Lecture Notes

Geometrical Anatomy of Theoretical Physics

Dr. Fredric P Schuller*

Piyush Kumar Singh †

December 2024

*YouTube playlist

†iamPiyushKrSingh.github.io

Structure of this Course

In theoretical physics we mainly deal with three big part of mathematics, namely, ‘Analysis’, ‘Algebra’ and ‘Geometry’. And at the mutual intersection of these three, we have different branches of physics like ‘Quantum Mechanics’, ‘General Relativity’, ‘Statistical Mechanics’, etc. see fig. 0.1.

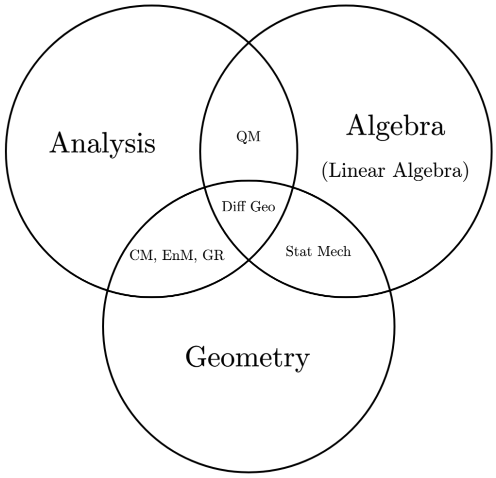

Figure 0.1: Structure of Theoretical Physics

This course mainly focuses on differential geometry and topology, and their applications in theoretical physics. And we will start with the basic propositional logic and set theory, and then move to the topology and geometry of manifolds, and then we will see the applications of these in physics. Following is the structure of this course:

- Logic
- Set Theory
- Topology
- Topological Manifolds
- Differential Manifolds
- Bundles
- Geometry: Symplectic Geometry, Metric Geometry, etc.
- Physics: Classical Mechanics, Electrodynamics, Quantum Mechanics, Statistical Mechanics, Special and General Relativity, etc.

Contents

1 Differentiable Manifolds

Lecture 7

§1.1 Adding Structure by refining the (maximal) $\mathcal{C}^0$-atlas

Definition 1.1 (❀-Atlas):
Let $(M, \mathcal{O})$ be a d-dimensional manifold. An atlas $\mathscr{A}$ is called ❀-atlas, if any two charts $(U, x), (V, y) \in \mathscr{A}$ are ❀-compatible.

In other words, either $U \cap V = \emptyset$ or if $U \cap V \neq \emptyset$, then the transition map $y \circ x^{-1}$ is ❀ as a map from $\mathbb{R}^d$ to $\mathbb{R}^d$.

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

Now, we can define the placeholder symbol ❀ as:

- ❀ $= \mathcal{C}^0$: see **??**.
- ❀ $= \mathcal{C}^k$: the transition map is k-times continuously differentiable as maps $\mathbb{R}^d \to \mathbb{R}^d$.
- ❀ $= \mathcal{C}^\infty$: the transition map is smooth (infinitely many times differentiable); i.e., k-times continuously differentiable for all $k \in \mathbb{N}$.
- ❀ $= \mathcal{C}^\omega$: the transition map is real-analytic; *i.e.*, it can be locally represented by a convergent power series.
- ❀ $= \mathcal{C}^\omega_\mathbb{C}$: the transition map is complex-analytic; equivalently, it satisfies the Cauchy-Riemann conditions.

Here for completeness, we need to define what are the Cauchy-Riemann conditions:
Set theoretical we know that $\mathbb{C} \cong_{\text{set}} \mathbb{R}^2$. Let $f : \mathbb{C} \to \mathbb{C}$ be a complex function defined as

$$\begin{aligned} f : \mathbb{C} &\to \mathbb{C} \\ x + \mathrm{i}y &\mapsto u(x, y) + \mathrm{i}v(x, y) \end{aligned} \tag{1.1}$$

where $u, v : \mathbb{R}^2 \to \mathbb{R}$ are real-valued functions. Then the Cauchy-Riemann conditions says that f is complex-analytic at $x_0 + \mathrm{i}y_0$ if and only if the following two conditions are satisfied:

1. All the partial derivatives of u and v exist at (x_0, y_0) and are continuous in a neighborhood of (x_0, y_0).
2. The following two equations are satisfied:

$$\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \quad \text{and} \quad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0). \tag{1.2}$$

Theorem 1.2 (Whitney)
Any maximal $\mathcal{C}^k$-atlas (for any $k \geq 1$) contains a $\mathcal{C}^\infty$-atlas. Moreover, any two maximal $\mathcal{C}^k$-atlases that contains the same $\mathcal{C}^\infty$-atlas are identical.

In other words, we refine (or remove) all the charts in a maximal $\mathcal{C}^k$-atlas which are not $\mathcal{C}^\infty$-compatible, and we get a maximal $\mathcal{C}^\infty$-atlas. This is the reason why we can always work with $\mathcal{C}^\infty$-atlases given that we are working with a differentiable manifold. Immediate consequence of this theorem is that if any result is true for a $\mathcal{C}^k$-atlas for any $k \geq 1$, then it is also true for a $\mathcal{C}^\infty$-atlas.

But in the case of $\mathcal{C}^0$-atlas, it may happen that it doesn't admit a $\mathcal{C}^1$-atlas, and hence we cannot refine it to a $\mathcal{C}^\infty$-atlas.

Hence, we will not make any distinction between $\mathcal{C}^k$-manifolds and $\mathcal{C}^\infty$-manifolds, and we will always work with $\mathcal{C}^\infty$-manifolds.

Definition 1.3 ($\mathcal{C}^k$-manifold):
A triple $(M, \mathcal{O}, \mathscr{A})$ is called a $\mathcal{C}^k$-manifold where

- *$(M, \mathcal{O})$ is a topological manifold.*
- *$\mathscr{A}$ is a maximal $\mathcal{C}^k$-atlas on M.*

Definition 1.4 (Incompatible Atlases):
Let two ❀-compactible atlases $\mathscr{A}_1$ and $\mathscr{A}_2$ on a topological manifold $(M, \mathcal{O})$ be called compatible if $\mathscr{A}_1 \cup \mathscr{A}_2$ is a ❀-atlas on M. Otherwise, they are called incompatible.

Remark 1.5. A given topological manifold $(M, \mathcal{O})$ can have different incompatible atlases.

A simple example of incompatible atlases,

Example 1.6
Let $M = \mathbb{R}$ with the standard topology, and let $\mathscr{A}_1 = \{(\mathbb{R}, \mathrm{id}_\mathbb{R})\}$ and $\mathscr{A}_2 = \{(\mathbb{R}, a \overset{x}{\mapsto} \sqrt[3]{a})\}$. Since each atlas contains only one chart, they are trivially $\mathcal{C}^\infty$-compatible as the transition map is the identity map in both cases. But $\mathscr{A}_1 \cup \mathscr{A}_2$ is not a $\mathcal{C}^\infty$-atlas on M because the transition maps $\mathrm{id}_\mathbb{R} \circ x^{-1} \equiv a \mapsto a^3$ which is a smooth map, but the transition map $x \circ \mathrm{id}_\mathbb{R}^{-1} \equiv a \mapsto \sqrt[3]{a}$ is not smooth as it is not differentiable at $a = 0$. Hence, $\mathscr{A}_1$ and $\mathscr{A}_2$ are incompatible atlases on M.

This example shows that we can equip the real line $\mathbb{R}$ with different incompatible $\mathcal{C}^\infty$-struictures. This sounds bad as we want to do physics on $\mathbb{R}$, and we want to have a unique $\mathcal{C}^\infty$-structure on it. But this is not a problem, as we are given an atlas by the definition of differentiable manifold.

Definition 1.7 (Differentiability):
Let $(M, \mathcal{O}_M, \mathscr{A}_M)$ and $(N, \mathcal{O}_N, \mathscr{A}_N)$ be two $\mathcal{C}^k$-manifolds of dimension m and n respectively. A map $f : M \to N$ is called $\mathcal{C}^k$-differentiable at a point $p \in M$ if there exists a chart $(U, x) \in \mathscr{A}_M$ around p and a chart $(V, y) \in \mathscr{A}_N$ around $f(p)$ such that the map $(y \circ f \circ x^{-1}) : x(U) \to y(V)$ is $\mathcal{C}^k$-differentiable at $x(p)$ as a map from $\mathbb{R}^m$ to $\mathbb{R}^n$.

$$\begin{array}{ccc} M \supseteq U & \xrightarrow{\quad f \quad} & N \subseteq V \\ \downarrow{\scriptstyle x} & & \downarrow{\scriptstyle y} \\ \mathbb{R}^m \supseteq x(U) & \xrightarrow{y \circ f \circ x^{-1}} & y(V) \subseteq \mathbb{R}^n \end{array}$$

If f is $\mathcal{C}^k$-differentiable at every point $p \in M$, then we say that f is a $\mathcal{C}^k$-differentiable map from M to N.

Proposition 1.8
The definition of $\mathcal{C}^k$-differentiability is independent of the choice of charts $(U,x) \in \mathscr{A}_M$ and $(V,y) \in \mathscr{A}_N$ *i.e.* the definition is well-defined.

Proof:
Consider two charts $(U,x),(\tilde{U},\tilde{x}) \in \mathscr{A}_M$ around p and two charts $(V,y),(\tilde{V},\tilde{y}) \in \mathscr{A}_N$ around $f(p)$. We need to show that if f is $\mathcal{C}^k$-differentiable at p with respect to the charts (U,x) and (V,y), then it should also be $\mathcal{C}^k$-differentiable with respect to the charts $(\tilde{U},\tilde{x})$ and $(\tilde{V},\tilde{y})$. Consider the following commutative diagram:

$$\begin{tikzcd}
\mathbb{R}^m \supseteq \tilde{x}(U\cap\tilde{U}) \arrow[r, "\tilde{y}\circ f\circ\tilde{x}^{-1}"] \arrow[dd, bend right=60, "x\circ\tilde{x}^{-1}"'] & \tilde{y}(V\cap\tilde{V}) \subseteq \mathbb{R}^n \\
M \supseteq U\cap\tilde{U} \ni p \arrow[r, "f"] \arrow[u, "\tilde{x}"] \arrow[d, "x"] & f(p) \in V\cap\tilde{V} \subseteq N \arrow[u, "\tilde{y}"'] \arrow[d, "y"'] \\
\mathbb{R}^m \supseteq x(U\cap\tilde{U}) \arrow[r, "y\circ f\circ x^{-1}"] & y(V\cap\tilde{V}) \subseteq \mathbb{R}^n \arrow[uu, bend right=60, "\tilde{y}\circ y^{-1}"']
\end{tikzcd}$$

We know that the transition maps $x\circ\tilde{x}^{-1}$ and $\tilde{y}\circ y^{-1}$ are $\mathcal{C}^k$-differentiable as they are transition maps between charts in the same atlas. So the composition of the maps

$$\tilde{y}\circ f\circ\tilde{x}^{-1} = (\tilde{y}\circ y^{-1})\circ(y\circ f\circ x^{-1})\circ(x\circ\tilde{x}^{-1}) \tag{1.3}$$

is also $\mathcal{C}^k$-differentiable as a composition of $\mathcal{C}^k$-differentiable maps. Hence, f is $\mathcal{C}^k$-differentiable at p with respect to the charts $(\tilde{U},\tilde{x})$ and $(\tilde{V},\tilde{y})$. Q.E.D.

Definition 1.9 (Diffeomorphism):
Let $(M,\mathcal{O}_M,\mathscr{A}_M)$ and $(N,\mathcal{O}_N,\mathscr{A}_N)$ be two $\mathcal{C}^k$-manifolds. A map $f: M\to N$ is called a $\mathcal{C}^k$-diffeomorphism if it is a bijection and both f and its inverse $f^{-1}: N\to M$ are $\mathcal{C}^k$-differentiable.

