Chapter 4

Manifolds, Lie Groups, and Lie Algebras; "Baby Case"

In this section we define precisely *embedded submanifolds*, *matrix Lie groups*, and their *Lie algebras*.

One of the reasons that Lie groups are nice is that they have a *differential structure*, which means that the notion of *tangent space* makes sense at any point of the group.

Furthermore, the tangent space at the identity happens to have some algebraic structure, that of a *Lie algebra*.

Roughly, the tangent space at the identity provides a "linearization" of the Lie group, and it turns out that many properties of a Lie group are reflected in its Lie algebra.

Fortunately, most of the Lie groups that we need to consider are subspaces of $\mathbb{R}^N$ for some sufficiently large N.

In fact, they are all isomorphic to subgroups of $\mathbf{GL}(N, \mathbb{R})$ for some suitable N, even $\mathbf{SE}(n)$, which is isomorphic to a subgroup of $\mathbf{SL}(n+1)$.

Such groups are called *linear Lie groups* (or *matrix groups*).

Since the groups under consideration are subspaces of $\mathbb{R}^N$, we do not need immediately the definition of an abstract manifold.

We just have to define *embedded submanifolds* (also called submanifolds) of $\mathbb{R}^N$ (in the case of $\mathbf{GL}(n, \mathbb{R})$, $N = n^2$).

In general, the difficult part in proving that a subgroup of $\mathbf{GL}(n, \mathbb{R})$ is a Lie group is to prove that it is a manifold.

Fortunately, there is simple a characterization of the linear groups.

This characterization rests on two theorems. First, a Lie subgroup H of a Lie group G (where H is an embedded submanifold of G) is closed in G.

Second, a theorem of Von Neumann and Cartan asserts that a closed subgroup of $\mathbf{GL}(n, \mathbb{R})$ is an embedded submanifold, and thus, a Lie group.

Thus, *a linear Lie group is a closed subgroup of* $\mathbf{GL}(n, \mathbb{R})$.

A small annoying technical arises in our approach, the problem with *discrete subgroups*.

If A is a subset of $\mathbb{R}^N$, recall that A inherits a topology from $\mathbb{R}^N$ called the *subspace topology*, and defined such that a subset V of A is open if

$$V = A \cap U$$

for some open subset U of $\mathbb{R}^N$.

A point $a \in A$ is said to be *isolated* if there is some open subset U of $\mathbb{R}^N$ such that

$$\{a\} = A \cap U,$$

in other words, if $\{a\}$ is an open set in A.

The group $\mathbf{GL}(n,\mathbb{R})$ of real invertible $n \times n$ matrices can be viewed as a subset of $\mathbb{R}^{n^2}$, and as such, it is a topological space under the subspace topology (in fact, a dense open subset of $\mathbb{R}^{n^2}$).

One can easily check that multiplication and the inverse operation are continuous, and in fact smooth (i.e., C^∞-continuously differentiable).

This makes $\mathbf{GL}(n,\mathbb{R})$ a *topological group*.

Any subgroup G of $\mathbf{GL}(n,\mathbb{R})$ is also a topological space under the subspace topology.

A subgroup G is called a *discrete subgroup* if it has some isolated point.

This turns out to be equivalent to the fact that *every point of G is isolated*, and thus, G has the discrete topology (every subset of G is open).

Now, because $\mathbf{GL}(n, \mathbb{R})$ is a topological group, it can be shown that *every discrete subgroup of $\mathbf{GL}(n, \mathbb{R})$ is closed*, and in fact countable.

Thus, *discrete subgroups of $\mathbf{GL}(n, \mathbb{R})$ are Lie groups!*

But these are not very interesting Lie groups so we will consider only closed subgroups of $\mathbf{GL}(n, \mathbb{R})$ that are not discrete.

We wish to define embedded submanifolds in $\mathbb{R}^N$.

For the sake of brevity, we use the terminology *manifold* (but other authors would say *embedded submanifold*, or something like that).

The intuition behind the notion of a smooth manifold in $\mathbb{R}^N$ is that a subspace M is a manifold of dimension m if every point $p \in M$ is contained in some open subset U of M (in the subspace topology) that can be parametrized by some function $\varphi\colon \Omega \to U$ from some open subset Ω of the origin in $\mathbb{R}^m$, and that φ has some nice properties that allow:

(1) The definition of *smooth functions on M* and

(2) The definition of the *tangent space at p*.

For this, φ has to be at least a homeomorphism, but more is needed: φ must be smooth, and the derivative $\varphi'(0_m)$ at the origin must be *injective* (letting $0_m = \underbrace{(0, \ldots, 0)}_{m}$).

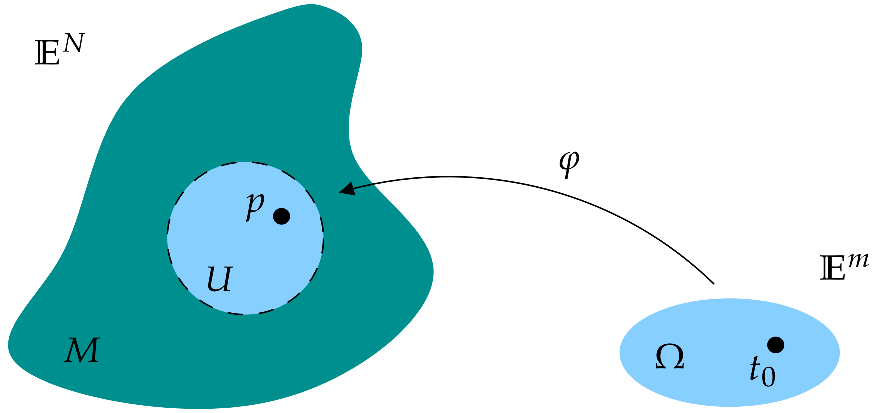

Figure 4.1: A manifold in $\mathbb{R}^N$

Definition 4.1. Given any integers N, m, with $N \geq m \geq 1$, an *m-dimensional smooth manifold in $\mathbb{R}^N$, for short a manifold*, is a nonempty subset M of $\mathbb{R}^N$ such that for every point $p \in M$ there are two open subsets $\Omega \subseteq \mathbb{R}^m$ and $U \subseteq M$, with $p \in U$, and a smooth function $\varphi\colon \Omega \to \mathbb{R}^N$ such that φ is a homeomorphism between Ω and $U = \varphi(\Omega)$, and $\varphi'(t_0)$ is injective, where $t_0 = \varphi^{-1}(p)$.

The function $\varphi\colon \Omega \to U$ is called a *(local) parametrization of M at p*. If $0_m \in \Omega$ and $\varphi(0_m) = p$, we say that $\varphi\colon \Omega \to U$ is *centered at p*.

Recall that $M \subseteq \mathbb{R}^N$ is a topological space under the subspace topology, and U is some open subset of M in the subspace topology, which means that $U = M \cap W$ for some open subset W of $\mathbb{R}^N$.

Since $\varphi\colon \Omega \to U$ is a homeomorphism, it has an inverse $\varphi^{-1}\colon U \to \Omega$ that is also a homeomorphism, called a *(local) chart*.

Since $\Omega \subseteq \mathbb{R}^m$, for every $p \in M$ and every parametrization $\varphi\colon \Omega \to U$ of M at p, we have $\varphi^{-1}(p) = (z_1, \ldots, z_m)$ for some $z_i \in \mathbb{R}$, and we call $z_1, \ldots, z_m$ the *local coordinates of p (w.r.t. φ^{-1})*.

We often refer to a manifold M without explicitly specifying its dimension (the integer m).

Intuitively, a chart provides a “flattened” local map of a region on a manifold.

Remark: We could allow $m = 0$ in definition 4.1. If so, a manifold of dimension 0 is just a set of isolated points, and thus it has the discrete topology.

In fact, it can be shown that a discrete subset of $\mathbb{R}^N$ is countable. Such manifolds are not very exciting, but they do correspond to discrete subgroups.

Example 4.1. The unit sphere S^2 in $\mathbb{R}^3$ defined such that

$$S^2 = \left\{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\right\}$$

is a smooth 2-manifold, because it can be parametrized using the following two maps φ_1 and φ_2:

$$\varphi_1 \colon (u, v) \mapsto \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{u^2 + v^2 - 1}{u^2 + v^2 + 1}\right)$$

and

$$\varphi_2 \colon (u, v) \mapsto \left(\frac{2u}{u^2 + v^2 + 1}, \frac{2v}{u^2 + v^2 + 1}, \frac{1 - u^2 - v^2}{u^2 + v^2 + 1}\right).$$

The map φ_1 corresponds to the inverse of the stereographic projection from the north pole $N = (0, 0, 1)$ onto the plane $z = 0$, and the map φ_2 corresponds to the inverse of the stereographic projection from the south pole $S = (0, 0, -1)$ onto the plane $z = 0$, as illustrated in Figure 4.2.

The reader should check that the map φ_1 parametrizes $S^2 - \{N\}$ and that the map φ_2 parametrizes $S^2 - \{S\}$ (and that they are smooth, homeomorphisms, etc.).

Using φ_1, the open lower hemisphere is parametrized by the open disk of center O and radius 1 contained in the plane $z = 0$.

The chart φ_1^{-1} assigns local coordinates to the points in the open lower hemisphere.

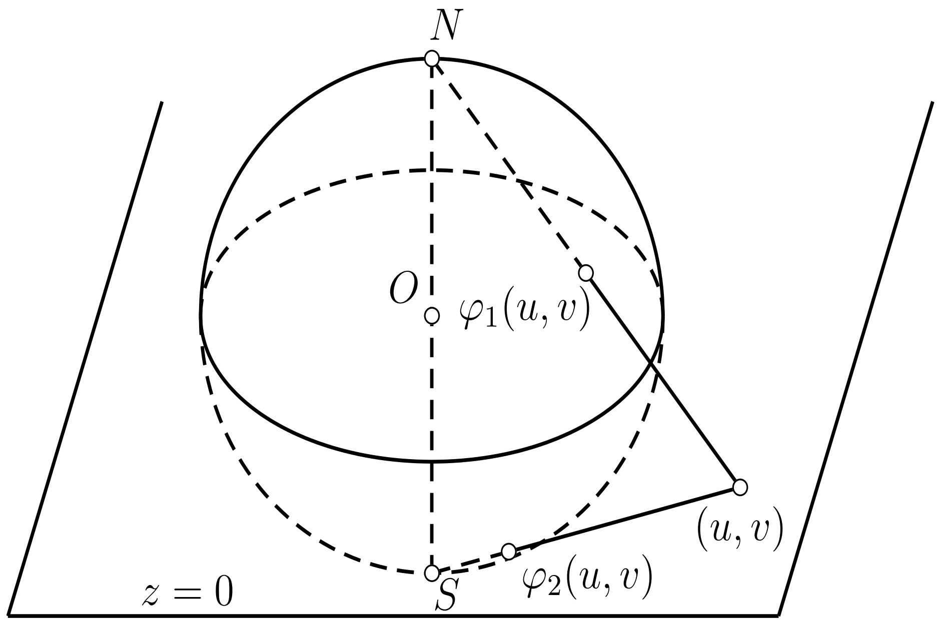

Figure 4.2: Inverse stereographic projections

We urge our readers to define a manifold structure on a torus. This can be done using four charts.

