Notes on *Riemannian Geometry*

Louis Bergamo Radial

January 26, 2024

Contents

Chapter 1

Topological Manifolds

1.1 Topology

In order to define the notion of smooth manifolds, we must first begin with some building blocks, such as topology and topological manifolds.

> **Definition 1.1: Topology**
>
> A *topology* on the set M is a family O of subsets of M satisfying
>
> (a) the empty set and the set M belong to O;
>
> (b) a finite intersection of elements of O is a member of O; and
>
> (c) an arbitrary union of members of O belongs to O.
>
> The pair (M, O_M) is named a *topological space*, elements of O are called *open sets* and elements of $M \setminus O$ are called *closed sets*. Additionally, given an element $p \in M$ an open set U that contains p is called a *neighborhood* of p.

In Propositions 1.1 to 1.3 we show a couple of important examples that illustrate how the axioms of topological spaces given in Definition 1.1 are used.

> **Proposition 1.1: Standard topology in $\mathbb{R}^n$**
>
> We define the *open ball* $B_n(r, p) \subset \mathbb{R}^n$ of radius $r > 0$ centered at $p = (p^1, \ldots, p^n)$ as the set
>
> $$B_n(r, p) = \left\{ q = (q^1, \ldots, q^n) \in \mathbb{R}^n : \sum_{i=1}^{n} (q^i - p^i)^2 < r^2 \right\}.$$
>
> Next, we define the *standard topology* O_{standard} of $\mathbb{R}^n$. A subset $U \subset \mathbb{R}^n$ is an open set if for every point $p \in U$ there exists $r > 0$ such that $B_n(r, p) \subset U$. Then, $(\mathbb{R}^n, O_{\text{standard}})$ is a topological space.

Proof. It is easy to see $\mathbb{R}^n \in O_{\text{standard}}$ and $\emptyset \in O_{\text{standard}}$.

Suppose $U, V \in O_{\text{standard}}$ and let $p \in U \cap V \neq \emptyset$. Then, there exists $r_U > 0$ and $r_V > 0$ such that $B_n(r_U, p) \subset U$ and $B_n(r_V, p) \subset V$. Setting $r = \min\{r_U, r_V\} > 0$ we have $B_n(r, p)$ as subset of both U and V, that is, $B_n(r, p) \subset U \cap V$. It follows that $U \cap V \in O_{\text{standard}}$.

Let $\{U_\alpha\}_{\alpha \in J}$ be a family of sets in O_{standard}. Let $p \in \bigcup_{\alpha \in J} U_\alpha$, that is, there exists $\beta \in J$ such that $p \in U_\beta$. Since $U_\beta \in O_{\text{standard}}$, there exists $r_\beta > 0$ such that $B_n(r, p) \subset U_\beta \subset \bigcup_{\alpha \in J} U_\alpha$. $\qquad \square$

> **Proposition 1.2: Subspace topology is a topology**
>
> Given a topological space (M, O_M) and a subset S of M, we define the *subspace topology* $O_M|_S$ as
> $$O_M|_S = \{U \cap S : U \in O_M\}.$$
> Then $(S, O_M|_S)$ is a topological space.

Proof. We must show the conditions (a), (b), and (c) of Definition 1.1 are satisfied.

(a) Since $S = M \cap S$ and $\emptyset = \emptyset \cap S$, we have $S \in O_M|_S$ and $\emptyset \in O_M|_S$.

(b) Let $U, V \in O_M|_S$. Then, there exists $\tilde{U}, \tilde{V} \in O_M$ such that $U = \tilde{U} \cap S$ and $V = \tilde{V} \cap S$. Then, $U \cap V = (\tilde{U} \cap S) \cap (\tilde{V} \cap S) = (\tilde{U} \cap \tilde{V}) \cap S$. Since $\tilde{U} \cap \tilde{V} \in O_M$, we have $U \cap V \in O_M|_S$.

(c) Let $\{U_\alpha\}_{\alpha \in J}$ be a family of open sets in $O_M|_S$. For each $\alpha \in J$, there exists a $\tilde{U}_\alpha \in O_M$ such that $U_\alpha = \tilde{U}_\alpha \cap S$. Then
$$\bigcup_{\alpha \in J} U_\alpha = \bigcup_{\alpha \in J} \tilde{U}_\alpha \cap S$$
$$= \{m \in S : \exists \alpha \in J \text{ such that } m \in \tilde{U}_\alpha\}$$
$$= \{m \in M : \exists \alpha \in J \text{ such that } m \in \tilde{U}_\alpha\} \cap S$$
$$= S \cap \bigcup_{\alpha \in J} \tilde{U}_\alpha.$$

Since arbitrary unions of open sets is an open set, it follows that $\bigcup_{\alpha \in J} U_\alpha \in O_M|_S$.

$\square$

> **Proposition 1.3: Product topology**
>
> Let (M, O_M) and (N, O_N) be topological spaces. Define the *product topology* $O_{M \times N}$ as the collection of subsets $U \subset M \times N$ such that for all $(m, n) \in U$, there exists neighborhoods $S \subset M$ and $T \subset N$ of $m \in M$ and $n \in N$ such that $S \times T \subset U$. Then $(M \times N, O_{M \times N})$ is a topological space.

Proof. Clearly, $M \times N$ and $\emptyset$ are open sets in the product topology.

Next, we consider open sets $U, V \in O_{M \times N}$ and an element $p \in U \cap V$. Let $p = (m, n) \in M \times N$, then there exists neighborhoods $S_U, S_V \subset M$ of m and $T_U, T_V \subset N$ of n such that $S_U \times T_U \subset U$ and $S_V \times T_V \subset V$. Let $S = S_U \cap S_V$ and $T = T_U \cap T_V$, then $S \in O_M$ and $T \in O_N$ are neighborhoods of m and n, respectively. Moreover, $S \times T \subset U \cap V$ is a neighborhood of p, from which follows $U \cap V \in O_{M \times N}$.

Let $\{U_\alpha\}_{\alpha \in J}$ be a family of open sets in the product topology. Let $p \in \bigcup_{\alpha \in J} U_\alpha$, then there exists $\beta \in J$ such that $p \in U_\beta$. By definition, there exists open sets $S \in O_M$ and $T \in O_N$ such that $S \times T \subset U_\beta \subset \bigcup_{\alpha \in J} U_\alpha$. Therefore, $\bigcup_{\alpha \in J} U_\alpha$ is an open set. $\square$

Along with the axioms of topological spaces described in Definition 1.1 one might add further restrictions to specify the space considered. Some common restrictions are called the *separation axioms*. Among these, we will make use of the T2 axiom, namely the Hausdorff property. Historically, Felix Hausdorff used this axiom in his original definition of a topological space, although the formulation of his other axioms was not exactly as those of Definition 1.1, but an equivalent one.

Definition 1.2: Hausdorff space

A topological space $(M, \mathcal{O}_M)$ is called a *Hausdorff space* if for any $p, q \in M$ with $p \neq q$, there exists a neighborhood U of p, i.e. $p \in U \in \mathcal{O}_M$, and a neighborhood V of q such that $U \cap V = \emptyset$.

1.2 Convergence

In analysis on $\mathbb{R}^n$ with the standard topology, we often consider sequences $x : \mathbb{N} \to \mathbb{R}^n$ and study whether it converges to a value. We say the sequence x converges to $y \in \mathbb{R}^n$ if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $x(i) - y \in B_n(\varepsilon, 0)$ for all $i > N$. We generalize the notion of a convergent sequence to any topological space in Definition 1.3.

Definition 1.3: Convergence of a sequence

A sequence $x : \mathbb{N} \to M$ on a topological space $(M, \mathcal{O}_M)$ is said to *converge* to a *limit point* $p \in M$ if for every neighborhood $U \in \mathcal{O}_M$ of p there exists $N \in \mathbb{N}$ such that $x(n) \in U$ for all $n > N$.

Theorem 1.1: Unique limit on Hausdorff spaces

Let $(M, \mathcal{O}_M)$ be a Hausdorff space. If a sequence x converges on M, its limit point is unique.

Proof. Let $p, q \in M$ be limit points of the sequence x. Suppose, by contradiction, that $p \neq q$. By the Hausdorff property, there exists neighborhoods $U, V \in \mathcal{O}_M$ of p and q, respectively, such that $U \cap V = \emptyset$. From the definition of convergence, there exists $N_p, N_q \in \mathbb{N}$ such that $x(n) \in U$ for all $n > N_p$ and $x(n) \in V$ for all $n > N_q$. Let $N = \min\{N_p, N_q\}$, then for all $n > N$, $x(n) \in U$ and $x(n) \in V$, that is, $x(n) \in U \cap V = \emptyset$. This contradiction proves the statement. $\square$

1.3 Homeomorphisms

With the notion of topological spaces, we may ask ourselves whether certain maps between topological spaces can preserve the topology. That is, a map that takes open sets in the domain topology into open sets in the target topology. To define such a map we define *continuity*.

Definition 1.4: Continuous map

Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be topological spaces. Then a map $f : M \to N$ is *continuous* (with respect to $\mathcal{O}_M$ and $\mathcal{O}_N$) if, for all $V \in \mathcal{O}_N$, the preimage $f^{-1}(V)$ is an open set in $\mathcal{O}_M$.

In short, a map is continuous if and only the preimages of (all) open sets are open sets. Now a map that preserves the topology is called a *homeomorphism*, which is defined as a continuous bijection with continuous inverse. We now prove such a map satisfies the condition required.

> ### Proposition 1.4: Homeomorphism maps open sets to open sets
>
> Let (M, O_M) and (N, O_N) be topological spaces. Suppose a map $f : M \to N$ is a homeomorphism, then f maps open sets in O_M into open sets in O_N.

Proof. Given a subset $U \in O_M$, we must show the image $V = f(U)$ is open in (N, O_N). Taking our attention to the inverse map $g = f^{-1} : N \to M$, we see the preimage $g^{-1}(U) = V$ must be open in (N, O_N), due to continuity. $\qquad\square$

If there exists a homeomorphism between two topological spaces, they are said to be homeomorphic to each other. This begs the question: if (M, O_M) is homeomorphic to (N, O_N) and (N, O_N) is homeomorphic to (P, O_P), are (M, O_M) and (P, O_P) homeomorphic? To answer this we must show whether the composition of continuous maps is itself continuous.

> ### Theorem 1.2: Composition of continuous maps
>
> Let (M, O_M), (N, O_N), and (P, O_P) be topological spaces. If the maps $f : M \to N$ and $g : N \to P$ are continuous (with respect to the appropriate topologies), then the map $g \circ f : M \to P$ is continuous with respect to O_M and O_P.

Proof. Let V be an open set of (P, O_P). We must show the preimage $(g \circ f)^{-1}(V)$ is an open set of (M, O_M). We have

$$(g \circ f)^{-1}(V) = \{m \in M : g \circ f(m) \in V\}$$
$$= \{m \in M : f(m) \in g^{-1}(V)\}$$
$$= f^{-1}\left(g^{-1}(V)\right).$$

Since the map g is continuous and V is an open set in (P, O_P), it follows that $g^{-1}(V)$ is open in (N, O_N). By the same argument, $f^{-1}\left(g^{-1}(V)\right)$ is an open set in (M, O_M). $\qquad\square$

Corollary 1.1. *If (M, O_M) is homeomorphic to (N, O_N) and (N, O_N) is homeomorphic to (P, O_P), then (M, O_M) is homeomorphic to (P, O_P).*

Proof. Let $f : M \to N$ and $g : N \to P$ be homeomorphisms from (M, O_M) to (N, O_N) and (N, O_N) to (P, O_P), respectively. Consider the composition $g \circ f : M \to P$.

$$
\begin{array}{ccc}
M & \xrightarrow{\ f\ } & N \\
 & {\scriptstyle g\circ f}\searrow & \downarrow {\scriptstyle g} \\
 & & P
\end{array}
$$

By Theorem 1.2, the map $g \circ f$ is a homeomorphism from (M, O_M) to (P, O_P). $\qquad\square$

As was done for the subspace topology, we prove a similar result for continuous maps.

> ### Proposition 1.5: Restriction of a continuous map
>
> Let (M, O_M) and (N, O_N) be topological spaces and let $f : M \to N$ be a continuous map. Let S be a subset of M and let (S, O_S) be the subspace topology, then $f|_S : S \to N$ is a continuous map with respect to O_S and O_N.

Proof. Let $V \in \mathcal{O}_N$. Then, by the definition of preimage, we have

$$
\begin{aligned}
f|_S^{-1}(V) &= \{s \in S : f|_S(s) \in V\} \\
&= \{s \in S : f(s) \in V\} \\
&= f^{-1}(V) \cap S.
\end{aligned}
$$

By hypothesis, the preimage $f^{-1}(V)$ is an open set in $(M, \mathcal{O}_M)$, so $f|_S^{-1}(V)$ is an open set in the subspace topology. $\qquad\square$

We can now define the notion of a topological space locally resembling Euclidean space.

> **Definition 1.5: Locally Euclidean topological space**
>
> A topological space $(M, \mathcal{O}_M)$ is *locally Euclidean* of dimension n if for all $m \in M$ there exists an open subset $U \in \mathcal{O}_M$ about m that is homeomorphic to $\mathbb{R}^n$ with respect to the subspace topology and the standard topology of $\mathbb{R}^n$.

It is sufficient to show the subspace topology $(U, \mathcal{O}_U)$ is homeomorphic to an open ball in $\mathbb{R}^n$, due to Proposition 1.6.

> **Proposition 1.6: Open ball is homeomorphic to the Euclidean space**
>
> Let $r > 0$, then the map $f : B_n(r, 0) \subset \mathbb{R}^n \to \mathbb{R}^n$ given by
>
> $$ f(x) = \frac{x}{r - \|x\|} $$
>
> is a homeomorphism with respect to the standard topology.

Proof. We begin by checking f is one-to-one and onto.

Suppose there exists $x_1, x_2 \in B_n(r, 0)$ such that $f(x_1) = f(x_2)$. It follows from

$$
\begin{aligned}
f(x_2) - f(x_1) &= \frac{x_2}{r - \|x_2\|} - \frac{x_1}{r - \|x_1\|} \\
&= \frac{(r - \|x_1\|) x_2 - (r - \|x_2\|) x_1}{(r - \|x_2\|)(r - \|x_1\|)}
\end{aligned}
$$

that $(r - \|x_1\|) x_2 = (r - \|x_2\|) x_1$. Taking the norm on both sides, we have $\|x_1\| = \|x_2\|$. Substituting back, we have $x_1 = x_2$, proving f is injective.

Suppose $y \in \mathbb{R}^n$ and consider $\xi = \frac{ry}{1 + \|y\|}$. Clearly, $\xi \in B_n(r, 0)$. We have

$$
\begin{aligned}
f(\xi) &= f\left(\frac{ry}{1 + \|y\|}\right) \\
&= \frac{ry}{1 + \|y\|} \frac{1}{r - \left\|\frac{ry}{1 + \|y\|}\right\|} \\
&= \frac{1}{(1 + \|y\|)\left(1 - \frac{\|y\|}{1 + \|y\|}\right)} y \\
&= y,
\end{aligned}
$$

so f is onto.

We have shown f is a bijection with inverse $f^{-1} : \mathbb{R}^n \to B_n(r, 0)$ defined by

$$f^{-1}(x) = \frac{rx}{1 + \|x\|}. \tag{1.3.1}$$

With the standard topology, continuity of f and f^{-1} follows from techniques of elementary calculus, and we conclude f is a homeomorphism.   □

 Suppose we are given a set M, a topological space $(N, \mathcal{O}_N)$ and a map $f : M \to N$. Without structure on the set M, we may not say whether f is a continuous map. Is there a way to induce a topology on M from the topology of N and the map f? In the Proposition 1.7 we construct a topology in M such that f is a continuous map from M to N.

> **Proposition 1.7: Initial topology**
>
> TODO

Proof.   □

1.4   Compactness and paracompactness

> **Definition 1.6: Compactness**
>
> A topological space $(M, \mathcal{O}_M)$ is *compact* if every *open cover* of M has a finite subcover. That is, the topological space is compact if for every family of open sets C that covers M, i.e. $\bigcup_{U \in C} U = M$ with $U \in \mathcal{O}_M$, there exists a finite family of open sets $F \subset C$ such that $\bigcup_{U \in F} U = M$.
>
> Additionally, in a topological space $(N, \mathcal{O}_N)$, a subset $S \subset N$ is called compact if the subspace topology is compact.

> **Theorem 1.3: Heine-Borel theorem**
>
> A subset $S \subset \mathbb{R}^n$ with the standard topology is compact if it is closed and bounded.

Proof. Refer to [Rud64].   □

> **Definition 1.7: Locally finite cover**
>
> A cover C of a topological space $(M, \mathcal{O}_M)$ is called *locally finite* if each point in the space has a neighborhood that intersects only finitely many sets in C. More precisely, for all $p \in M$ there exists a neighborhood $U \in \mathcal{O}_M$ about p such that $U \cap V \neq \emptyset$ only for finitely many $V \in C$.

> **Definition 1.8: Refinement**
>
> A *refinement* of a cover C of a topological space $(M, \mathcal{O}_M)$ is a cover D such that every set in D is contained in some set in C. Precisely, let $C = \{U_\alpha\}_{\alpha \in A}$ and $D = \{V_\beta\}_{\beta \in B}$ such that $\bigcup_{\alpha \in A} U_\alpha = M$ and $\bigcup_{\beta \in B} V_\beta = M$, then D is a refinement of C if for all $\beta \in B$ there exists $\alpha \in A$ such that $V_\beta \subset U_\alpha$.

> **Definition 1.9: Paracompactness**
>
> A topological space $(M, \mathcal{O}_M)$ is called *paracompact* if every open cover C has an *open refinement* $\tilde{C}$ that is *locally finite*.

> **Theorem 1.4: Stone's theorem**
>
> Any *metrizable space* is paracompact.

Proof. Refer to [Mun00]. □

> **Definition 1.10: Partition of unity subordinate to a cover**
>
> Let $C = \{U_\alpha\}_{\alpha \in J}$ be an open cover of the topological space $(M, \mathcal{O}_M)$. The family of functions $\mathcal{F}_C = \{f_\alpha\}_{\alpha \in J}$ from M to $[0, 1] \subset \mathbb{R}$ is a (continuous) *partition of unity subordinate to U_α* if
>
> (a) each $f_\alpha : X \to [0, 1]$ is continuous;
>
> (b) for each $\alpha \in J$, the support of f_α is contained in U_α, that is, for every $f \in \mathcal{F}_C$ there exists $U \in C$ such that $f(p) \neq 0 \implies p \in U$;
>
> (c) the family $\{\operatorname{supp} f_\alpha\}_{\alpha \in J}$ is locally finite, i.e. for any $p \in M$ there exists a neighborhood $U \in \mathcal{O}_M$ about p where all but finitely many functions f_α vanish on U;
>
> (d) for any point $p \in M$, $\sum_{\alpha \in J} f_\alpha = 1$.

> **Theorem 1.5: Paracompactness and partitions of unity**
>
> Let $(M, \mathcal{O}_M)$ be a Hausdorff space. Then it is paracompact if and only if every open cover C admits a partition of unity subordinate to that cover.

Proof. Refer to [Mun00]. □

1.5 Connectedness and path-connectedness

> **Definition 1.11: Connectedness**
>
> A topological space $(M, \mathcal{O}_M)$ is *connected* unless there exists two non-empty, non-intersecting open sets $A, B \in \mathcal{O}_M$ such that $M = A \cup B$.

> **Theorem 1.6: Interval is connected**
>
> Every interval $I \subset \mathbb{R}$ is connected with respect to the standard topology.

Proof. Suppose I is not connected, then $I = A \cup B$, where $A, B \subset I$ are non-empty, non-intersecting open sets. Let $a \in A$ and $b \in B$. Without loss of generality, we assume $a < b$.

Consider $\alpha = \sup\{x \in \mathbb{R} : [a, x) \cap I \subset A\}$. Then $a \leq \alpha \leq b$, so $\alpha \in I$. Since $B = I \setminus A$ is open, we have A closed, hence $\alpha \in A$. Since A is open, there exists $r > 0$ such that $(\alpha - r, \alpha + r) \cap I \subset A$. We conclude $(a, \alpha + r) \cap I \subset A$, which is a contradiction. □

> **Theorem 1.7: Open and closed subsets in a connected space**
>
> A topological space (M, O_M) is connected if and only if $\emptyset$ and M are the only subsets that are both open and closed.

Proof. Suppose the topological space is connected. Suppose, by contradiction, the non-empty subset $U \subsetneq M$ is open and closed. It follows from $M = U \cup (M \setminus U)$ that the topological space is not connected, a contradiction.

Suppose the empty set and M are the only subsets that are both open and closed. Suppose, by contradiction, that the topological space is not connected. Then there exists two non-empty non-intersecting open sets $A, B \in O_M$ such that $M = A \cup B$. Clearly, $B = M \setminus A$ is closed, and likewise, A is closed. By hypothesis, $A = B = M$ since they are both open and closed and are non-empty. We have thus arrived at a contradiction, since $A \cap B = M$ is non-empty, which proves the topological space is connected. $\qquad\square$

> **Definition 1.12: Path-connectedness**
>
> A topological space (M, O_M) is called *path-connected* if for every pair of points $p, q \in M$ there exists a continuous curve $\gamma : [0, 1] \to M$ such that $\gamma(0) = p$ and $\gamma(1) = q$.

> **Theorem 1.8: Path-connectedness implies connectedness**
>
> A path-connected topological space is connected.

Proof. Suppose, by contradiction, a topological space (M, O_M) is path-connected but not connected. Then, there exists non-empty, non-intersecting open sets $A, B \in O_M$ such that $M = A \cup B$. Choose $a \in A$ and $b \in B$. Since (M, O_M) is path-connected, there exists a continuous curve $\gamma : [0, 1] \to M$ such that $\gamma(0) = a$ and $\gamma(1) = b$. Consider the preimage $[0, 1] = \gamma^{-1}(M)$. If follows from

$$
\begin{aligned}
\gamma^{-1}(M) &= \gamma^{-1}(A \cup B) \\
&= \{ x \in [0, 1] : \gamma(x) \in A \cup B \} \\
&= \gamma^{-1}(A) \cup \gamma^{-1}(B)
\end{aligned}
$$

that $[0, 1]$ is the union of two non-empty sets, since $0 \in \gamma^{-1}(A)$ and $1 \in \gamma^{-1}(B)$. Suppose $\gamma^{-1}(A) \cap \gamma^{-1}(B)$ is non-empty. Then there exists $t \in [0, 1]$ such that $\gamma(t) \in A$ and $\gamma(t) \in B$, that is, $\gamma(t) \in A \cap B$. This contradiction shows $\gamma^{-1}(A)$ and $\gamma^{-1}(B)$ are non-intersecting. By continuity, these sets are both open and closed. Therefore, $[0, 1]$ is not connected. By Theorem 1.6, this is a contradiction, and the theorem follows. $\qquad\square$

1.6 Homotopic curves and the fundamental group

> **Definition 1.13: Group**
>
> A *group* is a non-empty set G, called the *underlying set*, closed under a map $\cdot : G \times G \to G$, named *group operation*, such that the *group axioms* are satisfied:
>
> (a) Associativity: For all $a, b, c \in G$, on has $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
>
> (b) Identity element: There exists a unique element $e \in G$, called the *identity element* or *neutral element* of the group, such that for every $g \in G$, $e \cdot g = g$ and $g \cdot e = g$.

(c) Inverse element: For each $g \in G$, there exists a unique element $g^{-1} \in G$ such that $g \cdot g^{-1} = e$ and $g^{-1} \cdot g = e$.

If the group operation is commutative, $G, \cdot$ is named an *abelian group*.

Remark 1.1. *If the group operation is notated as addition, the inverse element of g is typically denoted by $-g$, the identity element denoted by 0, and the group may be called an* additive group. *Similarly, a group may be called* multiplicative group *if the group operation is notated as multiplication, and the identity element is denoted by 1. Additionally, in a multiplicative group, the group operation may be denoted simply by juxtaposition, that is, $f \cdot g = fg$.*

Definition 1.14: Homotopic curves

Let $(M, \mathcal{O}_M)$ be a topological space and let $p, q \in M$. Two curves $\gamma, \eta : [0,1] \to M$ from p to q, i.e. $\gamma(0) = \eta(0) = p$ and $\gamma(1) = \eta(1) = q$, are *homotopic* if there exists a continuous map $h : [0,1] \times [0,1] \to M$ such that $h(0, \lambda) = \gamma(\lambda)$ and $h(1, \lambda) = \eta(\lambda)$, for all $\lambda \in [0,1]$.

It is easy to check that homotopic curves form a equivalence relation. (do that)

Definition 1.15: Loops at a point

Let $(M, \mathcal{O}_M)$ be a topological space and let $p \in M$. We define the *space of loops at p* as the set $\mathscr{L}_p$ of continuous curves $\gamma : [0,1] \to M$ such that $\gamma(0) = \gamma(1) = p$.
A *concatenation* is a binary operation $*_p : \mathscr{L}_p \times \mathscr{L}_p \to \mathscr{L}_p$ defined by

$$(\gamma *_p \eta)(\lambda) = \begin{cases} \gamma(2\lambda), & \text{for } \lambda \in \left[0, \frac{1}{2}\right] \\ \eta(2\lambda - 1), & \text{for } \lambda \in \left(\frac{1}{2}, 1\right] \end{cases}$$

for all $\gamma, \eta \in \mathscr{L}_p$.

Definition 1.16: Fundamental group

The *fundamental group* $(\pi_{1,p}, \cdot)$ of a topological space is the set

$$\pi_{1,p} = \mathscr{L}_p / \text{homotopy} = \{[\gamma]_{\text{homotopy}} : \gamma \in \mathscr{L}_p\}$$

together with the product $\cdot : \pi_{1,p} \times \pi_{1,p} \to \pi_{1,p}$ defined by

$$[\gamma] \cdot [\eta] = [\gamma *_p \eta].$$

Example 1.1. (a) *On the two-sphere S^2, the fundamental group has a single element, represented by the constant loop.*

(b) *For the cylinder $C = \mathbb{R} \times S^1$, the fundamental group is homomorphic to the group $(\mathbb{Z}, +)$. There exists a bijection $f : \pi_1 \to \mathbb{Z}$ with the property $f(\alpha \cdot \beta) = f(\alpha) + f(\beta)$.*

(c) *On the torus $T^2 = S^1 \times S^1$, we have π_1 isomorphic to $\mathbb{Z} \times \mathbb{Z}$.*

> **Definition 1.17: Simply connected**
>
> A topological space (M, O_M) is *simply connected* if it is path-connected and if for every point $p \in M$ the fundamental group $(\pi_{1,p}, \cdot)$ is the trivial group.

Remark 1.2. *Building up on the previous examples, we see that the two-sphere is simply connected, while the cylinder and the torus are not, although path-connected.*

1.7 Topological manifolds

We now finally define the notion of a manifold, which is a "well-behaving" topological space that *locally* looks like Euclidean space.

> **Definition 1.18: Topological manifold**
>
> A paracompact Hausdorff space (M, O_M) locally Euclidean of dimension n is called a *n-dimensional topological manifold*.

Example 1.2. *We list a few examples of constructions of new manifolds from other manifolds.*

(a) *As with topological spaces, it is clear that the subspace topology (N, O_N) is in its own right a manifold. This construction takes the name of* submanifold.

(b) *Let (M, O_M) be an m-dimensional manifold and (N, O_N) be an n-dimensional manifold. Then the product topology $(M \times N, O_{M \times N})$ is an $(n + m)$-dimensional manifold. As an example, the circle S^1 is a 1-dimensional manifold, so the torus $T^2 = S^1 \times S^1$ is a 2-dimensional manifold.*

1.7.1 Bundles

It is clear we may not always construct a manifold as a product manifold. To see this, take the Möbius strip. We generalize this concept with *bundles*.

> **Definition 1.19: Bundles of topological manifolds**
>
> A *bundle of topological manifolds* is a triple (E, π, M), where (E, O_E) is a topological manifold called the *total space*, (M, O_M) is a topological manifold called the *base space*, and $\pi : E \to M$ is a continuous surjective map called the *projection*. Let $p \in M$, then $\pi^{-1}(\{p\}) = F_p$ is the *fiber at p*.

Example 1.3. (a) *The first example is when the total space $E = M \times F$ is a product manifold of the base space M and a fiber F. We take the projection $\pi : M \times F \to M$ as the continuous map $(p, f) \mapsto p$.*

(b) *We now take the Möbius strip as the total space and the base space as S^1. For any point in S^1, the preimage of the projection is some interval on the real line.*

(c) *We consider $M = \mathbb{R}$ as the base space. We begin construct the total space by attaching to each point of M a circle. Then, we continuously deform the circles such that they collapse to a point on the positive numbers of the real line. Finally, we continuously stretch the point on that half of the line to increasing intervals. The fiber at any given point may be homeomorphic to a circle or to a point, or to an interval.*

In the last example, a bundle was constructed with a fiber space that is not the same at different points. This motivates the following definition.

> **Definition 1.20: Fiber bundle**
>
> A bundle (E, π, M) is a *fiber bundle with typical fiber F* if for all $p \in M$, the preimage $\pi^{-1}(\{p\})$ is homeomorphic to the manifold F. The diagram
>
> $$F \longrightarrow E \xrightarrow{\pi} M$$
>
> denotes a fiber bundle.

As an example, we consider the $\mathbb{C}$-*line bundle over M* as the bundle (E, π, M) with typical fiber $\mathbb{C}$.

> **Definition 1.21: Section of the bundle**
>
> Let (E, π, M) be a bundle. A map $\sigma : M \to E$ is called a section of the bundle if $\pi \circ \sigma = \mathrm{id}_M$.

As a special case we consider the "product bundle"

$$F \longrightarrow M \times F \xrightarrow{\pi} M,$$

where $\pi : M \times F \to M$ is the projection $(p, f) \mapsto p$. Given any function $s : M \to F$, we may construct a section $\sigma : M \to M \times F$ defined by $p \mapsto (p, s(p))$.

Given a bundle (E, π, M), we may consider the submanifolds $E' \subset E$ and $M' \subset M$ and the restriction $\pi' = \pi|_{E'}$, then we call (E', π', M') a *subbundle*. Similarly, we consider another submanifold $N \subset M$ and the preimage $G = \pi^{-1}(N)$, then $(G, \pi|_G, N)$ is called the *restricted bundle*.