Definition 1.10 (Diffeomorphic):
Two $\mathcal{C}^k$-manifolds $(M,\mathcal{O}_M,\mathscr{A}_M)$ and $(N,\mathcal{O}_N,\mathscr{A}_N)$ are called diffeomorphic if there exists a $\mathcal{C}^k$-diffeomorphism $f: M\to N$. Then we write

$$M \cong_{\mathit{diff}} N \quad \text{or} \quad (M,\mathcal{O}_M,\mathscr{A}_M) \cong_{\mathit{diff}} (N,\mathcal{O}_N,\mathscr{A}_N). \tag{1.4}$$

With this new notation, we want to finally answer the question: whether, for instance

$$(\mathbb{R},\mathcal{O}_{\text{std.}},\mathscr{A}_{1,\max}) \cong_{\text{diff}} (\mathbb{R},\mathcal{O}_{\text{std.}},\mathscr{A}_{2,\max}) \tag{1.5}$$

where $\mathscr{A}_{1,\max}$ and $\mathscr{A}_{2,\max}$ are the maximal $\mathcal{C}^\infty$-atlases on $\mathbb{R}$ defined in the previous example.

In principle, we want to know, how many different differentiable structures are there on a given topological manifold $(M,\mathcal{O})$ – up to diffeomorphism? The answer to this question is not known in general, but we know that it depends on the dimension of the manifold M.

Theorem 1.11 (Radon-Moise)
For $d\le 3$, any two $\mathcal{C}^\infty$-manifolds of dimension d are diffeomorphic if and only if they are homeomorphic. In other words, let $(M,\mathcal{O}_M,\mathscr{A}_M)$ and $(N,\mathcal{O}_N,\mathscr{A}_N)$ be two $\mathcal{C}^\infty$-manifolds of dimension d and $d\le 3$. Then

$$(M,\mathcal{O}_M,\mathscr{A}_M) \cong_{\text{diff}} (N,\mathcal{O}_N,\mathscr{A}_N) \Leftrightarrow (M,\mathcal{O}_M) \cong_{\text{top.}} (N,\mathcal{O}_N). \tag{1.6}$$

So in particular, if $(M,\mathcal{O}_M)$ and $(N,\mathcal{O}_N)$ are homeomorphic, then we have a unique $\mathcal{C}^\infty$-structure on M and N up to diffeomorphism.

From the above theorem, we can say that given a topological manifold $(M, \mathcal{O}_M)$, there is a unique $\mathcal{C}^\infty$-structure on it up to diffeomorphism if $\dim M \leq 3$.

The answer for $d > 4$ (specifically for $d \neq 4$) is that there are finitely many different differentiable structures on a given *compact* topological manifold $(M, \mathcal{O}_M)$ up to diffeomorphism. This answer is provided by *surgery theory* (or obstruction 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.

> Remark 1.12 (Good News for String Theorists). According to many string theorists, our space-time is a 10-dimensional manifold. Since we don't have a unique differentiable structure on a 10-dimensional manifold, so in principle, different differentiable structures can lead to different predictions in physics, which is not what we want. But the good news is that for $d = 10$, there are only finitely many different differentiable structures, so we can decide which one is the correct for our space-time by performing finite number of experiments.

For $d = 4$, the situation is very different. In fact, the problem of classifying all smooth differentiable structures is still open. But we know following partial results:

- Non-compact 4-manifolds:

  There are uncountably many non-diffeomorphic $\mathcal{C}^\infty$-structures, specifically on $\mathbb{R}^4$.

- Compact 4-manifolds:

  The classification is not yet complete, but one of the most interesting results is that there are countably many non-diffeomorphic $\mathcal{C}^\infty$-structures on a given compact 4-manifold with $b_2 > 18$ (where b_2 is the second Betti number, which is a topological invariant of the manifold).

> Remark 1.13 (Betti Numbers). Betti numbers are topological invariants defined using homology groups (notion of algebraic topology). But, intiuitively, the k-th Betti number b_k of a topological space is the number of k-dimensional holes in it.
>
> - b_0 is the number of connected components;
> - b_1 is the number of 1-dimensional or "circular" holes;
> - b_2 is the number of 2-dimensional "voids" or "cavities".
> - And so on.
>
> For example, the 2-sphere S^2 has $b_0 = 1$, $b_1 = 0$ and $b_2 = 1$ as it has one connected component, no circular holes and one 2-dimensional cavity. And the 2-torus T^2 has $b_0 = 1$, $b_1 = 2$ and $b_2 = 1$ as it has one connected component, two circular holes (one equitorial and one meridional) and one 2-dimensional cavity.

Key feature of a differentiable manifold is that there exists a "*tangent space*" at each point of the manifold.

LECTURE 8

§1.2 Review of Vector Spaces

Definition 1.14 ((Algebraic) Field):
Let K be a non-empty set with two binary operations $+$ and $\cdot$ (addition and multiplication) such that

- *$(K, +)$ is an abelian group with identity element 0.*
- *$(K \setminus \{0\}, \cdot)$ is an abelian group with identity element 1.*
- *Multiplication is distributive over addition, i.e. for all $a, b, c \in K$,*

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c. \tag{1.7}$$

Then K is called a field.

Later we will use a set R equipped with two binary operations $+$ and $\cdot$ but fewer axioms than a field, and we will call it a *ring*.

Definition 1.15 (Ring):
A ring is a set R equipped with two binary operations $+$ and $\cdot$ such that

- *$(R, +)$ is an abelian group with identity element 0.*
- *$(R, \cdot)$ is a monoid with identity element 1.*

  i.e. multiplication is associative and has an identity element, but it need not be commutative.
- *Multiplication is distributive over addition, i.e. for all $a, b, c \in R$,*

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c. \tag{1.8}$$

Some examples which we are using from our school days:

- $(\mathbb{Z}, +, \cdot)$ is a *commutative* ring but not a field as it has no multiplicative inverses for all non-zero elements.
- $(\mathbb{Q}, +, \cdot)$ and $(\mathbb{R}, +, \cdot)$ are fields.
- $(\mathcal{M}_n(\mathbb{R}), +, \circ)$ is a ring, where $\mathcal{M}_n(\mathbb{R})$ is the set of $n \times n$ real matrices with the usual matrix addition and multiplication. Matrix multiplication is associative but not commutative, and it has an identity element I_n (the identity matrix). But it does not have multiplicative inverses for all non-zero elements, so it is not a field.

Definition 1.16 (Vector Space over a Field K):
A vector space V over a field $(K, +, \cdot)$ is a set equipped with two operations $\oplus : V \times V \to V$ (vector addition) and $\odot : K \times V \to V$ (scalar multiplication) such that

- *$(V, \oplus)$ is an abelian group with identity element $\mathbf{0}$ (the zero vector).*
- *The scalar multiplication satisfies the following properties*
  - *(i)* *Let 1 be the multiplicative identity of the field K. Then for all $\mathbf{v} \in V : 1 \odot \mathbf{v} = \mathbf{v}$.*
  - *(ii)* $\forall a, b \in K : \forall \mathbf{v} \in V : a \odot (b \odot \mathbf{v}) = (a \cdot b) \odot \mathbf{v}$.
  - *(iii)* $\forall a \in K : \forall \mathbf{u}, \mathbf{v} \in V : a \odot (\mathbf{u} \oplus \mathbf{v}) = a \odot \mathbf{u} \oplus a \odot \mathbf{v}$.
  - *(iv)* $\forall a, b \in K : \forall \mathbf{v} \in V : (a + b) \odot \mathbf{v} = a \odot \mathbf{v} \oplus b \odot \mathbf{v}$.

Definition 1.17 ((Vector) Subspace):
A subset $U \subseteq V$ of a vector space V over a field K is called a subspace *if it is closed under the vector addition and scalar multiplication, i.e.*

$$\forall \mathbf{u}, \mathbf{v} \in U : \mathbf{u} \oplus \mathbf{v} \in U \quad \text{and} \quad \forall a \in K : \forall \mathbf{u} \in U : a \odot \mathbf{u} \in U. \tag{1.9}$$

Since we are already comfortable with vector spaces, we drop the special notation for vector addition and scalar multiplication and use the usual $+$ and $\cdot$ for these operations.

Now, continuing with our usual structure, now we will define structure preserving maps between vector spaces *i.e. linear maps.*

Definition 1.18 (Linear Map):
Let $(V, \oplus, \odot)$ *and* $(W, \boxplus, \boxdot)$ *be two vector spaces over the same field* K*. A map* $L : V \to W$ *is called a* linear map *if it satisfies the following properties:*

(i) $\forall \mathbf{u}, \mathbf{v} \in V : L(\mathbf{u} \oplus \mathbf{v}) = L(\mathbf{u}) \boxplus L(\mathbf{v})$ *(preserves vector addition).*

(ii) $\forall a \in K : \forall \mathbf{v} \in V : L(a \odot \mathbf{v}) = a \boxdot L(\mathbf{v})$ *(preserves scalar multiplication).*

If a linear map $L : V \to W$ *is bijective, it is called a* linear isomorphism*. We write* $V \cong_{vec} W$ *if there exists a linear isomorphism between* V *and* W*.*

We can compress the definition of a linear map into a single equation:

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in V, \forall a, b \in K : L(a \odot \mathbf{v}_1 \oplus b \odot \mathbf{v}_2) = a \boxdot L(\mathbf{v}_1) \boxplus b \boxdot L(\mathbf{v}_2). \tag{1.10}$$

Remark 1.19 (Inverse being a linear map). As in case of topological spaces, we need to check that the inverse of a continuous map is continuous, here linearity of the inverse map follows from the linearity of the original map. Thus, it is enough to check that a linear map $L : V \to W$ is bijective to conclude that its inverse $L^{-1} : W \to V$ is also a linear map.

Pf: Let $L : V \to W$ be a linear map which is bijective. We need to show that $L^{-1} : W \to V$ is linear.

(i) Since L is bijective, there exists a unique $\mathbf{v} \in V$ such that $L(\mathbf{v}) = \mathbf{w}$ for any $\mathbf{w} \in W$. Thus, $L^{-1}(\mathbf{w}) = \mathbf{v}$.

(ii) For any $\mathbf{w}_1, \mathbf{w}_2 \in W$, we have

$$L^{-1}(\mathbf{w}_1 + \mathbf{w}_2) = L^{-1}(L(\mathbf{v}_1) + L(\mathbf{v}_2)) = L^{-1}(L(\mathbf{v}_1 + \mathbf{v}_2)) = \mathbf{v}_1 + \mathbf{v}_2 = L^{-1}(\mathbf{w}_1) + L^{-1}(\mathbf{w}_2),$$

where we used the linearity of L and the fact that L^{-1} is the inverse of L.

(iii) For any $a \in K$ and $\mathbf{w} \in W$, we have

$$L^{-1}(a \cdot \mathbf{w}) = L^{-1}(a \cdot L(\mathbf{v})) = L^{-1}(L(a \cdot \mathbf{v})) = a \cdot \mathbf{v} = a \cdot L^{-1}(\mathbf{w}),$$

again using the linearity of L.

Thus, L^{-1} is a linear map. □

Let's consider the set of all linear maps from V to W, denoted by $\mathrm{Hom}(V, W)$.

$$\mathrm{Hom}(V, W) := \{L : V \to W \mid L \text{ is a linear map}\} \equiv \left\{L : V \xrightarrow{\sim} W\right\} \tag{1.11}$$

here $\xrightarrow{\sim}$ denotes that the map is a linear map.

Proposition 1.20
$\mathrm{Hom}(V, W)$ is a vector space over the field K with the following operations:

- $\oplus : \mathrm{Hom}(V, W) \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$ defined by

$$(L_1, L_2) \xmapsto{\oplus} L_1 \oplus L_2 \tag{1.12}$$

  where

$$L_1 \oplus L_2 : V \xrightarrow{\sim} W, \quad \mathbf{v} \mapsto (L_1 \oplus L_2)(\mathbf{v}) := L_1(\mathbf{v}) \boxplus L_2(\mathbf{v}). \tag{1.13}$$

- $\odot : K \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$ defined by

$$(a, L) \xmapsto{\odot} a \odot L \tag{1.14}$$

where

$$a \diamondplus L : V \xrightarrow{\sim} W, \quad \mathbf{v} \mapsto (a \diamondplus L)(\mathbf{v}) := a \boxdot L(\mathbf{v}). \tag{1.15}$$

To establish this, we need to verify that the operations defined above satisfy the linear map properties:

Proof:
We need to show that $\mathrm{Hom}(V, W)$ is a vector space over K with the operations $\diamondplus$ and $\diamonddot$.

