$$\langle \alpha x, y \rangle = \alpha \langle x, y \rangle, \qquad \langle x, y \rangle = \overline{\langle y, x \rangle},$$

$\langle x, x \rangle \geqq 0$ and equals 0 if and only if $x = 0$.

We have the **Schwartz inequality**:

$$|\langle x, y \rangle|^2 \leqq \langle x, x \rangle \langle y, y \rangle$$

whose proof is as follows. For all α, β complex,

$$0 \leqq \langle \alpha x + \beta y, \alpha x + \beta y \rangle = \alpha\bar{\alpha}\langle x, x \rangle + \beta\bar{\alpha}\langle x, y \rangle + \alpha\bar{\beta}\langle x, y \rangle + \beta\bar{\beta}\langle y, y \rangle.$$

We let $\alpha = \langle y, y \rangle$ and $\beta = -\langle x, y \rangle$. The inequality drops out.

We define the **norm** of a vector x to be $\langle x, x \rangle^{1/2}$ and denote it by $|x|$. Using the Schwartz inequality, one sees that $|x|$ defines a metric on **E**, the distance between x and y being $|x - y|$. The norm is continuous.

We write $x \perp y$ and say that x is **perpendicular** to y if $\langle x, y \rangle = 0$.

The following identities are useful and trivially proved.

Parallelogram Law. $|x+y|^2 + |x-y|^2 = 2|x|^2 + 2|y|^2$.

Pythagoras Theorem. *If* $x \perp y$, *then* $|x+y|^2 = |x|^2 + |y|^2$.

A **Hilbert space** is an inner product space which is complete under the induced metric. For the rest of this appendix, a **subspace** will always mean a closed subspace, with its structure of Hilbert space induced by that of **E**.

Lemma 1.1. *Let* **F** *be a subspace of* **E**, *let* $x \in \mathbf{E}$, *and let*

$$a = \inf|x - y|$$

the inf *taken over all* $y \in \mathbf{F}$. *Then there exists an element* $y_0 \in \mathbf{F}$ *such that* $a = |x - y_0|$.

Proof. Let y_n be a sequence in **F** such that $|y_n - x|$ tends to a. We must show that y_n is Cauchy. By the parallelogram law,

$$\begin{aligned} |y_n - y_m|^2 &= 2|y_n - x|^2 + 2|y_m - x|^2 - 4|\tfrac{1}{2}(y_n + y_m) - x|^2 \\ &\leqq 2|y_n - x|^2 + 2|y_m - x|^2 - 4a^2 \end{aligned}$$

which shows that y_n is Cauchy, converging to some vector y_0. The lemma follows by continuity.

Theorem 1.2. *If* **F** *is a subspace properly contained in* **E**, *then there exists a vector* z *in* **E** *which is perpendicular to* **F** (*and* $\neq 0$).

Proof. Let $x \in \mathbf{E}$ and $x \notin \mathbf{F}$. Let y_0 be an element of **F** which is at minimal distance from x (use Lemma 1.1). Let a be this distance and let $z = y_0 - x$. After a translation, we may assume that $z = x$, so that $|x| = a$. For any complex number α and $y \in \mathbf{F}$ we have $|x + \alpha y| \geqq a$, whence

$$\begin{aligned} \langle x + \alpha y, x + \alpha y\rangle &= |x|^2 + \bar{\alpha}\langle x, y\rangle + \alpha\overline{\langle x, y\rangle} + \alpha\bar{\alpha}|y|^2 \\ &\geqq a^2. \end{aligned}$$

Put $\alpha = \overline{t\langle x, y\rangle}$. We get a contradiction for small values of t.

APP., §2. FUNCTIONALS AND OPERATORS

A linear map A from a Hilbert space **E** to a Hilbert space **H** is **bounded** if there exists a positive real number α such that

$$|Ax| \leqq \alpha|x|$$

for all $x \in \mathbf{E}$. The norm of A, denoted by $|A|$ is the inf of all such α.

Proposition 2.1. *A linear map is bounded if and only if it maps the unit sphere on a bounded subset, if and only if it is continuous.*

Proof. Clear.

A **functional** is a continuous linear map into **C**. Functionals are bounded. We have the fundamental:

Representation Theorem. *A linear map* λ: $\mathbf{E} \to \mathbf{C}$ *is bounded if and only if there exists* $y \in \mathbf{E}$ *such that* $\lambda(x) = \langle x, y\rangle$ *for all* $x \in \mathbf{E}$. *If such a* y *exists, it is unique.*

Proof. If $\lambda(x) = \langle x, y\rangle$ then the Schwartz inequality shows that it is bounded, with bound $|y|$. It is obvious that y is unique.

Conversely, let λ be bounded. Let **F** be the kernel of λ. Then **F** is a subspace. If $\mathbf{E} = \mathbf{F}$ then everything is trivial. If $\mathbf{E} \neq \mathbf{F}$, then there exists $z \in \mathbf{E}$, $z \notin \mathbf{F}$ such that z is perpendicular to **F** by Theorem 1.2. We contend that some multiple $y = \alpha z$ does it. A necessary condition on α is that

$$\langle z, \alpha z\rangle = \bar{\alpha}|z|^2.$$

This is also sufficient. Namely, $x - (\lambda(x)/\lambda(z))z$ lies in **F**. Put $\alpha = \lambda(z)/|z|^2$. Then one sees at once that $\lambda(x) = \langle x, y\rangle$ as was to be shown.

By an **operator** we shall always mean a continuous linear map of a space into itself. It is straightforward to show that operators form a Banach space, and in fact a normed ring. In other words, in addition to the Banach space property, we have

$$|AB| \leqq |A|\,|B|.$$

We adopt the convention that a ring also has a unit element, which the algebra of operators does have. A **Banach algebra** is a Banach space, with a bilinear multiplication which is continuous. In our examples, it will also be a normed ring.

Proposition 2.2. *If* A *is an operator and* $\langle Ax, x\rangle = 0$ *for all* x, *then* $A = O$.

Proof. This follows from the polarization identity,

$$\langle A(x+y), (x+y)\rangle - \langle A(x-y), (x-y)\rangle = 2[\langle Ax, y\rangle + \langle Ay, x\rangle].$$

Replace x by ix. Then we get

$$\langle Ax, y\rangle + \langle Ay, x\rangle = 0,$$
$$i\langle Ax, y\rangle - i\langle Ay, x\rangle = 0,$$

for all x, y whence $\langle Ax, y\rangle = 0$ and $A = O$.

