INTRODUCTION TO BUNDLES, CONNECTIONS, METRICS AND CURVATURE

Clifford Henry Taubes
Department of Mathematics
Harvard University
Table of Contents

Prologue

Chapter 1: Smooth manifolds
   a) Smooth manifolds
   b) Inverse and implicit function theorem
   c) Submanifolds of \( \mathbb{R}^n \)
   d) Submanifolds of manifolds
   e) More constructions of manifolds
   f) More smooth manifolds: The Grassmannians
Appendix 1: How to prove the inverse and implicit function theorems
Appendix 2: Partitions of unity
Additional reading

Chapter 2: Matrices and Lie groups
   a) The general linear group
   b) Lie groups
   c) Examples of Lie groups
   d) Some complex Lie groups
   e) The groups \( \text{Sl}(n; \mathbb{C}), \text{U}(n) \) and \( \text{SU}(n) \)
   f) Notation with regards to matrices and differentials
Appendix: The transition functions for the Grassmannians
Additional reading

Chapter 3: Introduction to vector bundles
   a) The definition
   b) The standard definition
   c) The first examples of vector bundles
   d) The tangent bundle
   e) Tangent bundle examples
   f) The cotangent bundle
   g) Bundle homomorphisms
   h) Sections of bundles
   i) Sections of TM and T*M
Additional reading

Chapter 4: Algebra of vector bundles
   a) Subbundles
b) Quotient bundles

Chapter 5: Maps and vector bundles

a) The pull-back construction
b) Pull-backs and Grassmannians
c) Pull-back of differential forms and push-forward of vector fields
d) Invariant forms and vector fields on Lie groups
e) The exponential map on a matrix group
f) The exponential map and right/left invariance on Gl(n; C) and its subgroups
g) Immersions, submersions and transversality

Chapter 6: Vector bundles with fiber C^n

a) Definitions
b) Comparing definitions
c) Examples: The complexification
d) Complex bundles over surfaces in \( \mathbb{R}^3 \)
e) The tangent bundle to a surface in \( \mathbb{R}^3 \)
f) Bundles over 4-dimensional manifolds in \( \mathbb{R}^5 \)
g) Complex bundles over 4-dimensional manifolds
h) The complex Grassmannians
i) The exterior product construction
j) Algebraic operations
k) Pull-back

Chapter 7: Metrics on vector bundles

a) Metrics and transition functions
b) Metrics and transition functions for complex bundles
c) Metrics, algebra and maps
d) Metrics on TM

Additional reading
Chapter 8: Geodesics
   a) Riemannian metrics and distance
   b) Length minimizing curves
   c) The existence of geodesics
   d) First examples
   e) Geodesics on SO(n)
   f) Geodesics on U(n) and SU(n)
   g) Geodesics and matrix groups
Appendix: The proof of the vector field theorem
Additional reading

Chapter 9: Properties of geodesics
   a) The maximal extension of a geodesic
   b) The exponential map
   c) Gaussian coordinates
   d) The proof of the geodesic theorem
Additional reading

Chapter 10: Principal bundles
   a) The definition
   b) A cocycle definition
   c) Principal bundles from vector bundles
   d) Quotients of Lie groups by subgroups
   e) Examples of Lie group quotients
   f) Cocycle construction examples
   g) Pull-backs of principal bundles
   h) Reducible principal bundles
   i) Associated vector bundles
Appendix: The proof of the proposition in Chapter 10d
Additional reading

Chapter 11: Covariant derivatives and connections
   a) Covariant derivatives
   b) The space of covariant derivatives
   c) Another construction of covariant derivatives
   d) Principal bundles and connections
   e) Connections and covariant derivatives
   f) Horizontal lifts
   g) An application to the classification of principal bundles
   h) Connections, covariant derivatives and pull-back
Additional reading

Chapter 12: Covariant derivatives, connections and curvature
   a) Exterior derivative
   b) Closed forms, exact forms, diffeomorphisms and De Rham cohomology
   c) Lie derivative
   d) Curvature and covariant derivatives
   e) An example
   f) Curvature and commutators
   g) Connections and curvature
   h) The horizontal subbundle revisited

Additional reading

Chapter 13: Flat connections and holonomy
   a) Flat connections
   b) Flat connections on bundles over the circle
   c) Foliations
   d) Automorphisms of a principal bundle
   e) The fundamental group
   f) The flat connections on bundles over M
   g) The universal covering space
   h) Holonomy and curvature
   i) The proof of the classification theorem for flat connections

Appendix 1: Smoothing maps
Appendix 2: The proof of Frobenius’ theorem

Additional reading

Chapter 14: Curvature polynomials and characteristic classes
   a) The Bianchi identity
   b) Characteristic forms
   c) Characteristic classes: Part 1
   d) Characteristic classes: Part 2
   e) Characteristic classes for complex vector bundles and the Chern classes
   f) Characteristic classes for real vector bundles and the Pontrjagin classes
   g) Examples of bundles with non-zero Chern classes
   h) The degree of the map \( g \to g^m \) from SU(2) to itself
   i) A Chern-Simons form

Appendix 1: The ad-invariant functions on \( \mathbb{M}(n; \mathbb{C}) \)
Appendix 2: Integration on manifolds
Appendix 3: The degree of a map
Additional reading

Chapter 15: *Covariant derivatives and metrics*
   a) Metric compatible covariant derivatives
   b) Torsion free covariant derivatives on T*M
   c) The Levi-Civita connection/covariant derivative
   d) A formula for the Levi-Civita connection
   e) Covariantly constant sections
   f) An example of the Levi-Civita connection
   g) The curvature of the Levi-Civita connection

Additional reading

Chapter 16: *The Riemann curvature tensor*
   a) Spherical metrics, flat metrics and hyperbolic metrics
   b) The Schwarzschild metric
   c) Curvature conditions
   d) Manifolds in dimension 2: The Gauss-Bonnet formula
   e) Metrics on manifolds of dimension 2
   f) Conformal changes
   g) Sectional curvatures and universal covering spaces
   h) The Jacobi field equation
   i) Constant sectional curvature and the Jacobi field equation
   j) Manifolds in dimension 3
   i) The Riemann curvature of a compact matrix group

Additional reading

Chapter 17: *Complex manifolds*
   a) Some basics concerning holomorphic functions on \( \mathbb{C}^n \)
   b) The definition of a complex manifold
   c) First examples of complex manifolds
   d) The Newlander-Nirenberg theorem
   e) Metrics and almost complex structures on TM
   f) The almost Kähler 2-form
   g) Symplectic forms
   h) Kähler manifolds
   i) Complex manifolds with closed almost Kähler form
   j) Examples of Kähler manifolds

Appendix: Compatible almost complex structures

Additional reading
Chapter 18: *Holomorphic submanifolds, holomorphic sections and curvature*

a) Holomorphic submanifolds of a complex manifold  
b) Holomorphic submanifolds of projective spaces  
c) Proof of the proposition about holomorphic submanifolds in $\mathbb{CP}^n$  
d) The curvature of a Kähler metric  
e) Curvature with no $(0, 2)$ part  
f) Holomorphic sections  
g) Examples on $\mathbb{CP}^n$  

Additional reading

Chapter 19: *The Hodge star*

a) Definition of the Hodge star  
b) Representatives of De Rham cohomology  
c) A faery tale  
d) The Hodge theorem  
e) Self-duality  

Additional reading

Indexed list of propositions by subject

Index
This book is meant to be an introduction to the subject of vector bundles, principal bundles, metrics (Riemannian and otherwise), covariant derivatives, connections and curvature. I am imagining an audience of first year graduate students or advanced undergraduate students who have some familiarity with the basics of linear algebra and with the notion of a smooth manifold. Even so, I start with a review of the latter subject. I have tried to make the presentation as much as possible self-contained with proofs of basic results presented in full. In particular, I have supplied proofs for almost all of the background material either in the text or in the chapter appendices. Even so, you will most likely have trouble if you are not accustomed to matrices with real and complex number entries, in particular the notions of an eigenvalue and eigenvector. You should also be comfortable working with multi-variable calculus. At the very end of each chapter is a very brief list of other books with parts that cover some of the chapter’s subject matter.

I have worked out many examples in the text, because truth be told, the subject is not interesting to me in the abstract. I for one need to feel the geometry to understand what is going on. In particular, I present in detail many of the foundational examples.

I learned much of the material that I present here from a true master, Raoul Bott. In particular, I put into this book the topics that I recall Raoul covering in his first semester, graduate differential geometry class. Although the choice of topics are those I recall Raoul covering, the often idiosyncratic points of view and the exposition are my own.
CHAPTER 1: Smooth manifolds

Said briefly, differential geometry studies various constructions that can be built on a smooth manifold. Granted this, the task for this first chapter is to summarize various facts about smooth manifolds that are used, either explicitly or implicitly, in much of what follows. These facts are offered sans proof since proofs can be had in most any textbook on differential topology; the my favorite being Differential Topology by Victor Guillemin and Alan Pollack (Prentice Hall 1974).

1a) Smooth manifolds

The Ur smooth manifold of dimension $n$ is the $n$-dimensional Euclidean space, $\mathbb{R}^n$. I think of $\mathbb{R}^n$ as coming with a standard set of coordinates, $(x_1, \ldots, x_n)$. The Euclidean space is the local model for all other manifolds.

Topological manifolds: A manifold of dimension $n$ is a paracompact, Hausdorff space such that each point has a neighborhood that is homeomorphic to $\mathbb{R}^n$ or, what amounts to the same thing, to the interior of a ball in $\mathbb{R}^n$. Such a neighborhood is called a coordinate chart and the pull-back of the coordinate functions from $\mathbb{R}^n$ are called local coordinates. A collection

$$\mathcal{U} = \{(U, \varphi): U \subset M \text{ is open and } \varphi: U \to \mathbb{R}^n \text{ is an embedding}\}$$

is said to be a coordinate atlas for $M$ if $M = \bigcup_{(U, \varphi) \in \mathcal{U}} U$. A coordinate atlas is said to be locally finite if it is the case that each point $p \in M$ is contained in but a finite collection of its open sets.

Two topological manifolds, $M$ and $M'$ are said to be homeomorphic if there is a homeomorphism between them. This is a continuous, 1-1, surjective map from $M$ to $M'$ whose inverse is continuous.

Smooth manifolds: Let $\mathcal{U}$ denote a coordinate atlas for $M$. Suppose that $(U, \varphi_U)$ and $(U', \varphi_{U'})$ are two elements from $\mathcal{U}$. The map

$$\varphi_{U'} \circ \varphi_U^{-1}: \varphi_U(U' \cap U) \to \varphi_{U'}(U' \cap U)$$

is a homeomorphism between two open subsets of Euclidean space. This map is said to be the coordinate transition function for the pair of charts $(U, \varphi_U)$ and $(U', \varphi_{U'})$.

A smooth structure on $M$ is defined by an equivalence class of coordinate atlases with the following property: All transition functions are diffeomorphisms. This is to say that they have partial derivatives to all orders, as do their inverses. Coordinate atlases $\mathcal{U}$
and \( V \) are deemed to be equivalent when the following condition holds: Given any pairs \((U, \varphi_U) \in \mathcal{U}\) and \((V, \varphi_V) \in \mathcal{V}\), the compositions

\[
\varphi_U \circ \varphi_V^{-1} : \varphi_V(V \cap U) \to \varphi_U(V \cap U) \quad \text{and} \quad \varphi_V \circ \varphi_U^{-1} : \varphi_U(V \cap U) \to \varphi_V(V \cap U) \quad (\ast)
\]

are diffeomorphisms between open subsets of Euclidean space.

In what follows, a map with partial derivatives to all orders is said to be \textit{smooth}. A manifold \( M \) with a smooth structure is said to be a \textit{smooth} manifold. The point here is that one can do calculus on smooth manifolds for it makes sense to say that a function \( f : M \to \mathbb{R} \) is differentiable, or infinitely differentiable. The collection of smooth functions on \( M \) is denoted by \( C^\infty(M; \mathbb{R}) \) and consists of those functions with the following property: Let \( \mathcal{U} \) denote any given coordinate atlas from the equivalence class that defines the smooth structure. If \((U, \varphi) \in \mathcal{U}\), then \( f \circ \varphi^{-1} \) is a smooth function on \( \mathbb{R}^n \). This requirement defines the same set of smooth functions no matter the choice of representative atlas by virtue of the condition \((\ast)\) that defines the equivalence relation.

Maps between smooth manifolds: If \( M \) and \( N \) are smooth manifolds, then a map \( h : M \to N \) is said to be smooth if the following is true: Let \( \mathcal{U} \) denote a locally finite coordinate atlas from the equivalence class that gives the smooth structure to \( M \); and let \( \mathcal{V} \) denote a corresponding atlas for \( N \). Then each map in the collection

\[
\{\psi \circ h \circ \varphi^{-1} : h(U) \cap V \neq \emptyset\}_{(U, \varphi) \in \mathcal{U}, (V, \psi) \in \mathcal{V}}
\]

is infinitely differentiable as a map from one Euclidean space to another. Note again that the that the equivalence relation described by \((\ast)\) above guarantees that the definition of a smooth map depends only on the smooth structures of \( M \) and \( N \), but not on the chosen representative coordinate atlases.

Two smooth manifolds are said to be diffeomorphic when there exists a smooth homeomorphism \( h : M \to N \) with smooth inverse. This is to say that \( M \) is homeomorphic to \( N \) and that the given equivalence classes of atlases that define the respective smooth structures are one and the same.

1b) \textbf{The inverse function theorem and implicit function theorem}

Various constructions of smooth manifolds rely on two theorems from differential topology, these being the \textit{inverse function theorem} and the \textit{implicit function theorem}. What follows here is a digression to state these theorems. Their proofs are sketched in the first appendix to this chapter.

The statement of these theorems introduces the notion of the \textit{differential} of a map between one Euclidean space and another. Here is a reminder: Let \( U \subset \mathbb{R}^m \) denote a
given open set and let \( \psi: U \rightarrow \mathbb{R}^n \) denote a given map. The differential of \( \psi \) is denoted by \( \psi_* \); it is a matrix valued function on \( U \) with \( m \) columns and \( n \) rows whose entries come from the partial derivatives of the components of \( \psi \). This is to say that if \( \psi \) has components \((\psi_1, \ldots, \psi_n)\), then the entry in the \( i \)'th row and \( j \)'th column is \( \partial_j \psi_i = \frac{\partial \psi_i}{\partial x_j} \).

(Here and throughout, \( \partial_j \) is shorthand for the partial derivative with respect to the coordinate \( x_j \).)

The inverse function theorem: Let \( U \subset \mathbb{R}^m \) denote a neighborhood of the origin, and let \( \psi: U \rightarrow \mathbb{R}^m \) denote a smooth map. Suppose that \( p \in U \) and that the matrix \( \psi_* \), at \( p \) is invertible. Then there is a neighborhood \( V \subset \mathbb{R}^m \) of \( \psi(p) \) and a smooth map \( \sigma: V \rightarrow U \) such that \( \sigma(\psi(p)) = p \) and

- \( \sigma \circ \psi \) is the identity on some neighborhood \( U' \subset U \) of \( p \)
- \( \psi \circ \sigma \) is the identity on \( V \).

Conversely, if \( \psi_* \) is not invertible at \( p \), then there is no such set \( V \) and map \( \sigma \) with these properties.

The key conclusion here is that a given map has an inverse on a sufficiently small neighborhood of any given point if and only if its differential has an inverse at the point.

The statement of the implicit function theorem refers to a regular value of a map between Euclidean spaces. I remind you of the definition: Let \( U \subset \mathbb{R}^m \) denote a given open subset and let \( \psi: U \rightarrow \mathbb{R}^k \) denote a given smooth map. A point \( a \in \mathbb{R}^m \) is said to be a regular value of \( \psi \) if the matrix \( \psi_* \) is surjective at all points in \( \psi^{-1}(a) \).

The implicit function theorem: Fix non-negative integers \( m \geq n \). Suppose that \( U \subset \mathbb{R}^m \) is an open set, \( \psi: U \rightarrow \mathbb{R}^{m-n} \) is a smooth map, and \( a \in \mathbb{R}^{m-n} \) is a regular value of \( \psi \). Then \( \psi^{-1}(a) \subset U \) has the structure of a smooth, \( n \) dimensional manifold whose smooth structure is defined by coordinate charts of the following sort: Fix a point \( p \in \psi^{-1}(a) \). Then there is a ball \( B \subset \mathbb{R}^m \) centered at \( p \) such that the orthogonal projection from \( B \) to the kernel of \( \psi_*|_p \) restricts to \( \psi^{-1}(a) \cap B \) as a coordinate chart. In addition, there is a diffeomorphism, \( \varphi: B \rightarrow \mathbb{R}^m \) such that \( \varphi(B \cap \psi^{-1}(a)) \) is a neighborhood of the origin in the \( n \)-dimensional linear subspace of points \((x_1, \ldots, x_m)\) with \( x_{n+1} = \cdots = x_m = 0 \).

This theorem is, of course, vacuous if \( m < n \). By the way, this theorem is often used in conjunction with

Sard’s theorem: Suppose that \( U \subset \mathbb{R}^m \) is an open set, \( \psi: U \rightarrow \mathbb{R}^n \) is a smooth map.

Then the set of regular values of \( \psi \) have full measure.
This means the following: Fix any ball in $\mathbb{R}^n$. Delete the all non-regular values from this ball. The volume of the resulting set is the same as that of the ball. Thus, a randomly chosen point from the ball has probability 1 of being a regular value. (A special case of Sard’s theorem is proved in Appendix 2 of Chapter 14.)

1c) **Submanifolds of $\mathbb{R}^m$**

Let $f: \mathbb{R}^m \to \mathbb{R}$ denote a smooth function and let $a \in \mathbb{R}$ denote a regular value of $f$. Then $f^{-1}(a)$ is a smooth, (m-1)-dimensional manifold; this a consequence of the implicit function theorem. This is an example of an (m-1)-dimensional submanifold of $\mathbb{R}^m$. For a concrete example, take the function $x \to f(x) = |x|^2$ where $|x|$ denotes the Euclidean norm of $x$. Any non-zero number is a regular value, but of course only positive regular values are relevant. In this case $f^{-1}(r^2)$ for $r > 0$ is the sphere in $\mathbb{R}^{n+1}$ of radius $r$. Spheres of different radii are diffeomorphic.

**Definition:** A submanifold in $\mathbb{R}^m$ of dimension $n < m$ is a subset, $\Sigma$, with the following property: Let $p$ denote any given point in $\Sigma$. There is a ball $U_p \subset \mathbb{R}^m$ around $p$ and a map $\psi_p: U_p \to \mathbb{R}^{m-n}$ with 0 as a regular value and such that $\Sigma \cap U_p = \psi_p^{-1}(0)$.

The implicit function theorem says that any such $\Sigma$ is a smooth, n-dimensional manifold. The following lemma can be used to obtain examples of submanifolds.

**Lemma:** Suppose that $n \leq m$ and that $B \subset \mathbb{R}^n$ is an open ball centered on the origin. Let $\varphi: B \to \mathbb{R}^m$ denote a smooth, 1-1 map whose differential is everywhere injective. Let $W \subset B$ denote any given open set with compact closure. Then $\varphi(W)$ is a submanifold of $\mathbb{R}^m$ such that $\varphi|_W: W \to \varphi(W)$ is a diffeomorphism.

**Proof:** This can be proved using the inverse function theorem. To say more, fix any given point $p \in \varphi(W)$. Needed is a smooth map, $\psi_p$ from a ball centered at $p$ to $\mathbb{R}^{m-n}$ that maps $p$ to the origin, whose differential at $p$ is surjective, and is such that $\varphi(W)$ near $p$ is $\psi_p^{-1}(0)$. To obtain this data, let $z \in W$ denote the point mapped by $\varphi$ to $p$. Let $K \subset \mathbb{R}^m$ denote the kernel of the adjoint of $\varphi|_p$. This is an m-n dimensional subspace because the differential of $\varphi$ at $z$ is injective. Keeping this in mind, define $\lambda: W \times K \to \mathbb{R}^m$ by the rule $\lambda(x, v) = \varphi(x) + v$. Note that $\lambda(z, 0) = p$. As the differential of $\lambda$ at $(z, 0)$ is an isomorphism, the inverse function theorem finds a ball, $U_p$, of $p$ and a smooth map $\eta: U_p \to W \times K$ such that the following is true: First, $\eta(U_p)$ is an open neighborhood of $(z, 0)$. Second, $\varphi \circ \eta$ is the identity. Third, $\eta \circ \varphi$ is the identity on $\eta(U_p)$.
× K → K denote the orthogonal projection. Identify K with \( \mathbb{R}^{n-m} \) and define the map \( \psi_p: U_p \to K = \mathbb{R}^{n-m} \) by the rule \( \psi_p(x) = p\eta(x) \). By construction, \( \psi_p^{-1}(0) = \phi(\eta(U_p)) \) and \( \psi_p|_p \) is surjective.

Here is an example of a dimension 2 submanifold in \( \mathbb{R}^3 \). Fix \( \rho \in (0, 1) \). The submanifold is the set of points \((x_1, x_2, x_3) \in \mathbb{R}^3\) where

\[
f(x_1, x_2, x_3) = ((x_1^2 + x_2^2)^{1/2} - 1)^2 + x_3^2 = \rho^2.
\]

This is a torus of revolution; it is diffeomorphic to \( S^1 \times S^1 \) as can be seen by introducing an angle \( \phi \in \mathbb{R}/2\pi\mathbb{Z} \) for the left most \( S^1 \) factor, a second angle \( \varphi \) for the right most factor, and then using these to parametrize the locus by the map that sends

\[
(\phi, \varphi) \to (x_1 = (1 + \rho\cos\phi)\cos\varphi, x_2 = (1 + \rho\cos\phi)\sin\varphi, x_3 = \rho\sin\phi).
\]

A torus sitting in \( \mathbb{R}^3 \) looks like the surface of an inner tube of the sort that you would inflate for a bicycle. What follows is a depiction viewed from above

![Image of a torus]

A torus sitting in \( \mathbb{R}^3 \) looks like the surface of an inner tube of the sort that you would inflate for a bicycle. What follows is a depiction viewed from above can also appear as a submanifold of \( \mathbb{R}^3 \). The notion of what constitutes a hole has rigorous definition, see for example Chapter 12 in the book *Topology* by James Munkres (Prentice Hall, 2000). Chapter 14 gives two equivalent definitions using differential geometric notions. Suffice it to say here that the torus depicted above has genus 1, and a standard pretzel has genus 2. What follows depicts a surface in \( \mathbb{R}^3 \) with a given genus \( g \). This construction was explained to the author by Curt McMullen. To start, fix \( r \in (0, \frac{1}{100g}) \) and a function \( h: \mathbb{R}^2 \to \mathbb{R} \) with the following properties:

- \( h(x, y) \leq 0 \) where \( |x|^2 + |y|^2 \geq 1 \).
- *For each* \( k = \{1, \ldots, g\} \), *require* \( h(x, y) \leq 0 \) where \( |x - \frac{k}{2g}|^2 + |y|^2 \leq r^2 \).
- \( h(x, y) > 0 \) otherwise.
• dh = \frac{\partial}{\partial x} h \, dx + \frac{\partial}{\partial y} h \, dy \neq 0 \text{ at any point where } h = 0, \text{ thus on the circle of radius 1 about the origin, and on the g circles of radius r about the respective points in the set } \{(x = \frac{k}{2g}, y = 0)\}_{k=1,...,g}.

Below is a schematic of h might look like in the case g = 2. The blue area is where h > 0; the white is where h \leq 0.

The corresponding surface in \( \mathbb{R}^3 \) is the set of points where the function \( f = z^2 - h(x, y) \) is zero. The condition on h in the fourth bullet guarantees that 0 is a regular value of f and so \( f^{-1}(0) \) is a smooth submanifold in \( \mathbb{R}^3 \). It is compact because \( h(x, y) \leq 0 \) when \( x^2 + y^2 \) is greater than 1.

Any manifold can be realized as a submanifold of some \( m \gg n \) version of \( \mathbb{R}^m \).

What follows is an argument for the case when the manifold in question is compact manifold. To start, fix a finite coordinate atlas for M. Label the elements in this atlas as \{(U_1, \varphi_1), \ldots, (U_N, \varphi_N)\}. Fix a subordinate partition of unity, \( \{\chi_\alpha\}_{1 \leq \alpha \leq N} \). This is a set of smooth functions on M such that any given \( \chi_\alpha \) has support in the corresponding set \( U_\alpha \) and such that \( \sum_{1 \leq \alpha \leq N} \chi_\alpha = 1 \) at each point. The second appendix to this chapter gives a quick tutorial such thing.

Introduce the map

\[ \psi: M \to \mathbb{R}^{Nm+N} = \times_{1 \leq \alpha \leq N} (\mathbb{R} \times \mathbb{R}^m) \]

that is defined by the following rule: Send any given point \( x \in M \) to the point in the \( \alpha \)'th factor of \( \mathbb{R}^n \) given by \( (1 - \chi_\alpha(x), \chi_\alpha(x)\varphi_\alpha(x)) \). The differential of this map is injective because there exists near any given point at least one index \( \alpha \) such that \( \varphi_\alpha \neq 0 \). If this is the case for \( \alpha \), then the map to \( \mathbb{R}^{m+1} \) from a small neighborhood of this point given by \( x \to (1 - \chi_\alpha(x), \chi_\alpha(x)\varphi_\alpha(x)) \) is injective. This understood, the implicit function theorem can be used to infer that the image of \( \psi \) is a submanifold of \( \mathbb{R}^{N(m+1)} \) and that \( \psi \) is a diffeomorphism onto its image.
Coordinates near a submanifold: The implicit function theorems can be used to construct ‘nice’ coordinates for \( \mathbb{R}^n \) near any given point of a submanifold. Here is what can be done: Let \( Y \subset \mathbb{R}^n \) denote a submanifold of some dimension \( d \), and let \( p \) denote any given point in \( Y \). Then there is a neighborhood \( U \subset \mathbb{R}^n \) of \( p \) and a diffeomorphism \( \psi: U \rightarrow \mathbb{R}^n \) such that \( \psi(Y) \) near \( p \) is the locus of points \( (x_1, \ldots, x_n) \) where \( x_{d+1} = x_{d+2} = \cdots = x_n = 0 \).

1d) Submanifolds of manifolds

Suppose that \( M \) is a smooth manifold of dimension \( m \). A subset \( Y \subset M \) is a submanifold of some given dimension \( n \leq m \) if the following is true: Fix any point \( p \in Y \) and there is an open set \( U \subset M \) that contains \( p \), a coordinate chart map \( \varphi: U \rightarrow \mathbb{R}^m \) such that \( \varphi(Y \cap U) \) is an \( n \) dimensional submanifold of \( \mathbb{R}^m \).

Note that if \( \psi: M \rightarrow N \) is smooth, and if \( Y \subset M \) is a smooth submanifold, then the restriction \( \psi_Y: Y \rightarrow N \) is smooth.

Immersions and submersions: The notion of a submanifold is closely related to the notions of immersion and submersion. To say more, let \( M \) denote a manifold of dimension \( m \) and let \( Y \) denote one of dimension \( n \). Suppose first that \( n \leq m \). A smooth map \( \psi: Y \rightarrow M \) is said to be an immersion when it has the following property: Fix any point \( p \in Y \) and an open set \( V \subset Y \) that contains \( p \) with a coordinate map \( \varphi_V: V \rightarrow \mathbb{R}^n \). Let \( U \subset M \) denote an open set containing \( \psi(p) \) with a coordinate map \( \varphi_U: U \rightarrow \mathbb{R}^m \). Then the differential of \( \varphi_U \circ \psi \circ \varphi_V^{-1} \) at \( \varphi_V(p) \) is an injective linear map from \( \mathbb{R}^n \) to \( \mathbb{R}^m \). This implies that \( \psi(V) \) is a \( n \)-dimensional submanifold of \( M \) on some small neighborhood of \( \psi(p) \).

Suppose next that \( n \geq m \). The map \( \psi \) is said to be a submersion when the following is true: Fix \( p, V \) and \( U \) as above. Then the differential of \( \varphi_U \circ \psi \circ \varphi_V \) at \( p \) is a surjective linear map from \( \mathbb{R}^n \) to \( \mathbb{R}^m \). An application of the inverse function theorem leads to the conclusion that \( \psi^{-1}(\psi(p)) \) is a \( n-m \) dimensional submanifold of \( Y \) near \( p \).

1e) More constructions of manifolds

What follows are some simple ways to obtain new manifolds from old ones.

Products of smooth manifolds. The product of two circles has the following generalization: Suppose that \( M, N \) are smooth manifolds. Then \( M \times N \) has a canonical smooth structure. A coordinate atlas for the latter can be taken as follows: Fix a coordinate atlases, \( \mathcal{U} \) for \( M \) and \( \mathcal{V} \) for \( N \), that define the respective smooth structures. Then a coordinate atlas for \( M \times N \) is the collection \( \{(U \times V, \varphi \times \psi)\}_{(U,\varphi) \in \mathcal{U}, (V,\psi) \in \mathcal{V}} \).
The dimension of \( M \times N \) is the sum of the dimensions of \( M \) and \( N \). For example, the circle is \( S^1 \); the n-fold product of the circle is called the n-dimensional torus.

**Open subsets of manifolds:** Let \( M \) denote a smooth manifold and \( W \subset M \) any open subset. Then \( W \) inherits the structure smooth manifold. Indeed, fix a coordinate atlas for \( M \) that defines the smooth structure. Then a coordinate atlas the smooth structure on \( W \) can be taken to be the collection \( \{(U \cap W, \phi)\} \).

**Quotients that are manifolds:** Let \( M \) denote a smooth manifold. An equivalence relation on \( M \) is a collection of disjoint subsets in \( M \) whose union is the whole of \( M \). I will use \( M/\sim \) to denote this collection of sets. An element in the set \( M/\sim \) is said to be an equivalence class; and when \( x \) and \( x' \) are points in \( M \), I write \( x \sim x' \) when they lie in the same equivalence class. A map \( \pi: M \to M/\sim \) is defined by sending any given point to its equivalence class. The set \( M/\sim \) is given the quotient topology whereby a set \( U \subset M/\sim \) is declared to be an open set if and only if \( \pi^{-1}(U) \) is open in \( M \).

It is often the case that \( M/\sim \) inherits from \( M \) the structure of a smooth manifold. In particular, this occurs when the following is true: Let \( x \in M \). Then there is an open neighborhood \( V \subset M \) on \( x \) such that \( \pi|_V \) is a 1-1 and a homeomorphism from \( V \) to \( \pi(V) \). If this is the case, then \( M/\sim \) has a canonical smooth structure, this defined as follows: Let \( \mathcal{U} \) denote a given coordinate atlas for \( M \) that defines the smooth structure. Fix an open cover, \( \mathcal{V} \), of \( M \) such that \( \pi \) is 1-1 on any set from \( \mathcal{V} \). A coordinate atlas for \( M \) that in \( \mathcal{U} \)'s equivalence class is defined by the collection \( \{(U \cap V, \phi)\} \). With this understood, a coordinate atlas that defines the smooth structure for \( M/\sim \) is given by the collection \( \{(\pi(U \cap V), \phi): U \cap V \to \mathbb{R}^n\} \). Note in particular that the map \( \pi \) from \( M \) to \( M/\sim \) is a smooth map.

Here is a first example: The circle \( S^1 \) can be viewed as \( \mathbb{R}/\sim \) where the equivalence relation identifies points \( x \) and \( x' \) if and only if \( x = x' + 2\pi k \) for some \( k \in \mathbb{Z} \). With \( S^1 \) viewed as the unit circle in \( \mathbb{R}^2 \), the map \( \pi \) becomes the map sending \( t \in \mathbb{R} \) to the point \((\cos(t), \sin(t))\). When I think of \( S^1 \) in terms of this equivalence relation, I will write it as \( \mathbb{R}/2\pi \mathbb{Z} \).

By the same token, the n-dimensional torus \( \times_n S^1 \) can be viewed as \( \mathbb{R}^n/\sim \) where the equivalence relation identifies points \( x \) and \( x' \) if each entry of \( x \) differs from the corresponding entry of \( x' \) by \( 2\pi \) times an integer.

Here is a second example: The real projective space \( \mathbb{R}P^n = S^n/\sim \) where the equivalence relation identifies unit vectors \( x \) and \( x' \) in \( \mathbb{R}^{n+1} \) if and only if \( x = \pm x' \). As a parenthetical remark, note that \( \mathbb{R}P^n \) can be viewed as the set of 1-dimensional linear subspaces in \( \mathbb{R}^{n+1} \). The correspondence is as follows: A given 1-dimensional subspace has precisely two vectors of unit length, one is \(-1\) times the other. These two vectors
define a point in $\mathbb{R}P^n$. Conversely, a point in $\mathbb{R}P^n$ gives the linear subspace that is spanned by either of its corresponding points in $S^n$.

For the third example, $n \in \{2, \ldots, \}$, and fix a positive integer $m$. Define an equivalence relation $\sim$ on $S^{2n+1}$ as follows: The point $x = (x_1, x_2, \ldots, x_{2n})$ is equivalent to $x' = (x_1', x_2', \ldots, x_{2n}')$ if and only if for each $k \in \{1, \ldots, n\}$, the following is true:

\[ x_{2k+1}' = \cos(2\pi j/m) x_{2k+1} - \sin(2\pi j/m) x_{2k} \quad \text{and} \quad x_{2k}' = \cos(2\pi j/m) x_{2k} + \sin(2\pi j/m) x_{2k+1} \]

with $j$ some integer. Note that $\mathbb{R}P^{2n+1}$ is obtained when $m = 2$. Manifolds of this sort are examples of lens spaces.

If) More smooth manifolds: The Grassmannians

As noted above, the real projective space $\mathbb{R}P^n$ can be viewed as the set of 1-dimensional vector subspaces in $\mathbb{R}^{n+1}$. In general, if $m > n$, then the set of $n$-dimensional vector subspaces in $\mathbb{R}^m$ also has the structure of a smooth manifold. The latter manifold is denoted by $\text{Gr}(m; n)$ and is called the Grassmannian of $n$-planes in $\mathbb{R}^m$. The manifold structure for $\text{Gr}(m; n)$ can be obtained by exhibiting a coordinate atlas whose transition functions are diffeomorphisms. The first task, however, is to give $\text{Gr}(m; n)$ a suitable topology. The latter is the metric topology defined as follows: Let $V \subset \mathbb{R}^m$ denote a given $n$-dimensional subspace. Let $\Pi_V: \mathbb{R}^m \to V$ denote the orthogonal projection. An open neighborhood of $V$ is the set, $O_V$, of subspaces $V'$ such that $\Pi_V: V' \to V$ is an isomorphism. A basis of open neighborhoods of $V$ is indexed by the positive numbers; and the $\varepsilon > 0$ member, $O_{V, \varepsilon}$, consists of the $n$-dimensional subspaces $V' \subset \mathbb{R}^m$ such that $|\Pi_V v - v| < \varepsilon |v|$ for all $v \neq 0$ in $V'$.

As a manifold, $\text{Gr}(m, n)$ has dimension $n(m-n)$. The neighborhood $O_V$ will serve as a coordinate chart centered around $V$; thus a map from $O_V$ to $\mathbb{R}^{n(m-n)}$ that gives the local coordinates is needed. In order to give the coordinate map, first introduce first $M(m, n)$ to denote the vector space of $(m-n) \times m$ matrices. The entries of any given matrix $p \in M(m, n)$ serve as coordinates and identify this space with $\mathbb{R}^{n(m-n)}$. Now, let $V \subset \mathbb{R}^m$ denote a given $n$-dimensional subspace and let $V^\perp \subset \mathbb{R}^m$ denote its orthogonal complement. Use $\Pi_{V^\perp}: \mathbb{R}^m \to V^\perp$ in what follows to denote the orthogonal projection to $V^\perp$. Fix a basis for $V$ to identify it with $\mathbb{R}^m$ and fix a basis for $V^\perp$ to identify it with $\mathbb{R}^{m-n}$. Use these bases to identify the space of linear maps from $V$ to $V^\perp$ (this denoted by $\text{Hom}(V; V^\perp)$) with $\mathbb{R}^{n(m-n)}$. Now, write $\mathbb{R}^m = V \oplus V^\perp$. Suppose that $V' \in O_V$. Since $\Pi_{V'}: V' \to V$ is an isomorphism, there is a linear inverse $L_{V, V'}: V \to V'$. This understood, the composition $\Pi_{V^\perp} L_{V, V'}$ is a linear map from $V$ to $V^\perp$, thus an element in $\text{Hom}(V; V^\perp) = \mathbb{R}^{n(m-n)}$. This point is defined to be $\varphi_V(V')$. It is an exercise to check that the map $\varphi_V$ is
continuous. To see that it is a homeomorphism, it is enough to exhibit a continuous inverse: The inverse sends a given \( n \times (m-n) \) matrix \( p \in \text{Hom}(V, V^\perp) = \mathbb{R}^{n(m-n)} \) to the \( m \)-dimensional linear subspace \( \varphi^{-1}_V(p) = \{(v, pv): v \in V\} \).

It remains yet to verify that the transition function between any two intersecting charts defines a smooth map between domains in \( \mathbb{R}^{n(m-n)} \). This task is deferred to the appendix of Chapter 2.

It is left as a linear algebra exercise to verify that the transition functions between two charts \( O_V \) and \( O_{V'} \) are diffeomorphisms.

I also leave it as an exercise to verify that the smooth structure just defined for \( \text{Gr}(n+1,n) \) is the same as that given previously for \( \mathbb{R}P^n \).

**Appendix 1: How to prove the inverse function and implicit function theorems**

What follows are detailed outlines of proofs first the inverse function theorem and then of the implicit function theorem. See, for example *Differential Topology* by Victor Guillemin and Alan Pollack (Prentice Hall 1974) for unabridged proofs.

**Proof of the inverse function theorem:** The first point to make is that no generality is lost by taking \( p \) and \( \psi(p) \) to be the origin. Let \( m \) denote the \( n \times n \) matrix \( \psi_{\cdot 0} \). I look for an inverse, \( \sigma \), that has the form

\[
\sigma(x) = m^{-1}x + f(x)
\]

where \( \tau \) is such that \(|f(x)| \leq c_0|x|^2\). Here I have introduced a convention that is used in this chapter and all subsequent chapters: What is denoted by \( c_0 \) is always a constant, greater than 1, whose value increases on subsequent appearances. The use of this generic notation is meant to avoid the proliferation of different symbols or subscripts to denote positive numbers which are either greater than 1 or less than 1, but are such that the precise value is of no real consequence to the discussion at hand. The convention here takes \( c_0 \) to be greater than 1, so when it is important that a positive number have value less than 1, I use \( c_0^{-1} \).

To find \( f \), I use Taylor’s theorem with remainder to write \( \psi(x) = mx + \tau(x) \) where \( \tau \) obeys \(|\tau(x)| \leq c_0|x|^2\). Note also that any given partial derivative of \( \tau \) obeys \(|\partial_j \tau(x)| \leq c_0|x|\). In any event, the map \( \sigma \) is an inverse to \( \psi \) if and only if \( \psi(\sigma(x)) = x \) for \( x \) near 0 in \( \mathbb{R}^n \). In terms of \( f \), this demands that

\[
mf(x) + \tau(m^{-1}x + f(x)) = 0.
\]

This is to say that
This sort of equation can be solved for any given $x$ using the contraction mapping theorem. The point being that the desired solution, $f(x)$, to the preceding equation is a fixed point of the map from $\mathbb{R}^n$ to $\mathbb{R}^n$ that sends $y$ to $T(x) = -m^1 \tau (m^1 x + f(x))$. By way of reminder, the contraction mapping theorem asserts the following:

**The contraction mapping theorem:** Let $B \subset \mathbb{R}^n$ denote a ball and $T$ a map from $B$ to itself. The map $T$ has unique fixed point in $B$ if $|T(y) - T(y')| \leq (1 - \delta)|y - y'|$ for some positive $\delta$ and all $y, y' \in B$.

**Proof:** Start with $y_0 \in B$ and then construct a sequence $\{y_k = T(y_{k-1})\}_{k=1,2,...}$. This is a Cauchy sequence because $|y_k - y_{k-1}| = |T(y_{k-1}) - T(y_{k-2})| \leq (1 - \delta)|y_{k-1} - y_{k-2}|$. Let $y_\ell$ denote the limit point. This point is in $B$ and $T(y_\ell) = y_\ell$. There can’t be any other fixed point in $B$; were $y'$ such a point, then $|y_\ell - y'| = |T(y_\ell) - T(y')| \leq (1 - \delta)|y_\ell - y'|$ which can be true only if $y_\ell = y'$.

In the case at hand, $|T\ell(y)| \leq c_0(1 + |y|^2)$ because $\tau$ comes from the remainder term in the Taylor’s expansion for $\psi$. This implies that $T\ell$ maps any ball of radius $r < \frac{1}{2} c_0^{-1}$ to itself if $|x| \leq c_0^{-1} r$. For the same reason, $|T\ell(y) - T\ell(y')| \leq c_0(|x| + |y|)|y - y'|$, and so $T\ell$ is a contraction mapping on such a ball if $|x| \leq c_0^{-1} r$. The unique fixed point is $f(x)$ that obeys $|f(x)| \leq c_0 r^2$.

It remains yet to prove that the map $x \rightarrow f(x)$ is smooth. A proof that such is the case can also be had by successively using the contraction mapping theorem to analyze first the difference $|f(x) - f(x')|$, and then successively higher difference quotients.

For the first difference quotient, use the fact that $f(x)$ and $f(x')$ are respective fixed points of $T\ell$ and $T\ell$ to write their difference as

$$f(x) - f(x') = m^1 \tau (m^1 x + f(x)) - \tau (m^1 x' + f(x'))$$

Given that $\tau$ is the remainder term in Taylor’s theorem, and given that $x$ and $x'$ lie in the ball of radius $r$, and that $f(x)$ and $f(x')$ lie in the ball of radius $c_0 r^2$, this last equation implies that $|f(x) - f(x')| \leq c_0 \tau (|x - x'| + |f(x) - f(x')|)$. In particular, if $r < \frac{1}{2} c_0^{-1}$, this tells us that $|f(x) - f(x')| \leq c_0^{-1} |x - x'|$.

**Proof of the implicit function theorem:** What follows is a sketch of the proof of the implicit function theorem. The first step constructs a local coordinate chart centered on any given point in $\psi^{-1}(a)$. Let $p$ denote such a point. Introduce $K \subset \mathbb{R}^m$ to denote the kernel of $\psi_{|p}$, and let $p: \mathbb{R}^m \rightarrow K$ denote the orthogonal projection. Now define a map, $\phi$:
U → \mathbb{R}^m = K \times \mathbb{R}^{m-n} by sending any given point x ∈ \mathbb{R}^m to (p(x - p), ψ(x) - a). The differential of this map at p is an isomorphism, and so the implicit function theorem finds balls B_K ⊂ K and B´ ⊂ \mathbb{R}^{m-n} about the respective origins with a map σ: B_K × B´ → \mathbb{R}^m that obeys σ(0) = p and is such that \varphi \circ σ = identity. In particular, this has the following implication: Let y ∈ B_K. Then ψ(σ(y, 0)) = a. As a consequence, the map from B_K to \mathbb{R}^m that sends y to σ(y, 0) is 1-1 and invertible. The inverse is the map x → p(x - p) from \mathbb{R}^m to K = \mathbb{R}^n, and so the latter map gives local coordinates for \psi^{-1}(a) near p.

Granted the preceding, it remains yet to prove that the transition functions for these coordinate charts are smooth and have smooth inverses. There is nothing terribly complicated about this, and so it is left to the reader.

Appendix 2: Partitions of unity

This appendix says something about how to construct a locally finite partition of unity for a coordinate atlas. The story starts on [0, ∞). What follows directly describes a smooth function that is zero at points t ∈ [0, ∞) with t ≥ r and is non-zero at all points t ∈ [0, ∞) with t < 1. The function, χ, is defined by the rule whereby

- \chi(t) = e^{1/(t-1)} for t < 1.
- \chi(t) = 0 for t ≥ 1.

Granted this, suppose that \mathcal{U} is a locally finite coordinate atlas for a given smooth, n-dimensional manifold. Since \mathcal{U} is locally finite, one can assign to each (U, ϕ_U) ∈ \mathcal{U} a positive real number, r_U, such that the following is true: Let U_r denote the inverse image via ϕ_U of the ball of radius r_U about the origin in \mathbb{R}^n. Then the collection \{U_r\}_{(U,ϕ)∈\mathcal{U}} is an open cover of M. This understood, associate to any given (U, ϕ) ∈ \mathcal{U} the function σ_U on M that is defined as follows: If p is not in U, then σ_U(p) = 0. If p is in U, then σ_U(p) is equal to \chi(|ϕ_U(p)|/r_U). Note that σ_U(p) is non-zero on U, but zero on the complement of U.

The collection \{σ_U\}_{(U,ϕ)∈\mathcal{U}} is almost, but not quite the desired subordinate partition of unity function. It is certainly the case that \sum_{(U,ϕ)∈\mathcal{U}} σ_U is nowhere zero, but this sum is not necessarily equal to 1. The desired subordinate partition of unity consists of the set \{\chi_U\}_{(U,ϕ)∈\mathcal{U}} where \chi_U = (\sum_{(U,ϕ)∈\mathcal{U}} σ_U)^{-1} σ_U.

Additional Reading

CHAPTER 2: Matrices and Lie groups

Matrices and especially invertible matrices play a central role in most all of differential geometry. In hindsight, I don’t find this to be surprising. The role stems ultimately from Taylor’s theorem: A smooth map between two Euclidean spaces can be well approximated near any given point by a linear map, this the differential at the point. The inverse and implicit function theorems depend entirely on this observation. In any event, matrices and notions from linear algebra appear in all of the subsequent chapters. Constructions involving matrices also give some interesting and explicit examples of smooth manifolds.

2a) The general linear group

I use $\mathbb{M}(n; \mathbb{R})$ to denote the vector space of $n \times n$ matrices with real entries. This is a copy of the Euclidean space $\mathbb{R}^{n^2}$ with the entries of a matrix giving the coordinate functions. Matrices $m$ and $m'$ can be multiplied, and matrix multiplication gives a smooth map, $(m, m') \to mm'$, from $\mathbb{M}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R})$ to $\mathbb{M}(n; \mathbb{R})$. This is smooth by virtue of the fact that the entries of $mm'$ are linear functions of the coordinates of $m$, and also linear in the coordinates of $m'$. The space $\mathbb{M}(n; \mathbb{R})$ has two special functions. The first function, $\det: \mathbb{M}(n; \mathbb{R}) \to \mathbb{R}$ sends any given matrix to its determinant. This is a smooth function since it is a polynomial of degree $n$ in the entries. The determinant of a matrix, $m$, is denoted by $\det(m)$. The other function of interest is the trace. The trace is the sum of the diagonal entries and so a linear function of the entries. The trace of a matrix $m$ is denoted by $\text{trace}(m)$.

Let $\text{Gl}(n; \mathbb{R})$ denote the subspace in $\mathbb{M}(n; \mathbb{R})$ of invertible matrices. This is to say that $m \in \text{Gl}(n; \mathbb{R})$ if and only if the $\det(m) \neq 0$. This is an open subset of $\mathbb{M}(n; \mathbb{R})$, and so defacto $\text{Gl}(n; \mathbb{R})$ is a smooth manifold of dimension $n^2$. In particular, the $n^2$ entries of the matrix restrict to $\text{Gl}(n; \mathbb{R})$ near any given point to give local coordinates. The multiplication map restricts to $\text{Gl}(n; \mathbb{R})$ as a smooth map.

The manifold $\text{Gl}(n; \mathbb{R})$ has a canonical diffeomorphism, this the map to itself that sends any given matrix $m$ to the inverse matrix $m^{-1}$. This map is smooth because any given entry of $m^{-1}$ is the quotient by $\det(m)$ of a polynomial in the entries of $m$. For example, in the $2 \times 2$ case, write

$$m = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

and then its determinant is $ad - bc$ and its inverse is
\[
m^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.
\]

The form just asserted for the entries of the inverse in the \(n > 2\) case (a polynomial divided by the determinant) can be seen by using the reduced row echelon form algorithm to construct the inverse.

What follows is gives a very useful formula for the directional derivatives of the map \(m \to m^{-1}\). Let \(a \in M(n; \mathbb{R})\) denote any given matrix. Then the directional derivatives of the coordinates of map \(m \to m^{-1}\) in the direction \(a\) are the entries of the matrix \(-m^{-1}a m^{-1}\). Consider, for example, the coordinate given by the \((i,j)\) entry, \((m^{-1})_{ij}\).

The directional derivative in the direction \(a\) of this function on \(Gl(n; \mathbb{R})\) is \(- (m^{-1}a m^{-1})_{ij}\). In particular, the partial derivative of the function \(m \to (m^{-1})_{ij}\) with respect to the coordinate \(m_{rs}\) is \(- (m^{-1})_{ir}(m^{-1})_{sj}\).

2b) Lie groups

A group is a set with a fiducial point and two special maps. To elaborate, let \(G\) denote the set. The fiducial point is denoted in what follows by \(\iota\); it is called the identity.

The first of the maps is the multiplication map, this a map from \(G \times G\) to \(G\). It is customary to denote this map by \((g, g') \to gg'\). This map is constrained so that \(\iota g = g\) and also \(\iota g' = g\). The other map sends \(G\) to \(G\); it is the inverse. It is customary to write the inverse as \(g \to g^{-1}\). The inverse map is constrained so that \(g g^{-1} = \iota = g^{-1} g\).

A Lie group is a group with the structure of a smooth manifold such that both the multiplication map and the inverse map are smooth. The manifold \(Gl(n; \mathbb{R})\) as just described is the Ur example of a Lie group. Note that \(\mathbb{R}^n\) with the origin playing the role of \(\iota\), with addition playing the role of multiplication, and with the map \(x \to -x\) playing the role of inverse is a group. The latter is an example of an Abelian group, this a group with the property that \(gg' = g'g\) for all pairs \(g\) and \(g'\).

It is often the case that interesting groups occur as subgroups of larger groups. By way of reminder, a subgroup of a group \(G\) is a subset, \(H\), that contains \(\iota\), is mapped to itself by the inverse map, and is such that multiplication maps \(H \times H\) to \(H\).

**Lemma:** A subgroup of a Lie group that is also a submanifold is a Lie group with respect to the induced smooth structure.

**Proof:** This follows by virtue of the fact that the restriction to a submanifold of any smooth map to any given manifold defines a smooth map from the submanifold to the given manifold.

2c) Examples of Lie groups
The group $\text{Gl}(n; \mathbb{R})$ contains sub-Lie groups that arise most often in differential geometric constructions.

**The group $\text{Sl}(n; \mathbb{R})$:** Let $\text{Sl}(n; \mathbb{R}) \subset \text{Gl}(n; \mathbb{R})$ denote the subset of matrices whose determinant is 1. This is a subgroup by virtue of the fact that $\det(mm') = \det(m)\det(m')$. It is also a smooth submanifold of dimension $n^2 - 1$. This follows from the fact that the differential of this function can be written as

$$d(\det)_{m} = \det(m)\sum_{ij} (m^{-1})_{ij} dm_{ij} = \det(m) \text{trace}(m^{-1} dm).$$

Here, $dm$ denotes the $n \times n$ matrix whose $(i,j)$'th entry is the coordinate differential $dm_{ij}$. As can be seen from the preceding formula, the differential of $\det$ is not zero where $\det$ is non-zero.

As it turns out, the group $\text{Sl}(2; \mathbb{R})$ is diffeomorphic to $S^1 \times \mathbb{R}^2$. This can be seen by using a linear change of coordinates on $M(2; \mathbb{R})$ that writes the entries in terms of coordinates $(x, y, u, v)$ as follows:

$$m_{11} = x - u, \quad m_{22} = x + u, \quad m_{12} = v - y \quad \text{and} \quad m_{21} = v + y.$$  

The condition that $\det(m) = 1$ now says that $x^2 + y^2 = 1 + u^2 + v^2$. This understood, the diffeomorphism from $S^1 \times \mathbb{R}^2$ to $\text{Sl}(2; \mathbb{R})$ sends a triple $(\theta, a, b)$ to the matrix determined by $x = (a^2 + b^2)^{1/2} \cos(\theta)$, $y = (a^2 + b^2)^{1/2} \sin(\theta)$, $u = a$, $v = b$. Here, $\theta \in [0, 2\pi]$ is the angular coordinate for $S^1$.

**The orthogonal groups $O(n)$ and $SO(n)$:** The orthogonal group $O(n)$ is the set of matrices $m \in \text{Gl}(n; \mathbb{R})$ such that $m^T m = 1$. Here, $(m^T)_{ik} = m_{ki}$. Since $(ab)^T = b^T a^T$, the set of orthogonal matrices forms a subgroup. The group $SO(n)$ sits in $O(n)$ as the subgroup of orthogonal matrices with determinant 1. To see that $SO(n)$ and $O(n)$ are manifolds, introduce $\text{Sym}(n; \mathbb{R})$ to denote the vector subspace of $n \times n$ symmetric matrices. This is to say that a matrix $h$ is in $\text{Sym}(n; \mathbb{R})$ when $h^T = h$; thus when $h_{ij} = h_{ji}$. This is vector space has dimension $n(n+1)/2$. The vector space $\text{Sym}(n; \mathbb{R})$ is a linear subspace of the Euclidean space $M(n; \mathbb{R})$, and so a version of $\mathbb{R}^{n(n+1)/2}$. Let $\psi: M(n; \mathbb{R}) \rightarrow \text{Sym}(n; \mathbb{R})$ denote the map $m \mapsto m^T m$. This map is quadratic in the entries of the matrix, and so a smooth map between Euclidean spaces. As is explained in the next paragraph, the identity matrix $\iota \in \text{Sym}(n; \mathbb{R})$ is a regular value of $\psi$. This being the case, it follows from the implicit function theorem that $\psi^{-1}(\iota)$ is a submanifold of $\text{Gl}(n; \mathbb{R})$ and so a Lie group.

24
To see that $\iota$ is a regular value, one must verify that the differential of $\iota$ at any given matrix $m \in O(n)$ is surjective. To see this, note that the differential of $\psi$ at a given matrix $m$ can be written as $\psi_*|_m = dm^Tm + m^Tdm$, where $dm$ is the matrix of coordinate differentials with $(i, j)$ entry equal to $dm_{ij}$. To say explain this notation, I should tell you how to interpret $dm^Tm + m^Tdm$ as a linear map from the Euclidean space $\mathbb{M}(n; \mathbb{R})$ to the Euclidean space $\text{Sym}(n; \mathbb{R})$. This linear map acts as follows: Let $a \in \mathbb{M}(n; \mathbb{R})$ denote any given matrix. Then $\psi_*|_m a$ is the symmetric matrix $a^Tm + m^Ta$. The rule here is to replace $dm$ with $a$. Note that this notation is used throughout this book, so it is best to come to terms with it. (It was used above to write $d(\det)|_m$ as $\det(m)\,\text{trace}(m^{-1}dm)$.) In any event, granted this identification of $\psi_*|_m$ it is necessary to prove the following: Given a symmetric matrix $h$, there exists a matrix $a$ such that $a^Tm + m^Ta = h$. Given that $mm^T = \iota$, this equation is solved by taking $a = \frac{1}{2} m h$.

Note that $SO(3)$ is diffeomorphic to the quotient space $\mathbb{RP}^3 = S^3/\sim$ where $\sim$ equates any given $x$ with $\pm x$. Meanwhile, $SO(4)$ is diffeomorphic to $(S^1 \times S^3)/\sim$ where the equivalence relation equates a pair $(x, y) \in SO(3)$ only with itself and with $(-x, -y)$.

2d) Some complex Lie groups

The group $\text{Gl}(n; \mathbb{R})$ has its analog the group $\text{Gl}(n; \mathbb{C})$ of $n \times n$, invertible matrices with complex number enter entries. There are two equivalent ways to view $\text{Gl}(n; \mathbb{C})$ as a Lie group. The first deals with complex numbers right off the bat by introducing the vector space (over $\mathbb{C}$) of $n \times n$ complex matrices. The latter is denoted by $\mathbb{M}(n; \mathbb{C})$. This vector space can be viewed as a version of $\mathbb{R}^{2n^2}$ by using the real and imaginary parts of the $n^2$ entries as coordinates. Viewed in this light, multiplication of matrices defines a smooth map from $\mathbb{M}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C})$ to itself by virtue of the fact that the entries of a product $mm'$ is linear in the entries of $m$ and linear in those of $m'$. The determinant function maps $\mathbb{M}(n; \mathbb{C})$ to $\mathbb{C}$, and this is a smooth map since it is a polynomial of degree $n$ in the entries. The latter map is also denoted by $\det(\cdot)$ with it understood that this version has values in $\mathbb{C}$ while the version defined on the space of matrices with real entries has values in $\mathbb{R}$.

The group $\text{Gl}(n; \mathbb{C})$ appears as the open set in the Euclidean space $\mathbb{M}(n; \mathbb{C})$ where the determinant is non-zero. The matrix multiplication for matrices with complex number entries restricts to $\text{Gl}(n; \mathbb{C})$ as a smooth map from $\times_2 \text{Gl}(n; \mathbb{C})$ to itself. Meanwhile, the inverse $m \rightarrow m^{-1}$ is smooth because any given entry of $m^{-1}$ is the quotient of an $n$th order polynomial in the entries of $m$ by the non-zero, smooth, $\mathbb{C}$-valued function given by the determinant function $\det(\cdot)$. 

25
The second and equivalent way to view $\text{Gl}(n; \mathbb{C})$ as a manifold does not introduce complex numbers until late in the game. This view is ultimately useful because it introduces some constructions that play a role in later chapters. This approach views $\text{Gl}(n; \mathbb{C})$ as a subgroup in $\text{Gl}(2n; \mathbb{R})$. To see $\text{Gl}(n; \mathbb{C})$ in this light, I first introduce the notion of an *almost complex structure*, this an element $j \in \mathbb{M}(2n; \mathbb{R})$ with $j^2 = -\mathbb{I}$ with $\mathbb{I}$ again denoting the identity matrix. The canonical example is $j_0$ whose $n = 1$ and $n = 2$ versions are

$$
\begin{pmatrix}
0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0
\end{pmatrix},
\begin{pmatrix}
0 & -1 \\
1 & 0 \\
0 & 0 \\
0 & 1
\end{pmatrix}
$$

In the $n > 2$ case, $j_0$ is given as follows: Let $e_k$ denote the basis vector for $\mathbb{R}^{2n}$ that has 1 in the entry $k$ and zero in all other entries. Then $j_0$ is defined by the rules $j_0 e_{2n-1} = e_{2n}$ and $j_0 e_{2n} = -e_{2n-1}$. Use $M_j$ for the moment to denote the vector subspace of $\mathbb{M}(2n; \mathbb{R})$ of matrices $m$ such that $mj_0 - j_0 m = 0$. This is to say that the entries of $m$ are such that

$$m_{2k,2j} = m_{2k-1,2j-1} \quad \text{and} \quad m_{2k,2j-1} = m_{2k-1,2j}$$

for each pair $k, j \in \{1, \ldots, n\}$.

As is explained momentarily, the vector space $M_j$ is $\mathbb{M}(n; \mathbb{C})$. In any event, $M_j$ is a vector space over $\mathbb{R}$, and so a Euclidean space. Introduce for the moment $G_j \subset M_j$ to denote the subset of invertible elements. Note that $G_j$ is not empty because the identity matrix $\mathbb{I}$ and also $j_0$ are in $G_j$. As an open subset of a Euclidean space, $G_j$ is a smooth manifold. It is also a group by virtue of the fact that $mm'j_0 = mj_0m' = j_0mm'$ when $m$ and $m'$ are both in $G_j$. Thus it is a Lie group. I shall explain in a moment how to identify $G_j$ with $\text{Gl}(n; \mathbb{C})$.

