Notes On Group Actions
Manifolds, Lie Groups and Lie Algebras

Jean Gallier
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104, USA
e-mail: jean@saul.cis.upenn.edu

April 27, 2005

Contents

Chapter 1

Introduction

1.1 Motivations and Goals

The motivations for writing these notes arose while I was coteaching a seminar on Special Topics in Machine Perception with Kostas Daniilidis in the Spring 2004. The main theme of the seminar was group-theoretical methods in visual perception. In particular, Kostas decided to present some exciting results from Christopher Geyer's Ph.D. thesis [29] on scene reconstruction using two parabolic catadioptric cameras (Chapters 4 and 5). Catadioptric cameras are devices which use both mirrors (catioptric elements) and lenses (dioptric elements) to form images. Catadioptric cameras have been used in computer vision and robotics to obtain a wide field of view, often greater than 180°, unobtainable from perspective cameras. Applications of such devices include navigation, surveillance and vizualization, among others. Technically, certain matrices called *catadioptric fundamental matrices* come up. Geyer was able to give several equivalent characterizations of these matrices (see Chapter 5, Theorem 5.2). To my surprise, the Lorentz group $\mathbf{O}(3,1)$ (of the theory of special relativity) comes up naturally! The set of fundamental matrices turns out to form a manifold, $\mathcal{F}$, and the question then arises: What is the dimension of this manifold? Knowing the answer to this question is not only theoretically important but it is also practically very significant because it tells us what are the "degrees of freedom" of the problem.

Chris Geyer found an elegant and beautiful answer using some rather sophisticated concepts from the theory of group actions and Lie groups (Theorem 5.10): The space $\mathcal{F}$ is isomorphic to the quotient

$$\mathbf{O}(3,1) \times \mathbf{O}(3,1)/H_F,$$

where H_F is the stabilizer of any element, F, in $\mathcal{F}$. Now, it is easy to determine the dimension of H_F by determining the dimension of its Lie algebra, which is 3. As $\dim \mathbf{O}(3,1) = 6$, we find that $\dim \mathcal{F} = 2 \cdot 6 - 3 = 9$.

Of course, a certain amount of machinery is needed in order to understand how the above results are obtained: group actions, manifolds, Lie groups, homogenous spaces, Lorentz groups, etc. As most computer science students, even those specialized in computer vision or robotics, are not familiar with these concepts, we thought that it would be useful to give a fairly detailed exposition of these theories.

During the seminar, I also used some material from my book, Gallier [27], especially from Chapters 11, 12 and 14. Readers might find it useful to read some of this material beforehand or in parallel with these notes, especially Chapter 14, which gives a more elementary introduction to Lie groups and manifolds. In fact, during the seminar, I lectured on most of Chapter 2, but only on the "gentler" versions of Chapters 3, 4, as in [27] and not at all on Chapter 5, which was written after the course had ended.

One feature worth pointing out is that we give a complete proof of the surjectivity of the exponential map, $\exp\colon \mathfrak{so}(1,3) \to \mathbf{SO}_0(1,3)$, for the Lorentz group $\mathbf{SO}_0(3,1)$ (see Section 4.5, Theorem 4.21). Although we searched the literature quite thoroughly, we did not find a proof of this specific fact (the physics books we looked at, even the most reputable ones, seem to take this fact as obvious and there are also wrong proofs, see the Remark following Theorem 2.6). We are aware of two proofs of the surjectivity of $\exp\colon \mathfrak{so}(1,n) \to \mathbf{SO}_0(1,n)$ in the general case where where n is arbitrary: One due to Nishikawa [48] (1983) and an earlier one due to Marcel Riesz [52] (1957). In both cases, the proof is quite involved (40 pages or so). In the case of $\mathbf{SO}_0(1,3)$, a much simpler argument can be made using the fact that $\varphi\colon \mathbf{SL}(2,\mathbb{C}) \to \mathbf{SO}_0(1,3)$, is surjective and that its kernel is $\{I, -I\}$ (see Proposition 4.20). Actually, a proof of this fact is not easy to find in the literature either (and, beware there are wrong proofs, again, see the Remark following Theorem 2.6). We have made sure to provide all the steps of the proof of the surjectivity of $\exp\colon \mathfrak{so}(1,3) \to \mathbf{SO}_0(1,3)$. For more on this subject, see the discussion in Section 4.5, after Corollary 4.17.

We hope that our readers will not be put off by the level of abstraction in Chapters 3 and 5 and instead will be inspired to read more about these concepts, even fibre bundles!

Chapter 2

Review of Groups and Group Actions

2.1 Groups

Definition 2.1 A *group* is a set, G, equipped with an operation, $\cdot\colon G \times G \to G$, having the following properties: $\cdot$ is *associative*, has an *identity element*, $e \in G$, and every element in G is *invertible* (w.r.t. $\cdot$). More explicitly, this means that the following equations hold for all $a, b, c \in G$:

(G1) $a \cdot (b \cdot c) = (a \cdot b) \cdot c.$ (associativity);

(G2) $a \cdot e = e \cdot a = a.$ (identity);

(G3) For every $a \in G$, there is some $a^{-1} \in G$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$ (inverse).

A group G is *abelian* (or *commutative*) if

$$a \cdot b = b \cdot a$$

for all $a, b \in G$.

A set M together with an operation $\cdot\colon M \times M \to M$ and an element e satisfying only conditions (G1) and (G2) is called a *monoid*. For example, the set $\mathbb{N} = \{0, 1, \ldots, n \ldots\}$ of natural numbers is a (commutative) monoid. However, it is not a group.

Observe that a group (or a monoid) is never empty, since $e \in G$.

Some examples of groups are given below:

Example 2.1

1. The set $\mathbb{Z} = \{\ldots, -n, \ldots, -1, 0, 1, \ldots, n \ldots\}$ of integers is a group under addition, with identity element 0. However, $\mathbb{Z}^* = \mathbb{Z} - \{0\}$ is not a group under multiplication.

2. The set $\mathbb{Q}$ of rational numbers is a group under addition, with identity element 0. The set $\mathbb{Q}^* = \mathbb{Q} - \{0\}$ is also a group under multiplication, with identity element 1.

3. Similarly, the sets $\mathbb{R}$ of real numbers and $\mathbb{C}$ of complex numbers are groups under addition (with identity element 0), and $\mathbb{R}^* = \mathbb{R} - \{0\}$ and $\mathbb{C}^* = \mathbb{C} - \{0\}$ are groups under multiplication (with identity element 1).

4. The sets $\mathbb{R}^n$ and $\mathbb{C}^n$ of n-tuples of real or complex numbers are groups under componentwise addition:

$$(x_1, \ldots, x_n) + (y_1, \cdots, y_n) = (x_1 + y_n, \ldots, x_n + y_n),$$

with identity element $(0, \ldots, 0)$. All these groups are abelian.

5. Given any nonempty set S, the set of bijections $f \colon S \to S$, also called *permutations of S*, is a group under function composition (i.e., the multiplication of f and g is the composition $g \circ f$), with identity element the identity function id_S. This group is not abelian as soon as S has more than two elements.

6. The set of $n \times n$ matrices with real (or complex) coefficients is a group under addition of matrices, with identity element the null matrix. It is denoted by $\mathrm{M}_n(\mathbb{R})$ (or $\mathrm{M}_n(\mathbb{C})$).

7. The set $\mathbb{R}[X]$ of polynomials in one variable with real coefficients is a group under addition of polynomials.

8. The set of $n \times n$ invertible matrices with real (or complex) coefficients is a group under matrix multiplication, with identity element the identity matrix I_n. This group is called the *general linear group* and is usually denoted by $\mathbf{GL}(n, \mathbb{R})$ (or $\mathbf{GL}(n, \mathbb{C})$).

9. The set of $n \times n$ invertible matrices with real (or complex) coefficients and determinant $+1$ is a group under matrix multiplication, with identity element the identity matrix I_n. This group is called the *special linear group* and is usually denoted by $\mathbf{SL}(n, \mathbb{R})$ (or $\mathbf{SL}(n, \mathbb{C})$).

10. The set of $n \times n$ invertible matrices with real coefficients such that $RR^\top = I_n$ and of determinant $+1$ is a group called the *orthogonal group* and is usually denoted by $\mathbf{SO}(n)$ (where $R^\top$ is the *transpose* of the matrix R, i.e., the rows of $R^\top$ are the columns of R). It corresponds to the rotations in $\mathbb{R}^n$.

11. Given an open interval $]a, b[$, the set $C(]a, b[)$ of continuous functions $f \colon]a, b[\to \mathbb{R}$ is a group under the operation $f + g$ defined such that

$$(f + g)(x) = f(x) + g(x)$$

for all $x \in]a, b[$.

Given a group, G, for any two subsets $R, S \subseteq G$, we let

$$RS = \{r \cdot s \mid r \in R,\ s \in S\}.$$

In particular, for any $g \in G$, if $R = \{g\}$, we write

$$gS = \{g \cdot s \mid s \in S\}$$

and similarly, if $S = \{g\}$, we write

$$Rg = \{r \cdot g \mid r \in R\}.$$

From now on, we will drop the multiplication sign and write $g_1 g_2$ for $g_1 \cdot g_2$.

Definition 2.2 Given a group, G, a subset, H, of G is a *subgroup of* G iff

(1) The identity element, e, of G also belongs to H ($e \in H$);

(2) For all $h_1, h_2 \in H$, we have $h_1 h_2 \in H$;

(3) For all $h \in H$, we have $h^{-1} \in H$.

It is easily checked that a subset, $H \subseteq G$, is a subgroup of G iff H is nonempty and whenever $h_1, h_2 \in H$, then $h_1 h_2^{-1} \in H$.

If H is a subgroup of G and $g \in G$ is any element, the sets of the form gH are called *left cosets of* H *in* G and the sets of the form Hg are called *right cosets of* H *in* G. The left cosets (resp. right cosets) of H induce an equivalence relation, $\sim$, defined as follows: For all $g_1, g_2 \in G$,

$$g_1 \sim g_2 \quad \text{iff} \quad g_1 H = g_2 H$$

(resp. $g_1 \sim g_2$ iff $Hg_1 = Hg_2$).

Obviously, $\sim$ is an equivalence relation. Now, it is easy to see that $g_1 H = g_2 H$ iff $g_2^{-1} g_1 \in H$, so the equivalence class of an element $g \in G$ is the coset gH (resp. Hg). The set of left cosets of H in G (which, in general, is **not** a group) is denoted G/H. The "points" of G/H are obtained by "collapsing" all the elements in a coset into a single element.

It is tempting to define a multiplication operation on left cosets (or right cosets) by setting

$$(g_1 H)(g_2 H) = (g_1 g_2)H,$$

but this operation is not well defined in general, unless the subgroup H possesses a special property. This property is typical of the kernels of group homomorphisms, so we are led to

Definition 2.3 Given any two groups, G, G', a function $\varphi\colon G \to G'$ is a *homomorphism* iff

$$\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2), \quad \text{for all } g_1, g_2 \in G.$$

Taking $g_1 = g_2 = e$ (in G), we see that

$$\varphi(e) = e',$$

and taking $g_1 = g$ and $g_2 = g^{-1}$, we see that

$$\varphi(g^{-1}) = \varphi(g)^{-1}.$$

If $\varphi: G \to G'$ and $\psi: G' \to G''$ are group homomorphisms, then $\psi \circ \varphi: G \to G''$ is also a homomorphism. If $\varphi: G \to G'$ is a homomorphism of groups and $H \subseteq G$ and $H' \subseteq G'$ are two subgroups, then it is easily checked that

$$\text{Im } H = \varphi(H) = \{\varphi(g) \mid g \in H\} \quad \text{is a subgroup of } G'$$

(Im H is called the *image of H by φ*) and

$$\varphi^{-1}(H') = \{g \in G \mid \varphi(g) \in H'\} \quad \text{is a subgroup of } G.$$

In particular, when $H' = \{e'\}$, we obtain the *kernel*, Ker φ, of φ. Thus,

$$\text{Ker } \varphi = \{g \in G \mid \varphi(g) = e'\}.$$

It is immediately verified that $\varphi: G \to G'$ is injective iff Ker $\varphi = \{e\}$. (We also write Ker $\varphi = (0)$.) We say that φ is an *isomorphism* if there is a homomorphism, $\psi: G' \to G$, so that

$$\psi \circ \varphi = \text{id}_G \quad \text{and} \quad \varphi \circ \psi = \text{id}_{G'}.$$

In this case, ψ is unique and it is denoted φ^{-1}. When φ is an isomorphism we say the the groups G and G' are *isomorphic*. When $G' = G$, a group isomorphism is called an *automorphism*.

We claim that $H = \text{Ker } \varphi$ satisfies the following property:

$$gH = Hg, \quad \text{for all } g \in G. \qquad (*)$$

First, note that $(*)$ is equivalent to

$$gHg^{-1} = H, \quad \text{for all } g \in G,$$

and the above is equivalent to

$$gHg^{-1} \subseteq H, \quad \text{for all } g \in G. \qquad (**)$$

This is because $gHg^{-1} \subseteq H$ implies $H \subseteq g^{-1}Hg$, and this for all $g \in G$. But,

$$\varphi(ghg^{-1}) = \varphi(g)\varphi(h)\varphi(g^{-1}) = \varphi(g)e'\varphi(g)^{-1} = \varphi(g)\varphi(g)^{-1} = e',$$

for all $h \in H = \text{Ker } \varphi$ and all $g \in G$. Thus, by definition of $H = \text{Ker } \varphi$, we have $gHg^{-1} \subseteq H$.

Definition 2.4 For any group, G, a subgroup, $N \subseteq G$, is a *normal subgroup* of G iff

$$gNg^{-1} = N, \quad \text{for all } g \in G.$$

This is denoted by $N \lhd G$.

If N is a normal subgroup of G, the equivalence relation induced by left cosets is the same as the equivalence induced by right cosets. Furthermore, this equivalence relation, $\sim$, is a *congruence*, which means that: For all $g_1, g_2, g_1', g_2' \in G$,

(1) If $g_1N = g_1'N$ and $g_2N = g_2'N$, then $g_1g_2N = g_1'g_2'N$, and

(2) If $g_1N = g_2N$, then $g_1^{-1}N = g_2^{-1}N$.

As a consequence, we can define a group structure on the set $G/\sim$ of equivalence classes modulo $\sim$, by setting

$$(g_1N)(g_2N) = (g_1g_2)N.$$

This group is denoted G/N. The equivalence class, gN, of an element $g \in G$ is also denoted $\overline{g}$. The map $\pi\colon G \to G/N$, given by

$$\pi(g) = \overline{g} = gN,$$

is clearly a group homomorphism called the *canonical projection*.

Given a homomorphism of groups, $\varphi\colon G \to G'$, we easily check that the groups $G/\mathrm{Ker}\ \varphi$ and $\mathrm{Im}\ \varphi = \varphi(G)$ are isomorphic.

2.2 Group Actions and Homogeneous Spaces, I

If X is a set (usually, some kind of geometric space, for example, the sphere in $\mathbb{R}^3$, the upper half-plane, etc.), the "symmetries" of X are often captured by the action of a group, G, on X. In fact, if G is a Lie group and the action satisfies some simple properties, the set X can be given a manifold structure which makes it a projection (quotient) of G, a so-called "homogeneous space".

Definition 2.5 Given a set, X, and a group, G, a *left action of G on X* (for short, an *action of G on X*) is a function, $\varphi\colon G \times X \to X$, such that

(1) For all $g, h \in G$ and all $x \in X$,

$$\varphi(g, \varphi(h, x)) = \varphi(gh, x),$$

(2) For all $x \in X$,

$$\varphi(1, x) = x,$$

where $1 \in G$ is the identity element of G.

To alleviate the notation, we usually write $g \cdot x$ or even gx for $\varphi(g, x)$, in which case, the above axioms read:

(1) For all $g, h \in G$ and all $x \in X$,

$$g \cdot (h \cdot x) = gh \cdot x,$$

(2) For all $x \in X$,

$$1 \cdot x = x.$$

The set X is called a *(left) G-set*. The action φ is *faithful* or *effective* iff for any g, if $g \cdot x = x$ for all $x \in X$, then $g = 1$; the action φ is *transitive* iff for any two elements $x, y \in X$, there is some $g \in G$ so that $g \cdot x = y$.

Given an action, $\varphi\colon G \times X \to X$, for every $g \in G$, we have a function, $\varphi_g\colon X \to X$, defined by

$$\varphi_g(x) = g \cdot x, \quad \text{for all } x \in X.$$

Observe that φ_g has $\varphi_{g^{-1}}$ as inverse, since

$$\varphi_{g^{-1}}(\varphi_g(x)) = \varphi_{g^{-1}}(g \cdot x) = g^{-1} \cdot (g \cdot x) = (g^{-1}g) \cdot x = 1 \cdot x = x,$$

and similarly, $\varphi_g \circ \varphi_{g^{-1}} = \mathrm{id}$. Therefore, φ is a bijection of X, i.e., a permutation of X. Moreover, we check immediately that

$$\varphi_g \circ \varphi_h = \varphi_{gh},$$

so, the map $g \mapsto \varphi_g$ is a group homomorphism from G to $\mathfrak{S}_X$, the group of permutations of X. With a slight abuse of notation, this group homomorphism $G \longrightarrow \mathfrak{S}_X$ is also denoted φ.

Conversely, it is easy to see that any group homomorphism, $\varphi\colon G \to \mathfrak{S}_X$, yields a group action, $\cdot\colon G \times X \longrightarrow X$, by setting

$$g \cdot x = \varphi(g)(x).$$

Observe that an action, φ, is faithful iff the group homomorphism, $\varphi\colon G \to \mathfrak{S}_X$, is injective. Also, we have $g \cdot x = y$ iff $g^{-1} \cdot y = x$, since $(gh) \cdot x = g \cdot (h \cdot x)$ and $1 \cdot x = x$, for all $g, h \in G$ and all $x \in X$.

Definition 2.6 Given two G-sets, X and Y, a function, $f\colon X \to Y$, is said to be *equivariant*, or a *G-map* iff for all $x \in X$ and all $g \in G$, we have

$$f(g \cdot x) = g \cdot f(x).$$

Remark: We can also define a *right action*, $\cdot\colon X \times G \to X$, of a group G on a set X, as a map satisfying the conditions

(1) For all $g, h \in G$ and all $x \in X$,
$$(x \cdot g) \cdot h = x \cdot gh,$$

(2) For all $x \in X$,
$$x \cdot 1 = x.$$

Every notion defined for left actions is also defined for right actions, in the obvious way.

Here are some examples of (left) group actions.

Example 1: The unit sphere S^2 (more generally, S^{n-1}).

Recall that for any $n \geq 1$, the *(real) unit sphere*, S^{n-1}, is the set of points in $\mathbb{R}^n$ given by
$$S^{n-1} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 = 1\}.$$
In particular, S^2 is the usual sphere in $\mathbb{R}^3$. Since the group $\mathbf{SO}(3) = \mathbf{SO}(3, \mathbb{R})$ consists of (orientation preserving) linear isometries, i.e., *linear* maps that are distance preserving (and of determinant $+1$), and every linear map leaves the origin fixed, we see that any rotation maps S^2 into itself.

Beware that this would be false if we considered the group of *affine* isometries, $\mathbf{SE}(3)$, of $\mathbb{E}^3$. For example, a screw motion does *not* map S^2 into itself, even though it is distance preserving, because the origin is translated.

Thus, we have an action, $\cdot\colon \mathbf{SO}(3) \times S^2 \to S^2$, given by
$$R \cdot x = Rx.$$
The verification that the above is indeed an action is trivial. This action is transitive. This is because, for any two points x, y on the sphere S^2, there is a rotation whose axis is perpendicular to the plane containing x, y and the center, O, of the sphere (this plane is not unique when x and y are antipodal, i.e., on a diameter) mapping x to y.

Similarly, for any $n \geq 1$, we get an action, $\cdot\colon \mathbf{SO}(n) \times S^{n-1} \to S^{n-1}$. It is easy to show that this action is transitive.

Analogously, we can define the *(complex) unit sphere*, Σ^{n-1}, as the set of points in $\mathbb{C}^n$ given by
$$\Sigma^{n-1} = \{(z_1, \ldots, z_n) \in \mathbb{C}^n \mid z_1\overline{z}_1 + \cdots + z_n\overline{z}_n = 1\}.$$
If we write $z_j = x_j + iy_j$, with $x_j, y_j \in \mathbb{R}$, then
$$\Sigma^{n-1} = \{(x_1, \ldots, x_n, y_1, \ldots, y_n) \in \mathbb{R}^{2n} \mid x_1^2 + \cdots + x_n^2 + y_1^2 + \cdots + y_n^2 = 1\}.$$
Therefore, we can view the complex sphere, Σ^{n-1} (in $\mathbb{C}^n$), as the real sphere, S^{2n-1} (in $\mathbb{R}^{2n}$). By analogy with the real case, we can define an action, $\cdot\colon \mathbf{SU}(n) \times \Sigma^{n-1} \to \Sigma^{n-1}$, of the group, $\mathbf{SU}(n)$, of *linear* maps of $\mathbb{C}^n$ preserving the hermitian inner product (and the origin, as all linear maps do) and this action is transitive.

One should not confuse the unit sphere, Σ^{n-1}, with the hypersurface, $S^{n-1}_{\mathbb{C}}$, given by

$$S^{n-1}_{\mathbb{C}} = \{(z_1, \ldots, z_n) \in \mathbb{C}^n \mid z_1^2 + \cdots + z_n^2 = 1\}.$$

For instance, one should check that a line, L, through the origin intersects Σ^{n-1} in a circle, whereas it intersects $S^{n-1}_{\mathbb{C}}$ in exactly two points!

Example 2: The upper half-plane.

The *upper half-plane*, H, is the open subset of $\mathbb{R}^2$ consisting of all points, $(x, y) \in \mathbb{R}^2$, with $y > 0$. It is convenient to identify H with the set of complex numbers, $z \in \mathbb{C}$, such that $\Im z > 0$. Then, we can define an action, $\cdot\colon \mathbf{SL}(2, \mathbb{R}) \times H \to H$, of the group $\mathbf{SL}(2, \mathbb{R})$ on H, as follows: For any $z \in H$, for any $A \in \mathbf{SL}(2, \mathbb{R})$,

$$A \cdot z = \frac{az + b}{cz + d},$$

where

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

with $ad - bc = 1$. It is easily verified that $A \cdot z$ is indeed always well defined and in H when $z \in H$. This action is transitive (check this).

Maps of the form

$$z \mapsto \frac{az + b}{cz + d},$$

where $z \in \mathbb{C}$ and $ad - bc = 1$, are called *Möbius transformations*. Here, $a, b, c, d \in \mathbb{R}$, but in general, we allow $a, b, c, d \in \mathbb{C}$. Actually, these transformations are not necessarily defined everywhere on $\mathbb{C}$, for example, for $z = -d/c$ if $c \neq 0$. To fix this problem, we add a "point at infinity", ∞, to $\mathbb{C}$ and define Möbius transformations as functions $\mathbb{C} \cup \{\infty\} \longrightarrow \mathbb{C} \cup \{\infty\}$. If $c = 0$, the Möbius transformation sends ∞ to itself, otherwise, $-d/c \mapsto \infty$ and $\infty \mapsto a/c$. The space $\mathbb{C} \cup \{\infty\}$ can be viewed as the plane, $\mathbb{R}^2$, extended with a point at infinity. Using a stereographic projection from the sphere S^2 to the plane, (say from the north pole to the equatorial plane), we see that there is a bijection between the sphere, S^2, and $\mathbb{C} \cup \{\infty\}$. More precisely, the *stereographic projection* of the sphere S^2 from the north pole, $N = (0, 0, 1)$, to the plane $z = 0$ (extended with the point at infinity, ∞) is given by

$$(x, y, z) \in S^2 - \{(0, 0, 1)\} \mapsto \left(\frac{x}{1 - z}, \frac{y}{1 - z}\right) = \frac{x + iy}{1 - z} \in \mathbb{C}, \quad \text{with} \quad (0, 0, 1) \mapsto \infty.$$

The inverse stereographic projection is given by

$$(x, y) \mapsto \left(\frac{2x}{x^2 + y^2 + 1}, \frac{2y}{x^2 + y^2 + 1}, \frac{x^2 + y^2 - 1}{x^2 + y^2 + 1}\right), \quad \text{with} \quad \infty \mapsto (0, 0, 1).$$

Intuitively, the inverse stereographic projection "wraps" the equatorial plane around the sphere. The space $\mathbb{C} \cup \{\infty\}$ is known as the *Riemann sphere*. We will see shortly that

$\mathbb{C} \cup \{\infty\} \cong S^2$ is also the complex projective line, $\mathbb{CP}^1$. In summary, Möbius transformations are bijections of the Riemann sphere. It is easy to check that these transformations form a group under composition for all $a, b, c, d \in \mathbb{C}$, with $ad - bc = 1$. This is the *Möbius group*, denoted $\mathbf{M\ddot{o}b}^+$. The Möbius transformations corresponding to the case $a, b, c, d \in \mathbb{R}$, with $ad - bc = 1$ form a subgroup of $\mathbf{M\ddot{o}b}^+$ denoted $\mathbf{M\ddot{o}b}^+_{\mathbb{R}}$. The map from $\mathbf{SL}(2, \mathbb{C})$ to $\mathbf{M\ddot{o}b}^+$ that sends $A \in \mathbf{SL}(2, \mathbb{C})$ to the corresponding Möbius transformation is a surjective group homomorphism and one checks easily that its kernel is $\{-I, I\}$ (where I is the 2×2 identity matrix). Therefore, the Möbius group $\mathbf{M\ddot{o}b}^+$ is isomorphic to the quotient group $\mathbf{SL}(2, \mathbb{C})/\{-I, I\}$, denoted $\mathbf{PSL}(2, \mathbb{C})$. This latter group turns out to be the group of projective transformations of the projective space $\mathbb{CP}^1$. The same reasoning shows that the subgroup $\mathbf{M\ddot{o}b}^+_{\mathbb{R}}$ is isomorphic to $\mathbf{SL}(2, \mathbb{R})/\{-I, I\}$, denoted $\mathbf{PSL}(2, \mathbb{R})$.

The group $\mathbf{SL}(2, \mathbb{C})$ acts on $\mathbb{C} \cup \{\infty\} \cong S^2$ the same way that $\mathbf{SL}(2, \mathbb{R})$ acts on H, namely: For any $A \in \mathbf{SL}(2, \mathbb{C})$, for any $z \in \mathbb{C} \cup \{\infty\}$,

$$A \cdot z = \frac{az + b}{cz + d},$$

where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{with} \quad ad - bc = 1.$$

This action is clearly transitive.

One may recall from complex analysis that the (complex) Möbius transformation

$$z \mapsto \frac{z - i}{z + i}$$

is a biholomorphic isomorphism between the upper half plane, H, and the open unit disk,

$$D = \{z \in \mathbb{C} \mid |z| < 1\}.$$

As a consequence, it is possible to define a transitive action of $\mathbf{SL}(2, \mathbb{R})$ on D. This can be done in a more direct fashion, using a group isomorphic to $\mathbf{SL}(2, \mathbb{R})$, namely, $\mathbf{SU}(1, 1)$ (a group of complex matrices), but we don't want to do this right now.

Example 3: The set of $n \times n$ symmetric, positive, definite matrices, $\mathbf{SPD}(n)$.

The group $\mathbf{GL}(n) = \mathbf{GL}(n, \mathbb{R})$ acts on $\mathbf{SPD}(n)$ as follows: For all $A \in \mathbf{GL}(n)$ and all $S \in \mathbf{SPD}(n)$,

$$A \cdot S = ASA^\top.$$

It is easily checked that $ASA^\top$ is in $\mathbf{SPD}(n)$ if S is in $\mathbf{SPD}(n)$. This action is transitive because every SPD matrix, S, can be written as $S = AA^\top$, for some invertible matrix, A (prove this as an exercise).

Example 4: The projective spaces $\mathbb{RP}^n$ and $\mathbb{CP}^n$.

The *(real) projective space*, $\mathbb{RP}^n$, is the set of all lines through the origin in $\mathbb{R}^{n+1}$, i.e., the set of one-dimensional subspaces of $\mathbb{R}^{n+1}$ (where $n \geq 0$). Since a one-dimensional subspace, $L \subseteq \mathbb{R}^{n+1}$, is spanned by any nonzero vector, $u \in L$, we can view $\mathbb{RP}^n$ as the set of equivalence classes of vectors in $\mathbb{R}^{n+1} - \{0\}$ modulo the equivalence relation,

$$u \sim v \quad \text{iff} \quad v = \lambda u, \quad \text{for some} \quad \lambda \neq 0 \in \mathbb{R}.$$

In terms of this definition, there is a projection, $pr\colon \mathbb{R}^{n+1} - \{0\} \to \mathbb{RP}^n$, given by $pr(u) = [u]_\sim$, the equivalence class of u modulo $\sim$. Write $[u]$ for the line defined by the nonzero vector, u. Since every line, L, in $\mathbb{R}^{n+1}$ intersects the sphere S^n in two antipodal points, we can view $\mathbb{RP}^n$ as the quotient of the sphere S^n by identification of antipodal points. We write

$$S^n/\{I, -I\} \cong \mathbb{RP}^n.$$

We define an action of $\mathbf{SO}(n+1)$ on $\mathbb{RP}^n$ as follows: For any line, $L = [u]$, for any $R \in \mathbf{SO}(n+1)$,

$$R \cdot L = [Ru].$$

Since R is linear, the line $[Ru]$ is well defined, i.e., does not depend on the choice of $u \in L$. It is clear that this action is transitive.

The *(complex) projective space*, $\mathbb{CP}^n$, is defined analogously as the set of all lines through the origin in $\mathbb{C}^{n+1}$, i.e., the set of one-dimensional subspaces of $\mathbb{C}^{n+1}$ (where $n \geq 0$). This time, we can view $\mathbb{CP}^n$ as the set of equivalence classes of vectors in $\mathbb{C}^{n+1} - \{0\}$ modulo the equivalence relation,

$$u \sim v \quad \text{iff} \quad v = \lambda u, \quad \text{for some} \quad \lambda \neq 0 \in \mathbb{C}.$$

We have the projection, $pr\colon \mathbb{C}^{n+1} - \{0\} \to \mathbb{CP}^n$, given by $pr(u) = [u]_\sim$, the equivalence class of u modulo $\sim$. Again, write $[u]$ for the line defined by the nonzero vector, u.

Remark: Algebraic geometers write $\mathbb{P}^n_{\mathbb{R}}$ for $\mathbb{RP}^n$ and $\mathbb{P}^n_{\mathbb{C}}$ (or even $\mathbb{P}^n$) for $\mathbb{CP}^n$.

Recall that $\Sigma^n \subseteq \mathbb{C}^{n+1}$, the unit sphere in $\mathbb{C}^{n+1}$, is defined by

$$\Sigma^n = \{(z_1, \ldots, z_{n+1}) \in \mathbb{C}^{n+1} \mid z_1\overline{z}_1 + \cdots + z_{n+1}\overline{z}_{n+1} = 1\}.$$

For any line, $L = [u]$, where $u \in \mathbb{C}^{n+1}$ is a nonzero vector, writing $u = (u_1, \ldots, u_{n+1})$, a point $z \in \mathbb{C}^{n+1}$ belongs to L iff $z = \lambda(u_1, \ldots, u_{n+1})$, for some $\lambda \in \mathbb{C}$. Therefore, the intersection, $L \cap \Sigma^n$, of the line L and the sphere Σ^n is given by

$$L \cap \Sigma^n = \{\lambda(u_1, \ldots, u_{n+1}) \in \mathbb{C}^{n+1} \mid \lambda \in \mathbb{C},\ \lambda\overline{\lambda}(u_1\overline{u}_1 + \cdots + u_{n+1}\overline{u}_{n+1}) = 1\},$$

i.e.,

$$L \cap \Sigma^n = \left\{ \lambda(u_1, \ldots, u_{n+1}) \in \mathbb{C}^{n+1} \;\middle|\; \lambda \in \mathbb{C},\ |\lambda| = \frac{1}{\sqrt{|u_1|^2 + \cdots + |u_{n+1}|^2}} \right\}.$$

Thus, we see that there is a bijection between $L \cap \Sigma^n$ and the circle, S^1, i.e., geometrically, $L \cap \Sigma^n$ is a circle. Moreover, since any line, L, through the origin is determined by just one other point, we see that for any two lines L_1 and L_2 through the origin,

$$L_1 \neq L_2 \quad \text{iff} \quad (L_1 \cap \Sigma^n) \cap (L_2 \cap \Sigma^n) = \emptyset.$$

However, Σ^n is the sphere S^{2n+1} in $\mathbb{R}^{2n+2}$. It follows that $\mathbb{CP}^n$ is the quotient of S^{2n+1} by the equivalence relation, $\sim$, defined such that

$$y \sim z \quad \text{iff} \quad y, z \in L \cap \Sigma^n, \quad \text{for some line, } L, \text{ through the origin.}$$

Therefore, we can write

$$S^{2n+1}/S^1 \cong \mathbb{CP}^n.$$

The case $n = 1$ is particularly interesting, as it turns out that

$$S^3/S^1 \cong S^2.$$

This is the famous *Hopf fibration*. To show this, proceed as follows: As

$$S^3 \cong \Sigma^1 = \{(z, z') \in \mathbb{C}^2 \mid |z|^2 + |z'|^2 = 1\},$$

define a map, $\mathrm{HF}\colon S^3 \to S^2$, by

$$\mathrm{HF}((z, z')) = (2z\overline{z'}, |z|^2 - |z'|^2).$$

We leave as a homework exercise to prove that this map has range S^2 and that

$$\mathrm{HF}((z_1, z'_1)) = \mathrm{HF}((z_2, z'_2)) \quad \text{iff} \quad (z_1, z'_1) = \lambda(z_2, z'_2), \quad \text{for some } \lambda \text{ with } |\lambda| = 1.$$

In other words, for any point, $p \in S^2$, the inverse image, $\mathrm{HF}^{-1}(p)$ (also called *fibre* over p), is a circle on S^3. Consequently, S^3 can be viewed as the union of a family of disjoint circles. This is the *Hopf fibration*. It is possible to visualize the Hopf fibration using the stereographic projection from S^3 onto $\mathbb{R}^3$. This is a beautiful and puzzling picture. For example, see Berger [4]. Therefore, HF induces a bijection from $\mathbb{CP}^1$ to S^2, and it is a homeomorphism.

We define an action of $\mathbf{SU}(n+1)$ on $\mathbb{CP}^n$ as follows: For any line, $L = [u]$, for any $R \in \mathbf{SU}(n+1)$,

$$R \cdot L = [Ru].$$

Again, this action is well defined and it is transitive.

Example 5: Affine spaces.

If E is any (real) vector space and X is any set, a transitive and faithful action, $\cdot\colon E \times X \to X$, of the additive group of E on X makes X into an *affine space*. The intuition is that the members of E are translations.

Those familiar with affine spaces as in Gallier [27] (Chapter 2) or Berger [4] will point out that if X is an affine space, then, not only is the action of E on X transitive, but more is true: For any two points, $a, b \in E$, there is a *unique* vector, $u \in E$, such that $u \cdot a = b$. By the way, the action of E on X is usually considered to be a right action and is written additively, so $u \cdot a$ is written $a + u$ (the result of translating a by u). Thus, it would seem that we have to require more of our action. However, this is not necessary because E (under addition) is *abelian*. More precisely, we have the proposition

Proposition 2.1 *If G is an abelian group acting on a set X and the action $\cdot \colon G \times X \to X$ is transitive and faithful, then for any two elements $x, y \in X$, there is a unique $g \in G$ so that $g \cdot x = y$ (the action is simply transitive).*

Proof. Since our action is transitive, there is at least some $g \in G$ so that $g \cdot x = y$. Assume that we have $g_1, g_2 \in G$ with

$$g_1 \cdot x = g_2 \cdot x = y.$$

We shall prove that, actually,

$$g_1 \cdot z = g_2 \cdot z, \quad \text{for all } z \in X.$$

As our action is faithful we must have $g_1 = g_2$, and this proves our proposition.

Pick any $z \in X$. As our action is transitive, there is some $h \in G$ so that $z = h \cdot x$. Then, we have

$$\begin{aligned}
g_1 \cdot z &= g_1 \cdot (h \cdot x) \\
&= (g_1 h) \cdot x \\
&= (h g_1) \cdot x \qquad \text{(since } G \text{ is abelian)} \\
&= h \cdot (g_1 \cdot x) \\
&= h \cdot (g_2 \cdot x) \qquad \text{(since} \quad g_1 \cdot x = g_2 \cdot x) \\
&= (h g_2) \cdot x \\
&= (g_2 h) \cdot x \qquad \text{(since } G \text{ is abelian)} \\
&= g_2 \cdot (h \cdot x) \\
&= g_2 \cdot z.
\end{aligned}$$

Therefore, $g_1 \cdot z = g_2 \cdot z$, for all $z \in X$, as claimed. $\square$

More examples will be considered later.

The subset of group elements that leave some given element $x \in X$ fixed plays an important role.

Definition 2.7 Given an action, $\cdot \colon G \times X \to X$, of a group G on a set X, for any $x \in X$, the group G_x (also denoted $\mathrm{Stab}_G(x)$), called the *stabilizer* of x or *isotropy group at x* is given by

$$G_x = \{g \in G \mid g \cdot x = x\}.$$

We have to verify that G_x is indeed a subgroup of G, but this is easy. Indeed, if $g \cdot x = x$ and $h \cdot x = x$, then we also have $h^{-1} \cdot x = x$ and so, we get $gh^{-1} \cdot x = x$, proving that G_x is a subgroup of G. In general, G_x is **not** a normal subgroup.

Observe that

$$G_{g \cdot x} = gG_x g^{-1},$$

for all $g \in G$ and all $x \in X$.

Indeed,

$$\begin{aligned} G_{g \cdot x} &= \{h \in G \mid h \cdot (g \cdot x) = g \cdot x\} \\ &= \{h \in G \mid hg \cdot x = g \cdot x\} \\ &= \{h \in G \mid g^{-1}hg \cdot x = x\} \\ &= gG_x g^{-1}. \end{aligned}$$

Therefore, the stabilizers of x and $g \cdot x$ are conjugate of each other.

When the action of G on X is transitive, for any fixed $x \in G$, the set X is a quotient (as set, not as group) of G by G_x. Indeed, we can define the map, $\pi_x \colon G \to X$, by

$$\pi_x(g) = g \cdot x, \quad \text{for all } g \in G.$$

Observe that

$$\pi_x(gG_x) = (gG_x) \cdot x = g \cdot (G_x \cdot x) = g \cdot x = \pi_x(g).$$

This shows that $\pi_x \colon G \to X$ induces a quotient map, $\overline{\pi}_x \colon G/G_x \to X$, from the set, G/G_x, of (left) cosets of G_x to X, defined by

$$\overline{\pi}_x(gG_x) = g \cdot x.$$

Since

$$\pi_x(g) = \pi_x(h) \quad \text{iff} \quad g \cdot x = h \cdot x \quad \text{iff} \quad g^{-1}h \cdot x = x \quad \text{iff} \quad g^{-1}h \in G_x \quad \text{iff} \quad gG_x = hG_x,$$

we deduce that $\overline{\pi}_x \colon G/G_x \to X$ is injective. However, since our action is transitive, for every $y \in X$, there is some $g \in G$ so that $g \cdot x = y$ and so, $\overline{\pi}_x(gG_x) = g \cdot x = y$, i.e., the map $\overline{\pi}_x$ is also surjective. Therefore, the map $\overline{\pi}_x \colon G/G_x \to X$ is a bijection (of sets, not groups). The map $\pi_x \colon G \to X$ is also surjective. Let us record this important fact as

Proposition 2.2 *If $\cdot \colon G \times X \to X$ is a transitive action of a group G on a set X, for every fixed $x \in X$, the surjection, $\pi \colon G \to X$, given by*

$$\pi(g) = g \cdot x$$

induces a bijection

$$\overline{\pi} \colon G/G_x \to X,$$

where G_x is the stabilizer of x.

The map $\pi \colon G \to X$ (corresponding to a fixed $x \in X$) is sometimes called a *projection* of G onto X. Proposition 2.2 shows that for every $y \in X$, the subset, $\pi^{-1}(y)$, of G (called the *fibre above* y) is equal to some coset, gG_x, of G and thus, is in bijection with the group G_x itself. We can think of G as a moving family of fibres, G_x, parametrized by X. This point of view of viewing a space as a moving family of simpler spaces is typical in (algebraic) geometry, and underlies the notion of (principal) fibre bundle.

Note that if the action $\cdot \colon G \times X \to X$ is transitive, then the stabilizers G_x and G_y of any two elements $x, y \in X$ are isomorphic, as they as conjugates. Thus, in this case, it is enough to compute one of these stabilizers for a "convenient" x.

As the situation of Proposition 2.2 is of particular interest, we make the following definition:

Definition 2.8 A set, X, is said to be a *homogeneous space* if there is a transitive action, $\cdot \colon G \times X \to X$, of some group, G, on X.

We see that all the spaces of Example 1–5 are homogeneous spaces. Another example that will play an important role when we deal with Lie groups is the situation where we have a group, G, a subgroup, H, of G (not necessarily normal) and where $X = G/H$, the set of left cosets of G modulo H. The group G acts on G/H by left multiplication:

$$a \cdot (gH) = (ag)H,$$

where $a, g \in G$. This action is clearly transitive and one checks that the stabilizer of gH is gHg^{-1}. If G is a topological group and H is a closed subgroup of G (see later for an explanation), it turns out that G/H is Hausdorff (Recall that a topological space, X, is *Hausdorff* iff for any two distinct points $x \neq y \in X$, there exists two disjoint open subsets, U and V, with $x \in U$ and $y \in V$.) If G is a Lie group, we obtain a manifold.

Even if G and X are topological spaces and the action, $\cdot \colon G \times X \to X$, is continuous, the space G/G_x under the quotient topology is, in general, **not** homeomorphic to X.

We will give later sufficient conditions that insure that X is indeed a topological space or even a manifold. In particular, X will be a manifold when G is a Lie group.

In general, an action $\cdot \colon G \times X \to X$ is not transitive on X, but for every $x \in X$, it is transitive on the set

$$O(x) = G \cdot x = \{g \cdot x \mid g \in G\}.$$

Such a set is called the *orbit* of x. The orbits are the equivalence classes of the following equivalence relation:

Definition 2.9 Given an action, $\cdot \colon G \times X \to X$, of some group, G, on X, the equivalence relation, $\sim$, on X is defined so that, for all $x, y \in X$,

$$x \sim y \quad \text{iff} \quad y = g \cdot x, \quad \text{for some } g \in G.$$

For every $x \in X$, the equivalence class of x is the *orbit of* x, denoted $O(x)$ or $\mathrm{Orb}_G(x)$, with

$$O(x) = \{g \cdot x \mid g \in G\}.$$

The set of orbits is denoted X/G.

The orbit space, X/G, is obtained from X by an identification (or merging) process: For every orbit, all points in that orbit are merged into a single point. For example, if $X = S^2$ and G is the group consisting of the restrictions of the two linear maps I and $-I$ of $\mathbb{R}^3$ to S^2 (where $-I(x, y, z) = (-x, -y, -z)$), then

$$X/G = S^2/\{I, -I\} \cong \mathbb{RP}^2.$$

Many manifolds can be obtained in this fashion, including the torus, the Klein bottle, the Möbius band, etc.

Since the action of G is transitive on $O(x)$, by Proposition 2.2, we see that for every $x \in X$, we have a bijection

$$O(x) \cong G/G_x.$$

As a corollary, if both X and G are finite, for any set, $A \subseteq X$, of representatives from every orbit, we have the *orbit formula*:

$$|X| = \sum_{a \in A} [G \colon G_x] = \sum_{a \in A} |G|/|G_x|.$$

Even if a group action, $\cdot \colon G \times X \to X$, is not transitive, when X is a manifold, we can consider the set of orbits, X/G, and if the action of G on X satisfies certain conditions, X/G is actually a manifold. Manifolds arising in this fashion are often called *orbifolds*. In summary, we see that manifolds arise in at least two ways from a group action:

(1) As homogeneous spaces, G/G_x, if the action is transitive.

(2) As orbifolds, X/G.

Of course, in both cases, the action must satisfy some additional properties.

Let us now determine some stabilizers for the actions of Examples 1–4, and for more examples of homogeneous spaces.

(a) Consider the action, $\cdot \colon \mathbf{SO}(n) \times S^{n-1} \to S^{n-1}$, of $\mathbf{SO}(n)$ on the sphere S^{n-1} $(n \geq 1)$ defined in Example 1. Since this action is transitive, we can determine the stabilizer of any convenient element of S^{n-1}, say $e_1 = (1, 0, \ldots, 0)$. In order for any $R \in \mathbf{SO}(n)$ to leave e_1 fixed, the first column of R must be e_1, so R is an orthogonal matrix of the form

$$R = \begin{pmatrix} 1 & U \\ 0 & S \end{pmatrix}, \quad \text{with} \quad \det(S) = 1.$$

As the rows of R must be unit vector, we see that $U = 0$ and $S \in \mathbf{SO}(n-1)$. Therefore, the stabilizer of e_1 is isomorphic to $\mathbf{SO}(n-1)$, and we deduce the bijection

$$\mathbf{SO}(n)/\mathbf{SO}(n-1) \cong S^{n-1}.$$

Strictly speaking, $\mathbf{SO}(n-1)$ is not a subgroup of $\mathbf{SO}(n)$ and in all rigor, we should consider the subgroup, $\widetilde{\mathbf{SO}}(n-1)$, of $\mathbf{SO}(n)$ consisting of all matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & S \end{pmatrix}, \quad \text{with} \quad \det(S) = 1$$

and write

$$\mathbf{SO}(n)/\widetilde{\mathbf{SO}}(n-1) \cong S^{n-1}.$$

However, it is common practice to identify $\mathbf{SO}(n-1)$ with $\widetilde{\mathbf{SO}}(n-1)$.

When $n = 2$, as $\mathbf{SO}(1) = \{1\}$, we find that $\mathbf{SO}(2) \cong S^1$, a circle, a fact that we already knew. When $n = 3$, we find that $\mathbf{SO}(3)/\mathbf{SO}(2) \cong S^2$. This says that $\mathbf{SO}(3)$ is somehow the result of glueing circles to the surface of a sphere (in $\mathbb{R}^3$), in such a way that these circles do not intersect. This is hard to visualize!

A similar argument for the complex unit sphere, Σ^{n-1}, shows that

$$\mathbf{SU}(n)/\mathbf{SU}(n-1) \cong \Sigma^{n-1} \cong S^{2n-1}.$$

Again, we identify $\mathbf{SU}(n-1)$ with a subgroup of $\mathbf{SU}(n)$, as in the real case. In particular, when $n = 2$, as $\mathbf{SU}(1) = \{1\}$, we find that

$$\mathbf{SU}(2) \cong S^3,$$

i.e., the group $\mathbf{SU}(2)$ is topologically the sphere S^3! Actually, this is not surprising if we remember that $\mathbf{SU}(2)$ is in fact the group of unit quaternions.

(b) We saw in Example 2 that the action, $\cdot\colon \mathbf{SL}(2,\mathbb{R}) \times H \to H$, of the group $\mathbf{SL}(2,\mathbb{R})$ on the upper half plane is transitive. Let us find out what the stabilizer of $z = i$ is. We should have

$$\frac{ai+b}{ci+d} = i,$$

that is, $ai + b = -c + di$, i.e.,

$$(d-a)i = b + c.$$

Since a, b, c, d are real, we must have $d = a$ and $b = -c$. Moreover, $ad - bc = 1$, so we get $a^2 + b^2 = 1$. We conclude that a matrix in $\mathbf{SL}(2,\mathbb{R})$ fixes i iff it is of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{with} \quad a^2 + b^2 = 1.$$

Clearly, these are the rotation matrices in $\mathbf{SO}(2)$ and so, the stabilizer of i is $\mathbf{SO}(2)$. We conclude that

$$\mathbf{SL}(2,\mathbb{R})/\mathbf{SO}(2) \cong H.$$

This time, we can view $\mathbf{SL}(2,\mathbb{R})$ as the result of glueing circles to the upper half plane. This is not so easy to visualize. There is a better way to visualize the topology of $\mathbf{SL}(2,\mathbb{R})$ by making it act on the open disk, D. We will return to this action in a little while.

Now, consider the action of $\mathbf{SL}(2,\mathbb{C})$ on $\mathbb{C} \cup \{\infty\} \cong S^2$. As it is transitive, let us find the stabilizer of $z = 0$. We must have

$$\frac{b}{d} = 0,$$

and as $ad - bc = 1$, we must have $b = 0$ and $ad = 1$. Thus, the stabilizer of 0 is the subgroup, $\mathbf{SL}(2,\mathbb{C})_0$, of $\mathbf{SL}(2,\mathbb{C})$ consisting of all matrices of the form

$$\begin{pmatrix} a & 0 \\ c & a^{-1} \end{pmatrix}, \quad \text{where} \quad a \in \mathbb{C} - \{0\} \quad \text{and} \quad c \in \mathbb{C}.$$

We get

$$\mathbf{SL}(2,\mathbb{C})/\mathbf{SL}(2,\mathbb{C})_0 \cong \mathbb{C} \cup \{\infty\} \cong S^2,$$

but this is not very illuminating.

(c) In Example 3, we considered the action, $\cdot\colon \mathbf{GL}(n) \times \mathbf{SPD}(n) \to \mathbf{SPD}(n)$, of $\mathbf{GL}(n)$ on $\mathbf{SPD}(n)$, the set of symmetric positive definite matrices. As this action is transitive, let us find the stabilizer of I. For any $A \in \mathbf{GL}(n)$, the matrix A stabilizes I iff

$$AIA^{\top} = AA^{\top} = I.$$

Therefore, the stabilizer of I is $\mathbf{O}(n)$ and we find that

$$\mathbf{GL}(n)/\mathbf{O}(n) = \mathbf{SPD}(n).$$

Observe that if $\mathbf{GL}^+(n)$ denotes the subgroup of $\mathbf{GL}(n)$ consisting of all matrices with a strictly positive determinant, then we have an action $\cdot\colon \mathbf{GL}^+(n) \times \mathbf{SPD}(n) \to \mathbf{SPD}(n)$ of $\mathbf{GL}^+(n)$ on $\mathbf{SPD}(n)$. This action is transtive and we find that the stabilizer of I is $\mathbf{SO}(n)$; consequently, we get

$$\mathbf{GL}^+(n)/\mathbf{SO}(n) = \mathbf{SPD}(n).$$

(d) In Example 4, we considered the action, $\cdot\colon \mathbf{SO}(n+1) \times \mathbb{RP}^n \to \mathbb{RP}^n$, of $\mathbf{SO}(n+1)$ on the (real) projective space, $\mathbb{RP}^n$. As this action is transitive, let us find the stabilizer of the line, $L = [e_1]$, where $e_1 = (1, 0, \ldots, 0)$. For any $R \in \mathbf{SO}(n+1)$, the line L is fixed iff either $R(e_1) = e_1$ or $R(e_1) = -e_1$, since e_1 and $-e_1$ define the same line. As R is orthogonal with $\det(R) = 1$, this means that R is of the form

$$R = \begin{pmatrix} \alpha & 0 \\ 0 & S \end{pmatrix}, \quad \text{with} \quad \alpha = \pm 1 \quad \text{and} \quad \det(S) = \alpha.$$

But, S must be orthogonal, so we conclude $S \in \mathbf{O}(n)$. Therefore, the stabilizer of $L = [e_1]$ is isomorphic to the group $\mathbf{O}(n)$ and we find that

$$\mathbf{SO}(n+1)/\mathbf{O}(n) \cong \mathbb{RP}^n.$$

Strictly speaking, $\mathbf{O}(n)$ is not a subgroup of $\mathbf{SO}(n+1)$, so the above equation does not make sense. We should write

$$\mathbf{SO}(n+1)/\widetilde{\mathbf{O}}(n) \cong \mathbb{RP}^n,$$

where $\widetilde{\mathbf{O}}(n)$ is the subgroup of $\mathbf{SO}(n+1)$ consisting of all matrices of the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & S \end{pmatrix}, \quad \text{with} \quad S \in \mathbf{O}(n),\ \alpha = \pm 1 \quad \text{and} \quad \det(S) = \alpha.$$

However, the common practice is to write $\mathbf{O}(n)$ instead of $\widetilde{\mathbf{O}}(n)$.