The above proposition is valid only in the complex case.

In the real case, we shall need it only when A is symmetric (see below), in which case it is equally clear. A similar remark applies to the next result.

Lemma 2.3. *Let A be an operator, and c a number such that*

$$|\langle Ax, x\rangle| \leqq c|x|^2$$

for all $x \in \mathbf{E}$. Then for all x, y we have

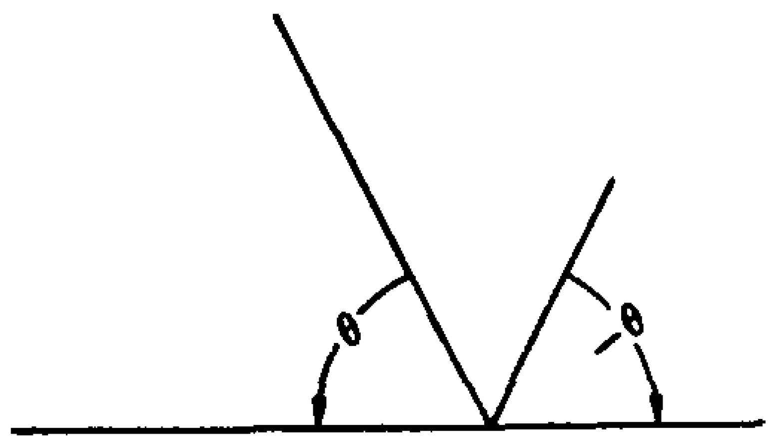
$$|\langle Ax, y\rangle| + |\langle x, Ay\rangle| \leqq 2c|x|\,|y|.$$

Proof. By the polarization identity,

$$2|\langle Ax, y\rangle + \langle Ay, x\rangle| \leqq c|x+y|^2 + c|x-y|^2 = 2c(|x|^2 + |y|^2).$$

Hence

$$|\langle Ax, y\rangle + \langle Ay, x\rangle| \leqq c(|x|^2 + |y|^2).$$



We multiply y by $e^{i\theta}$ and thus get on the left-hand side

$$|e^{-i\theta}\langle Ax, y\rangle + e^{i\theta}\langle Ay, x\rangle|.$$

The right-hand side remains unchanged, and for suitable θ, the left-hand side becomes

$$|\langle Ax, y\rangle| + |\langle Ay, x\rangle|.$$

(In other words, we are lining up two complex numbers by rotating one by θ and the other by $-\theta$.) Next we replace x by tx and y by y/t for t real and $t > 0$. Then the left-hand side remains unchanged, while the right-hand side becomes

$$g(t) = t^2|x|^2 + \frac{1}{t^2}|y|^2.$$

The point at which $g'(t) = 0$ is the unique minimum, and at this point t_0 we find that

$$g(t_0) = |x|\,|y|.$$

This proves our lemma.

In our applications, we need the lemma only when A is self-adjoint (i.e. symmetric, see below), in which case it is even more trivial.

For fixed y, the function of x given by $\langle Ax, y\rangle$ is a functional (bounded because of the Schwartz inequality). Hence by the representation theorem, there exists an element y^* such that $\langle Ax, y\rangle = \langle x, y^*\rangle$ for all x. We define A^*, the **adjoint** of A, by letting $A^*y = y^*$. Since y^* is unique, we see that A^* is the unique operator such that

$$\langle Ax, y\rangle = \langle x, A^*y\rangle$$

for all x, y in $\mathbf{E}$.

Theorem 2.4. *We have:*

$$(A+B)^* = A^* + B^*, \qquad A^{**} = A,$$
$$(\alpha A)^* = \bar{\alpha}A^*, \qquad |A^*| = |A|,$$
$$(AB)^* = B^*A^*, \qquad |AA^*| = |A|^2.$$

and the mapping $A \mapsto A^*$ is continuous.

Proof. Exercise for the reader.

APP., §3. HERMITIAN OPERATORS

We shall say that an operator A is **symmetric** (or **hermitian**) if $A = A^*$.

Proposition 3.1. *A is hermitian if and only if $\langle Ax, x\rangle$ is real for all x.*

Proof. Let A be hermitian. Then $\overline{\langle Ax, x\rangle} = \overline{\langle x, Ax\rangle} = \langle Ax, x\rangle$. Conversely, $\langle Ax, x\rangle = \overline{\langle Ax, x\rangle} = \langle x, Ax\rangle = \langle A^*x, x\rangle$ implies that

$$\langle (A - A^*)x, x\rangle = 0$$

whence $A = A^*$ by polarization.

Proposition 3.2. *Let A be a hermitian operator. Then $|A|$ is the greatest lower bound of all values c such that*

$$|\langle Ax, x\rangle| \leqq c|x|^2$$

for all x, or equivalently, the sup of all values $|\langle Ax, x\rangle|$ taken for x on the unit sphere in E.

Proof. When A is hermitian we obtain

$$|\langle Ax, y\rangle| \leqq c|x|\,|y|$$

for all $x, y \in E$, so that we get $|A| \leqq c$ in Lemma 2.3. On the other hand, $c = |A|$ is certainly a possible value for c by the Schwartz inequality. This proves our proposition.

Proposition 3.2 allows us to define an ordering in the space of hermitian operators. If A is hermitian, we define $A \geqq O$ and say that A is **semipositive** if $\langle Ax, x\rangle \geqq 0$ for all $x \in E$. If A, B are hermitian we define $A \geqq B$ if $A - B \geqq O$. This is indeed an ordering; the usual rules hold: If $A_1 \geqq B_1$ and $A_2 \geqq B_2$, then

$$A_1 + A_2 \geqq B_1 + B_2.$$

If c is a real number $\geqq 0$ and $A \geqq O$, then $cA \geqq O$. So far, however, we say nothing about a product of semipositive hermitian operators AB, even if $AB = BA$. We shall deal with this question later.

