

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](#).

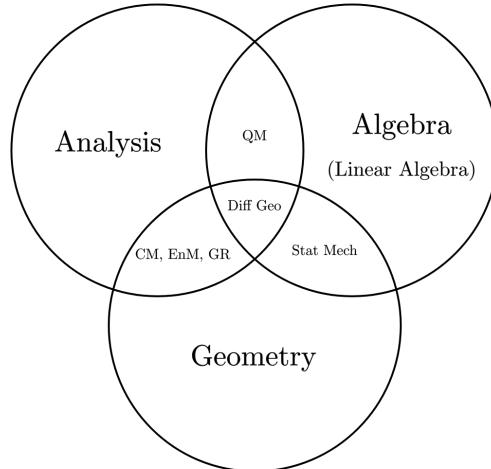


Figure 0.1: Structure of Theoretical Physics

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

Structure of this Course	ii
Contents	iii
1 Differentiable Manifolds	1
Lecture 8	1
1.1 Review of Vector Spaces	1
1.1.1 Tensors	4
1.1.2 Dimension of a Vector Space	5
1.1.3 Change of Basis	9
1.1.4 Determinants	11
1.1.4.1 Permutation Group	12
Lecture 9	14
1.2 Tangent Spaces to a Manifold	14
1.2.1 Algebras and Derivations	17
1.2.2 Basis and Dimension of Tangent Space	20
1.2.3 Change of Basis and Coordinate Transformation	22
1.2.3.1 Position Vector in Physics	23
Lecture 10: Construction of Tangent Bundle	23
1.3 Cotangent Spaces and Gradient	23
1.3.1 Basis of Cotangent Space	24
1.4 Push-Forward and Pull-Back	25

1 Differentiable Manifolds

LECTURE 8

§1.1 Review of Vector Spaces

Definition 1.1 ((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 $a, b, c \in K$,

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c. \quad (1.1)$$

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 1.2 (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 $a, b, c \in R$,

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c. \quad (1.2)$$

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 1.3 (Vector Space over a Field K):

A vector space V over a field $(K, +, \cdot)$ is a set equipped with two operations $\oplus : V \times V \rightarrow V$ (vector addition) and $\odot : K \times V \rightarrow V$ (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 1.4 ((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.

$$\forall \mathbf{u}, \mathbf{v} \in U : \mathbf{u} \oplus \mathbf{v} \in U \quad \text{and} \quad \forall a \in K : \forall \mathbf{u} \in U : a \odot \mathbf{u} \in U. \quad (1.3)$$

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 1.5 (Linear Map):

Let (V, \oplus, \odot) and (W, \boxplus, \boxdot) be two vector spaces over the same field K . A map $L : V \rightarrow W$ 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 $L : V \rightarrow W$ is bijective, it is called a linear isomorphism. We write $V \cong_{\text{vec}} W$ if there exists a linear isomorphism between V and W .

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

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in V, \forall a, b \in K : L(a \odot \mathbf{v}_1 \oplus b \odot \mathbf{v}_2) = a \boxdot L(\mathbf{v}_1) \boxplus b \boxdot L(\mathbf{v}_2). \quad (1.4)$$

Remark 1.6 (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 $L : V \rightarrow W$ is bijective to conclude that its inverse $L^{-1} : W \rightarrow V$ is also a linear map.

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

$$\text{Hom}(V, W) := \{L : V \rightarrow W \mid L \text{ is a linear map}\} \equiv \{L : V \xrightarrow{\sim} W\} \quad (1.5)$$

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

Proposition 1.7

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

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

$$(L_1, L_2) \mapsto L_1 \oplus L_2 \quad (1.6)$$

where

$$L_1 \oplus L_2 : V \xrightarrow{\sim} W, \quad \mathbf{v} \mapsto (L_1 \oplus L_2)(\mathbf{v}) := L_1(\mathbf{v}) \boxplus L_2(\mathbf{v}). \quad (1.7)$$

- $\diamond : K \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W)$ defined by

$$(a, L) \mapsto a \diamond L \quad (1.8)$$

where

$$a \diamond L : V \xrightarrow{\sim} W, \quad \mathbf{v} \mapsto (a \diamond L)(\mathbf{v}) := a \square L(\mathbf{v}). \quad (1.9)$$

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

Proof:

We need to show that $\text{Hom}(V, W)$ is a vector space over K with the operations \oplus and \diamond .