(i) Closure under addition: Let $L_1, L_2 \in \mathrm{Hom}(V, W)$. Then $L_1 \diamondplus L_2$ is defined as

$$(L_1 \diamondplus L_2)(\mathbf{v}) = L_1(\mathbf{v}) \boxplus L_2(\mathbf{v}). \tag{1.16}$$

So for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $a, b \in K$,

$$\begin{aligned}
(L_1 \diamondplus L_2)(a \odot \mathbf{v}_1 \oplus b \odot \mathbf{v}_2) &= L_1(a \odot \mathbf{v}_1 \oplus b \odot \mathbf{v}_2) \boxplus L_2(a \odot \mathbf{v}_1 \oplus b \odot \mathbf{v}_2) \\
&= (a \boxdot L_1(\mathbf{v}_1) \boxplus b \boxdot L_1(\mathbf{v}_2)) \boxplus (a \boxdot L_2(\mathbf{v}_1) \boxplus b \boxdot L_2(\mathbf{v}_2)) \\
&= a \boxdot (L_1(\mathbf{v}_1) \boxplus L_2(\mathbf{v}_1)) \boxplus b \boxdot (L_1(\mathbf{v}_2) \boxplus L_2(\mathbf{v}_2)) \\
&= a \boxdot (L_1 \diamondplus L_2)(\mathbf{v}_1) \boxplus b \boxdot (L_1 \diamondplus L_2)(\mathbf{v}_2).
\end{aligned}$$

Thus, $L_1 \diamondplus L_2$ is a linear map.

(ii) Closure under scalar multiplication: Let $a \in K$ and $L \in \mathrm{Hom}(V, W)$. Then $a \diamonddot L$ is defined as

$$(a \diamonddot L)(\mathbf{v}) = a \boxdot L(\mathbf{v}). \tag{1.17}$$

So for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $b \in K$,

$$\begin{aligned}
(a \diamonddot L)(b \odot \mathbf{v}_1 \oplus c \odot \mathbf{v}_2) &= a \boxdot L(b \odot \mathbf{v}_1 \oplus c \odot \mathbf{v}_2) \\
&= a \boxdot (b \boxdot L(\mathbf{v}_1) \boxplus c \boxdot L(\mathbf{v}_2)) \\
&= (a \cdot b) \boxdot L(\mathbf{v}_1) \boxplus (a \cdot c) \boxdot L(\mathbf{v}_2) \\
&= (b \cdot a) \boxdot L(\mathbf{v}_1) \boxplus (c \cdot a) \boxdot L(\mathbf{v}_2) \\
&= b \boxdot (a \boxdot L(\mathbf{v}_1)) \boxplus c \boxdot (a \boxdot L(\mathbf{v}_2)) \\
&= b \boxdot (a \diamonddot L)(\mathbf{v}_1) \boxplus c \boxdot (a \diamonddot L)(\mathbf{v}_2).
\end{aligned}$$

Thus, $a \diamonddot L$ is a linear map.

We have shown that $\mathrm{Hom}(V, W)$ is closed under both operations $\diamondplus$ and $\diamonddot$. Thus, $\mathrm{Hom}(V, W)$ is a vector space over the field K. Q.E.D.

Till now, we haven't used the inverse element of the field K in the definition of a vector space. So, we can define a similar structure on a set M over a (unital) ring R with two operations $\oplus : M \times M \to M$ (addition) and $\odot : R \times M \to M$ (scalar multiplication), we call this a *module* over the ring R.

> Remark 1.21 (Case of $\mathrm{Hom}(V, W)$ as a module). In case of modules, we can still define $\mathrm{Hom}(V, W)$ as a set of all linear maps from V to W over a ring R. But as in general, ring multiplication is not commutative, $\mathrm{Hom}(V, W)$ is not a module over R.

Some useful terminology:

- Endomorphism is a linear map $L : V \to V$ from a vector space V to itself.

$$\mathrm{End}(V) := \mathrm{Hom}(V, V) = \left\{ L : V \xrightarrow{\sim} V \right\} \tag{1.18}$$

- Automorphism is a linear isomorphism $L : V \to V$ from a vector space V to itself.

$$\mathrm{Aut}(V) := \left\{ L : V \xrightarrow{\sim} V \;\middle|\; L \text{ is a linear isomorphism} \right\}. \tag{1.19}$$

It is easy to see that $\mathrm{Aut}(V)$ is a subspace of $\mathrm{End}(V)$.

- A field K can be considered as a vector space over itself, *i.e.* K is a vector space over K with the usual addition and multiplication operations. With this view, we can define the notion of *linear functionals* as linear maps from V to K:

$$V^* := \mathrm{Hom}(V, K) \tag{1.20}$$

This set V^* is called the *dual space* of V.

§1.2.1 Tensors

Definition 1.22 (Tensor):
A type (p, q)-tensor *is a multilinear map (linear in each argument) of the form*

$$T : \underbrace{V^* \times \cdots \times V^*}_{p \text{ times}} \times \underbrace{V \times \cdots \times V}_{q \text{ times}} \to K \tag{1.21}$$

where V *is a vector space over a field* K *and* V^* *is its dual space.*

$$\mathsf{T}^p_q V := \underbrace{V \otimes \cdots \otimes V}_{p \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q \text{ times}} = \{T \mid T \text{ is type } (p, q)\text{-tensor}\} \tag{1.22}$$

Let's define the operations on tensors similar to the operations we defined on the set of linear maps $\mathrm{Hom}(V, W)$:

1. Addition of tensors:

$$\oplus : \mathsf{T}^p_q V \times \mathsf{T}^p_q V \to \mathsf{T}^p_q V \tag{1.23}$$

Let $T_1, T_2 \in \mathsf{T}^p_q V$ be two tensors. Then their sum $T_1 \oplus T_2$ is defined as

$$(T_1 \oplus T_2)(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) = T_1(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) + T_2(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) \tag{1.24}$$

2. Scalar multiplication of tensors:

$$\odot : K \times \mathsf{T}^p_q V \to \mathsf{T}^p_q V \tag{1.25}$$

Let $a \in K$ and $T \in \mathsf{T}^p_q V$ be a tensor. Then the scalar multiplication $a \odot T$ is defined as

$$(a \odot T)(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) = a \cdot T(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) \tag{1.26}$$

Proposition 1.23
$\mathsf{T}^p_q V$ is a vector space over the field K with the operations defined above.

Sometimes, we refer (p, q) as the valence of the tensor, and $p + q$ as the rank of the tensor.

Now, we define a binary operation on tensors, called the *tensor product* of two tensors.

Definition 1.24 (Tensor Product):
Let $T_1 \in \mathsf{T}^p_q V$ *and* $T_2 \in \mathsf{T}^r_s V$ *be two tensors. The tensor product* $T_1 \otimes T_2$ *is defined as a tensor of type* $(p + r, q + s)$ *given by*

$$\begin{aligned}(T_1 \otimes T_2)(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_{p+r}, \mathbf{v}_1, \ldots, \mathbf{v}_{q+s}) &= \\ T_1(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) &\cdot T_2(\boldsymbol{\omega}_{p+1}, \ldots, \boldsymbol{\omega}_{p+r}, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_{q+s}).\end{aligned} \tag{1.27}$$

Example 1.25
Now, let's look at some examples and interesting results involving tensors:

(a) The set $\mathsf{T}^0_0 V = K$ is the set of all scalar tensors, which is just the field K itself.

(b) The set $\mathsf{T}^0_1 V = V^* := \mathrm{Hom}(V, K)$.

(c) The set $\mathsf{T}^1_1 V \equiv V \otimes V^* = \{T \mid T : V^* \times V \to K \text{ is a linear in both argument}\} \cong_{\text{vec}} \text{End}(V^*)$.

Pf: Let $T \in \mathsf{T}^1_1 V$ be a tensor. We need to construct a linear map $L \in \text{End}(V^*)$ using T. For each $\boldsymbol{\omega} \in V^*$, we define a linear map $L_{\boldsymbol{\omega}} : V \to K \in V^*$ by

$$\forall \mathbf{v} \in V : L_{\boldsymbol{\omega}}(\mathbf{v}) = T(\boldsymbol{\omega}, \mathbf{v}). \tag{1.28}$$

Now, we can define a linear map $L : V^* \to V^*$ by

$$L(\boldsymbol{\omega}) = L_{\boldsymbol{\omega}} \qquad \Rightarrow \qquad \boldsymbol{\omega} \mapsto \underbrace{(\mathbf{v} \mapsto T(\boldsymbol{\omega}, \mathbf{v}))}_{L_{\boldsymbol{\omega}}}. \tag{1.29}$$

Now, let $L \in \text{End}(V^*)$ be a linear map. We need to show that there exists a tensor $T \in \mathsf{T}^1_1 V$ corresponding to L. Define $T : V^* \times V \to K$ by

$$T(\boldsymbol{\omega}, \mathbf{v}) = L(\boldsymbol{\omega})(\mathbf{v}) \tag{1.30}$$

as $L(\boldsymbol{\omega}) \in V^*$ is a linear map from V to K. $\square$

(d) From previous example, we expect that $\mathsf{T}^1_0 V \cong_{\text{vec}} V$; but this is not true in general. However, if V is finite-dimensional, then $\mathsf{T}^1_0 V \cong_{\text{vec}} V$.

(e) Similarly, $\mathsf{T}^1_1 V \cong_{\text{vec}} \text{End}(V)$ is also not true in general, but if V is finite-dimensional, then $\mathsf{T}^1_1 V \cong_{\text{vec}} \text{End}(V)$.

(f) All these examples which are not true in general, but true for finite-dimensional vector spaces, can be summarized as

$$(V^*)^* \ncong_{\text{vec}} V. \tag{1.31}$$

This is true only for finite-dimensional vector spaces.

§1.2.2 Dimension of a Vector Space

Now, we have mentioned dimensions of vector spaces several times, so let's define the notion of dimension of a vector space. But first, we need to define a basis of a vector space. Since till now, we haven't considered any additional structure on a vector space, we can only define a *Hamel basis* of a vector space.

Definition 1.26 (Hamel Basis):
Let $(V, +, \cdot)$ be a vector space over a field K. A subset $\mathcal{B} \subseteq V$ is called a Hamel basis *if it satisfies the following properties:*

(i) $\mathcal{B}$ is a linearly independent set, i.e. for any finite subset $\{\mathbf{b}_1, \ldots, \mathbf{b}_N\} \subseteq \mathcal{B}$, the only solution to the equation

$$\sum_{i=1}^{N} \lambda^i \cdot \mathbf{b}_i = \mathbf{0} \quad \textit{for } \lambda^i \in K \tag{1.32}$$

is $\lambda^i = 0$ for all i.

(ii) $\mathcal{B}$ is a spanning set, i.e. every element $\mathbf{v} \in V$ can be expressed as a finite linear combination of elements from $\mathcal{B}$, i.e.

$$\forall \mathbf{v} \in V : \exists M \in \mathbb{N} : \exists v^1, \ldots, v^M \in K : \exists \mathbf{b}_1, \ldots, \mathbf{b}_M \in \mathcal{B} : \mathbf{v} = \sum_{i=1}^{M} v^i \cdot \mathbf{b}_i. \tag{1.33}$$

Say we all have decided to use a single Hamel basis $\mathcal{B}$ for our vector space V. Then it is more convenient to talk about the elements of V as an array of coefficients *i.e.* $(v^1, \ldots, v^M)$ also called the *coordinates* of the vector with respect to the Hamel basis $\mathcal{B}$.

With this definition, we can define the dimension of a vector space. But first, we need to ensure that the dimension is well-defined, *i.e.* it does not depend on the choice of Hamel basis. This is guaranteed by the following proposition.

Proposition 1.27
Let $(V, +, \cdot)$ be a vector space over a field K and let $\mathcal{B}$ be a Hamel basis of V. Then $\mathcal{B}$ is the minimal spanning set of V, and maximal linearly independent set of V. In other words, let $S \subseteq V$.

(i) If $\operatorname{span}(S) = V$, then $|S| \geq |\mathcal{B}|$.

(ii) If S is linearly independent, then $|S| \leq |\mathcal{B}|$.

Definition 1.28 (Dimension):
The dimension *of a vector space V over a field K is defined as the cardinality of any Hamel basis of V. We denote the dimension of V by $\dim V$.*

$$\dim V := |\mathcal{B}| \quad \textit{for any Hamel basis } \mathcal{B} \subseteq V. \tag{1.34}$$

If V has a finite Hamel basis, we say that V is finite-dimensional *and $\dim V < \infty$. If V does not have a finite Hamel basis, we say that V is* infinite-dimensional *and $\dim V = \infty$.*

Now, we have proper tools to prove following theorem.

