Mathematical Notes

Contents

Part I

Basic Mathematics

Chapter 1

Lie Algebras

We already defined in the previous chapter that an algebra is a vector space A with an additional bilinear map (called binary operation or product) $\bullet\colon A \times A \to A$. A very important class of algebras, that we will also see later, are the so-called Lie algebras, in which the product $v \bullet w$ is called "Lie bracket" and denoted as $[v, w]$. In general Lie algebras are just a very specific class of algebras, hence we might have them introduced in the previous chapter under "algebras". However, since they are so important, and lengthy, we will introduce them separately in their own chapter.

Lie algebras are closely related to Lie groups, which are groups that are also smooth manifolds: any Lie group gives rise to a Lie algebra, which is its tangent space at the identity. Conversely, to any finite-dimensional Lie algebra over real or complex numbers, there is a corresponding connected Lie group unique up to finite coverings. This correspondence allows one to study the structure and classification of Lie groups in terms of Lie algebras (we will see all of that as we proceed in the notes).

In physics, Lie groups appear as symmetry groups of physical systems, and their Lie algebras (tangent vectors near the identity) may be thought of as infinitesimal symmetry motions. Thus Lie algebras and their representations are used extensively in physics, notably in quantum mechanics and particle physics.

1.1 Basic Definitions

Definition 1.1 (Lie Algebra)**.** *A* ***Lie algebra*** *A over a field K is an algebra whose product $[-,-]$, called* Lie bracket*, satisfies*

i) bilinearity: $A \times A \to A$: $[av + w, z] = a[v, w] + [v, z]$

ii) antisymmetry: $\forall\, v \in A : [v, v] = 0$;

iii) the Jacobi identity: $\forall\, v, w, z \in A : [v, [w, z]] + [w, [z, v]] + [z, [v, w]] = 0$.

Note that the zeros above represent the additive identity element in A, not the zero scalar

Some remarks are in order

Remark 1.1. The antisymmetry condition immediately implies $[v, w] = -[w, v]$ for all $v, w \in A$ since

$$[v + w, v + w] = [v, v] + [v, w] + [w, v] + [w, w] = [v, w] + [w, v] = 0 \implies [v, w] = -[w, v]$$

Remark 1.2. Notice that the Lie bracket is not defined as the usual commutator $[v, w] = vw - wv$, but is defined very abstractly by the 3 conditions. In other words, anything that satisfies these 3 conditions can be defined as a Lie bracket. Of course one example is the commutator (you can check it yourself)

Remark 1.3. Notice that we specifically defined the Lie algebra on top of a field K. One can construct an algebra over a ring, by imposing all the axioms on a module instead of a vector space. However, in this notes we will stick with Lie algebras on top of a vector space, and more specifically on top of a complex vector space (i.e where the K field is the complex and to the real numbers), since they are more related to our purposes. In general, same definitions apply for an algebra over a ring with the appropriate changes when needed.

Now let's give some examples of Lie algebras.

Example 1.1. The usual cross product between vectors $u \times w$ in $\mathbb{R}^3$ can be shown that satisfies all the requirements for a Lie bracket, hence the vector space $\mathbb{R}^3$ equipped with the cross product $[u, w] = u \times w$ is actually a Lie algebra.

Example 1.2. Let V be a vector space. Recall that we defined the set $\mathrm{End}(V)$ as the set of all endomorphisms of V, i.e the set of all linear maps that send V back to itself. Now we define the following Lie bracket:

$$\begin{aligned} [-,-]\colon \mathrm{End}(V) \times \mathrm{End}(V) &\to \mathrm{End}(V) \\ (\phi, \psi) &\mapsto [\phi, \psi] := \phi \circ \psi - \psi \circ \phi. \end{aligned}$$

It is instructive to check that this is actually a Lie bracket. Hecne, $(\mathrm{End}(V), +, \cdot, [-,-])$ is a Lie algebra. In this case, the Lie bracket is typically called the *commutator*. (Remember that after having chosen a basis then we can "represent" the elements of $\mathrm{End}(V)$ as $n \times n$ matrices over a field K, with their commutator $[v, w] = vw - wv$ where here the composition is the usual matrix multiplication).

As usual we can define the concept of homomorphism and isomorphism in the level of Lie algebras.

Definition 1.2 (Lie Algebra Homomorphism)**.** *A map ϕ between two Lie algebras that preserves both the vector space structure and the bracket structure is called a* ***Lie algebra homomorphism****.*

Definition 1.3 (Homomorphic Lie Algebras)**.** *Two Lie algebras over the same field K are said to be* ***homomorphic*** *if there exists a lie algebra homomorphism between them.*

Definition 1.4 (Lie Algebra Isomorphism)**.** *A bijective Lie algebra homomorphism is called a* ***Lie algebra isomorphism****.*

Definition 1.5 (Isomorphic Lie Algebras)**.** *Two Lie algebras over the same field K are said to be* ***isomorphic*** *if there exists a lie algebra isomorphism.*

In what follows we will make heavy use of the following notation that we will give in a form of definition:

Definition 1.6 (Bracket)**.** *Given two subsets A, B of a Lie algebra L we define the* ***bracket*** *of these two subsets $[A, B]$ as the subset defined by the span of all commutators $[x, y]$ where $x \in A$ and $y \in B$, i.e*

$$[A, B] := \mathrm{span}_K\big(\{[x, y] \in L \mid x \in A \textit{ and } y \in B\}\big)$$

In other words is just the set of all commutators $[x, y]$ where $x \in A$ and $y \in B$.

Now let's give some very basic definitions of Lie algebras.

Definition 1.7 (Abelian Lie Algebra)**.** *A Lie algebra L is said to be* ***abelian*** *if $\forall\, x, y \in L : \ [x, y] = 0$ or equivalently in bracket notation $[L, L] = 0$, where 0 denotes the trivial Lie algebra $\{0\}$.*

Abelian Lie algebras are highly non-interesting as Lie algebras: since the bracket is identically zero, it may as well not be there. On top of that, the vanishing of the bracket implies that, given any two abelian Lie algebras, every linear isomorphism between their underlying vector spaces is automatically a Lie algebra isomorphism. Therefore, for each $n \in \mathbb{N}$, there is (up to isomorphism) only one abelian n-dimensional Lie algebra.

Definition 1.8 (Subalgebra)**.** *We say L' is a* ***subalgebra*** *of L if L' is a vector subspace of L and $\forall x, y \in L' : [x, y] \in L'$ or equivalently in bracket notation $[L', L'] \in L'$.*

One can prove that if A, B are Lie subalgebras of a Lie algebra L over K, then the bracket $[A, B]$ is again a Lie subalgebra of L.

Definition 1.9 (Ideal)**.** *An* ***ideal*** *I of a Lie algebra L is a Lie subalgebra such that $\forall\, x \in I : \forall\, y \in L : [x, y] \in I$ or equivalently in bracket notation $[I, L] \subseteq I$.*

Note that no matter the Lie algebra, we can show that: $[0, L] = 0 \subseteq 0$ and $[L, L] \subseteq L$ hence both 0 and L are always ideals of any Lie algebra.

Remark 1.4. Recall from the definition of an algebra (any algebra) that the operation (or product) of the algebra $\bullet\colon A \times A \to A$ is a bilinear map with no need to be surjective. This means that applying the operation to every possible element of the algebra does not guarantee that will give us back the whole algebra (but it does guarantee to give us back a subalgebra). In other words, $[L, L] \subseteq L$ and not $[L, L] = L$.

Definition 1.10 (Trivial Ideals). *The ideals* 0 *and* L *are called the* ***trivial ideals*** *of* L*.*

Definition 1.11 (Simple Lie Algebra). *A Lie algebra* L *is said to be* ***simple*** *if it is non-abelian and it contains no non-trivial ideals.*

Definition 1.12 (Semi-Simple Lie Algebra). *A Lie algebra* L *is said to be* ***semi-simple*** *if it contains no non-trivial abelian ideals.*

Remark 1.5. Note that any simple Lie algebra is also semi-simple. The requirement that a simple Lie algebra be non-abelian is due to the 1-dimensional abelian Lie algebra, which would otherwise be the only simple Lie algebra which is not semi-simple.

Definition 1.13 (Derived Subalgebra). *Let* L *be a Lie algebra. The Lie subalgebra* $L' := [L, L]$ *is called the* ***derived subalgebra*** *of* L*.*

Hence, once we have a Lie algebra we can compute the derived subalgebra $L' := [L, L]$. However since L' is by itself an algebra we can compute its own derived subalgebra $L'' := [L', L']$ (which is the derived subalgebra of the derived subalgebra of L). And of course we can go on forever.

Definition 1.14 (Derived Series). *The sequence* $L \supseteq L' \supseteq L'' \supseteq \cdots \supseteq L^{(n)} \supseteq \cdots$ *of Lie subalgebras is called the* ***derived series*** *of* L *usually denoted by* $L^{(n)}$*.*

Definition 1.15 (Solvable Lie Algebra). *A Lie algebra* L *is* ***solvable*** *if there exists* $k \in \mathbb{N}$ *such that* $L^{(k)} = 0$*.*

Recall that the direct sum of vector spaces $V \oplus W$ has $V \times W$ as its underlying set and operations defined componentwise.

Definition 1.16 (Direct Sum Of Lie Algebras). *Let* L_1 *and* L_2 *be Lie algebras. The* ***direct sum*** $L_1 \oplus_{\text{Lie}} L_2$ *has* $L_1 \oplus L_2$ *as its underlying vector space and Lie bracket defined as*

$$[x_1 + x_2, y_1 + y_2]_{L_1 \oplus_{\text{Lie}} L_2} := [x_1, y_1]_{L_1} + [x_2, y_2]_{L_2}$$

for all $x_1, y_1 \in L_1$ *and* $x_2, y_2 \in L_2$*. Alternatively, by identifying* L_1 *and* L_2 *with the subspaces* $L_1 \oplus 0$ *and* $0 \oplus L_2$ *of* $L_1 \oplus L_2$ *respectively, we require*

$$[L_1, L_2]_{L_1 \oplus_{\text{Lie}} L_2} = 0.$$

In the following, we will drop the "Lie" subscript and understand $\oplus$ *to mean* $\oplus_{\text{Lie}}$ *whenever the summands are Lie algebras.*

There is a weaker notion than the direct sum, defined only for Lie algebras.

Definition 1.17 (Semi-Direct Sum Of Lie Algebras). *Let* R *and* L *be Lie algebras. The* ***semi-direct sum*** $R \oplus_s L$ *has* $R \oplus L$ *as its underlying vector space and Lie bracket satisfying*

$$[R, L]_{R \oplus_s L} \subseteq R,$$

i.e. R *is an ideal of* $R \oplus_s L$*.*

We are now ready to state Levi's decomposition theorem.

Theorem 1.1 (Levi). *Any finite-dimensional complex Lie algebra* L *can be decomposed as*

$$L = R \oplus_s (L_1 \oplus \cdots \oplus L_n)$$

where R *is a solvable Lie algebra and* $L_1, \ldots, L_n$ *are simple Lie algebras.*

As of today, no general classification of solvable Lie algebras is known, except for some special cases (e.g. in low dimensions). In contrast, the finite dimensional, simple, complex Lie algebras have been classified completely.

Proposition 1.1. *A Lie algebra is semi-simple if, and only if, it can be expressed as a direct sum of simple Lie algebras.*

Hence, the simple Lie algebras are the basic building blocks from which one can build any semi-simple Lie algebra. Then, by Levi's theorem, the classification of simple Lie algebras easily extends to a classification of all semi-simple Lie algebras.

In order to do computations, it is useful to introduce a basis $\{e_i\}$ on L. Recall that an algebra is nothing else but a vector space with an extra operation. Hence, we can simply pick a basis $\{e_i\}$ on the vector space, and then examine how the Lie bracket behaves when we plug in, not any random element of algebra (i.e of the vector space) but specifically the elements of the basis.

Definition 1.18 (Structure Constants)**.** *Let L be a Lie algebra over K and let $\{e_i\}$ be a basis of the underlying vector space. Then, we have*

$$[e_i, e_j] = C^k{}_{ij} e_k$$

for some $C^k{}_{ij} \in K$. The numbers $C^k{}_{ij}$ are called the ***structure constants*** *of L with respect to the basis $\{e_i\}$.*

Remark 1.6. Since the operation of the algebra $\bullet\colon A \times A \to A$, sends two elements of the algebra to an element of the algebra, this can be translated as sending two elements of the vector space to an element of the vector space, i.e $[e_i, e_j] = v \in V$ for some fixed i and j. However since the final result v is again an element of the vector space it can also be expressed as a linear combination of the basis $v = v^k e_k$. This v^k is actually the structure constants (again for some fixed i and j, if we do not fix them we have to include them on the v^k hence we obtain $v^k \to C^k{}_{ij}$). This is why it is guaranteed that the structure constants $C^k{}_{ij} \in K$ exist.

In terms of the structure constants, the anti-symmetry of the Lie bracket reads

$$[e_i, e_j] = -[e_j, e_i] \implies C^k{}_{ij} e_k = -C^k{}_{ji} e_k \implies C^k{}_{ij} = -C^k{}_{ji}$$

while after some trivial calculations one can show that the Jacobi identity becomes

$$C^n{}_{im} C^m{}_{jk} + C^n{}_{jm} C^m{}_{ki} + C^n{}_{km} C^m{}_{ij} = 0.$$

1.2 The Adjoint Map & The Killing Form

Definition 1.19 (Adjoint Map)**.** *Let L be a Lie algebra over K and let $x \in L$. The* adjoint map *with respect to x is the K-linear map*

$$\begin{aligned} \mathrm{ad}_x\colon L &\xrightarrow{\sim} L \\ y &\mapsto \mathrm{ad}_x(y) := [x, y]. \end{aligned}$$

The linearity of ad_x follows from the linearity of the bracket in the second argument, while the linearity in the first argument of the bracket implies that the map

$$\begin{aligned} \mathrm{ad}\colon L &\xrightarrow{\sim} \mathrm{End}(L) \\ x &\mapsto \mathrm{ad}(x) := \mathrm{ad}_x\,. \end{aligned}$$

itself is also linear. In fact, more is true. Recall that $\mathrm{End}(L)$ is a Lie algebra with bracket

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

Then, we have the following.

Proposition 1.2. *The map $\mathrm{ad}\colon L \xrightarrow{\sim} \mathrm{End}(L)$ is a Lie algebra homomorphism.*

Proof. It remains to check that ad preserves the brackets. Let $x, y, z \in L$. Then

$$\begin{aligned}
\mathrm{ad}_{[x,y]}(z) &:= [[x,y],z] && \text{(definition of ad)}\\
&= -[[y,z],x] - [[z,x],y] && \text{(Jacobi's identity)}\\
&= [x,[y,z]] - [y,[x,z]] && \text{(anti-symmetry)}\\
&= \mathrm{ad}_x(\mathrm{ad}_y(z)) - \mathrm{ad}_y(\mathrm{ad}_x(z))\\
&= (\mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x)(z)\\
&= [\mathrm{ad}_x, \mathrm{ad}_y](z).
\end{aligned}$$

Hence, we have $\mathrm{ad}([x,y]) = [\mathrm{ad}(x), \mathrm{ad}(y)]$. □

By choosing a basis for the vector space, we can express the adjoint map in terms of components with respect to the basis as follows. Start by noting that

$$\begin{aligned}
\mathrm{ad}\colon L &\xrightarrow{\sim} \mathrm{End}(L)\\
x &\mapsto \mathrm{ad}(x) := \mathrm{ad}_x\,.
\end{aligned}$$

which means that ad_x is an element of $\mathrm{End}(L)$ hence an endomorphism of L. Recall that for any vector space V: $\mathrm{End}(V) \cong_{\mathrm{vec}} T^1_1 V$ which means that if $\phi \in \mathrm{End}(V)$, we can think of $\phi \in T^1_1 V$, using the same symbol, as $\phi(\omega, v) := \omega(\phi(v))$ hence the components of $\phi \in \mathrm{End}(V)$ are $\phi^a{}_b := \epsilon^a(\phi(e_b))$.

So, in our case, let $\{e_i\}$ and $\{\varepsilon^i\}$ be a basis and its dual basis of the underlying vector space of a Lie algebra L. Then

$$\begin{aligned}
(\mathrm{ad}_{e_i})^k{}_j &:= \varepsilon^k(\mathrm{ad}_{e_i}(e_j))\\
&= \varepsilon^k([e_i, e_j])\\
&= \varepsilon^k(C^m{}_{ij} e_m)\\
&= C^m{}_{ij}\varepsilon^k(e_m)\\
&= C^k{}_{ij}.
\end{aligned}$$

In other words, the adjoint map represents the structure constants without the need of choosing a basis.

Definition 1.20 (Killing Form)**.** *Let L be a Lie algebra over K. The **Killing form** on L is the K-bilinear map*

$$\begin{aligned}
\kappa\colon L \times L &\to K\\
(x,y) &\mapsto \kappa(x,y) := \mathrm{tr}(\mathrm{ad}_x \circ \mathrm{ad}_y),
\end{aligned}$$

where tr *is the usual trace on the vector space* $\mathrm{End}(L)$.

Note that the Killing form is not a "form" in the sense that we defined previously. In fact, since L is finite-dimensional, the trace is cyclic and thus κ is symmetric, i.e.

$$\forall\, x, y \in L:\ \kappa(x,y) = \kappa(y,x).$$

An important property of κ is its associativity with respect to the bracket.

Proposition 1.3. *Let L be a Lie algebra. For any $x, y, z \in L$, we have*

$$\kappa([x,y],z) = \kappa(x,[y,z]).$$

Proof. This follows easily from the fact that ad is a homomorphism.

$$\begin{aligned}
\kappa([x,y],z) &:= \operatorname{tr}(\operatorname{ad}_{[x,y]} \circ \operatorname{ad}_z) \\
&= \operatorname{tr}([\operatorname{ad}_x, \operatorname{ad}_y] \circ \operatorname{ad}_z) \\
&= \operatorname{tr}((\operatorname{ad}_x \circ \operatorname{ad}_y - \operatorname{ad}_y \circ \operatorname{ad}_x) \circ \operatorname{ad}_z) \\
&= \operatorname{tr}(\operatorname{ad}_x \circ \operatorname{ad}_y \circ \operatorname{ad}_z) - \operatorname{tr}(\operatorname{ad}_y \circ \operatorname{ad}_x \circ \operatorname{ad}_z) \\
&= \operatorname{tr}(\operatorname{ad}_x \circ \operatorname{ad}_y \circ \operatorname{ad}_z) - \operatorname{tr}(\operatorname{ad}_x \circ \operatorname{ad}_z \circ \operatorname{ad}_y) \\
&= \operatorname{tr}(\operatorname{ad}_x \circ (\operatorname{ad}_y \circ \operatorname{ad}_z - \operatorname{ad}_z \circ \operatorname{ad}_y)) \\
&= \operatorname{tr}(\operatorname{ad}_x \circ [\operatorname{ad}_y, \operatorname{ad}_z]) \\
&= \operatorname{tr}(\operatorname{ad}_x \circ \operatorname{ad}_{[y,z]}) \\
&=: \kappa(x,[y,z]),
\end{aligned}$$

where we used the cyclicity of the trace. □

As we did for the adjoint map we can also express the Killing form in terms of components with respect to a basis.

Recall from linear algebra that if V is finite-dimensional, for any $\phi \in \operatorname{End}(V)$ we have $\operatorname{tr}(\phi) = \Phi^k{}_k$, where Φ is the matrix representing the linear map in any basis. Also, recall that the matrix representing $\phi \circ \psi$ is the product $\Phi\Psi$. Using these, by letting $\{e_i\}$ and $\{\varepsilon^i\}$ be a basis and its dual basis of the underlying vector space of a Lie algebra L we have

$$\begin{aligned}
\kappa_{ij} &:= \kappa(e_i, e_j) \\
&= \operatorname{tr}(\operatorname{ad}_{e_i} \circ \operatorname{ad}_{e_j}) \\
&= (\operatorname{ad}_{e_i} \circ \operatorname{ad}_{e_j})^k{}_k \\
&= (\operatorname{ad}_{e_i})^m{}_k (\operatorname{ad}_{e_j})^k{}_m \\
&= C^m{}_{ik} C^k{}_{jm},
\end{aligned}$$

where we used the same notation for the linear maps and their matrices.

We can use κ to give a further equivalent characterisation of semi-simplicity.

Proposition 1.4 (Cartan's criterion)**.** *A Lie algebra L is semi-simple if, and only if, the Killing form κ is non-degenerate, i.e.*

$$(\forall\, y \in L : \kappa(x,y) = 0) \Rightarrow x = 0.$$

Hence, if L is semi-simple, then κ is a pseudo inner product on L. Recall the following definition from linear algebra.

Definition 1.21. *A linear map $\phi\colon V \xrightarrow{\sim} V$ is said to be* symmetric *with respect to the pseudo inner product $B(-,-)$ on V if*

$$\forall\, v,w \in V : \; B(\phi(v),w) = B(v,\phi(w)).$$

If, instead, we have

$$\forall\, v,w \in V : \; B(\phi(v),w) = -B(v,\phi(w)),$$

then ϕ is said to be anti-symmetric *with respect to B.*

The associativity property of κ with respect to the bracket can be restated by saying that, for any $z \in L$, the linear map ad_z is anti-symmetric with respect to κ, i.e.

$$\forall\, x,y \in L : \; \kappa(\operatorname{ad}_z(x),y) = -\kappa(x,\operatorname{ad}_z(y)).$$

1.3 The Fundamental Roots & The Weyl Group

We will now focus on finite-dimensional semi-simple complex Lie algebras, whose classification hinges on the existence of a special type of subalgebra.

Definition 1.22 (Cartan Subalgebra)**.** *Let L be a d-dimensional Lie algebra. A* ***Cartan subalgebra*** *H of L is a maximal Lie subalgebra of L with the following property: there exists a basis $\{h_1, \dots, h_r\}$ of H which can be extended to a basis $\{h_1, \dots, h_r, e_1, \dots, e_{d-r}\}$ of L such that $e_1, \dots, e_{d-r}$ are eigenvectors of $\mathrm{ad}(h)$ for any $h \in H$, i.e.*

$$\forall\, h \in H : \exists\, \lambda_\alpha(h) \in \mathbb{C} : \ \mathrm{ad}(h)e_\alpha = \lambda_\alpha(h)e_\alpha,$$

for each $1 \leq \alpha \leq d - r$.

The basis $\{h_1, \dots, h_r, e_1, \dots, e_{d-r}\}$ is known as a *Cartan-Weyl basis* of L. Of course, we would like to know when we can find such a subalgebra.

Theorem 1.2. *Let L be a finite-dimensional semi-simple complex Lie algebra. Then*

i) L possesses a Cartan subalgebra;

ii) all Cartan subalgebras of L have the same dimension, called the rank *of L;*

iii) any of Cartan subalgebra H of L is abelian, i.e $[H, H] = 0$.

Note that we can think of the λ_α appearing above as a map $\lambda_\alpha \colon H \to \mathbb{C}$. Moreover, for any $z \in \mathbb{C}$ and $h, h' \in H$, we have

$$\begin{aligned}
\lambda_\alpha(zh + h')e_\alpha &= \mathrm{ad}(zh + h')e_\alpha \\
&= [zh + h', e_\alpha] \\
&= z[h, e_\alpha] + [h', e_\alpha] \\
&= z\lambda_\alpha(h)e_\alpha + \lambda_\alpha(h')e_\alpha \\
&= (z\lambda_\alpha(h) + \lambda_\alpha(h'))e_\alpha,
\end{aligned}$$

Hence λ_α is a $\mathbb{C}$-linear map $\lambda_\alpha \colon H \xrightarrow{\sim} \mathbb{C}$, and thus $\lambda_\alpha \in H^*$.

Definition 1.23 (Roots)**.** *The maps $\lambda_1, \dots, \lambda_{d-r} \in H^*$ are called the* ***roots*** *of L.*

Definition 1.24 (Root Set)**.** *The collection of the roots of an algebra*

$$\Phi := \{\lambda_\alpha \mid 1 \leq \alpha \leq d - r\} \subseteq H^*$$

is called the ***root set*** *of L.*

One can show that if λ_α were the zero map, then we would have $e_\alpha \in H$. Thus, we must have $0 \notin \Phi$. Note that a consequence of the anti-symmetry of each $\mathrm{ad}(h)$ with respect to the Killing form κ is that

$$\lambda \in \Phi \ \Rightarrow \ -\lambda \in \Phi.$$

Hence Φ is not a linearly independent subset of H^*.

Definition 1.25 (Fundamental Roots)**.** *A set of* ***fundamental roots*** *$\Pi := \{\pi_1, \dots, \pi_f\}$ is a subset $\Pi \subseteq \Phi$ such that*

a) Π is a linearly independent subset of H^*;

b) for each $\lambda \in \Phi$, there exist $n_1, \dots, n_f \in \mathbb{N}$ and $\varepsilon \in \{+1, -1\}$ such that

$$\lambda = \varepsilon \sum_{i=1}^{f} n_i \pi_i.$$

Since $n_i \in \mathbb{N}$ this means that they are all positive numbers (as they should be by the definition of a basis).By also picking an $\varepsilon \in \{+1, -1\}$ to be either +1 or -1, we are able, no matter the choice of fundamental roots, to obtain the opposite signed ones. That way, observe that, for any $\lambda \in \Phi$, the coefficients of $\pi_1, \dots, \pi_f$ in the expansion above always have the same sign. We can write the last equation more concisely as $\lambda \in \mathrm{span}_{\varepsilon,\mathbb{N}}(\Pi)$ where in general $\mathrm{span}_{\varepsilon,\mathbb{N}}(\Pi) \neq \mathrm{span}_{\mathbb{Z}}(\Pi)$.

Theorem 1.3. *Let L be a finite-dimensional semi-simple complex Lie algebra. Then* *i) a set $\Pi \subseteq \Phi$ of fundamental roots always exists;*

ii) we have $\operatorname{span}_{\mathbb{C}}(\Pi) = H^*$, that is, Π is a basis of H^*.

Corollary 1.1. *We have $|\Pi| = r$, where r is the rank of L.*

Proof. Since Π is a basis, $|\Pi| = \dim H^* = \dim H = r$. □

We would now like to use κ to define a pseudo inner product on H^*. We know from linear algebra that a pseudo inner product $B(-,-)$ on a finite-dimensional vector space V over K induces a linear isomorphism

$$\begin{aligned} i\colon V &\xrightarrow{\sim} V^* \\ v &\mapsto i(v) := B(v,-) \end{aligned}$$

which can be used to define a pseudo inner product $B^*(-,-)$ on V^* as

$$\begin{aligned} B^*\colon V^* \times V^* &\to K \\ (\phi,\psi) &\mapsto B^*(\phi,\psi) := B(i^{-1}(\phi), i^{-1}(\psi)). \end{aligned}$$

We would like to apply this to the restriction of κ to the Cartan subalgebra. However, a pseudo inner product on a vector space is not necessarily a pseudo inner product on a subspace, since the non-degeneracy condition may fail when considered on a subspace.

Proposition 1.5. *The restriction of κ to H is a pseudo inner product on H.*

Proof. Bilinearity and symmetry are automatically satisfied. It remains to show that κ is non-degenerate on H.

i) Let $\{h_1,\dots,h_r,e_{r+1},\dots,e_d\}$ be a Cartan-Weyl basis of L and let $\lambda_\alpha \in \Phi$. Then

$$\begin{aligned} \lambda_\alpha(h_j)\kappa(h_i,e_\alpha) &= \kappa(h_i,\lambda_\alpha(h_j)e_\alpha) \\ &= \kappa(h_i,[h_j,e_\alpha]) \\ &= \kappa([h_i,h_j],e_\alpha) \\ &= \kappa(0,e_\alpha) \\ &= 0. \end{aligned}$$

Since $\lambda_\alpha \neq 0$, there is some h_j such that $\lambda_\alpha(h_j) \neq 0$ and hence

$$\kappa(h_i,e_\alpha) = 0.$$

By linearity, we have $\kappa(h,e_\alpha) = 0$ for any $h \in H$ and any e_α.

ii) Let $h \in H \subseteq L$. Since κ is non-degenerate on L, we have

$$\big(\forall\, x \in L : \kappa(h,x) = 0\big) \Rightarrow h = 0.$$

Expand $x \in L$ in the Cartan-Weyl basis as

$$x = h' + e$$

where $h' := x^i h_i$ and $e := x^\alpha e_\alpha$. Then, we have

$$\kappa(h,x) = \kappa(h,h') + x^\alpha \kappa(h,e_\alpha) = \kappa(h,h').$$

Thus, the non-degeneracy condition reads

$$\big(\forall\, h' \in H : \kappa(h,h') = 0\big) \Rightarrow h = 0,$$

which is what we wanted. □

We can now define

$$\begin{aligned}\kappa^*\colon H^* \times H^* &\to \mathbb{C}\\ (\mu,\nu) &\mapsto \kappa^*(\mu,\nu) := \kappa(i^{-1}(\mu), i^{-1}(\nu)),\end{aligned}$$

where $i\colon H \xrightarrow{\sim} H^*$ is the linear isomorphism induced by κ.

Remark 1.7. If $\{h_i\}$ is a basis of H, the components of κ^* with respect to the dual basis satisfy

$$(\kappa^*)^{ij}\kappa_{jk} = \delta^i_k.$$

Hence, we can write

$$\kappa^*(\mu,\nu) = (\kappa^*)^{ij}\mu_i\nu_j,$$

where $\mu_i := \mu(h_i)$.