We should mention that $\mathbb{RP}^3$ and $\mathbf{SO}(3)$ are homeomorphic spaces. This is shown using the quaternions, for example, see Gallier [27], Chapter 8.

A similar argument applies to the action, $\cdot\colon \mathbf{SU}(n+1) \times \mathbb{CP}^n \to \mathbb{CP}^n$, of $\mathbf{SU}(n+1)$ on the (complex) projective space, $\mathbb{CP}^n$. We find that

$$\mathbf{SU}(n+1)/\mathbf{U}(n) \cong \mathbb{CP}^n.$$

Again, the above is a bit sloppy as $\mathbf{U}(n)$ is not a subgroup of $\mathbf{SU}(n+1)$. To be rigorous, we should use the subgroup, $\widetilde{\mathbf{U}}(n)$, consisting of all matrices of the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & S \end{pmatrix}, \quad \text{with} \quad S \in \mathbf{U}(n),\ |\alpha| = 1 \quad \text{and} \quad \det(S) = \overline{\alpha}.$$

The common practice is to write $\mathbf{U}(n)$ instead of $\widetilde{\mathbf{U}}(n)$. In particular, when $n = 1$, we find that

$$\mathbf{SU}(2)/\mathbf{U}(1) \cong \mathbb{CP}^1.$$

But, we know that $\mathbf{SU}(2) \cong S^3$ and, clearly, $\mathbf{U}(1) \cong S^1$. So, again, we find that $S^3/S^1 \cong \mathbb{CP}^1$ (but we know, more, namely, $S^3/S^1 \cong S^2 \cong \mathbb{CP}^1$.)

(e) We now consider a generalization of projective spaces (real and complex). First, consider the real case. Given any $n \geq 1$, for any k, with $0 \leq k \leq n$, let $G(k, n)$ be the set of all linear k-dimensional subspaces of $\mathbb{R}^n$ (also called *k-planes*). Any k-dimensional subspace, U, of $\mathbb{R}$ is spanned by k linearly independent vectors, $u_1, \ldots, u_k$, in $\mathbb{R}^n$; write $U = \mathrm{span}(u_1, \ldots, u_k)$. We can define an action, $\cdot\colon \mathbf{O}(n) \times G(k, n) \to G(k, n)$, as follows: For any $R \in \mathbf{O}(n)$, for any $U = \mathrm{span}(u_1, \ldots, u_k)$, let

$$R \cdot U = \mathrm{span}(Ru_1, \ldots, Ru_k).$$

We have to check that the above is well defined. If $U = \mathrm{span}(v_1, \ldots, v_k)$ for any other k linearly independent vectors, $v_1, \ldots, v_k$, we have

$$v_i = \sum_{j=1}^{k} a_{ij} u_j, \quad 1 \leq i \leq k,$$

for some $a_{ij} \in \mathbb{R}$, and so,

$$Rv_i = \sum_{j=1}^{k} a_{ij} Ru_j, \quad 1 \leq i \leq k,$$

which shows that

$$\mathrm{span}(Ru_1, \ldots, Ru_k) = \mathrm{span}(Rv_1, \ldots, Rv_k),$$

i.e., the above action is well defined. This action is transitive. This is because if U and V are any two k-planes, we may assume that $U = \mathrm{span}(u_1, \ldots, u_k)$ and $V = \mathrm{span}(v_1, \ldots, v_k)$, where the u_i's form an orthonormal family and similarly for the v_i's. Then, we can extend these families to orthonormal bases $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$ or $\mathbb{R}^n$, and w.r.t. the orthonormal basis $(u_1, \ldots, u_n)$, the matrix of the linear map sending u_i to v_i is orthogonal. Thus, it is enough to find the stabilizer of any k-plane. Pick $U = \mathrm{span}(e_1, \ldots, e_k)$, where $(e_1, \ldots, e_n)$ is the canonical basis of $\mathbb{R}^n$ (i.e., $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, with the 1 in the ith position). Now, any $R \in \mathbf{O}(n)$ stabilizes U iff R maps $e_1, \ldots, e_k$ to k linearly independent vectors in the subspace $U = \mathrm{span}(e_1, \ldots, e_k)$, i.e., R is of the form

$$R = \begin{pmatrix} S & 0 \\ 0 & T \end{pmatrix},$$

where S is $k \times k$ and T is $(n-k) \times (n-k)$. Moreover, as R is orthogonal, S and T must be orthogonal, i.e., $S \in \mathbf{O}(k)$ and $T \in \mathbf{O}(n-k)$. We deduce that the stabilizer of U is isomorphic to $\mathbf{O}(k) \times \mathbf{O}(n-k)$ and we find that

$$\mathbf{O}(n)/(\mathbf{O}(k) \times \mathbf{O}(n-k)) \cong G(k, n).$$

It turns out that this makes $G(k, n)$ into a smooth manifold of dimension $k(n-k)$ called a *Grassmannian*.

If we recall the projection $pr\colon \mathbb{R}^{n+1} - \{0\} \to \mathbb{RP}^n$, by definition, a *k-plane* in $\mathbb{RP}^n$ is the image under pr of any $(k+1)$-plane in $\mathbb{R}^{n+1}$. So, for example, a line in $\mathbb{RP}^n$ is the image of a 2-plane in $\mathbb{R}^{n+1}$, and a hyperplane in $\mathbb{RP}^n$ is the image of a hyperplane in $\mathbb{R}^{n+1}$. The advantage of this point of view is that the k-planes in $\mathbb{RP}^n$ are arbitrary, i.e., they do not have to go through "the origin" (which does not make sense, anyway!). Then, we see that we can interpret the Grassmannian, $G(k+1, n+1)$, as a space of "parameters" for the k-planes in $\mathbb{RP}^n$. For example, $G(2, n+1)$ parametrizes the lines in $\mathbb{RP}^n$. In this viewpoint, $G(k+1, n+1)$ is usually denoted $\mathbb{G}(k, n)$.

It can be proved (using some exterior algebra) that $G(k, n)$ can be embedded in $\mathbb{RP}^{\binom{n}{k}-1}$. Much more is true. For example, $G(k, n)$ is a projective variety, which means that it can be defined as a subset of $\mathbb{RP}^{\binom{n}{k}-1}$ equal to the zero locus of a set of homogeneous equations. There is even a set of quadratic equations, known as the *Plücker equations*, defining $G(k,n)$. In particular, when $n=4$ and $k=2$, we have $G(2,4)\subseteq \mathbb{RP}^5$ and $G(2,4)$ is defined by a single equation of degree 2. The Grassmannian $G(2,4)=\mathbb{G}(1,3)$ is known as the *Klein quadric*. This hypersurface in $\mathbb{RP}^5$ parametrizes the lines in $\mathbb{RP}^3$.

Complex Grassmannians are defined in a similar way, by replacing $\mathbb{R}$ by $\mathbb{C}$ throughout. The complex Grassmannian, $G_{\mathbb{C}}(k,n)$, is a complex manifold as well as a real manifold and we have

$$\mathbf{U}(n)/(\mathbf{U}(k)\times\mathbf{U}(n-k))\cong G_{\mathbb{C}}(k,n).$$

We now return to case (b) to give a better picture of $\mathbf{SL}(2,\mathbb{R})$. Instead of having $\mathbf{SL}(2,\mathbb{R})$ act on the upper half plane we define an action of $\mathbf{SL}(2,\mathbb{R})$ on the open unit disk, D. Technically, it is easier to consider the group, $\mathbf{SU}(1,1)$, which is isomorphic to $\mathbf{SL}(2,\mathbb{R})$, and to make $\mathbf{SU}(1,1)$ act on D. The group $\mathbf{SU}(1,1)$ is the group of 2×2 complex matrices of the form

$$\begin{pmatrix} a & b\\ \overline{b} & \overline{a}\end{pmatrix},\quad \text{with}\quad a\overline{a}-b\overline{b}=1.$$

The reader should check that if we let

$$g=\begin{pmatrix} 1 & -i\\ 1 & i\end{pmatrix},$$

then the map from $\mathbf{SL}(2,\mathbb{R})$ to $\mathbf{SU}(1,1)$ given by

$$A\mapsto gAg^{-1}$$

is an isomorphism. Observe that the Möbius transformation associated with g is

$$z\mapsto \frac{z-i}{z+1},$$

which is the holomorphic isomorphism mapping H to D mentionned earlier! Now, we can define a bijection between $\mathbf{SU}(1,1)$ and $S^1\times D$ given by

$$\begin{pmatrix} a & b\\ \overline{b} & \overline{a}\end{pmatrix}\mapsto (a/|a|, b/a).$$

We conclude that $\mathbf{SL}(2,\mathbb{R})\cong\mathbf{SU}(1,1)$ is topologically an open solid torus (i.e., with the surface of the torus removed). It is possible to further classify the elements of $\mathbf{SL}(2,\mathbb{R})$ into three categories and to have geometric interpretations of these as certain regions of the torus. For details, the reader should consult Carter, Segal and Macdonald [14] or Duistermatt and Kolk [25] (Chapter 1, Section 1.2).

The group $\mathbf{SU}(1,1)$ acts on D by interpreting any matrix in $\mathbf{SU}(1,1)$ as a Möbius tranformation, i.e.,

$$\begin{pmatrix} a & b\\ \overline{b} & \overline{a}\end{pmatrix}\mapsto \left(z\mapsto \frac{az+b}{\overline{b}z+\overline{a}}\right).$$

The reader should check that these transformations preserve D. Both the upper half-plane and the open disk are models of Lobachevsky's non-Euclidean geometry (where the parallel postulate fails). They are also models of hyperbolic spaces (Riemannian manifolds with constant negative curvature, see Gallot, Hulin and Lafontaine [28], Chapter III). According to Dubrovin, Fomenko, and Novikov [23] (Chapter 2, Section 13.2), the open disk model is due to Poincaré and the upper half-plane model to Klein, although Poincaré was the first to realize that the upper half-plane is a hyperbolic space.

2.3 The Lorentz Groups $\mathbf{O}(n,1)$, $\mathbf{SO}(n,1)$ and $\mathbf{SO}_0(n,1)$

The Lorentz group provides another interesting example. Moreover, the Lorentz group $\mathbf{SO}(3,1)$ shows up in an interesting way in computer vision.

Denote the $p \times p$-identity matrix by I_p, for $p, q, \geq 1$, and define

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

If $n = p + q$, the matrix $I_{p,q}$ is associated with the nondegenerate symmetric bilinear form

$$\varphi_{p,q}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) = \sum_{i=1}^{p} x_i y_i - \sum_{j=p+1}^{n} x_j y_j$$

with associated quadratic form

$$\Phi_{p,q}((x_1, \ldots, x_n)) = \sum_{i=1}^{p} x_i^2 - \sum_{j=p+1}^{n} x_j^2.$$

In particular, when $p = 1$ and $q = 3$, we have the *Lorentz metric*

$$x_1^2 - x_2^2 - x_3^2 - x_4^2.$$

In physics, x_1 is interpreted as time and written t and x_2, x_3, x_4 as coordinates in $\mathbb{R}^3$ and written x, y, z. Thus, the Lozentz metric is usually written a

$$t^2 - x^2 - y^2 - z^2,$$

although it also appears as

$$x^2 + y^2 + z^2 - t^2,$$

which is equivalent but slightly less convenient for certain purposes, as we will see later. The space $\mathbb{R}^4$ with the Lorentz metric is called *Minkowski space*. It plays an important role in Einstein's theory of special relativity.

The group $\mathbf{O}(p,q)$ is the set of all $n \times n$-matrices

$$\mathbf{O}(p,q) = \{A \in \mathbf{GL}(n,\mathbb{R}) \mid A^\top I_{p,q} A = I_{p,q}\}.$$

This is the group of all invertible linear maps of $\mathbb{R}^n$ that preserve the quadratic form, $\Phi_{p,q}$, i.e., the group of isometries of $\Phi_{p,q}$. Clearly, $I_{p,q}^2 = I$, so the condition $A^\top I_{p,q} A = I_{p,q}$ is equivalent to $I_{p,q} A^\top I_{p,q} A = I$, which means that

$$A^{-1} = I_{p,q} A^\top I_{p,q}.$$

Thus, $A I_{p,q} A^\top = I_{p,q}$ also holds, which shows that $\mathbf{O}(p,q)$ is closed under transposition (i.e., if $A \in \mathbf{O}(p,q)$, then $A^\top \in \mathbf{O}(p,q)$). We have the subgroup

$$\mathbf{SO}(p,q) = \{A \in \mathbf{O}(p,q) \mid \det(A) = 1\}$$

consisting of the isometries of $(\mathbb{R}^n, \Phi_{p,q})$ with determinant $+1$. It is clear that $\mathbf{SO}(p,q)$ is also closed under transposition. The condition $A^\top I_{p,q} A = I_{p,q}$ has an interpretation in terms of the inner product $\varphi_{p,q}$ and the columns (and rows) of A. Indeed, if we denote the jth column of A by A_j, then

$$A^\top I_{p,q} A = (\varphi_{p,q}(A_i, A_j)),$$

so $A \in \mathbf{O}(p,q)$ iff the columns of A form an "orthonormal basis" w.r.t. $\varphi_{p,q}$, i.e.,

$$\varphi_{p,q}(A_i, A_j) = \begin{cases} \delta_{ij} & \text{if } 1 \leq i, j \leq p; \\ -\delta_{ij} & \text{if } p+1 \leq i, j \leq p+q. \end{cases}$$

The difference with the usual orthogonal matrices is that $\varphi_{p,q}(A_i, A_i) = -1$, if $p + 1 \leq i \leq p + q$. As $\mathbf{O}(p,q)$ is closed under transposition, the rows of A also form an orthonormal basis w.r.t. $\varphi_{p,q}$.

It turns out that $\mathbf{SO}(p,q)$ has two connected components and the component containing the identity is a subgroup of $\mathbf{SO}(p,q)$ denoted $\mathbf{SO}_0(p,q)$. The group $\mathbf{SO}_0(p,q)$ turns out to be homeomorphic to $\mathbf{SO}(p) \times \mathbf{SO}(q) \times \mathbb{R}^{pq}$, but this is not easy to prove. (One way to prove it is to use results on pseudo-algebraic subgroups of $\mathbf{GL}(n,\mathbb{C})$, see Knapp [36] or Gallier's notes on Clifford algebras (on the web)).

We will now determine the polar decomposition and the SVD decomposition of matrices in the Lorentz groups $\mathbf{O}(n,1)$ and $\mathbf{SO}(n,1)$. Write $J = I_{n,1}$ and, given any $A \in \mathbf{O}(n,1)$, write

$$A = \begin{pmatrix} B & u \\ v^\top & c \end{pmatrix},$$

where B is an $n \times n$ matrix, u, v are (column) vectors in $\mathbb{R}^n$ and $c \in \mathbb{R}$. We begin with the polar decomposition of matrices in the Lorentz groups $\mathbf{O}(n,1)$.

Proposition 2.3 *Every matrix $A \in \mathbf{O}(n,1)$ has a polar decomposition of the form*

$$A = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

where $Q \in \mathbf{O}(n)$ and $c = \sqrt{\|v\|^2 + 1}$.

Proof. Write A in block form as above. As the condition for A to be in $\mathbf{O}(n,1)$ is $A^\top JA = J$, we get

$$\begin{pmatrix} B^\top & v \\ u^\top & c \end{pmatrix}\begin{pmatrix} B & u \\ -v^\top & -c \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ 0 & -1 \end{pmatrix},$$

i.e,.

$$\begin{aligned} B^\top B &= I + vv^\top \\ u^\top u &= c^2 - 1 \\ B^\top u &= cv. \end{aligned}$$

If we remember that we also have $AJA^\top = J$, then

$$Bv = cu,$$

which can also be deduced from the three equations above. From $u^\top u = \|u\|^2 = c^2 - 1$, we deduce that $|c| \geq 1$, and from $B^\top B = I + vv^\top$, we deduce that $B^\top B$ is symmetric, positive definite. Now, geometrically, it is well known that $vv^\top / v^\top v$ is the orthogonal projection onto the line determined by v. Consequently, the kernel of $vv^\top$ is the orthogonal complement of v and $vv^\top$ has the eigenvalue 0 with multiplicity $n-1$ and the eigenvalue $c^2 - 1 = \|v\|^2 = v^\top v$ with multiplicity 1. The eigenvectors associated with 0 are orthogonal to v and the eigenvectors associated with $c^2 - 1$ are proportional with v. It follows that $I + vv^\top$ has the eigenvalue 1 with multiplicity $n-1$ and the eigenvalue c^2 with multiplicity 1, the eigenvectors being as before. Now, B has polar form $B = QS_1$, where Q is orthogonal and S_1 is symmetric positive definite and $S_1^2 = B^\top B = I + vv^\top$. Therefore, if $c > 0$, then $S_1 = \sqrt{I + vv^\top}$ is a symmetric positive definite matrix with eigenvalue 1 with multiplicity $n-1$ and eigenvalue c with multiplicity 1, the eigenvectors being as before. If $c < 0$, then change c to $-c$.

Case 1: $c > 0$. Then, v is an eigenvector of S_1 for c and we must also have $Bv = cu$, which implies

$$Bv = QS_1 v = Q(cv) = cQv = cu,$$

so

$$Qv = u.$$

It follows that

$$A = \begin{pmatrix} B & u \\ v^\top & c \end{pmatrix} = \begin{pmatrix} QS_1 & Qv \\ v^\top & c \end{pmatrix} = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix}.$$

Therefore, the polar decomposition of $A \in \mathbf{O}(n,1)$ is

$$A = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

where $Q \in \mathbf{O}(n)$ and $c = \sqrt{\|v\|^2 + 1}$.

Case 2: $c < 0$. Then, v is an eigenvector of S_1 for $-c$ and we must also have $Bv = cu$, which implies

$$Bv = QS_1 v = Q(-cv) = cQ(-v) = cu,$$

so

$$Q(-v) = u.$$

It follows that

$$A = \begin{pmatrix} B & u \\ v^\top & c \end{pmatrix} = \begin{pmatrix} QS_1 & Q(-v) \\ v^\top & c \end{pmatrix} = \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & -v \\ -v^\top & -c \end{pmatrix}.$$

In this case, the polar decomposition of $A \in \mathbf{O}(n, 1)$ is

$$A = \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & -v \\ -v^\top & -c \end{pmatrix},$$

where $Q \in \mathbf{O}(n)$ and $c = -\sqrt{\|v\|^2 + 1}$. Therefore, we conclude that any $A \in \mathbf{O}(n, 1)$ has a polar decomposition of the form

$$A = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

where $Q \in \mathbf{O}(n)$ and $c = \sqrt{\|v\|^2 + 1}$. □

Thus, we see that $\mathbf{O}(n, 1)$ has four components corresponding to the cases:

(1) $Q \in \mathbf{O}(n)$; $\det(Q) < 0$; $+1$ as the lower right entry of the orthogonal matrix;

(2) $Q \in \mathbf{SO}(n)$; -1 as the lower right entry of the orthogonal matrix;

(3) $Q \in \mathbf{O}(n)$; $\det(Q) < 0$; -1 as the lower right entry of the orthogonal matrix;

(4) $Q \in \mathbf{SO}(n)$; $+1$ as the lower right entry of the orthogonal matrix.

Observe that $\det(A) = -1$ in cases (1) and (2) and that $\det(A) = +1$ in cases (3) and (4). Thus, (3) and (4) correspond to the group $\mathbf{SO}(n, 1)$, in which case the polar decomposition is of the form

$$A = \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

where $Q \in \mathbf{O}(n)$, with $\det(Q) = -1$ and $c = \sqrt{\|v\|^2 + 1}$ or

$$A = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix}$$

where $Q \in \mathbf{SO}(n)$ and $c = \sqrt{\|v\|^2 + 1}$. The components in (1) and (2) are not groups. We will show later that all four components are connected and that case (4) corresponds to a group (Proposition 2.8). This group is the connected component of the identity and it is denoted $\mathbf{SO}_0(n,1)$ (see Corollary 2.27). For the time being, note that $A \in \mathbf{SO}_0(n,1)$ iff $A \in \mathbf{SO}(n,1)$ and $a_{n+1\,n+1}$ $(= c) > 0$ (here, $A = (a_{i\,j})$.) In fact, we proved above that if $a_{n+1\,n+1} > 0$, then $a_{n+1\,n+1} \geq 1$.

Remark: If we let

$$\Lambda_P = \begin{pmatrix} I_{n-1,1} & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \Lambda_T = I_{n,1}, \quad \text{where} \quad I_{n,1} = \begin{pmatrix} I_n & 0 \\ 0 & -1 \end{pmatrix},$$

then we have the disjoint union

$$\mathbf{O}(n,1) = \mathbf{SO}_0(n,1) \cup \Lambda_P \mathbf{SO}_0(n,1) \cup \Lambda_T \mathbf{SO}_0(n,1) \cup \Lambda_P \Lambda_T \mathbf{SO}_0(n,1).$$

In order to determine the SVD of matrices in $\mathbf{SO}_0(n,1)$, we analyze the eigenvectors and the eigenvalues of the positive definite symmetric matrix

$$S = \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix}$$

involved in Proposition 2.3. Such a matrix is called a *Lorentz boost*. Observe that if $v = 0$, then $c = 1$ and $S = I_{n+1}$.

Proposition 2.4 *Assume $v \neq 0$. The eigenvalues of the symmetric positive definite matrix*

$$S = \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

where $c = \sqrt{\|v\|^2 + 1}$, are 1 with multiplicity $n-1$, and e^α and $e^{-\alpha}$ each with multiplicity 1 (for some $\alpha \geq 0$). An orthonormal basis of eigenvectors of S consists of vectors of the form

$$\begin{pmatrix} u_1 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} u_{n-1} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{v}{\sqrt{2}\|v\|} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{v}{\sqrt{2}\|v\|} \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

where the $u_i \in \mathbb{R}^n$ are all orthogonal to v and pairwise orthogonal.

Proof. Let us solve the linear system

$$\begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix} \begin{pmatrix} v \\ d \end{pmatrix} = \lambda \begin{pmatrix} v \\ d \end{pmatrix}.$$

We get

$$\begin{aligned} \sqrt{I + vv^\top}(v) + dv &= \lambda v \\ v^\top v + cd &= \lambda d, \end{aligned}$$

that is (since $c = \sqrt{\|v\|^2 + 1}$ and $\sqrt{I + vv^\top}(v) = cv$),

$$\begin{aligned} (c + d)v &= \lambda v \\ c^2 - 1 + cd &= \lambda d. \end{aligned}$$

Since $v \neq 0$, we get $\lambda = c + d$. Substituting in the second equation, we get

$$c^2 - 1 + cd = (c + d)d,$$

that is,

$$d^2 = c^2 - 1.$$

Thus, either $\lambda_1 = c + \sqrt{c^2 - 1}$ and $d = \sqrt{c^2 - 1}$, or $\lambda_2 = c - \sqrt{c^2 - 1}$ and $d = -\sqrt{c^2 - 1}$. Since $c \geq 1$ and $\lambda_1\lambda_2 = 1$, set $\alpha = \log(c + \sqrt{c^2 - 1}) \geq 0$, so that $-\alpha = \log(c - \sqrt{c^2 - 1})$ and then, $\lambda_1 = e^\alpha$ and $\lambda_2 = e^{-\alpha}$. On the other hand, if u is orthogonal to v, observe that

$$\begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix} \begin{pmatrix} u \\ 0 \end{pmatrix} = \begin{pmatrix} u \\ 0 \end{pmatrix},$$

since the kernel of $vv^\top$ is the orthogonal complement of v. The rest is clear. $\square$

Corollary 2.5 *The singular values of any matrix $A \in \mathbf{O}(n, 1)$ are 1 with multiplicity $n - 1$, e^α, and $e^{-\alpha}$, for some $\alpha \geq 0$.*

Note that the case $\alpha = 0$ is possible, in which case, A is an orthogonal matrix of the form

$$\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} Q & 0 \\ 0 & -1 \end{pmatrix},$$

with $Q \in \mathbf{O}(n)$. The two singular values e^α and $e^{-\alpha}$ tell us how much A deviates from being orthogonal.

We can now determine a convenient form for the SVD of matrices in $\mathbf{O}(n, 1)$.

Theorem 2.6 *Every matrix $A \in \mathbf{O}(n, 1)$ can be written as*

$$A = \begin{pmatrix} P & 0 \\ 0 & \epsilon \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cosh\alpha & \sinh\alpha \\ 0 & \cdots & 0 & \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} Q^\top & 0 \\ 0 & 1 \end{pmatrix}$$

with $\epsilon = \pm 1$, $P \in \mathbf{O}(n)$ and $Q \in \mathbf{SO}(n)$. When $A \in \mathbf{SO}(n, 1)$, we have $\det(P)\epsilon = +1$, and when $A \in \mathbf{SO}_0(n, 1)$, we have $\epsilon = +1$ and $P \in \mathbf{SO}(n)$, that is,

$$A = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cosh\alpha & \sinh\alpha \\ 0 & \cdots & 0 & \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} Q^\top & 0 \\ 0 & 1 \end{pmatrix}$$

with $P \in \mathbf{SO}(n)$ and $Q \in \mathbf{SO}(n)$.

Proof. By Proposition 2.3, any matrix $A \in \mathbf{O}(n)$ can be written as

$$A = \begin{pmatrix} R & 0 \\ 0 & \epsilon \end{pmatrix} \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix}$$

where $\epsilon = \pm 1$, $R \in \mathbf{O}(n)$ and $c = \sqrt{\|v\|^2 + 1}$. The case where $c = 1$ is trivial, so assume $c > 1$, which means that α from Proposition 2.4 is such that $\alpha > 0$. The key fact is that the eigenvalues of the matrix

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}$$

are e^α and $e^{-\alpha}$ and that

$$\begin{pmatrix} e^\alpha & 0 \\ 0 & e^{-\alpha} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

From this fact, we see that the diagonal matrix

$$D = \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & e^\alpha & 0 \\ 0 & \cdots & 0 & 0 & e^{-\alpha} \end{pmatrix}$$

of eigenvalues of S is given by

$$D = \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cosh\alpha & \sinh\alpha \\ 0 & \cdots & 0 & \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

By Proposition 2.4, an orthonormal basis of eigenvectors of S consists of vectors of the form

$$\begin{pmatrix} u_1 \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} u_{n-1} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{v}{\sqrt{2}\|v\|} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \frac{v}{\sqrt{2}\|v\|} \\ -\frac{1}{\sqrt{2}} \end{pmatrix},$$

where the $u_i \in \mathbb{R}^n$ are all orthogonal to v and pairwise orthogonal. Now, if we multiply the matrices

$$\begin{pmatrix} u_1 & \cdots & u_{n-1} & \frac{v}{\sqrt{2}\|v\|} & \frac{v}{\sqrt{2}\|v\|} \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \cdots & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix},$$

we get an orthogonal matrix of the form

$$\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}$$

where the columns of Q are the vectors

$$u_1, \cdots, u_{n-1}, \frac{v}{\|v\|}.$$

By flipping u_1 to $-u_1$ if necessary, we can make sure that this matrix has determinant $+1$. Consequently,

$$S = \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cosh\alpha & \sinh\alpha \\ 0 & \cdots & 0 & \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} Q^\top & 0 \\ 0 & 1 \end{pmatrix},$$

so

$$A = \begin{pmatrix} R & 0 \\ 0 & \epsilon \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cosh\alpha & \sinh\alpha \\ 0 & \cdots & 0 & \sinh\alpha & \cosh\alpha \end{pmatrix} \begin{pmatrix} Q^\top & 0 \\ 0 & 1 \end{pmatrix},$$

and if we let $P = RQ$, we get the desired decomposition. $\square$

Remark: We warn our readers about Chapter 6 of Baker's book [3]. Indeed, this chapter is seriously flawed. The main two Theorems (Theorem 6.9 and Theorem 6.10) are false and as consequence, the proof of Theorem 6.11 is wrong too. Theorem 6.11 states that the exponential map $\exp\colon \mathfrak{so}(n,1) \to \mathbf{SO}_0(n,1)$ is surjective, which is correct, but known proofs are nontrivial and quite lengthy (see Section 4.5). The proof of Theorem 6.12 is also false, although the theorem itself is correct (this is our Theorem 4.21, see Section 4.5). The main problem with Theorem 6.9 (in Baker) is that the existence of the normal form for matrices in $\mathbf{SO}_0(n,1)$ claimed by this theorem is unfortunately false on several accounts. Firstly, it would imply that every matrix in $\mathbf{SO}_0(n,1)$ can be diagonalized, but this is false for $n \geq 2$. Secondly, even if a matrix $A \in \mathbf{SO}_0(n,1)$ is diagonalizable as $A = PDP^{-1}$, Theorem 6.9 (and Theorem 6.10) miss some possible eigenvalues and the matrix P is not necessarily in $\mathbf{SO}_0(n,1)$ (as the case $n = 1$ already shows). For a thorough analysis of the eigenvalues of Lorentz isometries (and much more), one should consult Riesz [52] (Chapter III).

Clearly, a result similar to Theorem 2.6 also holds for the matrices in the groups $\mathbf{O}(1,n)$, $\mathbf{SO}(1,n)$ and $\mathbf{SO}_0(1,n)$. For example, every matrix $A \in \mathbf{SO}_0(1,n)$ can be written as

$$A = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} \cosh\alpha & \sinh\alpha & 0 & \cdots & 0 \\ \sinh\alpha & \cosh\alpha & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & Q^\top \end{pmatrix},$$

where $P, Q \in \mathbf{SO}(n)$.

In the case $n = 3$, we obtain the *proper orthochronous Lorentz group*, $\mathbf{SO}_0(1,3)$, also denoted $\mathbf{Lor}(1,3)$. By the way, $\mathbf{O}(1,3)$ is called the *(full) Lorentz group* and $\mathbf{SO}(1,3)$ is the *special Lorentz group*.

Theorem 2.6 (really, the version for $\mathbf{SO}_0(1,n)$) shows that the Lorentz group $\mathbf{SO}_0(1,3)$ is generated by the matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \quad \text{with } P \in \mathbf{SO}(3)$$

and the matrices of the form

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha & 0 & 0 \\ \sinh\alpha & \cosh\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

This fact will be useful when we prove that the homomorphism $\varphi\colon \mathbf{SL}(2,\mathbb{C}) \to \mathbf{SO}_0(1,3)$ is surjective.

Remark: Unfortunately, unlike orthogonal matrices which can always be diagonalized over $\mathbb{C}$, **not** every matrix in $\mathbf{SO}(1,n)$ can be diagonalized for $n \geq 2$. This has to do with the fact that the Lie algebra $\mathfrak{so}(1,n)$ has non-zero idempotents (see Section 4.5).

It turns out that the group $\mathbf{SO}_0(1,3)$ admits another interesting characterization involving the hypersurface

$$\mathcal{H} = \{(t,x,y,z) \in \mathbb{R}^4 \mid t^2 - x^2 - y^2 - z^2 = 1\}.$$

This surface has two sheets and it is not hard to show that $\mathbf{SO}_0(1,3)$ is the subgroup of $\mathbf{SO}(1,3)$ that preserves these two sheets (does not swap them). Actually, we will prove this fact for any n. In preparation for this we need some definitions and a few propositions.

Let us switch back to $\mathbf{SO}(n,1)$. First, as a matter of notation, we write every $u \in \mathbb{R}^{n+1}$ as $u = (\mathbf{u}, t)$, where $\mathbf{u} \in \mathbb{R}^n$ and $t \in \mathbb{R}$, so that the Lorentz inner product can be expressed as

$$\langle u, v\rangle = \langle (\mathbf{u}, t), (\mathbf{v}, s)\rangle = \mathbf{u}\cdot\mathbf{v} - ts,$$

where $\mathbf{u}\cdot\mathbf{v}$ is the standard Euclidean inner product (the Euclidean norm of x is denoted $\|x\|$). Then, we can classify the vectors in $\mathbb{R}^{n+1}$ as follows:

Definition 2.10 A nonzero vector, $u = (\mathbf{u}, t) \in \mathbb{R}^{n+1}$ is called

(a) *spacelike* iff $\langle u, u\rangle > 0$, i.e., iff $\|\mathbf{u}\|^2 > t^2$;

(b) *timelike* iff $\langle u, u\rangle < 0$, i.e., iff $\|\mathbf{u}\|^2 < t^2$;

(c) *lightlike* or *isotropic* iff $\langle u, u\rangle = 0$, i.e., iff $\|\mathbf{u}\|^2 = t^2$.

A spacelike (resp. timelike, resp. lightlike) vector is said to be *positive* iff $t > 0$ and *negative* iff $t < 0$. The set of all isotropic vectors

$$\mathcal{H}_n(0) = \{u = (\mathbf{u}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{u}\|^2 = t^2\}$$

is called the *light cone*. For every $r > 0$, let

$$\mathcal{H}_n(r) = \{u = (\mathbf{u}, t) \in \mathbb{R}^{n+1} \mid \|\mathbf{u}\|^2 - t^2 = -r\},$$

a hyperboloid of two sheets.

It is easy to check that $\mathcal{H}_n(r)$ has two connected components as follows: First, since $r > 0$ and

$$\|\mathbf{u}\|^2 + r = t^2,$$

we have $|t| \geq \sqrt{r}$. Now, for any $x = (x_1, \ldots, x_n, t) \in \mathcal{H}_n(r)$ with $t \geq \sqrt{r}$, we have the continuous path from $(0, \ldots, 0, \sqrt{r})$ to x given by

$$\lambda \mapsto (\lambda x_1, \ldots, \lambda x_n, \sqrt{r + \lambda^2(t^2 - r)}),$$

where $\lambda \in [0, 1]$, proving that the component of $(0, \ldots, 0, \sqrt{r})$ is connected. Similarly, when $t \leq -\sqrt{r}$, we have the continuous path from $(0, \ldots, 0, -\sqrt{r})$ to x given by

$$\lambda \mapsto (\lambda x_1, \ldots, \lambda x_n, -\sqrt{r + \lambda^2(t^2 - r)}),$$

where $\lambda \in [0, 1]$, proving that the component of $(0, \ldots, 0, -\sqrt{r})$ is connected. We denote the sheet containing $(0, \ldots, 0, \sqrt{r})$ by $\mathcal{H}_n^+(r)$ and sheet containing $(0, \ldots, 0, -\sqrt{r})$ by $\mathcal{H}_n^-(r)$

Since every Lorentz isometry, $A \in \mathbf{SO}(n, 1)$, preserves the Lorentz inner product, we conclude that A globally preserves every hyperboloid, $\mathcal{H}_n(r)$, for $r > 0$. We claim that every $A \in \mathbf{SO}_0(n, 1)$ preserves both $\mathcal{H}_n^+(r)$ and $\mathcal{H}_n^-(r)$. This follows immediately from

Proposition 2.7 *If $a_{n+1\,n+1} > 0$, then every isometry, $A \in \mathbf{SO}(n, 1)$, preserves all positive (resp. negative) timelike vectors and all positive (resp. negative) lightlike vectors. Moreover, if $A \in \mathbf{SO}(n, 1)$ preserves all positive timelike vectors, then $a_{n+1\,n+1} > 0$.*

Proof. Let $u = (\mathbf{u}, t)$ be a nonzero timelike or lightlike vector. This means that

$$\|\mathbf{u}\|^2 \leq t^2 \quad \text{and} \quad t \neq 0.$$

Since $A \in \mathbf{SO}(n, 1)$, the matrix A preserves the inner product; if $\langle u, u\rangle = \|\mathbf{u}\|^2 - t^2 < 0$, we get $\langle Au, Au\rangle < 0$, which shows that Au is also timelike. Similarly, if $\langle u, u\rangle = 0$, then $\langle Au, Au\rangle = 0$. As $A \in \mathbf{SO}(n, 1)$, we know that

$$\langle A_{n+1}, A_{n+1}\rangle = -1,$$

that is,

$$\|\mathbf{A}_{n+1}\|^2 - a_{n+1,\,n+1}^2 = -1,$$

where $A_{n+1} = (\mathbf{A}_{n+1}, a_{n+1,\,n+1})$ is the $(n+1)$th row of the matrix A. The $(n+1)$th component of the vector Au is

$$\mathbf{u} \cdot \mathbf{A}_{n+1} + a_{n+1,\,n+1}t.$$

By Cauchy-Schwarz,

$$(\mathbf{u} \cdot \mathbf{A}_{n+1})^2 \leq \|\mathbf{u}\|^2 \|\mathbf{A}_{n+1}\|^2,$$

so we get,

$$\begin{aligned}
(\mathbf{u} \cdot \mathbf{A}_{n+1})^2 &\leq \|\mathbf{u}\|^2 \|\mathbf{A}_{n+1}\|^2 \\
&\leq t^2(a_{n+1,\,n+1}^2 - 1) = t^2 a_{n+1,\,n+1}^2 - t^2 \\
&< t^2 a_{n+1,\,n+1}^2,
\end{aligned}$$

since $t \neq 0$. It follows that $\mathbf{u} \cdot \mathbf{A}_{n+1} + a_{n+1,\,n+1}t$ has the same sign as t, since $a_{n+1,\,n+1} > 0$. Consequently, if $a_{n+1,\,n+1} > 0$, we see that A maps positive timelike (resp. lightlike) vectors to positive timelike (resp. lightlike) vectors and similarly with negative timelight (resp. lightlike) vectors.

Conversely, as $e_{n+1} = (0, \ldots, 0, 1)$ is timelike and positive, if A preserves all positive timelike vectors, then Ae_{n+1} is timelike positive, which implies $a_{n+1,\,n+1} > 0$. □

Using Proposition 2.7, we can now show that $\mathbf{SO}_0(n,1)$ is a subgroup of $\mathbf{SO}(n,1)$. Recall that

$$\mathbf{SO}_0(n,1) = \{A \in \mathbf{SO}(n,1) \mid a_{n+1\,n+1} > 0\}.$$

Proposition 2.8 *The set* $\mathbf{SO}_0(n,1)$ *is a subgroup of* $\mathbf{SO}(n,1)$.

Proof. Let $A \in \mathbf{SO}_0(n,1) \subseteq \mathbf{SO}(n,1)$, so that $a_{n+1\,n+1} > 0$. The inverse of A in $\mathbf{SO}(n,1)$ is $JA^\top J$, where

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

which implies that $a_{n+1\,n+1}^{-1} = a_{n+1\,n+1} > 0$ and so, $A^{-1} \in \mathbf{SO}_0(n,1)$. If $A, B \in \mathbf{SO}_0(n,1)$, then, by Proposition 2.7, both A and B preserve all positive timelike vectors, so AB preserve all positive timelike vectors. By Proposition 2.7, again, $AB \in \mathbf{SO}_0(n,1)$. Therefore, $\mathbf{SO}_0(n,1)$ is a group. □

Since any matrix, $A \in \mathbf{SO}_0(n,1)$, preserves the Lorentz inner product and all positive timelike vectors and since $\mathcal{H}_n^+(1)$ consists of timelike vectors, we see that every $A \in \mathbf{SO}_0(n,1)$ maps $\mathcal{H}_n^+(1)$ into itself. Similarly, every $A \in \mathbf{SO}_0(n,1)$ maps $\mathcal{H}_n^-(1)$ into itself. Thus, we can define an action $\cdot\colon \mathbf{SO}_0(n,1) \times \mathcal{H}_n^+(1) \longrightarrow \mathcal{H}_n^+(1)$ by

$$A \cdot u = Au$$

and similarly, we have an action $\cdot\colon \mathbf{SO}_0(n,1) \times \mathcal{H}_n^-(1) \longrightarrow \mathcal{H}_n^-(1)$.

Proposition 2.9 *The group* $\mathbf{SO}_0(n,1)$ *is the subgroup of* $\mathbf{SO}(n,1)$ *that preserves* $\mathcal{H}_n^+(1)$ *(and* $\mathcal{H}_n^-(1)$*) i.e.,*

$$\mathbf{SO}_0(n,1) = \{A \in \mathbf{SO}(n,1) \mid A(\mathcal{H}_n^+(1)) = \mathcal{H}_n^+(1) \quad \textit{and} \quad A(\mathcal{H}_n^-(1)) = \mathcal{H}_n^-(1)\}.$$

Proof. We already observed that $A(\mathcal{H}_n^+(1)) = \mathcal{H}_n^+(1)$ if $A \in \mathbf{SO}_0(n,1)$ (and similarly, $A(\mathcal{H}_n^-(1)) = \mathcal{H}_n^-(1)$). Conversely, for any $A \in \mathbf{SO}(n,1)$ such that $A(\mathcal{H}_n^+(1)) = \mathcal{H}_n^+(1)$, as $e_{n+1} = (0,\ldots,0,1) \in \mathcal{H}_n^+(1)$, the vector Ae_{n+1} must be positive timelike, but this says that $a_{n+1,\,n+1} > 0$, i.e., $A \in \mathbf{SO}_0(n,1)$. $\square$

Next, we wish to prove that the action $\mathbf{SO}_0(n,1) \times \mathcal{H}_n^+(1) \longrightarrow \mathcal{H}_n^+(1)$ is transitive. For this, we need the next two propositions.

Proposition 2.10 *Let* $u = (\mathbf{u}, t)$ *and* $v = (\mathbf{v}, s)$ *be nonzero vectors in* $\mathbb{R}^{n+1}$ *with* $\langle u, v\rangle = 0$. *If* u *is timelike, then* v *is spacelike (i.e.,* $\langle v, v\rangle > 0$*).*

Proof. We have $\|\mathbf{u}\|^2 < t^2$, so $t \neq 0$. Since $\mathbf{u}\cdot\mathbf{v} - ts = 0$, we get

$$\langle v, v\rangle = \|\mathbf{v}\|^2 - s^2 = \|\mathbf{v}\|^2 - \frac{(\mathbf{u}\cdot\mathbf{v})^2}{t^2}.$$

But, Cauchy-Schwarz implies that $(\mathbf{u}\cdot\mathbf{v})^2 \leq \|\mathbf{u}\|^2\,\|\mathbf{v}\|^2$, so we get

$$\langle v, v\rangle = \|\mathbf{v}\|^2 - \frac{(\mathbf{u}\cdot\mathbf{v})^2}{t^2} > \|\mathbf{v}\|^2 - \frac{(\mathbf{u}\cdot\mathbf{v})^2}{\|\mathbf{u}\|^2} \geq 0,$$

as $\|\mathbf{u}\|^2 < t^2$. $\square$

Lemma 2.10 also holds if $u = (\mathbf{u}, t)$ is a nonzero isotropic vector and $v = (\mathbf{v}, s)$ is a nonzero vector that is not collinear with u: If $\langle u, v\rangle = 0$, then v is spacelike (i.e., $\langle v, v\rangle > 0$). The proof is left as an exercise to the reader.

Proposition 2.11 *The action* $\mathbf{SO}_0(n,1) \times \mathcal{H}_n^+(1) \longrightarrow \mathcal{H}_n^+(1)$ *is transitive.*

Proof. Let $e_{n+1} = (0,\ldots,0,1) \in \mathcal{H}_n^+(1)$. It is enough to prove that for every $u = (\mathbf{u}, t) \in \mathcal{H}_n^+(1)$, there is some $A \in \mathbf{SO}_0(n,1)$ such that $Ae_{n+1} = u$. By hypothesis,

$$\langle u, u\rangle = \|\mathbf{u}\|^2 - t^2 = -1.$$

We show that we can construct an orthonormal basis, $e_1, \ldots, e_n, u$, with respect to the Lorentz inner product. Consider the hyperplane

$$H = \{v \in \mathbb{R}^{n+1} \mid \langle u, v\rangle = 0\}.$$

Since u is timelike, by Proposition 2.10, every nonzero vector $v \in H$ is spacelike, i.e., $\langle v, v\rangle > 0$. Let $v_1, \ldots, v_n$ be a basis of H. Since all (nonzero) vectors in H are spacelike, we can apply the Gramm-Schmidt orthonormalization procedure and we get a basis $e_1, \ldots, e_n$, of H, such that

$$\langle e_i, e_j \rangle = \delta_{i,j}, \quad 1 \leq i, j \leq n.$$

Now, by construction, we also have

$$\langle e_i, u \rangle = 0, \quad 1 \leq i \leq n, \quad \text{and} \quad \langle u, u \rangle = -1.$$

Therefore, $e_1, \ldots, e_n, u$ are the column vectors of a Lorentz matrix, A, such that $Ae_{n+1} = u$, proving our assertion. □

Let us find the stabilizer of $e_{n+1} = (0, \ldots, 0, 1)$. We must have $Ae_{n+1} = e_{n+1}$, and the polar form implies that

$$A = \begin{pmatrix} P & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{with} \quad P \in \mathbf{SO}(n).$$

Therefore, the stabilizer of e_{n+1} is isomorphic to $\mathbf{SO}(n)$ and we conclude that $\mathcal{H}_n^+(1)$, as a homogeneous space, is

$$\mathcal{H}_n^+(1) \cong \mathbf{SO}_0(n, 1)/\mathbf{SO}(n).$$

We will show in Section 2.5 that $\mathbf{SO}_0(n, 1)$ is connected.

2.4 More on $\mathbf{O}(p, q)$

Recall from Section 2.3 that the group $\mathbf{O}(p, q)$ is the set of all $n \times n$-matrices

$$\mathbf{O}(p, q) = \{A \in \mathbf{GL}(n, \mathbb{R}) \mid A^\top I_{p,q} A = I_{p,q}\}.$$

We deduce immediately that $|\det(A)| = 1$ and we also know that $AI_{p,q}A^\top = I_{p,q}$ holds. Unfortunately, when $p \neq 0, 1$ and $q \neq 0, 1$, it does not seem possible to obtain a formula as nice as that given in Proposition 2.3. Nevertheless, we can obtain a formula for the polar form of matrices in $\mathbf{O}(p, q)$. First, recall (for example, see Gallier [27], Chapter 12) that if S is a symmetric positive definite matrix, then there is a unique symmetric positive definite matrix, T, so that

$$S = T^2.$$

We denote T by $S^{\frac{1}{2}}$ or $\sqrt{S}$. By $S^{-\frac{1}{2}}$, we mean the inverse of $S^{\frac{1}{2}}$. In order to obtain the polar form of a matrix in $\mathbf{O}(p, q)$, we begin with the following proposition:

Proposition 2.12 *Every matrix $X \in \mathbf{O}(p, q)$ can be written as*

$$X = \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} \alpha^{\frac{1}{2}} & \alpha^{\frac{1}{2}} Z^\top \\ \delta^{\frac{1}{2}} Z & \delta^{\frac{1}{2}} \end{pmatrix},$$

where $\alpha = (I - Z^\top Z)^{-1}$ and $\delta = (I - ZZ^\top)^{-1}$, for some orthogonal matrices $U \in \mathbf{O}(p)$, $V \in \mathbf{O}(q)$ and for some $q \times p$ matrix, Z, such that $I - Z^\top Z$ and $I - ZZ^\top$ are symmetric positive definite matrices. Moreover, U, V, Z are uniquely determined by X.

Proof. If we write

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

with A a $p \times p$ matrix, D a $q \times q$ matrix, B a $p \times q$ matrix and C a $q \times p$ matrix, then the equations $A^\top I_{p,q} A = I_{p,q}$ and $A I_{p,q} A^\top = I_{p,q}$ yield the (not independent) conditions

$$\begin{aligned} A^\top A &= I + C^\top C \\ D^\top D &= I + B^\top B \\ A^\top B &= C^\top D \\ AA^\top &= I + BB^\top \\ DD^\top &= I + CC^\top \\ AC^\top &= BD^\top. \end{aligned}$$

Since $C^\top C$ is symmetric and since it is easy to show that $C^\top C$ has nonnegative eigenvalues, we deduce that $A^\top A$ is symmetric positive definite and similarly for $D^\top D$. If we assume that the above decomposition of X holds, we deduce that

$$\begin{aligned} A &= U(I - Z^\top Z)^{-\frac{1}{2}} \\ B &= U(I - Z^\top Z)^{-\frac{1}{2}} Z^\top \\ C &= V(I - ZZ^\top)^{-\frac{1}{2}} Z \\ D &= V(I - ZZ^\top)^{-\frac{1}{2}}, \end{aligned}$$

which implies

$$Z = D^{-1}C \quad \text{and} \quad Z^\top = A^{-1}B.$$

Thus, we must check that

$$(D^{-1}C)^\top = A^{-1}B$$

i.e.,

$$C^\top (D^\top)^{-1} = A^{-1}B,$$

namely,

$$AC^\top = BD^\top,$$

which is indeed the last of our identities. Thus, we must have $Z = D^{-1}C = (A^{-1}B)^\top$. The above expressions for A and D also imply that

$$A^\top A = (I - Z^\top Z)^{-1} \quad \text{and} \quad D^\top D = (I - ZZ^\top)^{-1},$$

so we must check that the choice $Z = D^{-1}C = (A^{-1}B)^\top$ yields the above equations.

Since $Z^\top = A^{-1}B$, we have

$$\begin{aligned} Z^\top Z &= A^{-1}BB^\top (A^\top)^{-1} \\ &= A^{-1}(AA^\top - I)(A^\top)^{-1} \\ &= I - A^{-1}(A^\top)^{-1} \\ &= I - (A^\top A)^{-1}. \end{aligned}$$

Therefore,

$$(A^\top A)^{-1} = I - Z^\top Z,$$

i.e.,

$$A^\top A = (I - Z^\top Z)^{-1},$$

as desired. We also have, this time, with $Z = D^{-1}C$,

$$\begin{aligned} ZZ^\top &= D^{-1}CC^\top(D^\top)^{-1} \\ &= D^{-1}(DD^\top - I)(D^\top)^{-1} \\ &= I - D^{-1}(D^\top)^{-1} \\ &= I - (D^\top D)^{-1}. \end{aligned}$$

Therefore,

$$(D^\top D)^{-1} = I - ZZ^\top,$$

i.e.,

$$D^\top D = (I - ZZ^\top)^{-1},$$

as desired. Now, since $A^\top A$ and $D^\top D$ are positive definite, the polar form implies that

$$A = U(A^\top A)^{\frac{1}{2}} = U(I - Z^\top Z)^{-\frac{1}{2}}$$

and

$$D = V(D^\top D)^{\frac{1}{2}} = V(I - ZZ^\top)^{-\frac{1}{2}},$$

for some unique matrices, $U \in \mathbf{O}(p)$ and $V \in \mathbf{O}(q)$. Since $Z = D^{-1}C$ and $Z^\top = A^{-1}B$, we get $C = DZ$ and $B = AZ^\top$, but this is

$$\begin{aligned} B &= U(I - Z^\top Z)^{-\frac{1}{2}}Z^\top \\ C &= V(I - ZZ^\top)^{-\frac{1}{2}}Z, \end{aligned}$$

as required. Therefore, the unique choice of $Z = D^{-1}C = (A^{-1}B)^\top$, U and V does yield the formula of the proposition. □

It remains to show that the matrix

$$\begin{pmatrix} \alpha^{\frac{1}{2}} & \alpha^{\frac{1}{2}}Z^\top \\ \delta^{\frac{1}{2}}Z & \delta^{\frac{1}{2}} \end{pmatrix} = \begin{pmatrix} (I - Z^\top Z)^{-\frac{1}{2}} & (I - Z^\top Z)^{-\frac{1}{2}}Z^\top \\ (I - ZZ^\top)^{-\frac{1}{2}}Z & (I - ZZ^\top)^{-\frac{1}{2}} \end{pmatrix}$$

is symmetric. To prove this, we will use power series and a continuity argument.