- Closure under addition: Let $L_1, L_2 \in \text{Hom}(V, W)$. Then $L_1 \oplus L_2$ is defined as

$$(L_1 \oplus L_2)(\mathbf{v}) = L_1(\mathbf{v}) \boxplus L_2(\mathbf{v}). \quad (1.10)$$

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

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

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

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

$$(a \diamond L)(\mathbf{v}) = a \square L(\mathbf{v}). \quad (1.11)$$

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

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

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

We have shown that $\text{Hom}(V, W)$ is closed under both operations \diamondsuit and \diamond . Thus, $\text{Hom}(V, W)$ 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 $\oplus : M \times M \rightarrow M$ (addition) and $\odot : R \times M \rightarrow M$ (scalar multiplication), we call this a *module* over the ring R .

Remark 1.8 (Case of $\text{Hom}(V, W)$ as a module). In case of modules, we can still define $\text{Hom}(V, W)$ as a set of all linear maps from V to W over a ring R . But as in general, ring multiplication is not commutative, $\text{Hom}(V, W)$ is not a module over R .

Some useful terminology:

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

$$\text{End}(V) := \text{Hom}(V, V) = \{L : V \xrightarrow{\sim} V\} \quad (1.12)$$

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

$$\text{Aut}(V) := \{L : V \xrightarrow{\sim} V \mid L \text{ is a linear isomorphism}\}. \quad (1.13)$$

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

- 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 :

$$V^* := \text{Hom}(V, K) \quad (1.14)$$

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

§1.1.1 Tensors

Definition 1.9 (Tensor):

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

$$T : \underbrace{V^* \times \cdots \times V^*}_{p \text{ times}} \times \underbrace{V \times \cdots \times V}_{q \text{ times}} \rightarrow K \quad (1.15)$$

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

$$\mathbb{T}_q^p V := \underbrace{V \otimes \cdots \otimes V}_{p \text{ times}} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q \text{ times}} = \{T \mid T \text{ is type } (p, q)\text{-tensor}\} \quad (1.16)$$

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

1. Addition of tensors:

$$\oplus : \mathbb{T}_q^p V \times \mathbb{T}_q^p V \rightarrow \mathbb{T}_q^p V \quad (1.17)$$

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

$$(T_1 \oplus T_2)(\omega_1, \dots, \omega_p, v_1, \dots, v_q) = T_1(\omega_1, \dots, \omega_p, v_1, \dots, v_q) + T_2(\omega_1, \dots, \omega_p, v_1, \dots, v_q) \quad (1.18)$$

2. Scalar multiplication of tensors:

$$\odot : K \times \mathbb{T}_q^p V \rightarrow \mathbb{T}_q^p V \quad (1.19)$$

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

$$(a \odot T)(\omega_1, \dots, \omega_p, v_1, \dots, v_q) = a \cdot T(\omega_1, \dots, \omega_p, v_1, \dots, v_q) \quad (1.20)$$

Proposition 1.10

$T_q^p V$ 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 1.11 (Tensor Product):

Let $T_1 \in T_q^p V$ and $T_2 \in T_s^r V$ be two tensors. The tensor product $T_1 \otimes T_2$ is defined as a tensor of type $(p + r, q + s)$ given by

$$(T_1 \otimes T_2)(\omega_1, \dots, \omega_{p+r}, \mathbf{v}_1, \dots, \mathbf{v}_{q+s}) = T_1(\omega_1, \dots, \omega_p, \mathbf{v}_1, \dots, \mathbf{v}_q) \cdot T_2(\omega_{p+1}, \dots, \omega_{p+r}, \mathbf{v}_{q+1}, \dots, \mathbf{v}_{q+s}). \quad (1.21)$$