Given two bundles (E, π, M) and (E', π', M') and a pair of maps $u : E \to E'$ and $f : M \to M'$. Then the pair (u, f) is called a *bundle morphism* if the diagram

$$
\begin{array}{ccc}
E & \xrightarrow{u} & E' \\
\downarrow{\pi} & & \downarrow{\pi'} \\
M & \xrightarrow{f} & M'
\end{array}
$$

commutes. The bundles are called *isomorphic as bundles* if there exists bundle morphisms (u, f) and (u^{-1}, f^{-1}). In this case (u, f) are called *bundle isomorphisms* and are the structure-preserving maps for bundles.

We may weaken this condition by not requiring it globally. Two bundles (E, π, M) and (E', π', M') are *locally isomorphic as bundles* if for every $p \in M$ there exists a neighborhood $U \in \mathcal{O}_M$ such that the restricted bundle $(\pi^{-1}(U), \pi|_{\pi^{-1}(U)}, U)$ is isomorphic to (E', π', M').

A bundle is called *(locally) trivial* if it is (locally) isomorphic to a product bundle. As an example, the cylinder is trivial and thus locally trivial and a Möbius strip is not trivial, but it is locally trivial. From now on we will disregard bundles that are not locally trivial. Then, locally, any section can be represented as a map from the base space to fiber.

> **Definition 1.22: Pullback bundle**
>
> Let (E, π, M) be a bundle, M' be a manifold and $f : M' \to M$ be a continuous map.
>
> $$\begin{array}{ccc} f^*E & \xrightarrow{\;u\;} & E \\ \downarrow{\scriptstyle \pi'} & & \downarrow{\scriptstyle \pi} \\ M' & \xrightarrow{\;f\;} & M \end{array}$$
>
> Let $f^*E = \{(m', e) \in M' \times E : \pi(e) = f(m')\} \subset M' \times E$ equipped with the subspace topology, define the map $u : f^*E \to E$ by $(m', e) \mapsto e$ and the projection $\pi' : f^*E \to M'$ by $(m', e) \mapsto m'$ such that the diagram above commutes. The bundle (f^*E, π', M') is called the *pullback bundle by f* or the *bundle induced by f*.

Given a bundle (E, π, M) with section $\sigma : M \to E$ and a continuous map $f : M' \to M$, where M' is a manifold, we may define the pullback section $f^*\sigma : M' \to f^*E$ on the bundle induced by f by the map $m' \mapsto (m', \sigma \circ f(m'))$. Since $\pi \circ \sigma = \mathrm{id}_M$, it is clear the image of the pullback section is contained in f^*E, so the map is well-defined.

1.7.2 Atlas

Since a topological manifold is locally Euclidean, for any point there is a neighborhood homeomorphic to Euclidean space. That is, there exists a homeomorphism from one such neighborhood to Euclidean space. Moreover, the collection of all such neighborhoods must cover the manifold. These remarks motivate the Definitions 1.23 and 1.24.

> **Definition 1.23: Chart of a manifold**
>
> Let $(M, \mathcal{O}_M)$ be a topological manifold of dimension n. Then a pair (U, x) where $U \in \mathcal{O}_M$ and $x : U \to x(U) \subset \mathbb{R}^n$ is called a *chart* of the manifold.
> The component functions of x, the maps $x^i : U \to \mathbb{R}$ defined by $p \mapsto \mathrm{proj}_i(x(p))$, are called the *coordinates of the point $p \in U$ with respect to the chart (U, x)*.

> **Definition 1.24: Atlas of a manifold**
>
> Let $(M, \mathcal{O}_M)$ be a topological manifold. The *atlas* $\mathscr{A} = \{(U_\alpha, x_\alpha)\}_{\alpha \in J}$ is a family of charts of the manifold such that $\bigcup_{\alpha \in J} U_\alpha = M$.

It is easy to see that the domains of different charts may overlap. Let (U, x) and (V, y) be charts of a manifold $(M, \mathcal{O}_M)$ with $U \cap V \neq \emptyset$. Since $U \cap V \in \mathcal{O}_M$ is a non-empty open set, the pairs $(U \cap V, x)$ and $(U \cap V, y)$ are charts on the manifold.

$$\begin{array}{ccc} & U \cap V & \\ {\scriptstyle x}\swarrow & & \searrow{\scriptstyle y} \\ x(U \cap V) & \xrightarrow{\;y \circ x^{-1}\;} & y(U \cap V) \end{array}$$

As the maps x and y are bijections, the map $y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$ is well defined and it is called the *chart transition map*. By Theorem 1.2, the transition map is a homeomorphism.

The following definitions will seem redundant, but will serve as a framework of definitions as we require more and more structure on manifolds. Namely, later on we will replace the continuity requirement with a differentiability class.

> **Definition 1.25: C^0-compatible charts**
>
> Two charts (U,x) and (V,y) of a n-dimensional manifold $(M, \mathcal{O}_M)$ are C^0-*compatible* if either
>
> (a) $U \cap V = \emptyset$; or
>
> (b) $U \cap V \neq \emptyset$ and the transition map $y \circ x^{-1}$ is continuous as a map $\mathbb{R}^n \to \mathbb{R}^n$.

As discussed above, it is clear that any two charts on a manifold are C^0-compatible. However, as we can study, for example, the differentiability class of the transition map as a map $\mathbb{R}^n \to \mathbb{R}^n$, we may not conclude any such thing from the chart maps before adding structure to the manifold.

> **Definition 1.26: C^0-atlas**
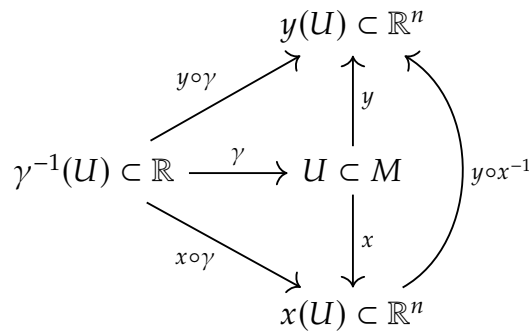>
> A C^0-*atlas* $\mathscr{A}$ is an atlas whose charts are pairwise C^0-compatible.

> **Definition 1.27: Maximal C^0-atlas**
>
> A C^0-atlas $\mathscr{A}$ is *maximal* if any chart (U,x) that is C^0-compatible with any $(V,y) \in \mathscr{A}$ is already contained in $\mathscr{A}$.

Even though every atlas is a C^0-atlas, not every atlas is maximal. As an example we consider $(M, \mathcal{O}_M)$ as the real line equipped with the standard topology. Then, the family $\{(\mathbb{R}, \mathrm{id}_{\mathbb{R}})\}$ is an atlas, but the chart $((-\infty, 0), \mathrm{id}_{\mathbb{R}})$ is not contained in it.

With an atlas and charts, one may study objects on an n-dimensional topological manifold $(M, \mathcal{O}_M)$ from different points of view. As an example, we consider a curve $\gamma : \mathbb{R} \to M$ and ask ourselves whether the curve is continuous. Clearly, by definition, we can check whether the preimage of open sets in M are open. From another perspective, we may employ charts.

$$
\begin{array}{ccc}
 & y(U) \subset \mathbb{R}^n & \\
 \overset{y \circ \gamma}{\nearrow} & \uparrow y & \nwarrow \\
\gamma^{-1}(U) \subset \mathbb{R} \xrightarrow{\ \gamma\ } U \subset M & & \Big) y \circ x^{-1} \\
 \underset{x \circ \gamma}{\searrow} & \downarrow x & \\
 & x(U) \subset \mathbb{R}^n &
\end{array}
$$

Let $U \subset M$ be an open set that contains the image of the curve, and let (U,x) be a chart. The *expression* of γ in this chart is the map $x \circ \gamma : \gamma^{-1}(U) \subset \mathbb{R} \to x(U) \subset \mathbb{R}^n$. Then, the curve is continuous if its expression is continuous as a function from $\mathbb{R} \to \mathbb{R}^n$, that is, if its components are continuous real-valued functions of a single variable. Additionally, if (U,y) is another chart, then the expression of the curve in this chart $y \circ \gamma : \gamma^{-1}(U) \to y(U)$ is continuous if and only if the expression in the other chart $x \circ \gamma$ is continuous, because the chart transition map $y \circ x^{-1} : x(U) \to y(U)$ is continuous. In this perspective, it is possible to

ignore altogether the inner workings of the manifolds and use only the coordinate systems given by the charts.

Analogously, a map $\phi : M \to N$, where (M, O_M) is an m-dimensional topological manifold and (N, O_N) is an n-dimensional topological manifold, is continuous if preimages of open sets in N are open sets in M. By employing charts (U, x) on M and (V, y) on N, the expression of the map is $y \circ \phi \circ x^{-1} : x(U) \subset \mathbb{R}^n \to y(V) \subset \mathbb{R}^n$, and its continuity can be determined as a function of $\mathbb{R}^m \to \mathbb{R}^n$.

$$
\begin{array}{ccc}
\tilde{x}(U) \subset \mathbb{R}^m & \xrightarrow{\tilde{y} \circ \phi \circ \tilde{x}^{-1}} & \tilde{y}(V) \subset \mathbb{R}^n \\
\uparrow{\scriptstyle \tilde{x}} & & \uparrow{\scriptstyle \tilde{y}} \\
U \subset M & \xrightarrow{\phi} & V \subset N \\
\downarrow{\scriptstyle x} & & \downarrow{\scriptstyle y} \\
x(U) \subset \mathbb{R}^m & \xrightarrow{y \circ \phi \circ x^{-1}} & y(V) \subset \mathbb{R}^n
\end{array}
$$

with $\tilde{x} \circ x^{-1}$ and $\tilde{y} \circ y^{-1}$.

Taking another pair of charts $(U, \tilde{x})$ and $(V, \tilde{y})$, it follows from the continuity of the chart transition maps that the expression of ϕ in these charts $\tilde{y} \circ \phi \circ \tilde{x}^{-1} : \tilde{x}(U) \subset \mathbb{R}^n \to \tilde{y}(V) \subset \mathbb{R}^n$ is a continuous function if and only if the expression $y \circ \phi \circ x^{-1}$ is continuous.

Chapter 2

Differentiable Manifolds

2.1 Smooth Atlas and Differentiable Manifolds

Even though for topological manifolds the atlas was mostly redundant, although useful, we may add a differentiable structure to a topological manifold by refining the maximal C^0-atlas. The following definitions are almost the same as Definitions 1.25 to 1.27, substituting the continuity trivial requirement to the differentiability class C^k, with $k \geq 1$.

Definition 2.1: Differentiability class on Euclidean spaces

A function $f : U \subset \mathbb{R}^n \to \mathbb{R}$ defined on an open set $U \subset \mathbb{R}^n$ (with respect to the standard topology) is of *differentiability class C^k on U* if all partial derivatives $\partial_\alpha f$ on U exist and are continuous, where α is a multi-index with $|\alpha| \leq k$. We denote the set of all functions from U to $\mathbb{R}$ that are of differentiability class C^k on U by $C^k(U)$ and we say $f \in C^k(U)$ if f is of that differentiability class.

Similarly, a function $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is of *differentiability class C^k on U* if its component functions $f_i = \pi_i \circ f$ are of that differentiability class, where $\pi_i : \mathbb{R}^m \to \mathbb{R}$ are the projections $(x^1, \dots, x^m) \mapsto x^i$. Likewise, $f \in C^k(U, \mathbb{R}^m)$ if $f : U \to \mathbb{R}^m$ if f is of differentiability class C^k on U.

Analogously, we may change the differentiability class C^k to the smoothness condition C^∞ or to real C^ω or complex analytic functions.

Definition 2.2: C^k-compatible charts

Let $(M, \mathcal{O}_M)$ be an n-dimensional topological manifold. Two charts (U, x) and (V, y) are C^k-compatible if

(a) $U \cap V = \emptyset$; or

(b) $U \cap V \neq \emptyset$ and the transition map $y \circ x^{-1}$ is of class C^k as a map $\mathbb{R}^n \to \mathbb{R}^n$.

Definition 2.3: C^k-atlas

A C^k-*compatible atlas* $\mathscr{A}$ is an atlas whose charts are pairwise C^k-compatible.

> **Definition 2.4: Maximal C^k-atlas**
>
> A C^k-atlas $\mathscr{A}$ is *maximal* if any chart (U, x) that is C^k-compatible with any $(V, y) \in \mathscr{A}$ is already contained in $\mathscr{A}$.

We now state a theorem [Hir12] that allows us to consider either maximal C^0-atlas or a maximal smooth atlas. That is, the construction of a maximal C^1-atlas is not weaker than the construction of a maximal C^∞-atlas.

> **Theorem 2.1: Any maximal differentiable atlas contains a smooth atlas**
>
> Any maximal C^k-atlas with $k \geq 1$ contains a smooth atlas. And two maximal C^k-atlases that contain the same smooth atlas are already identical.

> **Definition 2.5: C^k-manifold**
>
> A C^k-manifold is a triple $(M, O_M, \mathscr{A}_M)$ when (M, O_M) is a topological manifold and $\mathscr{A}_M$ is a maximal C^k-atlas.

Remark 2.1. *A given topological manifold can carry different incompatible atlases. As an example, take the manifold as the real line equipped with the standard topology. We consider two atlases $\mathscr{A}_1 = \{(\mathbb{R}, \mathrm{id}_{\mathbb{R}})\}$ and $\mathscr{A}_2 = \{(\mathbb{R}, x)\}$ where $x : \mathbb{R} \to \mathbb{R}$ is the map $p \mapsto p^{\frac{1}{3}}$. Clearly, $\mathscr{A}_1$ can be completed to be a maximal smooth atlas. The other atlas is a smooth atlas, as there is only one chart, so the chart transition map is the identity map, which is smooth. As such, it is also possible to complete this atlas to a maximal smooth atlas. We observe however that the atlas $\mathscr{A}_1 \cup \mathscr{A}_2$ is not even C^1-compatible, as the transition map is not differentiable at $p = 0$.*

> **Definition 2.6: Differentiable map between manifolds**
>
> Let $\phi : M \to N$ be a map, where $(M, O_M, \mathscr{A}_M)$ and $(N, O_N, \mathscr{A}_N)$ are C^k-manifolds of dimensions m and n respectively.
>
> $$\begin{array}{ccc} U \subset M & \xrightarrow{\phi} & V \subset N \\ \downarrow{\scriptstyle x} & & \downarrow{\scriptstyle y} \\ x(U) \subset \mathbb{R}^m & \xrightarrow{y \circ \phi \circ x^{-1}} & y(V) \subset \mathbb{R}^n \end{array}$$
>
> The map ϕ is *differentiable* at $p \in M$ if there exists charts $(U, x) \in \mathscr{A}_M$ and $(V, y) \in \mathscr{A}_N$, where U and V are neighborhoods of p and $\phi(p)$, such that the expression of ϕ in these charts, that is the map $y \circ \phi \circ x^{-1} : x(U) \subset \mathbb{R}^m \to y(V) \subset \mathbb{R}^n$ is a C^k map from $\mathbb{R}^m \to \mathbb{R}^n$.

Although the definition relies on the existence of charts, we must show the differentiability of a map between manifolds is independent of the choice of charts. Without loss of generality, we suppose there are another pair of charts $(U, \tilde{x}) \in \mathscr{A}_M$ and $(V, \tilde{y}) \in \mathscr{A}_N$.

$$\tilde{x}(U) \subset \mathbb{R}^m \xrightarrow{\ \tilde{y}\circ\phi\circ\tilde{x}^{-1}\ } \tilde{y}(V) \subset \mathbb{R}^n$$

Because the atlases are C^k-compatible, the chart transition maps $\tilde{x} \circ x^{-1}$ and $\tilde{y} \circ y^{-1}$ are C^k maps. Therefore, the expression $\tilde{y} \circ \phi \circ \tilde{x}^{-1} : \tilde{x}(U) \to \tilde{y}(V)$ is a C^k map if and only if $y \circ \phi \circ x^{-1} : x(U) \to y(V)$ is a C^k. This shows the definition is independent of the choice of charts.

We now define the maps that preserve the differentiable structure on manifolds.

Definition 2.7: Diffeomorphism

The map $\phi : M \to N$ is a *diffeomorphism* if it is is bijective and the maps ϕ and ϕ^{-1} are smooth.

Definition 2.8: Diffeomorphic manifolds

Two manifolds $(M, \mathcal{O}_M, \mathscr{A}_M)$ and $(N, \mathcal{O}_N, \mathscr{A}_N)$ are *diffeomorphic* if there exists a diffeomorphism between them.

It is custom to regard diffeomorphic manifolds to be the same up to diffeomorphism. A natural question arises: how many different differentiable structures can one add on a given n-dimensional topological manifold up to diffeomorphism? The answer differs for the dimension of the manifold.

(a) The case $1 \leq n \leq 3$. Radó-Moise theorems. There is a unique smooth manifold one can construct of a given topological manifold.

(b) The case $n = 4$. Depending on the structure of the topological manifold, there are possibly uncountably many different smooth manifolds.

(c) The case $n > 4$. Surgery theory (1960s). For compact manifolds, there are finitely many smooth manifolds one can make from a given topological manifold.

2.2 Tensors over a field

Before defining tensors, we first review vector spaces.

Definition 2.9: Field

A *field* $(\mathbb{K}, +, \cdot)$ is a set $\mathbb{K}$ equipped with two maps $+, \cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ called addition and multiplication that satisfy

(a) Associativity of addition and multiplication: For all $a, b, c \in \mathbb{K}$, $a + (b + c) = (a + b) + c$ and $a \cdot (b \cdot c) = (a \cdot b) \cdot c$;

(b) Commutativity of addition and multiplication: For all $a, b \in \mathbb{K}$, $a + b = b + a$ and

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

(c) Additive and multiplicative identity: There exists two distinct elements 0 and 1 in $\mathbb{K}$ such that for all $a \in \mathbb{K}$, $a + 0 = a$ and $a \cdot 1 = a$;

(d) Additive inverse: For every $a \in \mathbb{K}$ there exists an element in $-a \in \mathbb{K}$, called the additive inverse of a, such that $a + (-a) = 0$;

(e) Multiplicative inverse: For every $a \in \mathbb{K} \setminus \{0\}$, there exists an element in $a^{-1} \in \mathbb{K}$, called the multiplicative inverse of a, such that $a \cdot a^{-1} = 1$; and

(f) Distributivity of multiplication over addition: For all $a, b, c \in \mathbb{K}$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

Usually the multiplication $a \cdot b$ is denoted by ab.

Remark 2.2. *A field is a group under addition with 0 as the additive identity, and the nonzero elements are a group under multiplication with 1 as the multiplicative identity.*

Definition 2.10: Vector space over a field

A *vector space* $(V, +, \cdot)$ *over a field* $\mathbb{K}$ is a set V equipped with two maps $+ : V \times V \to V$, called vector addition, and $\cdot : \mathbb{K} \times V \to V$, called scalar multiplication, which satisfy

(a) Associativity of vector addition: For all $u, v, w \in V$, $u + (v + w) = (u + v) + w$;

(b) Commutativity of vector addition: For all $u, v \in V$, $u + v = v + u$;

(c) Identity element of vector addition: There exists an element $0 \in V$, called the zero vector, such that $v + 0 = v$ for all $v \in V$.

(d) Additive inverse: For every $v \in V$ there exists an element in $-v \in V$, called the additive inverse of v, such that $v + (-v) = 0$;

(e) Compatibility of scalar multiplication with field multiplication: For every $a, b \in \mathbb{K}$ and $v \in V$, $a \cdot (b \cdot v) = (ab) \cdot v$;

(f) Identity element of scalar multiplication: For all $v \in V$, $1 \cdot v = v$, where 1 is the multiplicative identity of $\mathbb{K}$;

(g) Distributivity of scalar multiplication with respect to vector addition: For all $u, v \in V$ and $a \in \mathbb{K}$, $a \cdot (u + v) = (a \cdot u) + (a \cdot v)$;

(h) Distributivity of scalar multiplication with respect to field addition: For all $a, b \in \mathbb{K}$ and $v \in V$, $(a + b) \cdot v = (a \cdot v) + (b \cdot v)$.

Usually the scalar multiplication $a \cdot v$ is denoted by av and it is clear from context that it is the scalar multiplication. A vector space over a field $\mathbb{K}$ may also be referred to a $\mathbb{K}$-vector space.

Remark 2.3. *It is easy to verify the field $\mathbb{K}$ is a vector space over $\mathbb{K}$.*

Remark 2.4. *A vector space is an abelian additive group under vector addition, with the extra structure of the scalar multiplication.*

> **Definition 2.11: Vector subspace**
>
> A subset $U \subset V$ is a *vector subspace* if the vector addition and scalar multiplication are closed in U. That is, for all $u, u_1, u_2 \in U$ and $\lambda \in \mathbb{K}$, we have $u_1 +|_U u_2 \in U$ and $\lambda \cdot|_U u \in U$.

> **Definition 2.12: Linear map**
>
> Let V, W be vector spaces over a field $\mathbb{K}$. Then a map $f : V \to W$ is a *linear map* if for all $v, v_1, v_2 \in V$ and $\lambda \in \mathbb{K}$, it satisfies
>
>   (a) $f(v_1 + v_2) = f(v_1) + f(v_2)$; and
>
>   (b) $f(\lambda v) = \lambda f(v)$.
>
> As a shorthand, we denote $f : V \xrightarrow{\sim} M$ if f is a linear map. The map may also be referred to a $\mathbb{K}$-linear map, if one wants to specify the underlying field of a vector space.

> **Definition 2.13: Vector space isomorphism**
>
> A *vector space isomorphism* is a bijective linear map $f : V \xrightarrow{\sim} W$ from vector spaces V, W over a field $\mathbb{K}$. If such a map exists, V and W are *isomorphic vector spaces*.

> **Definition 2.14: Vector space homomorphisms**
>
> The set of all linear maps between vector spaces V, W over a field $\mathbb{K}$ is denoted by $\mathrm{Hom}_{\mathbb{K}}(V, W)$, called the *vector space homomorphisms*.

> **Proposition 2.1: The set of vector space homomorphisms has a canonical a vector space structure**
>
> Defining the operations
>
>   (a) $+ : \mathrm{Hom}_{\mathbb{K}}(V, W) \times \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Hom}_{\mathbb{K}}(V, W)$ by the map $(f, g) \mapsto f + g$, where $f + g : V \xrightarrow{\sim} W$ is defined by $(f + g)(v) = f(v) + g(v)$ for all $v \in V$, and
>
>   (b) $\cdot : \mathbb{K} \times \mathrm{Hom}_{\mathbb{K}}(V, W) \to \mathrm{Hom}_{\mathbb{K}}(V, W)$ by the map $(\lambda, f) \mapsto \lambda f$, where $\lambda f : V \xrightarrow{\sim} W$ is defined by $(\lambda f)(v) = \lambda \cdot (f(v))$,
>
> Then $(\mathrm{Hom}_{\mathbb{K}}(V, W), +, \cdot)$ is a vector space.

Proof. We check these operations are indeed linear functions. For all $u, v \in V$ and $\lambda \in \mathbb{K}$, we have

$$\begin{aligned}
(f + g)(u + \lambda v) &= f(u + \lambda v) + g(u + \lambda v) \\
&= f(u) + \lambda f(v) + g(u) + \lambda g(v) \\
&= (f + g)(u) + \lambda (f + g)(v),
\end{aligned}$$

and

$$
\begin{aligned}
(\mu h)(u + \lambda v) &= \mu \cdot (h(u + \lambda v)) \\
&= \mu \cdot (h(u) + \lambda h(v)) \\
&= (\mu h)(u) + (\mu \lambda) h(v) \\
&= (\mu h)(u) + \lambda(\mu h(v))(u),
\end{aligned}
$$

where in this last step the commutativity of the field multiplication was used. That is, these operations are indeed closed in $\mathrm{Hom}_{\mathbb{K}}(V, W)$. The vector space axioms are then verified by computing the linear maps of $\mathrm{Hom}_{\mathbb{K}}(V, W)$ on arbitrary vectors of V and using the vector space axioms of W. $\qquad\square$

Definition 2.15: Kernel of a linear map

Let V, W be vector spaces over a field $\mathbb{K}$ and let $T \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. The *kernel* of the linear map T is the set $\ker T = \{v \in V : T(v) = 0\}$.

Proposition 2.2: Kernel and image of a linear map are vector subspaces

Let V, W be vector spaces over a field $\mathbb{K}$ and let $T \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. Then the kernel and the image of T are vector subspaces of V and W, respectively.

Proof. Let $u, v \in \ker T$ and let $\lambda \in \mathbb{K}$. Then, by linearity, $T(u + \lambda v) = T(u) + \lambda T(v) = 0$, therefore $u + \lambda v \in \ker T$. This shows the kernel is a vector subspace of V.

Let $x, y \in T(V)$. Then, there exists $u, v \in V$ such that $T(u) = x$ and $T(v) = y$. By linearity, if $\mu \in \mathbb{K}$, we have $T(u + \mu v) = T(u) + \mu T(v) = x + \mu y$. This shows the image is a vector subspace of W. $\qquad\square$

We now prove a result that will be needed in the next section.

Lemma 2.1: Trivial kernel and injective map

Let V, W be vector spaces over a field $\mathbb{K}$ and let $T \in \mathrm{Hom}_{\mathbb{K}}(V, W)$ be a linear map. The map T is injective if and only if the kernel is the trivial vector subspace, that is $\ker T = \{0\} \subset V$.

Proof. Suppose the map is injective. Then, for all $u, v \in V$, we have $T(u) = T(v) \implies u = v$. Clearly $T(0) = 0$, so $T(v) = 0 = T(0) \implies v = 0$ for all $v \in V$. It follows that $\ker T = \{0\}$.

Suppose $\ker T = \{0\}$. Suppose $T(u) = T(v)$ for some $u, v \in V$. Then $T(u) - T(v) = T(u - v) = 0$. It follows that $u - v \in \ker T$, so $u = v$. Then T is injective. $\qquad\square$

Definition 2.16: Endomorphism

Let V be a vector space. An *endomorphism* is a map $f : V \xrightarrow{\sim} V$ and it is an *automorphism* if it is a vector space isomorphism. The set of all endomorphisms on V is denoted by $\mathrm{End}(V)$, while the set of all automorphisms on V is denoted by $\mathrm{Aut}(V)$.

Remark 2.5. *Clearly,* $\mathrm{End}(V) = \mathrm{Hom}_{\mathbb{K}}(V, V)$ *and* $\mathrm{Aut}(V) \subset \mathrm{End}(V)$*. By Proposition 2.1, the set of endomorphisms on V also has a canonical vector space structure. However, the set of automorphisms on V is not a vector subspace of the vector space of endomorphisms.*

Definition 2.17: Dual space

The vector space $V^* = \text{Hom}(V, \mathbb{K})$ is the *dual* vector space to V.

Remark 2.6. *Regarding the vector space V as a base vector space, elements of V may be called vectors, while the dual vector space V^* elements may be called covectors or linear functionals on V.*

2.2.1 Basis and Dimension

Definition 2.18: Hamel Basis

Let V be a vector space. Then a subset $\mathcal{B} \subset V$ is a *(Hamel) basis* if

(a) every finite subset $\{b_1, \ldots b_N\} \subset \mathcal{B}$ is *linear independent*, that is,

$$\sum_{i=1}^{N} \lambda^i b_i = 0 \implies \lambda^1 = \cdots = \lambda^N = 0;$$

(b) for every vector $v \in V$, there exists a finite subset $\{b_1, \ldots, b_M\} \subset \mathcal{B}$ and a subset $\{v^1, \ldots, v^M\} \subset \mathbb{K}$ such that v is a *linear combination* of $\{b_1, \ldots, b_M\}$, that is

$$v = \sum_{i=1}^{M} v^i b_i.$$

We say $\mathcal{B}$ *spans* V or $\mathcal{B}$ is a *generating set* of V.

We assume a result that is proven later in a more general setting, namely Existence of Hamel basis.

Definition 2.19: Vector space dimension

Let V be a vector space with a basis $\mathcal{B}$. The *dimension* of V, denoted by $\dim V$, is equal to the cardinality of $\mathcal{B}$. If a basis has a finite number of elements, we say V is *finite dimensional*.

Remark 2.7. *It is not immediate that this is well-defined. In fact, this is motivated by a theorem that states any two different basis for a vector space have the same cardinality.*

In the following, we will show the relations between a finite-dimensional vector space to its dual and bidual spaces. To show these relations, we prove a lemma and an important theorem of linear maps.

Lemma 2.2: Dimension of a vector subspace

Let V be a finite-dimensional vector space over $\mathbb{K}$. If $U \subset V$ is a vector subspace of V, then $\dim U = \dim V$ if and only if $U = V$.

Proof. It is obvious that if $U = V$, then $\dim U = \dim V$. We now show the converse, if $\dim U = \dim V$, then $U = V$.

Let $\mathcal{B}$ be a basis for U. Then $\mathcal{B}$ is linear independent and spans U. Suppose, by contradiction, $\mathcal{B}$ is not a basis for V. This implies there exists $v \in V$ that is not a linear combination of the elements of $\mathcal{B}$. As a result, $\mathcal{B} \cup \{v\}$ is a linearly independent subset of V with $1 + \dim U > \dim V$ elements. This contradiction shows that $\mathcal{B}$ is a basis for V. $\square$

Remark 2.8. *This lemma guarantees that $U \subset V \implies \dim U \leq \dim V$, with equality implying $U = V$.*

We may now prove the theorem that relates the dimensions of the domain, kernel and image of a linear map.

> **Theorem 2.2: Rank-nullity theorem**
>
> Let V, W be vector spaces over a field $\mathbb{K}$ and let $T \in \mathrm{Hom}_{\mathbb{K}}(V, W)$. If V is finite dimensional, we have $\dim V = \dim \ker T + \dim T(V)$.

Proof. Since the kernel is a subspace of V, we have $\dim V \geq \dim \ker T$, by the lemma. We first consider $\dim V = \dim \ker T$. This implies $T(V) = \{0\} \subset W$ and the statement follows. We may now assume $\dim V > \dim \ker T$.

Let $\mathcal{B}_{\ker T} = \{e_1, \ldots, e_n\}$ be a basis for $\ker T$, where $n = \dim \ker T$ it the nullity of T. We may complete this basis such that the resulting set is a basis $\mathcal{B} = \mathcal{B}_{\ker T} \cup \{e_{n+1}, \ldots, e_{n+m}\}$ for V, with $\dim V = m + n$ and $m \geq 1$. In particular, for all $v \in V$ there exists a family $\{v^i\}_{i=1}^{n+m}$ in $\mathbb{K}$ such that

$$v = \sum_{i=1}^{n+m} v^i e_i,$$

since $\mathcal{B}$ is a generating set for V.

