Appendix B

Lie groups and Lie algebras

Contents

B.1 Topological and Lie groups

Almost every interesting bundle comes with some extra structure attached to its fibers, for instance the vector space structure for a vector bundle, or an inner product (bundle metric) or orientation. One of the useful things about the concept of a bundle is that it allows us to treat all of these structures as special cases of the same thing: in principle one need only specify the *structure group* of the bundle. This notion is explained in Chapter 2, and plays an important role in the theory of connections in Chapter 3. Since one must first understand the basic concepts of Lie groups and their associated Lie algebras, we shall now give a quick overview of these ideas. Besides, Lie groups are interesting in their own right.

Recall first that a *group* is a set G together with an *operation* $G \times G \to G$, usually denoted as "multiplication" $(a, b) \mapsto ab$.[1] The operation is required to be associative, $a(bc) = (ab)c$, and there is a special element $e \in G$, called the *identity*, such that $eg = ge = g$ for all $g \in G$. Moreover, every $g \in G$ has an *inverse* $g^{-1} \in G$ such that $g^{-1}g = gg^{-1} = e$. We expect

[1]Sometimes the group operation is denoted alternatively as "addition" $(a, b) \mapsto a+b$. This is the custom particularly when the operation is *commutative*, i.e. $a + b = b + a$, in which case we say G is *abelian*.

this notion is already familiar to the reader. From a geometric perspective, it becomes more interesting if we assume that G is also a topological space or a manifold.

Definition B.1. A *topological group* G is a group with a topology such that the maps

$$G \times G \to G : (a, b) \mapsto ab \qquad \text{and} \qquad G \to G : a \mapsto a^{-1}$$

are both continuous.

Similarly, a *Lie group* is a group that is also a smooth manifold, such that the two maps above are both smooth. A *Lie subgroup* is a subgroup $H \subset G$ that is also a submanifold, i.e. the inclusion map $H \hookrightarrow G$ is both a group homomorphism and a smooth embedding.

The groups that are of greatest interest in geometry and topology typically consist of bijective maps $\varphi : M \to M$ on some space M, with the group multiplication law given by composition,

$$\varphi\psi := \varphi \circ \psi,$$

and inversion being the obvious operation $\varphi \mapsto \varphi^{-1}$. If M is a topological space, then there are various possible topologies one can define on spaces of continuous maps $M \to M$, depending sometimes on extra (e.g. smooth) structures; these details will not be integral to our discussion, so we will suppress them. It will at least be clear in all the examples of interest that the groups under consideration can be regarded as topological groups.

The simplest example is

$$\mathrm{Homeo}(M) = \{\varphi : M \to M \mid \varphi \text{ is bijective, } \varphi \text{ and } \varphi^{-1} \text{ both continuous}\},$$

where M is any topological space. If M is also an orientable topological manifold, then $\mathrm{Homeo}(M)$ has an important subgroup

$$\mathrm{Homeo}^+(M) = \{\varphi \in \mathrm{Homeo}(M) \mid \varphi \text{ preserves orientation}\};$$

we refer to [Hat02] for the definition of "orientation preserving" on a topological manifold. If M is a smooth manifold, there are the corresponding subgroups

$$\mathrm{Diff}(M) = \{\varphi \in \mathrm{Homeo}(M) \mid \varphi \text{ and } \varphi^{-1} \text{ are smooth}\},$$

and in the orientable case,

$$\mathrm{Diff}^+(M) = \{\varphi \in \mathrm{Diff}(M) \mid \varphi \text{ preserves orientation}\}.$$

There are still other groups corresponding to extra structures on a smooth manifold M, such as a Riemannian metric g or a symplectic form ω. These two in particular give rise to the groups of *isometries* and *symplectomorphisms* respectively,

$$\mathrm{Isom}(M, g) = \{\varphi \in \mathrm{Diff}(M) \mid \varphi^* g \equiv g\},$$
$$\mathrm{Symp}(M, \omega) = \{\varphi \in \mathrm{Diff}(M) \mid \varphi^* \omega \equiv \omega\}.$$

Most of these examples are infinite dimensional,[2] which introduces some complications from an analytical point of view. They cannot be considered smooth Lie groups, though they clearly are topological groups.

Actual Lie groups are usually obtained by considering linear transformations on finite dimensional vector spaces. As always, let $\mathbb{F}$ denote either $\mathbb{R}$ or $\mathbb{C}$. Groups which are smooth submanifolds of the vector space

$$\mathbb{F}^{n\times n} := \{\text{all } n\text{-by-}n \text{ matrices with entries in } \mathbb{F}\}$$

are called *linear groups* or *matrix groups*. Denote by $\mathbb{1}_n$ (or simply $\mathbb{1}$ when there's no ambiguity) the n-by-n identity matrix

$$\mathbb{1} = \mathbb{1}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix};$$

this will be the identity element of every linear group. Now observe that the map

$$\mathbb{F}^{n\times n} \times \mathbb{F}^{n\times n} \to \mathbb{F}^{n\times n} : (\mathbf{A}, \mathbf{B}) \mapsto \mathbf{AB}$$

is smooth; indeed, it's a bilinear map, and *every* bilinear map on a finite dimensional vector space is smooth (prove it!). A slightly less obvious but very important fact is that if we define the open subset

$$\mathrm{GL}(n, \mathbb{F}) = \{\mathbf{A} \in \mathbb{F}^{n\times n} \mid \det(\mathbf{A}) \neq 0\},$$

then the map $\mathrm{GL}(n, \mathbb{F}) \to \mathbb{F}^{n\times n} : \mathbf{A} \mapsto \mathbf{A}^{-1}$ is also smooth. This is easy to see in the case $n = 2$ since you can write down $\mathbf{A}^{-1}$ explicitly; in fact it follows more generally from *Cramer's rule*, an explicit formula for inverses of arbitrary invertible matrices, but that's not the clever way to do it. The clever way is to observe first that the power series expansion $\frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots$ can be generalized to matrices: we have

$$(\mathbb{1} + \mathbf{B})^{-1} = \mathbb{1} - \mathbf{B} + \mathbf{B}^2 - \mathbf{B}^3 + \ldots \tag{B.1}$$

for any matrix $\mathbf{B}$ sufficiently close to 0. Indeed, we see that $(\mathbb{1} + \mathbf{B})$ is necessarily invertible if $\mathbf{B}$ is sufficiently small, since $\det(\mathbb{1}) = 1$ and

[2] The exception is $\mathrm{Isom}(M, g)$, which is generated infinitessimally by the finite dimensional space of *Killing vector fields*; see e.g. [GHL04].

$\det(\mathbb{1} + \mathbf{B})$ is a continuous function of $\mathbf{B}$. Defining a norm $|\mathbf{B}|$ on matrices by identifying $\mathbb{F}^{n\times n}$ with $\mathbb{F}^{n^2}$, it's not hard to show that there is a constant $C > 0$ such that $|\mathbf{AB}| \le C|\mathbf{A}||\mathbf{B}|$, and in particular $|\mathbf{B}^k| \le C^{k-1}|\mathbf{B}|^k = \frac{1}{C}\left(C|\mathbf{B}|\right)^k$. Thus by a standard argument from analysis, the infinite sum above can be bounded by

$$\left|\sum_{k=0}^{\infty}(-1)^k\mathbf{B}^k\right| \le \sum_{k=0}^{\infty}|\mathbf{B}^k| \le \frac{1}{C}\sum_{k=0}^{\infty}\left(C|\mathbf{B}|\right)^k < \infty$$

whenever $\mathbf{B}$ is small enough so that $C|\mathbf{B}| < 1$. Now one sees by an easy computation that

$$(\mathbb{1} + \mathbf{B})\sum_{k=0}^{\infty}(-1)^k\mathbf{B}^k = \left(\sum_{k=0}^{\infty}(-1)^k\mathbf{B}^k\right)(\mathbb{1} + \mathbf{B}) = \mathbb{1}.$$

We apply this as follows to show that the map $\iota : \mathrm{GL}(n,\mathbb{F}) \to \mathbb{F}^{n\times n} : \mathbf{A} \mapsto \mathbf{A}^{-1}$ is smooth: for $\mathbb{H} \in \mathbb{F}^{n\times n}$ sufficiently small, we have

$$\begin{aligned}\iota(\mathbf{A} + \mathbf{H}) &= (\mathbf{A} + \mathbf{H})^{-1} = (\mathbb{1} + \mathbf{A}^{-1}\mathbf{H})^{-1}\mathbf{A}^{-1} \\ &= \left(\mathbb{1} - \mathbf{A}^{-1}\mathbf{H} + (\mathbf{A}^{-1}\mathbf{H})^2 - \ldots\right)\mathbf{A}^{-1} \\ &= \iota(\mathbf{A}) - \mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1} + \mathcal{O}(|\mathbf{H}|^2),\end{aligned}$$

proving that ι is differentiable and $d\iota(\mathbf{A})\mathbf{H} = -\mathbf{A}^{-1}\mathbf{H}\mathbf{A}^{-1}$. (Notice that this generalizes the formula $\frac{d}{dx}\frac{1}{x} = -\frac{1}{x^2}$.) Next observe that the map $\mathrm{GL}(n,\mathbb{F}) \to \mathrm{Hom}(\mathbb{F}^{n\times n}, \mathbb{F}^{n\times n}) : \mathbf{A} \mapsto d\iota(\mathbf{A})$ is a *quadratic* function of the differentiable map $\mathbf{A} \mapsto \mathbf{A}^{-1}$, and is thus also differentiable. Repeating this argument inductively, we conclude that ι has infinitely many derivatives.

