## Chapter 3

## Manifolds

### 3.1 Charts and Manifolds

In Chapter 1 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, $U_{\alpha}$, where each $U_{\alpha}$ 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 earth 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 [147], Berger and Gostiaux [17], O’Neill [119], Do Carmo [50, 49], Gallot, Hulin and Lafontaine [60], Lang [95], Schwartz [135], Hirsch [76], Sharpe [139], Guillemin and Pollack [69], Lafontaine [92], Dubrovin, Fomenko and Novikov [52] and Boothby [18]. A nice (not very technical) exposition is given in Morita [114] (Chapter 1). The recent book by Tu [145] is also highly recommended for its clarity. Among the many texts on manifolds and differential geometry, the book by Choquet-Bruhat, DeWitt-Morette and Dillard-Bleick [37] 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, $p r_{i}: \mathbb{R}^{n} \rightarrow \mathbb{R}$, are defined by

$$
p r_{i}\left(x_{1}, \ldots, x_{n}\right)=x_{i}, \quad 1 \leq i \leq n
$$

For technical reasons (in particular, to ensure 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, $(U, \varphi)$, where $U$ is an open subset of $M$ and $\varphi: U \rightarrow \Omega$ 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, $(U, \varphi)$, is a chart at $p$ iff $p \in U$. If $(U, \varphi)$ is a chart, then the functions $x_{i}=p r_{i} \circ \varphi$ are called local coordinates and for every $p \in U$, the tuple $\left(x_{1}(p), \ldots, x_{n}(p)\right)$ is the set of coordinates of $p$ w.r.t. the chart. The inverse, $\left(\Omega, \varphi^{-1}\right)$, of a chart is called a local parametrization. Given any two charts, $\left(U_{i}, \varphi_{i}\right)$ and $\left(U_{j}, \varphi_{j}\right)$, if $U_{i} \cap U_{j} \neq \emptyset$, we have the transition maps, $\varphi_{i}^{j}: \varphi_{i}\left(U_{i} \cap U_{j}\right) \rightarrow \varphi_{j}\left(U_{i} \cap U_{j}\right)$ and $\varphi_{j}^{i}: \varphi_{j}\left(U_{i} \cap U_{j}\right) \rightarrow \varphi_{i}\left(U_{i} \cap U_{j}\right)$, 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, $\varphi_{j}^{i}=\left(\varphi_{i}^{j}\right)^{-1}$. Observe that the transition maps $\varphi_{i}^{j}$ (resp. $\varphi_{j}^{i}$ ) 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, $M$, given some integer $n \geq 1$ and given some $k$ such that $k$ is either an integer $k \geq 1$ or $k=\infty$, a $C^{k} n$-atlas (or $n$-atlas of class $C^{k}$ ), $\mathcal{A}$, is a family of charts, $\left\{\left(U_{i}, \varphi_{i}\right)\right\}$, such that
(1) $\varphi_{i}\left(U_{i}\right) \subseteq \mathbb{R}^{n}$ 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 $\operatorname{map} \varphi_{i}^{j}\left(\right.$ and $\left.\varphi_{j}^{i}\right)$ is a $C^{k}$-diffeomorphism. When $k=\infty$, the $\varphi_{i}^{j}$ are smooth diffeomorphisms.

We must ensure 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, $(U, \varphi)$, we say that $(U, \varphi)$ 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 $\mathcal{A}^{\prime}$ 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 some integer $n \geq 1$ and given some $k$ such that $k$ is either an integer $k \geq 1$ or $k=\infty$, 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 $\varphi_{i}^{j}$ 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, $n_{\varphi}$, 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. (Brouwer's Invariance of Domain Theorem states that if $U \subseteq \mathbb{R}^{n}$ is an open set and if $f: U \rightarrow \mathbb{R}^{n}$ is a continuous and injective map, then $f(U)$ is open in $\mathbb{R}^{n}$. Using Brouwer's Theorem, we can show the following fact: If $U \subseteq \mathbb{R}^{m}$ and $V \subseteq \mathbb{R}^{n}$ are two open subsets and if $f: U \rightarrow V$ is a homeomorphism between $U$ and $V$, then $m=n$. If $m>n$, then consider the injection, $i: \mathbb{R}^{n} \rightarrow \mathbb{R}^{m}$, where $i(x)=\left(x, 0_{m-n}\right)$. Clearly, $i$ is injective and continuous, so $f \circ i: U \rightarrow i(V)$ is injective and continuous and Brouwer's Theorem implies that $i(V)$ is open in $\mathbb{R}^{m}$, which is a contradiction, as $i(V)=V \times\left\{0_{m-n}\right\}$ is not open in $\mathbb{R}^{m}$. If $m<n$, consider the homeomorphism $f^{-1}: V \rightarrow U$.)

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!).

In order to get a better grasp of the notion of manifold it is useful to consider examples of non-manifolds. First, consider the curve in $\mathbb{R}^{2}$ given by the zero locus of the equation

$$
y^{2}=x^{2}-x^{3}
$$



Figure 3.1: A nodal cubic; not a manifold
namely, the set of points

$$
M_{1}=\left\{(x, y) \in \mathbb{R}^{2} \mid y^{2}=x^{2}-x^{3}\right\} .
$$

This curve showed in Figure 3.1 and called a nodal cubic is also defined as the parametric curve

$$
\begin{aligned}
x & =1-t^{2} \\
y & =t\left(1-t^{2}\right)
\end{aligned}
$$

We claim that $M_{1}$ is not even a topological manifold. The problem is that the nodal cubic has a self-intersection at the origin. If $M_{1}$ was a topological manifold, then there would be a connected open subset, $U \subseteq M_{1}$, containing the origin, $O=(0,0)$, namely the intersection of a small enough open disc centered at $O$ with $M_{1}$, and a local chart, $\varphi: U \rightarrow \Omega$, where $\Omega$ is some connected open subset of $\mathbb{R}$ (that is, an open interval), since $\varphi$ is a homeomorphism. However, $U-\{O\}$ consists of four disconnected components and $\Omega-\varphi(O)$ of two disconnected components, contradicting the fact that $\varphi$ is a homeomorphism.

Let us now consider the curve in $\mathbb{R}^{2}$ given by the zero locus of the equation

$$
y^{2}=x^{3}
$$

namely, the set of points

$$
M_{2}=\left\{(x, y) \in \mathbb{R}^{2} \mid y^{2}=x^{3}\right\} .
$$

This curve showed in Figure 3.2 and called a cuspidal cubic is also defined as the parametric curve

$$
\begin{aligned}
& x=t^{2} \\
& y=t^{3}
\end{aligned}
$$

Consider the map, $\varphi: M_{2} \rightarrow \mathbb{R}$, given by

$$
\varphi(x, y)=y^{1 / 3}
$$



Figure 3.2: A Cuspidal Cubic

Since $x=y^{2 / 3}$ on $M_{2}$, we see that $\varphi^{-1}$ is given by

$$
\varphi^{-1}(t)=\left(t^{2}, t^{3}\right)
$$

and clearly, $\varphi$ is a homeomorphism, so $M_{2}$ is a topological manifold. However, in the altas consisting of the single chart, $\left\{\varphi: M_{2} \rightarrow \mathbb{R}\right\}$, the space $M_{2}$ is also a smooth manifold! Indeed, as there is a single chart, condition (3) of Definition 3.2 holds vacuously.

This fact is somewhat unexpected because the cuspidal cubic is usually not considered smooth at the origin, since the tangent vector of the parametric curve, $c: t \mapsto\left(t^{2}, t^{3}\right)$, at the origin is the zero vector (the velocity vector at $t$, is $\left.c^{\prime}(t)=\left(2 t, 3 t^{2}\right)\right)$. However, this apparent paradox has to do with the fact that, as a parametric curve, $M_{2}$ is not immersed in $\mathbb{R}^{2}$ since $c^{\prime}$ is not injective (see Definition 3.20 (a)), whereas as an abstract manifold, with this single chart, $M_{2}$ is diffeomorphic to $\mathbb{R}$.

Now, we also have the chart, $\psi: M_{2} \rightarrow \mathbb{R}$, given by

$$
\psi(x, y)=y
$$

with $\psi^{-1}$ given by

$$
\psi^{-1}(u)=\left(u^{2 / 3}, u\right)
$$

Then, observe that

$$
\varphi \circ \psi^{-1}(u)=u^{1 / 3}
$$

a map that is not differentiable at $u=0$. Therefore, the atlas $\left\{\varphi: M_{2} \rightarrow \mathbb{R}, \psi: M_{2} \rightarrow \mathbb{R}\right\}$ is not $C^{1}$ and thus, with respect to that atlas, $M_{2}$ is not a $C^{1}$-manifold.

The example of the cuspidal cubic shows a peculiarity of the definition of a $C^{k}$ (or $C^{\infty}$ ) manifold: If a space, $M$, happens to be a topological manifold because it has an atlas consisting of a single chart, then it is automatically a smooth manifold! In particular, if $f: U \rightarrow \mathbb{R}^{m}$ is any continuous function from some open subset, $U$, of $\mathbb{R}^{n}$, to $\mathbb{R}^{m}$, then the graph, $\Gamma(f) \subseteq \mathbb{R}^{n+m}$, of $f$ given by

$$
\Gamma(f)=\left\{(x, f(x)) \in \mathbb{R}^{n+m} \mid x \in U\right\}
$$

is a smooth manifold with respect to the atlas consisting of the single chart, $\varphi: \Gamma(f) \rightarrow U$, given by

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

with its inverse, $\varphi^{-1}: U \rightarrow \Gamma(f)$, given by

$$
\varphi^{-1}(x)=(x, f(x)) .
$$

The notion of a submanifold using the concept of "adapted chart" (see Definition 3.19 in Section 3.4) gives a more satisfactory treatment of $C^{k}$ (or smooth) submanifolds of $\mathbb{R}^{n}$. The example of the cuspidal cubic also shows clearly that whether a topological space is a $C^{k}$-manifold or a smooth manifold depends on the choice of atlas.

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, $M$, given some integer $n \geq 1$ and given some $k$ such that $k$ is either an integer $k \geq 1$ or $k=\infty$, a $C^{k} n$-atlas (or $n$-atlas of class $C^{k}$ ), $\mathcal{A}$, is a family of charts, $\left\{\left(U_{i}, \varphi_{i}\right)\right\}$, such that
(1) Each $U_{i}$ is a subset of $M$ and $\varphi_{i}: U_{i} \rightarrow \varphi_{i}\left(U_{i}\right)$ is a bijection onto an open subset, $\varphi_{i}\left(U_{i}\right) \subseteq \mathbb{R}^{n}$, 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 sets $\varphi_{i}\left(U_{i} \cap U_{j}\right)$ and $\varphi_{j}\left(U_{i} \cap U_{j}\right)$ are open in $\mathbb{R}^{n}$ and the transition maps $\varphi_{i}^{j}$ and $\varphi_{j}^{i}$ are $C^{k}$-diffeomorphisms.

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 $\varphi_{i}: U_{i} \rightarrow \varphi_{i}\left(U_{i}\right)$ are homeomorphisms, so $M$ is a manifold according to Definition 3.3. 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. A sufficient condition of $M$ to be Hausdorff is that for all $p, q \in M$ with $p \neq q$, either $p, q \in U_{i}$ for some $U_{i}$ or $p \in U_{i}$ and $q \in U_{j}$ for some disjoint $U_{i}, U_{j}$. 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 of the atlases, $\mathcal{A}$, defining $M$ (not only the maximal one) by requiring $U \subseteq M$ to be open iff $\varphi_{i}\left(U \cap U_{i}\right)$ is open in $\mathbb{R}^{n}$, for every
chart, $\left(U_{i}, \varphi_{i}\right)$, 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 [17], Chapter 2.

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 $(U, \varphi)$ is a chart in $\mathcal{A}$, for any (nonempty) open subset, $V \subseteq U$, we get a chart, $(V, \varphi \upharpoonright V)$, and it is obvious that this chart is compatible with $\mathcal{A}$. Thus, $(V, \varphi \upharpoonright V)$ 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 $\sigma_{N}: S^{n}-\{N\} \rightarrow \mathbb{R}^{n}$ and $\sigma_{S}: S^{n}-\{S\} \rightarrow \mathbb{R}^{n}$, 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

$$
\sigma_{N}\left(x_{1}, \ldots, x_{n+1}\right)=\frac{1}{1-x_{n+1}}\left(x_{1}, \ldots, x_{n}\right) \quad \text { and } \quad \sigma_{S}\left(x_{1}, \ldots, x_{n+1}\right)=\frac{1}{1+x_{n+1}}\left(x_{1}, \ldots, x_{n}\right)
$$

The inverse stereographic projections are given by

$$
\sigma_{N}^{-1}\left(x_{1}, \ldots, x_{n}\right)=\frac{1}{\left(\sum_{i=1}^{n} x_{i}^{2}\right)+1}\left(2 x_{1}, \ldots, 2 x_{n},\left(\sum_{i=1}^{n} x_{i}^{2}\right)-1\right)
$$

and

$$
\sigma_{S}^{-1}\left(x_{1}, \ldots, x_{n}\right)=\frac{1}{\left(\sum_{i=1}^{n} x_{i}^{2}\right)+1}\left(2 x_{1}, \ldots, 2 x_{n},-\left(\sum_{i=1}^{n} x_{i}^{2}\right)+1\right)
$$

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

$$
\left(x_{1}, \ldots, x_{n}\right) \mapsto \frac{1}{\sum_{i=1}^{n} x_{i}^{2}}\left(x_{1}, \ldots, x_{n}\right)
$$

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 $\left(U_{N}, \sigma_{N}\right)$ and $\left(U_{S}, \sigma_{S}\right)$ form a smooth atlas for $S^{n}$.

Example 2. The projective space $\mathbb{R}^{\mathbb{P}^{n}}$.
To define an atlas on $\mathbb{R P}^{n}$ it is convenient to view $\mathbb{R P}^{n}$ 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 $p=\left[x_{1}, \ldots, x_{n+1}\right] \in \mathbb{R}^{p}$, we call $\left(x_{1}, \ldots, x_{n+1}\right)$ the homogeneous coordinates of $p$. It is customary to write $\left(x_{1}: \cdots: x_{n+1}\right)$ instead of $\left[x_{1}, \ldots, x_{n+1}\right]$. (Actually, in most books, the indexing starts with 0 , i.e., homogeneous coordinates for $\mathbb{R P}^{n}$ are written as $\left(x_{0}: \cdots: x_{n}\right)$.) For any $i$, with $1 \leq i \leq n+1$, let

$$
U_{i}=\left\{\left(x_{1}: \cdots: x_{n+1}\right) \in \mathbb{R}^{n} \mid x_{i} \neq 0\right\} .
$$

Observe that $U_{i}$ is well defined, because if $\left(y_{1}: \cdots: y_{n+1}\right)=\left(x_{1}: \cdots: x_{n+1}\right)$, then there is some $\lambda \neq 0$ so that $y_{j}=\lambda x_{j}$, for $j=1, \ldots, n+1$. We can define a homeomorphism, $\varphi_{i}$, of $U_{i}$ onto $\mathbb{R}^{n}$, as follows:

$$
\varphi_{i}\left(x_{1}: \cdots: x_{n+1}\right)=\left(\frac{x_{1}}{x_{i}}, \ldots, \frac{x_{i-1}}{x_{i}}, \frac{x_{i+1}}{x_{i}}, \ldots, \frac{x_{n+1}}{x_{i}}\right),
$$

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

$$
\psi_{i}\left(x_{1}, \ldots, x_{n}\right)=\left(x_{1}: \cdots: x_{i-1}: 1: x_{i}: \cdots: x_{n}\right)
$$

where the 1 goes in the $i$ th slot, for $i=1, \ldots, n+1$. One easily checks that $\varphi_{i}$ and $\psi_{i}$ are mutual inverses, so the $\varphi_{i}$ are homeomorphisms. On the overlap, $U_{i} \cap U_{j}$, (where $i \neq j$ ), as $x_{j} \neq 0$, we have

$$
\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)\left(x_{1}, \ldots, x_{n}\right)=\left(\frac{x_{1}}{x_{j}}, \ldots, \frac{x_{i-1}}{x_{j}}, \frac{1}{x_{j}}, \frac{x_{i}}{x_{j}}, \ldots, \frac{x_{j-1}}{x_{j}}, \frac{x_{j+1}}{x_{j}}, \ldots, \frac{x_{n}}{x_{j}}\right) .
$$

(We assumed that $i<j$; the case $j<i$ is similar.) This is clearly a smooth function from $\varphi_{i}\left(U_{i} \cap U_{j}\right)$ to $\varphi_{j}\left(U_{i} \cap U_{j}\right)$. As the $U_{i}$ cover $\mathbb{R}^{P} \mathbb{P}^{n}$, we conclude that the $\left(U_{i}, \varphi_{i}\right)$ are $n+1$ charts making a smooth atlas for $\mathbb{R} \mathbb{P}^{n}$. Intuitively, the space $\mathbb{R} \mathbb{P}^{n}$ 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, $\Lambda=\left(\lambda_{i j}\right)$, such that

$$
v_{j}=\sum_{i=1}^{k} \lambda_{i j} u_{i}, \quad 1 \leq j \leq k
$$

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, $S=\left\{i_{1}, \ldots, i_{k}\right\}$ of $\{1, \ldots, n\}$, let $U_{S}$ be the subset of $n \times k$ matrices, $A$, of rank $k$ whose rows of index in $S=\left\{i_{1}, \ldots, i_{k}\right\}$ form an invertible $k \times k$ matrix denoted $A_{S}$. Observe that the $k \times k$ matrix consisting of the rows of the matrix $A A_{S}^{-1}$ whose index belong to $S$ is the identity matrix, $I_{k}$. Therefore, we can define a map, $\varphi_{S}: U_{S} \rightarrow \mathbb{R}^{(n-k) \times k}$, where $\varphi_{S}(A)$ is equal to the $(n-k) \times k$ matrix obtained by deleting the rows of index in $S$ from $A A_{S}^{-1}$.

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 $\Lambda$ invertible, if we write

$$
A=\binom{A_{1}}{A_{2}} \quad \text { and } \quad B=\binom{B_{1}}{B_{2}}
$$

as $B=A \Lambda$, we get $B_{1}=A_{1} \Lambda$ and $B_{2}=A_{2} \Lambda$, from which we deduce that

$$
\binom{B_{1}}{B_{2}} B_{1}^{-1}=\binom{I_{k}}{B_{2} B_{1}^{-1}}=\binom{I_{k}}{A_{2} \Lambda \Lambda^{-1} A_{1}^{-1}}=\binom{I_{k}}{A_{2} A_{1}^{-1}}=\binom{A_{1}}{A_{2}} A_{1}^{-1} .
$$

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

$$
\psi_{S}(M)=P_{S}\binom{I_{k}}{M}
$$

The effect of $\psi_{S}$ 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 $\varphi_{S}\left(U_{S} \cap U_{U}\right)$ to $\varphi_{T}\left(U_{S} \cap U_{U}\right)$ is given by

$$
M \mapsto(C+D M)(A+B M)^{-1}
$$

where

$$
\left(\begin{array}{ll}
A & B \\
C & D
\end{array}\right)=P_{T} P_{S},
$$

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 $\left(U_{S}, \varphi_{S}\right)$ 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 the set of $k$-planes, $W$, supplementary to the $(n-k)$ plane spanned by the canonical basis vectors $e_{j_{k+1}}, \ldots, e_{j_{n}}$ (i.e., $\operatorname{span}\left(W \cup\left\{e_{j_{k+1}}, \ldots, e_{j_{n}}\right\}\right)=$ $\mathbb{R}^{n}$, where $S=\left\{i_{1}, \ldots, i_{k}\right\}$ and $\left.\left\{j_{k+1}, \ldots, j_{n}\right\}=\{1, \ldots, n\}-S\right)$.

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, $\left(U_{i}, \varphi_{i}\right)$ on $M_{1}$ and $\left(V_{j}, \psi_{j}\right)$ on $M_{2}$, we declare that $\left(U_{i} \times V_{j}, \varphi_{i} \times \psi_{j}\right)$ is a chart on $M_{1} \times M_{2}$, where $\varphi_{i} \times \psi_{j}: U_{i} \times V_{j} \rightarrow \mathbb{R}^{n_{1}+n_{2}}$ is defined so that

$$
\varphi_{i} \times \psi_{j}(p, q)=\left(\varphi_{i}(p), \psi_{j}(q)\right), \quad \text { for all }(p, q) \in U_{i} \times V_{j}
$$

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 is a continuous function, $h: M \rightarrow N$, satisfying the following property: For every $p \in M$, there is some chart, $(U, \varphi)$, at $p$ and some chart, $(V, \psi)$, at $q=h(p)$, with $f(U) \subseteq V$ and

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

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, $f: M \rightarrow \mathbb{R}$, is a $C^{k}$-map iff for every $p \in M$, there is some chart, $(U, \varphi)$, at $p$ so that

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

is a $C^{k}$-function. If $U$ is an open subset of $M$, the 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 $\gamma:] a, b[\rightarrow N$ is called a $C^{k}$-curve in $N$. One checks immediately that a function, $\left.\gamma:\right] a, b\left[\rightarrow N\right.$, is a $C^{k}$-map iff for every $q \in N$, there is some chart, $(V, \psi)$, at $q$ so that

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

is a $C^{k}$-function.
It is clear that the composition of $C^{k}$-maps is a $C^{k}$-map. A $C^{k}$-map, $h: M \rightarrow N$, between two manifolds is a $C^{k}$-diffeomorphism iff $h$ has an inverse, $h^{-1}: N \rightarrow M$ (i.e., $h^{-1} \circ h=\operatorname{id}_{M}$ and $h \circ h^{-1}=\operatorname{id}_{N}$ ), 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 $\gamma:]-\epsilon, \epsilon[\rightarrow M$ 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 $\gamma^{\prime}(0)$ does not make sense. However, for any chart, $(U, \varphi)$, at $p$, the map $\varphi \circ \gamma$ is a $C^{1}$-curve in $\mathbb{R}^{n}$ and the tangent vector $v=(\varphi \circ \gamma)^{\prime}(0)$ 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, $\left.\gamma_{1}:\right]-\epsilon_{1}, \epsilon_{1}\left[\rightarrow M\right.$ and $\left.\gamma_{2}:\right]-\epsilon_{2}, \epsilon_{2}\left[\rightarrow M\right.$, through $p$ (i.e., $\gamma_{1}(0)=\gamma_{2}(0)=p$ ) are equivalent iff there is some chart, $(U, \varphi)$, at $p$ so that

$$
\left(\varphi \circ \gamma_{1}\right)^{\prime}(0)=\left(\varphi \circ \gamma_{2}\right)^{\prime}(0)
$$

Now, the problem is that this definition seems to depend on the choice of the chart. Fortunately, this is not the case. For, if $(V, \psi)$ 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

$$
\begin{aligned}
\left(\psi \circ \gamma_{1}\right)^{\prime}(0) & =\left(\eta \circ \varphi \circ \gamma_{1}\right)^{\prime}(0) \\
& =\eta^{\prime}(\varphi(p))\left(\left(\varphi \circ \gamma_{1}\right)^{\prime}(0)\right) \\
& =\eta^{\prime}(\varphi(p))\left(\left(\varphi \circ \gamma_{2}\right)^{\prime}(0)\right) \\
& =\left(\eta \circ \varphi \circ \gamma_{2}\right)^{\prime}(0) \\
& =\left(\psi \circ \gamma_{2}\right)^{\prime}(0) .
\end{aligned}
$$

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 $\left.T_{p} M\right)$.

It is obvious that $T_{p}(M)$ is a vector space. If $u, v \in T_{p}(M)$ are defined by the curves $\gamma_{1}$ and $\gamma_{2}$, then $u+v$ is defined by the curve $\gamma_{1}+\gamma_{2}$ (we may assume by reparametrization that $\gamma_{1}$ and $\gamma_{2}$ have the same domain.) Similarly, if $u \in T_{p}(M)$ is defined by a curve $\gamma$ and $\lambda \in \mathbb{R}$, then $\lambda u$ 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 [135] (Chapter 3, Section 9) but also by Warner [147].

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

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p} f=\left.\frac{\partial\left(f \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}, \quad 1 \leq i \leq n
$$

(Here, $\left.\left(\partial g / \partial X_{i}\right)\right|_{y}$ denotes the partial derivative of a function $g: \mathbb{R}^{n} \rightarrow \mathbb{R}$ with respect to the $i$ th 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 not 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

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

However, we have to check that these definitions make sense, that is, that they don't depend on the choice of representatives chosen in the equivalence classes $[f]$ and $[g]$. Let us give the details of this verification for the sum of two germs, $[f]$ and $[g]$. For any two locally defined functions, $(f, U)$ and $(g, V)$, at $p$, let $f+g$ be the locally defined function at $p$ with domain $U \cap V$ and such that $(f+g)(x)=f(x)+g(x)$ for all $x \in U \cap V$. We need to check that for any locally defined functions $\left(U_{1}, f_{1}\right),\left(U_{2}, f_{2}\right),\left(V_{1}, g_{1}\right)$, and $\left(V_{2}, g_{2}\right)$, at $p$, if $\left(U_{1}, f_{1}\right)$ and $\left(U_{2}, f_{2}\right)$ are equivalent and if $\left(V_{1}, g_{1}\right)$ and $\left(V_{2}, g_{2}\right)$ are equivalent, then $\left(U_{1} \cap V_{1}, f_{1}+g_{1}\right)$ and $\left(U_{2} \cap V_{2}, f_{2}+g_{2}\right)$ are equivalent. However, as $\left(U_{1}, f_{1}\right)$ and $\left(U_{2}, f_{2}\right)$ are equivalent, there is some $W_{1} \subseteq U_{1} \cap U_{2}$ so that $f_{1} \upharpoonright W_{1}=f_{2} \upharpoonright W_{1}$ and as ( $V_{1}, g_{1}$ ) and $\left(V_{2}, g_{2}\right)$ are equivalent, there is some $W_{2} \subseteq V_{1} \cap V_{2}$ so that $g_{1} \upharpoonright W_{2}=g_{2} \upharpoonright W_{2}$. Then, observe that $\left(f_{1}+g_{1}\right) \upharpoonright\left(W_{1} \cap W_{2}\right)=\left(f_{2}+g_{2}\right) \upharpoonright\left(W_{1} \cap W_{2}\right)$, which means that $\left[f_{1}+g_{1}\right]=\left[f_{2}+g_{2}\right]$. Therefore, $[f+g]$ does not depend on the representatives chosen in the equivalence classes $[f]$ and $[g]$ and it makes sense to set

$$
[f]+[g]=[f+g] .
$$

We can proceed in a similar fashion to define $\lambda[f]$ and $[f][g]$. Therefore, the germs at $p$ form a ring. The ring of germs of $C^{k}$-functions at $p$ is denoted $\mathcal{O}_{M, p}^{(k)}$. When $k=\infty$, we usually drop the superscript $\infty$.

Remark: Most readers will most likely be puzzled by the notation $\mathcal{O}_{M, p}^{(k)}$. 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 $\mathcal{O}_{M}^{(k)}(U)$ (certainly by people with an algebraic geometry bent!). Then, it turns out that the map $U \mapsto \mathcal{O}_{M}^{(k)}(U)$ is a sheaf, denoted $\mathcal{O}_{M}^{(k)}$, and the ring $\mathcal{O}_{M, p}^{(k)}$ is the stalk of the sheaf $\mathcal{O}_{M}^{(k)}$ at $p$. Such rings are called local rings. Roughly speaking, all the "local" information about $M$ at $p$ is contained in the local ring $\mathcal{O}_{M, p}^{(k)}$. (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

$$
v_{i}: f \mapsto\left(\frac{\partial}{\partial x_{i}}\right)_{p} f
$$

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

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p}(f g)=f(p)\left(\frac{\partial}{\partial x_{i}}\right)_{p} g+g(p)\left(\frac{\partial}{\partial x_{i}}\right)_{p} f .
$$

Secondly, if $\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))=0$, 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 $\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))=0$, 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, $(U, \varphi)$, 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 $\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))=0$ for some chart, $(U, \varphi)$, at $p$ and some function, $f$, in the germ, $\mathbf{f}$. The $C^{k}$-stationary germs form a subring of $\mathcal{O}_{M, p}^{(k)}$ (but not an ideal!) denoted $\mathcal{S}_{M, p}^{(k)}$.

Remarkably, it turns out that the dual of the vector space, $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$, is isomorphic to the tangent space, $T_{p}(M)$. First, we prove that the subspace of linear forms on $\mathcal{O}_{M, p}^{(k)}$ that vanish on $\mathcal{S}_{M, p}^{(k)}$ 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 $(U, \varphi)$ at $p$, the $n$ functions, $\left(\frac{\partial}{\partial x_{1}}\right)_{p}, \ldots,\left(\frac{\partial}{\partial x_{n}}\right)_{p}$, defined on $\mathcal{O}_{M, p}^{(k)}$ by

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p} f=\left.\frac{\partial\left(f \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}, \quad 1 \leq i \leq n
$$

are linear forms that vanish on $\mathcal{S}_{M, p}^{(k)}$. Every linear form, $L$, on $\mathcal{O}_{M, p}^{(k)}$ that vanishes on $\mathcal{S}_{M, p}^{(k)}$ 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

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p}, \quad i=1, \ldots, n
$$

form a basis of the vector space of linear forms on $\mathcal{O}_{M, p}^{(k)}$ that vanish on $\mathcal{S}_{M, p}^{(k)}$.
Proof. The first part of the proposition is trivial, by definition of $\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))$ and of $\left(\frac{\partial}{\partial x_{i}}\right)_{p} f$.

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

$$
g(q)=f(q)-\sum_{i=1}^{n}\left(p r_{i} \circ \varphi\right)(q)\left(\frac{\partial}{\partial x_{i}}\right)_{p} f
$$

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

$$
\begin{aligned}
g(q)=\left(g \circ \varphi^{-1}\right)(\varphi(q)) & =\left(f \circ \varphi^{-1}\right)(\varphi(q))-\sum_{i=1}^{n}\left(p r_{i} \circ \varphi\right)(q)\left(\frac{\partial}{\partial x_{i}}\right)_{p} f \\
& =\left(f \circ \varphi^{-1}\right)\left(X_{1}(q) \ldots, X_{n}(q)\right)-\sum_{i=1}^{n} X_{i}(q)\left(\frac{\partial}{\partial x_{i}}\right)_{p} f
\end{aligned}
$$

with $X_{i}(q)=\left(p r_{i} \circ \varphi\right)(q)$. It follows that

$$
\left.\frac{\partial\left(g \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}=\left.\frac{\partial\left(f \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}-\left(\frac{\partial}{\partial x_{i}}\right)_{p} f=0
$$

But then, as $L$ vanishes on stationary germs, we get

$$
L(f)=\sum_{i=1}^{n} L\left(p r_{i} \circ \varphi\right)\left(\frac{\partial}{\partial x_{i}}\right)_{p} f
$$

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 $x_{i}=p r_{i} \circ \varphi$, as

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p} x_{j}=\delta_{i j}
$$

we get $\lambda_{i}=0$, for $i=1, \ldots, n$.
As the subspace of linear forms on $\mathcal{O}_{M, p}^{(k)}$ that vanish on $\mathcal{S}_{M, p}^{(k)}$ is isomorphic to the dual, $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{*}$, of the space $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$, we see that the

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p}, \quad i=1, \ldots, n
$$

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

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

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

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p}, \quad i=1, \ldots, n
$$

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.