Every open subset of $\mathbb{R}^N$ is a manifold in a trivial way. Indeed, we can use the inclusion map as a parametrization.

In particular, $\mathbf{GL}(n, \mathbb{R})$ is an open subset of $\mathbb{R}^{n^2}$, since its complement is closed (the set of invertible matrices is the inverse image of the determinant function, which is continuous).

Thus, $\mathbf{GL}(n, \mathbb{R})$ is a manifold. We can view $\mathbf{GL}(n, \mathbb{C})$ as a subset of $\mathbb{R}^{(2n)^2}$ using the embedding defined as follows:

For every complex $n \times n$ matrix A, construct the real $2n \times 2n$ matrix such that every entry $a + ib$ in A is replaced by the 2×2 block

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

where $a, b \in \mathbb{R}$.

It is immediately verified that this map is in fact a group isomorphism.

Thus, we can view $\mathbf{GL}(n, \mathbb{C})$ as a subgroup of $\mathbf{GL}(2n, \mathbb{R})$, and as a manifold in $\mathbb{R}^{(2n)^2}$.

A 1-manifold is called a *(smooth) curve*, and a 2-manifold is called a *(smooth) surface* (although some authors require that they also be connected).

The following two lemmas provide the link with the definition of an abstract manifold.

Lemma 4.1. *Given an m-dimensional manifold M in $\mathbb{R}^N$, for every $p \in M$ there are two open sets $O, W \subseteq \mathbb{R}^N$ with $0_N \in O$ and $p \in M \cap W$, and a smooth diffeomorphism $\varphi\colon O \to W$, such that $\varphi(0_N) = p$ and*

$$\varphi(O \cap (\mathbb{R}^m \times \{0_{N-m}\})) = M \cap W.$$

There is an open subset Ω of $\mathbb{R}^m$ such that

$$O \cap (\mathbb{R}^m \times \{0_{N-m}\}) = \Omega \times \{0_{N-m}\},$$

and the map $\psi\colon \Omega \to \mathbb{R}^N$ given by

$$\psi(x) = \varphi(x, 0_{N-m})$$

is an immersion and a homeomorphism onto $U = W \cap M$; so ψ is a parametrization of M at p.

We can think of φ as a promoted version of ψ which is actually a diffeomorphism between open subsets of $\mathbb{R}^N$; see Figure 4.3.

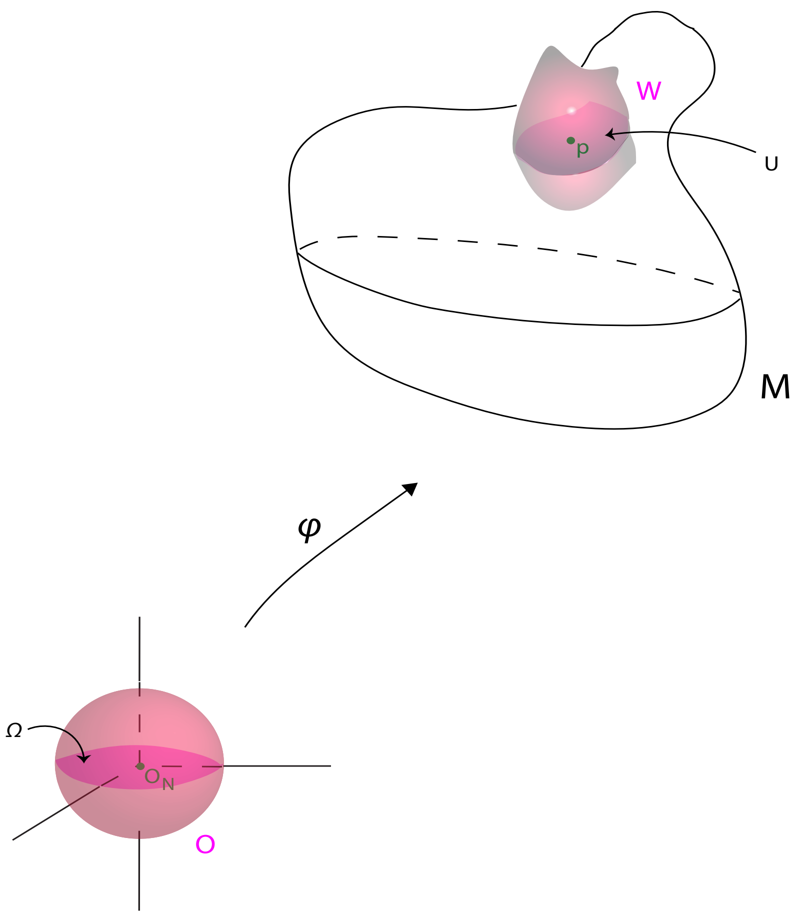

Figure 4.3: An illustration of Lemma 4.1, where M is a surface embedded in $\mathbb{R}^3$, namely $m = 2$ and $N = 3$.

The next lemma is easily shown from Lemma 4.1. It is a key technical result used to show that interesting properties of maps between manifolds do not depend on parametrizations.

Lemma 4.2. *Given an m-dimensional manifold M in $\mathbb{R}^N$, for every $p \in M$ and any two parametrizations $\varphi_1 \colon \Omega_1 \to U_1$ and $\varphi_2 \colon \Omega_2 \to U_2$ of M at p, if $U_1 \cap U_2 \neq \emptyset$, the map $\varphi_2^{-1} \circ \varphi_1 \colon \varphi_1^{-1}(U_1 \cap U_2) \to \varphi_2^{-1}(U_1 \cap U_2)$ is a smooth diffeomorphism.*

The maps $\varphi_2^{-1} \circ \varphi_1 \colon \varphi_1^{-1}(U_1 \cap U_2) \to \varphi_2^{-1}(U_1 \cap U_2)$ are called *transition maps*.

Lemma 4.2 is illustrated in Figure 4.4.

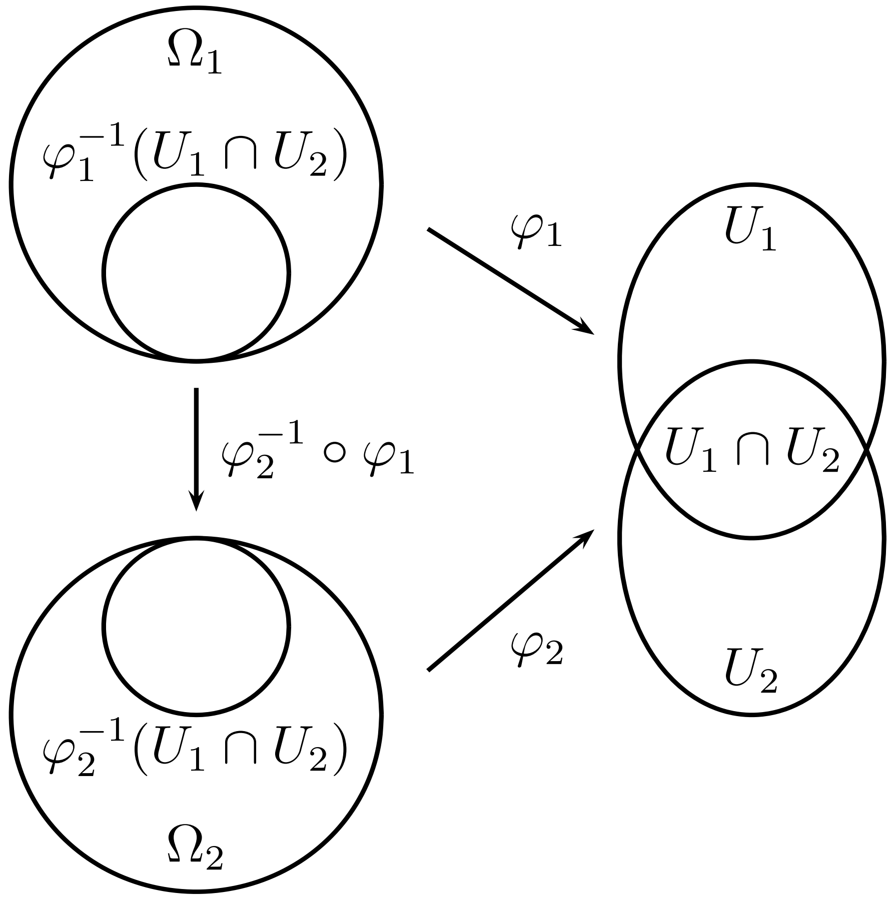

Figure 4.4: Parametrizations and transition functions

Using Definition 4.1, it may be quite hard to prove that a space is a manifold. Therefore, it is handy to have alternate characterizations such as those given in the next Proposition:

Proposition 4.3. *A subset, $M \subseteq \mathbb{R}^{m+k}$, is an m-dimensional manifold iff either*

(1) For every $p \in M$, there is some open subset, $W \subseteq \mathbb{R}^{m+k}$, with $p \in W$ and a (smooth) submersion, $f\colon W \to \mathbb{R}^k$, so that $W \cap M = f^{-1}(0)$,
or

(2) For every $p \in M$, there is some open subset, $W \subseteq \mathbb{R}^{m+k}$, with $p \in W$ and a (smooth) map, $f\colon W \to \mathbb{R}^k$, so that $f'(p)$ is surjective and $W \cap M = f^{-1}(0)$.

See Figure 4.5.