This fact lays the groundwork for the following examples.

Example B.2. The set $\mathrm{GL}(n,\mathbb{F})$ defined above is precisely the set of all invertible n-by-n matrices: this is called the *general linear group*. As an open subset of the vector space $\mathbb{F}^{n\times n}$, it is trivially also a smooth manifold (of dimension n^2 if $\mathbb{F} = \mathbb{R}$, or $2n^2$ if $\mathbb{F} = \mathbb{C}$), and the remarks above show that multiplication and inversion are smooth, so that $\mathrm{GL}(n,\mathbb{F})$ is a Lie group. It's useful to think of $\mathrm{GL}(n,\mathbb{F})$ as the set of invertible $\mathbb{F}$-linear transformations $\mathbf{A} : \mathbb{F}^n \to \mathbb{F}^n$, with group multiplication defined by composition of transformations. In this way $\mathrm{GL}(n,\mathbb{F})$ is naturally a subgroup of $\mathrm{Diff}(\mathbb{F}^n)$. The next several examples will all be Lie subgroups of either $\mathrm{GL}(n,\mathbb{R})$ or $\mathrm{GL}(n,\mathbb{C})$.

In fact, one can also regard $\mathrm{GL}(n,\mathbb{C})$ as a Lie subgroup of $\mathrm{GL}(2n,\mathbb{R})$. Indeed, if we identify $\mathbb{C}^n$ with $\mathbb{R}^{2n}$ via the bijection

$$\mathbb{R}^{2n} \to \mathbb{C}^n : (x^1, \ldots, x^n, y^1, \ldots, y^n) \mapsto (x^1 + iy^1, \ldots, x^n + iy^n),$$

then multiplication of i by vectors in $\mathbb{C}^n$ becomes a real linear transformation $\mathbf{J}_0 : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ represented by the matrix

$$\mathbf{J}_0 = \begin{pmatrix} 0 & -\mathbb{1}_n \\ \mathbb{1}_n & 0 \end{pmatrix},$$

and multiplication of a complex scalar $a + ib$ by vectors in $\mathbb{R}^{2n}$ is now the linear transformation $a\mathbb{1} + b\mathbf{J}_0 \in \mathbb{R}^{2n\times 2n}$. Any $\mathbf{A} \in \mathbb{C}^{n\times n}$ defines a complex linear transformation $\mathbb{C}^n \to \mathbb{C}^n$, and therefore also a *real* linear transformation $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$, but with the extra property of preserving *complex* scalar multiplication: this means that for all $v \in \mathbb{R}^{2n}$ and $a + ib \in \mathbb{C}$,

$$\mathbf{A}(a\mathbb{1} + b\mathbf{J}_0)v = (a\mathbb{1} + b\mathbf{J}_0)\mathbf{A}v.$$

It's not hard to see that this is true if and only if $\mathbf{A}$ *commutes* with $\mathbf{J}_0$, i.e. $\mathbf{A}\mathbf{J}_0 = \mathbf{J}_0\mathbf{A}$. We therefore can give an alternative definition of $\mathbb{C}^{n\times n}$: in addition to being the complex n^2-dimensional vector space of n-by-n matrices, it is also the real $2n^2$-dimensional vector space of matrices $\mathbf{A} \in \mathbb{R}^{2n\times 2n}$ such that $\mathbf{A}\mathbf{J}_0 = \mathbf{J}_0\mathbf{A}$. As such it is a linear subspace of $\mathbb{R}^{2n\times 2n}$, and we can similarly define $\mathrm{GL}(n,\mathbb{C})$ as a Lie subgroup

$$\mathrm{GL}(n,\mathbb{C}) = \{\mathbf{A} \in \mathrm{GL}(2n,\mathbb{R}) \mid \mathbf{A}\mathbf{J}_0 = \mathbf{J}_0\mathbf{A}\},$$

also of (real) dimension $2n^2$. Take a brief moment to convince yourself that if $\mathbf{A}$ and $\mathbf{B}$ both commute with $\mathbf{J}_0$, then so do $\mathbf{AB}$ and $\mathbf{A}^{-1}$.

Example B.3. It's easy to see that topologically, $\mathrm{GL}(n,\mathbb{R})$ is not *connected*: one can't join every pair of matrices $\mathbf{A}, \mathbf{B} \in \mathrm{GL}(n,\mathbb{R})$ by a continuous path of matrices in $\mathrm{GL}(n,\mathbb{R})$. In particular there is no such path if $\det(\mathbf{A}) > 0$ and $\det(\mathbf{B}) < 0$. The sign of the determinant defines an important subgroup of $\mathrm{GL}(n,\mathbb{R})$:

$$\mathrm{GL}_+(n,\mathbb{R}) = \{\mathbf{A} \in \mathrm{GL}(n,\mathbb{R}) \mid \det(\mathbf{A}) > 0\}.$$

This is clearly an open subset of $\mathrm{GL}(n,\mathbb{R})$, and is thus also a Lie group of dimension n^2. To interpret $\mathrm{GL}_+(n,\mathbb{R})$ geometrically, recall that the set of bases of $\mathbb{R}^n$ can be divided into *positive* and *negative* bases: a basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is positive (regarded as an *ordered* set of linearly independent vectors) if and only if it can be deformed through a continuous family of bases to the *standard* basis $(\mathbf{e}_1, \ldots, \mathbf{e}_n)$; we say otherwise that $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is negative. An invertible linear map $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ is then called *orientation preserving* if it maps positive bases to positive bases, otherwise it is called *orientation reversing*. By definition the standard basis is positive, and since $\det(\mathbb{1}) = 1$, we deduce that the sign of any basis $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is the same as the sign of

$$\det \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix},$$

where the vectors $\mathbf{v}_j$ form the columns of an n-by-n matrix. Thus $\mathrm{GL}_+(n, \mathbb{R})$ is precisely the set of orientation preserving linear transformations $\mathbb{R}^n \to \mathbb{R}^n$.

Example B.4. The *special linear group* is defined by

$$\mathrm{SL}(n, \mathbb{F}) = \{\mathbf{A} \in \mathbb{F}^{n\times n} \mid \det(\mathbf{A}) = 1\}.$$

This is obviously a subgroup, and we claim that it is also a smooth submanifold of $\mathrm{GL}(n, \mathbb{F})$, with dimension $n^2 - 1$ in the real case, $2n^2 - 2$ in the complex case. This follows from the implicit function theorem, using the map

$$\det : \mathbb{F}^{n\times n} \to \mathbb{F}.$$

As a polynomial function of the matrix entries, det is clearly a smooth map; one then has to show that its derivative $d(\det)(\mathbf{A}) : \mathbb{F}^{n\times n} \to \mathbb{F}$ is a surjective linear map whenever $\det(\mathbf{A}) = 1$.

Exercise B.5.

(a) If $\mathbf{A}(t) \in \mathbb{F}^{n\times n}$ is a smooth path of matrices with $\mathbf{A}(0) = \mathbb{1}$ and its time derivative is denoted by $\dot{\mathbf{A}}(t)$, show that

$$\left.\frac{d}{dt}\det(\mathbf{A}(t))\right|_{t=0} = \mathrm{tr}(\dot{\mathbf{A}}(0)). \tag{B.2}$$

Hint: think of $\mathbf{A}(t)$ as an n-tuple of column vectors

$$\mathbf{A}(t) = \begin{pmatrix} \mathbf{v}_1(t) & \cdots & \mathbf{v}_n(t) \end{pmatrix}$$

with $\mathbf{v}_j(0) = \mathbf{e}_j$, the standard basis vector. Then $\det(\mathbf{A}(t))$ is the evaluation of an alternating n-form on these vectors, which can be written using components as in Appendix A. Write it this way and use the product rule. *Note:* you won't actually need to know what the components are!