We now turn our attention to the real subalgebra $H^*_{\mathbb{R}} := \operatorname{span}_{\mathbb{R}}(\Pi)$. Note that we have the following chain of inclusions

$$\Pi \subseteq \Phi \subseteq \operatorname{span}_{\varepsilon,\mathbb{N}}(\Pi) \subseteq \underbrace{\operatorname{span}_{\mathbb{R}}(\Pi)}_{H^*_{\mathbb{R}}} \subseteq \underbrace{\operatorname{span}_{\mathbb{C}}(\Pi)}_{H^*}.$$

The restriction of κ^* to $H^*_{\mathbb{R}}$ leads to a surprising result.

Theorem 1.4. *i) For any $\alpha,\beta \in H^*_{\mathbb{R}}$, we have $\kappa^*(\alpha,\beta) \in \mathbb{R}$.*

ii) $\kappa^*\colon H^*_{\mathbb{R}} \times H^*_{\mathbb{R}} \to \mathbb{R}$ is an inner product on $H^*_{\mathbb{R}}$.

This is indeed a surprise! Upon restriction to $H^*_{\mathbb{R}}$, instead of being weakened, the non-degeneracy of κ^* gets strengthened to positive definiteness. Now that we have a proper real inner product, we can define some familiar notions from basic linear algebra, such as lengths and angles.

Definition 1.26 (Length & Angle)**.** *Let $\alpha,\beta \in H^*_{\mathbb{R}}$. Then, we define*

*i) the **length** of α as $|\alpha| := \sqrt{\kappa^*(\alpha,\alpha)}$;*

*ii) the **angle** between α and β as $\varphi := \cos^{-1}\left(\dfrac{\kappa^*(\alpha,\beta)}{|\alpha||\beta|}\right)$.*

We need one final ingredient for our classification result.

Definition 1.27 (Weyl Transformation)**.** *For any $\lambda \in \Phi \subseteq H^*_{\mathbb{R}}$, define the linear map s_λ called a **Weyl transformation***

$$\begin{aligned}s_\lambda\colon H^*_{\mathbb{R}} &\xrightarrow{\sim} H^*_{\mathbb{R}}\\ \mu &\mapsto s_\lambda(\mu),\end{aligned}$$

where

$$s_\lambda(\mu) := \mu - 2\frac{\kappa^*(\lambda,\mu)}{\kappa^*(\lambda,\lambda)}\,\lambda.$$

Definition 1.28 (Weyl Group)**.** *The set*

$$W := \{s_\lambda \mid \lambda \in \Phi\}$$

*is a group under composition of maps, an it is called the **Weyl group**.*

Theorem 1.5. *i) The Weyl group W is generated by the fundamental roots in Π, in the sense that for some $1 \le n \le r$, with $r = |\Pi|$,*

$$\forall\, w \in W : \exists\, \pi_1,\dots,\pi_n \in \Pi :\ w = s_{\pi_1} \circ s_{\pi_2} \circ \cdots \circ s_{\pi_n};$$

ii) Every root can be produced from a fundamental root by the action of W, i.e.

$$\forall\, \lambda \in \Phi : \exists\, \pi \in \Pi : \exists\, w \in W :\ \lambda = w(\pi);$$

iii) The Weyl group permutes the roots, that is,

$$\forall\, \lambda \in \Phi : \forall\, w \in W : \; w(\lambda) \in \Phi.$$

1.4 Dynkin Diagrams & The Cartan Classification

Consider, for any $\pi_i, \pi_j \in \Pi$, the action of the Weyl transformation

$$s_{\pi_i}(\pi_j) := \pi_j - 2\frac{\kappa^*(\pi_i, \pi_j)}{\kappa^*(\pi_i, \pi_i)}\, \pi_i.$$

However, since $s_{\pi_i}(\pi_j) \in \Phi$ and $\Phi \subseteq \operatorname{span}_{\varepsilon,\mathbb{N}}(\Pi)$ this means that it must be written in terms of the basis as:

$$s_{\pi_i}(\pi_j) \in \Phi = \left(\varepsilon \sum_{i=1}^{f} n_i \pi_i \right) = C_1 \pi_j + C_2 \pi_i$$

But it is already written in such from since

$$s_{\pi_i}(\pi_j) = \pi_j - 2\frac{\kappa^*(\pi_i, \pi_j)}{\kappa^*(\pi_i, \pi_i)}\, \pi_i = 1\, \pi_j + \left(-2\frac{\kappa^*(\pi_i, \pi_j)}{\kappa^*(\pi_i, \pi_i)} \right) \pi_i$$

and from the first coefficient (a.k.a the number 1) which is positive, we conclude that for all $1 \leq i \neq j \leq r$ we must have

$$-2\frac{\kappa^*(\pi_i, \pi_j)}{\kappa^*(\pi_i, \pi_i)} \in \mathbb{N}.$$

Definition 1.29 (Cartan Matrix)**.** *The **Cartan matrix** of a Lie algebra is the $r \times r$ matrix C with entries*

$$C_{ij} := 2\frac{\kappa^*(\pi_i, \pi_j)}{\kappa^*(\pi_i, \pi_i)}$$

Remark 1.8. The C_{ij} should not be confused with the structure constants ${C^k}_{ij}$

Theorem 1.6. *To every simple finite-dimensional complex Lie algebra there corresponds a unique Cartan matrix and vice versa (up to relabelling of the basis elements).*

Of course, not every matrix can be a Cartan matrix. For instance, since $C_{ii} = 2$ (no summation implied), the diagonal entries of C are all equal to 2, while the off-diagonal entries are either zero or negative. In general, $C_{ij} \neq C_{ji}$, so the Cartan matrix is not symmetric, but if $C_{ij} = 0$, then necessarily $C_{ji} = 0$.

We have thus reduced the problem of classifying the simple finite-dimensional complex Lie algebras to that of finding all the Cartan matrices. This can, in turn, be reduced to the problem of determining all the inequivalent Dynkin diagrams.

Definition 1.30 (Bond Number)**.** *Given a Cartan matrix C, the ij-th **bond number** is*

$$n_{ij} := C_{ij} C_{ji} \qquad \textit{(no summation implied).}$$

Note that we have

$$\begin{aligned} n_{ij} &= 4\, \frac{\kappa^*(\pi_i, \pi_j)}{\kappa^*(\pi_i, \pi_i)}\, \frac{\kappa^*(\pi_j, \pi_i)}{\kappa^*(\pi_j, \pi_j)} \\ &= 4 \left(\frac{\kappa^*(\pi_i, \pi_j)}{|\pi_i||\pi_j|} \right)^2 \\ &= 4 \cos^2 \varphi, \end{aligned}$$

where φ is the angle between π_i and π_j.

For $i \neq j$, the angle φ is neither zero nor $180°$, hence $0 \leq \cos^2 \varphi < 1$, and therefore

$$n_{ij} \in \{0, 1, 2, 3\}.$$

Since $C_{ij} \leq 0$ for $i \neq j$, the only possibilities are

C_{ij}	C_{ji}	n_{ij}
0	0	0
-1	-1	1
-1	-2	2
-1	-3	3

Note that while the Cartan matrices are not symmetric, swapping any pair of C_{ij} and C_{ji} gives a Cartan matrix which represents the same Lie algebra as the original matrix, with two elements from the Cartan-Weyl basis swapped. This is why we have not included $(-2, -1)$ and $(-3, -1)$ in the table above.

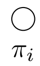
If $n_{ij} = 2$ or 3, then the corresponding fundamental roots have different lengths, i.e. either $|\pi_i| < |\pi_j|$ or $|\pi_i| > |\pi_j|$. We also have the following result.

Proposition 1.6. *The roots of a simple Lie algebra have, at most, two distinct lengths.*

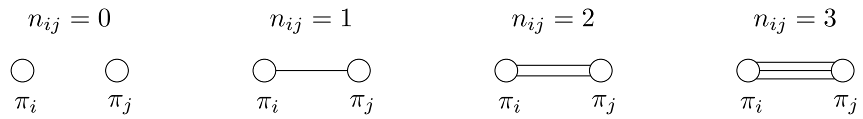
The redundancy of the Cartan matrices highlighted above is nicely taken care of by considering Dynkin diagrams.

Definition 1.31 (Dynkin Diagram)**.** *A **Dynkin diagram** associated to a Cartan matrix is constructed as follows.*

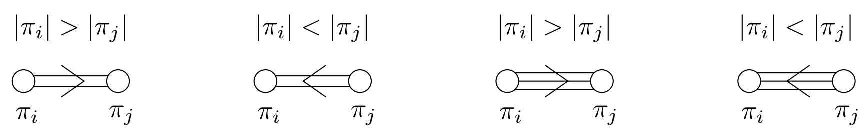
1. *Draw a circle for every fundamental root in $\pi_i \in \Pi$;*



2. *Draw n_{ij} lines between the circles representing the roots π_i and π_j;*



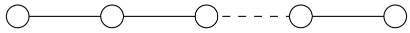
3. *If $n_{ij} = 2$ or 3, draw an arrow on the lines from the longer root to the shorter root.*

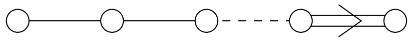


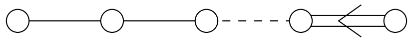
Dynkin diagrams completely characterise any set of fundamental roots, from which we can reconstruct the entire root set by using the Weyl transformations. The root set can then be used to produce a Cartan-Weyl basis. We are now finally ready to state the much awaited classification theorem.

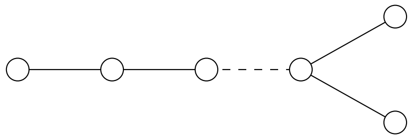
Theorem 1.7 (Killing, Cartan)**.** *Any simple finite-dimensional complex Lie algebra can be reconstructed from its set of fundamental roots Π, which only come in the following forms.*

i) There are 4 infinite families

A_n $\quad n \geq 1$ 

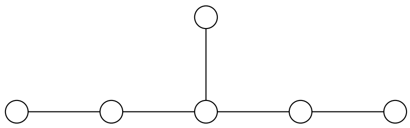
B_n $\quad n \geq 2$ 

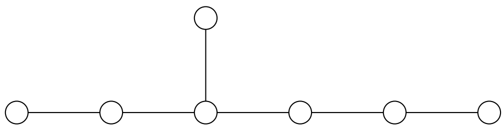
C_n $\quad n \geq 3$ 

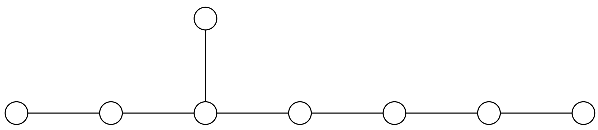
D_n $\quad n \geq 4$ 

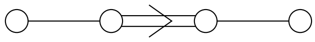
where the restrictions on n ensure that we don't get repeated diagrams (the diagram D_2 is excluded since it is disconnected and does not correspond to a simple Lie algebra)

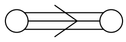
ii) five exceptional cases

E_6 

E_7 

E_8 

F_4 

G_2 

and no other. These are all the possible (connected) Dynkin diagrams.

At last, we have achieved a classification of all simple finite-dimensional complex Lie algebras. The finite-dimensional semi-simple complex Lie algebras are direct sums of simple Lie algebras, and correspond to disconnected Dynkin diagrams whose connected components are the ones listed above.

Chapter 2

Topological Manifolds

2.1 Topological Manifolds

Definition 2.1 (Topological Manifold)**.** *A paracompact, Hausdorff, topological space* $(M, \mathcal{O})$ *is called a* ***d-dimensional topological manifold*** *if for every point* $p \in M$ *there exist a neighbourhood* $U(p)$ *and a homeomorphism* $x\colon U(p) \to x(U(p)) \subseteq \mathbb{R}^d$*. We also write* $\dim M = d$*.*

Intuitively, a d-dimensional manifold is a topological space which locally (i.e. around each point) looks like $\mathbb{R}^d$. Note that, strictly speaking, what we have just defined are *real* topological manifolds. We could define *complex* topological manifolds as well, simply by requiring that the map x be a homeomorphism onto an open subset of $\mathbb{C}^d$.

Proposition 2.1. *Let* M *be a* d*-dimensional manifold and let* $U, V \subseteq M$ *be open, with* $U \cap V \neq \varnothing$*. If* x *and* y *are two homeomorphisms*

$$x\colon U \to x(U) \subseteq \mathbb{R}^d \qquad \text{and} \qquad y\colon V \to y(V) \subseteq \mathbb{R}^{d'},$$

then $d = d'$*.*

This ensures that the concept of dimension is indeed well-defined, i.e. it is the same at every point, at least on each connected component of the manifold.

Example 2.1. Trivially, $\mathbb{R}^d$ is a d-dimensional manifold for any $d \geq 1$. The space S^1 is a 1-dimensional manifold while the spaces S^2, C and T^2 are 2-dimensional manifolds.

Definition 2.2 (Topological Submanifold)**.** *Let* $(M, \mathcal{O})$ *be a topological manifold and let* $N \subseteq M$*. Then* $(N, \mathcal{O}|_N)$ *is called a* ***submanifold*** *of* $(M, \mathcal{O})$ *if it is a manifold in its own right.*

Example 2.2. The space S^1 is a submanifold of $\mathbb{R}^2$ while the spaces S^2, C and T^2 are submanifolds of $\mathbb{R}^3$.

Definition 2.3 (Product Manifold)**.** *Let* $(M, \mathcal{O}_M)$ *and* $(N, \mathcal{O}_N)$ *be topological manifolds of dimension* m *and* n*, respectively. Then,* $(M \times N, \mathcal{O}_{M\times N})$ *is a topological manifold of dimension* $m + n$ *called the* ***product manifold****.*

Example 2.3. We have $T^2 = S^1 \times S^1$ not just as topological spaces, but as topological manifolds as well. This is a special case of the n-torus:

$$T^n := \underbrace{S^1 \times S^1 \times \cdots \times S^1}_{n \text{ times}},$$

which is an n-dimensional manifold.

Example 2.4. The cylinder $C = S^1 \times \mathbb{R}$ is a 2-dimensional manifold.

2.2 Charts & Atlases

Definition 2.4 (Chart)**.** *Let* $(M, \mathcal{O})$ *be a* d*-dimensional manifold. Then, a pair* (U, x) *where* $U \in \mathcal{O}$ *and* $x\colon U \to x(U) \subseteq \mathbb{R}^d$ *is a homeomorphism, is said to be a* ***chart*** *of the manifold.*

Definition 2.5 (Components / Co-Ordinates Of A Chart)**.** *The **component functions (or maps)** of $x\colon U \to x(U) \subseteq \mathbb{R}^d$ are the maps:*

$$\begin{aligned} x^i \colon U &\to \mathbb{R} \\ p &\mapsto \mathrm{proj}_i(x(p)) \end{aligned}$$

*for $1 \leq i \leq d$, where $\mathrm{proj}_i(x(p))$ is the i-th component of $x(p) \in \mathbb{R}^d$. The $x^i(p)$ are called the **co-ordinates** of the point $p \in U$ with respect to the chart (U, x).*

Definition 2.6 (Atlas)**.** *An **atlas** of a manifold M is a collection $\mathscr{A} := \{(U_\alpha, x_\alpha) \mid \alpha \in \mathcal{A}\}$ of charts such that:*

$$\bigcup_{\alpha \in \mathcal{A}} U_\alpha = M.$$

Definition 2.7 ($\mathcal{C}^0$-Compatible Charts)**.** *Two charts (U, x) and (V, y) are said to be $\mathcal{C}^0$**-compatible** if either $U \cap V = \varnothing$ or the map:*

$$y \circ x^{-1} \colon x(U \cap V) \to y(U \cap V)$$

is continuous.

Note that $y \circ x^{-1}$ is a map from a subset of $\mathbb{R}^d$ to a subset of $\mathbb{R}^d$.

$$\begin{tikzcd} & U \cap V \subseteq M \arrow[dl, "x"'] \arrow[dr, "y"] & \\ x(U \cap V) \subseteq \mathbb{R}^d \arrow[rr, "y \circ x^{-1}"'] & & y(U \cap V) \subseteq \mathbb{R}^d \end{tikzcd}$$

Since the maps x and y are homeomorphisms, the composition map $y \circ x^{-1}$ is also a homeomorphism and hence continuous. Therefore, any two charts on a topological manifold are $\mathcal{C}^0$-compatible. This definition my thus seem redundant since it applies to every pair of charts. However, it is just a "warm up" since we will later refine this definition and define the *differentiability* of maps on a manifold in terms of $\mathcal{C}^k$-compatibility of charts.

Definition 2.8 (Chart Transition Map)**.** *The map $y \circ x^{-1}$ (and its inverse $x \circ y^{-1}$) is called the **chart transition map**.*

Definition 2.9 ($\mathcal{C}^0$-Atlas)**.** *A $\mathcal{C}^0$**-atlas** of a manifold is an atlas of pairwise $\mathcal{C}^0$-compatible charts.*

Note that any atlas is also a $\mathcal{C}^0$*-atlas.*

Definition 2.10 (Maximal Atlas)**.** *A $\mathcal{C}^0$-atlas $\mathscr{A}$ is said to be a **maximal atlas** if for every $(U, x) \in \mathscr{A}$, we have $(V, y) \in \mathscr{A}$ for all (V, y) charts that are $\mathcal{C}^0$-compatible with (U, x).*

Example 2.5. Not every $\mathcal{C}^0$-atlas is a maximal atlas. Indeed, consider $(\mathbb{R}, \mathcal{O}_{\mathrm{std}})$ and the atlas $\mathscr{A} := (\mathbb{R}, \mathrm{id}_{\mathbb{R}})$. Then $\mathscr{A}$ is not maximal since $((0,1), \mathrm{id}_{\mathbb{R}})$ is a chart which is $\mathcal{C}^0$-compatible with $(\mathbb{R}, \mathrm{id}_{\mathbb{R}})$ but $((0,1), \mathrm{id}_{\mathbb{R}}) \notin \mathscr{A}$.

We can now look at "objects on" topological manifolds from two points of view. For instance, consider a curve on a d-dimensional manifold M, i.e. a map $\gamma\colon \mathbb{R} \to M$. We now ask whether this curve is continuous, as it should be if models the trajectory of a particle on the "physical space" M.

A first answer is that $\gamma\colon \mathbb{R} \to M$ is continuous if it is continuous as a map between the topological spaces $\mathbb{R}$ and M.

However, the answer that may be more familiar to you from undergraduate physics is the following. We consider only a portion (open subset U) of the physical space M and, instead of studying the map $\gamma\colon \mathrm{preim}_\gamma(U) \to U$ directly, we study the map:

$$x \circ \gamma \colon \mathrm{preim}_\gamma(U) \to x(U) \subseteq \mathbb{R}^d,$$

where (U, x) is a chart of M. More likely, you would be checking the continuity of the co-ordinate maps $x^i \circ \gamma$, which would then imply the continuity of the "real" curve $\gamma \colon \mathrm{preim}_\gamma(U) \to U$ (real, as opposed to its co-ordinate representation).

$$
\begin{tikzcd}
 & y(U) \subseteq \mathbb{R}^d & \\
\mathrm{preim}_\gamma(U) \subseteq \mathbb{R} \arrow[ur, "y\circ\gamma"] \arrow[r, "\gamma"] \arrow[dr, "x\circ\gamma"'] & U \subseteq M \arrow[u, "y"'] \arrow[d, "x"] & \\
 & x(U) \subseteq \mathbb{R}^d \arrow[uu, bend right=80, "y\circ x^{-1}"'] &
\end{tikzcd}
$$

At some point you may wish to use a different "co-ordinate system" to answer a different question. In this case, you would chose a different chart (U, y) and then study the map $y \circ \gamma$ or its co-ordinate maps. Notice however that some results (e.g. the continuity of γ) obtained in the previous chart (U, x) can be immediately "transported" to the new chart (U, y) via the chart transition map $y \circ x^{-1}$. Moreover, the map $y \circ x^{-1}$ allows us to, intuitively speaking, forget about the inner structure (i.e. U and the maps γ, x and x) which, in a sense, is the real world, and only consider $\mathrm{preim}_\gamma(U) \subseteq \mathbb{R}$ and $x(U), y(U) \subseteq \mathbb{R}^d$ together with the maps between them, which is our representation of the real world.

As we already said, for a topological manifold $(M, \mathcal{O})$, the concept of a $\mathcal{C}^0$-atlas is fully redundant since every atlas is also a $\mathcal{C}^0$-atlas. We will now generalise the notion of a $\mathcal{C}^0$-atlas, or more precisely, the notion of $\mathcal{C}^0$-compatibility of charts, to something which is non-trivial and non-redundant.

Definition 2.11 (❀-Atlas)**.** *An atlas $\mathscr{A}$ for a topological manifold is called a* ***❀-atlas*** *if any two charts $(U, x), (V, y) \in \mathscr{A}$ are ❀-compatible, where the symbol ❀ is being used as a placeholder for any of the following:*

- *❀ $= \mathcal{C}^0$: this just reduces to the previous definition;*
- *❀ $= \mathcal{C}^k$: the transition maps are k-times continuously differentiable as maps between open subsets of $\mathbb{R}^{\dim M}$;*
- *❀ $= \mathcal{C}^\infty$: the transition maps are smooth (infinitely many times differentiable); equivalently, the atlas is $\mathcal{C}^k$ for all $k \geq 0$;*
- *❀ $= \mathcal{C}^\omega$: the transition maps are (real) analytic, which is stronger than being smooth;*
- *❀ = complex: if $\dim M$ is even, M is a* complex manifold *if the transition maps are continuous and satisfy the Cauchy-Riemann equations; its complex dimension is $\frac{1}{2}\dim M$.*

In other words, either $U \cap V = \varnothing$ or if $U \cap V \neq \varnothing$, then the transition map $y \circ x^{-1}$ from $x(U \cap V)$ to $y(U \cap V)$ must be ❀.

$$
\begin{tikzcd}
 & U \cap V \subseteq M \arrow[dl, "x"'] \arrow[dr, "y"] & \\
x(U \cap V) \subseteq \mathbb{R}^{\dim M} \arrow[rr, "y\circ x^{-1}"'] & & y(U \cap V) \subseteq \mathbb{R}^{\dim M}
\end{tikzcd}
$$

Theorem 2.1 (Whitney)**.** *Any maximal $\mathcal{C}^k$-atlas, with $k \geq 1$, contains a $\mathcal{C}^\infty$-atlas. Moreover, any two maximal $\mathcal{C}^k$-atlases that contain the same $\mathcal{C}^\infty$-atlas are identical.*

An immediate implication is that if we can find a $\mathcal{C}^1$-atlas for a manifold, then we can also assume the existence of a $\mathcal{C}^\infty$-atlas for that manifold. This is not the case for topological manifolds in general: a space with a $\mathcal{C}^0$-atlas may not admit any $\mathcal{C}^1$-atlas. But if we have at least a $\mathcal{C}^1$-atlas, then we can obtain a $\mathcal{C}^\infty$-atlas simply by removing charts, keeping only the ones which are $\mathcal{C}^\infty$-compatible.

Hence, for the purposes of this course, we will not distinguish between $\mathcal{C}^k$ $(k \geq 1)$ and $\mathcal{C}^\infty$-manifolds in the above sense.

We now give the explicit definition of a $\mathcal{C}^k$-manifold.

Definition 2.12 ($\mathcal{C}^k$-Manifold)**.** *A $\mathcal{C}^k$**-manifold** is a triple $(M, \mathcal{O}, \mathscr{A})$, where $(M, \mathcal{O})$ is a topological manifold and $\mathscr{A}$ is a maximal $\mathcal{C}^k$-atlas.*

Definition 2.13 (Smooth Manifold)**.** *A $\mathcal{C}^\infty$-manifold is called a **smooth manifold**.*

Remark 2.1. A given topological manifold can carry different incompatible atlases.

Note that while we only defined compatibility of charts, it should be clear what it means for two atlases of the same type to be compatible.

Definition 2.14 (Compatible / Incompatible Atlases)**.** *Two ❀-atlases $\mathscr{A}$, $\mathscr{B}$ are **compatible** if their union $\mathscr{A} \cup \mathscr{B}$ is again a ❀-atlas, and are **incompatible** otherwise.*

Alternatively, we can define the compatibility of two atlases in terms of the compatibility of any pair of charts, one from each atlas.

Example 2.6. Let $(M, \mathcal{O}) = (\mathbb{R}, \mathcal{O}_{\mathrm{std}})$. Consider the two atlases $\mathscr{A} = \{(\mathbb{R}, \mathrm{id}_\mathbb{R})\}$ and $\mathscr{B} = \{(\mathbb{R}, x)\}$, where $x\colon a \mapsto \sqrt[3]{a}$. Since they both contain a single chart, the compatibility condition on the transition maps is easily seen to hold (in both cases, the only transition map is $\mathrm{id}_\mathbb{R}$). Hence they are both $\mathcal{C}^\infty$-atlases. Consider now $\mathscr{A} \cup \mathscr{B}$. The transition map $\mathrm{id}_\mathbb{R} \circ x^{-1}$ is the map $a \mapsto a^3$, which is smooth. However, the other transition map, $x \circ \mathrm{id}_\mathbb{R}^{-1}$, is the map x, which is not even differentiable once (the first derivative at 0 does not exist). Consequently, $\mathscr{A}$ and $\mathscr{B}$ are not even $\mathcal{C}^1$-compatible.

The previous example shows that we can equip the real line with (at least) two different incompatible $\mathcal{C}^\infty$-structures. This looks like a disaster as it implies that there is an arbitrary choice to be made about which differentiable structure to use. Fortunately, the situation is not as bad as it looks, as we will see in the next sections.

2.3 Differentiable Manifolds

Definition 2.15 (Differentiable Map)**.** *Let $\phi\colon M \to N$ be a map, where $(M, \mathcal{O}_M, \mathscr{A}_M)$ and $(N, \mathcal{O}_N, \mathscr{A}_N)$ are $\mathcal{C}^k$-manifolds. Then ϕ is said to be ($\mathcal{C}^k$-)**differentiable at** $p \in M$ if for some charts $(U, x) \in \mathscr{A}_M$ with $p \in U$ and $(V, y) \in \mathscr{A}_N$ with $\phi(p) \in V$, the map $y \circ \phi \circ x^{-1}$ is k-times continuously differentiable at $x(p) \in x(U) \subseteq \mathbb{R}^{\dim M}$ in the usual sense.*

$$\begin{tikzcd}
U \subseteq M \arrow[rr, "\phi"] \arrow[d, "x"] & & V \subseteq N \arrow[d, "y"] \\
x(U) \subseteq \mathbb{R}^{\dim M} \arrow[rr, "y\circ\phi\circ x^{-1}"] & & y(V) \subseteq \mathbb{R}^{\dim N}
\end{tikzcd}$$

The above diagram shows a typical theme with manifolds. We have a map $\phi\colon M \to N$ and we want to define some property of ϕ at $p \in M$ analogous to some property of maps between subsets of $\mathbb{R}^d$. What we typically do is consider some charts (U, x) and (V, y) as above and define the desired property of ϕ at $p \in U$ in terms of the corresponding property of the downstairs map $y \circ \phi \circ x^{-1}$ at the point $x(p) \in \mathbb{R}^d$. Notice that in the previous definition we only require that *some* charts from the two atlases satisfy the stated property. So we should worry about whether this definition depends on which charts we pick. In fact, this "lifting" of the notion of differentiability from the chart representation of ϕ to the manifold level is well-defined.

Proposition 2.2. *The definition of differentiability is well-defined.*

Proof. We want to show that if $y \circ \phi \circ x^{-1}$ is differentiable at $x(p)$ for some $(U, x) \in \mathscr{A}_M$ with $p \in U$ and $(V, y) \in \mathscr{A}_N$ with $\phi(p) \in V$, then $\widetilde{y} \circ \phi \circ \widetilde{x}^{-1}$ is differentiable at $\widetilde{x}(p)$ for all charts $(\widetilde{U}, \widetilde{x}) \in \mathscr{A}_M$ with $p \in \widetilde{U}$ and $(\widetilde{V}, \widetilde{y}) \in \mathscr{A}_N$ with $\phi(p) \in \widetilde{V}$.

$$
\begin{tikzcd}
\widetilde{x}(U\cap\widetilde{U}) \subseteq \mathbb{R}^{\dim M} \arrow[r, "\widetilde{y}\circ\phi\circ\widetilde{x}^{-1}"] & \widetilde{y}(V\cap\widetilde{V}) \subseteq \mathbb{R}^{\dim N} \\
U\cap\widetilde{U}\subseteq M \arrow[u, "\widetilde{x}"'] \arrow[d, "x"'] \arrow[r, "\phi"] & V\cap\widetilde{V}\subseteq N \arrow[u, "\widetilde{y}"] \arrow[d, "y"] \\
x(U\cap\widetilde{U}) \subseteq \mathbb{R}^{\dim M} \arrow[r, "y\circ\phi\circ x^{-1}"] \arrow[uu, bend left=60, "\widetilde{x}\circ x^{-1}"] & y(V\cap\widetilde{V}) \subseteq \mathbb{R}^{\dim N} \arrow[uu, bend right=60, "\widetilde{y}\circ y^{-1}"']
\end{tikzcd}
$$

Consider the map $\widetilde{x} \circ x^{-1}$ in the diagram above. Since the charts (U, x) and $(\widetilde{U}, \widetilde{x})$ belong to the same $\mathcal{C}^k$-atlas $\mathscr{A}_M$, by definition the transition map $\widetilde{x} \circ x^{-1}$ is $\mathcal{C}^k$-differentiable as a map between subsets of $\mathbb{R}^{\dim M}$, and similarly for $\widetilde{y} \circ y^{-1}$. We now notice that we can write:

$$\widetilde{y} \circ \phi \circ \widetilde{x}^{-1} = (\widetilde{y} \circ y^{-1}) \circ (y \circ \phi \circ x^{-1}) \circ (\widetilde{x} \circ x^{-1})^{-1}$$

and since the composition of $\mathcal{C}^k$ maps is still $\mathcal{C}^k$, we are done. $\square$

This proof shows the significance of restricting to $\mathcal{C}^k$-atlases. Such atlases only contain charts for which the transition maps are $\mathcal{C}^k$, which is what makes our definition of differentiability of maps between manifolds well-defined.