To see that $M_j$ is $\mathbb{M}(n; \mathbb{C})$, note first that the eigenvalues of $j_0$ are $\pm i$, and any matrix that commutes with $j_0$ must preserve the eigenspaces. Indeed, let $\{e_1, \ldots, e_{2n}\}$ denote the standard basis for $\mathbb{R}^{2n}$, and then a basis of eigenvectors for the eigenvalue $i$ are

$$\{v_1 = e_1 - ie_2, \ldots, v_n = e_{2n-1} - ie_{2n}\}.$$  

The corresponding basis of eigenvectors for the eigenvalue $-i$ is given by the complex conjugates of $\{v_1, \ldots, v_n\}$. If $m$ commutes with $j_0$, then it must act as $m(v) = \tilde{a}_{kj}v_j$ (note the summation convention for repeated indices) where $a_{kj} \in \mathbb{C}$. Likewise, $m(\overline{v}) = \overline{\tilde{a}_{kj}} \overline{v_j}$. Here and in what follows, the overscore indicates complex conjugation. Note that the number $a_{kj} \in \mathbb{C}$ can be written in terms of the entries of $m$ as $a_{kj} = m_{2k-1,2j-1} + \text{im}_{2k,2j-1}$. 

26
The preceding implies that \( m \) is determined by \( n \) complex numbers and so the dimension of \( M_j \) is \( 2n^2 \). The identification between \( M_j \) and \( M(n; \mathbb{C}) \) comes about by associating to any given \( m \in M_j \) the \( n \times n \) matrix with complex entries given by the numbers \( \{a_{ij}\} \). The identification between \( M_j \) and \( M(n; \mathbb{C}) \) identifies the subset \( G_j \subset M_j \) with the group \( \text{Gl}(n; \mathbb{C}) \). To elaborate, suppose that \( m \) is a given element in \( M_j \) and let \( m_c \) denote the corresponding element in \( M(n; \mathbb{C}) \). As an element in \( M(2n; \mathbb{R}) \), the matrix \( m \) has a real valued determinant; and as an element in \( M(n; \mathbb{C}) \), the matrix \( m_c \) has a \( \mathbb{C} \)-valued determinant. These two notions of determinant are such that

\[
\det(m) = |\det(m_c)|^2,
\]

Thus, the real \( 2n \times 2n \) matrix \( m \) is and invertible element in \( M_j \) if and only if the corresponding complex, \( n \times n \) matrix \( m_c \) is an invertible element in \( M(n; \mathbb{C}) \). This correspondence between \( G_j \) and \( \text{Gl}(n; \mathbb{C}) \) has the following property: If \( m \) and \( m' \) are paired, then the correspondence pairs \( mm' \) with \( m_cm'_c \). As a consequence, if the correspondence pairs \( m \in G_j \) with \( m_c \in \text{Gl}(n; \mathbb{C}) \), it then pairs \( m^{-1} \) with \( m_c^{-1} \).

2e) The groups \( \text{Sl}(n; \mathbb{C}), \text{U}(n) \) and \( \text{SU}(n) \)

The three most commonly seen subgroups of \( \text{Gl}(n; \mathbb{C}) \) are the subject of this subchapter.

The group \( \text{Sl}(n; \mathbb{C}) \): This is the subset of matrices in \( M(n; \mathbb{C}) \) with determinant 1. To be precise, the determinant here is that of an \( n \times n \) matrix with complex entries. For example, if the matrix is an \( n \times n \) diagonal matrix, the determinant it the product of the diagonal entries. In particular, the determinant function on \( M(n; \mathbb{C}) \) is to be viewed as a map, \( \det: M(n; \mathbb{C}) \to \mathbb{C} \).

As a parenthetical remark, when \( M(n; \mathbb{C}) \) is viewed as the subvector space \( M_j \subset M(2n; \mathbb{R}) \), it inherits a determinant function from \( M(2n; \mathbb{R}) \), this mapping to \( \mathbb{R} \). Let me call the latter \( \det_R \) for the moment, and call the aforementioned determinant function \( \det_C \). These two functions on \( M_j \) are related by the rule \( \det_R = |\det_C|^2 \). This can be seen as follows: A matrix \( m \in M(n; \mathbb{C}) \) has \( n \) complex eigenvalues \( \{\lambda_1, \ldots, \lambda_n\} \). When viewed in as an element in \( M_j \) and thus in \( M(2n; \mathbb{R}) \), this matrix has \( 2n \) eigenvalues. The first \( n \) are \( \{\lambda_1, \ldots, \lambda_n\} \), and the remaining \( n \) eigenvalues are the latter’s complex conjugates.

In any event, to show that \( \text{Sl}(n; \mathbb{C}) \) is a submanifold, I should prove that the complex determinant \( \det: M(n; \mathbb{C}) \to \mathbb{C} \) has 1 as a regular value. Here is how to do this:
The directional derivative of det at an invertible matrix $m \in \mathbb{M}(n; \mathbb{C})$ in the direction of a some given matrix $a \in \mathbb{M}(n; \mathbb{C})$ is $\text{det}(m)\text{trace}(m^{-1}a)$. To see that any given value in $\mathbb{C}$ can be had by choosing an appropriate matrix $a$, let $c \in \mathbb{C}$ denote the desired value, and then take $a = c \text{det}(m)^{-1} m$.

The group $U(n)$: This is the subset of matrices $m \in \mathbb{M}(n; \mathbb{C})$ that obey $m^\dagger m = 1$. This is to say that $U(n)$ is the subgroup of matrices $m$ with $m^{-1} = m^\dagger$. Here, $m^\dagger$ is the matrix whose $(i, j)$ entry is the complex conjugate of the $(j, i)$ entry of $m$.

To prove that $U(n)$ is a Lie group, I need only verify that it is a submanifold on $\mathbb{M}(n; \mathbb{C})$.

To start, introduce the vector space $\text{Herm}(n)$ of $n \times n$ matrices $\mathfrak{h}$ that obey $\mathfrak{h}^\dagger = \mathfrak{h}$. I remind you that such a matrix is said to be Hermitian. This is a vector space over $\mathbb{C}$ of dimension $\frac{1}{2} n(n + 1)$; its complex coordinates can be taken to be the entries on and above the diagonal. Define a map $\psi: \mathbb{M}(n; \mathbb{C}) \rightarrow \text{Herm}(n)$ by the rule $\psi(m) = m^\dagger m$. Since $U(n) = \psi^{-1}(1)$, this subset is a submanifold if the identity matrix is a regular value of $\psi$. To see that such is the case, I note that the differential $\psi$, at a matrix $m$ sends a given matrix $a$ to $\psi_{m|a} = m^\dagger a + a^\dagger m$. The differential $\psi_{m|a}$ is surjective if and only if any given $\mathfrak{h} \in \text{Herm}(n)$ can be written as $m^\dagger a + a^\dagger m$ for some $a \in \mathbb{M}(n; \mathbb{C})$. Given that $m$ is invertible, then so is $m^\dagger$; its inverse is $(m^{-1})^\dagger$. This understood, take $a = \frac{1}{2} (m^\dagger)^{-1} \mathfrak{h}$.

The group $SU(n)$: This is the subgroup on $U(n)$ of matrices $m$ with $\text{det}(m) = 1$.

Note that if $m \in U(n)$, then $|\text{det}(m)| = 1$, since the equation $m^\dagger m = 1$ implies that $\text{det}(m^\dagger m) = \text{det}(1)$, and so $\text{det}(m^\dagger) \text{det}(m) = 1$. Given that the determinant of $m^\dagger$ is the complex conjugate of the determinant of $m$, this implies that $|\text{det}(m)|^2 = 1$. To see that $SU(n)$ is a Lie group, it is sufficient to prove that it is a submanifold of $\mathbb{M}(n; \mathbb{C})$. For this purpose, define a map $\phi: \mathbb{M}(n; \mathbb{C}) \rightarrow \text{Herm}(n) \times \mathbb{R}$ by the rule $\phi(m) = (m^\dagger m, -\frac{1}{2} (\text{det}(m) - \text{det}(m^\dagger)))$.

Noting that $\text{det}(m^\dagger)$ is the complex conjugate of $\text{det}(m)$, it follows that $SU(n)$ is the component of $\phi^{-1}((1, 0))$ that consists of matrices with determinant equal to 1. The other component consists of matrices with determinant equal to -1.

To see that $\phi$ is surjective, I must verify that its differential at any given $m \in SU(n)$ is surjective. The latter at a matrix $m$ sends any given matrix $a \in \mathbb{M}(n; \mathbb{C})$ to the pair in $\text{Herm}(n) \times \mathbb{R}$ whose first component is $m^\dagger a + a^\dagger m$ and whose second component is the imaginary part of $\text{det}(m) \text{trace}(m^{-1}a)$. Let $(\mathfrak{h}, t)$ denote a given element in $\text{Herm}_n \times \mathbb{R}$. Given that $m \in SU(2)$, then $a = \frac{1}{2} mh - itm$ is mapped by $\phi_{m|a}$ to $(\mathfrak{h}, t)$.

2f) Notation with regards to matrices and differentials

Suppose that $f$ is a function on $\mathbb{M}(n; \mathbb{R})$. Since the entries of a matrix serve as the coordinates, the differential of $f$ can be written as
\[ df = \sum_{1 \leq i,j \leq n} \left( \frac{\partial}{\partial m_{ij}} f \right) \, dm_{ij}. \]

I remind you that this differential form notation says the following: The directional derivative of \( f \) in the direction of a matrix \( a \) is

\[ \sum_{1 \leq i,j \leq n} \left( \frac{\partial}{\partial m_{ij}} f \right) \, a_{ij} \]

The latter can be written compactly as \( \text{trace}(\frac{\partial}{\partial m} f \, a^T) \) if it is agreed to introduce \( \frac{\partial}{\partial m} f \) to denote the matrix whose \((i,j)\) component is \( \frac{\partial}{\partial m_{ij}} f \). Here, \((\cdot)^T\) denotes the transpose of the indicated matrix.

The preceding observation observation suggests viewing the collection of coordinate differentials \( \{ dm_{ij} \} \) as the entries of a matrix of differentials, this denoted by \( dm \). Introducing this notation allows us to write \( df \) above as

\[ \text{trace}(\frac{\partial}{\partial m} f \, dm^T). \]

Here, \( dm^T \) is the matrix of differentials whose \((i,j)\) entry is \( dm_{ji} \). This notation is commonly used, and in particular, it will be used in subsequent parts of this book.

An analogous short hand is used when \( f \) is a function on \( \mathbb{M}(n; \mathbb{C}) \). In this case, the entries of a matrix serve as the complex coordinates for \( \mathbb{M}(n; \mathbb{C}) \). This is to say that their real and imaginary parts give \( \mathbb{M}(n; \mathbb{C}) \) its Euclidean coordinates as a manifold. The differential of the function \( f \) can be written concisely as

\[ \text{trace}(\frac{\partial}{\partial m} f \, dm^T) + \text{trace}(\frac{\partial}{\partial m} f \, dm^T). \]

Note that this notation makes sense even when \( f \) is complex valued. For that matter, it makes sense when \( f \) is replaced by a map, \( \psi \), from either \( \mathbb{M}(n; \mathbb{R}) \) or \( \mathbb{M}(n; \mathbb{C}) \) to a given vector space, \( V \). In the case of \( \mathbb{M}(n; \mathbb{R}) \), the differential of \( \psi \) can be written succinctly as

\[ \psi_* = \text{trace}(\frac{\partial}{\partial m} \psi \, dm^T). \]

To disentangle all of this, suppose that \( V \) has dimension \( q \), and a given a basis with respect to which \( \psi \) has coordinates \( \{ \psi^1, \ldots, \psi^q \} \). Then what is written above says the following: The differential of \( \psi \) at a given matrix \( m \) is the linear map from \( \mathbb{M}(n; \mathbb{R}) \) to \( V \) that sends a matrix \( a \) to the vector with coordinates
\[
(\sum_{1 \leq i, j \leq n} \frac{\partial}{\partial m_j} \psi^i a_{ij}, \ldots, \sum_{1 \leq i, j \leq n} \frac{\partial}{\partial m_j} \psi^i a_{ij})
\]

There is, of course, the analogous notation for the case when \( \psi \) maps \( \mathbb{M}(n; \mathbb{C}) \) to a vector space.

Appendix: The transition functions for the Grassmannians

This appendix ties up a loose end from Chapter 1f; it supplies a proof that the transition functions between any two coordinate charts in a given Grassmannian is a smooth map. To set the stage, fix integers \( m > 1 \) and \( n \in \{1, \ldots, m\} \) so to define the Grassmannian \( \text{Gr}(m; n) \) as described in Chapter 1f. Let \( V \) and \( V' \) denote two elements in \( \text{Gr}(m; n) \) and let \( \mathcal{O}_V \) and \( \mathcal{O}_{V'} \) denote the respective coordinate charts from Chapter 1f. Coordinate maps \( \phi_V^{-1}: \mathbb{R}^{n(m-n)} \rightarrow \mathcal{O}_V \) and \( \phi_{V'}^{-1}: \mathbb{R}^{n(m-n)} \rightarrow \mathcal{O}_{V'} \) are described. What follows explains why \( \phi_{V'} \circ \phi_V^{-1} \) is smooth on its domain of definition.

To start this task, introduce \( \{e_i\}_{1 \leq i \leq n} \) and \( \{u_a\}_{1 \leq a \leq m-n} \) to denote the respective orthonormal bases chosen for \( V \) and \( V' \). Use \( \{e'_i\}_{1 \leq i \leq n} \) and \( \{u'_a\}_{1 \leq a \leq m-n} \) to denote their respective \( V' \) counterparts. Now let \( W \subset \mathcal{O}_V \cap \mathcal{O}_{V'} \) and let \( p \in \mathbb{R}^{n(m-n)} \) and \( p' \in \mathbb{R}^{n(m-n)} \) denote respective images under \( \phi_V \) and \( \phi_{V'} \). Identify \( \mathbb{R}^{n(m-n)} \) with the vector space of matrices with \( n \) rows and \( m \) columns so as to write the components of \( p \) as \( \{p_{ia}\}_{1 \leq i \leq n, 1 \leq a \leq m-n} \) and those of \( p' \) as \( \{p'_{ia}\}_{1 \leq i \leq n, 1 \leq a \leq m-n} \). By definition, the coordinate \( p = \phi_V(W) \) is such that any given \( w \in W \) can be parametrized by \( \mathbb{R}^n \) as

\[
w = \sum_{1 \leq i \leq n} v_i (e_i + \sum_{1 \leq a \leq m-n} p_{ia} u_a)
\]

with \( v \in \mathbb{R}^n \). Likewise, the coordinate \( p' = \phi_{V'}(W) \) is defined so that

\[
w = \sum_{1 \leq i \leq n} v'_i (e'_i + \sum_{1 \leq a \leq m-n} p'_{ia} u'_a)
\]

with \( v' \in \mathbb{R}^n \). Note in particular that any given \( w \in W \) is parametrized in two ways by points in \( \mathbb{R}^n \). This being the case, the two vectors \( v \) and \( v' \) must be functionally related. To be explicit, the \( i \)'th component of \( v' \) can be written in terms of \( v \) by taking the inner product of \( e'_i \), with \( w \) and using the preceding equations to first identify the latter with \( v' \), and then using the other equation above to write this inner product in terms of \( v \). The result of equating these two expressions for \( e'_i \cdot w \) is

\[
v' = \sum_{1 \leq k, l \leq n} \sum_{1 \leq a \leq m-n} p_{ka} e'_l \cdot u_a
\]
where $\mathbf{x} \cdot \mathbf{x}'$ denotes the Euclidean inner product on $\mathbb{R}^m$. This last equation asserts that $v'_i = \sum_{1 \leq k \leq n} G_{ik} v_k$ where the numbers $\{G_{ik}\}_{1 \leq i,k \leq n}$ are

$$G_{ik} = e'_i \cdot e_k + \sum_{1 \leq a \leq n} p'_{ia} e'_i \cdot e_k.$$ 

These numbers can be viewed as the entries of an $n \times n$ matrix $G$. Note that this matrix $G$ is invertible when the respective projections $\Pi_V$ and $\Pi_{V'}$ map $W$ isomorphically to $V$ and $V'$. This is to say that $G \in \text{Gl}(n; \mathbb{R})$ when $W \in \mathcal{O}_V \cap \mathcal{O}_{V'}$.

The assignment of $p \in \mathbb{R}^{n(m-n)}$ to $G$ defines a smooth map from $\varphi_V(\mathcal{O}_V \cap \mathcal{O}_{V'})$ into $\mathbb{R}^{n(m-n)}$ into $\text{Gl}(n; \mathbb{R})$ since the expression above for $G$’s entries define an affine function of the coordinate functions on $\mathbb{R}^{n(m-n)}$. Save the matrix $G$ for the moment.

Use the expression that parametrizes $w \in W$ in terms of $v' \in \mathbb{R}^n$ to write $u'_a \cdot w$ in terms of $v'$ as

$$\sum_{1 \leq i \leq n} v'_i p'_{ia}.$$ 

Meanwhile, $u'_a \cdot w$ can also be written in terms of $v$ using the expression that writes $w$ in terms of $v$:

$$\sum_{1 \leq i \leq n} v_k (u'_a \cdot e_k + \sum_{1 \leq b \leq m-n} p_{kb} u'_a \cdot u_b).$$

Keeping in mind that these two expressions are equal, take $v'$ to be $i$’th basis vector in $\mathbb{R}^n$ to see that

$$p'_{ia} = \sum_{1 \leq k \leq n} (G^{-1})_{ik} (u'_a \cdot e_k + \sum_{1 \leq b \leq m-n} p_{kb} u'_a \cdot u_b).$$

Here, $G^{-1}$ is the inverse of the matrix $G$. With the entries of $G$ viewed as a function of $p \in \varphi_V(\mathcal{O}_V \cap \mathcal{O}_{V'})$, the latter expression writes $p'$ as a function of $p$. This function is the transition function $\varphi_{V'} \circ \varphi_V^{-1}$. This expression for $p'$ as a function on $\varphi_V(\mathcal{O}_V \cap \mathcal{O}_{V'})$ defines a smooth function on $\varphi_V(\mathcal{O}_V \cap \mathcal{O}_{V'})$ because the inverse map from $\text{Gl}(n; \mathbb{R})$ to itself is a smooth map.

**Additional reading**

• *Lie Groups, Beyond an Introduction*, Anthony W. Knapp; Birkhauser Boston, 1966
CHAPTER 3: Introduction to Vector Bundles

Euclidean space is a vector space, but the typical manifold is not. Vector bundles bring vector spaces back into the story. To a first approximation, a vector bundle (over a given manifold) is a family of vector spaces that is parametrized by the points in the manifold. Some sort of vector bundle appears in most every aspect of differential geometry. This chapter introduces these objects and gives the first few examples. By way of a warning, I give my favorite definition first; the standard definition is given subsequently in Chapter 3b. (These definitions describe the same objects).

3a) The definition

Let $M$ denote a smooth manifold of dimension $m$. A real vector bundle over $M$ of fiber dimension $n$ is a smooth manifold, $E$, of dimension $n + m$ with the following additional structure:

- There is a smooth map $\pi: E \to M$.
- There is a smooth map $\delta: M \to E$ such that $\pi \circ \delta$ is the identity.
- There is a smooth map $\mu: \mathbb{R} \times E \to E$ such that
  a) $\pi(\mu(r, v)) = \pi(v)$
  b) $\mu(r, \mu(r', v)) = \mu(rr', v)$
  c) $\mu(1, v) = v$
  d) $\mu(r, v) = v$ for $r \neq 1$ if and only if $v \in \text{image}(\delta)$.
- Let $p \in M$. There is a neighborhood, $U \subset M$, of $p$ and a map $\lambda_U: \pi^{-1}(U) \to \mathbb{R}^n$ such that $\lambda_U: \pi^{-1}(x) \to \mathbb{R}^n$ for each $x \in U$ is a diffeomorphism obeying $\lambda_U(\mu(r, v)) = r\lambda_U(v)$.

The first four remarks that follow concern notation. The remaining two remarks are more substantive.

Remark 1: The map from $E$ to $E$ given by $v \to \mu(r, v)$ for any given $r \in \mathbb{R}$ is called multiplication by $r$; this is because the map $\mu$ defines an action on $E$ of the multiplicative group of real numbers. This $\mathbb{R}$ action is reflected in the notation used henceforth that writes $\mu(r, v)$ as $rv$.

Remark 2: If $W \subset M$ is any given set, then $\pi^{-1}(W) \subset E$ is denoted by $E|_W$. The inverse images via $\pi$ of the points in $M$ are called the fibers of $E$; and if $p \in M$ is any given point, then the fiber $\pi^{-1}(p)$ is said to be the fiber over $p$. The map $\pi$ itself is called the bundle projection map, or just the projection map.
Remark 3: The map \( \hat{\delta} \) is called the zero section. The image of \( \hat{\delta} \) is also called the zero section; but it is almost always the case that these two meaning are distinguished by the context. This terminology reflects the fact that item d) of the third bullet and the fourth bullet require \( \lambda_{U} \circ \hat{\delta} = 0 \in \mathbb{R}^n \).

Remark 4: The map \( \lambda_{U} \) defines a diffeomorphism,

\[
\varphi_{U}: E|_{U} \to U \times \mathbb{R}^n
\]

given by \( \varphi_{U}(v) = (\pi(v), \lambda_{U}(v)) \). It is customary to talk about \( \varphi_{U} \) rather than \( \lambda_{U} \). The map \( \varphi_{U} \) is said to be a local trivialization for \( E \).

Remark 5: Let \( U \subset M \) denote an open set where there is a map \( \lambda_{U}: E|_{U} \to \mathbb{R}^n \) as above. The map \( \lambda_{U} \) can be used to give a local coordinate chart map for \( E \) in the following way: Let \( (W, \varphi: W \to \mathbb{R}^{\dim(M)}) \) denote a coordinate chart with \( W \subset U \). Then the pair

\[
(E|_{W}, (\varphi, \varphi_{U}): E|_{W} \to \mathbb{R}^{\dim(M)} \times \mathbb{R}^n = \mathbb{R}^{\dim(M)+n})
\]

is a coordinate chart for \( E \).

Remark 6: The definition given above endows each fiber of \( E \) with a canonical vector space structure. To see how this comes about, note first that each fiber is diffeomorphic to \( \mathbb{R}^n \); this being an obvious necessary condition. Second, the \( \mathbb{R} \) action on \( E \) gives an \( \mathbb{R} \) action on each fiber, and this gives \( \mathbb{R} \) action for the vector space structure. Third, the zero section \( \hat{\delta} \) gives each fiber an origin for the vector space structure.

Granted these points, I need only describe the notion of vector addition on each fiber. This is done using a local trivialization. To elaborate, let \( p \in M \) and let \( U \subset M \) denote an open set with a local trivializing map \( \varphi_{U}: E|_{U} \to U \times \mathbb{R}^n \). Write the latter map as \( v \to (\pi(v), \lambda(v)) \). If \( v \) and \( v' \) are in \( E|_{p} \), define \( v + v' \) to be \( \lambda_{U,p}^{-1}(\lambda_{U,p}(v) + \lambda_{U,p}(v')) \) where \( \lambda_{U,p}^{-1} \) is the inverse of the diffeomorphism from \( E|_{p} \) to \( \mathbb{R}^n \) that is defined by \( \lambda_{U,p} \).

To see that this makes sense, I must demonstrate that the addition so defined does not depend on the chosen set \( U \). The point here is that a given point \( p \) may lie in two sets, \( U \) and \( U' \), that come with corresponding maps \( \lambda_{U} \) and \( \lambda_{U'} \). I need to prove that the same point \( v + v' \) is obtained when using \( \lambda_{U'} \). Now, this follows if and only if

\[
\lambda_{U',p}(\lambda_{U,p}^{-1}(e + e')) = \lambda_{U',p}(\lambda_{U,p}^{-1}(e)) + \lambda_{U',p}(\lambda_{U,p}^{-1}(e'))
\]

for all pairs \( e \) and \( e' \in \mathbb{R}^n \). This is to say that the the map
\[
\psi_{U \cup U'} = \lambda_U \circ (\lambda_{U'}^{-1}): (U \cap U') \times \mathbb{R}^n \to \mathbb{R}^n
\]

must be linear in the \(\mathbb{R}^n\) factor; it must have the form \((p, v) \to g_{U \cup U'}(p)v\) where \(g_{U \cup U'}\) is a smooth map from \(U \cap U'\) to \(\text{Gl}(n; \mathbb{R})\). That such is the case is a consequence of the next lemma.

**Lemma:** A map \(\psi: \mathbb{R}^n \to \mathbb{R}^n\) with the property that \(\psi(rv) = r\psi(v)\) for all \(r \in \mathbb{R}\) is linear.

**Proof:** Use the chain rule to write \(\frac{d}{dt} \psi(tv) = \psi'_t |_{tv} \cdot v\). This is also equal to \(\psi(v)\) since \(\psi(tv) = t \psi(v)\). Since the right hand side of this last equality is independent of \(t\), so is the left hand side. Thus, the map \(t \to \psi'_t |_v \cdot v\) is independent of \(t\), and so equal to \(\psi'_0 |_v \cdot v\). As a consequence, \(\psi(v) = \psi'_0 |_v \cdot v\).

The map \(g_{U \cup U'}: U \cap U' \to \text{Gl}(n; \mathbb{R})\) is said to be a bundle transition function for the vector bundle.

### 3b) The standard definition

The observation that the map \(\psi_{U \cup U'}\) depicted above is linear in the \(\mathbb{R}^n\) factor suggests an alternate definition of a vector bundle. This one, given momentarily, is often the more useful of the two; and it is the one you will find in most other books.

In this definition, a real vector bundle of fiber dimension \(n\) is given by the following data: First, a locally finite open cover, \(\mathcal{U}\), of \(M\). Second, an assignment of a map, \(g_{U \cup U'}: U \cap U' \to \text{Gl}(n; \mathbb{R})\) to each pair \(U, U' \in \mathcal{U}\) subject to two constraints. First, \(g_{UU'} = g_{U'U}^{-1}\). Second, if \(U, U'\) and \(U''\) are any three sets from \(\mathcal{U}\) with \(U \cap U' \cap U'' \neq \emptyset\), then \(g_{U \cup U'}g_{UU''}g_{U''U} = 1\). By way of notation, the map \(g_{U \cup U'}\) is deemed the bundle transition function. Meanwhile, the constraint \(g_{U \cup U'}g_{UU''}g_{U''U} = 1\) is said to be a cocycle constraint.

Granted this data, the bundle \(E\) is defined as the quotient of the disjoint union \(\bigcup_{U \in \mathcal{U}} (U \times \mathbb{R}^n)\) by the equivalence relation that puts \((p', v') \in U' \times \mathbb{R}^n\) equivalent to \((p, v) \in U \times \mathbb{R}^n\) if and only if \(p = p'\) and \(v' = g_{U \cup U}(p)v\). To connect this definition with the previous one, define the map \(\pi\) to send the equivalence class of any given \((p, v)\) to \(p\). The multiplication map is defined by the rule that has \(r \in \mathbb{R}\) sending the equivalence class of the point \((p, v)\) to \((p, rv)\). Finally, any given \(U \in \mathcal{U}\) has the corresponding map \(\lambda_U\) that os defined as follows: If \(p \in U\), then any given equivalence class in \(\pi^{-1}(p)\) has a unique representative \((p, v) \in U \times \mathbb{R}^n\). This equivalence class is sent by \(\lambda_U\) to \(v \in \mathbb{R}^n\).

The smooth structure on a vector bundle \(E\) defined this is that associated to the following sort of coordinate atlas: If necessary, subdivide each set \(U \in \mathcal{U}\) so that the
resulting open cover consists of a set \( U' \) where each set \( U \in U' \) comes with a smooth coordinate map \( \phi_U: U \to \mathbb{R}^{\dim(M)} \) that embeds \( U \) as the interior of a ball. This is to say that the collection \( \{ (U, \phi_U: U \to \mathbb{R}^{\dim(M)}) \}_{U \in U'} \) consists of a coordinate atlas for \( M \). This understood, then the collection \( \{ (E|_U, (\phi_U, \lambda_U): E|_U \to \mathbb{R}^{\dim(M)} \times \mathbb{R}^n) \}_{U \in U'} \) defines a coordinate atlas for \( E \). Note that the coordinate transition function for the intersection of two such charts, thus for \( E|_{U' \cap U} \), sends a pair \((x, v) \in \text{image}(\phi_U) \times \mathbb{R}^n\) to the point with \( \mathbb{R}^\dim(M) \) coordinate \((\phi_{U'} \circ \phi_U^{-1})(x) \in \text{image}(\phi_U)\) and \( \mathbb{R}^n \) coordinate \( g_{U'U}(x)v \). Thus, the bundle transition functions \( \{ g_{U'U} \}_{U, U' \in U'} \) help to determine the smooth structure on \( E \).

The data consisting of the cover \( U \) with the bundle transition functions \( \{ g_{U'U} \}_{U, U' \in U} \) is called \textit{cocycle data} for the bundle \( E \).

3c. The first examples of vector bundles

What follows are some first examples of vector bundles.

The \textbf{trivial bundle}: The vector bundle \( M \times \mathbb{R}^n \) over \( M \) with fiber \( \mathbb{R}^n \) called the \textit{trivial} \( \mathbb{R}^n \) bundle for what I hope is an obvious reason. It is also often called the \textit{product} bundle.

The \textbf{Mobius bundle}: Introduce the standard angular coordinate \( \theta \in \mathbb{R}/(2\pi\mathbb{Z}) \) for \( S^1 \). With \( S^1 \) viewed as the unit radius circle about the origin in \( \mathbb{R}^2 \), a given angle \( \theta \) maps to the point \((\cos(\theta), \sin(\theta))\). Let \( E \subset S^1 \times \mathbb{R}^2 \) denote the subset of points \((\theta, (v_1, v_2))\) such that \( \cos(\theta)v_1 + \sin(\theta)v_2 = v_1 \) and \( \sin(\theta)v_1 - \cos(\theta)v_2 = v_2 \). This defines a bundle over \( S^1 \) with fiber dimension 1.

The \textbf{tautological bundle} over \( \mathbb{RP}^n \): Define \( \mathbb{RP}^n \) as \( S^n/\{\pm1\} \). A bundle \( E \to \mathbb{RP}^n \) with fiber \( \mathbb{R} \) is defined to be the quotient \( (S^n \times \mathbb{R})/\{\pm1\} \). Alternately, \( E \) is the set \((x, \gamma)\) such that \( x \in \mathbb{RP}^n \) and \( \gamma \) is the line through the origin with tangent vector proportional to \( \pm x \).

\textbf{Bundles over the Grassmannians}: Fix positive integer \( m > n \) and introduce from Chapter 1f the manifold \( \text{Gr}(m; n) \) whose points are the \( n \)-dimensional subspaces of \( \mathbb{R}^m \). Over \( \text{Gr}(m; n) \) sits an essentially tautological \( \mathbb{R}^n \) bundle that is defined given as a set by the set of pairs \( \{(V, v): V \subset \mathbb{R}^m \text{ is an } n \text{-dimensional subspace and } v \in V\} \). Let \( E \) denote this bundle. The projection map \( \pi: E \to \text{Gr}(m; n) \) sends a given pair \((V, v)\) to the \( n \)-dimensional subspace \( V \in \text{Gr}(m; n) \). The \( \mathbb{R} \) action on \( E \) has \( r \in \mathbb{R} \) sending the pair \((V, v)\) to the pair \((V, rv)\). The map \( \delta \) sends \( V \in \text{Gr}(m; n) \) to the pair \((V, 0)\) where 0 \in \mathbb{R}^m \) is the origin. Let \( O_V \subset \text{Gr}(m; n) \) denote the subspace of \( n \)-dimensional subspaces \( V' \) which are
mapped isomorphically via the orthogonal projection \( \Pi_V: \mathbb{R}^m \rightarrow V \). The set \( U = \mathcal{O}_V \) has the required map \( \lambda_u: \pi^1(U) \rightarrow \mathbb{R}^n \), this defined as follows: Fix a basis for \( V \) so as to identify \( V \) with \( \mathbb{R}^n \). Granted this identification, define the map \( \lambda_{OV} \) as so as to send any given pair \((V', v) \in E|_{OV}\) to \( \Pi_V v' \in V = \mathbb{R}^n \). I leave it as an exercise to verify that \( E \) as just defined has a suitable smooth structure. This bundle \( E \) is called the \textit{tautological} bundle over \( \text{Gr}(m; n) \).

**Bundles and maps from spheres:** The next example uses cocycle data to define a vector bundle. To start, let \( M \) denote the manifold in question and let \( m \) denote its dimension. Fix a finite set \( \Lambda \subset M \). Assign to each \( p \in \Lambda \) the following data: First, a smooth map \( g_p: S^{m-1} \rightarrow \text{Gl}(n; \mathbb{R}) \). Second, a coordinate chart \((U, \varphi_p): U \rightarrow \mathbb{R}^m\) with \( p \in U \) such that \( \varphi_p(p) = 0 \in \mathbb{R}^m \). Fix \( r > 0 \) and small enough so that the open ball, \( B \), of radius \( r \) about the origin in \( \mathbb{R}^m \) is in the image of each \( p \in \Lambda \) version of \( \varphi_p \); and so that the collection \( \{U_p = \varphi_p^{-1}(B)\}_{p \in \Lambda} \) consists of pairwise disjoint sets.

To construct a vector bundle from this data, use the cover given by the open sets \( U_0 = M - \Lambda \) and \( \{U_p\}_{p \in \Lambda} \). The only non-empty intersections are for this cover are those of the form \( U_0 \cap U_p = \varphi_p^{-1}(B - 0) \). Define the bundle transition function \( g_{U_p}: U_0 \cap U_p \rightarrow \text{Gl}(n; \mathbb{R}) \) on such a set to be the map \( x \rightarrow \varphi_p|_{\varphi_p^{-1}(x)}(B - 0) \). As there are no non-empty intersections between three distinct sets in the cover, there are no cocycle conditions to satisfy. As a consequence, this data is sufficient to define a vector bundle over \( M \) having fiber dimension \( n \).

**3d) The tangent bundle**

Every smooth manifold has comes a priori with certain canonical vector bundles, one of which is the \textit{tangent bundle}, this a bundle whose fiber dimension is that of the manifold in question. When \( M \) is the manifold in question, its tangent bundle is denoted by \( TM \). What follows is a cocycle definition of \( TM \).

Let \( n \) denote the dimension of \( M \). Fix a locally finite coordinate atlas, \( \mathcal{U} \), for \( M \). The collection of open sets from the pairs in \( \mathcal{U} \) is used to define the tangent bundle by specifying the bundle transition functions \( \{g_{U'U}: U' \cap U \rightarrow \text{Gl}(n; \mathbb{R})\}_{(U, \varphi), (U', \varphi') \in \mathcal{U}} \).

Fix a pair \((U, \varphi_U)\) and \((U', \varphi_U')\) from the atlas \( \mathcal{U} \). To define \( g_{U'U} \), introduce the coordinate transition function \( \psi_{U'U} = \varphi_{U'} \circ \varphi_U^{-1} \), a diffeomorphism from \( \varphi_U(U' \cap U) \subset \mathbb{R}^n \) to \( \varphi_U(U' \cap U) \subset \mathbb{R}^n \). The differential of this map, \( \psi_{U'U,*} \), is a map from \( \psi_U(U' \cap U) \) to \( \text{Gl}(n; \mathbb{R}) \). This understood, set \( g_{U'U}(p) = \psi_{U'U,|_{\varphi_U(p)}} \) for \( p \in U' \cap U \). The cocycle condition \( g_{U'U}g_{U'U''}g_{U''U} = \iota \) is guaranteed by the Chain rule by virtue of the fact that \( \psi_{U'U"} \circ \psi_{U'U} \circ \psi_{U''U'} \) is the identity map from \( \psi_{U'}(U' \cap U) \) to itself.

37
This definition appears to depend on the choice of the atlas $\mathcal{U}$ from the equivalence class that defines the smooth structure. As explained next, the same bundle $TM$ appears for any choice of atlas from this equivalence class. To see why this is, agree for the moment to use $\mathcal{T}\mathcal{U}$ to denote the version of $TM$ defined as above using the atlas $\mathcal{U}$. Suppose that $\mathcal{V}$ is a coordinate atlas that is equivalent to $U$. Let $\mathcal{T}\mathcal{V}$ denote the version of $M$’s tangent bundle that is defined as above from $V$. I now define a coordinate atlas $\mathcal{W}$ as follows: The collection $\mathcal{W}$ is the union of two coordinate atlases. The first, $\mathcal{W}_U$, consists of all pairs of the form $(U \cap V, \varphi_U)$ such that $(U, \varphi_U)$ is from $U$ and $(V, \varphi_V)$ is from $V$. The second atlas, $\mathcal{W}_V$, consists of all pairs of the form $(U \cap V, \varphi_V)$ with $U$, $V$ and $\varphi_V$ as before. Thus, if $(U, \varphi_U) \in U$ and $(V, \varphi_V) \in V$, then the open set $U \cap V$ appears in two pairs from $\mathcal{W}$, one with the map $\varphi_U$ and the other with the map $\varphi_V$. Use $\mathcal{T}\mathcal{W}_U$, $\mathcal{T}\mathcal{W}_V$ and $\mathcal{T}\mathcal{W}$ to denote the versions of the tangent bundle that are defined respectively by $\mathcal{W}_U$, $\mathcal{W}_V$ and $\mathcal{W}$. It follows directly from the definition of the equivalence relation that $\mathcal{T}\mathcal{W}_U = \mathcal{T}\mathcal{W}_V = \mathcal{T}\mathcal{W}$; indeed, such is the case because any given point in $M$ is contained in some version of $U \cap V$. It also follows directly from the equivalence relation that $\mathcal{T}\mathcal{W}_U = \mathcal{T}\mathcal{U}$ and $\mathcal{T}\mathcal{W}_V = \mathcal{T}\mathcal{V}$.

What follows is an important point to make with regards to the definition of $TM$: If $U \subset M$ is any given coordinate chart $\varphi_U: U \to \mathbb{R}^n$, then there is a canonically associated bundle trivialization map from $TM|_U$ to $U \times \mathbb{R}^n$. The latter is denoted by $\varphi_U^*$. 

A geometric interpretation of $TM$: Suppose that $m \geq n$ and that $\sigma: M \to \mathbb{R}^m$ is a smooth map. Suppose, in addition, that $a \in \mathbb{R}^{m-n}$ is a regular value of $\psi$. Recall from Chapter 1 that $M = \psi^\dagger(a)$ is a smooth, $n$-dimensional submanifold of $\mathbb{R}^m$. If $p \in M$, then the linear subspace $\Pi_p$ is the kernel of the $\psi|_p$. Thus, the tangent bundle $TM$ is the subset $\{(p, v): p \in M and \psi|_p(v) = 0 \}$. 

Suppose that $m \geq n$ and that $\psi: \mathbb{R}^m \to \mathbb{R}^{m-n}$ is a smooth map. Suppose, in addition, that $a \in \mathbb{R}^{m-n}$ is a regular value of $\psi$. Recall from Chapter 1 that $M = \psi^\dagger(a)$ is a smooth, n-dimensional submanifold of $\mathbb{R}^m$. If $p \in M$, then the linear subspace $\Pi_p$ is the kernel of the $\psi|_p$. Thus, the tangent bundle $TM$ is the subset $\{(p, v): p \in M and \psi|_p(v) = 0 \}$. 

38
As a parenthetical remark, note that the map from $\mathbb{R}^m \times \mathbb{R}^m$ to $\mathbb{R}^{m-n} \times \mathbb{R}^{m-n}$ that sends any given pair $(p, v)$ to $\Psi(p, v) = (\psi(p), \psi|_p v)$ has $(a, 0) \in \mathbb{R}^{m-n} \times \mathbb{R}^{m-n}$ as a regular value. As noted in the previous paragraph, $\Psi^{-1}(a, 0)$ is TM. Thus, TM sits in $\mathbb{R}^{m-n} \times \mathbb{R}^m$ as a submanifold. It is left as an exercise to check that its smooth structure as $M$'s tangent bundle is the same as that coming from the identification just given as the inverse image of a regular value of $\Psi$.

3e) Tangent bundle examples

What follows are some tangent bundle examples.

The manifold $\mathbb{R}^n$: The manifold $\mathbb{R}^n$ has a canonical coordinate atlas, this the atlas with one chart, and with the associated map to $\mathbb{R}^n$ being the identity. Using this coordinate atlas $U = \{(\mathbb{R}^n, \text{the identity map})\}$ as a fiducial atlas identifies $T\mathbb{R}^n$ with the product bundle $\mathbb{R}^n \times \mathbb{R}^n$ with the bundle projection $\pi$ given by projection to the left most $\mathbb{R}^n$ factor.

An open subset of $\mathbb{R}^n$: An open set $M \subset \mathbb{R}^n$ has the fiducial coordinate atlas $U = \{(M \cap \mathbb{R}^n, \text{the identity map } M \rightarrow \mathbb{R}^n)\}$. Using this chart identifies $M$'s tangent bundle with $M \times \mathbb{R}^n$ also. For example, the group $\text{Gl}(n; \mathbb{R})$ sits as an open set in the Euclidean space $\mathbb{M}(n; \mathbb{R})$. Thus, $T\text{Gl}(n; \mathbb{R})$ can be canonically identified with $\text{Gl}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R})$. Likewise, $T\text{Gl}(n; \mathbb{C})$ can be canonically identified with $\text{Gl}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C})$.

The sphere: The sphere $S^n \subset \mathbb{R}^{n+1}$ is the set of vectors $x \in \mathbb{R}^{n+1}$ with $|x| = 1$. This is $\psi^{-1}(1)$ where $\psi: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is the map sending $x$ to $|x|^2$. The differential of this map at a given point $x \in \mathbb{R}^{n+1}$ sends $v \in \mathbb{R}^{n+1}$ to $x \cdot v$ where $\cdot$ denotes the Euclidean inner product. This being the case, $TS^n$ sits in $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ as the set of pairs $\{(x, v): |x| = 1 \text{ and } v \text{ is orthogonal to } x\}$.

A surface in $\mathbb{R}^3$: Take $g \geq 0$ and construct a surface of genus $g$ embedded in $\mathbb{R}^3$ as described in Chapter 1c using a suitable function $h: \mathbb{R}^2 \rightarrow \mathbb{R}$. Recall that this surface is the level set with value 0 for the function $(x, y, z) \rightarrow z^2 - h(x, y)$. The tangent bundle to this surface sits in $\mathbb{R}^3 \times \mathbb{R}^3$ as the set of pairs $(p = (x, y, z), v = (v_x, v_y, v_z))$ with

$$z^2 - h(x, y) = 0 \quad \text{and also} \quad 2z v_z - (\frac{\partial h}{\partial x})(x, y) v_x - (\frac{\partial h}{\partial y})(x, y) v_y = 0.$$
The group SL(n; \mathbb{R}): This manifold sits inside \mathbb{M}(n; \mathbb{R}) as \psi^{-1}(1) where \psi is the map \psi from \mathbb{M}(n; \mathbb{R}) to \mathbb{R} given by \psi(m) = \det(m). The differential of this map at an invertible matrix \(m\) sends \(a \in \mathbb{M}(n; \mathbb{R})\) to \(\det(m) \trace(m^{-1} a)\). The kernel of this is the vector subspace in \(\mathbb{M}(n; \mathbb{R})\) consisting of matrices \(a = mc: \trace(c) = 0\). Thus, the tangent bundle to \(\text{SL}(n; \mathbb{R})\) sits in \(\mathbb{M}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R})\) as the submanifold consisting of the pairs \((m, mc): \det(m) = 1 \text{ and } \trace(c) = 0\).

The group O(n): The group O(n) sits in \(\mathbb{M}(n; \mathbb{R})\) as \(\psi^{-1}(\mathbb{I})\) where \(\psi\) is the map from \(\mathbb{M}(n; \mathbb{R})\) to the Euclidean space \(\text{Sym}(n; \mathbb{R})\) of \(n \times n\) symmetric matrices that sends a matrix \(m\) to \(m^T m\). The differential of \(\psi\) at \(m\) sends \(a \in \mathbb{M}(n; \mathbb{R})\) to \(m^T a + a^T m\). The kernel of this map at \(m \in \text{SO}(n)\) is the vector subspace \(\{a = mc: c^T = -c\}\). Thus, \(\text{TO}(n)\) sits in \(\mathbb{M}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R})\) as the submanifold of pairs \((m, mc): m^T m = 1 \text{ and } c^T = -c\).

The groups U(n) and SU(n): The group U(n) sits in \(\mathbb{M}(n; \mathbb{C})\) as the inverse image of \(\mathbb{I}\) via the map to the space \(\text{Herm}(n)\) of \(n \times n\) Hermitian matrices that sends a matrix \(m\) to \(m^\dagger m\). The kernel of the differential of this map at \(m \in \text{U}(n)\) consists of the matrices \(a \in \mathbb{M}(n; \mathbb{C})\) that can be written as \(a = mc\) where \(c^\dagger = -c\). This understood, \(\text{TU}(n)\) sits in \(\mathbb{M}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C})\) as the submanifold of pairs \((m, mc): m^\dagger m = \mathbb{I} \text{ and } c^\dagger = -c\). Meanwhile, \(\text{TSU}(n)\) sits in \(\mathbb{M}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C})\) as the submanifold of pairs \((m, mc): \text{both } m^\dagger m = \mathbb{I} \text{ and } \det(m) = 1, \text{ and both } c^\dagger = -c \text{ and } \trace(c) = 0\).

3f) The cotangent bundle

The cotangent bundle is the second of the two canonical bundles that are associated to \(M\). This one is denoted \(T^*M\); and it has fiber dimension \(n = \dim(M)\). What follows is the cocycle definition of \(T^*M\).

Fix a coordinate atlas \(\mathcal{U}\) for \(M\). As with the tangent bundle, the collection of open sets that appear in the pairs from \(\mathcal{U}\) supply the open cover for the cocycle definition. To define the bundle transition functions, suppose that \((U, \varphi_U)\) and \((U', \varphi_{U'})\) are two charts from \(\mathcal{U}\). Introduce the coordinate transition function \(\psi_{U'U} = \varphi_{U'}^{-1} \circ \varphi_U\). As noted a moment ago, this is a diffeomorphism from one open set in \(\mathbb{R}^n\) to another. In particular its differential \(\psi_{U'U}^*\) maps \(\psi_U(U' \cap U)\) to \(\text{Gl}(n; \mathbb{R})\). This understood, the bundle transition function \(g_{U'U}\) for the cotangent bundle is defined so that its value at any given point \(p \in U' \cap U\) is the transpose of the inverse of the matrix \(\psi_{U'U}^*|_{\varphi_U(p)}\). Said differently, any given bundle transition function for \(T^*M\) is obtained from the corresponding bundle transition function for \(TM\) by taking the transpose of the inverse of the latter.
The same argument that proves TM to be independent of the representative coordinate atlas proves that T*M is likewise independent of the representative atlas.

Here is a final remark: Let U ⊂ M denote any given coordinate chart, and let φ^U denote the corresponding map from U to \( \mathbb{R}^n \). It follows from the definition of T*M that the latter has a canonically associated bundle trivialization map from T*M|_U to U \times \mathbb{R}^n. The inverse of this trivialization map, a map from U \times \mathbb{R}^n to T*M|_U, is denoted by φ^*_U.

The only example of the cotangent bundle I’ll give now is that of \( \mathbb{R}^n \), in which case the canonical coordinate atlas \( \mathcal{U} = \{ (\mathbb{R}^n, \text{the identity map}) \} \) identifies T\( \mathbb{R}^n \) with the product bundle \( \mathbb{R}^n \times \mathbb{R}^n \). More cotangent bundles examples are given in later chapters.

3g) Bundle homomorphisms

Let \( \pi: E \to M \) and \( \pi': E' \to M \) denote a pair of vector bundles. A homomorphism \( h: E \to E' \) is a smooth map with the property that if \( p \in M \) is any given point, then \( h \) restricts to \( E|_p \) as a linear map to \( E'|_p \). An endomorphism of a given bundle \( E \) is a homomorphism from \( E \) to itself.

An isomorphism between given bundles \( E \) and \( E' \) is a bundle homomorphism that maps each \( p \in M \) version of \( E|_p \) isomorphically to the vector space \( E'|_p \). For example, the definition of a vector bundle given in Chapter 1a implies the following: If \( E \) is a vector bundle with fiber dimension \( n \) and point \( p \in M \) a given point, then \( p \) is contained in an open set \( U \) such that there is a bundle isomorphism \( \phi^U: E|_U \to U \times \mathbb{R}^n \).

These local trivializations can be used to give a somewhat different perspective on the notion of a bundle homomorphism. To set the background, suppose that \( E \) and \( E' \) are two vector bundles over \( M \) of fiber dimensions \( n \) and \( n' \), respectively. Let \( p \in M \) a given point. There is then an open neighborhood \( U \subset M \) of \( p \) with respective bundle isomorphisms \( \phi^U: E|_U \to U \times \mathbb{R}^n \) and \( \phi'|_U: E'|_U \to U \times \mathbb{R}^{n'} \). A given bundle homomorphisms \( h: E \to E' \) is such that the composition \( \phi'|_U \circ h \circ \phi^U \) is a map from \( U \times \mathbb{R}^n \) to \( U' \times \mathbb{R}^{n'} \) that can be written as \( (x, v) \to (x, h^U(x)v) \) where \( x \to h^U(x) \) is a smooth map from the set \( U \) to the vector space, Hom(\( \mathbb{R}^n; \mathbb{R}^{n'} \)), of linear maps from \( \mathbb{R}^n \) to \( \mathbb{R}^{n'} \).

Note that if \( U' \) is another open set in \( M \) with \( U' \cap U \neq \emptyset \) and with corresponding bundle isomorphisms \( \phi^U \) for \( E|_U \) and \( \phi'|_U \) for \( E'|_U \), then there is a corresponding \( h|_{U'} \). The latter is determined by \( h^U \) on \( U' \cap U \) as follows: Introduce the respective bundle transition functions, \( g^U_{U'}: U' \cap U \to \text{Gl}(n; \mathbb{R}) \) for \( E \) and \( g'|_{U':U}: U' \cap U \to \text{Gl}(n'; \mathbb{R}) \) for \( E' \). Then \( h^U = g^U_{U'} h^U (g'|_{U':U})^{-1} \) at the points in \( U' \cap U \).

Example 1: This second picture of a vector bundle isomorphism can be used to tie up a loose end that concerns Chapter 3b’s cocycle definition of a vector bundle. This loose end concerns the notion of the refinement of an open cover. To set the stage, let \( \mathcal{U} \)
denote a given open cover of M. A second open cover, $U'$, is said to be a refinement of $U$ when each set from $U'$ is contained entirely in at least one set from $U$. This notion of refinement is often used when comparing constructions that are defined using distinct open covers. The point being that if $U$ and $V$ are any two open covers, then the collection $\{U \cap V\}_{U \in U}$ is a simultaneous refinement of both.

In any event, suppose that $n \geq 0$ is a given integer, that $U$ is a finite open cover, and that $\{g_{U,U'}: U \cap U' \to \text{Gl}(n; \mathbb{R})\}_{U,U' \in U}$ supply the cocycle data needed to define a vector bundle over M with fiber $\mathbb{R}^n$. Let $E \to M$ denote the resulting vector bundle. Now suppose that $V$ is second finite open cover that refines $U$. Fix a map $u: V \to U$ such that $V \subset u(V)$ for all $V \in V$. This done, then the data $\{g_{u(V),u(V')}: V \cap V' \to \text{Gl}(n; \mathbb{R})\}_{V,V' \in V}$ supplies the cocyle data for a vector bundle, $E' \to M$. As it turns out, there is a canonical isomorphism from $E'$ to $E$. To see this isomorphism, recall that the cocycle definition of $E'$ defines the latter as the quotient of $\times_{V \in V}(V \times \mathbb{R}^n)$ by an equivalence relation. Likewise, $E$ is the quotient of $\times_{U \in U}(U \times \mathbb{R}^n)$. Let $V \in V$. The isomorphism in question takes the equivalence class defined by $(p, v) \in V \times \mathbb{R}^n$ to that defined by the pair $(p, v)$ with the latter viewed as a pair in $u(V) \times \mathbb{R}^n$. It is an exercise to unwind the definitions to see that this map is, as claimed, an isomorphism.

This canonical isomorphism is used implicitly to identify the bundle $E$ with the bundle $E'$ that is defined in this way using a refinement of the cover $U$.

**Example 2:** What follows describes a Lie group application of the notion of vector bundle isomorphism. To start, suppose that $G$ is one of the Lie groups that were introduced in Chapter 2e or Chapter 2e. In each case, the tangent bundle $TG$ is isomorphic to the product bundle $G \times \mathbb{R}^{\dim(G)}$. In the case $G = \text{Gl}(n; \mathbb{R})$ or $\text{Gl}(n; \mathbb{C})$, this comes about because $G$ is an open set in the Euclidean space $M(n; \mathbb{R})$ or $M(n; \mathbb{C})$ as the case may be.

The tangent bundle to $\text{Sl}(n; \mathbb{R})$ was identified in Chapter 3e with the submanifold in $\text{Sl}(n; \mathbb{R}) \times M(n; \mathbb{R})$ of pairs of the form $(m, mc)$ where $c$ has trace equal to 0. Let $M_0(n; \mathbb{R})$ denote the vector space of such matrices. This is a Euclidean space, and the map from $\text{Sl}(n; \mathbb{R}) \times M_0(n; \mathbb{R})$ to $T\text{Sl}(n; \mathbb{R})$ that sends a pair $(m, c)$ to $(m, mc)$ is a bundle isomorphism. The tangent space to $\text{SO}(n)$ was identified with the set of pairs $(m, mc) \in \text{SO}(n) \times M(n; \mathbb{R})$ such that $c^T = -c$. Let $A(n; \mathbb{R})$ denote the vector space of such $n \times n$, skew symmetric matrices. The map that sends a pair $(m, c)$ in the product vector bundle $SO(n) \times A(n; \mathbb{R})$ to $(m, mc) \in TSO(n)$ is also a bundle isomorphism.

There are analogous isomorphisms for the groups $U(n)$ and $SU(n)$. To describe the former case, let $A(n; \mathbb{C})$ denote the vector space of anti-Hermitian matrices. Thus, $c \in A(n; \mathbb{C})$ if and only if $c^\dagger = -c$. The map that sends $(m, c) \in U(n) \times A(n; \mathbb{C})$ to the pair
(m, mc) ∈ TU(n) is a bundle isomorphism. To consider the case of SU(n), introduce $A_0(n; \mathbb{C})$ to denote the subvector space in $\mathbb{A}(n; \mathbb{C})$ of trace zero matrices. The same map restricts to the subspace $SU(n) \times A_0(n; \mathbb{C})$ to give a bundle isomorphism from the latter space to TSU(n).

3h) Sections of vector bundles

Let $\pi: E \rightarrow M$ denote a given vector bundle. A section of $E$ is a smooth map $s: M \rightarrow E$ such that $\pi \circ s$ is the identity map on $M$. Thus a section assigns to each point $p \in M$ a point in $E_p$. For example, the zero section $\delta$ that appears in the definition given in Chapter 3a is an example. Note that the space of sections is a linear space. Moreover, it is linear over $C^\infty(M; \mathbb{R})$ in the following sense: If $s$ is any given section and $f: M \rightarrow \mathbb{R}$ any given function, then $fs$ is also a section; this is the section that assigns $f(p)s(p) \in E_p$ to $p \in M$. The space of sections of $E$ is denoted by $C^\infty(M; E)$ in what follows.

To see what a section looks like, suppose that $U \subset M$ is an open set with an isomorphism $\phi_U: E|_U \rightarrow U \times \mathbb{R}^n$. Here, $n$ is the fiber dimension of $E$. Let $s$ denote any given section of $E$. Then $\phi_U \circ s: U \rightarrow U \times \mathbb{R}^n$ has the form $x \mapsto (x, s_U(x))$ where $s_U$ is a smooth map from $U$ to $\mathbb{R}^n$. Note that if $U' \subset M$ is an open set with a trivialization $\phi_{U'}$ from $E|_{U'} \rightarrow U' \times \mathbb{R}^n$, then there is a corresponding map $s_{U'}$, with the latter determined by $s_U$ over $U' \cap U$ by the rule $s_{U'} = g_{U'U}s_U$ where $g_{U'U}: U' \cap U \rightarrow GL(n; \mathbb{R})$ again denotes the bundle transition function.

This identification can be used to see why the vector space of sections of any given bundle $E$ is infinite dimensional. Indeed, fix $U \subset M$ where there is a bundle isomorphism $\phi_U: E|_U \rightarrow U \times \mathbb{R}^n$. Let $s_U: U \rightarrow \mathbb{R}^n$ denote any given map which is zero on the complement of a compact set in $U$. A section, $s$, of $E$ can be defined now as follows: Set $s = \delta$ on $E|_{M-U}$ and set $s$ on $U$ to be the section given by $p \mapsto (s_U(p))$. By the way, a comparison of the definition of a section with that of a bundle homomorphism finds that a section of $E$ gives a bundle homomorphism from the trivial bundle $M \times \mathbb{R}$ to $E$, and vice versa. The correspondence identifies any given section $s$ with the homomorphism that send $(p, r) \in M \times \mathbb{R}$ to $rs(p)$. Meanwhile, a homomorphism $h: M \times \mathbb{R} \rightarrow E$ gives the section $p \mapsto h|_p \cdot 1$ where $1 \in \mathbb{R}$.

If $U \subset M$ is any open set, a set $\{s_k\}_{1 \leq k \leq n}$ of sections of $E|_U$ are said to define a basis of sections of $E|_U$ if the set $\{s_k|_p\}_{1 \leq k \leq n} \in E_p$ is linearly independent for each $p \in U$. Thus they define a basis of $E_p$. A basis of sections gives an isomorphism from the product bundle $U \times \mathbb{R}^n$ to $E|_U$, this the homomorphism that sends a given $(p, (r_1, \ldots, r_n)) \in U \times \mathbb{R}^n$ to the point $\sum_{1 \leq k \leq n} r_k s_k(p)$). Conversely, a basis of sections $\{e_k\}_{1 \leq k \leq n}$ for $E^*|_U$ defines an isomorphism from $E|_U$ to $U \times \mathbb{R}^n$, this the homomorphism that sends any given point $v \in E$
\(E|_p \) to \((p, (l_i(v), \ldots, l_n(v))) \in U \times \mathbb{R}^n\). A bundle \(E \to M\) is isomorphic to the product bundle if and only if there is a basis of sections for the whole of \(M\).

3i) Sections of TM and T*M

A section of the tangent bundle of \(M\) is called a \textit{vector field}. The space of sections of TM, thus \(C^\infty(M; TM)\), can be viewed as the vector space of \textit{derivations} on the algebra of functions on \(M\). To elaborate, remark that the space \(C^\infty(M; \mathbb{R})\) of smooth functions on \(M\) is an algebra on functions with addition given by \((f + g)(p) = f(p) + g(p)\) and multiplication given by \((fg)(p) = f(p)g(p)\). A derivation is by definition a map, \(L\), from \(C^\infty(M; \mathbb{R})\) to itself that obeys \(Lf + Lg = L(f + g)\) and \(Lf g = fLg\). Let \(\nu\) denote a given section of TM. The action on a function \(f\) of the derivation defined by \(\nu\) is denoted by \(\nu(f)\). To see what this looks like, go to a coordinate chart \(U \subset M\) with coordinate map \(\phi_U : U \to \mathbb{R}^n\) with \(n\) here set equal to the dimension of \(M\). Use \(\phi_U : TM|_U \to U \times \mathbb{R}^n\) to denote the associated isomorphism between \(TM|_U\) and the product bundle. Write \(\phi_U \circ \nu\) as the section of the product bundle given by \(p \to (p, \nu_U = (v_1, \ldots, v_n) : U \to \mathbb{R}^n)\).

Meanwhile, write \(f_U = f \circ \phi_U^{-1}\). The analogous \((\nu f)_U\) is given by \((\nu f)_U = \sum_1^\infty v_i (\frac{\partial}{\partial x^i} f_U)\).

The fact that this definition is consistent across coordinate charts follows using the chain rule and the bundle transition function for TM.

The function \(\nu f\) is often called the \textit{Lie derivative} of \(f\) along the \(\nu\). It is also called the \textit{directional} derivative of \(f\) along \(\nu\) because the formula that writes \((\nu f)_U\), as \(\sum_1^\infty v_i (\frac{\partial}{\partial x^i} f_U)\) identifies \((\nu f)_U\) with the directional derivative of the function \(f_U\) on \(\mathbb{R}^n\) along the vector field with components \((v_1, \ldots, v_n)\).