(b) Show that if $\mathbf{A} \in \mathrm{GL}(n, \mathbb{F})$ then the derivative of $\det : \mathbb{F}^{n\times n} \to \mathbb{F}$ at $\mathbf{A}$ is

$$d(\det)(\mathbf{A})\mathbf{H} = \det(\mathbf{A}) \cdot \mathrm{tr}(\mathbf{A}^{-1}\mathbf{H}).$$

(c) Show that the aforementioned derivative is surjective, implying that $\det^{-1}(1) \subset \mathbb{F}^{n\times n}$ is a smooth submanifold of dimension $n^2 - 1$ if $\mathbb{F} = \mathbb{R}$, or $2n^2 - 2$ if $\mathbb{F} = \mathbb{C}$.

In the real case, the special linear group has a simple geometric interpretation. Recall that if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are column vectors then

$$\det\begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}$$

gives the signed volume of the parallelopiped in $\mathbb{R}^n$ spanned by these vectors; in particular this is nonzero if and only if the vectors $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ form a basis of $\mathbb{R}^n$, and it's positive if and only if this basis is positive. It follows that $\mathrm{SL}(n, \mathbb{R})$ is the set of all linear transformations $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ that preserve both volume and orientation. It is precisely half of the larger group

$$\{\mathbf{A} \in \mathbb{F}^{n\times n} \mid \det(\mathbf{A}) = \pm 1\},$$

which consists of *all* volume preserving linear transformations on $\mathbb{R}^n$. While simpler to interpret geometrically, the larger group is less interesting in applications—we encounter $\mathrm{SL}(n, \mathbb{R})$ much more often.

It is not so simple to interpret $\mathrm{SL}(n, \mathbb{C})$ geometrically, but algebraically it can be characterized as the set of all linear transformations on $\mathbb{C}^n$ that preserve a certain natural alternating n-form. Indeed, defining $\omega \in \Lambda^n(\mathbb{C}^n)^*$ by

$$\omega(\mathbf{v}_1, \dots, \mathbf{v}_n) = \det \begin{pmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{pmatrix}, \tag{B.3}$$

a linear transformation $\mathbf{A} : \mathbb{C}^n \to \mathbb{C}^n$ is in $\mathrm{SL}(n, \mathbb{C})$ if and only if $\mathbf{A}^*\omega = \omega$.

Example B.6. The *orthogonal group* is the subgroup

$$\mathrm{O}(n) = \{\mathbf{A} \in \mathrm{GL}(n, \mathbb{R}) \mid \mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbb{1}\},$$

where $\mathbf{A}^{\mathrm{T}}$ denotes the transpose of the matrix. Its elements are called *orthogonal matrices*. These have a geometric interpretation as the set of all linear transformations $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$ that preserve the Euclidean dot product $\mathbf{v} \cdot \mathbf{w}$. Indeed, we have

$$\mathbf{A}\mathbf{v} \cdot \mathbf{A}\mathbf{w} = (\mathbf{A}\mathbf{v})^{\mathrm{T}}(\mathbf{A}\mathbf{w}) = \mathbf{v}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{w} = \mathbf{v}^{\mathrm{T}}\mathbf{w} = \mathbf{v} \cdot \mathbf{w}$$

for all $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ if and only if $\mathbf{A} \in \mathrm{O}(n)$.

Exercise B.7. Show that $\mathbf{A} \in \mathrm{O}(n)$ if and only if its columns $\mathbf{v}_1, \dots, \mathbf{v}_n$ form an *orthonormal* basis of $\mathbb{R}^n$, i.e.

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

And that the same remark applies to the rows of $\mathbf{A}$.

To show that $\mathrm{O}(n)$ is a smooth submanifold of $\mathbb{R}^{n\times n}$, denote by $\Sigma(n) \subset \mathbb{R}^{n\times n}$ the space of all *symmetric* real n-by-n matrices, i.e. those which satisfy $\mathbf{A}^{\mathrm{T}} = \mathbf{A}$. This is a linear subspace of $\mathbb{R}^{n\times n}$, and there is a smooth map

$$F : \mathbb{R}^{n\times n} \to \Sigma(n) : \mathbf{A} \mapsto \mathbf{A}^{\mathrm{T}}\mathbf{A},$$

such that $F^{-1}(\mathbb{1}) = \mathrm{O}(n)$.

Exercise B.8. Defining the map $F : \mathbb{R}^{n\times n} \to \Sigma(n)$ as above, show that the derivative $dF(\mathbf{A}) : \mathbb{R}^{n\times n} \to \Sigma(n)$ is surjective whenever $\mathbf{A}$ is invertible. Conclude that $\mathrm{O}(n) \subset \mathbb{R}^{n\times n}$ is a smooth manifold of dimension $\frac{n(n-1)}{2}$.

Observe that every matrix $\mathbf{A} \in \mathrm{O}(n)$ has $\det(\mathbf{A}) = \pm 1$. Taking only those with positive determinant defines the subgroup

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

called the *special orthogonal group*. Since $\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R})$, the linear transformations $\mathbf{A} \in \mathrm{SO}(n)$ can be characterized as those which preserve both orientation and the Euclidean dot product. These are the *rotations* of $\mathbb{R}^n$.

Exercise B.9. Show that every $\mathbf{A} \in \mathrm{O}(n)$ can be decomposed as $\mathbf{A} = \mathbf{RT}$ where $\mathbf{R} \in \mathrm{SO}(n)$ and $\mathbf{T}$ is the orientation reversing transformation

$$\mathbf{T} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}. \tag{B.4}$$

Proposition B.10. *The group* $\mathrm{SO}(n)$ *is connected for all* n*, i.e. for any pair of matrices* $\mathbf{A}_0, \mathbf{A}_1 \in \mathrm{SO}(n)$*, there is a continuous path* $\mathbf{A}(t) \in \mathrm{SO}(n)$ *for* $t \in [0, 1]$ *with* $\mathbf{A}(0) = \mathbf{A}_0$ *and* $\mathbf{A}(1) = \mathbf{A}_1$.

Proof. For each $n \geq 1$, there is an injective homomorphism $\mathrm{SO}(n) \hookrightarrow \mathrm{SO}(n+1)$ defined by

$$\mathbf{A} \mapsto \begin{pmatrix} 1 & \\ & \mathbf{A} \end{pmatrix}, \tag{B.5}$$

thus we can regard $\mathrm{SO}(n)$ as a subgroup of $\mathrm{SO}(n+1)$. We claim that for every $\mathbf{A}_0 \in \mathrm{SO}(n+1)$, there is a continuous path $\mathbf{A}(t) \in \mathrm{SO}(n+1)$ with $\mathbf{A}(0) = \mathbf{A}_0$ and $\mathbf{A}(1) \in \mathrm{SO}(n)$. The idea is to construct $\mathbf{A}(t)$ via column vectors

$$\mathbf{A}(t) = \begin{pmatrix} \mathbf{v}_1(t) & \cdots & \mathbf{v}_{n+1}(t) \end{pmatrix}$$

which form a continuous family of orthonormal bases. It's easy to see that such a family of bases (not necessarily orthonormal) can be constructed so that $\mathbf{v}_1(1)$ is the standard basis vector $\mathbf{e}_1$. One can then use the Gram-Schmidt orthgonalization process (cf. [Str80]) to turn this into an orthonormal basis for all t, without changing either $\mathbf{A}(0)$ or $\mathbf{v}_1(1)$. Now by construction, $\mathbf{A}(t)$ is orthogonal for all t, and since $\det(\mathbf{A}(0)) = 1$ and the determinant depends continuously on the matrix, $\mathbf{A}(t) \in \mathrm{SO}(n+1)$. Moreover the upper left entry of $\mathbf{A}(1)$ is 1, and since the rows of $\mathbf{A}(1)$ are all orthogonal unit vectors (by Exercise B.7), the top row of $\mathbf{A}(1)$ is also

$\mathbf{e}_1$, implying that $\mathbf{A}(1)$ has the same block form as in (B.5). One verifies easily that the lower right block must then be an orthogonal matrix with determinant 1, so $\mathbf{A}(1) \in \mathrm{SO}(n)$.