Let c be a bound for A. Then $|\langle Ax, x\rangle| \leqq c|x|^2$ and consequently

$$-cI \leqq A \leqq cI.$$

If we let

$$\alpha = \inf_{|x|=1} \langle Ax, x\rangle \quad \text{and} \quad \beta = \sup_{|x|=1} \langle Ax, x\rangle,$$

then we have

$$\alpha I \leqq A \leqq \beta I,$$

and from Proposition 3.1,

$$|A| = \max(|\alpha|, |\beta|).$$

Let p be a polynomial with real coefficients, and let A be a hermitian operator. Write

$$p(t) = a_n t^n + \cdots + a_0.$$

We define

$$p(A) = a_n A^n + \cdots + a_0 I.$$

We let $\mathbf{R}[A]$ be the algebra generated over $\mathbf{R}$ by A, that is the algebra of all operators $p(A)$, where $p(t) \in \mathbf{R}[t]$. We wish to investigate the closure of $\mathbf{R}[A]$ in the (real) Banach space of all operators. We shall show how to represent this closure as a ring of continuous functions on some compact subset of the reals. First, we observe that the hermitian operators form a closed subspace of $L(\mathbf{E}, \mathbf{E})$, and that $\overline{\mathbf{R}[A]}$ is a closed subspace of the space of hermitian operators.

We can find real numbers α, β such that

$$\alpha I \leqq A \leqq \beta I.$$

We shall prove that if p is a real polynomial which takes on values $\geqq 0$ on the interval $[\alpha, \beta]$, then $p(A)$ is a semipositive operator.

The fundamental theorem is the following.

Theorem 3.3. *Let α, β be real and $\alpha I \leqq A \leqq \beta I$. Let p be a real polynomial, semipositive in the interval $\alpha \leqq t \leqq \beta$. Then $p(A)$ is a semipositive operator.*

Proof. We shall need the following obvious facts.

If A, B are hermitian, A commutes with B, and $A \geqq O$, then AB^2 is semipositive.

If $p(t)$ is quadratic, of type $p(t) = t^2 + at + b$ and has imaginary roots, then

$$p(t) = \left(t + \frac{a}{2}\right)^2 + \left(b - \frac{a^2}{4}\right)$$

is a sum of squares.

A sum of squares times a sum of squares is a sum of squares (if they commute).

If $p(t)$ has a root γ in our interval, then the multiplicity of γ is even.

Our theorem now follows from the following purely algebraic statement.

Let $\alpha \leqq t \leqq \beta$ be a real interval, and $p(t)$ a real polynomial which is semipositive in this interval. Then $p(t)$ can be written:

$$p(t) = c\left[\sum Q_i^2 + \sum (t-\alpha)Q_j^2 + \sum (\beta - t)Q_k^2\right]$$

where Q^2 just denotes the square of some polynomial and c is a number $\geqq 0$.

In order to prove this, we split $p(t)$ over the real numbers into linear and quadratic factors. If a root γ is $\leqq \alpha$, then we write

$$(t-\gamma) = (t-\alpha) + (\alpha-\gamma)$$

and note that $(\alpha-\gamma)$ is a square. If a root γ is $\geqq \beta$, then we write

$$(\gamma-t) = (\gamma-\beta) + (\beta-t)$$

with $(\gamma-\beta)$ a square. We can then write, after expanding out the factorization of $p(t)$,

$$p(t) = c\left[\sum Q_i^2 + \sum (t-\alpha)Q_j^2 + \sum (\beta-t)Q_k^2 + \sum (t-\alpha)(\beta-t)Q_l^2\right]$$

with some constant c and Q^2 standing for the square of some polynomial. Note that c is $\geqq 0$ since $p(t)$ is semipositive on the interval. Our last step reduces the bad last term to the preceding ones by means of the identity

$$(t-\alpha)(\beta-t) = \frac{(t-\alpha)^2(\beta-t) + (t-\alpha)(\beta-t)^2}{\beta-\alpha}.$$

Corollary 3.4. *Suppose that $\alpha I \leqq A \leqq \beta I$. If $a \leqq p(t) \leqq b$ in the interval, then*

$$aI \leqq p(A) \leqq bI.$$

If $p(t)$ is a real polynomial, we define as usual

$$\|p\| = \sup |p(t)|$$

with t ranging over the interval.

Corollary 3.5. *Let $\alpha I \leqq A \leqq \beta I$. Let $p(t)$ be a real polynomial. Then $|p(A)| \leqq \|p\|$.*

Proof. Let $q(t) = \|p\| \pm p(t)$. Then $q(t)$ is $\geqq 0$ on the interval. Hence $q(A) \geqq O$ and our assertion follows at once.

As usual, we consider the continuous functions on the interval as a Banach space. If f is any continuous function on the interval, then by the Weierstrass approximation theorem, we can find a sequence of polynomials $\{p_n\}$ approaching f uniformly on this interval. We define $f(A)$ as the limit of $p_n(A)$. From Corollary 3.5 we deduce that $\{p_n(A)\}$ is a Cauchy sequence, and that its limit does not depend on the choice of the sequence $\{p_n\}$. Furthermore, by continuity, our corollary generalizes to continuous functions, so that $|f(A)| \leqq \|f\|$.

We see that the map $f \mapsto f(A)$ is a continuous homomorphism from the Banach algebra of continuous functions on the interval into the closure of the subalgebra generated by A.

Proposition 3.6. *Let A be a semipositive operator. Then there exists an operator B in the closure of the algebra generated by A such that $B^2 = A$.*

Proof. The continuous function $t^{1/2}$ maps on $A^{1/2}$.

Corollary 3.7. *The product of two semipositive, commuting hermitian operators is again semipositive.*

Proof. Let A, C be hermitian and $AC = CA$. If B is as in Proposition 3.6, then

$$\langle ACx, x\rangle = \langle B^2Cx, x\rangle = \langle BCx, Bx\rangle = \langle CBx, Bx\rangle \geqq 0.$$

The kernel of our homomorphism from the continuous functions to the operators is a closed ideal. Its zeros form a closed set called the **spectrum** of A and denoted by $\sigma(A)$.

Lemma 3.8. *Let X be a compact set, R the ring of continuous functions on X, and $\mathfrak{a}$ a closed ideal of R, $\mathfrak{a} \neq R$. Let C be the closed set of zeros of $\mathfrak{a}$. Then C is not empty and if a function $f \in R$ vanishes on C, then $f \in \mathfrak{a}$.*

Proof. Given ϵ, let U be the open set where $|f| < \epsilon$. Then $X - U$ is closed. For each point $t \in X - U$ there exists a function $g \in \mathfrak{a}$ such that $g(t) \neq 0$ in a neighborhood of t. These neighborhoods cover $X - U$, and so does a finite number of them, with functions $g_1, \dots, g_r$. Let $g = g_1^2 + \cdots + g_r^2$. Then $g \in \mathfrak{a}$. Our function g has a minimum on $X - U$ and for n large, the function

$$f\frac{ng}{1+ng}$$

is close to f on $X - U$ and is $< \epsilon$ on U, which proves what we wanted.