Theorem 1.29
Let V be a finite-dimensional vector space. Then

$$(V^*)^* \cong_{\text{vec}} V. \tag{1.35}$$

Proof:
Let V be an n-dimensional vector space over a field K. Define a linear map $L : V \to (V^*)^*$ by

$$\begin{aligned} L : V &\to (V^*)^* \\ \mathbf{v} &\mapsto L_{\mathbf{v}} \end{aligned} \tag{1.36}$$

where $L_{\mathbf{v}} : V^* \to K \in (V^*)^*$ is defined by

$$\begin{aligned} L_{\mathbf{v}} : V^* &\to K \\ \boldsymbol{\omega} &\mapsto L_{\mathbf{v}}(\boldsymbol{\omega}) := \boldsymbol{\omega}(\mathbf{v}) \end{aligned} \tag{1.37}$$

We need to show that L is a linear isomorphism.

(i) Linearity of L:
For any $\mathbf{v}_1, \mathbf{v}_2 \in V$ and $a, b \in K$,

$$\begin{aligned} L(a\mathbf{v}_1 + b\mathbf{v}_2)(\boldsymbol{\omega}) &= L_{\{a\mathbf{v}_1 + b\mathbf{v}_2\}}(\boldsymbol{\omega}) \\ &= \boldsymbol{\omega}(a\mathbf{v}_1 + b\mathbf{v}_2) \\ &= a \cdot \boldsymbol{\omega}(\mathbf{v}_1) + b \cdot \boldsymbol{\omega}(\mathbf{v}_2) \\ &= a \cdot L_{\mathbf{v}_1}(\boldsymbol{\omega}) + b \cdot L_{\mathbf{v}_2}(\boldsymbol{\omega}) \\ &= (aL_{\mathbf{v}_1} + bL_{\mathbf{v}_2})(\boldsymbol{\omega}). \end{aligned}$$

Thus, $L(a\mathbf{v}_1 + b\mathbf{v}_2) = aL_{\mathbf{v}_1} + bL_{\mathbf{v}_2}$, which shows that L is linear.

(ii) Injectivity of L:
Let $\mathbf{v} \neq \mathbf{0} \in V$. We need to show that $L(\mathbf{v}) \neq 0 \in (V^*)^*$. Since $\mathbf{v} \neq \mathbf{0}$, so $\operatorname{span}(\mathbf{v}) \subseteq V$ is a proper subspace of V, so we can extend this to a Hamel basis $\mathcal{B} = \{\mathbf{v}, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of V. Thus, any $\mathbf{u} \in V$ can be expressed as

$$\mathbf{u} = b \cdot \mathbf{v} + \sum_{i=2}^{n} b^i \cdot \mathbf{e}_i \tag{1.38}$$

for some $b, b^2, \ldots, b^n \in K$. Now, we can define a linear functional $\boldsymbol{\omega} \in V^*$ as follows:

$$\mathbf{u} \mapsto \boldsymbol{\omega}(\mathbf{u}) := b. \tag{1.39}$$

Thus, we have

$$L(\mathbf{v})(\boldsymbol{\omega}) = \boldsymbol{\omega}(\mathbf{v}) = 1 \neq 0. \tag{1.40}$$

This shows that $L(\mathbf{v}) \neq 0$, hence L is injective.

> Claim (Dual Basis): For a finite-dimensional vector space V, with a Hamel basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, there exists a unique dual basis $\mathcal{B}^* = \{\mathbf{f}^1, \ldots, \mathbf{f}^n\}$ such that
>
> $$\forall i, j \in \{1, \ldots, n\} : \mathbf{f}^i(\mathbf{e}_j) = \delta^i_j := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{1.41}$$

(iii) Surjectivity of L:
Let $F \in (V^*)^*$ be any linear functional. We need to show that there exists a $\mathbf{v} \in V$ such that $L(\mathbf{v}) = F$. Since V is finite-dimensional, we can choose a Hamel basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of V. Now, we can define a vector $\mathbf{v} \in V$ as follows:

$$\mathbf{v} = \sum_{i=1}^{n} F(\mathbf{f}^i) \cdot \mathbf{e}_i \tag{1.42}$$

where $\mathbf{f}^i$ is the i-th element of the dual basis $\mathcal{B}^* = \{\mathbf{f}^1, \ldots, \mathbf{f}^n\}$ corresponding to the Hamel basis $\mathcal{B}$. Now, we can check that

$$\begin{aligned} L(\mathbf{v})(\boldsymbol{\omega}) &= \sum_{i=1}^{n} F(\mathbf{f}^i) \cdot \boldsymbol{\omega}(\mathbf{e}_i) \\ &= \sum_{i=1}^{n} F(\mathbf{f}^i) \cdot \delta^j_i \\ &= F(\boldsymbol{\omega}). \end{aligned}$$

Thus, $L(\mathbf{v}) = F$, which shows that L is surjective.

Q.E.D.

Corollary 1.30
If V is a finite-dimensional vector space, then

(i) $\mathsf{T}^1_1 V \cong_{\text{vec}} \text{End}(V)$.

(ii) $\mathsf{T}^1_0 V \cong_{\text{vec}} V$.

Till now, we have only used basis to classify vector spaces, either finite-dimensional or infinite-dimensional. In vector thoery, we don't use basis to construct objects, but once we have the definition of the object without using basis, we can use basis to study the object in a more convenient way.

We have seen the components of a vector with respect to a basis, now we can define the components of a tensor with respect to a basis.

Definition 1.31 (Components of a Tensor):
Let V be a finite-dimensional vector space over a field K with a Hamel basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and the corresponding dual basis $\mathcal{B}^* = \{\boldsymbol{\epsilon}^1, \ldots, \boldsymbol{\epsilon}^n\}$. The components of a tensor $T \in \mathsf{T}^p_q V$ with respect to the basis $\mathcal{B}$ and $\mathcal{B}^*$ are defined as follows:

$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} := T(\boldsymbol{\epsilon}^{i_1}, \ldots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q}) \quad \textit{for } i_k, j_l \in \{1, \ldots, n\}. \tag{1.43}$$

With these tensor components, we can express the tensor T as a sum of its components multiplied by the basis 'tensors':

$$T = \underbrace{\sum_{i_1=1}^{n} \cdots \sum_{j_q=1}^{n}}_{p+q \text{ sums}} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \cdot \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_q}. \tag{1.44}$$

Here $\otimes$ denotes the tensor product of tensors, as $\mathbf{e}_{i_k} \in V \cong_{\text{vec}} \mathsf{T}^1_0 V$ and $\boldsymbol{\epsilon}^{j_l} \in V^* \cong_{\text{vec}} \mathsf{T}^0_1 V$. The components are field elements, so the implicit multiplication is the scalar multiplication of tensors of type (p, q).

From now on, we will do this big sums multiple times, so to reduce the clutter, we will use the Einstein summation convention. In this convention, we will omit the summation symbol and assume that repeated indices (one upper and one lower) are summed over. For example, the above expression can be written as

$$T = \underbrace{\sum_{i_1=1}^{n} \cdots \sum_{j_q=1}^{n}}_{p+q \text{ sums}} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \cdot \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_q} \equiv$$
$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \cdot \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_q}. \tag{1.45}$$

For Up-Down convection, we only need to remember the fundamental convention *i.e.*

- Basis vectors of V are denoted by lower indices, $\mathbf{e}_i$.
- Basis vectors of V^* are denoted by upper indices, $\boldsymbol{\epsilon}^i$.

So the indices of the components of a tensor are due to the basis vectors of V and V^* used to define it.

$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} := T(\boldsymbol{\epsilon}^{i_1}, \ldots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q}). \tag{1.46}$$

We can recover the indices for the coordinates of a vector and a covector as follows (given a choice of basis and dual basis)

$$\mathbf{v} = v^i \cdot \mathbf{e}_i \quad \text{where } v^i = \boldsymbol{\epsilon}^i(\mathbf{v}) \tag{1.47}$$
$$\boldsymbol{\omega} = \omega_j \cdot \boldsymbol{\epsilon}^j \quad \text{where } \omega_j = \mathbf{e}_j(\boldsymbol{\omega}). \tag{1.48}$$

Remark 1.32 (Caution). But we need to be careful about this convention, as it is only works for linear spaces and linear maps. For example, if we have a tensor $T : V^* \times V \to K$, then for any $\boldsymbol{\omega} \in V^*$ and $\mathbf{v} \in V$, we have

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

So, if we omit the summation symbol, then there is no way to distinguish between the first and second equation. But the second equation is only true for bilinear maps, and not for non-bilinear maps. So, we need to be careful about the context in which we are using the Einstein summation convention.

It is important to know that, how to go from one basis to another basis. This is called the *change of basis*.

§1.2.3 Change of Basis

Consider two bases $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\tilde{\mathcal{B}} = \{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ of a vector space V. Since each vector in $\tilde{\mathcal{B}}$ is a vector in V, we can express each $\tilde{\mathbf{e}}_i$ as a linear combination of the vectors in $\mathcal{B}$: (using the Einstein summation convention)

$$\tilde{\mathbf{e}}_i = A^j{}_i \mathbf{e}_j \tag{1.49}$$

where $A^j{}_i \in K$. Similarly, we can express each $\mathbf{e}_i$ as a linear combination of the vectors in $\tilde{\mathcal{B}}$:

$$\mathbf{e}_i = B^j{}_i \tilde{\mathbf{e}}_j \tag{1.50}$$

where $B^j{}_i \in K$. In linear algebra, we can define two matrices A and B with entries $A^j{}_i$ and $B^j{}_i$ respectively, such that

$$A^{-1} = B \quad \text{and} \quad B^{-1} = A. \tag{1.51}$$

But till now, we haven't explored the vector spaces using arrays, so we define following notation for using usual notion of matrices and (row and column) vectors from our linear algebra course.

> Remark 1.33 (Matrix Notation). Fix a basis $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of a vector space V and a dual basis $\mathcal{B}^* = \{\boldsymbol{\epsilon}^1, \ldots, \boldsymbol{\epsilon}^n\}$. Then we will express
>
> $$\mathbf{v} = v^i \mathbf{e}_i \quad \leftrightsquigarrow \quad \mathbf{v} \mathrel{\hat{=}} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \tag{1.52}$$
>
> and
>
> $$\boldsymbol{\omega} = \omega_j \boldsymbol{\epsilon}^j \quad \leftrightsquigarrow \quad \boldsymbol{\omega}^\top \mathrel{\hat{=}} \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix}. \tag{1.53}$$
>
> Now let $\phi \in \operatorname{End}(V) \cong_{\text{vec}} \mathsf{T}^1_1 V$ be a linear map. Then we can express ϕ as
>
> $$\phi = \phi^i{}_j \mathbf{e}_i \otimes \boldsymbol{\epsilon}^j \quad \leftrightsquigarrow \quad \Phi \mathrel{\hat{=}} \begin{pmatrix} \phi^1{}_1 & \cdots & \phi^1{}_n \\ \vdots & \ddots & \vdots \\ \phi^n{}_1 & \cdots & \phi^n{}_n \end{pmatrix}. \tag{1.54}$$

So in general, the convention is that we will use the upper indices for the rows and lower indices for the columns of a matrix.

We know for finite-dimensional vector space V, we have $\operatorname{End}(V) \cong_{\text{vec}} \mathsf{T}^1_1 V$. Let $\phi \in \operatorname{End}(V)$, so we can rethink ϕ as a type $(1,1)$-tensor in following sense:

$$\phi(\boldsymbol{\omega}, \mathbf{v}) := \boldsymbol{\omega}(\phi(\mathbf{v})). \tag{1.55}$$

So the components of tensor ϕ is defined as,

$$\phi^i{}_j := \phi(\boldsymbol{\epsilon}^i, \mathbf{e}_j) = \boldsymbol{\epsilon}^i(\phi(\mathbf{e}_j)). \tag{1.56}$$

Let $\phi, \psi \in \operatorname{End}(V)$. Let's compute the components of linear map $\phi \circ \psi$

$$\begin{aligned}
(\phi \circ \psi)^i{}_j &= \boldsymbol{\epsilon}^i((\phi \circ \psi)(\mathbf{e}_j)) \\
&= \boldsymbol{\epsilon}^i(\phi(\psi(\mathbf{e}_j))) \\
&= \boldsymbol{\epsilon}^i(\phi(\psi^a{}_j \mathbf{e}_a)) \\
&= \boldsymbol{\epsilon}^i(\psi^a{}_j \phi(\mathbf{e}_a)) \\
&= \boldsymbol{\epsilon}^i(\psi^a{}_j \phi^b{}_a \mathbf{e}_b) \\
&= \psi^a{}_j \cdot \phi^b{}_a \boldsymbol{\epsilon}^i(\mathbf{e}_b) \\
&= \psi^a{}_j \cdot \phi^b{}_a \cdot \delta^i_b \\
&= \psi^a{}_j \cdot \phi^i{}_a = \phi^i{}_a \cdot \psi^a{}_j.
\end{aligned}$$

With this, we have a definition for matrix multiplication *i.e.* say Φ, Ψ are matrices corresponding to linear maps ϕ, ψ, so the matrix for the composed linear map $\phi \circ \psi$ is given as

$$\Phi \cdot \Psi \,\hat{=}\, \left(\phi^i{}_a \cdot \psi^a{}_j\right)_{i,j=1}^{n} \tag{1.57}$$

We call this rule of producing matrices "*matrix-multiplication*."

