13.7 Finale: Lie Groups and Lie Algebras

In this section we attempt to define precisely Lie groups and 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, and that the loss of information is not too severe.

The challenge that we are facing is that, unless our readers are already familiar with manifolds, the amount of basic differential geometry required to define Lie groups and Lie algebras in full generality is overwhelming.

Fortunately, all 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 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 a characterization of the linear groups that obviates much of the work.

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 (see Warner [**?**], Chapter 3, Theorem 3.21, page 97).

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 (see Warner [**?**], Chapter 3, Theorem 3.42, page 110).

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

Since our Lie groups are subgroups (or isomorphic to subgroups) of $\mathbf{GL}(n, \mathbb{R})$ for some suitable n, it is easy to define the Lie algebra of a Lie group using curves.

Although we will not prove the theorem of Von Neumann and Cartan, we feel that it is important to make clear why the definitions make sense, i.e., why we are not dealing with trivial objects.

A small annoying technical problem will arise 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 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 Hausdorff, it can be shown that every discrete subgroup of $\mathbf{GL}(n, \mathbb{R})$ is closed (which means that its complement is open).

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

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

Let us now review the definition of an embedded submanifold. For simplicity, we restrict our attention to smooth manifolds.

For detailed presentations, see DoCarmo [**?**, **?**], Milnor [**?**], Marsden and Ratiu [**?**], Berger and Gostiaux [**?**], or Warner [**?**].

For the sake of brevity, we use the terminology *manifold* (but other authors would say *embedded submanifolds*, 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 set 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 the definition of smooth functions on M and 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 $0_m = \underbrace{(0, \ldots, 0)}_{m}$).

Definition 13.7.1 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 $\varphi: \Omega \to \mathbb{R}^N$ 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 $\varphi: \Omega \to U$ is called a *(local) parametrization of M at p*. If $0_m \in \Omega$ and $\varphi(0_m) = p$, we say that $\varphi: \Omega \to U$ 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.

For instance, in the case of surfaces (2-dimensional manifolds), a chart is analogous to a planar map of a region on the surface.

For a concrete example, consider a map giving a planar representation of a country, a region on the earth, a curved surface.

Remark: We could allow $m = 0$ in definition 13.7.1.

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 13.1 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 13.1.

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.

If we draw a grid of coordinate lines parallel to the x and y axes inside the open unit disk and map these lines onto the lower hemisphere using φ_1, we get curved lines on the lower hemisphere.

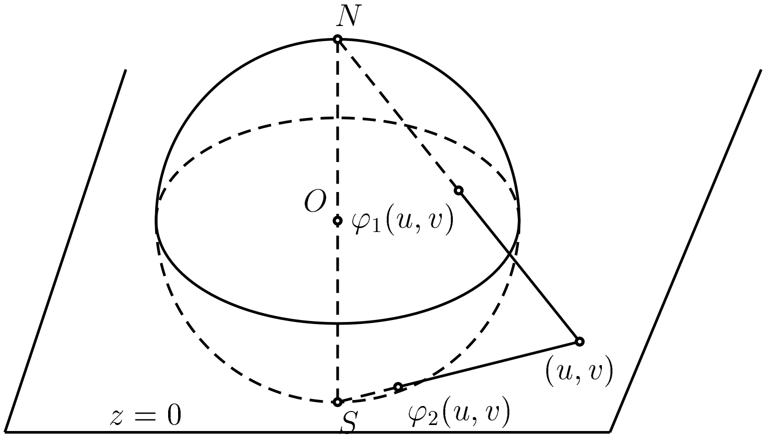

Figure 13.1: Inverse stereographic projections

These "coordinate lines" on the lower hemisphere provide local coordinates for every point on the lower hemisphere.

For this reason, older books often talk about *curvilinear coordinate systems* to mean the coordinate lines on a surface induced by a chart.

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.:

The first lemma is easily shown using the inverse function theorem.

Lemma 13.7.2 *Given an m-dimensional manifold M in $\mathbb{R}^N$, for every $p \in M$ there are two open sets $\Omega, W \subseteq \mathbb{R}^N$ with $0_N \in \Omega$ and $p \in M \cap W$, and a smooth diffeomorphism $\varphi\colon \Omega \to W$, such that $\varphi(0_N) = p$ and*

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

The next lemma is easily shown from Lemma 13.7.2. It is a key technical result used to show that interesting properties of maps between manifolds do not depend on parametrizations.

Lemma 13.7.3 *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 13.7.3 is illustrated in Figure 13.2.