We now redefine the norm of a continuous function f to be

$$\|f\|_A = \sup_{t\in\sigma(A)} |f(t)|.$$

Theorem 3.9. *The map*

$$f(t) \mapsto f(A)$$

induces a Banach-isomorphism (i.e. norm-preserving) of the Banach algebra of continuous functions on $\sigma(A)$ onto the closure of the algebra generated by A.

Proof. We have already proved that our map is an agebraic isomorphism and that $|f(A)| \leqq \|f\|_A$. In order to get the reverse inequality, we shall prove:

If $f(A) \geqq O$, then $f(t) \geqq 0$ on the spectrum of A. Indeed, if $f(c) < 0$ for some $c \in \sigma(A)$, we let $g(t)$ be a function which is 0 outside a small neighborhood of c, is $\geqq 0$ everywhere, and is > 0 at c. Then $g(A)$ and $g(A)f(A)$ are both $\geqq 0$ by Corollary 3.7. But $-g(t)f(t) \geqq 0$ gives $-g(A)f(A) \geqq O$ whence $g(A)f(A) = O$. Since $g(t)f(t)$ is not 0 on the spectrum of A, we get a contradiction.

Let now $s = |f(A)|$. Then $sI - f(A) \geqq O$ implies that $s - f(t) \geqq 0$, which proves the theorem.

From now on, the norm on continuous functions will refer to the spectrum. All that remains to do is identify our spectrum with what can be called the **general spectrum**, that is those complex values ξ such that $A - \xi$ is not invertible. (By invertible, we mean having an inverse which is an operator.)

Theorem 3.10. *The general spectrum is compact, and in fact, if ξ is in it, then $|\xi| \leqq |A|$. If A is hermitian, then the general spectrum is equal to $\sigma(A)$.*

Proof. The complement of the general spectrum is open, because if $A - \xi_0$ is invertible, and ξ is close to ξ_0, then $(A - \xi_0)^{-1}(A - \xi)$ is close to I, hence invertible, and hence $A - \xi$ is also invertible. Furthermore, if $\xi > |A|$, then $|A/\xi| < 1$ and hence $I - (A/\xi)$ is invertible (by the power series argument). So is $A - \xi$ and we are done. Finally, suppose that ξ is in the general spectrum. Then ξ is real. Otherwise, let

$$g(t) = (t - \xi)(t - \bar{\xi}).$$

Then $g(t) \neq 0$ on $\sigma(A)$ and $h(t) = 1/g(t)$ is its inverse. From this we see that $A - \xi$ is invertible.

Suppose ξ is not in the spectrum. Then $t - \xi$ is invertible and so is $A - \xi$.

Suppose ξ is in the spectrum. After a translation, we may suppose that

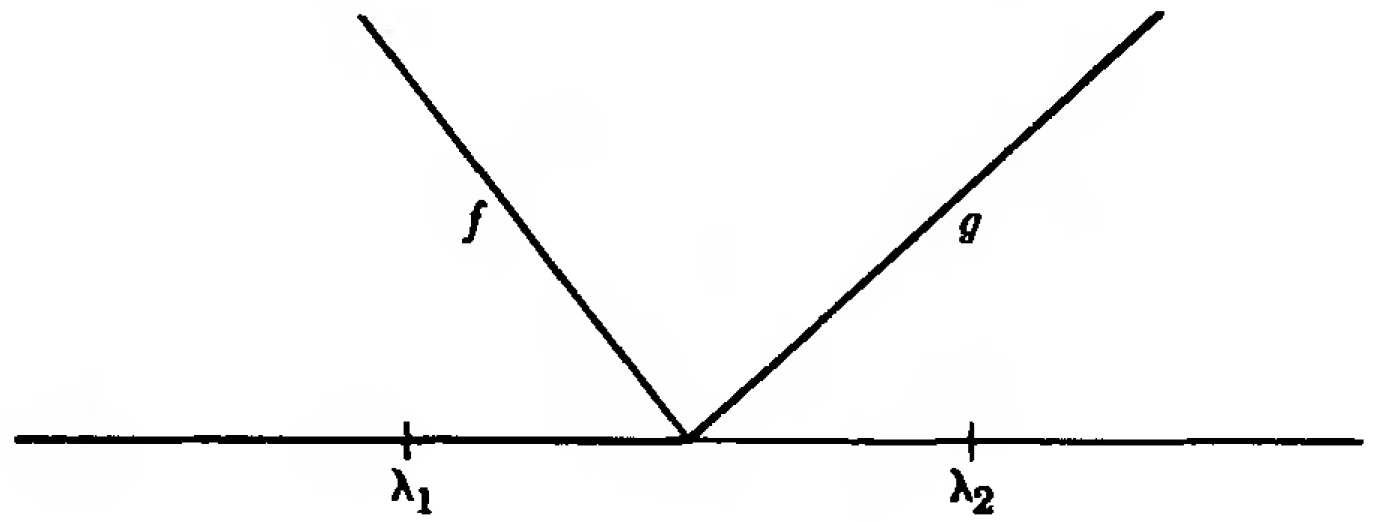
0 is in the spectrum. Consider the function $g(t)$ as follows:

$$g(t) = \begin{cases} 1/|t|, & |t| \geqq 1/N, \\ N, & |t| \leqq 1/N, \end{cases}$$

(g is positive and has a peak at 0.) If A is invertible, $BA = I$, then from $|tg(t)| \leqq 1$ we get $|Ag(A)| \leqq 1$ and hence $|g(A)| \leqq |B|$. But $g(A)$ becomes arbitrarily large as we take N large. Contradiction.

Theorem 3.11. *Let S be a set of operators of the Hilbert space $\mathbf{E}$, leaving no closed subspace invariant except 0 and $\mathbf{E}$ itself. Let A be a Hermitian operator such that $AB = BA$ for all $B \in S$. Then $A = \lambda I$ for some real number λ.*

Proof. It will suffice to prove that there is only one element in the spectrum of A. Suppose there are two, $\lambda_1 \neq \lambda_2$. There exist continuous functions f, g on the spectrum such that neither is 0 on the spectrum, but fg is 0 on the spectrum. For instance, one may take for f, g the functions whose graph is indicated on the next diagram.