Proposition 3.2 says that a linear form on $\mathcal{O}_{M, p}^{(k)}$ that vanishes on $\mathcal{S}_{M, p}^{(k)}$ is a linear derivation but in general, when $k \neq \infty$, a linear derivation on $\mathcal{O}_{M, p}^{(k)}$ does not necessarily vanish on $\mathcal{S}_{M, p}^{(k)}$. However, we will see in Proposition 3.6 that this is true for $k=\infty$.

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 $\left(\frac{\partial}{\partial t}\right)_{x}=\left.\frac{d}{d t}\right|_{x}$ is a basis vector of $T_{x}(\mathbb{R})$. For every $C^{k}$-function, $f$, locally defined at $x$, we have

$$
\left(\frac{\partial}{\partial t}\right)_{x} f=\left.\frac{d f}{d t}\right|_{x}=f^{\prime}(x) .
$$

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, $\gamma:]-\epsilon,+\epsilon\left[\rightarrow M\right.$, through $p$ (i.e., $p=\gamma(0)$ ). Then, the map $L_{u}$ defined on $\mathcal{O}_{M, p}^{(k)}$ by

$$
L_{u}(\mathbf{f})=(f \circ \gamma)^{\prime}(0)
$$

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 $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{*}$, 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 $\gamma$ and $\sigma$ are equivalent curves defining $u$, then $(\varphi \circ \sigma)^{\prime}(0)=(\varphi \circ \gamma)^{\prime}(0)$, so we get

$$
(f \circ \sigma)^{\prime}(0)=\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))\left((\varphi \circ \sigma)^{\prime}(0)\right)=\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))\left((\varphi \circ \gamma)^{\prime}(0)\right)=(f \circ \gamma)^{\prime}(0),
$$

which shows that $L_{u}(\mathbf{f})$ does not depend on the curve, $\gamma$, defining $u$. If $\mathbf{f}$ is a stationary germ, then pick any chart, $(U, \varphi)$, at $p$ and let $\psi=\varphi \circ \gamma$. We have

$$
L_{u}(\mathbf{f})=(f \circ \gamma)^{\prime}(0)=\left(\left(f \circ \varphi^{-1}\right) \circ(\varphi \circ \gamma)\right)^{\prime}(0)=\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))\left(\psi^{\prime}(0)\right)=0
$$

since $\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))=0$, 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 $\gamma$ and $\sigma$ through $p$ so that

$$
(\varphi \circ \gamma)^{\prime}(0) \neq(\varphi \circ \sigma)^{\prime}(0)
$$

Thus, there is some $i$, with $1 \leq i \leq n$, so that if we let $f=p r_{i} \circ \varphi$, then

$$
(f \circ \gamma)^{\prime}(0) \neq(f \circ \sigma)^{\prime}(0),
$$

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 $\mathcal{O}_{M, p}^{(k)}$ that vanishes on $\mathcal{S}_{M, p}^{(k)}$ can be uniquely expressed as

$$
L=\sum_{i=1}^{n} \lambda_{i}\left(\frac{\partial}{\partial x_{i}}\right)_{p}
$$

Define the curve, $\gamma$, on $M$ through $p$ by

$$
\gamma(t)=\varphi^{-1}\left(\varphi(p)+t\left(\lambda_{1}, \ldots, \lambda_{n}\right)\right)
$$

for $t$ in a small open interval containing 0 . Then, we have

$$
f(\gamma(t))=\left(f \circ \varphi^{-1}\right)\left(\varphi(p)+t\left(\lambda_{1}, \ldots, \lambda_{n}\right)\right),
$$

and we get

$$
(f \circ \gamma)^{\prime}(0)=\left(f \circ \varphi^{-1}\right)^{\prime}(\varphi(p))\left(\lambda_{1}, \ldots, \lambda_{n}\right)=\left.\sum_{i=1}^{n} \lambda_{i} \frac{\partial\left(f \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}=L(\mathbf{f})
$$

This proves that $T_{p}(M)$ and $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{*}$ are isomorphic.
In view of Proposition 3.3, we can identify $T_{p}(M)$ with $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{*}$. As the space $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$ is finite dimensional, $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{* *}$ is canonically isomorphic to $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$, so we can identify $T_{p}^{*}(M)$ with $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$. (Recall that if $E$ is a finite dimensional space, the map $i_{E}: E \rightarrow E^{* *}$ 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 $T_{p}(M)$ is the space of linear derivations on $\mathcal{O}_{M, p}^{(k)}$ that vanish on $\mathcal{S}_{M, p}^{(k)}$. Thus, $T_{p}(M)$ can be identified with $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{*}$. The space $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$ is called the cotangent space at $p$; it is isomorphic to the dual, $T_{p}^{*}(M)$, of $T_{p}(M)$. (For simplicity of notation we also denote $T_{p}(M)$ by $T_{p} M$ and $T_{p}^{*}(M)$ by $T_{p}^{*} M$.)

Even though this is just a restatement of Proposition 3.1, we state the following proposition because of its practical usefulness:

Proposition 3.4 Given any $C^{k}$-manifold, $M$, of dimension $n$, with $k \geq 1$, for any $p \in M$ and any chart $(U, \varphi)$ at $p$, the $n$ tangent vectors,

$$
\left(\frac{\partial}{\partial x_{1}}\right)_{p}, \ldots,\left(\frac{\partial}{\partial x_{n}}\right)_{p}
$$

form a basis of $T_{p} M$.

Observe that if $x_{i}=p r_{i} \circ \varphi$, as

$$
\left(\frac{\partial}{\partial x_{i}}\right)_{p} x_{j}=\delta_{i, j},
$$

the images of $x_{1}, \ldots, x_{n}$ in $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$ form the dual basis 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 $T_{p}^{*}(M)=\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$ by $d f_{p}$. This is the differential of $f$ at $p$. Using the isomorphism between $\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}$ and $\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{* *}$ described above, $d f_{p}$ corresponds to the linear map in $T_{p}^{*}(M)$ defined by $d f_{p}(v)=v(\mathbf{f})$, for all $v \in T_{p}(M)$. With this notation, we see that $\left(d x_{1}\right)_{p}, \ldots,\left(d x_{n}\right)_{p}$ 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, $\mathcal{O}_{M, p}^{(k)}$, 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, $(U, \varphi)$, 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 $v$. The directional derivative, $v(f)$, is also denoted $v[f]$.

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 [147]):

Proposition 3.5 If $g: \mathbb{R}^{n} \rightarrow \mathbb{R}$ 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

$$
g(q)=g(p)+\left.\sum_{i=1}^{n} \frac{\partial g}{\partial X_{i}}\right|_{p}\left(q_{i}-p_{i}\right)+\left.\sum_{i, j=1}^{n}\left(q_{i}-p_{i}\right)\left(q_{j}-p_{j}\right) \int_{0}^{1}(1-t) \frac{\partial^{2} g}{\partial X_{i} \partial X_{j}}\right|_{(1-t) p+t q} d t .
$$

In particular, if $g \in C^{\infty}(U)$, then the integral as a function of $q$ is $C^{\infty}$.
Proposition 3.6 Let $M$ be any $C^{\infty}$-manifold of dimension $n$. For any $p \in M$, any linear derivation on $\mathcal{O}_{M, p}^{(\infty)}$ vanishes on $\mathcal{S}_{M, p}^{(\infty)}$, the ring of stationary germs.

Proof. Pick some chart, $(U, \varphi)$, at $p$, where $U$ is convex (for instance, an open ball) and let f be any stationary germ. If we apply Proposition 3.5 to $f \circ \varphi^{-1}$ and then compose with $\varphi$, we get

$$
f=f(p)+\left.\sum_{i=1}^{n} \frac{\partial\left(f \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}\left(x_{i}-x_{i}(p)\right)+\sum_{i, j=1}^{n}\left(x_{i}-x_{i}(p)\right)\left(x_{j}-x_{j}(p)\right) h
$$

near $p$, where $h$ is $C^{\infty}$. Since $\mathbf{f}$ is a stationary germ, this yields

$$
f=f(p)+\sum_{i, j=1}^{n}\left(x_{i}-x_{i}(p)\right)\left(x_{j}-x_{j}(p)\right) h .
$$

If $v$ is any linear derivation, we get

$$
\begin{aligned}
v(f)=v(f(p))+ & \sum_{i, j=1}^{n}\left[\left(x_{i}-x_{i}(p)\right)(p)\left(x_{j}-x_{j}(p)\right)(p) v(h)\right. \\
& \left.+\left(x_{i}-x_{i}(p)\right)(p) v\left(x_{j}-x_{j}(p)\right) h(p)+v\left(x_{i}-x_{i}(p)\right)\left(x_{j}-x_{j}(p)\right)(p) h(p)\right]=0 .
\end{aligned}
$$

Thus, $v$ vanishes on stationary germs.
Proposition 3.6 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 $\mathcal{O}_{M, p}^{(\infty)}$. Let $\mathfrak{m}_{M, p} \subseteq \mathcal{O}_{M, p}^{(\infty)}$ be the ideal of germs that vanish at $p$. Then, we also have the ideal $\mathfrak{m}_{M, p}^{2}$, 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 $\mathfrak{m}_{M, p} / \mathfrak{m}_{M, p}^{2}$ (see Warner [147], Lemma 1.16).

Actually, if we let $\mathfrak{m}_{M, p}^{(k)}$ denote the $C^{k}$ germs that vanish at $p$ and $\mathfrak{s}_{M, p}^{(k)}$ denote the stationary $C^{k}$-germs that vanish at $p$, it is easy to show that

$$
\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)} \cong \mathfrak{m}_{M, p}^{(k)} / \mathfrak{s}_{M, p}^{(k)} .
$$

(Given any $\mathbf{f} \in \mathcal{O}_{M, p}^{(k)}$, send it to $\mathbf{f}-\mathbf{f}(\mathbf{p}) \in \mathfrak{m}_{M, p}^{(k)}$.) Clearly, $\left(\mathfrak{m}_{M, p}^{(k)}\right)^{2}$ consists of stationary germs (by the derivation property) and when $k=\infty$, Proposition 3.5 shows that every stationary germ that vanishes at $p$ belongs to $\mathfrak{m}_{M, p}^{2}$. Therefore, when $k=\infty$, we have $\mathfrak{s}_{M, p}^{(\infty)}=\mathfrak{m}_{M, p}^{2}$ and so,

$$
T_{p}^{*}(M)=\mathcal{O}_{M, p}^{(\infty)} / \mathcal{S}_{M, p}^{(\infty)} \cong \mathfrak{m}_{M, p} / \mathfrak{m}_{M, p}^{2}
$$

Remark: The ideal $\mathfrak{m}_{M, p}^{(k)}$ is in fact the unique maximal ideal of $\mathcal{O}_{M, p}^{(k)}$. This is because if $\mathbf{f} \in \mathcal{O}_{M, p}^{(k)}$ does not vanish at $p$, then it is an invertible element of $\mathcal{O}_{M, p}^{(k)}$ and any ideal containing $\mathfrak{m}_{M, p}^{(k)}$ and $\mathbf{f}$ would be equal to $\mathcal{O}_{M, p}^{(k)}$, which it absurd. Thus, $\mathcal{O}_{M, p}^{(k)}$ 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, $(U, \varphi, u)$, where $(U, \varphi)$ is any chart at $p$ and $u$ is any vector in $\mathbb{R}^{n}$. Say that two such triples $(U, \varphi, u)$ and $(V, \psi, v)$ are equivalent iff

$$
\left(\psi \circ \varphi^{-1}\right)_{\varphi(p)}^{\prime}(u)=v
$$

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 $(U, \varphi)$ is a chart on $M$ at $p$, we can define a natural isomorphism, $\theta_{U, \varphi, p}: \mathbb{R}^{n} \rightarrow T_{p}(M)$, between $\mathbb{R}^{n}$ and $T_{p}(M)$, as follows: For any $u \in \mathbb{R}^{n}$,

$$
\theta_{U, \varphi, p}: u \mapsto[(U, \varphi, u)] .
$$

One immediately checks 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 3.7 Let $M$ be any $C^{k}$-manifold of dimension $n$, with $k \geq 1$. For every $p \in M$, for every chart, $(U, \varphi)$, at $p$, if $x$ is any tangent vector (version 1) given by some equivalence class of $C^{1}$-curves, $\left.\gamma:\right]-\epsilon,+\epsilon[\rightarrow M$, through $p$ (i.e., $p=\gamma(0)$ ), then the map

$$
x \mapsto\left[\left(U, \varphi,(\varphi \circ \gamma)^{\prime}(0)\right)\right]
$$

is an isomorphism between $T_{p}(M)$-version 1 and $T_{p}(M)$-version 3.
Proof. If $\sigma$ is another curve equivalent to $\gamma$, then $(\varphi \circ \gamma)^{\prime}(0)=(\varphi \circ \sigma)^{\prime}(0)$, so the map is well-defined. It is clearly injective. As for surjectivity, define the curve, $\gamma$, on $M$ through $p$ by

$$
\gamma(t)=\varphi^{-1}(\varphi(p)+t u)
$$

Then, $(\varphi \circ \gamma)(t)=\varphi(p)+t u$ and

$$
(\varphi \circ \gamma)^{\prime}(0)=u
$$

After having explored thorougly the notion of tangent vector, we show how a $C^{k}$-map, $h: M \rightarrow N$, between $C^{k}$ manifolds, induces a linear map, $d h_{p}: T_{p}(M) \rightarrow T_{h(p)}(N)$, 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 $\mathcal{O}_{M, p}^{(k)}$ that vanishes on $\mathcal{S}_{M, p}^{(k)}$. We would like $d h_{p}(u)$ to be a linear derivation on $\mathcal{O}_{N, h(p)}^{(k)}$ that vanishes on $\mathcal{S}_{N, h(p)}^{(k)}$. So, for every germ, $\mathbf{g} \in \mathcal{O}_{N, h(p)}^{(k)}$, set

$$
d h_{p}(u)(\mathbf{g})=u(\mathbf{g} \circ \mathbf{h}) .
$$

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, $(V, \psi)$ on $N$ at $q=h(p)$, we have $\left(g \circ \psi^{-1}\right)^{\prime}(\psi(q))=0$ and, for any chart, $(U, \varphi)$, at $p$ on $M$, we get

$$
\left(g \circ h \circ \varphi^{-1}\right)^{\prime}(\varphi(p))=\left(g \circ \psi^{-1}\right)^{\prime}(\psi(q))\left(\left(\psi \circ h \circ \varphi^{-1}\right)^{\prime}(\varphi(p))\right)=0,
$$

which means that $\mathbf{g} \circ \mathbf{h}$ is stationary at $p$. Therefore, $d h_{p}(u) \in T_{h(p)}(M)$. It is also clear that $d h_{p}$ 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, $h: M \rightarrow N$, and for every $p \in M$, the differential of $h$ at $p$ or tangent map, $d h_{p}: T_{p}(M) \rightarrow T_{h(p)}(N)$, is the linear map defined so that

$$
d h_{p}(u)(\mathbf{g})=u(\mathbf{g} \circ \mathbf{h}),
$$

for every $u \in T_{p}(M)$ and every germ, $\mathbf{g} \in \mathcal{O}_{N, h(p)}^{(k)}$. The linear map $d h_{p}$ is also denoted $T_{p} h$ (and sometimes, $h_{p}^{\prime}$ or $D_{p} h$ ).

The chain rule is easily generalized to manifolds.

Proposition 3.8 Given any two $C^{k}$-maps $f: M \rightarrow N$ and $g: N \rightarrow P$ between smooth $C^{k}$ _manifolds, for any $p \in M$, we have

$$
d(g \circ f)_{p}=d g_{f(p)} \circ d f_{p}
$$

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 $d f_{p}$ is explicitly. Since $N=\mathbb{R}$, germs (of functions on $\mathbb{R}$ ) at $t_{0}=f(p)$ are just germs of $C^{k}$-functions, $g: \mathbb{R} \rightarrow \mathbb{R}$, locally defined at $t_{0}$. Then, for any $u \in T_{p}(M)$ and every germ $\mathbf{g}$ at $t_{0}$,

$$
d f_{p}(u)(\mathbf{g})=u(\mathbf{g} \circ \mathbf{f})
$$

If we pick a chart, $(U, \varphi)$, 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 $d f_{p}(u)(\mathbf{g})$ is when $u=\left(\frac{\partial}{\partial x_{i}}\right)_{p}$. In this case,

$$
d f_{p}\left(\left(\frac{\partial}{\partial x_{i}}\right)_{p}\right)(\mathbf{g})=\left.\frac{\partial\left(g \circ f \circ \varphi^{-1}\right)}{\partial X_{i}}\right|_{\varphi(p)}
$$

Using the chain rule, we find that

$$
d f_{p}\left(\left(\frac{\partial}{\partial x_{i}}\right)_{p}\right)(\mathbf{g})=\left.\left(\frac{\partial}{\partial x_{i}}\right)_{p} f \frac{d g}{d t}\right|_{t_{0}}
$$

Therefore, we have

$$
d f_{p}(u)=\left.u(\mathbf{f}) \frac{d}{d t}\right|_{t_{0}}
$$

This shows that we can identify $d f_{p}$ with the linear form in $T_{p}^{*}(M)$ defined by

$$
d f_{p}(u)=u(\mathbf{f}), \quad u \in T_{p} M
$$

by identifying $T_{t_{0}} \mathbb{R}$ with $\mathbb{R}$. This is consistent with our previous definition of $d f_{p}$ 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 $\left.\left(\mathcal{O}_{M, p}^{(k)} / \mathcal{S}_{M, p}^{(k)}\right)^{*}\right)$.

Again, even though this is just a restatement of facts we already showed, we state the following proposition because of its practical usefulness:

Proposition 3.9 Given any $C^{k}$-manifold, $M$, of dimension $n$, with $k \geq 1$, for any $p \in M$ and any chart $(U, \varphi)$ at $p$, the $n$ linear maps,

$$
\left(d x_{1}\right)_{p}, \ldots,\left(d x_{n}\right)_{p}
$$

form a basis of $T_{p}^{*} M$, where $\left(d x_{i}\right)_{p}$, the differential of $x_{i}$ at $p$, is identified with the linear form in $T_{p}^{*} M$ such that $\left(d x_{i}\right)_{p}(v)=v\left(\mathbf{x}_{\mathbf{i}}\right)$, for every $v \in T_{p} M$ (by identifying $T_{\lambda} \mathbb{R}$ with $\mathbb{R}$ ).

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, $\left.\gamma:\right] a, b\left[\rightarrow M\right.$, on a $C^{k}$-manifold, $M$, for any $\left.t_{0} \in\right] a, b\left[\right.$, we would like to define the tangent vector to the curve $\gamma$ at $t_{0}$ as a tangent vector to $M$ at $p=\gamma\left(t_{0}\right)$. We do this as follows: Recall that $\left.\frac{d}{d t}\right|_{t_{0}}$ is a basis vector of $T_{t_{0}}(\mathbb{R})=\mathbb{R}$. So, define the tangent vector to the curve $\gamma$ at $t_{0}$, denoted $\dot{\gamma}\left(t_{0}\right)$ (or $\gamma^{\prime}\left(t_{0}\right)$, or $\frac{d \gamma}{d t}\left(t_{0}\right)$ ) by

$$
\dot{\gamma}\left(t_{0}\right)=d \gamma_{t_{0}}\left(\left.\frac{d}{d t}\right|_{t_{0}}\right)
$$

Sometime, it is necessary to define curves (in a manifold) whose domain is not an open interval. A map, $\gamma:[a, b] \rightarrow M$, is a $C^{k}$-curve in $M$ if it is the restriction of some $C^{k}$-curve, $\widetilde{\gamma}:] a-\epsilon, b+\epsilon[\rightarrow M$, 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 continuous curve, $\gamma:[a, b] \rightarrow M$, is piecewise $C^{k}$ iff there a sequence, $a_{0}=a, a_{1}, \ldots, a_{m}=b$, so that the restriction, $\gamma_{i}$, of $\gamma$ to each $\left[a_{i}, a_{i+1}\right]$ is a $C^{k}$-curve, for $i=0, \ldots, m-1$. This implies that $\gamma_{i}^{\prime}\left(a_{i+1}\right)$ and $\gamma_{i+1}^{\prime}\left(a_{i+1}\right)$ are defined for $i=0, \ldots, m-1$, but there may be a jump in the tangent vector to $\gamma$ at $a_{i}$, that is, we may have $\gamma_{i}^{\prime}\left(a_{i+1}\right) \neq \gamma_{i+1}^{\prime}\left(a_{i+1}\right)$.

### 3.3 Tangent and Cotangent Bundles, Vector Fields, Lie Derivative

Let $M$ be a $C^{k}$-manifold (with $k \geq 2$ ). Roughly speaking, a vector field on $M$ is the assignment, $p \mapsto X(p)$, of a tangent vector, $X(p) \in T_{p}(M)$, 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 map, $X: M \rightarrow T(M)$, to be $C^{l}$.

If $M$ is a $C^{k}$-manifold of dimension $n$, then we can indeed make $T(M)$ into a $C^{k-1}$ manifold of dimension $2 n$ 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 triples $(U, \varphi, x)$, where $(U, \varphi)$ is a chart and $x \in \mathbb{R}^{n}$. 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,

$$
\pi: T(M) \rightarrow M, \quad \text { where } \quad \pi(v)=p \quad \text { if } \quad v \in T_{p}(M)
$$

We still have to give $T(M)$ a topology and to define a $C^{k-1}$-atlas. For every chart, $(U, \varphi)$, of $M$ (with $U$ open in $M$ ) we define the function, $\widetilde{\varphi}: \pi^{-1}(U) \rightarrow \mathbb{R}^{2 n}$, by

$$
\widetilde{\varphi}(v)=\left(\varphi \circ \pi(v), \theta_{U, \varphi, \pi(v)}^{-1}(v)\right),
$$

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 $\mathbb{R}^{2 n}$. 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 secondcountable in this topology. If $(U, \varphi)$ and $(V, \psi)$ are overlapping charts, then the transition map,

$$
\widetilde{\psi} \circ \widetilde{\varphi}^{-1}: \varphi(U \cap V) \times \mathbb{R}^{n} \longrightarrow \psi(U \cap V) \times \mathbb{R}^{n}
$$

is given by

$$
\tilde{\psi} \circ \widetilde{\varphi}^{-1}(z, x)=\left(\psi \circ \varphi^{-1}(z),\left(\psi \circ \varphi^{-1}\right)_{z}^{\prime}(x)\right), \quad(z, x) \in \varphi(U \cap V) \times \mathbb{R}^{n}
$$

It is clear that $\tilde{\psi} \circ \widetilde{\varphi}^{-1}$ is a $C^{k-1}$-map. Therefore, $T(M)$ is indeed a $C^{k-1}$-manifold of dimension $2 n$, called the tangent bundle.

Remark: Even if the manifold $M$ is naturally embedded in $\mathbb{R}^{N}$ (for some $N \geq n=\operatorname{dim}(M)$ ), it is not at all obvious how to view the tangent bundle, $T(M)$, as embedded in $\mathbb{R}^{N^{\prime}}$, for some suitable $N^{\prime}$. 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, $\pi: T^{*}(M) \rightarrow M$, and we can define charts in several ways. One method used by Warner [147] goes as follows: For any chart, $(U, \varphi)$, on $M$, we define the function, $\widetilde{\varphi}: \pi^{-1}(U) \rightarrow \mathbb{R}^{2 n}$, 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 $(U, \varphi)$. Again, one can make $T^{*}(M)$ into a $C^{k-1}$-manifold of dimension $2 n$, called the cotangent bundle. We leave the details as an exercise to the reader (Or, look at Berger and Gostiaux [17]). Another method using Version 3 of the definition of tangent vectors is presented in Section 7.2. For simplicity of notation, we also use the notation $T M$ for $T(M)$ (resp. $T^{*} M$ for $\left.T^{*}(M)\right)$.

Observe that for every chart, $(U, \varphi)$, on $M$, there is a bijection

$$
\tau_{U}: \pi^{-1}(U) \rightarrow U \times \mathbb{R}^{n}
$$

given by

$$
\tau_{U}(v)=\left(\pi(v), \theta_{U, \varphi, \pi(v)}^{-1}(v)\right)
$$

Clearly, $p r_{1} \circ \tau_{U}=\pi$, 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 $\tau_{U}$ 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 $\tau_{U}$ 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 $\tau_{U}$ ). For more on bundles, see Chapter 7, in particular, Section 7.2 on vector bundles where the construction of the bundles $T M$ and $T^{*} M$ is worked out in detail. See also the references in Chapter 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, $h: M \rightarrow N$, between two $C^{k}$-manifolds, we can define the function, $d h: T(M) \rightarrow T(N)$, (also denoted $T h$, or $h_{*}$, or $D h$ ) by setting

$$
d h(u)=d h_{p}(u), \quad \text { iff } \quad u \in T_{p}(M)
$$

We leave the next proposition as an exercise to the reader (A proof can be found in Berger and Gostiaux [17]).

Proposition 3.10 Given a $C^{k}$-map, $h: M \rightarrow N$, between two $C^{k}$-manifolds $M$ and $N$ (with $k \geq 1$ ), the map dh: $T(M) \rightarrow T(N)$ 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, $X$, of $T(M)$ over $U$, i.e., any function, $X: U \rightarrow T(M)$, such that $\pi \circ X=\operatorname{id}_{U}$ (i.e., $X(p) \in T_{p}(M)$, for every $\left.p \in U\right)$. We also say that $X$ is a lifting of $U$ into $T(M)$. We say that $X$ is a $C^{k}$-vector field on $U$ iff $X$ is a section over $U$ and a $C^{k}$-map. The set of $C^{k}$-vector fields over $U$ is denoted $\Gamma^{(k)}(U, T(M))$. Given a curve, $\gamma:[a, b] \rightarrow M$, a vector field, $X$, along $\gamma$ is any section of $T(M)$ over $\gamma$, i.e., a $C^{k}$-function, $X:[a, b] \rightarrow T(M)$, such that $\pi \circ X=\gamma$. We also say that $X$ lifts $\gamma$ into $T(M)$.

The above definition gives a precise meaning to the idea that a $C^{k}$-vector field on $M$ is an assignment, $p \mapsto X(p)$, of a tangent vector, $X(p) \in T_{p}(M)$, to a point, $p \in M$, so that $X(p)$ 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 $\left.\mathfrak{X}(M)\right)$.

Remark: We can also define a $C^{j}$-vector field on $U$ as a section, $X$, over $U$ which is a $C^{j}$-map, where $0 \leq j \leq k$. Then, we have the vector space, $\Gamma^{(j)}(U, T(M))$, etc .

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, $X$, such that

$$
X(p)=(p, u), \quad \text { with } \quad u \in \mathbb{R}^{n}
$$

for all $p \in U$. In other words, $X$ is defined by a function, $f: U \rightarrow \mathbb{R}^{n}$ (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, $X \in \Gamma^{(k)}(U, T(M))$, we define the vector field, $f X$, by

$$
(f X)(p)=f(p) X(p), \quad p \in U
$$

Obviously, $f X \in \Gamma^{(k)}(U, T(M))$, which shows that $\Gamma^{(k)}(U, T(M))$ is also a $\mathcal{C}^{k}(U)$-module. We also denote $X(p)$ by $X_{p}$. For any chart, $(U, \varphi)$, 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 $\frac{\partial}{\partial x_{i}}$.
Definition 3.15 Let $M$ be a $C^{k+1}$ manifold and let $X$ be a $C^{k}$ vector field on $M$. If $U$ is any open subset of $M$ and $f$ is any function in $\mathcal{C}^{k}(U)$, then the Lie derivative of $f$ with respect to $X$, denoted $X(f)$ or $L_{X} f$, is the function on $U$ given by

$$
X(f)(p)=X_{p}(f)=X_{p}(\mathbf{f}), \quad p \in U
$$

Observe that

$$
X(f)(p)=d f_{p}\left(X_{p}\right),
$$

where $d f_{p}$ is identified with the linear form in $T_{p}^{*}(M)$ defined by

$$
d f_{p}(v)=v(\mathbf{f}), \quad v \in T_{p} M
$$

by identifying $T_{t_{0}} \mathbb{R}$ with $\mathbb{R}$ (see the discussion following Proposition 3.8). The Lie derivative, $L_{X} f$, is also denoted $X[f]$.

As a special case, when $(U, \varphi)$ 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 $X(f) \in \mathcal{C}^{k-1}(U)$. As a consequence, every vector field $X \in$ $\Gamma^{(k)}(U, T(M))$ induces a linear map,

$$
L_{X}: \mathcal{C}^{k}(U) \longrightarrow \mathcal{C}^{k-1}(U)
$$

given by $f \mapsto X(f)$. It is immediate to check that $L_{X}$ has the Leibnitz property, i.e.,

$$
L_{X}(f g)=L_{X}(f) g+f L_{X}(g)
$$

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 [60] or Lafontaine [92]).

In the rest of this section, unless stated otherwise, we assume that $k \geq 1$. The following easy proposition holds (c.f. Warner [147]):

Proposition 3.11 Let $X$ be a vector field on the $C^{k+1}$-manifold, $M$, of dimension $n$. Then, the following are equivalent:
(a) $X$ is $C^{k}$.
(b) If $(U, \varphi)$ is a chart on $M$ and if $f_{1}, \ldots, f_{n}$ are the functions on $U$ uniquely defined by

$$
X \upharpoonright U=\sum_{i=1}^{n} f_{i} \frac{\partial}{\partial x_{i}}
$$

then each $f_{i}$ is a $C^{k}$-map.
(c) Whenever $U$ is open in $M$ and $f \in \mathcal{C}^{k}(U)$, then $X(f) \in \mathcal{C}^{k-1}(U)$.

Given any two $C^{k}$-vector field, $X, Y$, on $M$, for any function, $f \in \mathcal{C}^{k}(M)$, we defined above the function $X(f)$ and $Y(f)$. Thus, we can form $X(Y(f))$ (resp. $Y(X(f))$ ), which are in $\mathcal{C}^{k-2}(M)$. Unfortunately, even in the smooth case, there is generally no vector field, $Z$, such that

$$
Z(f)=X(Y(f)), \quad \text { for all } f \in \mathcal{C}^{k}(M)
$$

This is because $X(Y(f))$ (and $Y(X(f))$ ) involve second-order derivatives. However, if we consider $X(Y(f))-Y(X(f))$, then second-order derivatives cancel out and there is a unique vector field inducing the above differential operator. Intuitively, $X Y-Y X$ measures the "failure of $X$ and $Y$ to commute".