We consider $u \in T(V)$. Then there exists $v \in V$ such that $T(v) = u$, that is

$$u = T\left(\sum_{i=1}^{n+m} v^i e_i\right)$$

$$= \sum_{i=1}^{n+m} v^i T(e_i)$$

$$= \sum_{i=1}^{m} v^{i+n} T(e_{i+n}),$$

since $e_i \in \ker T$ for $i \leq n$. Equivalently, $\mathcal{B}_{T(V)} = \{T(e_{n+1}), \ldots, T(e_{n+m})\} \subset T(V)$ is a generating set for $T(V)$.

Consider $0 \in T(V) \subset W$. Since $\mathcal{B}_{T(V)}$ spans $T(V)$, there exists a family $\{\lambda^i\}_{i=1}^{m}$ in $\mathbb{K}$ such that

$$\sum_{i=1}^{m} \lambda^i T(e_{i+n}) = 0.$$

By linearity, this implies $w = \sum_{i=1}^{m} \lambda^i e_{i+n} \in \ker T$. Since e_{i+n} is not a linear combination of $\mathcal{B}_{\ker T}$, we must have $\lambda^i = 0$, for all $1 \leq i \leq m$. Therefore, $\mathcal{B}_{T(V)}$ is linearly independent and it is a basis of $T(V)$ with m elements and the theorem follows. $\qquad \square$

Remark 2.9. *The* nullity *and* rank *of a linear map are the dimensions of its kernel and image. We note* $\dim \ker T$ *and* $\dim T(V)$ *are well-defined due to* Proposition 2.2.

Remark 2.10. *It is not immediate it is possible to complete a basis given a linearly independent subset, this is related to* Existence of Hamel basis.

Corollary 2.1. *If both vector spaces are finite-dimensional with* $\dim V = \dim W$ *and if T is injective, then T is an isomorphism.*

Proof. Since T is one-to-one, its nullity is zero by Lemma 2.1. Then, by the Rank-nullity theorem, we have $\dim T(V) = \dim V = \dim W$. Since $T(V) \subset W$ is a vector subspace of W, it follows from Lemma 2.2 that $T(V) = W$, that is, T is onto. $\qquad\square$

> **Theorem 2.3: Dual vector space dimension**
>
> Let V be a finite-dimensional vector space over a field $\mathbb{K}$. Then $\dim V = \dim V^*$.

Proof. Let $n = \dim V$ and let $\mathcal{B} = \{e_1, \ldots, e_n\}$ be a basis for V. Then, define $\mathcal{B}^* = \{\epsilon^1, \ldots, \epsilon^n\}$, a subset of maps from $V \to \mathbb{K}$, by letting $\epsilon^i\left(\sum_{j=1}^n c^j e_j\right) = c^i$, where $c^i \in \mathbb{K}$ for $i = 1, \ldots, n$.

First, we show that ϵ^i is indeed an element of V^*. Let $x, y \in V$ with $x = \sum_{i=1}^n x^i e_i$ and $y = \sum_{i=1}^n y^i e_i$ and $x^i, y^i \in \mathbb{K}$. By definition of the maps ϵ^i, we have $\epsilon^i(x) = x^i$ and $\epsilon^i(y) = y^i$, for $i = 1, \ldots, n$. Let $\lambda \in \mathbb{K}$, then

$$\epsilon^i(x + \lambda y) = \epsilon^i\left[\sum_{j=1}^n (x^j + \lambda y^j)e_j\right]$$
$$= x^i + \lambda y^i$$
$$= \epsilon^i(x) + \lambda \epsilon^i(y),$$

that is, ϵ^i is a linear map. Therefore, $\mathcal{B}^* \subset V^*$.

We consider the linear combination $\omega = \sum_{i=1}^n \omega_i \epsilon^i \in V^*$, with $\omega_i \in \mathbb{K}$. The dual vector ω is the zero dual vector if $\omega(v) = 0$ for all v. We may choose v as each element of the basis $\mathcal{B}$, that is, if $v = e_j$, then $v^i = \delta_j^i$, where

$$\delta_j^i = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$$

is the *Kronecker delta*, for $j = 1, \ldots, n$. As consequence, we have $\omega_j = 0$ for $j = 1, \ldots, n$, therefore $\mathcal{B}^*$ is linearly independent.

We consider a dual vector $\varphi \in V^*$. Then, for all $u = \sum_{i=1}^n u^i e_i \in V$, we have $\epsilon^i(u) = u^i$ and

$$\varphi(u) = \varphi\left(\sum_{i=1}^n u^i e_i\right)$$
$$= \varphi\left(\sum_{i=1}^n \epsilon^i(u)e_i\right)$$
$$= \sum_{i=1}^n \varphi(e_i)\epsilon^i(u),$$

that is, $\varphi = \sum_{i=1}^n \varphi(e_i)\epsilon^i$. Then $\mathcal{B}^*$ is a generating set of V^*.

We have shown $\mathcal{B}^*$ is a basis for V^*, therefore $\dim V^* = n$ and the theorem follows. $\qquad\square$

Remark 2.11. *The construction used in this proof will be used extensively: in lieu of making arbitrary choices of basis for both V and V^*, only the choice of basis in V is needed, and we have an induced basis on V^*, henceforth named* dual basis.

Remark 2.12. *The proof that two finite-dimensional vector spaces are isomorphic if and only if their dimensions are equal is very similar.*

> **Theorem 2.4: Bidual vector space canonical linear isomorphism**
>
> Let V be a finite-dimensional vector space over a field $\mathbb{K}$. Then there exists a canonical linear isomorphism from V to the bidual vector space $(V^*)^*$.

Proof. We remind ourselves the bidual vector space is the set of linear maps from the dual space to the field vector space, that is

$$(V^*)^* = \left\{ \phi : V^* \to \mathbb{K} \text{ such that } \phi \text{ is linear} \right\}.$$

We consider the map

$$\psi : V \to (V^*)^*$$
$$v \mapsto \psi(v),$$

where $\psi(v) \in (V^*)^*$ is the linear map

$$\psi(v) : V^* \xrightarrow{\sim} \mathbb{K}$$
$$\omega \mapsto \omega(v).$$

Since this definition requires no additional structure, this map is canonically defined: no choice of basis in what follows taints this.

First, we show the map is linear. Let $u, v \in V$ and $\lambda \in \mathbb{K}$, then we let $\psi(u + \lambda v)$ act on a dual vector $\omega \in V^*$

$$\psi(u + \lambda v)(\omega) = \omega(u + \lambda v)$$
$$= \omega(u) + \lambda \omega(v)$$
$$= \psi(u)(\omega) + \lambda \psi(v)(\omega),$$

that is, $\psi(u + \lambda v) = \psi(u) + \lambda \psi(v)$. Therefore, $\psi : V \xrightarrow{\sim} (V^*)^*$ is a linear map.

We now show this map is injective. Let $v \in V$ such that $\psi(v)$ is the null map. Suppose, by contradiction, $v \neq 0$. The subset $\{v\} \subset V$ is linearly independent and we may complete this set to be a basis for V. Then, let $v^* \in V^*$ be the element in the dual basis such that $v^*(v) = 1$. We let the null map $\psi(v)$ act on v^*, arriving at at contradiction. This shows $v = 0$, that is, ψ is injective.

Noting $(V^*)^*$ is the dual space of V^*, we have $\dim V = \dim V^* = \dim(V^*)^*$ by Theorem 2.3. By Theorem 2.2, it follows that ψ is a bijection. $\qquad\square$

Remark 2.13. *Since for a finite-dimensional vector space V there is a natural isomorphism from the vector space to its bidual, if $v \in V$ and $\omega \in V^*$, one may write $\omega(v) = v(\omega)$, where $v(\omega) = \psi(v)(\omega)$.*

2.2.2 Tensor spaces

With the notion of vector spaces and dual vector spaces, we may construct functions of multiple vector variables, known as tensors.

Definition 2.20: Tensors on a vector space

Let V be a vector space over a field $\mathbb{K}$. An element of the vector space defined by the set

$$\underbrace{V \otimes \cdots \otimes V}_{r \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{s \text{ times}} = \left\{ \underbrace{V^* \times \cdots \times V^*}_{r \text{ times}} \times \underbrace{V \times \cdots \times V}_{s \text{ times}} \to \mathbb{K} : T \text{ is multilinear} \right\}$$

is a (r,s)-*tensor on the vector space* V. A map is *multilinear* if it is linear on each argument. The pair (r,s) is called the *valence* of the tensor. As a shorthand, we denote the set of (r,s)-tensors on the vector space V as $T^r_s V$.

Just as $\operatorname{Hom}_{\mathbb{K}}(V, W)$ had a canonical vector space structure, we show a similar result in Proposition 2.3.

Proposition 2.3: Tensors have a canonical vector space structure

Let V be a vector space over a field $\mathbb{K}$. The set $T^r_s V$ together with the operations

(a) $+ : T^r_s V \times T^r_s V \to T^r_s V$ defined by $(T, S) \mapsto T + S$ where

$$(T+S)(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s) = T(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s) + S(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s),$$

(b) $\cdot : \mathbb{K} \times T^r_s V \to T^r_s V$ defined by $(\lambda, T) \mapsto \lambda T$ where

$$(\lambda T)(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s) = \lambda \cdot (T(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s)),$$

is a vector space.

Proof. We verify the well-definitions of the operations above. Let $T, S \in T^r_s$ and $\lambda \in \mathbb{K}$. We consider a family $\{\omega_i\}^r_{i=1}$ of r covectors, a family of $\{v_i\}^s_{i=1}$ vectors. Without loss of generality, we verify the linearity on an arbitrary argument, say the first argument and we abbreviate $T(\omega_1, \ldots)$ to denote $T(\omega_1, \ldots, \omega_r, v_1, \ldots, v_s)$. Let $\sigma \in V^*$ and $\mu \in \mathbb{K}$. We have

$$\begin{aligned}(T + S)(\omega_1 + \mu\sigma, \ldots) &= T(\omega_1 + \mu\sigma, \ldots) + S(\omega_1 + \mu\sigma, \ldots)\\ &= T(\omega_1, \ldots) + S(\omega_1, \ldots) + \mu T(\sigma, \ldots) + \mu S(\sigma, \ldots)\\ &= (T + S)(\omega_1, \ldots) + \mu(T + S)(\sigma, \ldots),\end{aligned}$$

and

$$\begin{aligned}(\lambda T)(\omega_1 + \mu\sigma, \ldots) &= \lambda \cdot (T(\omega_1 + \mu\sigma, \ldots)),\\ &= \lambda \cdot (T(\omega_1, \ldots) + \mu T(\sigma, \ldots))\\ &= (\lambda T)(\omega_1, \ldots) + \mu(\lambda T)(\sigma, \ldots),\end{aligned}$$

where we have used the commutativity of field multiplication. This shows the operations yield indeed multilinear maps, and as such are well-defined. We could then verify the vector space axioms by letting (r,s)-tensors act on the families of vectors and covectors and using the vector space axioms on $\mathbb{K}$. $\square$

We may construct new tensors from given tensors on the same vector space with possibly different valences.

Definition 2.21: Tensor product

Let V be a vector space over $\mathbb{K}$. The *tensor product* is the map $\otimes : T^p_q V \times T^r_s V \to T^{p+r}_{q+s} V$ defined by

$$(T \otimes S) \left(\omega_1, \ldots, \omega_p, \omega_{p+1}, \ldots, \omega_{p+r}, v_1, \ldots, v_q, v_{q+1}, \ldots, v_{q+s} \right)$$
$$= T \left(\omega_1, \ldots, \omega_p, v_1, \ldots, v_q \right) \cdot S \left(\omega_{p+1}, \ldots, \omega_{p+r}, v_{q+1}, \ldots, v_{q+s} \right).$$

Example 2.1. *Let V be a vector space over a field $\mathbb{K}$.*

(a) *By convention, we set $T^0_0 V = \mathbb{K}$.*

(b) *$T^0_1 V = V^*$. That is, $(0, 1)$-tensors are elements of the dual vector space.*

(c) *$T^1_1 V = V \otimes V^* = \{V^* \times V \to \mathbb{K} : T \text{ is multilinear}\}$. We will show that this is isomorphic to $\mathrm{End}(V^*)$. That is, given $T \in V \otimes V^*$ we may construct $\tilde{T} \in \mathrm{End}(V^*)$ by setting $\tilde{T}(\omega) = T(\omega, \cdot)$. Similarly, given $\tilde{T} \in \mathrm{End}(V^*)$ we may reconstruct T by setting $T(\omega, v) = \tilde{T}(\omega)(v)$.*

(d) *Similarly, $T^1_1 V$ is isomorphic to $\mathrm{End}(V)$ if V is finite-dimensional, due to Theorem 2.4. In fact, one may check the map $\Psi : \mathrm{End}(V) \to T^1_1 V$ defined by $(\Psi(\phi))(\omega, v) = \omega(\phi(v))$ is linear and bijective.*

(e) *By Theorem 2.4, if V is finite-dimensional, then $T^1_0 V = \{V^* \to \mathbb{K} : T \text{ is multilinear}\} = (V^*)^*$ is isomorphic to V.*

Definition 2.22: Components of a tensor

Let $T \in T^r_s V$, where $\dim V = n$. Let $\{e_1, \ldots, e_n\}$ be a basis of V and let $\{\epsilon^1, \ldots, \epsilon^n\}$ be the dual basis of V^*. Then the *components of T with respect to the chosen basis* are

$$T^{i_1, \ldots, i_r}{}_{j_1, \ldots, j_s} = T \left(\epsilon^{i_1}, \ldots, \epsilon^{i_r}, e_{j_1}, \ldots, e_{j_r} \right) \in \mathbb{K},$$

with indices $i_k, j_l \in \{1, \ldots, n\}$ for all $k \in \{1, \ldots, r\}$ and $l \in \{1, \ldots, s\}$.

Proposition 2.4: Components determine a tensor

Let V be an n-dimensional vector space over a field $\mathbb{K}$. Let $\{e_1, \ldots, e_n\}$ be a basis of V and let $\{\epsilon^1, \ldots, \epsilon^n\}$ be the dual basis of V^*. The tensor $T \in T^r_s V$ with components $T^{i_1, \ldots, i_r}{}_{j_1, \ldots, j_s}$, with indices $i_k, j_l \in \{1, \ldots, n\}$ for all $k \in \{1, \ldots, r\}$ and $l \in \{1, \ldots, s\}$, is given by

$$T = \sum_{i_1=1}^{n} \cdots \sum_{i_s=1}^{n} \sum_{j_1=1}^{n} \cdots \sum_{j_s=1}^{n} T^{i_1, \ldots, i_r}{}_{j_1, \ldots, j_s} \, e_{i_1} \otimes \cdots \otimes e_{i_r} \otimes \epsilon^{j_1} \otimes \cdots \otimes \epsilon^{j_r}.$$

Proof. Let $\{a_i\}^r_{i=1}$ and $\{b_i\}^s_{i=1}$ be families of indices in $\{1, \ldots, n\}$. We verify the tensor T has

indeed those components by letting it act on the vector and dual basis elements:

$$
\begin{aligned}
T^{a_1,\dots,a_r}{}_{b_1,\dots,b_s} &= T(\epsilon^{a_1},\dots,\epsilon^{a_r},e_{b_1},\dots,e_{b_s}) \\
&= \sum_{i_1=1}^{n}\cdots\sum_{i_s=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_s=1}^{n} T^{i_1,\dots,i_r}{}_{j_1,\dots,j_s}\, e_{i_1}(\epsilon^{a_1})\cdot\ldots\cdot e_{i_r}(\epsilon^{a_r})\cdot \epsilon^{j_1}(e_{b_1})\cdot\ldots\cdot \epsilon^{j_r}(e_{b_r}) \\
&= \sum_{i_1=1}^{n}\cdots\sum_{i_s=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{j_s=1}^{n} T^{i_1,\dots,i_r}{}_{j_1,\dots,j_s}\, \delta^{a_1}_{i_1}\cdot\ldots\cdot \delta^{a_r}_{i_r}\cdot \delta^{j_1}_{b_1}\cdot\ldots\cdot \delta^{j_s}_{b_s} \\
&= T^{a_1,\dots,a_r}{}_{b_1,\dots,b_s},
\end{aligned}
$$

and we have recovered the desired components. $\qquad\square$

From now on, the *Einstein summation notation* will be used. In this notation, we use the convention that *basis vectors of V are labeled by lower indices* and *dual basis covectors are labeled by upper indices*, as was used in the previous definition. The summation convention is to omit the sum signs over an index whenever it appears as an upper and lower index in the same product. As an example, instead of writing $v = v^1 e_1 + \cdots + v^n e_n$, we simply write

$$
v = v^i e_i,
$$

and the summation over i from 1 to n is implied. Likewise, the expression in is simplified to

$$
T = T^{i_1,\dots,i_r}{}_{j_1,\dots,j_s}\, e_{i_1}\otimes\cdots\otimes e_{i_r}\otimes \epsilon^{j_1}\otimes \epsilon^{j_r},
$$

and the summation over all the indices are implied.

We note however that the convention only works with linear spaces and (multi)linear maps. We consider a $(1,1)$-tensor $\phi: V^* \times V \xrightarrow{\sim} K$ acting on $\omega \in V^*$ and $v \in V$, writing each step with and without the summation notation:

$$
\begin{aligned}
\phi(\omega,v) &= \phi\left(\sum_{i=1}^{n}\omega_i\epsilon^i, \sum_{j=1}^{n}v^j e_j\right) &&= \phi\left(\omega_i\epsilon^i, v^j e_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\phi\left(\omega_i\epsilon^i, v^j e_j\right) &&= \phi\left(\omega_i\epsilon^i, v^j e_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i v^j \phi\left(\epsilon^i, e_j\right) &&= \omega_i v^j \phi\left(\epsilon^i, e_j\right) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\omega_i v^j \phi^i{}_j &&= \omega_i v^j \phi^i{}_j.
\end{aligned}
$$

Notice that the first step relies on the multilinearity of ϕ, but in the summation notation it is not clear *where* the implied summations are happening. This example illustrates the use of the summation convention and serves as a warning that it is well-defined only for linear structures.

2.2.3 Change of Basis

Let V be an n-dimensional vector space over a field $\mathbb{K}$ with a basis $\mathcal{B} = \{e_1,\dots,e_n\}$ of V. We now consider another basis $\tilde{\mathcal{B}} = \{\tilde{e}_1,\dots,\tilde{e}_n\}$ of V. Since the elements of $\tilde{\mathcal{B}}$ are vectors of V,

there exists $A^i{}_j \in \mathbb{K}$ such that

$$\tilde{e}_j = A^i{}_j e_i,$$

for $j = 1, \ldots, n$. Likewise, elements of $\mathcal{B}$ can be expanded in terms of their components in the $\tilde{\mathcal{B}}$ basis, that is

$$e_j = B^i{}_j \tilde{e}_i,$$

for some $B^i{}_j \in \mathbb{K}$, for $j = 1, \ldots, n$. In this case, we must have $A^i{}_j B^j{}_k = \delta^i_k$. That is, there exists an automorphism $A : V \xrightarrow{\sim} V$ that relates the basis $\mathcal{B}$ to the basis $\tilde{\mathcal{B}}$ with components $A^i{}_j$ and its inverse $B = A^{-1} : V \xrightarrow{\sim} V$ has components $B^i{}_j$.

   We now see how the dual basis is modified under this change of basis. Let $\tilde{\epsilon}^i = C^i{}_j \epsilon^j$ and then determine the components $C^i{}_j$ with

$$\tilde{\epsilon}^i(\tilde{e}_j) = A^k{}_j C^i{}_l \epsilon^l(e_k)$$
$$\delta^i_j = A^k{}_j C^i{}_l \delta^l_k$$
$$\delta^i_j = C^i{}_k A^k{}_j,$$

which implies $C^i{}_k = B^i{}_k$. That is, $\tilde{\epsilon}^i = B^i{}_j \epsilon^j$ and $\epsilon^i = A^i{}_j \epsilon^j$.

   Let $\omega = \omega_i \epsilon^i \in V^*$ and $v = v^i e_i$ in the basis $\mathcal{B}$. We now determine the relation between the components of these objects in the basis $\tilde{\mathcal{B}}$. For the covector, we have

$$\omega_i = \omega(e_i) = \omega(B^j{}_i \tilde{e}_j) = B^j{}_i \tilde{\omega}_j,$$

and for the vector,

$$v^i = v(\epsilon^i) = \epsilon^i(v) = \epsilon^i(\tilde{v}^j \tilde{e}_j) = A^k{}_j \tilde{v}^j \epsilon^i(e_k) = A^i{}_j \tilde{v}^j.$$

We see the components of the covector ω change just like the basis vectors e_i do and the components of the vector v change as the dual basis vectors ϵ^i do. More generally, for a (r,s)-tensor, the upper indices change like vector components and lower indices like covector components:

$$T^{a_1, \ldots, a_r}{}_{b_1, \ldots, b_s} = A^{a_1}{}_{m_1} \ldots A^{a_r}{}_{m_r} B^{n_1}{}_{b_1} \ldots B^{n_s}{}_{b_s} \tilde{T}^{m_1, \ldots, m_r}{}_{n_1, \ldots, n_s}.$$

2.2.4   Determinants

Let V be a n-dimensional vector space over a field $\mathbb{K}$.

> **Definition 2.23: m-form on a vector space**
>
> An m-form is a $T^0_m V$ tensor ω that is totally anti-symmetric. That is, let π be a permutation of the permutation group S^m, then
>
> $$\omega(v_1, \ldots, v_m) = \text{sgn}(\pi) \cdot \omega(v_{\pi(1)}, \ldots, v_{\pi(m)}).$$

Remark 2.14. *In case $m = 0$, $\omega \in \mathbb{K}$ and in the case $m > d$, ω is the null tensor. The special case $m = n$ is called* volume form *and it can be shown that if ω and ω' are two non-vanishing volume forms, then there exists $\lambda \in \mathbb{K}$ such that $\omega' = \lambda \omega$.*

> **Definition 2.24: Volume form**
>
> A choice of a non-vanishing volume form ω is called a *choice of volume on V*. Let $\mathcal{F} = \{v_i\}_{i=1}^n$ be a family of n vectors in V, then
>
> $$\text{vol}(v_1, \ldots, v_n) = \omega(v_1, \ldots, v_n)$$
>
> is the *volume spanned by $\mathcal{F}$*.

Remark 2.15. *It follows from the anti-symmetries of the volume form ω that $\mathcal{F}$ is not linearly independent if and only if the volume spanned by $\mathcal{F}$ is zero. Equivalently, $\mathcal{F}$ is a basis if and only if the volume spanned by $\mathcal{F}$ is non-zero.*

> **Definition 2.25: Determinant of an endomorphism on V**
>
> Let $\phi \in \text{End}(V) \cong T_1^1 V$. The determinant of ϕ is defined as
>
> $$\det \phi = \frac{\text{vol}(\phi(e_1), \ldots, \phi(e_n))}{\text{vol}(e_1, \ldots, e_n)}$$
>
> for some choice of volume on V and some basis $\{e_1, \ldots, e_n\}$ of V.

> **Proposition 2.5: Determinant is well-defined**
>
> The determinant is well defined.

Proof. Let ω and ω' be two choices of volume. Then, there exists $\lambda \in \mathbb{K}$ such that $\omega' = \lambda\omega$. Then, for a basis $\mathcal{B} = \{e_1, \ldots, e_n\}$ and a family of n vectors $\{v_i\}_{i=1}^n$, we have

$$\frac{\omega(v_1, \ldots, v_n)}{\omega(e_1, \ldots, e_n)} = \frac{c\omega'(v_1, \ldots, v_n)}{c\omega'(e_1, \ldots, e_n)} = \frac{\omega'(v_1, \ldots, v_n)}{\omega'(e_1, \ldots, e_n)}.$$

In particular, the determinant is invariant by choice of volume.

To show independence from choice of basis, we first compute the determinant of an endomorphism $\phi : V \to V$ with components $\phi^i{}_j$ in the basis $\mathcal{B}$. We have

$$\begin{aligned}
\omega\left(\phi(e_1), \ldots, \phi(e_n)\right) &= \omega\left(\phi^{k_1}{}_1 e_{k_1}, \ldots, \phi^{k_n}{}_n e_{k_n}\right) \\
&= \phi^{k_1}{}_1 \cdot \cdots \cdot \phi^{k_n}{}_n \omega\left(e_{j_1}, \ldots, e_{j_n}\right) \\
&= \sum_{\sigma \in S_n} \text{sgn}(\sigma)\phi^{\sigma(1)}{}_1 \cdot \cdots \cdot \phi^{\sigma(n)}{}_n \omega(e_1, \ldots, e_n),
\end{aligned}$$

since the volume spanned by a linearly dependent family of vectors is zero. As a result, we have

$$\frac{\omega\left(\phi(e_1), \ldots, \phi(e_n)\right)}{\omega(e_1, \ldots, e_n)} = \sum_{\sigma \in S_n} \text{sgn}(\sigma)\phi^{\sigma(1)}{}_1 \cdot \cdots \cdot \phi^{\sigma(n)}{}_n,$$

where the choice of basis is implied by the components of the endomorphism.

Let $\tilde{\mathcal{B}} = \{\tilde{e}_1, \ldots, \tilde{e}_n\}$ be another basis. Then, there exists a linear map $A \in \text{Aut}(V)$ such that $\tilde{e}_i = A^j{}_i e_j$, and a linear map $B = A^{-1}$ such that $e_i = B^j{}_i \tilde{e}_j$ and $A^i{}_j B^j{}_k = \delta^i_k$. In the basis $\tilde{\mathcal{B}}$, the components of the map ϕ are given by

$$\tilde{\phi}^i{}_j = A^a{}_j B^i{}_b \phi^b{}_a,$$

and by the same computations done in the basis $\mathcal{B}$, we have

$$
\frac{\omega\left(\phi(\tilde{e}_1), \ldots, \phi(\tilde{e}_n)\right)}{\omega(\tilde{e}_1, \ldots, \tilde{e}_n)} = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \tilde{\phi}^{\sigma(1)}{}_1 \cdot \ldots \cdot \tilde{\phi}^{\sigma(n)}{}_n
$$

$$
= \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) A^{a_1}{}_1 B^{\sigma(1)}{}_{b_1} \phi^{b_1}{}_{a_1} \cdot \ldots \cdot A^{a_n}{}_n B^{\sigma(n)}{}_{b_n} \phi^{b_n}{}_{a_n}.
$$

TODO! $\qquad\qquad \square$

Find a way to calculate determinant of a bilinear map. Tensor density.

Remark 2.16. *Having chosen a basis, it is tempting to consider elements of vector spaces as collections of elements of $\mathbb{K}$ arranged in some matricial representation. As an example, we may choose the following convention*

$$
\omega = \omega_i \epsilon^i \in V^* \rightsquigarrow \quad \omega \doteq \begin{pmatrix} \omega^1 & \cdots & \omega^n \end{pmatrix}
$$

$$
v = v^i e_i \in V \rightsquigarrow \quad v \doteq \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix}
$$

$$
\phi = \phi^i{}_j e_i \otimes \epsilon^j \in T^1_1 V \rightsquigarrow \quad \phi \doteq \begin{pmatrix} \phi^1{}_1 & \phi^1{}_2 & \cdots & \phi^1{}_n \\ \phi^2{}_1 & \phi^2{}_2 & \cdots & \phi^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ \phi^n{}_1 & \phi^n{}_2 & \cdots & \phi^n{}_n \end{pmatrix},
$$

that is, covectors are represented by row matrices, vectors by column matrices, and the (1,1)-tensor $\phi : V \to V$ has its components $\phi^i{}_j$ arranged in a square matrix where the first (upper) index i indicates the row, and second (lower) index j, the column. First, we identify $\mathrm{End}(V)$ with $T^1_1 V$, that is $\phi(\epsilon^i, e_j) = \epsilon^i(\phi(e_j))$ by abuse of notation: in the left hand side, ϕ is understood as (1,1)-tensor and on the right hand side, as an endomorphism on V and therefore

$$
\phi^i{}_j = \phi(\epsilon^i, e_j) = \epsilon^i(\phi(e_j)) \implies \phi(e_j) = \phi^k{}_j e_k.
$$

To motivate this convention, if $\phi, \psi \in \mathrm{End}(V) \cong T^1_1$, then $\phi \circ \psi \in \mathrm{End}(V)$ and we have

$$
(\phi \circ \psi)(\epsilon^i, e_j) = \epsilon^i(\phi \circ \psi(e_j))
$$

$$
(\phi \circ \psi)^i{}_j = \epsilon^i(\phi(\psi(e_j)))
$$

$$
= \epsilon^i(\phi(\psi^k{}_j e_k))
$$

$$
= \psi^k{}_j \epsilon^i(\phi(e_k))
$$

$$
= \psi^k{}_j \phi^i{}_k
$$

$$
= \phi^i{}_k \psi^k{}_j,
$$

and this last line may be interpreted as the matrix multiplication of the matrix representations of ϕ and ψ, in this order. Similarly, if $\omega \in V^*$ and $v \in V$, then $\omega(v) = \omega_i v^i$ may be interpreted as the matrix multiplication of the row matrix that represents ω by the column matrix that represents v. Finally, we consider $\phi(\omega, v)$

$$
\phi(\omega, v) = \phi(\omega_i \epsilon^i, v^j e_j)
$$

$$
= \omega_i v^j \phi^i{}_j
$$

$$
= \omega_i \phi^i{}_j v^j,
$$

and the result may be interpreted as the matrix multiplication of a row matrix, by the square matrix and finally by the column matrix.

However, determinants TODO!

2.3   Tangent spaces to a manifold

We suppress the notation of a smooth manifold $(M, O_M, \mathscr{A}_M)$ to simply its set M, unless there are more manifolds and their underlying structure need be emphasized.

Let M be a smooth manifold. We equip the set of smooth functions from M to $\mathbb{R}$ with point-wise addition and scalar multiplication, and we claim $(C^\infty(M), +, \cdot)$ is a vector space. That is, if $f, g \in C^\infty(M)$ and $\lambda \in \mathbb{R}$, then $(f + g)(p) = f(p) + g(p)$ and $(\lambda f)(p) = \lambda \cdot f(p)$, where $p \in M$.

> **Definition 2.26: Directional derivative at a point along a curve**
>
> Let $I = (-\varepsilon, \varepsilon) \subset \mathbb{R}$ be an open interval, for some $\varepsilon > 0$. Let $\lambda : I \to M$ be a smooth curve through a point $p \in M$ such that $p = \gamma(0)$. The *directional derivative operator at the point p along the curve γ* is the linear map
>
> $$X_{\gamma,p} : C^\infty(M) \overset{\sim}{\to} \mathbb{R}$$
> $$f \mapsto (f \circ \gamma)'(0).$$
>
> In differential geometry, the operator $X_{\gamma,p}$ is usually called the *tangent vector to the curve γ at the point p*.