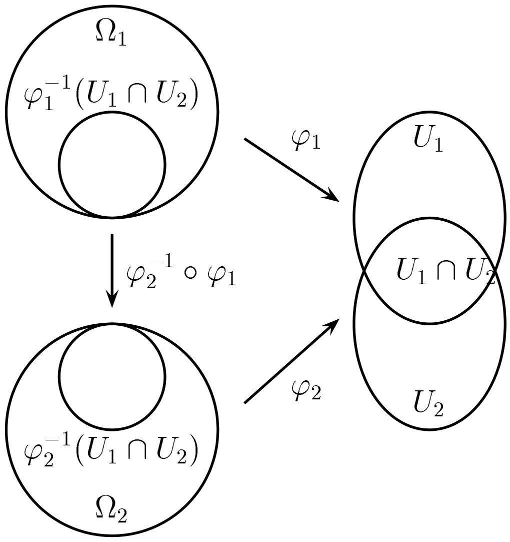

Figure 13.2: Parametrizations and transition functions

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

Definition 13.7.4 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.

Using Lemma 13.7.3, it is easily shown that Definition 13.7.4 does not depend on the choice of the parametrization $\varphi\colon \Omega \to U$ at p.

Lemma 13.7.3 also implies that γ viewed as a curve $\gamma\colon I \to \mathbb{R}^N$ is smooth.

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 $\mathbb{R}^m$.

The tangent vector at p to a curve γ on a manifold M is illustrated in Figure 13.3.

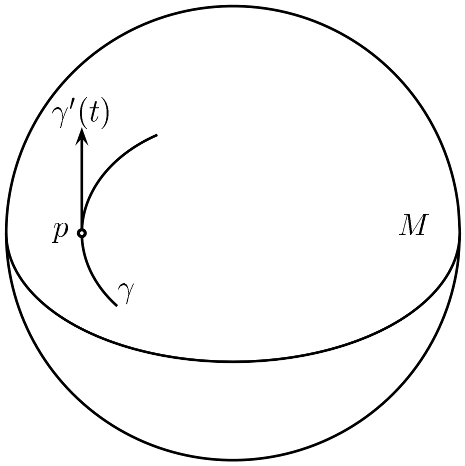

Figure 13.3: Tangent vector to a curve on a manifold

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

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

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 $\gamma(0) = p$).

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 13.7.5 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 $\gamma: I \to M$ 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 13.7.6 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 ${}^{-1}: G \to G$ are smooth.

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

This notion is explained in Definition 13.7.10, 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 13.7.7 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 $\mathbf{M}(n, \mathbb{R})$ 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 $\mathbf{M}(n, \mathbb{R})$, we have the following crucial theorem due to Von Neumann and Cartan:

Theorem 13.7.8 *A closed subgroup G of $\mathbf{GL}(n, \mathbb{R})$ is a linear Lie group. Furthermore, the set $\mathfrak{g}$ defined such that*

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

is a 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 $A, B \in \mathbf{M}(n, \mathbb{R})$.

Theorem 13.7.8 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 nonnull 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 nonnull rational numbers is a multiplicative subgroup of $\mathbb{R}^*$, but it is not closed, since $\mathbb{Q}$ is dense in $\mathbb{R}$.

The proof of Theorem 13.7.8 involves proving that when G is not a discrete subgroup, there is an open subset $\Omega \subseteq \mathbf{M}(n, \mathbb{R})$ such that $0_{n,n} \in \Omega$, an open subset $W \subseteq \mathbf{M}(n, \mathbb{R})$ such that $I \in W$, and that $\exp\colon \Omega \to W$ is a diffeomorphism such that

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

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

With the help of Theorem 13.7.8 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. 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 $\mathbf{SO}(n)$ and $\mathbb{R}^n$; see Chapter 2, Problem 2.19.

Returning to Theorem 13.7.8, the vector space $\mathfrak{g}$ is called the *Lie algebra* of the Lie group G.

Lie algebras are defined as follows.

Definition 13.7.9 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 13.7.8, 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 13.7.8 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 $\mathbb{R}^3$,

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

The exponential map $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ is given by Rodrigues's formula (see Lemma 13.2.3):

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

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 $\mathbf{M}(n, \mathbb{R})$, 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 13.4.

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.

Every such matrix is of the form

$$i(d\sigma_1 + c\sigma_2 + b\sigma_3) = \begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix},$$

where $b, c, d \in \mathbb{R}$, and $\sigma_1, \sigma_2, \sigma_3$ are the Pauli spin matrices (see Section 6.1), and thus the matrices $i\sigma_1, i\sigma_2, i\sigma_3$ form a basis of the Lie algebra $\mathfrak{su}(2)$.