The result now follows by induction, for in the $n = 1$ case, the argument above connects $\mathbf{A}(0)$ with $\mathbf{A}(1) \in \mathrm{SO}(1)$, and there is only one such matrix: the identity. This shows that there is a continuous path from any $\mathbf{A} \in \mathrm{SO}(n)$ to $\mathbb{1}$. □

Corollary B.11. $\mathrm{GL}_+(\mathbb{R}, n)$ *is also connected, and* $\mathrm{GL}(\mathbb{R}, n)$ *and* $\mathrm{O}(n)$ *each have exactly two connected components. In particular, there is a continuous path in* $\mathrm{GL}(\mathbb{R}, n)$ *or* $\mathrm{O}(n)$ *from any* $\mathbf{A}$ *to either* $\mathbb{1}$ *or the orientation reversing matrix* (B.4).[3]

Proof. The statement about connected components of $\mathrm{O}(n)$ follows immediately from Prop. B.10 and Exercise B.9. To apply this result to $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{GL}_+(n, \mathbb{R})$, we use the *polar decomposition*: recall that every $\mathbf{A} \in \mathrm{GL}(\mathbb{R}, n)$ can be written as

$$\mathbf{A} = \mathbf{P}\mathbf{Q}$$

where $\mathbf{P}$ is a positive definite symmetric matrix and $\mathbf{Q} \in \mathrm{O}(n)$; moreover $\mathbf{Q} \in \mathrm{SO}(n)$ if $\mathbf{A} \in \mathrm{GL}_+(\mathbb{R}, n)$. Indeed, since $\mathbf{A}$ is invertible, $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is a symmetric positive definite matrix, and the spectral theorem (see [Str80]) therefore guarantees that one can write

$$\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^{\mathrm{T}}$$

where $\mathbf{U} \in \mathrm{O}(n)$ and $\mathbf{D}$ is a diagonal matrix with only positive entries $\lambda_1, \ldots, \lambda_n$ along the diagonal. We can then define a "square root" $\sqrt{\mathbf{D}}$ as the diagonal matrix with entries $\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}$, and define

$$\mathbf{P} := \sqrt{\mathbf{A}^{\mathrm{T}}\mathbf{A}} := \mathbf{U}\sqrt{\mathbf{D}}\mathbf{U}^{\mathrm{T}},$$

a positive definite symmetric matrix whose square is $\mathbf{A}^{\mathrm{T}}\mathbf{A}$. It is now straightforward to verify that

$$\mathbf{Q} := \left(\sqrt{\mathbf{A}^{\mathrm{T}}\mathbf{A}}\right)^{-1}\mathbf{A}$$

is orthogonal, and by construction $\mathbf{A} = \mathbf{P}\mathbf{Q}$.

Observe next that one can easily construct a continuous path of positive definite symmetric matrices from $\mathbb{1}$ to $\mathbf{P}$: let

$$\mathbf{P}(t) = \mathbf{U}\begin{pmatrix} \lambda_1^{t/2} & & \\ & \ddots & \\ & & \lambda_n^{t/2} \end{pmatrix}\mathbf{U}^{\mathrm{T}}.$$

[3] An equivalent statement is that every basis of $\mathbb{R}^n$ can be deformed through a family of bases to one of two standard forms. This is why it makes sense to divide the set of all bases into two classes, positive and negative.

Then $\mathbf{A}(t) := \mathbf{P}(t)\mathbf{Q}$ is a continuous path from $\mathrm{O}(n)$ to $\mathbf{A}$. The statements about $\mathrm{GL}(n, \mathbb{R})$ and $\mathrm{GL}_+(n, \mathbb{R})$ therefore follow from the corresponding statements about $\mathrm{O}(n)$ and $\mathrm{SO}(n)$. $\square$

Example B.12. The complex analogue of $\mathrm{O}(n)$ is the *unitary* group

$$\mathrm{U}(n) = \{\mathbf{A} \in \mathbb{C}^{n \times n} \mid \mathbf{A}^\dagger \mathbf{A} = \mathbb{1}\},$$

where by definition $\mathbf{A}^\dagger$ is the complex conjugate of $\mathbf{A}^{\mathrm{T}}$. Its elements are called *unitary matrices*, and analogously to orthogonal matrices, their sets of columns and rows form orthonormal bases with respect to the standard Hermitian inner product

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^{n} \bar{\mathbf{v}}^j \mathbf{w}^j.$$

on $\mathbb{C}^n$. In fact, $\mathrm{U}(n)$ is precisely the set of linear transformations $\mathbf{A} : \mathbb{C}^n \to \mathbb{C}^n$ which preserve this inner product.

As with $\mathrm{O}(n)$, we can write $\mathrm{U}(n) = F^{-1}(\mathbb{1})$ for a smooth map $F : \mathbb{C}^{n \times n} \to H(n) : \mathbf{A} \mapsto \mathbf{A}^\dagger \mathbf{A}$, where $H(n) \subset \mathbb{C}^{n \times n}$ denotes the vector space of *Hermitian* matrices, satisfying $\mathbf{A}^\dagger = \mathbf{A}$. Then applying the implicit function theorem the same way as in Exercise B.8 shows that $\mathrm{U}(n)$ is a smooth submanifold of $\mathbb{C}^{n \times n}$ with dimension n^2.

Exercise B.13. Adapt the arguments of Prop. B.10 and Corollary B.11 to show that $\mathrm{U}(n)$ and $\mathrm{GL}(n, \mathbb{C})$ are both connected for all n. *Hint:* $\mathrm{U}(1)$ is diffeomorphic to S^1; in particular, it's connected (unlike $\mathrm{O}(1)$).[4]

Taking the determinant of $\mathbf{A}^\dagger \mathbf{A} = \mathbb{1}$ implies $|\det(\mathbf{A})| = 1$, but this allows many more possibilities than in the real case: $\det(\mathbf{A})$ can now lie anywhere on the unit circle in $\mathbb{C}$. The *special unitary group*

$$\mathrm{SU}(n) = \{\mathbf{A} \in \mathrm{U}(n) \mid \det(\mathbf{A}) = 1\}$$

is therefore a submanifold of dimension *less* than that of $\mathrm{U}(n)$. Indeed, writing the unit circle in $\mathbb{C}$ as S^1 and applying the implicit function theorem to the map

$$\det : \mathrm{U}(n) \to S^1,$$

one can show that $\mathrm{SU}(n)$ is a smooth manifold of dimension $n^2 - 1$. (Try it!) It consists of the linear transformations on $\mathbb{C}^n$ that preserve both the standard Hermitian inner product and the alternating n-form defined by (B.3).

[4]This is why the concept of *orientation* does not exist for complex bases.

Finally, we give two examples of Lie groups that are not matrix groups in any natural way (they do admit *representations* as matrix groups, but these are not canonical).

Example B.14. $\mathbb{F}^n$ itself is obviously an abelian Lie group, with vector addition as the group operation. We can regard it as a space of bijective maps as follows: for each $\mathbf{v} \in \mathbb{F}^n$, define the diffeomorphism

$$\varphi_{\mathbf{v}} : \mathbb{F}^n \to \mathbb{F}^n : \mathbf{w} \mapsto \mathbf{w} + \mathbf{v}.$$

Then clearly $\varphi_{\mathbf{v}+\mathbf{w}} = \varphi_{\mathbf{v}} \circ \varphi_{\mathbf{w}}$ and $\varphi_{-\mathbf{v}} = \varphi_{\mathbf{v}}^{-1}$. In this way $\mathbb{F}^n$ becomes a finite dimensional subgroup of $\mathrm{Diff}(\mathbb{F}^n)$.

Example B.15. The vector space structure of $\mathbb{F}^n$ defines a special group of diffeomorphisms of which $\mathrm{GL}(n, \mathbb{F})$ is a subgroup. Namely, a map $\varphi : \mathbb{F}^n \to \mathbb{F}^n$ is called *affine* if it has the form

$$\varphi(\mathbf{v}) = \mathbf{A}\mathbf{v} + \mathbf{b}$$

for any $\mathbf{A} \in \mathrm{GL}(n, \mathbb{F})$ and $\mathbf{b} \in \mathbb{F}^n$. These form the group of *affine transformations* on $\mathbb{F}^n$. It has both $\mathrm{GL}(n, \mathbb{F})$ and $\mathbb{F}^n$ as subgroups, and in group theoretic terms, is called the *semidirect product* of these. The same trick can be used to produce many Lie subgroups: for instance, the semidirect product of $\mathrm{O}(n)$ with $\mathbb{R}^n$ is the *Euclidean group*

$$\{\varphi \in \mathrm{Diff}(\mathbb{R}^n) \mid \varphi(\mathbf{v}) = \mathbf{A}\mathbf{v} + \mathbf{b} \text{ for some } \mathbf{A} \in \mathrm{O}(n) \text{ and } \mathbf{b} \in \mathbb{R}^n\}.$$

This is in fact $\mathrm{Isom}(\mathbb{R}^n, \langle\ ,\ \rangle)$, the group of *rigid motions* on $\mathbb{R}^n$, i.e. those diffeomorphisms of $\mathbb{R}^n$ which preserve the Riemannian metric defined by the standard Euclidean inner product. An important cousin of this arises in Einstein's theory of Special Relativity: the *Poincaré* group is the set of diffeomorphisms on $\mathbb{R}^4$ that preserve the *Minkowski metric*, a generalized version of an inner product which is not positive definite. See [Car] for details.