Proposition 2.13 *For any $q \times p$ matrix, Z, such that $I - Z^\top Z$ and $I - ZZ^\top$ are symmetric positive definite, the matrix*

$$S = \begin{pmatrix} \alpha^{\frac{1}{2}} & \alpha^{\frac{1}{2}}Z^\top \\ \delta^{\frac{1}{2}}Z & \delta^{\frac{1}{2}} \end{pmatrix}$$

is symmetric, where $\alpha = (I - Z^\top Z)^{-1}$ and $\delta = (I - ZZ^\top)^{-1}$.

Proof. The matrix S is symmetric iff

$$Z\alpha^{\frac{1}{2}} = \delta^{\frac{1}{2}} Z,$$

i.e., iff

$$Z(I - Z^\top Z)^{-\frac{1}{2}} = (I - ZZ^\top)^{-\frac{1}{2}} Z.$$

Consider the matrices

$$\beta(t) = (I - tZ^\top Z)^{-\frac{1}{2}} \quad \text{and} \quad \gamma(t) = (I - tZZ^\top)^{-\frac{1}{2}},$$

for any t with $0 \leq t \leq 1$. We claim that these matrices make sense. Indeed, since $Z^\top Z$ is symmetric, we can write

$$Z^\top Z = PDP^\top$$

where P is orthogonal and D is a diagonal matrix with nonnegative entries. Moreover, as

$$I - Z^\top Z = P(I - D)P^\top$$

and $I - Z^\top Z$ is positive definite, $0 \leq \lambda < 1$, for every eigenvalue in D. But then, as

$$I - tZ^\top Z = P(I - tD)P^\top,$$

we have $1 - t\lambda > 0$ for every λ in D and for all t with $0 \leq t \leq 1$, so that $I - tZ^\top Z$ is positive definite and thus, $(I - tZ^\top Z)^{-\frac{1}{2}}$ is also well defined. A similar argument applies to $(I - tZZ^\top)^{-\frac{1}{2}}$. Observe that

$$\lim_{t \to 1} \beta(t) = \alpha^{\frac{1}{2}}$$

since

$$\beta(t) = (I - tZ^\top Z)^{-\frac{1}{2}} = P(I - tD)^{-\frac{1}{2}} P^\top,$$

where $(I - tD)^{-\frac{1}{2}}$ is a diagonal matrix with entries of the form $(1 - t\lambda)^{-\frac{1}{2}}$ and these eigenvalues are continuous functions of t for $t \in [0, 1]$. A similar argument shows that

$$\lim_{t \to 1} \gamma(t) = \delta^{\frac{1}{2}}.$$

Therefore, it is enough to show that

$$Z\beta(t) = \gamma(t)Z,$$

with $0 \leq t < 1$ and our result will follow by continuity. However, when $0 \leq t < 1$, the power series for $\beta(t)$ and $\gamma(t)$ converge. Thus, we have

$$\beta(t) = 1 + \frac{1}{2} tZ^\top Z - \frac{1}{8} t^2 (Z^\top Z)^2 + \cdots + \frac{\frac{1}{2}\left(\frac{1}{2} - 1\right) \cdots \left(\frac{1}{2} - k + 1\right)}{k!} t^k (Z^\top Z)^k + \cdots$$

and

$$\gamma(t) = 1 + \frac{1}{2}tZZ^\top - \frac{1}{8}t^2(ZZ^\top)^2 + \cdots + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-k+1\right)}{k!}t^k(ZZ^\top)^k + \cdots$$

and we get

$$Z\beta(t) = Z + \frac{1}{2}tZZ^\top Z - \frac{1}{8}t^2 Z(Z^\top Z)^2 + \cdots + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-k+1\right)}{k!}t^k Z(Z^\top Z)^k + \cdots$$

and

$$\gamma(t)Z = Z + \frac{1}{2}tZZ^\top Z - \frac{1}{8}t^2(ZZ^\top)^2 Z + \cdots + \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-k+1\right)}{k!}t^k(ZZ^\top)^k Z + \cdots.$$

However

$$Z(Z^\top Z)^k = Z\underbrace{Z^\top Z\cdots Z^\top Z}_{k} = \underbrace{ZZ^\top\cdots ZZ^\top}_{k}Z = (ZZ^\top)^k Z,$$

which proves that $Z\beta(t) = \gamma(t)Z$, as required. $\square$

Another proof of Proposition 2.13 can be given using the SVD of Z. Indeed, we can write

$$Z = PDQ^\top$$

where P is a $q\times q$ orthogonal matrix, Q is a $p\times p$ orthogonal matrix and D is a $q\times p$ matrix whose diagonal entries are (strictly) positive and all other entries zero. Then,

$$I - Z^\top Z = I - QD^\top P^\top PDQ^\top = Q(I - D^\top D)Q^\top,$$

a symmetric positive definite matrix. We also have

$$I - ZZ^\top = I - PDQ^\top QD^\top P^\top = P(I - DD^\top)P^\top,$$

another symmetric positive definite matrix. Then,

$$Z(I - Z^\top Z)^{-\frac{1}{2}} = PDQ^\top Q(I - D^\top D)^{-\frac{1}{2}}Q^\top = PD(I - D^\top D)^{-\frac{1}{2}}Q^\top$$

and

$$(I - ZZ^\top)^{-\frac{1}{2}} = P(I - DD^\top)^{-\frac{1}{2}}P^\top PDQ^\top = P(I - DD^\top)^{-\frac{1}{2}}DQ^\top,$$

so it suffices to prove that

$$D(I - D^\top D)^{-\frac{1}{2}} = (I - DD^\top)^{-\frac{1}{2}}D.$$

However, D is essentially a diagonal matrix and the above is easily verified, as the reader should check.

Remark: The polar form can also be obtained *via* the exponential map and the Lie algebra, $\mathfrak{o}(p,q)$, of $\mathbf{O}(p,q)$, see Section 4.6.

We also have the following amusing property of the determinants of A and D:

Proposition 2.14 *For any matrix $X \in \mathbf{O}(p,q)$, if we write*

$$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

then

$$\det(X) = \det(A)\det(D)^{-1} \quad \textit{and} \quad |\det(A)| = |\det(D)| \geq 1.$$

Proof. Using the identities $A^\top B = C^\top D$ and $D^\top D = I + B^\top B$ proved earlier, observe that

$$\begin{pmatrix} A^\top & 0 \\ B^\top & -D^\top \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^\top A & A^\top B \\ B^\top A - D^\top C & B^\top B - D^\top D \end{pmatrix} = \begin{pmatrix} A^\top A & A^\top B \\ 0 & -I_q \end{pmatrix}.$$

If we compute determinants, we get

$$\det(A)(-1)^q \det(D)\det(X) = \det(A)^2(-1)^q.$$

It follows that

$$\det(X) = \det(A)\det(D)^{-1}.$$

From $A^\top A = I + C^\top C$ and $D^\top D = I + B^\top B$, we conclude that $\det(A) \geq 1$ and $\det(D) \geq 1$. Since $|\det(X)| = 1$, we have $|\det(A)| = |\det(D)| \geq 1$. □

Remark: It is easy to see that the equations relating A, B, C, D established in the proof of Proposition 2.12 imply that

$$\det(A) = \pm 1 \quad \text{iff} \quad C = 0 \quad \text{iff} \quad B = 0 \quad \text{iff} \quad \det(D) = \pm 1.$$

2.5 Topological Groups

Since Lie groups are topological groups (and manifolds), it is useful to gather a few basic facts about topological groups.

Definition 2.11 A set, G, is a *topological group* iff

(a) G is a Hausdorff topological space;

(b) G is a group (with identity 1);

(c) Multiplication, $\cdot\colon G \times G \to G$, and the inverse operation, $G \longrightarrow G\colon g \mapsto g^{-1}$, are continuous, where $G \times G$ has the product topology.

It is easy to see that the two requirements of condition (c) are equivalent to

(c′) The map $G \times G \longrightarrow G\colon (g,h) \mapsto gh^{-1}$ is continuous.

Given a topological group G, for every $a \in G$ we define *left translation* as the map, $L_a\colon G \to G$, such that $L_a(b) = ab$, for all $b \in G$, and *right translation* as the map, $R_a\colon G \to G$, such that $R_a(b) = ba$, for all $b \in G$. Observe that $L_{a^{-1}}$ is the inverse of L_a and similarly, $R_{a^{-1}}$ is the inverse of R_a. As multiplication is continuous, we see that L_a and R_a are continuous. Moreover, since they have a continuous inverse, they are homeomorphisms. As a consequence, if U is an open subset of G, then so is $gU = L_g(U)$ (resp. $Ug = R_gU$), for all $g \in G$. Therefore, the topology of a topological group (i.e., its family of open sets) is *determined* by the knowledge of the open subsets containing the identity, 1.

Given any subset, $S \subseteq G$, let $S^{-1} = \{s^{-1} \mid s \in S\}$; let $S^0 = \{1\}$ and $S^{n+1} = S^nS$, for all $n \geq 0$. Property (c) of Definition 2.11 has the following useful consequences:

Proposition 2.15 *If G is a topological group and U is any open subset containing* 1*, then there is some open subset, $V \subseteq U$, with $1 \in V$, so that $V = V^{-1}$ and $V^2 \subseteq U$. Furthermore, $\overline{V} \subseteq U$.*

Proof. Since multiplication $G \times G \longrightarrow G$ is continuous and $G \times G$ is given the product topology, there are open subsets, U_1 and U_2, with $1 \in U_1$ and $1 \in U_2$, so that $U_1U_2 \subseteq U$. Ley $W = U_1 \cap U_2$ and $V = W \cap W^{-1}$. Then, V is an open set containing 1 and, clearly, $V = V^{-1}$ and $V^2 \subseteq U_1U_2 \subseteq U$. If $g \in \overline{V}$, then gV is an open set containing g (since $1 \in V$) and thus, $gV \cap V \neq \emptyset$. This means that there are some $h_1, h_2 \in V$ so that $gh_1 = h_2$, but then, $g = h_2h_1^{-1} \in VV^{-1} = VV \subseteq U$. $\square$

A subset, U, containing 1 and such that $U = U^{-1}$, is called *symmetric*. Using Proposition 2.15, we can give a very convenient characterization of the Hausdorff separation property in a topological group.

Proposition 2.16 *If G is a topological group, then the following properties are equivalent:*

(1) G is Hausdorff;

(2) The set $\{1\}$ is closed;

(3) The set $\{g\}$ is closed, for every $g \in G$.

Proof. The implication (1) $\longrightarrow$ (2) is true in any Hausdorff topological space. We just have to prove that $G - \{1\}$ is open, which goes as follows: For any $g \neq 1$, since G is Hausdorff, there exists disjoint open subsets U_g and V_g, with $g \in U_g$ and $1 \in V_g$. Thus, $\bigcup U_g = G - \{1\}$, showing that $G - \{1\}$ is open. Since L_g is a homeomorphism, (2) and (3) are equivalent. Let us prove that (3) $\longrightarrow$ (1). Let $g_1, g_2 \in G$ with $g_1 \neq g_2$. Then, $g_1^{-1}g_2 \neq 1$ and if U and V are distinct open subsets such that $1 \in U$ and $g_1^{-1}g_2 \in V$, then $g_1 \in g_1U$ and $g_2 \in g_1V$, where g_1U and g_1V are still open and disjoint. Thus, it is enough to separate 1 and $g \neq 1$. Pick any $g \neq 1$. If every open subset containing 1 also contained g, then 1 would be in the closure of $\{g\}$, which is absurd, since $\{g\}$ is closed and $g \neq 1$. Therefore, there is some open subset, U, such that $1 \in U$ and $g \notin U$. By Proposition 2.15, we can find an open subset,

V, containing 1, so that $VV \subseteq U$ and $V = V^{-1}$. We claim that V and Vg are disjoint open sets with $1 \in V$ and $g \in gV$.

Since $1 \in V$, it is clear that $1 \in V$ and $g \in gV$. If we had $V \cap gV \neq \emptyset$, then we would have $g \in VV^{-1} = VV \subseteq U$, a contradiction. $\square$

If H is a subgroup of G (not necessarily normal), we can form the set of left cosets, G/H and we have the projection, $p\colon G \to G/H$, where $p(g) = gH = \overline{g}$. If G is a topological group, then G/H can be given the *quotient topology*, where a subset $U \subseteq G/H$ is open iff $p^{-1}(U)$ is open in G. With this topology, p is continuous. The trouble is that G/H is not necessarily Hausdorff. However, we can neatly characterize when this happens.

Proposition 2.17 *If G is a topological group and H is a subgroup of G then the following properties hold:*

(1) The map $p\colon G \to G/H$ is an open map, which means that $p(V)$ is open in G/H whenever V is open in G.

(2) The space G/H is Hausdorff iff H is closed in G.

(3) If H is open, then H is closed and G/H has the discrete topology (every subset is open).

(4) The subgroup H is open iff $1 \in \overset{\circ}{H}$ (i.e., there is some open subset, U, so that $1 \in U \subseteq H$).

Proof. (1) Observe that if V is open in G, then $VH = \bigcup_{h\in H} Vh$ is open, since each Vh is open (as right translation is a homeomorphism). However, it is clear that

$$p^{-1}(p(V)) = VH,$$

i.e., $p^{-1}(p(V))$ is open, which, by definition, means that $p(V)$ is open.

(2) If G/H is Hausdorff, then by Proposition 2.16, every point of G/H is closed, i.e., each coset gH is closed, so H is closed. Conversely, assume H is closed. Let $\overline{x}$ and $\overline{y}$ be two distinct point in G/H and let $x, y \in G$ be some elements with $p(x) = \overline{x}$ and $p(y) = \overline{y}$. As $\overline{x} \neq \overline{y}$, the elements x and y are not in the same coset, so $x \notin yH$. As H is closed, so is yH, and since $x \notin yH$, there is some open containing x which is disjoint from yH, and we may assume (by translation) that it is of the form Ux, where U is an open containing 1. By Proposition 2.15, there is some open V containing 1 so that $VV \subseteq U$ and $V = V^{-1}$. Thus, we have

$$V^2x \cap yH = \emptyset$$

and in fact,

$$V^2xH \cap yH = \emptyset,$$

since H is a group. Since $V = V^{-1}$, we get

$$VxH \cap VyH = \emptyset,$$

and then, since V is open, both VxH and VyH are disjoint, open, so $p(VxH)$ and $p(VyH)$ are open sets (by (1)) containing $\overline{x}$ and $\overline{y}$ respectively and $p(VxH)$ and $p(VyH)$ are disjoint (because $p^{-1}(p(VxH)) = VxHH = VxH$ and $p^{-1}(p(VyH)) = VyHH = VyH$ and $VxH \cap VyH = \emptyset$).

(3) If H is open, then every coset gH is open, so every point of G/H is open and G/H is discrete. Also, $\bigcup_{g \notin H} gH$ is open, i.e., H is closed.

(4) Say U is an open subset such that $1 \in U \subseteq H$. Then, for every $h \in H$, the set hU is an open subset of H with $h \in hU$, which shows that H is open. The converse is trivial. $\square$

Proposition 2.18 *If G is a connected topological group, then G is generated by any symmetric neighborhood, V, of 1. In fact,*

$$G = \bigcup_{n \geq 1} V^n.$$

Proof. Since $V = V^{-1}$, it is immediately checked that $H = \bigcup_{n \geq 1} V^n$ is the group generated by V. As V is a neighborhood of 1, there is some open subset, $U \subseteq V$, with $1 \in U$, and so $1 \in \overset{\circ}{H}$. From Proposition 2.17, the subgroup H is open and closed and since G is connected, $H = G$. $\square$

A subgroup, H, of a topological group G is *discrete* iff the induced topology on H is discrete, i.e., for every $h \in H$, there is some open subset, U, of G so that $U \cap H = \{h\}$.

Proposition 2.19 *If G is a topological group and H is discrete subgroup of G, then H is closed.*

Proof. As H is discrete, there is an open subset, U, of G so that $U \cap H = \{1\}$, and by Proposition 2.15, we may assume that $U = U^{-1}$. If $g \in \overline{H}$, as gU is an open set containing g, we have $gU \cap H \neq \emptyset$. Consequently, there is some $y \in gU \cap H = gU^{-1} \cap H$, so $g \in yU$ with $y \in H$. Thus, we have

$$g \in yU \cap \overline{H} \subseteq \overline{yU \cap H} = \overline{\{y\}} = \{y\},$$

since $U \cap H = \{1\}$, $y \in H$ and G is Hausdorff. Therefore, $g = y \in H$. $\square$

Proposition 2.20 *If G is a topological group and H is any subgroup of G, then the closure, $\overline{H}$, of H is a subgroup of G.*

Proof. This follows easily from the continuity of multiplication and of the inverse operation, the details are left as an exercise to the reader. $\square$

Proposition 2.21 *Let G be a topological group and H be any subgroup of G. If H and G/H are connected, then G is connected.*

Proof. It is a standard fact of topology that a space G is connected iff every continuous function, f, from G to the discrete space $\{0,1\}$ is constant. Pick any continuous function, f, from G to $\{0,1\}$. As H is connected and left translations are homeomorphisms, all cosets, gH, are connected. Thus, f is constant on every coset, gH. Thus, the function $f\colon G \to \{0,1\}$ induces a continuous function, $\overline{f}\colon G/H \to \{0,1\}$, such that $f = \overline{f} \circ p$ (where $p\colon G \to G/H$; the continuity of $\overline{f}$ follows immediately from the definition of the quotient topology on G/H). As G/H is connected, $\overline{f}$ is constant and so, $f = \overline{f} \circ p$ is constant. $\square$

Proposition 2.22 *Let G be a topological group and let V be any connected symmetric open subset containing* 1. *Then, if G_0 is the connected component of the identity, we have*

$$G_0 = \bigcup_{n \geq 1} V^n$$

and G_0 is a normal subgroup of G. Moreover, the group G/G_0 is discrete.

Proof. First, as V is open, every V^n is open, so the group $\bigcup_{n\geq 1} V^n$ is open, and thus closed, by Proposition 2.17 (3). For every $n \geq 1$, we have the continuous map

$$\underbrace{V \times \cdots \times V}_{n} \longrightarrow V^n : (g_1, \ldots, g_n) \mapsto g_1 \cdots g_n.$$

As V is connected, $V \times \cdots \times V$ is connected and so, V^n is connected. Since $1 \in V^n$ for all $n \geq 1$, and every V^n is connected, we conclude that $\bigcup_{n\geq 1} V^n$ is connected. Now, $\bigcup_{n\geq 1} V^n$ is connected, open and closed, so it is the connected component of 1. Finally, for every $g \in G$, the group gG_0g^{-1} is connected and contains 1, so it is contained in G_0, which proves that G_0 is normal. Since G_0 is open, the group G/G_0 is discrete. $\square$

A topological space, X, is *locally compact* iff for every point $p \in X$, there is a compact neighborhood, C of p, i.e., there is a compact, C, and an open, U, with $p \in U \subseteq C$. For example, manifolds are locally compact.

Proposition 2.23 *Let G be a topological group and assume that G is connected and locally compact. Then, G is countable at infinity, which means that G is the union of a countable family of compact subsets. In fact, if V is any symmetric compact neighborhood of* 1, *then*

$$G = \bigcup_{n \geq 1} V^n.$$

Proof. Since G is locally compact, there is some compact neighborhood, K, of 1. Then, $V = K \cap K^{-1}$ is also compact and a symmetric neigborhood of 1. By Proposition 2.18, we have

$$G = \bigcup_{n \geq 1} V^n.$$

An argument similar to the one used in the proof of Proposition 2.22 to show that V^n is connected if V is connected proves that each V^n compact if V is compact. $\square$

If a topological group, G acts on a topological space, X, and the action $\cdot\colon G \times X \to X$ is continuous, we say that G acts *continuously on* X. Under some mild assumptions on G and X, the quotient space, G/G_x, is homeomorphic to X. For example, this happens if X is a Baire space.

Recall that a *Baire space*, X, is a topological space with the property that if $\{F\}_{i\geq 1}$ is any countable family of closed sets, F_i, such that each F_i has empty interior, then $\bigcup_{i\geq 1} F_i$ also has empty interior. By complementation, this is equivalent to the fact that for every countable family of open sets, U_i, such that each U_i is dense in X (i.e., $\overline{U}_i = X$), then $\bigcap_{i\geq 1} U_i$ is also dense in X.

Remark: A subset, $A \subseteq X$, is *rare* if its closure, $\overline{A}$, has empty interior. A subset, $Y \subseteq X$, is *meager* if it is a countable union of rare sets. Then, it is immediately verified that a space, X, is a Baire space iff every nonempty open subset of X is not meager.

The following theorem shows that there are plenty of Baire spaces:

Theorem 2.24 *(Baire) (1) Every locally compact topological space is a Baire space.*

(2) Every complete metric space is a Baire space.

A proof of Theorem 2.24 can be found in Bourbaki [10], Chapter IX, Section 5, Theorem 1.

We can now greatly improve Proposition 2.2 when G and X are topological spaces having some "nice" properties.

Theorem 2.25 *Let G be a topological group which is locally compact and countable at infinity, X a Hausdorff topological space which is a Baire space and assume that G acts transitively and continuously on X. Then, for any $x \in X$, the map $\varphi\colon G/G_x \to X$ is a homeomorphism.*

By Theorem 2.24, we get the following important corollary:

Theorem 2.26 *Let G be a topological group which is locally compact and countable at infinity, X a Hausdorff locally compact topological space and assume that G acts transitively and continuously on X. Then, for any $x \in X$, the map $\varphi\colon G/G_x \to X$ is a homeomorphism.*

Proof of Theorem 2.25. We follow the proof given in Bourbaki [10], Chapter IX, Section 5, Proposition 6 (Essentially the same proof can be found in Mneimné and Testard [44], Chapter 2). First, observe that if a topological group acts continuously and transitively on a Hausdorff topological space, then for every $x \in X$, the stabilizer, G_x, is a closed subgroup of G. This is because, as the action is continuous, the projection $\pi\colon G \longrightarrow X\colon g \mapsto g \cdot x$ is continuous, and $G_x = \pi^{-1}(\{x\})$, with $\{x\}$ closed. Therefore, by Proposition 2.17, the quotient space, G/G_x, is Hausdorff. As the map $\pi\colon G \longrightarrow X$ is continuous, the induced map $\varphi\colon G/G_x \to X$ is continuous and by Proposition 2.2, it is a bijection. Therefore, to prove that φ is a homeomorphism, it is enough to prove that φ is an open map. For this, it suffices to show that π is an open map. Given any open, U, in G, we will prove that for any $g \in U$, the element $\pi(g) = g \cdot x$ is contained in the interior of $U \cdot x$. However, observe that this is equivalent to proving that x belongs to the interior of $(g^{-1} \cdot U) \cdot x$. Therefore, we are reduced to the case: If U is any open subset of G containing 1, then x belongs to the interior of $U \cdot x$.

Since G is locally compact, using Proposition 2.15, we can find a compact neighborhood of the form $W = \overline{V}$, such that $1 \in W$, $W = W^{-1}$ and $W^2 \subseteq U$, where V is open with $1 \in V \subseteq U$. As G is countable at infinity, $G = \bigcup_{i\geq 1} K_i$, where each K_i is compact. Since V is open, all the cosets gV are open, and as each K_i is covered by the gV's, by compactness of K_i, finitely many cosets gV cover each K_i and so,

$$G = \bigcup_{i\geq 1} g_i V = \bigcup_{i\geq 1} g_i W,$$

for countably many $g_i \in G$, where each $g_i W$ is compact. As our action is transitive, we deduce that

$$X = \bigcup_{i\geq 1} g_i W \cdot x,$$

where each $g_i W \cdot x$ is compact, since our action is continuous and the $g_i W$ are compact. As X is Hausdorff, each $g_i W \cdot x$ is closed and as X is a Baire space expressed as a union of closed sets, one of the $g_i W \cdot x$ must have nonempty interior, i.e., there is some $w \in W$, with $g_i w \cdot x$ in the interior of $g_i W \cdot x$, for some i. But then, as the map $y \mapsto g \cdot y$ is a homeomorphism for any given $g \in G$ (where $y \in X$), we see that x is in the interior of

$$w^{-1} g_i^{-1} \cdot (g_i W \cdot x) = w^{-1} W \cdot x \subseteq W^{-1} W \cdot x = W^2 \cdot x \subseteq U \cdot x,$$

as desired. □

As an application of Theorem 2.26 and Proposition 2.21, we show that the Lorentz group $\mathbf{SO}_0(n, 1)$ is connected. Firstly, it is easy to check that $\mathbf{SO}_0(n, 1)$ and $\mathcal{H}_n^+(1)$ satisfy the assumptions of Theorem 2.26 because they are both manifolds, although this notion has not been discussed yet (but will be in Chapter 3). Also, we saw at the end of Section 2.3 that the action $\cdot\colon \mathbf{SO}_0(n, 1) \times \mathcal{H}_n^+(1) \longrightarrow \mathcal{H}_n^+(1)$ of $\mathbf{SO}_0(n, 1)$ on $\mathcal{H}_n^+(1)$ is transitive, so that, as topological spaces

$$\mathbf{SO}_0(n, 1)/\mathbf{SO}(n) \cong \mathcal{H}_n^+(1).$$

Now, we already showed that $\mathcal{H}_n^+(1)$ is connected so, by Proposition 2.21, the connectivity of $\mathbf{SO}_0(n, 1)$ follows from the connectivity of $\mathbf{SO}(n)$ for $n \geq 1$. The connectivity of $\mathbf{SO}(n)$ is a consequence of the surjectivity of the exponential map (for instance, see Gallier [27], Chapter 14) but we can also give a quick proof using Proposition 2.21. Indeed, $\mathbf{SO}(n + 1)$ and S^n are both manifolds and we saw in Section 2.2 that

$$\mathbf{SO}(n + 1)/\mathbf{SO}(n) \cong S^n.$$

Now, S^n is connected for $n \geq 1$ and $\mathbf{SO}(1) \cong S^1$ is connected. We finish the proof by induction on n.

Corollary 2.27 *The Lorentz group* $\mathbf{SO}_0(n,1)$ *is connected; it is the component of the identity in* $\mathbf{O}(n,1)$.

Readers who wish to learn more about topological groups may consult Sagle and Walde [53] and Chevalley [16] for an introductory account, and Bourbaki [9], Weil [60] and Pontryagin [50, 51], for a more comprehensive account (especially the last two references).

Chapter 3

Manifolds, Tangent Spaces, Cotangent Spaces, Vector Fields, Flow, Integral Curves, Partitions of Unity, Manifolds with Boundary, Orientation of Manifolds

3.1 Manifolds

Elsewhere (in another set of notes and in Gallier [27], Chapter 14) we defined the notion of a manifold embedded in some ambient space, $\mathbb{R}^N$. In order to maximize the range of applications of the theory of manifolds it is necessary to generalize the concept of a manifold to spaces that are not a priori embedded in some $\mathbb{R}^N$. The basic idea is still that, whatever a manifold is, it is a topological space that can be covered by a collection of open subsets, U_α, where each U_α is isomorphic to some "standard model", *e.g.*, some open subset of Euclidean space, $\mathbb{R}^n$. Of course, manifolds would be very dull without functions defined on them and between them. This is a general fact learned from experience: Geometry arises not just from spaces but from spaces and interesting classes of functions between them. In particular, we still would like to "do calculus" on our manifold and have good notions of curves, tangent vectors, differential forms, etc. The small drawback with the more general approach is that the definition of a tangent vector is more abstract. We can still define the notion of a curve on a manifold, but such a curve does not live in any given $\mathbb{R}^n$, so it it not possible to define tangent vectors in a simple-minded way using derivatives. Instead, we have to resort to the notion of chart. This is not such a strange idea. For example, a geography atlas gives a set of maps of various portions of the earh and this provides a very good description of what the earth is, without actually imagining the earth embedded in 3-space.

The material of this chapter borrows from many sources, including Warner [59], Berger and Gostiaux [5], O'Neill [49], Do Carmo [22, 21], Gallot, Hulin and Lafontaine [28], Lang [38], Schwartz [56], Hirsch [33], Sharpe [57], Guillemin and Pollack [31], Lafontaine [37], Dubrovin, Fomenko and Novikov [24] and Boothby [6]. A nice (not very technical) exposition is given in Morita [45] (Chapter 1) and it should be said that among the many texts on manifolds and differential geometry, the book by Choquet-Bruhat, DeWitt-Morette and Dillard-Bleick [17] stands apart because it is one of the clearest and most comprehensive (many proofs are omitted, but this can be an advantage!) Being written for (theoretical) physicists, it contains more examples and applications than most other sources.

Given $\mathbb{R}^n$, recall that the projection functions, $pr_i\colon \mathbb{R}^n \to \mathbb{R}$, are defined by

$$pr_i(x_1, \ldots, x_n) = x_i, \quad 1 \leq i \leq n.$$

For technical reasons (in particular, to insure the existence of partitions of unity, see Section 3.6) and to avoid "esoteric" manifolds that do not arise in practice, from now on, all topological spaces under consideration will be assumed to be Hausdorff and second-countable (which means that the topology has a countable basis).

Definition 3.1 Given a topological space, M, a *chart* (or *local coordinate map*) is a pair, (U, φ), where U is an open subset of M and $\varphi\colon U \to \Omega$ is a homeomorphism onto an open subset, $\Omega = \varphi(U)$, of $\mathbb{R}^{n_\varphi}$ (for some $n_\varphi \geq 1$). For any $p \in M$, a chart, (U, φ), is a *chart at p* iff $p \in U$. If (U, φ) is a chart, then the functions $x_i = pr_i \circ \varphi$ are called *local coordinates* and for every $p \in U$, the tuple $(x_1(p), \ldots, x_n(p))$ is the set of *coordinates of p* w.r.t. the chart. The inverse, (Ω, φ^{-1}), of a chart is called a *local parametrization*. Given any two charts, (U_1, φ_1) and (U_2, φ_2), if $U_1 \cap U_2 \neq \emptyset$, we have the *transition maps*, $\varphi_i^j\colon \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ and $\varphi_j^i\colon \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$, defined by

$$\varphi_i^j = \varphi_j \circ \varphi_i^{-1} \quad \text{and} \quad \varphi_j^i = \varphi_i \circ \varphi_j^{-1}.$$

Clearly, $\varphi_j^i = (\varphi_i^j)^{-1}$. Observe that the transition maps φ_i^j (resp. φ_j^i) are maps between *open subsets of* $\mathbb{R}^n$. This is good news! Indeed, the whole arsenal of calculus is available for functions on $\mathbb{R}^n$, and we will be able to promote many of these results to manifolds by imposing suitable conditions on transition functions.

Definition 3.2 Given a topological space, M, and any two integers, $n \geq 1$ and $k \geq 1$, a *C^k n-atlas* (or *n-atlas of class C^k*), $\mathcal{A}$, is a family of charts, $\{(U_i, \varphi_i)\}$, such that

(1) $\varphi_i(U_i) \subseteq \mathbb{R}^n$ for all i;

(2) The U_i cover M, i.e.,

$$M = \bigcup_i U_i;$$

(3) Whenever $U_i \cap U_j \neq \emptyset$, the transition map φ_i^j (and φ_j^i) is a C^k-diffeomorphism.

We must insure that we have enough charts in order to carry out our program of generalizing calculus on $\mathbb{R}^n$ to manifolds. For this, we must be able to add new charts whenever necessary, provided that they are consistent with the previous charts in an existing atlas. Technically, given a C^k n-atlas, $\mathcal{A}$, on M, for any other chart, (U, φ), we say that (U, φ) is *compatible* with the altas $\mathcal{A}$ iff every map $\varphi_i \circ \varphi^{-1}$ and $\varphi \circ \varphi_i^{-1}$ is C^k (whenever $U \cap U_i \neq \emptyset$). Two atlases $\mathcal{A}$ and $\mathcal{A}'$ on M are *compatible* iff every chart of one is compatible with the other atlas. This is equivalent to saying that the union of the two atlases is still an atlas. It is immediately verified that compatibility induces an equivalence relation on C^k n-atlases on M. In fact, given an atlas, $\mathcal{A}$, for M, the collection, $\widetilde{\mathcal{A}}$, of all charts compatible with $\mathcal{A}$ is a maximal atlas in the equivalence class of charts compatible with $\mathcal{A}$. Finally, we have our generalized notion of a manifold.

Definition 3.3 Given any two integers, $n \geq 1$ and $k \geq 1$, a *C^k-manifold of dimension n* consists of a topological space, M, together with an equivalence class, $\overline{\mathcal{A}}$, of C^k n-atlases, on M. Any atlas, $\mathcal{A}$, in the equivalence class $\overline{\mathcal{A}}$ is called a *differentiable structure of class C^k (and dimension n) on M*. We say that M is *modeled on* $\mathbb{R}^n$. When $k = \infty$, we say that M is a *smooth manifold*.

Remark: It might have been better to use the terminology *abstract manifold* rather than manifold, to emphasize the fact that the space M is not a priori a subspace of $\mathbb{R}^N$, for some suitable N.

We can allow $k = 0$ in the above definitions. In this case, condition (3) in Definition 3.2 is void, since a C^0-diffeomorphism is just a homeomorphism, but φ_i^j is always a homeomorphism. In this case, M is called a *topological manifold of dimension n*. We do not require a manifold to be connected but we require all the components to have the same dimension, n. Actually, on every connected component of M, it can be shown that the dimension, n_φ, of the range of every chart is the same. This is quite easy to show if $k \geq 1$ but for $k = 0$, this requires a deep theorem of Brouwer. What happens if $n = 0$? In this case, every one-point subset of M is open, so every subset of M is open, i.e., M is any (countable if we assume M to be second-countable) set with the discrete topology!

Observe that since $\mathbb{R}^n$ is locally compact and locally connected, so is every manifold (check this!)

Remark: In some cases, M does not come with a topology in an obvious (or natural) way and a slight variation of Definition 3.2 is more convenient in such a situation:

Definition 3.4 Given a set, M, and any two integers, $n \geq 1$ and $k \geq 1$, a *C^k n-atlas* (or *n-atlas of class C^k*), $\mathcal{A}$, is a family of charts, $\{(U_i, \varphi_i)\}$, such that

(1) Each U_i is a subset of M and $\varphi_i \colon U_i \to \varphi_i(U_i)$ is a bijection onto an open subset, $\varphi_i(U_i) \subseteq \mathbb{R}^n$, for all i;

(2) The U_i cover M, i.e.,

$$M = \bigcup_i U_i;$$

(3) Whenever $U_i \cap U_j \neq \emptyset$, the set $\varphi_i(U_i \cap U_j)$ is open in $\mathbb{R}^n$ and the transition map φ_i^j (and φ_j^i) is a C^k-diffeomorphism.

Then, the notion of a chart being compatible with an atlas and of two atlases being compatible is just as before and we get a new definition of a manifold, analogous to Definition 3.3. But, this time, we give M the topology in which the open sets are arbitrary unions of domains of charts, U_i, more precisely, the U_i's of the maximal atlas defining the differentiable structure on M. It is not difficult to verify that the axioms of a topology are verified and M is indeed a topological space with this topology. It can also be shown that when M is equipped with the above topology, then the maps $\varphi_i \colon U_i \to \varphi_i(U_i)$ are homeomorphisms, so M is a manifold according to Definition 3.3. Thus, we are back to the original notion of a manifold where it is assumed that M is already a topological space.

One can also define the topology on M in terms of any the atlases, $\mathcal{A}$, defining M (not only the maximal one) by requiring $U \subseteq M$ to be open iff $\varphi_i(U \cap U_i)$ is open in $\mathbb{R}^n$, for every chart, (U_i, φ_i), in the altas $\mathcal{A}$. Then, one can prove that we obtain the same topology as the topology induced by the maximal atlas. For details, see Berger and Gostiaux [5], Chapter 2. We also require that under this topology, M is Hausdorff and second-countable. A sufficient condition (in fact, also necessary!) for being second-countable is that some atlas be countable.

If the underlying topological space of a manifold is compact, then M has some finite atlas. Also, if $\mathcal{A}$ is some atlas for M and (U, φ) is a chart in $\mathcal{A}$, for any (nonempty) open subset, $V \subseteq U$, we get a chart, $(V, \varphi \restriction V)$, and it is obvious that this chart is compatible with $\mathcal{A}$. Thus, $(V, \varphi \restriction V)$ is also a chart for M. This observation shows that if U is any open subset of a C^k-manifold, M, then U is also a C^k-manifold whose charts are the restrictions of charts on M to U.

Example 1. The sphere S^n.

Using the stereographic projections (from the north pole and the south pole), we can define two charts on S^n and show that S^n is a smooth manifold. Let $\sigma_N \colon S^n - \{N\} \to \mathbb{R}^n$ and $\sigma_S \colon S^n - \{S\} \to \mathbb{R}^n$, where $N = (0, \cdots, 0, 1) \in \mathbb{R}^{n+1}$ (the north pole) and $S = (0, \cdots, 0, -1) \in \mathbb{R}^{n+1}$ (the south pole) be the maps called respectively *stereographic projection from the north pole* and *stereographic projection from the south pole* given by

$$\sigma_N(x_1, \ldots, x_{n+1}) = \frac{1}{1 - x_{n+1}} (x_1, \ldots, x_n) \quad \text{and} \quad \sigma_S(x_1, \ldots, x_{n+1}) = \frac{1}{1 + x_{n+1}} (x_1, \ldots, x_n).$$

The inverse stereographic projections are given by

$$\sigma_N^{-1}(x_1, \ldots, x_n) = \frac{1}{(\sum_{i=1}^n x_i^2) + 1} \left(2x_1, \ldots, 2x_n, \left(\sum_{i=1}^n x_i^2\right) - 1\right)$$

and

$$\sigma_S^{-1}(x_1, \ldots, x_n) = \frac{1}{(\sum_{i=1}^n x_i^2) + 1} \left(2x_1, \ldots, 2x_n, -\left(\sum_{i=1}^n x_i^2\right) + 1\right).$$

Thus, if we let $U_N = S^n - \{N\}$ and $U_S = S^n - \{S\}$, we see that U_N and U_S are two open subsets covering S^n, both homeomorphic to $\mathbb{R}^n$. Furthermore, it is easily checked that on the overlap, $U_N \cap U_S = S^n - \{N, S\}$, the transition maps

$$\sigma_S \circ \sigma_N^{-1} = \sigma_N \circ \sigma_S^{-1}$$

are given by

$$(x_1, \ldots, x_n) \mapsto \frac{1}{\sum_{i=1}^n x_i^2} (x_1, \ldots, x_n),$$

that is, the inversion of center $O = (0, \ldots, 0)$ and power 1. Clearly, this map is smooth on $\mathbb{R}^n - \{O\}$, so we conclude that (U_N, σ_N) and (U_S, σ_S) form a smooth atlas for S^n.

Example 2. The projective space $\mathbb{RP}^n$.

To define an atlas on $\mathbb{RP}^n$ it is convenient to view $\mathbb{RP}^n$ as the set of equivalence classes of vectors in $\mathbb{R}^{n+1} - \{0\}$ modulo the equivalence relation,

$$u \sim v \quad \text{iff} \quad v = \lambda u, \quad \text{for some} \quad \lambda \neq 0 \in \mathbb{R}.$$

Given any $p = [x_1, \ldots, x_{n+1}] \in \mathbb{RP}^n$, we call $(x_1, \ldots, x_{n+1})$ the *homogeneous coordinates* of p. It is customary to write $(x_1 \colon \cdots \colon x_{n+1})$ instead of $[x_1, \ldots, x_{n+1}]$. (Actually, in most books, the indexing starts with 0, i.e., homogeneous coordinates for $\mathbb{RP}^n$ are written as $(x_0 \colon \cdots \colon x_n)$.) For any i, with $1 \leq i \leq n+1$, let

$$U_i = \{(x_1 \colon \cdots \colon x_{n+1}) \in \mathbb{RP}^n \mid x_i \neq 0\}.$$

Observe that U_i is well defined, because if $(y_1 \colon \cdots \colon y_{n+1}) = (x_1 \colon \cdots \colon x_{n+1})$, then there is some $\lambda \neq 0$ so that $y_i = \lambda x_i$, for $i = 1, \ldots, n+1$. We can define a homeomorphism, φ_i, of U_i onto $\mathbb{R}^n$, as follows:

$$\varphi_i(x_1 \colon \cdots \colon x_{n+1}) = \left(\frac{x_1}{x_i}, \ldots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \ldots, \frac{x_{n+1}}{x_i}\right),$$

where the ith component is omitted. Again, it is clear that this map is well defined since it only involves ratios. We can also define the maps, ψ_i, from $\mathbb{R}^n$ to $U_i \subseteq \mathbb{RP}^n$, given by

$$\psi_i(x_1, \ldots, x_n) = (x_1 \colon \cdots \colon x_{i-1} \colon 1 \colon x_i \colon \cdots \colon x_n),$$

where the 1 goes in the ith slot, for $i = 1, \ldots, n+1$. One easily checks that φ_i and ψ_i are mutual inverses, so the φ_i are homeomorphisms. On the overlap, $U_i \cap U_j$, (where $i \neq j$), as $x_j \neq 0$, we have

$$(\varphi_j \circ \varphi_i^{-1})(x_1, \ldots, x_n) = \left(\frac{x_1}{x_j}, \ldots, \frac{x_{i-1}}{x_j}, \frac{1}{x_j}, \frac{x_i}{x_j}, \ldots, \frac{x_{j-1}}{x_j}, \frac{x_{j+1}}{x_j}, \ldots, \frac{x_n}{x_j}\right).$$

(We assumed that $i < j$; the case $j < i$ is similar.) This is clearly a smooth function from $\varphi_i(U_i \cap U_j)$ to $\varphi_j(U_i \cap U_j)$. As the U_i cover $\mathbb{RP}^n$, see conclude that the (U_i, φ_i) are $n+1$ charts making a smooth atlas for $\mathbb{RP}^n$. Intuitively, the space $\mathbb{RP}^n$ is obtained by glueing the open subsets U_i on their overlaps. Even for $n = 3$, this is not easy to visualize!

Example 3. The Grassmannian $G(k, n)$.

Recall that $G(k, n)$ is the set of all k-dimensional linear subspaces of $\mathbb{R}^n$, also called k-planes. Every k-plane, W, is the linear span of k linearly independent vectors, $u_1, \ldots, u_k$, in $\mathbb{R}^n$; furthermore, $u_1, \ldots, u_k$ and $v_1, \ldots, v_k$ both span W iff there is an invertible $k \times k$-matrix, $\Lambda = (\lambda_{ij})$, such that

$$v_i = \sum_{j=1}^{k} \lambda_{ij} u_j, \quad 1 \leq i \leq k.$$

Obviously, there is a bijection between the collection of k linearly independent vectors, $u_1, \ldots, u_k$, in $\mathbb{R}^n$ and the collection of $n \times k$ matrices of rank k. Furthermore, two $n \times k$ matrices A and B of rank k represent the same k-plane iff

$$B = A\Lambda, \quad \text{for some invertible } k \times k \text{ matrix, } \Lambda.$$

(Note the analogy with projective spaces where two vectors u, v represent the same point iff $v = \lambda u$ for some invertible $\lambda \in \mathbb{R}$.) We can define the domain of charts (according to Definition 3.4) on $G(k, n)$ as follows: For every subset, $S = \{i_1, \ldots, i_k\}$ of $\{1, \ldots, n\}$, let U_S be the subset of $n \times k$ matrices, A, of rank k whose rows of index in $S = \{i_1, \ldots, i_k\}$ forms an invertible $k \times k$ matrix denoted A_S. Observe that the $k \times k$ matrix consisting of the rows of the matrix AA_S^{-1} whose index belong to S is the identity matrix, I_k. Therefore, we can define a map, $\varphi_S \colon U_S \to \mathbb{R}^{(n-k)\times k}$, where $\varphi_S(A) =$ the $(n-k) \times k$ matrix obtained by deleting the rows of index in S from AA_S^{-1}.

We need to check that this map is well defined, i.e., that it does not depend on the matrix, A, representing W. Let us do this in the case where $S = \{1, \ldots, k\}$, which is notationally simpler. The general case can be reduced to this one using a suitable permutation.

If $B = A\Lambda$, with Λ invertible, if we write

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

as $B = A\Lambda$, we get $B_1 = A_1\Lambda$ and $B_2 = A_2\Lambda$, from which we deduce that

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} B_1^{-1} = \begin{pmatrix} I_k \\ B_2 B_1^{-1} \end{pmatrix} = \begin{pmatrix} I_k \\ A_2 \Lambda \Lambda^{-1} A_1^{-1} \end{pmatrix} = \begin{pmatrix} I_k \\ A_2 A_1^{-1} \end{pmatrix} = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} A_1^{-1}.$$

Therefore, our map is indeed well-defined. It is clearly injective and we can define its inverse, ψ_S, as follows: Let π_S be the permutation of $\{1, \ldots, n\}$ swaping $\{1, \ldots, k\}$ and S and leaving every other element fixed (i.e., if $S = \{i_1, \ldots, i_k\}$, then $\pi_S(j) = i_j$ and $\pi_S(i_j) = j$, for $j = 1, \ldots, k$). If P_S is the permutation matrix associated with π_S, for any $(n-k) \times k$ matrix, M, let

$$\psi_S(M) = P_S \begin{pmatrix} I_k \\ M \end{pmatrix}.$$

The effect of ψ_S is to "insert into M" the rows of the identity matrix I_k as the rows of index from S. At this stage, we have charts that are bijections from subsets, U_S, of $G(k, n)$ to open subsets, namely, $\mathbb{R}^{(n-k)\times k}$. Then, the reader can check that the transition map $\varphi_T \circ \varphi_S^{-1}$ from $\varphi_S(U_S \cap U_U)$ to $\varphi_T(U_S \cap U_U)$ is given by

$$M \mapsto (C + DM)(A + BM)^{-1},$$

where

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = P_T P_S,$$

is the matrix of the permutation $\pi_T \circ \pi_S$ (this permutation "shuffles" S and T). This map is smooth, as it is given by determinants, and so, the charts (U_S, φ_S) form a smooth atlas for $G(k, n)$. Finally, one can check that the conditions of Definition 3.4 are satisfied, so the atlas just defined makes $G(k, n)$ into a topological space and a smooth manifold.

Remark: The reader should have no difficulty proving that the collection of k-planes represented by matrices in U_S is precisely set of k-planes, W, supplementary to the $(n-k)$-plane spanned by the $n-k$ canonical basis vectors $e_{j_{k+1}}, \ldots, e_{j_n}$ (i.e., $\mathrm{span}(W \cup \{e_{j_{k+1}}, \ldots, e_{j_n}\}) = \mathbb{R}^n$, where $S = \{i_1, \ldots, i_k\}$ and $\{j_{k+1}, \ldots, j_n\} = \{1, \ldots, n\} - S$).

Example 4. Product Manifolds.

Let M_1 and M_2 be two C^k-manifolds of dimension n_1 and n_2, respectively. The topological space, $M_1 \times M_2$, with the product topology (the opens of $M_1 \times M_2$ are arbitrary unions of sets of the form $U \times V$, where U is open in M_1 and V is open in M_2) can be given the structure of a C^k-manifold of dimension $n_1 + n_2$ by defining charts as follows: For any two charts, (U_i, φ_i) on M_1 and (V_j, ψ_j) on M_2, we declare that $(U_i \times V_j, \varphi_i \times \psi_j)$ is a chart on $M_1 \times M_2$, where $\varphi_i \times \psi_j \colon U_i \times V_j \to \mathbb{R}^{n_1+n_2}$ is defined so that

$$\varphi_i \times \psi_j(p, q) = (\varphi_i(p), \psi_j(q)), \quad \text{for all } (p, q) \in U_i \times V_j.$$

We define C^k-maps between manifolds as follows:

Definition 3.5 Given any two C^k-manifolds, M and N, of dimension m and n respectively, a *C^k-map* if a continuous functions, $h\colon M \to N$, so that for every $p \in M$, there is some chart, (U, φ), at p and some chart, (V, ψ), at $q = h(p)$, with $f(U) \subseteq V$ and

$$\psi \circ h \circ \varphi^{-1} \colon \varphi(U) \longrightarrow \psi(V)$$

a C^k-function.

It is easily shown that Definition 3.5 does not depend on the choice of charts. In particular, if $N = \mathbb{R}$, we obtain a *C^k-function on M*. One checks immediately that a function, $f\colon M \to \mathbb{R}$, is a C^k-map iff for every $p \in M$, there is some chart, (U, φ), at p so that

$$f \circ \varphi^{-1}\colon \varphi(U) \longrightarrow \mathbb{R}$$

is a C^k-function. If U is an open subset of M, set of C^k-functions on U is denoted by $\mathcal{C}^k(U)$. In particular, $\mathcal{C}^k(M)$ denotes the set of C^k-functions on the manifold, M. Observe that $\mathcal{C}^k(U)$ is a ring.

On the other hand, if M is an open interval of $\mathbb{R}$, say $M =]a, b[$, then $\gamma\colon]a, b[\to N$ is called a *C^k-curve* in N. One checks immediately that a function, $\gamma\colon]a, b[\to N$, is a C^k-map iff for every $q \in N$, there is some chart, (V, ψ), at q so that

$$\psi \circ \gamma\colon]a, b[\longrightarrow \psi(V)$$

is a C^k-function.

It is clear that the composition of C^k-maps is a C^k-map. A C^k-map, $h\colon M \to N$, between two manifolds is a *C^k-diffeomorphism* iff h has an inverse, $h^{-1}\colon N \to M$ (i.e., $h^{-1} \circ h = \mathrm{id}_M$ and $h \circ h^{-1} = \mathrm{id}_N$), and both h and h^{-1} are C^k-maps (in particular, h and h^{-1} are homeomorphisms). Next, we define tangent vectors.

3.2 Tangent Vectors, Tangent Spaces, Cotangent Spaces

Let M be a C^k manifold of dimension n, with $k \geq 1$. The most intuitive method to define tangent vectors is to use curves. Let $p \in M$ be any point on M and let $\gamma\colon]-\epsilon, \epsilon[\to M$ be a C^1-curve passing through p, that is, with $\gamma(0) = p$. Unfortunately, if M is not embedded in any $\mathbb{R}^N$, the derivative $\gamma'(0)$ does not make sense. However, for any chart, (U, φ), at p, the map $\varphi \circ \gamma$ is a C^1-curve in $\mathbb{R}^n$ and the tangent vector $v = (\varphi \circ \gamma)'(0)$ is well defined. The trouble is that different curves may yield the same v!