We have $f(A)B = Bf(A)$ for all $B \in S$ (because B commutes with real polynomials in A, hence with their limits). Hence $f(A)\mathbf{E}$ is invariant under S because

$$Bf(A)\mathbf{E} = f(A)B\mathbf{E} \subset f(A)\mathbf{E}.$$

Let $\mathbf{F}$ be the closure of $f(A)\mathbf{E}$. Then $\mathbf{F} \neq 0$ because $f(A) \neq O$. Furthermore, $\mathbf{F} \neq \mathbf{E}$ because $g(A)f(A)\mathbf{E} = 0$ and hence $g(A)\mathbf{F} = 0$. Since $\mathbf{F}$ is obviously invariant under S, we have a contradiction.

Corollary 3.12. *Let S be a set of operators of the Hilbert space $\mathbf{E}$, leaving no closed subspace invariant except 0 and $\mathbf{E}$ itself. Let A be an operator such that $AA^* = A^*A$, $AB = BA$, and $A^*B = BA^*$ for all $B \in S$. Then $A = \lambda I$ for some complex number λ.*

Proof. Write $A = A_1 + iA_2$ where A_1, A_2 are hermitian and commute (e.g. $A_1 = (A + A^*)/2$). Apply the theorem to each one of A_1 and A_2 to get the result.

Bibliography

[Ab 62] R. ABRAHAM, *Lectures of Smale on Differential Topology*, Columbia University, 1962

[AbM 78] R. ABRAHAM and J. MARSDEN, *Foundations of Mechanics*, second edition, Benjamin-Cummings, 1978

[Am 56] W. AMBROSE, Parallel translation of Riemannian curvature, *Ann. of Math.* **64** (1965) pp. 337–363

[Am 60] W. AMBROSE, The Cartan structural equations in classical Riemannian geometry, *J. Indian Math. Soc.* **24** (1960) pp. 23–75

[Am 61] W. AMBROSE, The index theorem in Riemannian geometry, *Ann. of Math.* **73** (1961) pp. 49–86

[APS 60] W. AMBROSE, R.S. PALAIS, and I.M. SINGER, Sprays, *Acad. Brasileira de Ciencias* **32** (1960) pp. 163–178

[ABP 75] M. ATIYAH, R. BOTT, and V.K. PATODI, On the heat equation and the index theorem, *Invent. Math.* **19** (1973) pp. 279–330; Errata, *Invent. Math.* **28** (1975) pp. 277–280

[AH 59] M. ATIYAH and F. HIRZEBRUCH, Riemann–Roch theorems for differentiable manifolds, *Bull. Amer. Math. Soc.* **65** (1959) pp. 276–281

[APS 75] M. ATIYAH, V.K. PATODI, and I. SINGER, Spectral asymmetry and Riemannian geometry: I. *Math. Proc. Camb. Philos. Soc.* **77** (1975) pp. 43–69

[AtP 83] M. ATIYAH and A.N. PRESSLEY, Convexity and loop groups, in *Arithmetic and Geometry* (dedicated to Shafarevich, M. Artin and J. Tate, eds.), Progress in Math. **36**, Birkhauser 1983 pp. 33–63

[Ba 95] W. BALLMAN, *Lectures on Spaces of Nonpositive Curvature*, Birkhäuser, 1995

[BGS 85] W. BALLMAN, M. GROMOV, and V. SCHROEDER, *Manifolds of Non-*

positive Curvature, Birkhäuser, 1985

[Be 58] M. Berger, Sur certaines variétés Riemanniennes à courbure positive, *C.R. Acad. Sci. Paris* **247** (1958) pp. 1165–1168

[BGM 71] M. Berger, P. Gauduchon, and E. Mazet, *Le Spectre d'une Variété Riemannienne*, Lecture Notes in Mathematics **195**, Springer-Verlag, 1971

[BGV 92] N. Berline, E. Getzler, and M. Vergne, *Heat Kernels and Dirac Operators*, Grundlehren der Math. Wiss. **298**, Springer-Verlag, 1992

[BiC 64] R. Bishop and R. Crittenden, *Geometry of Manifolds*, Academic Press, 1964

[BoY 53] S. Bochner and K. Yano, *Curvature and Betti Numbers*, Annals Studies **32**, Princeton University Press, 1953

[BoF 65] R. Bonic and J. Frampton, Differentiable functions on certain Banach spaces, *Bull. Amer. Math. Soc.* **71** (1965) pp. 393–395

[BoF 66] R. Bonic and J. Frampton, Smooth functions on Banach manifolds, *J. Math. Mech.* **15** No. 5 (1966) pp. 877–898

[Bo 60] R. Bott, *Morse Theory and its Applications to Homotopy Theory*, Lecture Notes by van de Ven, Bonn, 1960

[BoS 58] R. Bott and H. Samelson, Applications of the theory of Morse to symmetric spaces, *Amer. J. Math.* **80** (1958) pp. 964–1029

[BoT 82] R. Bott and L. Tu, *Differential Forms in Algebraic Topology*, Graduate Texts in Mathematics **82**, Springer-Verlag, 1982

[Bou 68] N. Bourbaki, *General Topology*, Addison-Wesley, 1968 (translated from the French, 1949)

[Bou 69] N. Bourbaki, *Fasicule de Résultats des Variétés*, Hermann, 1969

[Bou 81] N. Bourbaki, *Espaces Vectoriels Topologiques*, Masson, 1981

[BrH ??] M. Bridson and A. Haefliger, Metric spaces of non-positive curvature, in preparation, 1996

[Bro 89] K. Brown, *Buildings*, Springer-Verlag, 1989

[BrT 72] F. Bruhat and J. Tits, Groupes Réductifs sur un Corps Local I, *Pub. IHES* **41** (1972) pp. 5–251

[BuK 81] P. Buser and H. Karcher, Gromov's almost flat manifolds, *Asterisque* **81** (1981), see especially #7 on Lie groups

[Bu 96] S.V. Buyalo, Catching geodesics in Hadamard spaces, preprint, 1996

[Ca 27a] E. Cartan, Sur une classe remarquable d'espaces de Riemann, *Bull. Soc. Math. France* **54** (1927) pp. 114–134

[Ca 27b] E. Cartan, Sur certaines formes Riemanniennes remarquables des géometries à groupe fondamental simple, *Ann. Sci. Ecole Norm. Sup.* **44** (1927) pp. 345–467

[Ca 28] E. Cartan, *Leçons sur la Géometrie des Espaces de Riemann*, Gauthier-Villars, 1928; second edition, 1946