The same definition and proof work for smooth ($\mathcal{C}^\infty$) manifolds, in which case we talk about *smooth maps*. As we said before, this is the case we will be most interested in.

Example 2.7. Consider the smooth manifolds $(\mathbb{R}^d, \mathcal{O}_{\mathrm{std}}, \mathscr{A}_d)$ and $(\mathbb{R}^{d'}, \mathcal{O}_{\mathrm{std}}, \mathscr{A}_{d'})$, where $\mathscr{A}_d)$ and $\mathscr{A}_{d'})$ are the maximal atlases containing the charts $(\mathbb{R}^d, \mathrm{id}_{\mathbb{R}^d})$ and $(\mathbb{R}^{d'}, \mathrm{id}_{\mathbb{R}^{d'}})$ respectively, and let $f\colon \mathbb{R}^d \to \mathbb{R}^{d'}$ be a map. The diagram defining the differentiability of f with respect to these charts is

$$
\begin{tikzcd}
\mathbb{R}^d \arrow[r, "f"] \arrow[d, "\mathrm{id}_{\mathbb{R}^d}"] & \mathbb{R}^{d'} \arrow[d, "\mathrm{id}_{\mathbb{R}^{d'}}"] \\
\mathbb{R}^d \arrow[r, "\mathrm{id}_{\mathbb{R}^{d'}} \circ f \circ (\mathrm{id}_{\mathbb{R}^d})^{-1}"] & \mathbb{R}^{d'}
\end{tikzcd}
$$

and, by definition, the map f is smooth as a map between manifolds if, and only if, the map $\mathrm{id}_{\mathbb{R}^{d'}} \circ f \circ (\mathrm{id}_{\mathbb{R}^d})^{-1} = f$ is smooth in the usual sense.

Example 2.8. Let $(M, \mathcal{O}, \mathscr{A})$ be a d-dimensional smooth manifold and let $(U, x) \in \mathscr{A}$. Then $x\colon U \to x(U) \subseteq \mathbb{R}^d$ is smooth. Indeed, we have

$$
\begin{tikzcd}
U \arrow[r, "x"] \arrow[d, "x"] & x(U) \arrow[d, "\mathrm{id}_{x(U)}"] \\
x(U) \subseteq \mathbb{R}^d \arrow[r, "\mathrm{id}_{x(U)} \circ x \circ x^{-1}"] & x(U) \subseteq \mathbb{R}^d
\end{tikzcd}
$$

Hence $x\colon U \to x(U)$ is smooth if, and only if, the map $\mathrm{id}_{x(U)} \circ x \circ x^{-1} = \mathrm{id}_{x(U)}$ is smooth in the usual sense, which it certainly is.

The coordinate maps $x^i := \mathrm{proj}_i \circ x \colon U \to \mathbb{R}$ are also smooth. Indeed, consider the diagram

$$\begin{tikzcd} U \arrow[rr, "x^i"] \arrow[d, "x"] & & \mathbb{R} \arrow[d, "\mathrm{id}_{\mathbb{R}}"] \\ x(U) \subseteq \mathbb{R}^d \arrow[rr, "\mathrm{id}_{\mathbb{R}} \circ x^i \circ x^{-1}"] & & \mathbb{R} \end{tikzcd}$$

Then, x^i is smooth if, and only if, the map

$$\mathrm{id}_{\mathbb{R}} \circ x^i \circ x^{-1} = x^i \circ x^{-1} = \mathrm{proj}_i$$

is smooth in the usual sense, which it certainly is.

2.3.1 Classification Of Differentiable Structures

Definition 2.16 (Diffeomorphism)**.** *Let $\phi \colon M \to N$ be a bijective map between smooth manifolds. If both ϕ and ϕ^{-1} are smooth, then ϕ is said to be a **diffeomorphism**.*

Diffeomorphisms are the structure preserving maps between smooth manifolds.

Definition 2.17 (Diffeomorphic Manifolds)**.** *Two manifolds $(M, \mathcal{O}_M, \mathscr{A}_M)$, $(N, \mathcal{O}_N, \mathscr{A}_N)$ are said to be **diffeomorphic** if there exists a diffeomorphism $\phi \colon M \to N$ between them. We write $M \cong_{\mathit{diff}} N$.*

Note that if the differentiable structure is understood (or irrelevant), we typically write M instead of the triple $(M, \mathcal{O}_M, \mathscr{A}_M)$.

Remark 2.2. Being diffeomorphic is an equivalence relation. In fact, it is customary to consider diffeomorphic manifolds to be *the same* from the point of view of differential geometry. This is similar to the situation with topological spaces, where we consider homeomorphic spaces to be the same from the point of view of topology. This is typical of all structure preserving maps.

Armed with the notion of diffeomorphism, we can now ask the following question: how many smooth structures on a given topological space are there, up to diffeomorphism?
The answer is quite surprising: it depends on the dimension of the manifold!

Theorem 2.2 (Radon-Moise)**.** *Let M be a manifold with $\dim M = 1, 2,$ or 3. Then there is a unique smooth structure on M up to diffeomorphism.*

Recall that in a previous example, we showed that we can equip $(\mathbb{R}, \mathcal{O}_{\mathrm{std}})$ with two incompatible atlases $\mathscr{A}$ and $\mathscr{B}$. Let $\mathscr{A}_{\max}$ and $\mathscr{B}_{\max}$ be their extensions to maximal atlases, and consider the smooth manifolds $(\mathbb{R}, \mathcal{O}_{\mathrm{std}}, \mathscr{A}_{\max})$ and $(\mathbb{R}, \mathcal{O}_{\mathrm{std}}, \mathscr{B}_{\max})$. Clearly, these are different manifolds, because the atlases are different, but since $\dim \mathbb{R} = 1$, they must be diffeomorphic.

The answer to the case $\dim M > 4$ (we emphasize $\dim M \neq 4$) is provided by *surgery theory*. This is a collection of tools and techniques in topology with which one obtains a new manifold from given ones by performing surgery on them, i.e. by cutting, replacing and gluing parts in such a way as to control topological invariants like the fundamental group. The idea is to understand all manifolds in dimensions higher than 4 by performing surgery systematically. In particular, using surgery theory, it has been shown that there are only finitely many smooth manifolds (up to diffeomorphism) one can make from a topological manifold.

This is not as neat as the previous case, but since there are only finitely many structures, we can still enumerate them, i.e. we can write an exhaustive list.

While finding all the differentiable structures may be difficult for any given manifold, this theorem has an immediate impact on a physical theory that models spacetime as a manifold. For instance, some physicists believe that spacetime should be modelled as a 10-dimensional manifold (we are neither proposing nor condemning this view). If that is indeed the case, we need to worry about which differentiable structure we equip our 10-dimensional manifold with, as each different choice will likely lead to different predictions. But since there are only finitely many such structures, physicists can, at least in principle, devise and perform finitely many experiments to distinguish between them and determine which is the right one, if any.

We now turn to the special case $\dim M = 4$. The result is that if M is a non-compact topological manifold, then there are uncountably many non-diffeomorphic smooth structures that we can equip M with. In particular, this applies to $(\mathbb{R}^4, \mathcal{O}_{\text{std}})$.

2.4 Tangent Spaces

In this section, whenever we say "manifold", we mean a (real) d-dimensional differentiable manifold, unless we explicitly say otherwise. We will also suppress the differentiable structure in the notation.

Definition 2.18 ($\mathcal{C}^\infty(M)$ Vector Space)**.** *Let M be a manifold. We define the infinite-dimensional vector space over $\mathbb{R}$ of all smooth functions on M with underlying set*

$$\mathcal{C}^\infty(M) := \{f \colon M \to \mathbb{R} \mid f \text{ is smooth}\}$$

and operations defined pointwise, i.e. for any $p \in M$,

$$\begin{aligned}(f+g)(p) &:= f(p) + g(p)\\ (\lambda f)(p) &:= \lambda f(p).\end{aligned}$$

A routine check shows that this is indeed a vector space.

Definition 2.19 (Smooth Curve)**.** *A **smooth curve** on M is a smooth map $\gamma \colon \mathbb{R} \to M$, where $\mathbb{R}$ is understood as a 1-dimensional manifold.*

Definition 2.20 (Directional Derivative Operator)**.** *Let $\gamma \colon \mathbb{R} \to M$ be a smooth curve through $p \in M$; w.l.o.g. let $\gamma(0) = p$. The **directional derivative operator** at p along γ is the linear map*

$$\begin{aligned}X_{\gamma,p} \colon \mathcal{C}^\infty(M) &\xrightarrow{\sim} \mathbb{R}\\ f &\mapsto (f \circ \gamma)'(0),\end{aligned}$$

where $\mathbb{R}$ is understood as a 1-dimensional vector space over the field $\mathbb{R}$.

Note that $f \circ \gamma$ is a map $\mathbb{R} \to \mathbb{R}$, hence we can calculate the usual derivative and evaluate it at 0.

Remark 2.3. In differential geometry, $X_{\gamma,p}$ is called the *tangent vector* to the curve γ at the point $p \in M$. Intuitively, $X_{\gamma,p}$ is the velocity γ at p. Consider the curve $\delta(t) := \gamma(2t)$, which is the same curve parametrised twice as fast. We have, for any $f \in \mathcal{C}^\infty(M)$:

$$X_{\delta,p}(f) = (f \circ \delta)'(0) = 2(f \circ \gamma)'(0) = 2X_{\gamma,p}(f)$$

by using the chain rule. Hence $X_{\gamma,p}$ scales like a velocity should.

Definition 2.21 (Tangent Space)**.** *Let M be a manifold and $p \in M$. The **tangent space** to M at p is the vector space over $\mathbb{R}$ with underlying set*

$$T_pM := \{X_{\gamma,p} \mid \gamma \text{ is a smooth curve through } p\},$$

addition

$$\begin{aligned}\oplus \colon T_pM \times T_pM &\to T_pM\\ (X_{\gamma,p}, X_{\delta,p}) &\mapsto X_{\gamma,p} \oplus X_{\delta,p},\end{aligned}$$

and scalar multiplication

$$\begin{aligned}\odot \colon \mathbb{R} \times T_pM &\to T_pM\\ (\lambda, X_{\gamma,p}) &\mapsto \lambda \odot X_{\gamma,p},\end{aligned}$$

both defined pointwise, i.e. for any $f \in \mathcal{C}^\infty(M)$,

$$\begin{aligned}(X_{\gamma,p} \oplus X_{\delta,p})(f) &:= X_{\gamma,p}(f) + X_{\delta,p}(f)\\ (\lambda \odot X_{\gamma,p})(f) &:= \lambda X_{\gamma,p}(f).\end{aligned}$$

Note that the outputs of these operations do not look like elements in T_pM, because they are not of the form $X_{\sigma,p}$ for some curve σ. Hence, we need to show that the above operations are, in fact, well-defined.

Proposition 2.3. *Let $X_{\gamma,p}, X_{\delta,p} \in T_pM$ and $\lambda \in \mathbb{R}$. Then, we have $X_{\gamma,p} \oplus X_{\delta,p} \in T_pM$ and $\lambda \odot X_{\gamma,p} \in T_pM$.*

Since the derivative is a local concept, it is only the behaviour of curves near p that matters. In particular, if two curves γ and δ agree on a neighbourhood of p, then $X_{\gamma,p}$ and $X_{\delta,p}$ are the same element of T_pM. Hence, we can work *locally* by using a chart on M.

Proof. Let (U, x) be a chart on M, with U a neighbourhood of p.

i) Define the curve

$$\sigma(t) := x^{-1}((x \circ \gamma)(t) + (x \circ \delta)(t) - x(p)).$$

Note that σ is smooth since it is constructed via addition and composition of smooth maps and, moreover:

$$\begin{aligned}\sigma(0) &= x^{-1}(x(\gamma(0)) + x(\delta(0)) - x(p))\\ &= x^{-1}(x(p)) + x(p) - x(p))\\ &= x^{-1}(x(p))\\ &= p.\end{aligned}$$

Thus σ is a smooth curve through p. Let $f \in \mathcal{C}^\infty(U)$ be arbitrary. Then we have

$$\begin{aligned}X_{\sigma,p}(f) &:= (f \circ \sigma)'(0)\\ &= [f \circ x^{-1} \circ ((x \circ \gamma) + (x \circ \delta) - x(p))]'(0)\\ &= [\partial_a(f \circ x^{-1})(x(p))]\,((x^a \circ \gamma) + (x^a \circ \delta) - x^a(p))'(0)\\ &= [\partial_a(f \circ x^{-1})(x(p))]\,((x^a \circ \gamma)'(0) + (x^a \circ \delta)'(0))\\ &= (f \circ x^{-1} \circ x \circ \gamma)'(0) + (f \circ x^{-1} \circ x \circ \delta)'(0)\\ &= (f \circ \gamma)'(0) + (f \circ \delta)'(0)\\ &=: (X_{\gamma,p} \oplus X_{\delta,p})(f).\end{aligned}$$

Therefore $X_{\gamma,p} \oplus X_{\delta,p} = X_{\sigma,p} \in T_pM$.

ii) The second part is straightforward. Define $\sigma(t) := \gamma(\lambda t)$. This is again a smooth curve through p and we have:

$$\begin{aligned}X_{\sigma,p}(f) &:= (f \circ \sigma)'(0)\\ &= f'(\sigma(0))\,\sigma'(0)\\ &= \lambda f'(\gamma(0))\,\gamma'(0)\\ &= \lambda(f \circ \gamma)'(0)\\ &:= (\lambda \odot X_{\gamma,p})(f)\end{aligned}$$

for any $f \in \mathcal{C}^\infty(U)$. Hence $\lambda \odot X_{\gamma,p} = X_{\sigma,p} \in T_pM$. □

Hence indeed T_pM is a vector space.

The question is, what exactly $X_{\gamma,p}$ is mathematically speaking? Since it's **a** map of the form:

$$X_{\gamma,p} \colon \mathcal{C}^\infty(M) \xrightarrow{\sim} \mathbb{R}$$

it's clear that it's an element of $\mathrm{Hom}(\mathcal{C}^\infty(M), R)$, i.e an element of the dual vector space of $\mathcal{C}^\infty(M)$. Which subsequently makes T_pM a sub-vector space of the dual vector space of $\mathcal{C}^\infty(M)$. ($X_{\gamma,p}$ is a particular choice of a linear map, more specifically the derivative with respect to the parameter, and not **all** possible linear maps. This is why T_pM is not the whole dual vector space of $\mathcal{C}^\infty(M)$)

However, if we take the extra step and turn the $\mathcal{C}^\infty(M)$ from a vector space to an algebra (by defining an appropriate operation) then we can show that $X_{\gamma,p}$ is actually a derivation of the algebra.

More specifically we will define a product on $\mathcal{C}^\infty(M)$ by

$$\begin{aligned} \bullet \colon \mathcal{C}^\infty(M) \times \mathcal{C}^\infty(M) &\to \mathcal{C}^\infty(M) \\ (f,g) &\mapsto f \bullet g, \end{aligned}$$

where $f \bullet g$ is defined pointwise. Then $(\mathcal{C}^\infty(M), +, \cdot, \bullet)$ is an associative, unital and commutative algebra over $\mathbb{R}$.

Now that we have an algebra, let us remind ourselves what a derivation is and also try to combine the definition with our case:

Definition 2.22 (Derivation (On A Manifold))**.** *Let M be a manifold and let $p \in U \subseteq M$, where U is open. A* derivation on U at p *is an $\mathbb{R}$-linear map $D \colon \mathcal{C}^\infty(U) \xrightarrow{\sim} \mathbb{R}$ satisfying the Leibniz rule*

$$D(fg) = D(f)g(p) + f(p)D(g).$$

The usual derivative operator is a derivation on $\mathcal{C}^\infty(\mathbb{R})$, the algebra of smooth real functions, since it is linear and satisfies the Leibniz rule. (The second derivative operator, however, is not a derivation on $\mathcal{C}^\infty(\mathbb{R})$, since it does not satisfy the Leibniz rule. This shows that the composition of derivations need not be a derivation.) Hence, we managed to show that indeed $X_{\gamma,p}$ is actually a derivation of the algebra of smooth real functions on M.

2.4.1 Co-Ordinate Induced Basis For The Tangent Space

The following is a crucially important result about tangent spaces.

Theorem 2.3. *Let M be a manifold and let $p \in M$. Then*

$$\dim T_pM = \dim M.$$

Remark 2.4. Note carefully that, despite us using the same symbol, the two "dimensions" appearing in the statement of the theorem are, at least on the surface, entirely unrelated. Indeed, recall that $\dim M$ is defined in terms of charts (U,x), with $x \colon U \to x(U) \subseteq \mathbb{R}^{\dim M}$, while $\dim T_pM = |\mathcal{B}|$, where $\mathcal{B}$ is a Hamel basis for the vector space T_pM. The idea behind the proof is to construct a basis of T_pM from a chart on M.

Proof. W.l.o.g., let (U,x) be a chart *centred* at p, i.e. $x(p) = 0 \in \mathbb{R}^{\dim M}$. Define $(\dim M)$-many curves $\gamma_{(a)} \colon \mathbb{R} \to U$ through p by requiring $(x^b \circ \gamma_{(a)})(t) = \delta^b_a t$, i.e.

$$\begin{aligned} \gamma_{(a)}(0) &:= p \\ \gamma_{(a)}(t) &:= x^{-1} \circ (0, \ldots, 0, t, 0, \ldots, 0) \end{aligned}$$

where the t is in the a^{th} position, with $1 \leq a \leq \dim M$. Let us calculate the action of the tangent vector $X_{\gamma_{(a)},p} \in T_pM$ on an arbitrary function $f \in \mathcal{C}^\infty(U)$:

$$\begin{aligned} X_{\gamma_{(a)},p}(f) &:= (f \circ \gamma_{(a)})'(0) \\ &= (f \circ \mathrm{id}_U \circ \gamma_{(a)})'(0) \\ &= (f \circ x^{-1} \circ x \circ \gamma_{(a)})'(0) \\ &= [\partial_b(f \circ x^{-1})(x(p))]\, (x^b \circ \gamma_{(a)})'(0) \\ &= [\partial_b(f \circ x^{-1})(x(p))]\, (\delta^b_a t)'(0) \\ &= [\partial_b(f \circ x^{-1})(x(p))]\, \delta^b_a \\ &= \partial_a(f \circ x^{-1})(x(p)) \end{aligned}$$

We introduce a special notation for this last line, namely:

$$\partial_a(f \circ x^{-1})(x(p)) := \left(\frac{\partial}{\partial x^a}\right)_p (f)$$

Remark 2.5. While the symbol $\left(\frac{\partial}{\partial x^a}\right)_p$ has nothing to do with the idea of partial differentiation with respect to the variable x^a (since x refers to the chart map and no differentiation has been defined there), it is notationally consistent with it, in the following sense.

Let $M = \mathbb{R}^d$, $(U, x) = (\mathbb{R}^d, \mathrm{id}_{\mathbb{R}^d})$ and let $\left(\frac{\partial}{\partial x^a}\right)_p \in T_p\mathbb{R}^d$. If $f \in \mathcal{C}^\infty(\mathbb{R}^d)$, then

$$\left(\frac{\partial}{\partial x^a}\right)_p (f) = \partial_a(f \circ x^{-1})(x(p)) = \partial_a f(p),$$

since $x = x^{-1} = \mathrm{id}_{\mathbb{R}^d}$. Moreover, we have $\mathrm{proj}_a = x^a$. Thus, we can think of $x^1, \ldots, x^d$ as the independent variables of f, and we can then write

$$\left(\frac{\partial}{\partial x^a}\right)_p (f) = \frac{\partial f}{\partial x^a}(p).$$

Hence, up to this point we showed that:

$$X_{\gamma_{(a)},p}(f) = \left(\frac{\partial}{\partial x^a}\right)_p (f)$$

Or by removing the action on the function, simply:

$$X_{\gamma_{(a)},p} = \left(\frac{\partial}{\partial x^a}\right)_p$$

We now claim that

$$\mathcal{B} = \left\{ \left(\frac{\partial}{\partial x^a}\right)_p \in T_pM \;\middle|\; 1 \leq a \leq \dim M \right\}$$

is a basis of T_pM. First, we show that $\mathcal{B}$ spans T_pM.

Let $X_{\gamma,p} \in T_pM$. For any $f \in \mathcal{C}^\infty(U)$, we have

$$\begin{aligned} X_{\gamma,p}(f) &:= (f \circ \sigma)'(0) \\ &= (f \circ x^{-1} \circ x \circ \gamma)'(0) \\ &= \left[\partial_b(f \circ x^{-1})(x(p))\right](x^b \circ \gamma)'(0) \\ &= (x^b \circ \gamma)'(0) \left(\frac{\partial}{\partial x^b}\right)_p (f). \end{aligned}$$

Since $(x^b \circ \gamma)'(0) =: X^b \in \mathbb{R}$, we have:

$$X_{\gamma,p} = X^b \left(\frac{\partial}{\partial x^b}\right)_p,$$

i.e. any $X_{\gamma,p} \in T_pM$ is a linear combination of elements from $\mathcal{B}$.

To show linear independence, suppose that

$$\lambda^a \left(\frac{\partial}{\partial x^a}\right)_p = 0,$$

for some scalars λ^a. Note that this is an operator equation, and the zero on the right hand side is the zero operator $0 \in T_pM$.

Recall that, given the chart (U, x), the coordinate maps $x^b \colon U \to \mathbb{R}$ are smooth, i.e. $x^b \in \mathcal{C}^\infty(U)$. Thus, we can feed them into the left hand side to obtain

$$\begin{aligned} 0 &= \lambda^a \left(\frac{\partial}{\partial x^a}\right)_p (x^b) \\ &= \lambda^a\, \partial_a (x^b \circ x^{-1})(x(p)) \\ &= \lambda^a\, \partial_a (\mathrm{proj}_b)(x(p)) \\ &= \lambda^a \delta^b_a \\ &= \lambda^b \end{aligned}$$

i.e. $\lambda^b = 0$ for all $1 \leq b \leq \dim M$. So $\mathcal{B}$ is indeed a basis of T_pM, and since by construction $|\mathcal{B}| = \dim M$, the proof is complete. $\square$

Remark 2.6. While it is possible to define infinite-dimensional manifolds, in this course we will only consider finite-dimensional ones. Hence $\dim T_pM = \dim M$ will always be finite in this course.

Remark 2.7. Note that the basis that we have constructed in the proof is *not* chart-independent. Indeed, each different chart will induce a different tangent space basis, and we distinguish between them by keeping the chart map in the notation for the basis elements.

This is not a cause of concern for our proof however, since every basis of a vector space must have the same cardinality, and hence it suffices to find one basis to determine the dimension.

Definition 2.23 (Co-Ordinate Induced Basis). *Let $X_{\gamma,p} \in T_pM$ be a tangent vector and let (U, x) be a chart containing p. Then the basis $\{\left(\frac{\partial}{\partial x^a}\right)_p\}$ created by the usage of the chart is called a **co-ordinate induced basis**. In this basis an element $X_{\gamma,p} \in T_pM$ can be expressed as:*

$$X_{\gamma,p} = X^a \left(\frac{\partial}{\partial x^a}\right)_p,$$

*where the real numbers $X^1, \ldots, X^{\dim M}$ are called the **vector components** of $X_{\gamma,p}$ with respect to the co-ordinate induced basis by the chart (U, x).*

2.4.2 Change Of Vector Components Under A Change Of Chart

One of the most heavily used concepts is the transformation of the components of a vector under different co-ordinate systems (i.e under a chart transition map that subsequently changes the co-ordinate induced basis). Let's find out the rule.

Let $X_{\gamma,p} \in T_pM$ and let (U, x) and (V, y) be two charts containing p. Then $X_{\gamma,p}$ can be expressed in any of the two charts as:

$$X^a{}_{(y)} \left(\frac{\partial}{\partial y^a}\right)_p = X_{\gamma,p} = X^a{}_{(x)} \left(\frac{\partial}{\partial x^a}\right)_p$$

Let us act with $X_{\gamma,p}$ on some smooth function f of $\mathcal{C}^\infty(M)$ by using first the components of (U, x) chart:

$$\begin{aligned} X_{\gamma,p}(f) &= X^a{}_{(x)} \left(\frac{\partial}{\partial x^a}\right)_p (f) \\ &= X^a{}_{(x)} \partial_a (f \circ x^{-1})(x(p)) \\ &= X^a{}_{(x)} \partial_a (f \circ y^{-1} \circ y \circ x^{-1})(x(p)) \\ &= X^a{}_{(x)} \partial_a (y^b \circ x^{-1})(x(p))\, \partial_b (f \circ y^{-1})(y(p)) \\ &= X^a{}_{(x)} \frac{\partial y^b}{\partial x^a} \left(\frac{\partial}{\partial y^b}\right)_p (f) \end{aligned}$$

Similarly, let us now act with $X_{\gamma,p}$ on the smooth function f of $\mathcal{C}^\infty(M)$ by using the components of

(V, y) chart:

$$X_{\gamma,p}(f) = X^a{}_{(y)} \left(\frac{\partial}{\partial y^a}\right)_p (f)$$

These expressions are, of course, equal to each other so by suppressing now the action on the function f, we obtain:

$$X^a{}_{(x)} \frac{\partial y^b}{\partial x^a} \left(\frac{\partial}{\partial y^b}\right)_p = X^b{}_{(y)} \left(\frac{\partial}{\partial y^b}\right)_p$$
$$X^a{}_{(x)} \frac{\partial y^b}{\partial x^a} \left(\frac{\partial}{\partial y^b}\right)_p - X^b{}_{(y)} \left(\frac{\partial}{\partial y^b}\right)_p = 0$$
$$\left(X^a{}_{(x)} \frac{\partial y^b}{\partial x^a} - X^b{}_{(y)}\right) \left(\frac{\partial}{\partial y^b}\right)_p = 0$$

Finally, since the base vectors of $\{\left(\frac{\partial}{\partial y^a}\right)_p\}$ are linearly independent the only way for this equation to be zero is for the coefficients to be zero hence:

$$X^a{}_{(x)} \frac{\partial y^b}{\partial x^a} - X^b{}_{(y)} = 0$$

Of finally by solving w.r.t $X^b{}_{(y)}$ and renaming the indices:

$$X^a{}_{(y)} = \frac{\partial y^a}{\partial x^b} X^b{}_{(x)}$$

This equation shows as how the components of a vector transform under a chart transition map, i.e under the change of charts, i.e from one co-ordinate induced basis to another. Of course the formula agrees completely with the transformations of vector components under the change of basis that we showed in previous chapter: $\widetilde{v}^b = A^b{}_a v^a$.

The function $y^a = y^a(x^1, \ldots, x^{\dim M})$ expresses the new co-ordinates in terms of the old ones, and $A^b{}_a$ is the *Jacobian* matrix of this map, evaluated at $x(p)$. Note that no matter how non-linear the transformations of the co-ordinates are, the vectors always transform in a linear fashion. In a way, "vectors do not care about the non-linearity of co-ordinate transformations".

2.5 Cotangent Spaces

Since the tangent space is a vector space, we can do all the constructions we saw previously in the abstract vector space setting.

Definition 2.24 (Cotagent Space)**.** *Let M be a manifold and $p \in M$. The* ***cotangent space*** *to M at p is*

$$T_p^* M := (T_p M)^*$$

Since $\dim T_p M$ is finite, we have $T_p M \cong_{\text{vec}} T_p^* M$.

And of course, once we have the cotangent space, we can define the tensor spaces.

Definition 2.25 (Tensor Space)**.** *Let M be a manifold and $p \in M$. The* ***tensor space*** *$(T_s^r)_p M$ is defined as*

$$(T_s^r)_p M := T_s^r(T_p M) = \underbrace{T_p M \otimes \cdots \otimes T_p M}_{r \text{ copies}} \otimes \underbrace{T_p^* M \otimes \cdots \otimes T_p^* M}_{s \text{ copies}}.$$

2.5.1 Dual Basis For The Cotangent Space

Now let's give a very important definition that will help us to formalize elements, and subsequently a basis, for the cotangent space.