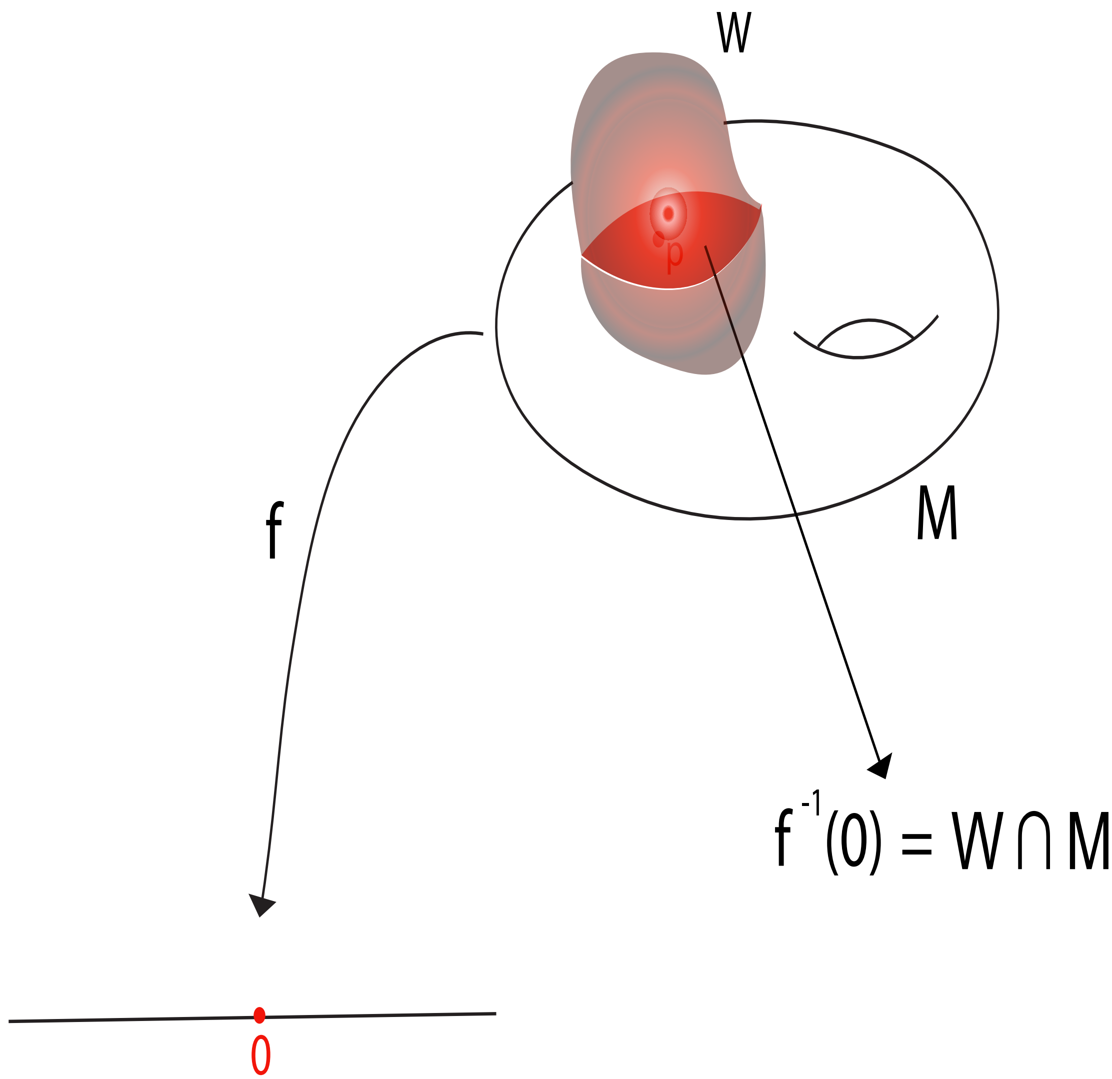

Figure 4.5: An illustration of Proposition 4.3, where M is the torus, $m = 2$, and $k = 1$. Note that $f^{-1}(0)$ is the pink patch of the torus, i.e. the zero level set of the open ball W.

Observe that condition (2), although apparently weaker than condition (1), is in fact equivalent to it, but more convenient in practice.

This is because to say that $f'(p)$ is surjective means that the Jacobian matrix of $f'(p)$ has rank k, which means that some determinant is nonzero, and because the determinant function is continuous this must hold in some open subset $W_1 \subseteq W$ containing p.

Consequently, the restriction, f_1, of f to W_1 is indeed a submersion and

$$f_1^{-1}(0) = W_1 \cap f^{-1}(0) = W_1 \cap W \cap M = W_1 \cap M.$$

The proof is based on two technical lemmas that are proved using the inverse function theorem.

Lemma 4.4. *Let $U \subseteq \mathbb{R}^m$ be an open subset of $\mathbb{R}^m$ and pick some $a \in U$. If $f \colon U \to \mathbb{R}^n$ is a smooth immersion at a, i.e., df_a is injective (so, $m \leq n$), then there is an open set, $V \subseteq \mathbb{R}^n$, with $f(a) \in V$, an open subset, $U' \subseteq U$, with $a \in U'$ and $f(U') \subseteq V$, an open subset $O \subseteq \mathbb{R}^{n-m}$, and a diffeomorphism, $\theta \colon V \to U' \times O$, so that*

$$\theta(f(x_1, \ldots, x_m)) = (x_1, \ldots, x_m, 0, \ldots, 0),$$

for all $(x_1, \ldots, x_m) \in U'$, as illustrated in the diagram below

$$\begin{tikzcd} U' \subseteq U \arrow[r, "f"] \arrow[dr, "in_1"'] & f(U') \subseteq V \arrow[d, "\theta"] \\ & U' \times O \end{tikzcd}$$

where $in_1(x_1, \ldots, x_m) = (x_1, \ldots, x_m, 0, \ldots, 0)$; see Figure 4.6.

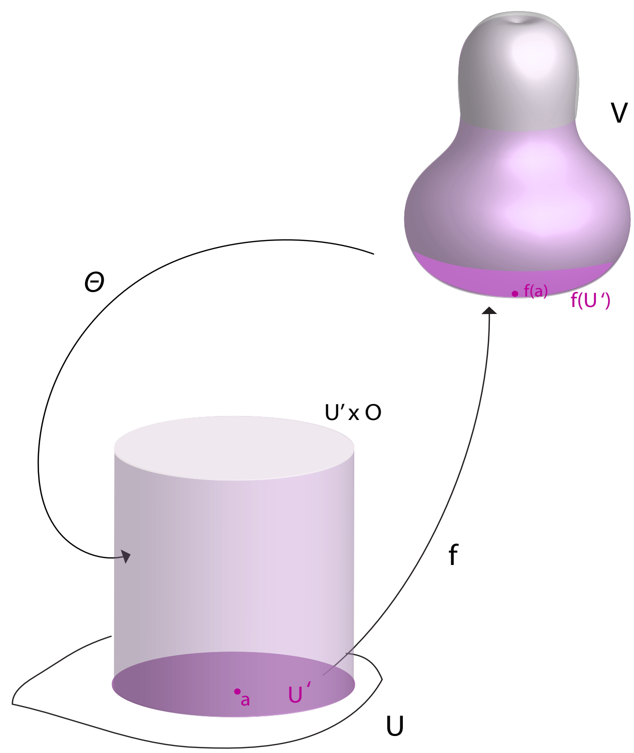

Figure 4.6: An illustration of Lemma 4.4, where $m = 2$ and $n = 3$. Note that U' is the base of the solid cylinder and θ is the diffeomorphism between the solid cylinder and the solid gourd shaped V. The composition $\theta \circ f$ injects U' into $U' \times O$.

Lemma 4.5. *Let $W \subseteq \mathbb{R}^m$ be an open subset of $\mathbb{R}^m$ and pick some $a \in W$. If $f\colon W \to \mathbb{R}^n$ is a smooth submersion at a, i.e., df_a is surjective (so, $m \geq n$), then there is an open set, $V \subseteq W \subseteq \mathbb{R}^m$, with $a \in V$, and a diffeomorphism, ψ, with domain $O \subseteq \mathbb{R}^m$, so that $\psi(O) = V$ and*

$$f(\psi(x_1, \ldots, x_m)) = (x_1, \ldots, x_n),$$

for all $(x_1, \ldots, x_m) \in O$, as illustrated in the diagram below

$$\begin{tikzcd} O \subseteq \mathbb{R}^m \arrow[r, "\psi"] \arrow[d, "\pi"'] & V \subseteq W \subseteq \mathbb{R}^m \arrow[dl, "f"] \\ \mathbb{R}^n, & \end{tikzcd}$$

where $\pi(x_1, \ldots, x_m) = (x_1, \ldots, x_n)$; see Figure 4.7.

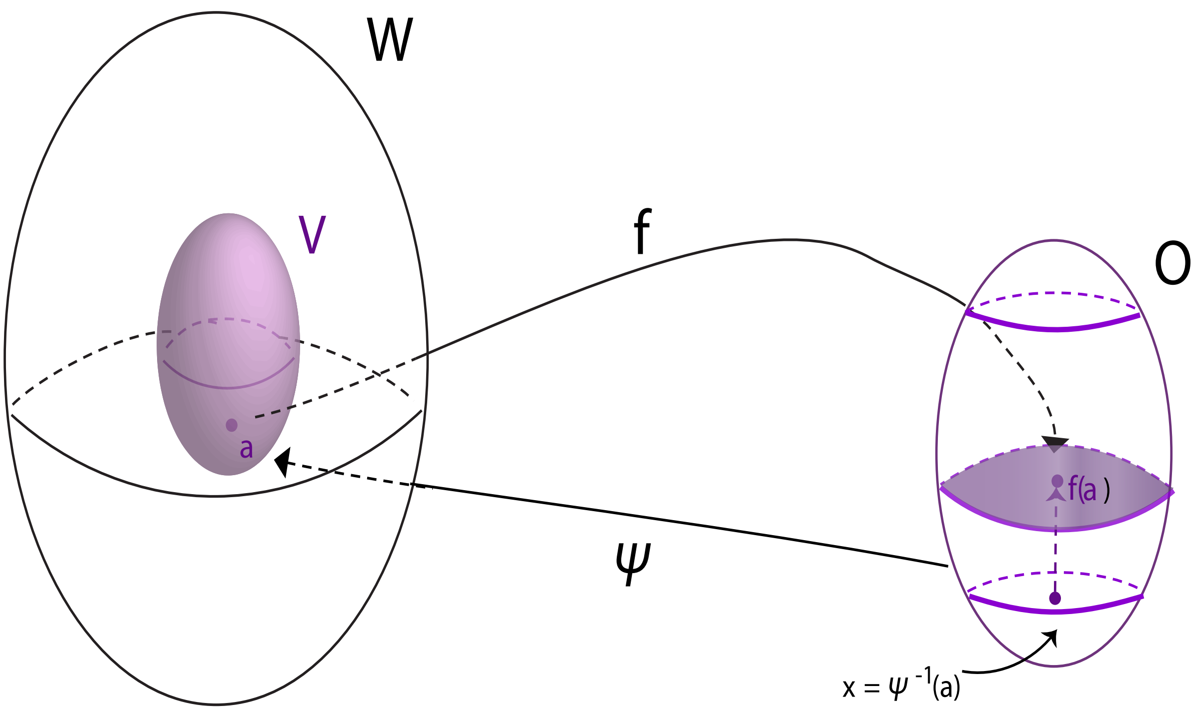

Figure 4.7: An illustration of Lemma 4.5, where $m = 3$ and $n = 2$. Note that ψ is the diffeomorphism between the 0 and the solid purple ball V. The composition $f \circ \psi$ projects O onto its equatorial pink disk.