Example 1.12

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

- (a) The set $T_0^0 V = K$ is the set of all scalar tensors, which is just the field K itself.
- (b) The set $T_1^0 V = V^* := \text{Hom}(V, K)$.
- (c) The set $T_1^1 V \equiv V \otimes V^* = \{T \mid T : V^* \times V \rightarrow K \text{ is a linear in both argument}\} \cong_{\text{vec}} \text{End}(V^*)$.

Pf: Let $T \in T_1^1 V$ be a tensor. We need to construct a linear map $L \in \text{End}(V^*)$ using T . For each $\omega \in V^*$, we define a linear map $L_\omega : V \rightarrow K \in V^*$ by

$$\forall \mathbf{v} \in V : L_\omega(\mathbf{v}) = T(\omega, \mathbf{v}). \quad (1.22)$$

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

$$L(\omega) = L_\omega \quad \Rightarrow \quad \omega \mapsto \underbrace{(\mathbf{v} \mapsto T(\omega, \mathbf{v}))}_{L_\omega}. \quad (1.23)$$

Now, let $L \in \text{End}(V^*)$ be a linear map. We need to show that there exists a tensor $T \in T_1^1 V$ corresponding to L . Define $T : V^* \times V \rightarrow K$ by

$$T(\omega, \mathbf{v}) = L(\omega)(\mathbf{v}) \quad (1.24)$$

as $L(\omega) \in V^*$ is a linear map from V to K . □

- (d) From previous example, we expect that $T_0^1 V \cong_{\text{vec}} V$; but this is not true in general. However, if V is finite-dimensional, then $T_0^1 V \cong_{\text{vec}} V$.
- (e) Similarly, $T_1^1 V \cong_{\text{vec}} \text{End}(V)$ is also not true in general, but if V is finite-dimensional, then $T_1^1 V \cong_{\text{vec}} \text{End}(V)$.
- (f) All these examples which are not true in general, but true for finite-dimensional vector spaces, can be summarized as

$$(V^*)^* \not\cong_{\text{vec}} V. \quad (1.25)$$

This is true only for finite-dimensional vector spaces.

§1.1.2 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 1.13 (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 $\{\mathbf{b}_1, \dots, \mathbf{b}_N\} \subseteq \mathcal{B}$, the only solution to the equation

$$\sum_{i=1}^N \lambda^i \cdot \mathbf{b}_i = \mathbf{0} \quad \text{for } \lambda^i \in K \quad (1.26)$$

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.

$$\forall \mathbf{v} \in V : \exists M \in \mathbb{N} : \exists v^1, \dots, v^M \in K : \exists \mathbf{b}_1, \dots, \mathbf{b}_M \in \mathcal{B} : \mathbf{v} = \sum_{i=1}^M v^i \cdot \mathbf{b}_i. \quad (1.27)$$

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, \dots, 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 1.14

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 $\text{span}(S) = V$, then $|S| \geq |\mathcal{B}|$.
- (ii) If S is linearly independent, then $|S| \leq |\mathcal{B}|$.

Definition 1.15 (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 $\dim V$.

$$\dim V := |\mathcal{B}| \quad \text{for any Hamel basis } \mathcal{B} \subseteq V. \quad (1.28)$$

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 1.16

Let V be a finite-dimensional vector space. Then

$$(V^*)^* \cong_{\text{vec}} V. \quad (1.29)$$

Proof:

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

$$\begin{aligned} L : V &\rightarrow (V^*)^* \\ \mathbf{v} &\mapsto L_{\mathbf{v}} \end{aligned} \quad (1.30)$$