Definition 2.26 (Gradient). *Let M be a manifold and let $f\colon M \to \mathbb{R}$ be smooth. The* ***gradient of*** f ***at*** $p \in M$ *is the* $\mathbb{R}$*-linear map*

$$\begin{aligned} \mathrm{d}_p\colon \mathcal{C}^\infty(M) &\xrightarrow{\sim} T_p^*M \\ f &\mapsto \mathrm{d}_p f, \end{aligned}$$

with $p \in U \subseteq M$*, defined by*

$$\mathrm{d}_p f(X_{\gamma,p}) := X_{\gamma,p}(f)$$

Remark 2.8. Note that since d_p is a map from $\mathcal{C}^\infty(M) \xrightarrow{\sim} T_p^*M$ that means that when it acts on a function of $\mathcal{C}^\infty(M)$ the final result $\mathrm{d}_p f$ is an element of T_p^*M hence a covector. By its turn, as an element of the dual space of T_pM it maps elements of T_pM to the real numbers (that's the definition of the dual space of a vector space). Hence the expression $\mathrm{d}_p f(X)$ must end up to a real number, which is indeed what $X_{\gamma,p}(f)$ is. By writing $\mathrm{d}_p f(X) := X_{\gamma,p}(f)$, we have committed a slight (but nonetheless real) abuse of notation, since $\mathrm{d}_p f(X) \in T_{f(p)}\mathbb{R}$ takes in a function and return a real number, but $X_{\gamma,p}(f)$ is already a real number! However by doing so we can now talk about $\mathrm{d}_p f$ without providing the vector that it acts on. In other words we can talk about covectors without the need of their actions on vectors.

Remark 2.9. The gradient of a function is a covector and **not** a vector.

Recall that if (U,x) is a chart on M, then the co-ordinate maps $x^a\colon U \to x(U) \subseteq \mathbb{R}^{\dim M}$ are smooth functions on U hence they belong to $\mathcal{C}^\infty(M)$. We can thus apply the gradient operator d_p (with $p \in U$) to each of them to obtain $(\dim M)$-many elements of T_p^*M.

Proposition 2.4. *Let (U,x) be a chart on M, with $p \in U$. The set $\mathcal{B} = \{\mathrm{d}_p x^a \mid 1 \leq a \leq \dim M\}$ forms the dual basis of T_p^*M.*

Proof. By simply acting on $\left(\frac{\partial}{\partial x^a}\right)_p$ with $\mathrm{d}_p x^a$ (in our notation, we have $(\mathrm{d}x^a)_p = \mathrm{d}_p x^a$) we obtain:

$$\begin{aligned} \mathrm{d}_p x^a \left(\left(\frac{\partial}{\partial x^b} \right)_p \right) &= \left(\frac{\partial}{\partial x^b} \right)_p (x^a) && \text{(definition of } \mathrm{d}_p x^a\text{)} \\ &= \partial_b (x^a \circ x^{-1})(x(p)) && \text{(definition of } \left(\tfrac{\partial}{\partial x^b}\right)_p\text{)} \\ &= \partial_b(\mathrm{proj}_a)(x(p)) \\ &= \delta^a_b \end{aligned}$$

Therefore, $\mathcal{B}$ is, in fact, the dual basis to $\{\left(\frac{\partial}{\partial x^a}\right)_p\}$. □

2.5.2 Change Of Covector Components Under A Change Of Chart

Once again, as we did in the vector case with the vector components, one needs to find the transformation of the components of a covector under different co-ordinate systems. We will follow exactly the same procedure.

Let $\omega_p \in T_p^*M$ and let (U,x) and (V,y) be two charts containing p. Then ω_p can be expressed in any of the two charts by using the dual basis as:

$$\omega_{(y)a}(dy^a)_p = \omega_p = \omega_{(x)a}(dx^a)_p$$

By repeating the same process as we did for the vectors it is very easy to show that covectors components transform as

$$\omega_{(y)a} = \left(\frac{\partial x^b}{\partial y^a} \right)_p \omega_{(x)b}$$

2.6 Push-Forward And Pull-Back

Definition 2.27 (Push-Forward)**.** *Let $\phi\colon M \to N$ be a smooth map between smooth manifolds. The* ***push-forward*** *(or* ***derivative****) of ϕ at $p \in M$ is the linear map $(\phi_*)_p$:*

$$\begin{aligned}(\phi_*)_p\colon T_pM &\xrightarrow{\sim} T_{\phi(p)}N\\ X_{\gamma,p} &\mapsto (\phi_*)_p(X_{\gamma,p})\end{aligned}$$

where $(\phi_*)_p(X_{\gamma,p})$ is defined as

$$\begin{aligned}(\phi_*)_p(X_{\gamma,p})\colon \mathcal{C}^\infty(N) &\xrightarrow{\sim} \mathbb{R}\\ f &\mapsto (\phi_*)_p(X_{\gamma,p})f := X_{\gamma,p}(f\circ\phi).\end{aligned}$$

In other words, since $(\phi_*)_p$ is a map from one tangent space to another this means that it acts on a tangent vector and produces another one, hence $(\phi_*)_p(X_{\gamma,p})$ is again a tangent vector (but on N). As a tangent vector it can act on a smooth function (again on N) and produce a real number, hence the action of a push-forward on a function is simply the one we wrote above.

Note that one has to define a push-forward $(\phi_*)_p$ for every point p of M. Although we have only one map ϕ we have many push-forward maps $(\phi_*)_p$.

Proposition 2.5. *Let $\phi\colon M \to N$ be smooth. The tangent vector $X_{\gamma,p} \in T_pM$ is pushed forward to the tangent vector $X_{\phi\circ\gamma,\phi(p)} \in T_{\phi(p)}N$, i.e.*

$$(\phi_*)_p(X_{\gamma,p}) = X_{\phi\circ\gamma,\phi(p)}.$$

Proof. Let $f \in \mathcal{C}^\infty(V)$, with (V,x) a chart on N and $\phi(p) \in V$. By applying the definitions, we have

$$\begin{aligned}(\phi_*)_p(X_{\gamma,p})(f) &= (X_{\gamma,p})(f\circ\phi) && \text{(definition of } (\phi_*)_p)\\ &= ((f\circ\phi)\circ\gamma)'(0) && \text{(definition of } X_{\gamma,p})\\ &= (f\circ(\phi\circ\gamma))'(0) && \text{(associativity of } \circ)\\ &= X_{\phi\circ\gamma,\phi(p)}(f) && \text{(definition of } X_{\phi\circ\gamma,\phi(p)})\end{aligned}$$

Since f was arbitrary, we have $(\phi_*)_p(X_{\gamma,p}) = X_{\phi\circ\gamma,\phi(p)}$. □

Related to the push-forward, there is the notion of pull-back of a smooth map.

Definition 2.28 (Pull-Back)**.** *Let $\phi\colon M \to N$ be a smooth map between smooth manifolds. The* ***pull-back*** *of ϕ at $p \in M$ is the linear map:*

$$\begin{aligned}(\phi^*)_p\colon T^*_{\phi(p)}N &\xrightarrow{\sim} T^*_pM\\ \omega_{\phi(p)} &\mapsto (\phi^*)_p(\omega_{\phi(p)}),\end{aligned}$$

where $(\phi^*)_p(\omega_{\phi(p)})$ is defined as

$$\begin{aligned}(\phi^*)_p(\omega_{\phi(p)})\colon T_pM &\xrightarrow{\sim} \mathbb{R}\\ X_{\gamma,p} &\mapsto (\phi^*)_p(\omega_{\phi(p)})(X_{\gamma,p}) := \omega_{\phi(p)}((\phi_*)_p(X_{\gamma,p})).\end{aligned}$$

In words, if $\omega_{\phi(p)}$ is a covector on N, its pull-back $(\phi^*)_p(\omega_{\phi(p)})$ is a covector on M. It acts on tangent vectors on M by first pushing them forward to tangent vectors on N, and then applying $\omega_{\phi(p)}$ to them to produce a real number.

Diagrammatically, what we've defined so far is the following

$$\begin{tikzcd} \mathcal{C}^\infty(M) \arrow[d, "X_{\gamma,p}"'] & \mathcal{C}^\infty(N) \arrow[l, "-\circ\phi"'] \arrow[dl, "(\phi_*)_p(X_{\gamma,p})"] \\ \mathbb{R} & \end{tikzcd} \qquad \begin{tikzcd} T_pM \arrow[r, "(\phi_*)_p"] \arrow[dr, "(\phi^*)_p(\omega_{\phi(p)})"'] & T_{\phi(p)}N \arrow[d, "\omega_{\phi(p)}"] \\ & \mathbb{R} \end{tikzcd}$$

Remark 2.10. It is quite easy to show that everything we have defined in this section is, in fact, linear.

Remark 2.11. We have seen that, given a smooth $\phi\colon M \to N$, we can push a vector $X_{\gamma,p} \in T_pM$ forward to a vector $(\phi_*)_p(X_{\gamma,p}) \in T_{\phi(p)}N$, and pull a covector $\omega_{\phi(p)} \in T^*_{\phi(p)}N$ back to a covector $(\phi^*)_p(\omega_{\phi(p)}) \in T^*_pM$. In other words both push-forward and pull-back work only in the direction of their definition. However, if $\phi\colon M \to N$ is a diffeomorphism (and only then), we can also pull a vector $Y_{\gamma,\phi(p)} \in T_{\phi(p)}N$ back to a vector $(\phi^*)_p(Y_{\gamma,\phi(p)}) \in T_pM$, and push a covector $\eta_p \in T^*_pM$ forward to a covector $(\phi_*)_p(\eta_p) \in T^*_{\phi(p)}N$, by using ϕ^{-1} as follows:

$$(\phi^*)_p(Y_{\gamma,\phi(p)}) := ((\phi^{-1})_*)_{\phi(p)}(Y_{\gamma,\phi(p)})$$
$$(\phi_*)_p(\eta_p) := ((\phi^{-1})^*)_{\phi(p)}(\eta_p).$$

In general, we should keep in mind that:

Vectors are pushed forward,
covectors are pulled back.

2.7 Immersions And Embeddings

We will now consider the question of under which circumstances a smooth manifold can "sit" in $\mathbb{R}^d$, for some $d \in \mathbb{N}$. There are, in fact, two notions of sitting inside another manifold, called immersion and embedding.

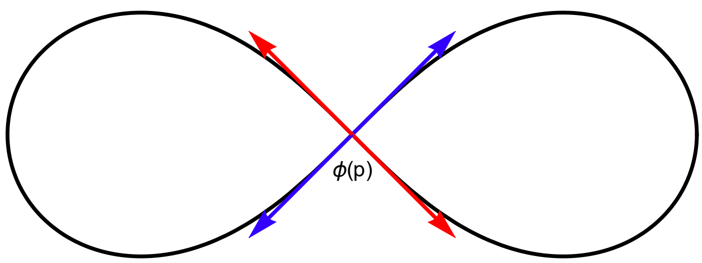
Definition 2.29 (Immersion)**.** *A smooth map $\phi\colon M \to N$ is said to be an* ***immersion*** *of M into N if the derivative*

$$(\phi_*)_p\colon T_pM \xrightarrow{\sim} T_{\phi(p)}N$$

is injective, for all $p \in M$. In that case, the manifold M is said to be an immersed submanifold *of N.*

From the theory of linear algebra, we immediately deduce that, for $\phi\colon M \to N$ to be an immersion, we must have $\dim M \leq \dim N$. A closely related notion is that of a *submersion*, where we require each $(\phi_*)_p$ to be surjective, and thus we must have $\dim M \geq \dim N$. However, we will not need this here.

Example 2.9. Consider the map $\phi\colon S^1 \to \mathbb{R}^2$ whose image is reproduced below.



The map ϕ is not injective, i.e. there are $p, q \in S^1$, with $p \neq q$ and $\phi(p) = \phi(q)$. Of course, this means that $T_{\phi(p)}\mathbb{R}^2 = T_{\phi(q)}\mathbb{R}^2$. However, the maps $(\phi_*)_p$ and $(\phi_*)_q$ are both injective, with their images being represented by the blue and red arrows, respectively. Hence, the map ϕ is immersion.

Definition 2.30 (Embedding)**.** *A smooth map $\phi\colon M \to N$ is said to be a* ***embedding*** *of M into N if*

- *$\phi\colon M \to N$ is an immersion;*
- *$M \cong_{\text{top}} \phi(M) \subseteq N$, where $\phi(M)$ carries the subset topology inherited from N.*

In that case the manifold M is said to be an embedded submanifold *of N.*

Remark 2.12. If a continuous map between topological spaces satisfies the second condition above, then it is called a *topological embedding*. Therefore, an embedding is a topological embedding which is also an immersion (as opposed to simply being a topological embedding).

In the early days of differential geometry there were two approaches to study manifolds. One was the extrinsic view, within which manifolds are defined as special subsets of $\mathbb{R}^d$, and the other was the intrinsic view, which is the view that we have adopted here.

Whitney's theorem, which we will state without proof, states that these two approaches are, in fact, equivalent.

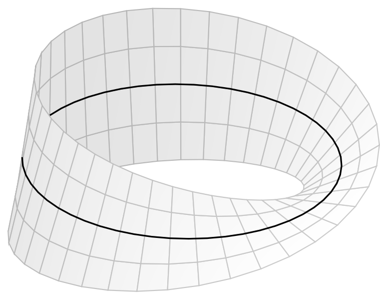
Theorem 2.4 (Whitney)**.** *Any smooth manifold M can be*

- *embedded in $\mathbb{R}^{2\dim M}$;*
- *immersed in $\mathbb{R}^{2\dim M-1}$.*

Example 2.10. The Klein bottle can be embedded in $\mathbb{R}^4$ but not in $\mathbb{R}^3$. It can, however, be immersed in $\mathbb{R}^3$.

2.8 Topological Bundles

While topological products are very useful, very often one intuitively thinks of the product of two manifolds as attaching a copy of the second manifold to each point of the first. However, not all interesting manifolds can be understood as products of manifolds. A classic example of this is the *Möbius strip*.



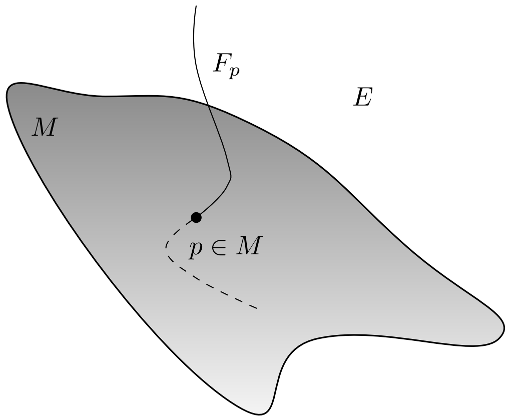
It looks locally like the finite cylinder $S^1 \times [0,1]$, which we can picture as the circle S^1 (the thicker line in figure) with the finite interval $[0,1]$ attached to each of its points in a "smooth" way. The Möbius strip has a "twist", which makes it globally different from the cylinder.

Definition 2.31 (Topological Bundles)**.** *A topological **bundle** (of topological manifolds) is a triple (E,π,M) where E and M are topological manifolds called the* total space *and the* base space *respectively, and π is a continuous, surjective map $\pi\colon E \to M$ called the* projection map*.*

We will often denote the bundle (E,π,M) by $E \xrightarrow{\pi} M$.

Definition 2.32 (Fiber)**.** *Let $E \xrightarrow{\pi} M$ be a bundle and let $p \in M$. Then, $F_p := \operatorname{preim}_\pi(\{p\})$ is called the **fiber** at the point p.*

Intuitively, the fiber at the point $p \in M$ is a set of points in E (represented below as a line) attached to the point p. The projection map sends all the points in the fiber F_p to the point p.



Example 2.11. A trivial example of a bundle is the *product bundle*. Let M and N be manifolds. Then, the triple $(M \times N, \pi, M)$, where:

$$\begin{aligned} \pi \colon M \times N &\to M \\ (p, q) &\mapsto p \end{aligned}$$

is a bundle since (one can easily check) π is a continuous and surjective map. Similarly, $(M \times N, \pi, N)$ with the appropriate π, is also a bundle.

Example 2.12. In a bundle, different points of the base manifold may have (topologically) different fibers. For example, consider the bundle $E \xrightarrow{\pi} \mathbb{R}$ where:

$$F_p := \mathrm{preim}_\pi(\{p\}) \cong_{\text{top}} \begin{cases} S^1 & \text{if } p < 0 \\ \{p\} & \text{if } p = 0 \\ [0, 1] & \text{if } p > 0 \end{cases}$$

Definition 2.33 (Fiber Bundle)**.** *Let $E \xrightarrow{\pi} M$ be a bundle and let F be a manifold. Then, $E \xrightarrow{\pi} M$ is called a **fiber bundle**, with (typical) fiber F, if:*

$$\forall\, p \in M : \mathrm{preim}_\pi(\{p\}) \cong_{\text{top}} F.$$

A fiber bundle is often represented diagrammatically as:

$$\begin{tikzcd} F \arrow[r] & E \arrow[d, "\pi"] \\ & M \end{tikzcd}$$

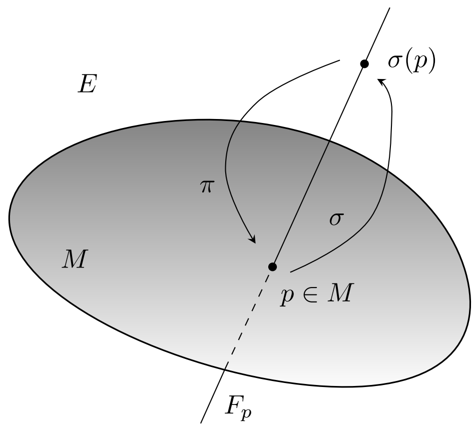
Example 2.13. The bundle $M \times N \xrightarrow{\pi} M$ is a fiber bundle with fiber $F := N$.

Example 2.14. The Möbius strip is a fiber bundle $E \xrightarrow{\pi} S^1$, with fiber $F := [0, 1]$, where $E \neq S^1 \times [0, 1]$, i.e. the Möbius strip is not a product bundle.

Example 2.15. A $\mathbb{C}$-line bundle over M is the fiber bundle (E, π, M) with fiber $\mathbb{C}$. Note that the product bundle $(M \times \mathbb{C}, \pi, M)$ is a $\mathbb{C}$-line bundle over M, but a $\mathbb{C}$-line bundle over M need not be a product bundle.

Definition 2.34 (Section)**.** *Let $E \xrightarrow{\pi} M$ be a bundle. A map $\sigma \colon M \to E$ is called a **section** of the bundle if $\pi \circ \sigma = \mathrm{id}_M$.*

Intuitively, a section is a map σ which sends each point $p \in M$ to *some* point $\sigma(p)$ in its fiber F_p, so that the projection map π takes $\sigma(p) \in F_p \subseteq E$ back to the point $p \in M$.



Example 2.16. Let $(M \times F, \pi, M)$ be a product bundle. Then, a section of this bundle is a map:

$$\begin{aligned} \sigma\colon M &\to M \times F \\ p &\mapsto (p, s(p)) \end{aligned}$$

where $s\colon M \to F$ is any map.

Definition 2.35 (Sub-Bundle)**.** *A* ***sub-bundle*** *of a bundle (E, π, M) is a triple (E', π', M') where $E' \subseteq E$ and $M' \subseteq M$ are submanifolds and $\pi' := \pi|_{E'}$.*

Definition 2.36 (Restricted Bundle)**.** *Let (E, π, M) be a bundle and let $N \subseteq M$ be a submanifold. The* ***restricted bundle*** *(to N) is the triple (E, π', N) where:*

$$\pi' := \pi|_{\mathrm{preim}_\pi(N)}$$

Definition 2.37 (Bundle Morphism)**.** *Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be bundles and let $u\colon E \to E'$ and $v\colon M \to M'$ be maps. Then (u, v) is called a* bundle morphism *if the following diagram commutes:*

$$\begin{tikzcd} E \arrow[r, "u"] \arrow[d, "\pi"] & E' \arrow[d, "\pi'"] \\ M \arrow[r, "v"] & M' \end{tikzcd}$$

i.e. if $\pi' \circ u = v \circ \pi$.

If (u, v) and (u, v') are both bundle morphisms, then $v = v'$. That is, given u, if there exists v such that (u, v) is a bundle morphism, then v is unique.

Definition 2.38 (Isomorphic Bundles)**.** *Two bundles $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ are said to be* ***isomorphic (as bundles)*** *if there exist bundle morphisms (u, v) and (u^{-1}, v^{-1}) satisfying:*

$$\begin{tikzcd} E \arrow[r, shift left, "u"] \arrow[d, "\pi"'] & E' \arrow[l, shift left, "u^{-1}"] \arrow[d, "\pi'"] \\ M \arrow[r, shift left, "v"] & M' \arrow[l, shift left, "v^{-1}"] \end{tikzcd}$$

Such a (u, v) is called a bundle isomorphism *and we write $E \xrightarrow{\pi} M \cong_{\mathrm{bdl}} E' \xrightarrow{\pi'} M'$.*

Bundle isomorphisms are the structure-preserving maps for bundles.

Definition 2.39 (Locally Isomorphic Bundles)**.** *A bundle $E \xrightarrow{\pi} M$ is said to be **locally isomorphic (as a bundle)** to a bundle $E' \xrightarrow{\pi'} M'$ if for all $p \in M$ there exists a neighbourhood $U(p)$ such that the restricted bundle:*

$$\mathrm{preim}_\pi(U(p)) \xrightarrow{\pi|_{\mathrm{preim}_\pi(U(p))}} U(p)$$

is isomorphic to the bundle $E' \xrightarrow{\pi'} M'$.

Definition 2.40 (Trivial / Locally Trivial Bundle)**.** *A bundle $E \xrightarrow{\pi} M$ is said to be:*

*i) **trivial** if it is isomorphic to a product bundle;*

*ii) **locally trivial** if it is locally isomorphic to a product bundle.*

Example 2.17. The cylinder C is trivial as a bundle, and hence also locally trivial.

Example 2.18. The Möbious strip is not trivial but it is locally trivial.

From now on, we will mostly consider locally trivial bundles.

Remark 2.13. In quantum mechanics, what is usually called a "wave function" is not a function at all, but rather a section of a $\mathbb{C}$-line bundle over physical space. However, if we assume that the $\mathbb{C}$-line bundle under consideration is locally trivial, then each section of the bundle can be represented (locally) by a map from the base space to the total space and hence it is appropriate to use the term "wave *function*".

Definition 2.41 (Pull-Back Bundle)**.** *Let $E \xrightarrow{\pi} M$ be a bundle and let $f\colon M' \to M$ be a map from some manifold M'. The **pull-back bundle of** $E \xrightarrow{\pi} M$ induced by f is defined as $E' \xrightarrow{\pi'} M'$, where:*

$$E' := \{(m', e) \in M' \times E \mid f(m') = \pi(e)\}$$

and $\pi'(m', e) := m'$.

If $E' \xrightarrow{\pi'} M'$ is the pull-back bundle of $E \xrightarrow{\pi} M$ induced by f, then one can easily construct a bundle morphism by defining:

$$\begin{aligned} u\colon \quad E' &\to E \\ (m', e) &\mapsto e \end{aligned}$$

This corresponds to the diagram:

$$\begin{tikzcd} E' \arrow[r, "u"] \arrow[d, "\pi'"] & E \arrow[d, "\pi"] \\ M' \arrow[r, "f"] & M \end{tikzcd}$$

Remark 2.14. Sections on a bundle pull back to the pull-back bundle. Indeed, let $E' \xrightarrow{\pi'} M'$ be the pull-back bundle of $E \xrightarrow{\pi} M$ induced by f.

$$\begin{tikzcd} E' \arrow[d, shift left, "\pi'"] & & E \arrow[d, shift left, "\pi"] \\ M' \arrow[u, shift left, "\sigma'"] \arrow[rru, "\sigma \circ f"] \arrow[rr, "f"] & & M \arrow[u, shift left, "\sigma"] \end{tikzcd}$$

If σ is a section of $E \xrightarrow{\pi} M$, then $\sigma \circ f$ determines a map from M' to E which sends each $m' \in M'$ to $\sigma(f(m')) \in E$. However, since σ is a section, we have:

$$\pi(\sigma(f(m')) = (\pi \circ \sigma \circ f)(m') = (\mathrm{id}_M \circ f)(m') = f(m')$$

and hence $(m', (\sigma \circ f)(m')) \in E'$ by definition of E'. Moreover:

$$\pi'(m', (\sigma \circ f)(m')) = m'$$

and hence the map:

$$\begin{aligned} \sigma' \colon M' &\to E' \\ m' &\mapsto (m', (\sigma \circ f)(m')) \end{aligned}$$

satisfies $\pi' \circ \sigma' = \mathrm{id}_{M'}$ and it is thus a section on the pull-back bundle $E' \xrightarrow{\pi'} M'$.

The reason of introducing the concept of a topological bundle, is because we need it in order to construct the so called "tangent bundle".

2.9 The Tangent Bundle

Up to this point, we have defined everything on the level of a point on the manifold. However, since we are interested in describing quantities as a whole in the entire manifold, we would like to define a vector field on a manifold M as a "smooth" map that assigns to each $p \in M$ a tangent vector in T_pM. However, since this would then be a "map" to a different space at each point, it is unclear how to define its smoothness.

The simplest solution is to merge all the tangent spaces into a unique set and equip it with a smooth structure, so that we can then define a vector field as a smooth map between smooth manifolds.

Definition 2.42 (Tangent Bundle)**.** *Given a smooth manifold M, the **tangent bundle** of M is the disjoint union of all the tangent spaces to M, i.e.*

$$TM := \dot{\bigcup}_{p \in M} T_pM,$$

equipped with the canonical projection map

$$\begin{aligned} \pi \colon TM &\to M \\ X_{\gamma,p} &\mapsto p, \end{aligned}$$

where p is the unique $p \in M$ such that $X_{\gamma,p} \in T_pM$.

In other word the projection map, takes a vector from TM and spits out the point that it belongs $\pi(X_{\gamma,p}) = p$.

Since TM is simpy a set (and not a smooth manifold), up to here what we have is a set bundle. In order for this set bundle to turn to a topological bundle as we defined it previously, we need to equip TM with the structure of a smooth manifold. We can achieve this by constructing a smooth atlas for TM from a smooth atlas on M, as follows.

Let $\mathscr{A}_M$ be a smooth atlas on M and let $(U, x) \in \mathscr{A}_M$. If $X_{\gamma,p} \in \mathrm{preim}_\pi(U) \subseteq TM$, then $X_{\gamma,p} \in T_{\pi(X_{\gamma,p})}M$, by definition of π. Moreover, since $\pi(X_{\gamma,p}) = p \in U$, we can expand $X_{\gamma,p}$ in terms of the basis induced by the chart (U, x):

$$X_{\gamma,p} = X^a \left(\frac{\partial}{\partial x^a}\right)_p = X^a \left(\frac{\partial}{\partial x^a}\right)_{\pi(X_{\gamma,p})}$$

where $X^1, \ldots, X^{\dim M} \in \mathbb{R}$. We can then define the map

$$\begin{aligned} \xi \colon \mathrm{preim}_\pi(U) &\to x(U) \times \mathbb{R}^{\dim M} \cong_{\mathrm{set}} \mathbb{R}^{2\dim M} \\ X_{\gamma,p} &\mapsto (x(\pi(X_{\gamma,p})), X^1, \ldots, X^{\dim M}). \end{aligned}$$

Assuming that TM is equipped with a suitable topology, for instance the initial topology (i.e. the coarsest topology on TM that makes π continuous), we claim that the pair $(\mathrm{preim}_\pi(U), \xi)$ is a chart on TM and

$$\mathscr{A}_{TM} := \{(\mathrm{preim}_\pi(U), \xi) \mid (U, x) \in \mathscr{A}_M\}$$

is a smooth atlas on TM. Note that, from its definition, it is clear that ξ is a bijection. We will not show that $(\mathrm{preim}_\pi(U), \xi)$ is a chart here, but we will show that $\mathscr{A}_{TM}$ is a smooth atlas.

Proposition 2.6. *Any two charts* $(\mathrm{preim}_\pi(U), \xi), (\mathrm{preim}_\pi(\widetilde{U}), \widetilde{\xi}) \in \mathscr{A}_{TM}$ *are* $\mathcal{C}^\infty$*-compatible.*

Proof. Let (U, x) and $(\widetilde{U}, \widetilde{x})$ be the two charts on M giving rise to $(\mathrm{preim}_\pi(U), \xi)$ and $(\mathrm{preim}_\pi(\widetilde{U}), \widetilde{\xi})$, respectively. We need to show that the map

$$\widetilde{\xi} \circ \xi^{-1} \colon x(U \cap \widetilde{U}) \times \mathbb{R}^{\dim M} \to \widetilde{x}(U \cap \widetilde{U}) \times \mathbb{R}^{\dim M}$$

is smooth, as a map between open subsets of $\mathbb{R}^{2\dim M}$. Recall that such a map is smooth if, and only if, it is smooth componentwise. On the first $\dim M$ components, $\widetilde{\xi} \circ \xi^{-1}$ acts as

$$\begin{aligned} \widetilde{x} \circ x^{-1} \colon x(U \cap \widetilde{U}) &\to \widetilde{x}(U \cap \widetilde{U}) \\ x(p) &\mapsto \widetilde{x}(p), \end{aligned}$$

while on the remaining $\dim M$ components it acts as the change of vector components we met previously, i.e.

$$X^a \mapsto \widetilde{X}^a = \partial_b(y^a \circ x^{-1})(x(p))\, X^b.$$

Hence, we have

$$\begin{aligned} \widetilde{\xi} \circ \xi^{-1} \colon \qquad x(U \cap \widetilde{U}) \times \mathbb{R}^{\dim M} &\to \widetilde{x}(U \cap \widetilde{U}) \times \mathbb{R}^{\dim M} \\ (x(\pi(X_{\gamma,p})), X^1, \dots, X^{\dim M}) &\mapsto (\widetilde{x}(\pi(X_{\gamma,p})), \widetilde{X}^1, \dots, \widetilde{X}^{\dim M}), \end{aligned}$$

which is smooth in each component, and hence smooth. $\square$

The tangent bundle of a smooth manifold M is therefore itself a smooth manifold of dimension $2\dim M$, and the projection $\pi \colon TM \to M$ is smooth with respect to this structure.