The observation that vector fields give derivations motivates the notation whereby \(\nu_U\) is written in terms of its components as \(\sum_1^\infty v_i \frac{\partial}{\partial x^i}\); and it motivates the notation that uses \(\{ \frac{\partial}{\partial x^i}\}_1^\infty\) as a basis of sections for \(T\mathbb{R}^n\). The map \(\phi_U^{-1}\) takes this basis to a basis of sections of TM over \(U\).

To how a derivation of \(C^\infty(M; \mathbb{R})\) gives a vector field, let \(\mathcal{L}\) denote a given derivation. Let \(U \subset M\) denote an open set where there is a coordinate map \(\phi_U \to \mathbb{R}^n\). For each \(k \in \{1, \ldots, n\}\), let \(f_k = x_k \circ \phi_U\). Define \(c_k\) to be the function on \(U\) given by \(L f_k\). It then follows using Taylor’s theorem with remainder and the Chain rule that \(\mathcal{L}\) acts on any given function, \(f\), on \(U\) as \(\mathcal{L}f = \sum_1^\infty c_k v_k\), where \(v_k\) is the vector field on \(U\) that corresponds via the trivialization \(\phi_U : TM|_U \to U \times \mathbb{R}^n\) to the basis \(\{ \frac{\partial}{\partial x^i}\}_1^\infty\).

A section of T*M is called a \textit{1-form}. Any given function, \(f\), on \(M\) defines a \(1\)-form, this denoted by \(df\). The latter is said to be the \textit{exterior} derivative of \(f\). The associated linear function on TM is defined as follows: Let \(p \in M\) and let \(v \in TM|_p\). Let \(v\) denote any given vector field on \(M\) with \(v|_p = v\). The linear function \(df\) send \(v\) to \((df)|_p\). To see what this looks like, let \(U \subset M\) denote a coordinate chart and let \(\phi_U : U \to \mathbb{R}^n\).
denote the corresponding map. Write $f \circ \varphi_U^{-1}$ as $f_U$ again, this a function on a domain in $\mathbb{R}^n$. Meanwhile, $\varphi_U$ defines a canonical trivialization of $T^*M|_U$. Then the image of the section $df$ via this trivialization is the section $x \mapsto (x, (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})|_x)$ of the product bundle $\varphi(U) \times \mathbb{R}^n$.

This last observation motivates the writing of $dh$ as $\sum_{1 \leq i \leq n} \frac{\partial h}{\partial x_i} dx^i$ when $h$ is a given function on $\mathbb{R}^n$. The notation indicates that the exterior derivatives of the coordinates functions $\{x_i\}_{1 \leq i \leq n}$ define the basis $\{dx_i\}_{1 \leq i \leq n}$ of sections for $T^*\mathbb{R}^n$. This basis corresponds via $\varphi_U$’s trivialization of $T^*M|_U$ to a basis of sections of $T^*M|_U$.

**Additional reading**

CHAPTER 4: Algebra of vector bundles

Any linear operations that can be done to generate a new vector space from some given set of initial vector spaces can be done fiberwise with an analogous set of vector bundles to generate a new vector bundle. What follows describes the most important examples.

4a) Subbundles

Let \( \pi: E \rightarrow M \) denote a vector bundle. A submanifold \( E' \subset E \) is said to be a subbundle of \( E \) if, for each \( p \in M \), the subspace \( E'_p = (E' \cap E)_p \) is a vector subspace.

An equivalent definition is as follows: The submanifold \( E' \subset E \) is a subbundle if it is the image via a bundle homorphism that is injective on each fiber.

To see that the structure of a subbundle \( E' \subset E \) near any given point \( p \in M \), fix an open set \( U \subset M \) that contains \( p \) and that comes with an isomorphism \( \varphi_U: E|_U \rightarrow U \times \mathbb{R}^n \).

The image via \( \varphi_U \) of \( E'_|_U \) is an \( n' \)-dimensional subbundle of the product bundle \( U \times \mathbb{R}^n \).

Let \( V' \subset \mathbb{R}^n \) denote the \( \varphi_U \) image of \( E'_|_p \), this an \( n' \) dimensional vector subspace of \( \mathbb{R}^n \).

Fix a basis for \( V' \) so as to identify it with \( \mathbb{R}^{n'} \). Let \( \pi_{V'}: \mathbb{R}^n \rightarrow V' = \mathbb{R}^{n'} \) denote the orthogonal projection. Then there is a neighborhood \( U' \subset U \) of \( p \) such that the composition \( \pi_{V'} \circ \varphi_U \) maps \( E'_|_{U'} \) isomorphically to the trivial bundle \( U' \times V' = U' \times \mathbb{R}^{n'} \).

By way of an example, suppose that \( M \) sits in \( \mathbb{R}^N \) as a submanifold. Then \( TM \) sits in the product bundle \( M \times \mathbb{R}^N \) as a subbundle. To see a second example, fix positive integers \( m > n \) and introduce from Section 1f the Grassmannian \( Gr(m; n) \) whose points are the \( n \)-dimensional vector subspaces of \( \mathbb{R}^m \). Introduce from Chapter 3c the tautological \( \mathbb{R}^n \)-bundle \( E \rightarrow Gr(m, n) \), this the set of pairs \( \{(V, v) \in Gr(m, n) \times \mathbb{R}^n: v \in V\} \). This bundle is defined to be a subbundle of the product bundle \( Gr(m; n) \times \mathbb{R}^n \).

As it turns out, any given bundle \( E \rightarrow M \) can be viewed as a subbundle of some product bundle \( M \times \mathbb{R}^N \) for \( N \) very large. What follows explains why this is true in the case when \( M \) is compact. To start, fix a finite open cover \( \mathcal{U} \) of \( M \) such that each set \( U \subset M \) has an associated bundle isomorphism \( \varphi_U: E|_U \rightarrow U \times \mathbb{R}^n \). By enlarging the set \( \mathcal{U} \) if necessary, one can also assume that there is an associated, subordinate partition of unity, this a set of functions \( \{\chi_U: U \rightarrow [0, \infty)\}_{U \in \mathcal{U}} \) such that any given \( \chi_U \) is zero on \( M \cup \mathbb{R}^n \).

Now define a bundle homomorphism \( \psi: E \rightarrow M \times \mathbb{R}^N \) by the following rule: Write \( j \in \{1, \ldots, Kn\} \) as \((k-1)n + i\) where \( k \in \{1, \ldots, K\} \) and where \( i \in \{1, \ldots, n\} \).

A given point \( v \in E \) is sent to the point in \( M \times \mathbb{R}^N \) whose \( j \)th coordinate in the \( \mathbb{R}^{Kn} \) factor is \( (\varphi_U^{-1}(v))_{\chi_U} |_{\psi(v)} \). By construction, the map so defined is fiberwise injective, and so \( \psi(E) \) sits in \( M \times \mathbb{R}^{Kn} \) as a subbundle.
4b) Quotient bundles

Suppose that $V$ is a vector space and $V' \subset V$ a vector subspace. One can then define the quotient space $V/V'$, this the set of equivalence classes that are defined by the rule that has $v \sim u$ when $v - u \in V'$. The quotient $V/V'$ is a vector space in its own right, its dimension is $\dim(V) - \dim(V')$. Note that in the case when $V$ is identified with $\mathbb{R}^n$, the orthogonal projection from $\mathbb{R}^n$ to the orthogonal complement of $V'$ identifies the quotient bundle $V/V'$ with this same orthogonal complement.

There is the analogous quotient construction for vector bundles over a given manifold. To elaborate, suppose that $\pi : E \to M$ is a vector bundle with fiber dimension $n$ and $E' \subset E$ is a subbundle with fiber dimension $n' \leq n$. The quotient bundle $E/E'$ is defined so that its fiber over any given point $p \in M$ is $E|_p/E'|_p$. The projection from $E/E'$ is induced by that of $E$ to $M$, and the action is induced from the latter’s action on $E$.

To see the structure of $E/E'$ near a given point $p \in M$, suppose that $U \subset M$ is an open set that contains $p$ and that comes with a bundle isomorphism $\phi_U : E|_U \to U \times \mathbb{R}^n$. The latter restricts to $E'|_U$ so as to map $E'|_U$ isomorphically onto some $n'$-dimensional subbundle of the product bundle $U \times \mathbb{R}^n$. Set $\phi_U(E'|_p) = V'$, this an $n'$-dimensional vector subspace of $\mathbb{R}^n$. Let $\pi_{V'}$ as before denote the orthogonal projection from $\mathbb{R}^n$ to $V'$. As noted in Chapter 4a, there is a neighborhood $U' \subset U$ of $p$ such that $\phi_U(E'|_p) = V'$ maps $E'|_U$ isomorphically to the trivial bundle $U \times V'$. Let $V'^\perp \subset \mathbb{R}^n$ denote the orthogonal complement of $V'$, and let $\pi_{V'}$ denote the orthogonal projection from $\mathbb{R}^n$ to $V'^\perp$. Introduce the $p : U \times V' \to V'^\perp$ to denote the composition $\pi_{V'} \circ \phi_U$ maps $E'|_U$ isomorphically to the trivial bundle $U \times V'$. Let $V'^\perp \subset \mathbb{R}^n$ denote the orthogonal complement of $V'$, and let $\pi_{V'}$ denote the orthogonal projection from $\mathbb{R}^n$ to $V'^\perp$. Introduce the $p : U \times V' \to V'^\perp$ to denote the composition $\pi_{V'} \circ \phi_U$ maps $E'|_U$ isomorphically to the trivial bundle $U \times V'$. By construction, this map sends vectors in $E'|_U$ to zero. Meanwhile, it is surjective at $p$, and this implies that it is surjective on some neighborhood $U'' \subset U'$ of $p$. This understood, it induces an isomorphism $(E/E')|_U \to U' \times V'^\perp$. The choose a basis for $V'^\perp$ to identify the latter with $\mathbb{R}^{n-n'}$ and the resulting map from $(E/E')|_U \to U' \times \mathbb{R}^{n-n'}$ gives the requisite local isomorphism.

By way of an example, suppose that $M$ is realized as a submanifold in $\mathbb{R}^m$ for some $m > n = \dim(M)$. As explained in Chapter 3d, the tangent bundle $TM$ appears as a subbundle of $M \times \mathbb{R}^n$. The quotient bundle $(M \times \mathbb{R}^n)/TM$ is said to be the normal bundle of $M$ in $\mathbb{R}^n$. Consider the case when $M \subset \mathbb{R}^3$ is a surface of genus $g$ as constructed in Chapter 1c using a function $h : \mathbb{R}^2 \to \mathbb{R}$. Recall that $M$ in this case is the locus of points where $z^2 - h(x, y) = 0$. The normal bundle in this case has fiber dimension 1, and orthogonal projection in $\mathbb{R}^3$ identifies it with the subbundle in $M \times \mathbb{R}^3$ given by

$$\{(x, y, z), v = r (z \frac{\partial}{\partial z} - (\frac{\partial h}{\partial x})|_{(x,y)} \frac{\partial}{\partial x} - (\frac{\partial h}{\partial y})|_{(x,y)} \frac{\partial}{\partial y}): r \in \mathbb{R}\}.$$
4c) The dual bundle

Let $V$ denote a vector space of dimension $n$. The dual vector space, $V^*$, is the vector space of linear maps from $V$ to $\mathbb{R}$. There is a corresponding notion of the dual of a vector bundle. To say more, let $\pi: E \to M$ denote the vector bundle in question. Its dual is denoted by $E^*$; it is defined so that the fiber of $E^*$ over any given point $p \in M$ is the dual of the vector space $E|_p$. To elaborate, $E^*$ is, first of all, the set of pairs $\{(p, l): p \in M \text{ and } l: E|_p \to \mathbb{R}\}$ is a linear map. The bundle projection $\pi: E^* \to M$ sends any given pair $(p, l) \in E^*$ to the point $p$. The $\mathbb{R}$ action is such that $r \in \mathbb{R}$ acts to send $(p, l) \in E^*$ to $(p, rl)$.

The definition also requires a suitable map to $\mathbb{R}^n$ near any given point; a map that restricts to each fiber as an $\mathbb{R}$-equivariant diffeomorphism. For this purpose, let $U \subset M$ denote an open set where there is a smooth map $\lambda_U: E|_U \to \mathbb{R}^n$ that restricts to each fiber as an $\mathbb{R}$-equivariant diffeomorphism. The corresponding map $\lambda^*_U$ for $E^*|_U$ is best viewed as a map to the vector space of linear functions on $\mathbb{R}^n$. (This, of course, is a copy of $\mathbb{R}^n$.) Granted this view, the map $\lambda^*_U$ is defined by the following requirement: Let $p \in U$ and let $e \in E|_p$. Then $\lambda^*_U$ maps $(p, l)$ to the linear functional on $\mathbb{R}^n$ whose value on $\lambda_U(e)$ is the same as that of $l$ on $e$.

Note that this definition of $\lambda^*_U$ has the following consequence: Let $U' \subset M$ denote an open set with $U \cap U' \neq \emptyset$ and with an analogous map $\lambda_U': E|_{U'} \to \mathbb{R}^n$. Introduce $g_{U:U'}: U' \cap U \to \text{Gl}(n; \mathbb{R})$ to denote the corresponding bundle transition function. Given $\lambda_U'$, there is a corresponding $\lambda^*_U': E^*|_{U'} \to \mathbb{R}^n$, and a corresponding bundle transition function for $E^*$ that maps $U' \cap U$ to $\text{Gl}(n; \mathbb{R})$. The latter is the map $(g_{U:U'}^{-1})^T$, the transpose of the inverse of $g_{U:U'}$.

The description of $E^*$ is not complete without saying something about its smooth structure. The latter comes about as follows: Let $\mathcal{U} = \{(U, \varphi: U \to \mathbb{R}^{\dim(M)})\}$ denote a coordinate atlas for $M$. No generality is lost by requiring that each set $U$ from this atlas come with a map $\lambda_U: E|_U \to \mathbb{R}^n$ that restricts to each fiber as an $\mathbb{R}$-equivariant diffeomorphism. As noted earlier, the set $E|_U$ with the map $(\varphi, \lambda_U)$ gives a coordinate chart for $E$. This understood, the smooth structure for $E^*$ is that defined by the atlas given by $\{(E^*|_U, (\varphi, \lambda^*_U))\}_{(U, \varphi) \in \mathcal{U}}$. The coordinate transition functions for this atlas are smooth. Indeed, if $(U', \varphi') \in \mathcal{U}$ and $U' \cap U \neq \emptyset$, then the coordinate transition function for the intersection of the charts $E^*|_U \cap E^*|_{U'} = E^*|_{U \cap U'}$ is the map

$$(\varphi' \circ (\varphi)^{-1}, (g_{U:U'}^{-1})^T(\cdot))$$
where $g_{U' U}: U' \cap U \to \text{Gl}(n; \mathbb{R})$ is the bundle transition function for $E$ over $U \cap U'$. This is a smooth map, and so the map to $\text{Gl}(n; \mathbb{R})$ given by the transpose of its inverse is also a smooth map.

Note that if $V$ is a vector space, then $(V^*)^*$ has a canonical identification with $V$. Indeed, a vector $v \in V$ gives the linear function $l \mapsto l(v)$ on $V^*$. By the same token, if $E$ is a vector bundle, then $(E^*)^* = E$.

Perhaps the canonical example is the case when $E$ is the tangent bundle $TM$; in this case, $E^*$ is the cotangent bundle, $T^*M$. This can be seen by comparing the bundle transition functions.

4d) Bundles of homomorphisms

The notion of the dual of a vector space, $\text{Hom}(V; \mathbb{R})$, has a generalization as follows: Let $V$ and $V'$ denote two vector spaces. Let $\text{Hom}(V, V')$ denote the set of linear maps from $V$ to $V'$. This also has the structure of a vector space, one with dimension equal to the product of the dimensions of $V$ and of $V'$. This can be seen by choosing basis for $V$ to identify it with $\mathbb{R}^{\dim(V)}$ and likewise for $V'$ to identify it with $\mathbb{R}^{\dim(V')}$. A linear map from $V$ to $V'$ appears now as a matrix with $\dim(V)$ columns and $\dim(V')$ rows.

The construction $\text{Hom}(\cdot, \cdot)$ can be applied fiberwise to a pair of vector bundles, $E$ and $E'$, over $M$, to give a new vector bundle, this denoted by $\text{Hom}(E, E')$. The fiber of $\text{Hom}(E, E')$ over any given point $p \in M$ is the vector space $\text{Hom}(E|_p, E'|_p)$. If $E$ has fiber dimension $n$ and $E'$ fiber dimension $n'$, then the fiber dimension of $\text{Hom}(E, E')$ is $nn'$.

To continue with the definition of the vector bundle structure, note that $\mathbb{R}$ action is such that $r \in \mathbb{R}$ acts on $\text{Hom}(E, E')$ so as to send a given point $(p, l: E|_p \to E'|_p)$ to $(p, rl)$. The required neighborhood maps to $\mathbb{R}^{nn'}$ are obtained as follows: Let $p \in M$ denote some given point and $U \subset M$ a neighborhood of $p$ with $\mathbb{R}$-equivariant maps $\lambda_U: E|_U \to \mathbb{R}^n$ and $\lambda'_U: E'|_U \to \mathbb{R}^{n'}$. The requisite map $\lambda_{\text{Hom},U}: \text{Hom}(E, E')|_U$ is defined so as to send any given point $(p', l)$ to the point matrix $\lambda_U^\circ (l \circ (\lambda_U^{-1}))$ where here $\lambda_U^{-1}$ is viewed as a linear map from $\mathbb{R}^n$ to $E|_p$.

As for the smooth structure, take $U$ as above so that it comes with a coordinate map $\varphi: U \to \mathbb{R}^{\dim(M)}$. This understood, then $\text{Hom}(E, E')|_U$ is a coordinate chart for $\text{Hom}(E, E')$ with the map that sends any given pair $(x, l)$ to $(\varphi(p), \lambda'_U \circ (l \circ (\lambda_U^{-1}))) \in \mathbb{R}^{\dim(M)} \times \mathbb{R}^{nn'}$. To see that the coordinate transition functions are smooth, it proves useful to view any given point $m \in \mathbb{R}^{nn'}$ as a matrix $m \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^{n'})$. This understood, suppose that $U' \subset M$ is another coordinate chart with map $\varphi': U' \to \mathbb{R}^{\dim(M)}$ and with corresponding bundle maps $\lambda_{U'}: E|_{U'} \to \mathbb{R}^n$ and $\lambda'_{U'}: E'|_{U'} \to \mathbb{R}^{n'}$. Then the transition
function sends a given point \((x, m) \in \varphi(U \cap U') \times \text{Hom}(\mathbb{R}^n, \mathbb{R}^n')\) to the point with coordinates \(((\varphi \circ \varphi^{-1})(x), g'_{U'}(x) m g_{U}(x)^{-1}) \in \varphi(U \cap U') \times \text{Hom}(\mathbb{R}^n, \mathbb{R}^n')\).

Note that a section of \(\text{Hom}(E, E')\) over \(M\) is nothing more than a bundle homomorphism from \(E\) to \(E'\).

4e) Tensor product bundles

Fix vector spaces \(V\) and \(V'\) with respective dimensions \(n\) and \(n'\). The vector space \(\text{Hom}(V^*; V')\) of linear maps from the dual space \(V^*\) to \(V'\) is denoted by \(V' \otimes V\). What follows gives a pedestrian way to view: Fix a basis \(\{v_a\}_a\) for \(V\) and a basis \(\{v'_a\}_{1 \leq a \leq n'}\) for \(V'\). A basis for the \(nn'\) dimensional vector space \(V' \otimes V\) can be written as \(\{v'_a \otimes v_b\}_{1 \leq a \leq n', 1 \leq b \leq n}\) where the element \(v'_a \otimes v_b\) represents the linear map from \(V^*\) to \(V'\) that assigns to any given linear function \(l \in V^*\) the element \(l(v_b)v'_a \in V'\). Note that a decomposable element in \(V' \otimes V\) is given by a pair \(v' \in V'\) and \(v \in V\); it assigns to any given \(l \in V^*\) the element \(l(v)v' \in V'\). A decomposable element of this sort is written as \(v' \otimes v\). The vector space \(V' \otimes V\) is spanned by its decomposable elements.

By analogy, if \(E' \to M\) and \(E \to M\) are any two vector bundles, the corresponding tensor product bundle is \(E' \otimes E = \text{Hom}(E^*; E')\).

4f) The direct sum

Let \(V\) and \(V'\) denote a pair of vector spaces with respective dimensions \(n\) and \(n'\). The direct sum \(V \oplus V'\) is the vectors space of pairs \((v, v') \in V \times V'\) where addition is defined by \((v, v') + (u, u') = (u+v, u'+v')\) and the \(\mathbb{R}\) action has \(r \in \mathbb{R}\) sending \((v, v')\) to \((rv, rv')\).

To see the bundle analog, let \(\pi: E \to M\) and \(\pi': E' \to M\) denote a pair of vector bundles. Their direct sum, \(E \oplus E'\), is also a bundle over \(M\). It can be viewed as the subset in \(E \times E'\) of pairs \((v, v')\): \(\pi(v) = \pi'(v')\). As explained momentarily, this subset is a submanifold and thus \(E \oplus E'\) is a smooth manifold. The bundle projection from \(E \oplus E'\) to \(M\) sends \((v, v')\) to \(\pi(v)\), this by definition being the same point as \(\pi'(v')\). The \(\mathbb{R}\) action has \(r \in \mathbb{R}\) acting on \((v, v')\) as \((rv, rv')\). To complete the definition, I need to exhibit a local trivialization on some neighborhood of any given point. To do this, let \(U \subset M\) denote a given open set with a pair of maps \(\lambda_U: E|_U \to \mathbb{R}^n\) and \(\lambda'_U: E'|_U \to \mathbb{R}^{n'}\) that restrict to the respective fibers as an \(\mathbb{R}\)-equivariant diffeomorphism. The corresponding map for \((E \oplus E')|_U\) sends the latter to \(\mathbb{R}^n \times \mathbb{R}^{n'}\) and is given by \((\lambda_U, \lambda'_U')\). It is a consequence of these definitions that the fiber over any given point \(p \in M\) of \(E \oplus E'\) is the vector space direct sum \(E|_p \oplus E'|_p\).

To tie up a loose end, I need to explain why \(E \oplus E'\) sits in \(E \times E'\) as a submanifold. For this purpose, take \(U\) as above, and let \(W \subset U\) denote an open subset with a smooth coordinate embedding \(\varphi: W \to \mathbb{R}^{\dim(M)}\). The pair \((\varphi, \lambda_U)\) gives a
coordinate embedding from $E|_{W} \to \mathbb{R}^{\dim(M)} \times \mathbb{R}^{n}$. Likewise, the pair $(\varphi, \lambda_{W})$ gives a coordinate embedding from $E|_{W} \to \mathbb{R}^{\dim(M)} \times \mathbb{R}^{n}$. This understood, then $(E|_{W} \times E|_{W})$ has the coordinate embedding to $(\mathbb{R}^{\dim(M)} \times \mathbb{R}^{n}) \times (\mathbb{R}^{\dim(M)} \times \mathbb{R}^{n})$ given by $((\varphi, \lambda_{W}), (\varphi, \lambda_{W}')).$

The composition of this map with the map to $\mathbb{R}^{\dim(n)}$ given by $((x, v), (x', v')) \mapsto x - x'$ has 0 as a regular value. This, the inverse image of 0 is a smooth submanifold. The latter is, by definition $(E \oplus E'|_{W})$. As a parenthetical remark, it follows as a consequence that map from $(E \oplus E'|_{W})$ to $(\mathbb{R}^{\dim(M)} \times \mathbb{R}^{n} \times \mathbb{R}^{n})$ given by $(\varphi, (\lambda_{W}), (\lambda_{W}'))$ is a coordinate embedding.

4g) Tensor powers

Let $V$ denote a given vector space. Recall that $V^{*}$ is the vector space of linear maps from $V$ to $\mathbb{R}$. Fix $k \geq 1$. As it turns out, the vector space $V^{*} \otimes V^{*}$ is canonically isomorphic to the vector space of bilinear maps from $V \times V$ to $\mathbb{R}$. This is to say the vector of maps $(v_{1}, v_{2}) \mapsto f(v_{1}, v_{2}) \in \mathbb{R}$ with the property that $f(\cdot, v_{2})$ is linear for any fixed $v_{2} \in V$ and $f(v_{1}, \cdot)$ is linear for any fixed $v_{1} \in V$. The identification is such as to send any given decomposable element $l_{1} \otimes l_{2} \in V^{*} \otimes V^{*}$ to the function $(v_{1}, v_{2}) \mapsto l_{1}(v_{1})l_{2}(v_{2})$.

By the same token, if $k \geq 1$ is any given integer, then $\otimes_{k} V^{*} = V^{*} \otimes \cdots \otimes V^{*}$ is canonically isomorphic as a vector space to the vector space of $k$-linear maps from $\times_{k} V$ to $\mathbb{R}$. The identification is such as to send a decomposable element $l_{1} \otimes \cdots \otimes l_{k}$ to the map that sends any given $k$-tuple $(v_{1}, \ldots, v_{k})$ to $l_{1}(v_{1}) \cdots l_{k}(v_{k})$.

There is an analogous way to view the $k$-fold tensor power of the dual of a given bundle $\pi: E \to M$. Indeed, a section $\otimes_{k} E^{*}$ defines a $k$-linear, fiber preserving map form $E \oplus \cdots \oplus E$ to $M \times \mathbb{R}$. Conversely, a map $f: E \oplus \cdots \oplus E \to M \times \mathbb{R}$ that preserves the fibers and is linear in each entry defines a section of $\otimes_{k} E^{*}$.

Symmetric powers: A $k$-bilinear map, $f$, from $\times_{k} V$ to $\mathbb{R}$ is said to be symmetric if

$$f(v_{1}, \ldots, v_{i}, \ldots v_{j}, \ldots, v_{k}) = f(v_{1}, \ldots, v_{j}, \ldots v_{i}, \ldots, v_{k})$$

for any pair of indices $i, j \in \{1, \ldots, k\}$ and vectors $(v_{1}, \ldots, v_{k}) \in \times_{k} V$. The vector space of symmetric, $k$-bilinear maps from $\times_{k} V$ to $\mathbb{R}$ is a vector subspace of $\otimes_{k} V^{*}$, this denoted by $\text{Sym}^{k}(V^{*})$. Note that this vector space is isomorphic to the vector space of homogeneous, $k$’th order polynomials on $V$. For example, an isomorphism sends asymmetric map $f$ to the polynomial $v \mapsto f(v, \ldots, v)$. The inverse of this isomorphism sends a homogeneous polynomial $\varrho$ to the symmetric, $k$-bilinear map that sends a given $(v_{1}, \ldots, v_{k}) \in \times_{k} V$ to $(\frac{\partial}{\partial t_{1}} \cdots \frac{\partial}{\partial t_{k}} \varrho(t_{1}v_{1} + \cdots t_{k}v_{k}))|_{t_{1}=\cdots=t_{k}=0}$.

By the way, a smooth map $f: V \to \mathbb{R}$ that is smooth and is such that $f(rv) = r^{k}f(v)$ for all $r \in \mathbb{R}$ and $v \in V$ is, defacto, a homogeneous polynomial of degree $k$. Indeed, any
such map must obey \( f(v) = \left. \frac{1}{k!} \left( \frac{\partial^k}{\partial t^k} f(tv) \right) \right|_{t=0} \), and so is given by the order \( k \) term in \( f \)'s Taylor’s expansion at the origin in \( V \).

Let \( \pi: E \to M \) again denote a vector bundle. The bundle \( \text{Sym}^k(E^*) \) is the subbundle of \( \otimes_k E^* \) whose restriction to any given fiber defines a symmetric, \( k \)-linear function on the fiber. Said differently, the fiber of \( \text{Sym}^k(E^*) \) over any given \( p \in M \) is the vector space of maps from \( E|_p \) to \( \mathbb{R} \) that obey \( f(rv) = r^k f(v) \) for each \( r \in \mathbb{R} \) and \( v \in E|_p \). A standard convention sets \( \text{Sym}^0(E) = M \times \mathbb{R} \).

**Anti-symmetric powers:** Fix \( k \in \{1, \ldots, n\} \). A \( k \)-linear map from \( \times_k V \) to \( \mathbb{R} \) is said to be **anti-symmetric** when

\[
f(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = -f(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)
\]

for any pair of indices \( i, j \in \{1, \ldots, k\} \) and vectors \((v_1, \ldots, v_k) \in \times_k V\). The vector space of anti-symmetric, \( k \)-linear maps from \( \times_k V \) to \( \mathbb{R} \) is also a vector subspace of \( \otimes_k V^* \), this denoted by \( \wedge^k V^* \). Note that \( \wedge^k V^* = \{0\} \) if \( k > n \), and that \( \wedge^n V^* = \mathbb{R} \). If \( k \in \{1, \ldots, n\} \), then \( \wedge^k V^* \) has dimension \( \frac{n!}{k!(n-k)!} \).

Suppose that \( \{v_1, \ldots, v_n\} \) is a given basis for \( V \). Introduce \( \{l_1, \ldots, l_n\} \) to denote the dual a basis for \( V^* \). This set determines a basis of sections of \( \wedge^k V^* \), where any given basis element is labeled by a \( k \)-tuple \( (\alpha_1, \ldots, \alpha_k) \) of integers that obey \( 1 \leq \alpha_1 < \cdots < \alpha_k \leq n \). The corresponding basis element is written as

\[
l_{\alpha_1} \wedge l_{\alpha_2} \wedge \cdots \wedge l_{\alpha_k};
\]

this corresponding to the antisymmetric map from \( \times_k V \) to \( \mathbb{R} \) that assigns 1 to the element \((v_{\alpha_1}, \ldots, v_{\alpha_k})\) and zero to elements that have an entry whose label is not from \( \{\alpha_1, \ldots, \alpha_k\} \). Note that there is a canonical homomorphism from \( \wedge^k V^* \otimes \wedge^k V^* \to \wedge^{k+k} V^* \) given by anti-symmetrization. The image of a pair \((\omega, \mu)\) under this homomorphism is written as \( \omega \wedge \mu \). This homomorphism is sometimes called the **wedge product** and other times the **exterior product**. The wedge/exterior product is such that \( \omega \wedge \mu = (-1)^{kk} \mu \wedge \omega \).

Let \( \pi: E \to M \) denote a given vector bundle. If \( k \in \{1, \ldots, n\} \), the bundle \( \wedge^k E^* \) is the subbundle in \( \otimes_k E^* \) whose fiber over any given point \( p \in M \) consists of the elements that define anti-symmetric, \( k \)-linear functions on \( E|_p \). The bundle \( \wedge^k E^* \) is the 0-dimensional bundle \( M \times \{0\} \) if \( k > n \). Meanwhile \( \wedge^n E^* \) is a real line bundle. However, it need not be isomorphic to the trivial line bundle! To say more about the latter, suppose that \( U \subseteq M \) is an open set and that \( \varphi_U: E|_U \to U \times \mathbb{R}^n \) is a local trivialization of \( E \) over \( U \). Let \( U' \subseteq M \) denote a second open set, and suppose that \( \varphi_{U'}: E|_{U'} \to U' \times \mathbb{R}^n \) is a local trivialization over \( U' \). Assume that \( U' \cap U \neq \emptyset \) so that there is a well defined bundle
transition function $g_{U'U}: U \cap U' \rightarrow \text{Gl}(n; \mathbb{R})$. The corresponding bundle transition function for $\wedge^n E^*$ is $\det(g_{U'U})$. The bundle $\wedge^n E$ is denoted by $\det(E)$.

The exterior/wedge product applied fiberwise defines a bundle homomorphism from $\wedge^k E^* \otimes \wedge^{k'} E^*$ to $\wedge^{k+k'} E^*$ which is also called the exterior or wedge product.

Of special interest is the case of when $E = TM$ in which case a section of $\wedge^k T^* M$ is called a $k$-form.

**Additional reading**

Chapter 5: Maps and vector bundles

Suppose that M, N are smooth manifolds. As explained in what follows, any given map \( \psi: M \to N \) can be used to associate a vector bundle on M to one on N.

5a) The pull-back construction

Let \( \pi: E \to N \) denote a vector bundle with fiber dimension \( n \). There is a canonical bundle, denoted \( \psi^*E \), over M, also with fiber dimension \( n \), which comes with a canonical map \( \psi^*: \psi^*E \to E \) that restricts to the fiber over any given point \( p \in M \) as a vector space isomorphism from \( (\psi^*E)|_p \) to \( E|_{\psi(p)} \). The bundle \( \psi^*E \) as the submanifold in \( M \times E \) of pairs \( \{(m, e): \psi(m) = \pi(e)\} \). This is to say that the fiber of \( \psi^*E \) over \( m \in M \) is the fiber of E over \( \psi(m) \). The map \( \psi^* \) sends a pair \( (m, e) \in \psi^*E \) (thus \( m \in M \) and \( e \in N \) with \( \pi(e) = \psi(m) \)) to the point \( \psi^*(m, e) = e \). To be somewhat more explicit, suppose that \( U \subset N \) is an open set and that \( \varphi_U: E|_U \to U \times \mathbb{R}^n \) is a bundle isomorphism. Then the \( \psi^*E \) is isomorphic over \( \psi^{-1}(U) \subset M \) to \( \psi^{-1}(U) \times \mathbb{R}^n \), the isomorphism taking a pair \( (m, e) \) to the pair \( (m, \lambda_U(e)) \) where \( \lambda_U: E|_U \to \mathbb{R}^n \) is defined by writing \( \varphi_U(e) \) as the pair \( (\pi(e), \lambda_U(e)) \).

Let \( U' \subset N \) denote a second open set with non-empty intersection with \( U \) and with a corresponding bundle trivialization \( \varphi_{U'}: E|_{U'} \to U' \times \mathbb{R}^n \). The associated transition function \( g_{U'U}: U' \cap U \to \text{Gl}(n; \mathbb{R}) \) determines the transition function of \( \psi^*E \) for the sets \( \psi^{-1}(U') \cap \psi^{-1}(U) \) as the latter is \( g_{U'U} \psi^* \).

Sections of bundles pull back also. To elaborate, suppose that \( \pi: E \to N \) is a smooth bundle and \( s: N \to E \) is a section. Let \( \psi: M \to N \) denote a smooth map. Then \( s \circ \psi \) defines a section of \( \psi^*E \).

Example: Let \( \Sigma \subset \mathbb{R}^3 \) denote a surface. Define a map from \( \Sigma \) to \( S^2 \) by associating to each point \( p \in \Sigma \) the vector normal to \( \Sigma \) (pointing out from the region that is enclosed by \( \Sigma \)). This map is called the Gauss map. Denote it by \( n \). Then \( TS^2 \) is isomorphic to \( n^*TS^2 \). The isomorphism is obtained as follows: View \( TS^2 \subset \mathbb{R}^3 \times \mathbb{R}^3 \) as the set of pairs of the form \( (v, w) \) with \( |v|^2 = 1 \) and \( v \cdot w = 0 \). Here, \( v \cdot w \) denotes the dot product in \( \mathbb{R}^3 \).

Granted this, then \( n^*TS^2 \) can be viewed as the set \( \{(p, w) \in \Sigma \times \mathbb{R}^3: n(p) \cdot w = 0\} \). As explained in Section 2d, this is identical to \( T\Sigma \).

A concrete example is had using a surface of genus \( g \) as constructed in Section 1c from a function, \( h \), on \( \mathbb{R}^2 \). Recall that the surface is given by the set of points in \( \mathbb{R}^3 \) where \( z^2 - h(x, y) = 0 \). The corresponding Gauss map is given by

\[
n|_{(s,y,z)} = (-\frac{a}{\partial x} h, \frac{a}{\partial y} h, 2z)/(|dh|^2 + 4z^2)^{1/2}\]
Example: Pick a point \( q \) in the region of \( \mathbb{R}^3 \) bounded by \( \Sigma \) and define a map, \( \psi \), from \( \Sigma \) to \( S^2 \) by taking the direction from the given point on \( \Sigma \) to \( q \). Is \( \psi^*TS^2 \) isomorphic to \( T\Sigma \)? What if the point \( q \) is outside \( \Sigma \) instead of inside?

Example: Let \( M \) denote a given \( m \)-dimensional manifold. What follows constructs a map from \( M \) to \( S^m \) that is relevant to one of the constructions in Section 1c. To begin the story, fix a finite set \( \Lambda \subset M \). Assign to each point \( p \in \Lambda \) a coordinate chart \( (U, \varphi_p: U \rightarrow \mathbb{R}^m) \) with \( p \in U \) such that \( \varphi_p(p) = 0 \in \mathbb{R}^m \). Fix \( r > 0 \) and small enough so that the open ball, \( B \), of radius \( r \) about the origin in \( \mathbb{R}^m \) is in the image of each \( p \in \Lambda \) version of \( \varphi_p \); and so that the collection \( \{U_p = \varphi_p^{-1}(B)\}_{p \in \Lambda} \) consists of pairwise disjoint sets. Use this data to define a map, \( \psi \), from \( M \) to \( S^m \subset \mathbb{R}^{m+1} \) as follows: The map \( \psi \) sends all of the set \( M - (\bigcup_p U_p) \) to the point \( (0, \ldots, 0, -1) \). To define \( \psi \) on a given version of \( U_p \), write \( \mathbb{R}^{m+1} = \mathbb{R}^m \times \mathbb{R} \). Introduce from Appendix 2 the function \( \chi \) and set \( \chi_p: U_p \rightarrow [0, 1] \) to equal \( \chi(\|\varphi_p\|) \). This done, the map \( \psi \) sends a \( p \in U_p \) to the point \( (2\chi_p\varphi_p, \chi_p^2 - (\varphi_p)^2)(\chi_p^2 + (\varphi_p)^2)^{-1} \).

To continue, fix a positive integer \( n \) and a map \( g: S^m \rightarrow \text{Gl}(n; \mathbb{R}) \). Define a rank \( n \) vector bundle \( E \rightarrow S^m \) by using the coordinate patches \( U_+ = S^m - (0, \ldots, -1) \) and \( U_- = S^m - (0, \ldots, +1) \) with the transition function that \( g_{U_+ U_-} \) send a point \( (x, x_{n+1}) \in S^n \subset \mathbb{R}^n \times \mathbb{R} \) to \( g(x/|x|) \). The resulting bundle \( \psi^*E \rightarrow M \) is isomorphic to the bundle constructed in Chapter 1c using the data consisting of \( \Lambda \) and the set \( \{(U_p, \varphi_p, g_p = g)\}_{p \in \Lambda} \).

5b) Pull-backs and Grassmannians

As it turns out, any given vector bundle is isomorphic to the pull-back of a tautological bundle over a Grassmannian, these as described in Chapter 1f. The next proposition makes this precise.

**Proposition:** Let \( M \) denote a smooth manifold, let \( n \) denote a positive integer, and let \( \pi: E \rightarrow M \) denote a given rank \( n \) vector bundle. If \( m \) is sufficiently large, there exists a map \( \psi_m: M \rightarrow \text{Gr}(m; n) \) and an isomorphism between \( E \) and the pull-back via \( \psi_m \) of the tautological bundle over \( \text{Gr}(m; n) \).

**Proof:** As noted in Chapter 4a, the bundle \( E \) is isomorphic to a subbundle of some \( N \gg 1 \) version of the product bundle \( M \times \mathbb{R}^N \). Use an isomorphism of this sort to view \( E \) as a subbundle in just such a product bundle. Let \( \lambda: M \times \mathbb{R}^N \rightarrow \mathbb{R}^N \) denote the projection map. The assignment to \( x \in M \) of the \( n \)-dimensional vector subspace \( \lambda(E_x) \subset \mathbb{R}^N \) defines a smooth map \( \psi: M \rightarrow \text{Gr}(N; n) \). Let \( E_N \rightarrow \text{Gr}(N; n) \) denote the tautological \( \mathbb{R}^n \) bundle. This is to say that the fiber of \( E_N \) over a given subspace \( \Pi \in \text{Gr}(N; m) \) is the vector space
Thus, the fiber of \( E_N \) over \( \psi(x) \) is \( E|_x \) and so \( E \) is the pull-back via \( \psi \) of the tautological bundle.

5c) Pull-back of differential forms and push-forward of vector fields

There is a canonical, vector space homomorphism that maps \( C^\infty(N; \wedge^k T^*N) \) to \( C^\infty(M; \wedge^k T^*M) \) for any \( k \in \{0, 1, \ldots, \} \), this also denoted by \( \psi^* \). It is defined so as to commute with wedge product, exterior multiplication and to factor with regards to compositions. This map is defined by the following rules

- \( \psi^*f = f \circ \psi \) and \( \psi^*df = d(\psi^*f) \) when \( f \in C^\infty(N) = C^\infty(N; \wedge^0 T^*N) \).
- \( \psi^*(f \omega) = (\psi^*f) \psi^*\omega \) when \( f \in C^\infty(N) \) and \( \omega \in C^\infty(N; \wedge^k T^*N) \) for any \( k \).

Induction on \( k \) using these two rules defines \( \psi^* \) as an algebra homomorphism from \( \bigoplus_{k\in\{0,1,\ldots,\}} C^\infty(N; \wedge^k T^*N) \) to \( \bigoplus_{k\in\{0,1,\ldots,\}} C^\infty(M; \wedge^k T^*M) \) that commutes with \( d \). Moreover, if \( Z \) is a third manifold and \( \varphi: N \to Z \) is a smooth map, then \( (\varphi \circ \psi)^* = \varphi^* \psi^* \). The homomorphism \( \psi^* \) is called the pull-back homomorphism, and \( \psi^*\omega \) is said to be the pull-back of \( \omega \).

The picture of the pull-back homomorphism in local coordinates shows that the latter is a fancy version of the Chain Rule. To elaborate, fix positive integers \( m \) and \( n \); and let \( \phi \) denote a smooth map between \( \mathbb{R}^m \) and \( \mathbb{R}^n \). Introduce coordinates \( (x^1, \ldots, x^m) \) for \( \mathbb{R}^m \) and coordinates \( (y^1, \ldots, y^n) \) for \( \mathbb{R}^n \). Use \( (\phi_1, \ldots, \phi_n) \) to denote the components of the map \( \phi \). Then

\[
\phi^*(dy^k) = \sum_{1 \leq i \leq m} \left( \frac{\partial \phi_k}{\partial y_i} \right)_x dx^i,
\]

Note that \( d(\phi^*dy^k) = 0 \) because \( \frac{\partial^2 \phi_k}{\partial x_i \partial x_j} = \frac{\partial^2 \phi_k}{\partial x_j \partial x_i} \) for any given function \( f \) and indices \( i \) and \( j \).

Use of local coordinate charts gives a picture of this map \( \psi^* \) which is essentially that given above for the cases \( M = \mathbb{R}^m \) and \( N = \mathbb{R}^n \). To elaborate, let \( M \) and \( N \) again be smooth manifolds and \( \psi: M \to N \) a smooth map. Let \( U \subset M \) denote a coordinate chart with coordinate map \( \varphi_U: U \to \mathbb{R}^{m=\dim(M)} \). Let \( (x_1, \ldots, x_m) \) the Euclidean coordinate functions on \( \mathbb{R}^m \), and recall from Section 3i that the map \( \varphi_U \) can be used to identify the basis of sections \( \{dx_i\}_{1 \leq i \leq m} \) of \( T^*\mathbb{R}^m \) with a corresponding basis of sections of \( T^*M|_U \). Use the notation \( \{dx^U_i\}_{1 \leq i \leq m} \) to denote the latter. Let \( V \subset N \) denote a coordinate chart whose intersection with \( \psi(U) \) is not empty, and let \( \varphi_V: V \to \mathbb{R}^{n=\dim(N)} \) denote the coordinate map. Use \( (y_1, \ldots, y_n) \) to denote the Euclidean coordinates of \( \mathbb{R}^n \) and use \( \varphi_V \) to define the basis of sections \( \{dy^V_i\}_{1 \leq i \leq n} \) for \( T^*N|_V \). The pull-back via \( \psi \) of a given \( dy^V_k \) is given by
\[(\psi^*dy^V)_x = \sum_{1 \leq i \leq m} (\frac{\partial \phi}{\partial x_i})_{\phi(x)}(dx^V)_x.\]

where \(\phi\) here denotes \(\varphi_u \circ \psi \circ (\varphi_r)^{-1}\), this a map from a domain in \(\mathbb{R}^m\) to one in \(\mathbb{R}^n\).

By the way, this pull-back map \(\psi^*\) is closely related to the previously defined notion of the pull-back bundle. Indeed, \(\psi^*\) can be defined as the composition of two linear maps. The first is map pulls back a section of \(\wedge^kT^*N\) to give a section of \(\psi^*(\wedge^kT^*N)\) over \(M\). This as described in Chapter 5a above. The second is induced by a certain canonical vector bundle homomorphism from \(\psi^*(\wedge^kT^*N)\) to \(\wedge^kT^*M\). To say more about the latter, let \(V \subset N\) and \(\varphi_V: V \to \mathbb{R}^{\dim(N)}\) be as described in the preceding paragraph. Then the sections \(\{(dy^V) \circ \psi\}_{1 \leq i \leq m}\) give a basis of sections for \((\psi^*T^*N)|_V\).

Now suppose that \(U \subset M\) and \(\varphi_U: U \to \mathbb{R}^{\dim(M)}\) is also as described in the preceding paragraph. Then the afore-mentioned bundle homomorphism from \(\psi^*T^*N\) to \(T^*M\) restricts over \(U\) so as to send any given \((dy^V) \circ \psi|_x\) to what is written on the left hand side of the equation in the preceding paragraph.

There is a dual of sorts to the pull-back map \(\psi^*: C^\infty(N; T^*N) \to C^\infty(M; T^*M)\), this canonical push-forward map from \(TM\) to \(TN\) that restricts to any given \(x \in M\) version of \(TM|_x\) so as a linear map to \(TN|_{\psi(x)}\). This push-forward is denoted by \(\psi_*\). To elaborate, suppose that \(v \in TM\) and \(\omega\) is a section of \(T^*N\). The pairing between \(\omega\) and \(\psi_*v\) is, by definition, equal to that between the pull-back \(\psi^*\omega\) and \(v\). Note that this push-forward can be viewed as the composition of a bundle homomorphism from \(TM\) to \(\psi^*TN\) followed by the map \(\tilde{\psi}: \psi^*TN \to TN\). The former is the dual homomorphism to that described in the preceding paragraph from \(\psi^*T^*N\) to \(T^*M\).

The next part of this chapter describes some important examples of pull-back and push-forward.

5d) Invariant forms and vector fields on Lie groups

Let \(G\) denote a Lie group. Then any given element \(g \in G\) defines two maps from \(G\) to itself, these being the maps that sent \(m \in G\) to \(l_g(m) = gm\) and to \(r_g(m) = mg\). The first is called left translation by \(g\), and the second is called right translation by \(g\). Both are diffeomorphism with inverses given respectively by left and right translation by \(g^{-1}\).

A 1-form \(\omega\) on \(G\) is said to be left invariant if \(l_g^*\omega = \omega\) for all \(g \in G\). The 1-form \(\omega\) is said to be right invariant if \(r_g^*\omega = \omega\) for all \(g \in G\). By the same token, a vector field \(v\) on \(G\) is left invariant or right invariant if \(l_g^*v = v\) or \(r_g^*v = v\) for all \(g \in G\) as the case may be.

Suppose that \(\omega\) is a left invariant 1-form. If \(\omega\) is non-zero on \(TG|_1\), then \(\omega\) is nowhere zero. This follows by virtue of the fact that \(l_g\) is a diffeomorphism. This implies that the restriction map from \(C^\infty(G; T^*G)\) to \(T^*G|_1\) identifies the vector space of left-invariant 1-forms with \(T^*G|_1\). This said, then a basis for \(T^*G|_1\) gives a basis for the left-
invariant 1-forms on $G$, and this basis gives a basis of sections for $T^*G$ over the whole of $G$. This then supplies a vector bundle isomorphism from the product bundle $G \times T^*G|_i$ to $T^*G$. By the same token, the vector space of right invariant 1-forms on $G$ is also isomorphic to $T^*G|_i$ and so a basis for $T^*G|_i$ extends to the whole of $G$ as a basis of right invariant 1-forms on $G$, which defines a basis of sections for $T^*G$. This then supplies a second isomorphism between $G \times T^*G|_i$ and $T^*G$. These two isomorphisms will not, in general agree.

Of course, a similar thing can be said with regards to $TG$ and the vector spaces of left and right invariant vector fields. A basis for $TG|_i$ extends as a basis of left invariant vector fields, and also as a basis of right invariant vector fields. Either basis gives a basis of sections of $TG$ over the whole of $G$ and thus an isomorphism from the product vector bundle $G \times TG|_i$ to $TG$.

**Example 1:** Introduce the $n^2$ dimensional Euclidean space $\mathbb{M}(n; \mathbb{R})$ of $n \times n$ real matrices. In what follows, I use notation whereby the entries $\{m_{ij}\}_{1 \leq i,j \leq n}$ of a given matrix define the Euclidean coordinates on $\mathbb{M}(n; \mathbb{R})$. This understood, the $n^2$ differentials $\{dm_{ij}\}_{1 \leq i,j \leq n}$ give a basis of sections of $T^*\mathbb{M}(n; \mathbb{R})$. Let $i: \text{Gl}(n \mathbb{R}) \to \mathbb{M}(n; \mathbb{R})$ denote the tautological embedding as the group of matrices with non-zero determinant. Fix a non-zero matrix $q$ and define the 1-form $\omega_q$ on $\text{Gl}(n; \mathbb{R})$ by the rule

$$\omega_q|_m = \text{tr}(qm^{-1}dm) .$$

(To review the notation, what is written above is $\omega_q = \sum_{i \leq j \leq n} q_{ij}(m^{-1})_{jk} dm_{ki}$.) This 1-form is left invariant: Indeed, let $g \in \text{Gl}(n; \mathbb{R})$. Then $g^*(m_{ij}) = \sum_{i \leq k \leq n} g_{ik} m_{ki}$ and so $g^*(dm_{ij}) = \sum_{i \leq k \leq n} g_{ik} dm_{kj}$. Thus, $(g^*\omega_q)|_m = \text{tr}(g^{-1}q^{-1}gdg^{-1}q^{-1}) = \omega_q|_{g^*m}$. The form $\omega_q$ is not right invariant as can be seen from the formula $(r^*\omega_q)|_m = \text{tr}(q^{-1}g^{-1}r^{-1}q^{-1}g^{-1}r^{-1})$, this being $\omega_{r^*q^{-1}g^*}$. A similar calculation finds that the form $\text{tr}(m^{-1}qdm)$ is right invariant but not left invariant.

The form $\omega_q$ at the identity $1 \in \text{Gl}(n; \mathbb{R})$ is non-zero since $q$ is non-zero, and so the form $\omega_q$ is nowhere zero on $\text{Gl}(n; \mathbb{R})$.

To see what a left-invariant vector field looks like, fix again some matrix $q$. This done, introduce

$$v_q|_m = -\text{tr}(qm^T \frac{\partial}{\partial m}) = -\sum_{i \leq j \leq n} \frac{\partial}{\partial m_{ij}} q_{ij} m_{ki} \frac{\partial}{\partial m_{ki}} .$$

Any such vector field is left invariant, and any left invariant vector field of this sort.

**5e) The exponential map on a matrix group**
Recall that the space $\mathbb{M}(n; \mathbb{R})$ of $n \times n$ real valued matrices is a Euclidean space with the coordinate functions given by the $n^2$ entries of a matrix. Fix an $n \times n$ matrix $m \in \mathbb{M}(n; \mathbb{R})$. Use this matrix to define the exponential map $\epsilon_m: \mathbb{M}(n; \mathbb{R}) \to \mathbb{M}(n; \mathbb{R})$ by the rule $\epsilon_m(a) = m e^a$, where

$$e^a = 1 + a + \frac{1}{2} a^2 + \frac{1}{3!} a^3 + \cdots$$

is the power series with the coefficient of $a^k$ equal to $1/k!$. This sequence converges absolutely and so defines a smooth map from $\mathbb{M}(n; \mathbb{R})$ to itself. Note that the matrix $\epsilon(a)$ is invertible, this by virtue of the fact that $\epsilon_i(a)^{-1} = \epsilon_i(-a)$. Thus $\epsilon_m(a) \in \text{Gl}(n; \mathbb{R})$ if $m$ is.

In particular, the map $\epsilon_i$ maps $\mathbb{M}(n; \mathbb{R})$ into $\text{Gl}(n; \mathbb{R})$. As the differential of the map $a \to e^a$ at $a = 0$ is the identity map on $\mathbb{M}(n; \mathbb{R})$, it follows as a consequence of the inverse function theorem that this restricts to some ball in $\mathbb{M}(n; \mathbb{R})$ as a diffeomorphism onto some neighborhood of $i$ in $\text{Gl}(n; \mathbb{R})$.

With the preceding understood, fix $q \in \mathbb{M}(n; \mathbb{R})$ and reintroduce the left-invariant vector field $\omega_q = \text{tr}(q m^{-1} dm)$ on $\text{Gl}(n; \mathbb{R})$. The pull-back of this form by $\epsilon_i$ is the 1-form

$$\epsilon_i^* \omega_q = \int_0^1 ds \text{tr}(e^{-sa}qe^{sa} da) .$$

To elaborate, view $\text{Gl}(n; \mathbb{R})$ as an open subset of $\mathbb{M}(n; \mathbb{R})$ so as to identify $T\text{Gl}(n; \mathbb{R})|_i$ with $\mathbb{M}(n; \mathbb{R})$. Granted this identification, then $\epsilon_i^* \omega_q$ is neither more nor less than a linear functional on $\mathbb{M}(n; \mathbb{R})$. This linear functional maps any given matrix $c$ to

$$\int_0^1 ds \text{tr}(e^{-sa}qe^{sa}c) .$$

What follows considers the exponential map in the context of the groups $\text{Sl}(n; \mathbb{R})$ and $\text{SO}(n)$.

The group $\text{Sl}(n; \mathbb{R})$: Recall that group $\text{Sl}(n; \mathbb{R})$ sits in $\text{Gl}(n; \mathbb{R})$ as the subgroup of matrices with determinant 1. As it turns out, $\epsilon_i$ maps the vector space of trace zero matrices in $\mathbb{M}(n; \mathbb{R})$ into $\text{Sl}(n; \mathbb{R})$. Indeed, this follows from the fact that

$$\frac{d}{ds} \ln(\det(\epsilon_i(sa))) = \text{tr}(a) .$$

As noted in Chapter 3e, the tangent space to $\text{Sl}(n; \mathbb{R})$ at the identity is the vector space of trace zero, $n \times n$ matrices. Given that the differential of $\epsilon_i$ at 0 is the identity map on
It follows directly from the inverse function theorem that \( \varepsilon \), maps a ball about the origin in the vector space of trace zero, \( n \times n \) matrices diffeomorphically onto an open neighborhood of the identity in \( \text{Sl}(n; \mathbb{R}) \). Likewise, \( \varepsilon_m \) for \( m \in \text{Sl}(n; \mathbb{R}) \) maps this same ball diffeomorphically onto an open neighborhood of \( m \) in \( \text{Sl}(n; \mathbb{R}) \).

Let \( i: \text{Sl}(n; \mathbb{R}) \rightarrow \text{Gl}(n; \mathbb{R}) \) denote the tautological inclusion map, and let \( q \) denote a given \( n \times n \) matrix. Then \( i^* \omega_q \) is non-zero unless \( q \) is a multiple of the identity, in which case \( i^* \omega_q = 0 \). The fact that the pull-back of the \( q = 1 \) version of \( \omega_q \) is zero on \( T\text{Sl}(n; \mathbb{R}) |_I \) can be seen using the formula given above for its pull-back via \( \varepsilon \). As \( i^* \omega_q \) is a left-invariant 1-form on \( \text{Sl}(n; \mathbb{R}) \), its vanishing on \( T\text{Sl}(n; \mathbb{R}) |_I \) implies that it must vanish everywhere.

The group \( \text{SO}(n) \): Let \( A(n; \mathbb{R}) \) again denote the vector space of \( n \times n \) anti-symmetric matrices. This is to say that \( a \in A(n; \mathbb{R}) \) when \( a \)'s transpose obeys \( a^T = -a \). The map \( \varepsilon \), as depicted above obeys \( \varepsilon^T(a) = \varepsilon(a^T) \) and so \( \varepsilon^T \varepsilon = 1 \) when \( a \in A(n; \mathbb{R}) \). It follows from this that \( \varepsilon \), restricts to \( A(n; \mathbb{R}) \) so as to map the letter subspace into \( \text{SO}(n) \). To say more, recall that Chapter 3e identifies \( T\text{SO}(n) |_I \) with \( A(n; \mathbb{R}) \). Granted this identification, and granted that the differential at \( a = 0 \) of \( \varepsilon \) is the identity map, it follows from the implicit function theorem that \( \varepsilon \), maps a ball in \( A(n; \mathbb{R}) \) diffeomorphically onto a neighborhood of \( I \in \text{SO}(n) \). By the same token if \( m \) is any given matrix in \( \text{SO}(n) \), then the map \( \varepsilon_m = meI \) restricts to this same ball as a diffeomorphism onto a neighborhood of \( m \) in \( \text{SO}(n) \).

Let \( q \) again be a matrix in \( M(n; \mathbb{R}) \). Let \( i: \text{SO}(n) \rightarrow \text{Gl}(n; \mathbb{R}) \) denote the tautological inclusion map. As explained next, the form \( i^* \omega_q \) is zero if \( q^T = q \) and never zero if \( q^T = -q \) unless \( q = 0 \). To see that \( i^* \omega_q = 0 \) when \( q \) is symmetric, consider its pull-back via the exponential map \( \varepsilon \), as depicted above. For any given \( a \in A(n; \mathbb{R}) \), this defines a linear form on \( A(n; \mathbb{R}) = T\text{SO}(n) |_I \) whose pairing with \( c \in A(n; \mathbb{R}) \) is

\[
\int_0^1 ds \, \text{tr}(e^{sa}qe^{sa}c).
\]

The term in the integrand is, for any given value of \( s \in [0, 1] \), the trace of the product of a symmetric matrix with an anti-symmetric matrix. Such a trace is always zero. To see that \( i^* \omega_q \neq 0 \) when \( q \in A(n; \mathbb{R}) \), it is enough to prove it non-zero on \( T\text{SO}(n) |_I \), since this form is invariant under left multiplication. The formula above shows that \( i^* \omega_q \) at the identity pairs with \( c \in A(n; \mathbb{R}) \) to give \( \text{tr}(qc) \). Taking \( c = -q = q^T \) shows that \( i^* \omega_q |_I \) is nonzero on \( A(n; \mathbb{R}) \) unless \( q \) is zero. Given that \( i^* \omega_q \) is left invariant, this implies that any \( q \neq 0 \) version is nowhere zero on \( \text{SO}(n) \).
The exponential map and right/left invariance on $\text{Gl}(n; \mathbb{C})$ and its subgroups

The group $\text{Gl}(n; \mathbb{C})$ is described in Chapter 2d. Recall that there are two descriptions. The first views $\text{Gl}(n; \mathbb{C})$ as the open subset in the Euclidean space $\mathbb{M}(n; \mathbb{C})$ of $n \times n$ matrices with complex entries that consists of the matrices with non-zero determinant. As a manifold, $\mathbb{M}(n; \mathbb{C})$ has (real) dimension $2n^2$, and its Euclidean coordinates are taken to be the real and imaginary parts of the entries of a given matrix $m$. This understood, view $dm$ as a matrix of $\mathbb{C}$-valued differential forms on $\mathbb{M}(n; \mathbb{C})$ whose real and imaginary parts supply a basis of sections for $T^*\mathbb{M}(n; \mathbb{C})$.

With the preceding understood, let $q \in \mathbb{M}(n; \mathbb{C})$. Then the $\mathbb{C}$-valued differential form $\omega_q = \text{tr}(qm^{-1}dm)$ on $\text{Gl}(n; \mathbb{C})$ is left invariant, and any $\mathbb{C}$-valued, left invariant 1-form on the group $\text{Gl}(n; \mathbb{C})$ can be written in this way, and any $\mathbb{R}$-valued 1-form is the real part of a form of some $\omega_q$. By the same token, the $\mathbb{C}$-valued 1-form $\text{tr}(m^{-1}qdm)$ is right-invariant, and any $\mathbb{C}$-valued, right right invariant 1-form on the group $\text{Gl}(n; \mathbb{C})$ is of this sort. Moreover, any $\mathbb{R}$-valued 1-form is given by the real part of some $\omega_q$.