To remedy this problem, we define an equivalence relation on curves through p as follows:

Definition 3.6 Given a C^k manifold, M, of dimension n, for any $p \in M$, two C^1-curves, $\gamma_1\colon]-\epsilon_1, \epsilon_1[\to M$ and $\gamma_2\colon]-\epsilon_2, \epsilon_2[\to M$, through p (i.e., $\gamma_1(0) = \gamma_2(0) = p$) are *equivalent* iff there is some chart, (U, φ), at p so that

$$(\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0).$$

Now, the problem is that this definition seems to depend on the choice of the chart. Fortunately, this is not the case. For, if (V, ψ) is another chart at p, as p belongs both to U

and V, we have $U \cap V \neq 0$, so the transition function $\eta = \psi \circ \varphi^{-1}$ is C^k and, by the chain rule, we have

$$\begin{aligned}
(\psi \circ \gamma_1)'(0) &= (\eta \circ \varphi \circ \gamma_1)'(0) \\
&= \eta'(\varphi(p))((\varphi \circ \gamma_1)'(0)) \\
&= \eta'(\varphi(p))((\varphi \circ \gamma_2)'(0)) \\
&= (\eta \circ \varphi \circ \gamma_2)'(0) \\
&= (\psi \circ \gamma_2)'(0).
\end{aligned}$$

This leads us to the first definition of a tangent vector.

Definition 3.7 *(Tangent Vectors, Version 1)* Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, a *tangent vector to M at p* is any equivalence class of C^1-curves through p on M, modulo the equivalence relation defined in Definition 3.6. The set of all tangent vectors at p is denoted by $T_p(M)$ (or T_pM).

It is obvious that $T_p(M)$ is a vector space. If $u, v \in T_p(M)$ are defined by the curves γ_1 and γ_2, then $u+v$ is defined by the curve $\gamma_1 + \gamma_2$ (we may assume by reparametrization that γ_1 and γ_2 have the same domain.) Similarly, if $u \in T_p(M)$ is defined by a curve γ and $\lambda \in \mathbb{R}$, then λu is defined by the curve $\lambda\gamma$. The reader should check that these definitions do not depend on the choice of the curve in its equivalence class. We will show that $T_p(M)$ is a vector space of dimension $n =$ dimension of M. One should observe that unless $M = \mathbb{R}^n$, in which case, for any $p, q \in \mathbb{R}^n$, the tangent space $T_q(M)$ is naturally isomorphic to the tangent space $T_p(M)$ by the translation $q - p$, for an arbitrary manifold, there is no relationship between $T_p(M)$ and $T_q(M)$ when $p \neq q$.

One of the defects of the above definition of a tangent vector is that it has no clear relation to the C^k-differential structure of M. In particular, the definition does not seem to have anything to do with the functions defined locally at p. There is another way to define tangent vectors that reveals this connection more clearly. Moreover, such a definition is more intrinsic, i.e., does not refer explicitly to charts. Our presentation of this second approach is heavily inspired by Schwartz [56] (Chapter 3, Section 9) but also by Warner [59].

As a first step, consider the following: Let (U, φ) be a chart at $p \in M$ (where M is a C^k-manifold of dimension n, with $k \geq 1$) and let $x_i = pr_i \circ \varphi$, the ith local coordinate $(1 \leq i \leq n)$. For any function, f, defined on $U \ni p$, set

$$\left(\frac{\partial}{\partial x_i}\right)_p f = \left.\frac{\partial(f \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)}, \qquad 1 \leq i \leq n.$$

(Here, $(\partial g/\partial X_i)|_y$ denotes the partial derivative of a function $g\colon \mathbb{R}^n \to \mathbb{R}$ with respect to the ith coordinate, evaluated at y.)

We would expect that the function that maps f to the above value is a linear map on the set of functions defined locally at p, but there is technical difficulty: The set of functions defined locally at p is **not** a vector space! To see this, observe that if f is defined on an open $U \ni p$ and g is defined on a different open $V \ni p$, then we do know how to define $f+g$. The problem is that we need to identify functions that agree on a smaller open. This leads to the notion of *germs*.

Definition 3.8 Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, a *locally defined function at* p is a pair, (U, f), where U is an open subset of M containing p and f is a function defined on U. Two locally defined functions, (U, f) and (V, g), at p are *equivalent* iff there is some open subset, $W \subseteq U \cap V$, containing p so that

$$f \upharpoonright W = g \upharpoonright W.$$

The equivalence class of a locally defined function at p, denoted $[f]$ or $\mathbf{f}$, is called a *germ at* p.

One should check that the relation of Definition 3.8 is indeed an equivalence relation. Of course, the value at p of all the functions, f, in any germ, $\mathbf{f}$, is $f(p)$. Thus, we set $\mathbf{f}(p) = f(p)$. One should also check that we can define addition of germs, multiplication of a germ by a scalar and multiplication of germs, in the obvious way: If $\mathbf{f}$ and $\mathbf{g}$ are two germs at p, and $\lambda \in \mathbb{R}$, then

$$\begin{aligned} [f]+[g] &= [f+g] \\ \lambda[f] &= [\lambda f] \\ [f][g] &= [fg]. \end{aligned}$$

(Of course, $f+g$ is the function locally defined so that $(f+g)(x) = f(x)+g(x)$ and similarly, $(\lambda f)(x) = \lambda f(x)$ and $(fg)(x) = f(x)g(x)$.) Therefore, the germs at p form a ring. The ring of germs of C^k-functions at p is denoted $\mathcal{O}^{(k)}_{M,p}$. When $k = \infty$, we usually drop the superscript ∞.

Remark: Most readers will most likely be puzzled by the notation $\mathcal{O}^{(k)}_{M,p}$. In fact, it is standard in algebraic geometry, but it is not as commonly used in differential geometry. For any open subset, U, of a manifold, M, the ring, $\mathcal{C}^k(U)$, of C^k-functions on U is also denoted $\mathcal{O}^{(k)}_M(U)$ (certainly by people with an algebraic geometry bent!). Then, it turns out that the map $U \mapsto \mathcal{O}^{(k)}_M(U)$ is a *sheaf*, denoted $\mathcal{O}^{(k)}_M$, and the ring $\mathcal{O}^{(k)}_{M,p}$ is the *stalk* of the sheaf $\mathcal{O}^{(k)}_M$ at p. Such rings are called *local rings*. Roughly speaking, all the "local" information about M at p is contained in the local ring $\mathcal{O}^{(k)}_{M,p}$. (This is to be taken with a grain of salt. In the C^k-case where $k < \infty$, we also need the "stationary germs", as we will see shortly.)

Now that we have a rigorous way of dealing with functions locally defined at p, observe that the map

$$v_i \colon f \mapsto \left(\frac{\partial}{\partial x_i}\right)_p f$$

yields the same value for all functions f in a germ $\mathbf{f}$ at p. Furthermore, the above map is linear on $\mathcal{O}^{(k)}_{M,p}$. More is true. Firstly for any two functions f, g locally defined at p, we have

$$\left(\frac{\partial}{\partial x_i}\right)_p (fg) = f(p)\left(\frac{\partial}{\partial x_i}\right)_p g + g(p)\left(\frac{\partial}{\partial x_i}\right)_p f.$$

Secondly, if $(f \circ \varphi^{-1})'(\varphi(p)) = 0$, then

$$\left(\frac{\partial}{\partial x_i}\right)_p f = 0.$$

The first property says that v_i is a *derivation*. As to the second property, when $(f \circ \varphi^{-1})'(\varphi(p)) = 0$, we say that f *is stationary at* p. It is easy to check (using the chain rule) that being stationary at p does not depend on the chart, (U, φ), at p or on the function chosen in a germ, $\mathbf{f}$. Therefore, the notion of a stationary germ makes sense: We say that $\mathbf{f}$ *is a stationary germ* iff $(f \circ \varphi^{-1})'(\varphi(p)) = 0$ for some chart, (U, φ), at p and some function, f, in the germ, $\mathbf{f}$. The C^k-stationary germs form a subring of $\mathcal{O}^{(k)}_{M,p}$ (but not an ideal!) denoted $\mathcal{S}^{(k)}_{M,p}$.

Remarkably, it turns out that the dual of the vector space, $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$, is isomorphic to the tangent space, $T_p(M)$. First, we prove that the subspace of linear forms on $\mathcal{O}^{(k)}_{M,p}$ that vanish on $\mathcal{S}^{(k)}_{M,p}$ has $\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_n}\right)_p$ as a basis.

Proposition 3.1 *Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$ and any chart (U, φ) at p, the n functions, $\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_n}\right)_p$, defined on $\mathcal{O}^{(k)}_{M,p}$ by*

$$\left(\frac{\partial}{\partial x_i}\right)_p f = \left.\frac{\partial(f \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)}, \qquad 1 \leq i \leq n$$

are linear forms that vanish on $\mathcal{S}^{(k)}_{M,p}$. Every linear form, L, on $\mathcal{O}^{(k)}_{M,p}$ that vanishes on $\mathcal{S}^{(k)}_{M,p}$ can be expressed in a unique way as

$$L = \sum_{i=1}^{n} \lambda_i \left(\frac{\partial}{\partial x_i}\right)_p,$$

where $\lambda_i \in \mathbb{R}$. Therefore, the

$$\left(\frac{\partial}{\partial x_i}\right)_p, \qquad i = 1, \ldots, n$$

form a basis of the vector space of linear forms on $\mathcal{O}^{(k)}_{M,p}$ that vanish on $\mathcal{S}^{(k)}_{M,p}$.

Proof. The first part of the proposition is trivial, by definition of $(f \circ \varphi^{-1})'(\varphi(p))$ and of $\left(\frac{\partial}{\partial x_i}\right)_p f$.

Next, assume that L is a linear form on $\mathcal{O}^{(k)}_{M,p}$ that vanishes on $\mathcal{S}^{(k)}_{M,p}$. Consider the locally defined function at p given by

$$g(x) = f(x) - f(p) - \sum_{i=1}^{n} (pr_i \circ \varphi)(x) \left(\frac{\partial}{\partial x_i}\right)_p f.$$

Observe that the germ of g is stationary at p, since

$$(g \circ \varphi^{-1})(\varphi(x)) = (f \circ \varphi^{-1})(\varphi(x)) - f(p) - \sum_{i=1}^{n} X_i \left(\frac{\partial}{\partial x_i}\right)_p f,$$

with $X_i = (pr_i \circ \varphi)(x)$. It follows that

$$\left.\frac{\partial(g \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)} = \left.\frac{\partial(f \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)} - \left(\frac{\partial}{\partial x_i}\right)_p f = 0.$$

But then, as constant functions have stationary germs and as L vanishes on stationary germs, we get

$$L(f) = \sum_{i=1}^{n} L(pr_i \circ \varphi) \left(\frac{\partial}{\partial x_i}\right)_p f,$$

as desired. We still have to prove linear independence. If

$$\sum_{i=1}^{n} \lambda_i \left(\frac{\partial}{\partial x_i}\right)_p = 0,$$

then, if we apply this relation to the functions $x_i = pr_i \circ \varphi$, as

$$\left(\frac{\partial}{\partial x_i}\right)_p x_j = \delta_{ij},$$

we get $\lambda_i = 0$, for $i = 1, \ldots, n$. $\square$

As the subspace of linear forms on $\mathcal{O}^{(k)}_{M,p}$ that vanish on $\mathcal{S}^{(k)}_{M,p}$ is isomorphic to the dual, $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^*$, of the space $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$, we see that the

$$\left(\frac{\partial}{\partial x_i}\right)_p, \qquad i = 1, \ldots, n$$

also form a basis of $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^*$.

To define our second version of tangent vectors, we need to define linear derivations.

Definition 3.9 Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, a *linear derivation at* p is a linear form, v, on $\mathcal{O}_{M,p}^{(k)}$, such that

$$v(\mathbf{fg}) = f(p)v(\mathbf{g}) + g(p)v(\mathbf{f}),$$

for all germs $\mathbf{f}, \mathbf{g} \in \mathcal{O}_{M,p}^{(k)}$. The above is called the *Leibnitz property*.

Recall that we observed earlier that the $\left(\frac{\partial}{\partial x_i}\right)_p$ are linear derivations at p. Therefore, we have

Proposition 3.2 *Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, the linear forms on $\mathcal{O}_{M,p}^{(k)}$ that vanish on $\mathcal{S}_{M,p}^{(k)}$ are exactly the linear derivations on $\mathcal{O}_{M,p}^{(k)}$ that vanish on $\mathcal{S}_{M,p}^{(k)}$.*

Proof. By Proposition 3.1, the

$$\left(\frac{\partial}{\partial x_i}\right)_p, \qquad i = 1, \ldots, n$$

form a basis of the linear forms on $\mathcal{O}_{M,p}^{(k)}$ that vanish on $\mathcal{S}_{M,p}^{(k)}$. Since each $\left(\frac{\partial}{\partial x_i}\right)_p$ is a also a linear derivation at p, the result follows. $\square$

Here is now our second definition of a tangent vector.

Definition 3.10 *(Tangent Vectors, Version 2)* Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, a *tangent vector to M at p* is any linear derivation on $\mathcal{O}_{M,p}^{(k)}$ that vanishes on $\mathcal{S}_{M,p}^{(k)}$, the subspace of stationary germs.

Let us consider the simple case where $M = \mathbb{R}$. In this case, for every $x \in \mathbb{R}$, the tangent space, $T_x(\mathbb{R})$, is a one-dimensional vector space isomorphic to $\mathbb{R}$ and $\left(\frac{\partial}{\partial t}\right)_x = \frac{d}{dt}\big|_x$ is a basis vector of $T_x(\mathbb{R})$. For every C^k-function, f, locally defined at x, we have

$$\left(\frac{\partial}{\partial t}\right)_x f = \left.\frac{df}{dt}\right|_x = f'(x).$$

Thus, $\left(\frac{\partial}{\partial t}\right)_x$ is: compute the derivative of a function at x.

We now prove the equivalence of the two definitions of a tangent vector.

Proposition 3.3 *Let M be any C^k-manifold of dimension n, with $k \geq 1$. For any $p \in M$, let u be any tangent vector (version 1) given by some equivalence class of C^1-curves, $\gamma\colon]-\epsilon, +\epsilon[\to M$, through p (i.e., $p = \gamma(0)$). Then, the map L_u defined on $\mathcal{O}_{M,p}^{(k)}$ by*

$$L_u(\mathbf{f}) = (f \circ \gamma)'(0)$$

is a linear derivation that vanishes on $\mathcal{S}_{M,p}^{(k)}$. Furthermore, the map $u \mapsto L_u$ defined above is an isomorphism between $T_p(M)$ and $(\mathcal{O}_{M,p}^{(k)}/\mathcal{S}_{M,p}^{(k)})^*$, the space of linear forms on $\mathcal{O}_{M,p}^{(k)}$ that vanish on $\mathcal{S}_{M,p}^{(k)}$.

Proof. Clearly, $L_u(\mathbf{f})$ does not depend on the representative, f, chosen in the germ, $\mathbf{f}$. If γ and σ are equivalent curves defining u, then $(\varphi \circ \sigma)'(0) = (\varphi \circ \gamma)'(0)$, so we get

$$(f \circ \sigma)'(0) = (f \circ \varphi^{-1})'(\varphi(p))((\varphi \circ \sigma)'(0)) = (f \circ \varphi^{-1})'(\varphi(p))((\varphi \circ \gamma)'(0)) = (f \circ \gamma)'(0),$$

which shows that $L_u(\mathbf{f})$ does not depend on the curve, γ, defining u. If $\mathbf{f}$ is a stationary germ, then pick any chart, (U, φ), at p and let $\psi = \varphi \circ \gamma$. We have

$$L_u(\mathbf{f}) = (f \circ \gamma)'(0) = ((f \circ \varphi^{-1}) \circ (\varphi \circ \gamma))'(0) = (f \circ \varphi^{-1})'(\varphi(p))(\psi'(0)) = 0,$$

since $(f \circ \varphi^{-1})'(\varphi(p)) = 0$, as $\mathbf{f}$ is a stationary germ. The definition of L_u makes it clear that L_u is a linear derivation at p. If $u \neq v$ are two distinct tangent vectors, then there exist some curves γ and σ through p so that

$$(\varphi \circ \gamma)'(0) \neq (\varphi \circ \sigma)'(0).$$

Thus, there is some i, with $1 \leq i \leq n$, so that if we let $f = pr_i \circ \varphi$, then

$$(f \circ \gamma)'(0) \neq (f \circ \sigma)'(0),$$

and so, $L_u \neq L_v$. This proves that the map $u \mapsto L_u$ is injective.

For surjectivity, recall that every linear map, L, on $\mathcal{O}^{(k)}_{M,p}$ that vanishes on $\mathcal{S}^{(k)}_{M,p}$ can be uniquely expressed as

$$L = \sum_{i=1}^{n} \lambda_i \left(\frac{\partial}{\partial x_i}\right)_p.$$

Define the curve, γ, on M through p by

$$\gamma(t) = \varphi^{-1}(\varphi(p) + t(\lambda_1, \ldots, \lambda_n)),$$

for t in a small open interval containing 0. Then, we have

$$f(\gamma(t)) = (f \circ \varphi^{-1})(\varphi(p) + t(\lambda_1, \ldots, \lambda_n)),$$

and we get

$$(f \circ \gamma)'(0) = (f \circ \varphi^{-1})'(\varphi(p))(\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} \lambda_i \left.\frac{\partial (f \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)} = L(\mathbf{f}).$$

This proves that $T_p(M)$ and $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^*$ are isomorphic. $\square$

In view of Proposition 3.3, we can identify $T_p(M)$ with $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^*$. As the space $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$ is finite dimensional, $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^{**}$ is canonically isomorphic to $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$, so we can identify $T^*_p(M)$ with $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$. (Recall that if E is a finite dimensional space, the map $i_E \colon E \to E^{**}$ defined so that, for any $v \in E$,

$$v \mapsto \widetilde{v}, \quad \text{where} \quad \widetilde{v}(f) = f(v), \quad \text{for all } f \in E^*$$

is a linear isomorphism.) This also suggests the following definition:

Definition 3.11 Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, the *tangent space at* p, denoted $T_p(M)$ (or T_pM) is the space of linear derivations on $\mathcal{O}^{(k)}_{M,p}$ that vanish on $\mathcal{S}^{(k)}_{M,p}$. Thus, $T_p(M)$ can be identified with $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^*$. The space $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$ is called the *cotangent space at* p; it is isomorphic to the dual, $T^*_p(M)$, of $T_p(M)$. (We also denote $T^*_p(M)$ by T^*_pM.)

Observe that if $x_i = pr_i \circ \varphi$, as

$$\left(\frac{\partial}{\partial x_i}\right)_p x_j = \delta_{i,j},$$

the images of $x_1, \ldots, x_n$ in $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$ are the dual of the basis $\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_n}\right)_p$ of $T_p(M)$. Given any C^k-function, f, on M, we denote the image of f in $T^*_p(M) = \mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$ by df_p. This is the *differential of* f *at* p. Using the isomorphism between $\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p}$ and $(\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p})^{**}$ described above, df_p corresponds to the linear map in $T^*_p(M)$ defined by $df_p(v) = v(\mathbf{f})$, for all $v \in T_p(M)$. With this notation, we see that $(dx_1)_p, \ldots, (dx_n)_p$ is a basis of $T^*_p(M)$, and this basis is dual to the basis $\left(\frac{\partial}{\partial x_1}\right)_p, \ldots, \left(\frac{\partial}{\partial x_n}\right)_p$ of $T_p(M)$. For simplicity of notation, we often omit the subscript p unless confusion arises.

Remark: Strictly speaking, a tangent vector, $v \in T_p(M)$, is defined on the space of germs, $\mathcal{O}^{(k)}_{M,p}$ at p. However, it is often convenient to define v on C^k-functions $f \in \mathcal{C}^k(U)$, where U is some open subset containing p. This is easy: Set

$$v(f) = v(\mathbf{f}).$$

Given any chart, (U, φ), at p, since v can be written in a unique way as

$$v = \sum_{i=1}^{n} \lambda_i \left(\frac{\partial}{\partial x_i}\right)_p,$$

we get

$$v(f) = \sum_{i=1}^{n} \lambda_i \left(\frac{\partial}{\partial x_i}\right)_p f.$$

This shows that $v(f)$ is the *directional derivative of* f *in the direction* v.

When M is a smooth manifold, things get a little simpler. Indeed, it turns out that in this case, every linear derivation vanishes on stationary germs. To prove this, we recall the following result from calculus (see Warner [59]):

Proposition 3.4 *If $g\colon \mathbb{R}^n \to \mathbb{R}$ is a C^k-function ($k \geq 2$) on a convex open, U, about $p \in \mathbb{R}^n$, then for every $q \in U$, we have*

$$g(q) = g(p) + \sum_{i=1}^{n} \frac{\partial g}{\partial X_i}\bigg|_p (q_i - p_i) + \sum_{i,j=1}^{n} (q_i - p_i)(q_j - p_j) \int_0^1 (1-t) \left.\frac{\partial^2 g}{\partial X_i \partial X_j}\right|_{(1-t)p+tq} dt.$$

In particular, if $g \in C^\infty(U)$, then the integral as a function of q is C^∞.

Proposition 3.5 *Let M be any C^∞-manifold of dimension n. For any $p \in M$, any linear derivation on $\mathcal{O}^{(\infty)}_{M,p}$ vanishes on stationary germs.*

Proof. Pick some chart, (U, φ), at p, where U is convex (for instance, an open ball) and let $\mathbf{f}$ be any stationary germ. If we apply Proposition 3.4 to $f \circ \varphi^{-1}$ and then compose with φ, we get

$$f = f(p) + \sum_{i=1}^{n} \left.\frac{\partial (f \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)} (x_i - x_i(p)) + \sum_{i,j=1}^{n} (x_i - x_i(p))(x_j - x_j(p))h,$$

near p, where h is C^∞. Since $\mathbf{f}$ is a stationary germ, this yields

$$f = f(p) + \sum_{i,j=1}^{n} (x_i - x_i(p))(x_j - x_j(p))h.$$

If v is any linear derivation, we get

$$\begin{aligned} v(f) = v(f(p)) + \sum_{i,j=1}^{n} \Big[&(x_i - x_i(p))(p)(x_j - x_j(p))(p)v(h) \\ &+ (x_i - x_i(p))(p)v(x_j - x_j(p))h(p) + v(x_i - x_i(p))(x_j - x_j(p))(p)h(p) \Big] = 0. \end{aligned}$$

Thus, v vanishes on stationary germs. $\square$

Proposition 3.5 shows that in the case of a smooth manifold, in Definition 3.10, we can omit the requirement that linear derivations vanish on stationary germs, since this is automatic. It is also possible to define $T_p(M)$ just in terms of $\mathcal{O}^{(\infty)}_{M,p}$. Let $\mathfrak{m}_{M,p} \subseteq \mathcal{O}^{(\infty)}_{M,p}$ be the ideal of germs that vanish at p. Then, we also have the ideal $\mathfrak{m}^2_{M,p}$, which consists of all finite sums of products of two elements in $\mathfrak{m}_{M,p}$, and it can be shown that $T^*_p(M)$ is isomorphic to $\mathfrak{m}_{M,p}/\mathfrak{m}^2_{M,p}$ (see Warner [59], Lemma 1.16).

Actually, if we let $\mathfrak{m}^{(k)}_{M,p}$ denote the C^k germs that vanish at p and $\mathfrak{s}^{(k)}_{M,p}$ denote the stationary C^k-germs that vanish at p, it is easy to show that

$$\mathcal{O}^{(k)}_{M,p}/\mathcal{S}^{(k)}_{M,p} \cong \mathfrak{m}^{(k)}_{M,p}/\mathfrak{s}^{(k)}_{M,p}.$$

(Given any $\mathbf{f} \in \mathcal{O}^{(k)}_{M,p}$, send it to $\mathbf{f} - \mathbf{f}(\mathbf{p}) \in \mathfrak{m}^{(k)}_{M,p}$.) Clearly, $(\mathfrak{m}^{(k)}_{M,p})^2$ consists of stationary germs (by the derivation property) and when $k = \infty$, Proposition 3.4 shows that every stationary germ that vanishes at p belongs to $\mathfrak{m}^2_{M,p}$. Therefore, when $k = \infty$, we have $\mathfrak{s}^{(\infty)}_{M,p} = \mathfrak{m}^2_{M,p}$ and so,

$$T^*_p(M) = \mathcal{O}^{(\infty)}_{M,p}/\mathcal{S}^{(\infty)}_{M,p} \cong \mathfrak{m}_{M,p}/\mathfrak{m}^2_{M,p}.$$

Remark: The ideal $\mathfrak{m}^{(k)}_{M,p}$ is in fact the unique maximal ideal of $\mathcal{O}^{(k)}_{M,p}$. This is because if $\mathbf{f} \in \mathcal{O}^{(k)}_{M,p}$ does not vanish at p, then it is an invertible element of $\mathcal{O}^{(k)}_{M,p}$ and any ideal containing $\mathfrak{m}^{(k)}_{M,p}$ and $\mathbf{f}$ would be equal to $\mathcal{O}^{(k)}_{M,p}$, which it absurd. Thus, $\mathcal{O}^{(k)}_{M,p}$ is a local ring (in the sense of commutative algebra) called the *local ring of germs of C^k-functions at p*. These rings play a crucial role in algebraic geometry.

Yet one more way of defining tangent vectors will make it a little easier to define tangent bundles.

Definition 3.12 *(Tangent Vectors, Version 3)* Given any C^k-manifold, M, of dimension n, with $k \geq 1$, for any $p \in M$, consider the triples, (U, φ, u), where (U, φ) is any chart at p and u is any vector in $\mathbb{R}^n$. Say that two such triples (U, φ, u) and (V, ψ, v) are *equivalent* iff

$$(\psi \circ \varphi^{-1})'_{\varphi(p)}(u) = v.$$

A *tangent vector* to M at p is an equivalence class of triples, $[(U, \varphi, u)]$, for the above equivalence relation.

The intuition behind Definition 3.12 is quite clear: The vector u is considered as a tangent vector to $\mathbb{R}^n$ at $\varphi(p)$. If (U, φ) is a chart on M at p, we can define a natural isomorphism, $\theta_{U,\varphi,p} \colon \mathbb{R}^n \to T_p(M)$, between $\mathbb{R}^n$ and $T_p(M)$, as follows: For any $u \in \mathbb{R}^n$,

$$\theta_{U,\varphi,p} \colon u \mapsto [(U, \varphi, u)].$$

One immediately check that the above map is indeed linear and a bijection.

The equivalence of this definition with the definition in terms of curves (Definition 3.7) is easy to prove.

Proposition 3.6 *Let M be any C^k-manifold of dimension n, with $k \geq 1$. For any $p \in M$, let x be any tangent vector (version 1) given by some equivalence class of C^1-curves, $\gamma \colon]-\epsilon, +\epsilon[\to M$, through p (i.e., $p = \gamma(0)$). The map*

$$x \mapsto [(U, \varphi, (\varphi \circ \gamma)'(0))]$$

is an isomorphism between $T_p(M)$-version 1 and $T_p(M)$-version 3.

Proof. If σ is another curve equivalent to γ, then $(\varphi \circ \gamma)'(0) = (\varphi \circ \sigma)'(0)$, so the map is well-defined. It is clearly injective. As for surjectivity, define the curve, γ, on M through p by

$$\gamma(t) = \varphi^{-1}(\varphi(p) + tu).$$

Then, $(\varphi \circ \gamma)(t) = \varphi(p) + tu$ and

$$(\varphi \circ \gamma)'(0) = u.$$

□

For simplicity of notation, we also use the notation T_pM for $T_p(M)$ (resp. T_p^*M for $T_p^*(M)$).

After having explored thorougly the notion of tangent vector, we show how a C^k-map, $h\colon M \to N$, between C^k manifolds, induces a linear map, $dh_p\colon T_p(M) \to T_{h(p)}(N)$, for every $p \in M$. We find it convenient to use Version 2 of the definition of a tangent vector. So, let $u \in T_p(M)$ be a linear derivation on $\mathcal{O}^{(k)}_{M,p}$ that vanishes on $\mathcal{S}^{(k)}_{M,p}$. We would like $dh_p(u)$ to be a linear derivation on $\mathcal{O}^{(k)}_{N,h(p)}$ that vanishes on $\mathcal{S}^{(k)}_{N,h(p)}$. So, for every germ, $\mathbf{g} \in \mathcal{O}^{(k)}_{N,h(p)}$, set

$$dh_p(u)(\mathbf{g}) = u(\mathbf{g} \circ \mathbf{h}).$$

For any locally defined function, g, at $h(p)$ in the germ, $\mathbf{g}$ (at $h(p)$), it is clear that $g \circ h$ is locally defined at p and is C^k, so $\mathbf{g} \circ \mathbf{h}$ is indeed a C^k-germ at p. Moreover, if $\mathbf{g}$ is a stationary germ at $h(p)$, then for some chart, (V, ψ) on N at $q = h(p)$, we have $(g \circ \psi^{-1})'(\psi(q)) = 0$ and, for some chart (U, φ) at p on M, we get

$$(g \circ h \circ \varphi^{-1})'(\varphi(p)) = (g \circ \psi^{-1})(\psi(q))((\psi \circ h \circ \varphi^{-1})'(\varphi(p))) = 0,$$

which means that $\mathbf{g} \circ \mathbf{h}$ is stationary at p. Therefore, $dh_p(u) \in T_{h(p)}(M)$. It is also clear that dh_p is a linear map. We summarize all this in the following definition:

Definition 3.13 Given any two C^k-manifolds, M and N, of dimension m and n, respectively, for any C^k-map, $h\colon M \to N$, and for every $p \in M$, the *differential of* h *at* p or *tangent map*, $dh_p\colon T_p(M) \to T_{h(p)}(N)$, is the linear map defined so that

$$dh_p(u)(\mathbf{g}) = u(\mathbf{g} \circ \mathbf{h}),$$

for every $u \in T_p(M)$ and every germ, $\mathbf{g} \in \mathcal{O}^{(k)}_{N,h(p)}$. The linear map dh_p is also denoted T_ph (and sometimes, h'_p or D_ph).

The chain rule is easily generalized to manifolds.

Proposition 3.7 *Given any two C^k-maps $f\colon M \to N$ and $g\colon N \to P$ between smooth C^k-manifolds, for any $p \in M$, we have*

$$d(g \circ f)_p = dg_{f(p)} \circ df_p.$$

In the special case where $N = \mathbb{R}$, a C^k-map between the manifolds M and $\mathbb{R}$ is just a C^k-function on M. It is interesting to see what df_p is explicitly. Since $N = \mathbb{R}$, germs (of functions on $\mathbb{R}$) at $t_0 = f(p)$ are just germs of C^k-functions, $g\colon \mathbb{R} \to \mathbb{R}$, locally defined at t_0. Then, for any $u \in T_p(M)$ and every germ $\mathbf{g}$ at t_0,

$$df_p(u)(\mathbf{g}) = u(\mathbf{g} \circ \mathbf{f}).$$

If we pick a chart, (U, φ), on M at p, we know that the $\left(\frac{\partial}{\partial x_i}\right)_p$ form a basis of $T_p(M)$, with $1 \leq i \leq n$. Therefore, it is enough to figure out what $df_p(u)(\mathbf{g})$ is when $u = \left(\frac{\partial}{\partial x_i}\right)_p$. In this case,

$$df_p\left(\left(\frac{\partial}{\partial x_i}\right)_p\right)(\mathbf{g}) = \left.\frac{\partial(g \circ f \circ \varphi^{-1})}{\partial X_i}\right|_{\varphi(p)}.$$

Using the chain rule, we find that

$$df_p\left(\left(\frac{\partial}{\partial x_i}\right)_p\right)(\mathbf{g}) = \left(\frac{\partial}{\partial x_i}\right)_p f \left.\frac{dg}{dt}\right|_{t_0}.$$

Therefore, we have

$$df_p(u) = u(\mathbf{f}) \left.\frac{d}{dt}\right|_{t_0}.$$

This shows that we can identify df_p with the linear form in $T_p^*(M)$ defined by

$$df_p(v) = v(\mathbf{f}).$$

This is consistent with our previous definition of df_p as the image of f in $T_p^*(M) = \mathcal{O}_{M,p}^{(k)}/\mathcal{S}_{M,p}^{(k)}$ (as $T_p(M)$ is isomorphic to $(\mathcal{O}_{M,p}^{(k)}/\mathcal{S}_{M,p}^{(k)})^*$).

In preparation for the definition of the flow of a vector field (which will be needed to define the exponential map in Lie group theory), we need to define the tangent vector to a curve on a manifold. Given a C^k-curve, $\gamma\colon]a, b[\to M$, on a C^k-manifold, M, for any $t_0 \in]a, b[$, we would like to define the tangent vector to the curve γ at t_0 as a tangent vector to M at $p = \gamma(t_0)$. We do this as follows: Recall that $\left.\frac{d}{dt}\right|_{t_0}$ is a basis vector of $T_{t_0}(\mathbb{R}) = \mathbb{R}$. So, define the *tangent vector to the curve* γ *at* t, denoted $\dot{\gamma}(t_0)$ (or $\gamma'(t)$, or $\frac{d\gamma}{dt}(t_0)$) by

$$\dot{\gamma}(t) = d\gamma_t\left(\left.\frac{d}{dt}\right|_{t_0}\right).$$

Sometime, it is necessary to define curves (in a manifold) whose domain is not an open interval. A map, $\gamma\colon [a, b] \to M$, is a *$C^k$-curve in M* if it is the restriction of some C^k-curve, $\widetilde{\gamma}\colon]a-\epsilon, b+\epsilon[\to M$, for some (small) $\epsilon > 0$. Note that for such a curve (if $k \geq 1$) the tangent vector, $\dot{\gamma}(t)$, is defined for all $t \in [a, b]$, A curve, $\gamma\colon [a, b] \to M$, is *piecewise* C^k iff there a sequence, $a_0 = a, a_1, \ldots, a_m = b$, so that the restriction of γ to each $[a_i, a_{i+1}]$ is a C^k-curve, for $i = 0, \ldots, m-1$.

3.3 Tangent and Cotangent Bundles, Vector Fields

Let M be a C^k-manifold (with $k \geq 2$). Roughly speaking, a vector field on M is the assignment, $p \mapsto \xi(p)$, of a tangent vector, $\xi(p) \in T_p(M)$, to a point $p \in M$. Generally, we would like such assignments to have some smoothness properties when p varies in M, for example, to be C^l, for some l related to k. Now, if the collection, $T(M)$, of all tangent spaces, $T_p(M)$, was a C^l-manifold, then it would be very easy to define what we mean by a C^l-vector field: We would simply require the maps, $\xi\colon M \to T(M)$, to be C^l.

If M is a C^k-manifold of dimension n, then we can indeed define make $T(M)$ into a C^{k-1}-manifold of dimension $2n$ and we now sketch this construction.

We find it most convenient to use Version 3 of the definition of tangent vectors, i.e., as equivalence classes of triple (U, φ, u). First, we let $T(M)$ be the disjoint union of the tangent spaces $T_p(M)$, for all $p \in M$. There is a *natural projection*,

$$\pi\colon T(M) \to M, \quad \text{where} \quad \pi(v) = p \quad \text{if} \quad v \in T_p(M).$$

We still have to give $T(M)$ a topology and to define a C^{k-1}-atlas. For every chart, (U, φ), of M (with U open in M) we define the function $\widetilde{\varphi}\colon \pi^{-1}(U) \to \mathbb{R}^{2n}$ by

$$\widetilde{\varphi}(v) = (\varphi \circ \pi(v), \theta^{-1}_{U,\varphi,\pi(v)}(v)),$$

where $v \in \pi^{-1}(U)$ and $\theta_{U,\varphi,p}$ is the isomorphism between $\mathbb{R}^n$ and $T_p(M)$ described just after Definition 3.12. It is obvious that $\widetilde{\varphi}$ is a bijection between $\pi^{-1}(U)$ and $\varphi(U) \times \mathbb{R}^n$, an open subset of $\mathbb{R}^{2n}$. We give $T(M)$ the weakest topology that makes all the $\widetilde{\varphi}$ continuous, i.e., we take the collection of subsets of the form $\widetilde{\varphi}^{-1}(W)$, where W is any open subset of $\varphi(U) \times \mathbb{R}^n$, as a basis of the topology of $T(M)$. One easily checks that $T(M)$ is Hausdorff and second-countable in this topology. If (U, φ) and (V, ψ) are overlapping charts, then the transition function

$$\widetilde{\psi} \circ \widetilde{\varphi}^{-1}\colon \varphi(U \cap V) \times \mathbb{R}^n \longrightarrow \psi(U \cap V) \times \mathbb{R}^n$$

is given by

$$\widetilde{\psi} \circ \widetilde{\varphi}^{-1}(p, u) = (\psi \circ \varphi^{-1}(p), (\psi \circ \varphi^{-1})'(u)).$$

It is clear that $\widetilde{\psi} \circ \widetilde{\varphi}^{-1}$ is a C^{k-1}-map. Therefore, $T(M)$ is indeed a C^{k-1}-manifold of dimension $2n$, called the *tangent bundle*.

Remark: Even if the manifold M is naturally embedded in $\mathbb{R}^N$ (for some $N \geq n = \dim(M)$), it is not at all obvious how to view the tangent bundle, $T(M)$, as embedded in $\mathbb{R}^{N'}$, for sone suitable N'. Hence, we see that the definition of an abtract manifold is unavoidable.

A similar construction can be carried out for the cotangent bundle. In this case, we let $T^*(M)$ be the disjoint union of the cotangent spaces $T^*_p(M)$. We also have a natural projection, $\pi: T^*(M) \to M$, and we can define charts as follows: For any chart, (U, φ), on M, we define the function $\widetilde{\varphi}: \pi^{-1}(U) \to \mathbb{R}^{2n}$ by

$$\widetilde{\varphi}(\tau) = \left(\varphi \circ \pi(\tau), \tau \left(\left(\frac{\partial}{\partial x_1} \right)_{\pi(\tau)} \right), \ldots, \tau \left(\left(\frac{\partial}{\partial x_n} \right)_{\pi(\tau)} \right) \right),$$

where $\tau \in \pi^{-1}(U)$ and the $\left(\frac{\partial}{\partial x_i}\right)_p$ are the basis of $T_p(M)$ associated with the chart (U, φ). Again, one can make $T^*(M)$ into a C^{k-1}-manifold of dimension $2n$, called the *cotangent bundle*. We leave the details as an exercise to the reader (Or, look at Berger and Gostiaux [5]). For simplicity of notation, we also use the notation TM for $T(M)$ (resp. T^*M for $T^*(M)$).

Observe that for every chart, (U, φ), on M, there is a bijection

$$\tau_U: \pi^{-1}(U) \to U \times \mathbb{R}^n,$$

given by

$$\tau_U(v) = (\pi(v), \theta^{-1}_{U,\varphi,\pi(v)}(v)).$$

Clearly, $pr_1 \circ \tau_U = \pi$, on $\pi^{-1}(U)$. Thus, locally, that is, over U, the bundle $T(M)$ looks like the product $U \times \mathbb{R}^n$. We say that $T(M)$ is *locally trivial* (over U) and we call τ_U a *trivializing map*. For any $p \in M$, the vector space $\pi^{-1}(p) = T_p(M)$ is called the *fibre above* p. Observe that the restriction of τ_U to $\pi^{-1}(p)$ is an isomorphism between $T_p(M)$ and $\{p\} \times \mathbb{R}^n \cong \mathbb{R}^n$, for any $p \in M$. All these ingredients are part of being a *vector bundle* (but a little more is required of the maps τ_U). For more on bundles, see Lang [38], Gallot, Hulin and Lafontaine [28], Lafontaine [37] or Bott and Tu [7].

When $M = \mathbb{R}^n$, observe that $T(M) = M \times \mathbb{R}^n = \mathbb{R}^n \times \mathbb{R}^n$, i.e., the bundle $T(M)$ is (globally) trivial.

Given a C^k-map, $h: M \to N$, between two C^k-manifolds, we can define the function, $dh: T(M) \to T(N)$, (also denoted Th, or h_*, or Dh) by setting

$$dh(u) = dh_p(u), \quad \text{iff} \quad u \in T_p(M).$$

We leave the next proposition as an exercise to the reader (A proof can be found in Berger and Gostiaux [5]).

Proposition 3.8 *Given a C^k-map, $h: M \to N$, between two C^k-manifolds M and N (with $k \geq 1$), the map $dh: T(M) \to T(N)$ is a C^{k-1}-map.*

We are now ready to define vector fields.

Definition 3.14 Let M be a C^{k+1} manifold, with $k \geq 1$. For any open subset, U of M, a *vector field on U* is any *section, ξ, of $T(M)$ over U*, i.e., any function, $\xi: U \to T(M)$, such that $\pi \circ \xi = \mathrm{id}_U$ (i.e., $\xi(p) \in T_p(M)$, for every $p \in U$). We also say that ξ is a *lifting of U into $T(M)$*. We say that ξ is a *C^h-vector field on U* iff ξ is a section over U and a C^h-map, where $0 \leq h \leq k$. The set of C^k-vector fields over U is denoted $\Gamma^{(k)}(U, T(M))$. Given a curve, $\gamma\colon [a,b] \to M$, a *vector field, ξ, along γ* is any section of $T(M)$ over γ, i.e., a C^k-function, $\xi\colon [a,b] \to T(M)$, such that $\pi \circ \xi = \gamma$. We also say that ξ *lifts γ into $T(M)$*.

The above definition gives a precise meaning to the idea that a C^k-vector field on M is an assignment, $p \mapsto \xi(p)$, of a tangent vector, $\xi(p) \in T_p(M)$, to a point, $p \in M$, so that $\xi(p)$ varies in a C^k-fashion in terms of p.

Clearly, $\Gamma^{(k)}(U, T(M))$ is a real vector space. For short, the space $\Gamma^{(k)}(M, T(M))$ is also denoted by $\Gamma^{(k)}(T(M))$ (or $\mathfrak{X}^{(k)}(M)$ or even $\Gamma(T(M))$ or $\mathfrak{X}(M)$). If $M = \mathbb{R}^n$ and U is an open subset of M, then $T(M) = \mathbb{R}^n \times \mathbb{R}^n$ and a section of $T(M)$ over U is simply a function, ξ, such that

$$\xi(p) = (p, u), \quad \text{with} \quad u \in \mathbb{R}^n,$$

for all $p \in U$. In other words, ξ is defined by a function, $f\colon U \to \mathbb{R}^n$ (namely, $f(p) = u$). This corresponds to the "old" definition of a vector field in the more basic case where the manifold, M, is just $\mathbb{R}^n$.

Given any C^k-function, $f \in \mathcal{C}^k(U)$, and a vector field, $\xi \in \Gamma^{(k)}(U, T(M))$, we define the vector field, $f\xi$, by

$$(f\xi)(p) = f(p)\xi(p), \quad p \in U.$$

Obviously, $f\xi \in \Gamma^{(k)}(U, T(M))$, which shows that $\Gamma^{(k)}(U, T(M))$ is also a $\mathcal{C}^k(U)$-module. We also denote $\xi(p)$ by ξ_p. For any chart, (U, φ), on M it is easy to check that the map

$$p \mapsto \left(\frac{\partial}{\partial x_i}\right)_p, \quad p \in U,$$

is a C^k-vector field on U (with $1 \leq i \leq n$). This vector field is denoted $\left(\frac{\partial}{\partial x_i}\right)$ or $\frac{\partial}{\partial x_i}$.

If U is any open subset of M and f is any function in $\mathcal{C}^k(U)$, then $\xi(f)$ is the function on U given by

$$\xi(f)(p) = \xi_p(f) = \xi_p(\mathbf{f}).$$

As a special case, when (U, φ) is a chart on M, the vector field, $\frac{\partial}{\partial x_i}$, just defined above induces the function

$$p \mapsto \left(\frac{\partial}{\partial x_i}\right)_p f, \quad p \in U,$$

denoted $\frac{\partial}{\partial x_i}(f)$ or $\left(\frac{\partial}{\partial x_i}\right) f$. It is easy to check that $\xi(f) \in \mathcal{C}^{k-1}(U)$. As a consequence, every vector field $\xi \in \Gamma^{(k)}(U, T(M))$ induces a linear map,

$$L_\xi\colon \mathcal{C}^k(U) \longrightarrow \mathcal{C}^{k-1}(U),$$

given by $f \mapsto \xi(f)$. It is immediate to check that L_ξ has the Leibnitz property, i.e.,

$$L_\xi(fg) = L_\xi(f)g + fL_\xi(g).$$

Linear maps with this property are called *derivations*. Thus, we see that every vector field induces some kind of differential operator, namely, a linear derivation. Unfortunately, not every linear derivation of the above type arises from a vector field, although this turns out to be true in the smooth case i.e., when $k = \infty$ (for a proof, see Gallot, Hulin and Lafontaine [28] or Lafontaine [37]).

In the rest of this section, unless stated otherwise, we assume that $k \geq 1$. The following easy proposition holds (c.f. Warner [59]):

Proposition 3.9 *Let ξ be a vector field on the C^{k+1}-manifold, M, of dimension n. Then, the following are equivalent:*

(a) ξ is C^k.

(b) If (U, φ) is a chart on M and if $f_1, \ldots, f_n$ are the functions on U uniquely defined by

$$\xi \restriction U = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i},$$

then each f_i is a C^k-map.

(c) Whenever U is open in M and $f \in \mathcal{C}^k(U)$, then $\xi(f) \in \mathcal{C}^{k-1}(U)$.

Given any two C^k-vector field, ξ, η, on M, for any function, $f \in \mathcal{C}^k(M)$, we defined above the function $\xi(f)$ and $\eta(f)$. Thus, we can form $\xi(\eta(f))$ (resp. $\eta(\xi(f))$), which are in $\mathcal{C}^{k-2}(M)$. Unfortunately, even in the smooth case, there is generally no vector field, ζ, such that

$$\zeta(f) = \xi(\eta(f)), \quad \text{for all } f \in \mathcal{C}^k(M).$$

This is because $\xi(\eta(f))$ (and $\eta(\xi(f))$) involve second-order derivatives. However, if we consider $\xi(\eta(f)) - \eta(\xi(f))$, then second-order derivatives cancel out and there is a unique vector field inducing the above differential operator. Intuitively, $\xi\eta - \eta\xi$ measures the "failure of ξ and η to commute".

Proposition 3.10 *Given any C^{k+1}-manifold, M, of dimension n, for any two C^k-vector fields, ξ, η, on M, there is a unique C^{k-1}-vector field, $[\xi, \eta]$, such that*

$$[\xi, \eta](f) = \xi(\eta(f)) - \eta(\xi(f)), \quad \textit{for all} \quad f \in \mathcal{C}^{k-1}(M).$$

Proof. First we prove uniqueness. For this it is enough to prove that $[\xi, \eta]$ is uniquely defined on $\mathcal{C}^k(U)$, for any chart, (U, φ). Over U, we know that

$$\xi = \sum_{i=1}^{n} \xi_i \frac{\partial}{\partial x_i} \quad \text{and} \quad \eta = \sum_{i=1}^{n} \eta_i \frac{\partial}{\partial x_i},$$

where $\xi_i, \eta_i \in \mathcal{C}^k(U)$. Then, for any $f \in \mathcal{C}^k(M)$, we have

$$\begin{aligned}\xi(\eta(f)) &= \xi\left(\sum_{j=1}^n \eta_j \frac{\partial}{\partial x_j}(f)\right) = \sum_{i,j=1}^n \xi_i \frac{\partial}{\partial x_i}(\eta_j)\frac{\partial}{\partial x_j}(f) + \sum_{i,j=1}^n \xi_i\eta_j \frac{\partial^2}{\partial x_j \partial x_i}(f)\\ \eta(\xi(f)) &= \eta\left(\sum_{i=1}^n \xi_i \frac{\partial}{\partial x_i}(f)\right) = \sum_{i,j=1}^n \eta_j \frac{\partial}{\partial x_j}(\xi_i)\frac{\partial}{\partial x_i}(f) + \sum_{i,j=1}^n \xi_i\eta_j \frac{\partial^2}{\partial x_i \partial x_j}(f).\end{aligned}$$

However, as $f \in \mathcal{C}^k(M)$, with $k \geq 2$, we have

$$\sum_{i,j=1}^n \xi_i\eta_j \frac{\partial^2}{\partial x_j \partial x_i}(f) = \sum_{i,j=1}^n \xi_i\eta_j \frac{\partial^2}{\partial x_i \partial x_j}(f),$$

and we deduce that

$$\xi(\eta(f)) - \eta(\xi(f)) = \sum_{i,j=1}^n \left(\xi_i \frac{\partial}{\partial x_i}(\eta_j) - \eta_i \frac{\partial}{\partial x_i}(\xi_j)\right)\frac{\partial}{\partial x_j}(f).$$

This proves that $[\xi, \eta] = \xi\eta - \eta\xi$ is uniquely defined on U and that it is C^{k-1}. Thus, if $[\xi, \eta]$ exists, it is unique.

To prove existence, we use the above expression to define $[\xi, \eta]_U$, locally on U, for every chart, (U, φ). On any overlap, $U \cap V$, by the uniqueness property that we just proved, $[\xi, \eta]_U$ and $[\xi, \eta]_V$ must agree. But then, the $[\xi, \eta]_U$ patch and yield a C^{k-1}-vector field defined on the whole of M. $\square$

Definition 3.15 Given any C^{k+1}-manifold, M, of dimension n, for any two C^k-vector fields, ξ, η, on M, the *Lie bracket*, $[\xi, \eta]$, of ξ and η, is the C^{k-1} vector field defined so that

$$[\xi, \eta](f) = \xi(\eta(f)) - \eta(\xi(f)), \quad \text{for all} \quad f \in \mathcal{C}^{k-1}(M).$$

We also have the following simple proposition whose proof is left as an exercise (or, see Do Carmo [22]):

Proposition 3.11 *Given any C^{k+1}-manifold, M, of dimension n, for any C^k-vector fields, ξ, η, ζ, on M, for all $f, g \in \mathcal{C}^k(M)$, we have:*

(a) $[[\xi, \eta], \zeta] + [[\eta, \zeta], \xi] + [[\zeta, \xi], \eta] = 0$ *(Jacobi identity).*

(b) $[\xi, \xi] = 0$.

(c) $[f\xi, g\eta] = fg[\xi, \eta] + f\xi(g)\eta - g\eta(f)\xi$.

(d) $[-, -]$ *is bilinear.*

As a consequence, for smooth manifolds ($k = \infty$), the space of vector fields, $\Gamma^{(\infty)}(T(M))$, is a vector space equipped with a bilinear operation, $[-,-]$, that satisfies the Jacobi identity. This makes $\Gamma^{(\infty)}(T(M))$ a *Lie algebra*.

One more notion will be needed when we deal with Lie algebras.