Now by using the smooth manifold M as the base space, the smooth manifold TM as the total space, and the smooth projection π we can define the topological tangent bundle as the triple:

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

Similarly, one can construct the *cotangent bundle* T^*M to M by defining

$$T^*M := \dot{\bigcup}_{p \in M} T_p^*M$$

and going through the above again, using the dual basis $\{(\mathrm{d}x^a)_p\}$ instead of $\{\big(\frac{\partial}{\partial x^a}\big)_p\}$.

2.10 Vector, Covector And Tensor Fields

Now that we have defined the tangent and cotangent bundles, we are ready to define fields.

Definition 2.43 (Vector Field). *Let M be a smooth manifold, and let $TM \xrightarrow{\pi} M$ be its tangent bundle. A **vector field** σ on M is a smooth section of the tangent bundle, i.e. a smooth map $\sigma \colon M \to TM$ such that $\pi \circ \sigma = \mathrm{id}_M$.*

$$\begin{tikzcd} TM \arrow[d, shift left, "\pi"] \\ M \arrow[u, shift left, "\sigma"] \end{tikzcd}$$

Definition 2.44 ($\Gamma(TM)$). *We denote the set of all vector fields on M by $\Gamma(TM)$, i.e.*

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

This is, in fact, the standard notation for the set of all sections on a bundle.

Remark 2.15. An equivalent definition is that a vector field σ on M is a derivation on the algebra $\mathcal{C}^\infty(M)$, i.e. an $\mathbb{R}$-linear map

$$\sigma \colon \mathcal{C}^\infty(M) \xrightarrow{\sim} \mathcal{C}^\infty(M)$$

satisfying the Leibniz rule (with respect to pointwise multiplication on $\mathcal{C}^\infty(M)$)

$$\sigma(fg) = g\,\sigma(f) + f\,\sigma(g).$$

This definition is better suited for some purposes, and later on we will switch from one to the other without making any notational distinction between them.

We can equip the set $\Gamma(TM)$ with the following operations. The first is our, by now familiar, pointwise addition:

$$\begin{aligned} \oplus \colon \Gamma(TM) \times \Gamma(TM) &\to \Gamma(TM) \\ (\sigma, \tau) &\mapsto \sigma \oplus \tau, \end{aligned}$$

where

$$\begin{aligned} \sigma \oplus \tau \colon M &\to \Gamma(TM) \\ p &\mapsto (\sigma \oplus \tau)(p) := \sigma(p) + \tau(p). \end{aligned}$$

Note that the $+$ on the right hand side above is the addition in T_pM.

More interestingly, we can define a multiplication operation not by a simple number (i.e an element of $\mathbb{R}$) but with a whole function (i.e an element of $\mathcal{C}^\infty(M)$) as follows:

$$\begin{aligned} \odot \colon \mathcal{C}^\infty(M) \times \Gamma(TM) &\to \Gamma(TM) \\ (f, \sigma) &\mapsto f \odot \sigma, \end{aligned}$$

where

$$\begin{aligned} f \odot \sigma \colon M &\to \Gamma(TM) \\ p &\mapsto (f \odot \sigma)(p) := f(p)\sigma(p). \end{aligned}$$

Note that since $f \in \mathcal{C}^\infty(M)$, we have $f(p) \in \mathbb{R}$ and hence the multiplication above is the scalar multiplication on T_pM.

Remark 2.16. Of course, we could have defined $\odot$ simply as pointwise *global* scaling, using the reals $\mathbb{R}$ instead of the real functions $\mathcal{C}^\infty(M)$. Then, since $(\mathbb{R}, +, \cdot)$ is an algebraic field, we would then have the obvious $\mathbb{R}$-vector space structure on $\Gamma(TM)$. There are two reasons why we don't do that:

- Since the vector field acts on the whole manifold M (it assigns a value $f(p)$ on every point p of the manifold) we want to be able to assign different values to different points. Otherwise we would only be able to assign the same value to every point (i.e having a constant vector field)
- A basis for the corresponding vector space would be necessarily uncountably infinite, and hence it would not provide a very useful decomposition for our vector fields. Instead, the operation $\odot$ that we have defined allows for *local* scaling, i.e. we can scale a vector field by a different value at each point, and a much more useful decomposition of vector fields.

The question now is, mathematically speaking, what exactly the triple $(\Gamma(TM), \oplus, \odot)$ is. Its nature of course depends on what the triple $(\mathcal{C}^\infty(M), +, \cdot)$ is. Let's recall that the triple $(\mathcal{C}^\infty(M), +, \cdot)$ can be viewed in two different ways:

- $(\mathcal{C}^\infty(M), +, \cdot)$, where $\cdot$ is scalar multiplication (by a real number), is an $\mathbb{R}$-vector space.
- $(\mathcal{C}^\infty(M), +, \bullet)$, where $\bullet$ is pointwise multiplication of maps, is a commutative, unital ring, but not a division ring since not every function has an inverse at every point (i.e at all points that a function is zero, we cannot define an inverse since we would divide by zero).

The first view is of no use since if the triple is seen as a vector space over the real numbers, there is nothing else we can do. However, if we consider the second view. i.e the triple $(\mathcal{C}^\infty(M), +, \bullet)$, where $\bullet$ is pointwise function multiplication as a ring, then the triple $(\Gamma(TM), \oplus, \odot)$ built on top of this ring satisfies

- $(\Gamma(TM), \oplus)$ is an abelian group, with $0 \in \Gamma(TM)$ being the section that maps each $p \in M$ to the zero tangent vector in T_pM;
- $\Gamma(TM) \setminus \{0\}$ satisfies:

  i) $\forall\, f \in \mathcal{C}^\infty(M) : \forall\, \sigma, \tau \in \Gamma(TM) \setminus \{0\} : f \odot (\sigma \oplus \tau) = (f \odot \sigma) \oplus (f \odot \tau)$;

  ii) $\forall\, f, g \in \mathcal{C}^\infty(M) : \forall\, \sigma \in \Gamma(TM) \setminus \{0\} : (f + g) \odot \sigma = (f \odot \sigma) \oplus (g \odot \sigma)$;

  iii) $\forall\, f, g \in \mathcal{C}^\infty(M) : \sigma \in \Gamma(TM) \setminus \{0\} : (f \bullet g) \odot \sigma = f \odot (g \odot \sigma)$;

  iv) $\forall\, \sigma \in \Gamma(TM) \setminus \{0\} : 1 \odot \sigma = \sigma$,

  where $1 \in \mathcal{C}^\infty(M)$ maps every $p \in M$ to $1 \in \mathbb{R}$.

which are precisely the axioms for a vector space! Hence given that the triple $(\mathcal{C}^\infty(M), +, \bullet)$ is a ring, that turns the triple $(\Gamma(TM), \oplus, \odot)$ to a $\mathcal{C}^\infty(M)$-module.

And this of course is of crucial importance since as we showed in previous chapters, if a ring R is not a division ring, then a R-module does not need to have a basis. And since as we already said $(\mathcal{C}^\infty(M), +, \bullet)$ is not a division ring, the vector fields as $\mathcal{C}^\infty(M)$-modules do not need to have a basis! And this is a shame, since if they would have a basis (let's say X_i) we would be able to write a vector field σ as:

$$\sigma = \sigma^i X_i$$

where σ^i would be functions acting as components of the vector field!

In a similar manner one can construct a covector field through the use of the cotangent bundle, and from there to define the set of all covector fields $\Gamma(T^*M)$ and subsequently a triple $(\Gamma(T^*M), \oplus, \odot)$.

Finally using $\Gamma(TM)$ and $\Gamma(T^*M)$ we can define a tensor field.

Definition 2.45 (Tensor Field). *Let M be a smooth manifold. A smooth (r, s)* ***tensor field*** *τ on M is a $\mathcal{C}^\infty(M)$-multilinear map*

$$\tau \colon \underbrace{\Gamma(T^*M) \times \cdots \times \Gamma(T^*M)}_{r \text{ copies}} \times \underbrace{\Gamma(TM) \times \cdots \times \Gamma(TM)}_{s \text{ copies}} \to \mathcal{C}^\infty(M).$$

The equivalence of this to the bundle definition is due to the pointwise nature of tensors. For instance, a covector field $\omega \in \Gamma(T^*M)$ can act on a vector field $X \in \Gamma(TM)$ to yield a smooth function $\omega(X) \in \mathcal{C}^\infty(M)$ by

$$(\omega(X))(p) := \omega(p)(X(p)).$$

Then, we see that for any $f \in \mathcal{C}^\infty(M)$, we have

$$(\omega(fX))(p) = \omega(p)(f(p)X(p)) = f(p)\omega(p)(X(p)) =: (f\omega(X))(p)$$

and hence, the map $\omega \colon \Gamma(TM) \xrightarrow{\sim} \mathcal{C}^\infty(M)$ is $\mathcal{C}^\infty(M)$-linear.

Similarly, the set $\Gamma(T^r_sM)$ of all (r, s) smooth tensor fields on M can be made into a $\mathcal{C}^\infty(M)$-module, with module operations defined pointwise.

We can also define the tensor product of tensor fields

$$\begin{aligned} \otimes \colon \Gamma(T^p_qM) \times \Gamma(T^r_sM) &\to \Gamma(T^{p+r}_{q+s}M) \\ (\tau, \sigma) &\mapsto \tau \otimes \sigma \end{aligned}$$

analogously to what we had with tensors on a vector space, i.e.

$$(\tau \otimes \sigma)(\omega_1, \dots, \omega_p, \omega_{p+1}, \dots, \omega_{p+r}, X_1, \dots, X_q, X_{q+1}, \dots, X_{q+s})$$
$$:= \tau(\omega_1, \dots, \omega_p, X_1, \dots, X_q)\, \sigma(\omega_{p+1}, \dots, \omega_{p+r}, X_{q+1}, \dots, X_{q+s}),$$

with $\omega_i \in \Gamma(T^*M)$ and $X_i \in \Gamma(TM)$.

Therefore, we can think of tensor fields on M either as sections of some tensor bundle on M, that is, as maps assigning to each $p \in M$ a tensor ($\mathbb{R}$-multilinear map) on the vector space T_pM, or as a $\mathcal{C}^\infty(M)$-multilinear map as above. We will always try to pick the most useful or easier to understand, based on the context.

To summarize, fields are the generalization of the definitions of vectors, covectors and tensors at a specific point p of M, to every possible point p of manifold M, hence to the whole manifold M. In a similar way we can generalize the concept of the gradient of f at $p \in M$ in the gradient of f at M.

Recall the definition of the gradient operator at a point $p \in M$. We can extend that definition to define the ($\mathbb{R}$-linear) operator:

$$\begin{aligned} \mathrm{d}\colon \mathcal{C}^\infty(M) &\xrightarrow{\sim} \Gamma(T^*M) \\ f &\mapsto \mathrm{d}f \end{aligned}$$

where, of course, $\mathrm{d}f\colon p \mapsto \mathrm{d}_pf$. Alternatively, we can think of $\mathrm{d}f$ as the $\mathbb{R}$-linear map

$$\begin{aligned} \mathrm{d}f\colon \Gamma(TM) &\xrightarrow{\sim} \mathcal{C}^\infty(M) \\ X &\mapsto \mathrm{d}f(X) = X(f). \end{aligned}$$

Remark 2.17. Locally on some chart (U, x) on M, the covector field $\mathrm{d}f$ can be expressed as

$$\mathrm{d}f = \lambda_a \, \mathrm{d}x^a$$

for some smooth functions $\lambda_i \in \mathcal{C}^\infty(U)$. To determine what they are, we simply apply both sides to the vector fields induced by the chart. We have

$$\mathrm{d}f\left(\frac{\partial}{\partial x^b}\right) = \frac{\partial}{\partial x^b}(f) = \partial_b f$$

and

$$\lambda_a \, \mathrm{d}x^a\left(\frac{\partial}{\partial x^b}\right) = \lambda_a \, \frac{\partial}{\partial x^b}(x^a) = \lambda_a \, \delta^a_b = \lambda_b.$$

Hence, the local expression of $\mathrm{d}f$ on (U, x) is

$$\mathrm{d}f = \partial_a f \, \mathrm{d}x^a.$$

Note that the operator d satisfies the Leibniz rule

$$\mathrm{d}(fg) = g \, \mathrm{d}f + f \, \mathrm{d}g.$$

Finally, we want to generalize the concepts of push-forward and pull-back from a point to the whole manifold (a.k.a from a vector/covector to a vector/covector field). For a good reason we will first start with the pull-back.

To avoid confusion, for this part we will denote a covector as W and a covector field as ω. Recall that given a smooth map $\phi\colon M \to N$ the definition of a pull-back for a covector was

$$\begin{aligned} (\phi^*)_p\colon T^*_{\phi(p)}N &\xrightarrow{\sim} T^*_pM \\ W &\mapsto (\phi^*)_p(W), \end{aligned}$$

where $(\phi^*)_p(W)$ is defined as

$$\begin{aligned}(\phi^*)_p(W) \colon T_pM &\xrightarrow{\sim} \mathbb{R} \\ X &\mapsto (\phi^*)_p(W)(X) := W((\phi_*)_p(X)).\end{aligned}$$

Now we can simply extend the definition of a pull-back for a covector at point p denoted $(\phi^*)_p$, to this of a pull-back for a covector field on a manifold M denoted ϕ^*, by simpy acting with $(\phi^*)_p$ at every point p of the manifold M

$$\begin{aligned}\phi^* \colon T^*N &\to T^*M \\ \omega &\mapsto \phi^*(\omega)\end{aligned}$$

where now ω is a covector field and not just a covector, and $\phi^*(\omega)$ is defined for every point p of M as

$$\phi^*(\omega)(p) := (\phi^*)_p(W)$$

where W is the the corresponding covector that the covector field σ produces at point p. It is a common thing to this point to drop the p from the pull-back of covectors and simply write:

$$(\phi^*\omega)|_p := \phi^*(W|_p)$$

which actually means that the pull-back of a covector field evaluated at point p is equal to the pull-back of the covector $W|_p$ generated by the covector field ω at point p (a.k.a $W|_p = \omega(p)$) . From now on we will be using this equation to switch from covector field to covector equations, although practically it's the same thing from a different perspective.

While the pull-back can be extended from covectros to covector fields without problems, the push-forward of a vector cannot be generalized to the push-forward of a vector field unless the underlying map ϕ is a diffeomorphism between the manifold M and N. Let's see why.

Let's start again with the pull-back that we have already defined. Observe that the pull back of a covector field includes the notion of the pull-back of a covector at a point p. Now, the map ϕ, as a map, maps every single point of its domain M to a single point of its target N. Hence the whole target M is hit by the map, but the whole target N is not (recall that the part of N that is hit by the map is called the image of ϕ). This means that in the case of a pull-back (after we have defined the tangent vectors in both M and N) every single point of the image of ϕ on N will have a corresponding point back on M hence the definition of the pull-back of a covector field will be well-defined.

On the other hand, in the case of a push forward we get two problems coming from the fact that, in general, the map ϕ may not be neither surjective nor injective. First of all if the map ϕ is not surjective that means that the image of ϕ is not equal to the entire domain M ($\text{im}_\phi(M) \neq N$), hence from a vector field defined on M we will never be able to define a vector field on N that lies outside the image of ϕ. Moreover, if ϕ is not injective, that means that distinct elements of the domain M are mapped to the same element in the target N hence it might be the case that the push-forward will create many different vectors for one given point p on N which will make it ill-defined.

Of course, if the map ϕ is both surjective and injective, hence bijective, hence has an inverse, then none of this problems exist any more, since then the case is similar to the case of pull-backs (both directions of the map behave similarly). But recall that a bijection between topological spaces is called a "homeomorphism", and moreover if the map is smooth (which in the case of smooth manifolds by definition always is) then the smooth "homeomorphism" is called a"diffeomorphism".

So we ended up to our initial conclusion that the push-forward of a vector can be generalized to the push-forward of a vector field only if the underlying map ϕ is a diffeomorphism between the manifold M and N.

Then we can simply follow the same procedure as with the pull-back and define the push-forward of a vector field by simpy acting with the push-forward at every point $\phi(p)$ on the manifold N. Recall that the push-forward of ϕ at $p \in M$ is the linear map:

$$\begin{aligned}(\phi_*)_p \colon T_pM &\xrightarrow{\sim} T_{\phi(p)}N \\ X &\mapsto (\phi_*)_p(X)\end{aligned}$$

where $(\phi_*)_p(X)$ is defined as

$$\begin{aligned}(\phi_*)_p(X) \colon \mathcal{C}^\infty(N) &\xrightarrow{\sim} \mathbb{R} \\ f &\mapsto (\phi_*)_p(X)f := X(f \circ \phi).\end{aligned}$$

Now we can simply extend this definition to a push-forward for a vector field by simpy acting with the push-forward at every point p of the manifold M

$$\begin{aligned}\phi_* \colon TM &\to TN \\ \sigma &\mapsto \phi_*(\sigma)\end{aligned}$$

where $\phi_*(\sigma)$ is defined for every point $\phi(p)$ of N as

$$\phi_*(\sigma)(\phi(p)) := (\phi_*)_p(X)$$

As with covectors, it is a common thing to this point to drop the p from the equation and simply write:

$$(\phi_*\sigma)|_{\phi(p)} := \phi_*(X|_p)$$

which actually means that the push-forward of a vector field evaluated at point $\phi(p)$ is equal to the push-forward of the vector $X|_p$ generated by the vector field σ at point p (a.k.a $X|_p = X(p)$). From now on we will be using this equation to switch from vector field to vector equations, although practically it's the same thing from a different perspective.

2.11 Differential Forms

Definition 2.46 (Differential Form)**.** *Let M be a smooth manifold. A **(differential) n-form** on M is a $(0, n)$ smooth tensor field ω which is totally antisymmetric, i.e.*

$$\omega(X_1, \dots, X_n) = \operatorname{sgn}(\pi)\, \omega(X_{\pi(1)}, \dots, X_{\pi(n)}),$$

for any $\pi \in S_n$, with $X_i \in \Gamma(TM)$. We call n the degree *of the form.*

Alternatively, we can define a differential form as a smooth section of the appropriate bundle on M, i.e. as a map assigning to each $p \in M$ an n-form on the vector space T_pM.

Of course, by definition, differential forms are nothing more but a very specific kind of tensors, hence it's a subset of the tensor space.

Example 2.19. The electromagnetic field strength F is a differential 2-form built from the electric and magnetic fields, which are also taken to be forms. We will define these later in some detail.

Definition 2.47 ($\Omega^n(M)$)**.** *We denote by $\Omega^n(M)$ the set of all differential n-forms on M, which then becomes a $\mathcal{C}^\infty(M)$-module by defining the addition and multiplication operations pointwise.*

We have $\Omega^0(M) \equiv \mathcal{C}^\infty(M)$ since they are $(0, 0)$ tensors a.k.a functions and $\Omega^1(M) \equiv \Gamma(T^0_1M) \equiv \Gamma(T^*M)$ sincey they are $(0, 1)$ tensors a.k.a covectors.

We can specialise the pull-back of tensors to differential forms.

Definition 2.48 (Pull-Back On Differential Forms)**.** *Let $\phi \colon M \to N$ be a smooth map and let $\omega \in \Omega^n(N)$. Then we define the **pull-back** $\Phi^*(\omega) \in \Omega^n(M)$ of ω as*

$$\begin{aligned}\Phi^*(\omega) \colon M &\to T^*M \\ p &\mapsto \Phi^*(\omega)(p),\end{aligned}$$

where

$$\Phi^*(\omega)(p)(X_1, \ldots, X_n) := \omega(\phi(p))\big(\phi_*(X_1), \ldots, \phi_*(X_n)\big),$$

for $X_i \in T_pM$.

The map $\Phi^* \colon \Omega^n(N) \to \Omega^n(M)$ is $\mathbb{R}$-linear, and its action on $\Omega^0(M)$ is simply

$$\begin{aligned}\Phi^* \colon \Omega^0(M) &\to \Omega^0(M)\\ f &\mapsto \Phi^*(f) := f \circ \phi.\end{aligned}$$

This works for any smooth map ϕ, and it leads to a slight modification of our mantra:

Vectors are pushed forward,
forms are pulled back.

The tensor product $\otimes$ does not interact well with forms, since the tensor product of two forms is not necessarily a form (it might be, for example, a symmetric $(0, n)$ tensor which, be definition, is not a form). Hence, we define the following.

Definition 2.49 (Wedge Product)**.** *Let M be a smooth manifold. We define the* ***wedge*** *(or* exterior*)* product *of forms as the map*

$$\begin{aligned}\wedge \colon \Omega^n(M) \times \Omega^m(M) &\to \Omega^{n+m}(M)\\ (\omega, \sigma) &\mapsto \omega \wedge \sigma,\end{aligned}$$

where

$$(\omega \wedge \sigma)(X_1, \ldots, X_{n+m}) := \frac{1}{n!\,m!} \sum_{\pi \in S_{n+m}} \operatorname{sgn}(\pi)(\omega \otimes \sigma)(X_{\pi(1)}, \ldots, X_{\pi(n+m)})$$

and $X_1, \ldots, X_{n+m} \in \Gamma(TM)$. By convention, for any $f, g \in \Omega^0(M)$ and $\omega \in \Omega^n(M)$, we set

$$f \wedge g := fg \qquad \textit{and} \qquad f \wedge \omega = \omega \wedge f = f\omega.$$

Example 2.20. Suppose that $\omega, \sigma \in \Omega^1(M)$. Then, for any $X, Y \in \Gamma(TM)$

$$\begin{aligned}(\omega \wedge \sigma)(X, Y) &= (\omega \otimes \sigma)(X, Y) - (\omega \otimes \sigma)(Y, X)\\ &= (\omega \otimes \sigma)(X, Y) - \omega(Y)\sigma(X)\\ &= (\omega \otimes \sigma)(X, Y) - (\sigma \otimes \omega)(X, Y)\\ &= (\omega \otimes \sigma - \sigma \otimes \omega)(X, Y).\end{aligned}$$

Hence

$$\omega \wedge \sigma = \omega \otimes \sigma - \sigma \otimes \omega.$$

The wedge product is bilinear over $\mathcal{C}^\infty(M)$, that is

$$(f\omega_1 + \omega_2) \wedge \sigma = f\,\omega_1 \wedge \sigma + \omega_2 \wedge \sigma,$$

for all $f \in \mathcal{C}^\infty(M)$, $\omega_1, \omega_2 \in \Omega^n(M)$ and $\sigma \in \Omega^m(M)$, and similarly for the second argument.

Remark 2.18. If (U, x) is a chart on M, then every n-form $\omega \in \Omega^n(U)$ can be expressed locally on U as

$$\omega = \omega_{a_1 \cdots a_n}\, \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n},$$

where $\omega_{a_1 \cdots a_n} \in \mathcal{C}^\infty(U)$ and $1 \le a_1 < \cdots < a_n \le \dim M$. The $\mathrm{d}x^{a_i}$ appearing above are the covector fields (1-forms)

$$\mathrm{d}x^{a_i} \colon p \mapsto \mathrm{d}_p x^{a_i}.$$

The pull-back distributes over the wedge product.

Theorem 2.5. *Let $\phi\colon M \to N$ be smooth, $\omega \in \Omega^n(N)$ and $\sigma \in \Omega^m(N)$. Then, we have*

$$\Phi^*(\omega \wedge \sigma) = \Phi^*(\omega) \wedge \Phi^*(\sigma).$$

Proof. Let $p \in M$ and $X_1, \ldots, X_{n+m} \in T_pM$. Then we have

$$\begin{aligned}
&\big(\Phi^*(\omega) \wedge \Phi^*(\sigma)\big)(p)(X_1, \ldots, X_{n+m}) \\
&\quad = \frac{1}{n!\,m!} \sum_{\pi \in S_{n+m}} \operatorname{sgn}(\pi)\big(\Phi^*(\omega) \otimes \Phi^*(\sigma)\big)(p)(X_{\pi(1)}, \ldots, X_{\pi(n+m)}) \\
&\quad = \frac{1}{n!\,m!} \sum_{\pi \in S_{n+m}} \operatorname{sgn}(\pi)\Phi^*(\omega)(p)(X_{\pi(1)}, \ldots, X_{\pi(n)}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \Phi^*(\sigma)(p)(X_{\pi(n+1)}, \ldots, X_{\pi(n+m)}) \\
&\quad = \frac{1}{n!\,m!} \sum_{\pi \in S_{n+m}} \operatorname{sgn}(\pi)\omega(\phi(p))\big(\phi_*(X_{\pi(1)}), \ldots, \phi_*(X_{\pi(n)})\big) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \sigma(\phi(p))\big(\phi_*(X_{\pi(n+1)}), \ldots, \phi_*(X_{\pi(n+m)})\big) \\
&\quad = \frac{1}{n!\,m!} \sum_{\pi \in S_{n+m}} \operatorname{sgn}(\pi)(\omega \otimes \sigma)(\phi(p))\big(\phi_*(X_{\pi(1)}), \ldots, \phi_*(X_{\pi(n+m)})\big) \\
&\quad = (\omega \wedge \sigma)(\phi(p))\big(\phi_*(X_1), \ldots, \phi_*(X_{n+m})\big) \\
&\quad = \Phi^*(\omega \wedge \sigma)(p)(X_1, \ldots, X_{n+m}).
\end{aligned}$$

Since $p \in M$ was arbitrary, the statement follows. □

2.11.1 The Grassmann Algebra

Note that the wedge product takes two differential forms and produces a differential form of a different type. It would be much nicer to have a space which is closed under the action of $\wedge$. In fact, such a space exists and it is called the Grassmann algebra of M.

Definition 2.50 (Grassmann Algebra)**.** *Let M be a smooth manifold. Define the $\mathcal{C}^\infty(M)$-module*

$$\operatorname{Gr}(M) \equiv \Omega(M) := \bigoplus_{n=0}^{\dim M} \Omega^n(M).$$

*The **Grassmann algebra** on M is the algebra $(\Omega(M), +, \cdot, \wedge)$, where*

$$\wedge\colon \Omega(M) \times \Omega(M) \to \Omega(M)$$

is the linear continuation of the previously defined $\wedge\colon \Omega^n(M) \times \Omega^m(M) \to \Omega^{n+m}(M)$.

Recall that the direct sum of modules has the Cartesian product of the modules as underlying set and module operations defined componentwise. Also, note that by "algebra" here we really mean "algebra over a module".

Example 2.21. Let $\psi = \omega + \sigma$, where $\omega \in \Omega^1(M)$ and $\sigma \in \Omega^3(M)$. Of course, this "+" is neither the addition on $\Omega^1(M)$ nor the one on $\Omega^3(M)$, but rather that on $\Omega(M)$ and, in fact, $\psi \in \Omega(M)$.

Let $\varphi \in \Omega^n(M)$, for some n. Then

$$\varphi \wedge \psi = \varphi \wedge (\omega + \sigma) = \varphi \wedge \omega + \varphi \wedge \sigma,$$

where $\varphi \wedge \omega \in \Omega^{n+1}(M)$, $\varphi \wedge \sigma \in \Omega^{n+3}(M)$, and $\varphi \wedge \psi \in \Omega(M)$.

Example 2.22. There is a lot of talk about *Grassmann numbers*, particularly in supersymmetry. One often hears that these are "numbers that do not commute, but anticommute". Of course, objects cannot be commutative or anticommutative by themselves. These qualifiers only apply to operations on the objects. In fact, the Grassmann numbers are just the elements of a Grassmann algebra.

The following result is about the anticommutative behaviour of $\wedge$.

Theorem 2.6. *Let $\omega \in \Omega^n(M)$ and $\sigma \in \Omega^m(M)$. Then*

$$\omega \wedge \sigma = (-1)^{nm}\, \sigma \wedge \omega.$$

We say that $\wedge$ is *graded commutative*, that is, it satisfies a version of anticommutativity which depends on the degrees of the forms.

Proof. First note that if $\omega, \sigma \in \Omega^1(M)$, then

$$\omega \wedge \sigma = \omega \otimes \sigma - \sigma \otimes \omega = -\,\sigma \wedge \omega.$$

Recall that is $\omega \in \Omega^n(M)$ and $\sigma \in \Omega^m(M)$, then locally on a chart (U, x) we can write

$$\begin{aligned}\omega &= \omega_{a_1 \cdots a_n} \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} \\ \sigma &= \sigma_{b_1 \cdots b_m} \mathrm{d}x^{b_1} \wedge \cdots \wedge \mathrm{d}x^{b_m}\end{aligned}$$

with $1 \leq a_1 < \cdots < a_n \leq \dim M$ and similarly for the b_i. The coefficients $\omega_{a_1 \cdots a_n}$ and $\sigma_{b_1 \cdots b_m}$ are smooth functions in $\mathcal{C}^\infty(U)$. Since $\mathrm{d}x^{a_i}, \mathrm{d}x^{b_j} \in \Omega^1(M)$, we have

$$\begin{aligned}\omega \wedge \sigma &= \omega_{a_1 \cdots a_n} \sigma_{b_1 \cdots b_m} \, \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} \wedge \mathrm{d}x^{b_1} \wedge \cdots \wedge \mathrm{d}x^{b_m} \\ &= (-1)^n \, \omega_{a_1 \cdots a_n} \sigma_{b_1 \cdots b_m} \, \mathrm{d}x^{b_1} \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} \wedge \mathrm{d}x^{b_2} \wedge \cdots \wedge \mathrm{d}x^{b_m} \\ &= (-1)^{2n} \, \omega_{a_1 \cdots a_n} \sigma_{b_1 \cdots b_m} \, \mathrm{d}x^{b_1} \wedge \mathrm{d}x^{b_2} \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} \wedge \mathrm{d}x^{b_3} \wedge \cdots \wedge \mathrm{d}x^{b_m} \\ &\;\;\vdots \\ &= (-1)^{nm} \, \omega_{a_1 \cdots a_n} \sigma_{b_1 \cdots b_m} \, \mathrm{d}x^{b_1} \wedge \cdots \wedge \mathrm{d}x^{b_m} \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} \\ &= (-1)^{nm} \, \sigma \wedge \omega\end{aligned}$$

since we have swapped 1-forms nm-many times. □

Remark 2.19. We should stress that this is only true when ω and σ are pure degree forms, rather than linear combinations of forms of different degrees. Indeed, if $\varphi, \psi \in \Omega(M)$, a formula like

$$\varphi \wedge \psi = \cdots \psi \wedge \varphi$$

does not make sense in principle, because the different parts of φ and ψ can have different commutation behaviours.