Proposition 3.12 Given any $C^{k+1}$-manifold, $M$, of dimension $n$, for any two $C^{k}$-vector fields, $X, Y$, on $M$, there is a unique $C^{k-1}$-vector field, $[X, Y]$, such that

$$
[X, Y](f)=X(Y(f))-Y(X(f)), \quad \text { for all } \quad f \in \mathcal{C}^{k-1}(M)
$$

Proof. First we prove uniqueness. For this it is enough to prove that $[X, Y]$ is uniquely defined on $\mathcal{C}^{k}(U)$, for any chart, $(U, \varphi)$. Over $U$, we know that

$$
X=\sum_{i=1}^{n} X_{i} \frac{\partial}{\partial x_{i}} \quad \text { and } \quad Y=\sum_{i=1}^{n} Y_{i} \frac{\partial}{\partial x_{i}}
$$

where $X_{i}, Y_{i} \in \mathcal{C}^{k}(U)$. Then, for any $f \in \mathcal{C}^{k}(M)$, we have

$$
\begin{aligned}
X(Y(f)) & =X\left(\sum_{j=1}^{n} Y_{j} \frac{\partial}{\partial x_{j}}(f)\right)=\sum_{i, j=1}^{n} X_{i} \frac{\partial}{\partial x_{i}}\left(Y_{j}\right) \frac{\partial}{\partial x_{j}}(f)+\sum_{i, j=1}^{n} X_{i} Y_{j} \frac{\partial^{2}}{\partial x_{j} \partial x_{i}}(f) \\
Y(X(f)) & =Y\left(\sum_{i=1}^{n} X_{i} \frac{\partial}{\partial x_{i}}(f)\right)=\sum_{i, j=1}^{n} Y_{j} \frac{\partial}{\partial x_{j}}\left(X_{i}\right) \frac{\partial}{\partial x_{i}}(f)+\sum_{i, j=1}^{n} X_{i} Y_{j} \frac{\partial^{2}}{\partial x_{i} \partial x_{j}}(f) .
\end{aligned}
$$

However, as $f \in \mathcal{C}^{k}(M)$, with $k \geq 2$, we have

$$
\sum_{i, j=1}^{n} X_{i} Y_{j} \frac{\partial^{2}}{\partial x_{j} \partial x_{i}}(f)=\sum_{i, j=1}^{n} X_{i} Y_{j} \frac{\partial^{2}}{\partial x_{i} \partial x_{j}}(f),
$$

and we deduce that

$$
X(Y(f))-Y(X(f))=\sum_{i, j=1}^{n}\left(X_{i} \frac{\partial}{\partial x_{i}}\left(Y_{j}\right)-Y_{i} \frac{\partial}{\partial x_{i}}\left(X_{j}\right)\right) \frac{\partial}{\partial x_{j}}(f)
$$

This proves that $[X, Y]=X Y-Y X$ is uniquely defined on $U$ and that it is $C^{k-1}$. Thus, if $[X, Y]$ exists, it is unique.

To prove existence, we use the above expression to define $[X, Y]_{U}$, locally on $U$, for every chart, $(U, \varphi)$. On any overlap, $U \cap V$, by the uniqueness property that we just proved, $[X, Y]_{U}$ and $[X, Y]_{V}$ must agree. But then, the $[X, Y]_{U}$ patch and yield a $C^{k-1}$-vector field defined on the whole of $M$.

Definition 3.16 Given any $C^{k+1}$-manifold, $M$, of dimension $n$, for any two $C^{k}$-vector fields, $X, Y$, on $M$, the Lie bracket, $[X, Y]$, of $X$ and $Y$, is the $C^{k-1}$ vector field defined so that

$$
[X, Y](f)=X(Y(f))-Y(X(f)), \quad \text { for all } \quad f \in \mathcal{C}^{k-1}(M)
$$

An an example, in $\mathbb{R}^{3}$, if $X$ and $Y$ are the two vector fields,

$$
X=\frac{\partial}{\partial x}+y \frac{\partial}{\partial z} \quad \text { and } \quad Y=\frac{\partial}{\partial y}
$$

then

$$
[X, Y]=-\frac{\partial}{\partial z}
$$

We also have the following simple proposition whose proof is left as an exercise (or, see Do Carmo [50]):

Proposition 3.13 Given any $C^{k+1}$-manifold, $M$, of dimension n, for any $C^{k}$-vector fields, $X, Y, Z$, on $M$, for all $f, g \in \mathcal{C}^{k}(M)$, we have:
(a) $[[X, Y], Z]+[[Y, Z], X]+[[Z, X], Y]=0 \quad$ (Jacobi identity).
(b) $[X, X]=0$.
(c) $[f X, g Y]=f g[X, Y]+f X(g) Y-g Y(f) X$.
(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.

Let $\varphi: M \rightarrow N$ be a diffeomorphism between two manifolds. Then, vector fields can be transported from $N$ to $M$ and conversely.
Definition 3.17 Let $\varphi: M \rightarrow N$ be a diffeomorphism between two $C^{k+1}$ manifolds. For every $C^{k}$ vector field, $Y$, on $N$, the pull-back of $Y$ along $\varphi$ is the vector field, $\varphi^{*} Y$, on $M$, given by

$$
\left(\varphi^{*} Y\right)_{p}=d \varphi_{\varphi(p)}^{-1}\left(Y_{\varphi(p)}\right), \quad p \in M
$$

For every $C^{k}$ vector field, $X$, on $M$, the push-forward of $X$ along $\varphi$ is the vector field, $\varphi_{*} X$, on $N$, given by

$$
\varphi_{*} X=\left(\varphi^{-1}\right)^{*} X
$$

that is, for every $p \in M$,

$$
\left(\varphi_{*} X\right)_{\varphi(p)}=d \varphi_{p}\left(X_{p}\right)
$$

or equivalently,

$$
\left(\varphi_{*} X\right)_{q}=d \varphi_{\varphi^{-1}(q)}\left(X_{\varphi^{-1}(q)}\right), \quad q \in N
$$

It is not hard to check that

$$
L_{\varphi_{*} X} f=L_{X}(f \circ \varphi) \circ \varphi^{-1}
$$

for any function $f \in C^{k}(N)$.
One more notion will be needed when we deal with Lie algebras.
Definition 3.18 Let $\varphi: M \rightarrow N$ be a $C^{k+1}$-map of manifolds. If $X$ is a $C^{k}$ vector field on $M$ and $Y$ is a $C^{k}$ vector field on $N$, we say that $X$ and $Y$ are $\varphi$-related iff

$$
d \varphi \circ X=Y \circ \varphi .
$$

The basic result about $\varphi$-related vector fields is:
Proposition 3.14 Let $\varphi: M \rightarrow N$ be a $C^{k+1}$-map of manifolds, let $X$ and $Y$ be $C^{k}$ vector fields on $M$ and let $X_{1}, Y_{1}$ be $C^{k}$ vector fields on $N$. If $X$ is $\varphi$-related to $X_{1}$ and $Y$ is $\varphi$-related to $Y_{1}$, then $[X, Y]$ is $\varphi$-related to $\left[X_{1}, Y_{1}\right]$.

Proof. Basically, one needs to unwind the definitions, see Warner [147], Chapter 1.

### 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 open 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

$$
\mathbb{R}^{m} \cong \mathbb{R}^{m} \times\{\underbrace{(0, \ldots, 0)}_{n-m}\} \hookrightarrow \mathbb{R}^{m} \times \mathbb{R}^{n-m}=\mathbb{R}^{n}, \quad\left(x_{1}, \ldots, x_{m}\right) \mapsto(x_{1}, \ldots, x_{m}, \underbrace{0, \ldots, 0}_{n-m}) .
$$

Definition 3.19 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, $(U, \varphi)$, of $M$, with $p \in U$, so that

$$
\varphi(U \cap N)=\varphi(U) \cap\left(\mathbb{R}^{m} \times\left\{0_{n-m}\right\}\right)
$$

(We write $0_{n-m}=\underbrace{(0, \ldots, 0)}_{n-m}$.)
The subset, $U \cap N$, of Definition 3.19 is sometimes called a slice of $(U, \varphi)$ and we say that $(U, \varphi)$ is adapted to $N$ (See O'Neill [119] or Warner [147]).

Other authors, including Warner [147], use the term submanifold in a broader sense than us and they use the word embedded submanifold for what is defined in Definition 3.19.

The following proposition has an almost trivial proof but it justifies the use of the word submanifold:

Proposition 3.15 Given a $C^{k}$-manifold, $M$, of dimension $n$, for any submanifold, $N$, of $M$ of dimension $m \leq n$, the family of pairs $(U \cap N, \varphi \upharpoonright U \cap N)$, where $(U, \varphi)$ 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 3.16 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, $(W, \eta)$, of $N$ at $p$, there is some chart, $(U, \varphi)$, of $M$ at $p$ so that

$$
\varphi(U \cap N)=\varphi(U) \cap\left(\mathbb{R}^{m} \times\left\{0_{n-m}\right\}\right) \quad \text { and } \quad \varphi \upharpoonright U \cap N=\eta \upharpoonright U \cap N
$$

where $p \in U \cap N \subseteq W$.

Proof. See Berger and Gostiaux [17] (Chapter 2).
It is also useful to define more general kinds of "submanifolds".
Definition 3.20 Let $\varphi: N \rightarrow M$ be a $C^{k}$-map of manifolds.
(a) The map $\varphi$ 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 $\varphi$ is an injective immersion.
(c) The map $\varphi$ is an embedding of $N$ into $M$ iff $\varphi$ 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, $\varphi$ 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 $\varphi$ 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 [147]) 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 [49], Berger and Gostiaux [17] or Gallier [58].

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): $x=t^{2}-1 ; y=t\left(t^{2}-1\right)$. Note that the cuspidal cubic, $t \mapsto\left(t^{2}, t^{3}\right)$, is an injective map, but it is not an immersion since its derivative at the origin is zero.

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 $\mathbb{R}^{2}$ :

$$
\begin{aligned}
& x=\frac{t\left(1+t^{2}\right)}{1+t^{4}} \\
& y=\frac{t\left(1-t^{2}\right)}{1+t^{4}}
\end{aligned}
$$

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

$$
t \mapsto(\cos t, \sin t, \cos c t, \sin c t),
$$

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 3.17 If $N$ is a submanifold of $M$, then the inclusion map, $j: N \rightarrow M$, is an embedding. Conversely, if $\varphi: N \rightarrow M$ is an embedding, then $\varphi(N)$ with the subspace topology is a submanifold of $M$ and $\varphi$ is a diffeomorphism between $N$ and $\varphi(N)$.

Proof. See O'Neill [119] (Chapter 1) or Berger and Gostiaux [17] (Chapter 2).
In summary, embedded submanifolds and (our) submanifolds coincide. Some authors refer to spaces of the form $\varphi(N)$, where $\varphi$ is an injective immersion, as immersed submanifolds and we have adopted this terminology. 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 [139] (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 3.21 Let $X$ 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 $X$ with initial condition $p_{0}$ is a $C^{k-1}$-curve, $\gamma: I \rightarrow M$, so that

$$
\dot{\gamma}(t)=X(\gamma(t)), \quad \text { for all } t \in I \quad \text { and } \quad \gamma(0)=p_{0}
$$

where $I=] a, b[\subseteq \mathbb{R}$ is an open interval containing 0 .

What definition 3.21 says is that an integral curve, $\gamma$, 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, $X(p)$, of the vector field $X$ at $p$.

Given a vector field, $X$, 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 3.22 Let $X$ 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 $X$ at $p_{0}$ is a map,

$$
\varphi: J \times U \rightarrow M
$$

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 $X$ with initial condition $p$.

Thus, a local flow for $X$ 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 [95] or Berger and Gostiaux [17].

Theorem 3.18 (Existence of a local flow) Let $X$ 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, $\varphi: J \times U \rightarrow M$, for $X$ at $p_{0}$. Furthermore, $\varphi$ is $C^{k-1}$.

Theorem 3.18 holds under more general hypotheses, namely, when the vector field satisfies some Lipschitz condition, see Lang [95] or Berger and Gostiaux [17].

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 $\gamma_{1}: I_{1} \rightarrow M$ and $\gamma_{2}: I_{2} \rightarrow M$ with initial condition $p_{0}$. This leads to the natural question: How do $\gamma_{1}$ and $\gamma_{2}$ differ on $I_{1} \cap I_{2}$ ? The next proposition shows they don't!

Proposition 3.19 Let $X$ 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 $\gamma_{1}: I_{1} \rightarrow M$ and $\gamma_{2}: I_{2} \rightarrow M$ are any two integral curves both with initial condition $p_{0}$, then $\gamma_{1}=\gamma_{2}$ on $I_{1} \cap I_{2}$.

Proof. Let $Q=\left\{t \in I_{1} \cap I_{2} \mid \gamma_{1}(t)=\gamma_{2}(t)\right\}$. 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

$$
Q=I_{1} \cap I_{2} \cap\left(\gamma_{1}, \gamma_{2}\right)^{-1}(\Delta)
$$

and $\gamma_{1}$ and $\gamma_{2}$ are continuous, we see that $Q$ is closed in $I_{1} \cap I_{2}$.
Pick any $u \in Q$ and consider the curves $\beta_{1}$ and $\beta_{2}$ 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 $\left.I=\right] a, b[$, we have $I-u=] a-u, b-u[$.$) Observe that$

$$
\dot{\beta}_{1}(t)=\dot{\gamma}_{1}(t+u)=X\left(\gamma_{1}(t+u)\right)=X\left(\beta_{1}(t)\right)
$$

and similarly, $\dot{\beta}_{2}(t)=X\left(\beta_{2}(t)\right)$. 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, $\beta_{1}:\left(I_{1}-u\right) \rightarrow M$ and $\beta_{2}:\left(I_{2}-u\right) \rightarrow M$ are two integral curves with the same initial condition, $q$. By Theorem 3.18, 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, $\gamma_{1}$ and $\gamma_{2}$ agree on $\widetilde{I}+u$, an open subset of $Q$, proving that $Q$ is indeed open in $I_{1} \cap I_{2}$.

Proposition 3.19 implies the important fact that there is a unique maximal integral curve with initial condition $p$. Indeed, if $\left\{\gamma_{j}: I_{j} \rightarrow M\right\}_{j \in K}$ is the family of all integral curves with initial condition $p$ (for some big index set, $K$ ), if we let $I(p)=\bigcup_{j \in K} I_{j}$, we can define a curve, $\gamma_{p}: I(p) \rightarrow M$, so that

$$
\gamma_{p}(t)=\gamma_{j}(t), \quad \text { if } \quad t \in I_{j} .
$$

Since $\gamma_{j}$ and $\gamma_{l}$ agree on $I_{j} \cap I_{l}$ for all $j, l \in K$, the curve $\gamma_{p}$ 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 $\gamma_{p}$ is called the maximal integral curve with initial condition $p$ and it is also denoted by $\gamma(p, t)$. Note that Proposition 3.19 implies that any two distinct integral curves are disjoint, i.e., do not intersect each other.

Consider the vector field in $\mathbb{R}^{2}$ given by

$$
X=-y \frac{\partial}{\partial x}+x \frac{\partial}{\partial y} .
$$

If we write $\gamma(t)=(x(t), y(t))$, the differential equation, $\dot{\gamma}(t)=X(\gamma(t))$, is expressed by

$$
\begin{aligned}
x^{\prime}(t) & =-y(t) \\
y^{\prime}(t) & =x(t)
\end{aligned}
$$

or, in matrix form,

$$
\binom{x^{\prime}}{y^{\prime}}=\left(\begin{array}{cc}
0 & -1 \\
1 & 0
\end{array}\right)\binom{x}{y} .
$$

If we write $X=\binom{x}{y}$ and $A=\left(\begin{array}{cc}0 & -1 \\ 1 & 0\end{array}\right)$, then the above equation is written as

$$
X^{\prime}=A X .
$$

Now, as

$$
e^{t A}=I+\frac{A}{1!} t+\frac{A^{2}}{2!} t^{2}+\cdots+\frac{A^{n}}{n!} t^{n}+\cdots,
$$

we get

$$
\frac{d}{d t}\left(e^{t A}\right)=A+\frac{A^{2}}{1!} t+\frac{A^{3}}{2!} t^{2}+\cdots+\frac{A^{n}}{(n-1)!} t^{n-1}+\cdots=A e^{t A}
$$

so we see that $e^{t A} p$ is a solution of the ODE $X^{\prime}=A X$ with initial condition $X=p$, and by uniqueness, $X=e^{t A} p$ is the solution of our ODE starting at $X=p$. Thus, our integral curve, $\gamma_{p}$, through $p=\binom{x_{0}}{y_{0}}$ is the circle given by

$$
\binom{x}{y}=\left(\begin{array}{cc}
\cos t & -\sin t \\
\sin t & \cos t
\end{array}\right)\binom{x_{0}}{y_{0}} .
$$

Observe that $I(p)=\mathbb{R}$, for every $p \in \mathbb{R}^{2}$.
The following interesting question now arises: Given any $p_{0} \in M$, if $\gamma_{p_{0}}: I\left(p_{0}\right) \rightarrow M$ is the maximal integral curve with initial condition $p_{0}$ and, for any $t_{1} \in I\left(p_{0}\right)$, if $p_{1}=\gamma_{p_{0}}\left(t_{1}\right) \in M$, then there is a maximal integral curve, $\gamma_{p_{1}}: I\left(p_{1}\right) \rightarrow M$, with initial condition $p_{1}$; what is the relationship between $\gamma_{p_{0}}$ and $\gamma_{p_{1}}$, if any? The answer is given by

Proposition 3.20 Let $X$ 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 $\gamma_{p_{0}}: I\left(p_{0}\right) \rightarrow M$ is the maximal integral curve with initial condition $p_{0}$, for any $t_{1} \in I\left(p_{0}\right)$, if $p_{1}=\gamma_{p_{0}}\left(t_{1}\right) \in M$ and $\gamma_{p_{1}}: I\left(p_{1}\right) \rightarrow M$ is the maximal integral curve with initial condition $p_{1}$, then

$$
I\left(p_{1}\right)=I\left(p_{0}\right)-t_{1} \quad \text { and } \quad \gamma_{p_{1}}(t)=\gamma_{\gamma_{p_{0}}\left(t_{1}\right)}(t)=\gamma_{p_{0}}\left(t+t_{1}\right), \quad \text { for all } t \in I\left(p_{0}\right)-t_{1} .
$$

Proof. Let $\gamma(t)$ be the curve given by

$$
\gamma(t)=\gamma_{p_{0}}\left(t+t_{1}\right), \quad \text { for all } t \in I\left(p_{0}\right)-t_{1} .
$$

Clearly, $\gamma$ is defined on $I\left(p_{0}\right)-t_{1}$ and

$$
\dot{\gamma}(t)=\dot{\gamma}_{p_{0}}\left(t+t_{1}\right)=X\left(\gamma_{p_{0}}\left(t+t_{1}\right)\right)=X(\gamma(t))
$$

and $\gamma(0)=\gamma_{p_{0}}\left(t_{1}\right)=p_{1}$. Thus, $\gamma$ is an integal curve defined on $I\left(p_{0}\right)-t_{1}$ with initial condition $p_{1}$. If $\gamma$ was defined on an interval, $\widetilde{I} \supseteq I\left(p_{0}\right)-t_{1}$ with $\widetilde{I} \neq I\left(p_{0}\right)-t_{1}$, then $\gamma_{p_{0}}$ would be defined on $\widetilde{I}+t_{1} \supset I\left(p_{0}\right)$, an interval strictly bigger than $I\left(p_{0}\right)$, contradicting the maximality of $I\left(p_{0}\right)$. Therefore, $I\left(p_{0}\right)-t_{1}=I\left(p_{1}\right)$.

It is useful to restate Proposition 3.20 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 $I\left(p_{0}\right)$ and get an integral curve, $\gamma_{p_{0}}$, with domain $I\left(p_{0}\right)$.

Instead of holding $p_{0} \in M$ fixed, we can hold $t \in \mathbb{R}$ fixed and consider the set

$$
\mathcal{D}_{t}(X)=\{p \in M \mid t \in I(p)\}
$$

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, $\gamma_{p}$, with initial condition $p$ (It is possible that $\mathcal{D}_{t}(X)=\emptyset$ ). By definition, if $\mathcal{D}_{t}(X) \neq \emptyset$, the point $\gamma_{p}(t)$ is well defined, and so, we obtain a map, $\Phi_{t}^{X}: \mathcal{D}_{t}(X) \rightarrow M$, with domain $\mathcal{D}_{t}(X)$, given by

$$
\Phi_{t}^{X}(p)=\gamma_{p}(t)
$$

The above suggests the following definition:
Definition 3.23 Let $X$ be a $C^{k-1}$ vector field on a $C^{k}$-manifold, $M,(k \geq 2)$. For any $t \in \mathbb{R}$, let

$$
\mathcal{D}_{t}(X)=\{p \in M \mid t \in I(p)\} \quad \text { and } \quad \mathcal{D}(X)=\{(t, p) \in \mathbb{R} \times M \mid t \in I(p)\}
$$

and let $\Phi^{X}: \mathcal{D}(X) \rightarrow M$ be the map given by

$$
\Phi^{X}(t, p)=\gamma_{p}(t)
$$

The map $\Phi^{X}$ is called the (global) flow of $X$ and $\mathcal{D}(X)$ is called its domain of definition. For any $t \in \mathbb{R}$ such that $\mathcal{D}_{t}(X) \neq \emptyset$, the map, $p \in \mathcal{D}_{t}(X) \mapsto \Phi^{X}(t, p)=\gamma_{p}(t)$, is denoted by $\Phi_{t}^{X}$ (i.e., $\Phi_{t}^{X}(p)=\Phi^{X}(t, p)=\gamma_{p}(t)$ ).

Observe that

$$
\mathcal{D}(X)=\bigcup_{p \in M}(I(p) \times\{p\})
$$

Also, using the $\Phi_{t}^{X}$ notation, the property of Proposition 3.20 reads

$$
\begin{equation*}
\Phi_{s}^{X} \circ \Phi_{t}^{X}=\Phi_{s+t}^{X} \tag{*}
\end{equation*}
$$

whenever both sides of the equation make sense. Indeed, the above says

$$
\Phi_{s}^{X}\left(\Phi_{t}^{X}(p)\right)=\Phi_{s}^{X}\left(\gamma_{p}(t)\right)=\gamma_{\gamma_{p}(t)}(s)=\gamma_{p}(s+t)=\Phi_{s+t}^{X}(p)
$$

Using the above property, we can easily show that the $\Phi_{t}^{X}$ are invertible. In fact, the inverse of $\Phi_{t}^{X}$ is $\Phi_{-t}^{X}$. First, note that

$$
\mathcal{D}_{0}(X)=M \quad \text { and } \quad \Phi_{0}^{X}=\mathrm{id}
$$

because, by definition, $\Phi_{0}^{X}(p)=\gamma_{p}(0)=p$, for every $p \in M$. Then, (*) implies that

$$
\Phi_{t}^{X} \circ \Phi_{-t}^{X}=\Phi_{t+-t}^{X}=\Phi_{0}^{X}=\mathrm{id},
$$

which shows that $\Phi_{t}^{X}: \mathcal{D}_{t}(X) \rightarrow \mathcal{D}_{-t}(X)$ and $\Phi_{-t}^{X}: \mathcal{D}_{-t}(X) \rightarrow \mathcal{D}_{t}(X)$ are inverse of each other. Moreover, each $\Phi_{t}^{X}$ is a $C^{k-1}$-diffeomorphism. We summarize in the following proposition some additional properties of the domains $\mathcal{D}(X), \mathcal{D}_{t}(X)$ and the maps $\Phi_{t}^{X}$ (for a proof, see Lang [95] or Warner [147]):

Theorem 3.21 Let $X$ 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 $\mathcal{D}_{t}(X) \neq \emptyset$, then $\mathcal{D}_{t}(X)$ is open (this is trivially true if $\mathcal{D}_{t}(X)=\emptyset$ ).
(b) The domain, $\mathcal{D}(X)$, of the flow, $\Phi^{X}$, is open and the flow is a $C^{k-1}$ map, $\Phi^{X}: \mathcal{D}(X) \rightarrow M$.
(c) Each $\Phi_{t}^{X}: \mathcal{D}_{t}(X) \rightarrow \mathcal{D}_{-t}(X)$ is a $C^{k-1}$-diffeomorphism with inverse $\Phi_{-t}^{X}$.
(d) For all $s, t \in \mathbb{R}$, the domain of definition of $\Phi_{s}^{X} \circ \Phi_{t}^{X}$ is contained but generally not equal to $\mathcal{D}_{s+t}(X)$. However, $\operatorname{dom}\left(\Phi_{s}^{X} \circ \Phi_{t}^{X}\right)=\mathcal{D}_{s+t}(X)$ if $s$ and $t$ have the same sign. Moreover, on $\operatorname{dom}\left(\Phi_{s}^{X} \circ \Phi_{t}^{X}\right)$, we have

$$
\Phi_{s}^{X} \circ \Phi_{t}^{X}=\Phi_{s+t}^{X} .
$$

## Remarks:

(1) We may omit the superscript, $X$, and write $\Phi$ instead of $\Phi^{X}$ if no confusion arises.
(2) The reason for using the terminology flow in referring to the map $\Phi^{X}$ can be clarified as follows: For any $t$ such that $\mathcal{D}_{t}(X) \neq \emptyset$, every integral curve, $\gamma_{p}$, with initial condition $p \in \mathcal{D}_{t}(X)$, 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, $\Phi^{X}(t, p)$ is the point reached by "flowing" for the amount of time $t$ on the integral curve $\gamma_{p}$ (through $p$ ) starting from $p$. Intuitively, we can imagine the flow of a fluid through $M$, and the vector field $X$ is the field of velocities of the flowing particles.

Given a vector field, $X$, as above, it may happen that $\mathcal{D}_{t}(X)=M$, for all $t \in \mathbb{R}$. In this case, namely, when $\mathcal{D}(X)=\mathbb{R} \times M$, we say that the vector field $X$ is complete. Then, the $\Phi_{t}^{X}$ are diffeomorphisms of $M$ and they form a group. The family $\left\{\Phi_{t}^{X}\right\}_{t \in \mathbb{R}}$ a called a 1-parameter group of $X$. In this case, $\Phi^{X}$ induces a group homomorphism, $(\mathbb{R},+) \longrightarrow \operatorname{Diff}(M)$, 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 $\mathcal{D}_{t}(X)=M$ for all $t$, the family $\left\{\Phi_{t}^{X}\right\}_{t \in \mathbb{R}}$ is called a local 1-parameter group generated by $X$, even though it is not a group, because the composition $\Phi_{s}^{X} \circ \Phi_{t}^{X}$ may not be defined.

If we go back to the vector field in $\mathbb{R}^{2}$ given by

$$
X=-y \frac{\partial}{\partial x}+x \frac{\partial}{\partial y},
$$

since the integral curve, $\gamma_{p}(t)$, through $p=\binom{x_{0}}{x_{0}}$ is given by

$$
\binom{x}{y}=\left(\begin{array}{cc}
\cos t & -\sin t \\
\sin t & \cos t
\end{array}\right)\binom{x_{0}}{y_{0}},
$$

the global flow associated with $X$ is given by

$$
\Phi^{X}(t, p)=\left(\begin{array}{cc}
\cos t & -\sin t \\
\sin t & \cos t
\end{array}\right) p
$$

and each diffeomorphism, $\Phi_{t}^{X}$, is the rotation,

$$
\Phi_{t}^{X}=\left(\begin{array}{cc}
\cos t & -\sin t \\
\sin t & \cos t
\end{array}\right)
$$

The 1-parameter group, $\left\{\Phi_{t}^{X}\right\}_{t \in \mathbb{R}}$, generated by $X$ is the group of rotations in the plane, $\mathrm{SO}(2)$.

More generally, if $B$ is an $n \times n$ invertible matrix that has a real logarithm, $A$ (that is, if $e^{A}=B$ ), then the matrix $A$ defines a vector field, $X$, in $\mathbb{R}^{n}$, with

$$
X=\sum_{i, j=1}^{n}\left(a_{i j} x_{j}\right) \frac{\partial}{\partial x_{i}},
$$

whose integral curves are of the form,

$$
\gamma_{p}(t)=e^{t A} p
$$

The one-parameter group, $\left\{\Phi_{t}^{X}\right\}_{t \in \mathbb{R}}$, generated by $X$ is given by $\left\{e^{t A}\right\}_{t \in \mathbb{R}}$.
When $M$ is compact, it turns out that every vector field is complete, a nice and useful fact.

Proposition 3.22 Let $X$ be a $C^{k-1}$ vector field on a $C^{k}$-manifold, $M,(k \geq 2)$. If $M$ is compact, then $X$ is complete, i.e., $\mathcal{D}(X)=\mathbb{R} \times M$. Moreover, the map $t \mapsto \Phi_{t}^{X}$ is a homomorphism from the additive group $\mathbb{R}$ to the group, $\operatorname{Diff}(M)$, of $\left(C^{k-1}\right)$ diffeomorphisms of $M$.

Proof. Pick any $p \in M$. By Theorem 3.18, there is a local flow, $\varphi_{p}: J(p) \times U(p) \rightarrow M$, 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 $J(p) \times U(p) \subseteq \mathcal{D}(X)$. 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 $t \in]-\epsilon,+\epsilon\left[\right.$ and, for all $p \in M$, if $\gamma_{p}$ is the maximal integral curve with initial condition $p$, then $]-\epsilon,+\epsilon[\subseteq I(p)$.

For any $t \in]-\epsilon,+\epsilon\left[\right.$, 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

$$
\gamma_{\gamma_{p}(t)}(t)=\gamma_{p}(t+t)=\gamma_{p}(2 t)
$$

which shows that $\gamma_{p}$ is in fact defined for all $\left.t \in\right]-2 \epsilon,+2 \epsilon[$. 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 $\mathcal{D}(X)=\mathbb{R} \times M$.

## Remarks:

(1) The proof of Proposition 3.22 also applies when $X$ is a vector field with compact support (this means that the closure of the set $\{p \in M \mid X(p) \neq 0\}$ is compact).
(2) If $\varphi: M \rightarrow N$ is a diffeomorphism and $X$ is a vector field on $M$, then it can be shown that the local 1-parameter group associated with the vector field, $\varphi_{*} X$, is

$$
\left(\varphi \circ \Phi_{t}^{X} \circ \varphi^{-1}\right)
$$

A point $p \in M$ where a vector field vanishes, i.e., $X(p)=0$, is called a critical point of $X$. 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 [106], Guillemin and Pollack [69] or DoCarmo [49], which contains an informal but very clear presentation of the PoincaréHopf index theorem). Another famous theorem about vector fields says that every smooth
vector field on a sphere of even dimension $\left(S^{2 n}\right)$ must vanish in at least one point (the socalled "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, $\gamma$, passes through a critical point, $p$, then $\gamma$ 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 $\gamma$ 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 $X$ and $Y$ on $M$. For any $p \in M$, we can flow along the integral curve of $X$ with initial condition $p$ to $\Phi_{t}(p)$ (for $t$ small enough) and then evaluate $Y$ there, getting $Y\left(\Phi_{t}(p)\right)$. Now, this vector belongs to the tangent space $T_{\Phi_{t}(p)}(M)$, but $Y(p) \in T_{p}(M)$. So to "compare" $Y\left(\Phi_{t}(p)\right)$ and $Y(p)$, we bring back $Y\left(\Phi_{t}(p)\right)$ to $T_{p}(M)$ by applying the tangent map, $d \Phi_{-t}$, at $\Phi_{t}(p)$, to $Y\left(\Phi_{t}(p)\right)$ (Note that to alleviate the notation, we use the slight abuse of notation $d \Phi_{-t}$ instead of $d\left(\Phi_{-t}\right)_{\Phi_{t}(p)}$.) Then, we can form the difference $d \Phi_{-t}\left(Y\left(\Phi_{t}(p)\right)\right)-Y(p)$, divide by $t$ and consider the limit as $t$ goes to 0 .