[Ca 28/46] E. Cartan, *Leçons sur la géometrie des espaces de Riemann*, Gauthiers-Villars, Paris, 1928; second edition, 1946

[Ca 51] E. Cartan, *Leçons sur la géometrie des espaces de Riemann* II, Gauthiers-Villars, Paris, 1951

[Cha 84] I. Chavel, *Eigenvalues in Riemannian Geometry*, Academic Press, 1984

[ChE 75] J. Cheeger and D. Ebin, *Comparison Theorems in Riemannian Geometry*, North-Holland, 1975

[Che 55] S. Chern, On curvature and characteristic classes of a Riemann manifold, *Abh. Math. Sem. Hamburg* **20** (1958) pp. 117–126

[doC 92] M.P. do Carmo, *Riemannian Geometry*, Birkhaüser, 1992

[Do 62] P. Dombrowski, On the geometry of the tangent bundle, *J. Reine Angew. Math.* **210** (1962) pp. 73–88

[Do 68] P. Dombrowski, Krümmungsgrössen gleichungsdefinierter Untermannigfaltigkeiten Riemannscher Mannigfaltigkeiten, *Math. Nachr.* **38** (1968) pp. 133–180

[Eb 70] D. Ebin, The manifold of Riemannian metrics, *Proc. Symp. Pure Math. XV* AMS **26** (1970) pp. 11–40; and *Bull. Amer. Math. Soc.* **4** (1970) pp. 1002–1004

[EbM 70] D.G. Ebin and J. Marsden, Groups of diffeomorphisms and the motion of an incompressible fluid, *Ann. of Math.* **92** (1970) pp. 102–163

[Ee 58] J. Eells, On the geometry of function spaces, *Symposium de Topologia Algebrica*, Mexico (1958) pp. 303–307

[Ee 59] J. Eells, On submanifolds of certain function spaces, *Proc. Nat. Acad. Sci. USA* **45** (1959) pp. 1520–1522

[Ee 61] J. Eells, Alexander–Pontrjagin duality in function spaces, *Proc. Symposia in Pure Math.* **3**, AMS (1961) pp. 109–129

[Ee 66] J. Eells, A setting for global analysis, *Bull. Amer. Math. Soc.* **72** (1966) pp. 751–807

[EeS 64] J. Eells and J. Sampson, Harmonic mappings of Riemannian manifolds, *Amer. J. Math.* **86** (1964) pp. 109–160

[Ei 26] L.P. Eisenhart, *Riemannian Geometry*, 1926; second edition, Princeton University Press, 1949

[El 67] H. Eliasson, Geometry of manifolds of maps, *J. Diff. Geom.* **1** (1967) pp. 169–194

[FrG 89] D. Fried and D. Groisser, The basic geometry of the manifold of Riemannian metrics and of its quotient by the diffeomorphism group, *Mich. Math. J.* **36** (1989) pp. 323–344

[GHL 87/93] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian Geometry*, second edition, Springer-Verlag, 1987, 1993 (corrected printing)

[Ga 87] F.P. Gardiner, *Teichmuller Theory and Quadratic Differentials*, Wiley–Interscience, 1987

[Ga 80] H. Garland, The arithmetic theory of loop groups, *Pub. IHES* **52** (1980) pp. 181–312

[Gi 95] P. Gilkey, *Invariance Theory, the Heat Equation and the Atiyah–Singer Index Theorem*, second edition, CRC Press, 1995

[God 58] R. Godement, *Topologie Algébrique et Théorie des Faisceaux*, Hermann, 1958

[Go 62] S.I. GOLDBERG, *Curvature and Homology*, Academic Press, 1962

[Gr 61] W. GRAEUB, Liesche Grupen und affin zusammenhangende Mannigfaltigkeiten, *Acta Math.* (1961) pp. 65–111

[Gray 73] A. GRAY, The volume of small geodesic balls, *Mich. J. Math.* **20** (1973) pp. 329–344

[GrV 79] A. GRAY and L. VANHECKE, Riemmanian geometry as determined by the volume of small geodesic balls, *Acta Math.* **142** (1979) pp. 157–198

[GriH 76] P. GRIFFITHS and J. HARRIS, *Principles of Algebraic Geometry*, Wiley, 1976

[GrKM 75] D. GROMOLL, W. KLINGENBERG, and W. MEYER, *Riemannsche Geometrie im Grössen*, Lecture Notes in Mathematics **55**, Springer-Verlag, 1975

[Gros 64] N. GROSSMAN, *Geodesics on Hilbert Manifolds*, Princeton University Press Notes, 1964 (referred to in McAlpin's thesis)

[Gu 91] P. GÜNTHER, Huygens' Principle and Hadamard's Conjecture, *Math. Intelligencer*, **13** No. 2 (1991) pp. 56–63

[Ha 86] J. HADAMARD, Les surfaces à courbures opposées et leur lignes géodesiques, *J. Math. Pures Appl.* **(5) 4** (1896) pp. 27–73

[He 61] S. HELGASON, Some remarks on the exponential mapping for an affine connections, *Math. Scand.* **9** (1961) pp. 129–146

[He 62] S. HELGASON, *Differential Geometry and Symmetric Spaces*, Academic Press, 1962

[He 72a] S. HELGASON, *Analysis on Lie Groups and Homogeneous Spaces*, Conf. Board Math. Sci. Series **14**, AMS, Providence, RI, 1972

[He 72b] S. HELGASON, A formula for the radial part of the Laplace–Beltrami operator, *J. Diff. Geom.* **6** (1972) pp. 411–419

[He 78] S. HELGASON, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, 1978

[He 84a] S. HELGASON, *Groups and Geometric Analysis*, Academic Press, 1984

[He 84b] S. HELGASON, *Wave Equations on Homogeneous Spaces*, Lecture Notes in Mathematics **1077**, Springer-Verlag 1984 pp. 254–287

[Hi 59] N. HICKS, A theorem on affine connexions, *Ill. J. Math.* **3** (1959) pp. 242–254

[HoR 31] H. HOPF and W. RINOW, Über den Begriff der vollständigen differentialgeometrischen Fläche, *Commentarii Math. Helv.* 3 (1931) pp. 209–225

[Ir 70] M.C. IRWIN, On the stable manifold theorem, *Bull. London Math. Soc.* (1970) pp. 68–70