In the next sections we will examine the relationship between the global algebraic structure of a Lie group and the algebraic structure that this induces locally on the tangent space to the identity. For this it will be helpful to know that for topological groups in general, the group structure is in some sense determined by a neighborhood of the identity. Recall that a subset of any topological space is called a *connected component* if it is both open and closed, and the space itself is called *connected* if it contains no connected components other than itself and the empty set.

Proposition B.16. *If G is a connected topological group and $\mathcal{U}$ is a neighborhood of the identity $e \in G$, then $\mathcal{U}$ generates G, i.e. every $g \in G$ can be written as a product of elements in $\mathcal{U}$.*

Proof. We shall give an outline of the proof and leave several details to the reader.[5] One begins by showing that if $H \subset G$ is a subgroup that is also an open set, then the subset $gH = \{gh \mid g \in H\}$ is also open for each g. One can then show that H is also a closed set, so in particular if G is connected, $H = G$. These are simple exercises in point set topology.

Now assuming G is connected and $\mathcal{U}$ is an open neighborhood of e, for each $n \in \mathbb{N}$ let

$$\mathcal{U}^n = \{a_1 \dots a_n \in G \mid a_1, \dots, a_n \in \mathcal{U}\}.$$

It is straightforward to show that $\mathcal{U}^{n+1}$ is a neighborhood of $\mathcal{U}^n$, i.e. any sequence $g_j \in G$ approaching a point in $\mathcal{U}^n$ satisfies $g_j \in \mathcal{U}^{n+1}$ for j sufficiently large. Then the union

$$\bigcup_{n=1}^{\infty} \mathcal{U}^n$$

is both a subgroup of G and an open subset, and is thus equal to G. $\square$

B.2 Lie algebras in general

Vector spaces are important in differential geometry mainly because they arise as "linearizations" (e.g. tangent spaces) of smooth manifolds. From this point of view, one should expect that the extra (algebraic) structure of a Lie group should contribute some extra algebraic structure to the corresponding vector spaces. This is known as a *Lie algebra* structure.

In general, a *Lie algebra* is any vector space V equipped with a *Lie bracket* $[\ ,\ ] : V \oplus V \to V$, which is a bilinear map satisfying the following two conditions:

1. Antisymmetry: $[u, v] = -[v, u]$
2. The Jacobi identity: $[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0$

Example B.17 (**The Lie algebra of vector fields**)**.** On a smooth manifold M, the space $\mathrm{Vec}(M)$ is a Lie algebra with respect to the usual bracket on vector fields.

Example B.18 (**The Lie algebra of matrices**)**.** The space $\mathbb{F}^{m\times m}$ of m-by-m matrices is a Lie algebra with respect to the *commutator* $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$.

As mentioned above, Lie brackets arise naturally as linearizations of group operations. A case that should already be familiar is the Lie algebra

[5] This outline is borrowed from some exercises in Chapter 10 of [Spi99].

of vector fields on a manifold M, where the bracket $[X, Y]$ in some sense measures the failure of the flows generated by X and Y to commute. Put another way, the operation

$$[\ ,\ ] : \mathrm{Vec}(M) \oplus \mathrm{Vec}(M) \to \mathrm{Vec}(M)$$

measures the extent to which the group $\mathrm{Diff}(M)$ is not abelian, at least in an infinitessimal neighborhood of the identity. This correspondence between the group $\mathrm{Diff}(M)$ and the Lie algebra $\mathrm{Vec}(M)$ is one example of a very general phenomenon which we will now explore further.

B.3 Lie algebras of matrix groups

To motivate the more general notions, we examine first a special case which is also in many ways the most important: assume G is a Lie subgroup of $\mathrm{GL}(n, \mathbb{F})$, e.g. one of the groups $\mathrm{O}(n)$, $\mathrm{SU}(n)$, or for that matter, $\mathrm{GL}(n, \mathbb{F})$ itself. Most such groups are not abelian, and the failure of commutativity can be measured by the failure of the equation

$$\mathbf{A}\mathbf{B}\mathbf{A}^{-1}\mathbf{B}^{-1} = \mathbb{1}. \tag{B.6}$$

Informally, consider two matrices $\mathbb{1} + \mathbf{A}$ and $\mathbb{1} + \mathbf{B}$ which we assume are infinitesimally close to the identity. Since $\mathbf{A}$ and $\mathbf{B}$ are small, we can express inverses via the expansion

$$(\mathbb{1} + \mathbf{A})^{-1} = \mathbb{1} - \mathbf{A} + \mathbf{A}^2 - \mathbf{A}^3 + \ldots$$

as in (B.1), and similarly for $\mathbf{B}$. Considering now the left side of (B.6) with $\mathbf{A}$ and $\mathbf{B}$ replaced by $\mathbb{1} + \mathbf{A}$ and $\mathbb{1} + \mathbf{B}$, we have

$$\begin{aligned}(\mathbb{1} + \mathbf{A})&(\mathbb{1} + \mathbf{B})(\mathbb{1} + \mathbf{A})^{-1}(\mathbb{1} + \mathbf{B})^{-1} \\ &= (\mathbb{1} + \mathbf{A})(\mathbb{1} + \mathbf{B})(\mathbb{1} - \mathbf{A} + \mathbf{A}^2 - \ldots)(\mathbb{1} - \mathbf{B} + \mathbf{B}^2 - \ldots) \\ &= \mathbb{1} + \mathbf{AB} - \mathbf{BA} + \ldots\end{aligned}$$

where the dots represent all terms that are of third degree or higher. We see that the commutator $\mathbf{AB} - \mathbf{BA}$ measures the failure of Equation (B.6) in the lowest order approximation.

To make this discussion precise, we replace "matrices infinitessimally close to $\mathbb{1}$" with smooth paths through $\mathbb{1}$, characterized by their tangent vectors in $T_{\mathbb{1}}G$. Recall that $\mathrm{GL}(n, \mathbb{F})$ is an open subset of the vector space $\mathbb{F}^{n \times n}$. Since G is a subgroup of $\mathrm{GL}(n, \mathbb{F})$ by assumption, the tangent space $T_{\mathbb{1}}G$ can be identified in a natural way with a (real) linear subspace of $\mathbb{F}^{n \times n}$. We denote this subspace by $\mathfrak{g}$ and call it the *Lie algebra of G*, for reasons that will shortly become apparent.

Consider two smooth paths $s \mapsto \mathbf{A}_s \in G$ and $t \mapsto \mathbf{B}_t \in G$ with $\mathbf{A}_0 = \mathbf{B}_0 = \mathbb{1}$, and denote their velocity vectors by

$$\dot{\mathbf{A}}_s = \frac{d}{ds}\mathbf{A}_s \in T_{\mathbf{A}_s}G \subset \mathbb{F}^{n\times n} \quad \text{and} \quad \dot{\mathbf{B}}_t = \frac{d}{dt}\mathbf{B}_t \in T_{\mathbf{B}_t}G \subset \mathbb{F}^{n\times n};$$

in particular, $\dot{\mathbf{A}}_0$ and $\dot{\mathbf{B}}_0$ are both elements of $\mathfrak{g}$. Now define a function

$$f(s,t) = \mathbf{A}_s\mathbf{B}_t\mathbf{A}_s^{-1}\mathbf{B}_t^{-1},$$

which maps a neighborhood of $(0,0)$ in $\mathbb{R}^2$ smoothly into a neighborhood of $\mathbb{1}$ in $G \subset \mathbb{F}^{n\times n}$. Obviously $f(0,0) = \mathbb{1}$, and we can measure the deviation of $f(s,t)$ from $\mathbb{1}$ via its partial derivatives at $(0,0)$. The following simple formula will come in useful.

Exercise B.19. Show that $\frac{d}{ds}\mathbf{A}_s^{-1} = -\mathbf{A}_s^{-1}\dot{\mathbf{A}}_s\mathbf{A}_s^{-1}$. *Hint:* the simplest approach is to apply the product rule to $\mathbf{A}_s^{-1}\mathbf{A}_s$, though this also follows from our discussion of the smoothness of $\iota(\mathbf{A}) = \mathbf{A}^{-1}$ in §B.1.