Definition 3.24 Let $M$ be a $C^{k+1}$ manifold. Given any two $C^{k}$ vector fields, $X$ and $Y$ on $M$, for every $p \in M$, the Lie derivative of $Y$ with respect to $X$ at $p$, denoted $\left(L_{X} Y\right)_{p}$, is given by

$$
\left(L_{X} Y\right)_{p}=\lim _{t \longrightarrow 0} \frac{d \Phi_{-t}\left(Y\left(\Phi_{t}(p)\right)\right)-Y(p)}{t}=\left.\frac{d}{d t}\left(d \Phi_{-t}\left(Y\left(\Phi_{t}(p)\right)\right)\right)\right|_{t=0}
$$

It can be shown that $\left(L_{X} Y\right)_{p}$ is our old friend, the Lie bracket, i.e.,

$$
\left(L_{X} Y\right)_{p}=[X, Y]_{p}
$$

(For a proof, see Warner [147] or O'Neill [119]).
In terms of Definition 3.17, observe that

$$
\left(L_{X} Y\right)_{p}=\lim _{t \longrightarrow 0} \frac{\left(\left(\Phi_{-t}\right)_{*} Y\right)(p)-Y(p)}{t}=\lim _{t \rightarrow 0} \frac{\left(\Phi_{t}^{*} Y\right)(p)-Y(p)}{t}=\left.\frac{d}{d t}\left(\Phi_{t}^{*} Y\right)(p)\right|_{t=0},
$$

since $\left(\Phi_{-t}\right)^{-1}=\Phi_{t}$.

### 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 gluing together items constructed on the domains of charts. Partitions of unity are a crucial technical tool in this gluing 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

$$
B(r)=\left\{\left(x_{1}, \ldots, x_{n}\right) \in \mathbb{R}^{n} \mid x_{1}^{2}+\cdots+x_{n}^{2}<r\right\}
$$

and by $\overline{B(r)}=\left\{\left(x_{1}, \ldots, x_{n}\right) \in \mathbb{R}^{n} \mid x_{1}^{2}+\cdots+x_{n}^{2} \leq r\right\}$, its closure.
Proposition 3.23 There is a smooth function, $b: \mathbb{R}^{n} \rightarrow \mathbb{R}$, 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, $h: \mathbb{R} \rightarrow \mathbb{R}$, 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 $C^{\infty}$ (but not analytic!). Then, define $b: \mathbb{R}^{n} \rightarrow \mathbb{R}$, by

$$
b\left(x_{1}, \ldots, x_{n}\right)=\frac{h\left(4-x_{1}^{2}-\cdots-x_{n}^{2}\right)}{h\left(4-x_{1}^{2}-\cdots-x_{n}^{2}\right)+h\left(x_{1}^{2}+\cdots+x_{n}^{2}-1\right)} .
$$

It is immediately verified that $b$ satisfies the required conditions.
Given a topological space, $X$, for any function, $f: X \rightarrow \mathbb{R}$, the support of $f$, denoted $\operatorname{supp} f$, is the closed set,

$$
\operatorname{supp} f=\overline{\{x \in X \mid f(x) \neq 0\}}
$$

Proposition 3.23 yields the following useful technical result:
Proposition 3.24 Let $M$ be a smooth manifold. For any open subset, $U \subseteq M$, any $p \in U$ and any smooth function, $f: U \rightarrow \mathbb{R}$, there exist an open subset, $V$, with $p \in V$ and a smooth function, $\widetilde{f}: M \rightarrow \mathbb{R}$, defined on the whole of $M$, so that $\bar{V}$ is compact,

$$
\bar{V} \subseteq U, \quad \operatorname{supp} \tilde{f} \subseteq U
$$

and

$$
\widetilde{f}(q)=f(q), \quad \text { for all } \quad q \in \bar{V}
$$

Proof. Using a scaling function, it is easy to find a chart, $(W, \varphi)$ 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 3.23. Then, $\widetilde{b}$ is a smooth function on $W$ with support $\varphi^{-1}(\overline{B(2)}) \subseteq W$. 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 $p, \bar{V}=\varphi^{-1}(\overline{B(1)}) \subseteq W$ is compact and, clearly, $\widetilde{b}=1$ on $\bar{V}$. 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 $\tilde{f}$ satisfies the required properties.
If $X$ is a (Hausdorff) topological space, a family, $\left\{U_{\alpha}\right\}_{\alpha \in I}$, of subsets $U_{\alpha}$ of $X$ is a cover (or covering) of $X$ iff $X=\bigcup_{\alpha \in I} U_{\alpha}$. A cover, $\left\{U_{\alpha}\right\}_{\alpha \in I}$, such that each $U_{\alpha}$ is open is an open cover. If $\left\{U_{\alpha}\right\}_{\alpha \in I}$ is a cover of $X$, for any subset, $J \subseteq I$, the subfamily $\left\{U_{\alpha}\right\}_{\alpha \in J}$ is a subcover of $\left\{U_{\alpha}\right\}_{\alpha \in I}$ if $X=\bigcup_{\alpha \in J} U_{\alpha}$, i.e., $\left\{U_{\alpha}\right\}_{\alpha \in J}$ is still a cover of $X$. Given two covers, $\left\{U_{\alpha}\right\}_{\alpha \in I}$ and $\left\{V_{\beta}\right\}_{\beta \in J}$, we say that $\left\{U_{\alpha}\right\}_{\alpha \in I}$ is a refinement of $\left\{V_{\beta}\right\}_{\beta \in J}$ iff there is a function, $h: I \rightarrow J$, so that $U_{\alpha} \subseteq V_{h(\alpha)}$, for all $\alpha \in I$.

A cover, $\left\{U_{\alpha}\right\}_{\alpha \in I}$, 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 has an open 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, $\left\{U_{i}\right\}_{i \geq 1}$, 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, $(U, \varphi)$, around $p$, then $\varphi(U)=\Omega$ for some open $\Omega \subseteq \mathbb{R}^{n}(n=\operatorname{dim} M)$. So, we can pick a small closed ball, $\overline{B(q, \epsilon)} \subseteq \Omega$, of center $q=\varphi(p)$ and radius $\epsilon$, and as $\varphi$ 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.

Definition 3.25 Let $M$ be a (smooth) manifold. A partition of unity on $M$ is a family, $\left\{f_{i}\right\}_{i \in I}$, of smooth functions on $M$ (the index set $I$ may be uncountable) such that
(a) The family of supports, $\left\{\operatorname{supp} f_{i}\right\}_{i \in I}$, 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 $\left\{U_{\alpha}\right\}_{\alpha \in J}$ is a cover of $M$, we say that the partition of unity $\left\{f_{i}\right\}_{i \in I}$ is subordinate to the cover $\left\{U_{\alpha}\right\}_{\alpha \in J}$ if $\left\{\operatorname{supp} f_{i}\right\}_{i \in I}$ is a refinement of $\left\{U_{\alpha}\right\}_{\alpha \in J}$. When $I=J$ and $\operatorname{supp} f_{i} \subseteq U_{i}$, we say that $\left\{f_{i}\right\}_{i \in I}$ is subordinate to $\left\{U_{\alpha}\right\}_{\alpha \in I}$ with the same index set as the partition of unity.

In Definition 3.25, 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, $\operatorname{supp} f_{i}$. 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 3.25 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 [147], Lemma 1.9).

The first step is to construct a sequence of open sets, $G_{i}$, such that

1. $\bar{G}_{i}$ is compact,
2. $\bar{G}_{i} \subseteq G_{i+1}$,
3. $X=\bigcup_{i=1}^{\infty} G_{i}$.

As $M$ is second-countable, there is a countable basis of open sets, $\left\{U_{i}\right\}_{i \geq 1}$, for $M$. Since $M$ is locally compact, we can find a subfamily of $\left\{U_{i}\right\}_{i \geq 1}$ 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, $\left\{U_{i}\right\}_{i \geq 1}$, 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 $\bar{G}_{k}$ 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

$$
\bar{G}_{k} \subseteq U_{1} \cup \cdots \cup U_{j_{k+1}}
$$

and set

$$
G_{k+1}=U_{1} \cup \cdots \cup U_{j_{k+1}} .
$$

Obviously, the family $\left\{G_{i}\right\}_{i \geq 1}$ satisfies (1)-(3).
Now, let $\left\{U_{\alpha}\right\}_{\alpha \in I}$ be an arbitrary open cover of $M$. For any $i \geq 3$, the set $\bar{G}_{i}-G_{i-1}$ is compact and contained in the open $G_{i+1}-\bar{G}_{i-2}$. For each $i \geq 3$, choose a finite subcover of the open cover $\left\{U_{\alpha} \cap\left(G_{i+1}-\bar{G}_{i-2}\right)\right\}_{\alpha \in I}$ of $\bar{G}_{i}-G_{i-1}$, and choose a finite subcover of the
open cover $\left\{U_{\alpha} \cap G_{3}\right\}_{\alpha \in I}$ of the compact set $\bar{G}_{2}$. 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 $\left\{U_{\alpha}\right\}_{\alpha \in I}$ and consists of open sets with compact closures.

## Remarks:

1. Proposition 3.25 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 $(\omega)$ has a countable basis of open sets. (The open subsets containing $\omega$ 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, $\left(U_{i}, \varphi_{i}\right)$, the map $\varphi_{i}$ 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 3.26 Let $M$ be a smooth manifold and let $\left\{U_{\alpha}\right\}_{\alpha \in I}$ be an open cover for $M$. Then, there is a countable partition of unity, $\left\{f_{i}\right\}_{i \geq 1}$, subordinate to the cover $\left\{U_{\alpha}\right\}_{\alpha \in I}$ and the support, supp $f_{i}$, of each $f_{i}$ is compact. If one does not require compact supports, then there is a partition of unity, $\left\{f_{\alpha}\right\}_{\alpha \in I}$, subordinate to the cover $\left\{U_{\alpha}\right\}_{\alpha \in I}$ with at most countably many of the $f_{\alpha}$ not identically zero. (In the second case, $\operatorname{supp} f_{\alpha} \subseteq U_{\alpha}$.)

Proof. Again, we reproduce Warner's proof (Warner [147], Theorem 1.11). As our manifolds are second-countable, Hausdorff and locally compact, from the proof of Proposition 3.25, we have the sequence of open subsets, $\left\{G_{i}\right\}_{i \geq 1}$ and we set $G_{0}=\emptyset$. For any $p \in M$, let $i_{p}$ be the largest integer such that $p \in M-\bar{G}_{i_{p}}$. Choose an $\alpha_{p}$ such that $p \in U_{\alpha_{p}}$; we can find a chart, $(U, \varphi)$, centered at $p$ such that $U \subseteq U_{\alpha_{p}} \cap\left(G_{i_{p}+2}-\bar{G}_{i_{p}}\right)$ 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 3.23. Then, $\psi_{p}$ is a smooth function on $M$ which has value 1 on some open subset, $W_{p}$, containing $p$ and has compact
support lying in $U \subseteq U_{\alpha_{p}} \cap\left(G_{i_{p}+2}-\bar{G}_{i_{p}}\right)$. For each $i \geq 1$, choose a finite set of points, $p \in M$, whose corresponding opens, $W_{p}$, cover $\bar{G}_{i}-G_{i-1}$. Order the corresponding $\psi_{p}$ functions in a sequence, $\psi_{j}, j=1,2, \ldots$. The supports of the $\psi_{j}$ 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, $\left\{f_{i}\right\}_{i \geq 1}$, is a partition of unity subordinate to the cover $\left\{U_{\alpha}\right\}_{\alpha \in I}$ and $\operatorname{supp} f_{i}$ is compact for all $i \geq 1$.

Now, when we don't require compact support, if we let $f_{\alpha}$ be identically zero if no $f_{i}$ has support in $U_{\alpha}$ and otherwise let $f_{\alpha}$ be the sum of the $f_{i}$ with support in $U_{\alpha}$, then we obtain a partition of unity subordinate to $\left\{U_{\alpha}\right\}_{\alpha \in I}$ with at most countably many of the $f_{\alpha}$ not identically zero. We must have supp $f_{\alpha} \subseteq U_{\alpha}$ because for any locally finite family of closed sets, $\left\{F_{\beta}\right\}_{\beta \in J}$, we have $\overline{\bigcup_{\beta \in J} F_{\beta}}=\bigcup_{\beta \in J} F_{\beta}$.

We close this section by stating a famous theorem of Whitney whose proof uses partitions of unity.

Theorem 3.27 (Whitney, 1935) Any smooth manifold (Hausdorff and second-countable), $M$, of dimension $n$ is diffeomorphic to a closed submanifold of $\mathbb{R}^{2 n+1}$.

For a proof, see Hirsch [76], 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,

$$
\mathbb{H}^{m}=\left\{\left(x_{1}, \ldots, x_{m}\right) \in \mathbb{R}^{m} \mid x_{m} \geq 0\right\} .
$$

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

$$
\partial \mathbb{H}^{m}=\left\{x \in \mathbb{H}^{m} \mid x_{m}=0\right\}
$$

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

$$
\operatorname{Int}\left(\mathbb{H}^{m}\right)=\mathbb{H}^{m}-\partial \mathbb{H}^{m}
$$

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 $f: U \rightarrow V$ is a continuous function, we need to explain what we mean by $f$ being smooth. We say that $f: U \rightarrow V$, as above, is smooth if it has an extension, $\widetilde{f}: \widetilde{U} \rightarrow \widetilde{V}$, 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, $(U, \varphi)$, where $U$ is an open subset of $M$ and $\varphi: U \rightarrow \Omega$ 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 3.26 Given some integer $n \geq 1$ and given some $k$ such that $k$ is either an integer $k \geq 1$ or $k=\infty$, 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, $\partial M$, of a manifold (with boundary), $M$, is the set of all points, $p \in M$, such that there is some chart, $\left(U_{\alpha}, \varphi_{\alpha}\right)$, with $p \in U_{\alpha}$ and $\varphi_{\alpha}(p) \in \partial \mathbb{H}^{n}$. We also let $\operatorname{Int}(M)=M-\partial M$ and call it the interior of $M$.

Do not confuse the boundary $\partial M$ and the interior $\operatorname{Int}(M)$ of a manifold with bound-
ary 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, $T_{p} M$, 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 [76], 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 [76], Chapter 1, Section 4, Exercise 12).

If $M$ is a manifold with boundary, we see that $\operatorname{Int}(M)$ is a manifold without boundary. What about $\partial M$ ? Interestingly, the boundary, $\partial M$, of a manifold with boundary, $M$, of dimension $n$, is a manifold of dimension $n-1$. For this, we need the following Proposition:

Proposition 3.28 If $M$ is a manifold with boundary of dimension $n$, for any $p \in \partial M$ on the boundary on $M$, for any chart, $(U, \varphi)$, 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, $(V, \psi)$, with $p \in V$ and $\psi(p) \in \partial \mathbb{H}^{n}$. Let $(U, \varphi)$ be any other chart, with $p \in M$ and assume that $q=\varphi(p) \in \operatorname{Int}\left(\mathbb{H}^{n}\right)$. The transition map, $\psi \circ \varphi^{-1}: \varphi(U \cap V) \rightarrow \psi(U \cap V)$, is a diffeomorphism and $q=\varphi(p) \in \operatorname{Int}\left(\mathbb{H}^{n}\right)$. By the inverse function theorem, there is some open, $W \subseteq \varphi(U \cap V) \cap \operatorname{Int}\left(\mathbb{H}^{n}\right) \subseteq \mathbb{R}^{n}$, with $q \in W$, so that $\psi \circ \varphi^{-1}$ maps $W$ homeomorphically onto some subset, $\Omega$, open $\operatorname{in} \operatorname{Int}\left(\mathbb{H}^{n}\right)$, with $\psi(p) \in \Omega$, contradicting the hypothesis, $\psi(p) \in \partial \mathbb{H}^{n}$.

Using Proposition 3.28, we immediately derive the fact that $\partial M$ is a manifold of dimension $n-1$. We obtain charts on $\partial M$ by considering the charts $(U \cap \partial M, L \circ \varphi)$, where $(U, \varphi)$ is a chart on $M$ such that $U \cap \partial M=\varphi^{-1}\left(\partial \mathbb{H}^{n}\right) \neq \emptyset$ and $L: \partial \mathbb{H}^{n} \rightarrow \mathbb{R}^{n-1}$ is the natural isomorphism (see see Hirsch [76], Chapter 1, Section 4).

### 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, $T_{p} M$, 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 $T_{p} M$ 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 $T_{p} M$ (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., $\left.g\left(u_{i}\right)=v_{i}, i=1, \ldots, n\right)$. Then, look at the determinant, $\operatorname{det}(g)$, of this map. We know that $\operatorname{det}(g)=\operatorname{det}(P)$, 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 $\operatorname{det}(g)$ 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, $h: E^{k} \rightarrow \mathbb{R}$, 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

$$
h\left(v_{1}, \ldots, v_{n}\right)=\operatorname{det}(g) h\left(u_{1}, \ldots, u_{n}\right)
$$

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

$$
\omega\left(u_{1}, \ldots, u_{n}\right) \text { and } \omega\left(v_{1}, \ldots, v_{n}\right) \text { have the same sign for all } \omega \in \bigwedge^{n} E^{*}-\{0\}
$$

(where 0 denotes 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), $\omega$, of $\bigwedge^{n} E^{*}$, and to say that $u_{1}, \ldots, u_{n}$ has positive orientation iff $\omega\left(u_{1}, \ldots, u_{n}\right)>0$.

Given an orientation (say, given by $\omega \in \bigwedge^{n} E^{*}$ ) of $E$, a linear map, $f: E \rightarrow E$, is orientation preserving iff $\omega\left(f\left(u_{1}\right), \ldots, f\left(u_{n}\right)\right)>0$ whenever $\omega\left(u_{1}, \ldots, u_{n}\right)>0$ (or equivalently, iff $\operatorname{det}(f)>0)$.

Now, to define the orientation of an $n$-dimensional manifold, $M$, we use charts. Given any $p \in M$, for any chart, $(U, \varphi)$, at $p$, the tangent map, $d \varphi_{\varphi(p)}^{-1}: \mathbb{R}^{n} \rightarrow T_{p} M$ makes sense. If $\left(e_{1}, \ldots, e_{n}\right)$ is the standard basis of $\mathbb{R}^{n}$, as it gives an orientation to $\mathbb{R}^{n}$, we can orient $T_{p} M$ by giving it the orientation induced by the basis $d \varphi_{\varphi(p)}^{-1}\left(e_{1}\right), \ldots, d \varphi_{\varphi(p)}^{-1}\left(e_{n}\right)$. Then, the consistency of orientations of the $T_{p} M$ 's is given by the overlapping of charts. We require that the Jacobian determinants of all $\varphi_{j} \circ \varphi_{i}^{-1}$ have the same sign, whenever $\left(U_{i}, \varphi_{i}\right)$ and $\left(U_{j}, \varphi_{j}\right)$ 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.

[^0]Definition 3.27 Given a smooth manifold, $M$, of dimension $n$, an orientation atlas of $M$ is any atlas so that the transition maps, $\varphi_{i}^{j}=\varphi_{j} \circ \varphi_{i}^{-1},\left(\right.$ from $\varphi_{i}\left(U_{i} \cap U_{j}\right)$ to $\left.\varphi_{j}\left(U_{i} \cap U_{j}\right)\right)$ all have a positive Jacobian determinant for every point in $\varphi_{i}\left(U_{i} \cap U_{j}\right)$. A manifold is orientable iff its has some orientation atlas.

Definition 3.27 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, $\Lambda^{n} T^{*} M$ (also denoted $\left.\bigwedge_{n}^{*} M\right)$ in much the same way that we defined the bundles $T M$ and $T^{*} M$. There is an obvious smooth projection map, $\pi: \bigwedge^{n} T^{*} M \rightarrow M$. Then, leaving the details of the fact that $\Lambda^{n} T^{*} M$ 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, $\omega: M \rightarrow \bigwedge^{n} T^{*} M$, so that $\pi \circ \omega=\mathrm{id}$.

Definition 3.28 If $M$ is an $n$-dimensional manifold, a smooth section, $\omega \in \Gamma\left(M, \bigwedge^{n} T^{*} M\right)$, is called a (smooth) $n$-form. The set of $n$-forms, $\Gamma\left(M, \bigwedge^{n} T^{*} M\right)$, is also denoted $\mathcal{A}^{n}(M)$. An $n$-form, $\omega$, is a nowhere-vanishing $n$-form on $M$ or volume form on $M$ iff $\omega_{p}$ is a nonzero form for every $p \in M$. This is equivalent to saying that $\omega_{p}\left(u_{1}, \ldots, u_{n}\right) \neq 0$, for all $p \in M$ and all bases, $u_{1}, \ldots, u_{n}$, of $T_{p} M$.

The determinant function, $\left(u_{1}, \ldots, u_{n}\right) \mapsto \operatorname{det}\left(u_{1}, \ldots, u_{n}\right)$, where the $u_{i}$ are expressed over the canonical basis $\left(e_{1}, \ldots, e_{n}\right)$ of $\mathbb{R}^{n}$, is a volume form on $\mathbb{R}^{n}$. We will denote this volume form by $\omega_{0}$. Another standard notation is $d x_{1} \wedge \cdots \wedge d x_{n}$, but this notation may be very puzzling for readers not familiar with exterior algebra. Observe the justification for the term volume form: the quantity $\operatorname{det}\left(u_{1}, \ldots, u_{n}\right)$ is indeed the (signed) volume of the parallelepiped

$$
\left\{\lambda_{1} u_{1}+\cdots+\lambda_{n} u_{n} \mid 0 \leq \lambda_{i} \leq 1,1 \leq i \leq n\right\} .
$$

A volume form on the sphere $S^{n} \subseteq \mathbb{R}^{n+1}$ is obtained as follows:

$$
\omega_{p}\left(u_{1}, \ldots u_{n}\right)=\operatorname{det}\left(p, u_{1}, \ldots u_{n}\right)
$$

where $p \in S^{n}$ and $u_{1}, \ldots u_{n} \in T_{p} S^{n}$. 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 $\omega$ is any $n$-form, then $f \omega$ is also an $n$-form.

Definition 3.29 Let $\varphi: M \rightarrow N$ 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 pull-back, $\varphi^{*} \omega$, of $\omega$ to $M$ is the $n$-form on $M$ given by

$$
\varphi^{*} \omega_{p}\left(u_{1}, \ldots, u_{n}\right)=\omega_{\varphi(p)}\left(d \varphi_{p}\left(u_{1}\right), \ldots, d \varphi_{p}\left(u_{n}\right)\right)
$$

for all $p \in M$ and all $u_{1}, \ldots, u_{n} \in T_{p} M$.

One checks immediately that $\varphi^{*} \omega$ is indeed an $n$-form on $M$. More interesting is the following Proposition:

Proposition 3.29 (a) If $\varphi: M \rightarrow N$ is a local diffeomorphism of manifolds, where $\operatorname{dim} M=$ $\operatorname{dim} N=n$, and $\omega \in \mathcal{A}^{n}(N)$ is a volume form on $N$, then $\varphi^{*} \omega$ is a volume form on $M$. (b) Assume $M$ has a volume form, $\omega$. Then, for every $n$-form, $\eta \in \mathcal{A}^{n}(M)$, there is a unique smooth function, $f \in C^{\infty}(M)$, so that $\eta=f \omega$. If $\eta$ is a volume form, then $f(p) \neq 0$ for all $p \in M$.

Proof. (a) By definition,

$$
\varphi^{*} \omega_{p}\left(u_{1}, \ldots, u_{n}\right)=\omega_{\varphi(p)}\left(d \varphi_{p}\left(u_{1}\right), \ldots, d \varphi_{p}\left(u_{n}\right)\right)
$$

for all $p \in M$ and all $u_{1}, \ldots, u_{n} \in T_{p} M$. As $\varphi$ 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 $d \varphi_{p}\left(u_{1}\right), \ldots, d \varphi_{p}\left(u_{n}\right)$, and as $\omega$ is nonzero at every point for every basis, $\varphi^{*} \omega_{p}\left(u_{1}, \ldots, u_{n}\right) \neq 0$.
(b) Pick any $p \in M$ and let $(U, \varphi)$ be any chart at $p$. As $\varphi$ is a diffeomorphism, by (a), we see that $\varphi^{-1^{*}} \omega$ is a volume form on $\varphi(U)$. But then, it is easy to see that $\varphi^{-1^{*}} \eta=g \varphi^{-1^{*}} \omega$, for some unique smooth function, $g$, on $\varphi(U)$ and so, $\eta=f_{U} \omega$, for some unique smooth function, $f_{U}$, on $U$. For any two overlapping charts, $\left(U_{i}, \varphi_{i}\right)$ and $\left(U_{j}, \varphi_{j}\right)$, for every $p \in U_{i} \cap U_{j}$, for every basis $u_{1}, \ldots, u_{n}$ of $T_{p} M$, we have

$$
\eta_{p}\left(u_{1}, \ldots, u_{n}\right)=f_{i}(p) \omega_{p}\left(u_{1}, \ldots, u_{n}\right)=f_{j}(p) \omega_{p}\left(u_{1}, \ldots, u_{n}\right)
$$

and as $\omega_{p}\left(u_{1}, \ldots, u_{n}\right) \neq 0$, 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, $\left\{U_{i}\right\}$, of $M$, and so, there is a smooth function, $f$, defined on the whole of $M$ and such that $f \upharpoonright U_{i}=f_{i}$. As the $f_{i}$ 's are unique, so is $f$. If $\eta$ is a volume form, then $\eta_{p}$ does not vanish for all $p \in M$ and since $\omega_{p}$ is also a volume form, $\omega_{p}$ does not vanish for all $p \in M$, so $f(p) \neq 0$ for all $p \in M$.

Remark: If $\varphi$ and $\psi$ 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 $\omega$ is any $n$-form.
The connection between Definition 3.27 and volume forms is given by the following important theorem whose proof contains a wonderful use of partitions of unity.

Theorem 3.30 A smooth manifold (Hausdorff and second-countable) is orientable iff it possesses a volume form.

Proof. First, assume that a volume form, $\omega$, exists on $M$, and say $n=\operatorname{dim} M$. For any atlas, $\left\{\left(U_{i}, \varphi_{i}\right)\right\}_{i}$, of $M$, by Proposition 3.29, each $n$-form, $\varphi_{i}^{-1^{*}} \omega$, is a volume form on $\varphi_{i}\left(U_{i}\right) \subseteq \mathbb{R}^{n}$ and

$$
\varphi_{i}^{-1^{*}} \omega=f_{i} \omega_{0}
$$

for some smooth function, $f_{i}$, never zero on $\varphi_{i}\left(U_{i}\right)$, where $\omega_{0}$ is a volume form on $\mathbb{R}^{n}$. By composing $\varphi_{i}$ with an orientation-reversing linear map if necessary, we may assume that for this new altlas, $f_{i}>0$ on $\varphi_{i}\left(U_{i}\right)$. We claim that the family $\left(U_{i}, \varphi_{i}\right)_{i}$ is an orientation atlas. This is because, on any (nonempty) overlap, $U_{i} \cap U_{j}$, as $\omega=\varphi_{j}^{*}\left(f_{j} \omega_{0}\right)$ and $\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)^{*}=\left(\varphi_{i}^{-1}\right)^{*} \circ \varphi_{j}^{*}$, we have

$$
\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)^{*}\left(f_{j} \omega_{0}\right)=f_{i} \omega_{0}
$$

and by the definition of pull-backs, we see that for every $x \in \varphi_{i}\left(U_{i} \cap U_{j}\right)$, if we let $y=\varphi_{j} \circ \varphi_{i}^{-1}(x)$, then

$$
\begin{aligned}
f_{i}(x)\left(\omega_{0}\right)_{x}\left(e_{1}, \ldots, e_{n}\right) & =\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}^{*}\left(f_{j} \omega_{0}\right)\left(e_{1}, \ldots, e_{n}\right) \\
& \left.=f_{j}(y)\left(\omega_{0}\right)_{y} d\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}\left(e_{1}\right), \ldots, d\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}\left(e_{n}\right)\right) \\
& =f_{j}(y) J\left(\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}\right)\left(\omega_{0}\right)_{y}\left(e_{1}, \ldots, e_{n}\right),
\end{aligned}
$$

where $e_{1}, \ldots, e_{n}$ is the standard basis of $\mathbb{R}^{n}$ and $J\left(\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}\right)$ 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 $J\left(\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}\right)>0$, as desired.

Conversely, assume that $J\left(\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)_{x}\right)>0$, for all $x \in \varphi_{i}\left(U_{i} \cap U_{j}\right)$, whenever $U_{i} \cap U_{j} \neq \emptyset$. We need to make a volume form on $M$. For each $U_{i}$, let

$$
\omega_{i}=\varphi_{i}^{*} \omega_{0}
$$

where $\omega_{0}$ is a volume form on $\mathbb{R}^{n}$. As $\varphi_{i}$ is a diffeomorphism, by Proposition 3.29, we see that $\omega_{i}$ is a volume form on $U_{i}$. Then, if we apply Theorem 3.26, we can find a partition of unity, $\left\{f_{i}\right\}$, subordinate to the cover $\left\{U_{i}\right\}$, with the same index set. Let,

$$
\omega=\sum_{i} f_{i} \omega_{i} .
$$

We claim that $\omega$ is a volume form on $M$.
It is clear that $\omega$ is an $n$-form on $M$. Now, since every $p \in M$ belongs to some $U_{i}$, check that on $\varphi_{i}\left(U_{i}\right)$, we have

$$
\varphi_{i}^{-1^{*}} \omega=\sum_{j \in \text { finite set }} \varphi_{i}^{-1^{*}}\left(f_{j} \omega_{j}\right)=\left(\sum_{j}\left(f_{j} \circ \varphi_{i}^{-1}\right) J\left(\varphi_{j} \circ \varphi_{i}^{-1}\right)\right) \omega_{0}
$$

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 $\varphi_{i}\left(U_{i}\right)$ 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 $\omega$ is a volume form on $M$.

Since we showed that there is a volume form on the sphere, $S^{n}$, by Theorem 3.30, the sphere $S^{n}$ is orientable. It can be shown that the projective spaces, $\mathbb{R} \mathbb{P}^{n}$, are non-orientable iff $n$ is even an thus, orientable iff $n$ is odd. In particular, $\mathbb{R} \mathbb{P}^{2}$ 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 $f: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ is a smooth submersion, then $M=f^{-1}(0)$ is a smooth orientable manifold. Another nice fact is that every Lie group is orientable.

By Proposition 3.29 (b), given any two volume forms, $\omega_{1}$ and $\omega_{2}$ on a manifold, $M$, there is a function, $f: M \rightarrow \mathbb{R}$, never 0 on $M$ such that $\omega_{2}=f \omega_{1}$. This fact suggests the following definition:

Definition 3.30 Given an orientable manifold, $M$, two volume forms, $\omega_{1}$ and $\omega_{2}$, on $M$ are equivalent iff $\omega_{2}=f \omega_{1}$ for some smooth function, $f: M \rightarrow \mathbb{R}$, such that $f(p)>0$ for all $p \in M$. An orientation of $M$ is the choice of some equivalence class of volume forms on $M$ and an oriented manifold is a manifold together with a choice of orientation. If $M$ is a manifold oriented by the volume form, $\omega$, for every $p \in M$, a basis, $\left(b_{1}, \ldots, b_{n}\right)$ of $T_{p} M$ is posively oriented iff $\omega_{p}\left(b_{1}, \ldots, b_{n}\right)>0$, else it is negatively oriented (where $n=\operatorname{dim}(M)$ ).

If $M$ is an orientable manifold, for any two volume forms $\omega_{1}$ and $\omega_{2}$ on $M$, as $\omega_{2}=f \omega_{1}$ for some function, $f$, on $M$ which is never zero, $f$ has a constant sign on every connected component of $M$. Consequently, a connected orientable manifold has two orientations.