2.11.2 The Exterior Derivative

Recall the "extended" definition of the gradient operator of a function d on the whole manifold M:

$$\begin{aligned}\mathrm{d} \colon \mathcal{C}^\infty(M) &\xrightarrow{\sim} \Gamma(T^*M) \\ f &\mapsto \mathrm{d}f\end{aligned}$$

Since $\Omega^0(M) \equiv \mathcal{C}^\infty(M)$ and $\Omega^1(M) \equiv \Gamma(T^0_1 M) \equiv \Gamma(T^*M)$, we can also understand this as an operator that takes in 0-forms and outputs 1-forms

$$\mathrm{d} \colon \Omega^0(M) \xrightarrow{\sim} \Omega^1(M).$$

This can then be extended to an operator which acts on any n-form. For this definition, we need to remind ourselves of the definition of commutator we gave in the algebra section of the notes. More precisely, if M is a smooth manifold and $X, Y \in \Gamma(TM)$ then the commutator (or Lie bracket) of X and Y is defined as

$$\begin{aligned}[X, Y] \colon \mathcal{C}^\infty(M) &\xrightarrow{\sim} \mathcal{C}^\infty(M) \\ f &\mapsto [X, Y](f) := X(Y(f)) - Y(X(f)),\end{aligned}$$

where we are using the definition of vector fields as $\mathbb{R}$-linear maps $\mathcal{C}^\infty(M) \xrightarrow{\sim} \mathcal{C}^\infty(M)$.

Using the commutator we can now extend the gradient as follows:

Definition 2.51 (Exterior Derivative). *The* ***exterior derivative*** *on M is the $\mathbb{R}$-linear operator*

$$\begin{aligned}\mathrm{d} \colon \Omega^n(M) &\xrightarrow{\sim} \Omega^{n+1}(M) \\ \omega &\mapsto \mathrm{d}\omega\end{aligned}$$

with $\mathrm{d}\omega$ *being defined as*

$$\begin{aligned}\mathrm{d}\omega(X_1,\dots,X_{n+1}) :=& \sum_{i=1}^{n+1}(-1)^{i+1}\,X_i\big(\omega(X_1,\dots,\widehat{X_i},\dots,X_{n+1})\big)\\ &+\sum_{i<j}(-1)^{i+j}\,\omega\big([X_i,X_j],X_1,\dots,\widehat{X_i},\dots,\widehat{X_j},\dots,X_{n+1}\big),\end{aligned}$$

where $X_i \in \Gamma(TM)$ *and the hat denotes omissions.*

Remark 2.20. Note that the operator d is only well-defined when it acts on forms. In order to define a derivative operator on general tensors we will need to add extra structure to our differentiable manifold.

Example 2.23. In the case $n = 1$, the form $\mathrm{d}\omega \in \Omega^2(M)$ is given by

$$\mathrm{d}\omega(X,Y) = X(\omega(Y)) - Y(\omega(X)) - \omega([X,Y]).$$

Let us check that this is indeed a 2-form, i.e. an antisymmetric, $\mathcal{C}^\infty(M)$-multilinear map

$$\mathrm{d}\omega\colon \Gamma(TM)\times\Gamma(TM)\to\mathcal{C}^\infty(M).$$

By using the antisymmetry of the Lie bracket, we immediately get

$$\mathrm{d}\omega(X,Y) = -\,\mathrm{d}\omega(Y,X).$$

Moreover, thanks to this identity, it suffices to check $\mathcal{C}^\infty(M)$-linearity in the first argument only. Additivity is easily checked

$$\begin{aligned}\mathrm{d}\omega(X_1+X_2,Y) &= (X_1+X_2)(\omega(Y)) - Y(\omega(X_1+X_2)) - \omega([X_1+X_2,Y])\\ &= X_1(\omega(Y)) + X_2(\omega(Y)) - Y(\omega(X_1)+\omega(X_2)) - \omega([X_1,Y]+[X_2,Y])\\ &= X_1(\omega(Y)) + X_2(\omega(Y)) - Y(\omega(X_1)) - Y(\omega(X_2)) - \omega([X_1,Y]) - \omega([X_2,Y])\\ &= \mathrm{d}\omega(X_1,Y) + \mathrm{d}\omega(X_2,Y).\end{aligned}$$

For $\mathcal{C}^\infty(M)$-scaling, first we calculate $[fX,Y]$. Let $g\in\mathcal{C}^\infty(M)$. Then

$$\begin{aligned}[fX,Y](g) &= fX(Y(g)) - Y(fX(g))\\ &= fX(Y(g)) - fY(X(g)) - Y(f)X(g)\\ &= f(X(Y(g)) - Y(X(g))) - Y(f)X(g)\\ &= f[X,Y](g) - Y(f)X(g)\\ &= (f[X,Y] - Y(f)X)(g).\end{aligned}$$

Therefore

$$[fX,Y] = f[X,Y] - Y(f)X.$$

Hence, we can calculate

$$\begin{aligned}\mathrm{d}\omega(fX,Y) &= fX(\omega(Y)) - Y(\omega(fX)) - \omega([fX,Y])\\ &= fX(\omega(Y)) - Y(f\omega(X)) - \omega(f[X,Y] - Y(f)X)\\ &= fX(\omega(Y)) - fY(\omega(X)) - Y(f)\omega(X) - f\omega([X,Y]) + \omega(Y(f)X)\\ &= fX(\omega(Y)) - fY(\omega(X)) - \not[gray]Y(f)\omega(X) - f\omega([X,Y]) + \not[gray]Y(f)\omega(X)\\ &= f\,\mathrm{d}\omega(X,Y),\end{aligned}$$

which is what we wanted.

The exterior derivative satisfies a graded version of the Leibniz rule with respect to the wedge product.

Theorem 2.7. *Let* $\omega\in\Omega^n(M)$ *and* $\sigma\in\Omega^m(M)$*. Then*

$$\mathrm{d}(\omega\wedge\sigma) = \mathrm{d}\omega\wedge\sigma + (-1)^n\,\omega\wedge\mathrm{d}\sigma.$$

Proof. We will work in local coordinates. Let (U, x) be a chart on M and write

$$\omega = \omega_{a_1 \cdots a_n} \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} =: \omega_A \mathrm{d}x^A$$
$$\sigma = \sigma_{b_1 \cdots b_m} \mathrm{d}x^{b_1} \wedge \cdots \wedge \mathrm{d}x^{b_m} =: \sigma_B \mathrm{d}x^B.$$

Locally, the exterior derivative operator d acts as

$$\mathrm{d}\omega = \mathrm{d}\omega_A \wedge \mathrm{d}x^A.$$

Hence

$$\begin{aligned}
\mathrm{d}(\omega \wedge \sigma) &= \mathrm{d}(\omega_A \sigma_B \, \mathrm{d}x^A \wedge \mathrm{d}x^B) \\
&= \mathrm{d}(\omega_A \sigma_B) \wedge \mathrm{d}x^A \wedge \mathrm{d}x^B \\
&= (\sigma_B \mathrm{d}\omega_A + \omega_A \mathrm{d}\sigma_B) \wedge \mathrm{d}x^A \wedge \mathrm{d}x^B \\
&= \sigma_B \mathrm{d}\omega_A \wedge \mathrm{d}x^A \wedge \mathrm{d}x^B + \omega_A \mathrm{d}\sigma_B \wedge \mathrm{d}x^A \wedge \mathrm{d}x^B \\
&= \sigma_B \mathrm{d}\omega_A \wedge \mathrm{d}x^A \wedge \mathrm{d}x^B + (-1)^n \omega_A \mathrm{d}x^A \wedge \mathrm{d}\sigma_B \wedge \mathrm{d}x^B \\
&= \sigma_B \mathrm{d}\omega \wedge \mathrm{d}x^B + (-1)^n \omega_A \mathrm{d}x^A \wedge \mathrm{d}\sigma \\
&= \mathrm{d}\omega \wedge \sigma + (-1)^n \, \omega \wedge \mathrm{d}\sigma
\end{aligned}$$

since we have "anticommuted" the 1-form $\mathrm{d}\sigma_B$ through the n-form $\mathrm{d}x^A$, picking up n minus signs in the process. □

An important property of the exterior derivative is the following.

Theorem 2.8. *Let $\phi\colon M \to N$ be smooth. For any $\omega \in \Omega^n(N)$, we have*

$$\Phi^*(\mathrm{d}\omega) = \mathrm{d}(\Phi^*(\omega)).$$

Remark 2.21. Informally, we can write this result as $\Phi^*\mathrm{d} = \mathrm{d}\Phi^*$, and say that the exterior derivative "commutes" with the pull-back.

However, you should bear in mind that the two d's appearing in the statement are two different operators. On the left hand side, it is $\mathrm{d}\colon \Omega^n(N) \to \Omega^{n+1}(N)$, while it is $\mathrm{d}\colon \Omega^n(M) \to \Omega^{n+1}(M)$ on the right hand side.

Remark 2.22. Of course, we could also combine the operators d into a single operator acting on the Grassmann algebra on M

$$\mathrm{d}\colon \Omega(M) \to \Omega(M)$$

by linear continuation.

2.11.3 De Rham Cohomology

Definition 2.52 (Closed / Exact Forms)**.** *Let M be a smooth manifold and let $\omega \in \Omega^n(M)$. We say that ω is*

- ***closed*** *if* $\mathrm{d}\omega = 0$*;*
- ***exact*** *if* $\exists\, \sigma \in \Omega^{n-1}(M) : \omega = \mathrm{d}\sigma$.

The question of whether every closed form is exact and vice versa, i.e. whether the implications

$$(\mathrm{d}\omega = 0) \Leftrightarrow (\exists\, \sigma : \omega = \mathrm{d}\sigma)$$

hold in general, belongs to the branch of mathematics called cohomology theory, to which we will now provide an introduction.

The answer for the $\Leftarrow$ direction is affirmative thanks to the following result.

Theorem 2.9. *Let M be a smooth manifold. The operator*

$$\mathrm{d}^2 \equiv \mathrm{d} \circ \mathrm{d}\colon \Omega^n(M) \to \Omega^{n+2}(M)$$

is identically zero, i.e. $\mathrm{d}^2 = 0$.

Proof. This can be shown directly using the definition of d. Here, we will instead show it by working in local coordinates.

Recall that, locally on a chart (U, x), we can write any form $\omega \in \Omega^n(M)$ as

$$\omega = \omega_{a_1 \cdots a_n} \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n}.$$

Then, we have

$$\begin{aligned} \mathrm{d}\omega &= \mathrm{d}\omega_{a_1 \cdots a_n} \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} \\ &= \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n}, \end{aligned}$$

and hence

$$\mathrm{d}^2\omega = \partial_c \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^c \wedge \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n}.$$

We can perform a little "trick" in the last equation and write it as twice the half expression

$$\mathrm{d}^2\omega = \frac{1}{2}\partial_c \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^c \wedge \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} + \frac{1}{2}\partial_c \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^c \wedge \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n}$$

Now we can inter-switch the c and b dummy indices in the second half part (we can do it since they are just dummy indices) and we get

$$\mathrm{d}^2\omega = \frac{1}{2}\partial_c \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^c \wedge \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} + \frac{1}{2}\partial_b \partial_c \omega_{a_1 \cdots a_n} \mathrm{d}x^b \wedge \mathrm{d}x^c \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n}$$

Since $\mathrm{d}x^b \wedge \mathrm{d}x^c = -\,\mathrm{d}x^c \wedge \mathrm{d}x^b$, and moreover, by Schwarz's theorem, we have $\partial_c \partial_b \omega_{a_1 \cdots a_n} = \partial_b \partial_c \omega_{a_1 \cdots a_n}$ we get

$$\mathrm{d}^2\omega = \frac{1}{2}\partial_c \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^c \wedge \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n} - \frac{1}{2}\partial_c \partial_b \omega_{a_1 \cdots a_n} \mathrm{d}x^c \wedge \mathrm{d}x^b \wedge \mathrm{d}x^{a_1} \wedge \cdots \wedge \mathrm{d}x^{a_n}$$

Hence

$$\mathrm{d}^2\omega = 0$$

Since this holds for any ω, we have $\mathrm{d}^2 = 0$. □

Corollary 2.1. *Every exact form is closed.*

We can extend the action of d to the zero vector space $0 := \{0\}$ by mapping the zero in 0 to the zero function in $\Omega^0(M)$. In this way, we obtain the chain of $\mathbb{R}$-linear maps

$$0 \xrightarrow{\mathrm{d}} \Omega^0(M) \xrightarrow{\mathrm{d}} \Omega^1(M) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^n(M) \xrightarrow{\mathrm{d}} \Omega^{n+1}(M) \xrightarrow{\mathrm{d}} \cdots \xrightarrow{\mathrm{d}} \Omega^{\dim M}(M) \xrightarrow{\mathrm{d}} 0,$$

where we now think of the spaces $\Omega^n(M)$ as $\mathbb{R}$-vector spaces.

Recall from linear algebra section in the notes that, given a linear map $\phi \colon V \to W$, one can define the subspace of V

$$\ker(\phi) := \{v \in V \mid \phi(v) = 0\},$$

called the *kernel* of ϕ, and the subspace of W

$$\mathrm{im}(\phi) := \{\phi(v) \mid v \in V\},$$

called the *image* of ϕ.

Going back to our chain of maps, the equation $\mathrm{d}^2 = 0$ is equivalent to

$$\mathrm{im}(\mathrm{d} \colon \Omega^n(M) \to \Omega^{n+1}(M)) \subseteq \ker(\mathrm{d} \colon \Omega^{n+1}(M) \to \Omega^{n+2}(M))$$

for all $0 \leq n \leq \dim M - 2$. Moreover, we have

$$\omega \in \Omega^n(M) \text{ is closed } \Leftrightarrow \ \omega \in \ker(\mathrm{d}\colon \Omega^n(M) \to \Omega^{n+1}(M))$$
$$\omega \in \Omega^n(M) \text{ is exact } \Leftrightarrow \ \omega \in \operatorname{im}(\mathrm{d}\colon \Omega^{n-1}(M) \to \Omega^n(M)).$$

The traditional notation for the spaces on the right hand side above is

$$Z^n := \ker(\mathrm{d}\colon \Omega^n(M) \to \Omega^{n+1}(M)),$$
$$B^n := \operatorname{im}(\mathrm{d}\colon \Omega^{n-1}(M) \to \Omega^n(M)),$$

so that Z^n is the space of closed n-forms and B^n is the space of exact n-forms.

Our original question can be restated as: does $Z^n = B^n$ for all n? We have already seen that $\mathrm{d}^2 = 0$ implies that $B^n \subseteq Z^n$ for all n (B^n is, in fact, a vector subspace of Z^n). Unfortunately the equality does not hold in general, but we do have the following result.

Lemma 2.1 (Poincaré)**.** *Let $M \subseteq \mathbb{R}^d$ be a simply connected domain. Then*

$$Z^n = B^n\,, \qquad \forall\, n > 0.$$

In the cases where $Z^n \neq B^n$, we would like to quantify by how much the closed n-forms fail to be exact. The answer is provided by the cohomology group.

Definition 2.53 (de Rham Cohomology Group)**.** *Let M be a smooth manifold. The n-th **de Rham cohomology group** on M is the quotient $\mathbb{R}$-vector space*

$$H^n(M) := Z^n/B^n.$$

You can think of the above quotient as $Z^n/\!\sim$, where $\sim$ is the equivalence relation

$$\omega \sim \sigma \ :\Leftrightarrow \ \omega - \sigma \in B^n.$$

The answer to our question as it is addressed in cohomology theory is: every exact n-form on M is also closed and vice versa if, only if,

$$H^n(M) \cong_{\text{vec}} 0.$$

Of course, rather than an actual answer, this is yet another restatement of the question. However, if we are able to determine the spaces $H^n(M)$, then we do get an answer.

A crucial theorem by de Rham states (in more technical terms) that $H^n(M)$ only depends on the global topology of M. In other words, the cohomology groups are topological invariants. This is remarkable because $H^n(M)$ is defined in terms of exterior derivatives, which have everything to do with the local differentiable structure of M, and a given topological space can be equipped with several inequivalent differentiable structures.

Example 2.24. Let M be any smooth manifold. We have

$$H^0(M) \cong_{\text{vec}} \mathbb{R}^{(\#\text{ of connected components of } M)}$$

since the closed 0-forms are just the locally constant smooth functions on M. As an immediate consequence, we have

$$H^0(\mathbb{R}) \cong_{\text{vec}} H^0(S^1) \cong_{\text{vec}} \mathbb{R}.$$

Example 2.25. By Poincaré lemma, we have

$$H^n(M) \cong_{\text{vec}} 0$$

for any simply connected $M \subseteq \mathbb{R}^d$.

2.12 Application - Part 1: $\mathrm{SL}(2,\mathbb{C})$

In this final chapter we will go through an application containing (almost) everything we have mentioned so far. More specifically, we we will examine in detail the special linear group of degree 2 over $\mathbb{C}$, also known as the relativistic spin group.

$\mathrm{SL}(2,\mathbb{C})$ **As A Set**

We define the following subset of $\mathbb{C}^4 := \mathbb{C}\times\mathbb{C}\times\mathbb{C}\times\mathbb{C}$

$$\mathrm{SL}(2,\mathbb{C}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{C}^4 \;\middle|\; ad-bc=1 \right\} \subseteq \mathbb{C}^4,$$

where the array is just an alternative notation for a quadruple (a,b,c,d). It's this extra constraint $ad-bc=1$ that removes one degree of freedom and makes it a subset and not the whole $\mathbb{C}^4$.

$\mathrm{SL}(2,\mathbb{C})$ **As A Group**

We define an operation

$$\begin{aligned} \bullet \colon \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C}) &\to \mathrm{SL}(2,\mathbb{C}) \\ (\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix}) &\mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \end{aligned}$$

where

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \begin{pmatrix} e & f \\ g & h \end{pmatrix} := \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}.$$

Formally, this operation is the same as matrix multiplication. We can check directly that the result of applying $\bullet$ lands back in $\mathrm{SL}(2,\mathbb{C})$, or simply recall that the determinant of a product is the product of the determinants. Moreover, the operation $\bullet$

i) is associative (straightforward but tedious to check);

ii) has an identity element, namely $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C})$;

iii) admits inverses: for each $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C})$, we have $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C})$ and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bullet \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \bullet \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence, we have $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Therefore, the pair $(\mathrm{SL}(2,\mathbb{C}), \bullet)$ is a (non-commutative) group.

$\mathrm{SL}(2,\mathbb{C})$ **As A Topological Space**

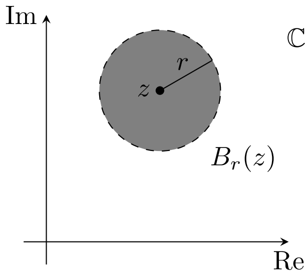
Recall that if N is a subset of M and $\mathcal{O}$ is a topology on M, then we can equip N with the subset topology inherited from M

$$\mathcal{O}|_N := \{U \cap N \mid U \in \mathcal{O}\}.$$

We begin by establishing a topology on $\mathbb{C}$ as follows. Let

$$B_r(z) := \{y \in \mathbb{C} \mid |z-y| < r\}$$

be the open ball of radius $r>0$ and centre $z\in\mathbb{C}$.



Define $\mathcal{O}_{\mathbb{C}}$ implicitly by

$$U \in \mathcal{O}_{\mathbb{C}} \;:\Leftrightarrow\; \forall\, z \in U : \exists\, r > 0 : B_r(z) \subseteq U.$$

Then, the pair $(\mathbb{C}, \mathcal{O}_{\mathbb{C}})$ is a topological space. In fact, we have

$$(\mathbb{C}, \mathcal{O}_{\mathbb{C}}) \cong_{\text{top}} (\mathbb{R}^2, \mathcal{O}_{\text{std}}).$$

We can then equip $\mathbb{C}^4$ with the product topology so that we can finally define

$$\mathcal{O} := (\mathcal{O}_{\mathbb{C}})|_{\mathrm{SL}(2,\mathbb{C})},$$

so that the pair $(\mathrm{SL}(2,\mathbb{C}), \mathcal{O})$ is a topological space. In fact, it is a connected topological space, and we will need this property later on.

$\mathrm{SL}(2,\mathbb{C})$ **As A Topological Manifold**

Recall that a topological space $(M, \mathcal{O})$ is a complex topological manifold if each point $p \in M$ has an open neighbourhood $U(p)$ which is homeomorphic to an open subset of $\mathbb{C}^d$. Equivalently, there must exist a $\mathcal{C}^0$-atlas, i.e. a collection $\mathscr{A}$ of charts (U_α, x_α), where the U_α are open and cover M and each x is a homeomorphism onto a subset of $\mathbb{C}^d$. In our case of $\mathrm{SL}(2,\mathbb{C})$ we will map it locally to $\mathbb{C}^3$ because as we said in the beginning of this section, $\mathrm{SL}(2,\mathbb{C})$ (as a set) is a subset of $\mathbb{C}^4 := \mathbb{C} \times \mathbb{C} \times \mathbb{C} \times \mathbb{C}$ due to the constraint $ad - bc = 1$ that removes one degree of freedom. This is why $\mathbb{C}^3$ suffices and we do not use $\mathbb{C}^4$.

In general (but not in our case of $\mathrm{SL}(2,\mathbb{C})$) nothing stops us from using just one chart (U, x) and cover the whole topological space. However, as we will see, here we need two charts to cover the whole $\mathrm{SL}(2,\mathbb{C})$ (in the same manner that a sphere cannot be covered with just one chart).

Let U be the set

$$U := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{C}) \;\middle|\; a \neq 0 \right\}$$

and define the map

$$\begin{aligned} x\colon \quad & U \to x(U) \subseteq \mathbb{C}^* \times \mathbb{C} \times \mathbb{C} \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, c), \end{aligned}$$

where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

Notice that given the mapping (a, b, c) one can reconstruct d from the constraint on the degree of freedom $d = \frac{1+bc}{a}$, hence the mapping to $\mathbb{C}^3$.

With a little more work on this direction, one can show that U is an open subset of $(\mathrm{SL}(2,\mathbb{C}), \mathcal{O})$ and x is a homeomorphism with inverse

$$\begin{aligned} x^{-1}\colon \quad & x(U) \to U \\ & (a, b, c) \mapsto \begin{pmatrix} a & b \\ c & \frac{1+bc}{a} \end{pmatrix}. \end{aligned}$$

This is the reason why we excluded the case $a = 0$ when we defined the set U of the chart (U, x), since if we hadn't, we wouldn't be able to divide with a and the map x wouldn't have an inverse.

However, this makes the chart (U, x) to not cover the whole $\mathrm{SL}(2, \mathbb{C})$ since U as a set takes care only the elements of $\mathrm{SL}(2, \mathbb{C})$ with $a \neq 0$. Hence we need at least one more chart. We thus define the set

$$V := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2, \mathbb{C}) \;\middle|\; b \neq 0 \right\}$$

and the map

$$\begin{aligned} y\colon \quad & V \to x(V) \subseteq \mathbb{C} \times \mathbb{C}^* \times \mathbb{C} \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, d). \end{aligned}$$

Similarly to the above, V is open and y is a homeomorphism with inverse

$$\begin{aligned} y^{-1}\colon \quad & x(V) \to V \\ & (a, b, d) \mapsto \begin{pmatrix} a & b \\ \frac{ad-1}{b} & d \end{pmatrix}. \end{aligned}$$

An element of $\mathrm{SL}(2, \mathbb{C})$ cannot have both a and b equal to zero, for otherwise $ad - bc = 0 \neq 1$. Hence $\mathscr{A}_{\mathrm{top}} := \{(U, x), (V, y)\}$ is an atlas, and we showed that x and y are homeomorphisms hence continuous and invertible. Since every atlas is automatically a $\mathcal{C}^0$-atlas, the triple $(\mathrm{SL}(2, \mathbb{C}), \mathcal{O}, \mathscr{A}_{\mathrm{top}})$ is a 3-dimensional, complex, topological manifold.

$\mathrm{SL}(2, \mathbb{C})$ As A Complex Differentiable Manifold

Recall that to obtain a $\mathcal{C}^1$-differentiable manifold from a topological manifold with atlas $\mathscr{A}$, we have to check that every transition map between charts in $\mathscr{A}$ is differentiable in the usual sense. Remember this picture:

$$\begin{tikzcd} & U \cap V \subseteq \mathrm{SL}(2, \mathbb{C}) \arrow[dl, "x"'] \arrow[dr, "y"] & \\ x(U \cap V) \subseteq \mathbb{C}^3 \arrow[rr, "y \circ x^{-1}"'] & & y(U \cap V) \subseteq \mathbb{C}^3 \end{tikzcd}$$

In our case, we have the atlas $\mathscr{A}_{\mathrm{top}} := \{(U, x), (V, y)\}$. We evaluate

$$(y \circ x^{-1})(a, b, c) = y\left(\begin{pmatrix} a & b \\ c & \frac{1+bc}{a} \end{pmatrix}\right) = (a, b, \tfrac{1+bc}{a}).$$

Hence we have the transition map

$$\begin{aligned} y \circ x^{-1}\colon & x(U \cap V) \to y(U \cap V) \\ & (a, b, c) \mapsto (a, b, \tfrac{1+bc}{a}). \end{aligned}$$

Similarly, we have

$$(x \circ y^{-1})(a, b, d) = y\left(\begin{pmatrix} a & b \\ \frac{ad-1}{b} & d \end{pmatrix}\right) = (a, b, \tfrac{ad-1}{b}).$$

Hence, the other transition map is

$$\begin{aligned} x \circ y^{-1}\colon & y(U \cap V) \to x(U \cap V) \\ & (a, b, c) \mapsto (a, b, \tfrac{ad-1}{b}). \end{aligned}$$

Since $a \neq 0$ and $b \neq 0$, the transition maps are complex differentiable.

Therefore, the atlas $\mathscr{A}_{\text{top}}$ is a differentiable atlas. By defining $\mathscr{A}$ to be the maximal differentiable atlas containing $\mathscr{A}_{\text{top}}$, we have that $(\mathrm{SL}(2,\mathbb{C}), \mathcal{O}, \mathscr{A})$ is a 3-dimensional, complex differentiable manifold.

Chapter 3

Lie Groups

3.1 Lie Groups

Definition 3.1 (Lie Group)**.** *A* ***Lie group*** *is a group $(G, \bullet)$, where G is a smooth manifold and the maps*

$$\begin{aligned} \mu\colon\ & G \times G \to G \\ & (g_1, g_2) \mapsto g_1 \bullet g_2 \end{aligned}$$

and

$$\begin{aligned} i\colon & G \to G \\ & g \mapsto g^{-1} \end{aligned}$$

are both smooth. Note that $G \times G$ inherits a smooth atlas from the smooth atlas of G.

Definition 3.2 (Dimension Of Lie Group)**.** *The* ***dimension*** *of a Lie group $(G, \bullet)$ is the dimension of G as a manifold.*

Example 3.1.

a) Consider $(\mathbb{R}^n, +)$, where $\mathbb{R}^n$ is understood as a smooth n-dimensional manifold. This is a commutative (or abelian) Lie group (since $\bullet$ is commutative), often called the n-dimensional translation group.

b) Let $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ and let $\cdot$ be the usual multiplication of complex numbers. Then $(S^1, \cdot)$ is a commutative Lie group usually denoted U(1).

c) As we discussed in the chapter of vector spaces, $\mathrm{Aut}(V) := \{\phi\colon V \xrightarrow{\sim} V \mid \det\phi \neq 0\}$ is the set of all of linear isomorphisms on V that we denoted as $GL(V)$. To make it more concrete, we can take as the vector space $V = \mathbb{R}^n$ hence $\mathrm{GL}(n, \mathbb{R}) = \{\phi\colon \mathbb{R}^n \xrightarrow{\sim} \mathbb{R}^n \mid \det\phi \neq 0\}$. This set can be endowed with the structure of a smooth n^2-dimensional manifold, by noting that there is a bijection between linear maps $\phi\colon \mathbb{R}^n \xrightarrow{\sim} \mathbb{R}^n$ and $\mathbb{R}^{2n}$. Then, $(\mathrm{GL}(n, \mathbb{R}), \circ)$ is a Lie group called the *general linear group*.