A precise intuition is $X_{\gamma,p}$ is the velocity of the curve γ at the point p. Given a curve $\gamma : (-\varepsilon, \varepsilon) \to M$, this notion is easily seen with a curve $\eta : \left(-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right) \to M$ with double the parameter speed $\eta(\lambda) = \gamma(2\lambda)$, for $|\lambda| < \varepsilon$, where one gets $X_{\eta,p} = 2X_{\gamma,p}$.

> **Definition 2.27: Tangent vector space at a point**
>
> The *tangent vector space at a point $p \in M$* is the set T_pM of all the directional derivatives operators at the point p along smooth curves $\gamma : I \to M$ with $\gamma(0) = p$ equipped with the addition
>
> $$+ : T_pM \times T_pM \to T_pM$$
> $$(X, Y) \mapsto X + Y$$
>
> and scalar multiplication
>
> $$\cdot : \mathbb{R} \times T_pM \to T_pM$$
> $$(\lambda, X) \mapsto \lambda X$$
>
> defined point-wise, that is, $(X + Y)f = Xf + Yf$ and $(\lambda X)f = \lambda \cdot (Xf)$ for all $f \in C^\infty(M)$.

We must now verify the claim that T_pM equipped with the above operations is indeed a vector space. The immediate concern is whether there exists curves such that $X + Y$ and λX are their tangent vectors. The vector space axioms follow from letting the tangent vectors act on smooth functions and by using the field properties of $\mathbb{R}$.

> **Proposition 2.6: Tangent vector operations are closed in the tangent space**
>
> Let $I = (-\varepsilon, \varepsilon) \subset \mathbb{R}$ be an interval and let $\gamma, \eta : I \to M$ be smooth curves with $\gamma(0) = p$ and $\eta(0) = p$, where $p \in M$. Let $X, Y \in T_p M$ be the directional derivative operator at p along the curves γ and η, respectively. Then, there exists smooth curves $\phi : I_\phi \to M$ and $\psi : I_\psi \to M$, where I_ϕ and I_ψ are intervals on $\mathbb{R}$, with $\phi(0) = p$ and $\psi(0) = p$, such that $X + Y$ is the tangent vector at p along ϕ and λX is the tangent vector at p along ψ, for some $\lambda \in \mathbb{R}$. That is, $X + Y \in T_p M$ and $\lambda X \in T_p M$.

Proof. We construct the curve ψ associated with scalar multiplication. If $\lambda = 0$, we set $I_\psi = \mathbb{R}$ and $\psi(t) = p$ for all $t \in I_\psi$, clearly yielding the zero directional derivative operator. For $\lambda \neq 0$, we set $I_\psi = \left(-\frac{\varepsilon}{|\lambda|}, \frac{\varepsilon}{|\lambda|}\right)$ and $\psi(t) = \gamma(\lambda t)$, for all $t \in I_\psi$. For any $f \in C^\infty(M)$, the map $f \circ \gamma : \mathbb{R} \to \mathbb{R}$ is smooth as a real valued function of a single variable. Composing $f \circ \gamma$ with the smooth map $t \mapsto \lambda t$ yields $f \circ \psi$. By the chain rule, we have

$$
(f \circ \psi)'(0) = \left(\frac{\mathrm{d}}{\mathrm{d}t} \underbrace{f \circ \gamma}_{\mathbb{R} \to \mathbb{R}} \underbrace{(\lambda t)}_{\mathbb{R} \to \mathbb{R}} \right)_{t=0}
$$
$$
= (f \circ \gamma)'(0) \cdot (\lambda t)'(0)
$$
$$
= \lambda X f,
$$

as desired.

Next we construct the curve ϕ associated with vector addition. We consider a chart (U, x) such that $p \in U$ and take a subset of $J \subset I$ such that its image by the curves γ and η is contained in U. Let $\phi : J \to M$ be defined by

$$
\phi = x^{-1} \circ (x \circ \gamma + x \circ \eta - x(p)),
$$

where $x(p)$ is the constant map $t \mapsto x(p)$ for all $t \in J$. We compute the directional derivative at p along ϕ of $f \in C^\infty(M)$ with the chain rule

$$
(f \circ \phi)'(0) = \left(\underbrace{f \circ x^{-1}}_{\mathbb{R}^n \to \mathbb{R}} \circ \underbrace{(x \circ \gamma + x \circ \eta - x(p))}_{\mathbb{R} \to \mathbb{R}^n} \right)'(0)
$$
$$
= \partial_i(f \circ x^{-1})(x(p)) \cdot \left(x^i \circ \gamma + x^i \circ \eta - x^i(p) \right)'(0)
$$
$$
= \partial_i(f \circ x^{-1})(x(p)) \cdot (x^i \circ \gamma)'(0) + \partial_j(f \circ x^{-1})(x(p)) \cdot (x^j \circ \eta)'(0)
$$
$$
= \left(f \circ x^{-1} \circ x \circ \gamma \right)'(0) + \left(f \circ x^{-1} \circ x \circ \eta \right)'(0)
$$
$$
= (f \circ \gamma)'(0) + (f \circ \eta)'(0)
$$
$$
= X f + Y f,
$$

as desired. We note the tangent vector to ϕ at p does not depend of the choice of chart, since the chart map was composed with its inverse. $\qquad \square$

2.3.1 A quick note on algebras and derivations

> **Definition 2.28: Algebra**
>
> An *algebra (over a field)* $\mathcal{A} = (V, +, \cdot, \bullet)$ is a vector space $(V, +, \cdot)$ over a field $\mathbb{K}$, equipped with a bilinear map $\bullet : V \times V \xrightarrow{\sim} V$, called a product.

Defining the product of two smooth functions on a manifold point-wise, we see $C^\infty(M)$ is an algebra over the real numbers, not just a vector space.

> **Definition 2.29: Derivation on an algebra**
>
> A *derivation D* on an algebra $\mathcal{A}$ is a linear map $D : \mathcal{A} \to \mathcal{A}$ that satisfies the *Leibniz rule*
> $$D(f \bullet g) = (Df) \bullet g + f \bullet (Dg),$$
> for all $f, g \in \mathcal{A}$.

Any $X \in T_p M$ is a *derivation at a point* on the algebra of smooth functions on a manifold, since the Leibniz rule is satisfied

$$X(fg) = (Xf)g(p) + f(p)(Xg).$$

As another example, we may set $\mathcal{A}$ as the algebra of endomorphisms on a vector space V with product $[\phi, \psi] = \phi \circ \psi - \psi \circ \phi$. It can be shown that this product satisfies the so called *Jacobi identity*,

$$[\phi, [\psi, \rho]] + [\psi, [\rho, \phi]] + [\rho, [\phi, \psi]] = 0.$$

In particular, this algebra is called a *Lie algebra*. Given $H \in \mathcal{A}$, we define a derivation

$$D_H : \mathcal{A} \xrightarrow{\sim} \mathcal{A}$$
$$\phi \mapsto [H, \phi].$$

Using the Jacobi identity, it's easy to verify this is indeed a derivation. We note that, in classical mechanics, the Poisson bracket is a derivation on the algebra of functions on phase space.

2.3.2 Chart induced basis of the tangent space

We now establish an important theorem that relates the dimension of the tangent space $T_p M$ with the dimension of the manifold. In this constructive proof, a basis for the tangent space will be determined from a chart. First, we prove a lemma.

> **Lemma 2.3: Component functions are smooth**
>
> Let $(M, \mathcal{O}_M, \mathcal{A}_M)$ be an n-dimensional smooth manifold. Let $(U, x) \in \mathcal{A}_M$ be a chart. Then, the component functions $x^i : U \to \mathbb{R}$ are smooth.

Proof. We recall that a map $f : M \to \mathbb{R}$ is smooth at a point $p \in U$ if and only if $f \circ x^{-1} \in C^\infty(\mathbb{R}^n, \mathbb{R})$.

In particular, $x^i : U \to \mathbb{R}$ is smooth if and only if $x^i \circ x^{-1} \in C^\infty(\mathbb{R}^n, \mathbb{R})$.

$$
\begin{array}{ccc}
U & \xrightarrow{\;x^i\;} & \mathbb{R} \\
{\scriptstyle x}\big\downarrow & \nearrow & \\
x(U) & {\scriptstyle x^i \circ x^{-1}} &
\end{array}
$$

By definition, we have $x^i \circ x^{-1} = \text{proj}_i$, which is smooth, since it is linear. $\qquad\square$

We may now construct a chart-induced basis for the tangent space.

> **Theorem 2.5: Dimension of the tangent space**
>
> Let M be an n-dimensional smooth manifold. Then, for all $p \in M$, the $\dim T_p M = n$.

Proof. Let (U, x) be a chart such that $p \in U$.

We consider the family of *coordinate curves* $\{\gamma_i : I \to U\}_{i=1}^n$, where $I \subset \mathbb{R}$ is an interval, with local expression satisfying

$$(x^j \circ \gamma_i)(\lambda) = x(p) + \delta_i^j \lambda,$$

for all $\lambda \in I$, $i, j \in \{1, \ldots, n\}$ and such that $\gamma_i(0) = p$. Intuitively, each curve is the image of a straight line in $x(U)$ under the map x^{-1}, where these straight lines are parallel to each axis of $\mathbb{R}^n$ and meet at $x(p)$.

Let $e_i \in T_p M$ be the tangent vector at p along γ_i and let $f \in C^\infty(M)$. By the chain rule,

$$
\begin{aligned}
e_i f &= (f \circ \gamma_i)'(0) \\
&= (f \circ x^{-1} \circ x \circ \gamma_i)'(0) \\
&= \partial_j(f \circ x^{-1})(x(p)) \cdot (x^j \circ \gamma_i)'(0) \\
&= \partial_j(f \circ x^{-1})(x(p)) \cdot (x(p) + \delta_i^j \lambda)'(0) \\
&= \delta_i^j \partial_j(f \circ x^{-1})(x(p)) \\
&= \partial_i(f \circ x^{-1})(x(p)).
\end{aligned}
$$

We write

$$
e_i f = \left. \frac{\partial}{\partial x^i} f \right|_p
$$

to denote $\partial_i(f \circ x^{-1})(x(p))$. That is, $e_i = \left. \frac{\partial}{\partial x^i} \right|_p$ and

$$
\mathcal{B} = \left\{ \left. \frac{\partial}{\partial x^1} \right|_p, \ldots, \left. \frac{\partial}{\partial x^n} \right|_p \right\}
$$

is the set of tangent vectors to the chart induced curves γ_i. We claim $\mathcal{B}$ is a generating set for $T_p M$. Indeed, let $\eta : I \to M$ be a smooth curve with $\eta(0) = p$ with tangent vector $X \in T_p M$ at p. For a smooth map $f \in C^\infty(M)$, we have

$$
\begin{aligned}
Xf &= (f \circ \eta)'(0) \\
&= (f \circ x^{-1} \circ x \circ \eta)'(0) \\
&= \partial_i(f \circ x^{-1})(x(p)) \cdot (x^i \circ \eta)'(0) \\
&= (x^i \circ \eta)'(0) \left. \frac{\partial}{\partial x^i} f \right|_p .
\end{aligned}
$$

This shows X is a linear combination of the elements of $\mathcal{B}$, with components $(x^i \circ \eta)'(0)$. Thus, $\mathcal{B}$ spans $T_p M$.

Let $\{\lambda^i\}_{i=1}^n$ be one family of components of a linear combination of the elements of $\mathcal{B}$ that results in the zero tangent vector. Due to Lemma 2.3, we may compute the directional derivative at p of component functions $x^i : U \to \mathbb{R}$. We have

$$\lambda^i \left.\frac{\partial}{\partial x^i} x^j\right|_p = \lambda^i \partial_i \left(x^j \circ x^{-1}\right)(x(p))$$
$$= \lambda^i \delta_i^j$$
$$= \lambda^j.$$

Since $\lambda^i \frac{\partial}{\partial x^i} = 0$, we have shown $\lambda^j = 0$, for all $j \in \{1, \ldots, n\}$. That is, $\mathcal{B}$ is linearly independent and, thus, a basis of $T_p M$. $\qquad \square$

Suppose $(U, x), (V, y) \in \mathscr{A}_M$ are charts of M where U and V are neighborhoods of p. Without loss of generality, we assume $U = U \cap V$ and restrict $y : U \to \mathbb{R}^n$. Then each chart induces a basis for $T_p M$, namely $\mathcal{B}_x = \left\{ \left.\frac{\partial}{\partial x^1}\right|_p, \ldots, \left.\frac{\partial}{\partial x^n}\right|_p \right\}$ and $\mathcal{B}_y = \left\{ \left.\frac{\partial}{\partial y^1}\right|_p, \ldots, \left.\frac{\partial}{\partial y^n}\right|_p \right\}$. We have already shown that for $X \in T_p M$ with an associated curve $\eta : I \to M$, then

$$X = (y^i \circ \eta)'(0) \left.\frac{\partial}{\partial y^i}\right|_p.$$

In particular, this is valid for the basis vectors in $\mathcal{B}_x$. As before, for a given basis vector, we look at its associated coordinate curve $\gamma_i : I \to M$ with $(x^j \circ \gamma_i)(\lambda) = x^j(p) + \delta_i^j \lambda$ for all $\lambda \in I$. Then, we have

$$\left.\frac{\partial}{\partial x^i}\right|_p = (y^j \circ \gamma_i)'(0) \left.\frac{\partial}{\partial y^j}\right|_p$$
$$= (y^j \circ x^{-1} \circ x \circ \gamma_i)'(0) \left.\frac{\partial}{\partial y^j}\right|_p$$
$$= \partial_k(y^j \circ x^{-1})(x(p)) \cdot (x^k \circ \gamma_i)'(0) \left.\frac{\partial}{\partial y^j}\right|_p$$
$$= \delta_i^k \left.\frac{\partial}{\partial x^k} y^j\right|_p \left.\frac{\partial}{\partial y^j}\right|_p$$
$$= \left.\frac{\partial}{\partial x^i} y^j\right|_p \left.\frac{\partial}{\partial y^j}\right|_p,$$

that is, the linear map with components $\left.\frac{\partial}{\partial x^i} y^j\right|_p$ is the map that governs the change of basis from $\mathcal{B}_x$ to $\mathcal{B}_y$. This map is usually called the *Jacobian*.

2.3.3 Cotangent spaces and the gradient

We now define the dual vector space to the tangent space.

> **Definition 2.30: Cotangent space at a point**
>
> Let M be a smooth manifold. The *cotangent space T_p^*M at a point $p \in M$* is the dual vector space to the tangent space T_pM, that is $T_p^*M = (T_pM)^*$.

Remark 2.17. *We recall that for a finite-dimensional vector space, its dimension is equal to the dimension of its dual vector space. Therefore, for a finite-dimensional manifold, the dimension of the cotangent space is equal to the tangent space.*

We may now define tensors over a tangent plane on a point of the manifold. As before, the set

$$T_s^r(T_pM) = \underbrace{T_pM \otimes \cdots \otimes T_pM}_{r \text{ times}} \otimes \underbrace{T_p^*M \otimes \cdots \otimes T_p^*M}_{s \text{ times}}$$

of multilinear maps

$$t : \underbrace{T_p^* \times \cdots \times T_p^*}_{r \text{ times}} \times \underbrace{T_pM \times \cdots \times T_pM}_{s \text{ times}} \tilde{\to} \mathbb{R}$$

is the set (vector space) of all (r, s)-tensors at the point $p \in M$.

> **Definition 2.31: Gradient operator at a point**
>
> The *gradient d_p operator at a point $p \in M$*, is a linear map
>
> $$d_p : C^\infty(M) \tilde{\to} T_p^*M$$
> $$f \mapsto d_p f$$
>
> defined by
>
> $$(d_p f)(X) = Xf,$$
>
> for all $X \in T_pM$.

Given a smooth map $f : M \to \mathbb{R}$, we may look at its level set $\{p \in M : f(p) = c\}$, for some constant $c \in \mathbb{R}$. Then, if $X \in T_pM$ is tangent to this level set, it follows that $d_p f(X) = 0$.

With the notion of the gradient operator, we construct a chart-induced basis for the cotangent space.

> **Theorem 2.6: Chart-induced covector basis**
>
> Let M be an n-dimensional smooth manifold. Let $(U, x) \in \mathscr{A}_M$ be a chart with $p \in U$. Then
>
> $$\mathcal{B} = \left\{ d_p x^1, \ldots, d_p x^n \right\}$$
>
> is the *chart-induced covector basis* for T_p^*M.

Proof. We recall Lemma 2.3 to show that, indeed, $d_p x^i \in T_p^*M$.

Consider $d_p x^i \left(\left. \frac{\partial}{\partial x^j} \right|_p \right)$. We have

$$d_p x^i \left(\left. \frac{\partial}{\partial x^j} \right|_p \right) = \left. \frac{\partial}{\partial x^j} x^i \right|_p$$

$$= \partial_j \underbrace{\left(x^i \circ x^{-1} \right)}_{\text{proj}_i : \, \mathbb{R}^n \to \mathbb{R}} (x(p))$$

$$= \delta^i_j.$$

From linearity, we have

$$d_p x^i \left(\lambda^j \left. \frac{\partial}{\partial x^j} \right|_p \right) = \lambda^i,$$

that is, $\mathcal{B}$ is the dual basis of $\left\{ \left. \frac{\partial}{\partial x^1} \right|_p, \ldots, \left. \frac{\partial}{\partial x^n} \right|_p \right\}$ for the dual space $T^*_p M$. $\hspace{1cm}$ □

2.3.4 Pushforward and pullback

Given a smooth map between smooth manifolds, we may define an object that takes tangent vectors on a manifold to tangent vectors on another manifold.

Definition 2.32: Pushforward at a point

Let $\phi : M \to N$ be a smooth map between smooth manifolds. The *pushforward ϕ_{*p} of the map ϕ at the point $p \in M$* is the linear map

$$\phi_{*p} : T_p M \xrightarrow{\sim} T_{\phi(p)} N$$
$$X \mapsto \phi_{*p}(X),$$

defined by

$$\phi_{*p}(X) f = X(f \circ \phi)$$

for all $f \in C^\infty(N)$.

Remark 2.18. *The pushforward ϕ_{*p} is often called the* differential *of ϕ at p.*

We note the tangent vector $X_{\gamma, p}$ at p of the smooth curve γ is *pushed forward* to the tangent vector $\phi_{*p}(X)$ of the smooth curve $\phi \circ \gamma$. Indeed, let $f \in C^\infty(N)$, then

$$\phi_{*p}(X) f = X(f \circ \phi)$$
$$= (f \circ \phi \circ \gamma)'(0)$$
$$= X_{\phi \circ \gamma, \phi(p)} f,$$

which implies $X_{\phi \circ \gamma, \phi(p)} = \phi_{*p}(X)$, as desired.

Similarly, we define an object that relates elements in cotangent spaces from one manifold to another.

Definition 2.33: Pullback at a point

Let $\phi : M \to N$ be a smooth map between smooth manifolds. The *pullback ϕ^*_p of the*

map ϕ at the point $\phi(p)$ is the linear map

$$\phi_p^* : T_{\phi(p)}^* N \xrightarrow{\sim} T_p^* M$$
$$\omega \mapsto \phi_p^*(\omega),$$

defined by

$$\phi_p^*(\omega)(Y) = \omega(\phi_{*p}(Y))$$

for all $Y \in T_p M$.

Remark 2.19. *Unless the map ϕ is a diffeomorphism, vectors are pushed forward and covectors are pulled back.*

2.3.5 Ambient space

So far, the tangent plane and its structure has been done intrinsically.

Decide the question under which circumstance some smooth manifold M can sit in some $\mathbb{R}^n$

Although the following definitions are done between smooth manifolds M and N we will turn our attention to the special case $N = \mathbb{R}^m$, equipped with the standard topology.

Definition 2.34: Immersion

An *immersion ϕ of M into N* is a smooth map $\phi : M \to N$ such that, for all $p \in M$, the pushforward ϕ_{*p} is injective.

Example 2.2. *(a) We consider a map $\phi : S^1 \to \mathbb{R}^2$ such that $\phi(M)$ is a closed simple smooth curve in $\mathbb{R}^2$. Then $\phi(M)$ is an immersion.*

(b) We consider the map $\tilde{\phi} : S^1 \to \mathbb{R}^2$ such that $\tilde{\phi}(M)$ is a lemniscate. We note $\tilde{\phi}$ is not injection, but $\tilde{\phi}_*$ is injective, and therefore $\tilde{\phi}_*$ is an immersion.

Definition 2.35: Embedding

An *embedding* is an immersion with image homeomorphic to its range.

Example 2.3. *The map ϕ defined in the previous example is an embedding, but the map $\tilde{\phi}$ is not.*

We may now quote two important theorems that answer the question considered in the beginning of this section.

Theorem 2.7: Whitney's immersion and embedding theorem

Any n-dimensional smooth manifold can be

(a) immersed in $\mathbb{R}^{2n-1}$;

(b) embedded in $\mathbb{R}^{2n}$.

Example 2.4. *The two-dimensional Klein bottle can be immersed into $\mathbb{R}^3$ but only embedded in $\mathbb{R}^4$. This is the "worst case" of the Whitney theorem. The theorem does not state if there is an immersion or embedding in lower dimensions other than the ones stated. One could check, for instance, the 2-sphere may be embedded into $\mathbb{R}^3$.*

With this theorem one could show the intrinsically definitions are equivalent to ones defined extrinsically. For instance, one could embed a manifold into a higher dimensional Euclidean space, named the *ambient space*, and define the tangent space at a point as a hyperplane in the ambient space.

Even though the extrinsic approach is equivalent to the intrinsic approach, in the purely abstract sense, it may not desirable to pursue in some cases. As an example, in General Relativity there would have to be a physical justification for an ambient space where spacetime is embedded.

2.4 Tangent bundle and tensor fields

Up until this point, the differentiable structure of the manifold was studied at one point at a time by virtue of the tangent space at such a point.

Definition 2.36: Tangent bundle

Let $(M, O_M, \mathcal{A}_M)$ be an n-dimensional smooth manifold. The *tangent bundle TM* is the disjoint union of all the tangent spaces of the manifold, that is

$$TM = \biguplus_{p \in M} T_p M$$

The *bundle projection π* is the map

$$\pi : TM \to M$$
$$X \mapsto p,$$

where $p \in M$ is *the* point for which $X \in T_p M$, often referred as the *base point*.

Remark 2.20. *We recall the disjoint union may be written as*

$$\biguplus_{p \in M} T_p M = \bigcup_{p \in M} \{p\} \times T_p M,$$

justifying the well definition of the projection map π.

The previous definition makes it clear that

$$TM \xrightarrow{\pi} M$$

is a *set bundle*. We wish to add structure to TM such that this is a bundle on smooth manifolds.

Let $(U, x) \in \mathcal{A}_M$ be a chart. We define the map

$$\xi : \pi^{-1}(U) \to \xi\left(\pi^{-1}(U)\right) \subset \mathbb{R}^{2n}$$
$$X \mapsto \left(x^1(\pi(X)), \dots, x^n(\pi(X)), X^1, \dots, X^n\right),$$

where $X^i = d_{\pi(X)} x^i(X)$ are the components of X in the basis $\left\{ \frac{\partial}{\partial x^1}\big|_{\pi(X)}, \dots, \frac{\partial}{\partial x^n}\big|_{\pi(X)} \right\}$. Suppose there exists $X, Y \in \pi^{-1}(U)$ such that $\xi(X) = \xi(Y)$. Since x is injective, we must have $\pi(X) =$

$\pi(Y)$, and because the components of X are the same as those of Y, we have $X = Y$, hence ξ is one-to-one. For any $(a^1, \ldots, a^n, b^1, \ldots, b^n) \in \pi^{-1}(\tilde{U})$, we have

$$\xi \left(b^i \left. \frac{\partial}{\partial x^i} \right|_{x^{-1}(a^1, \ldots, a^n)} \right) = (a^1, \ldots, a^n, b^1, \ldots, b^n),$$

therefore ξ is onto. We have shown ξ is a bijection, with inverse given by

$$\xi^{-1} : \xi \left(\pi^{-1}(U) \right) \to \pi^{-1}(U)$$

$$\left(a^1, \ldots, a^n, b^1, \ldots, b^n \right) \mapsto b^i \left. \frac{\partial}{\partial x^i} \right|_{x^{-1}(a^1, \ldots, a^n)},$$

where the base point is $x^{-1}(a^1, \ldots, a^n)$.

It is clear that if TM had a topology where $\pi^{-1}(U)$ was an open set, then ξ would be a homeomorphism from $\pi^{-1}(U)$ to $\mathbb{R}^{2n}$. Let $\mathscr{A}_M = \{(U_\alpha, x_\alpha)\}_{\alpha \in J}$ be the smooth atlas for M, where J is an indexing set. We define $\mathscr{A}_{TM} = \left\{ \left(\pi^{-1}(U_\alpha), \xi_\alpha \right) \right\}_{\alpha \in J}$, with ξ_α constructed analogously from x_α. We define the topology O_{TM} on TM with the maps ξ_α: a set $A \subset TM$ is open if and only if $\xi_\alpha(U_\alpha \cap A)$ is open in $x_\alpha(U_\alpha) \times \mathbb{R}^n$ in its subspace topology induced from the standard topology in $\mathbb{R}^{2n}$. I need to check if this is Hausdorff and second-countable, and that it is a topology.

We now check any two charts in $\mathscr{A}_{TM}$ are C^∞-compatible. Without loss of generality, let $(U, \tilde{x}) \in \mathscr{A}_M$ be another chart on M and let $\tilde{\xi} : \pi^{-1}(U) \to \tilde{\xi}\left(\pi^{-1}(U) \right)$ be constructed analogously to ξ, that is, $\left(\pi^{-1}(U), \tilde{\xi} \right) \in \mathscr{A}_{TM}$.

The chart transition map $\tilde{\xi} \circ \xi^{-1} : \xi \left(\pi^{-1}(U) \right) \to \tilde{\xi} \left(\pi^{-1}(U) \right)$ is given by

$$\tilde{\xi} \circ \xi^{-1} \left(a^1, \ldots, a^n, b^1, \ldots b^n \right) = \tilde{\xi} \left(b^i \left. \frac{\partial}{\partial x^i} \right|_{x^{-1}(a^1, \ldots, a^n)} \right)$$

$$= \left(\tilde{a}^1, \ldots, \tilde{a}^n, \tilde{b}^1, \ldots, \tilde{b}^n \right),$$

where $\tilde{a}^i = \tilde{x}^i \circ x^{-1} \left(a^1, \ldots, a^n \right)$ and $\tilde{b}^j = b^i \left. \frac{\partial}{\partial x^i} \tilde{x}^j \right|_{x^{-1}(a^1, \ldots, a^n)}$, since we have

$$b^k \left. \frac{\partial}{\partial x^k} \right|_p = b^i \left. \frac{\partial}{\partial x^i} \tilde{x}^j \right|_p \left. \frac{\partial}{\partial \tilde{x}^j} \right|_p,$$

for all $p \in U$. From the definition of the chart-induced basis, we have

$$\tilde{b}^j = b^i \partial_i(\tilde{x}^j \circ x^{-1})(a_1, \ldots, a^n),$$

so it follows from smoothness of the chart transition maps in $\mathscr{A}_M$ that $\tilde{\xi} \circ \xi^{-1}$ is smooth. Interpreting $\xi(\pi^{-1}(U)) = x(U) \times \mathbb{R}^n \subset \mathbb{R}^n \times \mathbb{R}^n$, we have shown

$$\tilde{\xi} \circ \xi^{-1}(a, b) = \left((\tilde{x} \circ x^{-1})(a), (\tilde{x} \circ x^{-1})_{*a}(b) \right),$$

for $(a, b) \in x(U) \times \mathbb{R}^n$.

Now that we have established that TM is a smooth manifold on its own right, we can ask whether the projection map $\pi : TM \to M$ is smooth. Let $(U, x) \in \mathscr{A}_M$ and $(\pi^{-1}(U), \xi) \in \mathscr{A}_{TM}$ as before.

$$
\begin{array}{ccc}
\pi^{-1}(U) & \xrightarrow{\ \pi\ } & U \\
\downarrow{\scriptstyle \xi} & & \downarrow{\scriptstyle x} \\
\xi\left(\pi^{-1}(U)\right) & \xrightarrow{x \circ \pi \circ \xi^{-1}} & x(U)
\end{array}
$$

It is easy to see the map $x \circ \pi \circ \xi^{-1}$ is the projection

$$
x \circ \pi \circ \xi^{-1} : \xi\left(\pi^{-1}(U)\right) \subset \mathbb{R}^n \times \mathbb{R}^n \to x(U) \subset \mathbb{R}^n
$$

$$
(a, b) \mapsto a,
$$

and therefore $x \circ \pi \circ \xi^{-1} \in C^\infty(\mathbb{R}^{2n}, \mathbb{R}^n)$. By the definition of a smooth map between manifolds, the bundle projection π is smooth.

2.4.1 Rings and modules

We now define vector fields on the manifold. Unlike the previous discussion on tangent spaces, vector fields will not be vector spaces, but *modules*.

Definition 2.37: Vector field

Let M be a smooth manifold and TM its tangent bundle. A *(smooth) vector field* σ is a (smooth) section of TM, that is, the (smooth) map $\sigma : M \to TM$ satisfies $\pi \circ \sigma = \mathrm{id}_M$. The set

$$\Gamma(TM) = \{\sigma : M \to TM \mid \sigma \text{ is smooth and } \pi \circ \sigma = \mathrm{id}_M\}$$

denotes the set of all smooth vector fields.

We equip the set $\Gamma(TM)$ with two operations defined point-wise, namely addition of vector fields

$$
+ : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)
$$

$$
(\sigma, \tau) \mapsto \sigma + \tau,
$$

where $(\sigma + \tau)(p) = \sigma(p) + \tau(p)$, and scaling by a smooth map

$$
\cdot : C^\infty(M) \times \Gamma(TM) \to \Gamma(TM)
$$

$$
(f, \sigma) \mapsto f \cdot \sigma,
$$

where $(f \cdot \sigma)(p) = f(p)\sigma(p)$.

The choice of scaling vector fields by functions and not just real numbers is a very intuitive notion, however the set of all smooth functions on a manifold, equipped with point-wise addition and multiplication, is not a field. To see this, one considers a map that has at least one zero at a point, so there is no guarantee of a multiplicative inverse.