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

$$\begin{aligned} L_{\mathbf{v}} : V^* &\rightarrow K \\ \omega &\mapsto L_{\mathbf{v}}(\omega) := \omega(\mathbf{v}) \end{aligned} \quad (1.31)$$

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 $\text{span}(\mathbf{v}) \subseteq V$ is a proper subspace of V , so we can extend this to a Hamel basis $\mathcal{B} = \{\mathbf{v}, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of V . Thus, any $\mathbf{u} \in V$ can be expressed as

$$\mathbf{u} = b \cdot \mathbf{v} + \sum_{i=2}^n b^i \cdot \mathbf{e}_i \quad (1.32)$$

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

$$\mathbf{u} \mapsto \boldsymbol{\omega}(\mathbf{u}) := b. \quad (1.33)$$

Thus, we have

$$L(\mathbf{v})(\boldsymbol{\omega}) = \boldsymbol{\omega}(\mathbf{v}) = 1 \neq 0. \quad (1.34)$$

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 $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, there exists a unique dual basis $\mathcal{B}^* = \{\mathbf{f}^1, \dots, \mathbf{f}^n\}$ such that

$$\forall i, j \in \{1, \dots, n\} : \mathbf{f}^i(\mathbf{e}_j) = \delta_j^i := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad (1.35)$$

(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 $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of V . Now, we can define a vector $\mathbf{v} \in V$ as follows:

$$\mathbf{v} = \sum_{i=1}^n F(\mathbf{f}^i) \cdot \mathbf{e}_i \quad (1.36)$$

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

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

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

Q.E.D.

Corollary 1.17

If V is a finite-dimensional vector space, then

$$(i) \quad T_1^1 V \cong_{\text{vec}} \text{End}(V).$$

$$(ii) \quad T_0^1 V \cong_{\text{vec}} V.$$

Till now, we have only used basis to classify vector spaces, either finite-dimensional or infinite-dimensional. In vector theory, 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 1.18 (Components of a Tensor):

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

$$T^{i_1 \dots i_p}_{j_1 \dots j_q} := T(\boldsymbol{\epsilon}^{i_1}, \dots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_q}) \quad \text{for } i_k, j_l \in \{1, \dots, n\}. \quad (1.37)$$

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

$$T = \underbrace{\sum_{i_1=1}^n \dots \sum_{j_q=1}^n}_{p+q \text{ sums}} T^{i_1 \dots i_p}_{j_1 \dots j_q} \cdot \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \dots \otimes \boldsymbol{\epsilon}^{j_q}. \quad (1.38)$$

Here \otimes denotes the tensor product of tensors, as $\mathbf{e}_{i_k} \in V \cong_{\text{vec}} T_0^1 V$ and $\boldsymbol{\epsilon}^{j_l} \in V^* \cong_{\text{vec}} T_1^0 V$. 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

$$\begin{aligned} T &= \underbrace{\sum_{i_1=1}^n \dots \sum_{j_q=1}^n}_{p+q \text{ sums}} T^{i_1 \dots i_p}_{j_1 \dots j_q} \cdot \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \dots \otimes \boldsymbol{\epsilon}^{j_q} \equiv \\ &\quad T^{i_1 \dots i_p}_{j_1 \dots j_q} \cdot \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \dots \otimes \boldsymbol{\epsilon}^{j_q}. \end{aligned} \quad (1.39)$$

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.

$$T^{i_1 \dots i_p}_{j_1 \dots j_q} := T(\boldsymbol{\epsilon}^{i_1}, \dots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_q}). \quad (1.40)$$

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

$$\mathbf{v} = v^i \cdot \mathbf{e}_i \quad \text{where } v^i = \boldsymbol{\epsilon}^i(\mathbf{v}) \quad (1.41)$$

$$\boldsymbol{\omega} = \omega_j \cdot \boldsymbol{\epsilon}^j \quad \text{where } \omega_j = \mathbf{e}_j(\boldsymbol{\omega}). \quad (1.42)$$

Remark 1.19 (Caution). But we need to be careful about this convention, as it only works for linear spaces and linear maps. For example, if we have a tensor $T : V^* \times V \rightarrow K$, then for any $\omega \in V^*$ and $\mathbf{v} \in V$, we have

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

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*.