The $\mathbb{C}$-valued, right and left invariant vector fields on $\text{Gl}(n; \mathbb{C})$ are given respectively by the rules $\text{tr}(qm^T \frac{\partial}{\partial m})$ and $\text{tr}(m^T q \frac{\partial}{\partial m})$ where the notation is as follows: What is written as $\frac{\partial}{\partial m}$ denotes here the $n \times n$ matrix whose $(k, j)$ entry is by definition $\frac{1}{2}(\frac{\partial}{\partial x_{kj}} - i \frac{\partial}{\partial y_{kj}})$ with $x_{kj}$ and $y_{kj}$ denoting the respective real and imaginary parts of the complex coordinate function $m_{kj}$.

Fix $m \in \mathbb{M}(n; \mathbb{C})$ and define the exponential map $e_m: \mathbb{M}(n; \mathbb{C}) \to \mathbb{M}(n; \mathbb{C})$ using the same formula as used to define the $\mathbb{M}(n; \mathbb{R})$ version. As in the case of $\mathbb{M}(n; \mathbb{R})$, the map $e_m$ sends $\mathbb{M}(n; \mathbb{C})$ into $\text{Gl}(n; \mathbb{C})$ when $m \in \text{Gl}(n; \mathbb{C})$, and it restricts to some $m$-independent ball about the origin in $\mathbb{M}(n; \mathbb{C})$ as a diffeomorphism onto an open neighborhood of $m$.

It is also the case that $e_m$ sends the complex, codimension 1 vector space of traceless matrices in $\mathbb{M}(n; \mathbb{C})$ into $\text{Sl}(n; \mathbb{C})$ when $m \in \text{Sl}(n; \mathbb{C})$; and it restricts to a ball about the origin in this subspace as a diffeomorphism onto an open neighborhood of $m$ in $\text{Sl}(n; \mathbb{C})$. Likewise, $e_m$ maps the vector subspace $\mathbb{A}(n; \mathbb{C}) \subset \mathbb{M}(n; \mathbb{C})$ of anti-Hermitian matrices into $\text{U}(n)$ when $m \in \text{U}(n)$ and it restricts to some $m$-independent ball about the origin in $\mathbb{A}(n; \mathbb{C})$ as a diffeomorphism onto a neighborhood of $m$ in $\text{U}(n)$. Finally, this same $e_m$ maps the vector subspace $\mathbb{A}_0(n; \mathbb{C}) \subset \mathbb{A}(n; \mathbb{C})$ of trace zero matrices into $\text{SU}(n)$ when $m \in \text{SU}(n)$; and it restricts to some $m$-independent ball about the origin in $\mathbb{A}_0(n; \mathbb{C})$ as a diffeomorphism onto an open neighborhood of $m$ in $\text{SU}(n)$. The proofs of all of these assertions are identical to the proofs of their analogs given in the preceding part of this chapter.
Let \( i \) denote the inclusion map from any of \( \text{Sl}(n; \mathbb{C}), \text{U}(n) \) or \( \text{SU}(n) \) into \( \text{Gl}(n; \mathbb{C}) \). In the case of \( \text{Sl}(n; \mathbb{C}) \), the pull-back \( i^* \omega_q \) is zero if \( q = c \) with \( c \in \mathbb{C} \) and nowhere zero otherwise. The form \( i^* \omega_q \) on \( \text{Sl}(n; \mathbb{C}) \) is a left-invariant, \( \mathbb{C} \)-valued 1-form, and any such 1-form can be written as \( i^* \omega_q \) with \( \text{tr}(q) = 0 \). Any \( \mathbb{R} \)-valued, left invariant 1-form is the real part of some \( i^* \omega_q \). In the case of \( \text{U}(n) \), the 1-form \( i^* \omega_q \) is zero if \( q^\dagger = q \), and it is nowhere zero if \( q^\dagger = -q \). Any such form is left-invariant, and any \( \mathbb{C} \)-valued, left invariant 1-form on \( \text{U}(n) \) is of this sort. The proofs of all the preceding statements are essentially identical to the proofs of the analogous statements in the preceding part of this chapter.

As noted at the outset, Chapter 2s provides two views of \( \text{Gl}(n; \mathbb{C}) \). The second view regards \( M(n) \subset M(n; 2\mathbb{R}) \) as the subvector space over \( \mathbb{R} \) of matrices that commute with a certain matrix, \( j_0 \), whose square is \(-1\). The latter subvector space is denoted by \( M_j \). \( \text{Gl}_j \subset \text{Gl}(2n; \mathbb{R}) \) denote the latter’s intersection with \( M_j \). Let \( a \rightarrow e^a \) denote the exponential map defined on \( M(n; 2\mathbb{R}) \). As can be seen from the formula, this map has the property that \( j_0 e^a j_0 = e^{j_0 a j_0} \). As a consequence, the \( M(n; 2\mathbb{R}) \) version of \( e^* \) restricts to \( M_j \) so as to map the latter vector space to the subgroup \( \text{Gl}_j \) when \( m \in M_j \). The identification of \( M_j \) with \( M(n; \mathbb{C}) \) that is obtained by writing \( m \in M_j \) as a linear combination of the eigenvalue +i eigenvectors of \( j_0 \) and their complex conjugates identifies the restriction to \( M_j \) of this \( M(2n; \mathbb{R}) \) version \( e^* \) with the \( M(n; \mathbb{C}) \) version of \( e^* \). This understood, all of the assertions given so far in this part of Chapter 5 that concern the exponential map on \( \text{Gl}(n; \mathbb{C}) \) and its subgroups can be reproved without recourse to complex numbers by considering the \( M(2n; \mathbb{R}) \) version of \( e^* \) on \( M_j \). The identification between \( M_j \) and \( M(n; \mathbb{C}) \) and between \( \text{Gl}_j \) and \( \text{Gl}(n; \mathbb{C}) \) can also be used to prove the assertions about left and right invariant 1-forms and vector fields with out recourse to complex numbers.

5g) Immersion, submersions and transversality

This part of the chapter uses the notion of vector push-forward to reinterpret and then elaborate on what is said in Chapter 1d. To set the stage, suppose that \( M \) and \( Y \) are smooth manifolds, and suppose that \( \psi: Y \rightarrow M \) is a smooth map. Chapter 1d introduced the descriptives immersion and submersion as applied to \( \psi \). These notions can be rephrased in terms of the push-forward \( \psi_* \) as follows:
Definition: The map $\psi$ is respectively an immersion or submersion when the vector bundle homormorphism $\psi_*: TY \to \psi^*TM$ is injective on each fiber or surjective on each fiber.

A useful notion that is intermediate to the immersion and immersion cases is described next. To set this stage, suppose that $Y, M$ and $\psi$ are as described above. Suppose in addition that $Z \subset Y$ is a submanifold. The map $\psi$ is said to be transversal to $Z$ when the following condition is met: Let $y \in Y$ denote any given point with $\psi(y) \in Z$. Then any vector $v \in TM|_{\psi(y)}$ can be written as $\psi_*v_Y + v_Z$ with $v_Y \in TY|_y$ and $v_Z \in TZ|_{\psi(y)}$.

The proposition that follows applies the notion of transversality to construct a submanifold of $Y$. This proposition uses $n, m$ and $d$ for the respective dimensions of $Y, M$ and $Z$

Proposition: Let $M, Y$ and $Z \subset M$ be as just described, and suppose that $\psi: Y \to M$ is transversal to $Z$. Then $\psi^{-1}(Z)$ is a smooth submanifold of $Y$ of dimension $n + d - m$.

The proof of this proposition is given momentarily. Note in the meantime that when $Z$ is a point, then $\psi$ is transversal to $Z$ if and only if $\psi_*$ is surjective at each point in $\psi^{-1}(z)$.

Proof: Let $y$ denote a given point in $\psi^{-1}(Z)$. Fix an open neighborhood, $U \subset Y$, of $y$ with a coordinate chart map $\phi: U \to \mathbb{R}^n$ that sends $y$ to the origin. Fix, in addition, an open neighborhood, $V \subset M$, of $\psi(y)$ with a coordinate chart map $\phi': V \to \mathbb{R}^m$ that sends $\psi(y)$ to the origin and sends $Z \cap V$ to the $d$-dimensional subspace of points $(u_1, \ldots, u_m) \in \mathbb{R}^m$ with $u_{d+1} = \cdots = u_m = 0$. Chapter 1c says something such a coordinate chart map. Let $f = (f_1, \ldots, f_m)$ denote the map from $\mathbb{R}^n$ to $\mathbb{R}^m$ given by $\phi' \circ \psi \circ \phi^{-1}$.

Let $f$ denote the map from a neighborhood of the origin in $\mathbb{R}^n$ to $\mathbb{R}^{m-d}$ given by

$$x \to f(x) = (f_{d+1}(y), \ldots, f_m(y)).$$

The transversality assumption guarantees that $f$ is a submersion near the origin in $\mathbb{R}^n$. As a consequence, $f^{-1}(0)$ is a submanifold near the origin in $\mathbb{R}^n$. This submanifold is, by construction, $\phi(\psi^{-1}(Z))$.

Additional reading
• *Introduction to Smooth Manifolds*, John M. Lee; Springer New York, 2003
• *A geometric approach to differential forms*, David Bachman; Birkhauser Boston 2006.
Chapter 6: Vector bundles with $\mathbb{C}^n$ as fiber

Just as there are vector spaces over $\mathbb{C}$, there are vector bundles whose fibers can be consistently viewed as $\mathbb{C}^n$ for some $n$. This chapter first defines these objects, and then supplies a number of examples.

6a) Definitions

What follows directly are three equivalent definitions of a complex vector bundle. The first definition highlights the fact that a complex vector bundle is a vector bundle of the sort introduced in Chapter 3a with some extra structure.

Definition 1: A complex vector bundle of rank $n$ is a vector bundle, $\pi: E \to M$, in the sense defined in Chapter 3a with fiber dimension $2n$, but equipped with a bundle endomorphism $j: E \to E$ such that $j^2 = -1$. This $j$ allows one to define an action of $\mathbb{C}$ on $E$ such that the real numbers in $\mathbb{C}$ act as $\mathbb{R}$ on the underlying real bundle of fiber dimension $2n$; and such that the number $i$ acts as $j$. Doing this identifies each fiber of $E$ with $\mathbb{C}^n$ up to the action of $\text{Gl}(n; \mathbb{C})$. The endomorphism $j$ is said to be an almost complex structure for $E$. The underlying real bundle with fiber $\mathbb{R}^{2n}$ is denoted by $E_\mathbb{R}$ when a distinction between the two is germane at any given time.

The preceding definition highlights the underlying real bundle structure. There are two equivalent definitions that put the complex numbers to the fore.

Definition 2: This definition gives the $\mathbb{C}$ analog of what is given in Chapter 3a: A complex vector bundle $E$ over $M$ of fiber dimension $n$ is a smooth manifold with the following additional structure:

- A smooth map $\pi: E \to M$.
- A smooth map $\delta: M \to E$ such that $\pi \circ \delta$ is the identity.
- A smooth map $\mu: \mathbb{C} \times E \to E$ such that
  a) $\pi(\mu(c, v)) = \pi(v)$
  b) $\mu(c, \mu(c', v)) = \mu(cc', v)$
  c) $\mu(1, v) = v$
  d) $\mu(c, v) = v$ for $c \neq 1$ if and only if $v \in \text{image}(\delta)$.
- Let $p \in M$. There is a neighborhood, $U \subseteq M$, of $p$ and a map $\lambda_u: \pi^{-1}(U) \to \mathbb{C}^n$ such that $\lambda_u: \pi^{-1}(x) \to \mathbb{C}^n$ for each $x \in U$ is a diffeomorphism obeying $\lambda_u(\mu(c, v)) = c\lambda_u(v)$.

Definition 3: This definition gives the $\mathbb{C}$ analog of the cocycle definition from Chapter 3b: A complex vector bundle of rank $n \geq 1$ is given by a locally finite open cover $\mathcal{U}$ of $M$ with the following additional data: An assignment to each pair $U, V \in \mathcal{U}$ a
map \( g_{UV} : U \cap V \to \text{Gl}(n; \mathbb{C}) \) such that \( g_{UV} \circ g_{VW} = g_{VU}^{-1} \) on \( U \cap V \cap W \). Given this data, define the bundle \( E \) to be the quotient of the disjoint union \( \bigcup_{U \in \mathcal{U}} (U \times \mathbb{C}^n) \) by the equivalence relation that puts \((p', v') \in U' \times \mathbb{C}^n\) equivalent to \((p, v) \in U \times \mathbb{C}^n\) if and only if \( p = p' \) and \( v = g_{U'U}(p)v \). The constraint involving the overlap of three sets guarantees that this does, indeed, specify an equivalence relation.

6b) Comparing definitions

The argument to equate the second and third definitions is almost verbatim what is given in Chapter 3b that equates the latter’s cocycle definition of a vector bundle with the definition given in Chapter 3a. Indeed, one need only change \( \mathbb{R} \) to \( \mathbb{C} \) in Chapter 3b’s argument.

To compare the first and second definitions, suppose first that \( E \) is a complex vector bundle in the sense of Definition 2. The needed data from Chapter 3a for the underlying real bundle \( E_\mathbb{R} \) consist of a projection map \( \pi : E_\mathbb{R} \to M \), a zero section \( \delta : M \to E_\mathbb{R} \), a multiplication map \( \mu : \mathbb{R} \times E \to E \). In addition, each \( p \in M \) should be contained in an open set \( U \) which comes with a certain map \( \lambda_U : E_\mathbb{R}|_U \to \mathbb{R}^{2n} \).

The \( E_\mathbb{R} \) is, first of all, declared equal to \( E \). Chapter 3a’s projection \( \pi \) is the projection given in Definition 2. Likewise, Chapter 3a’s zero section \( \delta \) is that given in Definition 2. Meanwhile, Chapter 3a’s map \( \mu \) from \( \mathbb{R} \times E_\mathbb{R} \) to \( E_\mathbb{R} \) is obtained from Definition 2’s map \( \mu \) by restricting the latter to the real numbers \( \mathbb{R} \subset \mathbb{C} \). Granted the preceding, any given version of Chapter 3a’s map \( \lambda_U \) can be taken equal to the corresponding version of Definition 2’s map \( \lambda_U \) after identifying \( \mathbb{C}^n \) as \( \mathbb{R}^{2n} \) by writing the complex coordinates \((z_1, \ldots, z_n)\) of \( \mathbb{C}^n \) in terms of their real and imaginary parts to obtain the Euclidean coordinates for \( \times_n \mathbb{R}^2 = \mathbb{R}^{2n} \). With \( E_\mathbb{R} \) so defined, Definition 1’s requires only the endomorphism \( j \). The latter is defined by the rule \( j(v) = \mu(i, v) \) with \( \mu \) here given by Definition 2.

Now suppose that \( E \) is a complex bundle according to the criteria from Definition 1. The data needed for Definition 2 are obtained as follows: First, the Definition 2’s projection \( \pi \) and Definition 2’s zero section \( \delta \) are those from the underlying real bundle \( E_\mathbb{R} \). Definition 2’s map \( \mu \) is defined as follows: Write a given complex number \( z \) in terms of its real and imaginary parts as \( z = a + ib \). Let \( v \in E \). Then \( \mu(a + ib, v) \) is defined to be \( a v + b jv \). It remains only to specify, for each \( p \in M \), an open set, \( U \), containing \( p \) with a suitable map \( \lambda_U : E_\mathbb{R}|_U \to \mathbb{C}^n \). This last task requires a preliminary lemma.

To set the stage for the required lemma, reintroduce the standard almost complex structure \( j_0 \) from Chapter 2d.
**Lemma:** Let $\pi: E_{\mathbb{R}} \rightarrow M$ denote a real vector bundle with fiber dimension $2n$ with an almost complex structure $j: E \rightarrow E$. There is a locally finite cover $\Omega$ for $M$ of the following sort: Each set $U \in \Omega$ comes with an isomorphism $\phi_U: E_{\mathbb{R}|U} \rightarrow U \times \mathbb{R}^{2n}$ that obeys $\phi_U j \phi_U^{-1} = j_0$ at each point of $U$.

This lemma is proved momentarily. It has the following consequence: Let $U$ denote a cover of the sort given by the lemma. Let $U \subset U$, and set $\lambda_U : E_{\mathbb{R}|U} \rightarrow \mathbb{R}^{2n}$ equal to the composition of $\phi_U$ followed by the projection to the $\mathbb{R}^n$ factor. Define the complex coordinates $(z_1, \ldots, z_n)$ for $\mathbb{C}^n$ from the real coordinates $(x_1, \ldots, x_{2n})$ of $\mathbb{R}^{2n}$ by writing $z_k = x_{2k-1} + ix_{2k}$. Doing so allows $\lambda_U$ to be viewed as a map, $\lambda_U$, from $E_{\mathbb{R}|U}$ to $\mathbb{C}^n$. Because $\phi_U$ intertwines $j$ with $j_0$, the latter map is equivariant with respect to the action of $\mathbb{C}$ on $E$ and $\mathbb{C}$ on $\mathbb{C}^n$. And, it restricts to any given fiber as an isomorphism between complex vector spaces.

**Proof of the lemma:** Fix $p \in M$ and an open set $V \subset M$ that contains $p$ with an isomorphism $\psi_V: E_{\mathbb{R}|V} \rightarrow V \times \mathbb{R}^{2n}$. Use $j_V$ to denote $\psi_V j \psi_V^{-1}$. At each point of $V$, this endomorphism of $\mathbb{R}^{2n}$ defines an almost complex structure on $\mathbb{R}^{2n}$. The plan in what follows is to find a smooth map $h$, from a neighborhood of $p$ in $V$ to $\text{Gl}(2n; \mathbb{R})$ such that $h j_V h^{-1} = j_0$. Given such a map, set $U$ to be this same neighborhood and define $\phi_U$ at any given point in $U$ to be $h \psi_V$.

To obtain $h$, digress momentarily and suppose that $j_1$ is any given almost complex structure. Then there exists $g \in \text{Gl}(2n; \mathbb{R})$ such that $g j_1 g^{-1} = j_0$. Indeed, this follows by virtue of the fact that $j_1$ and $j_0$ are diagonalizable and have eigenvalues only $\pm i$, each with multiplicity $n$. To elaborate, fix a basis of eigenvectors for $j_1$ with eigenvalue $+i$ and likewise a basis for $j_0$. Define $g$ by the rule that sends the respective the real and imaginary parts of the $k$'th basis vector for $j$ to those of the corresponding basis vector for $j_0$. Granted the preceding, it follows that $\psi_V$ can be changed by $\psi_V \mapsto g \psi_V$ with $g \in \text{Gl}(2n; \mathbb{R})$ so that the resulting version of $j_V$ is equal to $j_0$ at $p$.

With the preceding understood, reintroduce the exponential map $a \mapsto e^a$ as defined in Chapter 5e, and recall that the latter restricts to some ball about the origin in $\mathbb{M}(2n; \mathbb{R})$ as a diffeomorphism onto a neighborhood of $t$ in $\text{Gl}(2n; \mathbb{R})$. Granted the latter fact, it then follows that $j_V$ near $p$ can be written as $j_0 e^a$ where $a$ is a smooth map from a neighborhood of $p$ in $V$ to $\mathbb{M}(2n; \mathbb{R})$. Moreover, $a$ must be such that

$$e^{-j_0 a} e^a = 1,$$

67
for this condition is necessary and sufficient for $j_v^2$ to equal $-\mathbf{1}$. Because $(e^a)^{-1} = e^{-a}$, this last condition requires that $-j_0 \alpha j_0 = -\alpha$ on some neighborhood of $p$. With this last point in mind, use the power series definition of $e^a$ to write the latter $e^a = e^{a/2}e^{a/2}$, and then write

$$j_v = j_0 e^{a/2} e^{a/2} = e^{-a/2}j_0 e^{a/2}.$$ 

Then take $\mathfrak{h} = e^{a/2}$.

6c) Examples: The complexification

The simplest example of a complex vector bundle is the product bundle $M \times \mathbb{C}^n$. The underlying rank $2n$ real bundle, $E_R$, is $M \times \mathbb{R}^{2n}$; the almost complex structure is $j_0$. The examples given next are not much more complicated. These examples start with some given vector bundle $\pi: E \to M$ is a vector bundle with fiber $\mathbb{R}^n$. The resulting complex bundle has fiber $\mathbb{C}^n$; it is denoted $E_C$ and is said to be the complexification of $E$. The bundle $E_C$ is defined so that its fiber of $E_C$ at any given $x \in M$ is canonically isomorphic to $\mathbb{C} \otimes_{\mathbb{R}} E|_x$ as a complex vector space. What follows describes $E_C$ in the context of Definition 1 from Chapter 6a.

To start, introduce the tensor product bundle $(M \times \mathbb{R}^2) \otimes E$, this a bundle with fiber $\mathbb{R}^{2n}$. Define $j$ on this bundle as follows: Any element in the latter can be written as a finite sum of reducible elements, $\sum z_k \otimes e_k$, where $z_k \in \mathbb{R}^2$ and $e_k \in E$. This understood, define an almost complex structure on this tensor product bundle so as send $\sum z_k \otimes e_k$ to $\sum (j_0z_k) \otimes e_k$, where

$$j_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$ 

The resulting complex bundle is $E_C$. As advertised, $E_C$ is a rank $n$ complex bundle. The $\mathbb{C}$ action on $E_C$ can be seen by writing $\mathbb{R}^2$ as $\mathbb{C}$ so as to view $E_C$ as $(M \times \mathbb{C}) \otimes_{\mathbb{R}} E$. Write a typical element as $v = \sum z_k \otimes e_k$ but now view $z_k \in \mathbb{C}$. Then element $\sigma \in \mathbb{C}$ acts to send this element to $\sigma v = \sum z_k \sigma z_k \otimes e_k$.

Suppose that $E$ is a complex bundle already. This is to say that it can be viewed as a real bundle with an almost complex structure $j$. Then $E$ sits inside its complexification, $E_C$, as a complex, rank $n$ sub-bundle. In fact, $E_C$ has a direct sum decomposition, linear over $\mathbb{C}$, as $E \oplus \bar{E}$, where $\bar{E}$ is defined from the same real bundle as was $E$, but with $j$ replaced by $-j$. The bundle $\bar{E}$ is called the complex conjugate bundle to $E$.  

68
To see $E$ inside $E_C$, observe that the endomorphism $j$ acts as an endomorphism of $E_C$ sending any given $v = \sum_k z_k \otimes e_k$ to $jv = \sum_k z_k \otimes je_k$. As an endomorphism now of $E_C$, this $j$ obeys $j^2 v = -v$, and it also commutes with multiplication by elements in $\mathbb{C}$. This understood, the bundle $E$ sits inside the bundle $E_C$ as the set of vectors $v$ such that $jv = iv$. Meanwhile, the conjugate bundle $\overline{E}$ sits in $E_C$ as the set of vectors $v$ such that $jv = -iv$.

The inclusion of $E$ into $E_C$ as a direct summand sends any given vector $e \in E$ to the vector $\frac{1}{2} (1 \otimes e - i \otimes je)$. In general, $\overline{E}$ and $E$ are not isomorphic as complex bundles even though their underlying real bundles (i.e., $E_{\mathbb{R}}$ and $\overline{E}_{\mathbb{R}}$) are identical.

6d) Complex bundles over surfaces in $\mathbb{R}^3$

Introduce the Pauli matrices

$$
\tau_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}
$$

These generate the quaternion algebra as

$$\tau_1^2 = \tau_2^2 = \tau_3^3 = -i \quad \text{and} \quad \tau_1 \tau_2 = -\tau_2 \tau_1 = -\tau_3, \quad \tau_2 \tau_3 = -\tau_3 \tau_2 = -\tau_1, \quad \tau_3 \tau_1 = -\tau_1 \tau_3 = -\tau_2.$$

Here, $i$ denotes the $2 \times 2$ identity matrix. Let $\Sigma \subset \mathbb{R}^3$ denote a given embedded or immersed surface. Assign to each point $p \in \Sigma$ its normal vector $n(p) = (n_1, n_2, n_3)$. This assignment is used to define the map $n: \Sigma \to \mathbb{M}(2; \mathbb{C})$ given by $n = n_1 \tau_1 + n_2 \tau_2 + n_3 \tau_3$.

Note that $n^2 = -i$. Define a complex vector subbundle $E_2 \subset \Sigma \times \mathbb{C}^2$ to be the set of points $(p, v)$ such that $n(p)v = iv$. As explained momentarily, this is a complex vector bundle of rank 1 over $\Sigma$.

In general, if $M$ is a smooth manifold and $f = (f_1, f_2, f_3): M \to \mathbb{R}^3$ with image in $S^2$ is a smooth map, then there is the corresponding subbundle $E \subset M \times \mathbb{C}^2$ given by the set $(p, v): f(p) = iv$ where $f = \sum_j f_j \tau_j$. What follows uses Definition 2 of Chapter 1a to verify that $E$ does indeed define a complex vector bundle. Use for $\pi$ the restriction of the projection from $M \times \mathbb{C}^2$ to $M$. Likewise, use for $\mu$ the restriction of the multiplication on $\mathbb{C}^2$ by elements in $\mathbb{C}$. Take $\delta$ to be the map $x \to (x, 0) \in M \times \mathbb{C}^2$. To see about the final item, fix $p \in M$ and fix a unit length vector $v_p \in E|_p \subset \mathbb{C}^2$. This done, write $v_p$ the two components of $v_p$ as $(a_p, b_p)$ and write those of any other vector $v$ as $(a, b)$. Now define the $\mathbb{C}$-linear map $\eta_p: \mathbb{C}^2 \to \mathbb{C}$ by the rule that sends $(a, b) \in \mathbb{C}^2$ to the complex number $\overline{a}a + \overline{b}b$. As $\eta_p$ restricts to $E|_p$ as an isomorphism (over $\mathbb{C}$), so it restricts to a
neighborhood \( U \subset M \) of \( p \) so as to map each fiber of \( E|_U \) isomorphically to \( \mathbb{C} \). Define \( \lambda_U \) to be this restriction.

Here is another example of a map \( f \) from a compact surface to \( \mathbb{R}^3 \) to \( S^2 \): Let \( \Sigma \) denote the surface in question. Fix a point, \( x \in \mathbb{R}^3 \), that lies in the interior of the region bounded by \( \Sigma \). A map \( f: \Sigma \to S^2 \) is defined by taking for \( f(p) \) the unit vector \( (p - x)/|p - x| \). Alternately, one can take \( x \) to lie outside of \( \Sigma \) to define \( E \).

6e) The tangent bundle to a surface in \( \mathbb{R}^3 \)

The tangent bundle to a surface in \( \mathbb{R}^3 \) has an almost complex structure that is defined as follows: Let \( \Sigma \subset \mathbb{R}^3 \) denote the surface and let \( n = (n_1, n_2, n_3) \) again denote its normal bundle. As explained in Chapters 3d and 3e, the tangent bundle \( T\Sigma \) can be identified as the set of pair \((p, v) \in \Sigma \times \mathbb{R}^3 \) such that the dot product \( v \cdot n = 0 \). Define \( j: T\Sigma \to T\Sigma \) by the rule \( j(p, v) = (p, n \times v) \), where \( \times \) is the cross product on \( \mathbb{R}^3 \). Since \( n \times n \times v = -v \), it follows that \( j \) is an almost complex structure. As a consequence, we may view \( T\Sigma \) as a complex vector bundle of rank 1 over \( \Sigma \). This complex bundle is denoted often by \( T_{1,0}\Sigma \).

6f) Bundles over 4-dimensional submanifolds in \( \mathbb{R}^5 \)

Introduce a set of five matrices in \( M(4; \mathbb{C}) \) as follows: Written in 2 \( \times \) 2 block diagonal form, they are:

\[
\begin{align*}
\text{For } j = 1, 2, 3: \quad \gamma_j &= \begin{pmatrix} 0 & \tau_j \\ \tau_j & 0 \end{pmatrix}, \\
\gamma_4 &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \gamma_5 = i \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\end{align*}
\]

Here, \( \{\tau_a\}_{1,\ldots,3} \) again denote the Pauli matrices. Note that for each \( j \), one has \( \gamma_j^2 = -1 \). Note also that \( \gamma_i \gamma_j + \gamma_j \gamma_i = 0 \) for \( i \neq j \). Let \( M \subset \mathbb{R}^5 \) denote a dimension 4 submanifold; the sphere \( S^4 \) is an example. Let \( n = (n_1, \ldots, n_5) \) denote the normal vector to \( M \) at any given point. Define the map \( n: M \to M(4; \mathbb{C}) \) by sending \( p \) to \( n(p) = \sum \gamma_i n_i \). Then \( n^2 = -1 \). As a consequence, at each point in \( M \), the matrix \( n \) has a 2-dimensional (over \( \mathbb{C} \)) space of eigenvectors with eigenvalue \( +i \). Define the bundle \( E \subset M \times \mathbb{C}^4 \) to be the set \((p, v): n(p)v = iv \). This is a vector bundle over \( M \) with fiber \( \mathbb{C}^2 \).

6g) Complex bundles over 4-dimensional manifolds

What follows uses the cocycle definition of a complex vector bundle to construct some rank 2 complex bundles over a given 4-dimensional manifold. Let \( M \) denote a four dimensional manifold. Fix a set finite set \( \Lambda \subset M \) of distinct points. Assign to each \( p \in \Lambda \) an integer \( m(p) \in \mathbb{Z} \). Also assign to each \( p \in M \) a coordinate chart \( U_p \subset M \) that contains
p with diffeomorphism $\varphi_p: U_p \to \mathbb{R}^4$. Make these assignments so that the charts for distinct pairs $p$ and $p'$ are disjoint.

To define a vector bundle from this data, introduce $\mathcal{U}$ to denote the open cover of $M$ given by the sets $U_0 = M - \Lambda$ and the collection $\{U_p\}_{p \in \Lambda}$. As the only non empty intersections are between $U_0$ and the sets from $\{U_p\}_{p \in \Lambda}$, a vector bundle over $M$ with fiber $\mathbb{C}^2$ is defined by specifying a set of maps $\{g_{0p}: U_0 \cap U_p \to \text{Gl}(2, \mathbb{C})\}_{p \in \Lambda}$.

In order to specify these maps, first define a map $g: \mathbb{C}^2 \to \mathbb{M}(2; \mathbb{C})$ by the rule that sends $z = (z_1, z_2) \mapsto g(z) = \begin{pmatrix} z_1 & -\bar{z}_2 \\ z_2 & \bar{z}_1 \end{pmatrix}$.

Note that $\det(g(z)) = |z|^2$ so $g$ maps $\mathbb{C}^2 - 0$ to $\text{Gl}(2; \mathbb{C})$. Use $j_0$ to identify $\mathbb{R}^4$ with $\mathbb{C}^2$ so as to view $g$ as a map from $\mathbb{R}^4$ to $\mathbb{M}(2; \mathbb{C})$ that sends $\mathbb{R}^4 - 0$ to $\text{Gl}(2; \mathbb{C})$.

Now let $p \in \Lambda$. Since $\varphi_p(U_0 \cap U_p) = \mathbb{R}^4 - 0$, the composition $x \mapsto g(\varphi_p(x))^{m(p)}$ for $x \in U_0 \cap U_p$ defines a map from $U_0 \cap U_p$ to $\text{Gl}(2; \mathbb{C})$. The latter map is defined to be $g_{0p}$.

6h) Complex Grassmannians

Important examples of complex vector bundles involve the $\mathbb{C}^n$ analogs of the Grassmannians that are described in Chapter 1f. To say more, fix an integers $m > 1$ and $n \in \{1, \ldots, n-1\}$; then set $\text{Gr}_{\mathbb{C}}(m; n)$ denote the set of $m$ dimensional complex subvector spaces in $\mathbb{C}^n$. As is explained momentarily, $\text{Gr}_{\mathbb{C}}(m; n)$ is a smooth manifold of dimension $2n(m-n)$. Moreover, $\text{Gr}_{\mathbb{C}}(m; n)$ has over it a very natural, rank $n$, complex vector bundle. The latter sits in $\text{Gr}_{\mathbb{C}}(m; n) \times \mathbb{C}^m$ as the set

$$E = \{(V, v): V \text{ is an } n\text{-dimensional subspace of } \mathbb{C}^m \text{ and } v \in V\}.$$ 

The bundle $E$ is the $\mathbb{C}$-analog of what is described in Chapter 3c. This $\text{Gr}_{\mathbb{C}}(m; n)$ is called the Grassmannian of $n$-planes in $\mathbb{C}^m$, and the bundle $E$ is called the tautological $n$-plane bundle.

The argument to prove that $\text{Gr}_{\mathbb{C}}(m; n)$ has the structure of a smooth manifold is much the same as the argument used in Chapter 1f to prove that the $\mathbb{R}$ versions of the Grassmannians are smooth manifolds. Likewise, the argument that $E$ is a complex vector bundle over $\text{Gr}_{\mathbb{C}}(m; n)$ is much the same as that given in Section 3c for the $\mathbb{R}$ analog of $E$. Even so, it is worth summarizing both arguments because of a point that is referred to in a later chapter.
What follows directly is a brief interlude to set some notation. To start, suppose that \( v \) and \( w \) are vectors in \( \mathbb{C}^m \). Their Hermitian inner product is denoted by \( \overline{v} \cdot w \); it is defined using their respective components by

\[
\overline{v} \cdot w = \overline{v}_1 w_1 + \cdots + \overline{v}_m w_m.
\]

Thus, the norm of a vector \( v \in \mathbb{C}^m \) is \( |v| = (\overline{v} \cdot v)^{1/2} \). A basis \( \{e_i, \ldots, e_n\} \) for an \( n \)-dimensional subspace \( V \subset \mathbb{C}^m \) is orthonormal when \( \overline{e}_i e_j = 1 \) or \( 0 \) depending on whether \( i = j \) or \( i \neq j \).

The orthogonal projection \( \Pi_v : \mathbb{C}^m \to V \) can be defined as follows: Fix an orthonormal basis, \( \{e_1, \ldots, e_n\} \) for \( V \). Then

\[
\Pi_v w = \sum_{1 \leq i \leq n} (\overline{e}_i \cdot w) e_i.
\]

To continue setting notation, introduce \( M_{\mathbb{C}}(m; n) \) to denote the space of \( \mathbb{C} \)-linear homomorphisms from \( \mathbb{C}^n \) to \( \mathbb{C}^{m-n} \). This is to say that each point in \( M_{\mathbb{C}}(m; n) \) is a matrix with \( n \) columns and \( m-n \) rows whose entries are complex numbers. The entries serve as coordinates and so identify \( M_{\mathbb{C}}(m; n) \) with \( \mathbb{C}^{n(m-n)} \).

With this notation set, what follows next is a description of the smooth manifold structure on \( \text{Gr}_c(m; n) \). To start, define a topology on \( \text{Gr}_c(m; n) \) as follows: A basis of open neighborhoods for any given \( n \)-dimensional subspace \( V \) is labeled by the positive numbers, with the neighborhood labeled by a given \( \varepsilon > 0 \) consists of the set of subspace \( V' \) such that \( |\Pi_v v - v| < \varepsilon \).

A coordinate chart for a neighborhood of a given \( V \subset \text{Gr}_c(m; n) \) is the set \( \mathcal{O}_V \) of \( n \)-planes \( V' \) such that \( \Pi_v \) maps \( V' \) isomorphically to \( V \). A coordinate map, \( \varphi_v \), from \( \mathcal{O}_V \) to \( \mathbb{R}^{2n(m-n)} = \mathbb{C}^{n(m-n)} = M_{\mathbb{C}}(m; n) \) is defined as follows: Fix an orthonormal basis \( \{e_k\}_{1 \leq k \leq n} \) for \( V \) so as to identify this vector space with \( \mathbb{C}^n \). Let \( V^\perp \subset \mathbb{C}^m \) denote the Hermitian conjugate space, thus the kernel of \( \Pi_v \). Fix an orthonormal basis \( \{u_k\}_{1 \leq k \leq m-n} \) for \( V^\perp \) to make a \( \mathbb{C} \)-linear identification of the latter with \( \mathbb{C}^{m-n} \). Now let \( p \) denote a given point in \( \mathbb{C}^{n(m-n)} \). Viewed as a matrix in \( M_{\mathbb{C}}(m; n) \), it has components \( \{p_{ia}\}_{1 \leq i \leq n} \). The map \( \varphi^{-1}_v \) sends \( \mathbb{C}^{n(m-n)} \) to the \( n \)-dimensional subspace spanned by the set \( \{e_i + \sum_{1 \leq a \leq m-n} w_{ia} u_a\}_{1 \leq i \leq n} \).

Let \( V \) and \( V' \) denote any two points in \( \text{Gr}_c(m; n) \). Fix orthonormal bases \( \{e_i\}_{1 \leq i \leq n} \) and \( \{u_a\}_{1 \leq a \leq m-n} \) for \( V \) and \( V'^\perp \) to define the coordinate chart map \( \varphi_V : \mathcal{O}_{V'} \to M_{\mathbb{C}}(m; n) = \mathbb{C}^{n(m-n)} \). This done, then the corresponding transition function \( \varphi_{V'} \circ \varphi^{-1}_V \) sends a given point \( p \in \varphi_V(\mathcal{O}_V \cap \mathcal{O}_{V'}) \) to the point \( p' \in \varphi_{V'}(\mathcal{O}_V \cap \mathcal{O}_{V'}) \) with the latter determined by \( p \) using the rule

\[
p'_{ia} = \sum_{1 \leq k \leq n} (G^{-1})_{ak} (\overline{u}_a e_k + \sum_{1 \leq b \leq m-n} p_{kb} \overline{u}_b u_b),
\]
where the notation is as follows: First, any given \( i, k \in \{1, \ldots, n\} \) version of \( G_{ik} \) is given by

\[
G_{ik} = \bar{e}_i \cdot e_j + \sum_{1 \leq a \leq n} p_{ia} \bar{e}_i \cdot u_a.
\]

Second, \( G \) is the \( n \times n \) matrix with \((i, k)\) entry equal to \( G_{ik} \). Note that \( G \subset \text{Gl}(n; \mathbb{C}) \) by virtue of the fact that \( \phi_V^{-1}(p) \in \mathcal{O}_V \cap \mathcal{O}_{V'} \).

The simplest example is that of \( \text{Gr}_c(m; 1) \), the set of 1-dimensional complex subspaces in \( \mathbb{C}^m \). This manifold is usually called the \( m-1 \) dimensional complex projective space; and it is denoted by \( \mathbb{C}P^{m-1} \). This manifold has an open cover given by the collection of sets \( \{\mathcal{O}_1, \ldots, \mathcal{O}_m\} \) where \( \mathcal{O}_k \) is the version of what is denoted above by \( \mathcal{O}_V \) with \( V \) the complex line spanned by the vector in \( \mathbb{C}^m \) with \( k \)’th entry 1 and all other entries 0. The complex span of a vector \( v \in \mathbb{C}^m \) is in \( \mathcal{O}_k \) if and only if the \( k \)’th entry of \( v \) is non-zero. Noting that \( \mathbb{M}_c(m; 1) = \mathbb{C}^{m-1} \), the \( \mathcal{O}_k \) version of the coordinate map \( \phi_V^{-1} \) sends \((p_1, \ldots, p_{m-1}) \in \mathbb{C}^{m-1}\) to the line spanned by the vector \((p_1, \ldots, p_{k-1}, 1, p_k, \ldots, p_{m-1})\).

Thus, the \( i \)’th entry of this vector is equal to \( p_i \) when \( i < k \), it is equal to 1 when \( i = k \), and it is equal to \( p_{i+1} \) when \( k < i \). For \( k < k' \), the matrix \( G \) for the intersection of \( \mathcal{O}_k \) with \( \mathcal{O}_k' \) is a \( 1 \times 1 \) matrix, this the non-zero, \( \mathbb{C} \)-valued function on the set \( \mathcal{O}_k \cap \mathcal{O}_{k'} \) given by \( p_k \in \mathbb{C} \setminus \{0\} \). The transition function writes \((p'_1, \ldots, p'_{m-1})\) as the function of \((p_1, \ldots, p_{m-1})\) given as follows:

- If \( i < k' \ then \ p'_i = p_i/p_k \.
- If \( k' \leq i < k, then \ p'_i = p_{i+1}/p_k \.
- \ p'_k = 1/p_k.
- If \( k < i \ then \ p'_i = p_i/p_k \.

What follows next is a brief description of the tautological bundle \( E \to \text{Gr}_c(m; n) \). The data needed for Definition 2 of Chapter 6a are a projection map to \( \text{Gr}_c(m; n) \), a zero section, a multiplication by \( \mathbb{C} \) and a suitable map defined over a neighborhood of any given point from \( E \) to \( \mathbb{C}^n \). The bundle projection map \( \pi: E \to \text{Gr}_c(m; n) \) sends any given pair \((V, v)\) with \( v \in V \) and \( V \) an \( n \)-dimensional subspace of \( \mathbb{C}^m \) to \( V \). The zero section \( \delta \) sends this same \( V \in \text{Gr}_c(m; n) \) to the pair \((V, 0)\). Multiplication by \( \mathbb{C} \) on \( E \) by any given \( c \in \mathbb{C} \) sends \((V, v)\) to \((V, cv)\). For the final requirement, set \( U = \mathcal{O}_V \). Fix the basis \( \{e_i\}_{1 \leq i \leq n} \) for \( V \) as above. Let \( W \in \mathcal{O}_V \). Then the map \( \lambda_{\mathcal{O}_V} \) sends \((W, w) \in E|_W \) to the vector in \( \mathbb{C}^n \) with \( j \)’th entry equal to \( \bar{e}_j \cdot w \).
The map $\lambda_{O_V}$ just defined gives a $\mathbb{C}$-linear map from $E|_{O_V}$ to $O_V \times \mathbb{C}^n$ that respects the vector bundle structure of each fiber. These maps are used next to give a cocycle definition of $E$ in the manner of Definition 3 of Chapter 6a. A suitable chart for this consists of a finite collection of sets of the form $O_V$ for $V \in \text{Gr}_{\mathbb{C}}(m; n)$. To define the transition functions, choose for each point $V$ represented a basis, $\{e_i\}_{1 \leq i \leq n}$, for $V$. This done, suppose that $V$ and $V'$ are two points in $\text{Gr}_{\mathbb{C}}(m; n)$ that are represented. The corresponding bundle transition function $g_{O_V, O_{V'}} : O_V \cap O_{V'} \rightarrow \text{Gl}(n; \mathbb{C})$ is the map $G$ with entries $G_{ij}$ as depicted above.

6i) The exterior product construction

The notion of the exterior product for $\mathbb{C}^n$ can be used to define a complex vector bundle of rank 2 over any given 2n dimensional manifold.

To set the stage, let $\wedge^k \mathbb{C}^n = \oplus_{k=0,1,\ldots,n} (\wedge^k \mathbb{C}^n)$. Here, $\wedge^0 \mathbb{C}^n$ is defined to be $\mathbb{C}$. Meanwhile, $\wedge^1 \mathbb{C}^n = \mathbb{C}^n$ and any given $k > 1$ version of $\wedge^k \mathbb{C}^n$ is the $k$'th exterior power of $\mathbb{C}^n$. Fix a basis $\{e_1, \ldots, e_n\}$ for $\mathbb{C}^n$ and the vector space $\wedge^k \mathbb{C}^n$ has a corresponding basis whose elements are denoted by $\{e_{i_1} \wedge \cdots \wedge e_{i_k}\}_{1 \leq i_1 < \cdots < i_k \leq n}$.

Note that $\wedge^k \mathbb{C}^n$ has dimension $\frac{n!}{k!(n-k)!}$ and so $\dim(\wedge^* \mathbb{C}^n) = 2^n$.

There are two actions of $\mathbb{C}^n$ on this vector space. The first is exterior multiplication. This action has $z \in \mathbb{C}^n$ sending any given element $\omega \in \wedge^* \mathbb{C}^n$ to $v \wedge \omega$. By way of reminder, $z \wedge \omega$ the action is such as to sends $\wedge^k \mathbb{C}^n$ to $\wedge^{k+1} \mathbb{C}^n$. It can be defined by as follows: First, $z \wedge (\cdot)$ sends the element $1 \in \wedge^0 \mathbb{C}^n$ to $z \in \mathbb{C}^n = \wedge^1 \mathbb{C}^n$. It is defined on the $k > 0$ versions of $\wedge^k \mathbb{C}^n$ by its action on the basis vectors given above. To define the latter, write $z = z_1 e_1 + \cdots + z_n e_n$. Then

$$z \wedge e_{i_1} \wedge \cdots \wedge e_{i_k} = \sum_{1 \leq i_1 < \cdots < i_k \leq n} z_{i_1} (e_{i_1} \wedge (e_{i_2} \wedge \cdots \wedge e_{i_k}))$$

where $e_i \wedge (e_{i_1} \wedge \cdots \wedge e_{i_k})$ is zero if some $i_i = i$, and it is $\pm 1$ times the basis vector obtained by adjoining $i$ to the set $\{i_1, \ldots, i_k\}$ otherwise. The coefficient here is $+1$ if the ordering $(i, i_1, \ldots, i_k)$ requires an even number of interchanges to make an increasing sequence, and $-1$ otherwise. Note that this $\pm 1$ business guarantees that $z \wedge (z \wedge (\cdot)) = 0$.

The second action of $z$ is denoted in what follows by $I_z$. This homomorphism is zero on $\wedge^0 \mathbb{C}^n$ and maps any $k > 0$ version of $\wedge^k \mathbb{C}^n$ to $\wedge^{k+1} \mathbb{C}^n$. It is defined inductively as follows: First, $I_z w = \bar{z} \cdot w \in \wedge^0 \mathbb{C}^n = \mathbb{C}$. Now suppose that $I_z$ is defined on $\wedge^k \mathbb{C}^n$ for $k \geq 1$. 

74
To define it on $\wedge^{k+1}C^n$ it is enough to specify the action on elements of the form $w \wedge \omega$ with $w \in C^n$ and $\omega \in \wedge^k C^n$. This understood, the homomorphism $I_z$ sends $w \wedge \omega$ to $(\bar{z} \cdot w) \omega - w \wedge I_z(\omega)$.

Use $I(\cdot)$ and the exterior product to define the map $g: C^n \to \text{End}(\wedge^* C^n) = M(2^n; \mathbb{C})$

by the rule

$$g(z) = I_z + z \wedge$$

Noting that $g^2 = |z|^2 I$, it follows that $g$ maps $\mathbb{C}^n - 0$ to $\text{Gl}(2^n; \mathbb{C})$.

To continue with the preliminaries, suppose now that $M$ is a given $2n$ dimensional manifold. Fix a finite set, $\Lambda$, of distinct points in $M$. For each $p \in \Lambda$, fix a coordinate chart $U_p \subset M$ centered at $p$ with its map $\phi_p: U_p \to \mathbb{R}^{2n}$. Choose these charts so that they are pairwise disjoint. Use the almost complex structure $j_0$ to identify $\mathbb{R}^{2n}$ with $C^n$ so as to view any given $p \in \Lambda$ version of $\phi_p$ as a map to $C^n$.

The data just given will now be used to give a cocycle definition of a complex vector bundle with fiber dimension $2^n$. Take for the open cover the sets $\{U_p\}_{p \in \Lambda}$ with the extra set $U_0 = M - \Lambda$. The only relevant bundle transition functions are those for the sets $\{U_0 \cap U_p\}_{p \in \Lambda}$. For a given point $p \in \Lambda$, take the transition function $g_{U_0 U_p}$ to be the map from $U_0 \cap U_p$ to $\text{Gl}(2^n; \mathbb{C})$ that sends any given point $x$ to $g(\phi_p(x))$.

6j) **Algebraic operations**

All of the algebraic notions from Chapter 4 such as bundle isomorphisms, homomorphisms, subbundles, quotient bundles, tensor products, symmetric products, anti-symmetric projects subbundles, quotient bundles and direct sums have their analogs for complex vector bundles. The definitions are identical to those in Chapter 4 save that all maps are now required to be $\mathbb{C}$-linear.

There are two points that warrant an additional remark with regards to Definition 1 of Chapter 6a. The first involves the notion of a bundle homomorphism. In the context of Definition 1, a complex bundle homomorphism is a homomorphism between the underlying real bundles that intertwines the respective almost complex structures. The second point involves the notion of tensor product. To set the stage, suppose that $E$ and $E'$ are complex vector bundles over $M$ of rank $n$ and $n'$ respectively. Their tensor product is a complex vector bundle over $M$ of rank $nn'$. Note however, that the tensor product $E \otimes E'$ of their underlying real bundles has dimension $4nn'$ which is twice that of $(E \otimes E')_{\mathbb{R}}$. The latter sits in $E \otimes E'$ as the kernel of the homomorphism $j \otimes \iota' - \iota \otimes j'$ where $j$ and $j'$ denote the respective almost complex structures on $E$ and $E'$ while $\iota$ and $\iota'$ denote the respective identity homomorphisms on $E$ and $E'$. 

75
A third comment is warranted with regards to the exterior product. Recall that if \( E \to M \) is a real bundle of fiber dimension \( n \), then \( \wedge^n E \) is a vector bundle of fiber dimension 1 over \( M \). In the case when \( E \) is a complex vector bundle, then \( \wedge^n E \) is a vector bundle with fiber \( \mathbb{C} \). The latter is often denoted by \( \text{det}(E) \). The reason for this appellation is as follows: Suppose that \( E \) is given by the cocycle definition from Chapter 6a. This is to say that a locally finite open cover, \( \mathcal{U} \), of \( M \) is given with suitable bundle transition functions from overlapping sets to \( \text{Gl}(n; \mathbb{C}) \) with \( n \) the fiber dimension of \( E \). Then \( \wedge^n E \) is defined by using the same open cover and but with any given bundle transition function, given by the determinant of the corresponding transition function for \( E \). Note in this regard that any transition function for \( \wedge^n E \) must map to \( \text{Gl}(1; \mathbb{C}) = \mathbb{C} - 0 \).

6k) Pull-back

Let \( M \) and \( N \) denote smooth manifolds and let \( \psi: M \to N \) denote a smooth map. Chapter 5a explains how \( \psi \) is used to construct a real vector bundle over \( M \) from a given real vector bundle over \( N \). If \( \pi: E \to N \) is the given bundle over \( N \), then the corresponding bundle over \( M \) is denoted by \( \psi^*E \). If \( E \) has an almost complex structure, \( j \), then there is a unique almost complex structure for \( \psi^*E \) which is intertwined with \( j \) by the \( \hat{\psi}: \psi^*E \to E \). Indeed, with \((\psi^*E)\) viewed in \( M \times E \) as the set of pairs \((x, v)\) with \( \psi(x) = \pi(v) \), this almost complex structure acts to send \((x, v)\) to \((x, jv)\).

Granted the preceding observation, it follows that the pull-back of any given complex vector bundle \( \pi: E \to N \) can be viewed in a canonical way as a complex vector bundle over \( M \); and having done so, the covering map \( \hat{\psi}: \psi^*E \to E \) is \( \mathbb{C} \)-linear on each fiber.

By way of an example, return to the milieu of Chapter 6d. Let \( \Sigma \subset \mathbb{R}^3 \) denote a given embedded surface. As explained in Chapter 6d, the association to each point in \( \Sigma \) of its normal vector can be used to define a map \( n: \Sigma \to \mathbb{M}(2; \mathbb{C}) \) and the vector bundle subbundle \( E \subset \Sigma \times \mathbb{C}^2 \) given by the set of points \((p, v)\) with \( n(p)v = iv \). With this as background, remark that the assignment to \( p \) of its normal vector also defines a map from \( \Sigma \) to \( S^2 \). The latter, \( n(\cdot) \), is called the Gauss map. Let \( E_1 \to S^2 \) denote the subbundle of the product bundle \( S^2 \times \mathbb{C}^2 \) given by the set of points \((x_1, x_2, x_3, v)\) with \((\sum_{1 \leq i \leq 3} x_i^2\tau_i)v = iv \). The bundle \( E \) is the pull-back \( n^*E_1 \).

By way of a second example, suppose that \( \Sigma \subset \mathbb{R}^3 \) is an embedded surface. Chapter 6e defines the complex rank 1 bundle \( T_{1,0}\Sigma \to \Sigma \). The latter is \( n^*T_{1,0}S^2 \).

The following proposition gives a third application of pull-back. But for some notation, its proof is the same as that of its real bundle analog in Chapter 5b.

**Proposition:** Let \( M \) denote a smooth manifold, let \( n \) denote a positive integer, and let \( \pi: E \to M \) denote a given rank \( n \), complex vector bundle. If \( m \) is sufficiently large, there
exists a map $\psi_m: M \to \text{Gr}_C(m; n)$ and an isomorphism between $E$ and the pull-back via $\psi_m$ of the tautological bundle over $\text{Gr}_C(m; n)$.

**Additional reading**

Chapter 7: Metrics on vector bundles

A metric on \( \mathbb{R}^n \) is a positive definite symmetric bilinear form. This is to say that if \( g \) is a metric, then \( g(v, w) = g(w, v) \) and \( g(v, v) > 0 \) if \( v \neq 0 \). A metric on a real vector bundle \( E \to M \) is a section, \( g \), of \( \text{Sym}^2(E^*) \) that restricts to each fiber as symmetric, positive definite quadratic form.

Metrics always exist. To see this, let \( m = \dim(M) \) and let \( \{ \varphi_\alpha: U_\alpha \to \mathbb{R}^m \} \) denote a locally finite cover of \( M \) by coordinate charts such that \( \varphi_\alpha^{-1}(|x| < \frac{1}{2}) \) is also a cover of \( M \) and such that \( E|_{U_\alpha} \) over \( U_\alpha \) has an isomorphism, \( \nu_\alpha: E|_{U_\alpha} \to U_\alpha \times \mathbb{R}^n \). Use the function \( \chi \) from Appendix 2 in Chapter 1 to define \( \chi_\alpha: M \to [0, 1) \) so as to be zero on \( M \setminus U_\alpha \) and equal to \( \chi(|\varphi_\alpha|) \) on \( U_\alpha \). Let \( \delta \) denote a given metric on \( \mathbb{R}^n \), and view \( \delta \) as a metric on the product bundle \( U_\alpha \times \mathbb{R}^n \) for each \( \alpha \). A metric, \( g \), on \( E \) is \( g(v, w) = \sum_\alpha \chi_\alpha \delta(\nu_\alpha v, \nu_\alpha w) \).

Note in this regard that if \( g \) and \( g' \) are metrics on \( E \), then so is \( u g + u' g' \) for any functions \( u \) and \( v \) with both \( u \) and \( u' \) non-negative functions such that \( u + u' > 0 \) at each point.

A Hermitian metric on \( \mathbb{C}^n \) is a bilinear form, \( g \), on \( \mathbb{C}^n \) with the following properties: First, \( g(cu, cu') = c g(u, u') \) and \( g(cu, u') = \overline{c} g(u, u') \) for all \( u, u' \in \mathbb{C}^n \) and all \( c \in \mathbb{C} \). Second, \( g(u, u) > 0 \) for all \( u \in \mathbb{C}^n \neq 0 \). Finally, \( g(u, u') = g(u', u) \) for all \( u, u' \in \mathbb{C}^n \). A Hermitian metric on a complex vector bundle \( \pi: E \to M \) is a section of \( E^* \otimes E^* \) that restricts to each fiber as a Hermitian metric. Here is an equivalent definition: Let \( E_\mathbb{R} \) denote the underlying real bundle and let \( j \) denote its almost complex structure. A Hermitian on \( E \) is defined by a metric, \( g_\mathbb{R} \), on \( E_\mathbb{R} \) if \( g_\mathbb{R}(\cdot, j(\cdot)) = -g_\mathbb{R}(j(\cdot), \cdot) \). This condition asserts that \( j \) must define a skew-symmetric endomorphism on each fiber of \( E_\mathbb{R} \). To obtain \( g \) from \( g_\mathbb{R} \), view \( E \subset E_\mathbb{R} \) by writing \( E_\mathbb{R} \otimes_\mathbb{R} \mathbb{C} \) and then identifying \( E \) as the set of vectors of the form \( u = e - i(je) \) with \( e \in E_\mathbb{R} \). Then inner product \( g(u, u') \) is defined to be

\[
g(u, u') = g_\mathbb{R}(e, e') - i g_\mathbb{R}(e, je')
\]

Note that if \( c = c_1 + ic_2 \) is any given element in \( \mathbb{C} \), then \( cu = (c_1 e + c_2 je) - ij(c_1 u + c_2 ju) \). As a consequence, the preceding formula finds

\[
g(cu, u') = c_1 g_\mathbb{R}(e, e') + c_2 g(je, e') - i c_1 g_\mathbb{R}(e, je') - i c_2 g_\mathbb{R}(je, je')
\]

which is equal to \( \overline{c} g(u, u') \) by virtue of the fact that \( j \) is skew symmetric. A similar computation finds that \( g(u, cu') = c g(u, u') \) and also \( \overline{g(u, u')} = g(u', u) \).
Meanwhile, a given Hermitian metric, \( g \), on \( E \) defines the metric \( g_\mathbb{R} \) on \( E_\mathbb{R} \) by setting \( g_\mathbb{R}(e, e') \) to be the real part of \( g(u, u') \).

With only notational changes, the construction given above of a metric on a real bundle will give a Hermitian metric on a complex bundle. If \( g \) and \( g' \) are Hermitian metrics on a complex vector bundle, then so is \( u \cdot g + u' \cdot g' \) when \( u \) and \( u' \) are functions with \( u + u' > 0 \).

7a) Metrics and transition functions for real vector bundles

A real bundle \( E \to M \) of fiber dimension \( n \) with a metric can be used to construct a locally finite, open cover \( \mathcal{U} \) of \( M \) with two salient features: First, each \( U \in \mathcal{U} \) comes with an isomorphism from \( E|_U \) to \( U \times \mathbb{R}^n \). Second, the corresponding bundle transition functions map non-trivial intersections of sets from \( \mathcal{U} \) into \( O(n) \). To see why this is true, fix an locally finite cover, \( \mathcal{U} \), for \( E \) such that each \( U \in \mathcal{U} \) comes with an isomorphism from \( \varphi_U : E|_U \to U \times \mathbb{R}^n \). Use \( \varphi_U^{-1} \) to define a basis of sections of \( E|_U \) over \( U \), and then use the Gram-Schmidt procedure to construct a basis of sections of \( E|_U \) over \( U \) which is orthonormal at each point as defined by the fiber metric. Let \( \{ s_{\alpha j} \}_{1 \leq j \leq n} \) denote these sections. Define a new isomorphism \( E|_U \to U \times \mathbb{R}^n \) by declaring its inverse to be the map that sends any given point \( (x, (v_1, \ldots, v_n)) \in U \times \mathbb{R}^n \) to \( \sum_j v_j s_{\alpha j}|_x \). The corresponding bundle transition functions map any orthonormal vector in \( \mathbb{R}^n \) to an orthonormal vector. As a consequence, the all transition functions map to \( O(n) \).

The question arises as to whether a trivializing chart can be found whose corresponding transition functions map to \( SO(n) \).

**Definition:** A real vector bundle \( E \to M \) is said to be orientable if it has a trivializing cover such that the corresponding vector bundle transition functions on the overlaps have positive determinant.

If a vector bundle is orientable, and if it has a fiber metric, then its transition functions can be chosen to map to \( SO(n) \). The following lemma says something about when a bundle is orientable:

**Lemma:** A real vector bundle \( E \to M \) with some given fiber dimension \( n \geq 1 \) is orientable if and only if the real line bundle \( \text{det}(E) = \wedge^n E \) is isomorphic to the product bundle \( M \times \mathbb{R} \).