Definition 2.38: Ring

A ring $(R, +, \cdot)$ is a set R equipped with two maps $+, \cdot : R \times R \to R$ called addition and multiplication that satisfy

- $(R, +)$ is an abelian group under addition;

- Associativity of multiplication: For all $a, b, c \in R$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$;

- Distributivity of multiplication over addition: For all $a, b, c \in R$, $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$.

A *commutative ring* satisfies the commutativity of multiplication: for all $a, b \in R$, $a \cdot b = b \cdot a$.

If there exists a multiplicative identity, i.e., an element $1 \in R$ such that $1 \cdot a = a$ and $a \cdot 1 = a$ for all $a \in R$, then R is a *unital ring*. A unital ring is a *division ring* if there exists a multiplicative inverse, that is, for all $a \in R \setminus \{0\}$, there exists $a^{-1} \in R \setminus \{0\}$ such that $a^{-1} \cdot a = 1$ and $a \cdot a^{-1} = 1$.

Usually the multiplication $a \cdot b$ is denoted by ab.

Remark 2.21. *A field is a commutative unital division ring.*

It is easy to verify $(C^\infty(M), +, \cdot)$ is a commutative unital ring, but *not* a division ring. Thus we may not say $\Gamma(TM)$ is a vector space. We now generalize the concept of vector space over fields to rings.

Definition 2.39: Module over a commutative unital ring

An *R-module* $(V, +, \cdot)$ *over a commutative unital ring* R is a set V equipped with two maps $+ : V \times V \to V$, called addition, and $\cdot : R \times V \to V$, called scalar multiplication, which satisfy

- $(V, +)$ is an abelian group under addition;

- For all $r, s \in R$ and $x, y \in V$:

    (a) $r \cdot (x + y) = r \cdot x + r \cdot y$;

    (b) $(r + s) \cdot x = r \cdot x + s \cdot x$;

    (c) $(rs) \cdot x = r \cdot (s \cdot x)$;

    (d) $1 \cdot x = x$.

Usually the scalar multiplication $a \cdot v$ is denoted by av and it is clear from context that it is the scalar multiplication.

Remark 2.22. *If the ring is not commutative, one may define left and right modules. And if the ring is not unital, the axiom $1 \cdot x = x$ or $x \cdot 1 = x$ is omitted.*

Remark 2.23. *Since every field $\mathbb{K}$ is a commutative unital ring, it is trivial that every $\mathbb{K}$-vector space is a $\mathbb{K}$-module.*

It is easy to verify $(\Gamma(TM), +, \cdot)$ is a $C^\infty(M)$-module. We finally prove the existence of a Hamel basis for R-modules if the underlying ring is a *division* ring. As motivation, we consider the plane $\mathbb{R}^2$ and the 2-sphere S^2. It is easy to see the canonical basis of $\mathbb{R}^2$ provides

a basis for $\Gamma(T\mathbb{R}^2)$, however, there is no non-vanishing smooth vector field on the sphere [Car16], and thus there is no basis for $\Gamma(TS^2)$. Were $C^\infty(M)$ a division ring, there would exist a basis for vector fields in all manifolds.

We begin with a few definitions [SKM97] in order to state the Zorn's lemma, which is needed to prove the theorem.

> **Definition 2.40: Partially ordered set and chain**
>
> A *partially ordered set* X is a set X equipped with a *ordering* $\leq$ that is
>
> (a) reflexive: $\forall x \in X, x \leq x$;
>
> (b) transitive: $\forall x, y, z \in X, x \leq y \wedge y \leq z \implies x \leq z$; and
>
> (c) anti-symmetric: $\forall x, y \in X, x \leq y \wedge y \leq x \implies x = y$.
>
> If for any pair $x, y \in X$ either $x \leq y$ or $y \leq x$, then X is a *chain*.
> An *upper bound of a subset* $Y \subset X$ in a partially ordered set X is an element $x \in X$ such that $y \leq x$ for all $y \in Y$. The *greatest element of a partially ordered set* X is an element $n \in X$ such that $x \leq n$ for all $x \in X$. A *maximal element of a partially ordered set* X is an element $m \in X$ such that $x \in X : m \leq x \implies x = m$.

Example 2.5. *An example of a partially ordered set is the power set $\mathcal{P}(S)$ of some set S with the ordering given by the inclusion $\subset$.*

> **Theorem 2.8: Zorn's lemma**
>
> Let X be a non-empty partially ordered set, any chain in which has an upper bound in X. Then some chain has an upper bound that is simultaneously the maximal element in X.

Remark 2.24. *Within ZF-set theory, this lemma is equivalent to the axiom of choice, that is, this theorem follows from ZFC-set theory.*

> **Theorem 2.9: Existence of Hamel basis**
>
> If D is a division ring, then the D-module V has a Hamel basis.

Proof. We assume without loss of generality that V is a left D-module.

Trivially, V is a generating system for V. Let $S \subset V$ be a generating system for V, taking $S = V$, if necessary. We recall the definition of a generating system:

$$\forall v \in V, \exists N \in \mathbb{N}, \exists v^1, \ldots, v^N \in D, \exists e_1, \ldots, e_N \in S : v = v^i e_i.$$

We define the non-empty set X of linearly independent subsets of S

$$X = \big\{ U \in \mathcal{P}(S) : U \text{ is linearly independent} \big\},$$

which is partially ordered under inclusion. We recall a set U is linearly independent if any finite linear combination of elements in U equals 0_V, then the coefficients of the linear combination are 0_D.

Let $Z \subset X$ be any chain in X. Consider the union

$$T = \bigcup Z = \{ v \in V : \exists Y \in Z \text{ such that } v \in Y \}.$$

We consider a finite set $\{y_1, \ldots, y_m\}$ from T where $y_i \in Y_i \in Z$. Since Z is a chain, we either have $Y_i \subset Y_j$ or $Y_j \subset Y_i$, for all $i, j \in \{1, \ldots, m\}$, and we may ignore one of them, say Y_j, since if $Y_j \subset Y_i$, we have $y_j \in Y_i$. We may do this until only one subset Y remains and the entire set of m elements is contained in it. Since $Y \in Z$, the finite set is linearly independent, hence T is linearly independent and thus $T \in X$. It is clear $Y \subset T$, for all $Y \in Z$, therefore T is an upper bound of Z in X. By Zorn's lemma, X has a maximal element, one of which we name $\mathcal{B}$.

Since $\mathcal{B}$ is linearly independent, we want to show it is a basis for the generating system S. That is, we must show $\mathcal{B}$ is a generating system for S. Let $v \in S$. If $v \in \mathcal{B}$, then v is clearly a finite linear combination of elements in $\mathcal{B}$, so we may assume $v \notin \mathcal{B}$. In which case, we have $\mathcal{B} \cup \{v\}$ linearly dependent, since $\mathcal{B}$ is a maximal linearly independent set. This implies there exists $e_1, \ldots, e_N \in \mathcal{B}, a^1, \ldots, a^N, b \in D$ such that

$$a^i e_i + bv = 0$$

and not all $a^1, \ldots, a^N, b \in D$ vanish. Suppose $b = 0$, then $a^i e_i = 0$ is a finite linear combination of elements in $\mathcal{B}$ that results in zero, which would mean $a^i = 0$. Thus, b must be non-zero. Rearranging, we get

$$a^i e_i = -bv.$$

Since $b \neq 0$ and D is a division ring, we have

$$(-b)^{-1} a^i e_i = v,$$

hence $\mathcal{B}$ is a generating system for S. Since S spans V and $\mathcal{B}$ is a basis for S, it must follow that $\mathcal{B}$ is a basis for V. $\qquad\square$

Corollary 2.2. *Every vector space has a Hamel basis.*

Since components are related to the existence of a basis, then for a tensor field, which generically does not have a basis, one cannot think of it only about its components, since there is no such a thing globally.

Definition 2.41: Direct sum of modules

Let U, V be R-modules. The direct sum $U \oplus V$ denotes the R-module $(U \times V, +, \cdot)$, where

$$+ : (U \times V) \times (U \times V) \to U \times V$$
$$((u_1, v_1), (u_2, v_2)) \mapsto (u_1 + u_2, v_1 + v_2)$$

and

$$\cdot : R \times (U \times V) \to U \times V$$
$$(a, (u, v)) \mapsto (au, av).$$

Remark 2.25. *It is easy to verify $U \oplus V$ is an R-module.*

We may now state the theorem that classifies the sections of a vector fiber bundle, such as the tangent bundle, as a $C^\infty(M)$-module.

Definition 2.42: Free module

A *module* is free if it has a basis. It is said to be *finitely generated* if the cardinality of the basis is finite.

Remark 2.26. *It is possible to show that a finitely-generated R-module F is free, then F is isomorphic to $R \oplus \cdots \oplus R$.*

> **Definition 2.43: Projective module**
>
> An R-module Γ is *projective* if it is a direct summand of a free R-module F, that is
>
> $$\Gamma \oplus Q = F,$$
>
> for some R-module Q.

Remark 2.27. *Taking $Q = \{0\}$, we see every free module is projective.*

> **Theorem 2.10: Serre-Swan theorem**
>
> The set of all smooth sections of a vector fiber bundle E over a smooth manifold M is a finitely generated projective $C^\infty(M)$-module $\Gamma(E)$, that is
>
> $$\Gamma(E) \oplus Q = F,$$
>
> where F and Q are $C^\infty(M)$-modules and F is free.

> **Proposition 2.7: Set of R-modules of homomorphisms is a module**
>
> Let P, Q be finitely generated (projective) R-modules, where R is a *commutative* ring. The set of R-linear maps from P to Q
>
> $$\operatorname{Hom}_R (P, Q) = \left\{ \phi : P \xrightarrow{\sim} Q \text{ such that } \phi \text{ is } R\text{-linear} \right\}$$
>
> equipped with point-wise addition
>
> $$+ : \operatorname{Hom}_R (P, Q) \times \operatorname{Hom}_R (P, Q) \to \operatorname{Hom}_R (P, Q)$$
> $$(\phi, \psi) \mapsto \phi + \psi,$$
>
> where $(\phi + \psi)(a) = \phi(a) + \psi(a)$ and
>
> $$\cdot : R \times \operatorname{Hom}_R (P, Q) \to \operatorname{Hom}_R (P, Q)$$
> $$(b, \psi) \mapsto b \cdot \psi,$$
>
> where $(b \cdot \psi)(a) = b \cdot \psi(a)$, is a R-module.

Proof. We recall the proof of Proposition 2.1. The commutativity of field multiplication was necessary to prove the set of linear maps between vector spaces is a vector space. Doing the exact same computations, this proposition follows. □

In particular, the dual $C^\infty(M)$-module to $\Gamma(TM)$, the set

$$\Gamma(TM)^* = \operatorname{Hom}_{C^\infty(M)} (\Gamma(TM), C^\infty(M))$$

is a $C^\infty(M)$-module. It can be shown that $\Gamma(TM)^* = \Gamma(T^*M)$, called the cotangent bundle. I kinda want to show this, as it's not immediate to me, although intuitive.

> **Definition 2.44: Tensor field**
>
> A (r, s)-tensor field t on a smooth manifold M is a $C^\infty(M)$-multilinear map
>
> $$t : \underbrace{\Gamma(T^*M) \times \cdots \times \Gamma(T^*M)}_{r \text{ times}} \times \underbrace{\Gamma(TM) \times \cdots \times \Gamma(TM)}_{s \text{ times}} \overset{\sim}{\to} C^\infty(M).$$

2.4.2 Vector fields

Throughout this section, $(M, O_M, \mathscr{A}_M)$ is an n-dimensional smooth manifold. Even though there is no guarantee of a basis for the $C^\infty(M)$-module $\Gamma(TM)$, we may find *local* basis for submanifolds. Let $(U, x) \in \mathscr{A}_M$ be a chart, where we define the n vector fields

$$\frac{\partial}{\partial x^i} : U \to TU$$
$$p \mapsto \left. \frac{\partial}{\partial x^i} \right|_p,$$

then at each point $p \in U$, this set spans the tangent space $T_p U$. A vector field $X \in \Gamma(TM)$ may be expressed in U as a linear combination of the vector fields $\frac{\partial}{\partial x^i}$, that is

$$\frac{\partial}{\partial X} : U \to TU$$
$$p \mapsto a^i(p) \left. \frac{\partial}{\partial x^i} \right|_p,$$

may be written as $X = a^i \frac{\partial}{\partial x^i}$, where the component functions are maps $a^i : U \to \mathbb{R}$.

> **Lemma 2.4: Smooth vector field components are smooth**
>
> Let $(U, x) \in \mathscr{A}_M$ be a chart on M. A vector field $X = a^i \frac{\partial}{\partial x^i}$ on U is smooth if and only if the component functions a^i are smooth on U.

Proof. Let (TU, ξ) be the induced chart on TU by (U, x), defined by

$$\xi : TU \to x(U) \times \mathbb{R}^n$$
$$Y \mapsto (x^1 \circ \pi(Y), \ldots, x^n \circ \pi(Y), c^1(Y), \ldots, c^n(Y)),$$

where $c^i \in C^\infty(TU)$ are the smooth maps defined by

$$c^i : TU \to \mathbb{R}$$
$$Y \mapsto d_{\pi(Y)} x^i(Y),$$

that is, $Y = c^i(Y) \left. \frac{\partial}{\partial x^i} \right|_{\pi(Y)}$. Note that

$$c^i(X(p)) = a^i(p)$$

for all $p \in U$, then we have $a^i = c^i \circ X$.

$$
\begin{array}{ccc}
U & \overset{X}{\longrightarrow} & TU \\
& a^i \searrow \quad \swarrow c^i & \\
& \mathbb{R} &
\end{array}
$$

Since ξ is a diffeomorphism, X is smooth if and only if $\xi \circ X : U \to x(U) \times \mathbb{R}^n$ is smooth. For all $p \in U$, we have

$$\xi \circ X(p) = (x^1(p), \dots, x^n(p), a^1, \dots, a^1(p), \dots, a^n(p)).$$

Due to the component functions x^i being smooth, this implies $\xi \circ X$ is smooth if and only if the maps a^i are smooth. $\square$

Remark 2.28. *As a consequence[Tu11], a vector field X on M is smooth if and only if for any chart (U, x) the coefficients a^i of $X|_U = a^i \frac{\partial}{\partial x^i}$ are smooth.*

Just as tangent vectors are derivations at a point on the $\mathbb{R}$-algebra $C^\infty(M)$ of smooth maps, smooth vector fields are derivations on that algebra. Indeed, let $X \in \Gamma(TM)$ and $f \in C^\infty(M)$, then Xf is the map defined by

$$Xf : M \to \mathbb{R}$$
$$p \mapsto X(p)f.$$

Let $f, g \in C^\infty(M)$. If Xf and Xg are smooth, then it is easy to see that

$$X(fg) = (Xf)g + f(Xg),$$

from the fact $X(p)$ is a derivation at the point $p \in M$. It remains to show whether Xf is a smooth map, to conclude X is a derivation of the algebra.

> **Lemma 2.5: Smooth vector fields are derivations**
>
> Let $X \in \Gamma(TM)$ be a vector field and let $f \in C^\infty(M)$ be a smooth map. Then, the map Xf is smooth.

Proof. Suppose X is smooth. Then for any chart $(U, x) \in \mathcal{A}_M$, the coefficients a^i of $X = a^i \frac{\partial}{\partial x^i}$ are smooth on U. For all $p \in U$, we have

$$Xf(p) = a^i(p) \left. \frac{\partial}{\partial x^i} f \right|_p = a^i(p) \partial_i(f \circ x^{-1})(p),$$

that is, $Xf = a^i \cdot \partial_i(f \circ x^{-1})$ in U. By Lemma 2.4, $a^i \in C^\infty(U)$, and we conclude Xf is smooth in U. Since the charts cover M, $Xf \in C^\infty(M)$. $\square$

Remark 2.29. *Weakening the hypothesis from a smooth vector field X to any section of tangent bundle, the converse may be shown [Tu11]: if Xf is smooth, then X is smooth.*

From Lemma 2.5, we may consider the iterated action of vector fields on a smooth map. Let $X, Y \in \Gamma(TM)$ and $f \in C^\infty(M)$, then Xf and Yf are smooth maps, and as a consequence so too $X(Yf)$ and $Y(Xf)$ must be smooth. Let us denote the iterations by the compositions $XY = X \circ Y$ and $YX = Y \circ X$, regarding the vector fields as derivations. Let $f, g \in C^\infty(M)$, then

$$XY(fg) = X(Yf \cdot g + f \cdot Yg)$$
$$= XY(f) \cdot g + Yf \cdot Xg + Xf \cdot Yg + f \cdot XY(g)$$

and similarly

$$YX(fg) = YX(f) \cdot g + Xf \cdot Yg + Xf \cdot Yg + f \cdot YX(g).$$

It is clear the iterated derivations are not derivations, since the term $Xf \cdot Yg + Xf \cdot Yg$ appears in both.

Definition 2.45: Lie bracket of vector fields

Given two smooth vector fields $X, Y \in \Gamma(TM)$, their *Lie bracket* $[X, Y]$ or *commutator* is the vector field defined by

$$[X, Y](p)f = X(p)(Yf) - Y(p)(Xf),$$

for all $p \in M$ and $f \in C^\infty(M)$.

With a similar computation, we check $[X, Y](p)$ is a derivation at the point p, and thus a tangent vector of T_pM. For clarity, the evaluation of vector fields at a point are denoted by subscript, such as $X(p) = X_p$. For $f, g \in C^\infty(M)$, we have

$$X_p(Y(fg)) = X_p(gYf + fYg)$$
$$= (Yf)(p)X_pg + g(p)X_p(Yf) + f(p)X_p(Yg) + (Yg)(p)X_pf$$
$$= f(p)X_p(Yg) + g(p)X_p(Yf) + X_p((Yg)(p)f + (Yf)(p)g)$$

and similarly

$$Y_p(X(fg)) = f(p)Y_p(Xg) + g(p)Y_p(Xf) + Y_p((Xg)(p)f + (Xf)(p)g).$$

It is clear when we subtract this pair of equations, the first two terms will result in the desired outcome, however we must check the other subtraction vanishes. Indeed, let (U, x) be a chart such that $p \in U$, then $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^i \frac{\partial}{\partial y^i}$. We have

$$X_p((Yg)(p)f + (Yf)(p)g) = a^i(p)\frac{\partial}{\partial x^i}\left(b^j(p)\frac{\partial}{\partial x^j}g\Big|_p f + b^j(p)\frac{\partial}{\partial x^j}f\Big|_p g\right)\Bigg|_p$$
$$= a^i(p)b^j(p)\frac{\partial}{\partial x^i}\left(\partial_j(g \circ x^{-1})(p)f + \partial_j(f \circ x^{-1})(p)g\right)\Bigg|_p$$

and similarly

$$Y_p((Xg)(p)f + (Xf)(p)g) = b^i(p)\frac{\partial}{\partial x^i}\left(a^j(p)\frac{\partial}{\partial x^j}g\Big|_p f + a^j(p)\frac{\partial}{\partial x^j}f\Big|_p g\right)\Bigg|_p$$
$$= b^i(p)a^j(p)\frac{\partial}{\partial x^i}\left(\partial_j(g \circ x^{-1})(p)f + \partial_j(f \circ x^{-1})(p)g\right)\Bigg|_p,$$

which results in

$$X_p((Yg)(p)f + (Yf)(p)g) - Y_p((Xg)(p)f + (Xf)(p)g) =$$
$$= (a^ib^j - b^ia^j)(p)\frac{\partial}{\partial x^i}\left(\partial_j(g \circ x^{-1})(p)f + \partial_j(f \circ x^{-1})(p)g\right)\Bigg|_p$$
$$= (a^ib^j - b^ia^j)(p)\left(\partial_j(g \circ x^{-1})(p)\partial_i(f \circ x^{-1})(p) + \partial_j(f \circ x^{-1})(p)\partial_i(g \circ x^{-1})(p)\right)$$
$$= 0,$$

since the first term is anti-symmetric and the second is symmetric. In the end, we are left with

$$[X, Y]_p(fg) = X_p(Y(fg)) - Y_p(X(fg))$$
$$= f(p)\left(X_p(Yg) - Y_p(Xg)\right) + g(p)\left(X_p(Yf) - Y_p(Xf)\right)$$
$$= f(p)[X, Y]_pg + g(p)[X, Y]_pf,$$

that is, $[X, Y](p)$ is a derivation at p, so $[X, Y](p) \in T_p M$. Finally, $[X, Y] : M \to TM$ is a vector field.

> ### Theorem 2.11: Lie bracket of smooth vector fields is smooth
>
> If $X, Y \in \Gamma(TM)$ are smooth vector fields, then $[X, Y]$ is a smooth vector field.

Proof. Let $(U, x) \in \mathcal{A}_M$ be a chart, where $X = a^i \frac{\partial}{\partial x^i}$ and $Y = b^i \frac{\partial}{\partial x^i}$. Then, for any map $f \in C^\infty(M)$, we have

$$XYf = a^i \frac{\partial}{\partial x^i} \left(b^j \frac{\partial}{\partial x^j} f \right)$$
$$= a^i \frac{\partial}{\partial x^i} \left(b^j \partial_j (f \circ x^{-1}) \right)$$
$$= a^i \partial_i \left(b^j \circ x^{-1} \right) \partial_j \left(f \circ x^{-1} \right) + a^i b^j \partial^2_{ji} \left(f \circ x^{-1} \right)$$

and

$$YXf = b^i \partial_i \left(a^j \circ x^{-1} \right) \partial_j \left(f \circ x^{-1} \right) + b^i a^j \partial^2_{ji} \left(f \circ x^{-1} \right)$$

in U. Note the last term in each of the above equations is symmetric by Schwarz's theorem, then

$$[X, Y]f = a^i \partial_i \left(b^j \circ x^{-1} \right) \partial_j \left(f \circ x^{-1} \right) - b^i \partial_i \left(a^j \circ x^{-1} \right) \partial_j \left(f \circ x^{-1} \right)$$
$$= \left(a^i \frac{\partial}{\partial x^i} b^j - b^i \frac{\partial}{\partial x^i} a^j \right) \frac{\partial}{\partial x^j} f,$$

hence

$$[X, Y] = \left(a^i \frac{\partial}{\partial x^i} b^j - b^i \frac{\partial}{\partial x^i} a^j \right) \frac{\partial}{\partial x^j}.$$

By Lemma 2.4, the component functions a^i, b^i are smooth maps, so the components of $[X, Y]$ are smooth. Therefore, $[X, Y]$ is a smooth vector field. □

We may regard the Lie bracket as a product of vector fields. We now present properties of this map.

> ### Proposition 2.8: Properties of the Lie bracket of vector fields
>
> The Lie bracket is a map
>
> $$[\cdot, \cdot] : \Gamma(TM) \times \Gamma(TM) \to \Gamma(TM)$$
> $$(X, Y) \mapsto [X, Y]$$
>
> that satisfies the following properties
>
> (a) Anticommutativity: $[X, Y] = -[Y, X]$;
>
> (b) $\mathbb{R}$-bilinearity: $[\alpha X + \beta Y, Z] = \alpha [X, Z] + \beta [Y, Z]$;
>
> (c) Jacobi identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$,
>
> for all $\alpha, \beta \in \mathbb{R}$, and $X, Y, Z \in \Gamma(TM)$.

Remark 2.30. *This result establishes $\Gamma(TM)$ as a* Lie algebra *over the $\mathbb{R}$-vector space $\Gamma(TM)$.*

Proof. From the definition, it is clear that $[X, Y] = -[Y, X]$. We consider $f \in C^\infty(M)$, then

$$
\begin{aligned}
[\alpha X + \beta Y, Z]f &= (\alpha X + \beta Y)Zf - Z(\alpha X + \beta Y)f \\
&= \alpha XZf + \beta YZf - \alpha ZXf - \beta ZYf \\
&= \alpha(XZf - ZXf) + \beta(YZf - ZYf) \\
&= \alpha[X, Z]f + \beta[Y, Z]f.
\end{aligned}
$$

Linearity in the second argument follows similarly or by successive applications of the anti-commutativity property. Finally, we have

$$
\begin{aligned}
[X, [Y, Z]]f &= X[Y, Z]f - [Y, Z]Xf \\
&= X(YZf - ZYf) - (YZXf - ZYXf) \\
&= XYZf - XZYf - YZXf + ZYXf,
\end{aligned}
$$

and by cyclic permutations, we obtain

$$
[Y, [Z, X]]f = YZXf - YXZf - ZXYf + XZYf
$$

and

$$
[Z, [X, Y]]f = ZXYf - ZYXf - XYZf + YXZf.
$$

It is easy to see that the sum of the last three equations yields the Jacobi identity. $\qquad\square$

2.5 Differential forms and exterior algebra

Recall S_k is the group of permutations on the finite set of natural numbers $\{1, \dots, k\}$. If $\pi \in S_k$ is a permutation, then its parity $\operatorname{sgn}(\pi)$ is 1 if π has an even amount of inversions, and -1 otherwise.

Definition 2.46: Differential form

Let M be a smooth manifold. A *(differential) k-form* is a $(0, k)$-tensor field ω on M that is totally antisymmetric, that is

$$
\omega(X_1, \dots, X_k) = \operatorname{sgn}(\pi) \cdot \omega(X_{\pi(1)}, \dots, X_{\pi(k)}),
$$

where $X_i \in \Gamma(TM)$, and $\pi \in S_k$.

Example 2.6. *(a) Smooth functions on M and covector fields are trivially differential forms. We may now refer to covector fields as 1-forms.*

(b) If M is orientable, then there exists a nowhere vanishing volume form.

(c) Electromagnetic field strength F is a 2-form.

(d) In classical mechanics, if Q is the configuration space, then its cotangent bundle T^*Q is the phase space. On T^*Q there exists a canonically defined 2-form, called the symplectic form.

We denote the set of all k-forms on M by $\Lambda^k(M)$, which can be equipped to be a $C^\infty(M)$-module, as it is a submodule of the set of all $(0, k)$-tensor fields on M. It is easy to see the tensor product of differential forms does not yield a form. Indeed, let $\omega, \eta \in \Lambda^1(M)$ be 1-forms, then

$$(\omega \otimes \eta)(X, Y) = \omega(X) \cdot \eta(Y), \tag{2.5.1}$$

which generically is not the same as $-(\omega \otimes \eta)(Y, X)$.

We wish to construct a $C^\infty(M)$-algebra of forms, where the algebra over a ring is defined by changing "field" to "ring" and "vector field" to "module" in Definition 2.28. A product between forms that yields forms must be defined, but we have just seen the tensor product is not enough.

Definition 2.47: Wedge product

Let M be a smooth manifold. The *wedge product* between two differential forms is the map

$$\wedge : \Lambda^k(M) \times \Lambda^\ell(M) \to \Lambda^{k+\ell}(M)$$
$$(\omega, \eta) \mapsto \omega \wedge \eta,$$

defined by

$$(\omega \wedge \eta)(X_1, \ldots X_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \operatorname{sgn}(\pi) \cdot (\omega \otimes \eta)\left(X_{\pi(1)}, \ldots, X_{\pi(k+\ell)}\right),$$

where $X_i \in \Gamma(TM)$.

Example 2.7. *In the case of 1-forms, we have $\omega \wedge \eta = \omega \otimes \eta - \eta \otimes \omega$, and it is easy to see $\omega \wedge \eta$ indeed is a 2-form and that $\omega \wedge \eta = -\eta \wedge \omega$. A similar result is generalized for any pair of differential forms, which is shown in Proposition 2.10.*

We now check the assertion made in the previous definition. That is, we must check if the wedge product yields a differential form.

Proposition 2.9: Wedge product yields a differential form

Let $\omega \in \Lambda^k(M)$ and $\eta \in \Lambda^\ell(M)$. Then $\omega \wedge \eta \in \Lambda^{k+\ell}(M)$.

Proof. Consider the differential forms $\omega \in \Lambda^k(M)$ and $\eta \in \Lambda^\ell(M)$, we have. It is easy to see $\omega \wedge \eta$ is a multilinear map, since ω and η are tensors and the wedge product is defined with the tensor product. That is, $\omega \wedge \eta$ is a $(0, k + \ell)$-tensor field, and we must now show it is a differential form.

Let $\tau \in S_{k+\ell}$ be a permutation and let $X_i \in \Gamma(TM)$ be vector fields for $i \in \{1, \ldots, k + \ell\}$. From the definition of the wedge product, we have

$$(\omega \wedge \eta)\left(X_{\tau(1)}, \ldots, X_{\tau(k+\ell)}\right) = \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \operatorname{sgn}(\pi) \cdot (\omega \otimes \eta)\left(X_{\pi \circ \tau(1)}, \ldots, X_{\pi \circ \tau(k+\ell)}\right).$$

For any $\pi \in S_{k+\ell}$, $\pi \circ \tau \in S_{k+\ell}$, since the permutation group is closed under composition. It

is clear that $\text{sgn}(\pi \circ \tau) = \text{sgn}(\pi) \cdot \text{sgn}(\tau)$, and we have

$$\text{sgn}(\tau) \cdot (\omega \wedge \eta) \left(X_{\tau(1)}, \ldots, X_{\tau(k+\ell)} \right) = \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \text{sgn}(\pi \circ \tau) \cdot (\omega \otimes \eta) \left(X_{\pi \circ \tau(1)}, \ldots, X_{\pi \circ \tau(k+\ell)} \right)$$

$$= \frac{1}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma) \cdot (\omega \otimes \eta) \left(X_{\sigma(1)}, \ldots, X_{\sigma(k+\ell)} \right)$$

$$= (\omega \wedge \eta) \left(X_1, \ldots, X_{k+\ell} \right).$$

This shows $\omega \wedge \eta$ is an alternating tensor, and this proves our claim. $\square$

> **Proposition 2.10: Wedge product properties**
>
> The wedge product is associative and graded anticommutative, meaning
>
> $$\omega \wedge \eta = (-1)^{k\ell} \eta \wedge \omega,$$
>
> and
>
> $$(\omega \wedge \eta) \wedge \alpha = \omega \wedge (\eta \wedge \alpha),$$
>
> for all $\omega \in \Lambda^k(M)$, $\eta \in \Lambda^\ell(M)$, and $\alpha \in \Lambda^m(M)$.

Proof. As the proof for associativity is cumbersome, we refer to [Mun91]. To show graded anticommutativity, we consider the permutation $\tau \in S_{k+\ell}$ such that

$$(\tau(1), \ldots, \tau(k+\ell)) = (k+1, \ldots, k+\ell, 1, \ldots, k).$$

We verify the parity of this permutation by counting the number of inversions of τ. It is easy to see that the only inversions are between elements $k + i$ and j, where $1 \le i \le \ell$ and $1 \le j \le k$. Moreover, for a given element $k + i$, there are exactly k inversions. Then, there are $k\ell$ inversions, hence $\text{sgn}(\tau) = (-1)^{k\ell}$.