Definition 3.16 Let $\varphi: M \to N$ be a C^{k+1}-map of manifolds. If ξ is a C^k vector field on M and η is a C^k vector field on N, we say that *ξ and η are φ-related* iff

$$d\varphi \circ \xi = \eta \circ \varphi.$$

The basic result about φ-related vector fields is:

Proposition 3.12 *Let $\varphi: M \to N$ be a C^{k+1}-map of manifolds, let ξ and ξ_1 be C^k vector fields on M and let η, η_1 be C^k vector fields on N. If ξ is φ-related to ξ_1 and η is φ-related to η_1, then $[\xi, \eta]$ is φ-related to $[\xi_1, \eta_1]$.*

Proof. Basically, one needs to unwind the definitions, see Warner [59], Chapter 1. $\square$

3.4 Submanifolds, Immersions, Embeddings

Although the notion of submanifold is intuitively rather clear, technically, it is a bit tricky. In fact, the reader may have noticed that many different definitions appear in books and that it is not obvious at first glance that these definitions are equivalent. What is important is that a submanifold, N of a given manifold, M, not only have the topology induced M but also that the charts of N be somewhow induced by those of M. (Recall that if X is a topological space and Y is a subset of X, then the *subspace topology on Y* or *topology induced by X on Y* has for its open sets all subsets of the form $Y \cap U$, where U is an arbitary subset of X.).

Given m, n, with $0 \leq m \leq n$, we can view $\mathbb{R}^m$ as a subspace of $\mathbb{R}^n$ using the inclusion

$$\mathbb{R}^m \cong \mathbb{R}^m \times \{\underbrace{(0,\ldots,0)}_{n-m}\} \hookrightarrow \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n, \quad (x_1,\ldots,x_m) \mapsto (x_1,\ldots,x_m,\underbrace{0,\ldots,0}_{n-m}).$$

Definition 3.17 Given a C^k-manifold, M, of dimension n, a subset, N, of M is an *m-dimensional submanifold of M* (where $0 \leq m \leq n$) iff for every point, $p \in N$, there is a chart, (U, φ), of M, with $p \in U$, so that

$$\varphi(U \cap N) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\}).$$

(We write $0_{n-m} = \underbrace{(0,\ldots,0)}_{n-m}$.)

The subset, $U \cap N$, of Definition 3.17 is sometimes called a *slice* of (U, φ) and we say that (U, φ) is *adapted to N* (See O'Neill [49] or Warner [59]).

Other authors, including Warner [59], use the term submanifold in a broader sense than us and they use the word *embedded submanifold* for what is defined in Definition 3.17.

The following proposition has an almost trivial proof but it justifies the use of the word submanifold:

Proposition 3.13 *Given a C^k-manifold, M, of dimension n, for any submanifold, N, of M of dimension $m \leq n$, the family of pairs $(U \cap N, \varphi \restriction U \cap N)$, where (U, φ) ranges over the charts over any atlas for M, is an atlas for N, where N is given the subspace topology. Therefore, N inherits the structure of a C^k-manifold.*

In fact, every chart on N arises from a chart on M in the following precise sense:

Proposition 3.14 *Given a C^k-manifold, M, of dimension n and a submanifold, N, of M of dimension $m \leq n$, for any $p \in N$ and any chart, (W, η), of N at p, there is some chart, (U, φ), of M at p so that*

$$\varphi(U \cap N) = \varphi(U) \cap (\mathbb{R}^m \times \{0_{n-m}\}) \quad \textit{and} \quad \varphi \restriction U \cap N = \eta \restriction U \cap N,$$

where $p \in U \cap N \subseteq W$.

Proof. See Berger and Gostiaux [5] (Chapter 2). $\square$

It is also useful to define more general kinds of "submanifolds".

Definition 3.18 Let $\varphi\colon N \to M$ be a C^k-map of manifolds.

(a) The map φ is an *immersion* of N into M iff $d\varphi_p$ is injective for all $p \in N$.

(b) The set $\varphi(N)$ is an *immersed submanifold* of M iff φ is an injective immersion.

(c) The map φ is an *embedding* of N into M iff φ is an injective immersion such that the induced map, $N \longrightarrow \varphi(N)$, is a homeomorphism, where $\varphi(N)$ is given the subspace topology (equivalently, φ is an open map from N into $\varphi(N)$ with the subspace topology). We say that $\varphi(N)$ (with the subspace topology) is an *embedded submanifold* of M.

(d) The map φ is a *submersion* of N into M iff $d\varphi_p$ is surjective for all $p \in N$.

Again, we warn our readers that certain authors (such as Warner [59]) call $\varphi(N)$, in (b), a submanifold of M! We prefer the terminology *immersed submanifold*.

The notion of immersed submanifold arises naturally in the framewok of Lie groups. Indeed, the fundamental correspondence between Lie groups and Lie algebras involves Lie subgroups that are not necessarily closed. But, as we will see later, subgroups of Lie groups that are also submanifolds are always closed. It is thus necessary to have a more inclusive notion of submanifold for Lie groups and the concept of immersed submanifold is just what's needed.

Immersions of $\mathbb{R}$ into $\mathbb{R}^3$ are parametric curves and immersions of $\mathbb{R}^2$ into $\mathbb{R}^3$ are parametric surfaces. These have been extensively studied, for example, see DoCarmo [21], Berger and Gostiaux [5] or Gallier [27].

Immersions (i.e., subsets of the form $\varphi(N)$, where N is an immersion) are generally neither injective immersions (i.e., subsets of the form $\varphi(N)$, where N is an injective immersion) nor embeddings (or submanifolds). For example, immersions can have self-intersections, as the plane curve (nodal cubic): $x = t^2 - 1; y = t(t^2 - 1)$.

Injective immersions are generally not embeddings (or submanifolds) because $\varphi(N)$ may not be homeomorphic to N. An example is given by the Lemniscate of Bernoulli, an injective immersion of $\mathbb{R}$ into $\mathbb{R}^2$:

$$\begin{aligned} x &= \frac{t(1+t^2)}{1+t^4}, \\ y &= \frac{t(1-t^2)}{1+t^4}. \end{aligned}$$

Another interesting example is the immersion of $\mathbb{R}$ into the 2-torus, $T^2 = S^1 \times S^1 \subseteq \mathbb{R}^4$, given by

$$t \mapsto (\cos t, \sin t, \cos ct, \sin ct),$$

where $c \in \mathbb{R}$. One can show that the image of $\mathbb{R}$ under this immersion is closed in T^2 iff c is rational. Moreover, the image of this immersion is dense in T^2 but not closed iff c is irrational. The above example can be adapted to the torus in $\mathbb{R}^3$: One can show that the immersion given by

$$t \mapsto ((2 + \cos t)\cos(\sqrt{2}\, t), (2 + \cos t)\sin(\sqrt{2}\, t), \sin t),$$

is dense but not closed in the torus (in $\mathbb{R}^3$) given by

$$(s, t) \mapsto ((2 + \cos s)\cos t, (2 + \cos s)\sin t, \sin s),$$

where $s, t \in \mathbb{R}$.

There is, however, a close relationship between submanifolds and embeddings.

Proposition 3.15 *If N is a submanifold of M, then the inclusion map, $j\colon N \to M$, is an embedding. Conversely, if $\varphi\colon N \to M$ is an embedding, then $\varphi(N)$ with the subspace topology is a submanifold of M and φ is a diffeomorphism between N and $\varphi(N)$.*

Proof. See O'Neill [49] (Chapter 1) or Berger and Gostiaux [5] (Chapter 2). □

In summary, embedded submanifolds and (our) submanifolds coincide. Some authors refer to spaces of the form $\varphi(N)$, where φ is an injective immersion, as *immersed submanifolds*. However, in general, an immersed submanifold is *not* a submanifold. One case where this holds is when N is compact, since then, a bijective continuous map is a homeomorphism. For yet a notion of submanifold intermediate between immersed submanifolds and (our) submanifolds, see Sharpe [57] (Chapter 1).

Our next goal is to review and promote to manifolds some standard results about ordinary differential equations.

3.5 Integral Curves, Flow of a Vector Field, One-Parameter Groups of Diffeomorphisms

We begin with integral curves and (local) flows of vector fields on a manifold.

Definition 3.19 Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$ and let p_0 be a point on M. An *integral curve (or trajectory) for* ξ *with initial condition* p_0 is a C^{p-1}-curve, $\gamma\colon I \to M$, so that

$$\dot{\gamma}(t) = \xi(\gamma(t)), \quad \text{for all } t \in I \quad \text{and} \quad \gamma(0) = p_0,$$

where $I = \,]a, b[\, \subseteq \mathbb{R}$ is an open interval containing 0.

What definition 3.19 says is that an integral curve, γ, with initial condition p_0 is a curve on the manifold M passing through p_0 and such that, for every point $p = \gamma(t)$ on this curve, the tangent vector to this curve at p, i.e., $\dot{\gamma}(t)$, coincides with the value, $\xi(p)$, of the vector field ξ at p.

Given a vector field, ξ, as above, and a point $p_0 \in M$, is there an integral curve through p_0? Is such a curve unique? If so, how large is the open interval I? We provide some answers to the above questions below.

Definition 3.20 Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$ and let p_0 be a point on M. A *local flow for* ξ *at* p_0 is a map,

$$\varphi\colon J \times U \to M,$$

where $J \subseteq \mathbb{R}$ is an open interval containing 0 and U is an open subset of M containing p_0, so that for every $p \in U$, the curve $t \mapsto \varphi(t, p)$ is an integral curve of ξ with initial condition p.

Thus, a local flow for ξ is a family of integral curves for all points in some small open set around p_0 such that these curves all have the same domain, J, independently of the initial condition, $p \in U$.

The following theorem is the main existence theorem of local flows. This is a promoted version of a similar theorem in the classical theory of ODE's in the case where M is an open subset of $\mathbb{R}^n$. For a full account of this theory, see Lang [38] or Berger and Gostiaux [5].

Theorem 3.16 *(Existence of a local flow) Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$ and let p_0 be a point on M. There is an open interval, $J \subseteq \mathbb{R}$, containing 0 and an open subset, $U \subseteq M$, containing p_0, so that there is a* **unique** *local flow, $\varphi\colon J \times U \to M$, for ξ at p_0. Furthermore, φ is C^{k-1}.*

Theorem 3.16 holds under more general hypotheses, namely, when the vector field satisfies some *Lipschitz* condition, see Lang [38] or Berger and Gostiaux [5].

Now, we know that for any initial condition, p_0, there is some integral curve through p_0. However, there could be two (or more) integral curves $\gamma_1\colon I_1 \to M$ and $\gamma_2\colon I_2 \to M$ with initial condition p_0. This leads to the natural question: How do γ_1 and γ_2 differ on $I_1 \cap I_2$? The next proposition shows they don't!

Proposition 3.17 *Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$ and let p_0 be a point on M. If $\gamma_1\colon I_1 \to M$ and $\gamma_2\colon I_2 \to M$ are any two integral curves both with initial condition p_0, then $\gamma_1 = \gamma_2$ on $I_1 \cap I_2$.*

Proof. Let $Q = \{t \in I_1 \cap I_2 \mid \gamma_1(t) = \gamma_2(t)\}$. Since $\gamma_1(0) = \gamma_2(0) = p_0$, the set Q is nonempty. If we show that Q is both closed and open in $I_1 \cap I_2$, as $I_1 \cap I_2$ is connected since it is an open interval of $\mathbb{R}$, we will be able to conclude that $Q = I_1 \cap I_2$.

Since by definition, a manifold is Hausdorff, it is a standard fact in topology that the diagonal, $\Delta = \{(p, p) \mid p \in M\} \subseteq M \times M$, is closed, and since

$$Q = I_1 \cap I_2 \cap (\gamma_1, \gamma_2)^{-1}(\Delta)$$

and γ_1 and γ_2 are continuous, we see that Q is closed in $I_1 \cap I_2$.

Pick any $u \in Q$ and consider the curves β_1 and β_2 given by

$$\beta_1(t) = \gamma_1(t + u) \quad \text{and} \quad \beta_2(t) = \gamma_2(t + u),$$

where $t \in I_1 - u$ in the first case and $t \in I_2 - u$ in the second. (Here, if $I =]a, b[$, we have $I - u =]a - u, b - u[$.) Observe that

$$\dot{\beta}_1(t) = \dot{\gamma}_1(t + u) = \xi(\gamma_1(t + u)) = \xi(\beta_1(t))$$

and similarly, $\dot{\beta}_2(t) = \xi(\beta_2(t))$. We also have

$$\beta_1(0) = \gamma_1(u) = \gamma_2(u) = \beta_2(0) = q,$$

since $u \in Q$ (where $\gamma_1(u) = \gamma_2(u)$). Thus, $\beta_1\colon (I_1 - u) \to M$ and $\beta_2\colon (I_2 - u) \to M$ are two integral curves with the same initial condition, q. By Theorem 3.16, the uniqueness of local flow implies that there is some open interval, $\widetilde{I} \subseteq I_1 \cap I_2 - u$, such that $\beta_1 = \beta_2$ on $\widetilde{I}$. Consequently, γ_1 and γ_2 agree on $\widetilde{I} + u$, an open subset of Q, proving that Q is indeed open in $I_1 \cap I_2$. $\square$

Proposition 3.17 implies the important fact that there is a *unique maximal* integral curve with initial condition p. Indeed, if $\{\gamma_k\colon I_k \to M\}_{k\in K}$ is the family of all integral curves with initial condition p (for some big index set, K), if we let $I(p) = \bigcup_{k\in K} I_k$, we can define a curve, $\gamma_p\colon I(p) \to M$, so that

$$\gamma_p(t) = \gamma_k(t), \quad \text{if} \quad t \in I_k.$$

Since γ_k and γ_l agree on $I_k \cap I_l$ for all $k, l \in K$, the curve γ_p is indeed well defined and it is clearly an integral curve with initial condition p with the largest possible domain (the open interval, $I(p)$). The curve γ_p is called the *maximal integral curve with initial condition* p and it is also denoted $\gamma(t, p)$. Note that Proposition 3.17 implies that any two distinct integral curves are disjoint, i.e., do not intersect each other.

The following interesting question now arises: Given any $p_0 \in M$, if $\gamma_{p_0}\colon I(p_0) \to M$ is the maximal integral curve with initial condition p_0, for any $t_1 \in I(p_0)$, and if $p_1 = \gamma_{p_0}(t_1) \in M$, then there is a maximal integral curve, $\gamma_{p_1}\colon I(p_1) \to M$, with initial condition p_1. What is the relationship between γ_{p_0} and γ_{p_1}, if any? The answer is given by

Proposition 3.18 *Let ξ be a C^{k-1} vector field on a C^k-manifold, M, ($k \geq 2$) and let p_0 be a point on M. If $\gamma_{p_0}\colon I(p_0) \to M$ is the maximal integral curve with initial condition p_0, for any $t_1 \in I(p_0)$, if $p_1 = \gamma_{p_0}(t_1) \in M$ and $\gamma_{p_1}\colon I(p_1) \to M$ is the maximal integral curve with initial condition p_1, then*

$$I(p_1) = I(p_0) - t_1 \quad \textit{and} \quad \gamma_{p_1}(t) = \gamma_{\gamma_{p_0}(t_1)}(t) = \gamma_{p_0}(t + t_1), \quad \textit{for all } t \in I(p_0) - t_1.$$

Proof. Let $\gamma(t)$ be the curve given by

$$\gamma(t) = \gamma_{p_0}(t + t_1), \quad \text{for all } \; t \in I(p_0) - t_1.$$

Clearly, γ is defined on $I(p_0) - t_1$ and

$$\dot{\gamma}(t) = \dot{\gamma}_{p_0}(t + t_1) = \xi(\gamma_{p_0}(t + t_1)) = \xi(\gamma(t))$$

and $\gamma(0) = \gamma_{p_0}(t_1) = p_1$. Thus, γ is an integal curve defined on $I(p_0) - t_1$ with initial condition p_1. If γ was defined on an interval, $\widetilde{I} \supseteq I(p_0) - t_1$ with $\widetilde{I} \neq I(p_0) - t_1$, then γ_{p_0} would be defined on $\widetilde{I} + t_1 \supset I(p_0)$, an interval strictly bigger than $I(p_0)$, contradicting the maximality of $I(p_0)$. Therefore, $I(p_0) - t_1 = I(p_1)$. $\square$

It is useful to restate Proposition 3.18 by changing point of view. So far, we have been focusing on integral curves, i.e., given any $p_0 \in M$, we let t vary in $I(p_0)$ and get an integral curve, γ_{p_0}, with domain $I(p_0)$.

Instead of holding $p_0 \in M$ fixed, we can hold $t \in \mathbb{R}$ fixed and consider the set

$$\mathcal{D}_t(\xi) = \{p \in M \mid t \in I(p)\},$$

i.e., the set of points such that it is possible to "travel for t units of time from p" along the maximal integral curve, γ_p, with initial condition p (It is possible that $\mathcal{D}_t(\xi) = \emptyset$). By definition, if $\mathcal{D}_t(\xi) \neq \emptyset$, the point $\gamma_p(t)$ is well defined, and so, we obtain a map, $\Phi^\xi_t \colon \mathcal{D}_t(\xi) \to M$, with domain $\mathcal{D}_t(\xi)$, given by

$$\Phi^\xi_t(p) = \gamma_p(t).$$

The above suggests the following definition:

Definition 3.21 Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$. For any $t \in \mathbb{R}$, let

$$\mathcal{D}_t(\xi) = \{p \in M \mid t \in I(p)\} \quad \text{and} \quad \mathcal{D}(\xi) = \{(t, p) \in \mathbb{R} \times M \mid t \in I(p)\}$$

and let $\Phi^\xi \colon \mathcal{D}(\xi) \to M$ be the map given by

$$\Phi^\xi(t, p) = \gamma_p(t).$$

The map Φ^ξ is called the *(global) flow of* ξ and $\mathcal{D}(\xi)$ is called its *domain of definition*. For any $t \in \mathbb{R}$ such that $\mathcal{D}_t(\xi) \neq \emptyset$, the map, $p \in \mathcal{D}_t(\xi) \mapsto \Phi^\xi(t, p) = \gamma_p(t)$, is denoted by Φ^ξ_t (i.e., $\Phi^\xi_t(p) = \Phi^\xi(t, p) = \gamma_p(t)$).

Observe that

$$\mathcal{D}(\xi) = \bigcup_{p \in M} (I(p) \times \{p\}).$$

Also, using the Φ^ξ_t notation, the property of Proposition 3.18 reads

$$\Phi^\xi_s \circ \Phi^\xi_t = \Phi^\xi_{s+t}, \qquad (*)$$

whenever both sides of the equation make sense. Indeed, the above says

$$\Phi^\xi_s(\Phi^\xi_t(p)) = \Phi^\xi_s(\gamma_p(t)) = \gamma_{\gamma_p(t)}(s) = \gamma_p(s + t) = \Phi^\xi_{s+t}(p).$$

Using the above property, we can easily show that the Φ^ξ_t are invertible. In fact, the inverse of Φ^ξ_t is Φ^ξ_{-t}. First, note that

$$\mathcal{D}_0(\xi) = M \quad \text{and} \quad \Phi^\xi_0 = \mathrm{id},$$

because, by definition, $\Phi^\xi_0(p) = \gamma_p(0) = p$, for every $p \in M$. Then, $(*)$ implies that

$$\Phi^\xi_t \circ \Phi^\xi_{-t} = \Phi^\xi_{t+-t} = \Phi^\xi_0 = \mathrm{id},$$

which shows that $\Phi^\xi_t \colon \mathcal{D}_t(\xi) \to \mathcal{D}_{-t}(\xi)$ and $\Phi^\xi_{-t} \colon \mathcal{D}_{-t}(\xi) \to \mathcal{D}_t(\xi)$ are inverse of each other. Moreover, each Φ^ξ_t is a C^{k-1}-diffeomorphism. We summarize in the following proposition some additional properties of the domains $\mathcal{D}(\xi)$, $\mathcal{D}_t(\xi)$ and the maps Φ^ξ_t (for a proof, see Lang [38] or Warner [59]):

Theorem 3.19 *Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$. The following properties hold:*

(a) For every $t \in \mathbb{R}$, if $\mathcal{D}_t(\xi) \neq \emptyset$, then $\mathcal{D}_t(\xi)$ is open (this is trivially true if $\mathcal{D}_t(\xi) = \emptyset$).

(b) The domain, $\mathcal{D}(\xi)$, of the flow, Φ^ξ, is open and the flow is a C^{k-1} map, $\Phi^\xi\colon \mathcal{D}(\xi) \to M$.

(c) Each $\Phi^\xi_t\colon \mathcal{D}_t(\xi) \to \mathcal{D}_{-t}(\xi)$ is a C^{k-1}-diffeomorphism with inverse Φ^ξ_{-t}.

(d) For all $s, t \in \mathbb{R}$, the domain of definition of $\Phi^\xi_s \circ \Phi^\xi_t$ is contained but generally not equal to $\mathcal{D}_{s+t}(\xi)$. However, $\mathrm{dom}(\Phi^\xi_s \circ \Phi^\xi_t) = \mathcal{D}_{s+t}(\xi)$ if s and t have the same sign. Moreover, on $\mathrm{dom}(\Phi^\xi_s \circ \Phi^\xi_t)$, we have

$$\Phi^\xi_s \circ \Phi^\xi_t = \Phi^\xi_{s+t}.$$

The reason for using the terminology flow in referring to the map Φ^ξ can be clarified as follows: For any t such that $\mathcal{D}_t(\xi) \neq \emptyset$, every integral curve, γ_p, with initial condition $p \in \mathcal{D}_t(\xi)$, is defined on some open interval containing $[0, t]$, and we can picture these curves as "flow lines" along which the points p flow (travel) for a time interval t. Then, $\Phi^\xi(t, p)$ is the point reached by "flowing" for the amount of time t on the integral curve γ_p (through p) starting from p. Intuitively, we can imagine the flow of a fluid through M, and the vector field ξ is the field of velocities of the flowing particles.

Given a vector field, ξ, as above, it may happen that $\mathcal{D}_t(\xi) = M$, for all $t \in \mathbb{R}$. In this case, namely, when $\mathcal{D}(\xi) = \mathbb{R} \times M$, we say that the vector field ξ is *complete*. Then, the Φ^ξ_t are diffeomorphisms of M and they form a group. The family $\{\Phi^\xi_t\}_{t\in\mathbb{R}}$ a called a *1-parameter group of* ξ. In this case, Φ^ξ induces a group homomorphism, $(\mathbb{R}, +) \longrightarrow \mathrm{Diff}(M)$, from the additive group $\mathbb{R}$ to the group of C^{k-1}-diffeomorphisms of M.

By abuse of language, even when it is **not** the case that $\mathcal{D}_t(\xi) = M$ for all t, the family $\{\Phi^\xi_t\}_{t\in\mathbb{R}}$ is called a *local 1-parameter group of* ξ, even though it is **not** a group, because the composition $\Phi^\xi_s \circ \Phi^\xi_t$ may not be defined.

When M is compact, it turns out that every vector field is complete, a nice and useful fact.

Proposition 3.20 *Let ξ be a C^{k-1} vector field on a C^k-manifold, M, $(k \geq 2)$. If M is compact, then ξ is complete, i.e., $\mathcal{D}(\xi) = \mathbb{R} \times M$. Moreover, the map $t \mapsto \Phi^\xi_t$ is a homomorphism from the additive group $\mathbb{R}$ to the group, $\mathrm{Diff}(M)$, of (C^{k-1}) diffeomorphisms of M.*

Proof. Pick any $p \in M$. By Theorem 3.16, there is a local flow, $\varphi_p\colon J(p) \times U(p) \to M$, where $J(p) \subseteq \mathbb{R}$ is an open interval containing 0 and $U(p)$ is an open subset of M containing p, so that for all $q \in U(p)$, the map $t \mapsto \varphi(t, q)$ is an integral curve with initial condition q (where $t \in J(p)$). Thus, we have $J(p) \times U(p) \subseteq \mathcal{D}(\xi)$. Now, the $U(p)$'s form an open cover of M and since M is compact, we can extract a finite subcover, $\bigcup_{q\in F} U(q) = M$, for some finite subset, $F \subseteq M$. But then, we can find $\epsilon > 0$ so that $]-\epsilon, +\epsilon[\, \subseteq J(q)$, for all $q \in F$ and for all $t \in\,]-\epsilon, +\epsilon[$ and, for all $p \in M$, if γ_p is the maximal integral curve with initial condition p, then $]-\epsilon, +\epsilon[\, \subseteq I(p)$.

For any $t \in\,]-\epsilon, +\epsilon[$, consider the integral curve, $\gamma_{\gamma_p(t)}$, with initial condition $\gamma_p(t)$. This curve is well defined for all $t \in\,]-\epsilon, +\epsilon[$, and we have

$$\gamma_{\gamma_p(t)}(t) = \gamma_p(t+t) = \gamma_p(2t),$$

which shows that γ_p is in fact defined for all $t \in\,]-2\epsilon, +2\epsilon[$. By induction, we see that

$$]-2^n\epsilon, +2^n\epsilon[\, \subseteq I(p),$$

for all $n \geq 0$, which proves that $I(p) = \mathbb{R}$. As this holds for all $p \in M$, we conclude that $\mathcal{D}(\xi) = \mathbb{R} \times M$. $\square$

Remark: The proof of Proposition 3.20 also applies when ξ is a vector field with compact support (this means that the closure of the set $\{p \in M \mid \xi(p) \neq 0\}$ is compact).

A point $p \in M$ where a vector field vanishes, i.e., $\xi(p) = 0$, is called a *critical point of* ξ. Critical points play a major role in the study of vector fields, in differential topology (e.g., the celebrated Poincaré–Hopf index theorem) and especially in Morse theory, but we won't go into this here (curious readers should consult Milnor [42], Guillemin and Pollack [31] or DoCarmo [21], which contains an informal but very clear presentation of the Poincaré–Hopf index theorem). Another famous theorem about vector fields says that every smooth vector field on a sphere of even dimension (S^{2n}) must vanish in at least one point (the so-called "hairy-ball theorem". On S^2, it says that you can't comb your hair without having a singularity somewhere. Try it, it's true!).

Let us just observe that if an integral curve, γ, passes through a critical point, p, then γ is reduced to the point p, i.e., $\gamma(t) = p$, for all t. Indeed, such a curve is an integral curve with initial condition p. By the uniqueness property, it is the only one. Then, we see that if a maximal integral curve is defined on the whole of $\mathbb{R}$, either it is injective (it has no self-intersection), or it is simply periodic (i.e., there is some $T > 0$ so that $\gamma(t+T) = \gamma(t)$, for all $t \in \mathbb{R}$ and γ is injective on $[0, T[$), or it is reduced to a single point.

We conclude this section with the definition of the Lie derivative of a vector field with respect to another vector field.

Say we have two vector fields ξ and η on M. For any $p \in M$, we can flow along the integral curve of ξ with initial condition p to $\Phi_t^\xi(p)$ (for t small enough) and then evaluate η there, getting $\eta(\Phi_t^\xi(p))$. Now, this vector belongs to the tangent space $T_{\Phi_t^\xi(p)}(M)$, but $\eta(p) \in T_p(M)$. So to "compare" $\eta(\Phi_t^\xi(p))$ and $\eta(p)$, we bring back $\eta(\Phi_t^\xi(p))$ to $T_p(M)$ by applying the tangent map, $d\Phi_{-t}^\xi$, at $\Phi_t^\xi(p)$, to $\eta(\Phi_t^\xi(p))$ (Note that to alleviate the notation, we use the slight abuse of notation $d\Phi_{-t}^\xi$ instead of $d(\Phi_{-t}^\xi)_{\Phi_t^\xi(p)}$.) Then, we can form the difference $d\Phi^{\xi}_{-t}(\eta(\Phi^{\xi}_t(p))) - \eta(p)$, divide by t and consider the limit as t goes to 0. This is the *Lie derivative of η with respect to ξ at p*, denoted $(L_\xi\, \eta)_p$, and given by

$$(L_\xi\, \eta)_p = \lim_{t \longrightarrow 0} \frac{d\Phi^{\xi}_{-t}(\eta(\Phi^{\xi}_t(p))) - \eta(p)}{t} = \frac{d}{dt}\left(d\Phi^{\xi}_{-t}(\eta(\Phi^{\xi}_t(p)))\right)\bigg|_{t=0}.$$

It can be shown that $(L_\xi\, \eta)_p$ is our old friend, the Lie bracket, i.e.,

$$(L_\xi\, \eta)_p = [\xi, \eta]_p.$$

(For a proof, see Warner [59] or O'Neill [49]).

3.6 Partitions of Unity

To study manifolds, it is often necessary to construct various objects such as functions, vector fields, Riemannian metrics, volume forms, etc., by glueing together items constructed on the domains of charts. Partitions of unity are a crucial technical tool in this glueing process.

The first step is to define "bump functions" (also called plateau functions). For any $r > 0$, we denote by $B(r)$ the open ball

$$B(r) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 < r\},$$

and by $\overline{B(r)} = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \cdots + x_n^2 \leq r\}$, its closure.

Proposition 3.21 *There is a smooth function, $b\colon \mathbb{R}^n \to \mathbb{R}$, so that*

$$b(x) = \begin{cases} 1 & \text{if } x \in \overline{B(1)} \\ 0 & \text{if } x \in \mathbb{R}^n - B(2). \end{cases}$$

Proof. There are many ways to construct such a function. We can proceed as follows: Consider the function, $h\colon \mathbb{R} \to \mathbb{R}$, given by

$$h(x) = \begin{cases} e^{-1/x} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

It is easy to show that h is C^∞ (but **not** analytic!). Then, define $b\colon \mathbb{R}^n \to \mathbb{R}$, by

$$b(x_1, \ldots, x_n) = \frac{h(4 - x_1^2 - \cdots - x_n^2)}{h(4 - x_1^2 - \cdots - x_n^2) + h(x_1^2 + \cdots + x_n^2 - 1)}.$$

It is immediately verified that b satisfies the required conditions. $\square$

Proposition 3.21 yields the following useful technical result:

Proposition 3.22 *Let M be a smooth manifold. For any open subset, $U \subseteq M$, any $p \in U$ and any smooth function, $f\colon U \to \mathbb{R}$, there exist an open subset, V, with $p \in V$ and $\overline{V} \subseteq U$ and a smooth function, $\widetilde{f}\colon M \to \mathbb{R}$, defined on the whole of M, so that*

$$\widetilde{f}(q) = \begin{cases} f(q) & \text{if } q \in V \\ 0 & \text{if } q \in M - U. \end{cases}$$

Proof. Using a scaling function, it is easy to find a chart, (W, φ) at p, so that $W \subseteq U$, $B(3) \subseteq \varphi(W)$ and $\varphi(p) = 0$. Let $\widetilde{b} = b \circ \varphi$, where b is the function given by Proposition 3.21. Then, $\widetilde{b}$ is a smooth function on W and it is 0 outside of $\varphi^{-1}(B(2)) \subseteq W$. We can extend $\widetilde{b}$ outside W, by setting it to be 0 and we get a smooth function on the whole M. If we let $V = \varphi^{-1}(B(1))$, then V is an open subset around p and clearly, $\overline{V} \subseteq U$ and $\widetilde{b} = 1$ on V. Therefore, if we set

$$\widetilde{f}(q) = \begin{cases} \widetilde{b}(q)f(q) & \text{if } q \in W \\ 0 & \text{if } q \in M - W, \end{cases}$$

we see that $\widetilde{f}$ satisfies the required properties. □

If X is a (Hausdorff) topological space, a family, $\{U_\alpha\}_{\alpha \in I}$, of subsets U_α of X is a *cover* (or *covering*) of X iff $X = \bigcup_{\alpha \in I} U_\alpha$. A cover, $\{U_\alpha\}_{\alpha \in I}$, such that each U_α is open is an *open cover*. If $\{U_\alpha\}_{\alpha \in I}$ is a cover of X, for any subset, $J \subseteq I$, the subfamily $\{U_\alpha\}_{\alpha \in J}$ is a *subcover* of $\{U_\alpha\}_{\alpha \in I}$ if $X = \bigcup_{\alpha \in J} U_\alpha$, i.e., $\{U_\alpha\}_{\alpha \in J}$ is still a cover of X. Given two covers, $\{U_\alpha\}_{\alpha \in I}$ and $\{V_\beta\}_{\beta \in J}$, we say that $\{U_\alpha\}_{\alpha \in I}$ is a *refinement* of $\{V_\beta\}_{\beta \in J}$ iff there is a function, $h\colon I \to J$, so that $U_\alpha \subseteq V_{h(\alpha)}$, for all $\alpha \in I$.

A cover, $\{U_\alpha\}_{\alpha \in I}$, is *locally finite* iff for every point, $p \in X$, there is some open subset, U, with $p \in U$, so that $U \cap U_\alpha \neq \emptyset$ for only finitely many $\alpha \in I$. A space, X, is *paracompact* iff every open cover as a locally finite refinement.

Remark: Recall that a space, X, is *compact* iff it is Hausdorff and if every open cover has a *finite* subcover. Thus, the notion of paracompactess (due to Jean Dieudonné) is a generalization of the notion of compactness.

Recall that a topological space, X, is *second-countable* if it has a countable basis, i.e., if there is a countable family of open subsets, $\{U_i\}_{i \geq 1}$, so that every open subset of X is the union of some of the U_i's. A topological space, X, if *locally compact* iff it is Hausdorff and for every $a \in X$, there is some compact subset, K, and some open subset, U, with $a \in U$ and $U \subseteq K$. As we will see shortly, every locally compact and second-countable topological space is paracompact.

It is important to observe that every manifold (even not second-countable) is locally compact. Indeed, for every $p \in M$, if we pick a chart, (U, φ), around p, then $\varphi(U) = \Omega$ for some open $\Omega \subseteq \mathbb{R}^n$ ($n = \dim M$). So, we can pick a small closed ball, $\overline{B(q, \epsilon)} \subseteq \Omega$, of center $q = \varphi(p)$ and radius ϵ, and as φ is a homeomorphism, we see that

$$p \in \varphi^{-1}(B(q, \epsilon/2)) \subseteq \varphi^{-1}(\overline{B(q, \epsilon)}),$$

where $\varphi^{-1}(\overline{B(q,\epsilon)})$ is compact and $\varphi^{-1}(B(q,\epsilon/2))$ is open.

Finally, we define partitions of unity. Given a topological space, X, for any function, $f\colon X \to \mathbb{R}$, the *support of* f, denoted $\mathrm{supp}\, f$, is the closed set

$$\mathrm{supp}\, f = \overline{\{x \in X \mid f(x) \neq 0\}}.$$

Definition 3.22 Let M be a (smooth) manifold. A *partition of unity on* M is a family, $\{f_i\}_{i\in I}$, of smooth functions on M (the index set I may be uncountable) such that

(a) The family of supports, $\{\mathrm{supp}\, f_i\}_{i\in I}$, is locally finite.

(b) For all $i \in I$ and all $p \in M$, we have $0 \leq f_i(p) \leq 1$, and

$$\sum_{i\in I} f_i(p) = 1, \quad \text{for every } p \in M.$$

If $\{U_\alpha\}_{\alpha\in J}$ is a cover of M, we say that the partition of unity $\{f_i\}_{i\in I}$ is *subordinate* to the cover $\{U_\alpha\}_{\alpha\in J}$ if $\{\mathrm{supp}\, f_i\}_{i\in I}$ is a refinement of $\{U_\alpha\}_{\alpha\in J}$. When $I = J$ and $\mathrm{supp}\, f_i \subseteq U_i$, we say that $\{f_i\}_{i\in I}$ is *subordinate* to $\{U_\alpha\}_{\alpha\in I}$ *with the same index set as the partition of unity*.

In Definition 3.22, by (a), for every $p \in M$, there is some open set, U, with $p \in U$ and U meets only finitely many of the supports, $\mathrm{supp}\, f_i$. So, $f_i(p) \neq 0$ for only finitely many $i \in I$ and the infinite sum $\sum_{i\in I} f_i(p)$ is well defined.

Proposition 3.23 *Let X be a topological space which is second-countable and locally compact (thus, also Hausdorff). Then, X is paracompact. Moreover, every open cover has a countable, locally finite refinement consisting of open sets with compact closures.*

Proof. The proof is quite technical, but since this is an important result, we reproduce Warner's proof for the reader's convenience (Warner [59], Lemma 1.9).

The first step is to construct a sequence of open sets, G_i, such that

1. $\overline{G}_i$ is compact,
2. $\overline{G}_i \subseteq G_{i+1}$,
3. $X = \bigcup_{i=1}^{\infty} G_i$.

As M is second-countable, there is a countable basis of open sets, $\{U_i\}_{i\geq 1}$, for M. Since M is locally compact, we can find a subfamily of $\{U_i\}_{i\geq 1}$ consisting of open sets with compact closures such that this subfamily is also a basis of M. Therefore, we may assume that we start with a countable basis, $\{U_i\}_{i\geq 1}$, of open sets with compact closures. Set $G_1 = U_1$ and assume inductively that

$$G_k = U_1 \cup \cdots \cup U_{j_k}.$$

Since $\overline{G}_k$ is compact, it is covered by finitely many of the U_j's. So, let j_{k+1} be the smallest integer greater than j_k so that

$$\overline{G}_k = U_1 \cup \cdots \cup U_{j_{k+1}}$$

and set

$$G_{k+1} = U_1 \cup \cdots \cup U_{j_{k+1}}.$$

Obviously, the family $\{G_i\}_{i \geq 1}$ satisfies (1)–(3).

Now, let $\{U_\alpha\}_{\alpha \in I}$ be an arbitrary open cover of M. For any $i \geq 3$, the set $\overline{G}_i - G_{i-1}$ is compact and contained in the open $G_{i+1} - \overline{G}_{i-2}$. For each $i \geq 3$, choose a finite subcover of the open cover $\{U_\alpha \cap (G_{i+1} - \overline{G}_{i-2})\}_{\alpha \in I}$ of $\overline{G}_i - G_{i-1}$, and choose a finite subcover of the open cover $\{U_\alpha \cap G_3\}_{\alpha \in I}$ of the compact set $\overline{G}_2$. We leave it to the reader to check that this family of open sets is indeed a countable, locally finite refinement of the original open cover $\{U_\alpha\}_{\alpha \in I}$ and consists of open sets with compact closures. □

Remarks:

1. Proposition 3.23 implies that a second-countable, locally compact (Hausdorff) topological space is the union of countably many compact subsets. Thus, X is *countable at infinity*, a notion that we already encountered in Proposition 2.23 and Theorem 2.26. The reason for this odd terminology is that in the Alexandroff one-point compactification of X, the family of open subsets containing the point at infinity (ω) has a countable basis of open sets. (The open subsets containing ω are of the form $(M - K) \cup \{\omega\}$, where K is compact.)

2. A manifold that is countable at infinity has a countable open cover by domains of charts. This is because, if $M = \bigcup_{i \geq 1} K_i$, where the $K_i \subseteq M$ are compact, then for any open cover of M by domains of charts, for every K_i, we can extract a finite subcover, and the union of these finite subcovers is a countable open cover of M by domains of charts. But then, since for every chart, (U_i, φ_i), the map φ_i is a homeomorphism onto some open subset of $\mathbb{R}^n$, which is second-countable, so we deduce easily that M is second-countable. Thus, for manifolds, second-countable is equivalent to countable at infinity.

We can now prove the main theorem stating the existence of partitions of unity. Recall that we are assuming that our manifolds are Hausdorff and second-countable.

Theorem 3.24 *Let M be a smooth manifold and let $\{U_\alpha\}_{\alpha \in I}$ be an open cover for M. Then, there is a countable partition of unity, $\{f_i\}_{i \geq 1}$, subordinate to the cover $\{U_\alpha\}_{\alpha \in I}$ and the support,* $\mathrm{supp}\, f_i$, *of each f_i is compact. If one does not require compact supports, then there is a partition of unity, $\{f_\alpha\}_{\alpha \in I}$, subordinate to the cover $\{U_\alpha\}_{\alpha \in I}$ with at most countably many of the f_α not identically zero. (In the second case,* $\mathrm{supp}\, f_\alpha \subseteq U_\alpha$.*)*

Proof. Again, we reproduce Warner's proof (Warner [59], Theorem 1.11). As our manifolds are second-countable, Hausdorff and locally compact, from the proof of Proposition 3.23, we have the sequence of open subsets, $\{G_i\}_{i\geq 1}$ and we set $G_0 = \emptyset$. For any $p \in M$, let i_p be the largest integer such that $p \in M - \overline{G}_{i_p}$. Choose an α_p such that $p \in U_{\alpha_p}$; we can find a chart, (U, φ), centered at p such that $U \subseteq U_{\alpha_p} \cap (G_{i_p+2} - \overline{G}_{i_p})$ and such that $\overline{B(2)} \subseteq \varphi(U)$. Define

$$\psi_p = \begin{cases} b \circ \varphi & \text{on } U \\ 0 & \text{on } M - U, \end{cases}$$

where b is the bump function defined just before Proposition 3.21. Then, ψ_p is a smooth function on M which has value 1 on some open subset W_p containing p and has compact support lying in $U \subseteq U_{\alpha_p} \cap (G_{i_p+2} - \overline{G}_{i_p})$. For each $i \geq 1$, choose a finite set of points $p \in M$ whose corresponding opens W_p cover $\overline{G}_i - G_{i-1}$. Order the corresponding ψ_p functions in a sequence ψ_j, $j = 1, 2, \ldots$. The supports of the ψ_j form a locally finite family of subsets of M. Thus, the function

$$\psi = \sum_{j=1}^{\infty} \psi_j$$

is well-defined on M and smooth. Moreover, $\psi(p) > 0$ for each $p \in M$. For each $i \geq 1$, set

$$f_i = \frac{\psi_i}{\psi}.$$

Then, the family, $\{f_i\}_{i\geq 1}$, is a partition of unity subordinate to the cover $\{U_\alpha\}_{\alpha\in I}$ and $\operatorname{supp} f_i$ is compact for all $i \geq 1$.

Now, when we don't require compact support, if we let f_α be identically zero if no f_i has support in U_α and otherwise let f_α be the sum of the f_i with support in U_α, then we obtain a partition of unity subordinate to $\{U_\alpha\}_{\alpha\in I}$ with at most countably many of the f_α not identically zero. We must have $\operatorname{supp} f_\alpha \subseteq U_\alpha$ because for any locally finite family of closed sets, $\{F_\beta\}_{\beta\in J}$, we have $\overline{\bigcup_{\beta\in J} F_\beta} = \bigcup_{\beta\in J} F_\beta$. $\square$

We close this section by stating a famous theorem of Whitney whose proof uses partitions of unity.

Theorem 3.25 *(Whitney, 1935) Any smooth manifold (Hausdorff and second-countable), M, of dimension n is diffeomorphic to a closed submanifold of* $\mathbb{R}^{2n+1}$.

For a proof, see Hirsch [33], Chapter 2, Section 2, Theorem 2.14.

3.7 Manifolds With Boundary

Up to now, we have defined manifolds locally diffeomorphic to an open subset of $\mathbb{R}^m$. This excludes many natural spaces such as a closed disk, whose boundary is a circle, a closed ball,

$\overline{B(1)}$, whose boundary is the sphere, S^{m-1}, a compact cylinder, $S^1 \times [0,1]$, whose boundary consist of two circles, a Möbius strip, etc. These spaces fail to be manifolds because they have a boundary, that is, neighborhoods of points on their boundaries are not diffeomorphic to open sets in $\mathbb{R}^m$. Perhaps the simplest example is the (closed) upper half space,

$$\mathbb{H}^m = \{(x_1, \ldots, x_m) \in \mathbb{R}^m \mid x_m \geq 0\}.$$

Under the natural emdedding $\mathbb{R}^{m-1} \cong \mathbb{R}^{m-1} \times \{0\} \hookrightarrow \mathbb{R}^m$, the subset $\partial\mathbb{H}^m$ of $\mathbb{H}^m$ defined by

$$\partial\mathbb{H}^m = \{x \in \mathbb{H}^m \mid x_m = 0\}$$

is isomorphic to $\mathbb{R}^{m-1}$ and is called the *boundary of* $\mathbb{H}^m$. We also define the *interior* of $\mathbb{H}^m$ as

$$\mathrm{Int}(\mathbb{H}^m) = \mathbb{H}^m - \partial\mathbb{H}^m.$$

Now, if U and V are open subsets of $\mathbb{H}^m$, where $\mathbb{H}^m \subseteq \mathbb{R}^m$ has the subset topology, and if $f\colon U \to V$ is a continuous function, we need to explain what we mean by f being smooth. We say that $f\colon U \to V$, as above, is *smooth* if it has an extension, $\widetilde{f}\colon \widetilde{U} \to \widetilde{V}$, where $\widetilde{U}$ and $\widetilde{V}$ are open subsets of $\mathbb{R}^m$ with $U \subseteq \widetilde{U}$ and $V \subseteq \widetilde{V}$ and with $\widetilde{f}$ a smooth function. We say that f is a (smooth) *diffeomorphism* iff f^{-1} exists and if both f and f^{-1} are smooth, as just defined.

To define a *manifold with boundary*, we replace everywhere $\mathbb{R}$ by $\mathbb{H}$ in Definition 3.1 and Definition 3.2. So, for instance, given a topological space, M, a *chart* is now pair, (U, φ), where U is an open subset of M and $\varphi\colon U \to \Omega$ is a homeomorphism onto an open subset, $\Omega = \varphi(U)$, of $\mathbb{H}^{n_\varphi}$ (for some $n_\varphi \geq 1$), etc. Thus, we obtain

Definition 3.23 Given any two integers, $n \geq 1$ and $k \geq 1$, a *C^k-manifold of dimension n with boundary* consists of a topological space, M, together with an equivalence class, $\overline{\mathcal{A}}$, of C^k n-atlases, on M (where the charts are now defined in terms of open subsets of $\mathbb{H}^n$). Any atlas, $\mathcal{A}$, in the equivalence class $\overline{\mathcal{A}}$ is called a *differentiable structure of class C^k (and dimension n) on M*. We say that M is *modeled on* $\mathbb{H}^n$. When $k = \infty$, we say that M is a *smooth manifold with boundary*.

It remains to define what is the boundary of a manifold with boundary! By definition, the *boundary*, ∂M, of a manifold (with boundary), M, is the set of all points, $p \in M$, such that there is some chart, $(U_\alpha, \varphi_\alpha)$, with $p \in U_\alpha$ and $\varphi_\alpha(p) \in \partial\mathbb{H}^n$. We also let $\mathrm{Int}(M) = M - \partial M$ and call it the *interior* of M.

Do not confuse the boundary ∂M and the interior $\mathrm{Int}(M)$ of a manifold with boundary embedded in $\mathbb{R}^N$ with the topological notions of boundary and interior of M as a topological space. In general, they are different.

Note that manifolds as defined earlier (In Definition 3.3) are also manifolds with boundary: their boundary is just empty. We shall still reserve the word "manifold" for these, but for emphasis, we will sometimes call them "boundaryless".

The definition of tangent spaces, tangent maps, etc., are easily extended to manifolds with boundary. The reader should note that if M is a manifold with boundary of dimension n, the tangent space, T_pM, is defined for all $p \in M$ and has dimension n, *even* for boundary points, $p \in \partial M$. The only notion that requires more care is that of a submanifold. For more on this, see Hirsch [33], Chapter 1, Section 4. One should also beware that the product of two manifolds with boundary is generally **not** a manifold with boundary (consider the product $[0,1] \times [0,1]$ of two line segments). There is a generalization of the notion of a manifold with boundary called *manifold with corners* and such manifolds are closed under products (see Hirsch [33], Chapter 1, Section 4, Exercise 12).

If M is a manifold with boundary, we see that $\mathrm{Int}(M)$ is a manifold without boundary. What about ∂M? Interestingly, the boundary, ∂M, of a manifold with boundary, M, of dimension n, is a manifold of dimension $n-1$. For this, we need the following Proposition:

Proposition 3.26 *If M is a manifold with boundary of dimension n, for any $p \in \partial M$ on the boundary on M, for any chart, (U, φ), with $p \in M$, we have $\varphi(p) \in \partial \mathbb{H}^n$.*

Proof. Since $p \in \partial M$, by definition, there is some chart, (V, ψ), with $p \in V$ and $\psi(p) \in \partial \mathbb{H}^n$. Let (U, φ) be any other chart, with $p \in M$ and assume that $q = \varphi(p) \in \mathrm{Int}(\mathbb{H}^n)$. The transition map, $\psi \circ \varphi^{-1}\colon \varphi(U \cap V) \to \psi(U \cap V)$, is a diffeomorphism and $q = \varphi(p) \in \mathrm{Int}(\mathbb{H}^n)$. By the inverse function theorem, there is some open, $W \subseteq \varphi(U \cap V) \cap \mathrm{Int}(\mathbb{H}^n) \subseteq \mathbb{R}^n$, with $q \in W$, so that $\psi \circ \varphi^{-1}$ maps W homeomorphically onto some subset, Ω, open in $\mathrm{Int}(\mathbb{H}^n)$, with $\psi(p) \in \Omega$, contradicting the hypothesis, $\psi(p) \in \partial \mathbb{H}^n$. $\square$

Using Proposition 3.26, we immediately derive the fact that ∂M is a manifold of dimension $n-1$. We obtain charts on ∂M by considering the charts $(U \cap \partial M, L \circ \varphi)$, where (U, φ) is a chart on M such that $U \cap \partial M = \varphi^{-1}(\partial \mathbb{H}^n) \neq \emptyset$ and $L\colon \partial \mathbb{H}^n \to \mathbb{R}^{n-1}$ is the natural isomorphism.

3.8 Orientation of Manifolds

Although the notion of orientation of a manifold is quite intuitive it is technically rather subtle. We restrict our discussion to smooth manifolds (although the notion of orientation can also be defined for topological manifolds but more work is involved).

Intuitively, a manifold, M, is orientable if it is possible to give a consistent orientation to its tangent space, T_pM, at every point, $p \in M$. So, if we go around a closed curve starting at $p \in M$, when we come back to p, the orientation of T_pM should be the same as when we started. For exampe, if we travel on a Möbius strip (a manifold with boundary) dragging a coin with us, we will come back to our point of departure with the coin flipped. Try it!

To be rigorous, we have to say what it means to orient T_pM (a vector space) and what consistency of orientation means. We begin by quickly reviewing the notion of orientation of a vector space. Let E be a vector space of dimension n. If $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are two bases of E, a basic and crucial fact of linear algebra says that there is a unique linear map, g, mapping each u_i to the corresponding v_i (i.e., $g(u_i) = v_i$, $i = 1, \ldots, n$). Then, look at the determinant, $\det(g)$, of this map. We know that $\det(g) = \det(P)$, where P is the matrix whose j-th columns consist of the coordinates of v_j over the basis $u_1, \ldots, u_n$. Either $\det(g)$ is negative or it is positive. Thus, we define an equivalence relation on bases by saying that two bases have the *same orientation* iff the determinant of the linear map sending the first basis to the second has positive determinant. An *orientation* of E is the choice of one of the two equivalence classes, which amounts to picking some basis as an orientation frame.