**Proof:** Let \( \mathcal{U} \) denote a locally finite cover of \( M \) such that each chart \( U \in \mathcal{U} \) comes with an isomorphism \( \varphi_U : E|_U \to U \times \mathbb{R}^n \). As noted above, these isomorphisms can be chosen so that the corresponding bundle transition functions map to \( O(n) \). The isomorphism \( \varphi_U \)
induces an isomorphism between, \( \psi_U: \det(E)_{|U} \to U \times \mathbb{R} \). If \( U \) and \( U' \) are two intersecting sets from \( \mathcal{U} \), then the corresponding vector bundle transition function for \( E \) is a map \( g_{U'U}: U \cap U' \to O(n) \). The corresponding bundle transition function for \( \det(E) \) is the map from \( U \cap U' \) to \( \{ \pm 1 \} \) given by \( \det(g_{U'U}) \). This understood, if \( E \) is orientable, one can take \( \mathcal{U} \) and the corresponding isomorphisms \( \{ \phi_U \}_{U \in \mathcal{U}} \) such that \( \det(g_{U'U}) = 1 \) for all pairs \( U, U' \in \mathcal{U} \). As a consequence, an isomorphism \( \psi: \det(E) \to M \times \mathbb{R} \) is defined by declaring \( \psi \) on \( \det(E)_{|U} \) for any given \( U \in \mathcal{U} \) to be \( \psi_U \).

To prove the converse, suppose that \( \eta: \det(E) \to M \times \mathbb{R} \) is an isomorphism. If \( U \in \mathcal{U} \), then \( \eta \circ \psi_U^{-1} \) can be written as \( (x, r) \to (x, \eta_U r) \) where \( \eta_U \in \{ \pm 1 \} \). The latter is such that \( \eta_U \det(g_{U'U}) \eta_U^{-1} = 1 \) on \( U \cap U' \) when \( U \) and \( U' \) are any two sets from \( \mathcal{U} \) with non-empty overlap. This understood, define a new isomorphism, \( \phi_U: E_{|U} \to U \times \mathbb{R}^n \) by \( \phi_U = \sigma_U \circ \phi_U \) where \( \sigma_U \) at any given point in \( U \) is a diagonal matrix with entries \( (\eta_U, 1, \ldots, 1) \) on the diagonal. It then follows from the preceding equation that the corresponding transition functions for the data \( \{ \phi_U \}_{U \in \mathcal{U}} \) map to \( SO(n) \). This being the case, the Gram-Schmid procedure can be applied to obtain transition.

What follows are some examples.

The Mobius bundle: This bundle over \( S^1 \) is defined in Section 1c. It is not orientable. To see why, suppose to the contrary. Identify \( \theta \) this bundle would admit a section that has no intersections with the zero section. By definition, a section is a map \( v: S^1 \to \mathbb{R}^2 \) whose coordinate entries \( (v_1, v_2) \) obey the condition \( \cos(\theta)v_1 + \sin(\theta)v_2 = v_1 \) and \( \sin(\theta)v_1 - \cos(\theta)v_2 = v_2 \) at each \( \theta \in S^1 \). This requires that \( v_1 \) and \( v_2 \) can be written as \( v_1 = r(\theta) \cos(\frac{\theta}{2}) \) and \( v_2 = r(\theta) \sin(\frac{\theta}{2}) \) with \( r: S^1 \to \mathbb{R} - 0 \). Such a pair is not continuous on the whole of \( S^1 \).

The tautological bundle over \( \mathbb{R}P^n \): This bundle is also not orientable. To see why, consider the map from \( S^1 \) into \( \mathbb{R}P^n \) that sends the angle coordinate \( \theta \in \mathbb{R}/(2\pi \mathbb{Z}) \) to the line with tangent \( (\cos(\frac{\theta}{2}), \sin(\frac{\theta}{2}), 0, \ldots, 0) \in \mathbb{R}^{n+1} \). The pull-back via this map of the tautological bundle is the Mobius bundle. Given that the Mobius bundle is not oriented, it follows that neither is the tautological bundle over \( \mathbb{R}P^n \).

A bundle with an almost complex structure: As is explained next, a bundle with an almost complex structure is orientable. Indeed, let \( \pi: E \to M \) denote such a bundle.
and let 2n denote its fiber dimension. Recall from Chapter 6b that there is a locally finite cover \( \mathcal{U} \) for \( M \) such that each \( U \in \mathcal{U} \) comes with an isomorphism to the product bundle, and such that all transition functions map to the group \( G_j \subset \text{Gl}(2n; \mathbb{R}) \). As noted at the end of Chapter 2d, matrices in \( G_j \) have positive determinant.

7b) Metrics and transition functions for complex vector bundles

Let \( \pi : E \to M \) denote now a complex vector bundle with fiber dimension \( n \) (over \( \mathbb{C} \)). Suppose that \( g \) is a Hermitian metric on \( E \). Let \( U \subset M \) denote an open set and suppose that \( (s_1, \ldots, s_n) \) is a basis for \( E|_U \) is said to be orthonormal if \( g(s_i, s_i) = 1 \) when \( i = j \) and zero otherwise. As the Gram-Schmidt algorithm works just as well over \( \mathbb{C} \) as over \( \mathbb{R} \), it follows that \( M \) has a locally finite cover such that \( E \)'s restriction to each set from the cover has an orthonormal basis. The following is a consequence: There is a locally finite cover, \( \mathcal{U} \), of \( M \) with two salient features: First, each \( U \in \mathcal{U} \) comes with a bundle isomorphism from \( E|_U \) to \( U \times \mathbb{C}^n \). Second, all transition functions map to the group \( U(n) \).

Recall from Chapter 6j the definition of the complex bundle \( \det(E) = \wedge^n E \). This bundle has fiber \( \mathbb{C} \).

**Lemma:** The bundle \( \det(E) \) is isomorphic to the product bundle \( E \times \mathbb{C} \) if and only if the following is true: There is a locally finite cover, \( \mathcal{U} \), of \( M \) such that each \( U \in \mathcal{U} \) comes with an isomorphism from \( E|_U \) to \( U \times \mathbb{C}^n \) and such that all transition functions map to \( SU(n) \).

The proof of this lemma very much like that of its \( \mathbb{R} \) version in Chapter 7a and so left to the reader.

7c) Metrics, algebra and maps

Suppose that \( E \) is a vector bundle over a given manifold \( M \) with fiber \( \mathbb{R}^n \). Assume that \( E \) has a given fiber metric. The latter endows bundles the various algebraically related bundles such as \( E^*, \otimes_m E, \otimes_m E, \wedge^m E \) and \( \text{Sym}^m E \) with fiber metrics.

Consider, for example, the case of \( E^* \). To obtain an orthonormal basis for the fiber of the latter at a given point in \( M \), choose an orthonormal basis \( \{e_1, \ldots, e_n\} \) for \( E \) at the point in question. A corresponding basis for \( E^* \) at the given point is given by the elements \( \{\nu_1, \ldots, \nu_n\} \) with the property that \( \nu_i(e_k) = 0 \) if \( i \neq k \) and 1 if \( i = k \). Note that the metric on \( E \) at this point is given by the bilinear form \( \sum_{1 \leq i, k \leq n} \nu_i \otimes \nu_k \). With these corresponding metrics on \( E \) and \( E^* \), the metric on \( E \) can be viewed as an isometric isomorphism between \( E \) and \( E^* \).
Consider for a second example the bundle $\wedge^m E$. To obtain an orthonormal basis for the fiber of the latter at a given point in $M$, fix an orthonormal basis $\{e_i, \ldots, e_n\}$ for $E$ at the point in question. The set $\{e_{i_1} \wedge \cdots \wedge e_{i_m}\}_{1 \leq i_1 < \cdots < i_m \leq n}$ defines an orthonormal basis for $\wedge^m E$. Note that the metric so defined does not depend on the orthonormal basis.

By the same token, if $E \to M$ is a complex vector bundle with fiber $\mathbb{C}^n$, a Hermitian metric on $E$ endows all of the analogous algebraically related bundles. Note that with the given metric on $E$ and the corresponding metric on $E^*$, the metric on $E$ can be viewed as an isometry $\tau: E \to E^*$ with the property that $\tau(ce) = \overline{c} \tau(v)$ for any given complex number $c$. Alternately, the metric can be viewed as giving $\mathbb{C}$-linear isometry between the complex conjugate bundle $\overline{E}$ and $E^*$.

Suppose next that $E$ and $E'$ are vector bundles over $M$, either both real or both complex. Suppose in addition that each has a fiber metric. Then $E \otimes E'$ and $E \oplus E'$ inherit natural fiber metrics. If only $E'$ has a fiber metric, then an injective vector bundle homomorphism $f: E \to E'$ can be used to give $E$ a fiber metric. Indeed, the metric on $E$ is defined as follows: Let $g'$ denote the metric on $E'$. Then the induced metric, $g$, on $E$ is such that $g(e, e') = g'(f(e), f(e'))$. For example, this induced metric construction gives a fiber metric to any subbundle of $M \times \mathbb{R}^n$ or $M \times \mathbb{C}^n$ if one agrees to give the latter their obvious base point independent fiber metrics. Given that any vector bundle is isomorphic to a subbundle of a product bundle, this metric induction procedure gives another proof of the assertion that all vector bundles have fiber metrics.

If $f: E \to E'$ is an injective vector bundle homomorphism then the quotient bundle $E'/E$ also inherits a metric from a metric on $E'$. This comes about by writing $E'$ as $f(E) \oplus E^\perp$ where $E^\perp \subset E$ is the subbundle of vectors that are orthogonal to the image of $f$. As the map from $E'$ to $E'/E$ restricts to $E^\perp$ as an isomorphism, a metric on $E^\perp$ gives a metric on $E'/E$. Meanwhile, $E^\perp$ has its fiber metric as a subbundle of $E'$.

The final remark in this chapter concerns pull-backs: Suppose that $M$ and $N$ are smooth manifolds and $\psi: M \to N$ is a smooth map. If $\pi: E \to N$ is a vector bundle with fiber metric, then the pull-back bundle $\psi^* E$ inherits what is perhaps an obvious fiber metric. To be specific, the latter is defined by viewing $\psi^* E \subset M \times E$ as the set of pairs of the form $(p, v)$ with $\psi(v) = \pi(p)$. The metric assigns to pairs $(p, v)$ and $(p, w)$ the inner-product of $v$ with $w$.

7d) Metrics on $TM$  

A metric on $TM$ is called a Riemannian metric. To see what such a metric looks like, suppose that $U \subset M$ is an open set with a diffeomorphism $\varphi: U \to \mathbb{R}^n$. Let $g$ denote a metric on $TM$. Then $(\varphi^{-1})^* g$ defines a metric on $T\mathbb{R}^n$. Now, $T\mathbb{R}^n$ has its standard Euclidean basis of sections, this the basis given by the vector fields $\{\frac{\partial}{\partial x_i}\}_{i=1,2,\ldots,n}$. This understood, the metric $(\varphi^{-1})^* g$ can be viewed as a map from $\mathbb{R}^n$ to the space of $n \times n$, symmetric, positive definite matrices, this the map with entries.
\[ g_{ik} = g((\varphi^{-1})_*, \frac{\partial}{\partial x_i}), (\varphi^{-1})_*, \frac{\partial}{\partial x_k}) \]

Thus, \((\varphi^{-1})^* g\) is the section of \(\text{Sym}^2(T^*\mathbb{R}^n)\) given by \(\sum_{1 \leq i \leq k \leq n} g_{ik} \, dx^i \otimes dx^k\).

A manifold with a given metric on its tangent bundle is said to be a \textit{Riemannian} manifold.

Here is an often used way to obtain a Riemannian metric on a given manifold \(M\): Suppose that \(X\) is a Riemannian manifold (for example \(\mathbb{R}^N\) with its Euclidean metric), and suppose that \(M\) is given as a submanifold of \(X\). Then the Riemannian metric on \(X\) induces a Riemannian metric on \(M\) because the tangent bundle to \(M\) sits in \(TX\) as a subbundle. Said explicitly, the inner product between two given vectors in \(TM\) is obtained by viewing the latter as vectors in \(TX|_M\) and taking their inner product using the metric on \(TX\).

**Additional Reading**

- \textit{Riemannian Geometry}, Sylvestre Gallot, Dominque Hulin and Jacques Lafontaine; Springer, 2004
Chapter 8: Geodesics

Let $M$ denote a smooth manifold. A metric on $TM$ can be used to define a notion of the distance between any two points in $M$ and the distance traveled along any given path in $M$. What follows in this chapter first explains how this is done. The subsequent parts of this chapter concern the distance minimizing paths.

8a) Riemannian metrics and distance

A Riemannian metric on $M$ gives a distance function on $M$ which is defined as follows: Let $\gamma$ denote a smooth map from an interval $I \subset \mathbb{R}$ into $M$. Let $t$ the Euclidean coordinate on $\mathbb{R}$. Define $\dot{\gamma}$ to be the section of $\gamma^*TM$ that is defined by $\gamma, \frac{d}{dt}$, this the push forward of the tangent vector to $\mathbb{R}$. This has norm $|\dot{\gamma}| = g(\dot{\gamma}, \dot{\gamma})^{1/2}$ where $g$ here denotes the given Riemannian metric. The norm of $\dot{\gamma}$ is said to be the speed of $\gamma$. The length of $\gamma$ is defined to be

$$\ell_\gamma = \int_I g(\dot{\gamma}, \dot{\gamma})^{1/2} \, dt$$

Note that this does not depend on the parametrization of $\gamma$, so it is intrinsic to the image of $\gamma$ in $M$. In the case $M = \mathbb{R}^n$ with its Euclidean metric, what is defined here is what is usually deemed to be the length of $\gamma$.

Let $p$ and $q$ denote points in $M$. A path from $p$ to $q$ is defined to be the image in $M$ of a continuous, piecewise smooth map from an interval in $\mathbb{R}$ to $M$ that maps the smaller endpoint of the interval to $p$ and the larger endpoint to $q$. It follows from what was just said that each path in $M$ has a well defined length. Define

$$\text{dist}(p, q) = \inf_\gamma \ell_\gamma$$

where the infimum is over all paths in $M$ from $p$ to $q$. Note that $\text{dist}(p, q) = \text{dist}(q, p)$ and that the triangle inequality is obeyed:

$$\text{dist}(p, q) \leq \text{dist}(p, x) + \text{dist}(x, q) \quad \text{for any } x \in M.$$

Moreover, $\text{dist}(p, q) > 0$ unless $p = q$. To prove this last claim, go to a coordinate chart that contains $p$. Let $(x_1, \ldots, x_n)$ denote the coordinates in this chart, and view $p$ as a point, $z \in \mathbb{R}^n$. Since $q \neq p$, there is a ball in $\mathbb{R}^n$ of some radius $r > 0$ that has $z$ in its center, and does not contain the image of $q$ via the coordinate chart. (This is certainly the case if $q$ is not in the coordinate chart at all.) Then any path from $p$ to $q$ has image via the coordinate chart that enter this ball on its boundary and ends at the origin. The Euclidean
distance of this image is no less than r. Meanwhile, because g is a metric, when written in the coordinates via the coordinate chart, g appears as a symmetric, bilinear form with entries \( g_{ij} \). As it is positive definite, there is a constant, \( c_0 \), such that \( \sum_{i,j} g_{ij}(x) v^i v^j \geq c_0 |v|^2 \) for any vector \( v \in \mathbb{R}^n \) and any point \( x \) in the radius \( r \) ball about \( z \). This implies that the length of any path from \( q \) to \( p \) is at least \( c_0^{-1/2} r > 0 \).

The function \( \text{dist}(\cdot, \cdot) \) can be used to define a topology on \( M \) whereby the open neighborhoods of a given point \( p \in M \) are generated by sets of the form \( \{ O_p, \varepsilon \} \) \( \varepsilon > 0 \) where the set \( O_p, \varepsilon = \{ q \in M: \text{dist}(p, q) < \varepsilon \} \). The argument given in the preceding paragraph (with the fact that \( \sum_{i,j} g_{ij}(x) v^i v^j \leq c_0 |v|^2 \)) implies that this metric topology on \( M \) is identical to the given topology.

8b) Length minimizing curves

The following is the fundamental theorem that underlies much of the subject of Riemannian geometry. This theorem invokes two standard items of notation: First, repeated indices come with an implicit sum. For example, \( g_{ij} v^i v^j \) mean \( \sum_{1 \leq i \leq j \leq n} g_{ij} v^i v^j \). Second, \( g^{ij} \) denotes the components of the inverse, \( g^{-1} \), to a matrix \( g \) with components \( g_{ij} \).

**Theorem:** Suppose that \( M \) is a compact, manifold and that \( g \) is a Riemannian metric on \( M \). Fix any points \( p \) and \( q \) in \( M \)

- There is a smooth curve from \( p \) to \( q \) whose length is the distance between \( p \) and \( q \).
- Any length minimizing curve is an embedded, 1-dimensional submanifold that can be reparametrized to have constant speed; thus \( g(\dot{\gamma}, \dot{\gamma}) \).
- A length minimizing, constant speed curve is characterized as follows: Let \( U \subset M \) denote an open set with a diffeomorphism \( \Phi: U \to \mathbb{R}^n \). Introduce the Euclidean coordinates \( (x^1, \ldots, x^n) \) for \( \mathbb{R}^n \) and let \( g_{ij} \) denote the components of the metric in these coordinates. This is to say that the metric on \( U \) is given by \( g|_U = \Phi^*(g_{ij} \, dx^i \otimes dx^j) \).

Denote the coordinates of the \( \Phi \)-image of a length minimizing curve by \( \gamma = (\gamma^1, \ldots, \gamma^n) \). Then the latter curve in \( \mathbb{R}^n \) obeys the equation:

\[
\ddot{\gamma}^i + \Gamma^i_{km} \dot{\gamma}^k \dot{\gamma}^m = 0
\]

where \( \Gamma^i_{km} = \frac{1}{2} g^{ir} (\partial_j g_{rk} + \partial_r g_{jk} - \partial_k g_{jr}) \). Here, \( \dot{\gamma}^i = \frac{d}{dt} \gamma^i \) and \( \ddot{\gamma}^i = \frac{d^2}{dt^2} \gamma^i \). (Here, as above, repeated indices are implicitly summed.)

- Any curve that obeys this equation in coordinate charts is locally length minimizing. This is to say that there exists \( c_0 > 1 \) such that when \( p \) and \( q \) are two points on such a curve with \( \text{dist}(p, q) \leq c_0^{-1} \), then there is a segment of the curve with one endpoint \( p \) and the other \( q \) whose length is \( \text{dist}(p, q) \).
This theorem is called the \textit{geodesic theorem} in what follows. A proof is given at the end of the next chapter. The next part of this chapter supplies an existence theorem for geodesics, and the remaining parts consider various examples where the geodesics can be described in very explicit terms. What follows directly are some additional remarks about the statement of the theorem.

The equation depicted above in the theorem is called the \textit{geodesic equation}. This equation refers to a particular set of coordinates. Even so, the notion that a map from an interval to \( M \) obeys this equation is coordinate independent. This is to say the following: Let \( U \subset M \) denote an open set with a diffeomorphism, \( \varphi: U \to \mathbb{R}^n \). Now let \( I \subset \mathbb{R} \) denote an interval and let \( \gamma_M: I \to U \) denote a smooth map such that \( \gamma = \varphi \circ \gamma_M \) obeys the geodesic equation. Then so does \( \gamma' = \varphi' \circ \gamma_M \) if \( \varphi': U \to \mathbb{R}^n \) is any other diffeomorphism. You can use the chain rule to prove this assertion if you make sure to account for the fact that the metric coefficients \( g_{ij} \) depend on the choice of coordinates. To say more about this, write the coordinates of \( \psi = \varphi \circ \varphi'^{-1} \) as \( (\psi^1, \ldots, \psi^n) \). Let \( g_{ij} \) and \( g'_{ij} \) denote the components of the metric when written using the coordinates \( \varphi \) and \( \varphi' \). Then

\[
g'_{ij} = \left( \frac{\partial \psi^k}{\partial x_i} \frac{\partial \psi^m}{\partial x_j} g_{km}(\psi(\cdot))) \right)_{x^k}.
\]

Meanwhile the respective components of \( \gamma \) are given by \( \gamma'^{-1} = (\psi^1)'(\gamma) \). Granted all of this, then the argument for coordinate independence amounts to a tedious exercise with the chain rule. A \textit{geodesic} is a map \( \gamma \), from an interval \( I \subset \mathbb{R} \) into \( M \) that obeys the geodesic equation in one, and thus all coordinate charts that contain points in the image of \( \gamma \).

Here is a last bit of commonly used terminology: The functions

\[
\{ \Gamma^j_{km} = \frac{1}{2} g^{ip} (\partial_m g_{pk} + \partial_k g_{pm} - \partial_p g_{km}) \}_{i,j,k,m=1}^{n}
\]

that appear in the theorem’s statement are called the \textit{Christoffel symbols}.

\begin{itemize}
\item \textbf{8c) The existence of geodesics}
\end{itemize}

The proof given momentarily for the existence of geodesics invokes a fundamental theorem about differential equations.

\textbf{Theorem:} Let \( m \) denote a positive integer and let \( v: \mathbb{R}^m \to \mathbb{R}^m \) denote a given smooth map. Fix \( y_0 \in \mathbb{R}^m \) and there exists an interval \( I \subset \mathbb{R} \) centered on 0, a ball \( B \subset \mathbb{R}^m \) about the point \( y_0 \), and a smooth map \( \tilde{z}: I \times B \to \mathbb{R}^m \) with the following property: If \( y \in B \), then the map \( t \to z(t) = \tilde{z}(t, y) \) obeys
\[
\frac{d}{dt} z = v(z).
\]

with the initial condition \( z_{t=0} = y \). Moreover, there is only one solution of this equation that equals \( y \) at \( t = 0 \).

This theorem is called the vector field theorem in what follows. The name refers to the fact that \( v \) can be viewed as a vector field on \( \mathbb{R}^m \). When viewed in this light, the image of the map \( t \rightarrow z(t) \) is said to be an integral curve of \( v \).

This vector field theorem is proved in the Appendix to this chapter.

The first proposition below states a basic existence and uniqueness theorem for the geodesic equation.

**Proposition 1:** Let \( M \) be a smooth manifold and let \( g \) denote a Riemannian metric on \( M \). Let \( p \in M \) and let \( v \in T_{p}M \). There exists \( \varepsilon > 0 \) and a unique map from the interval \((-\varepsilon, \varepsilon)\) to \( M \) that obeys the geodesic equation, sends 0 to \( p \), and whose differential at 0 sends the vector \( \frac{d}{dt} \) to \( v \).

**Proof:** Fix a coordinate chart centered at \( p \) with coordinates \((x^1, \ldots, x^n)\) such that \( p \) is the origin. Let \( I \subset \mathbb{R} \) denote an open interval centered at 0, and view the geodesic equation as an equation for a map \( z: I \rightarrow \mathbb{R}^n \times \mathbb{R}^n \) that obeys the a certain system of differential equations that involve only first derivatives. To obtain this system of equations, write the coordinates of \( z \) as \((\gamma, u)\). Then

\[
\frac{d}{dt} (\gamma, u) = (u, -\Gamma^i_{jk}(\gamma) u^j u^k)
\]

It is left as an exercise to verify that \((\gamma, u)\) obeys the latter equation if and only if \( \gamma \) obeys the geodesic equation. This last equation has the schematic form \( \frac{d}{dt} z = v(z) \) for a function \( z: I \rightarrow \mathbb{R}^m \) where in this case \( m = 2n \). This understood, the assertions of Proposition 1 follow directly from the vector field theorem.

The upcoming Proposition 2 states a stronger version of Proposition 1. As with Proposition 1, it is also a direct consequence of the preceding theorem. To set the stage, introduce \( \pi: TM \rightarrow M \) to denote the vector bundle projection map. Suppose that \( O \subset TM \) is an open set. Proposition 1 above assigns the following data to each point \( v \in O \): First, a neighborhood of \( \pi(v) \), and second, a unique geodesic in this neighborhood that contains \( \pi(v) \) and whose tangent vector at \( \pi(v) \) is \( v \). Thus, the points in \( O \) parametrize a family of geodesics in \( M \). Among other things, the next proposition says that this family is smoothly parametrized by the points in \( O \).
Proposition 2: Let $M$ be a smooth manifold and let $g$ denote a Riemannian metric on $M$. Fix a point in $TM$. Then there exists an open neighborhood $O \subset TM$ of this point, a positive number $\varepsilon > 0$ and a smooth map $\gamma_0: (-\varepsilon, \varepsilon) \times O \to M$ with the following property: If $v \in O$, then the map $\gamma_0(\cdot, v): (-\varepsilon, \varepsilon) \to M$ is a geodesic with $\gamma_0(0, v) = \pi(v)$ and with $(\gamma_0(\cdot, v), \frac{\partial}{\partial t})_{|t=0}$ equal to the vector $v$.

Proof: Fix a coordinate chart for a neighborhood of $\pi(x)$ to identify the latter with a neighborhood of the origin in $\mathbb{R}^n$ and to identify $TM$ on this neighborhood with $\mathbb{R}^n \times \mathbb{R}^n$.

As noted at the outset of Chapter 8f, the geodesic equation can be viewed as an equation for a map, $z = (\gamma, u)$, from an interval in $\mathbb{R}$ about the origin to a neighborhood of $\{0\} \times \mathbb{R}^n$ in $\mathbb{R}^n \times \mathbb{R}^n$ that has the form

$$\frac{d}{dt} (\gamma, u) = (u, -\Gamma^i_{jk}(\gamma) u^j u^k).$$

As this equation has the form $\frac{d}{dt} z = v(z)$, an appeal to the theorem above proves Proposition 2.

8d) First examples

What follows gives some examples of metrics with their corresponding geodesics.

The standard metric on $\mathbb{R}^n$: This is the Euclidean metric $g = dx^i \otimes dx^i$. As you might expect, the geodesics are the straight lines.

The round metric on $S^n$: The round metric is that induced by its embedding in $\mathbb{R}^{n+1}$ as the set $\{x \in \mathbb{R}^{n+1}: |x| = 1\}$. This embedding identifies $TS^n$ as the set of pairs $(x, v) \in \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ with $|x| = 1$ and with $x \cdot v = 0$. Here, $x \cdot v = \sum_{1 \leq j \leq n+1} x_j v_j$. The round metric is such that the inner product between elements $(x, v)$ and $(x, w)$ in $TS^n$ is $v \cdot w$.

Suppose $\gamma: \mathbb{R} \to S^n \subset \mathbb{R}^{n+1}$ is a smooth curve. The geodesic equation when written in coordinates $\gamma(t) = (x^1(t), \ldots, x^{n+1}(t))$ asserts that

$$\ddot{x}^i + x^i \dot{x}^j \dot{x}_j = 0.$$ 

To see that the solutions to this equation are the geodesics, view $\mathbb{R}^{n+1}$ as $\mathbb{R}^n \times \mathbb{R}$ and introduce the coordinate embedding $y \to (y, (1 - |y|^2)^{1/2})$ of the ball of radius 1 in $\mathbb{R}^n$ into $S^n \subset \mathbb{R}^n \times \mathbb{R}$. The round metric on $S^n$ pulls back via this embedding as the metric on radius 1 ball in $\mathbb{R}^n$ with entries $g_{ij} = \delta_{ij} + y_i y_j (1 - |y|^2)^{1/2}$. Here, $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. Thus, $g_{ij} = \delta_{ij} + y_i y_j + O(|y|^3)$. As a consequence, $\Gamma^i_{jk} = \delta_{jk} y_i + O(|y|^3)$. The
equation asserted by the theorem is \( \dot{y}_j + y_j |\dot{y}|^2 + O(|y|^2) = 0 \). This agrees with what is written above to leading order in \( y \). Since the metric and the sphere are invariant under rotations of \( S^n \), as is the equation for \( x \) above, this verifies the equation at all points.

Note that this equation implies that \( t \to x(t) \) lies in a plane. Indeed, if \( |\dot{x}| = 1 \) at time \( t = 0 \), then this is true at all times \( t \). Then the equation above finds that

\[
x(t) = a^j \cos(t) + b^j \sin(t)
\]

where \( |a|^2 + |b|^2 = 1 \) and \( a^ja_j = 0 \). In particular, the geodesics are the curves in \( S^n \) that are obtained by intersecting \( S^n \) with a plane through the origin in \( \mathbb{R}^{n+1} \).

The induced metric on a codimension 1 submanifold in \( \mathbb{R}^{n+1} \): This last example where the geodesic equation for the sphere is written as an equation for a curve in \( \mathbb{R}^{n+1} \) has an analog of sorts for any \( n \)-dimensional submanifold in \( \mathbb{R}^{n+1} \). Let \( \Sigma \subset \mathbb{R}^{n+1} \) denote the submanifold in question. The relevant metric is defined as follows: Let \( x \to n(x) \in \mathbb{R}^{n+1} \) denote a smooth assignment to a point \( x \in \Sigma \) of a unit length unit normal vector to \( \Sigma \) at \( x \). View \( T\Sigma \subset \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \) as the set of pairs of the form \((x, v)\) with \( x \in \Sigma \) and with \( v \cdot n(x) = 0 \). The metric inner product between vectors \((x, v)\) and \((x, w)\) is declared to be \( v \cdot w \). To say something about the geodesics, write a curve in \( \Sigma \) as a map from an interval in \( \mathbb{R} \) to \( \mathbb{R}^{n+1} \) as a map \( t \to (x^1(t), \ldots, x^{n+1}(t)) \) whose image lies in \( \Sigma \). The geodesic equation is asks that the curve \( t \to x(t) \) in \( \mathbb{R}^{n+1} \) obey

\[
\dot{x}^j + n^j_k (\partial_k n_i)_x \dot{x}^k \dot{x}^i = 0.
\]

The proof that solutions to this equation give the geodesics in \( \Sigma \) is left to the reader. As a guide, remark that a proof can be had by choosing a very convenient coordinate chart about any given point, then writing this equation in the coordinate chart, and comparing it to the geodesic equation.

The hyperbolic metric on \( \mathbb{R}^n \): Let \( M \subset \mathbb{R}^{n+1} \) denote the branch of the hyperbola where \( |x_{n+1}|^2 - \sum_{1 \leq k \leq n} |x_k|^2 = 1 \) and \( x_{n+1} > 0 \). This manifold \( M \) is diffeomorphic to \( \mathbb{R}^n \) with diffeomorphism given by the map that sends \( y \in \mathbb{R}^n \) to

\[
\psi(y) = (y, (1 + |y|^2) \in \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}.
\]

In any event, view \( TM \subset TR^{n+1} = \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \) as the space of pairs \((x, v)\) with \( x \in M \) and
\[ v_{n+1}x_{n+1} - \sum_{1 \leq k \leq n} v_k x_k = 0. \]

Granted this identification of TM, define a Riemannian metric on TM is by the rule that makes the inner product between pairs \((x, v)\) and \((x, w)\) equal to

\[ \sum_{1 \leq k \leq n} v_k w_k - v_{n+1}w_{n+1}. \]

This metric is positive definite, as can be seen by writing \(v_{n+1} = x_{n+1} - \sum_{1 \leq k \leq n} v_k x_k\) and using the fact that \(x_{n+1} = (1 + \sum_{1 \leq k \leq n} x_k^2)^{1/2}\). I use \(g_H\) in what follows to denote this metric.

The metric \(g_H\) just defined is called the hyperbolic metric. To justify this name, remark that the geodesics are as follows: Fix any point \((x, v) \in TM \subset \mathbb{R}^{n+1} \times \mathbb{R}^{n+1}\) with \(\sum_{1 \leq k \leq n} v_k^2 - v_{n+1}^2 = 1\) and \(v \in TM\). Then the curve \(t \to x(t) = \cosh(t)x + \sinh(t)v\) is a geodesic curve in \(M\). Note that these curves are the intersections between \(M\) and the planes through the origin in \(\mathbb{R}^{n+1}\).

The proof that the curves just described are the geodesics can be had by using the coordinate chart given by the map \(\psi\) above. Note in this regard that it is enough to focus on the case where \(x = (0, 1) \in \mathbb{R}^n \times \mathbb{R}\) by invoking symmetry arguments that are much like those used in the discussion above for the round metric on \(S^n\). To say more about this, I need to introduce the Lorentz group, \(SO(n, 1)\). This is the subgroup of \(Gl(n+1; \mathbb{R})\) that preserves the indefinite quadratic form \(\eta(x, x) = |x_{n+1}|^2 - \sum_{1 \leq k \leq n} |x_k|^2\). This is to say that \(SO(n, 1)\) consists of the matrices \(m \in Gl(n+1; \mathbb{R})\) for which \(\eta(mx, mx) = \eta(x, x)\) for all \(x \in \mathbb{R}^{n+1}\). As it turns out, this is a Lie group whose dimension is \(\frac{1}{2} n(n+1)\).

In any event, \(mx \in M\) if \(x \in M\) and \(m \in SO(n, 1)\). This is to say that the linear map \(\mu_m: \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}\) given by \(x \to mx\) restricts to define a diffeomorphism of \(M\). Moreover, this diffeomorphism is such that \(\mu_m^* g_H = g_H\) and so the length of a curve and the length of its \(\mu_m\) image are the same. As a consequence, \(\mu_m\) maps geodesics to geodesics. As explained momentarily, given \(x \in M\) there exists \(m \in SO(n, 1)\) such that \(mx\) is the point \((0, 1) \in \mathbb{R}^n \times \mathbb{R}\). Thus, one need only identify the geodesics through the latter point.

What follows is a matrix \(m \in SO(n, 1)\) that maps a given point \(x \in M\) to the point \((0, 1)\): Write \(x = (\hat{x}, x_{n+1}) \in \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}\). Let \(o \in SO(n)\) denote a matrix which is such that \(o \hat{x} = (|\hat{x}|, 0, \ldots, 0)\). Let \(o \in SO(n, 1)\) denote the matrix that acts on any given \(v = (\hat{v}, v_{n+1}) \in \mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}\) to give \(ov = (o \hat{v}, v_{n+1})\). Now set \(p \in SO(n+1)\) to be the matrix that acts on \((v_1, v_2, \ldots, v_{n+1}) \in \mathbb{R}^{n+1}\) to give

\[ (x_{n+1}v_1 - |\hat{x}| v_{n+1}, v_2, \ldots, v_n, x_{n+1}v_{n+1} + |\hat{x}| v_1) . \]

The matrix \(m = po\) does the job.
8e) Geodesics on SO(n)

To define the relevant metric, first define an inner product on the space \( \mathbb{M}(n; \mathbb{R}) \) of \( n \times n \) matrices by \( \langle a, a' \rangle = \text{tr}(a^T a') \). This is a Euclidean metric. Fix an orthonormal basis. As I will be talking momentarily about SO(n) also, it is convenient to fix one so that the each basis element is either a symmetric matrix or an anti-symmetric one. (Note that the symmetric matrices are orthogonal to the anti-symmetric ones.) Let \( \{a_i\} \) denote the basis elements. Here, \( j \) runs from 1 to \( n^2 \). Assume that the first \( n(n-1)/2 \) basis elements are anti-symmetric.

For each \( j \in \{1, \ldots, n^2\} \), let \( \omega_j \) denote the 1-form on \( \text{Gl}(n; \mathbb{R}) \) given by

\[
\omega_j|_m = \text{tr}(a_j m^{-1} dm)
\]

As noted in Chapter 5c, these forms are left-invariant 1-forms. Since the collection spans \( T^*\text{Gl}(n; \mathbb{R}) \) at the identity, so these forms span \( T\text{GL}(n; \mathbb{R}) \) at any give point. Thus, they give an isomorphism of vector bundles \( T\text{Gl}(n; \mathbb{R}) \to \text{Gl}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R}) \) by the rule that associates to a pair \( m \in \text{Gl}(n; \mathbb{R}) \) and \( v \in T\text{Gl}(n; \mathbb{R})|_m \) the pair \( (m, \sum_i a_i \langle \omega_i|_m, v \rangle) \) in the product \( \text{Gl}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R}) \). This basis \( \{\omega_j\} \) gives the metric

\[
g = \sum_{1 \leq i \leq n^2} \omega^i \otimes \omega^i
\]

on \( \text{Gl}(n; \mathbb{R}) \). Note that this metric does not depend on the chosen orthonormal basis since any one basis can be obtained from any other by the action of the group \( \text{SO}(n^2) \) on the Euclidean space \( \mathbb{M}(n; \mathbb{R}) \).

The metric \( g \) is invariant under the action on \( \text{Gl}(n; \mathbb{R}) \) of left translation by elements in \( \text{Gl}(n; \mathbb{R}) \). This means the following: If \( u \in \text{Gl}(n; \mathbb{R}) \), then the diffeomorphism \( l_u: \text{Gl}(n; \mathbb{R}) \to \text{Gl}(n; \mathbb{R}) \) that sends \( m \) to \( um \) is such that \( l_u^* g = g \). Such is the case because \( l_u^* \omega^i = \omega^i \). (In general, metrics on a manifold \( M \) are sections of \( \text{Sym}^2(T^*M) \) and so pull-back under diffeomorphisms.)

As noted in Chapter 5e, the form \( \omega^i \) restricts as zero on \( T\text{SO}(n) \subset T\text{GL}(n; \mathbb{R})|_{\text{SO}(n)} \) when \( a_i = a_i^T \). Meanwhile, the collection \( \{\omega^i: a_i^T = -a_i\} \) span \( \text{SO}(n) \) along \( \text{SO}(n) \). This understood, introduce again \( A(n; \mathbb{R}) \subset \mathbb{M}(n; \mathbb{R}) \) to denote the vector space of anti-symmetric matrices. The set \( \{\omega^i\}_{1 \leq i \leq n(n-1)/2} \) restrict to \( \text{SO}(n) \) to give a bundle isomorphism \( T\text{SO}(n) \to \text{SO}(n) \times A(n; \mathbb{R}) \). It follows as a consequence that the restriction to \( T\text{SO}(n) \subset T\text{GL}(n; \mathbb{R})|_{\text{SO}(n)} \) of the metric defined above on \( T\text{GL}(n; \mathbb{R}) \) defines the metric.
\[
g = \sum_{i=\text{odd}(n-1)/2}^n \omega^i \otimes \omega^i
\]
on TSO(n). This metric is left invariant with respect to the left multiplication map on SO(n) because it is left invariant on Gl(n; \mathbb{R}) and, in addition, because left multiplication on Gl(n; \mathbb{R}) by an element in SO(n) maps SO(n) to itself.

As an aside, note that this metric \( g = \sum_{i=\text{odd}(n-1)/2}^n \omega^i \otimes \omega^i \) on TSO(n) is also invariant with respect to the right multiplication map on SO(n). This is to say that it pulls back to itself under the right multiplication map \( r_u \) on SO(n) that sends \( m \rightarrow mu^{-1} \) with \( u \) any given element in SO(n). To see why this is so, remark first that

\[
r_u^* \omega^i = \text{tr}(u^*a_i u m^{-1} dm).
\]

Note next that the map \( a \rightarrow u^*a u \) maps \( A(n; \mathbb{R}) \) to itself when \( u \in \text{SO}(n) \). Moreover, \( \text{tr}((u^*a u)^T (u^*a u)) = \text{tr}(a^T a) \) for any \( a \in \mathbb{M}(n; \mathbb{R}) \). Thus, the map \( a \rightarrow u^*a u \) acts as an isometry on \( A(n; \mathbb{R}) \) when \( u \in \text{SO}(n) \). This understood, it follows that

\[
u^{-1}a_i u = \sum_{i=\text{odd}(n-1)/2}^n S^i_j a_j
\]
when \( a_i \in A(n; \mathbb{R}) \) where \( S^i_j \) are the entries of a \( (n(n-1)/2) \times (n(n-1)/2) \) orthogonal matrix, \( S \). This implies that

\[
r_u^* g = \sum_i \sum_{j,k} S^i_j S^k_j \omega^i \otimes \omega^k
\]
where the sums range from 1 to \( n(n-1)/2 \). Since \( S \) is orthogonal, the sum on the right over \( i \) of \( S_i^i S^k_k \) is \( \delta^i_k \) where \( \delta^i_k \) here denotes the \( (n(n-1)/2) \times (n(n-1)/2) \) identity matrix. This is to say that \( r_u^* g = \sum_{i=\text{odd}(n-1)/2}^n \delta^i_i \omega^i \otimes \omega^i = g \). The metric \( g \) is said to be bi-invariant because it is both left and right invariant. As it turns out, any bi-invariant metric on SO(n) has the form \( cg \) with \( c > 0 \) a constant. Thus, the bi-invariant metrics are rather special.

Return now to the issue of the geodesics for this metric \( g \) for TSO(n). The proposition that follows describes these geodesics. The statement of the theorem refers to the exponential map from \( \mathbb{M}(n; \mathbb{R}) \) to Gl(n; \mathbb{R}) that sends a given matrix \( a \in \mathbb{M}(n; \mathbb{R}) \) to \( e^a = 1 + a + \frac{1}{2} a^2 + \frac{1}{3} a^3 + \cdots \). As noted in Chapter 5e, this map sends \( A(n; \mathbb{R}) \) into SO(n) and it restricts to some small radius ball about the origin in \( A(n; \mathbb{R}) \) as a diffeomorphism onto a neighborhood in SO(n) of \( 1 \).

**Proposition:** Identify TSO(n) with \( SO(n) \times A(n; \mathbb{R}) \) as just described. With the metric \( g \) as given above, any solution to the geodesic equation on SO(n) and has the form \( t \rightarrow \)
where \( m \in \text{SO}(n) \) and \( a \in \mathcal{A}(n; \mathbb{R}) \). Conversely, any such map from \( \mathbb{R} \) into \( \text{SO}(n) \) parametrizes a geodesic.

The proof of this proposition (and of others to come) make use of an important observation about geodesics for induced metrics. To set the stage, suppose that \( X \) is a given Riemannian manifold and that \( M \) is a submanifold of \( X \). Give \( M \) the induced metric. A geodesic in \( M \) need not be a geodesic in \( X \). Witness, for example, the case of the round sphere in \( \mathbb{R}^n \) as described in the previous part of this chapter. By contrast, a geodesic in \( X \) that lies entirely in \( M \) is apriori also a geodesic for the induced metric on \( M \). This follows from the fact that geodesics are locally length minimizing.

With the preceding understood, the proposition is proved by demonstrating that any path of the form \( m \to me^a \) for \( m \in \text{Gl}(n; \mathbb{R}) \) and \( a \in \mathcal{A}(n; \mathbb{R}) \) is apriori a geodesic in \( \text{Gl}(n; \mathbb{R}) \). Note that this last result implies the assertion made by the proposition for the following reason: If every path through a given matrix \( m \in \text{SO}(n) \) of the form \( t \to me^a \) with \( a \in \mathcal{A}(n; \mathbb{R}) \) is a geodesic, then Proposition 1 in the previous part of this chapter guarantees that such paths account for all of the geodesics that pass through \( m \).

By way of notation, a submanifold \( M \) in a Riemannian manifold \( X \) is said to be totally geodesic when all geodesics in \( M \) as defined by the induced metric from \( X \) are also geodesics for the metric on \( X \). The proof given here of the proposition demonstrates that \( \text{SO}(n) \) is a totally geodesic submanifold in \( \text{Gl}(n; \mathbb{R}) \).

Geodesics on \( \text{Gl}(n; \mathbb{R}) \): The proposition that follows describes certain of the geodesics on \( \text{Gl}(n; \mathbb{R}) \).

**Proposition:** Identify \( T\text{Gl}(n; \mathbb{R}) \) with \( \text{Gl}(n; \mathbb{R}) \times \mathbb{M}(n; \mathbb{R}) \) as described above. With the metric \( g \) as above, paths of the form \( t \to me^a \) with \( m \in \text{Gl}(n; \mathbb{R}) \) and \( a \in \mathbb{M}(n; \mathbb{R}) \) obeying \( a a^T - a^T a = 0 \) are geodesics. In particular, paths of the form \( t \to me^a \) are geodesics in \( \text{Gl}(n; \mathbb{R}) \) if \( a = \pm a^T \).

By way of notation, if \( a \) and \( a' \) are any two \( n \times n \) matrices, the matrix \( a a' - a' a \) is said to be their commutator and is written \([a, a']\).

**Proof:** The metric is left-invariant and this has the following consequence: If \( I \) is an interval containing \( 0 \) and \( \gamma: I \to \text{Gl}(n; \mathbb{R}) \) is a solution to the geodesic with \( \gamma(0) = m \), then the map \( t \to \bigm|_m \gamma(t) \) is a solution to the geodesic equation that hits the identity matrix \( t \) at \( t = 0 \). This understood, no generality is lost by showing that a solution, \( \gamma \), to the geodesic equation with \( \gamma(0) = t \) has the form \( t \to e^a \) if \([a, a^T] = 0 \).
To see this, remember that the map from $\mathbb{M}(n; \mathbb{R}) \to \text{Gl}(n; \mathbb{R})$ that sends $a$ to $e^a$ gives a diffeomorphism of some neighborhood of $t$ in $\text{Gl}(n; \mathbb{R})$ with a ball about the origin in $\mathbb{M}(n; \mathbb{R})$. This understood, it is enough to prove that any line through the origin in the vector space $\mathbb{M}(n; \mathbb{R})$ of the form $t \to ta$ with $[a, a^T] = 0$ is a geodesic for the metric that is obtained by pulling back the metric $g$ on $\text{Gl}(n; \mathbb{R})$. As a line obeys $\gamma^i = 0$, so it is sufficient to prove that the term $\Gamma^i_{jk} \gamma^j \gamma^k$ is also zero when $\gamma$ is the line $t \to ta$ with $[a, a^T] = 0$.

To show this, consider first the pull-back of $g$ via the map $a \to e^a$. Recall from Chapter 5e that the pull-back of the 1-form $\omega_q$ for $q \in \mathbb{M}(n; \mathbb{R})$ via the exponential map is the 1-form

$$\nu_q = \int_0^1 \text{tr}(e^{s a} q e^{-s a} da) ds$$

It follows as a consequence that the metric $g$ appears with respect to this coordinate chart as

$$g|_a = \int_0^1 \int_0^1 \text{tr}(e^{-s a} a^T e^{s a} + e^{s a} a e^{-s a} da e^{-s a}) ds ds'$$

This is to say that

$$g_{ij}|_a = \int_0^1 \int_0^1 \text{tr}(e^{-s a} a_i e^{s a} a_j e^{-s a}) ds ds'.$$

Granted the preceding, what follows explains why $\Gamma^i_{jk} \gamma^j \gamma^k$ is zero when $t \to \gamma(t) = ta$ with $a$ obeying $[a, a^T] = 0$. To show this, digress momentarily and fix elements $q_1, q_2 \in \mathbb{M}(n, \mathbb{R})$. Consider the function on $\mathbb{M}(n; \mathbb{R})$ given by

$$a \to \text{tr}(e^{-s a} q_1 e^{s a} q_2 e^{-s a}).$$

Note that the integrand above that defines the pull-back at $a$ of the metric $g$ involves this sort of expression. The directional derivative of the function just described in the direction of $q \in \text{Gl}(n; \mathbb{R})$ is given by

$$s \int_0^1 d\tau \text{tr}((e^{(1-\tau)s a} q e^{-s a} q_1 e^{s a} + e^{-s a} q_1 e^{(1-\tau)s a} q e^{-s a}) e^{s a} q_2 e^{-s a}).$$
This last observation is relevant because the evaluation \( \Gamma^{ij}_{jk} \) requires evaluating expressions of this sort when two of \( q_1, q_2 \) or \( q \) have the form \( \sum_i \gamma_i a_i \). Note in particular, that this last sum is 0 when \( \gamma \) is the line \( t \to ta \).

Granted the preceding consider the expression above first when \( q \) and either of \( q_1, q_2 \) are multiples of \( a \). If this is the case, then the cyclic property of the trace implies that the integrand above is zero. Meanwhile, if \( q_1 \) and \( q_2 \) are multiples of \( a \), then what is written above is proportional to

\[
\int_0^1 dt \ tr((s e^{t(1-t)a^T} q^T e^{t(1-t)a^T} + s' e^{t(1-t)a^T} q e^{t(1-t)a})(a, a^T))
\]

In particular, the integral of this last expression with respect to \( s \) and \( s' \) is zero since the result can be written as \( tr(b [a, a^T]) \). Thus \( \Gamma^{ij}_{jk} \gamma^k \) = 0 for the line \( t \to ta \) when \([a, a^T] = 0\).

8f) Geodesics on U(n) and SU(n)

The story for these groups is very much the \( \mathbb{C} \) analog of what was done just now for Gl(n; \( \mathbb{R} \)) and SO(n). To start, define a Hermitian inner product on the vector space \( \mathbb{M}(n; \mathbb{C}) \) of \( n \times n \) matrices with complex entries by the rule \((a, a') \to tr(a^\dagger a')\). Fix an orthonormal basis \( \{a_i\}_{1 \leq i \leq n^2} \) for \( \mathbb{M}(n; \mathbb{C}) \) with respect to this Hermitian inner product. For each integer \( i \in \{1, \ldots, n^2\} \), let \( \omega^i \) denote the \( \mathbb{C} \)-valued 1-form on \( \text{Gl}(n; \mathbb{C}) \) given by

\[
\omega^i_m = tr(a_m^1 dm).
\]

As noted in Chapter 5f, each of these forms is a left invariant 1-form on \( \text{Gl}(n; \mathbb{C}) \). As their real and imaginary parts span \( T^*\text{GL}(n; \mathbb{C})\), so their real and imaginary parts span \( T^*\text{GL}(n; \mathbb{C}) \) at each point. This understood, their real and imaginary parts define an isomorphism from \( T\text{GL}(n; \mathbb{C}) \) to the product bundle \( \text{Gl}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C}) \). The isomorphism in question sends a given tangent vector \( v \) at a given point \( m \in \text{Gl}(n; \mathbb{C}) \) to the pair \((m, \sum_{1 \leq i \leq n^2} a_i (\omega^i, v))\). Meanwhile, the collection \( \{\omega^i\} \) defines the metric

\[
g = \sum_{1 \leq i \leq n^2} \text{re}(\overline{\omega}^i \otimes \omega^i)
\]
on $\text{TGl}(n; \mathbb{C})$. Note that $g$ does not depend on the chosen Hermitian basis because any such basis is obtained from any other by the action of some element in $U(n^2)$.

As an aside, note that $\text{T}\mathbb{M}(n; \mathbb{C})$ has an almost complex structure such that the resulting complex, rank $n^2$ vector bundle is spanned at each point by the sections entries of the matrix $\frac{\partial}{\partial m}$. As an open set in $\mathbb{M}(n; \mathbb{C})$, the tangent bundle to $\text{Gl}(n; \mathbb{C})$ inherits this almost complex structure. This understood, the isomorphism defined by the forms $\{\omega^i\}$ from $\text{TGl}(n; \mathbb{C})$ to $\text{Gl}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C})$ is $\mathbb{C}$-linear, and thus an isomorphism of complex vector bundles. The metric $g$ just described is defined on the underlying real tangent bundle of $\text{Gl}(n; \mathbb{C})$ by the Hermitian metric $\sum_{i=1}^{n^2} \overline{\omega}^i \otimes \omega^i$ on the complex bundle.

To say more about the metric $g$, introduce from Chapter 2d the almost complex structure $J_0 \in \mathbb{M}(2n; \mathbb{R})$ so as to view $\mathbb{M}(n; \mathbb{C})$ as the subvector space $\mathbb{M}_j \subset \mathbb{M}(n; 2\mathbb{R})$ of matrices that obey $m_{k,1} J_0 - J_0 m = 0$. In particular, recall that $m \in \mathbb{M}_j$ if and only if its entries obey $m_{2k,2i} = m_{2k-1,2i-1}$ and $m_{2k-1,2i} = -m_{2k-1,2i-1}$ for any pair of $i, k$ from $\{1, \ldots, n\}$. This understood, the identification between $\mathbb{M}(n; \mathbb{C})$ and $\mathbb{M}_j$ associates to a given $a \in \mathbb{M}(n; \mathbb{C})$ the matrix $m \in \mathbb{M}_j$ with $m_{2k-1,2i-1}$ and $m_{2k-1,2i}$ given by the respective real and imaginary parts of $a$. This identification sends $\text{Gl}(n; \mathbb{C}) \subset \mathbb{M}(n; \mathbb{C})$ to the subspace $\mathbb{G}_j \subset \mathbb{M}_j$ of invertible elements. In particular, it follows directly from all of this that the metric $g$ defined above on $\text{Gl}(n; \mathbb{C})$ is the restriction to $\mathbb{G}_j \subset \text{Gl}(2n; \mathbb{R})$ of the metric that was defined above in the preceding part of this chapter.

The geodesics on $\text{Gl}(n; \mathbb{C})$: The next proposition describes some of the geodesics for the metric just described.

**Proposition:** Identify $\text{TGl}(n; \mathbb{C})$ with $\text{Gl}(n; \mathbb{C}) \times \mathbb{M}(n; \mathbb{C})$ as just described. With the metric $g$ as above, any path of the form $t \rightarrow m e^{\alpha t}$ with $m \in \text{Gl}(n; \mathbb{C})$ and $\alpha \in \mathbb{M}(n; \mathbb{C})$ obeying $[\alpha, \alpha^\dagger] = 0$ is a geodesic.

**Proof:** There are two ways to argue. Here is the first: Suppose that $\alpha \in \mathbb{M}_j$, and $m \in \mathbb{G}_j$. The path $t \rightarrow m e^{\alpha t}$ in $\text{Gl}(2n; \mathbb{R})$ lies entirely in $\mathbb{G}_j$, and so if it is a geodesic in $\text{Gl}(2n; \mathbb{R})$, then it is a geodesic in $\mathbb{G}_j$ as defined using the metric on the latter that is induced by the the inclusion $\mathbb{G}_j \subset \text{Gl}(2n; \mathbb{R})$. It is a geodesic in $\text{Gl}(2n; \mathbb{R})$, if it is the case that $[\alpha, \alpha^\dagger] = 0$. This understood, let $\alpha_c$ denote $\alpha$’s counterpart in $\mathbb{M}(n; \mathbb{C})$. As explained momentarily, the commutator $[\alpha, \alpha^\dagger] = 0$ in $\text{Gl}(2n; \mathbb{R})$ if and only if $[\alpha_c, \alpha_c^\dagger] = 0$ in $\text{Gl}(n; \mathbb{C})$. As noted in Chapter 2d, the product $b b'$ of matrices in $\mathbb{M}_j$ is in $\mathbb{M}_j$ and obeys $(b b')_c = b_c b'_c$. Thus $(m e^{\alpha})_c = m_c e^{\alpha_c}$, and so the paths in $\text{Gl}(n; \mathbb{C})$ as described by the proposition are geodesics.
To see that \([a, a^T] = 0\) in \(\text{Gl}(2n; \mathbb{R})\) if and only if \([a_c, a_c^\dagger] = 0\) in \(\text{Gl}(n; \mathbb{C})\), note first that the matrix \(j_0\) as defined in Chapter 2d obeys \(j_0^T = -j_0\). This implies that \(a^T \in M_i\) if \(a \in M_i\). It is left for the reader to verify that \((a^\dagger)_c = a_c^\dagger\). Granted these last points, it then follows \([a, a^T] \in M_i\) and \(([a, a^T])_c = [a_c, a_c^\dagger]\).

The second proof is the \(\mathbb{C}\) analog of the argument for the \(\text{Gl}(n; \mathbb{R})\) version of this proposition. The only substantive difference is that the expression for the metric in the case of \(\text{Gl}(n; \mathbb{C})\) replaces matrix transposes with Hermitian conjugates. To elaborate just a little, as with \(\text{Gl}(n; \mathbb{R})\), the fact that the metric is invariant with respect to the left translation diffeomorphisms implies that it is sufficient to prove that the paths \(t \to e^{a}\) are geodesics if \(a \in M(n; \mathbb{C})\) obeys \([a, a^\dagger] = 0\). To see if this is the case, recall from Chapter 5e that the exponential map \(a \to e^a\) from the vector space \(M(n; \mathbb{C})\) into \(\text{Gl}(n; \mathbb{C})\) restricts to a small radius ball about 0 in \(M(n; \mathbb{C})\) as a diffeomorphism from said ball onto a neighborhood of the identity matrix in \(\text{Gl}(n; \mathbb{C})\). The metric pulls back via this map to

\[
g|_a = -\text{re} \int_0^1 \int_0^1 \text{tr}(e^{as} da^\dagger e^{sa} e^{sa} da e^{-sa}) \, ds \, ds'
\]

Note in particular that Hermitian adjoints appear in this formula, not transposes. But for this replacement here and subsequently the argument for the proposition is identical to that used to prove the \(\text{Gl}(n; \mathbb{R})\) version of the proposition. As such, the details are left to the reader.

The geodesics on \(U(n)\) and \(SU(n)\): View \(U(n)\) as the subgroup in \(\text{Gl}(n; \mathbb{C})\) of matrices \(m\) with \(m^{-1} = m^\dagger\) and view \(SU(n) \subset U(n)\) as the subgroup of matrices with determinant 1. As such, the metric on \(\text{Gl}(n; \mathbb{C})\) described above induces metrics on both \(U(n)\) and \(SU(n)\). Note in this regard that the 1-from \(\text{tr}(a m^{-1} dm)\) restricts to zero on \(TU(n)\) if \(a^\dagger = a\), and it is nowhere zero and real valued on \(TU(n)\) if \(a \neq 0\) and \(a^\dagger = -a\). Such a form is non-zero on \(TSU(n)\) if and only if \(a\) is not a multiple of the identity.

Keeping this in mind, let \(\{a_i\}_{1 \leq i \leq n^2}\) denote an orthonormal basis for the vector space \(\mathbb{A}(n; \mathbb{C}) \subset M(n; \mathbb{C})\) denote the vector subspace (over \(\mathbb{R}\)) of skew Hermitian matrices. Thus \(a^\dagger = -a\) when \(a \in \mathbb{A}(n; \mathbb{C})\). Agree to take \(a_i\) to have trace zero when \(i < n^2\) and to take \(a^*_{n^2} = \frac{1}{\sqrt{n}}\). The induced metric on \(U(n)\) is \(g = \sum_{1 \leq i < n^2} \omega_i \otimes \omega_i^*,\) and that on \(SU(n)\) is \(g = \sum_{1 \leq i < n^2} \omega_i \otimes \omega_i^*\).

Let \(m \in U(n)\) and let \(a \in \mathbb{A}(n; \mathbb{C})\). As noted in Chapter 5f, the matrix \(m e^a\) is in \(U(n)\). Moreover, if \(m \in SU(n)\) and \(a\) has zero trace, then \(m e^a\) is in \(SU(n)\). This is background for the proposition that follows.
**Proposition:** The geodesics for the metric defined above on $U(n)$ are of the form $t \to me^a$ where $m \in U(n)$ and $a \in A(n; \mathbb{C})$. The geodesic for the metric just defined on $SU(n)$ are of this form with being in $m \in SU(n)$ and being $a$ a traceless matrix from $A(n; \mathbb{C})$.

**Proof:** Because the curves described are geodesics in $Gl(n; \mathbb{C})$, they are apriori geodesics in the subgroup in question. The first proposition in Chapter 8c above guarantees that these account for all of the geodesics.

8g) **Geodesics and matrix groups**

A Lie group $G$ is said in what follows to be a compact matrix group if it appears as a subgroup of some $n \geq 1$ version of either $SO(n)$ or $U(n)$. Thus, $G$ is a submanifold with the property that $m^{-1} \in G$ if $m$ is a matrix in $G$, and also $m\hbar \in G$ when both $m$ and $\hbar$ are in $G$. What was said above about the geodesics in $SO(n)$, $U(n)$ and $SU(n)$ has an analog for any compact matrix group. The metric here is assumed to be the one induced by the metric on $SO(n)$ or $U(n)$ that is discussed in the preceding Part e) or Part f) of this chapter. This is to say that a tangent vector to $G$ is apriori a tangent vector to the linear group, and so the metric used in Part e) or Part f) can be used to define its norm.

To elaborate, remark first that $T_G\hbar$ appears as a subvector space (over $\mathbb{R}$) in the vector space of rank $n \times n$ real, anti-symmetric matrices, or complex anti-hermitian matrices, as the case may be. It is customary to use $\text{lie}(G)$ to denote this vector space. The following proposition describes the geodesics in the general case:

**Proposition:** Suppose that $G$ is a compact matrix group, and let $m \in G$.

- The map $a \to me^a$ from $\text{lie}(G)$ to the general linear group has image in $G$.
- This map restricts to a some ball about the origin in $\text{lie}(G)$ as a diffeomorphism of the latter onto a neighborhood of $m$ in $G$.
- For any given $a \in \text{lie}(G)$, the map $t \to me^a$ from $\mathbb{R}$ to $G$ is a geodesic; and all geodesics through $m \in G$ are of this sort.