As a special case, $\frac{d}{ds}\mathbf{A}_s^{-1}\big|_{s=0} = -\dot{\mathbf{A}}_0$. Now we compute,

$$\begin{aligned}
\frac{\partial f}{\partial s} &= \dot{\mathbf{A}}_s\mathbf{B}_t\mathbf{A}_s^{-1}\mathbf{B}_t^{-1} - \mathbf{A}_s\mathbf{B}_t\mathbf{A}_s^{-1}\dot{\mathbf{A}}_s\mathbf{A}_s^{-1}\mathbf{B}_t^{-1},\\
\frac{\partial f}{\partial t} &= \mathbf{A}_s\dot{\mathbf{B}}_t\mathbf{A}_s^{-1}\mathbf{B}_t^{-1} - \mathbf{A}_s\mathbf{B}_t\mathbf{A}_s^{-1}\mathbf{B}_t^{-1}\dot{\mathbf{B}}_t\mathbf{B}_t^{-1}.
\end{aligned}$$

Notice that both derivatives vanish at $(s,t) = (0,0)$ since $\mathbf{A}_0 = \mathbf{B}_0 = \mathbb{1}$. We use these to compute the second partial derivatives at $(s,t) = (0,0)$,

$$\begin{aligned}
\left.\frac{\partial^2 f}{\partial s^2}\right|_{(s,t)=(0,0)} &= \ddot{\mathbf{A}}_0 - \dot{\mathbf{A}}_0\dot{\mathbf{A}}_0 - \dot{\mathbf{A}}_0\dot{\mathbf{A}}_0 - (-\dot{\mathbf{A}}_0)\dot{\mathbf{A}}_0 - \ddot{\mathbf{A}}_0 - \dot{\mathbf{A}}_0(-\dot{\mathbf{A}}_0) = 0,\\
\left.\frac{\partial^2 f}{\partial t^2}\right|_{(s,t)=(0,0)} &= \ddot{\mathbf{B}}_0 - \dot{\mathbf{B}}_0\dot{\mathbf{B}}_0 - \dot{\mathbf{B}}_0\dot{\mathbf{B}}_0 - (-\dot{\mathbf{B}}_0)\dot{\mathbf{B}}_0 - \ddot{\mathbf{B}}_0 - \dot{\mathbf{B}}_0(-\dot{\mathbf{B}}_0) = 0,\\
\left.\frac{\partial^2 f}{\partial s\partial t}\right|_{(s,t)=(0,0)} &= \dot{\mathbf{A}}_0\dot{\mathbf{B}}_0 - \dot{\mathbf{A}}_0\dot{\mathbf{B}}_0 - \dot{\mathbf{B}}_0\dot{\mathbf{A}}_0 - \dot{\mathbf{A}}_0(-\dot{\mathbf{B}}_0) = \dot{\mathbf{A}}_0\dot{\mathbf{B}}_0 - \dot{\mathbf{B}}_0\dot{\mathbf{A}}_0.
\end{aligned}$$

Thus the lowest order nonvanishing derivative of f at $(0,0)$ is the *commutator*, $[\dot{\mathbf{A}}_0, \dot{\mathbf{B}}_0] = \dot{\mathbf{A}}_0\dot{\mathbf{B}}_0 - \dot{\mathbf{B}}_0\dot{\mathbf{A}}_0$. This leads to a proof of the following rather non-obvious fact:

Proposition B.20. *For any* $\mathbf{A}, \mathbf{B} \in \mathfrak{g}$, *the commutator* $[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A}$ *is also in* $\mathfrak{g}$.

Proof. Referring to the construction above, it suffices to show that the mixed partial derivative $\partial_t\partial_s f(0,0) \in \mathbb{F}^{n\times n}$ lies in the subspace $\mathfrak{g} = T_{\mathbb{1}}G$. We can see this most easily by viewing it as the linearization of a vector field along a path (see Chapter 3). Namely, let $\gamma(t) = f(0,t)$ and define a vector field on G along γ by

$$V(t) = \partial_s f(0,t).$$

Since $V(0) = 0$, the expression $\nabla_t V(0) \in T_{\mathbb{1}}G$ is defined independently of any choice of connection on G; in particular, we can choose ∇ so that parallel transport is defined via the natural embedding of G in $\mathbb{F}^{n\times n}$, in which case $\nabla_t V(0) = \partial_t V(0) = \partial_t\partial_s f(0,0)$. □

In light of this result, $\mathfrak{g}$ inherits a Lie algebra structure from the commutator bracket on $\mathbb{F}^n$, and we say that $\mathfrak{g}$ is a *Lie subalgebra* of $\mathbb{F}^n$.

Definition B.21. A Lie algebra is called *abelian* if the bracket operation is trivial, i.e. $[v,w] = 0$ for all v and w.

It should now be clear where this terminology comes from: if G is an abelian subgroup of $\mathrm{GL}(n,\mathbb{F})$, then the map $(s,t) \mapsto \mathbf{A}_s\mathbf{B}_t\mathbf{A}_s^{-1}\mathbf{B}_t^{-1}$ is constant, so the above discussion implies that $[\mathbf{A},\mathbf{B}]$ must vanish for all $\mathbf{A},\mathbf{B} \in \mathfrak{g}$. There is a converse to this statement, though it's harder to prove; see Prop. B.30 below.

We've already seen the first example of the Lie algebra for a linear group: if $G = \mathrm{GL}(n,\mathbb{F})$ itself, the tangent space $T_{\mathbb{1}}\mathrm{GL}(n,\mathbb{F})$ is identified with $\mathbb{F}^{n\times n}$. There is a standard convention of denoting such algebras by the same letters as the corresponding groups, but in lowercase fraktur script, thus

$$\mathfrak{gl}(n,\mathbb{F}) = \mathbb{F}^{n\times n}.$$

Example B.22 (The Lie algebra of $\mathrm{O}(n)$). If $\mathbf{A}_t$ is a smooth path of orthogonal matrices with $\mathbf{A}_0 = \mathbb{1}$, then differentiating the relation $\mathbf{A}_t^{\mathrm{T}}\mathbf{A}_t \equiv \mathbb{1}$ at time 0 gives

$$\dot{\mathbf{A}}_0^{\mathrm{T}} + \dot{\mathbf{A}}_0 = 0.$$

Thus $\mathfrak{o}(n)$ is contained in the vector space of antisymmetric matrices in $\mathbb{R}^{n\times n}$; in fact, since the dimension of this space is $(n-1)n/2$ and thus matches $\dim \mathrm{O}(n)$, we see that $\mathfrak{o}(n)$ is precisely the space of antisymmetric matrices. This is indeed a Lie subalgebra of $\mathfrak{gl}(n,\mathbb{R})$, since for any two antisymmetric matrices $\mathbf{A}$ and $\mathbf{B}$, $[\mathbf{A},\mathbf{B}]^{\mathrm{T}} = (\mathbf{AB}-\mathbf{BA})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} - \mathbf{A}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}} = (-\mathbf{B})(-\mathbf{A}) - (-\mathbf{A})(-\mathbf{B}) = -(\mathbf{AB}-\mathbf{BA}) = -[\mathbf{A},\mathbf{B}]$.

Exercise B.23. Confirm the identity of each of the following Lie algebras and, just for reassurance, verify that each is closed under the commutator bracket. For some of these it will help to recall the formula of Exercise B.5, part (a).

(a) $\mathfrak{sl}(n,\mathbb{F}) = \{\mathbf{A} \in \mathfrak{gl}(n,\mathbb{F}) \mid \operatorname{tr}\mathbf{A} = 0\}$

(b) $\mathfrak{so}(n) = \mathfrak{o}(n) = \{\mathbf{A} \in \mathfrak{gl}(n, \mathbb{R}) \mid \mathbf{A}^{\mathrm{T}} = -\mathbf{A}\}$

(c) $\mathfrak{u}(n) = \{\mathbf{A} \in \mathfrak{gl}(n, \mathbb{C}) \mid \mathbf{A}^{\dagger} = -\mathbf{A}\}$

(d) $\mathfrak{su}(n) = \{\mathbf{A} \in \mathfrak{gl}(n, \mathbb{C}) \mid \mathbf{A}^{\dagger} = -\mathbf{A} \text{ and } \operatorname{tr} \mathbf{A} = 0\}$

(e) $\mathfrak{gl}(n, \mathbb{C}) = \mathbb{C}^{n \times n} \cong \{\mathbf{A} \in \mathbb{R}^{2n \times 2n} \mid \mathbf{A}\mathbf{J}_0 = \mathbf{J}_0\mathbf{A}\}$

It is a basic fact in the theory of linear differential equations that for any matrix $\mathbf{A} \in \mathbb{F}^{n \times n}$, the initial value problem

$$\begin{cases} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) \\ \mathbf{x}(0) &= \mathbf{x}_0 \end{cases}$$

has a unique solution $\mathbf{x} : \mathbb{R} \to \mathbb{F}^n$ in the form $\mathbf{x}(t) = e^{t\mathbf{A}}\mathbf{x}_0$. The matrix exponential is defined by the power series

$$\exp(\mathbf{A}) := e^{\mathbf{A}} := \mathbb{1} + \mathbf{A} + \frac{\mathbf{A}^2}{2!} + \frac{\mathbf{A}^3}{3!} + \dots$$

which converges for all $\mathbf{A} \in \mathbb{F}^{n \times n}$. The smooth path of matrices $\mathbf{\Phi}(t) = \exp(t\mathbf{A})$ can alternatively be characterized as the unique solution to the problem

$$\begin{cases} \dot{\mathbf{\Phi}}(t) &= \mathbf{A}\mathbf{\Phi}(t) \\ \mathbf{\Phi}(0) &= \mathbb{1} \end{cases}$$

Observe that $\mathbf{\Phi}(s + t) = \mathbf{\Phi}(s)\mathbf{\Phi}(t)$, so $\mathbf{\Phi}$ defines a group homomorphism $\mathbb{R} \to \mathrm{GL}(n, \mathbb{F})$. This turns out to be another characterization of the matrix exponential:

Proposition B.24. *Let G be a Lie subgroup of $\mathrm{GL}(n, \mathbb{F})$, and pick any $\mathbf{A} \in \mathfrak{g}$. Then $t \mapsto \exp(t\mathbf{A})$ is a smooth path through G, and it is the unique homomorphism $\mathbb{R} \to G$ with velocity vector $\mathbf{A}$ at $t = 0$.*

A corollary is, for example, the fact that $e^{t\mathbf{A}}$ is orthogonal for all t if $\mathbf{A}$ is antisymmetric. To prove the result in general, we first introduce the concept of left invariant vector fields on a Lie group.