We will also need the notion of orientation-preserving diffeomorphism.
Definition 3.31 Let $\varphi: M \rightarrow N$ be a diffeomorphism of oriented manifolds, $M$ and $N$, of dimension $n$ and say the orientation on $M$ is given by the volume form $\omega_{1}$ while the orientation on $N$ is given by the volume form $\omega_{2}$. We say that $\varphi$ is orientation preserving iff $\varphi^{*} \omega_{2}$ determines the same orientation of $M$ as $\omega_{1}$.

Using Definition 3.31 we can define the notion of a positive atlas.
Definition 3.32 If $M$ is a manifold oriented by the volume form, $\omega$, an atlas for $M$ is positive iff for every chart, $(U, \varphi)$, the diffeomorphism, $\varphi: U \rightarrow \varphi(U)$, is orientation preserving, where $U$ has the orientation induced by $M$ and $\varphi(U) \subseteq \mathbb{R}^{n}$ has the orientation induced by the standard orientation on $\mathbb{R}^{n}($ with $\operatorname{dim}(M)=n)$.

The proof of Theorem 3.30 shows
Proposition 3.31 If a manifold, $M$, has an orientation altas, then there is a uniquely determined orientation on $M$ such that this atlas is positive.

### 3.9 Covering Maps and Universal Covering Manifolds

Covering maps are an important technical tool in algebraic topology and more generally in geometry. This brief section only gives some basic definitions and states a few major facts. We apologize for his sketchy nature. Appendix A of O'Neill [119] gives a review of definitions and main results about covering manifolds. Expositions including full details can be found in Hatcher [71], Greenberg [65], Munkres [115], Fulton [56] and Massey [103, 104] (the most extensive).

We begin with covering maps.
Definition 3.33 A map, $\pi: M \rightarrow N$, between two smooth manifolds is a covering map (or cover) iff
(1) The map $\pi$ is smooth and surjective.
(2) For any $q \in N$, there is some open subset, $V \subseteq N$, so that $q \in V$ and

$$
\pi^{-1}(V)=\bigcup_{i \in I} U_{i}
$$

where the $U_{i}$ are pairwise disjoint open subsets, $U_{i} \subseteq M$, and $\pi: U_{i} \rightarrow V$ is a diffeomorphism for every $i \in I$. We say that $V$ is evenly covered.

The manifold, $M$, is called a covering manifold of $N$.
A homomorphism of coverings, $\pi_{1}: M_{1} \rightarrow N$ and $\pi_{2}: M_{2} \rightarrow N$, is a smooth map, $\varphi: M_{1} \rightarrow M_{2}$, so that

$$
\pi_{1}=\pi_{2} \circ \varphi,
$$

that is, the following diagram commutes:


We say that the coverings $\pi_{1}: M_{1} \rightarrow N$ and $\pi_{2}: M_{2} \rightarrow N$ are equivalent iff there is a homomorphism, $\varphi: M_{1} \rightarrow M_{2}$, between the two coverings and $\varphi$ is a diffeomorphism.

As usual, the inverse image, $\pi^{-1}(q)$, of any element $q \in N$ is called the fibre over $q$, the space $N$ is called the base and $M$ is called the covering space. As $\pi$ is a covering map, each fibre is a discrete space. Note that a homomorphism maps each fibre $\pi_{1}^{-1}(q)$ in $M_{1}$ to the fibre $\pi_{2}^{-1}(\varphi(q))$ in $M_{2}$, for every $q \in M_{1}$.

Proposition 3.32 Let $\pi: M \rightarrow N$ be a covering map. If $N$ is connected, then all fibres, $\pi^{-1}(q)$, have the same cardinality for all $q \in N$. Furthermore, if $\pi^{-1}(q)$ is not finite then it is countably infinite.

Proof. Pick any point, $p \in N$. We claim that the set

$$
S=\left\{q \in N| | \pi^{-1}(q)\left|=\left|\pi^{-1}(p)\right|\right\}\right.
$$

is open and closed.
If $q \in S$, then there is some open subset, $V$, with $q \in V$, so that $\pi^{-1}(V)$ is evenly covered by some family, $\left\{U_{i}\right\}_{i \in I}$, of disjoint open subsets, $U_{i}$, each diffeomorphic to $V$ under $\pi$. Then, every $s \in V$ must have a unique preimage in each $U_{i}$, so

$$
|I|=\left|\pi^{-1}(s)\right|, \quad \text { for all } s \in V
$$

However, as $q \in S,\left|\pi^{-1}(q)\right|=\left|\pi^{-1}(p)\right|$, so

$$
|I|=\left|\pi^{-1}(p)\right|=\left|\pi^{-1}(s)\right|, \quad \text { for all } s \in V
$$

and thus, $V \subseteq S$. Therefore, $S$ is open. Similary the complement of $S$ is open. As $N$ is connected, $S=N$.

Since $M$ is a manifold, it is second-countable, that is every open subset can be written as some countable union of open subsets. But then, every family, $\left\{U_{i}\right\}_{i \in I}$, of pairwise disjoint open subsets forming an even cover must be countable and since $|I|$ is the common cardinality of all the fibres, every fibre is countable.

When the common cardinality of fibres is finite it is called the multiplicity of the covering (or the number of sheets).

For any integer, $n>0$, the map, $z \mapsto z^{n}$, from the unit circle $S^{1}=\mathbf{U}(1)$ to itself is a covering with $n$ sheets. The map,

$$
t: \mapsto(\cos (2 \pi t), \sin (2 \pi t))
$$

is a covering, $\mathbb{R} \rightarrow S^{1}$, with infinitely many sheets.
It is also useful to note that a covering map, $\pi: M \rightarrow N$, is a local diffeomorphism (which means that $d \pi_{p}: T_{p} M \rightarrow T_{\pi(p)} N$ is a bijective linear map for every $\left.p \in M\right)$. Indeed, given any $p \in M$, if $q=\pi(p)$, then there is some open subset, $V \subseteq N$, containing $q$ so that $V$ is evenly covered by a family of disjoint open subsets, $\left\{U_{i}\right\}_{i \in I}$, with each $U_{i} \subseteq M$ diffeomorphic to $V$ under $\pi$. As $p \in U_{i}$ for some $i$, we have a diffeomorphism, $\pi \upharpoonright U_{i}: U_{i} \longrightarrow V$, as required.

The crucial property of covering manifolds is that curves in $N$ can be lifted to $M$, in a unique way. For any map, $\varphi: P \rightarrow N$, a lift of $\varphi$ through $\pi$ is a map, $\widetilde{\varphi}: P \rightarrow M$, so that

$$
\varphi=\pi \circ \widetilde{\varphi},
$$

as in the following commutative diagram:


We state without proof the following results:

Proposition 3.33 If $\pi: M \rightarrow N$ is a covering map, then for every smooth curve, $\alpha: I \rightarrow N$, in $N$ (with $0 \in I$ ) and for any point, $q \in M$, such that $\pi(q)=\alpha(0)$, there is a unique smooth curve, $\widetilde{\alpha}: I \rightarrow M$, lifting $\alpha$ through $\pi$ such that $\widetilde{\alpha}(0)=q$.

Proposition 3.34 Let $\pi: M \rightarrow N$ be a covering map and let $\varphi: P \rightarrow N$ be a smooth map. For any $p_{0} \in P$, any $q_{0} \in M$ and any $r_{0} \in N$ with $\pi\left(q_{0}\right)=\varphi\left(p_{0}\right)=r_{0}$, the following properties hold:
(1) If $P$ is connected then there is at most one lift, $\widetilde{\varphi}: P \rightarrow M$, of $\varphi$ through $\pi$ such that $\widetilde{\varphi}\left(p_{0}\right)=q_{0}$.
(2) If $P$ is simply connected, then such a lift exists.


Theorem 3.35 Every connected manifold, $M$, possesses a simply connected covering map, $\pi: \widetilde{M} \rightarrow M$, that is, with $\widetilde{M}$ simply connected. Any two simply connected coverings of $N$ are equivalent.

In view of Theorem 3.35, it is legitimate to speak of the simply connected cover, $\widetilde{M}$, of $M$, also called universal covering (or cover) of $M$.

Given any point, $p \in M$, let $\pi_{1}(M, p)$ denote the fundamental group of $M$ with basepoint $p$ (see any of the references listed above, in particular, Massey [103, 104]). If $\varphi: M \rightarrow N$ is a smooth map, for any $p \in M$, if we write $q=\varphi(p)$, then we have an induced group homomorphism

$$
\varphi_{*}: \pi_{1}(M, p) \rightarrow \pi_{1}(N, q) .
$$

Proposition 3.36 If $\pi: M \rightarrow N$ is a covering map, for every $p \in M$, if $q=\pi(p)$, then the induced homomorphism, $\pi_{*}: \pi_{1}(M, p) \rightarrow \pi_{1}(N, q)$, is injective.

The next proposition is a stronger version of part (2) of Proposition 3.34:
Proposition 3.37 Let $\pi: M \rightarrow N$ be a covering map and let $\varphi: P \rightarrow N$ be a smooth map. For any $p_{0} \in P$, any $q_{0} \in M$ and any $r_{0} \in N$ with $\pi\left(q_{0}\right)=\varphi\left(p_{0}\right)=r_{0}$, if $P$ is connected, then a lift, $\widetilde{\varphi}: P \rightarrow M$, of $\varphi$ such that $\widetilde{\varphi}\left(p_{0}\right)=q_{0}$ exists iff

$$
\varphi_{*}\left(\pi_{1}\left(P, p_{0}\right)\right) \subseteq \pi_{*}\left(\pi_{1}\left(M, q_{0}\right)\right)
$$

as illustrated in the diagram below

iff


Basic Assumption: For any covering, $\pi: M \rightarrow N$, if $N$ is connected then we also assume that $M$ is connected.

Using Proposition 3.36, we get
Proposition 3.38 If $\pi: M \rightarrow N$ is a covering map and $N$ is simply connected, then $\pi$ is a diffeomorphism (recall that $M$ is connected); thus, $M$ is diffeomorphic to the universal cover, $\widetilde{N}$, of $N$.

Proof. Pick any $p \in M$ and let $q=\varphi(p)$. As $N$ is simply connected, $\pi_{1}(N, q)=(0)$. By Proposition 3.36, since $\pi_{*}: \pi_{1}(M, p) \rightarrow \pi_{1}(N, q)$ is injective, $\pi_{1}(M, p)=(0)$ so $M$ is simply connected (by hypothesis, $M$ is connected). But then, by Theorem 3.35, $M$ and $N$ are diffeomorphic.

The following proposition shows that the universal covering of a space covers every other covering of that space. This justifies the terminology "universal covering".

Proposition 3.39 Say $\pi_{1}: M_{1} \rightarrow N$ and $\pi_{2}: M_{2} \rightarrow N$ are two coverings of $N$, with $N$ connected. Every homomorphism, $\varphi: M_{1} \rightarrow M_{2}$, between these two coverings is a covering map. As a consequence, if $\pi: \widetilde{N} \rightarrow N$ is a universal covering of $N$, then for every covering, $\pi^{\prime}: M \rightarrow N$, of $N$, there is a covering, $\varphi: \widetilde{N} \rightarrow M$, of $M$.

The notion of deck-transformation group of a covering is also useful because it yields a way to compute the fundamental group of the base space.

Definition 3.34 If $\pi: M \rightarrow N$ is a covering map, a deck-transformation is any diffeomorphism, $\varphi: M \rightarrow M$, such that $\pi=\pi \circ \varphi$, that is, the following diagram commutes:


Note that deck-transformations are just automorphisms of the covering map. The commutative diagram of Definition 3.34 means that a deck transformation permutes every fibre. It is immediately verified that the set of deck transformations of a covering map is a group denoted $\Gamma_{\pi}$ (or simply, $\Gamma$ ), called the deck-transformation group of the covering.

Observe that any deck transformation, $\varphi$, is a lift of $\pi$ through $\pi$. Consequently, if $M$ is connected, by Proposition 3.34 (1), every deck-transformation is determined by its value at
a single point. So, the deck-transformations are determined by their action on each point of any fixed fibre, $\pi^{-1}(q)$, with $q \in N$. Since the fibre $\pi^{-1}(q)$ is countable, $\Gamma$ is also countable, that is, a discrete Lie group. Moreover, if $M$ is compact, as each fibre, $\pi^{-1}(q)$, is compact and discrete, it must be finite and so, the deck-transformation group is also finite.

The following proposition gives a useful method for determining the fundamental group of a manifold.

Proposition 3.40 If $\pi: \widetilde{M} \rightarrow M$ is the universal covering of a connected manifold, $M$, then the deck-transformation group, $\widetilde{\Gamma}$, is isomorphic to the fundamental group, $\pi_{1}(M)$, of $M$.

Remark: When $\pi: \widetilde{M} \rightarrow M$ is the universal covering of $M$, it can be shown that the group $\widetilde{\Gamma}$ acts simply and transitively on every fibre, $\pi^{-1}(q)$. This means that for any two elements, $x, y \in \pi^{-1}(q)$, there is a unique deck-transformation, $\varphi \in \widetilde{\Gamma}$ such that $\varphi(x)=y$. So, there is a bijection between $\pi_{1}(M) \cong \widetilde{\Gamma}$ and the fibre $\pi^{-1}(q)$.

Proposition 3.35 together with previous observations implies that if the universal cover of a connected (compact) manifold is compact, then $M$ has a finite fundamental group. We will use this fact later, in particular, in the proof of Myers' Theorem.

## Chapter 4

## Construction of Manifolds From Gluing Data

### 4.1 Sets of Gluing Data for Manifolds

The definition of a manifold given in Chapter 3 assumes that the underlying set, $M$, is already known. However, there are situations where we only have some indirect information about the overlap of the domains, $U_{i}$, of the local charts defining our manifold, $M$, in terms of the transition functions,

$$
\varphi_{j i}: \varphi_{i}\left(U_{i} \cap U_{j}\right) \rightarrow \varphi_{j}\left(U_{i} \cap U_{j}\right)
$$

but where $M$ itself is not known. For example, this situation happens when trying to construct a surface approximating a 3D-mesh. If we let $\Omega_{i j}=\varphi_{i}\left(U_{i} \cap U_{j}\right)$ and $\Omega_{j i}=$ $\varphi_{j}\left(U_{i} \cap U_{j}\right)$, then $\varphi_{j i}$ can be viewed as a "gluing map",

$$
\varphi_{j i}: \Omega_{i j} \rightarrow \Omega_{j i}
$$

between two open subets of $\Omega_{i}$ and $\Omega_{j}$, respectively.
For technical reasons, it is desirable to assume that the images, $\Omega_{i}=\varphi_{i}\left(U_{i}\right)$ and $\Omega_{j}=$ $\varphi_{j}\left(U_{j}\right)$, of distinct charts are disjoint but this can always be achieved for manifolds. Indeed, the map

$$
\beta:\left(x_{1}, \ldots, x_{n}\right) \mapsto\left(\frac{x_{1}}{\sqrt{1+\sum_{i=1}^{n} x_{i}^{2}}}, \ldots, \frac{x_{n}}{\sqrt{1+\sum_{i=1}^{n} x_{i}^{2}}}\right)
$$

is a smooth diffeomorphism from $\mathbb{R}^{n}$ to the open unit ball $B(0,1)$ with inverse given by

$$
\beta^{-1}:\left(x_{1}, \ldots, x_{n}\right) \mapsto\left(\frac{x_{1}}{\sqrt{1-\sum_{i=1}^{n} x_{i}^{2}}}, \ldots, \frac{x_{n}}{\sqrt{1-\sum_{i=1}^{n} x_{i}^{2}}}\right)
$$

Since $M$ has a countable basis, using compositions of $\beta$ with suitable translations, we can make sure that the $\Omega_{i}$ 's are mapped diffeomorphically to disjoint open subsets of $\mathbb{R}^{n}$.

Remarkably, manifolds can be constructed using the "gluing process" alluded to above from what is often called sets of "gluing data." In this chapter, we are going to describe this construction and prove its correctness in details, provided some mild assumptions on the gluing data. It turns out that this procedure for building manifolds can be made practical. Indeed, it is the basis of a class of new methods for approximating 3D meshes by smooth surfaces, see Siqueira, Xu and Gallier [140].

It turns out that care must be exercised to ensure that the space obtained by gluing the pieces $\Omega_{i j}$ and $\Omega_{j i}$ is Hausdorff. Some care must also be exercised in formulating the consistency conditions relating the $\varphi_{j i}$ 's (the so-called "cocycle condition"). This is because the traditional condition (for example, in bundle theory) has to do with triple overlaps of the $U_{i}=\varphi_{i}^{-1}\left(\Omega_{i}\right)$ on the manifold, $M$, (see Chapter 7, especially Theorem 7.4) but in our situation, we do not have $M$ nor the parametrization maps $\theta_{i}=\varphi_{i}^{-1}$ and the cocycle condition on the $\varphi_{j i}$ 's has to be stated in terms of the $\Omega_{i}$ 's and the $\Omega_{j i}$ 's.

Finding an easily testable necessary and sufficient criterion for the Hausdorff condition appears to be a very difficult problem. We propose a necessary and sufficient condition, but it is not easily testable in general. If $M$ is a manifold, then observe that difficulties may arise when we want to separate two distinct point, $p, q \in M$, such that $p$ and $q$ neither belong to the same open, $\theta_{i}\left(\Omega_{i}\right)$, nor to two disjoint opens, $\theta_{i}\left(\Omega_{i}\right)$ and $\theta_{j}\left(\Omega_{j}\right)$, but instead, to the boundary points in $\left(\partial\left(\theta_{i}\left(\Omega_{i j}\right)\right) \cap \theta_{i}\left(\Omega_{i}\right)\right) \cup\left(\partial\left(\theta_{j}\left(\Omega_{j i}\right)\right) \cap \theta_{j}\left(\Omega_{j}\right)\right)$. In this case, there are some disjoint open subsets, $U_{p}$ and $U_{q}$, of $M$ with $p \in U_{p}$ and $q \in U_{q}$, and we get two disjoint open subsets, $V_{x}=\theta_{i}^{-1}\left(U_{p}\right) \subseteq \Omega_{i}$ and $V_{y}=\theta_{j}^{-1}\left(U_{q}\right) \subseteq \Omega_{j}$, with $\theta_{i}(x)=p, \theta_{j}(y)=q$, and such that $x \in \partial\left(\Omega_{i j}\right) \cap \Omega_{i}, y \in \partial\left(\Omega_{j i}\right) \cap \Omega_{j}$, and no point in $V_{y} \cap \Omega_{j i}$ is the image of any point in $V_{x} \cap \Omega_{i j}$ by $\varphi_{j i}$. Since $V_{x}$ and $V_{y}$ are open, we may assume that they are open balls. This necessary condition turns out to be also sufficient.

With the above motivations in mind, here is the definition of sets of gluing data.
Definition 4.1 Let $n$ be an integer with $n \geq 1$ and let $k$ be either an integer with $k \geq 1$ or $k=\infty$. A set of gluing data is a triple, $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$, satisfying the following properties, where $I$ is a (nonempty) countable set:
(1) For every $i \in I$, the set $\Omega_{i}$ is a nonempty open subset of $\mathbb{R}^{n}$ called a parametrization domain, for short, p-domain, and the $\Omega_{i}$ are pairwise disjoint (i.e., $\Omega_{i} \cap \Omega_{j}=\emptyset$ for all $i \neq j$ ).
(2) For every pair $(i, j) \in I \times I$, the set $\Omega_{i j}$ is an open subset of $\Omega_{i}$. Furthermore, $\Omega_{i i}=\Omega_{i}$ and $\Omega_{i j} \neq \emptyset$ iff $\Omega_{j i} \neq \emptyset$. Each nonempty $\Omega_{i j}($ with $i \neq j)$ is called a gluing domain.
(3) If we let

$$
K=\left\{(i, j) \in I \times I \mid \Omega_{i j} \neq \emptyset\right\}
$$

then $\varphi_{j i}: \Omega_{i j} \rightarrow \Omega_{j i}$ is a $C^{k}$ bijection for every $(i, j) \in K$ called a transition function (or gluing function) and the following condition holds:
(c) For all $i, j, k$, if $\Omega_{j i} \cap \Omega_{j k} \neq \emptyset$, then $\varphi_{j i}^{-1}\left(\Omega_{j i} \cap \Omega_{j k}\right) \subseteq \Omega_{i k}$ and

$$
\varphi_{k i}(x)=\varphi_{k j} \circ \varphi_{j i}(x), \quad \text { for all } \quad x \in \varphi_{j i}^{-1}\left(\Omega_{j i} \cap \Omega_{j k}\right)
$$

Condition (c) is called the cocycle condition.
(4) For every pair $(i, j) \in K$, with $i \neq j$, for every $x \in \partial\left(\Omega_{i j}\right) \cap \Omega_{i}$ and every $y \in \partial\left(\Omega_{j i}\right) \cap \Omega_{j}$, there are open balls, $V_{x}$ and $V_{y}$ centered at $x$ and $y$, so that no point of $V_{y} \cap \Omega_{j i}$ is the image of any point of $V_{x} \cap \Omega_{i j}$ by $\varphi_{j i}$.

## Remarks.

(1) In practical applications, the index set, $I$, is of course finite and the open subsets, $\Omega_{i}$, may have special properties (for example, connected; open simplicies, etc.).
(2) We are only interested in the $\Omega_{i j}$ 's that are nonempty but empty $\Omega_{i j}$ 's do arise in proofs and constructions and this is why our definition allows them.
(3) Observe that $\Omega_{i j} \subseteq \Omega_{i}$ and $\Omega_{j i} \subseteq \Omega_{j}$. If $i \neq j$, as $\Omega_{i}$ and $\Omega_{j}$ are disjoint, so are $\Omega_{i j}$ and $\Omega_{i j}$.
(4) The cocycle condition (c) may seem overly complicated but it is actually needed to guarantee the transitivity of the relation, $\sim$, defined in the proof of Proposition 4.1. Flawed versions of condition (c) appear in the literature, see the discussion after the proof of Proposition 4.1. The problem is that $\varphi_{k j} \circ \varphi_{j i}$ is a partial function whose domain, $\varphi_{j i}^{-1}\left(\Omega_{j i} \cap \Omega_{j k}\right)$, is not necessarily related to the domain, $\Omega_{i k}$, of $\varphi_{k i}$. To ensure the transitivity of $\sim$, we must assert that whenever the composition $\varphi_{k j} \circ \varphi_{j i}$ has a nonempty domain, this domain is contained in the domain of $\varphi_{k i}$ and that $\varphi_{k j} \circ \varphi_{j i}$ and $\varphi_{k i}$ agree. Since the $\varphi_{j i}$ are bijections, condition (c) implies the following conditions:
(a) $\varphi_{i i}=\operatorname{id}_{\Omega_{i}}$, for all $i \in I$.
(b) $\varphi_{i j}=\varphi_{j i}^{-1}$, for all $(i, j) \in K$.

To get (a), set $i=j=k$. Then, (b) follows from (a) and (c) by setting $k=i$.
(5) If $M$ is a $C^{k}$ manifold (including $k=\infty$ ), then using the notation of our introduction, it is easy to check that the open sets $\Omega_{i}, \Omega_{i j}$ and the gluing functions, $\varphi_{j i}$, satisfy the conditions of Definition 4.1 (provided that we fix the charts so that the images of distinct charts are disjoint). Proposition 4.1 will show that a manifold can be reconstructed from a set of guing data.

The idea of defining gluing data for manifolds is not new. André Weil introduced this idea to define abstract algebraic varieties by gluing irreducible affine sets in his book [148] published in 1946. The same idea is well-known in bundle theory and can be found in
standard texts such as Steenrod [141], Bott and Tu [19], Morita [114] and Wells [150] (the construction of a fibre bundle from a cocycle is given in Chapter 7, see Theorem 7.4).

The beauty of the idea is that it allows the reconstruction of a manifold, $M$, without having prior knowledge of the topology of this manifold (that is, without having explicitly the underlying topological space $M$ ) by gluing open subets of $\mathbb{R}^{n}$ (the $\Omega_{i}$ 's) according to prescribed gluing instructions (namely, glue $\Omega_{i}$ and $\Omega_{j}$ by identifying $\Omega_{i j}$ and $\Omega_{j i}$ using $\varphi_{j i}$ ). This method of specifying a manifold separates clearly the local structure of the manifold (given by the $\Omega_{i}$ 's) from its global structure which is specified by the gluing functions. Furthermore, this method ensures that the resulting manifold is $C^{k}$ (even for $k=\infty$ ) with no extra effort since the gluing functions $\varphi_{j i}$ are assumed to be $C^{k}$.

Grimm and Hughes $[67,68]$ appear to be the first to have realized the power of this latter property for practical applications and we wish to emphasize that this is a very significant discovery. However, Grimm [67] uses a condition stronger than our condition (4) to ensure that the resulting space is Hausdorff. The cocycle condition in Grimm and Hughes [67, 68] is also not strong enough to ensure transitivity of the relation $\sim$. We will come back to these points after the proof of Proposition 4.1.

Working with overlaps of open subsets of the parameter domain makes it much easier to enforce smoothness conditions compared to the traditional approach with splines where the parameter domain is subdivided into closed regions and where enforcing smoothness along boundaries is much more difficult.

Let us show that a set of gluing data defines a $C^{k}$ manifold in a natural way.
Proposition 4.1 For every set of gluing data, $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$, there is an $n$-dimensional $C^{k}$ manifold, $M_{\mathcal{G}}$, whose transition functions are the $\varphi_{j i}$ 's.

Proof. Define the binary relation, $\sim$, on the disjoint union, $\coprod_{i \in I} \Omega_{i}$, of the open sets, $\Omega_{i}$, as follows: For all $x, y \in \coprod_{i \in I} \Omega_{i}$,

$$
x \sim y \quad \text { iff } \quad(\exists(i, j) \in K)\left(x \in \Omega_{i j}, y \in \Omega_{j i}, y=\varphi_{j i}(x)\right)
$$

Note that if $x \sim y$ and $x \neq y$, then $i \neq j$, as $\varphi_{i i}=i d$. But then, as $x \in \Omega_{i j} \subseteq \Omega_{i}$, $y \in \Omega_{j i} \subseteq \Omega_{j}$ and $\Omega_{i} \cap \Omega_{j}=\emptyset$ when $i \neq j$, if $x \sim y$ and $x, y \in \Omega_{i}$, then $x=y$.

We claim that $\sim$ is an equivalence relation. This follows easily from the cocycle condition but to be on the safe side, we provide the crucial step of the proof. Clearly, condition (a) ensures reflexivity and condition (b) ensures symmetry. The crucial step is to check transitivity. Assume that $x \sim y$ and $y \sim z$. Then, there are some $i, j, k$ such that
(i) $x \in \Omega_{i j}, y \in \Omega_{j i} \cap \Omega_{j k}, z \in \Omega_{k j}$ and
(ii) $y=\varphi_{j i}(x)$ and $z=\varphi_{k j}(y)$.

Consequently, $\Omega_{j i} \cap \Omega_{j k} \neq \emptyset$ and $x \in \varphi_{j i}^{-1}\left(\Omega_{j i} \cap \Omega_{j k}\right)$, so by (c), we get $\varphi_{j i}^{-1}\left(\Omega_{j i} \cap \Omega_{j k}\right) \subseteq \Omega_{i k}$ and thus, $\varphi_{k i}(x)$ is defined and by (c) again,

$$
\varphi_{k i}(x)=\varphi_{k j} \circ \varphi_{j i}(x)=z,
$$

that is, $x \sim z$, as desired.
Since $\sim$ is an equivalence relation let

$$
M_{\mathcal{G}}=\left(\coprod_{i \in I} \Omega_{i}\right) / \sim
$$

be the quotient set and let $p: \coprod_{i \in I} \Omega_{i} \rightarrow M_{\mathcal{G}}$ be the quotient map, with $p(x)=[x]$, where $[x]$ denotes the equivalence class of $x$. Also, for every $i \in I$, let $\mathrm{in}_{i}: \Omega_{i} \rightarrow \coprod_{i \in I} \Omega_{i}$ be the natural injection and let

$$
\tau_{i}=p \circ \operatorname{in}_{i}: \Omega_{i} \rightarrow M_{\mathcal{G}}
$$

Since we already noted that if $x \sim y$ and $x, y \in \Omega_{i}$, then $x=y$, we conclude that every $\tau_{i}$ is injective.

We give $M_{\mathcal{G}}$ the coarsest topology that makes the bijections, $\tau_{i}: \Omega_{i} \rightarrow \tau_{i}\left(\Omega_{i}\right)$, into homeomorphisms. Then, if we let $U_{i}=\tau_{i}\left(\Omega_{i}\right)$ and $\varphi_{i}=\tau_{i}^{-1}$, it is immediately verified that the $\left(U_{i}, \varphi_{i}\right)$ are charts and this collection of charts forms a $C^{k}$ atlas for $M_{\mathcal{G}}$. As there are countably many charts, $M_{\mathcal{G}}$ is second-countable. Therefore, for $M_{\mathcal{G}}$ to be a manifold it only remains to check that the topology is Hausdorff. For this, we use the following:

Claim. For all $(i, j) \in I \times I$, we have $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \neq \emptyset$ iff $(i, j) \in K$ and if so,

$$
\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

Proof of Claim. Assume that $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \neq \emptyset$ and let $[z] \in \tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)$. Observe that $[z] \in \tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)$ iff $z \sim x$ and $z \sim y$, for some $x \in \Omega_{i}$ and some $y \in \Omega_{j}$. Consequently, $x \sim y$, which implies that $(i, j) \in K, x \in \Omega_{i j}$ and $y \in \Omega_{j i}$.

We have $[z] \in \tau_{i}\left(\Omega_{i j}\right)$ iff $z \sim x$ for some $x \in \Omega_{i j}$. Then, either $i=j$ and $z=x$ or $i \neq j$ and $z \in \Omega_{j i}$, which shows that $[z] \in \tau_{j}\left(\Omega_{j i}\right)$ and so,

$$
\tau_{i}\left(\Omega_{i j}\right) \subseteq \tau_{j}\left(\Omega_{j i}\right)
$$

Since the same argument applies by interchanging $i$ and $j$, we have

$$
\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

for all $(i, j) \in K$. Since $\Omega_{i j} \subseteq \Omega_{i}, \Omega_{j i} \subseteq \Omega_{j}$ and $\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)$ for all $(i, j) \in K$, we have

$$
\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right) \subseteq \tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)
$$

for all $(i, j) \in K$.

For the reverse inclusion, if $[z] \in \tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)$, then we know that there is some $x \in \Omega_{i j}$ and some $y \in \Omega_{j i}$ such that $z \sim x$ and $z \sim y$, so $[z]=[x] \in \tau_{i}\left(\Omega_{i j}\right),[z]=[y] \in \tau_{j}\left(\Omega_{j i}\right)$ and we get

$$
\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \subseteq \tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

This proves that if $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \neq \emptyset$, then $(i, j) \in K$ and

$$
\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

Finally, assume $(i, j) \in K$. Then, for any $x \in \Omega_{i j} \subseteq \Omega_{i}$, we have $y=\varphi_{j i}(x) \in \Omega_{j i} \subseteq \Omega_{j}$ and $x \sim y$, so that $\tau_{i}(x)=\tau_{j}(y)$, which proves that $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \neq \emptyset$ and our claim is proved.