Then, given vector fields $X_i \in \Gamma(TM)$, for $i \in \{1, \ldots, k+\ell\}$, it follows that

$$(\omega \otimes \eta)(X_1, \ldots, X_{k+\ell}) = (\eta \otimes \omega) \left(X_{\tau(1)}, \ldots, X_{\tau(k+\ell)} \right).$$

From the definition of the wedge product we have

$$(\omega \wedge \eta)(X_1, \ldots, X_{k+\ell}) = \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \text{sgn}(\pi) \cdot (\omega \otimes \eta) \left(X_{\pi(1)}, \ldots X_{\pi(k+\ell)} \right)$$

$$= \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \frac{\text{sgn}(\tau \circ \pi)}{\text{sgn}(\tau)} \cdot (\eta \otimes \omega) \left(X_{\tau \circ \pi(1)}, \ldots, X_{\tau \circ \pi(k+\ell)} \right)$$

$$= \frac{(-1)^{k\ell}}{k!\ell!} \sum_{\sigma \in S_{k+\ell}} \text{sgn}(\sigma)(\eta \otimes \omega) \left(X_{\sigma(1)}, \ldots, X_{\sigma(k+\ell)} \right)$$

$$= (-1)^{k\ell}(\eta \wedge \omega) \left(X_1, \ldots, X_{k+\ell} \right),$$

that is, $\omega \wedge \eta = (-1)^{k\ell} \eta \wedge \omega$, as desired. $\square$

Recall that a product in an algebra must not only be closed in the algebra, but it must be a bilinear map.

> **Proposition 2.11: Wedge product is $C^\infty(M)$-bilinear**
>
> Let $\alpha, \beta \in \Lambda^k(M)$, $\omega \in \Lambda^\ell(M)$ and $f \in C^\infty(M)$. Then
>
> $$(\alpha + f\beta) \wedge \omega = \alpha \wedge \omega + f(\beta \wedge \omega),$$
>
> that is, the wedge product is bilinear.

Proof. It is clear this property follows from the bilinearity of the tensor product, that is, we have $(\alpha + f\beta) \otimes \omega = \alpha \otimes \omega + f\beta \otimes \omega$. Let $X_i \in \Gamma(TM)$ be vector fields, with $i \in \{1, \ldots, k+\ell\}$, then

$$
\begin{aligned}
((\alpha + f\beta) \wedge \omega)(X_1, \ldots, X_{k+\ell}) &= \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \mathrm{sgn}(\pi) \cdot ((\alpha + f\beta) \otimes \omega)\left(X_{\pi(1)}, \ldots, X_{\pi(k+\ell)}\right) \\
&= \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \mathrm{sgn}(\pi) \cdot (\alpha \otimes \omega + f\beta \otimes \omega)\left(X_\pi(1), \ldots, X_{\pi(k+\ell)}\right) \\
&= (\alpha \wedge \omega + f\beta \wedge \omega)(X_1, \ldots, X_{k+\ell}),
\end{aligned}
$$

as desired. $\qquad\square$

We generalize the pullback of covectors to differential forms and show it is compatible with the wedge product.

> **Definition 2.48: Pullback of a differential form**
>
> Let $h : M \to N$ be a smooth map between smooth manifolds M and N. The *pullback of a k-form* $\omega \in \Lambda^k(N)$ *on* N is the *k*-form $h^*\omega \in \Lambda^k(M)$ *on* M defined by
>
> $$(h^*\omega)(p)(X_1(p), \ldots, X_k(p)) = \omega(h(p))\left(h_{*p}(X_1(p)), \ldots, h_{*p}(X_k(p))\right),$$
>
> for all points $p \in M$ and vector fields $X_i \in \Gamma(TM)$.

It is clear that $h^*\omega \in \Lambda^k(M)$, since it inherits both multilinearity and antisymmetry from the differential form $\omega \in \Lambda^k(N)$.

> **Proposition 2.12: Pullback is compatible with the wedge product**
>
> Let $h : M \to N$ be a smooth map between smooth manifolds M and N. Then
>
> $$h^*(\omega \wedge \eta) = (h^*\omega) \wedge (h^*\eta),$$
>
> for all $\omega \in \Lambda^k(N)$ and $\eta \in \Lambda^\ell(N)$.

Proof. We consider vector fields $X_i \in \Gamma(TM)$, for $i \in \{1, \ldots, k+\ell\}$ and differential forms $\omega \in \Lambda^k(N)$ and $\eta \in \Lambda^\ell(M)$. Then, for each point $p \in M$ we assign $Y_i = X_i(p) \in T_pM$ for aesthetic purposes. Then,

$$
\begin{aligned}
(h^*\omega \otimes h^*\eta)(p)(Y_1, \ldots, Y_{k+\ell}) &= h^*\omega(p)(Y_1, \ldots, Y_k) \cdot h^*\eta(p)(Y_{k+1}, \ldots, Y_{k+\ell}) \\
&= \omega(h(p))(h_{*p}Y_1, \ldots, h_{*p}Y_k) \cdot \eta(h(p))(h_{*p}Y_{k+1}, \ldots, H_{*p}Y_{k+\ell}) \\
&= (\omega \otimes \eta)(h(p))(h_{*p}Y_1, \ldots, h_{*p}Y_{k+\ell}),
\end{aligned}
$$

and as a result,

$$
\begin{aligned}
(h^*\omega \wedge h^*\eta)(p)(X_1(p), \dots, X_{k+\ell}(p)) &= \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \operatorname{sgn}(\pi)(h^*\omega \otimes h^*\eta)(p)(Y_{\pi(1)}, \dots, Y_{\pi(k+\ell)}) \\
&= \frac{1}{k!\ell!} \sum_{\pi \in S_{k+\ell}} \operatorname{sgn}(\pi)(\omega \otimes \eta)(h(p))(h_{*p}Y_{\pi(1)}, \dots, h_{*p}Y_{\pi(k+\ell)}) \\
&= (\omega \wedge \eta)(h(p))(h_{*p}Y_1, \dots, h_{*p}Y_{k+\ell}) \\
&= h^*(\omega \wedge \eta)(p)(X_1(p), \dots, X_{k+\ell}(p)),
\end{aligned}
$$

and we conclude $h^*\omega \wedge h^*\eta = h^*(\omega \wedge \eta)$ as desired. $\qquad \square$

We may now define the algebra with the $C^\infty(M)$-module of the direct sum of all differential forms, that is closed under the linear continuation of the wedge product.

Definition 2.49: Exterior algebra

Let M be an n-dimensional smooth manifold. The *exterior algebra* $\Lambda(M)$ is the $C^\infty(M)$-module

$$
\Lambda(M) = \bigoplus_{k=0}^{n} \Lambda^k(M)
$$

equipped with the *exterior product* $\wedge$, the bilinear map

$$
\begin{aligned}
\wedge : \Lambda(M) \times \Lambda(M) &\to \Lambda(M) \\
(\omega, \eta) &\mapsto \omega \wedge \eta,
\end{aligned}
$$

defined by linear continuation of the wedge product.

Remark 2.31. *Any differential form of rank greater than n is the zero tensor. To see this, consider a local chart, where the dimension of the tangent space is equal to n, that is, there are at most n linearly independent vectors, and thus any differential form of rank greater than n would only yield zero for any family of tangent vectors.*

Example 2.8. *We illustrate the meaning of the linear continuation of the wedge product. Recall an element u of a direct sum module $V \oplus W$ has a unique decomposition $u = v + w$, where $v \in V$ and $w \in W$. A differential form $\omega \in \Lambda(M)$ may be expressed as*

$$
\omega = \sum_{k=0}^{n} \omega^k,
$$

where $\omega^k \in \Lambda^k(M)$. Let $\eta \in \Lambda^\ell(M) \subset \Lambda(M)$. The exterior product $\omega \wedge \eta$ is then

$$
\omega \wedge \eta = \sum_{k=0}^{n} \omega^k \wedge \eta,
$$

where $\omega^k \wedge \eta$ is defined by the wedge form between a k-form and an ℓ-form, as before.

We warn, however, that the graded anticommutativity shown in Proposition 2.10 does not hold for a general element of the exterior algebra, and we may only use it for elements whose ranks are known, that is

$$
\eta \wedge \omega = \sum_{k=0}^{n} \eta \wedge \omega^k = \sum_{k=0}^{n} (-1)^{k\ell} \omega^k \wedge \eta,
$$

which is not, in general, related to $\omega \wedge \eta$.

Recall we have already defined the gradient operator

$$d : \Lambda^0(M) \to \Lambda^1(M)$$
$$f \mapsto df$$

where $(df)(p) = d_p f \in T_p^* M$, for all $p \in M$. This can be extended to differential forms of higher ranks.

Definition 2.50: Exterior derivative

The *exterior derivative* is the linear operator

$$d : \Lambda^k(M) \to \Lambda^{k+1}(M)$$
$$\omega \mapsto d\omega,$$

defined by

$$(d\omega)(X_1, \ldots, X_{k+1}) = \sum_{i=1}^{k+1} (-1)^{i+1} X_i \left(\omega(X_1, \ldots, \widehat{X}_i, \ldots, X_{k+1}) \right)$$
$$+ \sum_{i<j} (-1)^{i+j} \omega \left([X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_{k+1} \right),$$

where $X_i \in \Gamma(TM)$ are vector fields, with $\widehat{X}_m$ denoting the omission of X_m.

A few results must be obtained in order to show the exterior derivative is well-defined. Namely, the definition claims the operator is linear and that $d\omega$ is a $(k + 1)$-form. These are assumed without proof.

Theorem 2.12: Exterior derivative is an antiderivation

Let $\omega \in \Lambda^k(M)$ and $\sigma \in \Lambda^\ell(M)$, then

$$d(\omega \wedge \sigma) = d\omega \wedge \sigma + (-1)^k \omega \wedge d\sigma.$$

Theorem 2.13: Exterior differentiation commutes with the pullback

Let $h : M \to N$ be a smooth map between smooth manifolds M and N. Then

$$h^*(d\omega) = d(h^*\omega),$$

for all $\omega \in \Lambda^k(N)$.

By linear continuation, we may extend the exterior derivative to any differential form $\omega \in \Lambda(M)$.

Example 2.9. *We continue the examples given before.*

(a) *In Maxwell electrodynamics, the homogeneous equations may be expressed as $dF = 0$. In $\mathbb{R}^4$, this means $F = dA$, where A is the gauge potential.*

(b) *In classical mechanics, there exists a symplectic form $\omega \in \Lambda^2(T^*Q)$ with $d\omega = 0$. There is a natural 1-form defined by $\theta = p_i dq^i$, from which one defines the symplectic form $\omega = d\theta$. With this definition, one finds that $d\theta = 0$.*

Theorem 2.14: $d^2 = 0$

The operator $d \circ d : \Lambda^k(M) \to \Lambda^{k+2}(M)$ is the null operator. More succinctly, $d^2 = 0$.

Proof. In local coordinates,

$$d\omega = \partial_b \omega_{a_1 \dots a_k} dx^b \wedge dx^{a_1} \wedge \cdots \wedge dx^{a_k},$$

then

$$
\begin{aligned}
d(d\omega) &= \partial_c \partial_b \omega_{a_1 \dots a_k} dx^c \wedge dx^b \wedge dx^{a_1} \wedge \cdots \wedge dx^{a_k} \\
&= \partial_{(cb)} \omega_{[a_1 \dots a_k]} dx^{[c} \wedge dx^b \wedge dx^{a_1} \wedge \cdots \wedge dx^{a_k]} \\
&= 0.
\end{aligned}
$$

This result is extended for the entire manifold. $\qquad\qquad\square$

2.5.1 deRham cohomology

The last result implies that given the sequence of maps

$$\Lambda^0(M) \xrightarrow{d} \Lambda^1(M) \xrightarrow{d} \dots \xrightarrow{d} \Lambda^{k-1}(M) \xrightarrow{d} \Lambda^k(M) \xrightarrow{d} \Lambda^{k+1}(M) \dots \xrightarrow{d} \Lambda^n(M),$$

then the kernel of $d : \Lambda^k(M) \to \Lambda^{k+1}(M)$ contains the image of $d : \Lambda^{k-1}(M) \to \Lambda^k(M)$. A k-form is called *exact* if it lies in the image of d, and it is called *closed* if it lies in the kernel of d.

The submodule of exact k-forms is denoted by B^k and the submodule of closed k-forms is denoted by Z^k. The Poincaré lemma states that if the underlying manifold is $\mathbb{R}^n$, then $Z^k = B^k$, for $n > 0$. In cases where N^k is a proper subset of Z^k, how to relate these submodules?

Definition 2.51: deRham cohomology group

The *k-th deRham cohomology group* is the quotient space

$$H^k(M) = Z^k / B^k,$$

where the equivalence relation is given by

$$\omega \sim \sigma \iff \omega - \sigma \in B^k,$$

for $\omega, \sigma \in Z^k$.

Remarkably, the cohomology group $H^k(M)$ depends only on the *global* topology of M, by deRham's theorem. As an example, $H^0(M)$ is isomorphic to $\mathbb{R}^d$, where d is the number of connected pieces of M.

Chapter 3

Lie Groups and Lie Algebras

> **Definition 3.1: Lie group**
>
> A *Lie group* is a group $(G, \bullet)$, where G is a smooth manifold and the maps
>
> $$\mu : G \times G \to G$$
> $$(g_1, g_2) \mapsto g_1 \bullet g_2$$
>
> and
>
> $$i : G \to G$$
> $$g \mapsto g^{-1}$$
>
> are smooth. If the group $(G, \bullet)$ is commutative, then it is called a

Remark 3.1. *Usually the group operation will be simply denoted by $g_1 g_2 = g_1 \bullet g_2$.*

Example 3.1. *(a) The n-dimensional translation group $(\mathbb{R}^n, +)$ is a commutative Lie group.*

(b) The 1-sphere $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ is a commutative Lie group under complex multiplication, called $U(1)$.

(c) The set of invertible linear endomorphisms $GL(n, \mathbb{R}) = \{\phi \in \operatorname{End}(\mathbb{R}^n) : \det \phi \neq 0\}$ is a Lie group under composition, called the general linear group.

(d) Let V be a d-dimensional $\mathbb{R}$-vector space, equipped with a pseudo inner product *$(\cdot, \cdot) : V \times V \tilde{\to} \mathbb{R}$ satisfying*

- *bilinearity*
- *symmetry: $\forall v, w \in V : (v, w) = (w, v)$*
- *non-degeneracy: $\forall w \in V : (v, w) = 0 \implies v = 0$,*

then there are (up to isomorphism) only as many pseudo inner products on V as there are different signatures. The signature (p, q) of a pseudo inner product is the number of positive p numbers and negative n numbers on the main diagonal of the diagonalization of the matrix representation of the pseudo inner product. The set

$$O(p, q) = \left\{ \phi \in \operatorname{End}(V) : (\phi(v), \phi(w)) = (v, w), \forall v, w \in V \right\} \subset GL(p + q, \mathbb{R})$$

is a Lie group under composition, called the orthogonal group *with respect to the pseudo inner product $(\cdot, \cdot)$.*

3.1   The Lie algebra of a Lie group

For any $g \in G$ there is a map

$$\ell_g : G \to G$$
$$h \mapsto \ell_g(h),$$

where $\ell_g(h) = gh$, called the *left translation* with respect to g. One could define *right transla-tions* similarly, and the following results would follow analogously.

> **Proposition 3.1: Left translations are diffeomorphisms**
>
> Let $g \in G$. The left translation ℓ_g with respect to g is a diffeomorphism.

Proof. We check ℓ_g is an isomorphism. Let $h, h' \in G$ such that $\ell_g h = \ell_g h'$. Then, applying $\ell_{g^{-1}}$ to both sides, we get $h = h'$, that is, ℓ_g is injective. Let $f \in G$, and consider $\ell_{g^{-1}} f = g^{-1} f$. Then, $\ell_g(g^{-1} f) = \mathrm{id}_G f = f$, that is, ℓ_g is surjective. Note $\ell_g \circ \ell_{g^{-1}} = \mathrm{id}_G$ and $\ell_{g^{-1}} \circ \ell_g = \mathrm{id}_G$.

Recall $\mu : G \times G \to G$ is a smooth map from $G \times G$ to G. Then ℓ_g is the restriction of μ to $\{g\} \times G$, therefore it is smooth. □

Remark 3.2. *A group isomorphism is a map that preserves the group operation. Let $h_1, h_2 \in G$, then*

$$\ell_g(h_1 h_2) = g h_1 h_2$$

which is not generically equal to $(\ell_g h_1)(\ell_g h_2)$, that is, ℓ_g is not necessarily a group isomorphism.

Recall we could extend the pullback at a point of covectors to differential forms without problems. However, the same cannot be said to the pushforward, unless the underlying smooth map is a diffeomorphism. Indeed, let $h : M \to N$ be a smooth map between smooth manifolds M and N. Then $h(M) \subset N$ and to each point $p \in M$ it is assigned a point $h(p) \in N$. As such, we defined the pullback at p by taking the differential form at $h(p)$, thus defining a differential form throughout M. This is not the case for the pushforward, since $h(M)$ may be a proper subset of N, and therefore it is not possible to have a vector field throughout N. Even worse, h may not be injective, so two points in M may me mapped to the same point in N, and an attempt at establishing the pushforward of a vector field would be ill-defined.

Since the left translations are diffeomorphisms, it is possible to define the pushforward of vector fields in G. Namely, given $g \in G$, it is the map

$$\ell_{g_*} : \Gamma(TG) \to \Gamma(TG)$$
$$X \mapsto \ell_{g_*} X$$

where

$$\ell_{g_*} X : G \to TG$$
$$h \mapsto (\ell_{g_*} X)(h)$$

and $T_h G \ni (\ell_{g_*} X)(h) = \ell_{g_* g^{-1} h}(X(g^{-1} h)) \in T_{g^{-1} h} G$. Since ℓ_g is an isomorphism, a point $h' \in G$ may always be expressed as $h' = gh$ for some $h \in G$, and we may rewrite this last definition to

$$(\ell_{g_*} X)(gh) = \ell_{g_* h}(X(h)).$$

As an attempt to make the notation less busy, we note the points at which the vector field and the pushforward are evaluated must coincide, so we denote last equation as

$$\left(\ell_{g_*}X\right)_{gh} = \ell_{g_*}(X_h),$$

where the subscript after a vector field denotes $X_h = X(h)$.

Considering vector fields as derivations, we have

$$\left(\ell_{g_*}X\right)\varphi = X(\varphi \circ \ell_g),$$

for all $\varphi \in C^\infty(G)$. Since φ and ℓ_g are smooth, it follows that $\ell_{g_*}X$ is a smooth vector field.

Lemma 3.1: Composition of left translation pushforwards

Let $g, h \in G$. Then $\ell_{g_*} \circ \ell_{h_*} = \ell_{gh_*}$.

Proof. Let $X \in \Gamma(TG)$ be a smooth vector field and let $\varphi \in C^\infty(G)$ be a smooth map. Then, by the definition of the pushforward, we have

$$\begin{aligned}
\left(\ell_{g_*} \circ \ell_{h_*}X\right)\varphi &= \left(\ell_{h_*}X\right)\left(\varphi \circ \ell_g\right) \\
&= X\left(\varphi \circ \ell_g \circ \ell_h\right) \\
&= X\left(\varphi \circ \ell_{gh}\right) \\
&= \left(\ell_{gh_*}X\right)\varphi,
\end{aligned}$$

which proves the statement. $\qquad\square$

Definition 3.2: Left invariant vector field

A vector field is *left invariant* if the vector field is invariant by the pushforward of a left translation. Precisely, a vector field $X : M \to TM$ is left invariant if

$$\ell_{g_*}X = X,$$

for any $g \in G$.

Lemma 3.2: Left invariant vector fields are smooth

Let $X : G \to TG$ be a vector field. If X is left invariant, then X is smooth.

Proof. Let $\varphi \in C^\infty(G)$ be a smooth map and let $\gamma : (-\varepsilon, \varepsilon) \to G$ be a smooth curve such that $\gamma(0) = e$ and that its tangent vector at e is X_e. Then, at a point $g \in G$, we have

$$\begin{aligned}
X_g\varphi &= \ell_{g_*}(X_e)\varphi \\
&= X_e(\varphi \circ \ell_g) \\
&= (\varphi \circ \ell_g \circ \gamma)'(0),
\end{aligned}$$

thus $X_g\varphi$ depends smoothly on g, so $X \in \Gamma(TG)$. $\qquad\square$

The set of all left invariant vector fields on the Lie group G is denoted by $\mathfrak{g} \subset \Gamma(G)$.

> **Proposition 3.2: Equivalent definitions of left invariance**
>
> Let $X \in \Gamma(TG)$ be a vector field. The following statements are equivalent
>
> (a) X is left invariant.
>
> (b) $\ell_{g_*}(X_h) = X_{gh}$ for all $g, h \in G$.
>
> (c) $X(\varphi \circ \ell_g) = (X\varphi) \circ \ell_g$, for all $\varphi \in C^\infty(G)$ and $g \in G$.

Proof. Suppose X is left invariant, that is, $\ell_{g_*}X = X$ for all $g \in G$. Then, by the definition of pushforward of a vector field by a diffeomorphism, at $h \in G$ we have

$$\ell_{g_*}(X_h) = (\ell_{g_*}X)_{gh}$$
$$= X_{gh}.$$

Since h is arbitrary, (a) $\iff$ (b).
  Suppose X satisfies (b), then for all $g, h \in G$ and any smooth map $\varphi \in C^\infty(G)$

$$\ell_{g_*}(X_h)\varphi = X_h(\varphi \circ \ell_g)$$
$$= [X(\varphi \circ \ell_g)](h),$$

by the definition of pushforward at a point. Note that

$$X_{gh}\varphi = (X\varphi)(gh) = (X\varphi) \circ \ell_g(h),$$

then we may rewrite (b) as

$$[X(\varphi \circ \ell_g)](h) = (X\varphi) \circ \ell_g(h).$$

Since h is arbitrary, (b) $\iff$ (c). $\qquad\square$

  Restricting the $\mathbb{R}$-vector space operations of $\Gamma(TG)$ to $\mathfrak{g}$ we may check that $\mathfrak{g}$ is a $\mathbb{R}$-vector subspace of $\Gamma(TG)$. Indeed, let $X, Y \in \mathfrak{g}$ and $\alpha, \beta \in \mathbb{R}$, then

$$\ell_{g_*}(\alpha X + \beta Y)_{gh} = \ell_{g_*}(\alpha X_h + \beta Y_h)$$
$$= \alpha \ell_{g_*}(X_h) + \beta \ell_g(Y_h)_*$$
$$= \alpha X_{gh} + \beta Y_{gh}$$

for all $g, h \in G$. That is, $\alpha X + \beta Y \in \mathfrak{g}$. Even though the $\mathbb{R}$-vector space of smooth vector fields may be infinite dimensional, this is not the case for $\mathfrak{g}$, as will be shown.

> **Definition 3.3: Abstract Lie algebra**
>
> An *abstract Lie algebra* $(L, +, \cdot, [\cdot, \cdot])$ is a $\mathbb{K}$-vector space $(L, +, \cdot)$ equipped with an abstract *Lie bracket* $[\cdot, \cdot] : L \times L \to L$ that satisfies
>
> (a) antisymmetry: $[x, y] = -[y, x]$, for all $x, y \in L$.
>
> (b) $\mathbb{K}$-bilinearity: $[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z]$, for all $x, y, z \in L$ and $\alpha, \beta \in \mathbb{K}$.
>
> (c) Jacobi identity: $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

Example 3.2. *(a) We recall the commutator of two smooth vector fields satisfies these properties, so the $\mathbb{R}$-vector space of smooth vector fields is an infinite-dimensional Lie algebra.*

(b) *The three dimensional Euclidean space equipped with the vector product is a Lie algebra.*

(c) *The endomorphisms on a vector space equipped with the bracket defined by*

$$[\phi, \psi] = \phi \circ \psi - \psi \circ \phi$$

is a Lie algebra.

A Lie subalgebra is a vector subspace of a Lie algebra that is closed under the Lie bracket.

> **Theorem 3.1: The vector space of left invariant vector fields is a Lie subalgebra**
>
> The $\mathbb{R}$-vector space $(\mathfrak{g}, [\cdot, \cdot])$ is a Lie subalgebra of $(\Gamma(G), [\cdot, \cdot])$.

Proof. As noted, $\mathfrak{g} \subset \Gamma(TG)$, so it remains to be shown that the Lie bracket of left invariant vector fields is also left invariant. Let $X, Y \in \mathfrak{g}$ be left invariant vector fields. Then, for any $g \in G$ and $\varphi \in C^\infty(G)$,

$$\begin{aligned}
[X, Y](\varphi \circ \ell_g) &= X(Y(\varphi \circ \ell_g)) - Y(X(\varphi \circ \ell_g)) \\
&= X(Y\varphi \circ \ell_g) - Y(X\varphi \circ \ell_g) \\
&= X(Y\varphi) \circ \ell_g - Y(X\varphi) \circ \ell_g \\
&= [X, Y]\varphi \circ \ell_g,
\end{aligned}$$

that is, $[X, Y]$ is a left invariant vector field. $\qquad\square$

Remark 3.3. *The Lie algebra $\mathfrak{g}$ is called the* associated Lie algebra to the Lie group G.

> **Definition 3.4: Lie algebra homomorphism**
>
> Let $\mathfrak{g}_1$ and $\mathfrak{g}_2$ be Lie algebras over the same field $\mathbb{K}$. The vector space homomorphism $\phi : \mathfrak{g}_1 \xrightarrow{\sim} \mathfrak{g}_2$ is a *Lie algebra homomorphism* if
>
> $$\phi([x, y]) = [\phi(x), \phi(y)]$$
>
> for all $x, y \in \mathfrak{g}_1$. If ϕ is a bijection, then it is a *Lie algebra isomorphism*.

With a Lie algebra isomorphism, it would be possible to avoid dealing with left invariant vector fields by identifying another, possibly less complicated, Lie algebra with $\mathfrak{g}$. We now construct such an isomorphism with the tangent space at the identity.

> **Theorem 3.2: Left invariant vector fields are isomorphic to a tangent space**
>
> The vector space of left invariant vector fields is isomorphic to the tangent space $T_e G$.

Proof. Consider the map

$$\begin{aligned}
j : T_e G &\xrightarrow{\sim} \mathfrak{g} \\
A &\mapsto j(A)
\end{aligned}$$

where $j(A)_g = \ell_{g *e} A$ for all $g \in G$. It follows from the $\mathbb{R}$-linearity of the pushforward at a point that j is indeed a linear map.

We verify the vector field $j(A)$ is indeed smooth. Let $\varphi \in C^\infty(G)$ be a smooth map, then for all $g \in G$,

$$
\begin{aligned}
j(A)_g \varphi &= (\ell_{g_{*e}} A)\varphi \\
&= A(\varphi \circ \ell_g) \\
&= (\varphi \circ \ell_g \circ \gamma)'(0),
\end{aligned}
$$

where $\gamma : \mathbb{R} \to G$ is a smooth curve with $\gamma(0) = e$ and with tangent vector A at e. We consider the map

$$
\begin{aligned}
\psi : \mathbb{R} \times G &\to \mathbb{R} \\
(t, g) &\mapsto \varphi \circ \ell_g \circ \gamma(t),
\end{aligned}
$$

which is smooth, as a composition of smooth maps. Then, $j(A)_g \varphi = (\partial_1 \psi)(0, g)$, which depends smoothly on g, hence $j(A)\varphi$ is a smooth function and as a consequence, $j(A) \in \Gamma(TG)$.

To show $j(A)$ is left invariant, consider

$$
\ell_{g_*}(j(A)_h) = \ell_{g_*}(\ell_{h_{*e}} A).
$$

By Lemma 3.1, we have

$$
\begin{aligned}
\ell_{g_*}(j(A)_h) &= \ell_{gh_{*e}} A \\
&= j(A)_{gh}.
\end{aligned}
$$

By Proposition 3.2, $j(A) \in \mathfrak{g}$.

Let $A, B \in T_e G$ such that $j(A) = j(B)$. That is, for every $g \in G$, $j(A)_g = j(B)_g$. In particular, $j(A)_e = j(B)_e$, which by definition is equivalent to $\ell_{e_*} A = \ell_{e_*} B$. The left translation by the identity element is the identity map, thus $A = B$, so j is injective.

Let $X \in \mathfrak{g}$ and consider $X_e \in T_e G$. We have

$$
j(X_e)_g = \ell_{g_*}(X_e) = X_{ge} = X_g,
$$

therefore $j(X_e) = X$. Thus, j is surjective. □

Corollary 3.1. *The dimension of the vector space of left invariant vector fields is the same as the dimension of the manifold G.*

Remark 3.4. *This shows $\mathfrak{g}$ is a finite-dimensional $\mathbb{R}$-vector subspace of the infinite-dimensional $\mathbb{R}$-vector space $\Gamma(TG)$.*

Since $T_e G$ is isomorphic to $\mathfrak{g}$, we would like to equip $T_e G$ with a Lie bracket such that

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

in which case the isomorphism would not only identify uniquely elements from both vector spaces, but also preserve the Lie algebra structure of $\mathfrak{g}$. We verify the map

$$
\begin{aligned}
[\cdot, \cdot] : T_e G \times T_e G &\to T_e G \\
(A, B) &\mapsto j^{-1}([j(A), j(B)])
\end{aligned}
$$

is indeed a Lie bracket. Let $X, Y \in T_e G$, then

$$
\begin{aligned}
[Y, X] &= j^{-1}([j(Y), j(X)]) \\
&= j^{-1}(-[j(X), j(Y)]) \\
&= -j^{-1}([j(X), j(Y)]) \\
&= -[X, Y],
\end{aligned}
$$

that is, this map is antisymmetric. Let $Z \in T_e G$ and $\alpha, \beta \in \mathbb{R}$, then

$$
\begin{aligned}
[\alpha X + \beta Y, Z] &= j^{-1}([j(\alpha X + \beta Y), j(Z)]) \\
&= j^{-1}([j(\alpha X), j(Z)] + [j(\beta Y), j(Z)]) \\
&= j^{-1}(\alpha[j(X), j(Z)] + \beta[j(Y), j(Z)]) \\
&= \alpha j^{-1}([j(X), j(Z)]) + \beta j^{-1}(j(Y), j(Z)) \\
&= \alpha[X, Z] + \beta[Y, Z],
\end{aligned}
$$

thus the map is bilinear. Since j^{-1} is linear, the Jacobi identity follows from the property of the Lie bracket in $\mathfrak{g}$. This has shown the map j establishes a Lie algebra isomorphism between $T_e G$ and $\mathfrak{g}$.