Similarly, the action of a covector $\boldsymbol{\omega}(\mathbf{v}) = \omega_m v^m$, and this can be thought of as matrix multiplication as $\boldsymbol{\omega}^\top \cdot \mathbf{v}$. This notaion gives us a false picture that these are basis independent, but they are not. And in the similar spirit we also write,

$$\phi(\boldsymbol{\omega}, \mathbf{v}) = \omega_i \phi^i{}_j v^j \quad \leftrightsquigarrow \quad \boldsymbol{\omega}^\top \cdot \Phi \cdot \mathbf{v} \tag{1.58}$$

There is another issue with these representations, *i.e.* for tensor of type other than these three, we don't have nice picture to draw. So it is better to think them as multilinear maps instead of matrices.

Now, let's talk about the effect of change of basis on the vectors, covectors, and tensors. Recall we have two basis, $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\tilde{\mathcal{B}} = \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$ of a vector space V. Then the change of basis for vectors is given by

$$\tilde{\mathbf{e}}_i = A^j{}_i \mathbf{e}_j \qquad\qquad \mathbf{e}_j = B^i{}_j \tilde{\mathbf{e}}_i. \tag{1.59}$$

Since these relations are invertible, we have

$$A^i{}_m B^m{}_j = B^i{}_m A^m{}_j = \delta^i{}_j \tag{1.60}$$

For both the bases, we have the dual bases $\mathcal{B}^* = \{\boldsymbol{\epsilon}^1, \dots, \boldsymbol{\epsilon}^n\}$ and $\tilde{\mathcal{B}}^* = \{\tilde{\boldsymbol{\epsilon}}^1, \dots, \tilde{\boldsymbol{\epsilon}}^n\}$.

1. Covectors:
   Let $\boldsymbol{\omega} = \omega_j \boldsymbol{\epsilon}^j$ be a covector, where $\omega_j = \mathbf{e}_j(\boldsymbol{\omega})$. Here $\mathbf{e}_j$ is an element of $\mathsf{T}^1_0 V$. Then we have

$$\omega_j = \mathbf{e}_j(\boldsymbol{\omega}) = \boldsymbol{\omega}(\mathbf{e}_j) = \boldsymbol{\omega}(B^i{}_j \tilde{\mathbf{e}}_i) = B^i{}_j \boldsymbol{\omega}(\tilde{\mathbf{e}}_i) = B^i{}_j \tilde{\mathbf{e}}_i(\boldsymbol{\omega}) = B^i{}_j \tilde{\omega}_i. \tag{1.61}$$

2. Vectors:
   Let $\mathbf{v} = v^i \mathbf{e}_i$ be a vector, where $v^i = \boldsymbol{\epsilon}^i(\mathbf{v})$. Then we have

$$v^i = \boldsymbol{\epsilon}^i(\mathbf{v}) = \boldsymbol{\epsilon}^i(\tilde{v}^j \tilde{\mathbf{e}}_j) = \tilde{v}^j \boldsymbol{\epsilon}^i(\tilde{\mathbf{e}}_j) = \tilde{v}^j \boldsymbol{\epsilon}^i(A^k{}_j \mathbf{e}_k) = \tilde{v}^j A^k{}_j \boldsymbol{\epsilon}^i(\mathbf{e}_k) = \tilde{v}^j A^k{}_j \delta^i{}_k = A^i{}_j \tilde{v}^j. \tag{1.62}$$

Summarizing the above two equations, we have

$$v^i = A^i{}_j \tilde{v}^j \qquad\qquad \omega_j = B^i{}_j \tilde{\omega}_i \tag{1.63}$$

$$\tilde{v}^i = B^i{}_j v^j \qquad\qquad \tilde{\omega}_j = A^i{}_j \omega_i. \tag{1.64}$$

Now, let's see how the basis of dual space changes with respect to the change of basis of the vector space. Recall that the dual basis is defined as follows

$$\boldsymbol{\epsilon}^i(\mathbf{e}_j) = \delta^i_j \qquad\qquad \tilde{\boldsymbol{\epsilon}}^i(\tilde{\mathbf{e}}_j) = \delta^i_j. \tag{1.65}$$

Let $\tilde{\boldsymbol{\epsilon}}^i = C^i{}_j \boldsymbol{\epsilon}^j$, where $C^i{}_j \in K$. Then we have

$$\tilde{\boldsymbol{\epsilon}}^i(\tilde{\mathbf{e}}_j) = C^i{}_k \boldsymbol{\epsilon}^k(A^l{}_j \mathbf{e}_l) = C^i{}_k A^l{}_j \boldsymbol{\epsilon}^k(\mathbf{e}_l) = C^i{}_k A^l{}_j \delta^k_l = C^i{}_k A^k{}_j = \delta^i_j. \tag{1.66}$$

Thus, using unique existence of the inverse, we have

$$C^i{}_j = B^i{}_j. \tag{1.67}$$

So we have the following relations for the dual basis:

$$\tilde{\boldsymbol{\epsilon}}^i = B^i{}_j \boldsymbol{\epsilon}^j \qquad\qquad \boldsymbol{\epsilon}^j = A^j{}_i \tilde{\boldsymbol{\epsilon}}^i. \tag{1.68}$$

3. Tensors:
   Let $T \in \mathsf{T}^p_q V$ be a tensor, then the components of T with respect to the basis $\mathcal{B}$ and $\mathcal{B}^*$ are given by
$$T^{i_1\dots i_p}{}_{j_1\dots j_q} = T(\boldsymbol{\epsilon}^{i_1}, \dots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_q}). \tag{1.69}$$
   Now, we can express the components of T with respect to the basis $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{B}}^*$ as follows:
$$\begin{aligned}
\tilde{T}^{i_1\dots i_p}{}_{j_1\dots j_q} &= T(\tilde{\boldsymbol{\epsilon}}^{i_1}, \dots, \tilde{\boldsymbol{\epsilon}}^{i_p}, \tilde{\mathbf{e}}_{j_1}, \dots, \tilde{\mathbf{e}}_{j_q}) & (1.70)\\
&= T(B^{i_1}{}_{a_1}\boldsymbol{\epsilon}^{a_1}, \dots, B^{i_p}{}_{a_p}\boldsymbol{\epsilon}^{a_p}, A^{b_1}{}_{j_1}\mathbf{e}_{b_1}, \dots, A^{b_q}{}_{j_q}\mathbf{e}_{b_q}) & (1.71)\\
&= B^{i_1}{}_{a_1}\cdots B^{i_p}{}_{a_p}\cdot A^{b_1}{}_{j_1}\cdots A^{b_q}{}_{j_q}\cdot T(\boldsymbol{\epsilon}^{a_1}, \dots, \boldsymbol{\epsilon}^{a_p}, \mathbf{e}_{b_1}, \dots, \mathbf{e}_{b_q}) & (1.72)\\
&= B^{i_1}{}_{a_1}\cdots B^{i_p}{}_{a_p}\cdot A^{b_1}{}_{j_1}\cdots A^{b_q}{}_{j_q}\cdot T^{a_1\dots a_p}{}_{b_1\dots b_q}. & (1.73)
\end{aligned}$$
   Thus, the reversal of the basis is given by
$$T^{i_1\dots i_p}{}_{j_1\dots j_q} = A^{i_1}{}_{a_1}\cdots A^{i_p}{}_{a_p}\cdot B^{b_1}{}_{j_1}\cdots B^{b_q}{}_{j_q}\cdot \tilde{T}^{a_1\dots a_p}{}_{b_1\dots b_q}. \tag{1.74}$$

§1.2.4 Determinants

From our previous knowledge of linear algebra, we know that the determinant is a scalar-valued function that takes a square matrix and returns a scalar. But since we know that the square matrix is just a convention for a linear map, we need to define the determinant in a more general way, independent of the basis. But first, look at this weird result which is pruely due to the fact that we are using a witchcraft called *matrices* to represent linear maps.

Remark 1.34. Let $\phi \in \mathsf{T}^1_1 V$ be an endomorphism and $g \in \mathsf{T}^0_2 V$ be a bilinear form. We have
$$\phi = \phi^i{}_j \mathbf{e}_i \otimes \boldsymbol{\epsilon}^j \quad \text{and} \quad g = g_{ij}\boldsymbol{\epsilon}^i \otimes \boldsymbol{\epsilon}^j. \tag{1.75}$$
We can arrange the components of ϕ and g in a matrix form as follows:
$$\Phi = \begin{pmatrix} \phi^1{}_1 & \cdots & \phi^1{}_n \\ \vdots & \ddots & \vdots \\ \phi^n{}_1 & \cdots & \phi^n{}_n \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \cdots & g_{nn} \end{pmatrix}. \tag{1.76}$$
But say, we change the basis of V to $\tilde{\mathcal{B}}$, then the components of ϕ and g change as follows:
$$\tilde{\phi}^i{}_j = B^i{}_k \cdot \phi^k{}_l \cdot A^l{}_j \quad \text{and} \quad \tilde{g}_{ij} = A^k{}_i \cdot g_{kl} \cdot A^l{}_j. \tag{1.77}$$
Recall $(A^{-1})^i{}_j = B^i{}_j$, so we can write the above equations as (remember that $A^i{}_j = A_{ij}$, which give rise to transpose in the expression)
$$\tilde{\Phi} = A^{-1}\cdot\Phi\cdot A \quad \text{and} \quad \tilde{G} = A^\top\cdot G\cdot A. \tag{1.78}$$

Now, we keep the old notion of determinant, so we can find determinant of G. But that is not true, as determinant is only defined for endomorphisms. And to see, why this the case, we have to look at basis free definition of determinants.

§1.2.4.1 Permutation Group

Before we define the determinant, we need a few definitions. We will use the *permutation group* to define the determinant.

Definition 1.35 (Permutation and Permutation Group):
A permutation *of a non-empty set M is a bijective function $\sigma : M \to M$.*

Let the set M be finite with n elements, for purpose of this course, we will assume $M = \{1, 2, \dots, n\}$. The set of all permutations of M is denoted by S_n and is called the permutation group *of order n (or symmetric group of degree n). Binary operation on S_n is defined as the composition of two permutations.*

There are some special type of permutations, which are called *transpositions*. A transposition is a permutation that swaps two elements of the set and leaves the rest unchanged. For example, in S_3, the permutation $(1, 2, 3) \mapsto (2, 1, 3)$ is a transposition, usually denoted as $(1, 2)$.

There is a result from group theory, which states that every permutation can be expressed as a product of transpositions (precisely composition of transpositions). The number of transpositions in this product is called the *signature* or *sign* or *parity* of the permutation. Let $\sigma \in S_n$ be a permutation, then we define the sign of σ as follows:

$$\operatorname{sgn}(\sigma) := \begin{cases} +1 & \text{if } \sigma \text{ is the product of an even number of transpositions} \\ -1 & \text{if } \sigma \text{ is the product of an odd number of transpositions} \end{cases} \tag{1.79}$$

But there is an issue with this definition, *i.e.* the decomposition of a permutation into transpositions is not unique. For example, the permutation $(1, 2, 3)$ can be expressed as $(1, 2)(2, 3)$ or $(1, 3)(1, 2)$. But we can still show that the sign of a permutation is well-defined, *i.e.* it does not depend on the decomposition of the permutation into transpositions. Since the number of transpositions in the decomposition is either even or odd, and this property is preserved under composition of permutations.