Theorem 4.6. *A nonempty subset, $M \subseteq \mathbb{R}^N$, is an m-manifold (with $1 \leq m \leq N$) iff any of the following conditions hold:*

(1) For every $p \in M$, there are two open subsets $\Omega \subseteq \mathbb{R}^m$ and $U \subseteq M$, with $p \in U$, and a smooth function $\varphi\colon \Omega \to \mathbb{R}^N$ such that φ is a homeomorphism between Ω and $U = \varphi(\Omega)$, and $\varphi'(0)$ is injective, where $p = \varphi(0)$.

(2) For every $p \in M$, there are two open sets $O, W \subseteq \mathbb{R}^N$ with $0_N \in O$ and $p \in M \cap W$, and a smooth diffeomorphism $\varphi\colon O \to W$, such that $\varphi(0_N) = p$ and

$$\varphi(O \cap (\mathbb{R}^m \times \{0_{N-m}\})) = M \cap W.$$

(3) For every $p \in M$, there is some open subset, $W \subseteq \mathbb{R}^N$, with $p \in W$ and a smooth submersion $f\colon W \to \mathbb{R}^{N-m}$, so that $W \cap M = f^{-1}(0)$.

(4) *For every* $p \in M$*, there is some open subset,* $W \subseteq \mathbb{R}^N$*, and* $N - m$ *smooth functions,* $f_i \colon W \to \mathbb{R}$*, so that the linear forms* $df_1(p), \ldots, df_{N-m}(p)$ *are linearly independent and*

$$W \cap M = f_1^{-1}(0) \cap \cdots \cap f_{N-m}^{-1}(0).$$

See Figure 4.8.

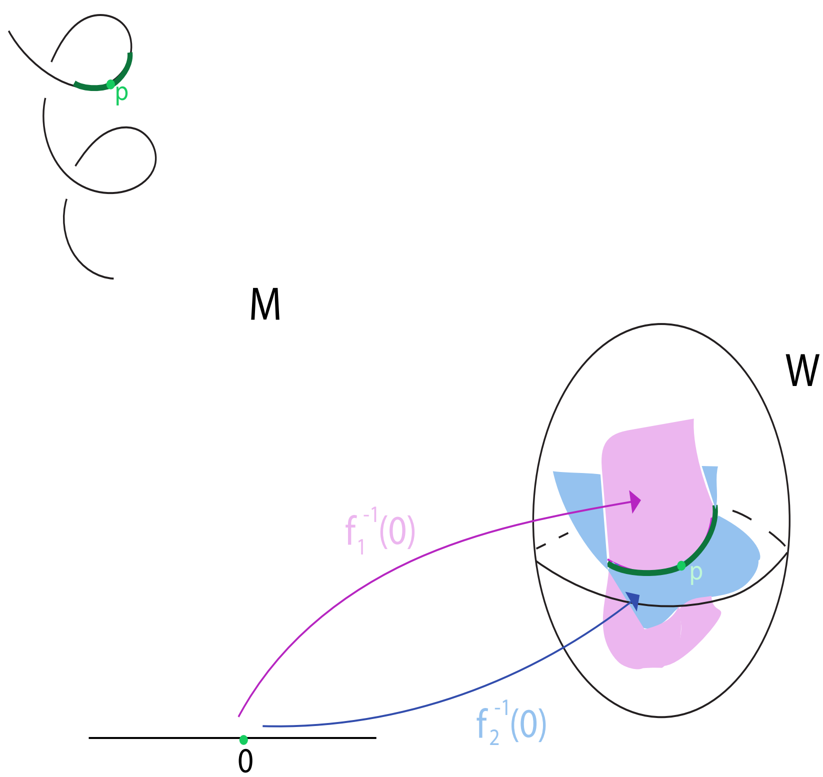

Figure 4.8: An illustration of Condition (4) in Theorem 4.6, where $N = 3$ and $m = 1$. The manifold M is the helix in $\mathbb{R}^3$. The dark green portion of M is magnified in order to show that it is the intersection of the pink surface, $f_1^{-1}(0)$, and the blue surface, $f_2^{-1}(0)$.

Condition (4) says that locally (that is, in a small open set of M containing $p \in M$), M is "cut out" by $N - m$ smooth functions, $f_i \colon W \to \mathbb{R}$, in the sense that the portion of the manifold $M \cap W$ is the intersection of the $N - m$ hypersurfaces, $f_i^{-1}(0)$, (the zero-level sets of the f_i) and that this intersection is "clean", which means that the linear forms $df_1(p), \ldots, df_{N-m}(p)$ are linearly independent.

As an illustration of Theorem 4.6, the sphere

$$S^n = \{x \in \mathbb{R}^{n+1} \mid \|x\|_2^2 - 1 = 0\}$$

is an n-dimensional manifold in $\mathbb{R}^{n+1}$.

Indeed, the map $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ given by $f(x) = \|x\|_2^2 - 1$ is a submersion, since

$$df(x)(y) = 2\sum_{k=1}^{n+1} x_k y_k.$$

The rotation group, $\mathbf{SO}(n)$, is an $\frac{n(n-1)}{2}$-dimensional manifold in $\mathbb{R}^{n^2}$.

Indeed, $\mathbf{GL}^+(n)$ is an open subset of $\mathbb{R}^{n^2}$ (recall, $\mathbf{GL}^+(n) = \{A \in \mathbf{GL}(n) \mid \det(A) > 0\}$) and if f is defined by

$$f(A) = A^\top A - I,$$

where $A \in \mathbf{GL}^+(n)$, then $f(A)$ is symmetric, so $f(A) \in \mathbf{S}(n) = \mathbb{R}^{\frac{n(n+1)}{2}}$.

We showed earlier that

$$df(A)(H) = A^\top H + H^\top A.$$

But then, $df(A)$ is surjective for all $A \in \mathbf{SO}(n)$, because if S is any symmetric matrix, we see that

$$df(A)\left(\frac{AS}{2}\right) = S.$$

As $\mathbf{SO}(n) = f^{-1}(0)$, we conclude that $\mathbf{SO}(n)$ is indeed a manifold.

A similar argument proves that $\mathbf{O}(n)$ is an $\frac{n(n-1)}{2}$-dimensional manifold.

Using the map, $f\colon \mathbf{GL}(n) \to \mathbb{R}$, given by $A \mapsto \det(A)$, we can prove that $\mathbf{SL}(n)$ is a manifold of dimension n^2-1.

Remark: We have $df(A)(B) = \det(A)\mathrm{tr}(A^{-1}B)$, for every $A \in \mathbf{GL}(n)$.

A class of manifolds generalizing the spheres and the orthogonal groups are the *Stiefel manifolds*.

For any $n \geq 1$ and any k with $1 \leq k \leq n$, let $S(k, n)$ be the set of all *orthonormal k-frames*; that is, of k-tuples of orthonormal vectors $(u_1, \ldots, u_k)$ with $u_i \in \mathbb{R}^n$.

Obviously $S(1, n) = S^{n-1}$, and $S(n, n) = \mathbf{O}(n)$.

Every orthonormal k-frame $(u_1, \ldots, u_k)$ can be represented by an $n \times k$ matrix Y over the canonical basis of $\mathbb{R}^n$, and such a matrix Y satisfies the equation

$$Y^\top Y = I.$$

Thus, $S(k, n)$ can be viewed as a subspace of $\mathrm{M}_{n,k}$. We claim that $S(k, n)$ is a manifold.

Let $W = \{A \in \mathrm{M}_{n,k} \mid \det(A^\top A) > 0\}$, an open subset of $\mathrm{M}_{n,k}$ such that $S(k, n) \subseteq W$.

Generalizing the situation involving $\mathbf{SO}(n)$, define the function $f\colon W \to \mathbf{S}(k)$ by

$$f(A) = A^{\top}A - I.$$

Basically the same computation as in the case of $\mathbf{SO}(n)$ yields

$$df(A)(H) = A^{\top}H + H^{\top}A.$$

The proof that $df(A)$ is surjective for all $A \in S(k, n)$ is the same as before, because only the equation $A^{\top}A = I$ is needed.

As $S(k, n) = f^{-1}(0)$, we conclude that $S(k, n)$ is a smooth manifold of dimension

$$nk - \frac{k(k+1)}{2} = k(n-k) + \frac{k(k-1)}{2}.$$

The third characterization of Theorem 4.6 suggests the following definition.

Definition 4.2. Let $f\colon \mathbb{R}^{m+k} \to \mathbb{R}^k$ be a smooth function. A point, $p \in \mathbb{R}^{m+k}$, is called a *critical point (of f)* iff df_p is *not* surjective and a point $q \in \mathbb{R}^k$ is called a *critical value (of f)* iff $q = f(p)$, for some critical point, $p \in \mathbb{R}^{m+k}$.

A point $p \in \mathbb{R}^{m+k}$ is a *regular point (of f)* iff p is not critical, i.e., df_p is surjective, and a point $q \in \mathbb{R}^k$ is a *regular value (of f)* iff it is not a critical value.

In particular, any $q \in \mathbb{R}^k - f(\mathbb{R}^{m+k})$ is a regular value and $q \in f(\mathbb{R}^{m+k})$ is a regular value iff *every* $p \in f^{-1}(q)$ is a regular point (but, in contrast, q is a critical value iff *some* $p \in f^{-1}(q)$ is critical).

Part (3) of Theorem 4.6 implies the following useful proposition:

Proposition 4.7. *Given any smooth function,* $f\colon \mathbb{R}^{m+k} \to \mathbb{R}^k$*, for every regular value,* $q \in f(\mathbb{R}^{m+k})$*, the preimage,* $Z = f^{-1}(q)$*, is a manifold of dimension* m*.*

Definition 4.2 and Proposition 4.7 can be generalized to manifolds

Regular and critical values of smooth maps play an important role in differential topology.

Firstly, given a smooth map, $f\colon \mathbb{R}^{m+k} \to \mathbb{R}^k$, almost every point of $\mathbb{R}^k$ is a regular value of f.

To make this statement precise, one needs the notion of a *set of measure zero*.

Then, *Sard's theorem* says that the set of critical values of a smooth map has measure zero.