Suppose G is any Lie group (not necessarily a subgroup of $\mathrm{GL}(n, \mathbb{F})$) with identity element $e \in G$, and denote $\mathfrak{g} = T_eG$. For every $X \in \mathfrak{g}$, there is a unique smooth vector field $\widetilde{X} \in \mathrm{Vec}(G)$ which satisfies $\widetilde{X}(e) = X$ and is *left invariant*, meaning

$$\widetilde{X}(L_g(h)) = (L_g)_*\widetilde{X}(h),$$

where L_g is the diffeomorphism $G \to G : h \mapsto gh$. A formula for $\widetilde{X}$ is easily found by plugging in $h = e$, thus $\widetilde{X}(g) = (L_g)_*X$.

Exercise B.25. Verify that for any fixed $X \in \mathfrak{g} = T_e G$, the vector field on G defined by $\widetilde{X}(g) = (L_g)_* X$ is left invariant.

For subgroups $G \subset \mathrm{GL}(n, \mathbb{F})$, the left invariant vector fields take an especially simple form: recall that we can naturally regard any such group as a submanifold of $\mathbb{F}^{n\times n}$, thus all tangent vectors are matrices in $\mathbb{F}^{n\times n}$. Now for any $\mathbf{B} \in G$, the diffeomorphism $L_{\mathbf{B}} : G \to G : \mathbf{A} \mapsto \mathbf{BA}$ is also a linear map on $\mathbb{F}^{n\times n}$, so the unique left invariant vector field $X_{\mathbf{A}} \in \mathrm{Vec}(G)$ with $X_{\mathbf{A}}(\mathbb{1}) = \mathbf{A} \in \mathfrak{g}$ takes the form

$$X_{\mathbf{A}}(\mathbf{B}) = (L_{\mathbf{B}})_* \mathbf{A} = \left.\frac{d}{dt} \mathbf{B} \exp(t\mathbf{A})\right|_{t=0} = \mathbf{BA}.$$

Notice that this vector field is globally defined on $\mathbb{F}^{n\times n}$ for every $\mathbf{A}$, but it is also tangent to G by construction if $\mathbf{A} \in \mathfrak{g}$; in particular, its flow then preserves G.

We are now in a position to prove Proposition B.24. The crucial observation is that $\mathbf{\Phi}(t) = \exp(t\mathbf{A})$ is an orbit of the left invariant vector field $X_{\mathbf{A}}$; indeed,

$$\frac{d}{dt} \exp(t\mathbf{A}) = \exp(t\mathbf{A})\mathbf{A} = X_{\mathbf{A}}(\exp(t\mathbf{A})).$$

Thus if $\mathbf{A} \in T_{\mathbb{1}} G \subset \mathbb{F}^{n\times n}$, the orbit is confined to $G \subset \mathbb{F}^{n\times n}$. The uniqueness of homomorphisms $\mathbb{R} \to G$ now follows from the fact that *any* such map must also be an orbit of some left invariant vector field. In particular, if $\mathbf{\Phi} : \mathbb{R} \to G$ satisfies $\mathbf{\Phi}(s + t) = \mathbf{\Phi}(s)\mathbf{\Phi}(t)$ and $\dot{\mathbf{\Phi}}(0) = \mathbf{A}$, then

$$\begin{aligned}\dot{\mathbf{\Phi}}(t) &= \left.\frac{d}{ds}\mathbf{\Phi}(t+s)\right|_{s=0} = \left.\frac{d}{ds}\mathbf{\Phi}(t)\mathbf{\Phi}(s)\right|_{s=0} = \mathbf{\Phi}(t)\dot{\mathbf{\Phi}}(0) = \mathbf{\Phi}(t)\mathbf{A} \\ &= X_{\mathbf{A}}(\mathbf{\Phi}(t)).\end{aligned}$$

Consequently $\mathbf{\Phi}(t) = \exp(t\mathbf{A})$.

B.4 General Lie groups and their Lie algebras

The above discussion can be extended to arbitrary Lie groups, at the cost of a little abstraction. In particular, the bracket operation $[\ ,\ ]$ on $\mathfrak{g}$ and the map $\exp : \mathfrak{g} \to G$ can be defined in the general case, though we will no longer have such convenient formulas for computing them.

Suppose G is a Lie group with identity element $e \in G$, and denote $\mathfrak{g} = T_e G$. The quickest definition of the Lie bracket on $\mathfrak{g}$ is via left invariant vector fields, using the following result.

Proposition B.26. *For any two left invariant vector fields $\widetilde{X}$ and $\widetilde{Y}$ on G, the vector field $[\widetilde{X}, \widetilde{Y}]$ is also left invariant.*

This follows easily from the formula $\varphi_*[X,Y] = [\varphi_* X, \varphi_* Y]$, true for any pair of vector fields and any diffeomorphism φ on a manifold M. As a result, the space of left invariant vector fields is a finite dimensional Lie subalgebra of $\mathrm{Vec}(G)$, and we can pull back this algebraic structure to define a bracket operation on $\mathfrak{g}$. Let us adopt the convention that for any $X, Y \in \mathfrak{g}$, the unique left invariant vector fields with these values at the identity are denoted by $\widetilde{X}$ and $\widetilde{Y}$ respectively. Then the bracket on $\mathfrak{g}$ is defined such that

$$[X,Y] = Z \qquad \Longleftrightarrow \qquad [\widetilde{X}, \widetilde{Y}] = \widetilde{Z}.$$

Proposition B.27. *The general definition of $[X,Y]$ for $X, Y \in \mathfrak{g}$ matches the previous definition in the case $G \subset \mathrm{GL}(n, \mathbb{F})$, i.e. for any $\mathbf{A}, \mathbf{B} \in \mathfrak{g}$, if $X_{\mathbf{A}}$ and $X_{\mathbf{B}}$ are the corresponding left invariant vector fields on G, then $[X_{\mathbf{A}}, X_{\mathbf{B}}] = X_{\mathbf{AB}-\mathbf{BA}}$.*

Proof. The following is essentially a coordinate proof, though we've tried to avoid letting it look unnecessarily ugly. Recall that in any local coordinate chart $x^1, \ldots, x^n$ on a manifold M, the Lie bracket of two vector fields $X = X^i \partial_i$ and $Y = Y^i \partial_i$ has components

$$[X,Y]^i = X^j \partial_j Y^i - Y^j \partial_j X^i.$$

If M is an open subset of $\mathbb{R}^n$, then there are global coordinates and every tangent vector can be identified with a vector in $\mathbb{R}^n$, so the vector fields are smooth maps $M \to \mathbb{R}^n$ and the above formula says that for all $p \in M$,

$$[X,Y](p) = dY(p)X(p) - dX(p)Y(p). \tag{B.7}$$

We apply this to the special case $M = \mathrm{GL}(n, \mathbb{R}) \subset \mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}$. The left invariant vector field defined by $\mathbf{A} \in \mathfrak{gl}(n, \mathbb{R})$ takes the form $X_{\mathbf{A}}(\mathbf{p}) = \mathbf{pA}$, and since this is a linear map $\mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}$, its derivative is

$$dX_{\mathbf{A}}(\mathbf{p}) : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n} : \mathbf{H} \mapsto \mathbf{HA}.$$

Now (B.7) gives

$$\begin{aligned} [X_{\mathbf{A}}, X_{\mathbf{B}}](\mathbf{p}) &= dX_{\mathbf{B}}(\mathbf{p})\mathbf{pA} - dX_{\mathbf{A}}(\mathbf{p})\mathbf{pB} = \mathbf{pAB} - \mathbf{pBA} \\ &= X_{\mathbf{AB}-\mathbf{BA}}(\mathbf{p}). \end{aligned}$$

This proves the result for all matrix groups, since they are all subgroups of $\mathrm{GL}(n, \mathbb{R})$; note that this is even true for $\mathrm{GL}(n, \mathbb{C})$ and its subgroups (e.g. $\mathrm{U}(n)$), which can be viewed as subgroups of $\mathrm{GL}(2n, \mathbb{R})$. $\square$

We next construct a generalization of the matrix exponential. One can show that on any Lie group, the flow of a left invariant vector field exists globally for all time (see, for example, [GHL04]). Then if $\Phi : \mathbb{R} \to G$ is a smooth homomorphism with $\dot{\Phi}(0) = X \in \mathfrak{g}$, we have

$$\dot{\Phi}(t) = \frac{d}{ds}\Phi(t+s)\Big|_{s=0} = \frac{d}{ds}\Phi(t)\Phi(s)\Big|_{s=0} = (L_{\Phi(t)})_*\dot{\Phi}(0) = \widetilde{X}(\Phi(t)),$$

so Φ is an orbit of $\widetilde{X}$. This proves both the existence and uniqueness of such homomorphisms.