3.2   Classification of Lie algebras

In this section, Lie algebras will be studied as their own objects.

> **Definition 3.5: Nilpotent and solvable Lie algebras**
>
> Let $\mathfrak{g}$ be a Lie algebra. We denote $\mathfrak{g}^{[n]}$ the sequence of sets given by $\mathfrak{g}^{[0]} = \mathfrak{g}$ and $\mathfrak{g}^{[n+1]} = [\mathfrak{g}, \mathfrak{g}^{[n]}]$ for all $n \in \mathbb{N}$. A *nilpotent* Lie algebra satisfies $\mathfrak{g}^{[m]} = \{0\}$ for some $m \in \mathbb{N}$. Similarly, we denote $\mathfrak{g}^{(n)}$ the sequence of sets given by $\mathfrak{g}^{(0)} = \mathfrak{g}$ and $\mathfrak{g}^{(n+1)} = [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}]$, for all $n \in \mathbb{N}$. A *solvable* Lie algebra satisfies $\mathfrak{g}^{(m)} = \{0\}$ for some $m \in \mathbb{N}$.

Remark 3.5. *It is easy to show that every nilpotent algebra is a solvable algebra: it suffices to show* $\mathfrak{g}^{(n)} \subset \mathfrak{g}^{[n]}$ *by induction. Indeed,* $\mathfrak{g}^{(1)} = \mathfrak{g}^{[1]}$ *and by the induction hypothesis,*

$$
\begin{aligned}
\mathfrak{g}^{(n+1)} = [\mathfrak{g}^{(n)}, \mathfrak{g}^{(n)}] &\subset [\mathfrak{g}, \mathfrak{g}^{(n)}] \\
&\subset [\mathfrak{g}, \mathfrak{g}^{[n]}] = \mathfrak{g}^{[n+1]}.
\end{aligned}
$$

The converse result, however, does not hold, that is, not every solvable Lie algebra is nilpotent.

> **Definition 3.6: Ideals, simple and semi-simple Lie algebras**
>
> A vector subspace I of a Lie algebra $\mathfrak{g}$ is an *ideal* if $[\mathfrak{g}, I] \subset I$. A non-abelian Lie algebra $\mathfrak{g}$ is *simple* if there are no non-trivial proper ideals. A Lie algebra $\mathfrak{g}$ is *semisimple* if it has no non-trivial solvable ideals.

Remark 3.6. *It is clear every ideal is a Lie subalgebra and that every simple Lie algebra is also semisimple.*

We now extend the concept of direct sum of vector spaces to Lie algebras.

> **Definition 3.7: Direct sum and semidirect sum of Lie algebras**
>
> A Lie algebra $\mathfrak{g}$ is a *direct sum* of two of its Lie subalgebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ if $[\mathfrak{g}_1, \mathfrak{g}_2] = \{0\}$ and if for every element $x \in \mathfrak{g}$ there exist unique $x_1 \in \mathfrak{g}_1$ and $x_2 \in \mathfrak{g}_2$ such that $x = x_1 + x_2$. In this case, we denote $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$.
> A Lie algebra $\mathfrak{g}$ is a *semidirect sum* of two of its Lie subalgebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$ if $[\mathfrak{g}_1, \mathfrak{g}_2] \subset \mathfrak{g}_1$ and if for every element $x \in \mathfrak{g}$ there exist unique $x_1 \in \mathfrak{g}_1$ and $x_2 \in \mathfrak{g}_2$ such that $x = x_1 + x_2$. In this case, we denote $\mathfrak{g} = \mathfrak{g}_1 \ltimes \mathfrak{g}_2$.

> **Proposition 3.3: Semisimple Lie algebra is a direct sum of simple Lie algebras**
>
> A Lie algebra $\mathfrak{g}$ is semisimple if and only if it may be expressed as
>
> $$\mathfrak{g} = \bigoplus_{i=1}^{n} \mathfrak{g}_i,$$
>
> where $\mathfrak{g}_i$ is a simple Lie algebra.

Proof. TODO! □

We now quote an important theorem that reduces the classification of finite-dimensional Lie algebras to the classification of solvable Lie algebras and semisimple Lie algebras.

> **Theorem 3.3: Levi's theorem**
>
> Every finite-dimensional Lie algebra $(\mathfrak{g}, [\cdot, \cdot])$ over a field $\mathbb{K}$ with characteristic zero can be decomposed as
>
> $$\mathfrak{g} = R \ltimes S,$$
>
> where R is a *solvable* ideal of $\mathfrak{g}$ and S is a semisimple Lie subalgebra.

Remark 3.7. *The characteristic of $\mathbb{R}$ and $\mathbb{C}$ is zero.*

Considering Proposition 3.3, we may classify semisimple Lie algebras with the simple Lie algebras as building blocks.

3.2.1 Adjoint endomorphism and Killing form

The Lie bracket defines a natural endomorphism, called the *adjoint endomorphism, adjoint action, or adjoint map*.

> **Definition 3.8: Adjoint endomorphism**
>
> Let $\mathfrak{g}$ be a Lie algebra and let $x \in \mathfrak{g}$ the linear map
>
> $$\mathrm{ad}_x : \mathfrak{g} \xrightarrow{\sim} \mathfrak{g}$$
> $$y \mapsto [x, y]$$
>
> is called the *adjoint endomorphism with respect to x*.

The linearity of ad_x follows directly from the bilinearity of the Lie bracket. By the same token, the map

$$\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$$
$$x \mapsto \mathrm{ad}_x$$

is linear. Recall the set of endomorphisms of a vector space has a natural Lie algebra defined with the commutator of endomorphisms. We may now show the map ad establishes a Lie algebra homomorphism between $\mathfrak{g}$ and $\mathrm{End}(\mathfrak{g})$.

Let $x, y, z \in \mathfrak{g}$, then

$$\mathrm{ad}_{[x,y]} z = [[x, y], z]$$
$$= [x, [y, z]] + [y, [z, x]],$$

by the Jacobi identity and the antisymmetric property. By rearranging with the antisymmetric property, we obtain the desired result,

$$
\begin{aligned}
\mathrm{ad}_{[x,y]}\, z &= [x, \mathrm{ad}_y\, z] - [y, \mathrm{ad}_x\, z] \\
&= \left(\mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x\right) z \\
&= \left[\mathrm{ad}_x, \mathrm{ad}_y\right] z,
\end{aligned}
$$

that is, ad is a Lie algebra homomorphism from $\mathfrak{g}$ to $\mathrm{End}(\mathfrak{g})$.

> **Definition 3.9: Killing form**
>
> The bilinear map
>
> $$
> \begin{aligned}
> K : \mathfrak{g} \times \mathfrak{g} &\tilde{\to} \mathbb{K} \\
> (x, y) &\mapsto \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_y\right)
> \end{aligned}
> $$
>
> is called the *Killing form*.

Remark 3.8. *If $\mathfrak{g}$ is finite-dimensional, then the trace is cyclic, and as a result the Killing form is symmetric.*

Remark 3.9. *A Lie algebra $\mathfrak{g}$ is semisimple if and only if K is non-degenerate:*

$$
\forall x \in \mathfrak{g} : K(x, y) = 0 \implies y = 0.
$$

Equivalently, in (semi)simple Lie algebras, the Killing form is a pseudo inner product. TODO!*prove this at least for simple algebras.*

Let $x, y, z \in \mathfrak{g}$, then for a finite-dimensional Lie algebra, we have

$$
\begin{aligned}
K([x, y], z) &= \mathrm{tr}\left(\mathrm{ad}_{[x,y]} \circ \mathrm{ad}_z\right) \\
&= \mathrm{tr}\left(\left[\mathrm{ad}_x, \mathrm{ad}_y\right] \circ \mathrm{ad}_z\right) \\
&= \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_y \circ \mathrm{ad}_z - \mathrm{ad}_y \circ \mathrm{ad}_x \circ \mathrm{ad}_z\right) \\
&= \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_y \circ \mathrm{ad}_z\right) - \mathrm{tr}\left(\mathrm{ad}_y \circ \mathrm{ad}_x \circ \mathrm{ad}_z\right) \\
&= \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_y \circ \mathrm{ad}_z\right) - \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_z \circ \mathrm{ad}_y\right) \\
&= \mathrm{tr}\left(\mathrm{ad}_x \circ [\mathrm{ad}_y, \mathrm{ad}_z]\right) \\
&= \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_{[y,z]}\right) \\
&= K(x, [y, z]).
\end{aligned}
$$

Recall a linear map $\varphi \in \mathrm{End}(V)$ is called symmetric with respect to a pseudo inner product B if

$$
B(\varphi(v), w) = B(v, \varphi(w))
$$

and antisymmetric if

$$
B(\varphi(v), w) = -B(v, \varphi(w)),
$$

for all $v, w \in V$. We may show the adjoint endomorphisms are antisymmetric with respect to the Killing form. Let $x, y, z \in \mathfrak{g}$, then as before, we have

$$
\begin{aligned}
K(\mathrm{ad}_x y, z) &= \mathrm{tr}\left(\mathrm{ad}_x \circ \mathrm{ad}_y \circ \mathrm{ad}_z\right) - \mathrm{tr}\left(\mathrm{ad}_y \circ \mathrm{ad}_x \circ \mathrm{ad}_z\right) \\
&= \mathrm{tr}\left(\mathrm{ad}_y \circ \mathrm{ad}_z \circ \mathrm{ad}_x\right) - \mathrm{tr}\left(\mathrm{ad}_y \circ \mathrm{ad}_x \circ \mathrm{ad}_z\right) \\
&= -\,\mathrm{tr}\left(\mathrm{ad}_y \circ \mathrm{ad}_{[x,z]}\right) \\
&= -K(y, \mathrm{ad}_x z)
\end{aligned}
$$

as desired.

We now turn our attention to the special case of a n-dimensional complex Lie algebra $\mathfrak{g}$. Let $\{E_1, \ldots, E_n\}$ be a basis of $\mathfrak{g}$, and let $\{\epsilon^1, \ldots, \epsilon^n\}$ be the dual basis. We define the *structure coefficients* $C^k{}_{ij} \in \mathbb{C}$ of $\mathfrak{g}$ with respect to the chosen basis by

$$[E_i, E_j] = C^k{}_{ij} E_k.$$

The antisymmetry of the Lie bracket may be expressed as

$$C^k{}_{ij} = -C^k{}_{ji}$$

and the Jacobi identity as

$$C^a{}_{ib} C^b{}_{jk} + C^a{}_{jb} C^b{}_{ki} + C^a{}_{kb} C^b{}_{ij} = 0.$$

Then, the components of the adjoint endomorphism with respect to the chosen basis are given by

$$
\begin{aligned}
(\mathrm{ad}_{E_i})^k{}_j &= \epsilon^k(\mathrm{ad}_{E_i}(E_j)) \\
&= \epsilon^k([E_i, E_j]) \\
&= \epsilon^k(C^m{}_{ij} E_m) \\
&= C^m{}_{ij} \delta^k_m \\
&= C^k{}_{ij},
\end{aligned}
$$

therefore the adjoint map has the same components as the structure coefficients. The components of the Killing form with respect to this basis are

$$
\begin{aligned}
K_{ij} &= K(E_i, E_j) \\
&= \mathrm{tr}\left(\mathrm{ad}_{E_i} \circ \mathrm{ad}_{E_j}\right) \\
&= C^k{}_{im} C^m{}_{jk}.
\end{aligned}
$$

3.2.2 The fundamental roots and the Weyl group

From now on, we concern ourselves with a n-dimensional semisimple complex Lie algebra $\mathfrak{g}$.

Definition 3.10: Cartan subalgebra

A Cartan subalgebra $\mathfrak{h}$ is a maximal Lie subalgebra of $\mathfrak{g}$ such that there exists a basis $\{h_1, \ldots, h_m\}$ of $\mathfrak{h}$ that can be extended to a basis $\{h_1, \ldots, h_m, e_1, \ldots, e_{n-m}\}$ of $\mathfrak{g}$, where e_α is an eigenvector of ad_h for any $h \in \mathfrak{h}$, that is

$$\forall h \in \mathfrak{h} : \exists \lambda_\alpha(h) \in \mathbb{C} : \mathrm{ad}_h\, e_\alpha = \lambda_\alpha(h) e_\alpha.$$

The basis $\{h_1, \ldots, h_m, e_1, \ldots, e_{n-m}\}$ is called the *Cartan-Weyl basis*.

Remark 3.10. *Any finite-dimensional Lie algebra possess a Cartan subalgebra. Moreover, if $\mathfrak{g}$ is simple, then $\mathfrak{h}$ is abelian, $[\mathfrak{h}, \mathfrak{h}] = \{0\}$.*

From the definition, we note $\lambda_\alpha : \mathfrak{h} \to \mathbb{C}$ is a linear map, that is $\lambda_\alpha \in \mathfrak{h}^*$. Also, we have $\lambda_\alpha \neq 0$, otherwise $e_\alpha \in \mathfrak{h}$, that is, $\mathfrak{h}$ wouldn't be maximal. TODO!

> **Definition 3.11: Roots of the Lie algebra**
>
> The linear functionals $\lambda_1, \ldots, \lambda_{n-m} \in \mathfrak{h}^*$ are called the *roots* of the Lie algebra $\mathfrak{g}$. The set
>
> $$\Phi = \{\lambda_1, \ldots, \lambda_{n-m}\} \subset \mathfrak{h}^*$$
>
> is called the *root set*.

Remark 3.11. *TODO!It follows from the antisymmetry of the adjoint endomorphisms with respect to the Killing form that if $\lambda \in \Phi$, then $-\lambda \in \Phi$. In particular, the set Φ is not linearly independent.*

> **Definition 3.12: Fundamental roots**
>
> A subset $\Pi \subset \Phi$ is a *set of fundamental roots* if $\Pi = \{\pi_1, \ldots, \pi_f\}$ is linearly independent and if, for all $\lambda \in \Phi$, there exists $N_1, \ldots, N_f \in \mathbb{N}$ and $\epsilon \in \{-1, 1\}$ such that
>
> $$\lambda = \epsilon \sum_{i=1}^{f} N_i \pi_i.$$
>
> This last expression may be denoted concisely by $\lambda \in \mathrm{span}_{\epsilon, \mathbb{N}} \Pi$ and we emphasize $\mathrm{span}_{\mathbb{Z}} \Pi \neq \mathrm{span}_{\epsilon, \mathbb{N}} \Pi$.

> **Theorem 3.4: Existence of fundamental roots**
>
> Let $\mathfrak{g}$ be a finite-dimensional complex Lie algebra with Cartan subalgebra $\mathfrak{h}$. Then, a set of fundamental roots Π exists and it is a basis for $\mathfrak{h}^*$, that is $\mathrm{span}_{\mathbb{C}} \Pi = \mathfrak{h}^*$.

Remark 3.12. *As with any basis, Π is not unique.*

We now construct a pseudo inner product on $\mathfrak{h}^*$ from the Killing form on $\mathfrak{g}$. First, the Killing form induces a linear isomorphism

$$\psi : \mathfrak{g} \xrightarrow{\sim} \mathfrak{g}^*$$
$$x \mapsto \psi(x),$$

where $\psi(x)(y) = K(x, y)$. With such a map, we may define

$$K^* : \mathfrak{h}^* \times \mathfrak{h}^* \xrightarrow{\sim} \mathbb{C}$$
$$(x, y) \mapsto K|_{\mathfrak{h}} \left(\psi^{-1}(x), \psi^{-1}(y) \right).$$

It is clear that, if $K|_{\mathfrak{h}}$ is a pseudo inner product, then K^* is a pseudo inner product, since ψ^{-1} is a linear isomorphism. By being a restriction of a pseudo inner product, we must only check non-degeneracy. Let $\{h_1, \ldots, h_m, e_1, \ldots, e_{n-m}\}$ be a Cartan-Weyl basis of $\mathfrak{g}$, and let $\lambda_\alpha \in \Phi$. Since λ_α is not the zero linear functional, there exists $h' \in \mathfrak{h}$ such that $\lambda_\alpha(h') \neq 0$, then

$$\begin{aligned}
\lambda_\alpha(h') K(h_i, e_\alpha) &= K(h_i, \lambda_\alpha(h') e_\alpha) \\
&= K(h_i, \mathrm{ad}_{h'} e_\alpha) \\
&= K(\mathrm{ad}_{h'} h_i, e_\alpha) \\
&= 0,
\end{aligned}$$

because $\mathfrak{h}$ is abelian. By bilinearity, this shows

$$\text{span}_{\mathbb{C}}\{e_1, \ldots, e_{n-m}\} = \mathfrak{h}^{\perp},$$

that is, for all $x \in \mathfrak{g}$ and all $h \in \mathfrak{h}$

$$K(h, x) = K(h, \text{proj}_{\mathfrak{h}} x).$$

Since K is non-degenerate,

$$(\forall x \in \mathfrak{h} : K(h, x) = 0) \implies h = 0,$$

that is, $K|_{\mathfrak{h}}$ is non-degenerate. This shows the bilinear map K_* is a pseudo inner product.

Consider $\mathfrak{h}_{\mathbb{R}}^* = \text{span}_{\mathbb{R}} \Pi \subset \mathfrak{h}$, then we have the chain of inclusions

$$\Pi \subset \Phi \subset \text{span}_{\epsilon, \mathbb{N}} \Pi \subset \mathfrak{h}_{\mathbb{R}}^* \subset \mathfrak{h}^*.$$

The restriction of K^* to $\mathfrak{h}_{\mathbb{R}}^*$ leads to the following surprising result.

> **Theorem 3.5: Inner product on $\mathfrak{h}_{\mathbb{R}}^*$**
>
> Let restriction of the pseudo inner product K^* to $\mathfrak{h}_{\mathbb{R}}^*$ be the bilinear map
>
> $$\kappa : \mathfrak{h}_{\mathbb{R}}^* \times \mathfrak{h}_{\mathbb{R}}^* \xrightarrow{\sim} \mathbb{R},$$
>
> then κ is an inner product on $\mathfrak{h}_{\mathbb{R}}^*$.

With such a structure, we may define lengths and angles between elements of $\mathfrak{h}_{\mathbb{R}}^*$. Namely, we define the length of $x \in \mathfrak{h}_{\mathbb{R}}^*$ as

$$\|\alpha\| = \sqrt{\kappa(x, x)}$$

and the angle between $x \neq 0$ and $y \in \mathfrak{h}_{\mathbb{R}}^* \setminus \{0\}$ as

$$\angle(x, y) = \cos^{-1}\left(\frac{\kappa(x, y)}{\|x\|\|y\|}\right).$$

The last piece to classify simple Lie algebras is the *Weyl group*, with which one may recover the set of roots of the Lie algebra from a set of fundamental roots.

> **Definition 3.13: Weyl group**
>
> For any $\lambda \in \Phi$, the map
>
> $$s_\lambda : \mathfrak{h}_{\mathbb{R}}^* \xrightarrow{\sim} \mathfrak{h}_{\mathbb{R}}^*$$
> $$x \mapsto x - 2\frac{\kappa(\lambda, x)}{\kappa(\lambda, \lambda)}\lambda$$
>
> is called a *Weyl transformation*. The set of all Weyl transformations
>
> $$W = \{s_\lambda : \lambda \in \Phi\}$$
>
> is a group under the composition of maps, called the *Weyl group*.

> **Theorem 3.6: Weyl groups properties**
>
> The Weyl group is *generated* by that fundamental roots in Π, that is,
>
> $$\forall w \in W, \exists \pi_1, \ldots, \pi_r \in \Pi : w = s_{\pi_1 \circ \cdots \circ \pi_r}.$$
>
> Every root $\lambda \in \Phi$ can be produced from a fundamental root $\pi \in \Pi$ by action of the Weyl group, that is,
> $$\forall \lambda \in \Phi, \exists w \in W, \pi \in \Pi : \lambda = w(\pi).$$
>
> The Weyl group merely permutes the roots, that is,
>
> $$\forall w \in W, \forall \lambda \in \Phi : w(\lambda) \in \Phi.$$

3.2.3 Dynkin diagrams

Consider $\pi_i, \pi_j \in \Pi$, then

$$s_{\pi_i}(\pi_j) = \pi_j - 2\frac{\kappa(\pi_i, \pi_j)}{\kappa(\pi_i, \pi_i)}\pi_i.$$

Since $s_{\pi_i}(\pi_j) \in \Phi$, there exists $n_i \in \mathbb{N}$ such that

$$s_{\pi_i}(\pi_j) = \epsilon \sum_{k=1}^{m} n_k \pi_k,$$

for $\epsilon \in \{-1, 1\}$. Then, for $i \neq j$

$$-2\frac{\kappa(\pi_i, \pi_j)}{\kappa(\pi_i, \pi_i)} \in \mathbb{N}.$$

We define the *Cartan matrix* $C_{ij} \in \mathbb{Z}$,

$$C_{ij} = 2\frac{\kappa(\pi_i, \pi_j)}{\kappa(\pi_i, \pi_i)}$$

for $i, j \in \{1, \ldots, m\}$. With the Cartan matrix, we may define a matrix whose elements are called *bond numbers*, given by

$$n_{ij} = C_{ij}C_{ji},$$

for all $i, j \in \{1, \ldots, m\}$.

Note that n_{ij} is related to the angle between two fundamental roots. Indeed, for $\pi_i, \pi_j \in \Pi$, we have

$$\begin{aligned}
n_{ij} &= \left(2\frac{\kappa(\pi_i, \pi_j)}{\kappa(\pi_i, \pi_i)}\right) \cdot \left(2\frac{\kappa(\pi_j, \pi_i)}{\kappa(\pi_j, \pi_j)}\right) \\
&= 4\frac{\kappa(\pi_i, \pi_j)^2}{\|\pi_i\|^2 \|\pi_j\|^2} \\
&= 4\cos^2\left(\angle(\pi_i, \pi_j)\right).
\end{aligned}$$

From the linear independence of Π, it follows that $0 \leq \cos^2\left(\angle(\pi_i, \pi_j)\right) < 1$ for $i \neq j$, therefore $n_{ij} \in \{0, 1, 2, 3\}$. Since C_{ij} is a negative integer for $i \neq j$, the only possible combinations are given in the following table.

C_{ij}	C_{ji}	n_{ij}
0	0	0
-1	-1	1
-1	-2	2
-2	-1	2
-3	-1	3
-1	-3	3

Let us consider the cases where $C_{ij} \neq C_{ji}$. If $C_{ij} < C_{ji}$, then

$$\frac{\kappa(\pi_i, \pi_j)}{\|\pi_i\|^2} < \frac{\kappa(\pi_j, \pi_i)}{\|\pi_j\|^2} \implies \left(\|\pi_j\|^2 - \|\pi_i\|^2 \right) \kappa(\pi_i, \pi_j) < 0,$$

which implies $\|\pi_j\| > \|\pi_i\|$, since $\kappa(\pi_i, \pi_j) < 0$. Analogously, if $C_{ij} > C_{ji}$, then $\|\pi_j\| < \|\pi_i\|$. If $C_{ij} = C_{ji}$, then either the fundamental roots have the same length or they are orthogonal with respect to the inner product κ.

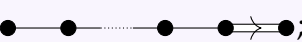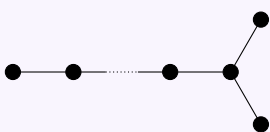**Definition 3.14: Dynkin diagrams**

A *Dynkin diagram* associated to a Cartan matrix is constructed by the following rules:

(a) For every fundamental root draw a circle.

(b) If two different circles represent $\pi_i, \pi_j \in \Pi$, draw n_{ij} lines between them.

(c) If $n_{ij} > 1$, draw an arrow on the lines from the largest root to the shorter one.

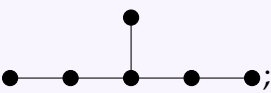**Theorem 3.7: (Killing, Cartan) Root systems of simple Lie algebras**

Any finite-dimensional simple complex Lie algebra can be reconstructed from its set Π of fundamental roots and the latter only come in the following forms:

(a) Four infinite families, named the classical Lie algebras,

- A_m, with $m \geq 1$, represented by ●——●······●——●;
- B_m, with $m \geq 2$, represented by ●——●······●——➤●;
- C_m, with $m \geq 3$, represented by ●——●······●——⬱●;
- D_m, with $m \geq 4$, represented by ●——●······●——●⟨ ;

where the restrictions on m are given to avoid repetition of diagrams, e.g. C_2 would be equivalent to B_2, and to ensure it represents a simple Lie algebra, e.g. D_2 would be semisimple.

(b) The five exceptional Lie algebras,

- E_6 represented by ●——●——●——●——●;

- E_7 represented by •——•——•——•——•——•;

- E_8 represented by •——•——•——•——•——•——•;

- F_4 represented by •——•⟩—•——•;

- G_2 represented by •⟩⟩•.

3.3 The Lie group $\mathrm{SL}(2, \mathbb{C})$ and its Lie algebra $\mathfrak{sl}(2, \mathbb{C})$

As an example, we study the *special linear* group $\mathrm{SL}(2, \mathbb{C})$ given by the set

$$\mathrm{SL}(2, \mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^4 : ad - bc = 1 \right\},$$

with group operation $\bullet$ defined by matrix multiplication, that is,

$$\bullet : \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \to \mathrm{SL}(2, \mathbb{C})$$
$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix} \right) \mapsto \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

It is straightforward to show, albeit tedious, that $\mathrm{SL}(2, \mathbb{C})$ is indeed a group, that is,

(a) $\bullet$ is associative,

(b) the element $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity element,

(c) for each $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C})$ there exists the inverse element given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}).$$

3.3.1 The topology of $\mathrm{SL}(2, \mathbb{C})$

We may identify $\mathbb{C}^4$ with $\mathbb{R}^8$ and equip $\mathbb{C}^4$ with the analogous standard topology. We may define its topology $\mathcal{O}$ as the subspace topology of the standard topology in $\mathbb{C}^4$.

First we show the properties of $(\mathrm{SL}(2, \mathbb{C}), \mathcal{O})$ as a topological space.

> **Proposition 3.4: The topological space $(\mathrm{SL}(2, \mathbb{C}), \mathcal{O})$**
>
> The topological space $(\mathrm{SL}(2, \mathbb{C}), \mathcal{O})$ is a paracompact, second countable, connected, Hausdorff space.

Proof. BIG TODO! □

Now we construct an atlas for this topological space, to show it is locally Euclidean, and thus a topological manifold.

> **Proposition 3.5:** $(\mathrm{SL}(2, \mathbb{C}), O)$ **is a topological manifold**
>
> The set $\mathrm{SL}(2, \mathbb{C})$ equipped with the topology O is a 3-dimensional complex topological manifold.

Proof. Consider the subset $U \subset \mathrm{SL}(2, \mathbb{C})$ defined by

$$U = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}) : a \neq 0 \right\}.$$

Clearly U is open in O, since it is the intersection of the open subset $\mathbb{C}^* \times \mathbb{C}^3$ of $\mathbb{C}^4$ with $\mathrm{SL}(2, \mathbb{C})$. We define the homeomorphism

$$x : U \to x(U) \subset \mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$$
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, c),$$

with inverse

$$x^{-1} : x(U) \to U$$
$$(\xi^1, \xi^2, \xi^3) \mapsto \begin{pmatrix} \xi^1 & \xi^2 \\ \xi^3 & \frac{1 + \xi^2 \xi^3}{\xi^1} \end{pmatrix}.$$

Similarly, we consider the open subset $V \subset \mathrm{SL}(2, \mathbb{C})$

$$V = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}) : b \neq 0 \right\},$$

where we define the homeomorphism

$$y : V \to y(V) \subset \mathbb{C} \times \mathbb{C}^* \times \mathbb{C}$$
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, d),$$

with inverse

$$y^{-1} : y(V) \to V$$
$$(\xi^1, \xi^2, \xi^3) \mapsto \begin{pmatrix} \xi^1 & \xi^2 \\ \frac{\xi^1 \xi^3 - 1}{\xi^2} & \xi^3 \end{pmatrix}.$$

Note that $U \cup V = \mathrm{SL}(2, \mathbb{C})$, since

$$a = 0 \wedge b = 0 \implies \begin{pmatrix} a & b \\ c & d \end{pmatrix} \notin \mathrm{SL}(2, \mathbb{C}).$$

Therefore, $\mathscr{A}_{\mathrm{top}} = \{(U, x), (V, y)\}$ is a topological atlas for $(\mathrm{SL}(2, \mathbb{C}), O)$. Thus, $\mathrm{SL}(2, \mathbb{C})$ is a topological manifold whose open subsets are homeomorphic to $\mathbb{C}^3$. $\qquad\square$

3.3.2 The differentiable structure of $\mathrm{SL}(2, \mathbb{C})$

We may now extend the topological atlas constructed to a differentiable atlas.

> **Proposition 3.6:** $\mathrm{SL}(2,\mathbb{C})$ **is a complex differentiable manifold**
>
> The set $\mathrm{SL}(2,\mathbb{C})$ is a complex differentiable manifold.

Proof. We consider the intersection $U \cap V$ and we check the differentiable compatibility of the chart transition maps. The chart transition maps are

$$y \circ x^{-1} : x(U \cap V) \subset \mathbb{C}^* \times \mathbb{C}^* \times \mathbb{C} \to y(U \cap V) \subset \mathbb{C}^* \times \mathbb{C}^* \times \mathbb{C}$$

$$(\xi^1, \xi^2, \xi^3) \mapsto \left(\xi^1, \xi^2, \frac{1 + \xi^2 \xi^3}{\xi^1} \right)$$

and

$$x \circ y^{-1} : y(U \cap V) \subset \mathbb{C}^* \times \mathbb{C}^* \times \mathbb{C} \to x(U \cap V) \subset \mathbb{C}^* \times \mathbb{C}^* \times \mathbb{C}$$

$$(\xi^1, \xi^2, \xi^3) \mapsto \left(\xi^1, \xi^2, \frac{\xi^1 \xi^3 - 1}{\xi^2} \right),$$

which are holomorphic in the open domains $x(U \cap V)$ and $y(U \cap V)$, respectively. We may then define $\mathscr{A}$ as the maximal complex differentiable atlas containing $\mathscr{A}_{\mathrm{top}}$, therefore $(\mathrm{SL}(2,\mathbb{C}), O, \mathscr{A})$ is a complex differentiable manifold. $\qquad\square$

From now own, we will denote this differentiable manifold simply by $\mathrm{SL}(2,\mathbb{C})$. We may now check that $\mathrm{SL}(2,\mathbb{C})$ is a Lie group.