Now we can define n-form for a vector space V of dimension d.

Definition 1.36 (n-form):
Let V be a finite-dimensional vector space of dimension d. An n-form *on V (where $0 \leq n \leq d$) is a type $(0, n)$-tensor $\omega \in \mathsf{T}^0_n V$ such that it is* totally antisymmetric*, i.e. for any n vectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ and any permutation $\sigma \in S_n$, we have*

$$\omega(\mathbf{v}_1, \dots, \mathbf{v}_n) = \operatorname{sgn}(\sigma) \cdot \omega(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(n)}). \tag{1.80}$$

For $n = 0$, we define the 0-form to be the scalar field K itself, which is trivially totally antisymmetric.

Special case of n-form is the *top form* on V, which is for $n = d$.

Proposition 1.37
A $(0, n)$-tensor ω is an n-form if and only if, $\omega(\mathbf{v}_1, \dots, \mathbf{v}_n) = 0$ whenever any $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly dependent.

Since any collection of $n > d$ vectors in V is linearly dependent, we can conclude that any n-form (for $n > d$) is identically zero.

We have a special symbol for the space of n-forms on V, which is denoted by $\Lambda^n V$. With this we have following important result

Theorem 1.38 (Dimension of n-forms)
Let V be a finite-dimensional vector space of dimension d. Then the dimension of the space of n-forms on V is given by

$$\dim \Lambda^n V = \begin{cases} \binom{d}{n} & \text{if } 0 \leq n \leq d \\ 0 & \text{if } n > d \end{cases} \tag{1.81}$$

In particular, the dimension of the space of top forms on V is 1, *i.e.* $\dim \Lambda^d V = 1$. This means that for any two top forms $\omega_1, \omega_2 \in \Lambda^d V$, there exists a scalar $\lambda \in K$ such that $\omega_1 = \lambda \cdot \omega_2$.

With this, we define *volume form* on V as a non-zero top form $\omega \in \Lambda^d V$. And a vector space V with a specified volume form is called a *vector space with volume form*.

Definition 1.39 (Volume Span):
Let V be a d-dimensional vector space with a volume form ω. Given a set of d vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_d\} \subseteq V$, the volume span *of these vectors is defined as*

$$\operatorname{vol}(\mathbf{v}_1, \dots, \mathbf{v}_d) := \omega(\mathbf{v}_1, \dots, \mathbf{v}_d). \tag{1.82}$$

Due to the antisymmetry of the volume form, we have $\mathrm{vol}(\mathbf{v}_1, \ldots, \mathbf{v}_d) = 0$ if and only if the vectors are linearly dependent. Indeed, in this case, these vectors at most span a $(d-1)$-dimensional hyperplane in V, which should have zero volume.

Now, we can define the determinant of a linear map $\phi \in \mathrm{End}(V)$ as follows:

Definition 1.40 (Determinant):
Let V be a d-dimensional vector space. The determinant *of a linear map $\phi \in \mathrm{End}(V)$ is defined as*

$$\det(\phi) := \frac{\omega(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\omega(\mathbf{e}_1, \ldots, \mathbf{e}_d)} \tag{1.83}$$

for some volume form ω on V and a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$ of V.

First and foremost, we need to check that this definition is independent of the choice of volume form and basis. For any other volume form $\tilde{\omega}$ on V, we can write

$$\tilde{\omega} = \lambda \cdot \omega \quad \Rightarrow \quad \det(\phi) = \frac{\tilde{\omega}(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\tilde{\omega}(\mathbf{e}_1, \ldots, \mathbf{e}_d)} = \frac{\lambda \cdot \omega(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\lambda \cdot \omega(\mathbf{e}_1, \ldots, \mathbf{e}_d)}. \tag{1.84}$$

So, the determinant is independent of the choice of volume form. Now, let $\{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_d\}$ be another basis of V. Then we can write

$$\begin{aligned}
\det(\phi) &= \frac{\omega(\phi(\tilde{\mathbf{e}}_1), \ldots, \phi(\tilde{\mathbf{e}}_d))}{\omega(\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_d)} \\
&= \frac{\omega(\phi(A^{i_1}{}_1\mathbf{e}_{i_1}), \ldots, \phi(A^{i_d}{}_d\mathbf{e}_{i_d}))}{\omega(A^{i_1}{}_1\mathbf{e}_{i_1}, \ldots, A^{i_d}{}_d\mathbf{e}_{i_d})} \\
&= \frac{A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\phi(\mathbf{e}_{i_1}), \ldots, \phi(\mathbf{e}_{i_d}))}{A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_d})}
\end{aligned}$$

say there is a permutation $\sigma_{i_1\ldots i_d}$ such that $\sigma_{i_1\ldots i_d}(i_1, \ldots, i_d) = (1, \ldots, d)$, then we can write

$$\begin{aligned}
&= \frac{A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\phi(\mathbf{e}_{\sigma_{i_1\ldots i_d}(1)}), \ldots, \phi(\mathbf{e}_{\sigma_{i_1\ldots i_d}(d)}))}{A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\mathbf{e}_{\sigma_{i_1\ldots i_d}(1)}, \ldots, \mathbf{e}_{\sigma_{i_1\ldots i_d}(d)})} \\
&= \frac{\mathrm{sgn}(\sigma_{i_1\ldots i_d}) \cdot A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\mathrm{sgn}(\sigma_{i_1\ldots i_d}) \cdot A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\mathbf{e}_1, \ldots, \mathbf{e}_d)}
\end{aligned}$$

call $\mathrm{sgn}(\sigma_{i_1\ldots i_d}) \cdot A^{i_1}{}_1 \cdots A^{i_d}{}_d = \lambda$, as all the indices are summed over, so we can write

$$= \frac{\lambda \cdot \omega(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\lambda \cdot \omega(\mathbf{e}_1, \ldots, \mathbf{e}_d)} = \det(\phi).$$

With this, we can conclude that the determinant is independent of the choice of basis. So we have

$$\det\left(\tilde{\Phi}\right) = \det\left(A^{-1} \bullet \Phi \bullet A\right) = \det\left(A^{-1}\right) \cdot \det(\Phi) \cdot \det(A) = \det(\Phi). \tag{1.85}$$

Recall the transformation of the bilinear form g under change of basis, we have

$$g \to A^\top g A. \tag{1.86}$$

So the determinant of the bilinear form g transforms as follows:

$$\det\left(A^\top g A\right) = \det\left(A^\top\right) \cdot \det(g) \cdot \det(A) = \det(A)^2 \cdot \det(g). \tag{1.87}$$

i.e. the determinant of a bilinear form transforms is not invariant under change of basis. But say there is another quantity, which transforms as follows:

$$X \to \frac{1}{\det(A)^2} X. \tag{1.88}$$

Then the quantity $\det(g)\cdot X$ is invariant under change of basis, *i.e.* it transforms as

$$\det(g)\cdot X \to \det(A)^2\cdot\det(g)\cdot\frac{1}{\det(A)^2}X = \det(g)\cdot X. \tag{1.89}$$

Here, it seems like two wrongs make a right, but this is not the case. In order to make this mathematically precise, we will have to introduce principal fibre bundles. Using them, we will be able to give a bundle definition of tensor and of tensor densities which are, loosely speaking, quantities that transform with powers of det A under a change of basis.

LECTURE 9

§1.3 Tangent Spaces to a Manifold

Let M be a smooth manifold (from now on, whenever we say a smooth manifold, the associated topology and atlas are always implied). Then we can construct the followwing vector space over $\mathbb{R}$:

$$(\mathcal{C}^\infty(M),+,\cdot) \tag{1.90}$$

where $\mathcal{C}^\infty(M) := \{f : M\to\mathbb{R} \mid f \text{ is smooth}\}$ is the set of all smooth real-valued functions on M. The notion of smoothness is via smooth charts in the atlas of M. The addition and scalar multiplication are defined pointwise, for all $f, g\in\mathcal{C}^\infty(M)$, $\lambda\in\mathbb{R}$, and $p\in M$, as

$$(f+g)(p) := f(p)+g(p) \tag{1.91}$$
$$(\lambda\cdot f)(p) := \lambda\cdot f(p). \tag{1.92}$$

It is easy to check that $(\mathcal{C}^\infty(M),+,\cdot)$ is indeed a vector space over $\mathbb{R}$.

Definition 1.41 (Directional Derivative):
Let $\gamma:\mathbb{R}\to M$ be a smooth curve[1] *through a point $p\in M$, and WLOG let $\gamma(0)=p$. Then the* directional derivative operator *along γ at p is a map*

$$X_{\gamma,p} : \mathcal{C}^\infty(M)\to\mathbb{R} \tag{1.93}$$

defined as

$$\mathcal{C}^\infty(M)\ni f\mapsto X_{\gamma,p}(f) := (f\circ\gamma)'(0)\in\mathbb{R}. \tag{1.94}$$

[1] Here the notion of smoothness is via charts in the atlas of M, *i.e.*, for all charts (U,x) of M such that $p\in U$, the composition $x\circ\gamma:\mathbb{R}\to\mathbb{R}^d$ is a smooth map in the usual sense.

Note the composition $f\circ\gamma$ is a map from $\mathbb{R}$ to $\mathbb{R}$, and hence the derivative is the usual derivative of real-valued functions of a real variable.

Proposition 1.42
The directional derivative operator $X_{\gamma,p}$ along a smooth curve γ through p is a linear map, *i.e.*, for all $f,g\in\mathcal{C}^\infty(M)$ and $\lambda,\mu\in\mathbb{R}$,

$$X_{\gamma,p}(\lambda f+\mu g) = \lambda X_{\gamma,p}(f)+\mu X_{\gamma,p}(g). \tag{1.95}$$

Proof:
This follows from the linearity of the usual derivative of real-valued functions of a real variable.

$$\begin{aligned} X_{\gamma,p}(\lambda f+\mu g) &= ((\lambda f+\mu g)\circ\gamma)'(0) \\ &= (\lambda(f\circ\gamma)+\mu(g\circ\gamma))'(0) \\ &= \lambda(f\circ\gamma)'(0)+\mu(g\circ\gamma)'(0) \\ &= \lambda X_{\gamma,p}(f)+\mu X_{\gamma,p}(g). \end{aligned}$$

Q.E.D.

In differential geometry, $X_{\gamma,p}$ is usually called a *tangent vector* to curve γ at point p. Physically, we can think of $X_{\gamma,p}$ as the velocity vector of a particle moving along the curve γ at point p. To see this, let $\gamma_1, \gamma_2 : \mathbb{R} \to M$ be two smooth curves through p such that $\gamma_1(0) = \gamma_2(0) = p$, and $\gamma_1(t) = \gamma_2(2t)$ for all $t \in \mathbb{R}$. Let $f \in \mathcal{C}^\infty(M)$ be a smooth function. Then

$$X_{\gamma_1,p}(f) = (f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0) \cdot 2 = 2(f \circ \gamma_2)'(0) = 2X_{\gamma_2,p}(f). \tag{1.96}$$

This means that $X_{\gamma_1,p} = 2X_{\gamma_2,p}$. So, if we think of γ_1 and γ_2 as the trajectories of two particles moving through point p on manifold M, then the velocity vector of the first particle at p is twice that of the second particle at p, which is consistent with our physical intuition.

Definition 1.43 (Tangent Vector Space):
Let M be a smooth manifold and $p \in M$. The tangent vector space *to M at p, denoted by $\mathcal{T}_pM$, is defined as*

$$\mathcal{T}_pM := \{X_{\gamma,p} \mid \gamma : \mathbb{R} \to M \text{ is a smooth curve with } \gamma(0) = p\}. \tag{1.97}$$

Equipped with following operations:

$$\begin{aligned}\oplus &: \mathcal{T}_pM \times \mathcal{T}_pM \to \mathcal{T}_pM,\\ \odot &: \mathbb{R} \times \mathcal{T}_pM \to \mathcal{T}_pM,\end{aligned}$$

defined pointwise as

$$\begin{aligned}(X_{\gamma_1,p} \oplus X_{\gamma_2,p})(f) &:= X_{\gamma_1,p}(f) + X_{\gamma_2,p}(f), \quad \forall f \in \mathcal{C}^\infty(M),\\ (\lambda \odot X_{\gamma,p})(f) &:= \lambda \cdot X_{\gamma,p}(f), \quad \forall f \in \mathcal{C}^\infty(M) \text{ and } \lambda \in \mathbb{R},\end{aligned}$$

$(\mathcal{T}_pM, \oplus, \odot)$ is a vector space over $\mathbb{R}$.