Definition 3.3 (Lie Group Homomorphism)**.** *Let $(G, \bullet)$ and $(H, \circ)$ be Lie groups. A map $\phi\colon G \to H$ is* ***Lie group homomorphism*** *if it is a group homomorphism and a smooth map.*

Definition 3.4 (Lie Group Isomorphism)**.** *A* ***Lie group isomorphism*** *is a Lie group homomorphism which is also a diffeomorphism of the underlying manifolds.*

3.2 The Left Translation Map

To every element of a Lie group there is associated a special map. Note that everything we will do here can be done equivalently by using right translation maps.

Definition 3.5 (Left Translation)**.** *Let $(G, \bullet)$ be a Lie group and let $g \in G$. The map*

$$\begin{aligned} \ell_g\colon & G \to G \\ & h \mapsto \ell_g(h) := g \bullet h \equiv gh \end{aligned}$$

*is called the **left translation** by g.*

One might thing that this is an overkill of notation since we already had the operation between two element from the group structure. However the left translation map is different, since we first have to fix an element of the group g (hence the index in ℓ_g) and then apply this element to the whole group (a.k.a to each element of the group).

If there is no danger of confusion, we usually suppress the $\bullet$ notation.

Proposition 3.1. *Let G be a Lie group. For any $g \in G$, the left translation map $\ell_g \colon G \to G$ is a isomorphism.*

Proof. Let $h, h' \in G$. Then, we have

$$\ell_g(h) = \ell_g(h') \;\Leftrightarrow\; gh = gh' \;\Leftrightarrow\; h = h'.$$

Moreover, for any $h \in G$, we have $g^{-1}h \in G$ and

$$\ell_g(g^{-1}h) = gg^{-1}h = h.$$

Therefore, ℓ_g is a bijection on G.

Note that

$$\ell_g = \mu(g, -)$$

and since $\mu \colon G \times G \to G$ is smooth by definition, so is ℓ_g.

The inverse map is $(\ell_g)^{-1} = \ell_{g^{-1}}$, since

$$\ell_{g^{-1}} \circ \ell_g = \ell_g \circ \ell_{g^{-1}} = \mathrm{id}_G\,.$$

Then, for the same reason as above with g replaced by g^{-1}, the inverse map $(\ell_g)^{-1}$ is also smooth. Hence, the map ℓ_g is indeed an isomorphism. □

Note that, ℓ_g is not an isomorphism of groups, i.e.

$$\ell_g(hh') \neq \ell_g(h)\,\ell_g(h')$$

in general. However that does not stop ℓ_g from being an isomorphism, which actually means that is a diffeomorphism of the underlying manifolds.

Since a lie group is a topological manifold, on top of being a group, this means that at any point g we can define the tangent space, and by following the analysis we did in the previous chapter to define fields on G. Recall from the previous chapter than once we have a diffeomorphism ϕ between two manifolds M and N, we can define the push-forward of a vector field X as

$$(\phi_* X)|_{\phi(p)} := \phi_*(X|_p)$$

where $X|_p$ is the vector created by the field X on point p.

Coming in our case, we just showed that the map $\ell_g \colon G \to G$ is a diffeomorphism so we can push-forward any vector field X on G to another vector field (again on G since the maps is between the same manifold). So in our case $\phi_*(X) = (\ell_g)_*(X)$ and for any point h in G: $\ell_g(h) = gh$ so the push-forward equation reads

$$(\ell_{g*} X)|_{gh} := (\ell_g)_*(X|_h)$$

3.3 The Lie Algebra Of A Lie Group

In Lie theory, we are typically not interested in general vector fields, but rather on special class of vector fields which are invariant under the induced push-forward of the left translation maps ℓ_g.

Definition 3.6 (Left Invariant Vector Fields)**.** *Let G be a Lie group. A vector field $X \in \Gamma(TG)$ is said to be **left-invariant** if*

$$\forall\, g \in G:\ (l_g)_*(X) = X.$$

Equivalently, we can require this to hold pointwise

$$\forall\, g, h \in G:\ (\ell_g)_*(X|_h) = X|_{gh}.$$

We can manipulate a bit the pointwise formulation to yield another reformulation. Since both sides are vectors we can let them act on a function f

$$(\ell_g)_*(X|_h)f = X|_{gh}f$$

By using the definition of a push-forward of a vector $(\phi_*)_p(X)f := X(f \circ \phi)$ the left part of the equation reads

$$(\ell_g)_*(X|_h)f = X|_h(f \circ \ell_g) = \big(X(f \circ \ell_g)\big)|_h$$

The right part can be manipulated as follows

$$X|_{gh}f = (Xf)|_{gh} = \Big((Xf) \circ \ell_g\Big)|_h$$

By substituting both final expressions back to the original one and discarding the point h since they must be true for any h we obtain the last reformulation of the push-forward

$$X(f \circ \ell_g) = X(f) \circ \ell_g.$$

Once again, let's emphasize that this equations holds only for left invariant vector fields and not all vector fields X of $\Gamma(TG)$.

Definition 3.7 ($\mathcal{L}(G)$ (As A Set))**.** *We denote the set of all left-invariant vector fields on G as $\mathcal{L}(G)$.*

Of course,

$$\mathcal{L}(G) \subseteq \Gamma(TG)$$

but, in fact, more is true. Recall that we equipped $(\Gamma(TG)$ with two operations and we showed that $(\Gamma(TG), +, \cdot)$ is in fact a $\mathcal{C}^\infty(G)$-submodule. One can check that $\mathcal{L}(G)$ is closed under

$$\begin{aligned} +&\colon \mathcal{L}(G) \times \mathcal{L}(G) \to \mathcal{L}(G) \\ \cdot&\colon \mathcal{C}^\infty(G) \times \mathcal{L}(G) \to \mathcal{L}(G), \end{aligned}$$

therefore, $\mathcal{L}(G)$ is a $\mathcal{C}^\infty(G)$-submodule of $\Gamma(TG)$.

On top of that we said that $(\Gamma(TG), +, \cdot)$ can also be seen as an $\mathbb{R}$-vector space. Up to now, we have refrained from thinking of $\Gamma(TG)$ as an $\mathbb{R}$-vector space since it is infinite-dimensional and, even worse, a basis is in general uncountable. A priori, this could be true for $\mathcal{L}(G)$ as well, but we will see that the situation is, in fact, much nicer as $\mathcal{L}(G)$ will turn out to be a finite-dimensional vector space over $\mathbb{R}$ (as an $\mathbb{R}$-vector subspace of $\Gamma(TG)$). Let's see why, since the reason why it is so it's of crucial importance.

Theorem 3.1. *Let G be a Lie group with identity element $e \in G$. Then $\mathcal{L}(G) \cong_{\text{vec}} T_eG$.*

Proof. We will construct a linear isomorphism $j\colon T_eG \xrightarrow{\sim} \mathcal{L}(G)$. Define

$$\begin{aligned} j\colon T_eG &\to \Gamma(TG) \\ A &\mapsto j(A), \end{aligned}$$

where $j(A)$ is define as

$$\begin{aligned} j(A)\colon G &\to TG \\ g &\mapsto j(A)|_g := (\ell_g)_*(A). \end{aligned}$$

Now we have to prove that this is actually a linear isomorphism, and we will do it in steps.

i) First, we show that for any $A \in T_eG$, $j(A)$ is a smooth vector field on G. It suffices to check that for any $f \in \mathcal{C}^\infty(G)$, we have $j(A)(f) \in \mathcal{C}^\infty(G)$. Indeed

$$\begin{aligned}(j(A)(f))(g) &= j(A)|_g(f)\\ &:= (\ell_g)_*(A)(f)\\ &= A(f \circ \ell_g)\\ &= (f \circ \ell_g \circ \gamma)'(0),\end{aligned}$$

where γ is a curve through $e \in G$ whose tangent vector at e is A. The map

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

is a composition of smooth maps, hence it is smooth. Then

$$(j(A)(f))(g) = (\partial_1 \varphi)(0, g)$$

depends smoothly on g and thus $j(A)(f) \in \mathcal{C}^\infty(G)$.

ii) We proved that $j(A)$ is indeed a smooth vector field, however now we need to prove that it is a left invariant vector field since it is an element of $\Gamma(TG)$. Let $g, h \in G$. Then, for every $A \in T_eG$, we have

$$\begin{aligned}(\ell_g)_*(j(A)|_h) &:= (\ell_g)_*((\ell_h)_*(A))\\ &= (\ell_{gh})_*(A)\\ &= j(A)|_{gh},\end{aligned}$$

so $j(A) \in \mathcal{L}(G)$. Hence, the map j is really $j \colon T_eG \to \mathcal{L}(G)$.

iii) We also need to check the linearity. Let $A, B \in T_eG$ and $\lambda \in \mathbb{R}$. Then, for any $g \in G$

$$\begin{aligned}j(\lambda A + B)|_g &= (\ell_g)_*(\lambda A + B)\\ &= \lambda(\ell_g)_*(A) + (\ell_g)_*(B)\\ &= \lambda j(A)|_g + j(B)|_g,\end{aligned}$$

since the push-forward is an $\mathbb{R}$-linear map. Hence, we have $j \colon T_eG \xrightarrow{\sim} \mathcal{L}(G)$.

iv) We also need to check that the map is injective. Let $A, B \in T_eG$. Then

$$\begin{aligned}j(A) = j(B) &\Leftrightarrow \forall\, g \in G : j(A)|_g = j(B)|_g\\ &\Rightarrow j(A)|_e = j(B)|_e\\ &\Leftrightarrow (\ell_e)_*(A) = (\ell_e)_*(B)\\ &\Leftrightarrow A = B,\end{aligned}$$

since $(\ell_e)_* = \mathrm{id}_{TG}$. Hence, the map j is injective.

v) Finally we nee to check that the map is surjective. Let $X \in \mathcal{L}(G)$. Define $A^X := X|_e \in T_eG$. Then, we have

$$j(A^X)|_g = (\ell_g)_*(A^X) = (\ell_g)_*(X|_e) = X_{ge} = X_g,$$

since X is left-invariant. Hence $X = j(A^X)$ and thus j is surjective.

Therefore, $j \colon T_eG \xrightarrow{\sim} \mathcal{L}(G)$ is indeed a linear isomorphism. □

Corollary 3.1. *It is* $\dim \mathcal{L}(G) = \dim G$*, hence* $\mathcal{L}(G)$ *turns out to be a finite-dimensional vector space over* $\mathbb{R}$ *(as an* $\mathbb{R}$*-vector subspace of* $\Gamma(TG)$*).*

So we proved that indeed $\mathcal{L}(G)$ is a finite-dimensional vector space, but as we said there is something more important here. Namely, since $j\colon T_eG \xrightarrow{\sim} \mathcal{L}(G)$ is a linear isomorphism this means that the spaces $\mathcal{L}(G)$ and T_eG are isomorphic which with its turn it means that the map j, first of all has an inverse, and most importantly maps each element of $\mathcal{L}(G)$ to exactly one element of T_eG. In other words there are as many left invariant fields as there are tangent vectors to the Lie group at the identity, so instead of studying the (quite complicated) left invariant vector fields we can simply study the (quite simpler) tangent vectors at the identity.

However, we can go one step further now. Recall from the Lie algebra chapter in the notes, that a Lie algebra over an algebraic field K is a vector space over K equipped with a Lie bracket $[-,-]$, i.e. a K-bilinear, antisymmetric map which satisfies the Jacobi identity.

Considering $\Gamma(TM)$ as an infinite-dimensional R-vector space, for two vector fields $X, Y \in \Gamma(TM)$, we can define their Lie bracket, or commutator, as

$$[X,Y](f) := X(Y(f)) - Y(X(f))$$

for any $f \in \mathcal{C}^\infty(M)$. Now we can check that indeed $[X,Y] \in \Gamma(TM)$, and that the bracket is $\mathbb{R}$-bilinear, antisymmetric and satisfies the Jacobi identity. Thus, $(\Gamma(TM), +, \cdot, [-,-])$ is an infinite-dimensional Lie algebra over $\mathbb{R}$. We suppress the $+$ and $\cdot$ when they are clear from the context.

Since $\mathcal{L}(G)$ is a subvector space (and a submodule) of $\Gamma(TM)$ we can inherit the commutator to $\mathcal{L}(G)$ and ask if $\mathcal{L}(G)$ is closed under the commutator so $(\mathcal{L}(G), +, \cdot, [-,-])$ is a subalgebra of $(\Gamma(TG), +, \cdot, [-,-])$. Indeed, this is the case.

Theorem 3.2. *Let G be a Lie group. Then $\mathcal{L}(G)$ is a Lie subalgebra of $\Gamma(TG)$.*

Proof. A Lie subalgebra of a Lie algebra is simply a vector subspace which is closed under the action of the Lie bracket. Therefore, we only need to check that

$$\forall\, X, Y \in \mathcal{L}(G) :\ [X,Y] \in \mathcal{L}(G).$$

Let $X, Y \in \mathcal{L}(G)$. For any $g \in G$ and $f \in \mathcal{C}^\infty(G)$, we have

$$\begin{aligned}
[X,Y](f \circ \ell_g) &:= X(Y(f \circ \ell_g)) - Y(X(f \circ \ell_g)) \\
&= X(Y(f) \circ \ell_g) - Y(X(f) \circ \ell_g) \\
&= X(Y(f)) \circ \ell_g - Y(X(f)) \circ \ell_g \\
&= \big(X(Y(f)) - Y(X(f))\big) \circ \ell_g \\
&= [X,Y](f) \circ \ell_g.
\end{aligned}$$

Hence, $[X,Y]$ is left-invariant. $\square$

To summarise, we began with $\mathcal{L}(G)$ as a set of all left invariant vector fields of G, which is a subset of $\Gamma(TG)$, then we inherited the $+$ and $\cdot$ of $\Gamma(TG)$ to $\mathcal{L}(G)$ and we showed that it is also a submodule and a subvector space of $\Gamma(TG)$, and finally we inherited the Lie bracket from $\Gamma(TG)$ and we showed that it is also a subalgebra of $\Gamma(TG)$. From now on when we will be referring to $\mathcal{L}(G)$, we will mean its algebra structure.

Definition 3.8 (The Lie Algebra Of A Lie Group)**.** *Let G be a Lie group. We call the algebra $\mathcal{L}(G)$ of all left invariant vector fields of G the* ***Lie algebra of the Lie group*** *G.*

We already showed that the underlying vector space of the Lie algebra of a Lie group $\mathcal{L}(G)$ is isomorphic to the tangent vector of the Lie group G at the identity T_eG. We will now see that the identification of $\mathcal{L}(G)$ and T_eG goes beyond the level of linear isomorphism as vector spaces, as they are isomorphic as algebras. Indeed, we can use the bracket on $L(G)$ to define a bracket on T_eG such that they be isomorphic as algebras.

Recall from the Lie algebras chapter in the notes that two algebras are called isomorphic if there exists an isomorphism between them, a.k.a a bijective map ϕ such that

$$\forall\, x, y \in L_1 :\ \phi([x,y]_{L_1}) = [\phi(x), \phi(y)]_{L_2}.$$

Well, we already have a bijective map j so by using the bracket $[-,-]_{\mathcal{L}(G)}$ on $\mathcal{L}(G)$ we can define, for any $A, B \in T_eG$

$$[A,B]_{T_eG} := j^{-1}\big([j(A), j(B)]_{\mathcal{L}(G)}\big),$$

where $j^{-1}(X) = X|_e$. Equipped with these brackets, we have

$$\mathcal{L}(G) \cong_{\text{Lie alg}} T_eG.$$

Hence, given a Lie group we have seen how we can construct its corresponding Lie algebra as the space of left-invariant vector fields and we also showed that this algebra is isomorphic to the algebra of tangent vectors at the identity. We will later explore the opposite direction, i.e. given a Lie algebra, we will see how to construct a Lie group whose associated Lie algebra is the one we started from.

3.4 Application - Part 2: $\mathrm{SL}(2,\mathbb{C})$

In the first part of of the application in the previous chapter, we defined the set $\mathrm{SL}(2,\mathbb{C})$ as a subset of $\mathbb{C}^4 := \mathbb{C}\times\mathbb{C}\times\mathbb{C}\times\mathbb{C}$. Then we showed that:

- $\mathrm{SL}(2,\mathbb{C})$ can be made into a group
- $\mathrm{SL}(2,\mathbb{C})$ can be made into a topological space
- $\mathrm{SL}(2,\mathbb{C})$ can be made into a topological manifold
- $\mathrm{SL}(2,\mathbb{C})$ can be made into a complex differentiable manifold

Hence we have left with $\mathrm{SL}(2,\mathbb{C})$ as a 3-dimensional, complex differentiable manifold.

$\mathrm{SL}(2,\mathbb{C})$ **As A Lie Group**

We equipped $\mathrm{SL}(2,\mathbb{C})$ with both a group and a manifold structure. In order to obtain a Lie group structure, we have to check that these two structures are compatible, that is, we have to show that the two maps

$$\begin{aligned}\mu\colon \mathrm{SL}(2,\mathbb{C})\times\mathrm{SL}(2,\mathbb{C}) &\to \mathrm{SL}(2,\mathbb{C})\\ \Big(\begin{pmatrix}a & b\\ c & d\end{pmatrix}, \begin{pmatrix}e & f\\ g & h\end{pmatrix}\Big) &\mapsto \begin{pmatrix}a & b\\ c & d\end{pmatrix}\bullet\begin{pmatrix}e & f\\ g & h\end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}i\colon \mathrm{SL}(2,\mathbb{C}) &\to \mathrm{SL}(2,\mathbb{C})\\ \begin{pmatrix}a & b\\ c & d\end{pmatrix} &\mapsto \begin{pmatrix}a & b\\ c & d\end{pmatrix}^{-1}\end{aligned}$$

are differentiable with respect to the differentiable structure on $\mathrm{SL}(2,\mathbb{C})$.

For instance, for the inverse map i, we have to show that the map $y\circ i\circ x^{-1}$ is differentiable in the usual for any pair of charts $(U,x),(V,y)\in\mathscr{A}$.

$$\begin{tikzcd}
U \subseteq \mathrm{SL}(2,\mathbb{C}) \arrow[r, "i"] \arrow[d, "x"] & V \subseteq \mathrm{SL}(2,\mathbb{C}) \arrow[d, "y"] \\
x(U) \subseteq \mathbb{C}^3 \arrow[r, "y\circ i\circ x^{-1}"] & y(V) \subseteq \mathbb{C}^3
\end{tikzcd}$$

where we remind that

$$\begin{aligned}x^{-1}\colon\quad x(U) &\to U\\ (a,b,c) &\mapsto \begin{pmatrix}a & b\\ c & \frac{1+bc}{a}\end{pmatrix}.\end{aligned}$$

and

$$\begin{aligned} y\colon \qquad & V \to x(V) \subseteq \mathbb{C} \times \mathbb{C}^* \times \mathbb{C} \\ & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a, b, d). \end{aligned}$$

Since $\mathrm{SL}(2,\mathbb{C})$ is connected, the differentiability of the transition maps in $\mathscr{A}$ implies that if $y \circ i \circ x^{-1}$ is differentiable for any two given charts, then it is differentiable for all charts in $\mathscr{A}$. Hence, we can simply let (U, x) and (V, y) be the two charts on $\mathrm{SL}(2,\mathbb{C})$ defined above. Then, we have

$$(y \circ i \circ x^{-1})(a, b, c) = (y \circ i)(\begin{pmatrix} a & b \\ c & \frac{1+bc}{a} \end{pmatrix}) = y(\begin{pmatrix} \frac{1+bc}{a} & -b \\ -c & a \end{pmatrix}) = (\tfrac{1+bc}{a}, -b, a)$$

which is certainly complex differentiable as a map between open subsets of $\mathbb{C}^3$ (recall that $a \neq 0$ on $x(U)$). We have to do the whole process again for $x \circ i \circ y^{-1}$ and show that is complex differentiable (which it is), hence we conclude that indeed the map i is is complex differentiable.

Checking that μ is complex differentiable is slightly more involved, since we first have to equip $\mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C})$ with a suitable "product differentiable structure" and then proceed as above. Once that is done, we can finally conclude that $((\mathrm{SL}(2,\mathbb{C}), \mathcal{O}, \mathscr{A}), \bullet)$ is a 3-dimensional complex Lie group.

The Lie Algebra Of The Lie Group $\mathrm{SL}(2,\mathbb{C})$

Recall that to every Lie group G, there is an associated Lie algebra $\mathcal{L}(G)$ of all left invariant vector fields of G a.k.a

$$\mathcal{L}(G) := \{X \in \Gamma(TG) \mid \forall\, g, h \in G : (\ell_g)_*(X|_h) = X|_{gh}\},$$

where the left translation map ℓ_g was given by

$$\ell_g(h) := g \bullet h \equiv gh$$

Coming to our case we have that $G = \mathrm{SL}(2,\mathbb{C})$ hence the corresponding Lie algebra of $\mathrm{SL}(2,\mathbb{C})$ usually denoted by small letters $sl(2,\mathbb{C})$ is

$$\mathfrak{sl}(2,\mathbb{C}) := \mathcal{L}(\mathrm{SL}(2,\mathbb{C})) := \{X \in \Gamma(T\,\mathrm{SL}(2,\mathbb{C})) \mid \forall\, g, h \in G : (\ell_g)_*(X|_h) = X|_{gh}\},$$

where the left translation map at a point $g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ of $\mathrm{SL}(2,\mathbb{C})$ is

$$\ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

Now, in order to find the Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ we need to find its structure constants. One can do that by computing the commutator (Lie bracket) relations of the underlying vector space. In this case, considering two vector fields $X, Y \in \Gamma(T\,\mathrm{SL}(2,\mathbb{C}))$

$$[X, Y] := X(Y) - Y(X)$$

However, as we proved earlier, we can always use the fact that the corresponding Lie algebra $\mathcal{L}(G)$ of a Lie group G is isomorphic to the Lie algebra T_eG with Lie bracket

$$[A, B]_{T_eG} := j^{-1}([j(A), j(B)]_{\mathcal{L}(G)})$$

induced by the Lie bracket on $\mathcal{L}(G)$ via the isomorphism j

$$j(A)|_g := (\ell_g)_*(A).$$

Hence, in our case instead the commutator in the algebra $\mathfrak{sl}(2,\mathbb{C})$ we can instead calculate it on $T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\mathrm{SL}(2,\mathbb{C})$, i.e

$$[A, B]_{T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\mathrm{SL}(2,\mathbb{C})} := j^{-1}([j(A), j(B)]_{\mathfrak{sl}(2,\mathbb{C})})$$

with

$$j(A)|_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} = \left(\ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}\right)_*(A)$$

First thing first, for any vector $A \in T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \mathrm{SL}(2,\mathbb{C})$ we have to compute $j(A)$.

Recall that if (U,x) is a chart on a manifold M and $p \in U$, then the chart (U,x) induces a basis of the tangent space T_pM. We shall use our previously defined chart (U,x) on $\mathrm{SL}(2,\mathbb{C})$, where $U := \{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{C}) \mid a \neq 0\}$ and

$$\begin{aligned} x\colon \qquad U &\to x(U) \subseteq \mathbb{C}^3 \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto (a,b,c). \end{aligned}$$

Note that the d appearing here is completely redundant, since the membership condition of $\mathrm{SL}(2,\mathbb{C})$ forces $d = \frac{1+bc}{a}$. However, we will keep writing the d to avoid having a fraction in a matrix in a subscript.

The chart (U,x) contains the identity $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ (we must include the identity since we are interested in the tangent space at the identity $T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \mathrm{SL}(2,\mathbb{C})$) and hence we get an induced co-ordinate basis

$$\left\{ \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \in T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \mathrm{SL}(2,\mathbb{C}) \;\middle|\; 1 \leq i \leq 3 \right\}$$

so that any $A \in T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \mathrm{SL}(2,\mathbb{C})$ can be written as

$$A = A^i \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} = \alpha \left(\frac{\partial}{\partial x^1}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} + \beta \left(\frac{\partial}{\partial x^2}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} + \gamma \left(\frac{\partial}{\partial x^3}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)},$$

for some $\alpha, \beta, \gamma \in \mathbb{C}$.

Let us now determine the image of these co-ordinate induced basis elements under the isomorphism j. The object

$$j\left(\left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right) \in \mathfrak{sl}(2,\mathbb{C})$$

is a left-invariant vector field on $\mathrm{SL}(2,\mathbb{C})$. It assigns to each point $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in U \subseteq \mathrm{SL}(2,\mathbb{C})$ the tangent vector

$$j\left(\left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right)\Bigg|_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} := \left(\ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}\right)_* \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \in T_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} \mathrm{SL}(2,\mathbb{C}).$$

This tangent vector is a $\mathbb{C}$-linear map $\mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C})) \xrightarrow{\sim} \mathbb{C}$, where $\mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C}))$ is the $\mathbb{C}$-vector space (in fact, the $\mathbb{C}$-algebra) of smooth complex-valued functions on $\mathrm{SL}(2,\mathbb{C})$ although, to be precise, since we are working in a chart we should only consider functions defined on U. For (the restriction to U of) any $f \in \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C}))$ by the definition of the push-forward we have, explicitly,

$$\begin{aligned} \left(\ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}\right)_* \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} (f) &= \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} \left(f \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}\right) \\ &= \partial_i \left(f \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} \circ x^{-1}\right)\left(x \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\right), \end{aligned}$$

where the argument of ∂_i in the last line is a map $x(U) \subseteq \mathbb{C}^3 \to \mathbb{C}$, hence ∂_i is simply the operation of complex differentiation with respect to the i-th (out of the 3) complex variable of the map $f \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} \circ x^{-1}$, which is then to be evaluated at $x \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathbb{C}^3$.

By inserting an identity in the composition, we have

$$\begin{aligned}
&= \partial_i\Big(f \circ \mathrm{id}_U \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1}\Big)(x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right))\\
&= \partial_i\Big(f \circ (x^{-1}\circ x) \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1}\Big)(x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right))\\
&= \partial_i\Big((f \circ x^{-1}) \circ (x \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})\Big)(x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right)),
\end{aligned}$$

where $f \circ x^{-1} \colon x(U) \subseteq \mathbb{C}^3 \to \mathbb{C}$ and $(x \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1}) \colon x(U) \subseteq \mathbb{C}^3 \to x(U) \subseteq \mathbb{C}^3$ and hence, we can use the multi-dimensional chain rule to obtain

$$= \Big(\partial_m(f\circ x^{-1})\big((x\circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})(x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right))\big)\Big)\Big(\partial_i(x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})(x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right))\Big),$$

with the summation going from $m = 1$ to $m = 3$. The first factor is simply

$$\begin{aligned}
\partial_m(f\circ x^{-1})\big((x\circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)})\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right)\big) &= \partial_m(f\circ x^{-1})(x\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right))\\
&=: \left(\frac{\partial}{\partial x^m}\right)_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)}(f).
\end{aligned}$$

To see what the second factor is, we first consider the map $x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1}$. This map acts on the triple $(e, f, g) \in x(U)$ as

$$\begin{aligned}
(x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})(e,f,g) &= (x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)})\begin{pmatrix} e & f \\ g & \frac{1+fg}{e}\end{pmatrix}\\
&= x^m(\begin{pmatrix} a & b \\ c & d\end{pmatrix} \bullet \begin{pmatrix} e & f \\ g & \frac{1+fg}{e}\end{pmatrix})\\
&= x^m(\begin{pmatrix} ae+bg & af+\frac{b(1+fg)}{e} \\ ce+dg & cf+\frac{d(1+fg)}{e}\end{pmatrix}),
\end{aligned}$$

and since $x^m := \mathrm{proj}_m \circ x$, with $m \in \{1,2,3\}$, we have

$$(x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})(e,f,g) = \mathrm{proj}_m(ae+bg, af+\tfrac{b(1+fg)}{e}, ce+dg),$$

the map proj_m simply picks the m-th component of the triple. We now have to apply ∂_i to this map, with $i \in \{1,2,3\}$, i.e. we have to differentiate with respect to each of the three complex variables e, f, and g. We can write the result as

$$\partial_i(x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})(e,f,g) = D(e,f,g)^m{}_i,$$

where m labels the rows and i the columns of the matrix

$$D(e,f,g) = \begin{pmatrix} a & 0 & b \\ -\frac{b(1+fg)}{e^2} & a+\frac{bg}{e} & \frac{bf}{e} \\ c & 0 & d\end{pmatrix}.$$

Finally, by evaluating this at $(e,f,g) = x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right) = (1,0,0)$, we obtain

$$\partial_i(x^m \circ \ell_{\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)} \circ x^{-1})(x\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1\end{smallmatrix}\right)) = D^m{}_i,$$

where, by recalling that $d = \frac{1+bc}{a}$,

$$D := D(1,0,0) = \begin{pmatrix} a & 0 & b \\ -b & a & 0 \\ c & 0 & \frac{1+bc}{a}\end{pmatrix}.$$

Putting the two factors back together yields

$$\left(\ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}\right)_* \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} (f) = D^m{}_i \left(\frac{\partial}{\partial x^m}\right)_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} (f).$$

Since this holds for an arbitrary $f \in \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C}))$, we have

$$j\left(\left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right)\Bigg|_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} := \left(\ell_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}\right)_* \left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)} = D^m{}_i \left(\frac{\partial}{\partial x^m}\right)_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)},$$

and since the point $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in U \subseteq \mathrm{SL}(2,\mathbb{C})$ is also arbitrary, we have

$$j\left(\left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right) = D^m{}_i \frac{\partial}{\partial x^m} \in \mathfrak{sl}(2,\mathbb{C}),$$

where D is now the corresponding matrix of co-ordinate functions

$$D := \begin{pmatrix} x^1 & 0 & x^2 \\ -x^2 & x^1 & 0 \\ x^3 & 0 & \frac{1+x^2x^3}{x^1} \end{pmatrix}.$$

Note that while the three vector fields

$$\begin{aligned} \frac{\partial}{\partial x^m} : \mathrm{SL}(2,\mathbb{C}) &\to T\,\mathrm{SL}(2,\mathbb{C}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \left(\frac{\partial}{\partial x^m}\right)_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)} \end{aligned}$$

are not individually left-invariant, their linear combination with coefficients $D^m{}_i$ is indeed left-invariant. Recall that these vector fields

i) are $\mathbb{C}$-linear maps

$$\begin{aligned} \frac{\partial}{\partial x^m} : \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C})) &\xrightarrow{\sim} \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C})) \\ f &\mapsto \partial_m(f \circ x^{-1}) \circ x; \end{aligned}$$

ii) satisfy the Leibniz rule

$$\frac{\partial}{\partial x^m}(fg) = f\frac{\partial}{\partial x^m}(g) + g\frac{\partial}{\partial x^m}(f);$$

iii) act on the coordinate functions $x^i \in \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C}))$ as

$$\frac{\partial}{\partial x^m}(x^i) = \partial_m(x^i \circ x^{-1}) \circ x = \partial_m(\mathrm{proj}_i \circ x \circ x^{-1}) \circ x = \delta^i_m \circ x = \delta^i_m,$$

since the composition of a constant function with any composable function is just the constant function.