Secondly, if we consider smooth functions, $f\colon \mathbb{R}^{m+1} \to \mathbb{R}$, a point $p \in \mathbb{R}^{m+1}$ is critical iff $df_p = 0$.

Then, we can use second order derivatives to further classify critical points. The *Hessian matrix* of f (at p) is the matrix of second-order partials

$$H_f(p) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p)\right)$$

and a critical point p is a *nondegenerate critical point* if $H_f(p)$ is a nonsingular matrix.

The remarkable fact is that, at a nondegenerate critical point, p, the local behavior of f is completely determined, in the sense that after a suitable change of coordinates (given by a smooth diffeomorphism)

$$f(x) = f(p) - x_1^2 - \cdots - x_\lambda^2 + x_{\lambda+1}^2 + \cdots + x_{m+1}^2$$

near p, where λ called the *index of f at p* is an integer which depends only on p (in fact, λ is the number of negative eigenvalues of $H_f(p)$).

This result is known as *Morse lemma* (after Marston Morse, 1892-1977).

Smooth functions whose critical points are all nondegenerate are called *Morse functions*.

It turns out that every smooth function, $f\colon \mathbb{R}^{m+1} \to \mathbb{R}$, gives rise to a large supply of Morse functions by adding a linear function to it.

More precisely, the set of $a \in \mathbb{R}^{m+1}$ for which the function f_a given by

$$f_a(x) = f(x) + a_1 x_1 + \cdots + a_{m+1} x_{m+1}$$

is not a Morse function has measure zero.

Morse functions can be used to study topological properties of manifolds.

In a sense to be made precise and under certain technical conditions, *a Morse function can be used to reconstuct a manifold by attaching cells, up to homotopy equivalence*.

However, these results are way beyond the scope of these notes.

Let us now review the definitions of a smooth curve in a manifold and the tangent vector at a point of a curve.

Definition 4.3. Let M be an m-dimensional manifold in $\mathbb{R}^N$. A *smooth curve γ in M* is any function $\gamma\colon I \to M$ where I is an open interval in $\mathbb{R}$ and such that for every $t \in I$, letting $p = \gamma(t)$, there is some parametrization $\varphi\colon \Omega \to U$ of M at p and some open interval $(t - \epsilon,\, t + \epsilon) \subseteq I$ such that the curve $\varphi^{-1} \circ \gamma\colon (t - \epsilon,\, t + \epsilon) \to \mathbb{R}^m$ is smooth. The notion of a smooth curve is illustrated in Figure 4.9.

Using Lemma 4.2, it is easily shown that Definition 4.3 does not depend on the choice of the parametrization $\varphi\colon \Omega \to U$ at p.

Lemma 4.2 also implies that γ viewed as a curve $\gamma\colon I \to \mathbb{R}^N$ is smooth.

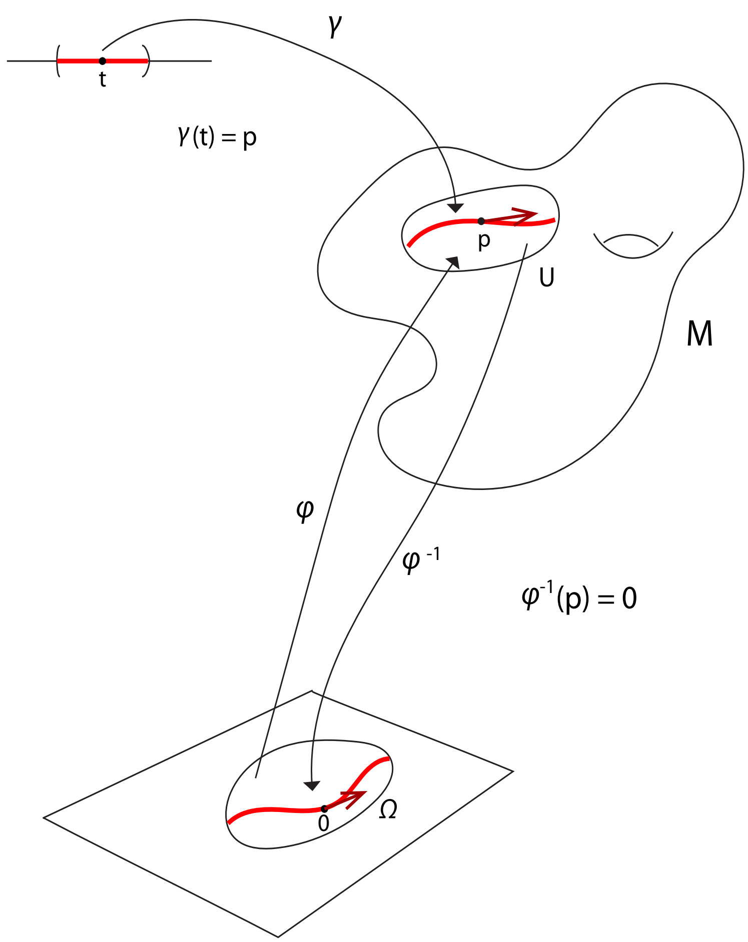

Figure 4.9: A smooth curve in a manifold M.

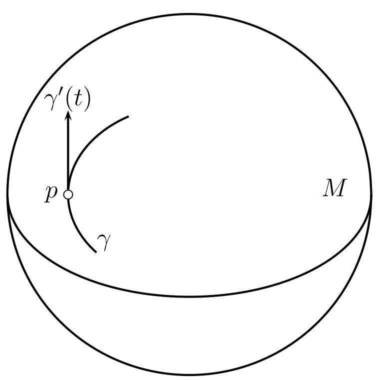

Figure 4.10: Tangent vector to a curve on a manifold.

Then the *tangent vector to the curve* $\gamma\colon I \to \mathbb{R}^N$ *at* t, denoted by $\gamma'(t)$, is the value of the derivative of γ at t (a vector in $\mathbb{R}^N$) computed as usual:

$$\gamma'(t) = \lim_{h \mapsto 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

Given any point $p \in M$, we will show that the set of tangent vectors to all smooth curves in M through p is a vector space isomorphic to the vector space $\mathbb{R}^m$.

Given a smooth curve $\gamma\colon I \to M$, for any $t \in I$, letting $p = \gamma(t)$, since M is a manifold, there is a parametrization $\varphi\colon \Omega \to U$ such that $\varphi(0_m) = p \in U$ and some open interval $J \subseteq I$ with $t \in J$ and such that the function

$$\varphi^{-1} \circ \gamma\colon J \to \mathbb{R}^m$$

is a smooth curve, since γ is a smooth curve.

Letting $\alpha = \varphi^{-1} \circ \gamma$, the derivative $\alpha'(t)$ is well-defined, and it is a vector in $\mathbb{R}^m$.

But $\varphi \circ \alpha\colon J \to M$ is also a smooth curve, which agrees with γ on J, and by the chain rule,

$$\gamma'(t) = \varphi'(0_m)(\alpha'(t)),$$

since $\alpha(t) = 0_m$ (because $\varphi(0_m) = p$ and $\gamma(t) = p$).

Observe that $\gamma'(t)$ is a vector in $\mathbb{R}^N$.

Now, for every vector $v \in \mathbb{R}^m$, the curve $\alpha\colon J \to \mathbb{R}^m$ defined such that

$$\alpha(u) = (u - t)v$$

for all $u \in J$ is clearly smooth, and $\alpha'(t) = v$.

This shows that the set of tangent vectors at t to all smooth curves (in $\mathbb{R}^m$) passing through 0_m is the entire vector space $\mathbb{R}^m$.

Since every smooth curve $\gamma\colon I \to M$ agrees with a curve of the form $\varphi \circ \alpha\colon J \to M$ for some smooth curve $\alpha\colon J \to \mathbb{R}^m$ (with $J \subseteq I$) as explained above, and since it is assumed that $\varphi'(0_m)$ is injective, $\varphi'(0_m)$ maps the vector space $\mathbb{R}^m$ injectively to the set of tangent vectors to γ at p, as claimed.

All this is summarized in the following definition.

Definition 4.4. Let M be an m-dimensional manifold in $\mathbb{R}^N$. For every point $p \in M$, the *tangent space T_pM at p* is the set of all vectors in $\mathbb{R}^N$ of the form $\gamma'(0)$, where $\gamma \colon I \to M$ is any smooth curve in M such that $p = \gamma(0)$.

The set T_pM is a vector space isomorphic to $\mathbb{R}^m$. Every vector $v \in T_pM$ is called a *tangent vector to M at p*.

We can now define Lie groups.

Definition 4.5. A *Lie group* is a nonempty subset G of $\mathbb{R}^N$ $(N \geq 1)$ satisfying the following conditions:

(a) G is a group.

(b) G is a manifold in $\mathbb{R}^N$.

(c) The group operation $\cdot : G \times G \to G$ and the inverse map ${}^{-1} \colon G \to G$ are smooth.

Actually, we haven't defined yet what a smooth map between manifolds is (in clause (c)).

This notion is explained in Definition 4.8, but we feel that most readers will appreciate seeing the formal definition a Lie group, as early as possible.

It is immediately verified that $\mathbf{GL}(n, \mathbb{R})$ is a Lie group. Since all the Lie groups that we are considering are subgroups of $\mathbf{GL}(n, \mathbb{R})$, the following definition is in order.

Definition 4.6. A *linear Lie group* is a subgroup G of $\mathbf{GL}(n, \mathbb{R})$ (for some $n \geq 1$) which is a smooth manifold in $\mathbb{R}^{n^2}$.

Let $\mathrm{M}_n(\mathbb{R})$ denote the set of all real $n \times n$ matrices (invertible or not). If we recall that the exponential map

$$\exp\colon A \mapsto e^A$$

is well defined on $\mathrm{M}_n(\mathbb{R})$, we have the following crucial theorem due to Von Neumann and Cartan:

Theorem 4.8. *(Von Neumann and Cartan, 1927) A nondiscrete closed subgroup G of $\mathbf{GL}(n, \mathbb{R})$ is a linear Lie group. Furthermore, the set $\mathfrak{g}$ defined such that*

$$\mathfrak{g} = \{X \in \mathrm{M}_n(\mathbb{R}) \mid e^{tX} \in G \text{ for all } t \in \mathbb{R}\}$$

is a nontrivial vector space equal to the tangent space $T_I G$ at the identity I, and $\mathfrak{g}$ is closed under the Lie bracket $[-, -]$ defined such that $[A, B] = AB - BA$ for all $A, B \in \mathrm{M}_n(\mathbb{R})$.

Theorem 4.8 applies even when G is a discrete subgroup, but in this case, $\mathfrak{g}$ is trivial (i.e., $\mathfrak{g} = \{0\}$).