This definition is still incomplete, as the pointwise addition and scalar multiplication doesn't guarantee that the results are still in $\mathcal{T}_pM$. So, we have to prove the following proposition.

Proposition 1.44
The operations $\oplus$ and $\odot$ defined above are closed in $\mathcal{T}_pM$, *i.e.*, for all $X_{\gamma_1,p}, X_{\gamma_2,p} \in \mathcal{T}_pM$ and $\lambda \in \mathbb{R}$,

$$X_{\gamma_1,p} \oplus X_{\gamma_2,p} \in \mathcal{T}_pM, \tag{1.98}$$
$$\lambda \odot X_{\gamma,p} \in \mathcal{T}_pM. \tag{1.99}$$

So we need to show that there exist smooth curves $\gamma_3, \gamma_4 : \mathbb{R} \to M$ such that $\gamma_3(0) = \gamma_4(0) = p$ and

$$\begin{aligned}X_{\gamma_3,p} &= X_{\gamma_1,p} \oplus X_{\gamma_2,p},\\ X_{\gamma_4,p} &= \lambda \odot X_{\gamma,p}.\end{aligned}$$

Since the notion of derivative is local, so if two curves agree on a neighborhood of $0 \in \mathbb{R}$, then they have the same derivative at 0 *i.e.*, if $\gamma_1(t) = \gamma_2(t)$ for all t in some open interval containing 0, then $X_{\gamma_1,p} = X_{\gamma_2,p}$. So, it is sufficient to construct γ_3 and γ_4 on some open interval containing 0.

Proof:
Let (U, x) be a chart of M around p, *i.e.*, $p \in U$ and $x : U \to x(U) \subseteq \mathbb{R}^d$ is a homeomorphism.

Let $I \subseteq \mathbb{R}$ be an open interval containing 0 such that $\gamma(t), \gamma_1(t), \gamma_2(t) \in U$ for all $t \in I$. Such an interval exists since $\gamma, \gamma_1, \gamma_2$ are continuous and $\gamma(0) = \gamma_1(0) = \gamma_2(0) = p \in U$, and U is open in M.

1. Construct a curve $\gamma_3 : I \to M$ using the chart (U, x) as follows:

$$\gamma_3(t) := x^{-1} \circ (x \circ \gamma_1(t) + x \circ \gamma_2(t) - x(p)), \quad \forall t \in I. \tag{1.100}$$

   Note that $x \circ \gamma_1(t), x \circ \gamma_2(t) \in \mathbb{R}^d$ and $x(p) \in \mathbb{R}^d$, so the addition and subtraction are well-defined. Also, since x is a diffeomorphism, $x \circ \gamma_1$ and $x \circ \gamma_2$ are smooth maps from I to $\mathbb{R}^d$, and hence their sum is also a smooth map from I to $\mathbb{R}^d$. Therefore, the composition γ_3 is a smooth map from I to M. Moreover, $\gamma_3(0) = x^{-1}(x(p) + x(p) - x(p)) = p$.

Now, for all $f \in \mathcal{C}^\infty(M)$,

$$\begin{aligned} X_{\gamma_3,p}(f) &= (f \circ \gamma_3)'(0) \\ &= \Big(f \circ x^{-1} \circ (x \circ \gamma_1 + x \circ \gamma_2 - x(p))\Big)'(0) \end{aligned}$$

map $f \circ x^{-1} : x(U) \subseteq \mathbb{R}^d \to \mathbb{R}$, and $x \circ \gamma_1 + x \circ \gamma_2 - x(p) : I \subseteq \mathbb{R} \to \mathbb{R}^d$, so we can apply the multivariate chain rule, taking derivative in j-th coordinate direction, $j = 1, \dots, d$:

$$= \Big[\partial_j(f \circ x^{-1})(x(p))\Big] \cdot \Big(x^j \circ \gamma_1 + x^j \circ \gamma_2 - x^j(p)\Big)'(0)$$

where x^j is the j-th coordinate function of chart x. Here sum over j from 1 to d is implied. Using linearity of the usual derivative, we have:

$$\begin{aligned} &= \Big[\partial_j(f \circ x^{-1})(x(p))\Big] \cdot \Big((x^j \circ \gamma_1)'(0) + (x^j \circ \gamma_2)'(0)\Big) \\ &= [\partial_j(f \circ x^{-1})(x(p)) \cdot (x^j \circ \gamma_1)'(0)] + [\partial_j(f \circ x^{-1})(x(p)) \cdot (x^j \circ \gamma_2)'(0)] \end{aligned}$$

combining the term in the square brackets, we get:

$$\begin{aligned} &= (f \circ \gamma_1)'(0) + (f \circ \gamma_2)'(0) \\ &= X_{\gamma_1,p}(f) + X_{\gamma_2,p}(f) \\ &= (X_{\gamma_1,p} \oplus X_{\gamma_2,p})(f). \end{aligned}$$

Since this is true for all $f \in \mathcal{C}^\infty(M)$, we have

$$X_{\gamma_3,p} = X_{\gamma_1,p} \oplus X_{\gamma_2,p}. \tag{1.101}$$

2. Construct a curve $\gamma_4 : I \to M$ using the chart (U, x) as follows:

$$\gamma_4(t) := x^{-1} \circ (x \circ \gamma(\lambda t)), \quad \forall t \in I. \tag{1.102}$$

Here it is tempting rewrite $\gamma_4(t) = \gamma(\lambda t)$, but this lead to a problem that how to define $f'(p)$ when we find $X_{\gamma_4,p}(f)$.

Note that since x is a diffeomorphism, $x \circ \gamma$ is a smooth map from I to $\mathbb{R}^d$, and hence the composition γ_4 is also a smooth map from I to M. Moreover, $\gamma_4(0) = x^{-1}(x(p)) = p$.

Now, for all $f \in \mathcal{C}^\infty(M)$,

$$\begin{aligned} X_{\gamma_4,p}(f) &= (f \circ \gamma_4)'(0) \\ &= \Big(f \circ x^{-1} \circ (x \circ \gamma(\lambda t))\Big)'(0) \end{aligned}$$

map $f \circ x^{-1} : x(U) \subseteq \mathbb{R}^d \to \mathbb{R}$, and $x \circ \gamma(\lambda t) : I \subseteq \mathbb{R} \to \mathbb{R}^d$, so we can apply the multivariate chain rule, taking derivative in j-th coordinate direction, $j = 1, \dots, d$:

$$= \Big[\partial_j(f \circ x^{-1})(x(p))\Big] \cdot \Big(x^j \circ \gamma(\lambda t)\Big)'(0)$$

where x^j is the j-th coordinate function of chart x. Using the chain rule for the usual derivative, we have:

$$\begin{aligned} &= \Big[\partial_j(f \circ x^{-1})(x(p))\Big] \cdot \Big[(x^j \circ \gamma)'(0) \cdot \lambda\Big] \\ &= \lambda[\partial_j(f \circ x^{-1})(x(p)) \cdot (x^j \circ \gamma)'(0)] \end{aligned}$$

combining the term in the square brackets, we get:

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

Since this is true for all $f \in \mathcal{C}^\infty(M)$, we have

$$X_{\gamma_4,p} = \lambda \odot X_{\gamma,p}. \tag{1.103}$$

Q.E.D.

Remark 1.45 (Independence of Chart Choice). The construction of γ_3 and γ_4 depends on the choice of chart (U, x). However, the resulting tangent vectors $X_{\gamma_3,p}$ and $X_{\gamma_4,p}$ do not depend on the choice of chart. This is because if we choose another chart (V, y) around p, then the transition map $y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$ is a diffeomorphism between open subsets of $\mathbb{R}^d$, and hence the construction of γ_3 and γ_4 using chart (V, y) will yield the same tangent vectors $X_{\gamma_3,p}$ and $X_{\gamma_4,p}$.

§1.3.1 Algebras and Derivations

Definition 1.46 (Algebra over a Field):
Let $(V, +, \cdot)$ be a vector space over a field K equipped with a "product" operation,

$$\bullet : V \times V \to V, \tag{1.104}$$

such that $\bullet$ is bilinear. Then $(V, +, \cdot, \bullet)$ is called an algebra over field K.

In the future, we will impose more conditions on the product operation $\bullet$, such as anti-symmetry to get something called a Lie algebra. A typical example for that is the cross product in $\mathbb{R}^3$.

Example 1.47 (Algebra of Smooth Functions)
We have already seen that $(\mathcal{C}^\infty(M), +, \cdot)$ is a vector space over $\mathbb{R}$. Now, we can define a product operation $\bullet$ on $\mathcal{C}^\infty(M)$ as follows:

$$\begin{aligned} \bullet : \mathcal{C}^\infty(M) \times \mathcal{C}^\infty(M) &\to \mathcal{C}^\infty(M), \\ (f, g) &\mapsto f \bullet g \end{aligned} \tag{1.105}$$

where $(f \bullet g)(p) := f(p) \cdot g(p)$ for all $p \in M$. It is easy to check that $\bullet$ is bilinear, and hence $(\mathcal{C}^\infty(M), +, \cdot, \bullet)$ is an algebra over $\mathbb{R}$.

Note the difference between $\cdot$ and $\bullet$: the former is scalar multiplication, while the latter is function multiplication, both at heart uses the field multiplication in $\mathbb{R}$.

Let's look at some special algebras where the product operation satisfies some special properties.

Definition 1.48:
Let $(V, +, \cdot, \bullet)$ be an algebra over a field K. The algebra is called:

- Associative *if for all $u, v, w \in V$,*

$$(u \bullet v) \bullet w = u \bullet (v \bullet w). \tag{1.106}$$

- Commutative *if for all $u, v \in V$,*

$$u \bullet v = v \bullet u. \tag{1.107}$$

- Unital *if there exists an element $\mathbf{1} \in V$ such that*

$$\mathbf{1} \bullet v = v \bullet \mathbf{1} = v, \quad \forall v \in V. \tag{1.108}$$

Now let's look at more important class of algebras, which are not necessarily associative or commutative.

Definition 1.49 (Lie Algebra):
A Lie algebra *over a field K is an algebra $(V, +, \cdot, [\star, \star])$ such that the product operation $[\star, \star]$, called the* Lie bracket*, satisfies the following properties:*

- Antisymmetry*: for all $u, v \in V$,*
$$[u, v] = -[v, u]. \tag{1.109}$$
- Jacobi Identity*: for all $u, v, w \in V$,*
$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0. \tag{1.110}$$

Note that the 0 here is the additive identity of the vector space $(V, +, \cdot)$.

It is easy to see that for a non-trivial Lie bracket, the algebra cannot be unital.

Definition 1.50 (Derivation):
Let $(V, +, \cdot, \bullet)$ be an algebra over a field K. A derivation *on V is a linear map $D : V \to V$ such that it satisfies the Leibniz rule:*
$$D(u \bullet v) = D(u) \bullet v + u \bullet D(v), \quad \forall u, v \in V. \tag{1.111}$$

Example 1.51 (Derivation on Smooth Functions)
We have already seen that $(\mathcal{C}^\infty(M), +, \cdot, \bullet)$ is an algebra over $\mathbb{R}$. Fix a point $p \in M$, take any tangent vector $X_{\gamma,p} \in \mathcal{T}_p M$. We know from the definition that $X_{\gamma,p} : \mathcal{C}^\infty(M) \to \mathbb{R}$ is a linear map. Now let's check if it satisfies the Leibniz rule, for all $f, g \in \mathcal{C}^\infty(M)$,
$$\begin{aligned} X_{\gamma,p}(f \bullet g) &= ((f \bullet g) \circ \gamma)'(0) \\ &= ((f \circ \gamma) \cdot (g \circ \gamma))'(0) \\ &= (f \circ \gamma)'(0) \cdot (g \circ \gamma)(0) + (f \circ \gamma)(0) \cdot (g \circ \gamma)'(0) \\ &= X_{\gamma,p}(f) \cdot g(p) + f(p) \cdot X_{\gamma,p}(g). \end{aligned}$$
So, $X_{\gamma,p}$ satisfies the Leibniz rule. However, note that $X_{\gamma,p}(f)$ is a real number, not a smooth function on M. So, $X_{\gamma,p}$ is not a derivation on the algebra $(\mathcal{C}^\infty(M), +, \cdot, \bullet)$, usually called a derivation at point p.