Hence, we have an expansion of the images of the basis of $T_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\mathrm{SL}(2,\mathbb{C})$ under j:

$$\begin{aligned} j\left(\left(\frac{\partial}{\partial x^1}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right) &= x^1\frac{\partial}{\partial x^1} - x^2\frac{\partial}{\partial x^2} + x^3\frac{\partial}{\partial x^3} \\ j\left(\left(\frac{\partial}{\partial x^2}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right) &= x^1\frac{\partial}{\partial x^2} \\ j\left(\left(\frac{\partial}{\partial x^3}\right)_{\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)}\right) &= x^2\frac{\partial}{\partial x^1} + \tfrac{1+x^2x^3}{x^1}\frac{\partial}{\partial x^3}. \end{aligned}$$

We now have to calculate the bracket (in $\mathfrak{sl}(2,\mathbb{C})$) of every pair of these. We can also do them all at once, which is a good exercise in index gymnastics. We have

$$\left[j\left(\left(\frac{\partial}{\partial x^i}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\right), j\left(\left(\frac{\partial}{\partial x^k}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\right)\right] = \left[D^m{}_i\,\frac{\partial}{\partial x^m}, D^n{}_k\,\frac{\partial}{\partial x^n}\right].$$

Letting this act on an arbitrary $f \in \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C}))$, by definition

$$\left[D^m{}_i\,\frac{\partial}{\partial x^m}, D^n{}_k\,\frac{\partial}{\partial x^n}\right](f) := D^m{}_i\,\frac{\partial}{\partial x^m}\Big(D^n{}_k\,\frac{\partial}{\partial x^n}(f)\Big) - D^n{}_k\,\frac{\partial}{\partial x^n}\Big(D^m{}_i\,\frac{\partial}{\partial x^m}(f)\Big).$$

The first term gives

$$\begin{aligned}
D^m{}_i\,\frac{\partial}{\partial x^m}\Big(D^n{}_k\,\frac{\partial}{\partial x^n}(f)\Big) &= D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k\,\partial_n(f\circ x^{-1})\circ x)\\
&= D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k)\,(\partial_n(f\circ x^{-1})\circ x) + D^m{}_i D^n{}_k\,\frac{\partial}{\partial x^m}(\partial_n(f\circ x^{-1})\circ x)\\
&= D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k)\,(\partial_n(f\circ x^{-1})\circ x) + D^m{}_i D^n{}_k\,\partial_m(\partial_n(f\circ x^{-1})\circ x\circ x^{-1})\circ x\\
&= D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k)\,(\partial_n(f\circ x^{-1})\circ x) + D^m{}_i D^n{}_k\,\partial_m\partial_n(f\circ x^{-1})\circ x.
\end{aligned}$$

Similarly, we have

$$D^n{}_k\,\frac{\partial}{\partial x^n}\Big(D^m{}_i\,\frac{\partial}{\partial x^m}(f)\Big) = D^n{}_k\,\frac{\partial}{\partial x^n}(D^m{}_i)\,(\partial_m(f\circ x^{-1})\circ x) + D^n{}_k D^m{}_i\,\partial_n\partial_m(f\circ x^{-1})\circ x.$$

Hence, recalling that $\partial_m\partial_n = \partial_n\partial_m$ by Schwarz's theorem, we have

$$\begin{aligned}
\left[D^m{}_i\,\frac{\partial}{\partial x^m}, D^n{}_k\,\frac{\partial}{\partial x^n}\right](f) &= D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k)\,(\partial_n(f\circ x^{-1})\circ x) + [gray]D^m{}_i D^n{}_k\,\partial_m\partial_n(f\circ x^{-1})\circ x\\
&\qquad - D^n{}_k\,\frac{\partial}{\partial x^n}(D^m{}_i)\,(\partial_m(f\circ x^{-1})\circ x) - [gray]D^n{}_k D^m{}_i\,\partial_n\partial_m(f\circ x^{-1})\circ x\\
&= \Big(D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k) - D^m{}_k\,\frac{\partial}{\partial x^m}(D^n{}_i)\Big)\partial_n(f\circ x^{-1})\circ x\\
&= \Big(D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k) - D^m{}_k\,\frac{\partial}{\partial x^m}(D^n{}_i)\Big)\frac{\partial}{\partial x^n}(f),
\end{aligned}$$

where we relabelled some dummy indices. Since the $f \in \mathcal{C}^\infty(\mathrm{SL}(2,\mathbb{C}))$ was arbitrary,

$$\left[D^m{}_i\,\frac{\partial}{\partial x^m}, D^n{}_k\,\frac{\partial}{\partial x^n}\right] = \Big(D^m{}_i\,\frac{\partial}{\partial x^m}(D^n{}_k) - D^m{}_k\,\frac{\partial}{\partial x^m}(D^n{}_i)\Big)\frac{\partial}{\partial x^n}.$$

We can now evaluate this explicitly. For $i = 1$ and $k = 2$, we have

$$\begin{aligned}
\left[D^m{}_1\frac{\partial}{\partial x^m}, D^n{}_2\frac{\partial}{\partial x^n}\right] &= \Big([gray]D^m{}_1\frac{\partial}{\partial x^m}(D^1{}_2) - D^m{}_2\frac{\partial}{\partial x^m}(D^1{}_1)\Big)\frac{\partial}{\partial x^1}\\
&\qquad + \Big(D^m{}_1\frac{\partial}{\partial x^m}(D^2{}_2) - D^m{}_2\frac{\partial}{\partial x^m}(D^2{}_1)\Big)\frac{\partial}{\partial x^2}\\
&\qquad + \Big([gray]D^m{}_1\frac{\partial}{\partial x^m}(D^3{}_2) - D^m{}_2\frac{\partial}{\partial x^m}(D^3{}_1)\Big)\frac{\partial}{\partial x^3}\\
&= -D^1{}_2\frac{\partial}{\partial x^1} + (D^1{}_1 + D^2{}_2)\frac{\partial}{\partial x^2} - D^3{}_2\frac{\partial}{\partial x^3}\\
&= 2x^1\frac{\partial}{\partial x^2}.
\end{aligned}$$

Similarly, we compute

$$\begin{aligned}
\left[D^m{}_1\frac{\partial}{\partial x^m}, D^n{}_3\frac{\partial}{\partial x^n}\right] &= \Big(D^m{}_1\frac{\partial}{\partial x^m}(D^1{}_3) - D^m{}_3\frac{\partial}{\partial x^m}(D^1{}_1)\Big)\frac{\partial}{\partial x^1} \\
&\quad + \Big(\text{[gray]}D^m{}_1\frac{\partial}{\partial x^m}(D^2{}_3) - D^m{}_3\frac{\partial}{\partial x^m}(D^2{}_1)\Big)\frac{\partial}{\partial x^2} \\
&\quad + \Big(D^m{}_1\frac{\partial}{\partial x^m}(D^3{}_3) - D^m{}_3\frac{\partial}{\partial x^m}(D^3{}_1)\Big)\frac{\partial}{\partial x^3} \\
&= -2x^2\frac{\partial}{\partial x^1} - 2\big(\tfrac{1+x^2x^3}{x^1}\big)\frac{\partial}{\partial x^3}
\end{aligned}$$

and

$$\begin{aligned}
\left[D^m{}_2\frac{\partial}{\partial x^m}, D^n{}_3\frac{\partial}{\partial x^n}\right] &= \Big(D^m{}_2\frac{\partial}{\partial x^m}(D^1{}_3) - \text{[gray]}D^m{}_3\frac{\partial}{\partial x^m}(D^1{}_2)\Big)\frac{\partial}{\partial x^1} \\
&\quad + \Big(\text{[gray]}D^m{}_2\frac{\partial}{\partial x^m}(D^2{}_3) - D^m{}_3\frac{\partial}{\partial x^m}(D^2{}_2)\Big)\frac{\partial}{\partial x^2} \\
&\quad + \Big(D^m{}_2\frac{\partial}{\partial x^m}(D^3{}_3) - \text{[gray]}D^m{}_3\frac{\partial}{\partial x^m}(D^3{}_2)\Big)\frac{\partial}{\partial x^3} \\
&= (D^2{}_1 - D^1{}_3)\frac{\partial}{\partial x^1} + D^2{}_3\frac{\partial}{\partial x^2} - D^3{}_2\frac{\partial}{\partial x^3} \\
&= x^1\frac{\partial}{\partial x^1} - x^2\frac{\partial}{\partial x^2} + x^3\frac{\partial}{\partial x^3},
\end{aligned}$$

where the differentiation rules that we have used come from the definition of the vector field $\frac{\partial}{\partial x^m}$, the Leibniz rule, and the action on co-ordinate functions.
By applying j^{-1}, which is just evaluation at the identity, to these vector fields, we finally see that the induced Lie bracket on $T_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\operatorname{SL}(2,\mathbb{C})$ satisfies

$$\begin{aligned}
\left[\left(\frac{\partial}{\partial x^1}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}, \left(\frac{\partial}{\partial x^2}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\right] &= 2\left(\frac{\partial}{\partial x^2}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)} \\
\left[\left(\frac{\partial}{\partial x^1}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}, \left(\frac{\partial}{\partial x^3}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\right] &= -2\left(\frac{\partial}{\partial x^3}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)} \\
\left[\left(\frac{\partial}{\partial x^2}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}, \left(\frac{\partial}{\partial x^3}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\right] &= \left(\frac{\partial}{\partial x^1}\right)_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}.
\end{aligned}$$

Hence, the structure constants of $T_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\operatorname{SL}(2,\mathbb{C})$ with respect to the co-ordinate basis are

$$C^2{}_{12} = 2, \qquad C^3{}_{13} = -2, \qquad C^1{}_{23} = 1,$$

with all other being either zero or related to these by anti-symmetry.

3.4.1 The simplicity of $\mathfrak{sl}(2,\mathbb{C})$

We have seen that the non-zero structure constants of $T_{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)}\operatorname{SL}(2,\mathbb{C})$ are

$$C^2{}_{12} = 2, \qquad C^3{}_{13} = -2, \qquad C^1{}_{23} = 1,$$

plus those related by anti-symmetry.

Proposition 3.2. *Two Lie algebras A and B are isomorphic if, and only if, there exists a basis of A and a basis of B in which the structure constants of A and B are the same.*

Since we have already proved that $T_eG \cong_{\text{Lie alg}} \mathcal{L}(G)$ for any Lie group G, we can deduce the existence of a basis $\{X_1, X_2, X_2\}$ of $\mathfrak{sl}(2,\mathbb{C})$ with respect to which the structure constants are those listed above.

In other words, we have

$$
\begin{aligned}
[X_1, X_2] &= 2X_2, \\
[X_1, X_3] &= -2X_3, \\
[X_2, X_3] &= X_1.
\end{aligned}
$$

In this basis, the Killing form of $\mathfrak{sl}(2,\mathbb{C})$ has components

$$\kappa_{ij} = C^m{}_{in} C^n{}_{jm},$$

with all indices ranging from 1 to 3. Explicitly, we have

$$
\begin{aligned}
\kappa_{11} &= C^m{}_{1n} C^n{}_{1m} \\
&= \cancel{C^1{}_{1n} C^n{}_{11}} + C^2{}_{1n} C^n{}_{12} + C^3{}_{1n} C^n{}_{13} \\
&= C^2{}_{12} C^2{}_{12} + C^3{}_{13} C^3{}_{13} \\
&= 8.
\end{aligned}
$$

Since κ is symmetric, we only need to determine κ_{ij} for $i \leq j$. By writing the components in a 3×3 array, we find

$$[\kappa_{ij}] = \begin{pmatrix} 8 & 0 & 0 \\ 0 & -8 & 0 \\ 0 & 0 & 8 \end{pmatrix},$$

which is just shorthand for

$$\kappa(X_1, X_1) = 8, \qquad \kappa(X_2, X_2) = -8, \qquad \kappa(X_3, X_3) = 8,$$

and $\kappa(X_i, X_j) = 0$ whenever $i \neq j$.

Proposition 3.3. *The Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ is semi-simple.*

Proof. Since the diagonal entries of κ are all non-zero, the Killing form is non-degenerate. By Cartan's criterion, this implies that $\mathfrak{sl}(2,\mathbb{C})$ is semi-simple. □

Remark 3.1. There is one more thing that can be read off from the components of κ, namely, that it is an *indefinite* form, i.e. the sign of $\kappa(X, X)$ can be positive or negative depending on which $X \in \mathfrak{sl}(2,\mathbb{C})$ we pick.
A result from Lie theory states that the Killing form on the Lie algebra of a compact Lie group is always negative semi-definite, i.e. $\kappa(X, X)$ is always negative or zero, for all X in the Lie algebra. Hence, we can conclude that $\mathrm{SL}(2,\mathbb{C})$ is not a compact Lie group.

In fact, $\mathfrak{sl}(2,\mathbb{C})$ is more than just semi-simple.

Proposition 3.4. *The Lie algebra $\mathfrak{sl}(2,\mathbb{C})$ is simple.*

Recall that a Lie algebra is said to be simple if it contains no non-trivial ideals, and that an ideal I of a Lie algebra L is a Lie subalgebra of L such that

$$\forall\, x \in I : \forall\, y \in L : \; [x, y] \in I.$$

Proof. Consider the ideal of $\mathfrak{sl}(2,\mathbb{C})$

$$I := \{\alpha X_1 + \beta X_2 + \gamma X_3 \mid \alpha, \beta, \gamma \text{ restricted so that } I \text{ is an ideal}\}.$$

Since the bracket is bilinear, it suffices to check the result of bracketing an arbitrary element of I with each of the basis vectors of $\mathfrak{sl}(2,\mathbb{C})$. We find

$$
\begin{aligned}
[\alpha X_1 + \beta X_2 + \gamma X_3, X_1] &= -2\beta X_1 + 2\gamma X_3, \\
[\alpha X_1 + \beta X_2 + \gamma X_3, X_2] &= 2\alpha X_2 - \gamma X_1, \\
[\alpha X_1 + \beta X_2 + \gamma X_3, X_3] &= -2\alpha X_3 + \beta X_1.
\end{aligned}
$$

We need to choose α, β, γ so that the results always land back in I. Of course, we can choose $\alpha, \beta, \gamma \in \mathbb{C}$ and $\alpha = \beta = \gamma = 0$, which correspond respectively to the trivial ideals $\mathfrak{sl}(2, \mathbb{C})$ and 0. If none of α, β, γ is zero, then you can check that the right hand sides above are linearly independent, so that I contains three linearly independent vectors. Since the only n-dimensional subspace of an n-dimensional vector space is the vector space itself, we have $I = L$. Thus, we are left with the following cases:

i) if $\alpha = 0$, then $I \subseteq \operatorname{span}_{\mathbb{C}}(\{X_2, X_3\})$ and hence we must have $\beta = \gamma = 0$ as well;

ii) if $\beta = 0$, then $I \subseteq \operatorname{span}_{\mathbb{C}}(\{X_1, X_3\})$, hence we must have $\alpha = 0$, so that in fact $I \subseteq \operatorname{span}_{\mathbb{C}}(\{X_3\})$, and hence $\gamma = 0$ as well;

iii) if $\gamma = 0$, then $I \subseteq \operatorname{span}_{\mathbb{C}}(\{X_1, X_2\})$, hence we must have $\alpha = 0$, so that in fact $I \subseteq \operatorname{span}_{\mathbb{C}}(\{X_2\})$, and hence $\beta = 0$ as well.

In all cases, we have $I = 0$. Therefore, there are no non-trivial ideals of $\mathfrak{sl}(2, \mathbb{C})$. □

3.4.2 The roots and Dynkin diagram of $\mathfrak{sl}(2, \mathbb{C})$

By observing the bracket relations of the basis elements of $\mathfrak{sl}(2, \mathbb{C})$, we can see that

$$H := \operatorname{span}_{\mathbb{C}}(\{X_1\})$$

is a Cartan subalgebra of $\mathfrak{sl}(2, \mathbb{C})$. Indeed, for any $h \in H$, there exists a $\xi \in \mathbb{C}$ such that $h = \xi X_1$, and hence we have

$$\begin{aligned} \operatorname{ad}(h) X_2 &= \xi [X_1, X_2] = 2\xi X_2, \\ \operatorname{ad}(h) X_3 &= \xi [X_1, X_3] = -2\xi X_3. \end{aligned}$$

Recall that in the section on Lie algebras, we re-interpreted these eigenvalue equations in terms of functionals $\lambda_2, \lambda_3 \in H^*$

$$\begin{aligned} \lambda_2 \colon \quad H &\xrightarrow{\sim} \mathbb{C} \\ \xi X_1 &\mapsto 2\xi, \end{aligned} \qquad\qquad \begin{aligned} \lambda_3 \colon \quad H &\xrightarrow{\sim} \mathbb{C} \\ \xi X_1 &\mapsto -2\xi \end{aligned}$$

whereby

$$\begin{aligned} \operatorname{ad}(h) X_2 &= \lambda_2(h) X_2, \\ \operatorname{ad}(h) X_3 &= \lambda_3(h) X_3. \end{aligned}$$

Then, λ_2 and λ_3 are called the roots of $\mathfrak{sl}(2, \mathbb{C})$, so that the root set is $\Phi = \{\lambda_2, \lambda_3\}$. Of course, we are mainly interested in a subset $\Pi \subset \Phi$ of fundamental roots, which satisfies

i) Π is a linearly independent subset of H^*;

ii) for any $\lambda \in \Phi$, we have $\lambda \in \operatorname{span}_{\epsilon, \mathbb{N}}(\Pi)$.

We can choose $\Pi := \{\lambda_2\}$, even though $\Pi := \{\lambda_3\}$ would work just as well. Since $|\Pi| = 1$, the Weyl group is generated by the single Weyl transformation

$$\begin{aligned} s_{\lambda_2} \colon H^*_{\mathbb{R}} &\to H^*_{\mathbb{R}} \\ \mu &\mapsto \mu - 2 \frac{\kappa^*(\lambda_2, \mu)}{\kappa^*(\lambda_2, \lambda_2)} \lambda_2. \end{aligned}$$

Recall that we can recover the entire root set Φ by acting on the fundamental roots with Weyl transformations. Indeed, we have

$$s_{\lambda_2}(\lambda_2) = \lambda_2 - 2 \frac{\kappa^*(\lambda_2, \lambda_2)}{\kappa^*(\lambda_2, \lambda_2)} \lambda_2 = \lambda_2 - 2\lambda_2 = -\lambda_2 = \lambda_3,$$

as expected. Since there is only one fundamental root, the Cartan matrix is actually just a 1×1 matrix. Its only entry is a diagonal entry, and since $\mathfrak{sl}(2, \mathbb{C})$ is simple, we have

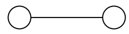
$$C = (2).$$

The Dynkin diagram of $\mathfrak{sl}(2,\mathbb{C})$ is simply

Hence, with reference to the Cartan classification, we have $A_1 = \mathfrak{sl}(2,\mathbb{C})$.

3.4.3 Reconstruction of A_2 from its Dynkin diagram

We have seen an example of how to construct the Dynkin diagram of a Lie algebra, albeit the simplest of this kind. Let us now consider the opposite direction. We will start from the Dynkin diagram



We immediately see that we have two fundamental roots, i.e. $\Pi = \{\pi_1, \pi_2\}$, since there are two circles in the diagram. The bond number is $n_{12} = 1$, so the two fundamental roots have the same length. Moreover, by definition

$$1 = n_{12} = C_{12}C_{21}$$

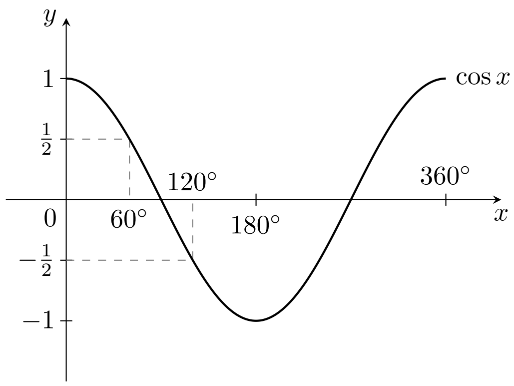
and since the off-diagonal entries of the Cartan matrix are non-positive integers, the only possibility is $C_{12} = C_{21} = -1$, so that we have

$$C = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

To determine the angle φ between π_1 and π_2, recall that

$$1 = n_{12} = 4\cos^2\varphi,$$

and hence $|\cos\varphi| = \frac{1}{2}$. There are two solutions, namely $\varphi = 60°$ and $\varphi = 120°$.



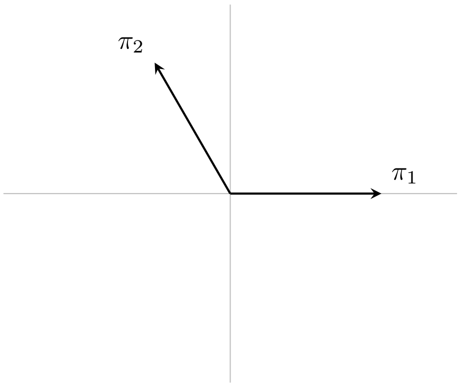
By definition, we have

$$\cos\varphi = \frac{\kappa^*(\pi_1,\pi_2)}{|\pi_1|\,|\pi_2|},$$

and therefore

$$0 > C_{12} = 2\frac{\kappa^*(\pi_1,\pi_2)}{\kappa^*(\pi_1,\pi_1)} = 2\frac{|\pi_1|\,|\pi_2|\cos\varphi}{\kappa^*(\pi_1,\pi_1)} = 2\frac{|\pi_2|}{|\pi_1|}\cos\varphi.$$

It follows that $\cos\varphi < 0$, and hence $\varphi = 120°$. We can thus plot the two fundamental roots in a plane as follows.



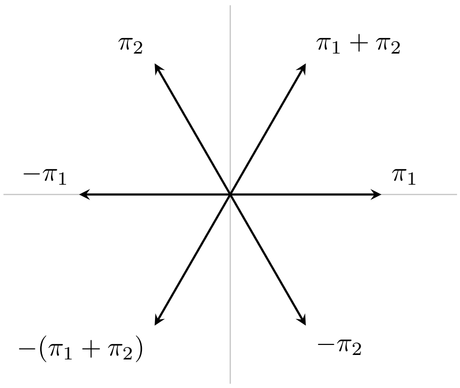
We can determine all the other roots in Φ by repeated action of the Weyl group. For instance, we easily find that $s_{\pi_1}(\pi_1) = -\pi_1$ and $s_{\pi_2}(\pi_2) = -\pi_2$. We also have

$$s_{\pi_1}(\pi_2) = \pi_2 - 2\frac{\kappa^*(\pi_1, \pi_2)}{\kappa^*(\pi_1, \pi_1)}\pi_1 = \pi_2 - 2(-\tfrac{1}{2})\pi_1 = \pi_1 + \pi_2.$$

Finally, we have $s_{\pi_1+\pi_2}(\pi_1 + \pi_2) = -(\pi_1 + \pi_2)$. Any further action by Weyl transformations simply permutes these roots. Hence, we have

$$\Phi = \{\pi_1, -\pi_1, \pi_2, -\pi_2, \pi_1 + \pi_2, -(\pi_1 + \pi_2)\}$$

and these are all the roots.



Since $H^* = \mathrm{span}_{\mathbb{C}}(\Pi)$, we have $\dim H^* = 2$, thus the dimension of the Cartan subalgebra is also 2. Since $|\Phi| = 6$, we know that any Cartan-Weyl basis of the Lie algebra A_2 must have $2+6 = 8$ elements. Hence, the dimension of A_2 is 8.

To complete our reconstruction of A_2, we would now like to understand how its bracket behaves. This amounts to finding its structure constants. Note that since $\dim A_2 = 8$, the structure constants ${C^k}_{ij}$ consist of $8^3 = 512$ complex numbers (not all unrelated, of course).

Denote by $\{h_1, h_2, e_3, \ldots, e_8\}$ a Cartan-Weyl basis of A_2, so that $H = \mathrm{span}_{\mathbb{C}}(\{h_1, h_2\})$ and the e_α are eigenvectors of every $h \in H$. Since A_2 is simple, H is abelian and hence

$$[h_1, h_2] = 0 \quad \Rightarrow \quad {C^k}_{12} = {C^k}_{21} = 0, \quad \forall\, 1 \leq k \leq 8.$$

To each e_α, for $3 \leq \alpha \leq 8$, there is an associated $\lambda_\alpha \in \Phi$ such that

$$\forall\, h \in H: \ \mathrm{ad}(h)e_\alpha = \lambda_\alpha(h)e_\alpha.$$

In particular, for the basis elements h_1, h_2,

$$[h_1, e_\alpha] = \mathrm{ad}(h_1)e_\alpha = \lambda_\alpha(h_1)e_\alpha,$$
$$[h_2, e_\alpha] = \mathrm{ad}(h_2)e_\alpha = \lambda_\alpha(h_2)e_\alpha,$$

so that we have

$$\begin{aligned}
{C^1}_{1\alpha} = {C^2}_{1\alpha} = 0, \quad {C^\alpha}_{1\alpha} = \lambda_\alpha(h_1), \quad \forall\, 3 \le \alpha \le 8, \\
{C^1}_{2\alpha} = {C^2}_{2\alpha} = 0, \quad {C^\alpha}_{2\alpha} = \lambda_\alpha(h_2), \quad \forall\, 3 \le \alpha \le 8.
\end{aligned}$$

Finally, we need to determine $[e_\alpha, e_\beta]$. By using the Jacobi identity, we have

$$\begin{aligned}
[h_i, [e_\alpha, e_\beta]] &= -[e_\alpha, [e_\beta, h_i]] - [e_\beta, [h_i, e_\alpha]] \\
&= -[e_\alpha, -\lambda_\beta(h_i)e_\beta] - [e_\beta, \lambda_\alpha(h_i)e_\alpha] \\
&= \lambda_\beta(h_i)[e_\alpha, e_\beta] + \lambda_\alpha(h_i)[e_\alpha, e_\beta] \\
&= (\lambda_\alpha(h_i) + \lambda_\beta(h_i))[e_\alpha, e_\beta],
\end{aligned}$$

that is,

$$\mathrm{ad}(h_i)[e_\alpha, e_\beta] = (\lambda_\alpha(h_i) + \lambda_\beta(h_i))[e_\alpha, e_\beta].$$

If $\lambda_\alpha + \lambda_\beta \in \Phi$, we have $[e_\alpha, e_\beta] = \xi e_\gamma$ for some $3 \le \gamma \le 8$ and $\xi \in \mathbb{C}$. Let us label the roots in our previous plot as

λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
π_1	π_2	$\pi_1 + \pi_2$	$-\pi_1$	$-\pi_2$	$-(\pi_1 + \pi_2)$

Then, for example

$$\mathrm{ad}(h)[e_3, e_4] = (\pi_1 + \pi_2)(h)[e_3, e_4],$$

and hence $[e_3, e_4]$ is an eigenvector of $\mathrm{ad}(h)$ with eigenvalues $(\pi_1 + \pi_2)(h)$. But so is e_5! Hence, we must have $[e_3, e_4] = \xi e_5$ for some $\xi \in \mathbb{C}$. Similarly, $[e_5, e_7] = \xi e_3$, and so on.
If $\lambda_\alpha + \lambda_\beta \notin \Phi$, then in order for the equation above to hold, we must have either $[e_\alpha, e_\beta] = 0$ (so both sides are zero), or $\lambda_\alpha(h) + \lambda_\beta(h) = 0$ for all h, i.e. $\lambda_\alpha + \lambda_\beta = 0$ as a functional. In the latter case, we must have $[e_\alpha, e_\beta] \in H$. This follows from a stronger version of the maximality property of the Cartan subalgebra H of a simple Lie algebra L, namely that

$$\big(\forall\, h \in H : [h, x] = 0\big) \Rightarrow x \in H.$$

Summarising, we have

$$[e_\alpha, e_\beta] = \begin{cases} \xi e_\gamma & \text{if } \lambda_\alpha + \lambda_\beta \in \Phi \\ \in H & \text{if } \lambda_\alpha + \lambda_\beta = 0 \\ 0 & \text{otherwise} \end{cases}$$

and these relations con be used to determine the remaining structure constants of A_2.