We now prove that the topology of $M_{\mathcal{G}}$ is Hausdorff. Pick $[x],[y] \in M_{\mathcal{G}}$ with $[x] \neq[y]$, for some $x \in \Omega_{i}$ and some $y \in \Omega_{j}$. Either $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\emptyset$, in which case, as $\tau_{i}$ and $\tau_{j}$ are homeomorphisms, $[x]$ and $[y]$ belong to the two disjoint open sets $\tau_{i}\left(\Omega_{i}\right)$ and $\tau_{j}\left(\Omega_{j}\right)$. If not, then by the Claim, $(i, j) \in K$ and

$$
\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

There are several cases to consider:
(a) If $i=j$, then $x$ and $y$ can be separated by disjoint opens, $V_{x}$ and $V_{y}$, and as $\tau_{i}$ is a homeomorphism, $[x]$ and $[y]$ are separated by the disjoint open subsets $\tau_{i}\left(V_{x}\right)$ and $\tau_{i}\left(V_{y}\right)$.
(b) If $i \neq j, x \in \Omega_{i}-\overline{\Omega_{i j}}$ and $y \in \Omega_{j}-\overline{\Omega_{j i}}$, then $\tau_{i}\left(\Omega_{i}-\overline{\Omega_{i j}}\right)$ and $\tau_{j}\left(\Omega_{j}-\overline{\Omega_{j i}}\right)$ are disjoint opens subsets separating $[x]$ and $[y]$.
(c) If $i \neq j, x \in \Omega_{i j}$ and $y \in \Omega_{j i}$, as $[x] \neq[y]$ and $y \sim \varphi_{i j}(y)$, then $x \neq \varphi_{i j}(y)$. We can separate $x$ and $\varphi_{i j}(y)$ by disjoint open subsets, $V_{x}$ and $V_{y}$ and $[x]$ and $[y]=\left[\varphi_{i j}(y)\right]$ are separated by the disjoint open subsets $\tau_{i}\left(V_{x}\right)$ and $\tau_{i}\left(V_{y}\right)$.
(d) If $i \neq j, x \in \partial\left(\Omega_{i j}\right) \cap \Omega_{i}$ and $y \in \partial\left(\Omega_{j i}\right) \cap \Omega_{j}$, then we use condition (4). This condition yields two disjoint open subsets $V_{x}$ and $V_{y}$ with $x \in V_{x}$ and $y \in V_{y}$ such that no point of $V_{x} \cap \Omega_{i j}$ is equivalent to any point of $V_{y} \cap \Omega_{j i}$, and so, $\tau_{i}\left(V_{x}\right)$ and $\tau_{j}\left(V_{y}\right)$ are disjoint open subsets separating $[x]$ and $[y]$.

Therefore, the topology of $M_{\mathcal{G}}$ is Hausdorff and $M_{\mathcal{G}}$ is indeed a manifold.
Finally, it is trivial to verify that the transition functions of $M_{\mathcal{G}}$ are the original gluing functions, $\varphi_{i j}$.

It should be noted that as nice as it is, Proposition 4.1 is a theoretical construction that yields an "abstract" manifold but does not yield any information as to the geometry of this manifold. Furthermore, the resulting manifold may not be orientable or compact, even if we start with a finite set of $p$-domains.

Here is an example showing that if condition (4) of Definition 4.1 is omitted then we may get non-Hausdorff spaces. Cindy Grimm uses a similar example in her dissertation [67] (Appendix C2, page 126), but her presentation is somewhat confusing because her $\Omega_{1}$ and $\Omega_{2}$ appear to be two disjoint copies of the real line in $\mathbb{R}^{2}$, but these are not open in $\mathbb{R}^{2}$ !

Let $\Omega_{1}=(-3,-1), \Omega_{2}=(1,3), \Omega_{12}=(-3,-2), \Omega_{21}=(1,2)$ and $\varphi_{21}(x)=x+4$. The resulting space, $M$, is a curve looking like a "fork", and the problem is that the images of -2 and 2 in $M$, which are distinct points of $M$, cannot be separated. Indeed, the images of any two open intervals $(-2-\epsilon,-2+\epsilon)$ and $(2-\eta, 2+\eta)$ (for $\epsilon, \eta>0$ ) always intersect since $(-2-\min (\epsilon, \eta),-2)$ and $(2-\min (\epsilon, \eta), 2)$ are identified. Clearly, condition (4) fails.

Cindy Grimm [67] (page 40) uses a condition stronger than our condition (4) to ensure that the quotient, $M_{\mathcal{G}}$ is Hausdorff, namely, that for all $(i, j) \in K$ with $i \neq j$, the quotient $\left(\Omega_{i} \amalg \Omega_{j}\right) / \sim$ should be embeddable in $\mathbb{R}^{n}$. This is a rather strong condition that prevents obtaining a 2 -sphere by gluing two open discs in $\mathbb{R}^{2}$ along an annulus (see Grimm [67], Appendix C2, page 126).

Grimm uses the following cocycle condition in [67] (page 40) and [68] (page 361):
(c') For all $x \in \Omega_{i j} \cap \Omega_{i k}$,

$$
\varphi_{k i}(x)=\varphi_{k j} \circ \varphi_{j i}(x)
$$

This condition is not strong enough to imply transitivity of the relation $\sim$, as shown by the following counter-example:

Let $\Omega_{1}=(0,3), \Omega_{2}=(4,5), \Omega_{3}=(6,9), \Omega_{12}=(0,1), \Omega_{13}=(2,3), \Omega_{21}=\Omega_{23}=(4,5)$, $\Omega_{32}=(8,9), \Omega_{31}=(6,7), \varphi_{21}(x)=x+4, \varphi_{32}(x)=x+4$ and $\varphi_{31}(x)=x+4$.

Note that the pairwise gluings yield Hausdorff spaces. Obviously, $\varphi_{32} \circ \varphi_{21}(x)=x+8$, for all $x \in \Omega_{12}$, but $\Omega_{12} \cap \Omega_{13}=\emptyset$. Thus, $0.5 \sim 4.5 \sim 8.5$, but $0.5 \nsim 8.5$ since $\varphi_{31}(0.5)$ is undefined.

Here is another counter-example in which $\Omega_{12} \cap \Omega_{13} \neq \emptyset$, using a disconnected open, $\Omega_{2}$.
Let $\Omega_{1}=(0,3), \Omega_{2}=(4,5) \cup(6,7), \Omega_{3}=(8,11), \Omega_{12}=(0,1) \cup(2,3), \Omega_{13}=(2,3)$, $\Omega_{21}=\Omega_{23}=(4,5) \cup(6,7), \Omega_{32}=(8,9) \cup(10,11), \Omega_{31}=(8,9), \varphi_{21}(x)=x+4, \varphi_{32}(x)=x+2$ on $(6,7), \varphi_{32}(x)=x+6$ on $(4,5), \varphi_{31}(x)=x+6$.

Note that the pairwise gluings yield Hausdorff spaces. Obviously, $\varphi_{32} \circ \varphi_{21}(x)=x+6=$ $\varphi_{31}(x)$ for all $x \in \Omega_{12} \cap \Omega_{13}=(2,3)$. Thus, $0.5 \sim 4.5 \sim 8.5$, but $0.5 \nsim 8.5$ since $\varphi_{31}(0.5)$ is undefined.

It is possible to give a construction, in the case of a surface, which builds a compact manifold whose geometry is "close" to the geometry of a prescribed 3D-mesh (see Siqueira, Xu and Gallier [140]). Actually, we are not able to guarantee, in general, that the parametrization functions, $\theta_{i}$, that we obtain are injective, but we are not aware of any algorithm that achieves this.

Given a set of gluing data, $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$, it is natural to consider the collection of manifolds, $M$, parametrized by maps, $\theta_{i}: \Omega_{i} \rightarrow M$, whose domains are the $\Omega_{i}$ 's and whose transitions functions are given by the $\varphi_{j i}$, that is, such that

$$
\varphi_{j i}=\theta_{j}^{-1} \circ \theta_{i}
$$

We will say that such manifolds are induced by the set of gluing data, $\mathcal{G}$.
The proof of Proposition 4.1 shows that the parametrization maps, $\tau_{i}$, satisfy the property: $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \neq \emptyset$ iff $(i, j) \in K$ and if so,

$$
\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

Furthermore, they also satisfy the consistency condition:

$$
\tau_{i}=\tau_{j} \circ \varphi_{j i}
$$

for all $(i, j) \in K$. If $M$ is a manifold induced by the set of gluing data, $\mathcal{G}$, because the $\theta_{i}$ 's are injective and $\varphi_{j i}=\theta_{j}^{-1} \circ \theta_{i}$, the two properties stated above for the $\tau_{i}$ 's also hold for the $\theta_{i}$ 's. We will see in Section 4.2 that the manifold, $M_{\mathcal{G}}$, is a "universal" manifold induced by $\mathcal{G}$ in the sense that every manifold induced by $\mathcal{G}$ is the image of $M_{\mathcal{G}}$ by some $C^{k}$ map.

Interestingly, it is possible to characterize when two manifolds induced by the same set of gluing data are isomorphic in terms of a condition on their transition functions.

Proposition 4.2 Given any set of gluing data, $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$, for any two manifolds $M$ and $M^{\prime}$ induced by $\mathcal{G}$ given by families of parametrizations $\left(\Omega_{i}, \theta_{i}\right)_{i \in I}$ and $\left(\Omega_{i}, \theta_{i}^{\prime}\right)_{i \in I}$, respectively, if $f: M \rightarrow M^{\prime}$ is a $C^{k}$ isomorphism, then there are $C^{k}$ bijections, $\rho_{i}: W_{i j} \rightarrow W_{i j}^{\prime}$, for some open subsets $W_{i j}, W_{i j}^{\prime} \subseteq \Omega_{i}$, such that

$$
\varphi_{j i}^{\prime}(x)=\rho_{j} \circ \varphi_{j i} \circ \rho_{i}^{-1}(x), \quad \text { for all } \quad x \in W_{i j}
$$

with $\varphi_{j i}=\theta_{j}^{-1} \circ \theta_{i}$ and $\varphi_{j i}^{\prime}=\theta_{j}^{\prime-1} \circ \theta_{i}^{\prime}$. Furthermore, $\rho_{i}=\left(\theta_{i}^{\prime-1} \circ f \circ \theta_{i}\right) \upharpoonright W_{i j}$ and if $\theta_{i}^{\prime-1} \circ f \circ \theta_{i}$ is a bijection from $\Omega_{i}$ to itself and $\theta_{i}^{\prime-1} \circ f \circ \theta_{i}\left(\Omega_{i j}\right)=\Omega_{i j}$, for all $i, j$, then $W_{i j}=W_{i, j}^{\prime}=\Omega_{i}$.

Proof. The composition $\theta_{i}^{\prime-1} \circ f \circ \theta_{i}$ is actually a partial function with domain

$$
\operatorname{dom}\left(\theta_{i}^{\prime-1} \circ f \circ \theta_{i}\right)=\left\{x \in \Omega_{i} \mid \theta_{i}(x) \in f^{-1} \circ \theta_{i}^{\prime}\left(\Omega_{i}\right)\right\}
$$

and its "inverse" $\theta_{i}^{-1} \circ f^{-1} \circ \theta_{i}^{\prime}$ is a partial function with domain

$$
\operatorname{dom}\left(\theta_{i}^{-1} \circ f^{-1} \circ \theta_{i}^{\prime}\right)=\left\{x \in \Omega_{i} \mid \theta_{i}^{\prime}(x) \in f \circ \theta_{i}\left(\Omega_{i}\right)\right\}
$$

The composition $\theta_{j}^{\prime-1} \circ f \circ \theta_{j} \circ \varphi_{j i} \circ \theta_{i}^{-1} \circ f^{-1} \circ \theta_{i}^{\prime}$ is also a partial function and we let

$$
W_{i j}=\Omega_{i j} \cap \operatorname{dom}\left(\theta_{j}^{\prime-1} \circ f \circ \theta_{j} \circ \varphi_{j i} \circ \theta_{i}^{-1} \circ f^{-1} \circ \theta_{i}^{\prime}\right), \quad \rho_{i}=\left(\theta_{i}^{\prime-1} \circ f \circ \theta_{i}\right) \upharpoonright W_{i j}
$$

and $W_{i j}^{\prime}=\rho_{i}\left(W_{i j}\right)$. Observe that $\theta_{j} \circ \varphi_{j i}=\theta_{j} \circ \theta_{j}^{-1} \circ \theta_{i}=\theta_{i}$, that is,

$$
\theta_{i}=\theta_{j} \circ \varphi_{j i}
$$

Using this, on $W_{i j}$, we get

$$
\begin{aligned}
\rho_{j} \circ \varphi_{j i} \circ \rho_{i}^{-1} & =\theta_{j}^{\prime-1} \circ f \circ \theta_{j} \circ \varphi_{j i} \circ\left(\theta_{i}^{\prime-1} \circ f \circ \theta_{i}\right)^{-1} \\
& =\theta_{j}^{\prime-1} \circ f \circ \theta_{j} \circ \varphi_{j i} \circ \theta_{i}^{-1} \circ f^{-1} \circ \theta_{i}^{\prime} \\
& =\theta_{j}^{\prime-1} \circ f \circ \theta_{i} \circ \theta_{i}^{-1} \circ f^{-1} \circ \theta_{i}^{\prime} \\
& =\theta_{j}^{\prime-1} \circ \theta_{i}^{\prime}=\varphi_{j i}^{\prime},
\end{aligned}
$$

as claimed. The last part of the proposition is clear.
Proposition 4.2 suggests defining a notion of equivalence on sets of gluing data which yields a converse of this proposition.

Definition 4.2 Two sets of gluing data, $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$ and $\mathcal{G}^{\prime}=$ $\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}^{\prime}\right)_{(i, j) \in K}\right)$, over the same sets of $\Omega_{i}$ 's and $\Omega_{i j}$ 's are equivalent iff there is a family of $C^{k}$ bijections, $\left(\rho_{i}: \Omega_{i} \rightarrow \Omega_{i}\right)_{i \in I}$, such that $\rho_{i}\left(\Omega_{i j}\right)=\Omega_{i j}$ and

$$
\varphi_{j i}^{\prime}(x)=\rho_{j} \circ \varphi_{j i} \circ \rho_{i}^{-1}(x), \quad \text { for all } \quad x \in \Omega_{i j},
$$

for all $i, j$.

Here is the converse of Proposition 4.2. It is actually nicer than Proposition 4.2 because we can take $W_{i j}=W_{i j}^{\prime}=\Omega_{i}$.

Proposition 4.3 If two sets of gluing data $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$ and $\mathcal{G}^{\prime}=$ $\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}^{\prime}\right)_{(i, j) \in K}\right)$ are equivalent, then there is a $C^{k}$ isomorphism, $f: M_{\mathcal{G}} \rightarrow$ $M_{\mathcal{G}^{\prime}}$, between the manifolds induced by $\mathcal{G}$ and $\mathcal{G}^{\prime}$. Furthermore, $f \circ \tau_{i}=\tau_{i}^{\prime} \circ \rho_{i}$, for all $i \in I$.

Proof. Let $f_{i}: \tau_{i}\left(\Omega_{i}\right) \rightarrow \tau_{i}^{\prime}\left(\Omega_{i}\right)$ be the $C^{k}$ bijection given by

$$
f_{i}=\tau_{i}^{\prime} \circ \rho_{i} \circ \tau_{i}^{-1}
$$

where the $\rho_{i}: \Omega_{i} \rightarrow \Omega_{i}$ 's are the maps giving the equivalence of $\mathcal{G}$ and $\mathcal{G}^{\prime}$. If we prove that $f_{i}$ and $f_{j}$ agree on the overlap, $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)$, then the $f_{i}$ patch and yield a $C^{k}$ isomorphism, $f: M_{\mathcal{G}} \rightarrow M_{\mathcal{G}^{\prime}}$. The conditions of Proposition 4.3 imply that

$$
\varphi_{j i}^{\prime} \circ \rho_{i}=\rho_{j} \circ \varphi_{j i}
$$

and we know that

$$
\tau_{i}^{\prime}=\tau_{j}^{\prime} \circ \varphi_{j i}^{\prime}
$$

Consequently, for every $[x] \in \tau_{j}\left(\Omega_{j i}\right)=\tau_{i}\left(\Omega_{i j}\right)$, with $x \in \Omega_{i j}$, we have

$$
\begin{aligned}
f_{j}([x]) & =\tau_{j}^{\prime} \circ \rho_{j} \circ \tau_{j}^{-1}([x]) \\
& =\tau_{j}^{\prime} \circ \rho_{j} \circ \tau_{j}^{-1}\left(\left[\varphi_{j i}(x)\right]\right) \\
& =\tau_{j}^{\prime} \circ \rho_{j} \circ \varphi_{j i}(x) \\
& =\tau_{j}^{\prime} \circ \varphi_{j i}^{\prime} \circ \rho_{i}(x) \\
& =\tau_{i}^{\prime} \circ \rho_{i}(x) \\
& =\tau_{i}^{\prime} \circ \rho_{i} \circ \tau_{i}^{-1}([x]) \\
& =f_{i}([x]),
\end{aligned}
$$

which shows that $f_{i}$ and $f_{j}$ agree on $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)$, as claimed.
In the next section, we describe a class of spaces that can be defined by gluing data and parametrization functions, $\theta_{i}$, that are not necessarily injective. Roughly speaking, the gluing data specify the topology and the parametrizations define the geometry of the space. Such spaces have more structure than spaces defined parametrically but they are not quite manifolds. Yet, they arise naturally in practice and they are the basis of efficient implementations of very good approximations of 3D meshes.

### 4.2 Parametric Pseudo-Manifolds

In practice, it is often desirable to specify some $n$-dimensional geometric shape as a subset of $\mathbb{R}^{d}$ (usually for $d=3$ ) in terms of parametrizations which are functions, $\theta_{i}$, from some subset of $\mathbb{R}^{n}$ into $\mathbb{R}^{d}$ (usually, $n=2$ ). For "open" shapes, this is reasonably well understood but dealing with a "closed" shape is a lot more difficult because the parametrized pieces should overlap as smoothly as possible and this is hard to achieve. Furthermore, in practice, the parametrization functions, $\theta_{i}$, may not be injective. Proposition 4.1 suggests various ways of defining such geometric shapes. For the lack of a better term, we will call these shapes, parametric pseudo-manifolds.

Definition 4.3 Let $n, k, d$ be three integers with $d>n \geq 1$ and $k \geq 1$ or $k=\infty$. A parametric $C^{k}$ pseudo-manifold of dimension $n$ in $\mathbb{R}^{d}$ is a pair, $\mathcal{M}=\left(\mathcal{G},\left(\theta_{i}\right)_{i \in I}\right)$, where $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$ is a set of gluing data for some finite set, $I$, and each $\theta_{i}$ is a $C^{k}$ function, $\theta_{i}: \Omega_{i} \rightarrow \mathbb{R}^{d}$, called a parametrization such that the following property holds:
(C) For all $(i, j) \in K$, we have

$$
\theta_{i}=\theta_{j} \circ \varphi_{j i} .
$$

For short, we use terminology parametric pseudo-manifold. The subset, $M \subseteq \mathbb{R}^{d}$, given by

$$
M=\bigcup_{i \in I} \theta_{i}\left(\Omega_{i}\right)
$$

is called the image of the parametric pseudo-manifold, $\mathcal{M}$. When $n=2$ and $d=3$, we say that $\mathcal{M}$ is a parametric pseudo-surface.

Condition (C) obviously implies that

$$
\theta_{i}\left(\Omega_{i j}\right)=\theta_{j}\left(\Omega_{j i}\right),
$$

for all $(i, j) \in K$. Consequently, $\theta_{i}$ and $\theta_{j}$ are consistent parametrizations of the overlap, $\theta_{i}\left(\Omega_{i j}\right)=\theta_{j}\left(\Omega_{j i}\right)$. Thus, the shape, $M$, is covered by pieces, $U_{i}=\theta_{i}\left(\Omega_{i}\right)$, not necessarily open, with each $U_{i}$ parametrized by $\theta_{i}$ and where the overlapping pieces, $U_{i} \cap U_{j}$, are parametrized consistently. The local structure of $M$ is given by the $\theta_{i}$ 's and the global structure is given by the gluing data. We recover a manifold if we require the $\theta_{i}$ to be bijective and to satisfy the following additional conditions:
(C') For all $(i, j) \in K$,

$$
\theta_{i}\left(\Omega_{i}\right) \cap \theta_{j}\left(\Omega_{j}\right)=\theta_{i}\left(\Omega_{i j}\right)=\theta_{j}\left(\Omega_{j i}\right)
$$

(C") For all $(i, j) \notin K$,

$$
\theta_{i}\left(\Omega_{i}\right) \cap \theta_{j}\left(\Omega_{j}\right)=\emptyset .
$$

Even if the $\theta_{i}$ 's are not injective, properties (C') and (C") would be desirable since they guarantee that $\theta_{i}\left(\Omega_{i}-\Omega_{i j}\right)$ and $\theta_{j}\left(\Omega_{j}-\Omega_{j i}\right)$ are parametrized uniquely. Unfortunately, these properties are difficult to enforce. Observe that any manifold induced by $\mathcal{G}$ is the image of a parametric pseudo-manifold.

Although this is an abuse of language, it is more convenient to call $M$ a parametric pseudo-manifold, or even a pseudo-manifold.

We can also show that the parametric pseudo-manifold, $M$, is the image in $\mathbb{R}^{d}$ of the abstract manifold, $M_{\mathcal{G}}$.

Proposition 4.4 Let $\mathcal{M}=\left(\mathcal{G},\left(\theta_{i}\right)_{i \in I}\right)$ be parametric $C^{k}$ pseudo-manifold of dimension $n$ in $\mathbb{R}^{d}$, where $\mathcal{G}=\left(\left(\Omega_{i}\right)_{\in I},\left(\Omega_{i j}\right)_{(i, j) \in I \times I},\left(\varphi_{j i}\right)_{(i, j) \in K}\right)$ is a set of gluing data for some finite set, $I$. Then, the parametrization maps, $\theta_{i}$, induce a surjective map, $\Theta: M_{\mathcal{G}} \rightarrow M$, from the abstract manifold, $M_{\mathcal{G}}$, specified by $\mathcal{G}$ to the image, $M \subseteq \mathbb{R}^{d}$, of the parametric pseudo-manifold, $\mathcal{M}$, and the following property holds: For every $\Omega_{i}$,

$$
\theta_{i}=\Theta \circ \tau_{i},
$$

where the $\tau_{i}: \Omega_{i} \rightarrow M_{\mathcal{G}}$ are the parametrization maps of the manifold $M_{\mathcal{G}}$ (see Proposition 4.1). In particular, every manifold, $M$, induced by the gluing data $\mathcal{G}$ is the image of $M_{\mathcal{G}}$ by a map $\Theta: M_{\mathcal{G}} \rightarrow M$.

Proof. Recall that

$$
M_{\mathcal{G}}=\left(\coprod_{i \in I} \Omega_{i}\right) / \sim,
$$

where $\sim$ is the equivalence relation defined so that, for all $x, y \in \coprod_{i \in I} \Omega_{i}$,

$$
x \sim y \quad \text { iff } \quad(\exists(i, j) \in K)\left(x \in \Omega_{i j}, y \in \Omega_{j i}, y=\varphi_{j i}(x)\right)
$$

The proof of Proposition 4.1 also showed that $\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right) \neq \emptyset$ iff $(i, j) \in K$ and if so,

$$
\tau_{i}\left(\Omega_{i}\right) \cap \tau_{j}\left(\Omega_{j}\right)=\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)
$$

In particular,

$$
\tau_{i}\left(\Omega_{i}-\Omega_{i j}\right) \cap \tau_{j}\left(\Omega_{j}-\Omega_{j i}\right)=\emptyset
$$

for all $(i, j) \in I \times I\left(\Omega_{i j}=\Omega_{j i}=\emptyset\right.$ when $\left.(i, j) \notin K\right)$. These properties with the fact that the $\tau_{i}$ 's are injections show that for all $(i, j) \notin K$, we can define $\Theta_{i}: \tau_{i}\left(\Omega_{i}\right) \rightarrow \mathbb{R}^{d}$ and $\Theta_{j}: \tau_{i}\left(\Omega_{j}\right) \rightarrow \mathbb{R}^{d}$ by

$$
\Theta_{i}([x])=\theta_{i}(x), x \in \Omega_{i} \quad \Theta_{j}([y])=\theta_{j}(y), y \in \Omega_{j}
$$

For $(i, j) \in K$, as the the $\tau_{i}$ 's are injections we can define $\Theta_{i}: \tau_{i}\left(\Omega_{i}-\Omega_{i j}\right) \rightarrow \mathbb{R}^{d}$ and $\Theta_{j}: \tau_{i}\left(\Omega_{j}-\Omega_{j i}\right) \rightarrow \mathbb{R}^{d}$ by

$$
\Theta_{i}([x])=\theta_{i}(x), x \in \Omega_{i}-\Omega_{i j} \quad \Theta_{j}([y])=\theta_{j}(y), y \in \Omega_{j}-\Omega_{j i} .
$$

It remains to define $\Theta_{i}$ on $\tau_{i}\left(\Omega_{i j}\right)$ and $\Theta_{j}$ on $\tau_{j}\left(\Omega_{j i}\right)$ in such a way that they agree on $\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)$. However, condition (C) in Definition 4.3 says that for all $x \in \Omega_{i j}$,

$$
\theta_{i}(x)=\theta_{j}\left(\varphi_{j i}(x)\right) .
$$

Consequently, if we define $\Theta_{i}$ on $\tau_{i}\left(\Omega_{i j}\right)$ and $\Theta_{j}$ on $\tau_{j}\left(\Omega_{j i}\right)$ by

$$
\Theta_{i}([x])=\theta_{i}(x), x \in \Omega_{i j}, \quad \Theta_{j}([y])=\theta_{j}(y), y \in \Omega_{j i},
$$

as $x \sim \varphi_{j i}(x)$, we have

$$
\Theta_{i}([x])=\theta_{i}(x)=\theta_{j}\left(\varphi_{j i}(x)\right)=\Theta_{j}\left(\left[\varphi_{j i}(x)\right]\right)=\Theta_{j}([x]),
$$

which means that $\Theta_{i}$ and $\Theta_{j}$ agree on $\tau_{i}\left(\Omega_{i j}\right)=\tau_{j}\left(\Omega_{j i}\right)$. But then, the functions, $\Theta_{i}$, agree whenever their domains overlap and so, they patch to yield a function, $\Theta$, with domain $M_{\mathcal{G}}$ and image $M$. By construction, $\theta_{i}=\Theta \circ \tau_{i}$ and as a manifold induced by $\mathcal{G}$ is a parametric pseudo-manifold, the last statement is obvious.

The function, $\Theta: M_{\mathcal{G}} \rightarrow M$, given by Proposition 4.4 shows how the parametric pseudomanifold, $M$, differs from the abstract manifold, $M_{\mathcal{G}}$. As we said before, a practical method for approximating 3D meshes based on parametric pseudo surfaces is described in Siqueira, Xu and Gallier [140].

## Chapter 5

## Lie Groups, Lie Algebras and the Exponential Map

### 5.1 Lie Groups and Lie Algebras

In Chapter 1 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 [34], Knapp [89], Warner [147], Duistermaat and Kolk [53], Bröcker and tom Dieck [25], Sagle and Walde [129], Helgason [73], Serre [137, 136], Kirillov [86], Fulton and Harris [57] and Bourbaki [22], one should be aware of more introductory sources and surveys such as Hall [70], Sattinger and Weaver [134], Carter, Segal and Macdonald [31], Curtis [38], Baker [13], Rossmann [127], Bryant [26], Mneimné and Testard [111] and Arvanitoyeogos [8].

Definition 5.1 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, $\cdot: G \times G \rightarrow G$, and the inverse map, ${ }^{-1}: G \rightarrow G$, are smooth.

We have already met a number of Lie groups: $\mathbf{G L}(n, \mathbb{R}), \mathbf{G L}(n, \mathbb{C}), \mathbf{S L}(n, \mathbb{R}), \mathbf{S L}(n, \mathbb{C})$, $\mathbf{O}(n), \mathbf{S O}(n), \mathbf{U}(n), \mathbf{S U}(n), \mathbf{E}(n, \mathbb{R})$. Also, every linear Lie group (i.e., a closed subgroup of $\mathbf{G L}(n, \mathbb{R}))$ 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 5.2 A (real) Lie algebra, $\mathcal{A}$, is a real vector space together with a bilinear map, $[\cdot, \cdot]: \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$, 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 5.3 Given two Lie groups $G_{1}$ and $G_{2}$, a homomorphism (or map) of Lie groups is a function, $f: G_{1} \rightarrow G_{2}$, 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, $f: \mathcal{A}_{1} \rightarrow \mathcal{A}_{2}$, 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, $L_{a}: G \rightarrow G$, such that $L_{a}(b)=a b$, for all $b \in G$, and right translation as the map, $R_{a}: G \rightarrow G$, such that $R_{a}(b)=b a$, 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 $\mathbf{A d}_{a}$ ) also play an important role. Note that

$$
R_{a^{-1}} L_{a}(b)=a b a^{-1} .
$$

The derivative

$$
d\left(R_{a^{-1}} L_{a}\right)_{1}: T_{1} G \rightarrow T_{1} G
$$

of $R_{a^{-1}} L_{a}: G \rightarrow G$ at 1 is an isomorphism of Lie algebras, and since $T_{1} G=\mathfrak{g}$, we get a map denoted

$$
\operatorname{Ad}_{a}: \mathfrak{g} \rightarrow \mathfrak{g}
$$

The map $a \mapsto \operatorname{Ad}_{a}$ is a map of Lie groups

$$
\mathrm{Ad}: G \rightarrow \mathbf{G L}(\mathfrak{g}),
$$

called the adjoint representation of $G$ (where $\mathbf{G L}(\mathfrak{g})$ denotes the Lie group of all bijective linear maps on $\mathfrak{g}$ ).

In the case of a linear group, one can verify that

$$
\operatorname{Ad}(a)(X)=\operatorname{Ad}_{a}(X)=a X a^{-1}
$$

for all $a \in G$ and all $X \in \mathfrak{g}$.
The derivative

$$
d \mathrm{Ad}_{1}: \mathfrak{g} \rightarrow \mathfrak{g l}(\mathfrak{g})
$$

of Ad: $G \rightarrow \mathbf{G L}(\mathfrak{g})$ at 1 is map of Lie algebras, denoted by

$$
\operatorname{ad}: \mathfrak{g} \rightarrow \mathfrak{g l}(\mathfrak{g}),
$$

called the adjoint representation of $\mathfrak{g}$. (Recall that Theorem 1.28 immediately implies that the Lie algebra, $\mathfrak{g l}(\mathfrak{g})$, of $\mathbf{G L}(\mathfrak{g})$ is the vector space, $\operatorname{End}(\mathfrak{g}, \mathfrak{g})$, of all endomorphisms of $\mathfrak{g}$, that is, the vector space of all linear maps on $\mathfrak{g}$ ).

In the case of a linear group, it can be verified that

$$
\operatorname{ad}(A)(B)=[A, B]=A B-B A
$$

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:

$$
\operatorname{ad}([u, v])=[\operatorname{ad}(u), \operatorname{ad}(v)],
$$

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 5.4 Given a Lie group, $G$, the tangent space, $\mathfrak{g}=T_{1} G$, at the identity with the Lie bracket defined by

$$
[u, v]=\operatorname{ad}(u)(v), \quad \text { for all } u, v \in \mathfrak{g}
$$

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 5.1 Given a Lie group, G, the Lie bracket, $[u, v]=\operatorname{ad}(u)(v)$, of Definition 5.4 satisfies the axioms of a Lie algebra (given in Definition 5.2). Therefore, $\mathfrak{g}$ with this bracket is a Lie algebra.