Now define a map
$$\begin{aligned} D : \mathcal{C}^\infty(M) &\to \mathcal{C}^\infty(M), \\ f &\mapsto D(f) \end{aligned} \tag{1.112}$$
where $D(f)(p) := X_{\gamma,p}(f)$ for all $p \in M$. Then D is a derivation on the algebra $(\mathcal{C}^\infty(M), +, \cdot, \bullet)$.

Example 1.52
Let V be the vector space over $\mathbb{R}$, define $A := \mathrm{End}(V)$, we know that $(A, +, \cdot)$ is a vector space over $\mathbb{R}$. Now define a product operation on A as follows:
$$\begin{aligned} [\star, \star] : A \times A &\to A, \\ (\phi, \psi) &\mapsto [\phi, \psi] := \phi \circ \psi - \psi \circ \phi, \end{aligned} \tag{1.113}$$
where $\circ$ is the composition of linear maps. It is easy to see that $[\star, \star]$ is bilinear, and hence $(A, +, \cdot, [\star, \star])$ is an algebra over $\mathbb{R}$. Moreover, $[\star, \star]$ is antisymmetric, and for all $\phi, \psi, \rho \in A$,
$$[\phi, [\psi, \rho]] + [\psi, [\rho, \phi]] + [\rho, [\phi, \psi]] = 0, \tag{1.114}$$
which is called the *Jacobi identity*. So, $(A, +, \cdot, [\star, \star])$ is a Lie algebra over $\mathbb{R}$.

Now fix $H \in A$, define a map

$$\begin{aligned} D_H : A &\to A, \\ \phi &\mapsto D_H(\phi) := [H, \phi]. \end{aligned} \tag{1.115}$$

Let's check if D_H is a derivation, for all $\phi, \psi \in A$,

$$D_H([\phi, \psi]) = [H, [\phi, \psi]]$$

using the Jacobi identity, we have:

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

rearranging the terms and use antisymmetry, we get:

$$\begin{aligned} &= [[H, \phi], \psi] + [\phi, [H, \psi]] \\ &= [D_H(\phi), \psi] + [\phi, D_H(\psi)]. \end{aligned}$$

So, D_H is a derivation on the Lie algebra $(A, +, \cdot, [\star, \star])$.

With this example, we can see the algebraic structure of Poisson brackets in classical mechanics, and the commutator in quantum mechanics.

Remark 1.53 (Poisson Bracket). In classical mechanics, the state of a system is represented by a point in phase space (which is a symplectic manifold), and observables are represented by smooth functions on the phase space. The Poisson bracket defines a Lie algebra structure on the space of observables. If we fix an observable H (the Hamiltonian), then the map $D_H(f) := \{H, f\}$ is a derivation on the Lie algebra of observables, which generates the time evolution of the system according to Hamilton's equations.

Remark 1.54 (Commutator in Quantum Mechanics). Similarly, in quantum mechanics, the state of a system is represented by a vector in a Hilbert space, and observables are represented by self-adjoint operators on that space. The commutator defines a Lie algebra structure on the space of observables. If we fix an observable H (the Hamiltonian operator), then the map $D_H(\phi) := [H, \phi]$ is a derivation on the Lie algebra of observables, which generates the time evolution of the system according to the Heisenberg equation of motion.

§1.3.2 Basis and Dimension of Tangent Space

We have shown that for a smooth manifold M and a point $p \in M$, the set of tangent vectors $\mathcal{T}_pM$ is a vector space over $\mathbb{R}$. Now we will prove a very crucial theorem in differential geometry, which states that the dimension of the tangent space $\mathcal{T}_pM$ is equal to the dimension of the manifold M.

Theorem 1.55 (Dimension of Tangent Space)
Let M be a smooth manifold of dimension d, then for all $p \in M$, the tangent vector space $\mathcal{T}_pM$ is a vector space over $\mathbb{R}$ of dimension d.

$$\dim(\mathcal{T}_pM) = \dim(M) = d. \tag{1.116}$$

Note that we have used the same symbol dim for the dimension of a manifold and the dimension of a vector space, but they are different concepts. The dimension of a manifold is defined as the dimension of the Euclidean space that it locally resembles, while the dimension of a vector space is defined as the cardinality of its basis.

Proof:
Fix a point $p \in M$, and fix a chart (U, x) of M around p.

To prove this theorem, we will construct a basis of $\mathcal{T}_pM$ consisting of d tangent vectors.

Define d curves $\gamma_j : \mathbb{R} \to U$, $j = 1, \ldots, d$, such that

$$\gamma_j(0) = p; \qquad x^i \circ \gamma_j(t) = \delta^i_j t, \quad \forall t \in \mathbb{R}, \tag{1.117}$$

where x^i is the i-th coordinate function of chart x, and δ^i_j is the Kronecker delta. So pictorially, γ_j is a curve that moves along the j-th coordinate axis in the Euclidean space $\mathbb{R}^d$ under the chart x.

Name the corresponding tangent vectors at p as

$$e_j := X_{\gamma_j,p}, \quad j = 1, \ldots, d. \tag{1.118}$$

Let's look at how e_j acts on a smooth function $f \in \mathcal{C}^\infty(M)$:

$$e_j(f) = (f \circ \gamma_j)'(0) = (f \circ \mathrm{id}_U \circ \gamma_j)'(0)$$

insert the identity map $\mathrm{id}_U = x^{-1} \circ x$ on U:

$$\begin{aligned}
&= \Big(f \circ x^{-1} \circ (x \circ \gamma_j)\Big)'(0) \\
&= \Big[\partial_i (f \circ x^{-1})(x(p))\Big] \cdot \Big(x^i \circ \gamma_j\Big)'(0) \\
&= \Big[\partial_i (f \circ x^{-1})(x(p))\Big] \cdot \delta^i_j \\
&= \partial_j (f \circ x^{-1})(x(p)).
\end{aligned}$$

Define a formal symbol as

$$\left(\frac{\partial}{\partial x^j}\right)_p (f) := \partial_j (f \circ x^{-1})(x(p)), \quad \forall f \in \mathcal{C}^\infty(M). \tag{1.119}$$

Don't confuse this notation with the usual partial derivative of a function of several real variables.

$$\begin{aligned}
&(\partial_j)_p : \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}) \to \mathbb{R}, \\
&\left(\frac{\partial}{\partial x^j}\right)_p : \mathcal{C}^\infty(M) \to \mathbb{R}.
\end{aligned}$$

So we have

$$e_j = \left(\frac{\partial}{\partial x^j}\right)_p, \quad j = 1, \ldots, d. \tag{1.120}$$

Define the set

$$\mathcal{B} := \{e_1, \ldots, e_d\} = \left\{\left(\frac{\partial}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial}{\partial x^d}\right)_p\right\}. \tag{1.121}$$

We will show that $\mathcal{B}$ is a basis of $\mathcal{T}_pM$, *i.e.*, for any $X \in \mathcal{T}_pM$, there exist unique real numbers $X^1, \ldots, X^d$ such that

$$X = X^j e_j = X^j \left(\frac{\partial}{\partial x^j}\right)_p. \qquad \text{(sum over } j \text{ from 1 to } d \text{ is implied)} \tag{1.122}$$

1. **Spanning**: We know $\exists \gamma : \mathbb{R} \to M$ such that $X = X_{\gamma,p}$. For all $f \in \mathcal{C}^\infty(M)$,

   $$X(f) = (f \circ \gamma)'(0) = (f \circ \mathrm{id}_U \circ \gamma)'(0)$$

   insert the identity map $\mathrm{id}_U = x^{-1} \circ x$ on U:

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

   Note that $x^i \circ \gamma : \mathbb{R} \to \mathbb{R}$ is a smooth function of a real variable, so $(x^i \circ \gamma)'(0) \in \mathbb{R}$. Define

   $$X^i := (x^i \circ \gamma)'(0) \in \mathbb{R}, \quad i = 1, \ldots, d. \tag{1.123}$$

   So we have

   $$X(f) = X^i \left(\frac{\partial}{\partial x^i}\right)_p (f) = X^i e_i(f), \quad \forall f \in \mathcal{C}^\infty(M). \tag{1.124}$$

   Since this is true for all $f \in \mathcal{C}^\infty(M)$, we have

   $$X = X^i e_i. \tag{1.125}$$

   Thus, $\mathcal{T}_pM = \mathrm{span}(\mathcal{B})$.

> Remark 1.56 (Smoothness of Chart map and co-ordinate functions). In general to talk about smooth of any function $f : M \to \mathbb{R}$, we have used charts such that f is smooth if and only if $f \circ x^{-1} : x(U) \subseteq \mathbb{R}^d \to \mathbb{R}$.
>
> So by this definition, the chart map $x : U \to x(U) \subseteq \mathbb{R}^d$ is trivially smooth, since $x \circ x^{-1} = \mathrm{id}_{x(U)}$ is smooth. Similarly, the coordinate functions $x^i : U \to \mathbb{R}$ are also smooth, since $x^i \circ x^{-1} : x(U) \subseteq \mathbb{R}^d \to \mathbb{R}$ is just the projection onto the i-th coordinate, which is a linear map and hence smooth.

2. **Linear Independence**: Suppose that $\mathcal{B}$ is linearly dependent, then there exist real numbers $X^1, \dots, X^d$, not all zero, such that

$$X^j e_j = 0. \tag{1.126}$$

So for all $f \in \mathcal{C}^\infty(M)$, we have $X^j e_j(f) = 0$. In particular, take $f = x^i$, the i-th coordinate function of chart x, then

$$\begin{aligned} 0 = X^j e_j(x^i) = X^j \left(\frac{\partial}{\partial x^j}\right)_p (x^i) &= X^j \partial_j (x^i \circ x^{-1})(x(p)) \\ = X^j \partial_j (\mathrm{proj}^i)(x(p)) = X^j \delta^i_j = X^i. \end{aligned}$$

Since this is true for all $i = 1, \dots, d$, we have $X^1 = X^2 = \dots = X^d = 0$, which contradicts our assumption. Hence, $\mathcal{B}$ is linearly independent.

Therefore, $\mathcal{B}$ is a basis of $\mathcal{T}_p M$, and $\dim(\mathcal{T}_p M) = d$. Q.E.D.

Terminology: Let $X \in \mathcal{T}_p M$, then we have

$$X = X^j \left(\frac{\partial}{\partial x^j}\right)_p, \tag{1.127}$$

where $X^j = X(x^j) = (x^j \circ \gamma)'(0)$ are called the *components* of X with respect to the basis $\mathcal{B}$ induced by the chart (U, x).

> Remark 1.57 (Dependence on Chart Choice). The basis $\mathcal{B}$ of $\mathcal{T}_p M$ constructed above depends on the choice of chart (U, x). If we choose another chart (V, y) around p, then the transition map $y \circ x^{-1} : x(U \cap V) \to y(U \cap V)$ is a diffeomorphism between open subsets of $\mathbb{R}^d$. The new basis induced by chart (V, y) will be
>
> $$\tilde{\mathcal{B}} = \left\{ \left(\frac{\partial}{\partial y^1}\right)_p, \dots, \left(\frac{\partial}{\partial y^d}\right)_p \right\}. \tag{1.128}$$
>
> The relationship between the two bases can be expressed using the Jacobian matrix of the transition map. Specifically, for each j,
>
> $$\left(\frac{\partial}{\partial y^j}\right)_p = \left[\frac{\partial y^i}{\partial x^j}(x(p))\right] \left(\frac{\partial}{\partial x^i}\right)_p, \tag{1.129}$$
>
> where $\frac{\partial y^i}{\partial x^j}$ is the (i, j)-th entry of the Jacobian matrix of the transition map $y \circ x^{-1}$ evaluated at $x(p)$. This shows how the basis of the tangent space transforms under a change of coordinates.

Usually in physics, when we talk about position vector, we mean the coordinates functions x^i in some chart. Say we go to a different chart y (coordinate transformation) which has high non-linear dependence on x, then the position vector in the new chart y is not simply related to the old position vector in chart x by a linear transformation, but rather by a non-linear transformation. However, the tangent vectors (velocity vectors) transform linearly under the change of coordinates, as shown above. So the notion of position vector, and its transformation is ill-defined in general.

This becomes very important when we study general relativity, where the spacetime is modeled as a 4-dimensional Lorentzian manifold. In some older physics literature, you may find the term position vector being used for the coordinates x^μ, which is not a well-defined concept.