For example, the set of nonzero reals $\mathbb{R}^* = \mathbb{R} - \{0\} = \mathbf{GL}(1, \mathbb{R})$ is a Lie group under multiplication, and the subgroup

$$H = \{2^n \mid n \in \mathbb{Z}\}$$

is a discrete subgroup of $\mathbb{R}^*$. Thus, H is a Lie group.

On the other hand, the set $\mathbb{Q}^* = \mathbb{Q} - \{0\}$ of nonzero rational numbers is a multiplicative subgroup of $\mathbb{R}^*$, but it is not closed, since $\mathbb{Q}$ is dense in $\mathbb{R}$.

If G is closed and not discrete, we must have $n \geq 1$, and $\mathfrak{g}$ has dimension n.

The first step to prove Theorem 4.8 is this:

Proposition 4.9. *Given any closed subgroup G in $\mathbf{GL}(n, \mathbb{R})$, the set*

$$\begin{aligned}\mathfrak{g} = \{X \in \mathrm{M}_n(\mathbb{R}) \mid X = \gamma'(0), \gamma\colon J \to G \\ \text{is a } C^1 \text{ curve in } \mathrm{M}_n(\mathbb{R}) \text{ such that } \gamma(0) = I\}\end{aligned}$$

satisfies the following properties:

(1) $\mathfrak{g}$ *is a vector subspace of* $\mathrm{M}_n(\mathbb{R})$.

(2) For every $X \in \mathrm{M}_n(\mathbb{R})$, we have $X \in \mathfrak{g}$ iff $e^{tX} \in G$ for all $t \in \mathbb{R}$.

(3) For every $X \in \mathfrak{g}$ and for every $g \in G$, we have $gXg^{-1} \in \mathfrak{g}$.

(4) $\mathfrak{g}$ *is closed under the Lie bracket.*

The second step to prove Theorem 4.8 is this:

Proposition 4.10. *Let G be a subgroup of $\mathbf{GL}(n, \mathbb{R})$, and assume that G is closed and not discrete. Then,* $\dim(\mathfrak{g}) \geq 1$*, and the exponential map is a diffeomorphism of a neighborhood of* 0 *in* $\mathfrak{g}$ *onto a neighborhood of I in G. Furthermore, there is an open subset $\Omega \subseteq$* $\mathrm{M}_n(\mathbb{R})$ *with $0_{n,n} \in \Omega$, an open subset $W \subseteq \mathbf{GL}(n, \mathbb{R})$ with $I \in W$, and a diffeomorphism $\Phi \colon \Omega \to W$ such that*

$$\Phi(\Omega \cap \mathfrak{g}) = W \cap G.$$

With the help of Theorem 4.8 it is now very easy to prove that $\mathbf{SL}(n)$, $\mathbf{O}(n)$, $\mathbf{SO}(n)$, $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{U}(n)$, and $\mathbf{SU}(n)$ are Lie groups. It suffices to show that these subgroups of $\mathbf{GL}(n, \mathbb{R})$ ($\mathbf{GL}(2n, \mathbb{R})$ in the case of $\mathbf{SL}(n, \mathbb{C})$, $\mathbf{U}(n)$, and $\mathbf{SU}(n)$) are closed.

We can also prove that $\mathbf{SE}(n)$ is a Lie group as follows.

Recall that we can view every element of $\mathbf{SE}(n)$ as a real $(n+1)\times(n+1)$ matrix

$$\begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix}$$

where $R \in \mathbf{SO}(n)$ and $U \in \mathbb{R}^n$.

In fact, such matrices belong to $\mathbf{SL}(n+1)$.

This embedding of $\mathbf{SE}(n)$ into $\mathbf{SL}(n+1)$ is a group homomorphism, since the group operation on $\mathbf{SE}(n)$ corresponds to multiplication in $\mathbf{SL}(n+1)$:

$$\begin{pmatrix} RS & RV + U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R & U \\ 0 & 1 \end{pmatrix} \begin{pmatrix} S & V \\ 0 & 1 \end{pmatrix}.$$

Note that the inverse is given by

$$\begin{pmatrix} R^{-1} & -R^{-1}U \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R^{\top} & -R^{\top}U \\ 0 & 1 \end{pmatrix}.$$

Also note that the embedding shows that as a manifold, $\mathbf{SE}(n)$ is diffeomorphic to $\mathbf{SO}(n) \times \mathbb{R}^n$ (given a manifold M_1 of dimension m_1 and a manifold M_2 of dimension m_2, the product $M_1 \times M_2$ can be given the structure of a manifold of dimension $m_1 + m_2$ in a natural way).

Thus, $\mathbf{SE}(n)$ is a Lie group with underlying manifold $\mathbf{SO}(n) \times \mathbb{R}^n$, and in fact, a subgroup of $\mathbf{SL}(n+1)$.

Even though $\mathbf{SE}(n)$ is diffeomorphic to $\mathbf{SO}(n) \times \mathbb{R}^n$ as a manifold, it is *not* isomorphic to $\mathbf{SO}(n) \times \mathbb{R}^n$ as a group, because the group multiplication on $\mathbf{SE}(n)$ is not the multiplication on $\mathbf{SO}(n) \times \mathbb{R}^n$. Instead, $\mathbf{SE}(n)$ is a *semidirect product* of $\mathbf{SO}(n)$ by $\mathbb{R}^n$.

Returning to Theorem 4.8, the vector space $\mathfrak{g}$ is called the *Lie algebra* of the Lie group G.

Lie algebras are defined as follows.

Definition 4.7. A *(real) Lie algebra* $\mathcal{A}$ is a real vector space together with a bilinear map $[\cdot, \cdot] \colon \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ called the *Lie bracket* on $\mathcal{A}$ such that the following two identities hold for all $a, b, c \in \mathcal{A}$:

$$[a,\, a] = 0,$$

and the so-called *Jacobi identity*

$$[a,\, [b,\, c]] + [c,\, [a,\, b]] + [b,\, [c,\, a]] = 0.$$

It is immediately verified that $[b, a] = -[a, b]$.

In view of Theorem 4.8, the vector space $\mathfrak{g} = T_I G$ associated with a Lie group G is indeed a Lie algebra. Furthermore, the exponential map $\exp \colon \mathfrak{g} \to G$ is well-defined.

In general, exp is neither injective nor surjective, as we observed earlier.

Theorem 4.8 also provides a kind of recipe for "computing" the Lie algebra $\mathfrak{g} = T_I G$ of a Lie group G.

Indeed, $\mathfrak{g}$ is the tangent space to G at I, and thus we can use curves to compute tangent vectors.

Actually, for every $X \in T_I G$, the map

$$\gamma_X \colon t \mapsto e^{tX}$$

is a smooth curve in G, and it is easily shown that $\gamma_X'(0) = X$. Thus, we can use these curves.

As an illustration, we show that the Lie algebras of $\mathbf{SL}(n)$ and $\mathbf{SO}(n)$ are the matrices with null trace and the skew symmetric matrices.

Let $t \mapsto R(t)$ be a smooth curve in $\mathbf{SL}(n)$ such that $R(0) = I$. We have $\det(R(t)) = 1$ for all $t \in (-\epsilon, \epsilon)$.

Using the chain rule, we can compute the derivative of the function

$$t \mapsto \det(R(t))$$

at

$t = 0$, and we get

$$\det{}'_I(R'(0)) = 0.$$

It is an easy exercise to prove that

$$\det{}'_I(X) = \mathrm{tr}(X),$$

and thus $\mathrm{tr}(R'(0)) = 0$, which says that the tangent vector $X = R'(0)$ has null trace.

Another proof consists in observing that $X \in \mathfrak{sl}(n, \mathbb{R})$ iff

$$\det(e^{tX}) = 1$$

for all $t \in \mathbb{R}$. Since $\det(e^{tX}) = e^{\mathrm{tr}(tX)}$, for $t = 1$, we get $\mathrm{tr}(X) = 0$, as claimed.

Clearly, $\mathfrak{sl}(n, \mathbb{R})$ has dimension $n^2 - 1$.

Let $t \mapsto R(t)$ be a smooth curve in $\mathbf{SO}(n)$ such that $R(0) = I$. Since each $R(t)$ is orthogonal, we have

$$R(t)\, R(t)^{\top} = I$$

for all $t \in (-\epsilon, \epsilon)$.

Taking the derivative at $t = 0$, we get

$$R'(0)\, R(0)^\top + R(0)\, R'(0)^\top = 0,$$

but since $R(0) = I = R(0)^\top$, we get

$$R'(0) + R'(0)^\top = 0,$$

which says that the tangent vector $X = R'(0)$ is skew symmetric.

Since the diagonal elements of a skew symmetric matrix are null, the trace is automatically null, and the condition $\det(R) = 1$ yields nothing new.

This shows that $\mathfrak{o}(n) = \mathfrak{so}(n)$. It is easily shown that $\mathfrak{so}(n)$ has dimension $n(n-1)/2$.

As a concrete example, the Lie algebra $\mathfrak{so}(3)$ of $\mathbf{SO}(3)$ is the real vector space consisting of all 3×3 real skew symmetric matrices. Every such matrix is of the form

$$\begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix}$$

where $b, c, d \in \mathbb{R}$.

The Lie bracket $[A, B]$ in $\mathfrak{so}(3)$ is also given by the usual commutator, $[A, B] = AB - BA$.

We can define an isomorphism of Lie algebras $\psi\colon (\mathbb{R}^3, \times) \to \mathfrak{so}(3)$ by the formula

$$\psi(b, c, d) = \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix}.$$

It is indeed easy to verify that

$$\psi(u \times v) = [\psi(u),\, \psi(v)].$$

It is also easily verified that for any two vectors $u = (b, c, d)$ and $v = (b', c', d')$ in $\mathbb{R}^3$,

$$\psi(u)(v) = u \times v.$$

In robotics and in computer vision, $\psi(u)$ is often denoted by $u_\times$.

The exponential map $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is given by Rodrigues's formula (see Lemma 1.12):

$$e^A = \cos\theta\, I_3 + \frac{\sin\theta}{\theta} A + \frac{(1-\cos\theta)}{\theta^2} B,$$

or equivalently by

$$e^A = I_3 + \frac{\sin\theta}{\theta} A + \frac{(1-\cos\theta)}{\theta^2} A^2$$

if $\theta \neq 0$, where

$$A = \begin{pmatrix} 0 & -d & c \\ d & 0 & -b \\ -c & b & 0 \end{pmatrix},$$

$\theta = \sqrt{b^2 + c^2 + d^2}$, $B = A^2 + \theta^2 I_3$, and with $e^{0_3} = I_3$.