[Ka 77] H. KARCHER, Riemannian center of mass and mollifier smoothing, *Comm. Pure App. Math.* **XXX** (1977) pp. 509–541

[Ka 89] H. KARCHER, Riemannian Comparison Constructions, Studies in Mathematics, Vol. 27, *Global Differential Geometry* (S.S. Chern ed.), Math. Ass. Amer. 1989 pp. 170–222

[Ke 55] J. KELLEY, *General Topology*, Van Nostrand, 1955

[Ker 83] S.P. KERCKHOFF, The Nielsen realization problem, *Ann. Math.* **117** (1983) pp. 235–265

[Ki 1891] W. KILLING, Über die Clifford-Kleinschen Raumformen *Math. Ann.* **39** (1891) pp. 257–278

[Kl 83/95] W. KLINGENBERG, *Riemannian Geometry*, de Gruyter, 1983; second edition 1995

[Ko 57] S. KOBAYASHI, Theory of connections, *Annali di Mat.* **43** (1957) pp. 119–194

[Ko 87] S. KOBAYASHI, *Differential Geometry of Complex Vector Bundles*, Iwanami Shoten and Princeton University Press, 1987

[KoY 66] S. KOBAYASHI and K. YANO, Prolongations of tensor fields and connections of tangent bundles I, *J. Math Soc. Japan* **18** (1966) pp. 194–210; II, pp. 236–246; III, **19** (1967) pp. 486–488

[KoN 63] S. KOBAYASHI and K. NOMIZU, *Foundations of Differential Geometry* I, Wiley, 1963 and 1969

[KoN 69] S. KOBAYASHI and K. NOMIZU, *Foundations of Differential Geometry* II, Wiley, 1969

[Kos 57] J.L. KOSZUL, *Variétés Kähleriennes*, Ist. Mat. Pura e Apl., Sao Paulo, 1957

[Kr 72] N. KRIKORIAN, Differentiable structure on function spaces, *Trans. Amer. Math. Soc.* **171** (1972) pp. 67–82

[La 61] S. LANG, Fonctions implicites et plongements Riemanniens, *Séminaire Bourbaki* 1961–62

[La 62] S. LANG, *Introduction to Differentiable Manifolds*, Addison-Wesley, 1962

[La 71] S. LANG, *Differential Manifolds*, Addison-Wesley, 1971; Springer-Verlag, 1985

[La 75] S. LANG, $SL_2(\mathbf{R})$, Addison-Wesley, 1975; reprinted by Springer-Verlag, 1985

[La 83] S. LANG, *Undergraduate Analysis*, Springer-Verlag, 1983; Second Edition, 1997

[La 87] S. LANG, *Introduction to Complex Hyperbolic Spaces*, Springer-Verlag, 1987

[La 93] S. LANG, *Real and Functional Analysis*, third edition, Springer-Verlag, 1993

[La 95] S. LANG, *Differential and Riemannian Manifolds*, Springer-Verlag, 1995

[Le 67] J. LESLIE, On a differential structure for the group of diffeomorphisms, *Topology* **6** (1967) pp. 263–271

[LoS 68] L. LOOMIS and S. STERNBERG, *Advanced Calculus*, Addison-Wesley, 1968

[Lo 69] O. LOOS, *Symmetric Spaces*, I and II, Benjamin, 1969

[vM 81] H. von MANGOLDT, Über diejenigen Punkte auf positiv gekrümmten Flächen, welche die Eigenschaft haben, dass die von ihnen ausgehenden geodätischen Linien nie aufhören, kurzeste Linien zu sein, *J. Reine Angew. Math.* **91** (1881) pp. 23–52

[Mar 74] J. MARSDEN, *Applications of Global Analysis to Mathematical Physics*, Publish or Perish, 1974; reprinted in "Berkeley Mathematics Department Lecture Note Series", available from the UCB Math. Dept.

[MaW 74] J. MARSDEN and A. WEINSTEIN, Reduction of symplectic manifolds with symmetry, *Rep. Math. Phys.* **5** (1974) pp. 121–130

[Maz 61] B. MAZUR, Stable equivalence of differentiable manifolds, *Bull. Amer. Math. Soc.* **67** (1961) pp. 377–384

[McA 65] J. MCALPIN, *Infinite Dimensional Manifolds and Morse Theory*, thesis, Columbia University, 1965

[McK-S 67] H.P. MCKEAN and I. SINGER, Curvature and the eigenvalues of the Laplacian, *J. Diff. Geom.* **1** (1967) pp. 43–69

[Mi 58] J. MILNOR, *Differential Topology*, Princeton University Press, 1968

[Mi 59] J. MILNOR, *Der Ring der Vectorraumbundel eines topologischen Raümes*, Bonn, 1959

[Mi 61] J. MILNOR, *Differentiable Structures*, Princeton University Press, 1961

[Mi 63] J. MILNOR, *Morse Theory*, Ann. Math. Studies **53**, Princeton University Press, 1963

[Mi 64] J. MILNOR, Eigenvalues of the Laplace operator on certain manifolds, *Proc. Nat. Acad. Sci. USA* **51** (1964) p. 542

[Min 49] S. MINAKSHISUNDARAM, A generalization of Epstein zeta functions, *Can. J. Math.* **1** (1949) pp. 320–329

[MinP 49] S. MINAKSHISUNDARAM and A. PLEIJEL, Some properties of the eigenfunctions of the Laplace operator on Riemannian manifolds, *Can. J. Math.* **1** (1949) pp. 242–256

[Mink 1884] H. MINKOWSKI, *Grundlagen für eine Theorie der quadratischen Formen mit ganzzahligen Koeffizienten*, Mémoire, Académie des Sciences, 1884, *Collected Works* I, pp. 3–144

[MokSY 93] N. MOK, Y-T. SIU, and S-K. YEUNG, Geometric superrigidity, *Invent. Math.* **113** (1993) pp. 57–83

[Mo 61] J. MOSER, A new technique for the construction of solutions for nonlinear differential equations, *Proc. Nat. Acad. Sci. USA* **47** (1961) pp. 1824–1831

[Mo 65] J. MOSER, On the volume element of a manifold, *Trans. Amer. Math. Soc.* **120** (1965) pp. 286–294

[Mo 53] D. MOSTOW, *Some new decomposition theorems for semisimple groups*, Memoirs AMS, 1953

[Nar 68] R. NARASIMHAN, *Analysis on Real and Complex Manifolds*, North-Holland, 1968