Definition B.28. For any Lie group G, the map $\exp : \mathfrak{g} \to G$ is defined so that $t \mapsto \exp(tX)$ is the unique smooth homomorphism $\mathbb{R} \to G$ with

$$\frac{d}{dt}\exp(tX)\Big|_{t=0} = X.$$

It turns out that exp is a smooth immersion, and becomes an embedding if we restrict to a small enough neighborhood of zero in $\mathfrak{g}$. More importantly for our present purposes, it provides a convenient formula for the flow of any left invariant vector field.

Exercise B.29. Show that if $\varphi^t_{\widetilde{X}} : G \to G$ is the flow of the left invariant vector field $\widetilde{X}$ with $\widetilde{X}(e) = X \in \mathfrak{g}$, then

$$\varphi^t_{\widetilde{X}}(g) = g\exp(tX). \tag{B.8}$$

We can now generalize the discussion by which the matrix Lie bracket was derived. Given $X, Y \in \mathfrak{g}$, we consider the smooth paths through $e \in G$ defined by

$$g_s = \exp(sX) \qquad \text{and} \qquad h_t = \exp(tY),$$

so $\dot{g}_0 = X$ and $\dot{h}_0 = Y$. Define a smooth "commutator" map $f : \mathbb{R}^2 \to G$ by

$$f(s,t) = g_s h_t g_s^{-1} h_t^{-1} = \exp(sX)\exp(tY)\exp(-sX)\exp(-tY). \tag{B.9}$$

As in the matrix case, the extent to which $f(s,t)$ deviates from e for (s,t) near $(0,0)$ measures the noncommutativity of G. We can use Exercise B.29, to reexpress this in terms of the flows of the left invariant vector fields $\widetilde{X}$ and $\widetilde{Y}$: in particular, if we define a map $F : \mathbb{R}^2 \times G \to G$ by

$$F(s,t,g) = \varphi^{-t}_{\widetilde{Y}} \circ \varphi^{-s}_{\widetilde{X}} \circ \varphi^{t}_{\widetilde{Y}} \circ \varphi^{s}_{\widetilde{X}}(g),$$

then $F(s,t,e) = f(s,t)$. Now, it is a familiar fact from differential geometry that the lowest order nonvanishing derivative of $F(s,t,\cdot)$ at $(s,t) = 0$ is precisely the Lie bracket of $\widetilde{X}$ and $\widetilde{Y}$: in particular

$$\nabla_s \partial_t F(0,0,g) = [\widetilde{X},\widetilde{Y}](g).$$

(A version of this is proved in [Spi99].) Consequently,

$$\nabla_s \partial_t f(0,0) = [X, Y].$$

One can take this as a coordinate free proof that our new definition of the bracket on $\mathfrak{g}$ matches the old one for $G \subset \mathrm{GL}(n, \mathbb{F})$.

Proposition B.30. *A connected Lie group G is abelian if and only if its Lie algebra $\mathfrak{g}$ is abelian.*

Proof. The above discussion shows that $[\ ,\ ]$ vanishes if G is abelian. To show the converse, recall from Prop. B.16 that every connected Lie group is generated by a neighborhood of $e \in G$, thus it suffices to show that

$$\exp(X)\exp(Y) = \exp(Y)\exp(X)$$

for all $X, Y \in \mathfrak{g}$, or equivalently using (B.9), $f(s,t) = e$ for all $s, t \in \mathbb{R}$. This follows from (B.8) and the fact that if $[\widetilde{X}, \widetilde{Y}] \equiv 0$, then the flows of $\widetilde{X}$ and $\widetilde{Y}$ commute. $\square$

Exercise B.31. Show that all connected 1-dimensional Lie groups are abelian.

There is one more piece of general Lie group theory that will be important in the discussions to come. For two Lie groups G and H, a map $\varphi : G \to H$ is called a *Lie group homomorphism* if it is both a group homomorphism and a smooth map. Then the tangent map at the identity is a *Lie algebra homomorphism*

$$d\varphi(e) : \mathfrak{g} \to \mathfrak{h},$$

i.e. a linear map $A : \mathfrak{g} \to \mathfrak{h}$ that satisfies $A([X,Y]) = [A(X), A(Y)]$. This follows easily from the above discussion, together with the fact that

$$\varphi \circ \exp(tX) = \exp[t(\varphi_* X)]$$

for every $X \in \mathfrak{g}$.

An important special case of a Lie group homomorphism is a representation. In particular, given G and a vector space V, a *representation of G on V* is a Lie group homomorphism

$$\rho : G \to \mathrm{Aut}(V),$$

where $\mathrm{Aut}(V)$ is the group of invertible linear maps on V. This induces a *Lie algebra representation*

$$\bar{\rho} : \mathfrak{g} \to \mathrm{End}(V),$$

satisfying $\bar{\rho}([X,Y]) = [\bar{\rho}(X), \bar{\rho}(Y)]$, with the bracket on $\mathrm{End}(V)$ defined as the commutator $[A,B] := AB - BA$. This fact is crucial in the study of group representations: the first step in the effort to classify all representations of a given group G is to classify the representations of its Lie algebra, usually a simpler problem since it is essentially linear in nature.

Every Lie group has a special representation on its own Lie algebra, called the *adjoint representation*,

$$\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g}).$$

This is defined via the conjugation map $C_g : G \to G : h \mapsto ghg^{-1}$, thus

$$\mathrm{Ad}_g(X) := (C_g)_* X = \frac{d}{dt} g \exp(tX) g^{-1}\Big|_{t=0}.$$

We observe that conjugation is a smooth left action of G on itself, i.e. $C_{gh} = C_g \circ C_h$. It follows easily that Ad is a group homomorphism since

$$\mathrm{Ad}_{gh} = (C_{gh})_* = (C_g \circ C_h)_* = (C_g)_*(C_h)_* = \mathrm{Ad}_g \circ \mathrm{Ad}_h .$$

Exercise B.32. Show that if G is a Lie subgroup of $\mathrm{GL}(n, \mathbb{F})$, the adjoint representation takes the form $\mathrm{Ad}_{\mathbf{B}}(\mathbf{A}) = \mathbf{B}\mathbf{A}\mathbf{B}^{-1}$.

The corresponding Lie algebra representation is denoted

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}) : X \mapsto \mathrm{ad}_X .$$

This is simplest to compute in the case of a matrix group $G \subset \mathrm{GL}(n, \mathbb{F})$, where

$$\mathrm{ad}_{\mathbf{A}}(\mathbf{B}) = \frac{d}{dt} \mathrm{Ad}_{e^{t\mathbf{A}}}(\mathbf{B})\Big|_{t=0} = \frac{d}{dt} e^{t\mathbf{A}} \mathbf{B} e^{-t\mathbf{A}}\Big|_{t=0} = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} = [\mathbf{A}, \mathbf{B}].$$

Exercise B.33. Use the Jacobi identity to verify that the map $\mathfrak{gl}(n, \mathbb{F}) \to \mathrm{End}(\mathfrak{gl}(n, \mathbb{F}))$ sending $\mathbf{A}$ to the linear map $\mathbf{B} \mapsto [\mathbf{A}, \mathbf{B}]$ is a Lie algebra homomorphism.

References

[Car] S. M. Carroll, *Lecture Notes on General Relativity*. available at http://pancake.uchicago.edu/~carroll/notes/.

[GHL04] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian geometry*, 3rd ed., Springer-Verlag, Berlin, 2004.

[Hat02] A. Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002.

[Spi99] M. Spivak, *A comprehensive introduction to differential geometry*, 3rd ed., Vol. 1, Publish or Perish Inc., Houston, TX, 1999.

[Str80] G. Strang, *Linear algebra and its applications*, 2nd ed., Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1980.