Using the above methods, it is easy to verify that the Lie algebras $\mathfrak{gl}(n, \mathbb{R})$, $\mathfrak{sl}(n, \mathbb{R})$, $\mathfrak{o}(n)$, and $\mathfrak{so}(n)$, are respectively $\mathrm{M}_n(\mathbb{R})$, the set of matrices with null trace, and the set of skew symmetric matrices (in the last two cases).

A similar computation can be done for $\mathfrak{gl}(n, \mathbb{C})$, $\mathfrak{sl}(n, \mathbb{C})$, $\mathfrak{u}(n)$, and $\mathfrak{su}(n)$, confirming the claims of Section 1.5.

It is easy to show that $\mathfrak{gl}(n, \mathbb{C})$ has dimension $2n^2$, $\mathfrak{sl}(n, \mathbb{C})$ has dimension $2(n^2 - 1)$, $\mathfrak{u}(n)$ has dimension n^2, and $\mathfrak{su}(n)$ has dimension $n^2 - 1$.

For example, the Lie algebra $\mathfrak{su}(2)$ of $\mathbf{SU}(2)$ (or S^3) is the real vector space consisting of all 2×2 (complex) skew Hermitian matrices of null trace.

As we showed, $\mathbf{SE}(n)$ is a Lie group, and its lie algebra $\mathfrak{se}(n)$ described in Section 1.7 is easily determined as the subalgebra of $\mathfrak{sl}(n+1)$ consisting of all matrices of the form

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix}$$

where $B \in \mathfrak{so}(n)$ and $U \in \mathbb{R}^n$.

Thus, $\mathfrak{se}(n)$ has dimension $n(n+1)/2$. The Lie bracket is given by

$$\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix}\begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} C & V \\ 0 & 0 \end{pmatrix}\begin{pmatrix} B & U \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} BC - CB & BV - CU \\ 0 & 0 \end{pmatrix}.$$

We conclude by indicating the relationship between homomorphisms of Lie groups and homomorphisms of Lie algebras.

Definition 4.8. Let M_1 (m_1-dimensional) and M_2 (m_2-dimensional) be manifolds in $\mathbb{R}^N$. A function $f\colon M_1 \to M_2$ is *smooth* if for every $p \in M_1$ there are parametrizations $\varphi\colon \Omega_1 \to U_1$ of M_1 at p and $\psi\colon \Omega_2 \to U_2$ of M_2 at $f(p)$ such that $f(U_1) \subseteq U_2$ and

$$\psi^{-1} \circ f \circ \varphi\colon \Omega_1 \to \mathbb{R}^{m_2}$$

is smooth; see Figure 4.11.

Using Lemma 4.2, it is easily shown that Definition 4.8 does not depend on the choice of the parametrizations $\varphi\colon \Omega_1 \to U_1$ and $\psi\colon \Omega_2 \to U_2$.

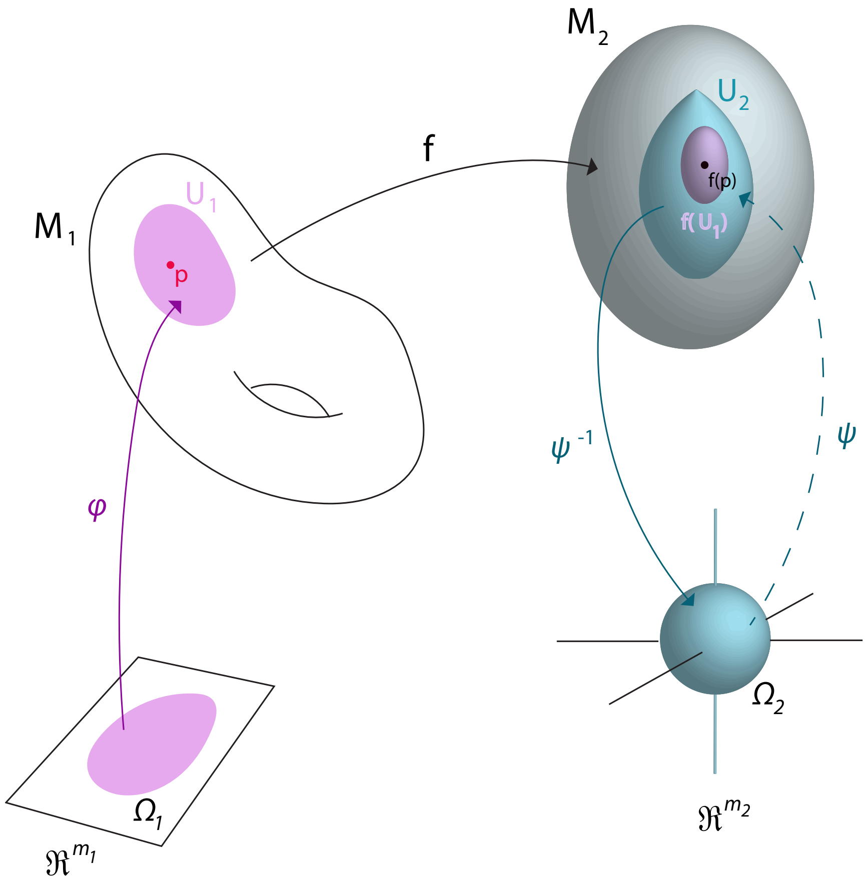

Figure 4.11: An illustration of a smooth map from the torus, M_1, to the solid ellipsoid M_2. The pink patch on M_1 is mapped into interior pink ellipsoid of M_2.

A smooth map f between manifolds is a *smooth diffeomorphism* if f is bijective and both f and f^{-1} are smooth maps.

Definition 4.9. Let M_1 (m_1-dimensional) and M_2 (m_2-dimensional) be manifolds in $\mathbb{R}^N$. For any smooth function $f\colon M_1 \to M_2$ and any $p \in M_1$, the function $f'_p\colon T_pM_1 \to T_{f(p)}M_2$, called the *tangent map of f at p, or derivative of f at p, or differential of f at p*, is defined as follows: For every $v \in T_pM_1$ and every smooth curve $\gamma\colon I \to M_1$ such that $\gamma(0) = p$ and $\gamma'(0) = v$,

$$f'_p(v) = (f \circ \gamma)'(0).$$

See Figure 4.12.

The map f'_p is also denoted by df_p or T_pf.

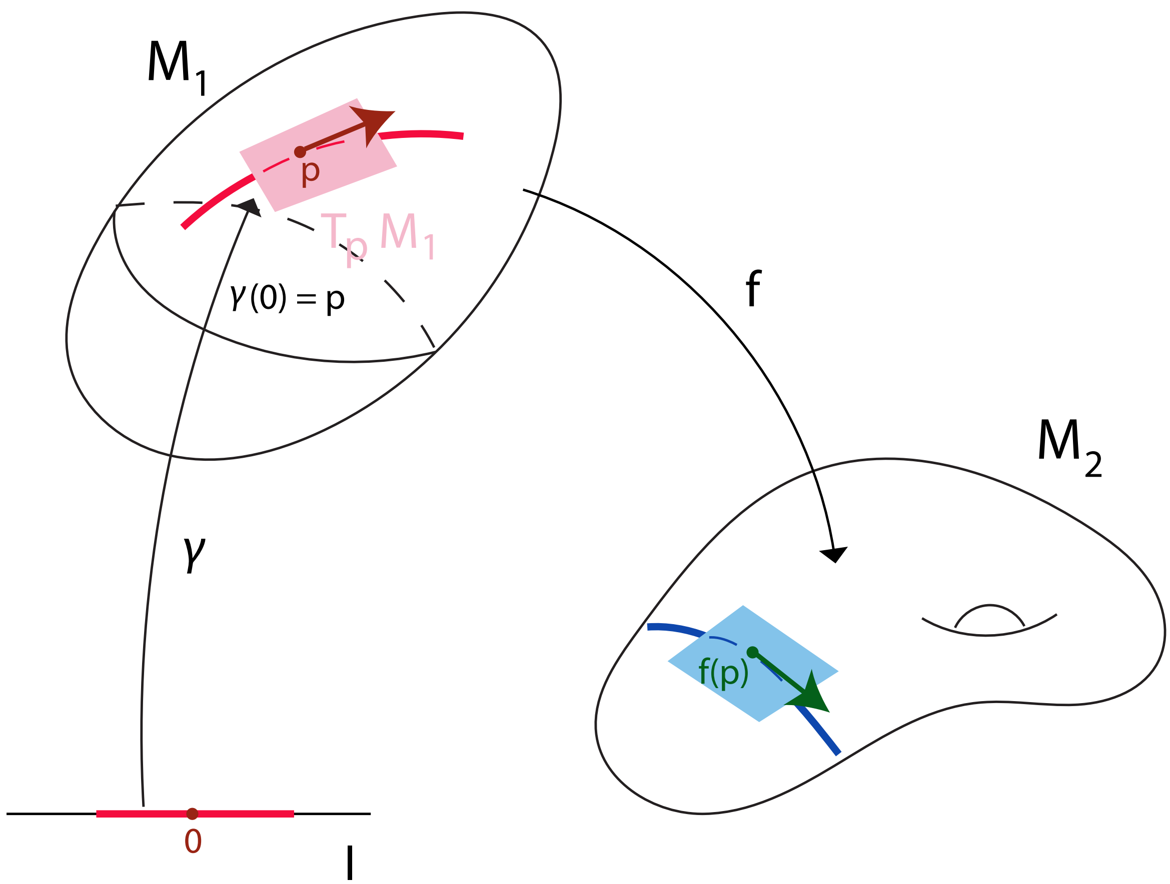

Figure 4.12: An illustration of the tangent map from T_pM_1 to $T_{f(p)}M_2$.

Doing a few calculations involving the facts that

$$f \circ \gamma = (f \circ \varphi) \circ (\varphi^{-1} \circ \gamma) \quad \text{and} \quad \gamma = \varphi \circ (\varphi^{-1} \circ \gamma)$$

and using Lemma 4.2, it is not hard to show that $f'_p(v)$ does not depend on the choice of the curve γ. It is easily shown that f'_p is a linear map.

Finally, we define homomorphisms of Lie groups and Lie algebras and see how they are related.

Definition 4.10. Given two Lie groups G_1 and G_2, a *homomorphism (or map) of Lie groups* is a function $f: G_1 \to G_2$ that is a homomorphism of groups and a smooth map (between the manifolds G_1 and G_2).

Given two Lie algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, a *homomorphism (or map) of Lie algebras* is a function $f\colon \mathcal{A}_1 \to \mathcal{A}_2$ that is a linear map between the vector spaces $\mathcal{A}_1$ and $\mathcal{A}_2$ and that preserves Lie brackets, i.e.,

$$f([A, B]) = [f(A), f(B)]$$

for all $A, B \in \mathcal{A}_1$.