The above definition is perfectly fine but it turns out that it is more convenient, in the long term, to use a definition of orientation in terms of alternate multi-linear maps (in particular, to define the notion of integration on a manifold). Recall that a function, $h \colon E^k \to \mathbb{R}$, is *alternate multi-linear* (or *alternate k-linear*) iff it is linear in each of its arguments (holding the others fixed) and if

$$h(\ldots, x, \ldots, x, \ldots) = 0,$$

that is, h vanishes whenever two of its arguments are identical. Using multi-linearity, we immediately deduce that h vanishes for all k-tuples of arguments, $u_1, \ldots, u_k$, that are linearly dependent and that h is *skew-symmetric*, i.e.,

$$h(\ldots, y, \ldots, x, \ldots) = -h(\ldots, x, \ldots, y, \ldots).$$

In particular, for $k = n$, it is easy to see that if $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are two bases, then

$$h(v_1, \ldots, v_n) = \det(g) h(u_1, \ldots, u_n),$$

where g is the unique linear map sending each u_i to v_i. This shows that any alternating n-linear function is a multiple of the determinant function and that the space of alternating n-linear maps is a one-dimensional vector space that we will denote $\bigwedge^n E^*$.[1] We also call an alternating n-linear map an *n-form*. But then, observe that two bases $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ have the same orientation iff

$$\omega(u_1, \ldots, u_n) \quad \text{and} \quad \omega(v_1, \ldots, v_n) \quad \text{have the same sign for all } \omega \in \bigwedge^n E^* - \{0\}$$

(where 0 denote the zero n-form) . As $\bigwedge^n E^*$ is one-dimensional, picking an orientation of E is equivalent to picking a generator (a one-element basis), ω, of $\bigwedge^n E^*$, and to say that $u_1, \ldots, u_n$ has positive orientation iff $\omega(u_1, \ldots, u_n) > 0$.

Given an orientation (say, given by $\omega \in \bigwedge^n E^*$) of E, a linear map, $f \colon E \to E$, is *orientation preserving* iff $\omega(f(u_1), \ldots, f(u_n)) > 0$ (or equivalently, iff $\det(f) > 0$).

[1] We are using the wedge product notation of exterior calculus even though we have not defined alternating tensors and the wedge product. This is standard notation and we hope that the reader will not be confused. In fact, in finite dimension, the space of alternating n-linear maps and $\bigwedge^n E^*$ are isomorphic.

Now, to define the orientation of an n-dimensional manifold, M, we use charts. Given any $p \in M$, for any chart, (U, φ), at p, the tangent map, $d\varphi^{-1}_{\varphi(p)}: \mathbb{R}^n \to T_pM$ makes sense. If $(e_1, \ldots, e_n)$ is the standard basis of $\mathbb{R}^n$, as it gives an orientation to $\mathbb{R}^n$, we can orient T_pM by giving it the orientation induced by the basis $d\varphi^{-1}_{\varphi(p)}(e_1), \ldots, d\varphi^{-1}_{\varphi(p)}(e_n)$. Then, the consistency of orientations of the T_pM's is given by the overlapping of charts. We require that the Jacobian determinants of all $\varphi_i \circ \varphi_j^{-1}$ have the same sign, whenever (U_i, φ_i) and (U_j, φ_j) are any two overlapping charts. Thus, we are led to the definition below. All definitions and results stated in the rest of this section apply to manifolds with or without boundary.

Definition 3.24 Given a smooth manifold, M, of dimension n, an *orientation atlas* of M is any atlas so that the transition maps, $\varphi_i^j = \varphi_j \circ \varphi_i^{-1}$, (from $\varphi_i(U_i \cap U_j)$ to $\varphi_j(U_i \cap U_j)$) all have a positive Jacobian determinant for every point in $\varphi_i(U_i \cap U_j)$. A manifold is *orientable* iff its has some orientation atlas.

Definition 3.24 can be hard to check in practice and there is an equivalent criterion is terms of n-forms which is often more convenient. The idea is that a manifold of dimension n is orientable iff there is a map, $p \mapsto \omega_p$, assigning to every point, $p \in M$, a nonzero n-form, $\omega_p \in \bigwedge^n T_p^*M$, so that this map is smooth. In order to explain rigorously what it means for such a map to be smooth, we can define the *exterior n-bundle*, $\bigwedge^n T^*M$ (also denoted $\bigwedge_n^* M$) in much the same way that we defined the bundles TM and T^*M. There is an obvious smooth projection map $\pi: \bigwedge^n T^*M \to M$. Then, leaving the details of the fact that $\bigwedge^n T^*M$ can be made into a smooth manifold (of dimension n) as an exercise, a smooth map, $p \mapsto \omega_p$, is simply a smooth section of the bundle $\bigwedge^n T^*M$, i.e., a smooth map, $\omega: M \to \bigwedge^n T^*M$, so that $\pi \circ \omega = \mathrm{id}$.

Definition 3.25 If M is an n-dimensional manifold, a smooth section, $\omega \in \Gamma(M, \bigwedge^n T^*M)$, is called a (smooth) *$n$-form*. The set of n-forms, $\Gamma(M, \bigwedge^n T^*M)$, is also denoted $\mathcal{A}^n(M)$. An n-form, ω, is a *nowhere-vanishing n-form on M* or *volume form on M* iff ω_p is a nonzero form for every $p \in M$. This is equivalent to saying that $\omega_p(u_1, \ldots, u_n) \neq 0$, for all $p \in M$ and all bases, $u_1, \ldots, u_n$, of T_pM.

The determinant function, $(u_1, \ldots, u_n) \mapsto \det(u_1, \ldots, u_n)$, where the u_i are expressed over the canonical basis $(e_1, \ldots, e_n)$ of $\mathbb{R}^n$, is a volume form on $\mathbb{R}^n$. We will denote this volume form by ω_0. Another standard notation is $dx_1 \wedge \cdots \wedge dx_n$, but this notation may be very puzzling for readers not familiar with exterior algebra. Observe the justification for the term volume form: the quantity $\det(u_1, \ldots, u_n)$ is indeed the (signed) volume of the parallelepiped

$$\{\lambda_1 u_1 + \cdots + \lambda_n u_n \mid 0 \leq \lambda_i \leq 1,\ 1 \leq i \leq n\}.$$

A volume form on the sphere $S^n \subseteq \mathbb{R}^{n+1}$ is obtained as follows:

$$\omega_p(u_1, \ldots u_n) = \det(p, u_1, \ldots u_n),$$

where $p \in S^n$ and $u_1, \ldots u_n \in T_pM$. As the u_i are orthogonal to p, this is indeed a volume form.

Observe that if f is a smooth function on M and ω is any n-form, then $f\omega$ is also an n-form.

Definition 3.26 Let $\varphi\colon M \to N$ be a smooth map of manifolds of the same dimension, n, and let $\omega \in \mathcal{A}^n(N)$ be an n-form on N. The *pullback, $\varphi^*\omega$, of ω to M* is the n-form on M given by

$$\varphi^*\omega_p(u_1, \ldots, u_n) = \omega_{\varphi(p)}(d\varphi_p(u_1), \ldots, d\varphi_p(u_n)),$$

for all $p \in M$ and all $u_1, \ldots, u_n \in T_pM$.

One checks immediately that $\varphi^*\omega$ is indeed an n-form on M. More interesting is the following Proposition:

Proposition 3.27 *(a) If $\varphi\colon M \to N$ is a local diffeomorphism of manifolds, where $\dim M = \dim N = n$, and $\omega \in \mathcal{A}^n(N)$ is a volume form on N, then $\varphi^*\omega$ is a volune form on M. (b) Assume M has a volume form, ω. Then, for every n-form, $\eta \in \mathcal{A}^n(M)$, there is a unique smooth function, f, never zero on M so that $\eta = f\omega$.*

Proof. (a) By definition,

$$\varphi^*\omega_p(u_1, \ldots, u_n) = \omega_{\varphi(p)}(d\varphi_p(u_1), \ldots, d\varphi_p(u_n)),$$

for all $p \in M$ and all $u_1, \ldots, u_n \in T_pM$. As φ is a local diffeomorphism, $d_p\varphi$ is a bijection for every p. Thus, if $u_1, \ldots, u_n$ is a basis, then so is $d\varphi_p(u_1), \ldots, d\varphi_p(u_n)$, and as ω is nonzero at every point for every basis, $\varphi^*\omega_p(u_1, \ldots, u_n) \neq 0$.

(b) Pick any $p \in M$ and let (U, φ) be any chart at p. As φ is a diffeomorphism, by (a), we see that $\varphi^{-1^*}\omega$ and $\varphi^{-1^*}\eta$ are volume forms on $\varphi(U)$. But then, it is easy to see that $\varphi^{-1^*}\eta = g\varphi^{-1^*}\omega$, for some unique smooth never zero function, g, on $\varphi(U)$ and so, $\eta = f_U\omega$, for some unique smooth never zero function, f_U, on U. For any two overlapping charts, (U_i, φ_i) and (U_j, φ_j), for every $p \in U_i \cap U_j$, for every basis $u_1, \ldots, u_n$ of T_pM, we have

$$\eta_p(u_1, \ldots, u_n) = f_i(p)\omega_p(u_1, \ldots, u_n) = f_j(p)\omega_p(u_1, \ldots, u_n),$$

and as $\eta_p(u_1, \ldots, u_n) \neq 0$, we deduce that f_i and f_j agree on $U_i \cap U_j$. But, then the f_i's patch on the overlaps of the cover, $\{U_i\}$, of M, and so, there is a smooth function, f, defined on the whole of M and such that $f \restriction U_i = f_i$. As the f_i's are unique, so is f. $\square$

Remark: if φ and ψ are smooth maps of manifolds, it is easy to prove that

$$(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$$

and that

$$\varphi^*(f\omega) = (f \circ \varphi)\varphi^*\omega,$$

where f is any smooth function on M and ω is any n-form.

The connection between Definition 3.24 and volume forms is given by the following important theorem whose proof contains a wonderful use of partitions of unity.

Theorem 3.28 *A smooth manifold (Hausdorff and second-countable) is orientable iff it possesses a volume form.*

Proof. First, assume that a volume form, ω, exists on M, and say $n = \dim M$. For any atlas, $(U_i, \varphi_i)_i$, of M, by Proposition 3.27, each n-form, ${\varphi_i^{-1}}^*\omega$, is a volume form on $\varphi_i(U_i) \subseteq \mathbb{R}^n$ and

$${\varphi_i^{-1}}^*\omega = f_i\omega_0,$$

for some smooth function, f_i, never zero on $\varphi_i(U_i)$, where ω_0 is a volume form on $\mathbb{R}^n$. By composing φ_i with an orientation-reversing linear map if necessary, we may assume that for this new altlas, $f_i > 0$ on $\varphi_i(U_i)$. We claim that the family $(U_i, \varphi_i)_i$ is an orientation atlas. This is because, on any (nonempty) overlap, $U_i \cap U_j$, we have

$$(\varphi_j \circ \varphi_i^{-1})^*(f_j\omega_0) = f_i\omega_0,$$

and by the definition of pullbacks, we see that for every $x \in \varphi_i(U_i \cap U_j)$, if we let $y = \varphi_j \circ \varphi_i^{-1}(x)$, then

$$\begin{aligned}(\varphi_j \circ \varphi_i^{-1})^*_x(f_j\omega_0)(e_1, \ldots, e_n) &= f_j(y)\omega_0(d(\varphi_j \circ \varphi_i^{-1})_x(e_1), \ldots, d(\varphi_j \circ \varphi_i^{-1})_x(e_n)) \\ &= f_j(y)J((\varphi_j \circ \varphi_i^{-1})_x)\omega_0,\end{aligned}$$

where $e_1, \ldots, e_n$ is the standard basis of $\mathbb{R}^n$ and $J((\varphi_j \circ \varphi_i^{-1})_x)$ is the Jacobian determinant of $\varphi_j \circ \varphi_i^{-1}$ at x. As both $f_j(y) > 0$ and $f_i(x) > 0$, we have $J((\varphi_j \circ \varphi_i^{-1})_x) > 0$, as desired.

Conversely, assume that $J((\varphi_j \circ \varphi_i^{-1})_x) > 0$, for all $x \in \varphi_i(U_i \cap U_j)$, whenever $U_i \cap U_j \neq \emptyset$. We need to make a volume form on M. For each U_i, let

$$\omega_i = \varphi_i^*\omega_0.$$

As φ_i is a diffeomorphism, by Proposition 3.27, we see that ω_i is a volume form on U_i. Then, if we apply Theorem 3.24, we can find a partition of unity, $\{f_i\}$, subordinate to the cover $\{U_i\}$, with the same index set. Let,

$$\omega = \sum_i f_i\omega_i.$$

We claim that ω is a volume form on M.

It is clear that ω is an n-form on M. Now, since every $p \in M$ belongs to some U_i, check that on $\varphi_i(U_i)$, we have

$$\varphi_i^{-1^*}\omega = \sum_{j \in \text{finite set}} \varphi_i^{-1^*}(f_j\omega_j) = \left(\sum_j (f_j \circ \varphi_i^{-1}) J(\varphi_j \circ \varphi_i^{-1})\right)\omega_0$$

and this sum is strictly positive because the Jacobian determinants are positive and as $\sum_j f_j = 1$ and $f_j \geq 0$, some term must be strictly positive. Therefore, $\varphi_i^{-1^*}\omega$ is a volume form on $\varphi_i(U_i)$ and so, $\varphi_i^* \varphi_i^{-1^*}\omega = \omega$ is a volume form on U_i. As this holds for all U_i, we conclude that ω is a volume form on M. $\square$

Since we showed that there is a volume form on the sphere, S^n, by Theorem 3.28, the sphere S^n is orientable. It can be shown that the projective spaces, $\mathbb{RP}^n$, are non-orientable iff n is even an thus, orientable iff n is odd. In particular, $\mathbb{RP}^2$ is not orientable. Also, even though M may not be orientable, its tangent bundle, $T(M)$, is always orientable! (Prove it). It is also easy to show that if $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ is a smooth submersion, then $M = f^{-1}(0)$ is a smooth orientable manifold. Another nice fact is that every Lie group is orientable.

Chapter 4

Lie Groups, Lie Algebras and the Exponential Map

4.1 Lie Groups and Lie Algebras

In Gallier [27], Chapter 14, we defined the notion of a Lie group as a certain type of manifold embedded in $\mathbb{R}^N$, for some $N \geq 1$. Now that we have the general concept of a manifold, we can define Lie groups in more generality. Besides classic references on Lie groups and Lie Algebras, such as Chevalley [16], Knapp [36], Warner [59], Duistermaat and Kolk [25], Bröcker and tom Dieck [11], Sagle and Walde [53], Fulton and Harris [26] and Bourbaki [8], one should be aware of more introductory sources and surveys such as Hall [32], Sattinger and Weaver [55], Carter, Segal and Macdonald [14], Curtis [18], Baker [3], Bryant [12], Mneimné and Testard [44] and Arvanitoyeogos [1].

Definition 4.1 A *Lie group* is a nonempty subset, G, satisfying the following conditions:

(a) G is a group (with identity element denoted e or 1).

(b) G is a smooth manifold.

(c) G is a topological group. In particular, the group operation, $\cdot : G \times G \to G$, and the inverse map, ${}^{-1}: G \to G$, are smooth.

We have already met a number of Lie groups: $\mathbf{GL}(n, \mathbb{R})$, $\mathbf{GL}(n, \mathbb{C})$, $\mathbf{SL}(n, \mathbb{R})$, $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{O}(n)$, $\mathbf{SO}(n)$, $\mathbf{U}(n)$, $\mathbf{SU}(n)$, $\mathbf{E}(n, \mathbb{R})$. Also, every linear Lie group (i.e., a closed subgroup of $\mathbf{GL}(n, \mathbb{R})$) is a Lie group.

We saw in the case of linear Lie groups that the tangent space to G at the identity, $\mathfrak{g} = T_1 G$, plays a very important role. In particular, this vector space is equipped with a (non-associative) multiplication operation, the Lie bracket, that makes $\mathfrak{g}$ into a Lie algebra. This is again true in this more general setting.

Recall that Lie algebras are defined as follows:

Definition 4.2 A *(real) Lie algebra*, $\mathcal{A}$, is a real vector space together with a bilinear map, $[\cdot,\cdot]\colon \mathcal{A}\times\mathcal{A}\to\mathcal{A}$, called the *Lie bracket* on $\mathcal{A}$ such that the following two identities hold for all $a,b,c\in\mathcal{A}$:

$$[a,\, a] = 0,$$

and the so-called *Jacobi identity*

$$[a,\, [b,\, c]] + [c,\, [a,\, b]] + [b,\, [c,\, a]] = 0.$$

It is immediately verified that $[b,a] = -[a,b]$.

Let us also recall the definition of homomorphisms of Lie groups and Lie algebras.

Definition 4.3 Given two Lie groups G_1 and G_2, a *homomorphism (or map) of Lie groups* is a function, $f\colon G_1\to G_2$, that is a homomorphism of groups and a smooth map (between the manifolds G_1 and G_2). Given two Lie algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, a *homomorphism (or map) of Lie algebras* is a function, $f\colon\mathcal{A}_1\to\mathcal{A}_2$, that is a linear map between the vector spaces $\mathcal{A}_1$ and $\mathcal{A}_2$ and that preserves Lie brackets, i.e.,

$$f([A,B]) = [f(A), f(B)]$$

for all $A,B\in\mathcal{A}_1$.

An *isomorphism of Lie groups* is a bijective function f such that both f and f^{-1} are maps of Lie groups, and an *isomorphism of Lie algebras* is a bijective function f such that both f and f^{-1} are maps of Lie algebras.

The Lie bracket operation on $\mathfrak{g}$ can be defined in terms of the so-called adjoint representation.

Given a Lie group G, for every $a\in G$ we define *left translation* as the map, $L_a\colon G\to G$, such that $L_a(b) = ab$, for all $b\in G$, and *right translation* as the map, $R_a\colon G\to G$, such that $R_a(b) = ba$, for all $b\in G$. Because multiplication and the inverse maps are smooth, the maps L_a and R_a are diffeomorphisms, and their derivatives play an important role. The inner automorphisms $R_{a^{-1}}\circ L_a$ (also written $R_{a^{-1}}L_a$ or $\mathbf{Ad}_a$) also play an important role. Note that

$$R_{a^{-1}}L_a(b) = aba^{-1}.$$

The derivative

$$d(R_{a^{-1}}L_a)_1\colon \mathfrak{g}\to\mathfrak{g}$$

of $R_{a^{-1}}L_a$ at 1 is an isomorphism of Lie algebras, denoted by $\mathrm{Ad}_a\colon\mathfrak{g}\to\mathfrak{g}$. The map $a\mapsto\mathrm{Ad}_a$ is a map of Lie groups

$$\mathrm{Ad}\colon G\to\mathbf{GL}(\mathfrak{g}),$$

called the *adjoint representation of* G (where $\mathbf{GL}(\mathfrak{g})$ denotes the Lie group of all bijective linear maps on $\mathfrak{g}$).

In the case of a linear group, one can verify that

$$\mathrm{Ad}(a)(X) = \mathrm{Ad}_a(X) = aXa^{-1}$$

for all $a \in G$ and all $X \in \mathfrak{g}$. The derivative

$$d\mathrm{Ad}_1 \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

of Ad at 1 is map of Lie algebras, denoted by

$$\mathrm{ad} \colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}),$$

called the *adjoint representation of* $\mathfrak{g}$ (where $\mathfrak{gl}(\mathfrak{g})$ denotes the Lie algebra, $\mathrm{End}(\mathfrak{g}, \mathfrak{g})$, of all linear maps on $\mathfrak{g}$).

In the case of a linear group, it can be verified that

$$\mathrm{ad}(A)(B) = [A,\, B]$$

for all $A, B \in \mathfrak{g}$.

One can also check (in general) that the Jacobi identity on $\mathfrak{g}$ is equivalent to the fact that ad preserves Lie brackets, i.e., ad is a map of Lie algebras:

$$\mathrm{ad}([u,\, v]) = [\mathrm{ad}(u),\, \mathrm{ad}(v)],$$

for all $u, v \in \mathfrak{g}$ (where on the right, the Lie bracket is the commutator of linear maps on $\mathfrak{g}$).

This is the key to the definition of the Lie bracket in the case of a general Lie group (not just a linear Lie group).

Definition 4.4 Given a Lie group, G, the tangent space, $\mathfrak{g} = T_1G$, at the identity with the Lie bracket defined by

$$[u,\, v] = \mathrm{ad}(u)(v), \quad \text{for all } u, v \in \mathfrak{g},$$

is the *Lie algebra of the Lie group G*.

Actually, we have to justify why $\mathfrak{g}$ really is a Lie algebra. For this, we have

Proposition 4.1 *Given a Lie group, G, the Lie bracket, $[u,\, v] = \mathrm{ad}(u)(v)$, of Definition 4.4 satisfies the axioms of a Lie algebra (given in Definition 4.2). Therefore, $\mathfrak{g}$ with this bracket is a Lie algebra.*

Proof. The proof requires Proposition 4.8, but we prefer to defer the proof of this Proposition until section 4.3. Since

$$\mathrm{Ad} \colon G \to \mathbf{GL}(\mathfrak{g})$$

is a Lie group homomorphism, by Proposition 4.8, the map $\mathrm{ad} = d\mathrm{Ad}_1$ is a homomorphism of Lie algebras, $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, which means that

$$\mathrm{ad}([u,v]) = \mathrm{ad}(u) \circ \mathrm{ad}(v) - \mathrm{ad}(v) \circ \mathrm{ad}(u), \quad \text{for all } u, v \in \mathfrak{g},$$

since the bracket in $\mathfrak{gl}(\mathfrak{g}) = \mathrm{End}(\mathfrak{g}, \mathfrak{g})$, is just the commutator. Applying the above to $w \in \mathfrak{g}$, we get the Jacobi identity. We still have to prove that $[u, u] = 0$, or equivalently, that $[v, u] = -[u, v]$. For this, following Duistermaat and Kolk [25] (Chapter 1, Section 1), consider the map

$$G \times G \longrightarrow G\colon (a, b) \mapsto aba^{-1}b^{-1}.$$

It is easy to see that its differential at $(1, 1)$ is the zero map. We can then compute the differential w.r.t. b at $b = 1$ and evaluate at $v \in \mathfrak{g}$, getting $(\mathrm{Ad}_a - \mathrm{id})(v)$. Then, the second derivative w.r.t. a at $a = 1$ evaluated at $u \in \mathfrak{g}$ is $[u, v]$. On the other hand if we differentiate first w.r.t. a and then w.r.t. b, we first get $(\mathrm{id} - \mathrm{Ad}_b)(u)$ and then $-[v, u]$. As our original map is smooth, the second derivative is bilinear symmetric, so $[u, v] = -[v, u]$. □

Remark: After proving that $\mathfrak{g}$ is isomorphic to the vector space of left-invariant vector fields on G, we get another proof of Proposition 4.1.

4.2 Left and Right Invariant Vector Fields, the Exponential Map

A fairly convenient way to define the exponential map is to use left-invariant vector fields.

Definition 4.5 If G is a Lie group, a vector field, ξ, on G is *left-invariant* (resp. *right-invariant*) iff

$$d(L_a)_b(\xi(b)) = \xi(L_a(b)) = \xi(ab), \quad \text{for all } a, b \in G.$$

(resp.

$$d(R_a)_b(\xi(b)) = \xi(R_a(b)) = \xi(ba), \quad \text{for all } a, b \in G.)$$

Equivalently, a vector field, ξ, is left-invariant iff the following diagram commutes (and similarly for a right-invariant vector field):

$$\begin{array}{ccc} T_bG & \xrightarrow{d(L_a)_b} & T_{ab}G \\ \xi \uparrow & & \uparrow \xi \\ G & \xrightarrow[L_a]{} & G \end{array}$$

If ξ is a left-invariant vector field, setting $b = 1$, we see that

$$\xi(a) = d(L_a)_1(\xi(1)),$$

which shows that ξ is determined by its value, $\xi(1) \in \mathfrak{g}$, at the identity (and similarly for right-invariant vector fields).

Conversely, given any $v \in \mathfrak{g}$, we can define the vector field, v^L, by

$$v^L(a) = d(L_a)_1(v), \quad \text{for all } a \in G.$$

We claim that v^L is left-invariant. This follows by an easy application of the chain rule:

$$\begin{aligned} v^L(ab) &= d(L_{ab})_1(v) \\ &= d(L_a \circ L_b)_1(v) \\ &= d(L_a)_b(d(L_b)_1(v)) \\ &= d(L_a)_b(v^L(b)). \end{aligned}$$

Furthermore, $v^L(1) = v$. Therefore, we showed that the map, $\xi \mapsto \xi(1)$, establishes an isomorphism between the space of left-invariant vector fields on G and $\mathfrak{g}$. In fact, the map $G \times \mathfrak{g} \longrightarrow TG$ given by $(a, v) \mapsto v^L(a)$ is an isomorphism between $G \times \mathfrak{g}$ and the tangent bundle, TG.

Remark: Given any $v \in \mathfrak{g}$, we can also define the vector field, v^R, by

$$v^R(a) = d(R_a)_1(v), \quad \text{for all } a \in G.$$

It is easily shown that v^R is right-invariant and we also have an isomorphism $G \times \mathfrak{g} \longrightarrow TG$ given by $(a, v) \mapsto v^R(a)$.

Another reason left-invariant (resp. right-invariant) vector fields on a Lie group are important is that they are complete, i.e., they define a flow whose domain is $\mathbb{R} \times G$. To prove this, we begin with the following easy proposition:

Proposition 4.2 *Given a Lie group, G, if ξ is a left-invariant (resp. right-invariant) vector field and Φ is its flow, then*

$$\Phi(t, g) = g\Phi(t, 1) \quad (\textit{resp.} \quad \Phi(t, g) = \Phi(t, 1)g), \quad \textit{for all } (t, g) \in \mathcal{D}(\xi).$$

Proof. Write

$$\gamma(t) = g\Phi(t, 1) = L_g(\Phi(t, 1)).$$

Then, $\gamma(0) = g$ and, by the chain rule

$$\dot{\gamma}(t) = d(L_g)_{\Phi(t,1)}(\dot{\Phi}(t, 1)) = d(L_g)_{\Phi(t,1)}(\xi(\Phi(t, 1))) = \xi(L_g(\Phi(t, 1))) = \xi(\gamma(t)).$$

By the uniqueness of maximal integral curves, $\gamma(t) = \Phi(t, g)$ for all t, and so,

$$\Phi(t, g) = g\Phi(t, 1).$$

A similar argument applies to right-invariant vector fields. $\square$

Proposition 4.3 *Given a Lie group, G, for every $v \in \mathfrak{g}$, there is a unique smooth homomorphism, $h_v \colon (\mathbb{R}, +) \to G$, such that $\dot{h}_v(0) = v$. Furthermore, $h_v(t)$ is the maximal integral curve of both v^L and v^R with initial condition 1 and the flows of v^L and v^R are defined for all $t \in \mathbb{R}$.*

Proof. Let $\Phi_t^v(g)$ denote the flow of v^L. As far as defined, we know that

$$\Phi_{s+t}^v(1) = \Phi_s^v(\Phi_t^v(1)) = \Phi_t^v(1)\Phi_s^v(1),$$

by Proposition 4.2. Now, if $\Phi_t^v(1)$ is defined on $]-\epsilon, \epsilon[$, setting $s = t$, we see that $\Phi_t^v(1)$ is actually defined on $]-2\epsilon, 2\epsilon[$. By induction, we see that $\Phi_t^v(1)$ is defined on $]-2^n\epsilon, 2^n\epsilon[$, for all $n \geq 0$, and so, $\Phi_t^v(1)$ is defined on $\mathbb{R}$ and the map $t \mapsto \Phi_t^v(1)$ is a homomorphism $h_v \colon (\mathbb{R}, +) \to G$ with $\dot{h}_v(0) = v$. Since $\Phi_t^v(g) = g\Phi_t^v(1)$, the flow, $\Phi_t^v(g)$, is defined for all $(t, g) \in \mathbb{R} \times G$. A similar proof applies to v^R. To show that h_v is smooth, consider the map

$$\mathbb{R} \times G \times \mathfrak{g} \longrightarrow G \times \mathfrak{g}, \quad \text{where} \quad (t, g, v) \mapsto (g\Phi_t^v(1), v).$$

It is immediately seen that the above is the flow of the vector field

$$(g, v) \mapsto (v(g), 0),$$

and thus, it is smooth. Consequently, the restriction of this smooth map to $\mathbb{R} \times \{1\} \times \{v\}$, which is just $t \mapsto \Phi_t^v(1) = h_v(t)$, is also smooth.

Assume $h_v \colon (\mathbb{R}, +) \to G$ is a smooth homomorphism with $\dot{h}(0) = v$. From

$$h(t+s) = h(t)h(s) = h(s)h(t),$$

if we differentiate with respect to s at $s = 0$, we get

$$\frac{dh}{dt}(t) = d(L_{h(t)})_1(v) = v^L(h(t))$$

and

$$\frac{dh}{dt}(t) = d(R_{h(t)})_1(v) = v^R(h(t)).$$

Therefore, $h(t)$ is an integral curve for v^L and v^R with initial condition $h(0) = 1$ and $h = \Phi_t^v(1)$. □

Proposition 4.3 yields the definition of the exponential map.

Definition 4.6 Given a Lie group, G, the *exponential map*, $\exp \colon \mathfrak{g} \to G$, is given by

$$\exp(v) = h_v(1) = \Phi_1^v(1), \quad \text{for all } v \in \mathfrak{g}.$$

We can see that exp is smooth as follows. As in the proof of Proposition 4.3, we have the smooth map

$$\mathbb{R} \times G \times \mathfrak{g} \longrightarrow G \times \mathfrak{g}, \quad \text{where} \quad (t, g, v) \mapsto (g\Phi_t^v(1), v),$$

which is the flow of the vector field

$$(g, v) \mapsto (v(g), 0).$$

Consequently, the restriction of this smooth map to $\{1\} \times \{1\} \times \mathfrak{g}$, which is just $v \mapsto \Phi_1^v(1) = \exp(v)$, is also smooth.

Observe that for any fixed $t \in \mathbb{R}$, the map

$$s \mapsto h_v(st)$$

is a smooth homomorphism, h, such that $\dot{h}(0) = tv$. By uniqueness, we have

$$h_v(st) = h_{tv}(s).$$

Setting $s = 1$, we find that

$$h_v(t) = \exp(tv), \quad \text{for all } v \in \mathfrak{g} \text{ and all } t \in \mathbb{R}.$$

Then, differentiating with respect to t at $t = 0$, we get

$$v = d\exp_0(v),$$

i.e., $d\exp_0 = \mathrm{id}_{\mathfrak{g}}$. By the inverse function theorem, exp is a local diffeomorphism at 0. This means that there is some open subset, $U \subseteq \mathfrak{g}$, containing 0, such that the restriction of exp to U is a diffeomorphism onto $\exp(U) \subseteq G$, with $1 \in \exp(U)$. In fact, by left-translation, the map $v \mapsto g\exp(v)$ is a local diffeomorphism between some open subset, $U \subseteq \mathfrak{g}$, containing 0 and the open subset, $\exp(U)$, containing g. The exponential map is also natural in the following sense:

Proposition 4.4 *Given any two Lie groups, G and H, for every Lie group homomorphism, $f\colon G \to H$, the following diagram commutes:*

$$\begin{array}{ccc} G & \xrightarrow{f} & H \\ {\scriptstyle \exp}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[df_1]{} & \mathfrak{h} \end{array}$$

Proof. Observe that the map $h\colon t \mapsto f(\exp(tv))$ is a homomorphism from $(\mathbb{R}, +)$ to G such that $\dot{h}(0) = df_1(v)$. Proposition 4.3 shows that $f(\exp(v)) = \exp(df_1(v))$. $\square$

As useful corollary of Proposition 4.4 is:

Proposition 4.5 *Let G be a connected Lie group and H be any Lie group. For any two homomorphisms, $\varphi_1: G \to H$ and $\varphi_2: G \to H$, if $d(\varphi_1)_1 = d(\varphi_2)_1$, then $\varphi_1 = \varphi_2$.*

Proof. We know that the exponential map is a diffeomorphism on some small open subset, U, containing 0. Now, by Proposition 4.4, for all $a \in \exp_G(U)$, we have

$$\varphi_i(a) = \exp_H(d(\varphi_i)_1(\exp_G^{-1}(a))), \quad i = 1, 2.$$

Since $d(\varphi_1)_1 = d(\varphi_2)_1$, we conclude that $\varphi_1 = \varphi_2$ on $\exp_G(U)$. However, as G is connected, Proposition 2.18 implies that G is generated by $\exp_G(U)$ (we can easily find a symmetric neighborhood of 1 in $\exp_G(U)$). Therefore, $\varphi_1 = \varphi_2$ on G. □

The above proposition shows that if G is connected, then a homomorphism of Lie groups, $\varphi: G \to H$, is uniquely determined by the Lie algebra homomorphism, $d\varphi_1: \mathfrak{g} \to \mathfrak{h}$.

Since the Lie algebra $\mathfrak{g} = T_1 G$ is isomorphic to the vector space of left-invariant vector fields on G and since the Lie bracket of vector fields makes sense (see Definition 3.15), it is natural to ask if there is any relationship between, $[u, v]$, where $[u, v] = \mathrm{ad}(u)(v)$, and the Lie bracket, $[u^L, v^L]$, of the left-invariant vector fields associated with $u, v \in \mathfrak{g}$. The answer is: Yes, they coincide (*via* the correspondence $u \mapsto u^L$). This fact is recorded in the proposition below whose proof involves some rather acrobatic uses of the chain rule found in Warner [59] (Chapter 3), Bröcker and tom Dieck [11] (Chapter 1, Section 2), or Marsden and Ratiu [40] (Chapter 9).

Proposition 4.6 *Given a Lie group, G, we have*

$$[u^L, v^L](1) = \mathrm{ad}(u)(v), \quad \textit{for all } u, v \in \mathfrak{g}.$$

We can apply Proposition 2.22 and use the exponential map to prove a useful result about Lie groups. If G is a Lie group, let G_0 be the connected component of the identity. We know G_0 is a topological normal subgroup of G and it is a submanifold in an obvious way, so it is a Lie group.

Proposition 4.7 *If G is a Lie group and G_0 is the connected component of 1, then G_0 is generated by $\exp(\mathfrak{g})$. Moreover, G_0 is countable at infinity.*

Proof. We can find a symmetric open, U, in $\mathfrak{g}$ in containing 0, on which exp is a diffeomorphism. Then, apply Proposition 2.22 to $V = \exp(U)$. That G_0 is countable at infinity follows from Proposition 2.23. □

4.3 Homomorphisms of Lie Groups and Lie Algebras, Lie Subgroups

If G and H are two Lie groups and $\varphi: G \to H$ is a homomorphism of Lie groups, then $d\varphi_1: \mathfrak{g} \to \mathfrak{h}$ is a linear map between the Lie algebras $\mathfrak{g}$ and $\mathfrak{h}$ of G and H. In fact, it is a Lie algebra homomorphism, as shown below.

Proposition 4.8 *If G and H are two Lie groups and $\varphi\colon G \to H$ is a homomorphism of Lie groups, then*

$$d\varphi_1 \circ \mathrm{Ad}_g = \mathrm{Ad}_{\varphi(g)} \circ d\varphi_1, \quad \textit{for all } g \in G$$

and $d\varphi_1\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism.

Proof. Recall that

$$R_{a^{-1}}L_a(b) = aba^{-1}, \quad \text{for all } a, b \in G$$

and that the derivative

$$d(R_{a^{-1}}L_a)_1\colon \mathfrak{g} \to \mathfrak{g}$$

of $R_{a^{-1}}L_a$ at 1 is an isomorphism of Lie algebras, denoted by $\mathrm{Ad}_a\colon \mathfrak{g} \to \mathfrak{g}$. The map $a \mapsto \mathrm{Ad}_a$ is a map of Lie groups

$$\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g}),$$

(where $\mathbf{GL}(\mathfrak{g})$ denotes the Lie group of all bijective linear maps on $\mathfrak{g}$) and the derivative

$$d\mathrm{Ad}_1\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

of Ad at 1 is map of Lie algebras, denoted by

$$\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}),$$

called the adjoint representation of $\mathfrak{g}$ (where $\mathfrak{gl}(\mathfrak{g})$ denotes the Lie algebra of all linear maps on $\mathfrak{g}$). Then, the Lie bracket is defined by

$$[u,\, v] = \mathrm{ad}(u)(v), \quad \text{for all } u, v \in \mathfrak{g}.$$

Now, as φ is a homomorphism, we have

$$\varphi(R_{a^{-1}}L_a(b)) = \varphi(aba^{-1}) = \varphi(a)\varphi(b)\varphi(a)^{-1} = R_{\varphi(a)^{-1}}L_{\varphi(a)}(\varphi(b)),$$

and by differentiating w.r.t. b at $b = 1$ and evaluating at $v \in \mathfrak{g}$, we get

$$d\varphi_1(\mathrm{Ad}_a(v)) = \mathrm{Ad}_{\varphi(a)}(d\varphi_1(v)),$$

proving the first part of the proposition. Differentiating again with respect to a at $a = 1$ and evaluating at $u \in \mathfrak{g}$ (and using the chain rule), we get

$$d\varphi_1(\mathrm{ad}(u)(v)) = \mathrm{ad}(d\varphi_1(u))(d\varphi_1(v)),$$

i.e.,

$$d\varphi_1[u, v] = [d\varphi_1(u), d\varphi_1(v)],$$

which proves that $d\varphi_1$ is indeed a Lie algebra homomorphism. $\square$

Remark: If we identify the Lie algebra, $\mathfrak{g}$, of G with the space of left-invariant vector fields on G, the map $d\varphi_1\colon \mathfrak{g} \to \mathfrak{h}$ is viewed as the map such that, for every left-invariant vector field, ξ, on G, the vector field $d\varphi_1(\xi)$ is the unique left-invariant vector field on H such that

$$d\varphi_1(\xi)(1) = d\varphi_1(\xi(1)),$$

i.e., $d\varphi_1(\xi) = d\varphi_1(\xi(1))^L$. Then, we can give another proof of the fact that $d\varphi_1$ is a Lie algebra homomorphism using the notion of φ-related vector fields.

Proposition 4.9 *If G and H are two Lie groups and $\varphi\colon G \to H$ is a homomorphism of Lie groups, if we identify $\mathfrak{g}$ (resp. $\mathfrak{h}$) with the space of left-invariant vector fields on G (resp. left-invariant vector fields on H), then,*

(a) ξ and $d\varphi_1(\xi)$ are φ-related, for every left-invariant vector field, ξ, on G;

(b) $d\varphi_1\colon \mathfrak{g} \to \mathfrak{h}$ is a Lie algebra homomorphism.

Proof. The proof uses Proposition 3.12. For details, see Warner [59].

We now consider Lie subgroups. As a preliminary result, note that if $\varphi\colon G \to H$ is an injective Lie group homomorphism, then $d\varphi_g\colon T_gG \to T_{\varphi(g)}H$ is injective for all $g \in G$. As $\mathfrak{g} = T_1G$ and T_gG are isomorphic for all $g \in G$ (and similarly for $\mathfrak{h} = T_1H$ and T_hH for all $h \in H$), it is sufficient to check that $d\varphi_1\colon \mathfrak{g} \to \mathfrak{h}$ is injective. However, by Proposition 4.4, the diagram

$$\begin{array}{ccc} G & \xrightarrow{\varphi} & H \\ \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow[d\varphi_1]{} & \mathfrak{h} \end{array}$$

commutes, and since the exponential map is a local diffeomorphism at 0, as φ is injective, then $d\varphi_1$ is injective, too. Therefore, if $\varphi\colon G \to H$ is injective, it is automatically an immersion.

Definition 4.7 Let G be a Lie group. A set, H, is an *immersed (Lie) subgroup* of G iff

(a) H is a Lie group;

(b) There is an injective Lie group homomorphism, $\varphi\colon H \to G$ (and thus, φ is an immersion, as noted above).

We say that H is a *Lie subgroup* (or *closed Lie subgroup*) of G iff H is a Lie group that is a subgroup of G and also a submanifold of G.

Observe that an immersed Lie subgroup, H, is an immersed submanifold, since φ is an injective immersion. However, $\varphi(H)$ may *not* have the subspace topology inherited from G and $\varphi(H)$ may not be closed.

An example of this situation is provided by the 2-torus, $T^2 \cong \mathbf{SO}(2) \times \mathbf{SO}(2)$, which can be identified with the group of 2×2 complex diagonal matrices of the form

$$\begin{pmatrix} e^{i\theta_1} & 0 \\ 0 & e^{i\theta_2} \end{pmatrix}$$

where $\theta_1, \theta_2 \in \mathbb{R}$. For any $c \in \mathbb{R}$, let S_c be the subgroup of T^2 consisting of all matrices of the form

$$\begin{pmatrix} e^{it} & 0 \\ 0 & e^{ict} \end{pmatrix}, \quad t \in \mathbb{R}.$$

It is easily checked that S_c is an immersed Lie subgroup of T^2 iff c is irrational. However, when c is irrational, one can show that S_c is dense in T^2 but not closed.

As we will see below, a Lie subgroup, is always closed. We borrowed the terminology "immersed subgroup" from Fulton and Harris [26] (Chapter 7), but we warn the reader that most books call such subgroups "Lie subgroups" and refer to the second kind of subgroups (that are submanifolds) as "closed subgroups".

Theorem 4.10 *Let G be a Lie group and let (H, φ) be an immersed Lie subgroup of G. Then, φ is an embedding iff $\varphi(H)$ is closed in G. As as consequence, any Lie subgroup of G is closed.*

Proof. The proof can be found in Warner [59] (Chapter 1, Theorem 3.21) and uses a little more machinery than we have introduced. However, we prove that a Lie subgroup, H, of G is closed. The key to the argument is this: Since H is a submanifold of G, there is chart, (U, φ), of G, with $1 \in U$, so that

$$\varphi(U \cap H) = \varphi(U) \cap (R^m \times \{0_{n-m}\}).$$

By Proposition 2.15, we can find some open subset, $V \subseteq U$, with $1 \in V$, so that $V = V^{-1}$ and $\overline{V} \subseteq U$. Observe that

$$\varphi(\overline{V} \cap H) = \varphi(\overline{V}) \cap (R^m \times \{0_{n-m}\})$$

and since $\overline{V}$ is closed and φ is a homeomorphism, it follows that $\overline{V} \cap H$ is closed. Thus, $\overline{V} \cap H = \overline{V} \cap \overline{H}$ (as $\overline{\overline{V} \cap H} = \overline{V} \cap \overline{H}$). Now, pick any $y \in \overline{H}$. As $1 \in V^{-1}$, the open set yV^{-1} contains y and since $y \in \overline{H}$, we must have $yV^{-1} \cap H \neq \emptyset$. Let $x \in yV^{-1} \cap H$, then $x \in H$ and $y \in xV$. Then, $y \in xV \cap \overline{H}$, wich implies $x^{-1}y \in V \cap \overline{H} \subseteq \overline{V} \cap \overline{H} = \overline{V} \cap H$. Therefore, $x^{-1}y \in H$ and since $x \in H$, we get $y \in H$ and H is closed. $\square$

We also have the following important and useful theorem: If G is a Lie group, say that a subset, $H \subseteq G$, is an *abstract subgroup* iff it is just a subgroup of the underlying group of G (i.e., we forget the topology and the manifold structure).

Theorem 4.11 *Let G be a Lie group. An abstract subgroup, H, of G is a submanifold (i.e., a Lie subgroup) of G iff H is closed (i.e, H with the induced topology is closed in G).*

Proof. We proved the easy direction of this theorem above. Conversely, we need to prove that if the subgroup, H, with the induced topology is closed in G, then it is a manifold. This can be done using the exponential map, but it is harder. For details, see Bröcker and tom Dieck [11] (Chapter 1, Section 3) or Warner [59], Chapter 3. □

4.4 The Correspondence Lie Groups–Lie Algebras

Historically, Lie was the first to understand that a lot of the structure of a Lie group is captured by its Lie algebra, a simpler object (since it is a vector space). In this short section, we state without proof some of the "Lie theorems", although not in their original form.

Definition 4.8 If $\mathfrak{g}$ is a Lie algebra, a *subalgebra*, $\mathfrak{h}$, of $\mathfrak{g}$ is a (linear) subspace of $\mathfrak{g}$ such that $[u, v] \in \mathfrak{h}$, for all $u, v \in \mathfrak{h}$. If $\mathfrak{h}$ is a (linear) subspace of $\mathfrak{g}$ such that $[u, v] \in \mathfrak{h}$ for all $u \in \mathfrak{h}$ and all $v \in \mathfrak{g}$, we say that $\mathfrak{h}$ is an *ideal* in $\mathfrak{g}$.

For a proof of the theorem below, see Warner [59] (Chapter 3) or Duistermaat and Kolk [25] (Chapter 1, Section 10).

Theorem 4.12 *Let G be a Lie group with Lie algebra, $\mathfrak{g}$, and let (H, φ) be an immersed Lie subgroup of G with Lie algebra $\mathfrak{h}$, then $d\varphi_1\mathfrak{h}$ is a Lie subalgebra of $\mathfrak{g}$. Conversely, for each subalgebra, $\widetilde{\mathfrak{h}}$, of $\mathfrak{g}$, there is a unique connected immersed subgroup, (H, φ), of G so that $d\varphi_1\mathfrak{h} = \widetilde{\mathfrak{h}}$. In fact, as a group, $\varphi(H)$ is the subgroup of G generated by $\exp(\widetilde{\mathfrak{h}})$. Furthermore, normal subgroups correspond to ideals.*

Theorem 4.12 shows that there is a one-to-one correspondence between connected immersed subgroups of a Lie group and subalgebras of its Lie algebra.

Theorem 4.13 *Let G and H be Lie groups with G connected and simply connected and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. For every homomorphism, $\psi\colon \mathfrak{g} \to \mathfrak{h}$, there is a unique Lie group homomorphism, $\varphi\colon G \to H$, so that $d\varphi_1 = \psi$.*

Again a proof of the theorem above is given in Warner [59] (Chapter 3) or Duistermaat and Kolk [25] (Chapter 1, Section 10).

Corollary 4.14 *If G and H are connected and simply connected Lie groups, then G and H are isomorphic iff $\mathfrak{g}$ and $\mathfrak{h}$ are isomorphic.*

It can also be shown that for every finite-dimensional Lie algebra, $\mathfrak{g}$, there is a connected and simply connected Lie group, G, such that $\mathfrak{g}$ is the Lie algebra of G. This is a consequence of deep theorem (whose proof is quite hard) known as *Ado's theorem*. For more on this, see Knapp [36], Fulton and Harris [26], or Bourbaki [8].

In summary, following Fulton and Harris, we have the following two principles of the Lie group/Lie algebra correspondence:

First Principle: If G and H are Lie groups, with G connected, then a homomorphism of Lie groups, $\varphi\colon G \to H$, is uniquely determined by the Lie algebra homomorphism, $d\varphi_1\colon \mathfrak{g} \to \mathfrak{h}$.

Second Principle: Let G and H be Lie groups with G connected and simply connected and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. A linear map, $\psi\colon \mathfrak{g} \to \mathfrak{h}$, is a Lie algebra map iff there is a unique Lie group homomorphism, $\varphi\colon G \to H$, so that $d\varphi_1 = \psi$.

4.5 More on the Lorentz Group $\mathbf{SO}_0(n, 1)$

In this section, we take a closer look at the Lorentz group $\mathbf{SO}_0(n, 1)$ and, in particular, at the relationship between $\mathbf{SO}_0(n, 1)$ and its Lie algebra, $\mathfrak{so}(n, 1)$. The Lie algebra of $\mathbf{SO}_0(n, 1)$ is easily determined by computing the tangent vectors to curves, $t \mapsto A(t)$, on $\mathbf{SO}_0(n, 1)$ through the identity, I. Since $A(t)$ satisfies

$$A^\top J A = J,$$

differentiating and using the fact that $A(0) = I$, we get

$$A'^\top J + J A' = 0.$$

Therefore,

$$\mathfrak{so}(n, 1) = \{A \in \mathrm{Mat}_{n+1,n+1}(\mathbb{R}) \mid A^\top J + J A = 0\}.$$

This means that JA is skew-symmetric and so,

$$\mathfrak{so}(n, 1) = \left\{ \begin{pmatrix} B & u \\ u^\top & 0 \end{pmatrix} \in \mathrm{Mat}_{n+1,n+1}(\mathbb{R}) \mid u \in \mathbb{R}^n, \quad B^\top = -B \right\}.$$

Observe that every matrix $A \in \mathfrak{so}(n, 1)$ can be written uniquely as

$$\begin{pmatrix} B & u \\ u^\top & 0 \end{pmatrix} = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix},$$

where the first matrix is skew-symmetric, the second one is symmetric and both belong to $\mathfrak{so}(n, 1)$. Thus, it is natural to define

$$\mathfrak{k} = \left\{ \begin{pmatrix} B & 0 \\ 0^\top & 0 \end{pmatrix} \mid B^\top = -B \right\}$$

and

$$\mathfrak{p} = \left\{ \begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix} \mid u \in \mathbb{R}^n \right\}.$$

It is immediately verified that both $\mathfrak{k}$ and $\mathfrak{p}$ are subspaces of $\mathfrak{so}(n,1)$ (as vector spaces) and that $\mathfrak{k}$ is a Lie subalgebra isomorphic to $\mathfrak{so}(n)$, but $\mathfrak{p}$ is *not* a Lie subalgebra of $\mathfrak{so}(n,1)$ because it is not closed under the Lie bracket. Still, we have

$$[\mathfrak{k},\mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{k},\mathfrak{p}] \subseteq \mathfrak{p}, \quad [\mathfrak{p},\mathfrak{p}] \subseteq \mathfrak{k}.$$

Clearly, we have the direct sum decomposition

$$\mathfrak{so}(n,1) = \mathfrak{k} \oplus \mathfrak{p},$$

known as *Cartan decomposition*. There is also an automorphism of $\mathfrak{so}(n,1)$ known as the *Cartan involution*, namely,

$$\theta(A) = -A^\top,$$

and we see that

$$\mathfrak{k} = \{A \in \mathfrak{so}(n,1) \mid \theta(A) = A\} \quad \text{and} \quad \mathfrak{p} = \{A \in \mathfrak{so}(n,1) \mid \theta(A) = -A\}.$$

Unfortunately, there does not appear to be any simple way of obtaining a formula for $\exp(A)$, where $A \in \mathfrak{so}(n,1)$ (except for small n—there is such a formula for $n = 3$ due to Chris Geyer). However, it is possible to obtain an explicit formula for the matrices in $\mathfrak{p}$. This is because for such matrices, A, if we let $\omega = \|u\| = \sqrt{u^\top u}$, we have

$$A^3 = \omega^2 A.$$

Thus, we get

Proposition 4.15 *For every matrix, $A \in \mathfrak{p}$, of the form*

$$A = \begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix},$$

we have

$$e^A = \begin{pmatrix} I + \dfrac{(\cosh\omega - 1)}{\omega^2} uu^\top & \dfrac{\sinh\omega}{\omega} u \\ \dfrac{\sinh\omega}{\omega} u^\top & \cosh\omega \end{pmatrix} = \begin{pmatrix} \sqrt{I + \dfrac{\sinh^2\omega}{\omega^2} uu^\top} & \dfrac{\sinh\omega}{\omega} u \\ \dfrac{\sinh\omega}{\omega} u^\top & \cosh\omega \end{pmatrix}.$$

Proof. Using the fact that $A^3 = \omega^2 A$, we easily prove that

$$e^A = I + \frac{\sinh\omega}{\omega} A + \frac{\cosh\omega - 1}{\omega^2} A^2,$$

which is the first equation of the proposition, since

$$A^2 = \begin{pmatrix} uu^\top & 0 \\ 0 & u^\top u \end{pmatrix} = \begin{pmatrix} uu^\top & 0 \\ 0 & \omega^2 \end{pmatrix}.$$

We leave as an exercise the fact that

$$\left(I + \frac{(\cosh\omega - 1)}{\omega^2} uu^\top\right)^2 = I + \frac{\sinh^2\omega}{\omega^2} uu^\top.$$

$\square$

Now, it clear from the above formula that each e^B, with $B \in \mathfrak{p}$ is a Lorentz boost. Conversely, every Lorentz boost is the exponential of some $B \in \mathfrak{p}$, as shown below.