**Proof:** The second and third bullets of the proposition follow with the verification of the first bullet using the same arguments given above for the $G = SO(n)$ or $U(n)$ or $SU(n)$ cases. To see about the first bullet, note first that it is enough to consider the case when $m$ is the identity. To prove the assertion in the latter case, it is enough to prove that the map $a \to e^a$ maps $\text{lie}(G)$ into $G$. Indeed, as the differential of this map at the origin is the identity on $\text{lie}(G)$, it then follows from the inverse function theorem that this map restricts to a ball about the origin in $\text{lie}(G)$ as a diffeomorphism onto a neighborhood in $G$ of 1.
Granted the preceding, fix some very large integer, N, and write $e^a$ as $(e^{aN})^N$.

Given the definition of $\text{lie}(G)$ as $\text{TG}|_{\iota}$, it follows that $e^{aN}$ can be written as $\mathfrak{g}_N + \tau_N$ where $\mathfrak{g}_N \in \mathfrak{g}$ and where $|\tau_N| \leq c_0 N^2$. As a consequence, $e^a = (\mathfrak{g}_N)^N + \mathfrak{m}_N$ where $|\mathfrak{m}_N| \leq c_0 N^{-1}$. As each $\mathfrak{g}_N \in \mathfrak{g}$, so $(\mathfrak{g}_N)^N \in G$. Taking $N \to \infty$ shows that $e^a \in G$ also.

A) Appendix: The proof of the vector field theorem

The primary purpose of this appendix is to prove the vector field theorem in Chapter 8c. By way of reminder, this theorem concerns solutions to an equation for a map from a neighborhood of the origin in $\mathbb{R}$ to $\mathbb{R}^m$ of the following sort: Let $v: \mathbb{R}^m \to \mathbb{R}^m$ denote a given smooth map, and let $z$ denote a map from a neighborhood of the origin in $\mathbb{R}$ to $\mathbb{R}^m$. The map $z$ is asked to obey the equation

$$\frac{dz}{dt} = v(z).$$

(1)

**Proof of the vector field theorem**: Let $B \subset \mathbb{R}^n$ denote the radius 2 ball. Fix $\varepsilon > 0$ and let $I = [-\varepsilon, \varepsilon]$. Let $C^0,(I; \mathbb{R}^m)$ denote the space of continuous maps from $I$ to $\mathbb{R}^m$ that equal 0 at $0 \in I$. Keep in mind that this is a complete metric space where the topology is defined by the distance function that assigns to maps $u$ and $w$ the distance $||u - w|| = \sup_{t \in I} |u(t) - w(t)|$. Define a map, $T: B \times C^0,(I; \mathbb{R}^m) \to C^0,(I; \mathbb{R}^m)$ by the rule

$$T(y, u) = \int_{[0,t]} v(y + u(s)) \, ds$$

(2)

If $T(y, u) = u$ and if $u$ is a smooth map, then the map $t \to z(t) = y + u(t)$ obeys (1) with initial condition $z(0) = y$. Conversely, a solution to (1) with $z(0) = y$ is a fixed point of the map $u \to T_y(u) = T(y; u)$.

**Existence of a fixed point**: To see that there is a fixed point, let $c_0 \geq 1$ denote a constant such that $|v(z)| \leq c_0$ when $|z| \leq 1$. Let $B_1$ denote the ball in $C^0,(I; \mathbb{R}^m)$ with radius 1 centered about the constant map. On this ball,

$$|T_y(u)| \leq |t| c_0 \leq \varepsilon c_0.$$

Thus $T_y$ maps $B_1$ to itself if $\varepsilon < c_0^{-1}$. To see that $T_y$ is a contraction for small $\varepsilon$, let $c_0 \geq 1$ be such that $|v(z) - v(z')| \leq c_0 |z - z'|$ when $|z|$ and $|z'|$ are both less than 1. Then

$$|(T_y(u) - T_y(u'))| \leq c_0 t \sup_{|z| \leq 1} |u(s) - u'(s)| \leq c_0 t ||u - u'||.$$
Thus if $\varepsilon < c_0^{-1}$, then $\| T_y(u) - T_y(u') \| \leq \frac{1}{2} \| u - u' \|$ and $T_y$ is uniformly contracting. As a consequence, it has a unique fixed point on the ball $B_1$.

*Fixed points are smooth maps:* To see that a fixed point is smooth, note that what is written on the right hand side (2) is a $C^1$ function. So, if $u$ is equal to what is written on the right hand side of (2), then $u$ is a $C^0$ function. This implies that the right hand side of (2) is a $C^2$ function. But then $u$ is a $C^2$ function and so the right hand side of (2) is a $C^3$ function. Continuing in this vein shows that $u$ is infinitely differentiable. It follows, in particular, that any fixed point of (2) obeys (1).

*There is a unique solution with $z(0) = y$:*
To prove this, suppose that $t \mapsto z(t)$ is a solution to (1) for $t \in I$. I need to prove that $u = z - y$ is in $B_1$, for then the result follows from the contraction mapping theorem. To see that $u$ is in $B_1$, I need to prove that $|u(t)|$ is small for $t \in I$. To see that such is the case, let $s \in (0, \infty)$ denote the smallest value of $t$ such that $|u(t)| = 1$. Since $u$ is a fixed point of (2), it follows that

$$|u(\pm s)| \leq c_0 |s|$$

and so if the constant $\varepsilon$ that defines the interval $I$ is chosen so that $\varepsilon < c_0^{-1}$, then this shows that $s > \varepsilon$ and so $u \in B_1$.

*The dependence on $y \in B$:*
Let $y \mapsto u_y$ now denote the mapping that associates to each $y \in B$ the unique fixed point of (2) in $B_1$. It is a consequence of (2) and the fact that $|v(z) - v(z')| \leq c_0 |z - z'|$ when both $|z|$ and $|z'|$ are less than $1$ that

$$\| u_y - u_{y'} \| \leq c_0 \varepsilon (|y - y'| + \| u_y - u_{y'} \|)$$

thus, given that $\varepsilon < c_0^{-1}$, this tells us that the assignment $y \mapsto u_y$ is Lipschitz as a map to the metric space $C^{0,\varepsilon}(I, \mathbb{R}^m)$. In particular, for each $t \in I$, one has $|u_y(t) - u_{y'}(t)| \leq c_0 |y - y'|$.

To say something about higher order derivatives, one takes successively more involved difference quotients. This is straightforward, but tedious and I leave it to you.

**Additional reading**

Chapter 9: Properties of geodesics

This chapter discusses various notions that involve geodesics. The final parts of this chapter use some of these to prove the geodesic theorem of Chapter 8b.

9a) The maximal extension of a geodesic

Let $I \subset \mathbb{R}$ denote a closed interval that $I$ is not the whole of $\mathbb{R}$, and let $s$ denote an end point of $I$, for the sake of the discussion, the maximum point. Let $\gamma: I \rightarrow M$ denote a geodesic. It follows from Proposition 1 in Chapter 8 that there exists an extension of $\gamma$ to an interval with maximal endpoint $s + \varepsilon$ for some $\varepsilon > 0$. To see this, take a point $q = \gamma(s')$ for $s'$ nearly $s$. Then $q$ is very close to $p = \gamma(s)$. Proposition 1 in Chapter 8 tells us that the geodesic $t \rightarrow \gamma(t) = \gamma(s' + t)$ is defined for $|t| \leq \varepsilon$ for some $\varepsilon$ that depends only on $p$. This extends $\gamma$. It follows as a consequence that each geodesic in $M$ has a maximum extension. In particular, if $M$ is compact, then each geodesic extends as a geodesic mapping $\mathbb{R}$ into $M$.

A manifold $M$ is called geodesically complete if each geodesic can be extended to have domain the whole of $\mathbb{R}$. All compact manifolds are geodesically complete. To see a manifold with metric that is not geodesically complete, take $M$ to be the complement in $\mathbb{R}^m$ of any non-empty set but take the Euclidean metric from $\mathbb{R}^n$. In general, the complement of any point in a geodesically complete manifold is not geodesically complete.

9b) The exponential map

The following proposition holds when $M$ is geodesically complete:

**Proposition:** There exists a smooth map, $\exp: \mathbb{R} \times TM \rightarrow M$ with the following property: Fix $v \in TM$. Then the corresponding map $\exp_v = \exp(\cdot, v): \mathbb{R} \rightarrow M$ is the unique solution to the geodesic equation with the property that $\exp_v(0) = \pi(v)$ and $(\exp_v')_{t=0} = v$.

**Proof:** The map $\exp$ is well defined because $M$ is geodesically complete. The issue is whether $\exp$ is smooth. This is a local question in the sense the following sense: The map $\exp$ is smooth if and only if it is smooth near any given point in its domain. To see that such is the case, fix a point $v \in TM$, a neighborhood $O \subset TM$ of $v$ and $\varepsilon > 0$ such that $\exp$ is smooth on $(-\varepsilon, \varepsilon) \times O$. Proposition 2 in Chapter 8 can be used to obtain $\varepsilon$. Let $T \geq \varepsilon$ denote either $\infty$ or the smallest number such that $\exp$’s restriction to $[-T, T] \times \hat{O}$ for any open set $\hat{O} \subset TM$ that contains $v$. The claim made by the proposition follows with a proof that $T = \infty$.

To see that $T = \infty$, suppose to the contrary that $T$ is finite so as to derive some nonsense. Let $p = \exp_v(T) = \exp(T, v)$ and let $v_\tau \in TM|_p$ denote the vector $(\exp_v')_{\tau} \big|_{\tau=0}$ at $t = T$. Fix
an open set $O_v \subset \text{TM}_p$ containing $v$. According to Proposition 2 in Chapter 8c, there exists $\delta > 0$ and a smooth map, $\gamma_{O_v}$, from $(-\delta, \delta) \times O_v$ into $M$ such that if $v' \in O_v$, then $\gamma_{O_v}(\cdot, v')$ is a geodesic that starts at $\pi(v') \in M$ and has initial velocity $v'$.

With the preceding understood, fix an open neighborhood $O' \subset \text{TM}$ of $v$ such that $e$ is smooth on $(T - \frac{1}{2} \delta, T)$, $\text{TM}$, $\text{TM}_p$. Define the map $\hat{e}$ to map this same domain into $\text{TM}$ by the rule $\hat{e}(t, v') = e^* \gamma_{O_v}(t, v')$. Thus, $\pi \circ \hat{e} = e_*$. There is then a neighborhood $O'' \subset O'$ such that $\hat{e}$ maps $(T - \frac{1}{4} \delta, T + T)$ into $O_v$. This understood, the uniqueness assertion in Proposition 1 of Chapter 3c implies that

$$e(T + s - \frac{1}{4} \delta, v') = \gamma_{O_v}(s, e(T - \frac{1}{4} \delta, v'))$$

This exhibits $\hat{e}$ on $(0, T + \frac{1}{4} \delta) \times O''$ as the composition of two smooth maps, and hence smooth. This is nonsense given the definition of $T$.

Given the map $e$, define the exponential map,

$$\exp: \text{TM} \to M$$

by the rule $\exp(v) = e(1, v)$. This is to say that $\exp(v)$ is the point in $M$ that is obtained from $\pi(v)$ by traveling for time $t = 1$ from $\pi(v)$ along the geodesic that starts at $\pi(v)$ at time 0, has initial direction $v$ at $\pi(p)$ and has speed $|v|$. When $p \in M$ has been specified, the restriction of $\exp$ to $\text{TM}_p$ is denoted in what follows by $\exp_p$.

The map $\exp$ is called the exponential map because this map is given by the exponential of a matrix when $M$ is $\text{SO}(n)$ or $\text{U}(n)$ or one of their subgroups. This identification of maps follows from what is said about the geodesics on such groups in Chapters 8e-8g.

**9c) Gaussian coordinates**

The exponential map can be used to obtain a preferred coordinate chart for a neighborhood of any given point in $M$. This is because the exponential map restricts to a small radius ball in the fiber of $\text{TM}$ over any given point as a diffeomorphism. To see that such is the case, fix a point $p \in M$. Then it is enough to prove that the differential of $\exp_p$ at the $v = 0$ in $\text{TM}_p$ is an isomorphism. In fact, it is an isometry. To elaborate, remark that $\pi'(\text{TM}_p)$ is canonically isomorphic to $\pi^*(\text{TM}_p)$ where $\pi: \text{TM} \to M$ is the projection map. This follows from the fact that the pull-back $\pi^*(\text{TM}_p) \to \text{TM}_p$ consist of pairs $(v, v') \in \text{TM}_p$.

The differential at $0 \in \exp_p$ sends $v \in \text{TM}_p$ to the push-forward of the vector $\frac{\partial}{\partial s}$ at $s = 0$ via the map $s \to \exp_p(sv)$. At any given $s$, the point $\exp_p(sv) \in M$ is the point
given by \( \epsilon(1, sv) \) with \( \epsilon \) as defined in the preceding part of this Chapter. This is the same point as \( \epsilon(s, v) \). As a consequence, \( ((\exp_p)_t, v)|_{t=0}^{TM_p} = (\epsilon(\cdot, v), \frac{\partial}{\partial t})|_{t=0}^{p} = v. \)

Fix an isometric identification between \( TM_p \) and \( \mathbb{R}^n \). The image via a small radius ball in \( \mathbb{R}^n \) via the exponential map gives a diffeomorphism between this ball and an open neighborhood of \( p \) in \( M \). The inverse of this diffeomorphism gives what are called Gaussian or normal coordinates centered at \( p \).

You may see the term injectivity radius at \( p \) used with regards to Gaussian coordinates. This is the least upper bound on the radii of a ball in \( \mathbb{R}^n \) about the origin which is mapped diffeomorphically onto its image by \( \exp_p \).

As is demonstrated next, the metric \( g_{ij} \) in these coordinates and the geodesics have a particularly nice form.

**Proposition:** Let \((x^1, \ldots, x^n)\) denote Gaussian coordinates centered at any given point in \( M \). Then the metric in these coordinates has the form \( g_{ij} = \delta_{ij} + K_{ij}(x) \) where \( K_{ij}(x)x^i = 0 \) and also \( |K_{ij}| \leq c_0|x|^2 \). Moreover, the geodesics through the given point appear in these coordinates as the straight lines through the origin.

**Proof:** The fact that the straight lines through the origin are the geodesics through the given point follows directly from the definition of the exponential map. As noted above, the differential of the exponential map at the origin in \( TM_p \) is the identity, and this implies that \( g_{ij} = \delta_{ij} + K_{ij} \) where \( |K| \leq c_0|x| \). To see that it is \( O(|x|^2) \), suppose that \( K_{ij} = K_{ij} + O(|x|^2) \). Note that \( K_{ij} = K_{ji} \). Then \( \Gamma_{jk}^i = \frac{1}{2} (K_{ij,k} + K_{ik,j} - K_{jk,i}) + O(|x|) \). This being the case, then the line \( t \to tv \) with \( v \) a constant vector can be a geodesic only if \( K_{ij,k} + K_{ik,j} - K_{jk,i} = 0 \). Taking all possible \( v \), this says that \( K_{ij,k} + K_{ik,j} - K_{jk,i} = 0 \). A little linear algebra will show that this equation can be satisfied for all \( i, j \) and \( k \) if \( K_{ij,k} = K_{ik,j} = K_{jk,i} = 0 \). Thus, all cases where two of \( i, j \) and \( k \) agree are zero. Now consider the case where \( i, j, \) and \( k \) are distinct, say \( 1, 2, 3 \). Then \( i = 1, j = 2 \) and \( k = 3 \) finds that \( K_{12,1} = 0 \). By symmetry, \( K_{12,1} = 0 \). Thus, all cases where two of \( i, j \) and \( k \) agree are zero.

To see that \( K_{ij}(x)x^i = 0 \), fix \( \epsilon > 0 \) but with \( \epsilon < c_0^{-1} \). Fix \( v \in \mathbb{R}^n \) with small norm; so that the point \( x = v \) lies in the Gaussian coordinate chart. Let \( t \to \eta(t) \) denote a smooth map from \([0, \infty)\) to \( \mathbb{R}^n \) such that \( \eta(0) = 0 \) and \( \eta = 0 \) for \( t \geq \epsilon \). For each \( s \in (-\epsilon, \epsilon) \) consider the path \( \gamma(s) \) given by \( t \to tv + s\eta(t) \). If \( \epsilon \) is small, then this will be a path in the coordinate chart for all \( t \) with \( |t| \leq 1 \). The \( s = 0 \) version is a geodesic, and as a consequence, the function
\[ s \to \ell_{[\gamma]} = \frac{1}{2} \int_0^1 (g_{ij}(\gamma[s])\dot{\gamma}[s]^i\dot{\gamma}[s]^j)^{1/2} \, dt \]

Has derivative zero at \( s = 0 \). The derivative of the integrand at \( s = 0 \) is

\[ |v|^i \left( K_{ij}(tv)v^j \dot{\eta}^i + \frac{1}{2} \eta^i \partial_k K_{ji} v^j v^i \right). \]

Integrating this expression, and then integrating by parts finds that

\[ (\partial_t K_{ij} - \frac{1}{2} \partial_i K_{jk}) |_v v^k w^i = 0 \]

for all \( t \in [0, 1] \) and for all vectors \( v \) and \( w \) in \( \mathbb{R}^n \). To see what this implies, consider first taking \( v = w \). Then this says that \( \partial_t (K_{ij}(tv)v^j v^i) = 0 \). Since \( K_{ij}(0) = 0 \), this implies that \( K_{ij}(tv)v^j v^i = 0 \) for all \( t \). This understood, take \( w \) so that \( w^i v^i = 0 \). (Thus, \( w \) is Euclidean orthogonal to \( v \).) Then the preceding equation asserts that

\[ \partial_t K_{ij}(tv)v^j w^i = \frac{1}{2} w^k (\partial_k K_{ij}) |_v v^j v^i. \]

To see what to make of this equation, differentiate with respect to \( s \) the equation

\[ K_{ij}(t(v + sw))(v^i + sw^i)(v^j + sw^j) = 0 \]

and set \( s = 0 \) to see that \( tw^k (\partial_k K_{ij}) |_v v^i v^j = -2K_{ij}(tv)v^i w^j \). Using this fact tells us that

\[ t \partial_t (K_{ij}(tv)v^i w^j) = -K_{ij}(tv)v^i w^j. \]

Integration of this tells us the following: If \( t > t' \), then \( K_{ij}(tv)v^i w^j = K_{ij}(t'v)v^i w^j e^{t-t'} \). Since \( K_{ij}(0) = 0 \), taking \( t' \to 0 \) gives what we want, that \( K_{ij}(tv)v^i w^j = 0 \) for all \( t, v \) and \( w \).

This last proposition has the following consequence: Introduce ‘spherical’ coordinates \((r, \theta^1, \ldots, \theta^{n-1})\) for \( \mathbb{R}^n \). This is to say that \( r = |x| \) and the coordinates \( \{\theta^a\}_{1 \leq a \leq n-1} \) are angle coordinates on \( S^{n-1} \). The metric in Gaussian coordinates when written in terms of these spherical coordinates for \( \mathbb{R}^n \) has the form

\[ g = dr^2 + r^2 k_{ab} d\theta^a d\theta^b \]

where \( k_{ab} \) is a function of \( r \) and the coordinates for \( S^{n-1} \). Here, \( dr^2 \) is a common shorthand for \( dr \otimes dr \), and \( d\theta^a \, d\theta^b \) is a common shorthand of \( d\theta^a \otimes d\theta^b \)
9d) The proof of the geodesic theorem

The proof in what follows of Chapter 8b’s geodesic theorem is broken into various parts

**Part 1**: A map, \( \gamma \), from an interval \( I \subset \mathbb{R} \) to \( M \) is said to be **locally length minimizing** when the following is true: Fix any point \( t_0 \in I \) and there is a neighborhood in \( I \) of \( t_0 \) such that if \( t \) is in this neighborhood, then the length of \( \gamma \)'s restriction to the interval between \( t \) and \( t_0 \) is equal to \( \text{dist}(\gamma(t_0), \gamma(t)) \).

This part of the proof states and then proves a lemma to the effect that geodesics are locally length minimizing. The proof uses the Gaussian coordinates that were just introduced in the previous part of this Chapter.

**Lemma 1**: Let \( p \in M \) denote a given point. There exists \( \varepsilon > 0 \) such that if \( q \in M \) has distance \( \varepsilon \) or less from \( p \), then there is a unique geodesic that starts at \( p \), ends at \( q \) and has length equal to the distance between \( p \) and \( q \). Moreover, this is the only path in \( M \) between \( p \) and \( q \) with this length.

**Proof**: If \( q \) is close to \( p \), then \( q \) will lie in a Gaussian coordinate chart centered at \( p \). Such a chart is given by a diffeomorphism, \( \varphi \), from an open neighborhood, \( U \), of \( p \) to the interior of a ball of some radius centered at 0 in \( \mathbb{R}^n \). Let \( \rho \) denote the radius of this ball. Take \( x = (x^1, \ldots, x^n) \) to be the Euclidean coordinates for \( \mathbb{R}^n \), thus the image of \( \varphi \) is the ball where \( |x| < \rho \). With these coordinates understood, suppose that \( \varphi(q) \) is the point \( x = v \) where \( 0 < |v| < \rho \). The part of the geodesic between \( p \) and \( q \) is the path \( t \rightarrow x(t) = tv \), this parametrized by \( t \in [0, 1] \). What follows explains why this path has length equal to the distance between \( p \) and \( q \); and why it is the only such path in \( M \).

To start, suppose that \( \gamma \colon [0, 1] \rightarrow M \) is any other path between \( p \) and \( q \). Let \( \gamma_p \) denote the component of \( \gamma \cap U \) that starts at \( p \) as \( \gamma(0) \). Write the \( \varphi \) image of this part of \( \gamma \) in polar coordinates \( t \rightarrow (r(t), \theta^1(t), \ldots, \theta^n(t)) \). Write the metric \( g_{ij} \, dx^i \otimes dx^j \) in polar coordinates as at the end of the preceding section. Having done so, one can write the square of the norm of \( \dot{\gamma} \) for this part of \( \gamma \) as

\[
|\dot{\gamma}|^2 = t^2 + r^2 k_{ab} \dot{\theta}^a \dot{\theta}^b.
\]

It follows from this that \( |\dot{\gamma}| \geq |t| \) at points where \( \gamma \) is in \( U \), with equality if and only if \( \gamma(t) = f(t) v_s \) where \( v_s \in \mathbb{R}^n \) has norm 1 and where \( f \) is a smooth function on an \( [0, t_0) \subset [0, 1] \) that contains 0. Note that this interval is a proper subset of \( [0, 1] \) if \( \gamma \) exits the coordinate patch, in which case \( f(t) \rightarrow \rho \) as \( t \rightarrow t_0 \). If \( t_0 = 1 \), then \( v_s \) is the unit vector in the direction
of v, and f(t) → |v| as t → 1. In any event, it follows from this that the length of γ is greater than |v| unless the image of γ is the arc between 0 and v.

Part 2: Lemma 1 has the following corollary: A path in M whose length is the distance between its endpoints parametrizes a subset of a geodesic. The next lemma asserts this and a bit more.

**Lemma 2:** Let p, q be points in M and suppose that γ ⊂ M is a path from p to q whose length is the distance between p and q. Then γ is an embedded closed subset of a geodesic.

**Proof:** It follows from Lemma 1 that γ is part of a geodesic. As can be seen from the Proposition in Chapter 1d, a geodesic is locally embedded. Thus, γ is embedded unless it crosses itself. This is to say that there is a positive length subpath γ′ ⊂ γ such that starts and ends at the same point. However, if such a subpath exists, then the complement in γ of the interior of γ′ would be a path from p to q with length less than that of γ. By assumption, no such path exists.

Part 3: Suppose M is a manifold with a complete Riemannian metric, g. Fix any two points, p and q, in M. This part proves that there is geodesic from p to q whose length is equal to the distance between p and q.

To see why this is, suppose that L is the distance between p and q. Fix ε > 0 and a map γ: [0, 1] → M with length less than L + ε and such that γ(0) = p and γ(1) = q. I can find N ≥ 1 (depending only on p and q) such that the following is true: There is a sequence of times

\[0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = 1\]

for which the points γ(t_k) and γ(t_{k+1}) for k ≥ 0 have distance 2L/N from each other and lie in a Gaussian coordinate chart centered at γ(t_k). I can also assume that γ(t_{k+2}) also lies in this coordinate chart if k ≤ N - 2. It follows that the shortest path from γ(t_k) to γ(t_{k+1}) is the geodesic arc between them. This understood, I can find a continuous, piece-wise smooth path from p to q with N smooth segments, with length less than L + ε, and such that each segment is a geodesic arc between its endpoints, and in particular a geodesic arc that lies entirely in a single Gaussian coordinate chart.

Now, consider a sequence ε_1 > ε_2 > \cdots > ε_m with limit zero, and for each m, a piece-wise smooth path from p to q of the sort just described. Let γ_m denote this path. There are N points, P_\alpha = \{p_{1,\alpha}, \ldots, p_{N-1,\alpha}\} in M where the segments that comprise γ_m join.
This ordered set defines a point in \( \times_{N-1} M \). Thus, we have a sequence \( \{P_m\}_{m=1,2,\ldots} \subset \times_{N-1} M \). Give the latter set the metric topology from the product of the Riemannian metric topologies on the factors. The sequence \( \{P_m\}_{m=1,2,\ldots} \) then lies in a bounded subset, and so in a compact subset. This is so because \( M \) is geodesically complete. This sequence has a convergent subsequence because \( M \) is geodesically complete. Let \( (p_1, \ldots, p_{N-1}) \) denote the limit. It follows from the way things were defined that each \( k \geq 0 \) version of \( p_{k+1} \) is contained in a Gaussian coordinate chart about \( p_k \). Here, \( p_0 = p \) and \( p_{N-1} = q \). Likewise, \( p_{k+2} \) is also contained in this chart if \( k \leq N-2 \). The sum of the distances between \( p_k \) and \( p_{k+1} \) for \( k \in \{0, \ldots, N-1\} \) is equal to \( L \). This understood, it follows that the union of the geodesic arcs between \( p_k \) and \( p_{k+1} \) for \( k \in \{0, \ldots, N-1\} \) is a length minimizing, piece-wise smooth path from \( p \) to \( q \). To see that this curve has no kinks, use the fact that \( p_{k+1} \) and \( p_{k+2} \) are both in the Gaussian coordinate chart centered at \( p_k \). It follows that the shortest path between \( p_k \) and \( p_{k+2} \) is the arc between them, and this must then be the union of the geodesic segments between \( p_k \) and \( p_{k+1} \) and between \( p_{k+1} \) and \( p_{k+2} \).

Note that the assumption of geodesic completeness is necessary. Consider, for example the case where \( M = \mathbb{R}^2 - \{0\} \) with its Euclidean metric. There is no shortest path between the points \((1, 0)\) and \((-1, 0)\) in \( M \).

**Part 4:** Given what is said in the preceding parts of the proof, there is only one loose end left to tie so as to complete the proof of the geodesic theorem. This is the assertion that the length minimizing path can be parametrized to have constant speed. Since this path is a geodesic, it is sufficient to prove that a solution to the geodesic equation has constant speed. That such is the follows directly from the proposition in Chapter 1c. One can also prove this by writing the the function \( t \to g(\gamma, \dot{\gamma}) \) with \( \gamma \) a geodesic in any local coordinate chart, and differentiating the resulting expression with respect to \( t \). Some algebraic manipulations with the geodesic equation can be used to prove that this derivative is zero.

**Additional reading**

- Riemannian Geometry, Peter Petersen; Springer, 2006.
Chapter 10: Principal bundles

A principal bundle is the Lie group analog of a vector bundle, and they are, in any event, intimately related to vector bundles. The definition requires the specification of a Lie group $G$. This chapter contains the beginnings of the principal $G$-bundle discussion. As was the case with vector bundles, my favorite definition is given first; the standard definition (which defines the same object) is the second definition.

10a) The definition

Fix a smooth manifold $M$, and a Lie group $G$. A principal $G$ bundle is a smooth manifold, $P$, with the following extra data:

• A smooth action of $G$ by diffeomorphisms; thus a map $m: G \times P \to P$ with the property that $m(1, p) = p$ and $m(h, m(g, p)) = m(hg, p)$. It is customary to write this action as $(g, p) \to pg^{-1}$.

• A surjective map $\pi: P \to M$ that is $G$-invariant. Thus, $\pi(pg^{-1}) = p$. The map $\pi$ is called the projection from $P$ to $M$.

• Any given point in $M$ has an open neighborhood, $U$, with a $G$-equivariant diffeomorphism $\phi: P|_U \to U \times G$ that intertwines the map $\pi$ with the evident projection from $U \times G$ to $U$. This is to say that if $\phi(p) = (\pi(p), h(p))$, with $h(p) \in G$, then $\phi(pg^{-1}) = (\pi(p), h(p)g^{-1})$.

Principal $G$-bundles $P$ and $P’$ over $M$ are said to be isomorphic when there is a $G$-equivariant diffeomorphism $\psi: P \to P’$ that intertwines the respective projection maps. A principal $G$-bundle is said to be trivial when it is isomorphic to $M \times G$, where the latter has the obvious projection and the obvious $G$-action. The examples discussed here concern solely the cases where $G = \text{Gl}(n; \mathbb{R})$, $\text{Gl}(n; \mathbb{C})$ or a matrix group such as $\text{SO}(n)$, $U(n)$ or $\text{SU}(n)$.

If $N \subset M$ is any subset, then $P|_N$ is used to denote $\pi^{-1}(N) \subset P$. If $N$ is a submanifold, then the restriction of $\pi$ and the multiplication map to $P|_N$ defines $P|_N$ as a principal $G$-bundle over $N$.

Note that the third bullet of the definition asserts that any given point in $M$ has a neighborhood, $U$, such that $P|_U$ is isomorphic to the product bundle.

Examples of principal bundles are given momentarily.

10b) A cocycle definition

The cocycle definition of vector bundles given in Chapter 3b has a principal bundle analog. This definition requires the specification of the following data: The first thing needed is a locally finite, open cover $\mathcal{U}$ of $M$. Needed next is a collection of
smooth maps from the intersections of the various sets in \( \mathcal{U} \) to the group \( G \); thus a set 
\( \{ g_{U'}: U \cap U' \to G \}_{U,U' \in \mathcal{U}} \). This data is required to obey the following conditions:

- \( g_{UU} \) is the constant map to the identity, \( \iota \in G \).
- \( g_{UU}^{-1} = g_{U'U} \).
- If \( U, U' \) and \( U'' \) are any three sets from \( \mathcal{U} \) with \( U \cap U' \cap U'' \neq \emptyset \), then the condition 
\( g_{UU'}g_{U'U}g_{U''U} = \iota \) must hold.

The map \( g_{UU} \) is said to be a principal bundle transition function and the conditions listed above are called cocyle constraints. Given the data just described, the principal \( G \) bundle is defined to be the quotient of the disjoint union \( \bigcup U \subset \mathcal{U} (U \times G) \) by the equivalence relation that puts \( (x, g) \in U \times G \) equivalent to \( (x', g') \in U' \times G \) if and only if \( x = x' \) and \( g = g_{UU}(x)g' \).

To see that this definition is equivalent to the previous, consider first a principal bundle that is defined by this cocycle data. The group \( G \) acts so as to have \( h \in G \) send the equivalence class of a pair \( (x, g) \) to \( (x, gh^{-1}) \). Meanwhile, the projection to \( M \) sends this same equivalence class to \( p \). Finally, the isomorphism required by the third bullet in Part a) of this chapter is as follows: Let \( x \in M \). Fix a set \( U \subset \mathcal{U} \) that contains \( x \). The map to \( U \times G \) of the equivalence class of \( U \times G \) assigns to this class none other but the product \( U \times G \). The cocycle conditions guarantee that all of this is well defined.

To go the other way, suppose that \( P \to M \) is a given principal bundle as defined in Part a) of this chapter. It is a consequence of the third bullet of the definition in Part a) that there is a locally finite cover, \( \mathcal{U} \), for \( M \) with the following property: If \( U \subset \mathcal{U} \), then there is an isomorphism, \( \phi_U: P|_U \to U \times G \). This understood, suppose that \( U \) and \( U' \) are any two intersecting sets from \( \mathcal{U} \). Then the composition \( \phi_U \circ \phi_U^{-1} \) maps \( U \cup U' \) to \( G \).

10c) Principal bundles constructed from vector bundles

What follows explains how to construct a principal bundle from a vector bundle.

Frame bundles: Let \( \pi: E \to M \) denote any given, rank \( n \)-vector bundle. Introduce \( P_{\text{Gl}(E)} \to M \) to denote the submanifold in \( \bigoplus E \) that consists of the \( n \)-tuples \( (e_1, \ldots, e_n) \) that give an orthonormal basis over their common base point in \( M \). This manifold \( P_{\text{Gl}(E)} \) is a principal \( \text{Gl}(n; \mathbb{R}) \) bundle over \( M \). To say more, remark that the projection to \( M \) sends a frame to its base point. Meanwhile, a matrix \( g \in \text{Gl}(n; \mathbb{R}) \) acts to send any given frame \( (e_1, \ldots, e_n) \) to \( (\sum_k g_{k1}e_k, \sum_k g_{k2}e_k, \ldots, \sum_k g_{kn}e_k) \). To see the required isomorphisms to
trivial bundles over suitable neighborhoods of points, fix attention on an open set $U \subset M$ where there is a vector bundle isomorphism $\varphi_U: E|_U \to U \times \mathbb{R}^n$. This isomorphism identifies $(\otimes_n E)|_U$ with $U \times (\otimes_n \mathbb{R}^n)$ and so identifies $P_{Gl|U}$ as $U \times Gl(n; \mathbb{R})$ where the embedding $Gl(n; \mathbb{R}) \to \otimes_n \mathbb{R}^n$ sends any given $n \times n$ matrix to the $n$-tuple given by its $n$ columns.

By the way, the transition functions for $P_{Gl|U}$ are the maps to $Gl(n; \mathbb{R})$ that are given by the transition functions for the vector bundle $E$. To see why this is, consider the open set $U$ described in the previous paragraph, and a second open set $U' \subset M$ that intersects $U$ and comes with a trivializing isomorphism $\varphi_{U'}: E|_{U'} \to U' \times \mathbb{R}^n$. Introduce $g_{UU'}: U \cap U' \to Gl(n; \mathbb{R})$ to denote the corresponding transition function for $E|_{U \cap U'}$. Now, let $(x, g = (v_1, \ldots, v_n)) \in U' \times Gl(n; \mathbb{R}) \subset U' \times (\otimes_n \mathbb{R}^n)$. The transition function for $E$ identifies this point with $(x, (g_{UU'}(x)v_1, \ldots, g_{UU'}(x)v_n)) = (x, g_{UU'}(x)g) \subset U \times Gl(n; \mathbb{R})$.

The principal bundle $P_{Gl|E}$ is called the frame bundle of $E$. In the case when $E = TM$, the corresponding $P_{GL(TM)}$ is called the frame bundle of $M$.

**Orthonormal frame bundles:** Suppose that $E \to M$ is a rank $n$-vector bundle with a fiber metric. The metric gives the principal $O(n)$ bundle $P_{O|E} \to M$. This is the submanifold of $P_{Gl|E}$ whose $n$-tuples supply an orthonormal frame for $E$ over their common base point. To elaborate, recall from Chapter 7a that there is a locally finite, cover $\mathcal{U}$ for $M$ with the following property: Each set $U \in \mathcal{U}$ comes with an isometric isomorphism $\varphi_U: E|_U \to U \times \mathbb{R}^n$. This isometry induces the principal bundle isomorphism $\varphi_U: P_{O|E} \to U \times O(n)$ by viewing $P_{O|E}$ as the submanifold in $\otimes_n E$ of $n$-tuples that give an orthonormal frame for $E$ over their common base point while viewing $O(n)$ as the subset of $\otimes_n \mathbb{R}^n$ of $n$-tuples of vectors that define an orthonormal basis for $\mathbb{R}^n$.

Recall from Chapter 7a that the bundle $E$ is said to be orientable if it can be defined using transition functions that take values in $SO(n)$. The lemma in Chapter 7a asserts that this is the case if and only if the real line bundle $\wedge^n E$ is isomorphic to the product line bundle. In any event, if $E$ is orientable, then there is the analogous principal $SO(n)$ bundle $P_{SO|E} \subset P_{O|E} \subset \otimes_n E$ that consists of the $n$-tuples that define an oriented frame over their common base point.

A natural case to consider arises when $M$ has a Riemannian metric and $E = TM$. The bundle $P_{O|TM}$ consists of the subspace in $\otimes_n TM$ of $n$-tuples that give an orthonormal basis for the tangent bundle to $M$ over their common base point. The manifold $M$ is said to be orientable when $\wedge^n TM$ is the trivial line bundle. One has in this case the principal $SO(n)$ bundle of oriented, orthonormal frames in $TM$.

As example of the latter, consider the case of $S^n$ with its round metric. The oriented orthonormal frame bundles for $TS^n$ turns out to be the group $SO(n+1)$ with the projection map to $S^n$ identifying the latter with the quotient space $SO(n+1)/SO(n)$. To see
how this comes about, view $S^n$ as the unit sphere in $\mathbb{R}^{n+1}$. A point in the oriented, orthonormal frame bundle consists of an $(n+1)$ tuple $(x, v_1, \ldots, v_n)$ of orthonormal vectors in $\mathbb{R}^{n+1}$; where the projection to $S^n$ gives the first element, $x$. The identification with $\text{SO}(n+1)$ takes the this $n+1$ tuple to the matrix with these $n+1$ vectors as its columns.

**Complex frame bundles:** Suppose that $\pi: E \to M$ is a complex vector bundle with fiber $\mathbb{C}^n$. Sitting inside $\oplus_n E$ is the principal $\text{Gl}(n; \mathbb{C})$ bundle $P_{\text{Gl}(E)} \to M$ of $n$-tuples that give a $\mathbb{C}$-linear basis for the fiber at each point. If $E$ has a hermitian metric, then one can define inside $P_{\text{Gl}(E)}$ the principal $\text{U}(n)$ bundle $P_{\text{U}(E)} \to M$ of $n$-tuples that define an orthonormal frame at each point. Recall in this regard the lemma in Chapter 7b that finds such transition functions if and only if the complex rank 1 bundle $\wedge^n E \to M$ is isomorphic to the product bundle $M \times \mathbb{C}$.

What follows is an example of $P_{\text{U}(E)}$. View $S^2$ as the unit radius ball about the origin in $\mathbb{R}^3$. Reintroduce the Pauli matrices $\{\tau^1, \tau^2, \tau^3\}$ from Chapter 6d, and let $E \to S^2$ denote the subbundle in $S^2 \times \mathbb{C}^2$ of pairs $(x, v)$ such that $x_j \tau^j v = iv$. This bundle has its Hermitian metric from $\mathbb{C}^2$. Let $(z, w)$ denote the complex coordinates of $\mathbb{C}^2$. Let $\pi: S^3 \to S^2$ denote the map that sends $(z, w)$ with norm 1 to $x = (2\text{re}(z \bar{w}), 2\text{im}(z \bar{w}), |w|^2 - |z|^2)$. The matrix $x_j \tau^j$ is

$$
\begin{pmatrix}
|w|\bar{\tau}^j - |z|\bar{\tau}^j & 2\bar{z}w \\
2z\bar{w} & -(|w|\bar{\tau}^j - |z|\bar{\tau}^j)
\end{pmatrix}
$$

Note in particular that the vector $v = \begin{pmatrix} w \\ z \end{pmatrix}$ is such that $x_4 \tau^4 v = iv$. Thus, $P_{\text{U}(E)} = S^3$.

**10d) Quotients of Lie groups by subgroups**

The example described above with the oriented orthonormal frame bundle of $S^n$ is but one application of a much more general construction that is summarized in the central proposition of this part of Chapter 10. The statement of the proposition requires introducing four new notions and some notation. These are introduced in what follows directly.

**Quotients of Lie groups:** Suppose that $G$ is a Lie group and $H \subset G$ is a Lie subgroup. This is to say that $H$ is a smooth submanifold of $G$ that has the following three properties: First, the identity element is in $H$. Second, the inverse of an element in $H$ is also in $H$. Finally, the product of any two elements in $H$ is in $H$. For example, $\text{SO}(n)$ is a subgroup of $\text{Gl}(n; \mathbb{R})$. Likewise, $\text{SU}(n)$ is a subgroup of $\text{U}(n)$.
If \( H \) is a subgroup of \( G \), use \( G/H \) to denote the space of equivalence classes that are defined by the rule \( g \sim g' \) if and only if \( g' = gh \) for some \( h \in H \). The topology here is the quotient topology.

The **Lie algebra**: The Lie algebra of the group \( G \) is the tangent space to \( G \) at the identity \( 1 \). This vector space is denoted by \( \text{lie}(G) \) in what follows. For example, the Lie algebra of \( \text{Gl}(m; \mathbb{R}) \) is the vector space of \( n \times n \) real matrices, and that of \( \text{Gl}(m; \mathbb{C}) \) is the vector space of \( n \times n \) complex matrices. If \( G = \text{SO}(n) \), then \( \text{lie}(G) \) is the vector space of \( n \times n \), anti-symmetric matrices. If \( G = \text{U}(n) \), then \( \text{lie}(G) \) is the vector space of \( n \times n \), complex, anti-hermitian matrices. If \( G = \text{SU}(n) \), then \( \text{lie}(G) \) is the codimension 1 subspace in \( \mathbb{A}(n; \mathbb{C}) \) of trace zero matrices.

The algebra designation for the vector space \( \text{lie}(G) \) signifies the existence of a bilinear map from \( \text{lie}(G) \times \text{lie}(G) \) of a certain sort. This map is called the **Lie bracket** and the Lie bracket of vectors \( \sigma, \tau \in \text{lie}(G) \) is traditionally written as \( [\sigma, \tau] \). It obeys the **Jacobi relation**

\[
[\sigma, [\tau, \mu]] + [\mu, [\sigma, \tau]] + [\tau, [\mu, \sigma]] = 0.
\]

In the examples used in these lecture, \( G \) is given as a subgroup of some \( n \geq 1 \) version of \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \). If \( G \) is a subgroup of some \( n \geq 1 \) version of \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \), then \( \text{lie}(G) \) appears as a vector subspace of the corresponding vector space of real or complex \( n \times n \) matrices. The Lie bracket in this case sets \( [\tau, \sigma] = \tau \sigma - \sigma \tau \). This Lie bracket can be thought of as a measure of the failure of elements in \( G \) to commute because \( e^\sigma e^\tau e^{-\tau} e^{-\sigma} = 1 + [\sigma, \tau] + \varepsilon \) where the remainder, \( \varepsilon \), has norm that is the order of \( |\sigma||\tau|(|\sigma| + |\tau|) \). Note in this regard what is said in Chapters 8e-f: The map \( \sigma \rightarrow e^\sigma \) sends \( \text{lie}(G) \) into \( G \).

If \( G \) is not given as a subgroup of some \( n \geq 1 \) version of \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \), then the Lie bracket can still be defined in the following way: Extend a given pair \( \sigma, \tau \in TG_1 \) as vector fields, \( \nu_\sigma \) and \( \nu_\tau \), that are defined in a neighborhood of the identity of \( G \). View these as derivations of the algebra \( C^\infty(G) \). The commutator of two derivations is also a derivation; in particular, this is the case for the commutator of \( \nu_\sigma \) with \( \nu_\tau \). View the latter derivation as a vector field defined near \( 1 \in G \). Its value at \( 1 \) in \( TG_1 \) is \( [\sigma, \tau] \).

If \( G \) appears as a subgroup of some \( n \geq 1 \) version of \( \text{Gl}(n; \mathbb{R}) \), then its lie algebra inherits the inner product that assigns the trace of \( \sigma^\tau \) to a given pair \( \sigma, \tau \in \text{lie}(G) \). If \( G \subset \text{Gl}(n; \mathbb{C}) \), then the inner product of a given pair \( \sigma \) and \( \tau \) from \( \text{lie}(G) \) is assumed to be given by the real part of the trace of \( \sigma^\tau \). These inner products are used implicitly in what follows to define norms and orthogonal complements. If \( G \) is not given as a subgroup of one of these groups, agree beforehand to fix a convenient inner product on \( TG_1 \).
**Representations of Lie groups:** Let $V$ denote a vector space, either $\mathbb{R}^m$ or $\mathbb{C}^m$. Introduce $\text{Gl}(V)$ to denote either the group $\text{Gl}(m; \mathbb{R})$ or $\text{Gl}(m; \mathbb{C})$ as the case may be. Now suppose that $G$ is a given Lie group. A *representation* of $G$ on $V$ is defined to be a smooth map $\rho: G \rightarrow \text{Gl}(V)$ with the property that $\rho(1) = 1$ and $\rho(gg') = \rho(g)\rho(g')$ for any pair $g, g' \in G$. Suppose that $v \in V$. The *stabilizer* of $v$ is the subset $\{g \in G: \rho(g)v = v\}$.

**Lie algebras and Representations:** Suppose that $\rho: G \rightarrow \text{Gl}(V)$ is a representation. Introduce $\rho_*: \text{lie}(G) \rightarrow \text{lie}(\text{Gl}(V))$ to denote the differential of $\rho$ at the identity. In the case when $G$ is a matrix group, the map $\rho_*$ is given by the rule $\rho_*(q) = \left( \frac{d}{dt} \rho(e^{tq}) \right) \big|_{t=0}$. The map $\rho_*$ is a linear map that intertwines the Lie bracket on $\text{lie}(G)$ with the matrix Lie bracket on $\text{lie}(\text{Gl}(V))$ in the sense that $\rho_*(\sigma, \tau) = [\rho_*(\sigma), \rho_*(\tau)]$ for any given pair $\sigma, \tau \in \text{lie}(G)$. That such is the case follows from the fact that $\rho$ intertwines group multiplication on $G$ with matrix multiplication.

This intertwining property can be verified when $G$ is a subgroup of some $n \geq 1$ version of either $\text{Gl}(n; \mathbb{R})$ or $\text{Gl}(n; \mathbb{C})$ as follows: Let $a, a' \in \text{lie}(G)$. For any given pair $t, s \in \mathbb{R}$, both $e^{ta}$ and $e^{ta'}$ are in $G$. As a consequence, the assignment to $(t, s)$ of $e^{ta}e^{ta'}e^{-ta}e^{-ta'}$ defines a smooth map from a neighborhood of the origin in $\mathbb{R}^2$ to $G$. It follows from what was said above about the commutator that this map has the form $(t, s) \mapsto t + ts [a, a'] + \epsilon$ where $|\epsilon|$ is bounded by some multiple of $|t||s|(|t| + |s|)$. Given that $\rho$ is a representation, it follows that

$$\rho(e^{ta}e^{ta'}e^{-ta}e^{-ta'}) = \rho(e^{ta})\rho(e^{ta'})\rho(e^{-ta})\rho(e^{-ta'}) .$$

Now use Taylor’s expansion with remainder to write this as

$$t + ts\rho_*([a, a']) = t + ts [\rho_*([a], \rho_*([a']) + \epsilon$$

where $|\epsilon|$ is also bounded by a multiple of $|t||s|(|t| + |s|)$. Given such bound for $\epsilon$, what is written above is possible for all $t, s$ near $0 \in \mathbb{R}^2$ only if $\rho_*$ intertwines the Lie bracket on $\text{lie}(G)$ with the commutator bracket on the lie algebra of $\text{Gl}(V)$.

The stage is now set. In the proposition below, $G$ is a Lie group, $V$ is a vector space, either $\mathbb{R}^m$ or $\mathbb{C}^m$ for some $m \geq 1$ and $v \in V$ is a given non-zero element. Meanwhile, $\rho: G \rightarrow \text{Gl}(V)$ is a representation.

**Proposition:** The stabilizer of $v \in V$ is a Lie subgroup $H \subset G$ whose tangent space at the identity is the subspace $\mathfrak{h} = \{q \in \text{lie}(G): (\rho_*q)v = 0\}$. If $G$ is compact, then the following is also true:
• The subspace \( M_v = \{ \rho(g)v : g \in G \} \subset V \) is a smooth manifold, homeomorphic to the quotient space \( G/H \).
• The tangent space to \( M_v \) at \( v \) is canonically isomorphic to the orthogonal complement of \( \mathfrak{h} \) in \( \text{lie}(G) \) with the map that sends any given \( \zeta \) in this orthogonal complement to \( \rho.(\zeta)v \).
• The map \( \pi : G \to M_v \) defines a principal \( H \) bundle.

A proof of this proposition is given in the Appendix of this chapter; but only in the special case when \( G \) is a subgroup of some \( n \geq 1 \) version of either \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \).

10e) Examples of Lie group quotients

This part of the chapter gives some examples of principal bundles that arise using the preceding proposition.

Example 1: This first case concerns \( \text{SO}(n) \). Use the inner product on \( \mathbb{M}(n; \mathbb{R}) \) given by \( \langle m, m' \rangle = \text{trace}(m^Tm') \) to view \( \mathbb{M}(n; \mathbb{R}) \) as a Euclidean space of dimension \( n^2 \). A representation \( \rho : \text{SO}(n) \to \text{SO}(n^2) \subset \text{Gl}(n^2; \mathbb{R}) \) is defined by having \( m \in \text{SO}(n) \) act on \( a \in \mathbb{M}(n; \mathbb{R}) \) via \( m \cdot a \cdot m^{-1} \). Given a non-zero \( a \in \mathbb{M}(n; \mathbb{R}) \), define the subset \( M_a \subset \mathbb{M}(n; \mathbb{R}) \) to be the set of elements of the form \( m \cdot a \cdot m^{-1} \) for \( m \in \text{SO}(n) \). Use \( H \subset \text{SO}(n) \) to denote the subgroup of matrices \( m \) such that \( m \cdot a \cdot m^{-1} = a \). The proposition asserts that \( M_a \) is a smooth manifold and that the the map \( \text{SO}(n) \to M_a \) defines a principal \( H \)-bundle.

For a concrete example, choose \( a \) to be a diagonal matrix with distinct eigenvalues. The group \( H \) in this case is a finite group, being the group of diagonal matrices in \( \text{SO}(n) \). For a somewhat more interesting example, fix a set \( n_1 \leq n_2 \cdots \leq n_k \) of positive integers that sum to \( n \). Take \( a \) to be diagonal with \( m \) distinct eigenvalues, these of multiplicities \( n_1, n_2, \ldots, n_k \). The subgroup \( H \) in this case is isomorphic to the subgroup \( \text{SO}(n_1) \times \cdots \times \text{SO}(n_k) \subset \text{SO}(n_1) \times \cdots \times \text{SO}(n_k) \) consisting of the \( m \)-tuples of matrices for which the corresponding product of determinants is equal to 1.

In the case when \( k = 2 \), the resulting manifold \( M_a \) is diffeomorphic to the Grassmannian \( \text{Gr}(n; n_1) \) of \( n_1 \)-dimensional subspaces in \( \mathbb{R}^n \), this defined in Chapter 1f.

To see why this is, remark that no generality is lost by taking \( a \) to be the matrix where the top \( n_1 \) elements on the diagonal are equal to 1 and the remaining are equal to 0. This understood, define a map \( \Phi : M_a \to \text{Gr}(n; n_1) \) by setting \( \Phi(v) \) for \( v \in M_a \) to be the span of the eigenvectors of \( v \) with eigenvalue 1. To see about an inverse, suppose that \( V \subset \mathbb{R}^n \) is a given \( n_1 \)-dimensional subspace. Let \( \Pi_V \) denote the orthogonal projection to \( V \). This is a self-adjoint matrix with \( n_1 \) eigenvalues equal to 1 and \( n - n_1 \) equal to zero. This matrix is diagonalizable, and so there exists \( m \in \text{SO}(n) \) such that \( m^T \Pi_V m = a \). Thus, \( \Pi_V \in M_a \). The assignment of \( \Pi_V \) to \( V \) inverts the map \( \Phi \). To summarize: The proposition identifies
Gr(n; n_1) with the manifold SO(n)/S(O(n_1) × O(n-n_1)). The principal bundle defined by the quotient map from SO(n) to SO(n)/S(O(n_1) × O(n-n_1)) can be viewed in the context of Gr(n; n_1) as follows: The total space P is SO(n), and the fiber over a given subspace V ⊆ Gr(n; n_1) consists of the set of matrices m ∈ SO(n) that rotate \( R^n \) so as to map V to itself.

In the case n_1 = 1, the space Gr(n; 1) is the real projective space \( \mathbb{RP}^{n-1} \). What is said above thus identifies \( \mathbb{RP}^{n-1} \) with SO(n)/S(O(1) × O(n-1)). Noting that O(1) is the two element group \{±1\}, and noting the identification SO(n)/SO(n-1) = S^{n-1}, the proposition in this case says not much more than the fact that \( \mathbb{RP}^{n-1} = S^{n-1}/\{±1\} \) where \{±1\} acts on \( R^n \) by sending any given vector x to -x.

**Example 2:** A very much analogous example involves U(n). To say more, introduce the Hermitian inner product on M(n; \( \mathbb{C}^n \)) given by \( \langle m, m' \rangle = \text{trace}(m^\dagger m') \) to view M(n; \( \mathbb{C}^n \)) as a copy of \( \mathbb{C}^{n^2} \). This done, the action of U(n) on M(n; \( \mathbb{C}^n \)) that sends a given matrix a to mam^{-1} defines a representation in U(n^2) ⊂ Gl(n^2; \( \mathbb{C}^n \)). If a ∈ M(n; \( \mathbb{C}^n \)) is any given non-zero matrix, then the proposition finds M_a = \{v ∈ M(n; \( \mathbb{C}^n \)): v = mam^{-1} \} is a smooth manifold. Moreover, the set H ⊂ U(n) of matrices m such that mam^{-1} = a is a subgroup, and the map from U(n) to M_a that sends m to mam^{-1} defines a principal H bundle.

A concrete example takes a to be diagonal with some n_1 ∈ \{1, …, n-1\} entries equal to 1 and the remaining equal to 0. As in the SO(n) example above, the space M_a is diffeomorphic to the complex Grassmannian Gr_C(n; n_1) as defined in Chapter 6h. In particular, Gr_C(n; n_1) is diffeomorphic to U(n)/(U(n_1) × U(n-n_1)). When U(n) is viewed as a principal U(n_1) × U(n-n_1) bundle over Gr_C(n; n_1), its fiber over a given complex n_1-dimensional subspace consists of the set of matrices in U(n) that act on \( \mathbb{C}^n \) so as to map the subspace to itself.

Note by the way, that if 1 ≤ n_1 ≤ … ≤ n_k are integers that sum to n, then the quotient U(n)/(\times_{i=1}^{k} U(m)) is the same as SU(n)/S(\times_{i=1}^{k} U(m)) where S(\times_{i=1}^{k} U(m)) is the subgroup of k-tuples in \( \times_{i=1}^{k} U(k) \) whose corresponding product of determinants is equal to 1.

**Example 3:** This case involves SU(2) acting on \( \mathbb{C}^2 \) via matrix multiplication. Take the vector \( v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and set M_v to be the set of vectors of the form mv ∈ \( \mathbb{C}^2 \) with m ∈ SU(2). The group H in this case is U(1) = S^1 and M_v = S^2. To be more explicit, view SU(2) as S^3, the unit sphere in \( \mathbb{C}^2 \). This is done by the usual correspondence that pairs a
given \( \begin{pmatrix} w \\ z \end{pmatrix} \in S^3 \) with the matrix \( \begin{pmatrix} w & -z \\ z & w \end{pmatrix} \). View \( S^1 \) as the subgroup \( U(1) \) of matrices \( \begin{pmatrix} u & 0 \\ 0 & \bar{u} \end{pmatrix} \) with \( |u| = 1 \), then the quotient \( SU(2)/U(1) \) is \( S^2 \). Thus, \( S^2 = SU(2)/U(1) \).

This generalizes to higher dimensions as follows: View \( S^{2n+1} \) as the unit sphere in \( C^{n+1} \). Then \( S^{2n+1} = U(n+1)/U(n) \) and \( U(n+1) \) can be viewed as the total space of a principal \( U(n) \) bundle over \( S^{2n+1} \). Note that there is a residual \( U(1) \) action on \( S^{2n+1} \), this the action that sends a given \( u \in U(1) \) and \( (z_1, \ldots, z_{n+1}) \in S^{2n+1} \) to \( (uz_1, \ldots, uz_n) \). This action is free, and the quotient is \( n \)-dimensional complex projective space \( \mathbb{CP}^n \); this the space of 1-dimensional vector subspaces in \( C^{n+1} \) from Chapter 6. Thus the quotient space projection, \( S^{2n+1} \rightarrow \mathbb{CP}^n \), defines a principal \( U(1) \) bundle. Another view of this bundle is given in Example 5 in the next part of Chapter 10.

10f) Cocycle construction examples

What follows describes some principle bundles that are constructed using the cocycle definition in Chapter 10.

Example 1: Principle \( U(1) \) bundles over \( S^2 \) can be constructed using the cocycle definition as follows: View \( S^2 \) as the set of unit length vectors in \( \mathbb{R}^3 \), and this done, decompose it as \( S^2 = U_+ \cup U_- \) where \( U_+ \) consists of vectors \( x = (x_1, x_2, x_3) \) with \( |x| = 1 \) and \( x_3 > -1 \), and where \( U_- \) consists of vectors with norm 1 and with \( x_3 < 1 \). These two sets intersect in the cylindrical region \( S^2 \) that consists of the points in \( \mathbb{R}^3 \) with norm 1 and with \( x_3 \in (-1, 1) \). This understood, the complex number \( z = x_1 + ix_2 \) is nowhere zero on this region, and so the assignment \( x \rightarrow |z|^{-1} z \) defines a smooth map from \( U_+ \cap U_- \) into the unit circle in \( C \), thus the group \( U(1) \). Let \( g: U_+ \cap U_- \) denote this map.

Now, let \( m \) denote any given integer. A principal \( U(1) \) bundle is defined by taking for cocycle data the open cover \( \mathcal{U} = (U_+, U_-) \) and principal bundle transition function \( g^m: U_+ \cap U_- \rightarrow U(1) \).

There is an analogous construction for any given surface. Let \( \Sigma \) denote the surface in question. Fix a finite set \( \Lambda \subset \Sigma \) of distinct points. Let \( p \in \Lambda \). Fix an open set \( U_p \subset \Sigma \) containing \( p \) and with a diffeomorphism \( \varphi_p: U_p \rightarrow \mathbb{R}^2 \) that sends \( p \) to the origin. Do this for each point in \( \Lambda \), but make sure to choose the sets in question so that their closures are pairwise disjoint. Assign an integer, \( m(p) \) to each \( p \in \Lambda \).

The cover \( \mathcal{U} = \{U_0 = \Sigma - \Lambda\} \cup \{U_p\}_{p \in \Lambda} \) is used for the cocycle data set. A bundle transition function \( g_{0p}: U_0 \cap U_p \rightarrow U(1) \) is defined as follows: If \( x \in U_0 \cap U_p \), write \( \varphi_p(x) \) as a complex number, \( z(x) \). This done, set \( g_{0p}(x) = (|z(x)| z(x))^{m(p)} \). As no three distinct sets from the cover intersect, this data defines a principal \( U(1) \) bundle over \( \Sigma \).
Example 2: Principal SU(2) bundles over $S^4$ can be constructed as follows: Identify the equatorial $S^3$ as SU(2). Now define a principal bundle SU(2) over $S^4$ by declaring that it be isomorphic to the trivial bundle over the complement of the south pole, and also trivial over the complement of the north pole. Use $U$ to denote the former set and $V$ to denote the latter. The intersection of these two sets can be described as follows: View $S^4$ as the set $y = (x_1, \ldots, x_4, x_5) \in \mathbb{R}^5$ with $|y| = 1$. Then $U$ is the set $x_5 > -1$ and $V$ is the set $x_5 < 1$. Thus, their overlap is the set where $x_5 \neq \pm 1$. Define a map from this set to $S^3$ by sending $y$ to the point $g(x) = (x_1, x_2, x_3, x_4)/(1 - x_5^2)^{1/2}$. Identify $S^3$ with SU(2) as done previously. Fix an integer $p$, and define a principal bundle SU(2) bundle $P^p \to S^4$ by declaring its transition function to be $g(x)^p$. As it turns out the bundle $P^{(1)}$ is diffeomorphic to $S^7$. So is $P^{(1)}$, but these bundles are not isomorphic as principal SU(2) bundles.