An *isomorphism of Lie groups* is a bijective function f such that both f and f^{-1} are maps of Lie groups, and an *isomorphism of Lie algebras* is a bijective function f such that both f and f^{-1} are maps of Lie algebras.

If $f\colon G_1 \to G_2$ is a homomorphism of Lie groups, then $f'_I\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a homomorphism of Lie algebras, but in order to prove this, we need the adjoint representation Ad, so we postpone the proof.

The notion of a one-parameter group plays a crucial role in Lie group theory.

Definition 4.11. A smooth homomorphism $h\colon (\mathbb{R}, +) \to G$ from the additive group $\mathbb{R}$ to a Lie group G is called a *one-parameter group* in G.

All parameter groups of a linear Lie group can be determined explicitly.

Proposition 4.11. *Let G be any linear Lie group.*

1. *For every $X \in \mathfrak{g}$, the map $h(t) = e^{tX}$ is a one-parameter group in G.*
2. *Every one-parameter group $h\colon \mathbb{R} \to G$ is of the form $h(t) = e^{tZ}$, with $Z = h'(0)$.*

In summary, for every $Z \in \mathfrak{g}$, there is a unique one-parameter group h such that $h'(0) = Z$ given by $h(t) = e^{Zt}$.

The exponential map is natural in the following sense:

Proposition 4.12. *Given any two linear Lie groups G and H, for every Lie group homomorphism $f\colon G \to H$, the following diagram commutes:*

$$\begin{tikzcd}
G \arrow[r, "f"] & H \\
\mathfrak{g} \arrow[u, "\exp"] \arrow[r, "df_I"'] & \mathfrak{h} \arrow[u, "\exp"']
\end{tikzcd}$$

Alert readers must have noticed that we only defined the Lie algebra of a linear group.

In the more general case, we can still define the Lie algebra $\mathfrak{g}$ of a Lie group G as the tangent space $T_I G$ at the identity I.

The tangent space $\mathfrak{g} = T_I G$ is a vector space, but we need to define the Lie bracket.

This can be done in several ways. We explain briefly how this can be done in terms of so-called *adjoint representations*.

This has the advantage of not requiring the definition of left-invariant vector fields, but it is still a little bizarre!

Given a Lie group G, for every $a \in G$ we define *left translation* as the map $L_a \colon G \to G$ such that $L_a(b) = ab$ for all $b \in G$, and *right translation* as the map $R_a \colon G \to G$ such that $R_a(b) = ba$ for all $b \in G$.

The maps L_a and R_a are diffeomorphisms, and their derivatives play an important role.

The inner automorphisms

$$\mathbf{Ad}_a = R_{a^{-1}} \circ L_a \; (= R_{a^{-1}} L_a)$$

of G also play an important role.

Note that

$$\mathbf{Ad}_a(b) = R_{a^{-1}} L_a(b) = aba^{-1}.$$

The derivative

$$(\mathbf{Ad}_a)'_I \colon T_I G \to T_I G$$

of $\mathbf{Ad}_a \colon G \to G$ at I is an isomorphism of Lie algebras, and since $T_I G = \mathfrak{g}$, if we denote $(\mathbf{Ad}_a)'_I$ by Ad_a, we get a map

$$\mathrm{Ad}_a \colon \mathfrak{g} \to \mathfrak{g}.$$

The map $a \mapsto \mathrm{Ad}_a$ is a map of Lie groups

$$\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g}),$$

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

In the case of a linear group, we have

$$\mathrm{Ad}(a)(X) = \mathrm{Ad}_a(X) = aXa^{-1}$$

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

We are now almost ready to prove that if $f\colon G_1 \to G_2$ is a homomorphism of Lie groups, then $f'_I\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ is a homomorphism of Lie algebras.

What we need is to express the Lie bracket $[A, B]$ in terms of the derivative of an expression involving the adjoint representation Ad.

For any $A, B \in \mathfrak{g}$, we have

$$\big(\mathrm{Ad}_{e^{tA}}(B)\big)'(0) = (e^{tA}Be^{-tA})'(0) = AB - BA = [A, B].$$

Proposition 4.13. *If $f\colon G_1 \to G_2$ is a homomorphism of linear Lie groups, then the linear map $df_I\colon \mathfrak{g}_1 \to \mathfrak{g}_2$ satisfies the equation*

$$df_I(\mathrm{Ad}_a(X)) = \mathrm{Ad}_{f(a)}(df_I(X)),$$

for all $a \in G$ and all $X \in \mathfrak{g}_1$; that is, the following diagram commutes

$$\begin{tikzcd}
\mathfrak{g}_1 \arrow[r, "df_I"] \arrow[d, "\mathrm{Ad}_a"'] & \mathfrak{g}_2 \arrow[d, "\mathrm{Ad}_{f(a)}"] \\
\mathfrak{g}_1 \arrow[r, "df_I"'] & \mathfrak{g}_2
\end{tikzcd}$$

Furthermore, df_I is a homomorphism of Lie algebras.

If some additional assumptions are made about G_1 and G_2 (for example, connected, simply connected), it can be shown that f is pretty much determined by f'_I.

The derivative

$$\mathrm{Ad}'_I\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

of $\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g})$ at I is map of Lie algebras, and if we denote Ad'_I by ad, it is a map

$$\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}),$$

called the *adjoint representation of* $\mathfrak{g}$.

Recall that Theorem 4.8 immediately implies that the Lie algebra $\mathfrak{gl}(\mathfrak{g})$ of $\mathbf{GL}(\mathfrak{g})$ is the vector space $\mathrm{Hom}(\mathfrak{g}, \mathfrak{g})$ of all linear maps on $\mathfrak{g}$.

If we apply Proposition 4.12 to $\mathrm{Ad}\colon G \to \mathbf{GL}(\mathfrak{g})$, we obtain the equation

$$\mathrm{Ad}_{e^A} = e^{\mathrm{ad}_A} \quad \text{for all } A \in \mathfrak{g},$$

which is a generalization of the identity of Proposition 3.1.

In the case of a linear group, we have

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

for all $A, B \in \mathfrak{g}$.

This can be shown using Propositions 3.1 and 4.12

One can also check that the Jacobi identity on $\mathfrak{g}$ is equivalent to the fact that ad preserves Lie brackets, i.e., ad is a map of Lie algebras:

$$\mathrm{ad}([A,\, B]) = [\mathrm{ad}(A),\, \mathrm{ad}(B)] \quad \text{for all } A, B \in \mathfrak{g}$$

(where on the right, the Lie bracket is the commutator of linear maps on $\mathfrak{g}$).

Thus, we recover the Lie bracket from ad.

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).

We define the Lie bracket on $\mathfrak{g}$ as

$$[A,\, B] = \mathrm{ad}(A)(B).$$

To be complete, we still have to define the exponential map $\exp\colon \mathfrak{g} \to G$ for a general Lie group.

For this we need to introduce some left-invariant vector fields induced by the derivatives of the left translations and integral curves associated with such vector fields.

We conclude this section by computing explicitly the adjoint representations ad of $\mathfrak{so}(3)$ and Ad of $\mathbf{SO}(3)$.

Recall that for every $X \in \mathfrak{so}(3)$, ad_X is a linear map $\mathrm{ad}_X\colon \mathfrak{so}(3) \to \mathfrak{so}(3)$.

Also, for every $R \in \mathbf{SO}(3)$, the map $\mathrm{Ad}_R\colon \mathfrak{so}(3) \to \mathfrak{so}(3)$ is an invertible linear map of $\mathfrak{so}(3)$.

Now, as we saw earlier, $\mathfrak{so}(3)$ is isomorphic to $(\mathbb{R}^3, \times)$, where $\times$ is the cross-product on $\mathbb{R}^3$, *via* the isomorphism $\psi\colon (\mathbb{R}^3, \times) \to \mathfrak{so}(3)$ given by the formula

$$\psi(a, b, c) = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}.$$

In robotics and in computer vision, $\psi(u)$ is often denoted by $u_\times$.

The image of the canonical basis (e_1, e_2, e_3) of $\mathbb{R}^3$ is the following basis of $\mathfrak{so}(3)$:

$$\left(E_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \right.$$
$$\left. E_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right).$$

Observe that

$$[E_1, E_2] = E_3, \quad [E_2, E_3] = E_1, \quad [E_3, E_1] = E_2.$$

Using the isomorphism ψ, we obtain an isomorphism Ψ between $\mathrm{Hom}(\mathfrak{so}(3), \mathfrak{so}(3))$ and $\mathrm{M}_3(\mathbb{R}) = \mathfrak{gl}(3, \mathbb{R})$ given by

$$\Psi(f) = \psi^{-1} \circ f \circ \psi,$$

where $\Psi(f)$ is expressed in the basis (e_1, e_2, e_3).

By restricting Ψ to $\mathbf{GL}(\mathfrak{so}(3))$, we obtain an isomorphism between $\mathbf{GL}(\mathfrak{so}(3))$ and $\mathbf{GL}(3, \mathbb{R})$.

It turns out that if we use the basis (E_1, E_2, E_3) in $\mathfrak{so}(3)$, for every $X \in \mathfrak{so}(3)$, the matrix representing $\mathrm{ad}_X \in \mathrm{Hom}(\mathfrak{so}(3), \mathfrak{so}(3))$ is X itself, and for every $R \in \mathbf{SO}(3)$, the matrix representing $\mathrm{Ad}_R \in \mathbf{GL}(\mathfrak{so}(3))$ is R itself.

Proposition 4.14. *For all* $X \in \mathfrak{so}(3)$ *and all* $R \in \mathbf{SO}(3)$*, we have*

$$\Psi(\mathrm{ad}_X) = X, \qquad \Psi(\mathrm{Ad}_R) = R,$$

which means that $\Psi \circ \mathrm{ad}$ *is the inclusion map from* $\mathfrak{so}(3)$ *to* $\mathrm{M}_3(\mathbb{R}) = \mathfrak{gl}(3, \mathbb{R})$*, and that* $\Psi \circ \mathrm{Ad}$ *is the inclusion map from* $\mathbf{SO}(3)$ *to* $\mathbf{GL}(3, \mathbb{R})$*. Equivalently, for all* $u \in \mathbb{R}^3$*, we have*

$$\mathrm{ad}_X(\psi(u)) = \psi(Xu), \quad \mathrm{Ad}_R(\psi(u)) = \psi(Ru).$$

These equations can also be written as

$$[X, u_\times] = (Xu)_\times, \qquad R u_\times R^{-1} = (Ru)_\times.$$