Proposition 4.16 *Every Lorentz boost,*

$$A = \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

with $c = \sqrt{\|v\|^2 + 1}$, *is of the form* $A = e^B$, *for* $B \in \mathfrak{p}$, *i.e., for some* $B \in \mathfrak{so}(n,1)$ *of the form*

$$B = \begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix}.$$

Proof. We need to find some

$$B = \begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix}$$

solving the equation

$$\begin{pmatrix} \sqrt{I + \frac{\sinh^2\omega}{\omega^2} uu^\top} & \frac{\sinh\omega}{\omega} u \\ \frac{\sinh\omega}{\omega} u^\top & \cosh\omega \end{pmatrix} = \begin{pmatrix} \sqrt{I + vv^\top} & v \\ v^\top & c \end{pmatrix},$$

with $\omega = \|u\|$ and $c = \sqrt{\|v\|^2 + 1}$. When $v = 0$, we have $A = I$, and the matrix $B = 0$ corresponding to $u = 0$ works. So, assume $v \neq 0$. In this case, $c > 1$. We have to solve the equation $\cosh\omega = c$, that is,

$$e^{2\omega} - 2ce^\omega + 1 = 0.$$

The roots of the corresponding algebraic equation $X^2 - 2cX + 1 = 0$ are

$$X = c \pm \sqrt{c^2 - 1}.$$

As $c > 1$, both roots are strictly positive, so we can solve for ω, say $\omega = \log(c + \sqrt{c^2 - 1}) \neq 0$. Then, $\sinh\omega \neq 0$, so we can solve the equation

$$\frac{\sinh\omega}{\omega} u = v,$$

which yields a $B \in \mathfrak{so}(n,1)$ of the right form with $A = e^B$. $\square$

Remarks:

(1) It is easy to show that the eigenvalues of matrices

$$B = \begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix}$$

are 0, with multiplicity $n-1$, $\|u\|$ and $-\|u\|$. Eigenvectors are also easily determined.

(2) The matrices $B \in \mathfrak{so}(n,1)$ of the form

$$B = \begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \alpha \\ 0 & \cdots & \alpha & 0 \end{pmatrix}$$

are easily seen to form an abelian Lie subalgebra, $\mathfrak{a}$, of $\mathfrak{so}(n,1)$ (which means that for all $B, C \in \mathfrak{a}$, $[B,C] = 0$, i.e., $BC = CB$). One will easily check that for any $B \in \mathfrak{a}$, as above, we get

$$e^B = \begin{pmatrix} 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & \cosh\alpha & \sinh\alpha \\ 0 & \cdots & 0 & \sinh\alpha & \cosh\alpha \end{pmatrix}$$

The matrices of the form e^B, with $B \in \mathfrak{a}$, form an abelian subgroup, A, of $\mathbf{SO}_0(n,1)$ isomorphic to $\mathbf{SO}_0(1,1)$. As we already know, the matrices $B \in \mathfrak{so}(n,1)$ of the form

$$\begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix},$$

where B is skew-symmetric, form a Lie subalgebra, $\mathfrak{k}$, of $\mathfrak{so}(n,1)$. Clearly, $\mathfrak{k}$ is isomorphic to $\mathfrak{so}(n)$ and using the exponential, we get a subgroup, K, of $\mathbf{SO}_0(n,1)$ isomorphic to $\mathbf{SO}(n)$. It is also clear that $\mathfrak{k} \cap \mathfrak{a} = (0)$, but $\mathfrak{k} \oplus \mathfrak{a}$ is *not* equal to $\mathfrak{so}(n,1)$. What is the missing piece? Consider the matrices $N \in \mathfrak{so}(n,1)$ of the form

$$N = \begin{pmatrix} 0 & -u & u \\ u^\top & 0 & 0 \\ u^\top & 0 & 0 \end{pmatrix},$$

where $u \in \mathbb{R}^{n-1}$. The reader should check that these matrices form an abelian Lie subalgebra, $\mathfrak{n}$, of $\mathfrak{so}(n,1)$ and that

$$\mathfrak{so}(n,1) = \mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}.$$

This is the *Iwasawa decomposition* of the Lie algebra $\mathfrak{so}(n,1)$. Furthermore, the reader should check that every $N \in \mathfrak{n}$ is nilpotent; in fact, $N^3 = 0$. (It turns out that $\mathfrak{n}$ is a nilpotent Lie algebra, see Knapp [36]). The connected Lie subgroup of $\mathbf{SO}_0(n,1)$

associated with $\mathfrak{n}$ is denoted N and it can be shown that we have the *Iwasawa decomposition* of the Lie group $\mathbf{SO}_0(n,1)$:

$$\mathbf{SO}_0(n,1) = KAN.$$

It is easy to check that $[\mathfrak{a},\mathfrak{n}] \subseteq \mathfrak{n}$, so $\mathfrak{a}\oplus\mathfrak{n}$ is a Lie subalgebra of $\mathfrak{so}(n,1)$ and $\mathfrak{n}$ is an ideal of $\mathfrak{a}\oplus\mathfrak{n}$. This implies that N is normal in the group corresponding to $\mathfrak{a}\oplus\mathfrak{n}$, so AN is a subgroup (in fact, solvable) of $\mathbf{SO}_0(n,1)$. For more on the Iwasawa decomposition, see Knapp [36]. Observe that the image, $\overline{\mathfrak{n}}$, of $\mathfrak{n}$ under the Cartan involution, θ, is the Lie subalgebra

$$\overline{\mathfrak{n}} = \left\{ \begin{pmatrix} 0 & u & u \\ -u^\top & 0 & 0 \\ u^\top & 0 & 0 \end{pmatrix} \mid u \in \mathbb{R}^{n-1} \right\}.$$

It is easy to see that the centralizer of $\mathfrak{a}$ is the Lie subalgebra

$$\mathfrak{m} = \left\{ \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_{n+1,n+1}(\mathbb{R}) \mid B \in \mathfrak{so}(n-1) \right\}$$

and the reader should check that

$$\mathfrak{so}(n,1) = \mathfrak{m}\oplus\mathfrak{a}\oplus\mathfrak{n}\oplus\overline{\mathfrak{n}}.$$

We also have

$$[\mathfrak{m},\mathfrak{n}] \subseteq \mathfrak{n},$$

so $\mathfrak{m}\oplus\mathfrak{a}\oplus\mathfrak{n}$ is a subalgebra of $\mathfrak{so}(n,1)$. The group, M, associated with $\mathfrak{m}$ is isomorphic to $\mathbf{SO}(n-1)$ and it can be shown that $B = MAN$ is a subgroup of $\mathbf{SO}_0(n,1)$. In fact,

$$\mathbf{SO}_0(n,1)/(MAN) = KAN/MAN = K/M = \mathbf{SO}(n)/\mathbf{SO}(n-1) = S^{n-1}.$$

It is customary to denote the subalgebra $\mathfrak{m}\oplus\mathfrak{a}$ by $\mathfrak{g}_0$, the algebra $\mathfrak{n}$ by $\mathfrak{g}_1$ and $\overline{\mathfrak{n}}$ by $\mathfrak{g}_{-1}$, so that $\mathfrak{so}(n,1) = \mathfrak{m}\oplus\mathfrak{a}\oplus\mathfrak{n}\oplus\overline{\mathfrak{n}}$ is also written

$$\mathfrak{so}(n,1) = \mathfrak{g}_0\oplus\mathfrak{g}_{-1}\oplus\mathfrak{g}_1.$$

By the way, if $N \in \mathfrak{n}$, then

$$e^N = I + N + \frac{1}{2}N^2,$$

and since $N + \frac{1}{2}N^2$ is also nilpotent, e^N can't be diagonalized when $N \neq 0$. This provides a simple example of matrices in $\mathbf{SO}_0(n,1)$ that can't be diagonalized.

Combining Proposition 2.3 and Proposition 4.16, we have the corollary:

Corollary 4.17 *Every matrix $A \in \mathbf{O}(n,1)$ can be written as*

$$A = \begin{pmatrix} Q & 0 \\ 0 & \epsilon \end{pmatrix} e^{\begin{pmatrix} 0 & u \\ u^\top & 0 \end{pmatrix}}$$

where $Q \in \mathbf{O}(n)$, $\epsilon = \pm 1$ and $u \in \mathbb{R}^n$.

Observe that Corollary 4.17 proves that every matrix, $A \in \mathbf{SO}_0(n,1)$, can be written as

$$A = Pe^S, \quad \text{with } P \in K \cong \mathbf{SO}(n) \text{ and } S \in \mathfrak{p},$$

i.e.,

$$\mathbf{SO}_0(n,1) = K\exp(\mathfrak{p}),$$

a version of the polar decomposition for $\mathbf{SO}_0(n,1)$.

Now, it is known that the exponential map, $\exp\colon \mathfrak{so}(n) \to \mathbf{SO}(n)$, is surjective. So, when $A \in \mathbf{SO}_0(n,1)$, since then $Q \in \mathbf{SO}(n)$ and $\epsilon = +1$, the matrix

$$\begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}$$

is the exponential of some skew symmetric matrix

$$C = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix} \in \mathfrak{so}(n,1),$$

and we can write $A = e^C e^Z$, with $C \in \mathfrak{k}$ and $Z \in \mathfrak{p}$. Unfortunately, C and Z generally don't commute, so it is generally not true that $A = e^{C+Z}$. Thus, we don't get an "easy" proof of the surjectivity of the exponential $\exp\colon \mathfrak{so}(n,1) \to \mathbf{SO}_0(n,1)$. This is not too surprising because, to the best of our knowledge, proving surjectivity for all n is not a simple matter. One proof is due to Nishikawa [48] (1983). Nishikawa's paper is rather short, but this is misleading. Indeed, Nishikawa relies on a classic paper by Djokovic [20], which itself relies heavily on another fundamental paper by Burgoyne and Cushman [13], published in 1977. Burgoyne and Cushman determine the conjugacy classes for some linear Lie groups and their Lie algebras, where the linear groups arise from an inner product space (real or complex). This inner product is nondegenerate, symmetric, or hermitian or skew-symmetric of skew-hermitian. Altogether, one has to read over 40 pages to fully understand the proof of surjectivity.

In his introduction, Nishikawa states that he is not aware of any other proof of the surjectivity of the exponential for $\mathbf{SO}_0(n,1)$. However, such a proof was also given by Marcel Riesz as early as 1957, in some lectures notes that he gave while visiting the University of Maryland in 1957-1958. These notes were probably not easily available until 1993, when they were published in book form, with commentaries, by Bolinder and Lounesto [52].

Interestingly, these two proofs use very different methods. The Nishikawa–Djokovic–Burgoyne and Cushman proof makes heavy use of methods in Lie groups and Lie algebra, although not far beyond linear algebra. Riesz's proof begins with a deep study of the structure of the minimal polynomial of a Lorentz isometry (Chapter III). This is a beautiful argument that takes about 10 pages. The story is not over, as it takes most of Chapter IV (some 40 pages) to prove the surjectivity of the exponential (actually, Riesz proves other things along the way). In any case, the reader can see that both proofs are quite involved.

It is worth noting that Milnor (1969) also uses techniques very similar to those used by Riesz (in dealing with minimal polynomials of isometries) in his paper on isometries of inner product spaces [41].

What we will do to close this section is to give a relatively simple proof that the exponential map, $\exp\colon \mathfrak{so}(1,3) \to \mathbf{SO}_0(1,3)$, is surjective. In the case of $\mathbf{SO}_0(1,3)$, we can use the fact that $\mathbf{SL}(2,\mathbb{C})$ is a two-sheeted covering space of $\mathbf{SO}_0(1,3)$, which means that there is a homomorphism, $\varphi\colon \mathbf{SL}(2,\mathbb{C}) \to \mathbf{SO}_0(1,3)$, which is surjective and that $\mathrm{Ker}\,\varphi = \{-I, I\}$. Then, the small miracle is that, although the exponential, $\exp\colon \mathfrak{sl}(2,\mathbb{C}) \to \mathbf{SL}(2,\mathbb{C})$, is *not* surjective, for every $A \in \mathbf{SL}(2,\mathbb{C})$, *either A or $-A$ is in the image of the exponential!*

Proposition 4.18 *Given any matrix*

$$B = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathfrak{sl}(2,\mathbb{C}),$$

let ω be any of the two complex roots of $a^2 + bc$. If $\omega \neq 0$, then

$$e^B = \cosh\omega\, I + \frac{\sinh\,\omega}{\omega}\, B,$$

and $e^B = I + B$, if $a^2 + bc = 0$. Furthermore, every matrix $A \in \mathbf{SL}(2,\mathbb{C})$ is in the image of the exponential map, unless $A = -I + N$, where N is a nonzero nilpotent (i.e., $N^2 = 0$ with $N \neq 0$). Consequently, for any $A \in \mathbf{SL}(2,\mathbb{C})$, either A or $-A$ is of the form e^B, for some $B \in \mathfrak{sl}(2,\mathbb{C})$.

Proof. Observe that

$$A^2 = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}\begin{pmatrix} a & b \\ c & -a \end{pmatrix} = (a^2 + bc)I.$$

Then, it is straighforward to prove that

$$e^B = \cosh\omega\, I + \frac{\sinh\,\omega}{\omega}\, B,$$

where ω is a square root of $a^2 + bc$ is $\omega \neq 0$, otherwise, $e^B = I + B$.

Let

$$A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

be any matrix in $\mathbf{SL}(2,\mathbb{C})$. We would like to find a matrix, $B \in \mathfrak{sl}(2,\mathbb{C})$, so that $A = e^B$. In view of the above, we need to solve the system

$$\begin{aligned} \cosh\omega + \frac{\sinh\omega}{\omega}a &= \alpha \\ \cosh\omega - \frac{\sinh\omega}{\omega}a &= \delta \\ \frac{\sinh\omega}{\omega}b &= \beta \\ \frac{\sinh\omega}{\omega}c &= \gamma. \end{aligned}$$

From the first two equations, we get

$$\begin{aligned}\cosh\omega &= \frac{\alpha+\delta}{2}\\ \frac{\sinh\omega}{\omega}a &= \frac{\alpha-\delta}{2}.\end{aligned}$$

Thus, we see that we need to know whether complex cosh is surjective and when complex sinh is zero. We claim:

(1) cosh is surjective.

(2) $\sinh z = 0$ iff $z = n\pi i$, where $n \in \mathbb{Z}$.

Given any $c \in \mathbb{C}$, we have $\cosh\omega = c$ iff

$$e^{2\omega} - 2e^{\omega}c + 1 = 0.$$

The corresponding algebraic equation

$$Z^2 - 2cZ + 1 = 0$$

has discriminant $4(c^2-1)$ and it has two complex roots

$$Z = c \pm \sqrt{c^2-1}$$

where $\sqrt{c^2-1}$ is some square root of c^2-1. Observe that these roots are *never zero*. Therefore, we can find a complex log of $c+\sqrt{c^2-1}$, say ω, so that $e^{\omega} = c+\sqrt{c^2-1}$ is a solution of $e^{2\omega} - 2e^{\omega}c + 1 = 0$. This proves the surjectivity of cosh.

We have $\sinh\omega = 0$ iff $e^{2\omega} = 1$; this holds iff $2\omega = n2\pi i$, i.e., $\omega = n\pi i$.

Observe that

$$\frac{\sinh n\pi i}{n\pi i} = 0 \quad \text{if } n \neq 0, \text{ but} \quad \frac{\sinh n\pi i}{n\pi i} = 1 \quad \text{when } n = 0.$$

We know that

$$\cosh\omega = \frac{\alpha+\delta}{2}$$

can always be solved.

Case 1. If $\omega \neq n\pi i$, with $n \neq 0$, then

$$\frac{\sinh\omega}{\omega} \neq 0$$

and the other equations can be solved, too (this includes the case $\omega = 0$). Therefore, in this case, the exponential is surjective. It remains to examine the other case.

Case 2. Assume $\omega = n\pi i$, with $n \neq 0$. If n is even, then $e^\omega = 1$, which implies

$$\alpha + \delta = 2.$$

However, $\alpha\delta - \beta\gamma = 1$ (since $A \in \mathbf{SL}(2, \mathbb{C})$), so we deduce that A has the double eigenvalue, 1. Thus, $N = A - I$ is nilpotent (i.e., $N^2 = 0$) and has zero trace; but then, $N \in \mathfrak{so}(2, \mathbb{C})$ and

$$e^N = I + N = I + A - I = A.$$

If n is odd, then $e^\omega = -1$, which implies

$$\alpha + \delta = -2.$$

In this case, A has the double eigenvalue -1 and $A + I = N$ is nilpotent. So, $A = -I + N$, where N is nilpotent. If $N \neq 0$, then A cannot be diagonalized. We claim that there is no $B \in \mathfrak{so}(2, \mathbb{C})$ so that $e^B = A$.

Indeed, any matrix $B \in \mathfrak{so}(2, \mathbb{C})$ has zero trace, which means that if λ_1 and λ_2 are the eigenvalues of B, then $\lambda_1 = -\lambda_2$. If $\lambda_1 \neq 0$, then $\lambda_1 \neq \lambda_2$ so B can be diagonalized, but then e^B can also be diagonalized, contradicting the fact that A can't be diagonalized. If $\lambda_1 = \lambda_2 = 0$, then e^B has the double eigenvalue $+1$, but A has eigenvalues -1. Therefore, the only matrices $A \in \mathbf{SL}(2, \mathbb{C})$ that are not in the image of the exponential are those of the form $A = -I + N$, where N is a nonzero nilpotent. However, note that $-A = I - N$ *is* in the image of the exponential. $\square$

Remark: If we restrict our attention to $\mathbf{SL}(2, \mathbb{R})$, then we have the following proposition that can be used to prove that the exponential map $\exp\colon \mathfrak{so}(1, 2) \to \mathbf{SO}_0(1, 2)$ is surjective:

Proposition 4.19 *Given any matrix*

$$B = \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathfrak{sl}(2, \mathbb{R}),$$

if $a^2 + b > 0$, then let $\omega = \sqrt{a^2 + bc} > 0$ and if $a^2 + b < 0$, then let $\omega = \sqrt{-(a^2 + bc)} > 0$ (i.e., $\omega^2 = -(a^2 + bc)$). In the first case ($a^2 + bc > 0$), we have

$$e^B = \cosh\omega\, I + \frac{\sinh\,\omega}{\omega}\, B,$$

and in the second case ($a^2 + bc < 0$), we have

$$e^B = \cos\omega\, I + \frac{\sin\,\omega}{\omega}\, B.$$

If $a^2 + bc = 0$, then $e^B = I + B$. Furthermore, every matrix $A \in \mathbf{SL}(2, \mathbb{R})$ whose trace satisfies $\mathrm{tr}(A) \geq -2$ *in the image of the exponential map. Consequently, for any $A \in \mathbf{SL}(2, \mathbb{R})$, either A or $-A$ is of the form e^B, for some $B \in \mathfrak{sl}(2, \mathbb{R})$.*

We now return to the relationship between $\mathbf{SL}(2, \mathbb{C})$ and $\mathbf{SO}_0(1, 3)$. In order to define a homomorphism $\varphi\colon \mathbf{SL}(2, \mathbb{C}) \to \mathbf{SO}_0(1, 3)$, we begin by defining a linear bijection, h, between $\mathbb{R}^4$ and $\mathbf{H}(2)$, the set of complex 2×2 Hermitian matrices, by

$$(t, x, y, z) \mapsto \begin{pmatrix} t + x & y - iz \\ y + iz & t - x \end{pmatrix}.$$

Those familiar with quantum physics will recognize a linear combination of the Pauli matrices! The inverse map is easily defined and we leave it as an exercise. For instance, given a Hermitian matrix

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

we have

$$t = \frac{a + d}{2}, \; x = \frac{a - d}{2}, \quad \text{etc.}$$

Next, for any $A \in \mathbf{SL}(2, \mathbb{C})$, we define a map, $l_A\colon \mathbf{H}(2) \to \mathbf{H}(2)$, *via*

$$S \mapsto ASA^*.$$

(Here, $A^* = \overline{A}^{\top}$.) Using the linear bijection $h\colon \mathbb{R}^4 \to \mathbf{H}(2)$ and its inverse, we obtain a map $\mathrm{lor}_A\colon \mathbb{R}^4 \to \mathbb{R}^4$, where

$$\mathrm{lor}_A = h^{-1} \circ l_A \circ h.$$

As ASA^* is hermitian, we see that l_A is well defined. It is obviously linear and since $\det(A) = 1$ (recall, $A \in \mathbf{SL}(2, \mathbb{C})$) and

$$\det \begin{pmatrix} t + x & y - iz \\ y + iz & t - x \end{pmatrix} = t^2 - x^2 - y^2 - z^2,$$

we see that lor_A preserves the Lorentz metric! Furthermore, it is not hard to prove that $\mathbf{SL}(2, \mathbb{C})$ is connected (use the polar form or analyze the eigenvalues of a matrix in $\mathbf{SL}(2, \mathbb{C})$, for example, as in Duistermatt and Kolk [25] (Chapter 1, Section 1.2)) and that the map

$$\varphi\colon A \mapsto \mathrm{lor}_A$$

is a continuous group homomorphism. Thus, the range of φ is a connected subgroup of $\mathbf{SO}_0(1, 3)$. This shows that $\varphi\colon \mathbf{SL}(2, \mathbb{C}) \to \mathbf{SO}_0(1, 3)$ is indeed a homomorphism. It remains to prove that it is surjective and that its kernel is $\{I, -I\}$.

Proposition 4.20 *The homomorphism, $\varphi\colon \mathbf{SL}(2, \mathbb{C}) \to \mathbf{SO}_0(1, 3)$, is surjective and its kernel is $\{I, -I\}$.*

Proof. Recall that from Theorem 2.6, the Lorentz group $\mathbf{SO}_0(1,3)$ is generated by the matrices of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \quad \text{with } P \in \mathbf{SO}(3)$$

and the matrices of the form

$$\begin{pmatrix} \cosh\alpha & \sinh\alpha & 0 & 0 \\ \sinh\alpha & \cosh\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Thus, to prove the surjectivity of φ, it is enough to check that the above matrices are in the range of φ. For matrices of the second kind, the reader should check that

$$A = \begin{pmatrix} e^{\frac{1}{2}\alpha} & 0 \\ 0 & e^{-\frac{1}{2}\alpha} \end{pmatrix}$$

does the job. For matrices of the first kind, we recall that the group of unit quaternions, $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, can be viewed as $\mathbf{SU}(2)$, *via* the correspondence

$$a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mapsto \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix},$$

where $a, b, c, d \in \mathbb{R}$ and $a^2 + b^2 + c^2 + d^2 = 1$. Moreover, the algebra of quaternions, $\mathbb{H}$, is the real algebra of matrices as above, without the restriction $a^2 + b^2 + c^2 + d^2 = 1$ and $\mathbb{R}^3$ is embedded in $\mathbb{H}$ as the *pure quaternions*, i.e., those for which $a = 0$. Observe that when $a = 0$,

$$\begin{pmatrix} ib & c+id \\ -c+id & -ib \end{pmatrix} = i \begin{pmatrix} b & d-ic \\ d+ic & -b \end{pmatrix} = ih(0, b, d, c).$$

Therefore, we have a bijection between the pure quaternions and the subspace of the hermitian matrices

$$\begin{pmatrix} b & d-ic \\ d+ic & -b \end{pmatrix}$$

for which $a = 0$, the inverse being division by i, i.e., multiplication by $-i$. Also, when q is a unit quaternion, let $\overline{q} = a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$, and observe that $\overline{q} = q^{-1}$. Using the embedding $\mathbb{R}^3 \hookrightarrow \mathbb{H}$, for every unit quaternion, $q \in \mathbf{SU}(2)$, define the map, $\rho_q \colon \mathbb{R}^3 \to \mathbb{R}^3$, by

$$\rho_q(X) = qX\overline{q} = qXq^{-1},$$

for all $X \in \mathbb{R}^3 \hookrightarrow \mathbb{H}$. Then, it is well known that ρ_q is a rotation (i.e., $\rho_q \in \mathbf{SO}(3)$) and, moreover, the map $q \mapsto \rho_q$, is a surjective homomorphism, $\rho \colon \mathbf{SU}(2) \to \mathbf{SO}(3)$, and Ker $\varphi = \{I, -I\}$ (For example, see Gallier [27], Chapter 8).

Now, consider a matrix, A, of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} \quad \text{with } P \in \mathbf{SO}(3).$$

We claim that we can find a matrix, $B \in \mathbf{SL}(2,\mathbb{C})$, such that $\varphi(B) = \mathrm{lor}_B = A$. We claim that we can pick $B \in \mathbf{SU}(2) \subseteq \mathbf{SL}(2,\mathbb{C})$. Indeed, if $B \in \mathbf{SU}(2)$, then $B^* = B^{-1}$, so

$$B \begin{pmatrix} t+x & y-iz \\ y+iz & t-x \end{pmatrix} B^* = t \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - iB \begin{pmatrix} ix & z+iy \\ -z+iy & -ix \end{pmatrix} B^{-1}.$$

The above shows that lor_B leaves the coordinate t invariant. The term

$$B \begin{pmatrix} ix & z+iy \\ -z+iy & -ix \end{pmatrix} B^{-1}$$

is a pure quaternion corresponding to the application of the rotation ρ_B induced by the quaternion B to the pure quaternion associated with (x, y, z) and multiplication by $-i$ is just the corresponding hermitian matrix, as explained above. But, we know that for any $P \in \mathbf{SO}(3)$, there is a quaternion, B, so that $\rho_B = P$, so we can find our $B \in \mathbf{SU}(2)$ so that

$$\mathrm{lor}_B = \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = A.$$

Finally, assume that $\varphi(A) = I$. This means that

$$ASA^* = S,$$

for all hermitian matrices, S, defined above. In particular, for $S = I$, we get $AA^* = I$, i.e., $A \in \mathbf{SU}(2)$. We have

$$AS = SA$$

for all hermitian matrices, S, defined above, so in particular, this holds for diagonal matrices of the form

$$\begin{pmatrix} t+x & 0 \\ 0 & t-x \end{pmatrix},$$

with $t + x \neq t - x$. We deduce that A is a diagonal matrix, and since it is unitary, we must have $A = \pm I$. Therefore, $\mathrm{Ker}\,\varphi = \{I, -I\}$. $\square$

Remark: The group $\mathbf{SL}(2,\mathbb{C})$ is isomorphic to the group $\mathbf{Spin}(1,3)$, which is a (simply-connected) double-cover of $\mathbf{SO}_0(1,3)$. This is a standard result of Clifford algebra theory, see Bröcker and tom Dieck [11] or Fulton and Harris [26]. What we just did is to provide a direct proof of this fact.

We just proved that there is an isomorphism

$$\mathbf{SL}(2,\mathbb{C})/\{I, -I\} \cong \mathbf{SO}_0(1,3).$$

However, the reader may recall that $\mathbf{SL}(2,\mathbb{C})/\{I, -I\} = \mathbf{PSL}(2,\mathbb{C}) \cong \mathbf{M\ddot{o}b}^+$. Therefore, the Lorentz group is isomorphic to the Möbius group.

We now have all the tools to prove that the exponential map, $\exp\colon \mathfrak{so}(1,3) \to \mathbf{SO}_0(1,3)$, is surjective.

Theorem 4.21 *The exponential map,* $\exp\colon \mathfrak{so}(1,3) \to \mathbf{SO}_0(1,3)$*, is surjective.*

Proof. First, recall from Proposition 4.4 that the following diagram commutes:

$$\begin{array}{ccc} \mathbf{SL}(2,\mathbb{C}) & \xrightarrow{\varphi} & \mathbf{SO}_0(1,3) \\ {\scriptstyle\exp}\uparrow & & \uparrow{\scriptstyle\exp} \\ \mathfrak{sl}(2,\mathbb{C}) & \xrightarrow[d\varphi_1]{} & \mathfrak{so}(1,3) \end{array}.$$

Pick any $A \in \mathbf{SO}_0(1,3)$. By Proposition 4.20, the homomorphism φ is surjective and as $\mathrm{Ker}\,\varphi = \{I, -I\}$, there exists some $B \in \mathbf{SL}(2,\mathbb{C})$ so that

$$\varphi(B) = \varphi(-B) = A.$$

Now, by Proposition 4.18, for any $B \in \mathbf{SL}(2,\mathbb{C})$, either B or $-B$ is of the form e^C, for some $C \in \mathfrak{sl}(2,\mathbb{C})$. By the commutativity of the diagram, if we let $D = d\varphi_1(C) \in \mathfrak{so}(1,3)$, we get

$$A = \varphi(\pm e^C) = e^{d\varphi_1(C)} = e^D,$$

with $D \in \mathfrak{so}(1,3)$, as required. $\square$

Remark: We can restrict the bijection $h\colon \mathbb{R}^4 \to \mathbf{H}(2)$ defined earlier to a bijection between $\mathbb{R}^3$ and the space of real symmetric matrices of the form

$$\begin{pmatrix} t+x & y \\ y & t-x \end{pmatrix}.$$

Then, if we also restrict ourselves to $\mathbf{SL}(2,\mathbb{R})$, for any $A \in \mathbf{SL}(2,\mathbb{R})$ and any symmetric matrix, S, as above, we get a map

$$S \mapsto ASA^{\top}.$$

The reader should check that these transformations correspond to isometries in $\mathbf{SO}_0(1,2)$ and we get a homomorphism, $\varphi\colon \mathbf{SL}(2,\mathbb{R}) \to \mathbf{SO}_0(1,2)$. Then, we have a version of Proposition 4.20 for $\mathbf{SL}(2,\mathbb{R})$ and $\mathbf{SO}_0(1,2)$:

Proposition 4.22 *The homomorphism,* $\varphi\colon \mathbf{SL}(2,\mathbb{R}) \to \mathbf{SO}_0(1,2)$*, is surjective and its kernel is* $\{I, -I\}$*.*

Using Proposition 4.22 and Proposition 4.19, we get a version of Theorem 4.21 for $\mathbf{SO}_0(1,2)$:

Theorem 4.23 *The exponential map,* $\exp\colon \mathfrak{so}(1,2) \to \mathbf{SO}_0(1,2)$*, is surjective.*

Also observe that $\mathbf{SO}_0(1,1)$ consists of the matrices of the form

$$A = \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}$$

and a direct computation shows that

$$e^{\begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}} = \begin{pmatrix} \cosh\alpha & \sinh\alpha \\ \sinh\alpha & \cosh\alpha \end{pmatrix}.$$

Thus, we see that the map $\exp\colon \mathfrak{so}(1,1) \to \mathbf{SO}_0(1,1)$ is also surjective. Therefore, we have proved that $\exp\colon \mathfrak{so}(1,n) \to \mathbf{SO}_0(1,n)$ is surjective for $n = 1, 2, 3$. This actually holds for all $n \geq 1$, but the proof is much more involved, as we already discussed earlier.

4.6 More on the Topology of $\mathbf{O}(p,q)$ and $\mathbf{SO}(p,q)$

It turns out that the topology of the group, $\mathbf{O}(p,q)$, is completely determined by the topology of $\mathbf{O}(p)$ and $\mathbf{O}(q)$. This result can be obtained as a simple consequence of some standard Lie group theory. The key notion is that of a pseudo-algebraic group.

Consider the group, $\mathbf{GL}(n,\mathbb{C})$, of invertible $n \times n$ matrices with complex coefficients. If $A = (a_{kl})$ is such a matrix, denote by x_{kl} the real part (resp. y_{kl}, the imaginary part) of a_{kl} (so, $a_{kl} = x_{kl} + iy_{kl}$).

Definition 4.9 A subgroup, G, of $\mathbf{GL}(n,\mathbb{C})$ is *pseudo-algebraic* iff there is a finite set of polynomials in $2n^2$ variables with real coefficients, $\{P_i(X_1,\ldots,X_{n^2},Y_1,\ldots,Y_{n^2})\}_{i=1}^t$, so that

$$A = (x_{kl} + iy_{kl}) \in G \quad \text{iff} \quad P_i(x_{11},\ldots,x_{nn},y_{11},\ldots,y_{nn}) = 0, \quad \text{for } i = 1,\ldots,t.$$

Recall that if A is a complex $n \times n$-matrix, its *adjoint*, A^*, is defined by $A^* = (\overline{A})^\top$. Also, $\mathbf{U}(n)$ denotes the group of unitary matrices, i.e., those matrices $A \in \mathbf{GL}(n,\mathbb{C})$ so that $AA^* = A^*A = I$, and $\mathbf{H}(n)$ denotes the vector space of Hermitian matrices, i.e., those matrices A so that $A^* = A$. Then, we have the following theorem which is essentially a refined version of the polar decomposition of matrices:

Theorem 4.24 *Let G be a pseudo-algebraic subgroup of $\mathbf{GL}(n,\mathbb{C})$ stable under adjunction (i.e., we have $A^* \in G$ whenever $A \in G$). Then, there is some integer, $d \in \mathbb{N}$, so that G is homeomorphic to $(G \cap \mathbf{U}(n)) \times \mathbb{R}^d$. Moreover, if $\mathfrak{g}$ is the Lie algebra of G, the map*

$$(\mathbf{U}(n) \cap G) \times (\mathbf{H}(n) \cap \mathfrak{g}) \longrightarrow G, \quad \textit{given by} \quad (U,H) \mapsto Ue^H,$$

is a homeomorphism onto G.

Proof. A proof can be found in Knapp [36], Chapter 1, or Mneimné and Testard [44], Chapter 3. □

We now apply Theorem 4.24 to determine the structure of the space $\mathbf{O}(p,q)$. We know that $\mathbf{O}(p,q)$ consists of the matrices, A, in $\mathbf{GL}(p+q,\mathbb{R})$ such that

$$A^\top I_{p,q} A = I_{p,q},$$

and so, $\mathbf{O}(p,q)$ is clearly pseudo-algebraic. Using the above equation, it is easy to determine the Lie algebra, $\mathfrak{o}(p,q)$, of $\mathbf{O}(p,q)$. We find that $\mathfrak{o}(p,q)$ is given by

$$\mathfrak{o}(p,q) = \left\{ \begin{pmatrix} X_1 & X_2 \\ X_2^\top & X_3 \end{pmatrix} \;\middle|\; X_1^\top = -X_1,\ X_3^\top = -X_3,\ X_2 \text{ arbitrary} \right\}$$

where X_1 is a $p \times p$ matrix, X_3 is a $q \times q$ matrix and X_2 is a $p \times q$ matrix. Consequently, it immediately follows that

$$\mathfrak{o}(p,q) \cap \mathbf{H}(p+q) = \left\{ \begin{pmatrix} 0 & X_2 \\ X_2^\top & 0 \end{pmatrix} \;\middle|\; X_2 \text{ arbitrary} \right\},$$

a vector space of dimension pq.

Some simple calculations also show that

$$\mathbf{O}(p,q) \cap \mathbf{U}(p+q) = \left\{ \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \;\middle|\; X_1 \in \mathbf{O}(p),\ X_2 \in \mathbf{O}(q) \right\} \cong \mathbf{O}(p) \times \mathbf{O}(q).$$

Therefore, we obtain the structure of $\mathbf{O}(p,q)$:

Proposition 4.25 *The topological space $\mathbf{O}(p,q)$ is homeomorphic to $\mathbf{O}(p) \times \mathbf{O}(q) \times \mathbb{R}^{pq}$.*

Since $\mathbf{O}(p)$ has two connected components when $p \geq 1$, we see that $\mathbf{O}(p,q)$ has four connected components when $p,q \geq 1$. It is also obvious that

$$\mathbf{SO}(p,q) \cap \mathbf{U}(p+q) = \left\{ \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \;\middle|\; X_1 \in \mathbf{O}(p),\ X_2 \in \mathbf{O}(q),\ \det(X_1)\det(X_2) = 1 \right\}.$$

This is a subgroup of $\mathbf{O}(p) \times \mathbf{O}(q)$ that we denote $S(\mathbf{O}(p) \times \mathbf{O}(q))$. Furthermore, it is easy to show that $\mathfrak{so}(p,q) = \mathfrak{o}(p,q)$. Thus, we also have

Proposition 4.26 *The topological space $\mathbf{SO}(p,q)$ is homeomorphic to $S(\mathbf{O}(p) \times \mathbf{O}(q)) \times \mathbb{R}^{pq}$.*

Observe that the dimension of all these space depends only on $p+q$: It is $(p+q)(p+q-1)/2$. Also, $\mathbf{SO}(p,q)$ has two connected components when $p,q \geq 1$. The connected component of I_{p+q} is the group $\mathbf{SO}_0(p,q)$. This latter space is homeomorphic to $\mathbf{SO}(q) \times \mathbf{SO}(q) \times \mathbb{R}^{pq}$.

Theorem 4.24 gives the polar form of a matrix $A \in \mathbf{O}(p, q)$: We have

$$A = Ue^S, \quad \text{with} \quad U \in \mathbf{O}(p) \times \mathbf{O}(q) \quad \text{and} \quad S \in \mathfrak{so}(p, q) \cap \mathbf{S}(p + q),$$

where U is of the form

$$U = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix}, \quad \text{with} \quad P \in \mathbf{O}(p) \quad \text{and} \quad Q \in \mathbf{O}(q)$$

and $\mathfrak{so}(p, q) \cap \mathbf{S}(p + q)$ consists of all $(p + q) \times (p + q)$ symmetric matrices of the form

$$S = \begin{pmatrix} 0 & X \\ X^\top & 0 \end{pmatrix},$$

with X an arbitrary $p \times q$ matrix. It turns out that it is not very hard to compute explicitly the exponential, e^S, of such matrices (see Mneimné and Testard [44]). Recall that the functions cosh and sinh also make sense for matrices (since the exponential makes sense) and are given by

$$\cosh(A) = \frac{e^A + e^{-A}}{2} = I + \frac{A^2}{2!} + \cdots + \frac{A^{2k}}{(2k)!} + \cdots$$

and

$$\sinh(A) = \frac{e^A - e^{-A}}{2} = A + \frac{A^3}{3!} + \cdots + \frac{A^{2k+1}}{(2k+1)!} + \cdots.$$

We also set

$$\frac{\sinh(A)}{A} = I + \frac{A^2}{3!} + \cdots + \frac{A^{2k}}{(2k+1)!} + \cdots,$$

which is defined for all matrices, A (even when A is singular). Then, we have

Proposition 4.27 *For any matrix S of the form*

$$S = \begin{pmatrix} 0 & X \\ X^\top & 0 \end{pmatrix},$$

we have

$$e^S = \begin{pmatrix} \cosh((XX^\top)^{\frac{1}{2}}) & \dfrac{\sinh((XX^\top)^{\frac{1}{2}})X}{(XX^\top)^{\frac{1}{2}}} \\ \dfrac{\sinh((X^\top X)^{\frac{1}{2}})X^\top}{(X^\top X)^{\frac{1}{2}}} & \cosh((X^\top X)^{\frac{1}{2}}) \end{pmatrix}.$$

Proof. By induction, it is easy to see that

$$S^{2k} = \begin{pmatrix} (XX^\top)^k & 0 \\ 0 & (X^\top X)^k \end{pmatrix}$$

and
$$S^{2k+1} = \begin{pmatrix} 0 & (XX^\top)^k X \\ (X^\top X)^k X^\top & 0 \end{pmatrix}.$$
The rest is left as an exercise. □

Remark: Although at first glance, e^S does not look symmetric, but it is!

As a consequence of Proposition 4.27, every matrix $A \in \mathbf{O}(p,q)$ has the polar form
$$A = \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} \begin{pmatrix} \cosh((XX^\top)^{\frac{1}{2}}) & \dfrac{\sinh((XX^\top)^{\frac{1}{2}})X}{(XX^\top)^{\frac{1}{2}}} \\ \dfrac{\sinh((X^\top X)^{\frac{1}{2}})X^\top}{(X^\top X)^{\frac{1}{2}}} & \cosh((X^\top X)^{\frac{1}{2}}) \end{pmatrix},$$
with $P \in \mathbf{O}(p)$, $Q \in \mathbf{O}(q)$ and X an arbitrary $p \times q$ matrix.

Chapter 5

Principal Fibre Bundles and Homogeneous Spaces, II

5.1 Fibre Bundles, Vector Bundles

We saw in Section 2.2 that a transitive action, $\cdot\colon G \times X \to X$, of a group, G, on a set, X, yields a description of X as a quotient G/G_x, where G_x is the stabilizer of any element, $x \in X$. In Theorem 2.26, we saw that if X is a "well-behaved" topological space, G is a "well-behaved" topological group and the action is continuous, then G/G_x is homeomorphic to X. In particular the conditions of Theorem 2.26 are satisfied if G is a Lie group and X is a manifold. Intuitively, the above theorem says that G can be viewed as a family of "fibres", G_x, all isomorphic to G, these fibres being parametrized by the "base space", X, and varying smoothly when x moves in X. We have an example of what is called a fibre bundle, in fact, a principal fibre bundle. Now that we know about manifolds and Lie groups, we can be more precise about this situation.

Although we will not make extensive use of it, we begin by reviewing the definition of a fibre bundle because we believe that it clarifies the notion of principal fibre bundle, the concept that is our primary concern. The following definition is not the most general but it is sufficient for our needs:

Definition 5.1 A *fibre bundle with (typical) fibre F and structure group G* is a tuple $\xi = (E, \pi, B, F, G)$, where E, B, F are smooth manifolds, $\pi\colon E \to B$ is a smooth surjective map, G is a Lie group of diffeomorphisms of F and there is some open cover, $\mathcal{U} = (U_\alpha)_{\alpha\in I}$, of B and a family, $\varphi = (\varphi_\alpha)_{\alpha\in I}$, of diffeomorphisms,

$$\varphi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times F,$$

such that the following properties hold:

(a) The diagram

$$\begin{array}{ccc} \pi^{-1}(U_\alpha) & \xrightarrow{\varphi_\alpha} & U_\alpha \times F \\ & \searrow^{\pi} \quad \swarrow_{p_1} & \\ & U_\alpha & \end{array}$$

commutes for all $\alpha \in I$, where $p_1 \colon U_\alpha \times F \to U_\alpha$ is the first projection. Equivalently, for all $(b, y) \in U_\alpha \times F$,

$$\pi \circ \varphi_\alpha^{-1}(b, y) = b.$$

The space B is called the *base space*, E is called the *total space*, F is called the *(typical) fibre*, and each φ_α is called a *(local) trivialization*. The pair $(U_\alpha, \varphi_\alpha)$ is called a *bundle chart* and the family $\{(U_\alpha, \varphi_\alpha)\}$ is a *trivializing cover*. For each $b \in B$, the space $\pi^{-1}(b)$ is called the *fibre above* b; it is also denoted by E_b, and $\pi^{-1}(U_\alpha)$ is also denoted by $E \restriction U_\alpha$.

For every $(U_\alpha, \varphi_\alpha)$ and every $b \in U_\alpha$, we have the diffeomorphism

$$(p_2 \circ \varphi_\alpha) \restriction E_b \colon E_b \to F,$$

where $p_2 \colon U_\alpha \times F \to F$ is the second projection, which we denote by $\varphi_{\alpha,b}$; hence, we have the diffeomorphism $\varphi_{\alpha,b} \colon \pi^{-1}(b)\ (= E_b) \to F$. Furthermore, for all U_α, U_β in $\mathcal{U}$ such that $U_\alpha \cap U_\beta \neq \emptyset$, for every $b \in U_\alpha \cap U_\beta$, there is a relationship between $\varphi_{\alpha,b}$ and $\varphi_{\beta,b}$ which gives the twisting of the bundle:

(b) The diffeomorphism

$$\varphi_{\alpha,b} \circ \varphi_{\beta,b}^{-1} \colon F \to F$$

is an element of the group G.

(c) The map $g_{\alpha\beta} \colon U_\alpha \cap U_\beta \to G$ defined by

$$g_{\alpha\beta}(b) = \varphi_{\alpha,b} \circ \varphi_{\beta,b}^{-1}$$

is smooth. The maps $g_{\alpha\beta}$ are called the *transition maps* of the fibre bundle.

A fibre bundle $\xi = (E, \pi, B, F, G)$ is also called a *fibre bundle over* B. Observe that the bundle charts, $(U_\alpha, \varphi_\alpha)$, are similar to the charts of a manifold. Actually, Definition 5.1 is too restrictive because it does not allow for the addition of compatible bundle charts, for example, when considering a refinement of the cover, $\mathcal{U}$. This problem can easily be fixed using a notion of equivalence of trivializing covers analogous to the equivalence of atlases for manifolds (see Remark (2) below). Also Observe that (b) and (c) imply that there is some smooth map, $g_{\alpha\beta} \colon U_\alpha \cap U_\beta \to G$, so that

$$\varphi_\alpha \circ \varphi_\beta^{-1}(b, x) = (b, g_{\alpha\beta}(b)(x)),$$

for all $b \in U_\alpha \cap U_\beta$ and all $x \in F$. Note that the isomorphism $g_{\alpha,\beta}\colon (U_\alpha \cap U_\beta) \times F \to (U_\alpha \cap U_\beta) \times F$ describes how the fibres viewed over U_β are viewed over U_α. Thus, it might have been better to denote $g_{\alpha,\beta}$ by g^α_β, so that

$$g^\beta_\alpha = \varphi_{\beta,b} \circ \varphi^{-1}_{\alpha,b},$$

where the subscript, α, indicates the source and the superscript, β, indicates the target.

Intuitively, a fibre bundle over B is a family $E = (E_b)_{b\in B}$ of spaces E_b (fibres) indexed by B and varying smoothly as b moves in B, such that every E_b is diffeomorphic to F. The bundle $E = B \times F$ where π is the first projection is called the *trivial bundle* (over B). The local triviality condition (a) says that *locally*, that is, over every subset, U_α, from some open cover of the base space, B, the bundle $\xi \restriction U_\alpha$ is trivial. Note that if G is the trivial one-element group, then a fibre bundle is trivial.

A Möbius strip is an example of a nontrivial fibre bundle where the base space B is the circle S^1 and the fibre space F is the closed interval $[-1, 1]$ and the structural group is $G = \{1, -1\}$, where -1 is the reflection of the interval $[-1, 1]$ about its midpoint, 0. The total space E is the strip obtained by rotating the line segment $[-1, 1]$ around the circle, keeping its midpoint in contact with the circle, and gradually twisting the line segment so that after a full revolution, the segment has been tilted by π. The reader should work out the transition functions for an open cover consisting of two open intervals on the circle. A Klein bottle is also a fibre bundle for which both the base space and the fibre are the circle S^1. Again, the reader should work out the details for this example.

Remark:

(1) The above definition is slightly different (but equivalent) to the definition given in Bott and Tu [7], page 47-48. Definition 5.1 is closer to the one given in Hirzebruch [34]. Bott and Tu and Hirzebruch assume that G acts effectively on the left on the fibre F. This means that there is a smooth action, $\cdot\colon G \times F \to F$, and recall that G acts effectively on F iff for every $g \in G$,

$$\text{if} \quad g \cdot x = x \quad \text{for all } x \in F, \quad \text{then} \quad g = 1.$$

Every $g \in G$ induces a diffeomorphism, $\varphi_g\colon F \to F$, defined by

$$\varphi_g(x) = g \cdot x$$

for all $x \in F$. The fact that G acts effectively on F means that the map $g \mapsto \varphi_g$ is injective. This justifies viewing G as a group of diffeomorphims of F, and from now on, we will denote $\varphi_g(x)$ by $g(x)$.

(2) We observed that Definition 5.1 is too restrictive because it does not allow for the addition of compatible bundle charts. We can fix this problem as follows: Given a trivializing cover, $\{(U_\alpha, \varphi_\alpha)\}$, for any open, U, of B and any diffeomorphism,

$$\varphi\colon \pi^{-1}(U) \to U \times F,$$

we say that (U, φ) is *compatible with the trivializing cover,* $\{(U_\alpha, \varphi_\alpha)\}$, iff whenever $U \cap U_\alpha \neq \emptyset$, there is some smooth map, $g_\alpha \colon U \cap U_\alpha \to G$, so that

$$\varphi \circ \varphi_\alpha^{-1}(b, x) = (b, g_\alpha(b)(x)),$$

for all $b \in U \cap U_\alpha$ and all $x \in F$. Two trivializing covers are *equivalent* iff every bundle chart of one cover is compatible with the other cover. This is equivalent to saying that the union of two trivializing covers is a trivializing cover. Then, we can define a fibre bundle as a tuple, $(E, \pi, B, F, G, \{(U_\alpha, \varphi_\alpha)\})$, where $\{(U_\alpha, \varphi_\alpha)\}$ is an equivalence class of trivializing covers. As for manifolds, given a trivializing cover, $\{(U_\alpha, \varphi_\alpha)\}$, the set of all bundle charts compatible with $\{(U_\alpha, \varphi_\alpha)\}$ is a maximal trivializing cover equivalent to $\{(U_\alpha, \varphi_\alpha)\}$.

A special case of the above occurs when we have a trivializing cover, $\{(U_\alpha, \varphi_\alpha)\}$, with $\mathcal{U} = \{U_\alpha\}$ an open cover of B and another open cover, $\mathcal{V} = (V_\beta)_{\beta \in J}$, of B which is a refinement of $\mathcal{U}$. This means that there is a map, $\tau \colon J \to I$, such that $V_\beta \subseteq U_{\tau(j)}$ for all $\beta \in J$. Then, for every $V_\beta \in \mathcal{V}$, since $V_\beta \subseteq U_{\tau(\beta)}$, the restriction of $\varphi_{\tau(\beta)}$ to V_β is a trivialization

$$\varphi'_\beta \colon \pi^{-1}(V_\beta) \to V_\beta \times F$$

and conditions (b) and (c) are still satisfied, so $(V_\beta, \varphi'_\beta)$ is compatible with $\{(U_\alpha, \varphi_\alpha)\}$.

(3) (For readers familiar with sheaves) Hirzebruch defines the sheaf G_∞ such that $\Gamma(U, G_\infty)$ is the group of smooth functions $g \colon U \to G$, where U is some open subset of B and G is a Lie group acting effectively (on the left) on the fibre F. The group operation on $\Gamma(U, G_\infty)$ is induced by multiplication in G, that is, given two (smooth) functions $g \colon U \to G$ and $h \colon U \to G$,

$$gh(b) = g(b)h(b)$$

for all $b \in U$.

Beware that gh is **not** function composition, unless G itself is a group of functions, which is the case for vector bundles.

Our conditions (b) and (c) are then replaced by the following equivalent condition: For all U_α, U_β in $\mathcal{U}$ such that $U_\alpha \cap U_\beta \neq \emptyset$, there is some $g_{\alpha\beta} \in \Gamma(U_\alpha \cap U_\beta, G_\infty)$ such that

$$\varphi_\alpha \circ \varphi_\beta^{-1}(b, x) = (b, g_{\alpha\beta}(b)(x)),$$

for all $b \in U_\alpha \cap U_\beta$ and all $x \in F$.