The Lie bracket $[A, B]$ in $\mathfrak{su}(2)$ is given by the usual commutator, $[A, B] = AB - BA$.

It is easily checked that the vector space $\mathbb{R}^3$ is a Lie algebra if we define the Lie bracket on $\mathbb{R}^3$ as the usual cross product $u \times v$ of vectors.

Then we can define an isomorphism of Lie algebras $\varphi\colon (\mathbb{R}^3, \times) \to \mathfrak{su}(2)$ by the formula

$$\varphi(b, c, d) = \frac{i}{2}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{2}\begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix}.$$

It is indeed easy to verify that

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

Returning to $\mathfrak{su}(2)$, letting $\theta = \sqrt{b^2 + c^2 + d^2}$, we can write

$$d\sigma_1 + c\sigma_2 + b\sigma_3 = \begin{pmatrix} b & -ic + d \\ ic + d & -b \end{pmatrix} = \theta A,$$

where

$$A = \frac{1}{\theta}(d\sigma_1 + c\sigma_2 + b\sigma_3) = \frac{1}{\theta}\begin{pmatrix} b & -ic + d \\ ic + d & -b \end{pmatrix},$$

so that $A^2 = I$, and it can be shown that the exponential map $\exp\colon \mathfrak{su}(2) \to \mathbf{SU}(2)$ is given by

$$\exp(i\theta A) = \cos\theta\,\mathbf{1} + i\sin\theta\,A.$$

In view of the isomorphism $\varphi\colon (\mathbb{R}^3, \times) \to \mathfrak{su}(2)$, where

$$\varphi(b, c, d) = \frac{1}{2}\begin{pmatrix} ib & c + id \\ -c + id & -ib \end{pmatrix} = i\frac{\theta}{2}A,$$

the exponential map can be viewed as a map $\exp\colon (\mathbb{R}^3, \times) \to \mathbf{SU}(2)$ given by the formula

$$\exp(\theta v) = \left[\cos\frac{\theta}{2}, \sin\frac{\theta}{2}\, v\right],$$

for every vector θv, where v is a unit vector in $\mathbb{R}^3$ and $\theta \in \mathbb{R}$.

In this form, $\exp(\theta v)$ is a quaternion corresponding to a rotation of axis v and angle θ.

As we showed, $\mathbf{SE}(n)$ is a Lie group, and its lie algebra $\mathfrak{se}(n)$ described in Section 13.6 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.

First, we need to explain what is meant by a smooth map between manifolds.

Definition 13.7.10 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.

Using Lemma 13.7.3, it is easily shown that Definition 13.7.10 does not depend on the choice of the parametrizations $\varphi\colon \Omega_1 \to U_1$ and $\psi\colon \Omega_2 \to U_2$.

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

We now define the derivative of a smooth map between manifolds.

Definition 13.7.11 Let M_1 (m_1-dimensional) and M_2 (m_2-dimensional) be manifolds in $\mathbb{R}^N$. For any smooth function $f: M_1 \to M_2$ and any $p \in M_1$, the function $f'_p: T_pM_1 \to T_{f(p)}M_2$, 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 $\gamma: I \to M_1$ such that $\gamma(0) = p$ and $\gamma'(0) = v$,

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

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

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 13.7.3, 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 13.7.12 Given two Lie groups G_1 and G_2, a *homomorphism (or map) of Lie groups* is a function $f\colon G_1 \to G_2$ that is a homomorphism of groups and a smooth map (between the manifolds G_1 and G_2).

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

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

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

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

It is immediately verified that if $f: G_1 \to G_2$ is a homomorphism of Lie groups, then $f'_I: \mathfrak{g}_1 \to \mathfrak{g}_2$ 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.

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 $R_{a^{-1}} \circ L_a$ (also written as $R_{a^{-1}} L_a$) also play an important role.

Note that

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

The derivative

$$(R_{a^{-1}}L_a)'_I\colon \mathfrak{g} \to \mathfrak{g}$$

of $R_{a^{-1}}L_a$ at I is an isomorphism of Lie algebras, denoted by $\mathrm{Ad}_a\colon \mathfrak{g} \to \mathfrak{g}$.

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

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

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

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

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

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

The derivative

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

of Ad at I is map of Lie algebras, denoted by $\mathrm{ad}\colon \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$, called the *adjoint representation of* $\mathfrak{g}$ (where $\mathfrak{gl}(\mathfrak{g})$ denotes the Lie algebra of all linear maps on $\mathfrak{g}$).