To see the relation to $S^7$, it is useful to introduce $\mathbb{H}$ to denote the set of $2 \times 2$ complex matrices that can be written as

$$
\begin{pmatrix}
a & -\bar{b} \\
b & \bar{a}
\end{pmatrix}.
$$

If I reintroduce the Pauli matrices $\{\tau_1, \tau_2, \tau_3\}$ from Chapter 6d, then an element in $\mathbb{H}$ can be written as $x_4 + x_1\tau_1 + x_2\tau_2 + x_3\tau_3$. Here, $a = x_4 + ix_3$ and $b = x_2 + ix_1$. Thus, $\mathbb{H}$ can be viewed as $\mathbb{R}^4$, or as $\mathbb{C}^2$. Given the multiplication rule described in Chapter 6d, this last way of writing $\mathbb{H}$ identifies it with the vectors space (and algebra via matrix multiplication) of quaternions.

In any event, the 7-sphere is the set of pairs $(a, b) \in \mathbb{H} \times \mathbb{H}$ with $|a|^2 + |b|^2 = 1$. Here, $|a|^2 = \text{tr}(a^\dagger a)$. The group SU(2) sits in $\mathbb{H}$ as the set of unit vectors. It acts on $S^7$ so that $m \in \text{SU}(2)$ sends $(a, b)$ to $(am^{-1}, bm^{-1})$. It can also act to send $(a, b)$ to $(ma, mb)$. The quotient of either action is $S^4$, and one quotient defined $P^{(1)}$, the other defines $P^{(1)}$.

To see that the quotient is $S^4$, define a map from $S^7$ to $\mathbb{R}^5$ by sending a given vector $(a, b)$ to the vector $(2ab^\dagger, |a|^2 - |b|^2) \in \mathbb{H} \times \mathbb{R} = \mathbb{H}^5$. This vector has unit length, so defines a point $S^4$. This map is constant along orbits of SU(2) by the action that sends $m \in \text{SU}(2)$ and $(a, b) \in S^7$ to $(am^{-1}, bm^{-1})$. As exhibited momentarily, it is the projection to $S^4$ that defines the principal bundle $P^{(1)}$.

The following description of the cocycle data for $P^{(1)}$ verifies that the map just given from $S^7$ to $S^4$ is indeed the principal bundle projection. To start, identify $U_\ast$ with $\mathbb{H}$ via the map that sends any given point $\zeta \in \mathbb{H}$ to $(2\zeta, 1 - |\zeta|^2)/(1 + |\zeta|^2)$. Define the inverse map from $\mathbb{H} \times \text{SU}(2)$ to $S^7$ so as to send $(\zeta, m)$ to $(a = \zeta m, b = m)/(1 + |\zeta|^2)^{1/2}$. Over the set $U_\ast$, define a map from $\mathbb{H}$ to $U_\ast$ by sending $r \in \mathbb{H}$ to $(2r^\dagger, |r|^2 - 1)/(1 + |r|^2)^{1/2}$.

118
Now define a map from $\mathbb{H} \times \text{SU}(2)$ to $S^7$ by sending $(r, n)$ to $(a = n, b = rn)/(1 + |r|^2)^{1/2}$. The equator is the set $|z| = |r| = 1$ and $\tau = \frac{1}{2}$. Thus, the points $m$ and $n$ are sent to the same point if $n = zm$. This is the transition function for $\mathbb{P}^{(1)}$.

There is a very much analogous story for $\mathbb{P}^{(2)}$ that identifies the bundle projection with the map from $S^7$ that sends $(a, b)$ to $(2a^*b, |a|^2 - |b|^2)$.

Here is a parenthetical remark: One can define a fiber bundle over $S^4$ with fiber $\text{SU}(2)$ (but not a principal bundle) as follows: Choose integers $(p, q)$ and declare the space $X_{p,q}$ to be diffeomorphic over $U_+$ to $U_+ \times \text{SU}(2)$ and also diffeomorphic over $U_-$ to $U_- \times \text{SU}(2)$. Over the intersection of these sets, define the transition function so as to send $(x, m) \in U_+ \times \text{SU}(2)$ to $U_- \times \text{SU}(2)$ using the rule $(x, m) \rightarrow (x, g(x)^p m g(x)^q)$. This is a principal bundle if and only if one of $p$ or $q$ is zero. Now restrict to the case where $p+q = 1$, and set $k = p-q$. John Milnor proved that $X_{p,q}$ in this case is homeomorphic to $S^7$ but not diffeomorphic to $S^7$ if $k^2 - 1 \neq 0 \mod(7)$. In fact, each residue class $\mod(7)$ gives a different smooth structure on $S^7$. As it turns out, there are 28 different smooth structures on $S^7$. All can be realized as follows: Intersect the unit sphere about the origin in $\mathbb{C}^5$ with the submanifold in $\mathbb{C}^5 - \{0\}$ where the complex coordinates $(z_1, \ldots, z_5)$ obey the equation

$$z_1^2 + z_2^2 + z_3^2 + z_4^3 + z_5^{6k-1} = 0$$

for $k \in \{1, 2, \ldots, 28\}$. (These are known as Brieskorn spheres.) Note that this submanifold is a complex manifold, and that it intersects $S^9$ transversely.


**Example 3:** Let $M$ denote any given 4-dimensional manifold. Choose a finite set, $\Lambda$, of distinct points in $M$ and associate to each an integer. For each $p \in \Lambda$, fix a coordinate chart $U$, centered at $p$, with coordinate map $\phi: U \to \mathbb{R}^4$. Take $U$ so that charts given by distinct points from $\Lambda$ are disjoint. Define a principal $\text{SU}(2)$ bundle $P \to M$ as follows: Set $P_{M,\Lambda}$ to be isomorphic to $(M-\Lambda) \times \text{SU}(2)$. For each $p \in \Lambda$, set $P_{U,p} = U \times \text{SU}(2)$ also. To define the transition functions, first agree to identify the unit sphere in $\mathbb{R}^4$ with $S^3$. This done, let $q$ denote the integer associated to the point $p$, and then the transition function, $g_{U,M-\Lambda}$ to be the map that sends $x \in U \cap (M-p) = U-p$ to $(\phi(x)/|\phi(x)|)^q \in \text{SU}(2).$
Example 4: Something along the lines of what is done in Example 3 can be done any time you have an interesting map from \(S^{n-1}\) to \(SO(k)\), \(U(k)\), or some other group \(G\). Let \(f\) denote your map. Pick a set \(\Lambda\) of distinct points in \(M^n\), and associate to each an integer. Choose for each point a coordinate chart as in the previous example. Define the bundle to be trivial over \(M - \Lambda\) and over each coordinate chart. If \(p \in \Lambda\), the transition function for \(U \cap (M - p) = U - p\) sends \(x\) to \(f(x)^q\). For example, recall from Chapter 6f the definition of the matrices \(\{\gamma_\alpha\}_{\alpha=1,\ldots,5}\) in \(M(4; \mathbb{C})\). Define a map from \(R^5\) to \(M(4; \mathbb{C})\) by sending a given point \(x = (x_1, \ldots, x_5)\) to \(f(x) = \sum_j x_j \gamma_j\). Note that \(f^* f = |x|^2\), so \(f\) maps \(S^n\) to \(U(4)\).

Example 5: This example returns to the principal \(U(1)\) bundle that is defined by the projection from \(S^{2n+1}\) to \(\mathbb{CP}^n\) as described at the end of the previous part of this chapter. Recall from Chapter 6h that \(\mathbb{CP}^n\) has an open cover by \(n+1\) sets with each set in the cover, \(\mathcal{U} = \{O_1, \ldots, O_{n+1}\}\) diffeomorphic to \(\mathbb{C}^n\). By way of reminder, any given point \((z_1, \ldots, z_n) \in O_k = \mathbb{C}^n\) parametrizes the complex line in \(\mathbb{C}^{n+1}\) that is spanned (over \(\mathbb{C}\)) by the vector with \(k\)'th entry is 1, whose \(i\)'th entry for \(i < k\) is \(z_i\), and whose \(i\)'th entry for \(i > k\) is \(z_{i-1}\).

This understood, the intersection between \(O_k\) and \(O_{k'}\) consists of the complex lines in \(\mathbb{C}^{n+1}\) spanned by vectors of the form \((p_1, \ldots, p_{n+1})\) with both \(p_k\) and \(p_{k'}\) non-zero. This understood, a map from \(O_k \cap O_{k'}\) to \(U(1)\) assigns to any such line the complex number \(g_{kk'} = |p_k|^{-1} p_k / p_{k'}\). The collection \(\{g_{kk'}\}_{1 \leq k < n+1}\) defines principal \(U(1)\) bundle transition functions for the open cover \(\mathcal{U}\). To see that such is the case, it is enough to verify that the cocycle condition is obeyed on any given triple intersection. This understood, suppose that \(k, k', k''\) are distinct. Then \(g_{kk''} g_{k'k''} g_{k''k'}\) is observably \(1\) as required.

By the way, the identification between the trivial bundle \(O_k \times U(1)\) with a subset of \(S^{2n+1}\) is as follows: Use \(z = (z_1, \ldots, z_n)\) as before for the coordinates on \(O_k\). Let \(u \in U(1)\). The point \((z, u) \in O_k \times U(1)\) gives the point whose \(k\)'th coordinate is \((1 + |z|^2)^{-1/2} u\), whose \(i\)'th coordinate for \(i < k\) is \((1 + |z|^2)^{-1/2} z_i u\), and whose \(i\)'th coordinate for \(i > k\) is given by \((1 + |z|^2)^{-1/2} z_{i-1} u\).

Example 6: This example constructs what are known as quaternionic projective spaces. To start, fix a positive integer, \(n\). Consider \(\mathbb{R}^{4(n+1)}\) as \(\mathbb{H}^{n+1}\). Let \(S^{4n+3}\) denote the unit sphere in this Euclidean space; and define an action of \(SU(2)\) on \(S^{4n+2}\) by sending any give \(m \in SU(2)\) and \((a_1, \ldots, a_{n+1}) \in S^{4n+3}\) to \((a_1 m^{-1}, \ldots, a_{n+1} m^{-1})\). The space of orbits, \(S^{4n+3}/SU(2)\) is a smooth manifold which is usually denoted by \(\mathbb{HP}^n\), the quaternionic projective space. It has dimension \(4n\). The quotient map defines a principal \(SU(2)\) bundle. The proof that such is the case copies in an almost verbatim fashion what is said.
in the preceding example about the principle U(1) bundle $S^{2n+1} \to \mathbb{C}P^n$. The details are left to you the reader to work out.

10g) **Pull backs of principal bundles**

Let $N$ and $M$ denote a pair of smooth manifolds, and let $f: M \to N$ denote a smooth map. As explained in Chapter 5a, the map $f$ can be used to pull-back a vector bundle over $N$ to obtain a vector bundle over $M$. Principal bundles can also be pulled back using $f$. To elaborate, suppose that $G$ is a compact Lie group and $\pi: P \to N$ is a principal $G$ bundle. The pull-back bundle $f^*P$ is a principal bundle over $M$ that is defined as follows: It sits in $M \times P$ as the subset of pairs $(x, p)$ with $f(x) = \pi(p)$. Here is a neat way to see $f^*P$ as a manifold: Let $\Delta_N \subset N \times N$ denote the diagonal, thus the set of pairs $(y, y)$ with $y \in N$. This is a submanifold. Map $M \times P$ to $N \times N$ using the map that sends a pair $(x, p)$ to $(f(x), \pi(p))$. It is a straightforward exercise to verify that this map is transversal to the diagonal in the sense used in Chapter 5g. This understood, the proposition in Chapter 5g guarantees that that the inverse image of $\Delta_N$ in $M \times P$ is a submanifold. This inverse image is $f^*P$.

To see that $f^*P$ meets the principal bundle requirements listed in Chapter 10a, note to start that $G$ acts so that any given $g \in G$ sends a point $(x, p) \in f^*P$ to $(x, pg^{-1})$. Meanwhile, the map $\pi$ to $M$ sends $(x, p)$ to $x$. The last of the three requirements listed in Chapter 10a is the following: Given $x \in M$, needed is an open set, $U \subset M$, that contains $x$ and comes with a $G$-equivariant diffeomorphism from $P|_U$ to $U \times G$ that intertwines $\pi$ with the projection from $U \times G$ to $U$. To obtain this data, go to $N$ and select an open set $V \subset N$ that contains $f(x)$ and comes with a $G$-equivariant diffeomorphism, $\varphi_V$, from $P|_V$ to $V \times G$ that intertwines $\pi$ with the projection from $P|_V$ to $V$. Let $\psi_V: P|_V \to G$ denote the composition of first $\varphi_V$ and then projection to $G$. Now take $U = f^{-1}(V)$ and set $\varphi_U: f^*P|_U \to U \times G$ to be the map that sends a pair $(x, p)$ to $(x, \psi_V(p))$.

It is perhaps redundant to say that the bundle $f^*P$ comes with the tautological, $G$-equivariant map to $P$, this the map that sends a given pair $(x, p)$ to $p$.

**Example 1:** Chapter 10a declared a pair of principal $G$-bundles $\pi: P \to M$ and $\pi': P' \to M$ to be isomorphic if there exists a $G$-equivariant diffeomorphism $f: P \to P'$ such that $\pi = \pi' \circ f$. Keeping this in mind, suppose that $\pi: P \to M$ defines a principal $U(1)$ bundle. As it turns out, there exists an integer $n \geq 1$, a map $f: M \to \mathbb{C}P^n$, and an isomorphism between $P$ and the pull-back $f^*S^{2n+1}$ with $S^{2n+1}$ viewed here as a principal $U(1)$ bundle over $\mathbb{C}P^n$; this as described by the fifth example in preceding part of Chapter 10. What follows gives a construction of this data when $M$ is compact.

To start, fix a finite cover, $\{U_i\}$ of $M$ such that each set $U$ from $\{U_i\}$ comes with a bundle isomorphism $\varphi_U: P|_U \to U \times U(1)$. Let $\psi_U: P \to U(1)$ denote the composition of
first $\varphi_U$ and then the projection to the $U(1)$ factor. By refining this cover, I can assume that each open set is also a coordinate chart, and so there is a subordinate partition of unity. Recall from the appendix to Chapter 1 that this is a collection, $\{\chi_U\}_{U \in \mathcal{U}}$, of non-negative functions such that any given $U \in \mathcal{U}$ version of $\chi_U$ has support only in $U$, and such that $\sum_{U \in \mathcal{U}} \chi_U = 1$ everywhere. Set $h_U(\cdot)$ to denote the function on $P$ that is obtained by composing $\pi$ with $\chi_U(\sum_{U \in \mathcal{U}} \chi_U)^{1/2}$. Write the number of sets in the cover $\mathcal{U}$ as $n+1$, and label the sets in $\mathcal{U}$ from 1 to $n+1$.

Define now a map, $\hat{f}$, from $P$ to $S^{2n+1}$ to the point in $S^{2n+1} \subset \mathbb{C}^{n+1}$ whose $k$'th coordinate is zero on the complement of $U_k$ and equal to $h_{U_k} \psi_{U_k}$ on $U_k$. This map is $U(1)$ equivariant, and so the quotient of the image by $U(1)$ defines a map, $f$, from $M$ to $\mathbb{C}P^n$. The bundle isomorphism between $P$ and $f^*S^{2n+1}$ sends $p \in P$ to $(\pi(p), \hat{f}(p)) \in M \times S^{2n+1}$.

**Example 2:** The sixth example from the preceding part of this Chapter describes the quaternionic projective spaces $\{\mathbb{H}P^n\}_{n=1,2,\ldots}$; and it explains how to view the map from $S^{4n+3}$ to $\mathbb{H}P^n$ as a principal $SU(2)$ bundle. Let $\pi: P \to M$ denote a given principal $SU(2)$ bundle. A construction that is very much like that in the preceding example will give an integer, $n \geq 1$, a map $f: M \to \mathbb{H}P^n$, and a principal bundle isomorphism between $P$ and the pull-back bundle $f^*S^{4n+3}$.

**Example 3:** Let $N$ denote a smooth manifold and suppose that $E \to N$ is a vector bundle with fiber dimension $n$, either real or complex. As explained in Chapter 10c, one can construct from $E$ the principal bundle of frames for $E$. This is a principal bundle with group $G$ either $GL(n; \mathbb{R})$ or $GL(n; \mathbb{C})$ as the case may be. If $E$ also has a fiber metric, one can use the metric to construct the principal $G = O(n)$, $SO(n)$, $U(n)$ or $SU(n)$ bundle depending on the circumstances. I use $P_E$ in what follows to denote any one of these and use $G$ to denote the relevant Lie group. Now let $M$ denote a second smooth manifold and suppose that $f: M \to N$ is a smooth map. As explained in Chapter 5a, the map $f$ can be used to pull $E$ back so as to give a vector bundle, $f^*E$, over $M$. Metrics on bundles pull-back as well, the norm of a vector $(x, v) \in f^*E \subset M \times E$ being $|v|$. If $E$ has a metric, then give $f^*E$ this pull-back fiber metric. In any event, one has the principal $G$ bundle $P_{f^*E}$ over $M$. Meanwhile, principal $G$-bundles pull-back as explained above, so one also has the principal $G$-bundle $f^*P_E$. I hope it is no surprise that $P_{f^*E}$ is the same as $f^*P_E$. Those who doubt can unravel the definitions and so verify the assertion.

**10h) Reducible principal bundles**

Suppose that $G$ is a Lie group, $H \subset G$ is a subgroup, and $P \to M$ is a principal $H$ bundle. A principal $G$ bundle over $M$ can be constructed from $P$ as follows:
\[ P_G = (P \times G)/H = P \times_G G, \]

where the equivalence relation identifies pairs \((p, g)\) with \((ph^{-1}, hg)\). Multiplication by an element \(g \in G\) sends the equivalence class \((p, g)\) to that of \((pg^{-1}, g)\). To see that this is a principal bundle (and a manifold), go to any given set \(U \subset M\) where \(P|_U\) admits an isomorphism, \(\varphi\), with \(U \times H\). This isomorphism identifies \(P_G\) with \(U \times (H \times G)/H\) which is \(U \times G\). If \(g_{U_U} : U \cap U' \to H\) is the transition function for \(P\) over an intersection of charts, then this is also the transition function for \(P_G\) with the understanding that \(H\) is to be viewed as a subgroup in \(G\).

A principal \(G\) bundle \(P' \to M\) that is isomorphic to \(P \times_G G\) with \(P \to M\) some principal \(H\) bundle and with \(H \subset G\) a proper subgroup is said to be reducible. For example, if \(E \to M\) is a complex vector bundle with fiber \(\mathbb{C}^n\) with Hermitian metric, then one can construct the \(U(n)\) principal bundle \(P_{U(E)} \to M\). This bundle is reducible to a principal \(SU(n)\) bundle if the vector bundle \(\wedge^n E = \det(E)\) is trivial.

Every bundle has a fiber metric, and this implies that every principal \(GL(n; \mathbb{R})\) bundle is reducible to a principal \(O(n)\) bundle. (If the transition functions can be chose so as to have positive determinant, then the bundle is reducible to an \(SO(n)\) bundle.) By the same token, any given principal \(GL(n; \mathbb{C})\) bundle is reducible to a principal \(U(n)\) bundle. This is because the metric can be used to find local trivializations for the bundle whose transition functions map to \(O(n)\) or \(U(n)\) as the case may be.

Here is another example: If a \(2n\) dimensional manifold \(M\) admits an almost complex structure, \(j\), then its principal \(GL(2n; \mathbb{R})\) frame bundle is reducible to a \(GL(n; \mathbb{C})\) bundle. Here, \(GL(n; \mathbb{C}) \subset GL(2n; \mathbb{R})\) is identified with the subgroup, \(G_j\), of matrices \(m\) such that \(mj = j_0m\) where \(j_0\) is the standard almost complex structure on \(\mathbb{R}^{2n}\) as defined in Chapter 2d.

10i) Associated vector bundles

Chapter 10c explains how to construct a principal bundle from a given vector bundle. As explained here, this construction has an inverse of sorts. To set the stage, suppose that \(G\) is a Lie group and \(\pi : P \to M\) is a principal \(G\) bundle. The construction of a vector bundle from \(P\) requires as additional input a representation of the group \(G\) into either \(GL(n; \mathbb{R})\) or \(GL(n; \mathbb{C})\). Use \(V\) here to denote either \(\mathbb{R}^n\) or \(\mathbb{C}^n\) as the case may be; and use \(GL(V)\) to denote the corresponding general linear group, either \(GL(n; \mathbb{R})\) or \(GL(n; \mathbb{C})\). Let \(\rho\) denote the representation in question. Recall from Chapter 10d that \(\rho\) is a smooth map from \(G\) to \(GL(V)\) two special properties: First, \(\rho(1)\) is the identity in the matrix group \(GL(V)\). Second, \(\rho(gg') = \rho(g)\rho(g')\) for any given pair of elements \(g, g'\) in \(G\).

The corresponding vector bundle has fiber \(V\) and is denoted by \(P \times_G V\). It is the quotient of \(P \times V\) by the equivalence relation equates a given pair \((p, v)\) to all pairs of the form \((pg^{-1}, \rho(g)v)\) with \(g \in G\). Said differently, \(G\) acts on \(P \times V\) so that any given \(g \in G\)
sends any given pair $(p, v)$ to $(pg^{-1}, \rho(g)v)$. The space $P \times_\rho V$ is the space of $G$-orbits via this action.

To verify that this is, indeed, a vector bundle, note first that the projection to $M$ is that sending the equivalence class of a given $(p, v)$ to $\pi(p)$. The action of $\mathbb{R}$ or $\mathbb{C}$ has a given real or complex number $z$ sending $(p, v)$ to $(p, zv)$. The zero section, $\delta$, is the equivalence class of $(p, 0) \in P \times V$. To see the local structure, let $x \in M$ denote any given point and let $U \subset M$ denote an open set with a principal bundle isomorphism, $\varphi$, from $P|_U$ to $U \times G$. Use $\psi: P \to G$ to denote the composition of first $\varphi$ and then projection from $U \times G$ to $G$. Now define $\varphi_V: (P \times_\rho V)|_U \to U \times V$ to be the map that sends the equivalence class of $(p, v)$ to $(\pi(p), \rho(\psi(p))v)$. Note that this map is 1-1 and invertible; its inverse sends $(x, v) \in U \times V$ to the equivalence class of $\varphi^{-1}(x, t), v)$.

This same map $\varphi_V$ defines the smooth structure for $(P \times_\rho V)|_U$. I claim that this smooth structure is compatible with that defined for $(P \times_\rho V)|_{U'}$ when $U$ and $U'$ are intersecting open sets. To verify this claim, let $\varphi: P|_U \to U \times G$ and $\varphi': P|_{U'} \to U' \times G$ denote the associated principal bundle isomorphisms. The composition $\varphi' \circ \varphi^{-1}$ on maps $(U \cap U') \times G$ as a map, $(x, g) \to (x, g_{U\cap U}(x)g)$ where $g_{U\cap U}$ maps $U \cap U'$ to $G$. The corresponding vector bundle transition function $\varphi_V \circ \varphi^{-1}$ sends $(x, v)$ to $(x, \rho(g_{U\cap U}(x))v)$, and this, of course, a smooth, invertible map that is linear over $\mathbb{R}$ or $\mathbb{C}$ as the case may be.

What follows are examples of associated vector bundles.

**Example 1:** If $E \to M$ is a given vector bundle, with fiber $V = \mathbb{R}^n$ or $\mathbb{C}^n$, then we defined in Chapter 10c the bundle $P_{\text{Gl}(E)}$ which is a principal $G = \text{Gl}(n; \mathbb{R})$ or $G = \text{Gl}(n; \mathbb{C})$ bundle as the case may be. Let $\rho$ now denote the defining representation of $G$ on $V$. Then $P_{\text{Gl}(E)} \times_\rho V$ is canonically isomorphic to $E$. Of course, they have the same transition functions, so they are isomorphic. Another way to see this is to view $P_{\text{Gl}(E)} \subset \mathcal{O}_n E$ as the set of elements $(e_1, \ldots, e_n)$ that span the fiber of $E$ at each point. Define $f: (P_{\text{Gl}(E)} \times V) \to E$ by sending $e = (e_1, \ldots, e_\rho)$ and $v$ to $f(e, v) = \sum_j v_j e_j$. This map is invariant with respect to the $G$ action on $P_{\text{Gl}(E)} \times V$ whose quotient gives $P_{\text{Gl}(E)} \times_\rho V$, and so the map descends to the desired bundle isomorphism between $P_{\text{Gl}} \times_\rho V$ and $E$.

If $V = \mathbb{R}^n$ and $E$ has a fiber metric, or if $V = \mathbb{C}^n$ and $E$ has a hermitian metric, we defined in Chapter 10c the respective principal $\text{SO}(n)$ or $\text{U}(n)$ bundles $P_{\text{O}(E)}$ and $P_{\text{U}(E)}$. Let $\rho$ denote the standard inclusion homomorphism from $\text{SO}(n)$ into $\text{Gl}(n; \mathbb{R})$ denote the standard inclusion homomorphism when $V = \mathbb{R}^n$, and let it denote the corresponding homomorphism from $\text{U}(n)$ into $\text{Gl}(n; \mathbb{C})$ when $V = \mathbb{C}^n$. Reasoning as in the previous paragraph leads to the conclusion that $P_{\text{O}(E)} \times_\rho \mathbb{R}^n = E$ in the real case, and $P_{\text{U}(E)} \times_\rho \mathbb{C}^n = E$ in the complex case.
Example 2: This example describes a sort of converse to what is said in the previous example. To set the stage, let $V = \mathbb{R}^n$ or $\mathbb{C}^n$. In the real case, use $G$ to denote either $\text{Gl}(n; \mathbb{R})$ or $\text{O}(n)$; and in the complex case, use $G$ to denote either $\text{Gl}(n; \mathbb{C})$ or $\text{U}(n)$. In any case, use $\rho$ to denote the standard representation of the relevant group acting on $V$. Suppose that $P \to M$ is a principal $G$ bundle, and let $E = P \times_{\rho} V$. In the real case, $P$ is then isomorphic to $P_{\text{Gl}(E)}$ or $P_{\text{SO}(E)}$ as the case may be. In the complex case, $P$ is isomorphic to $P_{\text{GL}(E)}$ or $P_{\text{U}(E)}$.

Now recall from Chapters 5b in the real case that there exists $m > n$ and a smooth map from $M$ to the Grassmannian $\text{Gr}(m; n)$ with the property that its pull-back of the tautological $\mathbb{R}^n$ bundle is isomorphic to $E$. Chapter 6k tells a similar story in the complex case: There exists $m > n$ and a map from $M$ to $\text{Gr}_{\mathbb{C}}(m; n)$ whose pull-back of the tautological $\mathbb{C}^n$ bundle is isomorphic to $E$. Keeping these facts in mind, let $\text{Gr}$ denote $\text{Gr}(m; n)$ in the real case and $\text{Gr}_{\mathbb{C}}(m; n)$ in the complex case, and let $f: M \to \text{Gr}$ denote the relevant map.

Introduce $E_{\tau} \to \text{Gr}$ to denote the aforementioned tautological $G$ bundle, and use $P_{\tau} \to \text{Gr}$ to denote the bundle of frames in $E_{\tau}$. Take these to be orthonormal or unitary in the case that $G$ is $\text{SO}(n)$ or $\text{U}(n)$ respectively. Then the map that $f$ also pulls back $P_{\tau}$, and $f^*P_{\tau} = P_{\text{Gl}(f^*E_{\tau})}$ or $P_{\text{O}(f^*E_{\tau})}$ or $P_{\text{U}(f^*E_{\tau})}$ as the case may be. It follows as a consequence that the original bundle $P$ is isomorphic to the appropriate version of $f^*P_{\tau}$. Here is a formal statement of this last observation:

**Proposition:** Let $G$ denote either $\text{Gl}(n; \mathbb{R})$ or $\text{SO}(n)$ in the real case, and either $\text{Gl}(n; \mathbb{C})$ or $\text{U}(n)$ in the complex case. Let $M$ denote a smooth manifold and let $P \to M$ denote a principal $G$ bundle. There exists $m \geq n$ and a smooth map from $M$ to the Grassmannian $\text{Gr}$ whose pull-back of the principal $G$ bundle $P_{\tau} \to \text{Gr}$ is isomorphic to $P$.

Example 3: The neat thing about this associated vector bundle construction is that all bundles that are constructed from $E$ via various algebraic operations on the fiber vectors space $V = \mathbb{C}^n$ or $\mathbb{R}^n$, such as direct sum, tensor product, dualizing, $\text{Hom}$, etc, can be viewed as arising from $P_{\text{Gl}(E)}$ through an associated bundle construction. For example, $\bigotimes_n E$, $\wedge^p E$, $\text{Sym}^p(E)$, $E^* = \text{Hom}(E; \mathbb{R} \text{ or } \mathbb{C})$, etc correspond to representations of $\text{Gl}(n; \mathbb{R})$ or $\text{Gl}(n; \mathbb{C})$ into various vector spaces. Thus, all of the latter bundles can be studied at once by focussing on the one principal bundle $P_{\text{Gl}(E)}$.

One example is as follows: Let $g$ be a Riemannian metric on a manifold $M$, and so that there exists the principal $\text{O}(n)$ bundle $P_{\text{O}(\text{TM})} \to M$ of oriented, orthonormal frames in $\text{TM}$. Let $\rho$ denote the representation of $\text{O}(n)$ on the vector space $\text{A}(n; \mathbb{R})$ of $n \times n$ anti-symmetric matrices that has $g \in \text{O}(n)$ act so as to send $a \in \text{A}$ to $\rho(g)a = gag^{-1}$. This
representation is called the *adjoint* representation, and the associated vector bundle is isomorphic to $\wedge^2 T^*M$.

A second example along these same lines concerns the bundle $\wedge^n T^*M$. The latter is associated to $P_{O(TM)}$ via the representation of $O(n)$ in the two group $O(1) = \{-1, +1\}$ that sends a matrix $g$ to its determinant.

Vector bundles that are associated to the orthonormal frame bundle of a Riemannian manifold are called *tensor* bundles, and a section of such a bundle is said to be a *tensor* or sometimes a *tensor field*.

**Appendix: Proof of the proposition in Chapter 10d**

Assume below that $G$ is a subgroup of some $n \geq 1$ version of either $GL(n; \mathbb{R})$ or $GL(n; \mathbb{C})$.

To see that $H$ is a subgroup, note that if $h$ and $h'$ are in $H$, then $\rho(h'h')v$ is the same as $\rho(h)\rho(h')v = \rho(h)v = v$. To see that $h^{-1} \in H$, note that if $\rho(\iota)v = v$, and because $\iota = h^{-1}h$, so $\rho(h^{-1}h)v = v$. As $\rho(h^{-1}h) = \rho(h')\rho(h)$, this means that $\rho(h^{-1})\rho(h)v = v$. If $\rho(h)v = v$, then this last identity requires that $\rho(h^{-1})v = v$ also.

To see that $H$ is a submanifold and a Lie group, it is sufficient to give a local coordinate chart near any given element. To start, remark that a chart for a neighborhood of the identity $\iota$ supplies one for a neighborhood of any other element in $H$. Here is why: If $U \subset H$ is an open set containing $\iota$ with a diffeomorphism $\varphi: U \rightarrow \mathbb{R}^m$ for some $m$, then the set $U_h = \{h': h'h^{-1} \in U\}$ is a chart near $h$. This understood, the diffeomorphism $\varphi_h$ that sends $h'$ to $\varphi(h'h^{-1})$ gives the desired diffeomorphism for $U_h$.

Granted all of this, what follows constructs coordinates for some neighborhood of $\iota$. To do this, introduce $\mathfrak{h} \subset \text{lie}(G)$ to denote the kernel of the differential of $\rho$ at the identity. This is a Euclidean space whose dimension depends, in general, on $v$. Let $m$ denote this dimension. As is argued next, the exponential map $h \rightarrow e^h$ restricts $\mathfrak{h}$ so as to map the latter into $H$. For a proof, fix a large integer $N >> 1$ and write $e^h = e^{h/N} \cdots e^{h/N}$ as the product of $N$ factors. This allows $\rho(e^h)v$ to be written as $\rho(e^{h/N}) \cdots \rho(e^{h/N})v$. Meanwhile, Taylor’s theorem with remainder finds that

$$\rho(e^{h/N})v = \rho(\iota)v + \frac{1}{N} \rho_*(h)v + \epsilon_1 = v + \epsilon_1.$$

where $|\epsilon| \leq c_0 N^{-2}$. This understood, it follows by iterating this last result $N$ times that

$$\rho(e^h)v = v + \epsilon_N$$

where $|\epsilon_N| \leq c_0 N^{-1}$. As $N$ can be as large as desired, this means that $\rho(e^h)v = v$.  

126
It follows from what was just said that the map \( h \to e^h \) embeds a ball about 0 in \( \mathcal{F} \) into \( H \). What follows argues that this embedding restricts to some ball about 0 in \( \mathcal{F} \) so as to map onto a neighborhood of 1. Note in the mean time that this result implies that the map \( h \to e^h \) gives local coordinates for a neighborhood of the identity in \( H \). It also implies that \( H \) is a totally geodesic submanifold of \( G \) when given the metric that is induced from the left invariant metric on \( G \) that comes from the latter’s identification as a submanifold of \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \) as the case may be.

To see that this map \( h \to e^h \) is onto a neighborhood of the identity, suppose \( h \in H \) is near the identity. As noted in Chapter 8g, this element \( h \) can be written in any event as \( e^n \) for some \( n \in \text{Lie}(G) \). Write \( m = h + z \) with \( h \in \mathcal{F} \) and with \( z \) orthogonal to \( \mathcal{F} \) and \( |m| \) small. Keep in mind that with \( |m| \) being small, then both \( |h| \) and \( |z| \) are small as both are less than \( |m| \). This understood, note that \( e^{h+z} = e^h e^z + e \) where \( |e| \leq c_0 |h||z| \). It follows as a consequence, that \( \rho(e^{h+z})v = \rho(e^h)v + v` \) with \( |v`| \leq c_0 |h||z| \). Meanwhile,

\[
\rho(e^h) = 1 + \rho_z l_3 + O(|z|^2)
\]

and so \( \rho(e^h)v = v + (\rho_z l_3) v + v` \) where \( |v`| \leq c_0 |m||z| \). Given that \( u \) is orthogonal to \( \mathcal{F} \), it follows that \( |\rho(e^h)v - v| \geq c_i |g||v| \) when \( |m| \) is small. Given that \( h \) is in \( H \), so it follows that \( z = 0 \) if \( h \) is in some neighborhood of 1.

A suitable cover of \( M_v \) by charts will give it a manifold structure. Keep in mind when defining the charts that any given element in \( M_v \) is mapped to \( v \) by the action of \( G \) on \( V \). As a consequence, it is sufficient to construct a local chart near \( v \). Indeed, if \( U \subset M_v \) is a neighborhood of \( v \) with a homeomorphism \( \varphi: U \to \mathbb{R}^d \) for some \( d \), and if \( v` = \rho(g)v \) for some \( g \in G \), then \( U_{v`} = \{ b \in M_v : \rho(g^{-1})b \in U \} \) is a neighborhood of \( v` \). Moreover, the map \( \varphi_v \) that sends \( b \in U_{v`} \) to \( \varphi(\rho(g^{-1})b) \) is then a local homeomorphism from \( U_{v`} \) to \( \mathbb{R}^d \).

Let \( \mathcal{F}^\perp \subset \text{Lie}(G) \) denote the orthogonal complement of \( \mathcal{F} \). The chart for a neighborhood of \( v \) with its homeomorphism is obtained by defining the inverse, this a homeomorphism, \( \psi \), from a ball in \( \mathcal{F}^\perp \) to a neighborhood of \( v \) that sends the origin to \( v \). This homeomorphism sends \( z \in \mathcal{F}^\perp \) to \( \psi(z) = \rho(e^z)v \). This map is 1-1 on a small ball about the origin. This because \( e^z - e^z = z - z` + \tau \) with \( |\tau| \leq c_0 (|z| + |z`|) |v - v`| \). Thus,

\[
\rho(e^z) - \rho(e^z`) = \rho_z l_3 - \rho_z l_3` + \tau` \text{ where } \tau` \leq c_0 (|z| + |z`|) |v - v`| \cdot
\]

As a consequence, \( \rho(e^z)v = \rho(e^z`)v \) if and only if \( z - z` \in \mathcal{F} \) when both \( |z| \) and \( |z`| \) are small. This is possible only if \( z - z` = 0 \).

To see why this map is onto a neighborhood of \( v \) in \( M_v \), suppose for the sake of argument that there is a sequence \( \{ v_k \}_{k=1,2,\ldots} \subset M_v \) that converges to \( v \), but is such that no element from this sequence can be written as \( \rho(e^z)v \) with \( z \in \mathcal{F}^\perp \) having small norm. By
definition, each $v_k$ can be written as $\rho(g_k)v$ with $g_k \in G$. Moreover, there exists $\delta > 0$ such that each $g_k$ has distance at least $\delta$ from any point in $H$. Here, the metric is that used in Chapter 8g. To see why this is, note that if such were not the case, then $g_k$ could be modified by multiplying on the right by the closest element in $H$ so that the result is close to the origin in $G$. The result could then be written as $\exp(\mathfrak{h})\exp(\mathfrak{g})$ with $\mathfrak{h} \in \mathfrak{f}$ and $\mathfrak{g} \in \mathfrak{f}^\perp$; and this would exhibit $v_k$ as $\rho(e^z)v$. Meanwhile, the fact that $G$ is compact implies that the sequence $\{g_k\}_{k=1,2,\ldots}$ has a convergent subsequence. Let $g$ denote the limit. Given that the elements in the sequence $\{g_k\}$ are uniformly far from $H$, this matrix $g$ is not in $H$. Even so $\rho(g)v = v$. As these last two conclusions contradict each other, there is no sequence $\{v_k\}_{k=1,2,\ldots}$ as described above. Note that this is the only place where the compactness of $G$ is used.

To finish the proof of the proposition, consider the claim in the last bullet of the proposition. This claim follows with a proof that the map from $G$ to $M_v$ has the property that is described by the third bullet in Chapter 10a’s definition of a principal bundle. This understood, note that it is sufficient to verify the following: There is an open set $U \subset M_v$ containing $v$ with a diffeomorphism $\varphi: G|_U \to U \times H$ that is suitably equivariant. To prove that such is the case, fix a small radius ball $B \subset \mathfrak{f}^\perp$ on which the map $\mathfrak{z} \to \rho(e^z)v$ is a diffeomorphism onto its image. Let $U$ denote the image of this diffeomorphism, and set $\psi: U \to B$ to be the inverse of this diffeomorphism. Thus, $\rho(e^{\psi(\eta)})v = \eta$ for all $\eta \in U$. Define $\varphi$ on $G|_U$ to be the map that sends $g \in G|_U$ to $\varphi(g) = (\rho(g)v,e^{-\psi(\varphi(g))}g)$. This gives the required $H$-equivariant diffeomorphism.

**Additional reading**

Chapter 11: Covariant derivatives and connections

Let \( \pi: E \to M \) denote a vector bundle; with fiber \( V = \mathbb{R}^n \) or \( \mathbb{C}^n \). Use \( C^\infty(M; E) \) to denote the vector space of smooth sections of \( E \). Recall here that a section, \( s \), is a map from \( M \) to \( E \) such that \( \pi \circ s \) is the identity. This is an infinite dimensional vector space. The question arises as to how to take the derivative of a section of \( E \) in a given direction. Of course, one can take the differential of a map, this giving a linear map \( s_\ast: TM \to s^*TE \). This notion of derivative turns out to be less than satisfactory. Needed is a derivative that defines a bundle homomorphism from \( TM \) to \( E \), not to \( s^*TE \). At issue here is how to define such a derivative.

As explained in what follows, there are various ways to proceed, though all give the same thing at the end of the day. One such definition involves principal Lie group bundles and the notion of a connection on such a bundle. The related notions of covariant derivative and connection are the focus of this chapter.

11a) Covariant derivatives

The space \( C^\infty(M; E) \) is a module for the action of the algebra, \( C^\infty(M) \), of smooth functions with values in \( \mathbb{R} \) if \( V = \mathbb{R}^n \) or values in \( \mathbb{C} \) if \( V = \mathbb{C}^n \). The action is such that a given function \( f \) acts on \( s \in C^\infty(M; E) \) to give \( fs \). A covariant derivative for \( C^\infty(M; E) \) is a map \( \nabla: C^\infty(M; E) \to C^\infty(M; E \otimes T^*M) \) with the following two properties: First, it respects the vector space structure. This is to say that \( \nabla(cs) = c\nabla s \) and \( \nabla(s + s') = \nabla s + \nabla s' \) and when \( c \) is in \( \mathbb{R} \) or \( \mathbb{C} \) (as the case may be), while \( s \) and \( s' \) are sections of \( E \).

Second, it obeys the analog of Leibnitz’ rule:

\[
\nabla(f s) = f \nabla s + s \otimes df
\]

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

To see that there exist covariant derivatives, it is enough to exhibit at least one. To obtain one, take a locally finite open cover, \( \Lambda \), of \( M \) such that any given open set \( U \in \Lambda \) comes with a vector bundle isomorphism \( \varphi_U: E|_U \to U \times V \). Let \( \{\chi_U\}_{U \in \Lambda} \) denote a subordinate partition of unity. Thus, \( \chi_U: M \to [0, 1] \) has support only in \( U \), and \( \sum_{U \in \Lambda} \chi_U = 1 \) at each point. Save this partition for a moment.

For each \( U \in \Lambda \), define the covariant derivative, \( d \), on \( C^\infty(U; U \times V) \) as follows: Write a given section, \( u \), of the trivial bundle \( U \times V \) as a map \( x \to (x, v(x)) \). Now define the covariant derivative \( du \) as the map \( x \to (x, dv|_x) \).

Now, suppose that \( s \) is any given section of \( E \). Define

\[
\nabla s = \sum_{U \in \Lambda} \chi_U \varphi_U^{-1}(d(\varphi_U s))
\]
Note that this obeys Leibnitz rule by virtue of the fact that \( \sum_{\lambda \in \Lambda} \chi_{\lambda} = 1 \).

11b) **The space of covariant derivatives**

There are lots of covariant derivatives. As is explained next, the space of covariant derivatives is an affine space modeled on \( C^\infty(M; \text{Hom}(E; E \otimes T^*M)) \). To see this, first note that if \( a \in C^\infty(M; \text{Hom}(E; E \otimes T^*M)) \) and if \( \nabla \) is any given covariant derivative, then \( \nabla + a \) is also a covariant derivative. Meanwhile, if \( \nabla \) and \( \nabla' \) are both covariant derivatives, then their difference, \( \nabla - \nabla' \), is a section of \( \text{Hom}(E; E \otimes T^*M) \). This is because their difference is linear over the action of \( C^\infty(M) \).

The following lemma puts this last remark in a more general context. This lemma applies to the situation at hand using for \( E' \) the bundle \( E \otimes T^*M \) and taking \( L = \nabla' - \nabla \).

**Lemma:** Suppose that \( E \) and \( E' \) are vector bundles (either real or complex) and that \( L \) is an \( \mathbb{R} \) or \( \mathbb{C} \) linear map (as the case may be) that takes a section of \( E \) to one of \( E' \). Suppose in addition that \( L(f \cdot) = f L(\cdot) \) for all functions \( f \). Then there exists a unique section, \( L \), of \( \text{Hom}(E; E') \) such that \( L(\cdot) = L(\cdot) \).

**Proof:** To find \( L \), fix an open set \( U \subset M \) where both \( E \) and \( E' \) has a basis of sections. Denote these respective basis as \( \{ e_a \}_{a=1}^d \) and \( \{ e'_b \}_{b=1}^{d'} \) where \( d \) and \( d' \) denote here the respective fiber dimensions of \( E \) and \( E' \). Since \( \{ e'_b \}_{b=1}^{d'} \) is a basis of sections of \( E' \), any given \( a \in \{ 1, \ldots, d \} \) version of \( L e_a \) can be written as a linear combination of this basis. This is to say that there are functions \( \{ L_{ab} \}_{1 \leq b \leq d'} \) on \( U \) such that

\[
L e_a = \sum_{1 \leq b \leq d'} L_{ab} e'_b
\]

This understood, the homomorphism \( L \) is defined over \( U \) as follows: Let \( s \) denote a section of \( E \) over \( U \), and write \( s = \sum_{1 \leq a \leq d} s_a e_a \) in terms of the basis \( \{ e_a \}_{1 \leq a \leq d} \). Then \( L s \) is defined to be the section of \( E' \) given by \( L s = \sum_{1 \leq a \leq d} \sum_{1 \leq b \leq d'} L_{ab} s_a e'_b \). The identity \( L s = L s \) is guaranteed by the fact that \( L(f \cdot) = f L(\cdot) \) when \( f \) is a function. The same identity guarantees that the homomorphism \( L \) does not depend on the choice of the basis of sections. This is any two choices give the same section of \( \text{Hom}(E; E') \).

It is traditional to view \( \nabla - \nabla' \) as a section of \( \text{End}(E) \otimes T^*M \) rather than as a section of \( \text{Hom}(E; E \otimes T^*M) \); these bundles being canonically isomorphic.

What was just said about the affine nature of the space of covariant derivatives has the following implication: Let \( \nabla \) denote a covariant derivative on \( C^\infty(M; E) \), and let \( s \) denote a section of \( E \). Suppose that \( U \) is an open set in \( M \) and \( \phi_U: E|_U \to U \times V \) is an bundle isomorphism. Write \( \phi_U s \) as \( x \to (x, s_U(x)) \) with \( s_U: U \to V \). Then \( \phi_U(\nabla s) \) appears as the section
\[ x \mapsto (x, (\nabla s)_U) \] where \( \nabla s_U = ds_U + a_U s_U \)

and \( a_U \) is some \( s \)-independent map from \( U \) to \( \text{End}(V) \otimes T^*M \).

Be forewarned that the assignment \( U \mapsto a_U \) does not define a section over \( M \) of \( \text{End}(E) \otimes T^*M \). Here is why: Suppose that \( U' \) is another open set with a bundle isomorphism \( \phi_{U'}: E|_{U'} \to U' \times V \) that overlaps with \( U \). Let \( g_{U'U}: U \cap U' \to \text{End}(V) \) denote the transition function. Thus, \( s_{U'} = g_{U'U} s_U \). Meanwhile, \( \nabla s \) is a bonafide section of \( E \otimes T^*M \) and so \( (\nabla s)_{U'} = g_{U'U} (\nabla s)_U \). This requires that

\[ a_{U'} = g_{U'U} a_U g_{U'U}^{-1} - (dg_{U'U}) g_{U'U}^{-1}. \]

Conversely, a covariant derivative on \( \mathcal{C}^\infty(M; E) \) is defined by the following data: First, a locally finite, open cover \( \Lambda \) of \( M \) such that each \( U \in \Lambda \) comes with a bundle isomorphism \( \phi_U: E|_U \to U \times V \). Second, for each \( U \in \Lambda \), a map \( a_U: U \to \text{End}(V) \otimes T^*M|_U \). The collection \( \{a_U\}_{U \in \Lambda} \) define a covariant derivative if and only if the condition above holds for any pair \( U, U' \in \Lambda \) that overlap.

11c) Another construction of a covariant derivatives

What follows describes a relatively straightforward construction of a covariant derivative on a given vector bundle \( E \to M \) with fiber \( \mathbb{V}^n = \mathbb{R}_n \) or \( \mathbb{C}_n \). I assume here that \( M \) is compact, but with care, you can make this construction work in general. Let \( \mathbb{V} \) denote either \( \mathbb{R} \) or \( \mathbb{C} \) as the case may be.

This construction exploits the fact that the product bundle \( M \times \mathbb{V}^N \) has one very simple covariant derivative that is defined as follows: Let \( x \mapsto f = (x, (f_1(x), \ldots, f_N(x))) \) denote a section of the bundle \( M \times \mathbb{V}^N \to M \). Define the section \( df \) of \( (M \times \mathbb{V}^N) \otimes T^*M \) by

\[ x \mapsto (x, (df_1|_x, \ldots, df_N|_x)). \]

This gives a covariant derivative. Now, suppose that \( N \geq 1 \) and that \( E \) is a subbundle of \( M \times \mathbb{V}^N \). Introduce \( \Pi: \text{End}(M; M \times \mathbb{V}^N) \) to denote the fiberwise orthogonal projection in \( \mathbb{V}^N \) onto \( E \). Let \( s: M \to E \) now denote a section of \( E \). Since \( E \) sits in \( M \times \mathbb{R}^N \), I can view \( s \) as a section of \( M \times \mathbb{R}^N \) and so make sense of \( ds \). This understood, then

\[ \nabla s = \Pi ds \]

is a section of \( E \otimes T^*M \), and the assignment \( s \to \nabla s \) is a covariant derivative on \( E \).
Now suppose we are given the bundle $E \to M$ in isolation. Here is how to view $E$ as a subbundle of some trivial bundle: Let $E^* \to M$ denote the bundle $\text{Hom}(E; \mathbb{V})$. Recall that one can find some integer $N \geq 1$ and a set $\{s_1, \ldots, s_N\}$ of sections of $E^*$ with the property that this set spans the fiber of $E$ at each point in $M$. One way to do this is to use the open cover, $\Lambda$, of $M$ that was introduced above, with its associated partition of unity $\{\chi_U\}_{U \in \Lambda}$. Fix a basis $\{v_1, \ldots, v_n\}$ for $V$. Each $U \in \Lambda$ determines the set of $n$ sections $\{s_{1U}, \ldots, s_{nU}\}$ where $s_{kU} = \varphi_U^{-1}(\chi_U v_k)$. Here $\varphi_U : E|_U \to U \times V$ is the associated isomorphism. Also, $v_k$ is viewed here as a section of $U \times V$. Take the set $\{s_1, \ldots, s_n\}$ to be an ordering of the set $\cup_{U \in \Lambda} \{s_{1U}, \ldots, s_{nU}\}$.

Define a bundle homomorphism $\psi : E \to M \times \mathbb{V}^N$ by the following rule: Let $x \in M$ and $v \in E|_x$. Then $\psi(v) = (x, s_1|_x v, \ldots, s_n|_x v)$. This realizes $E$ as a subbundle of the trivial bundle. This puts $E$ inside $M \times \mathbb{V}^N$ as a subbundle, and so we can use the definition of $\mathbb{V}$ above to define a covariant derivative for sections of $E$.

Here is a relatively concrete example: View $S^n$ as the unit sphere in $\mathbb{R}^{n+1}$. Now view $TS^n$ as the set of pairs $\{(x, v) : x, v \in \mathbb{R}^{n+1} \text{ with } |x| = 1 \text{ and } x^T v = 0\}$. This identifies $TS^n$ with a subbundle in $S^n \times \mathbb{R}^{n+1}$. Let $s$ now denote a section of $TS^n$, thus a vector field on $S^n$. Define the section $\nabla s$ of $TS^n \otimes T^*S^n$ by the rule $\nabla s = \Pi ds$. Note that in this case $\Pi_k = 1 - xx^T$ where $1$ is the identity $(n+1) \times (n+1)$ matrix. Since $x^T s(x) = 0$, another way to write this is $\nabla s = \nabla s = ds + xdx^T s$. For instance, consider the following vector field: Let $e \in \mathbb{R}^{n+1}$ denote a constant vector. Then $x \to s(x) = e - xx^T e$ is a section of $TS^n$. Its covariant derivative is $(xx^T - 1)dx^T e$.

11d) Principal bundles and connections

This part the chapter defines the notion of a connection on a principal bundle. This notion is of central importance in its own right. In any event, connections are used momentarily to give an alternate and very useful definition of the covariant derivative. The discussion here has seven parts.

Part 1: Suppose that $E$ is associated to some principal bundle $\pi : P \to M$. (We know that $E$ is always associated to $P_{G(E)}$, and to perhaps other principal bundles.) In particular, suppose that $G$ is a Lie group, $V = \mathbb{R}^m$ or $\mathbb{C}^m$ as the case may be, and $\rho$ is a representation of $G$ in $\text{GL}(V)$. Suppose further that $E$ is given as $P \times \rho V$. Recall from Chapter 10i that the latter is the quotient of $P \times V$ by the relation $(p, v) \sim (pg^{-1}, \rho(g)v)$ when $g \in G$. A section $s$ of $E$ appears in this context as a $G$-equivariant map from $P$ to $V$. Said differently, a section $s$ defines, and is conversely defined by a smooth map $s^p : P \to V$ that obeys $s^p(pg^{-1}) = \rho(g)s^p(p)$. Indeed, suppose first that one is given such a map. The corresponding section $s$ associates to any given point $x \in M$ the equivalence class of $(p, s^p(p))$ where $p$ can be any point in $P|_x$. To see that this makes sense, remark that any other point in $P|_x$ can be written as $pg^{-1}$ with $G \in G$. Because $s^p(pg^{-1}) = \rho(g)s^p(p)$, the pair
(pg^\text{-1}, s^p(pg^\text{-1})) defines the same equivalence class in P \times \rho V as does the original pair (p, s^p(p)). To go in reverse, suppose that s is a section of P \times \rho V. By definition, s associates an equivalence class in P \times V to every point in M. An equivalence class is an assignment to each p \in P of a point v \in V such that pg^\text{-1} is assigned the point \rho(g)p. But this is just another way to say that such an assignment is a smooth map from P to V that obeys the required condition.

This view from P of a section of P \times \rho V has the following useful generalization: Suppose that p: E' \rightarrow M is a second vector bundle either real or complex, and perhaps unrelated to P. The examples that follow take E' to be either TM, T^*M or some exterior power \wedge^p T^*M. In any event, a section over M of (P \times \rho V) \otimes E' appears upstairs on P as a suitably G-equivariant, fiberwise linear map from \pi^*E' to V. The notion of G-equivariance in this context is as in the preceding paragraph: If s denotes the section of (P \times \rho V) \otimes E' in question, then the corresponding map, s^p, from \pi^*E' to V is such that s^p(pg, e') = \rho(g)s^p(p, e') for each g \in G and pair (p, e') \in \pi^*E' \subset P \times E'. Meanwhile, the notion of fiberwise linear means the following: If r is in either \mathbb{R} or \mathbb{C} as the case may be, then s^p(p, re') = rs(p, e'); and if (p, e_1') and (p, e_2') are points in \pi^*E' over the same point in P, then s^p(p, e_1' + e_2') = s^p(p, e_1') + s^p(p, e_2').

Part 2: We already know how to take the derivative of a map from a manifold to a vector space. This Y is the manifold in question, and \nu the map from Y to the vector space, then the derivative is a vector of differential forms on Y; any given component is the exterior derivative of the corresponding component of \nu. In the case when Y = P and \nu is the map s^p that corresponds to a section, s, of P \times \rho V, then this vector of differential forms is denoted by, (s^p). As a vector of differential forms, it defines a fiberwise linear map from TP to V. Although this map is suitably G-equivariant, it does not by itself define a covariant derivative. Indeed, a covariant derivative appears upstairs on P as a G-equivariant, fiberwise linear map from \pi^*TM to V. Said differently, a covariant derivative appears on P as a suitably G-equivariant section over P of the tensor product of \pi^*TM with the product bundle P \times V.

Part 3: To see what (s^p), is missing, it is important to keep in mind that TP has some special properties that arise from the fact that P is a principal bundle over M. In particular, there exists over P the sequence of vector bundle homomorphisms

\[ 0 \rightarrow \ker(\pi_\ast) \rightarrow TP \rightarrow \pi^*TM \rightarrow 0, \]

\( (*) \)
where the notation is as follows: First, $\pi; TP \to TM$ is the differential of the projection map $\pi$ and $\ker(\pi_\ast)$ designates the subbundle in $TP$ that is sent by $\pi_\ast$ to the zero section in $TM$. This is to say that the vectors in the kernel $\pi_\ast$ are those that are tangent to the fibers of $\pi$. Thus, $\ker(\pi_\ast)$ over $P|_x$ is canonically isomorphic to $T(P|_x)$. The arrows in (*) are meant to indicate the following homomorphisms: That from $\ker(\pi_\ast)$ to $TP$ is the tautological inclusion as a subbundle. That from $TP$ to $\pi_\ast TM$ sends a given vector $v \in TP$ to the pair $(\pi(p), \pi_\ast v)$ in $\pi_\ast TM \subset P \times TM$.

What follows are three key observations about (*). Here is the first: The sequence in (*) is exact in the sense that the image of any one of the homomorphisms is the kernel of the homomorphism to its right. Here is the second: The action of $G$ on $P$ lifts to give an action of $G$ on each of the bundles that appear in (*). To elaborate, suppose that $g$ is a given element in $G$. Introduce for the moment $m_g: P \to P$ to denote the action of $g \in G$ that sends $p$ to $pg^{-1}$. This action lifts to $TP$ so as to send as the push-forward map $m_g_\ast$ as defined in Chapter 5c. If $v \in TP$ is in the kernel of $\pi_\ast$, then so is $m_g_\ast v$ because $\pi \circ m_g = \pi$. Thus, $m_g_\ast$ acts also on $\ker(\pi_\ast)$. The action lift of $m_g_\ast$ to $\pi_\ast TM$ is defined by viewing the latter in the manner described above as a subset of $P \times TM$. Viewed in this way, $m_g_\ast$ act so as to send a pair $(p, v) \in P \times TM$ to the pair $(pg^{-1}, v)$. Here, is the final observation: The homomorphisms in (*) are equivariant with respect to the lifts just described of the $G$ action on $P$. Indeed, this follows automatically for the inclusion map from $\ker(\pi_\ast)$ to $TP$, and it follows for the map from $TP$ to $\pi_\ast TM$ because $\pi \circ m_g = \pi$.

A very much related fact is that $\ker(\pi_\ast)$ is canonically isomorphic to the trivial bundle $P \times \text{lie}(G)$. This isomorphism is given by the map $\psi: P \times \text{lie}(G) \to \ker(\pi_\ast)$ that sends a pair $(p, m) \in P \times \text{lie}(G)$ to the tangent vector at $t = 0$ of the path $t \to p \exp(t \cdot m)$. This map $\psi$ is equivariant with respect to the action of $G$ on $\text{lie}(G)$ that sends $g \in G$ and $m$ to $gmg^{-1}$. This is to say that the differential of the map $m_g_\ast: P \to P$ that sends $p$ to $pg^{-1}$ act so that

$$m_g_\ast \psi(p, v) = \psi(pg^{-1}, gvg^{-1}) .$$

Part 4: A connection on the principal bundle $P$ is neither more nor less than a $G$-equivariant splitting of the exact sequence (*). Thus, a connection, $A$, is by definition a linear map

$$A: TP \to \ker(\pi_\ast)$$

that equals the identity on the kernel of $\pi_\ast$ and is equivariant with respect to the action of $G$ on $P$. This is to say that if $g \in G$ and $v \in TP$, then $m_g_\ast(Av) = A(m_g_\ast v)$.  

134
The isomorphism \( \psi: P \times \text{lie}(G) \to \ker(\pi) \) is used to view a connection, \( A \), as a map

\[
A: TP \to \text{lie}(G)
\]

with the following properties:

- \( A(\psi(p, m)) = m \).
- If \( g \in G \) and \( v \in TP \), then \( A(m_g^*v) = gA(v)g^{-1} \).

This last view of \( A \) is quite useful. Viewed this way, a connection is a section of the tensor product of \( P \)'s cotangent bundle with the product bundle \( P \times \text{lie}(G) \) with certain additional properties that concern the \( G \) action on \( P \) and the sequence (*)

The kernel of the homomorphism \( A \) is a subbundle in \( TP \) which is isomorphic via the right most arrow in (*) to \( \pi^*TM \). This subbundle \( \ker(A) \) is often called the horizontal subbundle of \( TP \) and denoted by \( H_A \subset TP \). Meanwhile, the bundle \( \ker(\pi) \) is just as often called the vertical subbundle.

With regards to the notation in what follows, it is customary not to distinguish by notation the aforementioned two views of a connection, one as a bundle homomorphism from \( TP \) to \( \ker(\pi) \) and the other as a fiberwise linear map from \( TP \) to \( \text{lie}(G) \). In any event, the latter view is used primarily in what follows. Another notational quirk with regards to the second of these two views is uses \( T^*P \otimes \text{lie}(G) \) to denote the tensor product of \( T^*P \) with the product vector bundle \( P \times \text{lie}(G) \).