(4) The family of transition functions $(g_{\alpha\beta})$ satisfies the *cocycle condition*

$$g_{\alpha\beta}(b) g_{\beta\gamma}(b) = g_{\alpha\gamma}(b),$$

for all α, β, γ such that $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$ and all $b \in U_\alpha \cap U_\beta \cap U_\gamma$. Setting $\alpha = \beta = \gamma$, we get

$$g_{\alpha\alpha} = \mathrm{id},$$

and setting $\gamma = \alpha$, we get

$$g_{\beta\alpha} = g_{\alpha\beta}^{-1}.$$

Again, beware that this means that $g_{\beta\alpha}(b) = g_{\alpha\beta}^{-1}(b)$, where $g_{\alpha\beta}^{-1}(b)$ is the inverse of $g_{\beta\alpha}(b)$ in G. In general, $g_{\alpha\beta}^{-1}$ is **not** the functional inverse of $g_{\beta\alpha}$.

The classic source on fibre bundles is Steenrod [58]. The most comprehensive treatment of fibre bundles and vector bundles is probably given in Husemoller [35]. However, we can hardly recommend this book. We find the presentation overly formal, and intuitions are absent. A more extensive list of references is given at the end of Section 5.2.

Remark: (The following paragraph is intended for readers familiar with Čech cohomology.) The cocycle condition makes it possible to view a fibre bundle over B as a member of a certain (Čech) cohomology set $\check{H}^1(B, \mathcal{G})$, where $\mathcal{G}$ denotes a certain sheaf of functions from the manifold B into the Lie group G, as explained in Hirzebruch [34], Section 3.2. However, this requires defining a noncommutative version of Čech cohomology (at least, for $\check{H}^1$), and clarifying when two open covers and two trivializations define the same fibre bundle over B, or equivalently, defining when two fibre bundles over B are equivalent. If the bundles under considerations are line bundles (see Definition 5.5), then $\check{H}^1(B, \mathcal{G})$ is actually a group. In this case, $G = \mathrm{GL}(1, \mathbb{R}) \approx \mathbb{R}^*$ in the real case and $G = \mathrm{GL}(1, \mathbb{C}) \approx \mathbb{C}^*$ in the complex case (where $\mathbb{R}^* = \mathbb{R} - \{0\}$ and $\mathbb{C}^* = \mathbb{C} - \{0\}$), and the sheaf $\mathcal{G}$ is the sheaf of smooth (real-valued or complex-valued) functions vanishing nowhere. The group $\check{H}^1(B, \mathcal{G})$ plays an important role, especially when the bundle is a holomorphic line bundle over a complex manifold. In the latter case, it is called the *Picard group* of B.

A map of fibre bundles is defined as follows:

Definition 5.2 Given two fibre bundles $\xi_1 = (E_1, \pi_1, B_1, F, G)$ and $\xi_2 = (E_2, \pi_2, B_2, F, G)$ with the same typical fibre F and the same structure group G, a *bundle map (or bundle morphism)* $f \colon \xi_1 \to \xi_2$ is a pair $f = (f_E, f_B)$ of smooth maps $f_E \colon E_1 \to E_2$ and $f_B \colon B_1 \to B_2$ such that the following diagram commutes:

$$\begin{array}{ccc} E_1 & \xrightarrow{f_E} & E_2 \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ B_1 & \xrightarrow[f_B]{} & B_2 \end{array}$$

A bundle map $f \colon \xi_1 \to \xi_2$ is an *isomorphism* if there is some bundle map $g \colon \xi_2 \to \xi_1$ called the *inverse of f* such that

$$g_E \circ f_E = \mathrm{id}, \qquad f_E \circ g_E = \mathrm{id}.$$

The bundles ξ_1 and ξ_2 are called *isomorphic*. Given two fibre bundles $\xi_1 = (E_1, \pi_1, B, F, G)$ and $\xi_2 = (E_2, \pi_2, B, F, G)$ over the same base space B, a *bundle map (or bundle morphism)* $f\colon \xi_1 \to \xi_2$ is a pair $f = (f_E, f_B)$ where $f_B = \mathrm{id}$ (the identity map). Such a bundle map is an *isomorphism* if it has an inverse as defined above. In this case, we say that the bundles ξ_1 and ξ_2 over B are *isomorphic*.

When f is an isomorphism, the surjectivity of π_1 and π_2 implies that

$$g_B \circ f_B = \mathrm{id}, \qquad f_B \circ g_B = \mathrm{id}.$$

Thus, when $f = (f_E, f_B)$ is an isomorphism, both f_E and f_B are diffeomorphisms. Some authors require the "preservation" of fibres, that is, for every $b \in B_1$, the map of fibres

$$f_E \restriction \pi_1^{-1}(b)\colon \pi_1^{-1}(b) \to \pi_2^{-1}(f_B(b))$$

must be a diffeomorphism. This is automatic for isomorphisms.

We can also define the notion of equivalence for fibre bundles over the same base space B (see Hirzebruch [34], Section 3.2, Chern [15], Section 5, and Husemoller [35], Chapter 5). Equivalence of bundles implies that they are isomorphic.

Definition 5.3 Given two fibre bundles $\xi_1 = (E_1, \pi_1, B, F, G)$ and $\xi_2 = (E_2, \pi_2, B, F, G)$ over the same base space B, we say that ξ_1 and ξ_2 are *equivalent* if there is an open cover $\mathcal{U} = (U_\alpha)_{\alpha \in I}$ for B, a family $\varphi = (\varphi_\alpha)_{\alpha \in I}$ of trivializations

$$\varphi_\alpha\colon \pi_1^{-1}(U_\alpha) \to U_\alpha \times F$$

for ξ_1, a family $\varphi' = (\varphi'_\alpha)_{\alpha \in I}$ of trivializations

$$\varphi'_\alpha\colon \pi_2^{-1}(U_\alpha) \to U_\alpha \times F$$

for ξ_2, and a family $(\rho_\alpha)_{\alpha \in I}$ of smooth maps $\rho_\alpha\colon U_\alpha \to G$ such that

$$\varphi_\alpha \circ {\varphi'_\alpha}^{-1}(b, x) = (b, \rho_\alpha(b)(x)),$$

for all $b \in U_\alpha$ and all $x \in F$.

Since the trivializations are bijections, the family $(\rho_\alpha)_{\alpha \in I}$ is unique. The following lemma shows that equivalent fibre bundles are isomorphic:

Proposition 5.1 *If two fibre bundles $\xi_1 = (E_1, \pi_1, B, F, G)$ and $\xi_2 = (E_2, \pi_2, B, F, G)$ over the same base space B are equivalent, then*

$$g'_{\alpha\beta}(b) = \rho_\alpha(b)^{-1} g_{\alpha\beta}(b) \rho_\beta(b),$$

for all $b \in U_\alpha \cap U_\beta$, where the $g_{\alpha\beta}$s are the transition functions associated with the φ_αs and the $g'_{\alpha\beta}$s are the transition functions associated with the φ'_αs. Furthermore, ξ_1 and ξ_2 are isomorphic.

Proof. We only check the first part, leaving the second as an exercise (or consult Husemoller [35], Chapter 5). Recall that

$$\varphi_\alpha \circ \varphi_\beta^{-1}(b, x) = (b, g_{\alpha\beta}(b)(x)),$$

for all $b \in U_\alpha \cap U_\beta$ and all $x \in F$. This is equivalent to

$$\varphi_\beta^{-1}(b, x) = \varphi_\alpha^{-1}(b, g_{\alpha\beta}(b)(x)),$$

and it is desirable to introduce ψ_α such that $\psi_\alpha = \varphi_\alpha^{-1}$. Then, we have

$$\psi_\beta(b, x) = \psi_\alpha(b, g_{\alpha\beta}(b)(x)),$$

and

$$\varphi_\alpha \circ {\varphi'_\alpha}^{-1}(b, x) = (b, \rho_\alpha(b)(x))$$

becomes

$$\psi'_\alpha(b, x) = \psi_\alpha(b, \rho_\alpha(b)(x)).$$

We have

$$\psi'_\beta(b, x) = \psi_\beta(b, \rho_\beta(b)(x)) = \psi_\alpha(b, g_{\alpha\beta}(b)(\rho_\beta(b)(x)))$$

and

$$\psi'_\alpha(b, g'_{\alpha\beta}(b)(x)) = \psi_\alpha(b, \rho_\alpha(b)(g'_{\alpha\beta}(b)(x))).$$

Since we also have

$$\psi'_\beta(b, x) = \psi'_\alpha(b, g'_{\alpha\beta}(b)(x)),$$

we get

$$\psi_\alpha(b, g_{\alpha\beta}(b)(\rho_\beta(b)(x))) = \psi_\alpha(b, \rho_\alpha(b)(g'_{\alpha\beta}(b)(x))),$$

which implies that

$$g_{\alpha\beta}(b)(\rho_\beta(b)(x)) = \rho_\alpha(b)(g'_{\alpha\beta}(b)(x))),$$

i.e.

$$g'_{\alpha\beta}(b) = \rho_\alpha(b)^{-1} g_{\alpha\beta}(b) \rho_\beta(b).$$

□

In general, isomorphic fibre bundles over the same base B may not be equivalent, because a smooth map $h: U \times F \to F$ may not arise from a smooth map $\rho: U \to G$, in the sense that

$$h(b, x) = \rho(b)(x),$$

for all $b \in U$ and all $x \in F$. However, this will be the case when $G = \mathrm{GL}(n, \mathbb{R}))$ (or $G = \mathrm{GL}(n, \mathbb{C})))$ and when h is a smooth map linear in x for every fixed b. This is the case for vector bundles, to be considered shortly. Following Hirzebruch [34], it is possible to modify the notion of a map of fibre bundles over B so that isomorphism of bundles implies bundle equivalence.

Definition 5.4 Given two fibre bundles $\xi_1 = (E_1, \pi_1, B, F, G)$ and $\xi_2 = (E_2, \pi_2, B, F, G)$ over the same base space B, a *strong bundle map* $f\colon \xi_1 \to \xi_2$ is a bundle map as in Definition 5.2 such that the following additional conditions hold:

(a) There is an open cover $\mathcal{U} = (U_\alpha)_{\alpha\in I}$ for B, a family $\varphi = (\varphi_\alpha)_{\alpha\in I}$ of trivializations

$$\varphi_\alpha\colon \pi_1^{-1}(U_\alpha) \to U_\alpha \times F$$

for ξ_1, a family $\varphi' = (\varphi'_\alpha)_{\alpha\in I}$ of trivializations

$$\varphi'_\alpha\colon \pi_2^{-1}(U_\alpha) \to U_\alpha \times F$$

for ξ_2.

(b) For every $b \in B$, there are some trivializations $\varphi_\alpha\colon \pi_1^{-1}(U_\alpha) \to U_\alpha \times F$ and $\varphi'_\alpha\colon \pi_2^{-1}(U_\alpha) \to U_\alpha \times F$, with $b \in U_\alpha$, and some smooth map $\rho_\alpha\colon U_\alpha \to G$ such that

$$\varphi'_\alpha \circ f \circ {\varphi_\alpha}^{-1}(u, x) = (u, \rho_\alpha(b)(x)),$$

for all $u \in U_\alpha$ and all $x \in F$.

A strong bundle map is an isomorphism if it has an inverse as in Definition 5.2. In this case, we say that the bundles ξ_1 and ξ_2 over B are *strongly isomorphic*.

The following lemma is not hard to prove (see Husemoller [35], Chapter 5):

Proposition 5.2 *Two fibre bundles $\xi_1 = (E_1, \pi_1, B, F, G)$ and $\xi_2 = (E_2, \pi_2, B, F, G)$ over the same base space B are equivalent iff they are strongly isomorphic.*

Given a fibre bundle $\xi = (E, \pi, B, F, G)$, we observed that the family $g = (g_{\alpha\beta})$ of transition maps $g_{\alpha\beta}\colon U_\alpha \cap U_\beta \to G$ induced by a trivializing family $\varphi = (\varphi_\alpha)_{\alpha\in I}$ relative to the open cover $\mathcal{U} = (U_\alpha)_{\alpha\in I}$ for B satisfies the *cocycle condition*

$$g_{\alpha\beta}(b)g_{\beta\gamma}(b) = g_{\alpha\gamma}(b),$$

for all α, β, γ such that $U_\alpha \cap U_\beta \cap U_\gamma \neq \emptyset$ and all $b \in U_\alpha \cap U_\beta \cap U_\gamma$. Without altering anything, we may assume that $g_{\alpha\beta}$ is the (unique) function from $\emptyset$ to G when $U_\alpha \cap U_\beta = \emptyset$. Then, we call a family $g = (g_{\alpha\beta})_{(\alpha,\beta)\in I\times I}$ as above a *$\mathcal{U}$-cocycle*, or simply, a *cocycle*. Remarkably, given such a cocycle g relative to $\mathcal{U}$, a fibre bundle ξ_g over B with fibre F and structure group G having g as family of transition functions can be constructed. In view of Proposition 5.1, we say that two cocycles $g = (g_{\alpha\beta})_{(\alpha,\beta)\in I\times I}$ and $g' = (g_{\alpha\beta})_{(\alpha,\beta)\in I\times I}$ are *equivalent* if there is a family $(\rho_\alpha)_{\alpha\in I}$ of smooth maps $\rho_\alpha\colon U_\alpha \to G$ such that

$$g'_{\alpha\beta}(b) = \rho_\alpha(b)^{-1} g_{\alpha\beta}(b)\rho_\beta(b),$$

for all $b \in U_\alpha \cap U_\beta$.

Theorem 5.3 *Given two smooth manifolds B and F, a Lie group G acting effectively on F, an open cover $\mathcal{U} = (U_\alpha)_{\alpha\in I}$ of B, and a cocycle $g = (g_{\alpha\beta})_{(\alpha,\beta)\in I\times I}$, there is a fibre bundle $\xi_g = (E, \pi, B, F, G)$ whose transition maps are the maps in the cocycle g. Furthermore, if g and g' are equivalent cocycles, then ξ_g and $\xi_{g'}$ are equivalent.*

Proof sketch. First, we define the space Z as the disjoint sum

$$Z = \coprod_{\alpha\in I} U_\alpha \times F.$$

We define the relation $\simeq$ on $Z \times Z$ as follows: For all $(b, x) \in U_\beta \times F$ and $(b, y) \in U_\alpha \times F$, if $U_\alpha \cap U_\beta \neq \emptyset$,

$$(b, x) \simeq (b, y) \quad \text{iff} \quad y = g_{\alpha\beta}(b)(x).$$

We let $E = Z/\simeq$, and we give E the largest topology such that the the injections $\eta_\alpha \colon U_\alpha \times F \to Z$ are smooth. The cocycle condition insures that $\simeq$ is indeed an equivalence relation. We define $\pi\colon E \to B$ by $\pi([b, x]) = b$. If $p\colon Z \to E$ is the the quotient map, observe that the maps $p \circ \eta_\alpha \colon U_\alpha \times F \to E$ are injective, and that

$$\pi \circ p \circ \eta_\alpha(b, x) = b.$$

Thus,

$$p \circ \eta_\alpha \colon U_\alpha \times F \to \pi^{-1}(U_\alpha)$$

is a bijection, and we define the trivializing maps by setting

$$\varphi_\alpha = (p \circ \eta_\alpha)^{-1}.$$

It is easily verified that the corresponding transition functions are the original $g_{\alpha\beta}$. There are some details to check, for instance, see Husemoller [35], Chapter 5, or Wells [61]. The fact that ξ_g and $\xi_{g'}$ are equivalent when g and g' are equivalent follows from Proposition 5.1 (see Husemoller [35], Chapter 5). □

Remark: (The following paragraph is intended for readers familiar with Čech cohomology.) Obviously, if we start with a fibre bundle $\xi = (E, \pi, B, F, G)$ whose transition maps are the cocycle $g = (g_{\alpha\beta})$ and form the fibre bundle ξ_g, the bundles ξ and ξ_g are equivalent. This leads to a characterization of the set of equivalence classes of fibre bundles over a base space B as the cohomology set $\check{H}^1(B, \mathcal{G})$. In the present case, the sheaf $\mathcal{G}$ is defined such that $\Gamma(U, \mathcal{G})$ is the group of smooth maps from the open subset U of B to the Lie group G. Since G is not abelian, the coboundary maps have to be interpreted multiplicatively. It is natural to define

$$\delta_0 \colon C^0(\mathcal{U}, \mathcal{G}) \to C^1(\mathcal{U}, \mathcal{G})$$

by

$$(\delta_0 g)_{\alpha\beta} = g_\alpha^{-1} g_\beta,$$

for any $g = (g_\alpha)$, with $g_\alpha \in \Gamma(U_\alpha, \mathcal{G})$. As to

$$\delta_1 \colon C^1(\mathcal{U}, \mathcal{G}) \to C^2(\mathcal{U}, \mathcal{G}),$$

since the cocycle condition in the usual case is

$$g_{\alpha\beta} + g_{\beta\gamma} = g_{\alpha\gamma},$$

we set

$$(\delta_1 g)_{\alpha\beta\gamma} = g_{\alpha\beta} g_{\beta\gamma} g_{\alpha\gamma}^{-1},$$

for any $g = (g_{\alpha\beta})$, with $g_{\alpha\beta} \in \Gamma(U_\alpha \cap U_\beta, \mathcal{G})$. Note that a cocycle $g = (g_{\alpha\beta})$ is indeed an element of $Z^1(\mathcal{U}, \mathcal{G})$, and the condition for being in the kernel of

$$\delta_1 \colon C^1(\mathcal{U}, \mathcal{G}) \to C^2(\mathcal{U}, \mathcal{G})$$

is the cocycle condition

$$g_{\alpha\beta}(b) g_{\beta\gamma}(b) = g_{\alpha\gamma}(b),$$

for all $b \in U_\alpha \cap U_\beta \cap U_\gamma$. In the commutative case, two cocycles g and g' are equivalent if their difference is a boundary, which can be stated as

$$g_{\alpha\beta} + \rho_\beta = g'_{\alpha\beta} + \rho_\alpha = \rho_\alpha + g'_{\alpha\beta},$$

where $\rho_\alpha \in \Gamma(U_\alpha, \mathcal{G})$, for all $\alpha \in I$. In the present case, two cocycles g and g' are equivalent iff there is a family $(\rho_\alpha)_{\alpha \in I}$, with $\rho_\alpha \in \Gamma(U_\alpha, \mathcal{G})$, such that

$$g'_{\alpha\beta}(b) = \rho_\alpha(b)^{-1} g_{\alpha\beta}(b) \rho_\beta(b),$$

for all $b \in U_\alpha \cap U_\beta$. This is the same condition of equivalence defined earlier. Thus, it is easily seen that if $g, h \in Z^1(\mathcal{U}, \mathcal{G})$, then ξ_g and ξ_h are equivalent iff g and h correspond to the same element of the cohomology set $\check{H}^1(\mathcal{U}, \mathcal{G})$. As usual, $\check{H}^1(B, \mathcal{G})$ is defined as the direct limit of the directed system of sets $\check{H}^1(\mathcal{U}, \mathcal{G})$, over the preordered directed family of open covers. For details, see Hirzebruch [34], Section 3.1. In summary, there is a bijection between the equivalence classes of fibre bundles over B (with fibre F and structure group G) and the cohomology set $\check{H}^1(B, \mathcal{G})$. In the case of line bundles, it turns out that $\check{H}^1(B, \mathcal{G})$ is in fact a group.

There are two particularly interesting special cases of fibre bundles:

(1) *Vector bundles*, which are fibre bundles for which the typical fibre is a finite-dimensional vector space V and the structure group is a subgroup of the group of linear isomorphisms ($\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$, where $n = \dim V$).

(2) *Principal fibre bundles*, which are fibre bundles for which the fibre, F, is equal to the structure group G, with G acting on itself by left translation.

Although this is not our main concern, we briefly discuss vector bundles.

Definition 5.5 A *rank n real smooth vector bundle with fibre V* is a tuple $\xi = (E, \pi, B, V)$ such that $(E, \pi, B, V, \mathrm{GL}(n, \mathbb{R}))$ is a smooth fibre bundle and the fibre V is a real vector space of dimension n.

A *rank n complex smooth vector bundle with fibre V* is a tuple $\xi = (E, \pi, B, V)$ such that $(E, \pi, B, V, \mathrm{GL}(n, \mathbb{C}))$ is a smooth fibre bundle and the fibre V is an n-dimensional complex vector space (viewed as a real smooth manifold). When $n = 1$, a vector bundle is called a *line bundle*.

Maps of vector bundles are maps of fibre bundles such that the isomorphisms between fibres are linear.

Definition 5.6 Given two vector bundles $\xi_1 = (E_1, \pi_1, B_1, V)$ and $\xi_2 = (E_2, \pi_2, B_2, V)$ with the same typical fibre V a *bundle map (or bundle morphism)* $f\colon \xi_1 \to \xi_2$ is a pair $f = (f_E, f_B)$ of smooth maps $f_E\colon E_1 \to E_2$ and $f_B\colon B_1 \to B_2$ such that the following diagram commutes:

$$\begin{array}{ccc} E_1 & \xrightarrow{f_E} & E_2 \\ {\scriptstyle \pi_1}\downarrow & & \downarrow{\scriptstyle \pi_2} \\ B_1 & \xrightarrow[f_B]{} & B_2 \end{array}$$

and such that for every $b \in B_1$, the map of fibres

$$f_E \upharpoonright \pi_1^{-1}(b)\colon \pi_1^{-1}(b) \to \pi_2^{-1}(f_B(b))$$

is a linear map. A bundle map *isomorphism* $f\colon \xi_1 \to \xi_2$ is defined as in Definition 5.2. Given two vector bundles $\xi_1 = (E_1, \pi_1, B, V)$ and $\xi_2 = (E_2, \pi_2, B, V)$ over the same base space B, a *bundle map (or bundle morphism)* $f\colon \xi_1 \to \xi_2$ and *isomorphism of vector bundles over B* are also defined as in Definition 5.2.

Some authors require the preservation of fibres, that is, the map

$$f_E \upharpoonright \pi_1^{-1}(b)\colon \pi_1^{-1}(b) \to \pi_2^{-1}(f_B(b))$$

is a bijective linear map.

A *holomorphic vector bundle* is a fibre bundle where E, B are complex manifolds, V is a complex vector space of dimension n, the map π is holomorphic, the φ_α are biholomorphic, and the transistion functions $g_{\alpha\beta}$ are holomorphic. When $n = 1$, a holomorphic vector bundle is called a *holomorphic line bundle*.

Definition 5.3 also applies to vector bundles and defines the notion of equivalence of vector bundles over B. Proposition 5.1 also holds for equivalent vector bundles. This time, because the fibre is a finite-dimensional vector space, two vector bundles over the same base space B are isomorphic iff they are equivalent.

Proposition 5.4 *Two vector bundles $\xi_1 = (E_1, \pi_1, B, F, G)$ and $\xi_2 = (E_2, \pi_2, B, F, G)$ over the same base space B are equivalent iff they are isomorphic.*

5.2 Principal Fibre Bundles

We now consider principal bundles. Such bundles arise in terms of Lie groups acting on manifolds.

Definition 5.7 Let G be a Lie group. A *principal fibre bundle* or for short, a *principal bundle*, is a fibre bundle, $\xi = (E, \pi, B, G, G)$, in which the fibre is equal to the stucture group, G, and G acts on itself by left translation (multiplication on the left). This means that every transition function, $g_{\alpha\beta} \colon U_\alpha \cap U_\beta \to G$, satisfies

$$g_{\alpha\beta}(b)(h) = g_{\alpha\beta}(b)h,$$

for all $b \in U_\alpha \cap U_\beta$ and all $h \in G$. A principal G-bundle is denoted $\xi = (E, \pi, B, G)$.

When we want to emphasize that a principal bundle has structure group, G, we use the locution *principal G-bundle*.

It turns out that if $\xi = (E, \pi, B, G)$ is a principal bundle, then G acts on the total space, E, on the right. For the next proposition, recall that a right action, $\cdot \colon X \times G \to X$, is *free* iff for every $x \in X$, if $x \cdot g = x$, then $g = 1$.

Proposition 5.5 *If $\xi = (E, \pi, B, G)$ is a principal bundle, then there is a right action of G on E. This action takes each fibre to itself and is free. Moreover, E/G is diffeomorphic to B.*

Proof. We show how to define the right action and leave the rest as as exercise. Let $\{(U_\alpha, \varphi_\alpha)\}$ be the equivalence class of trivializing covers defining ξ. For every $z \in E$, pick some U_α so that $\pi(z) \in U_\alpha$ and let $\varphi_\alpha(z) = (b, h)$, where $b = \pi(z)$ and $h \in G$. For any $g \in G$, we set

$$z \cdot g = \varphi_\alpha^{-1}(b, hg).$$

If we can show that this action does not depend on the choice of U_α, then it is clear that it is a free action. Suppose that we also have $b = \pi(z) \in U_\beta$ and that $\varphi_\beta(z) = (b, h')$. By definition of the transition functions, we have

$$h' = g_{\beta\alpha}(b)h \quad \text{and} \quad \varphi_\beta(z \cdot g) = (b, g_{\beta\alpha}(b)(hg)).$$

However,

$$g_{\beta\alpha}(b)(hg) = (g_{\beta\alpha}(b)h)g = h'g,$$

hence

$$z \cdot g = \varphi_\beta^{-1}(b, h'g),$$

which proves that our action does not depend on the choice of U_α. □

Observe that the action of Proposition 5.5 is defined by

$$z \cdot g = \varphi_{\alpha,b}^{-1}(\varphi_{\alpha,b}(z)g), \quad \text{with} \quad b = \pi(z),$$

for all $z \in E$ and all $g \in G$. It is clear that this action satisfies the two properties: For every $(U_\alpha, \varphi_\alpha)$,

(1) $\pi(z \cdot g) = \pi(z)$ and

(2) $\varphi_\alpha(z \cdot g) = \varphi(z) \cdot g$, for all $z \in E$ and all $g \in G$,

where we define the right action of G on $U_\alpha \times G$ so that $(b, h) \cdot g = (b, hg)$. We say that φ_α is *G-equivariant* (or *equivariant*).

The following proposition shows that it is possible to define a principal G-bundle using a suitable right action and equivariant trivializations.

Proposition 5.6 *Let E be a smooth manifold, G a Lie group and let $\cdot\colon E \times G \to E$ be a smooth right action of G on E and assume that*

(a) The right action of G on E is free;

(b) The orbit space, $B = E/G$, is a smooth manifold under the quotient topology and the projection, $\pi\colon E \to E/G$, is smooth;

(c) There is a family of local trivializations, $\{(U_\alpha, \varphi_\alpha)\}$, where $\{U_\alpha\}$ is an open cover for $B = E/G$ and each

$$\varphi_\alpha\colon \pi^{-1}(U_\alpha) \to U_\alpha \times G$$

is an equivariant diffeomorphism, which means that

$$\varphi_\alpha(z \cdot g) = \varphi_\alpha(z) \cdot g,$$

for all $z \in \pi^{-1}(U_\alpha)$ and all $g \in G$, where the right action of G on $U_\alpha \times G$ is $(b, h) \cdot g = (b, hg)$.

Then, $\xi = (E, \pi, E/G, G)$ is a principal G-bundle.

Proof. Since the action of G on E is free, every orbit, $b = z \cdot G$, is isomorphic to G and so, every fibre, $\pi^{-1}(b)$, is isomorphic to G. Thus, given that we have trivializing maps, we just have to prove that G acts by left translation on itself. Pick any (b, h) in $U_\beta \times G$ and let $z \in \pi^{-1}(U_\beta)$ be the unique element such that $\varphi_\beta(z) = (b, h)$. Then, as

$$\varphi_\beta(z \cdot g) = \varphi_\beta(z) \cdot g, \quad \text{for all } g \in G,$$

we have

$$\varphi_\beta(z \cdot g) = \varphi_\beta(\varphi_\beta^{-1}(b, h) \cdot g) = \varphi_\beta(z) \cdot g = (b, h) \cdot g,$$

which implies that

$$\varphi_\beta^{-1}(b, h) \cdot g = \varphi_\beta^{-1}((b, h) \cdot g).$$

Consequently,

$$\varphi_\alpha \circ \varphi_\beta^{-1}(b, h) = \varphi_\alpha \circ \varphi_\beta^{-1}((b, 1) \cdot h) = \varphi_\alpha(\varphi_\beta^{-1}(b, 1) \cdot h) = \varphi_\alpha \circ \varphi_\beta^{-1}(b, 1) \cdot h,$$

and since

$$\varphi_\alpha \circ \varphi_\beta^{-1}(b, h) = (b, g_{\alpha\beta}(b)(h)) \quad \text{and} \quad \varphi_\alpha \circ \varphi_\beta^{-1}(b, 1) = (b, g_{\alpha\beta}(b)(1))$$

we get

$$g_{\alpha\beta}(b)(h) = g_{\alpha\beta}(b)(1)h.$$

This shows that $g_{\alpha\beta}(b)$ is multiplication on the left by $g_{\alpha\beta}(b)(1)$ and ξ is indeed a principal G-bundle. □

Bröcker and tom Dieck [11] (Chapter I, Section 4) and Duistermaat and Kolk [25] (Appendix A) define principal bundles using the conditions of Lemma 5.6. Propositions 5.5 and 5.6 show that this alternate definition is equivalent to ours (Definition 5.7).

Even though we are not aware of any practical applications in computer vision or robotics, we wish to digress briefly on the issue of the triviality of bundles and the existence of sections.

It is certainly a natural question, and it does come up in in physics (field theory), to ask whether a fibre bundle, ξ, is isomorphic to a trivial bundle (if so, we say that ξ is trivial). Generally, this is a very difficult question, but a first step can be made by showing that it reduces to the question of triviality for principal bundles.

Indeed, if $\xi = (E, \pi, B, F, G)$ is a fibre bundle with fibre, F, using Theorem 5.3, we can construct a principal fibre bundle, $P(\xi)$, using the transition functions, $\{g_{\alpha\beta}\}$, of ξ, but using G itself as the fibre (acting on itself by left translation) instead of F. We obtain the *principal bundle, $P(\xi)$, associated to* ξ. Then, given two fibre bundles ξ and ξ', we see that ξ and ξ' are isomorphic iff $P(\xi)$ and $P(\xi')$ are isomorphic. More is true: The fibre bundle ξ is trivial iff the principal fibre bundle $P(\xi)$ is trivial (this is easy to prove, do it!). Morever, there is a test for the triviality of a principal bundle, the existence of a (global) section.

Definition 5.8 Given a fibre bundle, $\xi = (E, \pi, B, F, G)$, a *smooth section* of ξ is a smooth map, $s\colon B \to E$, so that $\pi \circ s = \mathrm{id}_B$. Given an open subset, U, of B, a *(smooth) section of ξ over U* is a smooth map, $s\colon U \to E$, so that $\pi \circ s(b) = b$, for all $b \in U$; we say that s is a *local section* of ξ. The set of all sections over U is denoted $\Gamma(U, \xi)$ and $\Gamma(B, \xi)$ is the set of *global sections* of ξ.

The following proposition, although easy to prove, it crucial:

Proposition 5.7 *If ξ is a principal bundle, then ξ is trivial iff it possesses some global section.*

Generally, in geometry, many objects of interest arise as global sections of some suitable bundle (or sheaf): vector fields, differential forms, tensor fields, etc.

There is also a construction that takes us from principal bundles to fibre bundles. Given a principal bundle, $\xi = (E, \pi, B, G)$, and given a manifold, F, if G acts effectively on F from the left, we can define a fibre bundle, $\xi[F]$, from ξ, with F as typical fibre. As ξ is a principal bundle, recall that G acts on E from the right, so we have a right action of G on $E \times F$, *via*

$$(z, f) \cdot g = (z \cdot g, g^{-1} \cdot f).$$

Consequently, we obtain the orbit set, $E \times F/\sim$, denoted $E \times_G F$, where $\sim$ is the equivalence relation

$$(z, f) \sim (z', f') \quad \text{iff} \quad (\exists g \in G)(z' = z \cdot g,\ g' = g^{-1} \cdot f).$$

Note that the composed map

$$E \times F \xrightarrow{pr_1} E \xrightarrow{\pi} B$$

factors through $E \times_G F$, since

$$\pi(pr_1(z, f)) = \pi(z) = \pi(z \cdot g) = \pi(pr_1(z \cdot g, g^{-1} \cdot f)).$$

Let $p: E \times_G F \to B$ be the corresponding map. The following proposition is not hard to show:

Proposition 5.8 *If $\xi = (E, \pi, B, G)$ is a principal bundle and F is any manifold such that G acts effectively on F from the left, then, $\xi[F] = (E \times_G F, p, B, F, G)$ is a fibre bundle with fibre F and structure group G.*

Let us verify that the charts of ξ yield charts for $\xi[F]$. For any U_α in an open cover for B, we have a diffeomorphism

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

Observe that we have an isomorphism

$$(U_\alpha \times G) \times_G F \cong U_\alpha \times F,$$

where, as usual, G acts on $U_\alpha \times G$ *via* $(z, h) \cdot g = (z, hg)$, an isomorphism

$$p^{-1}(U_\alpha) \cong \pi^{-1}(U_\alpha) \times_G F,$$

and that φ_α induces an isomorphism

$$\pi^{-1}(U_\alpha) \times_G F \xrightarrow{\varphi_\alpha} (U_\alpha \times G) \times_G F.$$

So, we get the commutative diagram

$$\begin{array}{ccc} p^{-1}(U_\alpha) & \xrightarrow{\sim} & U_\alpha \times F \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle pr_1} \\ U_\alpha & = & U_\alpha, \end{array}$$

which yields a local trivialization for $\xi[F]$. In fact, it is easy to see that the transition functions of $\xi[F]$ are the same as the transition functions of ξ.

The fibre bundle $\xi[F]$ is called the fibre bundle *induced by* ξ. Now, if we start with a fibre bundle, ξ, with fibre F and structure group G, make the associated principal bundle, $P(\xi)$, and then the induced fibre bundle, $P(\xi)[F]$, what is the relationship between ξ and $P(\xi)[F]$?

The answer is: ξ and $P(\xi)[F]$ are *equivalent* (this is because the transition functions are the same.)

Now, if we start with a principal G-bundle, ξ, make the fibre bundle, $\xi[F]$, as above, and then the principal bundle, $P(\xi[F])$, we get a principal bundle equivalent to ξ. Therefore, the maps

$$\xi \mapsto \xi[F] \quad \text{and} \quad \xi \mapsto P(\xi),$$

are mutual inverses and they set up a bijection between equivalence classes of principal G-bundles over B and equivalence classes of fibre bundles over B (with structure group, G). Moreover, this map extends to morphisms, so it is functorial. As a consequence, in order to "classify" equivalence classes of fibre bundles (assuming B and G fixed), it is enough to know how to classify principal G-bundles over B. Given some reasonable conditions on the coverings of B, Milnor solved this classification problem, but this is taking us way beyond the scope of these notes!

For more on fibre bundles, vector bundles and principal bundles, see Steenrod [58], Bott and Tu [7], Madsen and Tornehave [39], Griffith and Harris [30], Wells [61], Hirzebruch [34], Milnor and Stasheff [43], Davis and Kirk [19], Atiyah [2], Chern [15], Choquet-Bruhat, DeWitt-Morette and Dillard-Bleick [17], Hirsh [33], Sato [54], Narasimham [47], Sharpe [57] and also Husemoller [35], which covers much more, including characteristic classes.

Proposition 5.6 shows that principal bundles are induced by suitable right actions but we still need sufficient conditions to guarantee conditions (a), (b) and (c). Such conditions are given in the next section.

5.3 Homogeneous Spaces, II

Now that we know about manifolds and Lie groups, we can revisit the notion of homogeneous space given in Definition 2.8, which only applied to groups and sets without any topology or differentiable structure.

Definition 5.9 A *homogeneous space* is a smooth manifold, M, together with a smooth transitive action, $\cdot\colon G \times M \to M$, of a Lie group, G, on M.

In this section, we prove that G is the total space of a principal bundle with base space M and structure group, G_x, the stabilizer of any $x \in M$.

If M is a manifold, G is a Lie group and $\cdot\colon M \times G \to M$ is a smooth right action, in general, M/G is not even Hausdorff. A sufficient condition can be given using the notion of a proper map. If X and Y are two Hausdorff topological spaces,[1] a continuous map, $\varphi\colon X \to Y$, is *proper* iff for every topological space, Z, the map $\varphi \times \mathrm{id}\colon X \times Z \to Y \times Z$ is a *closed map* (A map, f, is a closed map iff the image of any closed set by f is a closed set). If we let Z be a one-point space, we see that a proper map is closed. It can be shown (see Bourbaki, General Topology [9], Chapter 1, Section 10) that a continuous map, $\varphi\colon X \to Y$, is proper iff φ is closed and if $\varphi^{-1}(y)$ is compact for every $y \in Y$. If φ is proper, it is easy to show that $\varphi^{-1}(K)$ is compact in X whenever K is compact in Y. Moreover, if Y is also locally compact, then Y is compactly generated, which means that a subset, C, of Y is closed iff $K \cap C$ is closed in C for every compact subset K of Y (see Munkres [46]). In this case (Y locally compact), φ is a closed map iff $\varphi^{-1}(K)$ is compact in X whenever K is compact in Y (see Bourbaki, General Topology [9], Chapter 1, Section 10).[2] In particular, this is true if Y is a manifold since manifolds are locally compact. Then, we say that the action $\cdot\colon M \times G \to M$ is *proper* iff the map

$$M \times G \longrightarrow M \times M, \quad (x, g) \mapsto (x, x \cdot g)$$

is proper.

If G and M are Hausdorff and G is locally compact, then it can be shown (see Bourbaki, General Topology [9], Chapter 3, Section 4) that the action $\cdot\colon M \times G \to M$ is proper iff for all $x, y \in M$, there exist some open sets, V_x and V_y in M, with $x \in V_x$ and $y \in V_y$, so that the closure, $\overline{K}$, of the set $K = \{g \in G \mid V_x \cdot g \cap V_y \neq \emptyset\}$ is compact in G. In particular, if G has the discrete topology, this conditions holds iff the sets $\{g \in G \mid V_x \cdot g \cap V_y \neq \emptyset\}$ are finite. Also, if G is compact, then $\overline{K}$ is automatically compact, so every compact group acts properly. If the action $\cdot\colon M \times G \to M$ is proper, then the orbit equivalence relation is closed since it is the image of $M \times G$ in $M \times M$, and so, M/G is Hausdorff. We then have the following theorem proved in Duistermaat and Kolk [25] (Chapter 1, Section 11):

Theorem 5.9 *Let M be a smooth manifold, G a Lie group and let $\cdot\colon M \times G \to M$ be a right smooth action which is proper and free. Then, M/G is a principal G-bundle of dimension* $\dim M - \dim G$.

Theorem 5.9 has some interesting corollaries. Let G be a Lie group and let H be a closed subgroup of G. Then, there is a right action of H on G, namely

$$G \times H \longrightarrow G, \quad (g, h) \mapsto gh,$$

and this action is clearly free and proper. Because a closed subgroup of a Lie group is a Lie group, we get the following result whose proof can be found in Bröcker and tom Dieck [11] (Chapter I, Section 4) or Duistermaat and Kolk [25] (Chapter 1, Section 11):

---

[1] It is not necessary to assume that X and Y are Hausdorff but, if X and/or Y are not Hausdorff, we have to replace "compact" by "quasi-compact." We have no need for this extra generality.

[2] Duistermaat and Kolk [25] seem to have overlooked the fact that a condition on Y (such as local compactness) is needed in their remark on lines 5-6, page 53, just before Lemma 1.11.3.

Corollary 5.10 *If G be a Lie group and H a closed subgroup of G, then, the right action of H on G defines a principal H-bundle, $\xi = (G, \pi, G/H, H)$, where $\pi\colon G \to G/H$ is the canonical projection. Moreover, π is a submersion, which means that $d\pi_g$ is surjective for all $g \in G$ (equivalently, the rank of $d\pi_g$ is constant and equal to $\dim G/H$, for all $g \in G$).*

Now, if $\cdot\colon G \times M \to M$ is a smooth transitive action of a Lie group, G, on a manifold, M, we know that the stabilizers, G_x, are all isomorphic and closed (see Section 2.5, Remark after Theorem 2.26). Then, we can let $H = G_x$ and apply Corollary 5.10 to get the following result (mostly proved in in Bröcker and tom Dieck [11] (Chapter I, Section 4):

Proposition 5.11 *Let $\cdot\colon G \times M \to M$ be smooth transitive action of a Lie group, G, on a manifold, M. Then, G/G_x and M are diffeomorphic and G is the total space of a principal bundle, $\xi = (G, \pi, M, G_x)$, where G_x is the stabilizer of any element $x \in M$.*

Thus, we finally see that homogeneous spaces induce principal bundles. Going back to some of the examples of Section 2.2, we see that

(1) $\mathbf{SO}(n+1)$ is a principal $\mathbf{SO}(n)$-bundle over the sphere S^n (for $n \geq 0$).

(2) $\mathbf{SU}(n+1)$ is a principal $\mathbf{SU}(n)$-bundle over the sphere S^{2n+1} (for $n \geq 0$).

(3) $\mathbf{SL}(2,\mathbb{R})$ is a principal $\mathbf{SO}(2)$-bundle over the upper-half space, H.

(4) $\mathbf{GL}(n,\mathbb{R})$ is a principal $\mathbf{O}(n)$-bundle over the space $\mathbf{SPD}(n)$ of symmetric, positive definite matrices.

(5) $\mathbf{SO}(n+1)$ is a principal $\mathbf{O}(n)$-bundle over the real projective space $\mathbb{RP}^n$ (for $n \geq 0$).

(6) $\mathbf{SU}(n+1)$ is a principal $\mathbf{U}(n)$-bundle over the complex projective space $\mathbb{CP}^n$ (for $n \geq 0$).

(7) $\mathbf{O}(n)$ is a principal $\mathbf{O}(k) \times \mathbf{O}(n-k)$-bundle over the Grassmannian, $G(k,n)$.

(8) From Section 2.5, we see that the Lorentz group, $\mathbf{SO}_0(n,1)$, is a principal $\mathbf{SO}(n)$-bundle over the space, $\mathcal{H}_n^+(1)$, consisting of one sheet of the hyperbolic paraboloid $\mathcal{H}_n(1)$.

Thus, we see that both $\mathbf{SO}(n+1)$ and $\mathbf{SO}_0(n,1)$ are principal $\mathbf{SO}(n)$-bundles, the difference being that the base space for $\mathbf{SO}(n+1)$ is the sphere, S^n, which is compact, whereas the base space for $\mathbf{SO}_0(n,1)$ is the (connected) surface, $\mathcal{H}_n^+(1)$, which is not compact. More examples can be given, for example, see Arvanitoyeogos [1].

Bibliography

[1] Andreas Arvanitoyeogos. *An Introduction to Lie Groups and the Geometry of Homogeneous Spaces*. SML, Vol. 22. AMS, first edition, 2003.

[2] Michael F. Atiyah. *K-Theory*. Addison Wesley, first edition, 1988.

[3] Andrew Baker. *Matrix Groups. An Introduction to Lie Group Theory*. SUMS. Springer, 2002.

[4] Marcel Berger. *Géométrie 1*. Nathan, 1990. English edition: Geometry 1, Universitext, Springer Verlag.

[5] Marcel Berger and Bernard Gostiaux. *Géométrie différentielle: variétés, courbes et surfaces*. Collection Mathématiques. Puf, second edition, 1992. English edition: Differential geometry, manifolds, curves, and surfaces, GTM No. 115, Springer Verlag.

[6] William M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, second edition, 1986.

[7] Raoul Bott and Tu Loring W. *Differential Forms in Algebraic Topology*. GTM No. 82. Springer Verlag, first edition, 1986.

[8] Nicolas Bourbaki. *Elements of Mathematics. Lie Groups and Lie Algebras, Chapters 1–3*. Springer, first edition, 1989.

[9] Nicolas Bourbaki. *Topologie Générale, Chapitres 1-4*. Eléments de Mathématiques. Masson, 1990.

[10] Nicolas Bourbaki. *Topologie Générale, Chapitres 5-10*. Eléments de Mathématiques. CCLS, 1990.

[11] T. Bröcker and T. tom Dieck. *Representation of Compact Lie Groups*. GTM, Vol. 98. Springer Verlag, first edition, 1985.

[12] R.L. Bryant. An introduction to Lie groups and symplectic geometry. In D.S. Freed and K.K. Uhlenbeck, editors, *Geometry and Quantum Field Theory*, pages 5–181. AMS, Providence, Rhode Island, 1995.

[13] N. Burgoyne and R. Cushman. Conjugacy classes in linear groups. *Journal of Algebra*, 44:339–362, 1977.

[14] Roger Carter, Graeme Segal, and Ian Macdonald. *Lectures on Lie Groups and Lie Algebras*. Cambridge University Press, first edition, 1995.

[15] Shiing-shen Chern. *Complex Manifolds without Potential Theory*. Universitext. Springer Verlag, second edition, 1995.

[16] Claude Chevalley. *Theory of Lie Groups I*. Princeton Mathematical Series, No. 8. Princeton University Press, first edition, 1946. Eighth printing.

[17] Yvonne Choquet-Bruhat, Cécile DeWitt-Morette, and Margaret Dillard-Bleick. *Analysis, Manifolds, and Physics, Part I: Basics*. North-Holland, first edition, 1982.

[18] Morton L. Curtis. *Matrix Groups*. Universitext. Springer Verlag, second edition, 1984.

[19] James F. Davis and Paul Kirk. *Lecture Notes in Algebraic Topology*. GSM, Vol. 35. AMS, first edition, 2001.

[20] Dragomir Djokovic. On the exponential map in classical lie groups. *Journal of Algebra*, 64:76–88, 1980.

[21] Manfredo P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice Hall, 1976.

[22] Manfredo P. do Carmo. *Riemannian Geometry*. Birkhäuser, second edition, 1992.

[23] B.A. Dubrovin, A.T. Fomenko, and S.P. Novikov. *Modern Geometry–Methods and Applications. Part I*. GTM No. 93. Springer Verlag, second edition, 1985.

[24] B.A. Dubrovin, A.T. Fomenko, and S.P. Novikov. *Modern Geometry–Methods and Applications. Part II*. GTM No. 104. Springer Verlag, first edition, 1985.

[25] J.J. Duistermaat and J.A.C. Kolk. *Lie Groups*. Universitext. Springer Verlag, first edition, 2000.

[26] William Fulton and Joe Harris. *Representation Theory, A first course*. GTM No. 129. Springer Verlag, first edition, 1991.

[27] Jean H. Gallier. *Geometric Methods and Applications, For Computer Science and Engineering*. TAM, Vol. 38. Springer, first edition, 2000.

[28] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Universitext. Springer Verlag, second edition, 1993.

[29] Christopher Michael Geyer. *Catadioptric Projective Geometry: Theory and Applications*. PhD thesis, University of Pennsylvania, 200 South 33rd Street, Philadelphia, PA 19104, 2002. Dissertation.

[30] Phillip Griffiths and Joseph Harris. *Principles of Algebraic Geometry*. Wiley Interscience, first edition, 1978.

[31] Victor Guillemin and Alan Pollack. *Differential Topology*. Prentice Hall, first edition, 1974.

[32] Brian Hall. *Lie Groups, Lie Algebras, and Representations. An Elementary Introduction*. GTM No. 222. Springer Verlag, first edition, 2003.

[33] Morris W. Hirsch. *Differential Topology*. GTM No. 33. Springer Verlag, first edition, 1976.

[34] Friedrich Hirzebruch. *Topological Methods in Algebraic Geometry*. Springer Classics in Mathematics. Springer Verlag, second edition, 1978.

[35] Dale Husemoller. *Fiber Bundles*. GTM No. 20. Springer Verlag, third edition, 1994.

[36] Anthony W. Knapp. *Lie Groups Beyond an Introduction*. Progress in Mathematics, Vol. 140. Birkhäuser, second edition, 2002.

[37] Jacques Lafontaine. *Introduction Aux Variétés Différentielles*. PUG, first edition, 1996.

[38] Serge Lang. *Fundamentals of Differential Geometry*. GTM No. 191. Springer Verlag, first edition, 1999.

[39] Ib Madsen and Jorgen Tornehave. *From Calculus to Cohomology. De Rham Cohomology and Characteristic Classes*. Cambridge University Press, first edition, 1998.

[40] Jerrold E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry*. TAM, Vol. 17. Springer Verlag, first edition, 1994.

[41] John Milnor. On isometries of inner product spaces. *Inventiones Mathematicae*, 8:83–97, 1969.

[42] John W. Milnor. *Morse Theory*. Annals of Math. Series, No. 51. Princeton University Press, third edition, 1969.

[43] John W. Milnor and James D. Stasheff. *Characteristic Classes*. Annals of Math. Series, No. 76. Princeton University Press, first edition, 1974.

[44] R. Mneimné and F. Testard. *Introduction à la Théorie des Groupes de Lie Classiques*. Hermann, first edition, 1997.

[45] Shigeyuki Morita. *Geometry of Differential Forms*. Translations of Mathematical Monographs No 201. AMS, first edition, 2001.

[46] James R. Munkres. *Topology, a First Course*. Prentice Hall, first edition, 1975.

[47] Raghavan Narasimham. *Compact Riemann Surfaces*. Lecture in Mathematics ETH Zürich. Birkhäuser, first edition, 1992.

[48] Mitsuru Nishikawa. On the exponential map of the group $\mathbf{O}(p,q)_0$. *Memoirs of the Faculty of Science, Kyushu University, Ser. A*, 37:63–69, 1983.

[49] Barrett O'Neill. *Semi-Riemannian Geometry With Applications to Relativity*. Pure and Applies Math., Vol 103. Academic Press, first edition, 1983.

[50] L. Pontryagin. *Topological Groups*. Princeton University Press, first edition, 1939.

[51] L. Pontryagin. *Topological Groups*. Gordon and Breach, second edition, 1960.

[52] Marcel Riesz. *Clifford Numbers and Spinors*. Kluwer Academic Press, first edition, 1993. Edited by E. Folke Bolinder and Pertti Lounesto.

[53] Arthur A. Sagle and Ralph E. Walde. *Introduction to Lie Groups and Lie Algebras*. Academic Press, first edition, 1973.

[54] Hajime Sato. *Algebraic Topology: An Intuitive Approach*. Mathematical Monographs No 183. AMS, first edition, 1999.

[55] D.H. Sattinger and O.L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Applied Math. Science, Vol. 61. Springer Verlag, first edition, 1986.

[56] Laurent Schwartz. *Analyse II. Calcul Différentiel et Equations Différentielles*. Collection Enseignement des Sciences. Hermann, 1992.

[57] Richard W. Sharpe. *Differential Geometry. Cartan's Generalization of Klein's Erlangen Program*. GTM No. 166. Springer Verlag, first edition, 1997.

[58] Norman Steenrod. *The Topology of Fibre Bundles*. Princeton Math. Series, No. 14. Princeton University Press, 1956.

[59] Frank Warner. *Foundations of Differentiable Manifolds and Lie Groups*. GTM No. 94. Springer Verlag, first edition, 1983.

[60] André Weil. *L'Intégration dans les Groupes Topologiques et ses Applications*. Hermann, second edition, 1979.

[61] R.O. Wells. *Differential Analysis on Complex Manifolds*. GTM No. 65. Springer Verlag, second edition, 1980.