[Nas 56] J. NASH, The embedding problem for Riemannian manifolds, *Ann. of Math.* **63** (1956) pp. 20–63

[Om 70] H. OMORI, On the group of diffeomorphisms of a compact manifold, *Proc. Symp. Pure Math. XV*, AMS (1970) pp. 167–184

[O'N 66] B. O'NEILL, The fundamental equations of a submersion, *Mich. Math. J.* **13** (1966) pp. 459–469

[O'N 83] B. O'NEILL, *Elementary Differential Geometry*, Academic Press, 1966

[O'N 97] B. O'NEILL, *Elementary Differential Geometry*, second edition, Academic Press, 1997

[Pa 63] R. PALAIS, Morse theory on Hilbert maifolds, *Topology* **2** (1963) pp. 299–340

[Pa 66] R. PALAIS, Lusternik–Schnirelman theory on Banach manifolds, *Topology* **5** (1966) pp. 115–132

[Pa 68] R. PALAIS, *Foundations of Global Analysis*, Benjamin, 1968

[Pa 69] R. PALAIS, The Morse lemma on Banach spaces, *Bull. Amer. Math. Soc.* **75** (1969) pp. 968–971

[Pa 70] R. PALAIS, Critical point theory and the minimax principle, *Proc. Symp. Pure Math. AMS* **XV** (1970) pp. 185–212

[PaS 64] R. PALAIS and S. SMALE, A generalized Morse theory, *Bull. Amer. Math. Soc.* **70** (1964) pp. 165–172

[PaT 88] R. PALAIS and C. TERNG, *Critical Point Theory and Submanifold Geometry*, Lecture Notes Mathematics **1353**, Springer-Verlag, 1988

[Pat 71a] V.K. PATODI, Curvature and the eigenforms of the Laplace operator, *J. Diff. Geom.* **5** (1971) pp. 233–249

[Pat 71b] V.K. PATODI, An analytic proof of the Riemann–Roch–Hirzebruch theorem, *J. Diff. Geom.* **5** (1971) pp. 251–283

[Po 62] F. POHL, Differential geometry of higher order, *Topology* **1** (1962) pp. 169–211

[Pro 61] *Proceedings of Symposia in Pure Mathematics III, Differential Geometry, AMS*, 1961

[Pro 70] *Proceedings of the Conference on Global Analysis, Berkeley, Calif. 1968*, AMS, 1970

[Ra 51] H. RAUCH, A contribution to differential geometry in the large, *Ann. of Math.* **34** (1951) pp. 38–55

[RaT 91] T. RATIU and A. TODOROV, An infinite-dimensional point of view on the Weil–Petersson metric, *Proc. Symposia in Pure Math.* Vol. 52 (1991), Part 2, pp. 467–476

[Re 64] G. RESTREPO, Differentiable norms in Banach spaces, *Bull. Amer. Math. Soc.* **70** (1964) pp. 413–414

[Ro 68] J. ROBBIN, On the existence theorem for differential equations, *Proc. Amer. Math. Soc.* **19** (1968) pp. 1005–1006

[Sa 96] T. SAKAI, *Riemannian Geometry*, AMS Translations of Mathematical Monographs #149, 1996

[Sas 58] S. SASAKI, On the differential geometry of tangent bundles of Riemannian manifolds, *Tohoku Math. J.* **10** (1958) pp. 338–354

[ScTZ 96] M. SCHOENBEK, A. TODOROV, and G. ZUBELLI, Grassmanians, geodesic flows and the KdV equation, preprint, paper in preparation

[Sc 60] J. SCHWARZ, On Nash's implicit functional theorem, *Commun. Pure Appl. Math.* **13** (1960) pp. 509–530

[Sm 61] S. SMALE, Generalized Poincaré's conjecture in dimensions greater than four, *Ann. of Math.* **74** (1964) pp. 391–406

[Sm 63] S. SMALE, Stable manifolds for differential equations and diffeomorphism, *Ann. Scuola Normale Sup. Pisa* **III** Vol. XVII (1963) pp. 97–116

[Sm 64] S. SMALE, Morse theory and a non-linear generalization of the Dirichlet problem, *Ann. of Math.* **80** No. 2 (1964) pp. 382–396

[Sm 67] S. SMALE, Differentiable dynamical systems, *Bull. Amer. Math. Soc.* **73**, No. 6 (1967) pp. 747–817

[Sp 70/79] M. SPIVAK, *Differential Geometry* (5 volumes), Publish or Perish, 1970–1979

[Ste 51] N. STEENROD, *The Topology of Fiber Fundles*, Princeton University Press, 1951

[Str 64] S. STERNBERG, *Lectures on Differential Geometry*, Prentice-Hall, 1964

[Sy 36] J.L. SYNGE, On the connectivity of spaces of positive curvature, *Quart. J. Math. Oxford Series* **7** (1936) pp. 316–320

[To 89] A. TODOROV, The Weil–Petersson geometry of the moduli space of SU($n \geqq 3$) (Calabi–Yau) manifolds I, *Comm. Math. Physics* **126** (1989) pp. 325–346

[To 97] A. TODOROV, Global Properties of the Moduli of Calabi–Yau Manifolds II (Teichmuller theory), preprint, 1997

[Tr 92] A. TROMBA, *Teichmuller theory in Riemannian Geometry*, Lectures in Mathematics, ETH Zurich, Birkhauser, 1992

[Vi 73] L. VIRTANEN, *Quasi Conformal Mappings in the Plane*, Springer-Verlag, 1973

[We 67] A. WEINSTEIN, A fixed point theorem for positively curved manifolds, *J. Math. Mech.* **18** (1968) pp. 149–153

[Wel 80] R. WELLS, *Differential Analysis on Complex Manifolds*, Graduate Texts in Mathematics **65**, Springer-Verlag, 1980

[Wey 18] H. WEYL, Reine infinitesimalgeometrie, *Math. Z.* **2** (1918) pp. 384–411

[Wh 32] J.H.C. WHITEHEAD, Convex regions in the geometry of paths, *Quaterly J. Math. (Oxford)* **3** (1932) pp. 33–42

[Wo 87] S.A. WOLPERT, Geodesic length functions and the Nielsen problem, *J. Diff. Geom.* **25** (1987) pp. 275–296

[Wu 65] H.-S. WU, Two remarks on sprays, *J. Math. Mech.* **14** No. 5 (1965) pp. 873–880

Index

A

B

C

D

E

F

G

H

I

J

K

L

M