> **Proposition 3.7:** $\mathrm{SL}(2,\mathbb{C})$ **is a Lie group**
>
> The maps
>
> $$i : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(2,\mathbb{C})$$
> $$g \mapsto g^{-1}$$
>
> and
>
> $$\mu : \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C}) \to \mathrm{SL}(2,\mathbb{C})$$
> $$(g, h) \mapsto gh$$
>
> are complex differentiable, that is, $\mathrm{SL}(2,\mathbb{C})$ is a Lie group.

Proof. From the definition of the inverse element, it is clear that $i(U) = W$ and $i(W) = U$, where U is defined as before and

$$W = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C}) : d \neq 0 \right\}$$

is an open subset of $\mathrm{SL}(2,\mathbb{C})$. Similarly, $i(V) = V$ and analogously for the other entry. Then, we must check differentiability of i with respect to elements in U and V, since $U \cup V$ covers $\mathrm{SL}(2,\mathbb{C})$.

$$
\begin{array}{ccc}
U \subset \mathrm{SL}(2,\mathbb{C}) & \xrightarrow{\ i\ } & W \subset \mathrm{SL}(2,\mathbb{C}) \\
\downarrow{\scriptstyle x} & & \downarrow{\scriptstyle z} \\
x(U) & \xrightarrow{\ z \circ i \circ x^{-1}\ } & z(W)
\end{array}
$$

Consider the chart (U, x), defined as before, and $(W, z) \in \mathscr{A}$, where the homeomorphism z is defined by

$$z : W \to z(W) \subset \mathbb{C} \times \mathbb{C} \times \mathbb{C}^*$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (b, c, d),$$

with inverse

$$z^{-1} : z(W) \to W$$

$$(\xi^1, \xi^2, \xi^3) \mapsto \begin{pmatrix} \frac{1+\xi^1 \xi^2}{\xi^3} & \xi^1 \\ \xi^2 & \xi^3 \end{pmatrix}.$$

In these charts, the local expression of i from U to W is given by

$$z \circ i \circ x^{-1} : x(U) \to z(W)$$

$$(\xi^1, \xi^2, \xi^3) \mapsto \left(-\xi^2, -\xi^3, \xi^1 \right),$$

which is obviously holomorphic in $x(U)$. Similarly, the local expression of i from V to V is given by

$$y \circ i \circ y^{-1} : y(V) \to y(V)$$

$$(\xi^1, \xi^2, \xi^3) \mapsto \left(\xi^3, -\xi^2, \xi^1 \right),$$

which is holomorphic in $y(V)$. This shows the map $i : \text{SL}(2, \mathbb{C}) \to \text{SL}(2, \mathbb{C})$ is a diffeomorphism.

We may equip $\text{SL}(2, \mathbb{C}) \times \text{SL}(2, \mathbb{C})$ with a differentiable atlas by virtue of the differentiable atlas $\mathscr{A}$ on $\text{SL}(2, \mathbb{C})$. With this, we may show that the map μ is differentiable. TODO!   □

3.3.3   The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ of the Lie group $\text{SL}(2, \mathbb{C})$

The set of left invariant vector fields on $\text{SL}(2, \mathbb{C})$ is the Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. Recall that $\mathfrak{sl}(2, \mathbb{C})$ is isomorphic to the Lie algebra of the tangent space $T_I \text{SL}(2, \mathbb{C})$ equipped with the Lie bracket

$$[\cdot, \cdot] : T_I \text{SL}(2, \mathbb{C}) \times T_I \text{SL}(2, \mathbb{C}) \to T_I \text{SL}(2, \mathbb{C})$$

$$(A, B) \mapsto j^{-1} \left([j(A), j(B)] \right),$$

where $j : T_I \text{SL}(2, \mathbb{C}) \to \mathfrak{sl}(2, \mathbb{C})$ was the vector space isomorphism defined by the pushforward of the left translations, namely

$$j : T_I \text{SL}(2, \mathbb{C}) \to \mathfrak{sl}(2, \mathbb{C})$$

$$A \mapsto j(A),$$

where $j(A)_g = \ell_{g*I} A$ for all $g \in \text{SL}(2, \mathbb{C})$ and $j^{-1}(X) = X_I$.

We would like to determine explicitly the Lie bracket on the tangent space at the identity element. For this task, we employ the (U, x) chart, since $I \in U$, therefore we may express any element A in $T_I \text{SL}(2, \mathbb{C})$ as

$$A = A^i \left. \frac{\partial}{\partial x^i} \right|_I ,$$

for coefficients $A^i \in \mathbb{C}$. Since j is linear, we may consider each basis vector individually. For any $f \in C^\infty(\mathrm{SL}(2, \mathbb{C}))$ and $g \in \mathrm{SL}(2, \mathbb{C})$, we have

$$\left[\ell_{g*} \left. \frac{\partial}{\partial x^i} \right|_I \right]_g f = \left. \frac{\partial}{\partial x^i} f \circ \ell_g \right|_I = \partial_i \left(f \circ \ell_g \circ x^{-1} \right) (x(I)),$$

where $f \circ \ell_g \circ x^{-1} : x(U) \to \mathbb{C}$ is a smooth map. Let $(\tilde{U}, \tilde{x}) \in \mathscr{A}$ be a chart, where $\tilde{U}$ is a neighborhood of g. By the chain rule,

$$\left[\ell_{g*} \left. \frac{\partial}{\partial x^i} \right|_I \right]_g f = \partial_i \left(f \circ \tilde{x}^{-1} \circ \tilde{x} \circ \ell_g \circ x^{-1} \right) (x(I))$$

$$= \partial_m \left(f \circ \tilde{x}^{-1} \right) (\tilde{x}(g)) \cdot \partial_i \left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (x(I))$$

$$= \partial_i \left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (x(I)) \cdot \left. \frac{\partial}{\partial \tilde{x}^m} f \right|_g ,$$

that is,

$$j \left(\left. \frac{\partial}{\partial x^i} \right|_I \right)_g = \partial_i \left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (x(I)) \cdot \left. \frac{\partial}{\partial \tilde{x}^m} \right|_g .$$

Let $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in \mathrm{SL}(2, \mathbb{C})$ and $(\xi^1, \xi^2, \xi^3) \in x(U) \subset \mathbb{C}^* \times \mathbb{C} \times \mathbb{C}$. Then,

$$\left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (\xi^1, \xi^2, \xi^3) = \tilde{x}^m \begin{pmatrix} a\xi^1 + b\xi^3 & a\xi^2 + b\frac{1+\xi^2\xi^3}{\xi^1} \\ c\xi^1 + d\xi^3 & c\xi^2 + d\frac{1+\xi^2\xi^3}{\xi^1} \end{pmatrix}$$

We first consider the case $\tilde{U} = U$. Then $\tilde{x} = x$, and

$$\left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (\xi^1, \xi^2, \xi^3) = \left(a\xi^1 + b\xi^3, a\xi^2 + b\frac{1+\xi^2\xi^3}{\xi^1}, c\xi^1 + d\xi^3 \right)^m .$$

We may now compute the partial derivatives and evaluate them at $x(I) = (1, 0, 0)$,

$$D_U(g)^m{}_i = \partial_i \left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (x(I)) = \begin{bmatrix} a & 0 & b \\ -b & a & 0 \\ c & 0 & d \end{bmatrix}^m{}_i .$$

This allows to compute $j \left(\left. \frac{\partial}{\partial x^i} \right|_I \right)_g = D_U(g)^m{}_i \left. \frac{\partial}{\partial x^m} \right|_g$ for all $g \in U$, namely

$$j \left(\left. \frac{\partial}{\partial x^1} \right|_I \right)_g = a \left. \frac{\partial}{\partial x^1} \right|_g \quad -b \left. \frac{\partial}{\partial x^2} \right|_g + c \left. \frac{\partial}{\partial x^3} \right|_g ,$$

$$j \left(\left. \frac{\partial}{\partial x^2} \right|_I \right)_g = \qquad a \left. \frac{\partial}{\partial x^2} \right|_g ,$$

$$j \left(\left. \frac{\partial}{\partial x^3} \right|_I \right)_g = b \left. \frac{\partial}{\partial x^1} \right|_g \qquad\qquad + d \left. \frac{\partial}{\partial x^3} \right|_g .$$

Since $U \cup V$ covers $\mathrm{SL}(2, \mathbb{C})$, the other case is when $\tilde{U} = V$. In this case, $\tilde{x} = y$ and

$$\left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right) (\xi^1, \xi^2, \xi^3) = \left(a\xi^1 + b\xi^3, a\xi^2 + b\frac{1+\xi^2\xi^3}{\xi^1}, c\xi^2 + d\frac{1+\xi^2\xi^3}{\xi^1} \right)^m .$$

Computing the partial derivatives at $x(I)$ yields

$$D_V(g)^m{}_i = \partial_i \left(\tilde{x}^m \circ \ell_g \circ x^{-1} \right)(x(I)) = \begin{bmatrix} a & 0 & b \\ -b & a & 0 \\ -d & c & 0 \end{bmatrix}^m_i.$$

This allows to compute $j \left(\frac{\partial}{\partial x^i}\big|_I \right)_g = D_V(g)^m{}_i \frac{\partial}{\partial y^m}\big|_g$ for all $g \in V$, namely

$$j \left(\frac{\partial}{\partial x^1}\Big|_I \right)_g = a \frac{\partial}{\partial y^1}\Big|_g \quad -b \frac{\partial}{\partial y^2}\Big|_g - d \frac{\partial}{\partial y^3}\Big|_g,$$

$$j \left(\frac{\partial}{\partial x^2}\Big|_I \right)_g = \quad a \frac{\partial}{\partial y^2}\Big|_g + c \frac{\partial}{\partial y^3}\Big|_g,$$

$$j \left(\frac{\partial}{\partial x^3}\Big|_I \right)_g = b \frac{\partial}{\partial y^1}\Big|_g.$$

These two expressions for $j \left(\frac{\partial}{\partial x^i}\big|_I \right)_g$ must coincide for $g \in U \cap V$. Recall the change of basis

$$\frac{\partial}{\partial x^i}\Big|_g = \frac{\partial}{\partial x^i} y^j\Big|_g \frac{\partial}{\partial y^j}\Big|_g = \partial_i \left(y^j \circ x^{-1} \right)(x(g)) \frac{\partial}{\partial y^j}\Big|_g.$$

We compute the partial derivatives: let $(\xi^1, \xi^2, \xi^3) \in x(U \cap V)$, then

$$\partial_i \left(y^j \circ x^{-1} \right)(\xi^1, \xi^2, \xi^3) = \partial_i \left(\xi^1, \xi^2, \frac{1 + \xi^2 \xi^3}{\xi^1} \right)^j (\xi^1, \xi^2, \xi^3)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{1+\xi^2\xi^3}{(\xi^1)^2} & \frac{\xi^3}{\xi^1} & \frac{\xi^2}{\xi^1} \end{bmatrix}^j_i.$$

The coefficients of the change of basis linear map at g is

$$\frac{\partial}{\partial x^i} y^j\Big|_g = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{d}{a} & \frac{c}{a} & \frac{b}{a} \end{bmatrix}^j_i,$$

since $ad - bc = 1$ and $a \neq 0$ for $g \in U \cap V$. It's straightforward to verify that

$$D_U(g)^k{}_i \frac{\partial}{\partial x^k} y^j\Big|_g = D_V(g)^j{}_i,$$

then

$$j \left(\frac{\partial}{\partial x^i}\Big|_I \right)_g = D_U(g)^k{}_i \frac{\partial}{\partial x^k}\Big|_g$$

$$= D_U(g)^k{}_i \frac{\partial}{\partial x^k} y^j\Big|_g \frac{\partial}{\partial y^j}\Big|_g$$

$$= D_V(g)^j{}_i \frac{\partial}{\partial y^j}\Big|_g,$$

as desired. We denote the vector field by

$$j\left(\left.\frac{\partial}{\partial x^i}\right|_I\right) = D^j{}_i \frac{\partial}{\partial \tilde{x}^j},$$

meaning to take the appropriate expression in each chart.

Let $f \in C^\infty(\mathrm{SL}(2, \mathbb{C}))$. Since vector fields are derivations, we have from the Leibniz rule

$$\frac{\partial}{\partial \tilde{x}^j}\left(D^k{}_m \frac{\partial}{\partial \tilde{x}^k} f\right) = \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right)\left(\frac{\partial}{\partial \tilde{x}^k} f\right) + D^k{}_m \frac{\partial}{\partial \tilde{x}^j}\left(\frac{\partial}{\partial \tilde{x}^k} f\right)$$

$$= \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right)\left(\frac{\partial}{\partial \tilde{x}^k} f\right) + D^k{}_m \frac{\partial}{\partial \tilde{x}^j}\left(\partial_k(f \circ \tilde{x}^{-1}) \circ \tilde{x}\right)$$

$$= \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right)\left(\frac{\partial}{\partial \tilde{x}^k} f\right) + D^k{}_m \partial_j\left(\partial_k(f \circ \tilde{x}^{-1}) \circ \tilde{x} \circ \tilde{x}^{-1}\right) \circ \tilde{x}$$

$$= \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right)\left(\frac{\partial}{\partial \tilde{x}^k} f\right) + D^k{}_m \partial_j \partial_k\left(f \circ \tilde{x}^{-1}\right) \circ \tilde{x}.$$

Similarly, we have

$$\frac{\partial}{\partial \tilde{x}^k}\left(D^j{}_i \frac{\partial}{\partial \tilde{x}^j} f\right) = \left(\frac{\partial}{\partial \tilde{x}^k} D^j{}_i\right)\left(\frac{\partial}{\partial \tilde{x}^j} f\right) + D^j{}_i \partial_k \partial_j\left(f \circ \tilde{x}^{-1}\right) \circ \tilde{x}$$

$$= \left(\frac{\partial}{\partial \tilde{x}^k} D^j{}_i\right)\left(\frac{\partial}{\partial \tilde{x}^j} f\right) + D^j{}_i \partial_j \partial_k\left(f \circ \tilde{x}^{-1}\right) \circ \tilde{x},$$

from Schwartz's theorem. It follows that

$$\left[j\left(\left.\frac{\partial}{\partial x^i}\right|_I\right), j\left(\left.\frac{\partial}{\partial x^m}\right|_I\right)\right] f = D^j{}_i \frac{\partial}{\partial \tilde{x}^j}\left(D^k{}_m \frac{\partial}{\partial \tilde{x}^k} f\right) - D^k{}_m \frac{\partial}{\partial \tilde{x}^k}\left(D^j{}_i \frac{\partial}{\partial \tilde{x}^j} f\right)$$

$$= D^j{}_i \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right)\left(\frac{\partial}{\partial \tilde{x}^k} f\right) + D^j{}_i D^k{}_m \partial_j \partial_k\left(f \circ \tilde{x}^{-1}\right) \circ \tilde{x}$$

$$\quad - D^k{}_m \left(\frac{\partial}{\partial \tilde{x}^k} D^j{}_i\right)\left(\frac{\partial}{\partial \tilde{x}^j} f\right) - D^k{}_m D^j{}_i \partial_j \partial_k\left(f \circ \tilde{x}^{-1}\right) \circ \tilde{x}$$

$$= D^j{}_i \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right)\left(\frac{\partial}{\partial \tilde{x}^k} f\right) - D^k{}_m \left(\frac{\partial}{\partial \tilde{x}^k} D^j{}_i\right)\left(\frac{\partial}{\partial \tilde{x}^j} f\right)$$

$$= \left(D^j{}_i \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right) - D^j{}_m \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_i\right)\right) \frac{\partial}{\partial \tilde{x}^k} f,$$

where we have relabeled dummy indices in the last step. Since f is arbitrary,

$$\left[j\left(\left.\frac{\partial}{\partial x^i}\right|_I\right), j\left(\left.\frac{\partial}{\partial x^m}\right|_I\right)\right] = \left(D^j{}_i \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_m\right) - D^j{}_m \left(\frac{\partial}{\partial \tilde{x}^j} D^k{}_i\right)\right) \frac{\partial}{\partial \tilde{x}^k}$$

is the commutator of left invariant vector fields. Evaluating at the identity element yields the structure coefficients of the Lie algebra $T_I \mathrm{SL}(2, \mathbb{C}) \cong \mathfrak{sl}(2, \mathbb{C})$. At the identity, we have

$$\left[\left.\frac{\partial}{\partial x^i}\right|_I, \left.\frac{\partial}{\partial x^m}\right|_I\right] = \left(D^j{}_i(I)\left(\left.\frac{\partial}{\partial x^j}\right|_I D^k{}_m\right) - D^j{}_m(I)\left(\left.\frac{\partial}{\partial x^j}\right|_I D^k{}_i\right)\right)\left.\frac{\partial}{\partial x^k}\right|_I$$

$$= \left(\delta^j{}_i\left(\left.\frac{\partial}{\partial x^j}\right|_I D^k{}_m\right) - \delta^j{}_m\left(\left.\frac{\partial}{\partial x^j}\right|_I D^k{}_i\right)\right)\left.\frac{\partial}{\partial x^k}\right|_I$$

$$= \left(\left.\frac{\partial}{\partial x^i}\right|_I D^k{}_m - \left.\frac{\partial}{\partial x^m}\right|_I D^k{}_i\right)\left.\frac{\partial}{\partial x^k}\right|_I,$$

where we have used that $D^i_{\ j}(I) = \delta^i_{\ j}$. Recall that

$$D^i_{\ j} \circ x^{-1}(\xi^1, \xi^2, \xi^3) = \begin{bmatrix} \xi^1 & 0 & \xi^2 \\ -\xi^2 & \xi^1 & 0 \\ \xi^3 & 0 & \frac{1+\xi^2\xi^3}{\xi^1} \end{bmatrix}^i_{\ j},$$

then

$$\left. \frac{\partial}{\partial x^1} \right|_I D^i_{\ j} = \partial_1 \left(D^i_{\ j} \circ x^{-1} \right)(x(I)) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}^i_{\ j},$$

$$\left. \frac{\partial}{\partial x^2} \right|_I D^i_{\ j} = \partial_2 \left(D^i_{\ j} \circ x^{-1} \right)(x(I)) = \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}^i_{\ j},$$

and

$$\left. \frac{\partial}{\partial x^3} \right|_I D^i_{\ j} = \partial_3 \left(D^i_{\ j} \circ x^{-1} \right)(x(I)) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}^i_{\ j}.$$

Finally, we have

$$\left[\left. \frac{\partial}{\partial x^1} \right|_I, \left. \frac{\partial}{\partial x^2} \right|_I \right] = 2 \left. \frac{\partial}{\partial x^2} \right|_I,$$

$$\left[\left. \frac{\partial}{\partial x^1} \right|_I, \left. \frac{\partial}{\partial x^3} \right|_I \right] = -2 \left. \frac{\partial}{\partial x^3} \right|_I,$$

$$\left[\left. \frac{\partial}{\partial x^2} \right|_I, \left. \frac{\partial}{\partial x^3} \right|_I \right] = \left. \frac{\partial}{\partial x^1} \right|_I.$$

3.3.4 The simple Lie algebra $\mathfrak{sl}(2, \mathbb{C})$

Let $X_i = j \left(\left. \frac{\partial}{\partial x^i} \right|_I \right)$, then $\{X_1, X_2, X_3\}$ is a basis for $\mathfrak{sl}(2, \mathbb{C})$ and the structure coefficients are

$$C^2_{\ 12} = 2, \quad C^3_{\ 13} = -2, \quad C^1_{\ 23} = 1,$$

and the other coefficients are either zero or obtained by antisymmetry in the lower indices.

> **Proposition 3.8: $\mathfrak{sl}(2, \mathbb{C})$ is semisimple**
>
> The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is semisimple.

Proof. We may compute the components of the Killing form

$$K_{ij} = C^m_{\ in} C^n_{\ jm}$$

from the structure coefficients. For $i = j = 1$, we have

$$
\begin{aligned}
K_{11} &= C^m_{\ 1n} C^n_{\ 1m} \\
&= C^1_{\ 1n} C^n_{\ 11} + C^2_{\ 1n} C^n_{\ 12} + C^3_{\ 1n} C^n_{\ 13} \\
&= C^2_{\ 12} C^2_{\ 12} + C^3_{\ 13} C^3_{\ 13} \\
&= 8.
\end{aligned}
$$

We repeat the same computation for $i \geq j$ and use the symmetric property of the Killing form, yielding

$$K_{ij} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 0 & 4 \\ 0 & 4 & 0 \end{bmatrix}_{ij}.$$

This shows K is non-degenerate, thus $\mathfrak{sl}(2, \mathbb{C})$ is semisimple. □

We have shown that

$$K(X_1, X_1) = 8, \quad K(X_2, X_2) = 0, \quad K(X_3, X_3) = 0.$$

From this we see K is an indefinite form, so $\mathrm{SL}(2, \mathbb{C})$ is not a compact Lie group.

> **Proposition 3.9:** $\mathfrak{sl}(2, \mathbb{C})$ **is simple**
>
> The Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ is simple.

Proof. Let $\mathfrak{I}$ be an ideal of $\mathfrak{sl}(2, \mathbb{C})$. Let $\alpha, \beta, \gamma \in \mathbb{C}$ such that

$$Y = \alpha X_1 + \beta X_2 + \gamma X_3 \in \mathfrak{I}.$$

We have

$$\begin{aligned}
[Y, X_1] &= & - 2\beta X_2 & + 2\gamma X_3 \\
[Y, X_2] &= -\gamma X_1 & + 2\alpha X_2 & \\
[Y, X_3] &= \beta X_1 & & - 2\alpha X_3
\end{aligned}$$

and all of them lie in $\mathfrak{I}$. TODO! □

3.3.5   Roots and fundamental roots of $\mathfrak{sl}(2, \mathbb{C})$

Notice the structure coefficient equations may be expressed with the adjoint map as

$$\mathrm{ad}_{X_1} X_2 = 2X_2, \quad \mathrm{ad}_{X_1} X_3 = -2X_3, \quad \mathrm{ad}_{X_2} X_3 = X_1.$$

In particular, we notice X_2 and X_3 are eigenvectors of the linear map ad_{X_1}, and no other basis vector is an eigenvector of the other adjoint maps, so $\mathfrak{h} = \mathrm{span}_{\mathbb{C}}\{X_1\}$ is a Cartan subalgebra of $\mathfrak{sl}(2, \mathbb{C})$. Therefore $\{X_1, X_2, X_3\}$ is a Cartan-Weyl basis and we have the roots $\lambda_2, \lambda_3 \in \mathfrak{h}^*$ defined by $\lambda_2(X) = 2\epsilon^1(X)$ and $\lambda_3(X) = -2\epsilon^1(X)$ for all $X \in \mathfrak{sl}(2, \mathbb{C})$, and $\{\epsilon^1, \epsilon^2, \epsilon^3\}$ is the dual basis.

Since the root space is just $\Phi = \{\lambda_2, \lambda_3\}$ and $\lambda_3 = -\lambda_2$, it's easy to see one choice of set of fundamental roots is $\Pi = \{\lambda_2\}$. Therefore, the Dynkin diagram of $\mathfrak{sl}(2, \mathbb{C})$ is A_1, ●.

3.4   Reconstruction of of the Lie algebra from the Dynkin diagram A_2

In the previous example we have constructed the Lie algebra associated to the Lie group $\mathrm{SL}(2, \mathbb{C})$ and we have arrived at the simplest possible Dynkin diagram, A_1. In this section, we consider the Dynkin diagram A_2

●——●

and aim to reconstruct the Lie algebra $\mathfrak{g}$ from this diagram.

From the diagram, we have the set of fundamental roots $\Pi = \{\pi^1, \pi^2\} \in \mathfrak{h}^* \subset \mathfrak{g}^*$ and bond number $n_{12} = 1$. Therefore, the Cartan matrix is given by

$$C_{ij} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}_{ij},$$

since $-C_{ij} \in \mathbb{N}$ for $i \neq j$ and $C_{12}C_{21} = n_{12}$. We recall the bond number is related to the angle between the fundamental roots,

$$n_{ij} = \left(2 \cos \angle(\pi_i, \pi_j)\right)^2.$$

Since $-C_{ij} \in \mathbb{N}$, we have $\angle(\pi_i, \pi_j) \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$, and we obtain $\angle(\pi^1, \pi^2) = \frac{2\pi}{3}$. We also know $\|\pi^1\| = \|\pi^2\|$ because the Cartan matrix is symmetric in this case.

Since $\mathfrak{h}^*_\mathbb{R}$ has an inner product, we may define an orthonormal basis $\{\epsilon^1, \epsilon^2\}$ obtained by the Gram-Schmidt process on Π. Since $\text{span}_\mathbb{C} \Pi = \mathfrak{h}^*$, it is clear that $\text{span}_\mathbb{C}\{\epsilon^1, \epsilon^2\} = \mathfrak{h}^*$. In this frame, $\pi^1 = \epsilon_1$ and $\pi^2 = -\frac{1}{2}\epsilon^1 + \frac{\sqrt{3}}{2}\epsilon^2$.
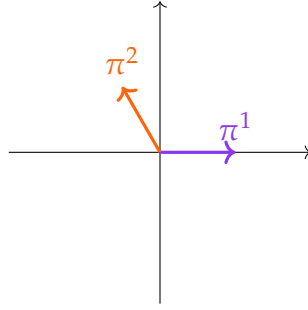

Figure 3.1: Fundamental roots in $\mathfrak{h}^*_\mathbb{R}$ found from Dynkin diagram A_2

In order to find the other roots, we repeatedly apply the Weyl transformations,

$$s_\lambda(\pi_j) = \pi_j - 2\frac{\kappa(\lambda, \pi_j)}{\kappa(\lambda, \lambda)}\lambda,$$

for all $\lambda \in \Phi$, the root space. Since $\Pi \subset \Phi$, we begin with $s_{\pi_i}(\pi_j)$,

$$s_{\pi_i}(\pi_j) = \pi_j - 2\frac{\kappa(\pi_i, \pi_j)}{\kappa(\pi_i, \pi_i)}\pi_i$$
$$= \pi_j - C_{ij}\pi_i,$$

that is, $\{\pi^1, -\pi^1, \pi^2, -\pi^2, \pi^1 + \pi^2\} \subset \Phi$. Since $\lambda \in \Phi \implies -\lambda \in \Phi$, we have $-\pi^1 - \pi^2 \in \Phi$. It is easy to verify the roots found are all the roots, that is, the root space is

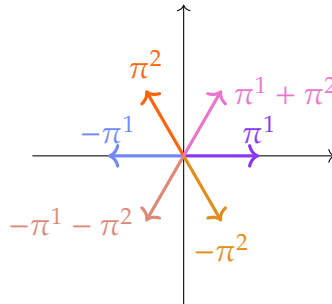$$\Phi = \{-\pi^1, \pi^1, -\pi^2, \pi^2, -\pi^1 - \pi^2, \pi^1 + \pi^2\}.$$

Figure 3.2: Roots recovered from Dynkin diagram A_2

Since there are six roots and the Cartan subalgebra has a basis with two elements, we may conclude the Lie algebra $\mathfrak{g}$ with Dynkin diagram A_2 is eight-dimensional. To further investigate this Lie algebra, we must compute the structure coefficients $C^k{}_{ij}$, with $i, j, k \in \{1, \dots, 8\}$.

Let $\{e_1, \dots, e_8\}$ be a Cartan-Weyl basis for $\mathfrak{g}$, where $\{e_1, e_2\} \subset \mathfrak{h}$ is the dual basis to $\{\epsilon^1, \epsilon^2\}$. We have $[e_1, e_2] = 0$, since $\mathfrak{g}$ is simple and

$$\mathrm{ad}_{e_i} e_j = \lambda_j(e_i) e_j,$$

for $i \in \{1, 2\}$, $j \in \{3, 4, 5, 6, 7, 8\}$ and λ_j is the $(j-2)$-th element of Φ. From this, we have twelve unique non-zero structure coefficients, given by

$$C^j{}_{ij} = \lambda_j(e_i),$$

where no summation is implied and i, j range over the aforementioned values, which in total yield 24 non-zero coefficients, by antisymmetry. In this construction, $\lambda_j(e_i)$ is the coefficient that appears in front of ϵ^i of λ_j, for example, $\pi_2(e_i) = -\frac{1}{2}\delta_i^1 + \frac{\sqrt{3}}{2}\delta_i^2$.

It remains to compute $C^k{}_{ij}$ for $i, j \in \{3, 4, 5, 6, 7, 8\}$ and $k \in \{1, \dots, 8\}$. For this, we use the Jacobi identity with the Lie brackets we have already computed. That is, for $\alpha \in \{1, 2\}$,

$$\begin{aligned}
\mathrm{ad}_{e_\alpha}[e_i, e_j] &= -\mathrm{ad}_{e_i}[e_j, e_\alpha] - \mathrm{ad}_{e_j}[e_\alpha, e_i] \\
&= \lambda_j(e_\alpha)\,\mathrm{ad}_{e_i} e_j - \lambda_i(e_\alpha)\,\mathrm{ad}_{e_j} e_i \\
&= \big((\lambda_i + \lambda_j)(e_\alpha)\big)\,[e_i, e_j],
\end{aligned}$$

where no summation is implied.

If $\lambda_i + \lambda_j \in \Phi$, then there exists $k \in \{3, 4, 5, 6, 7, 8\}$ such that $\lambda_k = \lambda_i + \lambda_j$ and $[e_i, e_j] = e_k$. If $\lambda_i + \lambda_j = 0$, all we may conclude is that $[e_i, e_j] \in \mathfrak{h}$. The only other case is when $\lambda_i + \lambda_j \notin \Phi \cup \{0\}$, which allows us to conclude $[e_i, e_j] = 0$.

Prove this TODO!

Also, not clear what to do when $[e_i, e_j] \in \mathfrak{h}$. TODO!

3.5 Representation theory of Lie groups and Lie algebras

Bibliography

[Rud64]   Walter Rudin. *Principles of Mathematical Analysis*. 1964.

[Mun91]   J.R. Munkres. *Analysis on Manifolds*. 1991.

[SKM97]   P.K. Suetin, A.I. Kostrikin, and Y.I. Manin. *Linear Algebra and Geometry*. 1997.

[Mun00]   J.R. Munkres. *Topology*. 2000.

[Tu11]   Loring W. Tu. *An Introduction to Manifolds*. 2011.

[Hir12]   M.W. Hirsch. *Differential Topology*. 2012.

[Car16]   M.P. do Carmo. *Differential Geometry of Curves and Surfaces: Revised and Updated Second Edition*. 2016.