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

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

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

One can also check 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:

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

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 would have to define the exponential map $\exp\colon \mathfrak{g} \to G$ for a general Lie group.

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

This is not hard, but we feel that it is now time to stop our introduction to Lie groups and Lie algebras, even though we have not even touched many important topics, for instance vector fields and differential foms.

Readers who wish to learn more about Lie groups and Lie algebras should consult (more or less listed in order of difficulty) Curtis [**?**], Sattinger and Weaver [**?**], and Marsden and Ratiu [**?**].

The excellent lecture notes by Carter, Segal, and Macdonald [**?**] constitute a very efficient (although somewhat terse) introduction to Lie algebras and Lie groups.

Classics such as Weyl [**?**] and Chevalley [**?**] are definitely worth consulting, although the presentation and the terminology may seem a bit old fashioned.

For more advanced texts, one may consult Abraham and Marsden [?], Warner [?], Sternberg [?], Bröcker and tom Dieck [?], and Knapp [?].

For those who read French, Mneimné and Testard [?] is very clear and quite thorough, and uses very little differential geometry, although it is more advanced than Curtis.

Chapter 1, by Bryant, in Freed and Uhlenbeck [?] is also worth reading, but the pace is fast, and Chapters 7 and 8 of Fulton and Harris [?] are very good, but familiarity with manifolds is assumed.

13.8 Applications of Lie Groups and Lie Algebras

Some applications of Lie groups and Lie algebras to robotics and motion planning are discussed in Selig [**?**] and Murray, Li, and Sastry [**?**].

Applications to physics are discussed in Sattinger and Weaver [**?**] and Marsden and Ratiu [**?**].

The fact that the exponential maps $\exp\colon \mathfrak{so}(3) \to \mathbf{SO}(3)$ and $\exp\colon \mathfrak{se}(3) \to \mathbf{SE}(3)$ are surjective is important in robotics applications.

Indeed, some matrices associated with joints arising in robot kinematics can be written as exponentials $e^{\theta \mathbf{s}}$, where θ is a joint angle and $\mathbf{s} \in \mathfrak{se}(3)$ is the so-called *joint screw* (see Selig [**?**], Chapter 4).

One should also observe that if a rigid motion (R, b) is used to define the position of a rigid body, then the velocity of a point p is given by $(R'p + b')$.

In other words, the element (R', b') of the Lie algebra $\mathfrak{se}(3)$ is a sort of velocity vector.

The surjectivity of the exponential map
$\exp\colon \mathfrak{se}(3) \to \mathbf{SE}(3)$ implies that there is a map
$\log\colon \mathbf{SE}(3) \to \mathfrak{se}(3)$, although it is multivalued.

Still, this log "function" can be used to perform motion interpolation.

For instance, given two rigid motions $B_1, B_2 \in \mathbf{SE}(3)$ specifying the position of a rigid body B, we can compute $\log(B_1)$ and $\log(B_2)$, which are just elements of the Euclidean space $\mathfrak{se}(3)$, form the linear interpolant $(1-t)\log(B_1) + t\log(B_2)$, and then apply the exponential map to get an interpolating rigid motion

$$e^{(1-t)\log(B_1) + t\log(B_2)}.$$

Of course, this can also be done for a sequence of rigid motions $B_1, \ldots, B_n$, where $n > 2$, and instead of using affine interpolation between two consecutive positions, a polynomial spline can be used to interpolate between the $\log(B_i)$'s in $\mathfrak{se}(3)$.

This approach has been investigated by Kim, M.-J., Kim, M.-S. and Shin [**?**, **?**], and Park and Ravani [**?**, **?**].

R.S. Ball published a treatise on the theory of screws in 1900 [**?**].

Basically, Ball's screws are rigid motions, and his instantaneous screws correspond to elements of the Lie algebra $\mathfrak{se}(3)$ (they are rays in $\mathfrak{se}(3)$). A *screw system* is simply a subspace of $\mathfrak{se}(3)$.

Such systems were first investigated by Ball [**?**].

The first heuristic classification of screw systems was given by Hunt [**?**]. Screw systems play an important role in kinematics, see McCarthy [**?**] and Selig [**?**], Chapter 8.

Lie groups and Lie algebras are also a key ingredient in the use of symmetries in motion, to reduce the number of parameters in the equations of motion, and in optimal control.

Such applications are described in a very exciting paper by Marsden and Ostrowski [**?**] (see also the references in this paper).