Proof. The proof requires Proposition 5.9, but we prefer to defer the proof of this Proposition until section 5.3. Since

$$
\mathrm{Ad}: G \rightarrow \mathbf{G} \mathbf{L}(\mathfrak{g})
$$

is a Lie group homomorphism, by Proposition 5.9, the map ad $=d \mathrm{Ad}_{1}$ is a homomorphism of Lie algebras, ad: $\mathfrak{g} \rightarrow \mathfrak{g l}(\mathfrak{g})$, which means that

$$
\operatorname{ad}([u, v])=\operatorname{ad}(u) \circ \operatorname{ad}(v)-\operatorname{ad}(v) \circ \operatorname{ad}(u), \quad \text { for all } u, v \in \mathfrak{g}
$$

since the bracket in $\mathfrak{g l}(\mathfrak{g})=\operatorname{End}(\mathfrak{g}, \mathfrak{g})$, 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 [53] (Chapter 1, Section 1), consider the map

$$
G \times G \longrightarrow G:(a, b) \mapsto a b a^{-1} b^{-1}
$$

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 $\left(\operatorname{Ad}_{a}-\mathrm{id}\right)(v)$. 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 $\left(\operatorname{id}-\operatorname{Ad}_{b}\right)(u)$ 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 5.1.

### 5.2 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 5.5 If $G$ is a Lie group, a vector field, $X$, on $G$ is left-invariant (resp. rightinvariant) iff

$$
d\left(L_{a}\right)_{b}(X(b))=X\left(L_{a}(b)\right)=X(a b), \quad \text { for all } a, b \in G
$$

(resp.

$$
\left.d\left(R_{a}\right)_{b}(X(b))=X\left(R_{a}(b)\right)=X(b a), \quad \text { for all } a, b \in G .\right)
$$

Equivalently, a vector field, $X$, is left-invariant iff the following diagram commutes (and similarly for a right-invariant vector field):


If $X$ is a left-invariant vector field, setting $b=1$, we see that

$$
X(a)=d\left(L_{a}\right)_{1}(X(1))
$$

which shows that $X$ is determined by its value, $X(1) \in \mathfrak{g}$, 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

$$
v^{L}(a)=d\left(L_{a}\right)_{1}(v), \quad \text { for all } a \in G
$$

We claim that $v^{L}$ is left-invariant. This follows by an easy application of the chain rule:

$$
\begin{aligned}
v^{L}(a b) & =d\left(L_{a b}\right)_{1}(v) \\
& =d\left(L_{a} \circ L_{b}\right)_{1}(v) \\
& =d\left(L_{a}\right)_{b}\left(d\left(L_{b}\right)_{1}(v)\right) \\
& =d\left(L_{a}\right)_{b}\left(v^{L}(b)\right)
\end{aligned}
$$

Furthermore, $v^{L}(1)=v$. Therefore, we showed that the map, $X \mapsto X(1)$, establishes an isomorphism between the space of left-invariant vector fields on $G$ and $\mathfrak{g}$. In fact, the map $G \times \mathfrak{g} \longrightarrow T G$ given by $(a, v) \mapsto v^{L}(a)$ is an isomorphism between $G \times \mathfrak{g}$ and the tangent bundle, $T G$.

Remark: Given any $v \in \mathfrak{g}$, we can also define the vector field, $v^{R}$, by

$$
v^{R}(a)=d\left(R_{a}\right)_{1}(v), \quad \text { for all } a \in G
$$

It is easily shown that $v^{R}$ is right-invariant and we also have an isomorphism $G \times \mathfrak{g} \longrightarrow T G$ given by $(a, v) \mapsto v^{R}(a)$.

Another reason why 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 5.2 Given a Lie group, $G$, if $X$ is a left-invariant (resp. right-invariant) vector field and $\Phi$ is its flow, then

$$
\Phi(t, g)=g \Phi(t, 1) \quad(\text { resp. } \quad \Phi(t, g)=\Phi(t, 1) g), \quad \text { for all }(t, g) \in \mathcal{D}(X)
$$

Proof. Write

$$
\gamma(t)=g \Phi(t, 1)=L_{g}(\Phi(t, 1)) .
$$

Then, $\gamma(0)=g$ and, by the chain rule

$$
\dot{\gamma}(t)=d\left(L_{g}\right)_{\Phi(t, 1)}(\dot{\Phi}(t, 1))=d\left(L_{g}\right)_{\Phi(t, 1)}(X(\Phi(t, 1)))=X\left(L_{g}(\Phi(t, 1))\right)=X(\gamma(t)) .
$$

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.

Proposition 5.3 Given a Lie group, $G$, for every $v \in \mathfrak{g}$, there is a unique smooth homomorphism, $h_{v}:(\mathbb{R},+) \rightarrow G$, 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

$$
\Phi_{s+t}^{v}(1)=\Phi_{s}^{v}\left(\Phi_{t}^{v}(1)\right)=\Phi_{t}^{v}(1) \Phi_{s}^{v}(1)
$$

by Proposition 5.2. Now, if $\Phi_{t}^{v}(1)$ is defined on $]-\epsilon, \epsilon\left[\right.$, setting $s=t$, we see that $\Phi_{t}^{v}(1)$ is actually defined on $]-2 \epsilon, 2 \epsilon\left[\right.$. 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, $h_{v}:(\mathbb{R},+) \rightarrow G$, 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

$$
\mathbb{R} \times G \times \mathfrak{g} \longrightarrow G \times \mathfrak{g}, \quad \text { where } \quad(t, g, v) \mapsto\left(g \Phi_{t}^{v}(1), v\right)
$$

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 $h:(\mathbb{R},+) \rightarrow G$ 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

$$
\frac{d h}{d t}(t)=d\left(L_{h(t)}\right)_{1}(v)=v^{L}(h(t))
$$

and

$$
\frac{d h}{d t}(t)=d\left(R_{h(t)}\right)_{1}(v)=v^{R}(h(t))
$$

Therefore, $h(t)$ is an integral curve for $v^{L}$ and $v^{R}$ with initial condition $h(0)=1$ and $h=\Phi_{t}^{v}(1)$.

Since $h_{v}:(\mathbb{R},+) \rightarrow G$ is a homomorphism, the integral curve, $h_{v}$, if often referred to as a one-parameter group. Proposition 5.3 yields the definition of the exponential map.

Definition 5.6 Given a Lie group, $G$, the exponential map, exp: $\mathfrak{g} \rightarrow G$, 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 5.3, we have the smooth map

$$
\mathbb{R} \times G \times \mathfrak{g} \longrightarrow G \times \mathfrak{g}, \quad \text { where } \quad(t, g, v) \mapsto\left(g \Phi_{t}^{v}(1), v\right)
$$

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

$$
s \mapsto h_{v}(s t)
$$

is a smooth homomorphism, $h$, such that $\dot{h}(0)=t v$. By uniqueness, we have

$$
h_{v}(s t)=h_{t v}(s)
$$

Setting $s=1$, we find that

$$
h_{v}(t)=\exp (t v), \quad \text { for all } v \in \mathfrak{g} \text { and all } t \in \mathbb{R}
$$

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 5.4 Given any two Lie groups, $G$ and $H$, for every Lie group homomorphism, $f: G \rightarrow H$, the following diagram commutes:


Proof. Observe that the map $h: t \mapsto f(\exp (t v))$ is a homomorphism from $(\mathbb{R},+)$ to $G$ such that $\dot{h}(0)=d f_{1}(v)$. Proposition 5.3 shows that $f(\exp (v))=\exp \left(d f_{1}(v)\right)$.

A useful corollary of Proposition 5.4 is:

Proposition 5.5 Let $G$ be a connected Lie group and $H$ be any Lie group. For any two homomorphisms, $\varphi_{1}: G \rightarrow H$ and $\varphi_{2}: G \rightarrow H$, if $d\left(\varphi_{1}\right)_{1}=d\left(\varphi_{2}\right)_{1}$, 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 5.4, for all $a \in \exp _{G}(U)$, we have

$$
\varphi_{i}(a)=\exp _{H}\left(d\left(\varphi_{i}\right)_{1}\left(\exp _{G}^{-1}(a)\right)\right), \quad i=1,2
$$

Since $d\left(\varphi_{1}\right)_{1}=d\left(\varphi_{2}\right)_{1}$, 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 $\left.\exp _{G}(U)\right)$. Therefore, $\varphi_{1}=\varphi_{2}$ on $G$.

The above proposition shows that if $G$ is connected, then a homomorphism of Lie groups, $\varphi: G \rightarrow H$, is uniquely determined by the Lie algebra homomorphism, $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$.

We obtain another useful corollary of Proposition 5.4 when we apply it to the adjoint representation of $G$,

$$
\mathrm{Ad}: G \rightarrow \mathbf{G L}(\mathfrak{g})
$$

and to the conjugation map,

$$
\mathbf{A d}_{a}: G \rightarrow G,
$$

where $\operatorname{Ad}_{a}(b)=a b a^{-1}$. In the first case, $d \operatorname{Ad}_{1}=$ ad, with ad: $\mathfrak{g} \rightarrow \mathfrak{g l}(\mathfrak{g})$ and in the second case, $d\left(\mathbf{A d}_{a}\right)_{1}=\operatorname{Ad}_{a}$.

Proposition 5.6 Given any Lie group, $G$, the following properties hold:

$$
\begin{equation*}
\operatorname{Ad}(\exp (u))=e^{\operatorname{ad}(u)}, \quad \text { for all } u \in \mathfrak{g} \tag{1}
\end{equation*}
$$

where exp: $\mathfrak{g} \rightarrow G$ is the exponential of the Lie group, $G$, and $f \mapsto e^{f}$ is the exponential map given by

$$
e^{f}=\sum_{k=0}^{\infty} \frac{f^{k}}{k!},
$$

for any linear map (matrix), $f \in \mathfrak{g l}(\mathfrak{g})$. Equivalently, the following diagram commutes:

(2)

$$
\exp \left(t \operatorname{Ad}_{g}(u)\right)=g \exp (t u) g^{-1}
$$

for all $u \in \mathfrak{g}$, all $g \in G$ and all $t \in \mathbb{R}$. Equivalently, the following diagram commutes:


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 3.16), it is natural to ask if there is any relationship between, $[u, v]$, where $[u, v]=\operatorname{ad}(u)(v)$, and the Lie bracket, $\left[u^{L}, v^{L}\right]$, 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 [147] (Chapter 3), Bröcker and tom Dieck [25] (Chapter 1, Section 2), or Marsden and Ratiu [102] (Chapter 9).

Proposition 5.7 Given a Lie group, G, we have

$$
\left[u^{L}, v^{L}\right](1)=\operatorname{ad}(u)(v), \quad \text { for all } u, v \in \mathfrak{g}
$$

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 5.8 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.

### 5.3 Homomorphisms of Lie Groups and Lie Algebras, Lie Subgroups

If $G$ and $H$ are two Lie groups and $\varphi: G \rightarrow H$ is a homomorphism of Lie groups, then $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$ 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 5.9 If $G$ and $H$ are two Lie groups and $\varphi: G \rightarrow H$ is a homomorphism of Lie groups, then

$$
d \varphi_{1} \circ \operatorname{Ad}_{g}=\operatorname{Ad}_{\varphi(g)} \circ d \varphi_{1}, \quad \text { for all } g \in G,
$$

that is, the following diagram commutes

and $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$ is a Lie algebra homomorphism.
Proof. Recall that

$$
R_{a^{-1}} L_{a}(b)=a b a^{-1}, \quad \text { for all } a, b \in G
$$

and that the derivative

$$
d\left(R_{a^{-1}} L_{a}\right)_{1}: \mathfrak{g} \rightarrow \mathfrak{g}
$$

of $R_{a^{-1}} L_{a}$ at 1 is an isomorphism of Lie algebras, denoted by $\operatorname{Ad}_{a}: \mathfrak{g} \rightarrow \mathfrak{g}$. The map $a \mapsto \operatorname{Ad}_{a}$ is a map of Lie groups

$$
\mathrm{Ad}: G \rightarrow \mathbf{G L}(\mathfrak{g}),
$$

(where $\mathbf{G L}(\mathfrak{g})$ denotes the Lie group of all bijective linear maps on $\mathfrak{g}$ ) and the derivative

$$
d \mathrm{Ad}_{1}: \mathfrak{g} \rightarrow \mathfrak{g l}(\mathfrak{g})
$$

of $\operatorname{Ad}$ at 1 is map of Lie algebras, denoted by

$$
\operatorname{ad}: \mathfrak{g} \rightarrow \mathfrak{g l}(\mathfrak{g})
$$

called the adjoint representation of $\mathfrak{g}$ (where $\mathfrak{g l}(\mathfrak{g})$ denotes the Lie algebra of all linear maps on $\mathfrak{g}$ ). Then, the Lie bracket is defined by

$$
[u, v]=\operatorname{ad}(u)(v), \quad \text { for all } u, v \in \mathfrak{g} .
$$

Now, as $\varphi$ is a homomorphism, we have

$$
\varphi\left(R_{a^{-1}} L_{a}(b)\right)=\varphi\left(a b a^{-1}\right)=\varphi(a) \varphi(b) \varphi(a)^{-1}=R_{\varphi(a)^{-1}} L_{\varphi(a)}(\varphi(b)),
$$

and by differentiating w.r.t. $b$ at $b=1$ in the direction, $v \in \mathfrak{g}$, we get

$$
d \varphi_{1}\left(\operatorname{Ad}_{a}(v)\right)=\operatorname{Ad}_{\varphi(a)}\left(d \varphi_{1}(v)\right)
$$

proving the first part of the proposition. Differentiating again with respect to $a$ at $a=1$ in the direction, $u \in \mathfrak{g}$, (and using the chain rule), we get

$$
d \varphi_{1}(\operatorname{ad}(u)(v))=\operatorname{ad}\left(d \varphi_{1}(u)\right)\left(d \varphi_{1}(v)\right)
$$

i.e.,

$$
d \varphi_{1}[u, v]=\left[d \varphi_{1}(u), d \varphi_{1}(v)\right],
$$

which proves that $d \varphi_{1}$ is indeed a Lie algebra homomorphism.
Remark: If we identify the Lie algebra, $\mathfrak{g}$, of $G$ with the space of left-invariant vector fields on $G$, the map $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$ is viewed as the map such that, for every left-invariant vector field, $X$, on $G$, the vector field $d \varphi_{1}(X)$ is the unique left-invariant vector field on $H$ such that

$$
d \varphi_{1}(X)(1)=d \varphi_{1}(X(1)),
$$

i.e., $d \varphi_{1}(X)=d \varphi_{1}(X(1))^{L}$. Then, we can give another proof of the fact that $d \varphi_{1}$ is a Lie algebra homomorphism using the notion of $\varphi$-related vector fields.

Proposition 5.10 If $G$ and $H$ are two Lie groups and $\varphi: G \rightarrow H$ 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) $X$ and $d \varphi_{1}(X)$ are $\varphi$-related, for every left-invariant vector field, $X$, on $G$;
(b) $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$ is a Lie algebra homomorphism.

Proof. The proof uses Proposition 3.14. For details, see Warner [147].
We now consider Lie subgroups. As a preliminary result, note that if $\varphi: G \rightarrow H$ is an injective Lie group homomorphism, then $d \varphi_{g}: T_{g} G \rightarrow T_{\varphi(g)} H$ is injective for all $g \in G$. As $\mathfrak{g}=T_{1} G$ and $T_{g} G$ are isomorphic for all $g \in G$ (and similarly for $\mathfrak{h}=T_{1} H$ and $T_{h} H$ for all $h \in H)$, it is sufficient to check that $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$ is injective. However, by Proposition 5.4, the diagram

commutes, and since the exponential map is a local diffeomorphism at 0 , as $\varphi$ is injective, then $d \varphi_{1}$ is injective, too. Therefore, if $\varphi: G \rightarrow H$ is injective, it is automatically an immersion.

Definition 5.7 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, $\varphi: H \rightarrow G$ (and thus, $\varphi$ 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 $\varphi$ 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, $T^{2} \cong \mathbf{S O}(2) \times \mathbf{S O}(2)$, which can be identified with the group of $2 \times 2$ complex diagonal matrices of the form

$$
\left(\begin{array}{cc}
e^{i \theta_{1}} & 0 \\
0 & e^{i \theta_{2}}
\end{array}\right)
$$

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

$$
\left(\begin{array}{cc}
e^{i t} & 0 \\
0 & e^{i c t}
\end{array}\right), \quad t \in \mathbb{R}
$$

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 [57] (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 5.11 Let $G$ be a Lie group and let $(H, \varphi)$ be an immersed Lie subgroup of $G$. Then, $\varphi$ 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 [147] (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, $(U, \varphi)$, of $G$, with $1 \in U$, so that

$$
\varphi(U \cap H)=\varphi(U) \cap\left(R^{m} \times\left\{0_{n-m}\right\}\right)
$$

By Proposition 2.15, we can find some open subset, $V \subseteq U$, with $1 \in V$, so that $V=V^{-1}$ and $\bar{V} \subseteq U$. Observe that

$$
\varphi(\bar{V} \cap H)=\varphi(\bar{V}) \cap\left(R^{m} \times\left\{0_{n-m}\right\}\right)
$$

and since $\bar{V}$ is closed and $\varphi$ is a homeomorphism, it follows that $\bar{V} \cap H$ is closed. Thus, $\bar{V} \cap H=\bar{V} \cap \bar{H}$ (as $\overline{\bar{V} \cap H}=\bar{V} \cap \bar{H})$. Now, pick any $y \in \bar{H}$. As $1 \in V^{-1}$, the open set $y V^{-1}$ contains $y$ and since $y \in \bar{H}$, we must have $y V^{-1} \cap H \neq \emptyset$. Let $x \in y V^{-1} \cap H$, then $x \in H$ and $y \in x V$. Then, $y \in x V \cap \bar{H}$, which implies $x^{-1} y \in V \cap \bar{H} \subseteq \bar{V} \cap \bar{H}=\bar{V} \cap H$. Therefore, $x^{-1} y \in H$ and since $x \in H$, we get $y \in H$ and $H$ is closed.

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 5.12 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 [25] (Chapter 1, Section 3) or Warner [147], Chapter 3.

### 5.4 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 5.8 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 [147] (Chapter 3) or Duistermaat and Kolk [53] (Chapter 1, Section 10).

Theorem 5.13 Let $G$ be a Lie group with Lie algebra, $\mathfrak{g}$, and let $(H, \varphi)$ 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, $(H, \varphi)$, 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 5.13 shows that there is a one-to-one correspondence between connected immersed subgroups of a Lie group and subalgebras of its Lie algebra.

Theorem 5.14 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, $\psi: \mathfrak{g} \rightarrow \mathfrak{h}$, there is a unique Lie group homomorphism, $\varphi: G \rightarrow H$, so that $d \varphi_{1}=\psi$.

Again a proof of the theorem above is given in Warner [147] (Chapter 3) or Duistermaat and Kolk [53] (Chapter 1, Section 10).

Corollary 5.15 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 [89], Fulton and Harris [57], or Bourbaki [22].

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, $\varphi: G \rightarrow H$, is uniquely determined by the Lie algebra homomorphism, $d \varphi_{1}: \mathfrak{g} \rightarrow \mathfrak{h}$.

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, $\psi: \mathfrak{g} \rightarrow \mathfrak{h}$, is a Lie algebra map iff there is a unique Lie group homomorphism, $\varphi: G \rightarrow H$, so that $d \varphi_{1}=\psi$.

### 5.5 More on the Lorentz Group $\mathrm{SO}_{0}(n, 1)$

In this section, we take a closer look at the Lorentz group $\mathbf{S O}_{0}(n, 1)$ and, in particular, at the relationship between $\mathbf{S O}_{0}(n, 1)$ and its Lie algebra, $\mathfrak{s o}(n, 1)$. The Lie algebra of $\mathbf{S O}_{0}(n, 1)$ is easily determined by computing the tangent vectors to curves, $t \mapsto A(t)$, on $\mathbf{S O}_{0}(n, 1)$ through the identity, $I$. Since $A(t)$ satisfies

$$
A^{\top} J A=J
$$

differentiating and using the fact that $A(0)=I$, we get

$$
A^{\prime \top} J+J A^{\prime}=0
$$

Therefore,

$$
\mathfrak{s o}(n, 1)=\left\{A \in \operatorname{Mat}_{n+1, n+1}(\mathbb{R}) \mid A^{\top} J+J A=0\right\} .
$$

This means that $J A$ is skew-symmetric and so,

$$
\mathfrak{s o}(n, 1)=\left\{\left.\left(\begin{array}{cc}
B & u \\
u^{\top} & 0
\end{array}\right) \in \operatorname{Mat}_{n+1, n+1}(\mathbb{R}) \right\rvert\, u \in \mathbb{R}^{n}, \quad B^{\top}=-B\right\}
$$

Observe that every matrix $A \in \mathfrak{s o}(n, 1)$ can be written uniquely as

$$
\left(\begin{array}{cc}
B & u \\
u^{\top} & 0
\end{array}\right)=\left(\begin{array}{cc}
B & 0 \\
0^{\top} & 0
\end{array}\right)+\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right),
$$

where the first matrix is skew-symmetric, the second one is symmetric and both belong to $\mathfrak{s o}(n, 1)$. Thus, it is natural to define

$$
\mathfrak{k}=\left\{\left.\left(\begin{array}{cc}
B & 0 \\
0^{\top} & 0
\end{array}\right) \right\rvert\, B^{\top}=-B\right\}
$$

and

$$
\mathfrak{p}=\left\{\left.\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right) \right\rvert\, u \in \mathbb{R}^{n}\right\}
$$

It is immediately verified that both $\mathfrak{k}$ and $\mathfrak{p}$ are subspaces of $\mathfrak{s o}(n, 1)$ (as vector spaces) and that $\mathfrak{k}$ is a Lie subalgebra isomorphic to $\mathfrak{s o}(n)$, but $\mathfrak{p}$ is not a Lie subalgebra of $\mathfrak{s o}(n, 1)$ 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

$$
\mathfrak{s o}(n, 1)=\mathfrak{k} \oplus \mathfrak{p}
$$

known as Cartan decomposition. There is also an automorphism of $\mathfrak{s o}(n, 1)$ known as the Cartan involution, namely,

$$
\theta(A)=-A^{\top}
$$

and we see that

$$
\mathfrak{k}=\{A \in \mathfrak{s o}(n, 1) \mid \theta(A)=A\} \quad \text { and } \quad \mathfrak{p}=\{A \in \mathfrak{s o}(n, 1) \mid \theta(A)=-A\} .
$$

Unfortunately, there does not appear to be any simple way of obtaining a formula for $\exp (A)$, where $A \in \mathfrak{s o}(n, 1)$ (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 5.16 For every matrix, $A \in \mathfrak{p}$, of the form

$$
A=\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right)
$$

we have

$$
e^{A}=\left(\begin{array}{cc}
I+\frac{(\cosh \omega-1)}{\omega^{2}} u u^{\top} & \frac{\sinh \omega}{\omega} u \\
\frac{\sinh \omega}{\omega} u^{\top} & \cosh \omega
\end{array}\right)=\left(\begin{array}{cc}
\sqrt{I+\frac{\sinh ^{2} \omega}{\omega^{2}} u u^{\top}} & \frac{\sinh \omega}{\omega} u \\
\frac{\sinh \omega}{\omega} u^{\top} & \cosh \omega
\end{array}\right) .
$$

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

$$
A^{2}=\left(\begin{array}{cc}
u u^{\top} & 0 \\
0 & u^{\top} u
\end{array}\right)=\left(\begin{array}{cc}
u u^{\top} & 0 \\
0 & \omega^{2}
\end{array}\right)
$$

We leave as an exercise the fact that

$$
\left(I+\frac{(\cosh \omega-1)}{\omega^{2}} u u^{\top}\right)^{2}=I+\frac{\sinh ^{2} \omega}{\omega^{2}} u u^{\top}
$$

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 5.17 Every Lorentz boost,

$$
A=\left(\begin{array}{cc}
\sqrt{I+v v^{\top}} & v \\
v^{\top} & c
\end{array}\right)
$$

with $c=\sqrt{\|v\|^{2}+1}$, is of the form $A=e^{B}$, for $B \in \mathfrak{p}$, i.e., for some $B \in \mathfrak{s o}(n, 1)$ of the form

$$
B=\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right)
$$

Proof. We need to find some

$$
B=\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right)
$$

solving the equation

$$
\left(\begin{array}{cc}
\sqrt{I+\frac{\sinh ^{2} \omega}{\omega^{2}} u u^{\top}} & \frac{\sinh \omega}{\omega} u \\
\frac{\sinh \omega}{\omega} u^{\top} & \cosh \omega
\end{array}\right)=\left(\begin{array}{cc}
\sqrt{I+v v^{\top}} & v \\
v^{\top} & c
\end{array}\right)
$$

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,

$$
e^{2 \omega}-2 c e^{\omega}+1=0
$$

The roots of the corresponding algebraic equation $X^{2}-2 c X+1=0$ are

$$
X=c \pm \sqrt{c^{2}-1}
$$

As $c>1$, both roots are strictly positive, so we can solve for $\omega$, say $\omega=\log \left(c+\sqrt{c^{2}-1}\right) \neq 0$. Then, $\sinh \omega \neq 0$, so we can solve the equation

$$
\frac{\sinh \omega}{\omega} u=v
$$

which yields a $B \in \mathfrak{s o}(n, 1)$ of the right form with $A=e^{B}$.

## Remarks:

(1) It is easy to show that the eigenvalues of matrices

$$
B=\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right)
$$

are 0 , with multiplicity $n-1,\|u\|$ and $-\|u\|$. Eigenvectors are also easily determined.
(2) The matrices, $B \in \mathfrak{s o}(n, 1)$, of the form

$$
B=\left(\begin{array}{cccc}
0 & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 0 & \alpha \\
0 & \cdots & \alpha & 0
\end{array}\right)
$$

are easily seen to form an abelian Lie subalgebra, $\mathfrak{a}$, of $\mathfrak{s o}(n, 1)$ (which means that for all $B, C \in \mathfrak{a},[B, C]=0$, i.e., $B C=C B)$. One will easily check that for any $B \in \mathfrak{a}$, as above, we get

$$
e^{B}=\left(\begin{array}{ccccc}
1 & \cdots & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \cdots & 1 & 0 & 0 \\
0 & \cdots & 0 & \cosh \alpha & \sinh \alpha \\
0 & \cdots & 0 & \sinh \alpha & \cosh \alpha
\end{array}\right)
$$

The matrices of the form, $e^{B}$, with $B \in \mathfrak{a}$, form an abelian subgroup, $A$, of $\mathbf{S O}_{0}(n, 1)$ isomorphic to $\mathbf{S O}_{0}(1,1)$. As we already know, the matrices, $B \in \mathfrak{s o}(n, 1)$, of the form

$$
\left(\begin{array}{ll}
B & 0 \\
0 & 0
\end{array}\right),
$$

where $B$ is skew-symmetric, form a Lie subalgebra, $\mathfrak{k}$, of $\mathfrak{s o}(n, 1)$. Clearly, $\mathfrak{k}$ is isomorphic to $\mathfrak{s o}(n)$ and using the exponential, we get a subgroup, $K$, of $\mathbf{S O}_{0}(n, 1)$ isomorphic to $\mathbf{S O}(n)$. It is also clear that $\mathfrak{k} \cap \mathfrak{a}=(0)$, but $\mathfrak{k} \oplus \mathfrak{a}$ is not equal to $\mathfrak{s o}(n, 1)$. What is the missing piece? Consider the matrices, $N \in \mathfrak{s o}(n, 1)$, of the form

$$
N=\left(\begin{array}{ccc}
0 & -u & u \\
u^{\top} & 0 & 0 \\
u^{\top} & 0 & 0
\end{array}\right)
$$

where $u \in \mathbb{R}^{n-1}$. The reader should check that these matrices form an abelian Lie subalgebra, $\mathfrak{n}$, of $\mathfrak{s o}(n, 1)$ and that

$$
\mathfrak{s o}(n, 1)=\mathfrak{k} \oplus \mathfrak{a} \oplus \mathfrak{n}
$$

This is the Iwasawa decomposition of the Lie algebra $\mathfrak{s o}(n, 1)$. 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 [89]). The connected Lie subgroup of $\mathbf{S O}_{0}(n, 1)$ associated with $\mathfrak{n}$ is denoted $N$ and it can be shown that we have the Iwasawa decomposition of the Lie group $\mathbf{S O}_{0}(n, 1)$ :

$$
\mathbf{S O}_{0}(n, 1)=K A N
$$

It is easy to check that $[\mathfrak{a}, \mathfrak{n}] \subseteq \mathfrak{n}$, so $\mathfrak{a} \oplus \mathfrak{n}$ is a Lie subalgebra of $\mathfrak{s o}(n, 1)$ 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 $A N$ is a subgroup (in fact, solvable) of $\mathbf{S O}_{0}(n, 1)$. For more on the Iwasawa decomposition, see Knapp [89]. Observe that the image, $\overline{\mathfrak{n}}$, of $\mathfrak{n}$ under the Cartan involution, $\theta$, is the Lie subalgebra

$$
\overline{\mathfrak{n}}=\left\{\left.\left(\begin{array}{ccc}
0 & u & u \\
-u^{\top} & 0 & 0 \\
u^{\top} & 0 & 0
\end{array}\right) \right\rvert\, u \in \mathbb{R}^{n-1}\right\} .
$$

It is easy to see that the centralizer of $\mathfrak{a}$ is the Lie subalgebra

$$
\mathfrak{m}=\left\{\left.\left(\begin{array}{cc}
B & 0 \\
0 & 0
\end{array}\right) \in \operatorname{Mat}_{n+1, n+1}(\mathbb{R}) \right\rvert\, B \in \mathfrak{s o}(n-1)\right\}
$$

and the reader should check that

$$
\mathfrak{s o}(n, 1)=\mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n} \oplus \overline{\mathfrak{n}} .
$$

We also have

$$
[\mathfrak{m}, \mathfrak{n}] \subseteq \mathfrak{n}
$$

so $\mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n}$ is a subalgebra of $\mathfrak{s o}(n, 1)$. The group, $M$, associated with $\mathfrak{m}$ is isomorphic to $\mathbf{S O}(n-1)$ and it can be shown that $B=M A N$ is a subgroup of $\mathbf{S O}_{0}(n, 1)$. In fact,

$$
\mathbf{S O}_{0}(n, 1) /(M A N)=K A N / M A N=K / M=\mathbf{S O}(n) / \mathbf{S O}(n-1)=S^{n-1}
$$

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 $\mathfrak{s o}(n, 1)=\mathfrak{m} \oplus \mathfrak{a} \oplus \mathfrak{n} \oplus \overline{\mathfrak{n}}$ is also written

$$
\mathfrak{s o}(n, 1)=\mathfrak{g}_{0} \oplus \mathfrak{g}_{-1} \oplus \mathfrak{g}_{1}
$$

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 $\mathbf{S O}_{0}(n, 1)$ that can't be diagonalized.