**Part 5:** If \( A \) and \( A' \) are connections on \( P \), then \( a^p = A - A' \) annihilates \( \ker(\pi) \). As a consequence, it defines a fiberwise linear, \( G \)-equivariant map from \( \pi^*TM \) to \( \text{lie}(G) \). Given that \( a^p \) is \( G \)-equivariant and linear on the fibers of \( \pi^*TM \), it corresponds (as described in Part 2 above) to a section over \( M \) of \( (P \times_{ad} \text{lie}(G)) \otimes T^*M \). Here, \( ad \) denotes the representation of \( G \) on its Lie algebra that has \( g \in G \) sending any given matrix \( a \in \text{lie}(G) \) to \( gag^{-1} \). As explained momentarily, the latter matrix is also in \( \text{lie}(G) \). This representation of \( G \) on \( \text{lie}(G) \) is called the adjoint representation.

To say more about the adjoint representation, recall that the Lie algebra of \( G \) was defined to be the tangent space to the identity element \( \iota \) in \( G \). Given that \( G \) is a matrix group, this vector space is a subspace of the relevant \( n \in \{1, 2, \ldots\} \) version of either \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \). What follows explains how to see that \( gag^{-1} \in \text{lie}(G) \) if \( a \in \text{lie}(G) \) and \( g \in G \). As explained in Chapter 8g, the exponential map

\[
a \to e^a = \iota + a + \frac{1}{2} a^2 \ldots
\]
maps \text{lie}(G) to \text{G}. Moreover, this map restricts to a ball about the origin in \text{lie}(G) so as to define a coordinate chart in \text{G} for a neighborhood of the origin. This implies in particular the following: Suppose that \( a \) is any given matrix. If \( t \in (0, 1) \) is not too big, the the matrix \( e^{ta} \) is in \text{G} if and only if \( a \in \text{lie}(G) \). With the preceding as background, suppose that \( a \in \text{lie}(G) \) and \( g \in \text{G} \). Fix some small \( t > 0 \). Use the formula above to see that the matrix \( e^{tgag^{-1}} = ge^{ta}g^{-1} \). Since all three matrices in this product are in \text{G}, the result is in \text{G}. If \( t \) is not too large, this implies that \( tgag^{-1} \in \text{lie}(G) \) and thus \( gag^{-1} \in \text{lie}(G) \).

To continue the discussion about connections, suppose now that \( a \) is a section over \( M \) of the bundle \( (P \times_{\text{ad}} \text{lie}(G)) \otimes T^*M \) and that \( A \) is a connection on \( P \). Just as a section of \( P \times_{\text{ad}} \text{lie}(G) \) can be viewed as a \text{G}-equivariant map from \( P \) to \text{lie}(G), so \( a \) can be viewed on \( P \) as a \text{G}-equivariant, fiberwise linear map from \( \pi^*TM \) to \text{lie}(G). Let \( a^p \) denote the latter incarnation of \( a \). Then \( A + a^p \) defines another connection on \( P \).

These last observations have the following corollary: If \( P \) has one connection, then it has infinitely many, and the space of smooth connections on \( P \) is an affine space based on \( C^\infty(M; (P \times_{\text{ad}} \text{lie}(G)) \otimes T^*M) \).

Part 6: What follows constructs a connection on \( P \). As a preamble to the construction, suppose that \( U \subset M \) is an open set with a principal bundle isomorphism \( \varphi_U : P|_U \rightarrow U \times G \). The product principal bundle \( U \times G \) has a tautological connection 1-form, this defined as follows when \( G \) is a subgroup in \( \text{Gl}(n; \mathbb{R}) \) or \( \text{Gl}(n; \mathbb{C}) \) as the case may be. The connection 1-form at a point \( (x, g) \in U \times G \) is

\[
A_0 = g^{-1}dg,
\]

this a matrix of 1-forms as it should be. Here, and in what follows, I view a connection on a principal \text{G} bundle \( P \) as a \text{lie}(G)-valued 1-form \( P \) with the properties listed at the end of the last lecture.

To see that \( A_0 \) has the correct properties, fix a matrix \( m \in \text{lie}(G) \) with \( \text{lie}(G) \) viewed as a subvector space in \( M(n; \mathbb{R}) \) or \( M(n; \mathbb{C}) \) as the case may be. Introduce the map \( t \rightarrow ge^{tm} \) from \( \mathbb{R} \) into \( G \), this as described in Chapter 5e. The pull-back of \( A_0 \) by this map at \( t = 0 \) is the 1-form on \( \mathbb{R} \) given by \( (A_0|_U)(\varphi(g, m)) \) \( dt \). This is \( mdt \), as it should be. Note also that \( A_0 \) is suitably equivariant under the action of \( G \) on \( G \) by right translation.

The \( \text{lie}(G) \) valued 1-form \( A_0 \) annihilates vectors that are tangent to the \( U \) factor of \( U \times G \). This understood, the corresponding horizontal subbundle \( H_{A_0} \subset TP \) are precisely the tangents to the \( U \) factor, this the \( TU \) in the obvious splitting of \( T(U \times G) = TU \oplus TG \).

With the preamble in mind, now fix a locally finite cover \( \mathcal{U} \) of \( M \) such that each set \( U \in \mathcal{U} \) comes with a principal \text{G}-bundle isomorphism \( \varphi_U : P|_U \rightarrow U \times G \). Let \( \{ \chi_U \}_{U \in \mathcal{U}} \) denote a subordinate, partition of unity. Such a partition is constructed in Appendix 2 of Chapter 1. Set
A = \sum_{U \in \mathcal{U}} \chi_U \varphi_U^* A_0 .

This \text{lie}(G) valued 1-form on P is a connection. The verification is left as an exercise.

**Part 7:** This last part of the story gives another way to view a connection on P. To start, let \( U \subset M \) again denote an open set with a principal G-bundle isomorphism \( \varphi: P|_U \to U \times G \). Now let \( A \) is a given connection on \( G \). Then \( (\varphi^{-1}_U)^* A \) is a connection on the product bundle \( U \times G \). As such, it can be written as

\[
(\varphi^{-1}_U)^* A = g^{-1} dg + g^{-1} a_U g
\]

where \( a_U \) is a 1-form on \( U \) with values in \text{lie}(G). The horizontal space \( H_A \) can be described as follows: Identify \( T(U \times G) = TU \oplus TG \) and identify \( TG \) with \text{lie}(G) as done above. Then

\[
\varphi_{U*}(H_A)|_{(x,g)} = \{(v, -g^{-1} a_U v g) \in TU \oplus \text{lie}(G): v \in TU|_x\}.
\]

11e) **Connections and covariant derivatives**

Suppose that \( E \to M \) is a vector bundle which is given by \( P \times_P V \) where \( \rho \) is a representation of \( G \) on \( V \). Let \( A \) denote a connection on \( P \). Then \( A \) defines a covariant derivative, \( \nabla_A \), on \( C^\infty(M; E) \) as follows: View a section \( s \) of \( E = P \times_P V \) as a \( G \)-equivariant map, \( s^p: P \to V \). The corresponding covariant derivative \( \nabla_A s \) will be viewed as a \( G \)-equivariant section, \( (\nabla_A s)^p \), of \( (P \times V) \otimes \pi^*TM \). To define this section, introduce as in Part 2 of the previous part of this chapter, the differential, \( (s^p)_* \), of the map \( s^p \) from \( P \) to \( V \). This differential maps \( TP \) to \( V \). Now, \( (\nabla_A s)^p \) is supposed to be a \( G \)-equivariant section over \( P \) of \( (P \times V) \otimes \pi^*TM \), which is to say a \( G \)-equivariant homomorphism from \( \pi^*TM \) to \( V \). To obtain such a thing, recall first that the connection’s horizontal subbundle, \( H_A \subset TP \) is canonically isomorphic to \( \pi^*TM \). This understood, the restriction to \( H_A \) of the homomorphism \( (s^p)_*: TP \to V \) defines a covariant derivative of \( s \). The latter is, by definition, the covariant derivative \( \nabla_A s \).

Here is a reinterpretation of this definition. Let \( x \in M \) and let \( v \in TM|_x \). The covariant derivative \( \nabla_A s \) is supposed to assign to \( v \) a section of \( E \). To obtain this section, pick any point \( p \in P|_x \). There is a unique horizontal vector \( v_A \in H_A|_p \) such that \( \pi_* v_A = v \). The covariant derivative \( \nabla_A s \) sends \( v \) to the equivalence class in \( E|_x = (P|_x \times_P V) \) of the pair \( (p, (s^p)_* v_A) \). To see that this is well defined, I need to check that the a different choice for \( p \in P|_x \) gives the same equivalence class. That such is the case is a consequence of the fact that \( A \) and \( s^p \) are suitably equivariant with resect to the \( G \) action on \( P|_x \). This is proved by unwinding all of the definitions—a task that I will leave to you.
To see that this defines a covariant derivative, one need only check Leibnitz’ rule. To do so, suppose that \( f \) is a smooth function on \( M \). Then \((f\pi)^p = \pi^*(f)\pi^p\) and as a consequence, \((f\pi)^p = \pi^*(f)\pi^p + \pi^p \otimes \pi^*(df)\). This implies that

\[
(\nabla_A (f\pi))^p = (f\nabla_A \pi)^p + \pi^p \otimes \pi^*(df).
\]

The latter asserts Leibnitz rule.

If \( A \) and \( A' \) are connections on \( P \), then the difference \( \nabla_{A'} - \nabla_A \) is supposed to be a section over \( M \) of \( \text{End}(E) \otimes T^*M \). What follows identifies this section: Recall from Part 5 of the previous part of this chapter that the difference \( A' - A = a \) can be viewed as a section over \( M \) of \((P \times_{ad} \text{lie}(G)) \otimes T^*M\). Meanwhile, the representation \( \rho \) induces a bundle homomorphism from \( P \times_{ad} \text{lie}(G) \) to \( \text{End}(E) \) as follows: The bundle \( \text{End}(E) \) is the associated vector bundle \( P \times_{ad(\rho)} \text{Gl}(V) \) where \( \text{ad}(\rho) \) here is the representation of \( G \) on \( \text{Gl}(V) \) that has any given \( g \in G \) act on \( m \in \text{Gl}(V) \) as \( \rho(g)m \rho(g^{-1}) \). This understood, the differential of \( \rho \) at the identity in \( G \) defines a \( G \) equivariant homomorphism \( \rho_* : \text{lie}(G) \to \text{Gl}(V) \); this intertwines the representation \( \text{ad} \) on \( \text{lie}(G) \) with the representation \( \text{ad}(\rho) \). The latter homomorphism gives a corresponding homomorphism from \((P \times_{ad} \text{lie}(G)) \otimes T^*M\) to \( \text{End}(E) \otimes T^*M \). This corresponding homomorphism is also denoted by \( \rho_* \) because it is defined by ignoring the \( T^*M \) factor. To elaborate, it is defined so as to send a given decomposable element \( m \otimes \nu \) to \( \rho_*(m) \otimes \nu \); and as any given element is a linear combination of decomposable elements, the latter rule is all that is needed. Granted all of the above, it then follows by unwinding the definitions that

\[
\nabla_{A'} - \nabla_A = \rho_*(a) .
\]

Note that if \( P \) is the bundle \( P_{\text{Gl}(E)} \) of orthonormal frames in \( E \), then \( P \times_{ad} \text{lie}(G) \) is the bundle \( \text{End}(E) \). This and the preceding equation gives a 1-1 correspondence between covariant derivatives on \( E \) and connections on \( P_{\text{Gl}(E)} \).

**11f) Horizontal lifts**

Let \( \pi : P \to M \) denote a principal \( G \) bundle. Fix a connection, \( A \) on \( P \). The connection gives a way to lift any given smooth map \( \gamma : \mathbb{R} \to M \) to a map \( \gamma_A : \mathbb{R} \to P \) such that \( \pi \circ \gamma_A = \gamma \). This is done as follows: Fix a point \( p_0 \) in the fiber of \( P \) over the point \( \gamma(0) \). The lift \( \gamma_A \) is defined by the following two conditions:

- \( \gamma_A(0) = p_0 \)
- \( \text{Let} \ \partial_t \text{ denote the tangent vector to } \mathbb{R}. \text{ Then } \gamma_A \partial_t \in H_A. \)

Thus, the lift starts at \( p_0 \) and its tangent vectors are everywhere horizontal as defined by the connection \( A \). This is to say that \( A(\gamma_A \partial_t) = 0 \).
To see what this looks like, suppose that $\gamma$ crosses a given open set $U \subset M$ where there is a bundle isomorphism $\varphi_U: P|_U \to U \times G$. Then $\varphi_U(\gamma_A)$ is the map $t \to (\gamma(t), g(t))$ where $t \to g(t)$ makes the tangent vector to the path $t \to (\gamma(t), g(t))$ horizontal as defined using the connection $(\varphi_U^{-1})^*A = g^{-1}dg + g^{-1}a_\gamma g$ on $U \times G$. With $T(U \times G)$ identified as before with $TU \oplus \text{Lie}(G)$, this tangent vector is

$$(\gamma, \partial_t, g^{-1} \frac{d}{dt} g)$$

The latter is horizontal if and only if it has the form $(v, -g^{-1}a(v)g)$ with $v \in TU$. This understood, the map $t \to g(t)$ must obey

$$g^{-1} \frac{d}{dt} g = -g^{-1}a_\gamma(\gamma, \partial_t)|_{\gamma(t)} g$$

which is to say that

$$\frac{d}{dt} g + a_\gamma(\gamma, \partial_t)|_{\gamma(t)} g = 0.$$

If you recall the vector field theorem from Chapter 8c about integrating vector fields, you will remember that an equation of this sort has a unique solution given its starting value at some $t = t_0$. This being the case, horizontal lifts always exist.

11g) An application to the classification of principal $G$-bundles up to isomorphism

The notion of a horizontal lift can be used to say something about the set of isomorphism classes of principal $G$ bundles over any given manifold. To this end, suppose that $X$ is a smooth manifold and $\pi: P_X \to X$ is a principal $G$-bundle. Let $M$ denote another smooth manifold.

**Theorem:** Let $f_0$ and $f_1$ denote a pair of smooth maps from $M$ to $X$. The principal $G$ bundles $f_0^*P_X$ and $f_1^*P_X$ are isomorphic if $f_0$ is homotopic to $f_1$.

**Proof:** There is an interval $I \subset \mathbb{R}$ containing $[0, 1]$ and a smooth map, $F$, from $I \times M$ to $X$ such that $F(0, \cdot) = f_0$ and $F(1, \cdot) = f_1$. Keeping this in mind, fix a connection, $A$, on the pull-back bundle $F^*P$. Define a map $\psi_A: f_0^*P_X \to f_1^*P_X$ as follows: Let $p \in f_0^*P$ and let $t \to \gamma_{A,p}(t) \in F^*P$ denote the horizontal lift of the path $t \to (t, \pi(p)) \in I \times M$ that starts at $p$ when $t = 0$. Set $\psi_A(p)$ to be $\gamma_{A,p}(1)$. This is an invertible, fiber preserving map. It also obeys $\psi_A(pg^{-1}) = \psi_A(p)g^{-1}$ because of the equivariance properties of the connection. In particular, these guarantee that $\gamma_{A_pg^{-1}} = \gamma_{A,p} g^{-1}$ at all times $t \in I$. 

139
The classification problem is completed with a second theorem which asserts the existence, for any given lie group G, of a universal classifying space, this a smooth manifold X such that any given principal G-bundle over any given manifold is the pull-back of a fixed principal bundle $P_X \to X$. One can take dim(X) to be finite if a bound has been given apriori on the dimension of the manifolds under consideration. In any event, the result is that there is a 1-1 correspondence

$$\{\text{principal G bundles on M up to isomorphism}\} \leftrightarrow \{\text{homotopy classes of maps from M to X}\}.$$ 

For example, the proposition in Example 2 of Chapter 10i asserts that the large m versions of real Grassmannian $\text{Gr}(m; n)$ and the complex Grassmannian $\text{Gr}_C(m; n)$ can serve as respective classifying spaces for principle $O(n)$ and $U(n)$ bundles over manifolds of a given dimension.

The preceding theorem has the following as a corollary.

**Corollary:** Let $P \to M$ denote a principal G-bundle and let $U$ denote a contractible, open set in $M$. Then there is an isomorphism $\varphi: P|_U \to U \times G$.

What follows is another corollary with regards to the case when M has dimension 2. To set the stage, recall that Chapters 6d, 6e and 6k described various complex, rank 1 vector bundles over surfaces in $\mathbb{R}^3$ that were pull-backs of a fixed bundle over $S^2$ that was defined using the Pauli matrices.

**Corollary:** Let $M$ denote a compact, 2-dimensional manifold and let $f_0$ and $f_1$ denote homotopic maps from $M$ to $S^2$. Let $\pi: E \to S^2$ denote the complex, rank 1 bundle. Then $f_0^*E$ is isomorphic to $f_1^*E$ if $f_0$ is homotopic to $f_1$.

11h) Connections, covariant derivatives and pull-back bundles.

Suppose that M and N are manifolds, that $\phi: M \to N$ is a smooth map and that $\pi: P \to N$ is a principal bundle with fiber some group G. As explained in Chapter 10b, the map $\phi$ can be used to define pull-back principal bundle, $\pi_M: \phi^*P \to M$, this the subspace in $M \times P$ of points $(x, p)$ with $\phi(x) = \pi(p)$. The projection, $\pi_M$ is induced by the obvious projection from $M \times P$ to its M factor. Meanwhile, there exists a canonical G-equivariant map $\hat{\phi}: \phi^*P \to P$ that covers $\phi$, this induced by the projection from $M \times P$ to its P factor.

By the way, another way to view $\phi^*P$ is to consider $\mathcal{P} = M \times P$ as a principal G-bundle over the product $M \times N$. Then M embeds in $M \times N$ as the graph of $\phi$, this the subset of pairs $(x, \phi(x))$. The bundle $\phi^*P$ is the restriction of the principal bundle $\mathcal{P}$ to the graph of $\phi$. 

140
Suppose that \( A \) is a connection on \( P \); thus a \( \text{lie}(G) \) valued 1-form on \( P \) that behaves in a specified manner when pulled back by the action of any given element in \( G \). Maps pull-back 1-forms, and the pull-back \( \hat{\phi}^*A \) is a \( \text{lie}(G) \)-valued 1-form on \( \phi^*P \). This 1-form satisfies all of the required conditions for a connection because \( \hat{\phi} \) is \( G \)-equivariant. You can check directly using the definition of pull-back from Chapter 5c; or you can see this by viewing \( \mathcal{P} = M \times P \) as a principal \( G \)-bundle over \( M \times N \). Then \( A \) defines a connection on \( \mathcal{P} \), since it is a \( \text{lie}(G) \) valued 1-form on \( \mathcal{P} \) with all of the required properties. It just happens to be a 1-form with no dependence on the \( M \) factor in \( M \times P \). Granted all of this, then \( \hat{\phi}^*A \) is the restriction of \( A \) (viewed as a connection on the bundle \( \mathcal{P} \to M \times N \)) to the graph of the map \( \phi \).

Now suppose that \( V \) is either \( \mathbb{R}^n \) or \( \mathbb{C}^n \) and \( \rho: G \to \text{Gl}(V) \) is a representation. The connection \( A \) on \( P \) gives the bundle \( E = (P \times_{\rho} V) \) a covariant derivative, \( \nabla \). The pull-back bundle, \( \hat{\phi}^*E = (\hat{\phi}^*P \times_{\rho} V) \) then has the covariant derivative, \( \hat{\phi}^*\nabla \), this defined by the connection \( \hat{\phi}^*A \). Here is another way to see this: View \( A \) again as a connection on the bundle \( \mathcal{P} = M \times P \) over \( M \times N \). Let \( \mathcal{E} \) denote the vector bundle \( M \times E \to M \times N \). Of course, this is \( P \times_{\rho} V \). The restriction of \( \mathcal{E} \) to the graph of \( \phi \) is the bundle \( \phi^*E \). The connection \( A \), now viewed as living on \( \mathcal{P} \), defines a covariant derivative for \( \mathcal{E} \), and hence for the restriction of \( \mathcal{E} \) to the graph of \( \phi \). This is covariant derivative is \( \phi^*\nabla \).

What was just said about vector bundles can be said with no reference to connections. To elaborate, suppose that \( \pi: E \to N \) is a vector bundle, with no particular principal bundle in view. View \( M \times E \) as a vector bundle, \( \mathcal{E} \), over \( M \times N \). As before, the pull-back bundle \( \phi^*E \) is the restriction of \( \mathcal{E} \) to the graph of \( \phi \). If \( V \) is a covariant derivative for sections of \( E \), then its restriction to sections of \( \mathcal{E} \) over the graph of \( \phi \) defines the covariant derivative \( \phi^*V \) for sections of \( \phi^*E \).

**Additional reading**

This chapter continues the discussion of connections and covariant derivatives with the introduction of the notion of the \textit{curvature} of a covariant derivative or connection. This notion of curvature is defined in the upcoming Part d) of this chapter. The first three parts of the chapter constitute a digression of sorts to introduce two notions involving differentiation of differential forms and vector fields that require \textit{no} auxiliary choices. Both are very important in their own right. In any event, they are used to define curvature when covariant derivatives reappear in the story.

\textbf{12a) Exterior derivative}

This first part of the digression introduce the \textit{exterior derivative}. As explained below, the exterior derivative is a natural extension to differential forms of the notion of the differential of a function.

To start, suppose that \( \mathcal{g} \) is a function on \( M \). Chapter 3i describes how its differential, \( d \mathcal{g} \), can be viewed as a section of \( T^*M \). The association of any given function to its differential defines a linear map from \( C^\infty(M) \) to \( C^\infty(M; T^*M) \). This understood, the exterior derivative is an extends this linear map so as to be a linear map from \( C^\infty(M; \wedge^p T^*M) \) to \( C^\infty(M; \wedge^{p+1} T^*M) \) for any given \( p \in \{0, \ldots, n = \dim(M)\} \). Note in this regard that \( \wedge^p T^*M \) is the 0-dimensional bundle \( M \times \{0\} \) for \( p > n \). This extension is also denoted by \( d \).

What follows defined \( d \) by induction on the degree, \( p \), of the differential form. If \( f \in C^\infty(M) = C^\infty(M; \wedge^0 T^*M) \), define \( df \) as in Chapter 3i. (Recall that in a coordinate patch with coordinates \( (x_1, \ldots, x_n) \), this 1-form is \( df = \sum \frac{\partial}{\partial x_i} f \, dx_i \).) Granted the preceding, suppose next that \( d: C^\infty(M; \wedge^p T^*M) \rightarrow C^\infty(M; \wedge^{p+1} T^*M) \) has been defined for all integers \( p = 0, \ldots, q-1 \) for \( 1 \leq q \leq n \). The following rules specify how \( d \) is to act on \( C^\infty(M; \wedge^q T^*M) \):

- \textbf{If} \( \omega \in C^\infty(M; \wedge^q T^*M) \) and \( \omega = d\alpha \) for some \( \alpha \in C^\infty(M; \wedge^{q-1} T^*M) \), then \( d\omega = 0 \).
- \textbf{If} \( \omega = f \, d\alpha \) for some \( \alpha \) as before, and with \( f \) a smooth function, then \( d\omega = df \wedge d\alpha \).
- \textbf{If} \( \omega_1 \) and \( \omega_2 \) are two \( q \)-forms, then \( d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2 \).

Note that this induction covers all cases. This is because any given point in \( M \) has a neighborhood on which \( \wedge^p T^*M \) is spanned by a set of sections of the form \( dg_{i_1} \wedge \cdots \wedge dg_{i_p} \) for \( g_{i_1}, \ldots, g_{i_p} \) chosen from a set of \( \dim(M) \) functions. (To see this, take a coordinate chart centered on the point. Then the \( p \)-fold wedge products of the differentials of the coordinate functions span \( \wedge^p T^*M \) on this coordinate neighborhood.)

This definition must be checked for consistency. The point being that a given \( p \)-form \( \omega \) can be written as a sum of terms, each of the form \( g_{p+1} \, dg_{i_1} \wedge \cdots \wedge dg_{i_p} \), in many
different ways. The consistency is equivalent to the following hallmark of partial derivatives: They commute: If \( f \) is a smooth function on \( \mathbb{R}^n \), then
\[
\frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f = \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f.
\]

To why this is key, fix a coordinate chart so as to identify a neighborhood of a given point in \( M \) with a neighborhood of the origin in \( \mathbb{R}^n \). Let \( f \) denote a function on \( M \). Then \( df \) appears in this coordinate chart as
\[
df = \sum_j \frac{\partial}{\partial x_j} f \, dx_j.
\]
Now, if \( d(df) \) is to be zero, and also \( d(dx_j) \) is to be zero, then the consistency of the definition requires that
\[
\sum_{k,j} \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f \, dx_k \wedge dx_j = 0.
\]
The left hand side above is, indeed zero. The reason is that \( dx_k \wedge dx_j = -dx_j \wedge dx_k \) and so what is written here is the same as
\[
\frac{1}{2} \sum_{k,j} \left( \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_k} f \right) \, dx_k \wedge dx_j,
\]
as can be seen by changing the order of summation. Meanwhile, what is written above is zero because partial derivatives in different directions commute.

By the same token, suppose that \( \alpha \) is a p-1 form for \( p > 1 \). In this coordinate patch,
\[
\alpha = \sum_{1 \leq i_1 < \cdots < i_{p-1} \leq n} \alpha_{i_1 \cdots i_{p-1}} \, dx_{i_1} \wedge \cdots \wedge dx_{i_{p-1}}
\]
where each \( \alpha_{\cdots} \) is a function. This understood, then
\[d\alpha = \sum_{1 \leq i_1 < \cdots < i_{p-1} \leq n} \sum_j \frac{\partial}{\partial x_j} \alpha_{i_1 \cdots i_{p-1}} \, dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_{p-1}},\]
\[d(d\alpha) = \sum_{1 \leq i_1 < \cdots < i_{p-1} \leq n} \sum_{k,j} \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} \alpha_{i_1 \cdots i_{p-1}} \, dx_k \wedge dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_{p-1}};\]
and so the fact that \( d(d\alpha) = 0 \) follows as before from the fact that the matrix of second derivatives is symmetric.

Here is a related fact: The assertion that the matrix of second derivatives of a given function vanishes is not a coordinate independent statement. Indeed, suppose that \( \frac{\partial}{\partial x_k} \frac{\partial}{\partial x_j} f = 0 \) at a given point. Suppose now that \( x_j = \phi_j(y) \) where \( \phi : \mathbb{R}^n \to \mathbb{R}^n \) is a
diffeomorphism on a neighborhood of the origin. Let \( f' = \phi^* f \). This is to say that \( f'(y) = f(\phi(y)) \). The chain rule tells us that

\[
\frac{\partial}{\partial y_j} f' = \sum_m \left( \frac{\partial \phi_m}{\partial y_j} \right) \left( \frac{\partial}{\partial \phi_m} f \right) \left|_{\phi(y)} \right.,
\]

and then that

\[
\frac{\partial}{\partial y_k} \frac{\partial}{\partial y_j} f = \sum_{p,m} \frac{\partial \phi_k}{\partial y_p} \frac{\partial \phi_m}{\partial y_j} \frac{\partial^2}{\partial y_p \partial y_m} f + \sum_m \frac{\partial^2 \phi_m}{\partial y_k \partial y_j} \frac{\partial}{\partial \phi_m} f
\]

Thus, the vanishing of the vanishing of the matrix of second derivatives of \( f \) in all coordinate charts requires both the first and second derivatives of \( f \) to vanish in all coordinate charts.

By the way, the preceding identity implies that the matrix of second derivatives of a function does not define by itself a section over \( M \) of a vector bundle. This because of the right most contribution to its transformation law.

Here is one additional observation about the exterior derivative: Suppose that \( \alpha \) is a \( p \)-form and \( \beta \) is a \( q \)-form. Then

\[
d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^p \alpha \wedge d\beta
\]

This is consistent with the fact that \( \alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha \).

12b) Closed forms, exact forms, diffeomorphisms and De Rham cohomology

A \( p \)-form \( \omega \) on \( M \) is said to be closed if \( d\omega = 0 \). Such a form is said to be exact of \( \omega = d\alpha \). The \( p \)'th De Rham cohomology of \( M \) is, by definition, the vector space quotient of the linear space of closed forms by the linear subspace of exact forms. This vector space is denoted by \( H^p_{\text{De Rham}}(M) \). Thus,

\[
H^p_{\text{De Rham}}(M) = \ker(d) / \text{image}(d)
\]

As is explained momentarily, these spaces are, apriori, invariants of the differentiable structure on \( M \). (Even so, they are actually dependent only on the topological structure. If \( M \) and \( M' \) are homeomorphic, smooth manifolds, then they have isomorphic DeRham cohomologies.) For example, \( H^p_{\text{De Rham}}(S^n) \) is zero if \( p \neq 0, n \) and \( \mathbb{R} \) if \( p = 0 \) or \( n \).

To see why the De Rham cohomology is a differentiable structure invariants, it is enough to prove that the exterior derivative, \( d \), commutes with pull-back by a map. This is to say the following: If \( M \) and \( N \) are smooth manifolds, if \( f: M \to N \) is a smooth map, and if \( \alpha \) is a \((p-1)\)-form on \( N \), then
\[ f^*(d\alpha) = d(f^*\alpha). \]

The proof is an exercise with the chain rule. To get you going on this task, suppose that \( p \in M \) and that \( (y_1, \ldots, y_m) \) are coordinates centered at \( p \). Let \( (x_1, \ldots, x_n) \) denote coordinates for a neighborhood of \( f(p) \) centered at \( f(p) \). Near \( p \), the map \( f \) appears in these coordinates as a map \( x \to (f_1(x), \ldots, f_n(x)) \) of a neighborhood of \( 0 \in \mathbb{R}^n \) to a neighborhood of \( 0 \in \mathbb{R}^m \). Use the formula given in the just completed Chapter 12a for \( \alpha \) and \( d\alpha \). Then use the chain rule to compute their pull-backs to the via \( f \).

On a related topic, suppose that \( f_0 \) and \( f_1 \) are smooth maps from \( M \) to a manifold \( N \) which are homotopic in the usual sense: There is a smooth map \( \psi: [0, 1] \times M \to N \) with \( \psi(0, \cdot) = f_0 \) and \( \psi(1, \cdot) = f_1 \). Let \( \omega \) denote a closed form on \( N \). Then both \( f_0^*\omega \) and \( f_1^*\omega \) are closed forms. However, they define the same De Rham cohomology class as they differ by the image of \( d \). To see why this is the case, write \( T^* (\mathbb{R} \times M) = \mathbb{R} \oplus T^* M \). This is to say that a cotangent vector at \((t, x)\) can be written as \( \alpha = \alpha_0 dt + \alpha_M \) where \( \alpha_M \in T^*_M|_x \) and where \( \alpha_0 \in \mathbb{R} \). By the same token \( \wedge^p T^* (\mathbb{R} \times M) = \wedge^p T^* M \oplus \wedge^p T^*_M \); this is to say that a \( p \)-form \( \alpha \) on \( \mathbb{R} \times M \) at any given point \((t, x)\) can be written as \( \alpha = \alpha_0 dt + \alpha_M \) where \( \alpha_M \in T^*_M|_x \) and where \( \alpha_0 \in \mathbb{R} \). By the same token

\[ \wedge^p T^* (\mathbb{R} \times M) = \wedge^p T^* M \oplus \wedge^p T^*_M; \]

these are \( p \)-forms on \( \mathbb{R} \times M \) at any given point \((t, x)\) can be written as

\[ dt \wedge \alpha_{p-1} + \alpha_p \]

where \( \alpha_{p-1} \in \wedge^{p-1} T^*_M|_x \) and \( \alpha_p \in \wedge^p T^*_M|_x \). The form \( \alpha_{p-1} \) can be viewed as a time \( t \in \mathbb{R} \) dependent section of \( \wedge^{p-1} T^*_M|_x \) and the form \( \alpha_p \) is a time \( t \in \mathbb{R} \) dependent section of \( \wedge^p T^*_M \). The exterior derivative \( d\alpha \) is a \( p+1 \)-form on \( \mathbb{R} \times M \) and so it has corresponding components \( dt \wedge (d\alpha)_p + (d\alpha)_{p+1} \). These are given by

\[ (d\alpha)_p = -d^i \alpha_{p-1} + \frac{\partial}{\partial t} \alpha_p \quad \text{and} \quad (d\alpha)_{p+1} = d^i \alpha_p \]

where \( d^i \) is used here to denote the exterior derivative at any fixed \( t \in \mathbb{R} \) of a differential form on \( \{t\} \times M \).

Now suppose that \( \alpha = \psi^*\omega \) where \( \psi: [0, 1] \times M \to N \) is the homotopy from \( f_0 \) to \( f_1 \). If \( \omega \) is closed, then \( d\psi^*\omega = \psi^*d\omega = 0 \), so \( \psi^*\omega \) is a closed form. This understood, the preceding formula for \( (d\psi^*\omega)_{p+1} \) says that the pull-back of \( \omega \) by \( \psi(t, \cdot) \) is closed for any given \( t \in [0, 1] \), and the formula for \( (d\psi^*\omega)_p \) can be integrated to see that

\[ f_1^*\omega = f_0^*\omega + d \int_0^1 (\psi^*\omega)_{p+1}. \]
This proves that $f_1^*\omega$ differs from $f_0^*\omega$ by an exact form.

This fact just stated has the following corollary:

**Proposition:** Let $U \subset M$ denote a contractible open set. This is to say that there exists a smooth map $\psi: [0, 1] \times U \rightarrow M$ such that $\psi(1, \cdot)$ is the identity map on $U$ and $\psi(0, \cdot)$ maps $U$ to a point $p \in M$. Let $p \geq 1$ and let $\omega$ denote a closed $p$-form on $M$. Then there exists a $p$-1 form $\alpha$ on $U$ such that $d\alpha = \omega$.

**Proof:** It follows that $\omega|_U = \psi(1, \cdot)^*\omega$ and this differs from $\psi(0, \cdot)^*\omega$ by an exact form. Meanwhile, $\psi(0, \cdot)$ maps $U$ to a point, and since $\wedge^p T^*\{\text{point}\} = 0$ for $p > 0$, so $\psi(0, \cdot)^*\omega$ is equal to zero.

The proposition has as its corollary what is usually called the Poincaré lemma:

**Corollary:** Let $U$ denote a contractible manifold, for instance $\mathbb{R}^n$. Then any closed form $p$-form on $U$ for $p \geq 1$ is the exterior derivative of a $p$-1 form.

This corollary can be used to prove the De Rham cohomology of a compact manifold is a finite dimensional vector space over $\mathbb{R}$. See, for example, Chapter ?? in [BT?].

**12c) Lie derivative**

The preceding discussion about the pull-back of a differential form via a homotopy gives a useful formula when the homotopy is a map $\psi: \mathbb{R} \times M \rightarrow M$ that is obtained by moving the points in $M$ some given time along the integral curves of a vector field. To elaborate, suppose that $M$ is compact and that $v$ is a vector field on $M$. The vector field theorem in Chapter 8c supplies a smooth map $\psi: \mathbb{R} \times M \rightarrow M$ with the following properties:

- $\psi(0, x) = x$.
- $\psi, \partial_t = v|_\psi$.

Let $\omega$ now denote a $p$-form on $M$. Then $\psi^*\omega$ is a $p$-form on $\mathbb{R} \times M$. What can be said about the components $(\psi^*\omega)_p$ and $(\psi^*\omega)_\nu$? The time derivative of $(\psi^*\omega)_p$ at $t = 0$ is said to be the Lie derivative of $\omega$ with respect to $v$. It is denoted $\mathcal{L}_v\omega$ and is such that

$$\mathcal{L}_v\omega|_{t=0} = \frac{\partial}{\partial t} (\psi^*\omega)_p|_{t=0} = \mathcal{L}_v\omega = (d\omega)(v, \cdot) + d(\omega(v, \cdot))$$
where the notation is as follows: If \( \nu \) is a p-form on \( M \) and \( v \) is a vector field, then \( \nu(v, \cdot) \) is the p-1 form that is obtained by the homomorphism \( TM \otimes (\otimes^p T^*M) \to \otimes^{p-1} T^*M \) that comes by writing \( TM \otimes (\otimes^p T^*M) \to (TM \otimes T^*M) \otimes (\otimes^{p-1} T^*M) \) and then using the defining homomorphism from \( (TM \otimes T^*M) \to \mathbb{R} \). The form \( (\psi^* \omega)_{p,1} \) at \( t = 0 \) is equal to \( \omega(v, \cdot) \). The values of \( (\psi^* \omega)_p \) and \( (\psi^* \omega)_{p,1} \) at times \( t \neq 0 \) are obtained by using these \( t = 0 \) formulae with \( \omega \) replaced its pullback via the diffeomorphism \( \psi(t, \cdot): M \to M \).

The Lie derivative acting on p-forms can be viewed as a derivation of an extension of the algebra of the algebra of smooth functions. Recall that \( C^\infty(M; \mathbb{R}) \) is an algebra with addition and multiplication being pointwise addition and multiplication of functions. The derivations of this algebra are the vector fields, thus the sections of \( TM \). Addition of forms and wedge product give \( \Omega_M = \bigoplus_{0 \leq p \leq n} C^\infty(M; \wedge^p T^*M) \) the structure of an algebra, albeit one in which multiplication is not commutative. This is an example of what is called a super algebra, which is a vector space with a \( \mathbb{Z}/2\mathbb{Z} = \{0, 1\} \) grading and a multiplication whereby \( \alpha \beta = (-1)^{\text{degree}(\alpha) \text{degree}(\beta)} \beta \alpha \). The grading for \( \Omega_M \) takes the \( p = \) even forms equal to have grading degree zero and the \( p = \) odd forms to have grading degree one. The point is that \( C^\infty(M) \) is a subalgebra of \( \Omega_M \) and the Lie derivative, \( v \to \mathcal{L}_v \) is an extension of the derivation that \( v \) defines on \( C^\infty(M) \) to give a derivation of \( \Omega_M \). This is to say that \( \mathcal{L}_v(\alpha \wedge \beta) = \mathcal{L}_v \alpha \wedge \beta + \alpha \wedge \mathcal{L}_v \beta \). You can check that this formula holds using the definition given above for \( \mathcal{L}_v \).

12d) Curvature and covariant derivatives

Suppose that \( \pi: E \to M \) is a vector bundle with fiber \( V = \mathbb{R}^n \) or \( \mathbb{C}^n \). Also suppose that \( \nabla \) is a covariant derivative on \( C^\infty(M; E) \). As explained below, there is a nice extension of \( \nabla \) that maps \( C^\infty(M; E \otimes (\wedge^p T^*M)) \) to \( C^\infty(M; E \otimes (\wedge^{p+1} T^*M)) \) for each \( p \in \{0, 1, \ldots, \dim(M)\} \). Here, as before, \( \wedge^p T^*M \) is the vector bundle \( M \times \mathbb{R}^p \), and \( \wedge^{\dim(M)+1} T^*M \) is the bundle with 0-dimensional fiber \( M \times \{0\} \). This extension is called the exterior covariant derivative, it is denoted as \( d\nabla \), and it is defined by the following rules:

- If \( \omega \) is a p-form and \( s \) is a section of \( E \), then \( d\nabla(s \omega) = \nabla s \wedge \omega + s d\omega \).
- If \( \mathfrak{w}_1 \) and \( \mathfrak{w}_2 \) are sections of \( E \otimes \wedge^p T^*M \), then \( d\nabla(\mathfrak{w}_1 + \mathfrak{w}_2) = d\nabla \mathfrak{w}_1 + d\nabla \mathfrak{w}_2 \).

These rules are sufficient to define \( d\nabla \) on all sections of \( E \otimes \wedge^p T^*M \).

Although \( d^2 = 0 \), this is generally not the case for \( d\nabla \). It is the case, however, that \( d\nabla^2 \) defines a section of \( C^\infty(M; \text{End}(E) \otimes \wedge^2 T^*M) \). This section is denoted by \( F_V \) and it is characterized by the fact that \( d\nabla^2 \mathfrak{w} = F_V \wedge \mathfrak{w} \). The section \( F_V \) is called the curvature of the covariant derivative. It measures the extent to which the covariant derivatives in different directions fail to commute. More is said about this momentarily.
What follows directly explains why $d_v^2 \omega$ can be written as $F_v \wedge \omega$ with $F_v$ an $\text{End}(E)$ valued 2-form. Here is a first justification: Suppose that $\omega$ is a smooth p-form and $s$ is a section of $E$. As $d_v(s \omega) = d_v s \wedge \omega + s \wedge d_v \omega$, so

$$d_v^2(s \omega) = d_v(d_v s \wedge \omega) + d_v(s \wedge d_v \omega).$$

Now, remember that if $\alpha$ is a 1-form, then $d(\alpha \wedge \beta) = d\alpha \wedge \beta - \alpha \wedge d\beta$. This being the case, the left most term on the right hand side is $d_v s \wedge \omega - d_v s \wedge d\omega$. Meanwhile, the right most term on the right hand side is $d_v s \wedge d\omega + s \wedge d_v^2 \omega$. Since $d_v^2 \omega$ is zero, the two terms on the right side above add to $d_v^2 s \wedge \omega$. In particular, if $\omega = f$, a function this says that $d_v^2 f s$ commutes with multiplication by a function. This understood, apply the lemma in Chapter 11b to see that $d_v^2$ is given by the action of a section of $\text{End}(E) \otimes \wedge^2 T^*M$.

This claim can be seen explicitly on a set $U \subset M$ where there is a bundle isomorphism $\phi: E|_U \to U \times V$. Let $d$ denote the covariant derivative that acts on a section $x \rightarrow v(x) = (x, v(x))$ of $U \times V$ so as to send it to the section $x \rightarrow dv|_x = (x, dv|_x)$ of the bundle $(U \times V) \otimes T^*M|_U$. Then $\phi d_v \phi^{-1} v$ acts to send the section $\nu$ to the section given by $x \rightarrow (x, dv + a v)$ where $a$ is a section of $\text{End}(E) \otimes T^*M|_U$. This understood, $\phi d_v^2 \phi^{-1} v$ is the section

$$x \rightarrow (x, (da + a \wedge a) v).$$

This identifies the curvature 2-form $F_v$ on $U$:

$$\phi(F_v) = da + a \wedge a.$$

What follows says more about the meaning of $a \wedge a$ and, more generally, $a \wedge b$ for $\text{End}(E)$-valued $p$ and $q$ forms $a$ and $b$. To start, fix any given point in $M$ and fix a basis $\{\eta_\alpha\}$ for $\wedge^p T^*M$ and a basis $\{\sigma_\beta\}$ for $\wedge^q T^*M$ at this point. Here, the index $\alpha$ runs from 1 to $d!/p!(d-p)!$ where $d = \dim(M)$. Likewise, the index $\beta$ runs from 1 to $d!/q!(d-q)!$. Write $a = \sum_{\alpha} a_\alpha \eta_\alpha$ and $b = \sum_{\beta} b_\beta \sigma_\beta$ where $\{a_\alpha\}$ and $\{b_\beta\}$ are in $\text{End}(E)$. This done, then

$$a \wedge b = \sum_{\alpha, \beta} a_\alpha b_\beta \eta_\alpha \wedge \sigma_\beta,$$

where $a_\alpha b_\beta \in \text{End}(E)$ denotes the composition of the indicated endomorphisms. (This is just matrix multiplication.) In the case where $a = b$ and both are 1-forms, then what is written above is equivalent to the formula

$$a \wedge a = \sum_{1 \leq \alpha < \beta \leq d} \{a_\alpha, a_\beta\} \eta_\alpha \wedge \eta_\beta.$$
In particular, \( a \land a = 0 \) if and only if the various components of \( a \) pairwise commute as endomorphisms of \( E \).

If \( U' \) is a second open set with an isomorphism \( \varphi': E|_{U'} \to U' \times V \), then \( \varphi' \cdot \varphi'^{-1} \) can be written as \( d + a' \). On \( U' \cap U \), the \( \text{End}(E) \) valued 1-from \( a' \) can be obtained from \( a \) by the rule

\[
a' = g^{-1}_{U,U}ag_{U,U}^{-1} + dg_{U,U} g_{U,U}^{-1}
\]

where \( g_{U,U}: U' \cap U \to \text{Gl}(V) \) is the transition function. This transformation rule is exactly what is needed so as to guarantee that

\[
da' + a' \land a' = g_{U,U}(da + a \land a)g_{U,U}^{-1}.
\]

This last formula says that the identification of \( \varphi(F_V) \) with the section

\[
(x, da + a \land a)
\]

of \( (U \times \text{Gl}(V) \otimes \wedge^2 T^* M)|_U \) does indeed specify consistently over the whole of \( M \) an \( \text{End}(E) \) valued 2-form.

12e) An example

Let \( V = \mathbb{R}^N \) or \( \mathbb{C}^N \) and suppose that \( E \to M \) is a subbundle of the product vector bundle \( M \times V \). Let \( \Pi: M \to \text{End}(V) \) denote the map that sends any given point \( x \in M \) to the orthogonal projection of \( V \) onto \( E_x \). Chapter 11c uses this setting to define a covariant derivative on sections of \( E \). By way of a reminder, this is done as follows: View a section \( s: M \to E \) as a map to \( V \) whose value at each \( x \in M \) happens to land in \( E_x \). Then define

\[
d \nabla s = d(s) = \Pi ds
\]

where \( d \) is the usual exterior derivative on maps from \( M \) to the fixed vector space \( V \). Thus, \( ds \) is a \( N \)-dimensional vector at each point whose entries are 1-forms on \( M \); any given entry is the differential of the function that defines the corresponding entry of \( s \). To compute \( d\nabla s \), use the fact that \( \Pi s = s \) to write \( d\nabla s \) as \( ds - (d\Pi)s \). Here and in what follows, \( d\Pi \) is the \( N \times N \) matrix of 1-forms that is obtained by differentiating the entries of \( \Pi \). Use this rewriting of \( d\nabla s \) with the fact that \( d^2 = 0 \) to see that

\[
d\nabla^2 s = \Pi d(\Pi) \land ds = \Pi d\Pi \land d(\Pi^2 s) = (\Pi d\Pi \land d\Pi) s.
\]
Here I have used two facts about $\Pi$. First, $\Pi d\Pi = 0$ which can be seen by first differentiating the identity $\Pi^2 = \Pi$ to get $d\Pi + \Pi d\Pi = d\Pi$ and then multiplying on the right by $\Pi$. Second, $\Pi s = s$. The preceding equation identifies the curvature 2-form:

$$F_\nu = \Pi d\Pi \wedge d\Pi.$$

Note that what is written on the right hand side here is bracketed fore and aft by $\Pi$ and so it defines a bonafide section of $\text{End}(E) \otimes \wedge^2 T^*M$.

By way of a concrete example, suppose that $M$ is an $n$-1 dimensional submanifold in $\mathbb{R}^n$. As explained in Chapter 3d, the tangent bundle $TM$ can be viewed as the set of pairs $(x, v) \in M \times \mathbb{R}^n$ with $n(x)^T v = 0$ where $n(x)$ is the unit length normal vector to $M$ at $x$. The corresponding covariant derivative is defined by the projection $\Pi = t - n n^T$ where $t$ is the $n \times n$ identity matrix. Then $\Pi d\Pi = dn n^T$ and $d\Pi = n d^T$. This understood, it follows that

$$F_\nu = dn \wedge dn^T.$$

On the face of it, this defines at each point of $M$ a skew-symmetric, $n \times n$ matrix of 1-forms whose $(i, j)$ component is $dn^i \wedge dn^j$ where $n^i$ is the $i$'th component of the normal vector with the latter viewed as a map from $M$ to $\mathbb{R}^n$. The fact that $|n| = 1$ at each point of $M$ implies that this $n \times n$ matrix of 1-forms annihilates the vector $n$. Indeed, multiplying this matrix against $n$ gives the $\mathbb{R}^n$-valued 1-form $dn \wedge (\sum_i dn^i n^i) = dn \wedge \frac{1}{2} d(|n|^2) = 0$. This being the case, what is written above does indeed define a 2-form on $M$ with values in $\text{End}(TM)$.

For example, if $M$ is the sphere $S^{n-1}$ in $\mathbb{R}^n$, then $n(x) = x$. Near the north pole where $x_n = 1$, the collection of functions $y = (x_1, \ldots, x_{n-1})$ can serve as coordinates for $S^{n-1}$. This understood, then $n(x(y))$ has $i$'th coordinate $y_i$ if $i < n$ and $n$'th coordinate $(1 - |y|^2)^{1/2}$. Thus, $dn$ is the vector of 1-forms whose $i$'th coordinate is $dy_i$ if $i < n$ and whose $n$'th coordinate is $-(1 - |y|^2)^{1/2} y_i dy_i$; where I am summing over repeated indices. Note that this vanishes at the north pole, as it should since $n$ at the north pole has $i$'th coordinate zero for $i < n$ and $n$'th coordinate 1. Meanwhile, $dn \wedge dn^T$ is apriori the $n \times n$ matrix of 2-forms whose $i, j$ entry for $i$ and $j$ both less than $n$ is $dy_i \wedge dy_j$, whose $(i, n)$ entry for $i < n$ is $-(1 - |y|^2)^{1/2} dy_i \wedge y^i dy_j$, whose $(n, i)$ entry is minus what was just written, and whose $(n, n)$ entry is zero. Note in particular that this matrix of 1-forms at the north pole has zero in its $(i, n)$ and $(n, i)$ entries for any $i$, and so defines (as it should) a 2-form with values in the endomorphisms of $TS^{n-1}$ at the north pole. This 2-form has $i, j$ entry $dy^i \wedge dy^j$.

**12f) Curvature and commutators**
Let \( v \) and \( u \) denote vector fields on \( M \). Viewed as derivations on \( C^\infty(M) \), they have a commutator, denoted \([v, u]\), which is also a derivation, and thus also a vector field on \( M \). The derivation \([v, u]\) sends a given function \( f \) to \( v(uf) - u(vf) \). Here, \( vf \) is shorthand for the action of \( v \) on the function \( f \). It is also equal to the pairing between the 1-form \( df \), a section of \( T^*M \), and the vector field \( v \), a section of \( TM \). This expression for \([v, u] = v(uf) - u(vf) \) obeys the required rule of a derivation: \( D(fg) = Df \, g + fDg \). In a local coordinate chart where \( v = v^i(x) \frac{\partial}{\partial x^i} \) and \( u = u^i(x) \frac{\partial}{\partial x^i} \), then

\[
[v, u] = (v^j \frac{\partial}{\partial x^j} (u^i)) \frac{\partial}{\partial x^i} - (u^j \frac{\partial}{\partial x^j} (v^i)) \frac{\partial}{\partial x^i} .
\]

Here, I am summing over repeated indices.

The commutator and exterior derivative are related in the following way: Let \( \alpha \) denote a 1-form on \( M \). Then

\[
(d\alpha)(v, u) = \alpha([v, u]) + v(\alpha(u)) - u(\alpha(v)) .
\]

This formula views a section of \( \wedge^p T^*M \) as a homomorphism from \( \otimes^p TM \) to \( \mathbb{R} \) which happens to be completely anti-symmetric with respect to permutations of the factors in the \( p \)-fold tensor product. For example, \( \alpha(v) \) denotes the function on \( M \) that is obtained by evaluation of \( \alpha \) on the vector field \( v \) and \( (d\alpha)(v, u) \) that obtained by evaluating \( d\alpha \) on \( v \otimes u \).

The preceding formula has the following application: Let \( E \rightarrow M \) denote a vector bundle and let \( \nabla \) denote a covariant derivative for \( C^\infty(M; E) \). Let \( v \) and \( u \) denote vector fields on \( M \), and let \( \nabla_v \) and \( \nabla_u \) denote the respective directional derivations on \( C^\infty(M; E) \) that are obtained from \( v \) and \( u \) using \( \nabla \). This is to say that if \( s \) is a section of \( E \), then \( \nabla_v s \) is the section of \( E \) that is obtained by pairing the section \( \nabla s \) of \( C^\infty(M; E \otimes T^*M) \) with \( v \). The commutator of these directional covariant derivatives is given by

\[
[\nabla_v, \nabla_u] - \nabla_{[v,u]} = F_v(v, u) .
\]

This gives another way to define \( F_v \). Here, \( F_v \) is viewed as an endomorphism from \( \otimes^2 T^*M \) to \( P \times_{ad} \text{lie}(G) \).

**12g) Connections and curvature**

Suppose that \( P \rightarrow M \) is a principal \( G \) bundle and \( A \) is a connection on \( P \). The connection defines a covariant derivative on sections of any vector bundle that is associated to \( P \) via a representation of \( G \) into a general linear group. Each such covariant derivative has an associated curvature 2-form. All of these 2-forms are derived from a single section, \( F_A \), of the bundle \((P \times_{ad} \text{lie}(G)) \otimes \wedge^2 T^*M \). The latter section is called the
Curvature 2-form or just curvature of the connection. What follows defines $F_A$ and explains how gives the curvatures of all of the various covariant derivatives.

A first way to proceed invokes what has been said about $P$ over open sets in $M$ where $P$ is isomorphic to the product principal $G$ bundle. To start, suppose that $U \subset M$ is an open set and that there is a given bundle isomorphism $\varphi: P|_{U} \rightarrow U \times G$. As explained in Part 7 of Chapter 11d, the $\text{Lie}(G)$ valued 1-form $(\varphi^{-1})^*A$ on $U \times G$ at a given point $(x, g) \in U \times G$ can be written as

$$(\varphi^{-1})^*A = g^{-1}dg + g^{-1}a_Ug,$$

where $a_U$ is a 1-form on $U$ with values in $\text{Lie}(G)$. Here, as before, I am assuming that $G$ is a subgroup of $\text{Gl}(n; \mathbb{V})$ with $\mathbb{V} = \mathbb{R}$ or $\mathbb{C}$ as the case may be. Thus, $\text{Lie}(G)$ is a subvector space of $n \times n$ matrices and $dg$ is an $n \times n$ matrix whose entries are the coordinate 1-forms on $\text{Gl}(n; \mathbb{V})$.

Now, let $V$ denote a vector space, and let $\rho: G \rightarrow \text{Gl}(V)$ denote a representation of $G$. As explained in Chapter 11e, the connection $A$ induces a covariant derivative, $\nabla_A$, on the sections of $P \times_\rho V$. Meanwhile, the isomorphism $\varphi_U$ supplies a bundle isomorphism (also denoted by $\varphi$) from $(P \times_\rho V)|_{U} \rightarrow U \times V$. Let $s$ denote a section of $P \times_\rho V$. Then $\varphi(s)$ sends $x \in U$ to a section of the form $(x, s_U(x))$, where $s_U$ is a smooth map to the vector space $V$. In addition, $\varphi(\nabla_A s)$ is the section over $U$ of $(U \times V) \otimes T^*M|_U$ given by

$$x \rightarrow (x, \rho_*(d\alpha_U + \alpha_U \wedge \alpha_U)), $$

where $\rho_*: \text{Lie}(G) \rightarrow \text{End}(V)$ is the differential of $\rho$ at the identity, $1$. It follows from this equation that the section $\varphi(F_{\nabla_A})$ of $(U \times \text{End}(V)) \otimes \wedge^2 T^*M$ is has the form

$$x \rightarrow (x, \rho_*(da_U + a_U \wedge a_U)).$$

Note in particular that how this expression for the curvature of $\nabla_A$ is using the $\text{Lie}(G)$ valued 1-form $a_U$ that defines the connection $A$.

With the preceding in mind, introduce the $\text{Lie}(G)$-valued 2-form

$$(F_A)_U = da_U + a_U \wedge a_U$$

The assignment of $x \in U$ to $(x, (F_A)_U|_x)$ is a section over $U$ of $(U \times \text{Lie}(G)) \otimes \wedge^2 T^*M$. The curvature 2-form $F_A$ can be defined by declaring that this section be its image under $\varphi$. Note in this regard that if $U' \subset M$ is another open set with an isomorphism $\varphi': P|_{U'} \rightarrow U' \times G$ and such that $U \cap U' \neq \emptyset$, then $a_{U'}$ and $a_U$ are related by the rule...
\[
\alpha_U = g_{UU} \alpha_U (g_{UU})^{-1} - dg_{UU} (g_{UU})^{-1}
\]

where \( g_{UU} = \varphi \circ \varphi^{-1} \): \( U \cap U' \rightarrow G \) is the transition function. This transformation rule guarantees that \( d\alpha_U + \alpha_U \wedge \alpha_U = g_{UU}(d\alpha_U + \alpha_U \wedge \alpha_U)g_{UU}^{-1} \). This transformation rule asserts the following: The assignment of \( F_A \) to any given open set \( U \rightarrow M \) with isomorphism \( \varphi: P|_U \rightarrow U \times G \) defines the image via \( \varphi \) of a bonafide section of \( (P \times_{ad} \text{lie}(G)) \otimes \wedge^2 T^*M \). This section is the curvature 2-form \( F_A \).

Granted this definition of \( F_A \), what was said two paragraphs back tells us that the curvature 2-form \( F_A \) of the covariant derivative \( \nabla_A \) as defined on sections of \( P \times \rho \) is equal to \( \rho_* (F_A) \).

12h) The horizontal subbundle revisited

What follows describes a definition of \( F_A \) that does not introduce isomorphisms of \( P \) with the product bundle over open sets in \( M \). To start, go up to \( P \) and reintroduce the horizontal subbundle \( H_A \subset TP \). Suppose that \( v_A \) and \( u_A \) are a pair of sections of \( H_A \). These are, of course, sections of \( TP \) also. This understood, one can consider their commutator \([v_A, u_A]\). This may or may not be a section of \( H_A \). If it lies in \( H_A \), then it’s image via \( A \) in \( \text{lie}(G) \) is zero.

Let \( v, u \) denote vector fields on \( M \) and let \( v_A, u_A \) denote horizontal lifts as sections over \( P \) as defined in Chapter 11f. This is to say that \( v_A|_{p \in P} \) is the unique lift of \( v|_{\pi(p)} \) to \( TP|_p \) that lies in the horizontal subspace \( H_A|_p \). Then the assignment of the ordered pair \((v, u) \rightarrow A([v_A, u_A]): P \rightarrow \text{lie}(G) \) has the following properties:

- \( A([v_A, u_A])|_{pg^{-1}} = g_A([u_A, v_A])g^{-1} \).
- If \( f, f' \in C^\infty(M) \), then \( A([(fv)_A, (f'u)_A]) = ff'A([v_A, u_A]) \).
- \( A([v_A, u_A]) = -A([u_A, v_A]) \).

The first properties assert the fact that the assignment of the ordered pair \((v, u) \) to \( A([v_A, u_A]) \) defines a section over \( M \) of the associated vector bundle \( P \times_{ad} \text{lie}(G) \). The second property implies that the assignment of the ordered pair \((v, u) \) to \( A([v_A, u_A]) \) does not involve the derivatives of \( v \) and \( u \). This the case, it can be viewed as the result of evaluating a section over \( M \) of \( (P \times_{ad} \text{lie}(G)) \otimes (\otimes^2 T^*M) \) on the pair \((v, u) \). The third property asserts that this section of \( (P \times_{ad} \text{lie}(G)) \otimes (\otimes^2 T^*M) \) changes sign when \( v \) and \( u \) are interchanged and so defines a section of \( (P \times_{ad} \text{lie}(G)) \otimes \wedge^2 T^*M \). This section is \(-F_A \).

To see that this section of \( (P \times_{ad} \text{lie}(G)) \otimes \wedge^2 T^*M \) is the same as that defined previously, one need only consider this just completed definition over an open set \( U \subset M \) where there is an isomorphism \( \varphi: P|_U \rightarrow U \times G \). As explained in Chapter 11f, the push-forward of the horizontal lift, \( v_A \), of a vector field \( v \) is the vector field on \( U \times G \) whose components at any given \((x, g)\) with respect to the identification \( T(U \times G) = TU \oplus \text{lie}(G) \)
are \((v, -g_a(v)g^{-1})\). This understood, it is an exercise with the definition of the Lie bracket to verify that the two definitions of \(F_A\) agree on any given pair of vectors \((v, u) \in TU\).

Additional reading