Combining Proposition 2.3 and Proposition 5.17, we have the corollary:

Corollary 5.18 Every matrix, $A \in \mathbf{O}(n, 1)$, can be written as

$$
A=\left(\begin{array}{cc}
Q & 0 \\
0 & \epsilon
\end{array}\right) e^{\left(\begin{array}{cc}
0 & u \\
u^{\top} & 0
\end{array}\right)}
$$

where $Q \in \mathbf{O}(n), \epsilon= \pm 1$ and $u \in \mathbb{R}^{n}$.

Observe that Corollary 5.18 proves that every matrix, $A \in \mathbf{S O}_{0}(n, 1)$, can be written as

$$
A=P e^{S}, \quad \text { with } P \in K \cong \mathbf{S O}(n) \text { and } S \in \mathfrak{p}
$$

i.e.,

$$
\mathbf{S O}_{0}(n, 1)=K \exp (\mathfrak{p})
$$

a version of the polar decomposition for $\mathbf{S O}_{0}(n, 1)$.
Now, it is known that the exponential map, exp: $\mathfrak{s o}(n) \rightarrow \mathbf{S O}(n)$, is surjective. So, when $A \in \mathbf{S O}_{0}(n, 1)$, since then $Q \in \mathbf{S O}(n)$ and $\epsilon=+1$, the matrix

$$
\left(\begin{array}{cc}
Q & 0 \\
0 & 1
\end{array}\right)
$$

is the exponential of some skew symmetric matrix

$$
C=\left(\begin{array}{cc}
B & 0 \\
0 & 0
\end{array}\right) \in \mathfrak{s o}(n, 1),
$$

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: $\mathfrak{s o}(n, 1) \rightarrow \mathbf{S O}_{0}(n, 1)$. 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 [118] (1983). Nishikawa's paper is rather short, but this is misleading. Indeed, Nishikawa relies on a classic paper by Djokovic [48], which itself relies heavily on another fundamental paper by Burgoyne and Cushman [27], 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 $\mathbf{S O}_{0}(n, 1)$. 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 [126].

Interestingly, these two proofs use very different methods. The Nishikawa-DjokovicBurgoyne 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 [107].

What we will do to close this section is to give a relatively simple proof that the exponential map, exp: $\mathfrak{s o}(1,3) \rightarrow \mathbf{S O}_{0}(1,3)$, is surjective. In the case of $\mathbf{S O}_{0}(1,3)$, we can use the fact that $\mathbf{S L}(2, \mathbb{C})$ is a two-sheeted covering space of $\mathbf{S O}_{0}(1,3)$, which means that there is a homomorphism, $\varphi: \mathbf{S L}(2, \mathbb{C}) \rightarrow \mathbf{S O}_{0}(1,3)$, which is surjective and that $\operatorname{Ker} \varphi=\{-I, I)$. Then, the small miracle is that, although the exponential, $\exp : \mathfrak{s l}(2, \mathbb{C}) \rightarrow \mathbf{S L}(2, \mathbb{C})$, is not surjective, for every $A \in \mathbf{S L}(2, \mathbb{C})$, either $A$ or $-A$ is in the image of the exponential!

Proposition 5.19 Given any matrix

$$
B=\left(\begin{array}{cc}
a & b \\
c & -a
\end{array}\right) \in \mathfrak{s l}(2, \mathbb{C})
$$

let $\omega$ be any of the two complex roots of $a^{2}+b c$. If $\omega \neq 0$, then

$$
e^{B}=\cosh \omega I+\frac{\sinh \omega}{\omega} B
$$

and $e^{B}=I+B$, if $a^{2}+b c=0$. Furthermore, every matrix $A \in \mathbf{S L}(2, \mathbb{C})$ 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 $A \in \mathbf{S L}(2, \mathbb{C})$, either $A$ or $-A$ is of the form $e^{B}$, for some $B \in \mathfrak{s l}(2, \mathbb{C})$.

Proof. Observe that

$$
A^{2}=\left(\begin{array}{cc}
a & b \\
c & -a
\end{array}\right)\left(\begin{array}{cc}
a & b \\
c & -a
\end{array}\right)=\left(a^{2}+b c\right) I
$$

Then, it is straighforward to prove that

$$
e^{B}=\cosh \omega I+\frac{\sinh \omega}{\omega} B
$$

where $\omega$ is a square root of $a^{2}+b c$ is $\omega \neq 0$, otherwise, $e^{B}=I+B$.
Let

$$
A=\left(\begin{array}{ll}
\alpha & \beta \\
\gamma & \delta
\end{array}\right)
$$

be any matrix in $\mathbf{S L}(2, \mathbb{C})$. We would like to find a matrix, $B \in \mathfrak{s l}(2, \mathbb{C})$, 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

$$
e^{2 \omega}-2 e^{\omega} c+1=0
$$

The corresponding algebraic equation

$$
Z^{2}-2 c Z+1=0
$$

has discriminant $4\left(c^{2}-1\right)$ 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 $\omega$, so that $e^{\omega}=c+\sqrt{c^{2}-1}$ is a solution of $e^{2 \omega}-2 e^{\omega} c+1=0$. This proves the surjectivity of cosh.

We have $\sinh \omega=0$ iff $e^{2 \omega}=1$; this holds iff $2 \omega=n 2 \pi i$, 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 $A \in \mathbf{S L}(2, \mathbb{C})$ ), 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, $N \in \mathfrak{s l}(2, \mathbb{C})$ 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 $B \in \mathfrak{s l}(2, \mathbb{C})$ so that $e^{B}=A$.

Indeed, any matrix, $B \in \mathfrak{s l}(2, \mathbb{C})$, has zero trace, which means that if $\lambda_{1}$ and $\lambda_{2}$ 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 $A \in \mathbf{S L}(2, \mathbb{C})$ 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.

Remark: If we restrict our attention to $\mathbf{S L}(2, \mathbb{R})$, then we have the following proposition that can be used to prove that the exponential map, $\exp : \mathfrak{s o}(1,2) \rightarrow \mathbf{S O}_{0}(1,2)$, is surjective:

Proposition 5.20 Given any matrix

$$
B=\left(\begin{array}{cc}
a & b \\
c & -a
\end{array}\right) \in \mathfrak{s l}(2, \mathbb{R}),
$$

if $a^{2}+b>0$, then let $\omega=\sqrt{a^{2}+b c}>0$ and if $a^{2}+b<0$, then let $\omega=\sqrt{-\left(a^{2}+b c\right)}>0$ (i.e., $\omega^{2}=-\left(a^{2}+b c\right)$ ). In the first case $\left(a^{2}+b c>0\right)$, we have

$$
e^{B}=\cosh \omega I+\frac{\sinh \omega}{\omega} B
$$

and in the second case $\left(a^{2}+b c<0\right)$, we have

$$
e^{B}=\cos \omega I+\frac{\sin \omega}{\omega} B
$$

If $a^{2}+b c=0$, then $e^{B}=I+B$. Furthermore, every matrix $A \in \mathbf{S L}(2, \mathbb{R})$ whose trace satisfies $\operatorname{tr}(A) \geq-2$ in the image of the exponential map. Consequently, for any $A \in \mathbf{S L}(2, \mathbb{R})$, either $A$ or $-A$ is of the form $e^{B}$, for some $B \in \mathfrak{s l}(2, \mathbb{R})$.

We now return to the relationship between $\mathbf{S L}(2, \mathbb{C})$ and $\mathbf{S O}_{0}(1,3)$. In order to define a homomorphism, $\varphi: \mathbf{S L}(2, \mathbb{C}) \rightarrow \mathbf{S O}_{0}(1,3)$, we begin by defining a linear bijection, $h$, between $\mathbb{R}^{4}$ and $\mathbf{H}(2)$, the set of complex $2 \times 2$ Hermitian matrices, by

$$
(t, x, y, z) \mapsto\left(\begin{array}{cc}
t+x & y-i z \\
y+i z & t-x
\end{array}\right)
$$

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

$$
\left(\begin{array}{ll}
a & b \\
c & d
\end{array}\right)
$$

we have

$$
t=\frac{a+d}{2}, x=\frac{a-d}{2}, \quad \text { etc. }
$$

Next, for any $A \in \mathbf{S L}(2, \mathbb{C})$, we define a map, $l_{A}: \mathbf{H}(2) \rightarrow \mathbf{H}(2)$, via

$$
S \mapsto A S A^{*}
$$

(Here, $A^{*}=\bar{A}^{\top}$.) Using the linear bijection, $h: \mathbb{R}^{4} \rightarrow \mathbf{H}(2)$, and its inverse, we obtain a $\operatorname{map}, \operatorname{lor}_{A}: \mathbb{R}^{4} \rightarrow \mathbb{R}^{4}$, where

$$
\operatorname{lor}_{A}=h^{-1} \circ l_{A} \circ h .
$$

As $A S A^{*}$ is hermitian, we see that $l_{A}$ is well defined. It is obviously linear and since $\operatorname{det}(A)=1$ (recall, $A \in \mathbf{S L}(2, \mathbb{C}))$ and

$$
\operatorname{det}\left(\begin{array}{cc}
t+x & y-i z \\
y+i z & t-x
\end{array}\right)=t^{2}-x^{2}-y^{2}-z^{2}
$$

we see that $\operatorname{lor}_{A}$ preserves the Lorentz metric! Furthermore, it is not hard to prove that $\mathbf{S L}(2, \mathbb{C})$ is connected (use the polar form or analyze the eigenvalues of a matrix in $\mathbf{S L}(2, \mathbb{C})$, for example, as in Duistermatt and Kolk [53] (Chapter 1, Section 1.2)) and that the map

$$
\varphi: A \mapsto \operatorname{lor}_{A}
$$

is a continuous group homomorphism. Thus, the range of $\varphi$ is a connected subgroup of $\mathbf{S O}_{0}(1,3)$. This shows that $\varphi: \mathbf{S L}(2, \mathbb{C}) \rightarrow \mathbf{S O}_{0}(1,3)$ is indeed a homomorphism. It remains to prove that it is surjective and that its kernel is $\{I,-I\}$.

Proposition 5.21 The homomorphism, $\varphi: \mathbf{S L}(2, \mathbb{C}) \rightarrow \mathbf{S O}_{0}(1,3)$, is surjective and its kernel is $\{I,-I\}$.

Proof. Recall that from Theorem 2.6, the Lorentz group $\mathbf{S O}_{0}(1,3)$ is generated by the matrices of the form

$$
\left(\begin{array}{ll}
1 & 0 \\
0 & P
\end{array}\right) \quad \text { with } P \in \mathbf{S O}(3)
$$

and the matrices of the form

$$
\left(\begin{array}{cccc}
\cosh \alpha & \sinh \alpha & 0 & 0 \\
\sinh \alpha & \cosh \alpha & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{array}\right)
$$

Thus, to prove the surjectivity of $\varphi$, it is enough to check that the above matrices are in the range of $\varphi$. For matrices of the second kind, the reader should check that

$$
A=\left(\begin{array}{cc}
e^{\frac{1}{2} \alpha} & 0 \\
0 & e^{-\frac{1}{2} \alpha}
\end{array}\right)
$$

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 $\mathbf{S U}(2)$, via the correspondence

$$
a \mathbf{1}+b \mathbf{i}+c \mathbf{j}+d \mathbf{k} \mapsto\left(\begin{array}{cc}
a+i b & c+i d \\
-c+i d & a-i b
\end{array}\right)
$$

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 $a=0$,

$$
\left(\begin{array}{cc}
i b & c+i d \\
-c+i d & -i b
\end{array}\right)=i\left(\begin{array}{cc}
b & d-i c \\
d+i c & -b
\end{array}\right)=i h(0, b, d, c) .
$$

Therefore, we have a bijection between the pure quaternions and the subspace of the hermitian matrices

$$
\left(\begin{array}{cc}
b & d-i c \\
d+i c & -b
\end{array}\right)
$$

for which $a=0$, the inverse being division by $i$, i.e., multiplication by $-i$. Also, when $q$ is a unit quaternion, let $\bar{q}=a \mathbf{1}-b \mathbf{i}-c \mathbf{j}-d \mathbf{k}$, and observe that $\bar{q}=q^{-1}$. Using the embedding $\mathbb{R}^{3} \hookrightarrow \mathbb{H}$, for every unit quaternion, $q \in \mathbf{S U}(2)$, define the map, $\rho_{q}: \mathbb{R}^{3} \rightarrow \mathbb{R}^{3}$, by

$$
\rho_{q}(X)=q X \bar{q}=q X q^{-1}
$$

for all $X \in \mathbb{R}^{3} \hookrightarrow \mathbb{H}$. Then, it is well known that $\rho_{q}$ is a rotation (i.e., $\left.\rho_{q} \in \mathbf{S O}(3)\right)$ and, moreover, the map $q \mapsto \rho_{q}$, is a surjective homomorphism, $\rho: \mathbf{S U ( 2 )} \rightarrow \mathbf{S O}(3)$, and Ker $\varphi=\{I,-I\}$ (For example, see Gallier [58], Chapter 8).

Now, consider a matrix, $A$, of the form

$$
\left(\begin{array}{ll}
1 & 0 \\
0 & P
\end{array}\right) \quad \text { with } P \in \mathbf{S O}(3)
$$

We claim that we can find a matrix, $B \in \mathbf{S L}(2, \mathbb{C})$, such that $\varphi(B)=\operatorname{lor}_{B}=A$. We claim that we can pick $B \in \mathbf{S U}(2) \subseteq \mathbf{S L}(2, \mathbb{C})$. Indeed, if $B \in \mathbf{S U}(2)$, then $B^{*}=B^{-1}$, so

$$
B\left(\begin{array}{cc}
t+x & y-i z \\
y+i z & t-x
\end{array}\right) B^{*}=t\left(\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right)-i B\left(\begin{array}{cc}
i x & z+i y \\
-z+i y & -i x
\end{array}\right) B^{-1}
$$

The above shows that $\operatorname{lor}_{B}$ leaves the coordinate $t$ invariant. The term

$$
B\left(\begin{array}{cc}
i x & z+i y \\
-z+i y & -i x
\end{array}\right) B^{-1}
$$

is a pure quaternion corresponding to the application of the rotation $\rho_{B}$ 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 $P \in \mathbf{S O}(3)$, there is a quaternion, $B$, so that $\rho_{B}=P$, so we can find our $B \in \mathbf{S U}(2)$ so that

$$
\operatorname{lor}_{B}=\left(\begin{array}{ll}
1 & 0 \\
0 & P
\end{array}\right)=A
$$

Finally, assume that $\varphi(A)=I$. This means that

$$
A S A^{*}=S
$$

for all hermitian matrices, $S$, defined above. In particular, for $S=I$, we get $A A^{*}=I$, i.e., $A \in \mathbf{S U}(2)$. We have

$$
A S=S A
$$

for all hermitian matrices, $S$, defined above, so in particular, this holds for diagonal matrices of the form

$$
\left(\begin{array}{cc}
t+x & 0 \\
0 & t-x
\end{array}\right)
$$

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, $\operatorname{Ker} \varphi=\{I,-I\}$.

Remark: The group $\mathbf{S L}(2, \mathbb{C})$ is isomorphic to the group $\operatorname{Spin}(1,3)$, which is a (simplyconnected) double-cover of $\mathbf{S O}_{0}(1,3)$. This is a standard result of Clifford algebra theory, see Bröcker and tom Dieck [25] or Fulton and Harris [57]. What we just did is to provide a direct proof of this fact.

We just proved that there is an isomorphism

$$
\mathbf{S L}(2, \mathbb{C}) /\{I,-I\} \cong \mathbf{S O}_{0}(1,3)
$$

However, the reader may recall that $\mathbf{S L}(2, \mathbb{C}) /\{I,-I\}=\mathbf{P S L}(2, \mathbb{C}) \cong$ Möb $^{+}$. Therefore, the Lorentz group is isomorphic to the Möbius group.

We now have all the tools to prove that the exponential map, $\exp : \mathfrak{s o}(1,3) \rightarrow \mathbf{S O}_{0}(1,3)$, is surjective.

Theorem 5.22 The exponential map, exp: $\mathfrak{s o}(1,3) \rightarrow \mathbf{S O}_{0}(1,3)$, is surjective.
Proof. First, recall from Proposition 5.4 that the following diagram commutes:


Pick any $A \in \mathbf{S O}_{0}(1,3)$. By Proposition 5.21, the homomorphism $\varphi$ is surjective and as $\operatorname{Ker} \varphi=\{I,-I\}$, there exists some $B \in \mathbf{S L}(2, \mathbb{C})$ so that

$$
\varphi(B)=\varphi(-B)=A
$$

Now, by Proposition 5.19 , for any $B \in \mathbf{S L}(2, \mathbb{C})$, either $B$ or $-B$ is of the form $e^{C}$, for some $C \in \mathfrak{s l}(2, \mathbb{C})$. By the commutativity of the diagram, if we let $D=d \varphi_{1}(C) \in \mathfrak{s o}(1,3)$, we get

$$
A=\varphi\left( \pm e^{C}\right)=e^{d \varphi_{1}(C)}=e^{D}
$$

with $D \in \mathfrak{s o}(1,3)$, as required.

Remark: We can restrict the bijection, $h: \mathbb{R}^{4} \rightarrow \mathbf{H}(2)$, defined earlier to a bijection between $\mathbb{R}^{3}$ and the space of real symmetric matrices of the form

$$
\left(\begin{array}{cc}
t+x & y \\
y & t-x
\end{array}\right)
$$

Then, if we also restrict ourselves to $\mathbf{S L}(2, \mathbb{R})$, for any $A \in \mathbf{S L}(2, \mathbb{R})$ and any symmetric matrix, $S$, as above, we get a map

$$
S \mapsto A S A^{\top}
$$

The reader should check that these transformations correspond to isometries in $\mathbf{S O}_{0}(1,2)$ and we get a homomorphism, $\varphi: \mathbf{S L}(2, \mathbb{R}) \rightarrow \mathbf{S O}_{0}(1,2)$. Then, we have a version of Proposition 5.21 for $\mathbf{S L}(2, \mathbb{R})$ and $\mathbf{S O}_{0}(1,2)$ :

Proposition 5.23 The homomorphism, $\varphi: \mathbf{S L}(2, \mathbb{R}) \rightarrow \mathbf{S O}_{0}(1,2)$, is surjective and its kernel is $\{I,-I\}$.

Using Proposition 5.23 and Proposition 5.20, we get a version of Theorem 5.22 for $\mathbf{S O}_{0}(1,2)$ :

Theorem 5.24 The exponential map, $\exp : \mathfrak{s o}(1,2) \rightarrow \mathbf{S O}_{0}(1,2)$, is surjective.

Also observe that $\mathbf{S O}_{0}(1,1)$ consists of the matrices of the form

$$
A=\left(\begin{array}{ll}
\cosh \alpha & \sinh \alpha \\
\sinh \alpha & \cosh \alpha
\end{array}\right)
$$

and a direct computation shows that

$$
e^{\left(\begin{array}{ll}
0 & \alpha \\
\alpha & 0
\end{array}\right)}=\left(\begin{array}{cc}
\cosh \alpha & \sinh \alpha \\
\sinh \alpha & \cosh \alpha
\end{array}\right)
$$

Thus, we see that the map exp: $\mathfrak{s o}(1,1) \rightarrow \mathbf{S O}_{0}(1,1)$ is also surjective. Therefore, we have proved that $\exp : \mathfrak{s o}(1, n) \rightarrow \mathbf{S O}_{0}(1, n)$ 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.

### 5.6 More on the Topology of $\mathbf{O}(p, q)$ and $\mathbf{S O}(p, q)$

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, $\mathbf{G L}(n, \mathbb{C})$, of invertible $n \times n$ matrices with complex coefficients. If $A=\left(a_{k l}\right)$ is such a matrix, denote by $x_{k l}$ the real part (resp. $y_{k l}$, the imaginary part) of $a_{k l}$ (so, $\left.a_{k l}=x_{k l}+i y_{k l}\right)$.

Definition 5.9 A subgroup, $G$, of $\mathbf{G L}(n, \mathbb{C})$ is pseudo-algebraic iff there is a finite set of polynomials in $2 n^{2}$ variables with real coefficients, $\left\{P_{i}\left(X_{1}, \ldots, X_{n^{2}}, Y_{1}, \ldots, Y_{n^{2}}\right)\right\}_{i=1}^{t}$, so that

$$
A=\left(x_{k l}+i y_{k l}\right) \in G \quad \text { iff } \quad P_{i}\left(x_{11}, \ldots, x_{n n}, y_{11}, \ldots, y_{n n}\right)=0, \quad \text { for } i=1, \ldots, t .
$$

Recall that if $A$ is a complex $n \times n$-matrix, its adjoint, $A^{*}$, is defined by $A^{*}=(\bar{A})^{\top}$. Also, $\mathbf{U}(n)$ denotes the group of unitary matrices, i.e., those matrices, $A \in \mathbf{G L}(n, \mathbb{C})$, so that $A A^{*}=A^{*} A=I$, 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 5.25 Let $G$ be a pseudo-algebraic subgroup of $\mathbf{G L}(n, \mathbb{C})$ 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

$$
(\mathbf{U}(n) \cap G) \times(\mathbf{H}(n) \cap \mathfrak{g}) \longrightarrow G, \quad \text { given by } \quad(U, H) \mapsto U e^{H}
$$

is a homeomorphism onto $G$.

Proof. A proof can be found in Knapp [89], Chapter 1, or Mneimné and Testard [111], Chapter 3.

We now apply Theorem 5.25 to determine the structure of the space $\mathbf{O}(p, q)$. We know that $\mathbf{O}(p, q)$ consists of the matrices, $A$, in $\mathbf{G L}(p+q, \mathbb{R})$ 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

$$
\mathfrak{o}(p, q)=\left\{\left.\left(\begin{array}{ll}
X_{1} & X_{2} \\
X_{2}^{\top} & X_{3}
\end{array}\right) \right\rvert\, X_{1}^{\top}=-X_{1}, \quad X_{3}^{\top}=-X_{3}, \quad X_{2} \text { arbitrary }\right\}
$$

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

$$
\mathfrak{o}(p, q) \cap \mathbf{H}(p+q)=\left\{\left.\left(\begin{array}{cc}
0 & X_{2} \\
X_{2}^{\top} & 0
\end{array}\right) \right\rvert\, X_{2} \text { arbitrary }\right\}
$$

a vector space of dimension $p q$.
Some simple calculations also show that

$$
\mathbf{O}(p, q) \cap \mathbf{U}(p+q)=\left\{\left.\left(\begin{array}{cc}
X_{1} & 0 \\
0 & X_{2}
\end{array}\right) \right\rvert\, X_{1} \in \mathbf{O}(p), \quad X_{2} \in \mathbf{O}(q)\right\} \cong \mathbf{O}(p) \times \mathbf{O}(q)
$$

Therefore, we obtain the structure of $\mathbf{O}(p, q)$ :
Proposition 5.26 The topological space $\mathbf{O}(p, q)$ is homeomorphic to $\mathbf{O}(p) \times \mathbf{O}(q) \times \mathbb{R}^{p q}$.

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

$$
\mathbf{S O}(p, q) \cap \mathbf{U}(p+q)=\left\{\left.\left(\begin{array}{cc}
X_{1} & 0 \\
0 & X_{2}
\end{array}\right) \right\rvert\, X_{1} \in \mathbf{O}(p), \quad X_{2} \in \mathbf{O}(q), \quad \operatorname{det}\left(X_{1}\right) \operatorname{det}\left(X_{2}\right)=1\right\}
$$

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 $\mathfrak{s o}(p, q)=\mathfrak{o}(p, q)$. Thus, we also have

Proposition 5.27 The topological space $\mathbf{S O}(p, q)$ is homeomorphic to $S(\mathbf{O}(p) \times \mathbf{O}(q)) \times \mathbb{R}^{p q}$.

Observe that the dimension of all these spaces depends only on $p+q$ : It is $(p+q)(p+$ $q-1) / 2$. Also, $\mathbf{S O}(p, q)$ has two connected components when $p, q \geq 1$. The connected component of $I_{p+q}$ is the group $\mathbf{S O}_{0}(p, q)$. This latter space is homeomorphic to $\mathbf{S O}(p) \times$ $\mathbf{S O}(q) \times \mathbb{R}^{p q}$.

Theorem 5.25 gives the polar form of a matrix $A \in \mathbf{O}(p, q)$ : We have

$$
A=U e^{S}, \quad \text { with } \quad U \in \mathbf{O}(p) \times \mathbf{O}(q) \quad \text { and } \quad S \in \mathfrak{s o}(p, q) \cap \mathbf{S}(p+q)
$$

where $U$ is of the form

$$
U=\left(\begin{array}{ll}
P & 0 \\
0 & Q
\end{array}\right), \quad \text { with } \quad P \in \mathbf{O}(p) \quad \text { and } \quad Q \in \mathbf{O}(q)
$$

and $\mathfrak{s o}(p, q) \cap \mathbf{S}(p+q)$ consists of all $(p+q) \times(p+q)$ symmetric matrices of the form

$$
S=\left(\begin{array}{cc}
0 & X \\
X^{\top} & 0
\end{array}\right)
$$

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 [111]). Recall that the functions cosh and sinh also make sense for matrices (since the exponential makes sense) and are given by

$$
\cosh (A)=\frac{e^{A}+e^{-A}}{2}=I+\frac{A^{2}}{2!}+\cdots+\frac{A^{2 k}}{(2 k)!}+\cdots
$$

and

$$
\sinh (A)=\frac{e^{A}-e^{-A}}{2}=A+\frac{A^{3}}{3!}+\cdots+\frac{A^{2 k+1}}{(2 k+1)!}+\cdots .
$$

We also set

$$
\frac{\sinh (A)}{A}=I+\frac{A^{2}}{3!}+\cdots+\frac{A^{2 k}}{(2 k+1)!}+\cdots
$$

which is defined for all matrices, $A$ (even when $A$ is singular). Then, we have

Proposition 5.28 For any matrix $S$ of the form

$$
S=\left(\begin{array}{cc}
0 & X \\
X^{\top} & 0
\end{array}\right)
$$

we have

$$
e^{S}=\left(\begin{array}{cc}
\cosh \left(\left(X X^{\top}\right)^{\frac{1}{2}}\right) & \frac{\sinh \left(\left(X X^{\top}\right)^{\frac{1}{2}}\right) X}{\left(X X^{\top}\right)^{\frac{1}{2}}} \\
\frac{\sinh \left(\left(X^{\top} X\right)^{\frac{1}{2}}\right) X^{\top}}{\left(X^{\top} X\right)^{\frac{1}{2}}} & \cosh \left(\left(X^{\top} X\right)^{\frac{1}{2}}\right)
\end{array}\right)
$$

Proof. By induction, it is easy to see that

$$
S^{2 k}=\left(\begin{array}{cc}
\left(X X^{\top}\right)^{k} & 0 \\
0 & \left(X^{\top} X\right)^{k}
\end{array}\right)
$$

and

$$
S^{2 k+1}=\left(\begin{array}{cc}
0 & \left(X X^{\top}\right)^{k} X \\
\left(X^{\top} X\right)^{k} X^{\top} & 0
\end{array}\right)
$$

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 5.28, every matrix, $A \in \mathbf{O}(p, q)$, has the polar form

$$
A=\left(\begin{array}{cc}
P & 0 \\
0 & Q
\end{array}\right)\left(\begin{array}{cc}
\cosh \left(\left(X X^{\top}\right)^{\frac{1}{2}}\right) & \frac{\sinh \left(\left(X X^{\top}\right)^{\frac{1}{2}}\right) X}{\left(X X^{\top}\right)^{\frac{1}{2}}} \\
\frac{\sinh \left(\left(X^{\top} X\right)^{\frac{1}{2}}\right) X^{\top}}{\left(X^{\top} X\right)^{\frac{1}{2}}} & \cosh \left(\left(X^{\top} X\right)^{\frac{1}{2}}\right)
\end{array}\right),
$$

with $P \in \mathbf{O}(p), Q \in \mathbf{O}(q)$ and $X$ an arbitrary $p \times q$ matrix.

### 5.7 Universal Covering Groups

Every connected Lie group, $G$, is a manifold and, as such, from results in Section 3.9, it has a universal cover, $\pi: \widetilde{G} \rightarrow G$, where $\widetilde{G}$ is simply connected. It is possible to make $\widetilde{G}$ into a group so that $\widetilde{G}$ is a Lie group and $\pi$ is a Lie group homomorphism. We content ourselves with a sketch of the construction whose details can be found in Warner [147], Chapter 3.

Consider the map, $\alpha: \widetilde{G} \times \widetilde{G} \rightarrow G$, given by

$$
\alpha(\widetilde{a}, \widetilde{b})=\pi(\widetilde{a}) \pi(\widetilde{b})^{-1}
$$

for all $\widetilde{a}, \widetilde{b} \in \widetilde{G}$, and pick some $\widetilde{e} \in \pi^{-1}(e)$. Since $\widetilde{G} \times \widetilde{G}$ is simply connected, it follows by Proposition 3.34 that there is a unique map, $\widetilde{\alpha}: \widetilde{G} \times \widetilde{G} \rightarrow \widetilde{G}$, such that

$$
\alpha=\pi \circ \widetilde{\alpha} \quad \text { and } \quad \widetilde{e}=\widetilde{\alpha}(\widetilde{e}, \widetilde{e}) .
$$

For all $\widetilde{a}, \widetilde{b} \in \widetilde{G}$, define

$$
\begin{equation*}
\widetilde{b}^{-1}=\widetilde{\alpha}(\widetilde{e}, \widetilde{b}), \quad \widetilde{a} \widetilde{b}=\widetilde{\alpha}\left(\widetilde{a}, \widetilde{b}^{-1}\right) \tag{*}
\end{equation*}
$$

Using Proposition 3.34, it can be shown that the above operations make $\widetilde{G}$ into a group and as $\widetilde{\alpha}$ is smooth, into a Lie group. Moreover, $\pi$ becomes a Lie group homomorphism. We summarize these facts as

Theorem 5.29 Every connected Lie group has a simply connected covering map, $\pi: \widetilde{G} \rightarrow G$, where $\widetilde{G}$ is a Lie group and $\pi$ is a Lie group homomorphism.

The group, $\widetilde{G}$, is called the universal covering group of $G$. Consider $D=\operatorname{ker} \pi$. Since the fibres of $\pi$ are countable The group $D$ is a countable closed normal subgroup of $\widetilde{G}$, that is, a discrete normal subgroup of $\widetilde{G}$. It follows that $G \cong \widetilde{G} / D$, where $\widetilde{G}$ is a simply connected Lie group and $D$ is a discrete normal subgroup of $\widetilde{G}$.

We conclude this section by stating the following useful proposition whose proof can be found in Warner [147] (Chapter 3, Proposition 3.26):

Proposition 5.30 Let $\varphi: G \rightarrow H$ be a homomorphism of connected Lie groups. Then $\varphi$ is a covering map iff $d \varphi_{e}: \mathfrak{g} \rightarrow \mathfrak{h}$ is an isomorphism of Lie algebras.

For example, we know that $\mathfrak{s u}(2)=\mathfrak{s o}(3)$, so the homomorphism from $\mathbf{S U}(2)$ to $\mathbf{S O}(3)$ provided by the representation of 3 D rotations by the quaternions is a covering map.


[^0]:    ${ }^{1}$ We are using the wedge product notation of exterior calculus even though we have not defined alternating tensors and the wedge product yet. 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. A thorough treatment of tensor algebra, including exterior algebra, and of differential forms, will be given in Chapters 22 and 8.