§1.1.3 Change of Basis

Consider two bases $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\tilde{\mathcal{B}} = \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_n\}$ of a vector space V . 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)

$$\tilde{\mathbf{e}}_i = A^j_i \mathbf{e}_j \quad (1.43)$$

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

$$\mathbf{e}_i = B^j_i \tilde{\mathbf{e}}_j \quad (1.44)$$

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

$$A^{-1} = B \quad \text{and} \quad B^{-1} = A. \quad (1.45)$$

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 1.20 (Matrix Notation). Fix a basis $\mathcal{B} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of a vector space V and a dual basis $\mathcal{B}^* = \{\epsilon^1, \dots, \epsilon^n\}$. Then we will express

$$\mathbf{v} = v^i \mathbf{e}_i \quad \rightsquigarrow \quad \mathbf{v} \hat{=} \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \quad (1.46)$$

and

$$\omega = \omega_j \epsilon^j \quad \rightsquigarrow \quad \omega^\top \hat{=} (\omega_1 \ \dots \ \omega_n). \quad (1.47)$$

Now let $\phi \in \text{End}(V) \cong_{\text{vec}} \mathsf{T}_1^1 V$ be a linear map. Then we can express ϕ as

$$\phi = \phi^i_j \mathbf{e}_i \otimes \epsilon^j \quad \rightsquigarrow \quad \Phi \hat{=} \begin{pmatrix} \phi^1_1 & \cdots & \phi^1_n \\ \vdots & \ddots & \vdots \\ \phi^n_1 & \cdots & \phi^n_n \end{pmatrix}. \quad (1.48)$$

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 $\text{End}(V) \cong_{\text{vec}} T_1^1 V$. Let $\phi \in \text{End}(V)$, so we can rethink ϕ as a type $(1, 1)$ -tensor in following sense:

$$\phi(\omega, \mathbf{v}) := \omega(\phi(\mathbf{v})). \quad (1.49)$$

So the components of tensor ϕ is defined as,

$$\phi^i_j := \phi(\epsilon^i, \mathbf{e}_j) = \epsilon^i(\phi(\mathbf{e}_j)). \quad (1.50)$$

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

$$\begin{aligned} (\phi \circ \psi)^i_j &= \epsilon^i((\phi \circ \psi)(\mathbf{e}_j)) \\ &= \epsilon^i(\phi(\psi(\mathbf{e}_j))) \\ &= \epsilon^i(\phi(\psi^a_j \mathbf{e}_a)) \\ &= \epsilon^i(\psi^a_j \phi(\mathbf{e}_a)) \\ &= \epsilon^i(\psi^a_j \phi^b_a \mathbf{e}_b) \\ &= \psi^a_j \cdot \phi^b_a \epsilon^i(\mathbf{e}_b) \\ &= \psi^a_j \cdot \phi^b_a \cdot \delta^i_b \\ &= \psi^a_j \cdot \phi^i_a = \phi^i_a \cdot \psi^a_j. \end{aligned}$$

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

$$\Phi \cdot \Psi \hat{=} \left(\phi^i_a \cdot \psi^a_j \right)_{i,j=1}^n \quad (1.51)$$

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

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

$$\phi(\omega, \mathbf{v}) = \omega_i \phi^i_j v^j \quad \rightsquigarrow \quad \omega^\top \cdot \Phi \cdot \mathbf{v} \quad (1.52)$$

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

$$\tilde{\mathbf{e}}_i = A^j_i \mathbf{e}_j \quad \mathbf{e}_j = B^i_j \tilde{\mathbf{e}}_i. \quad (1.53)$$

Since these relations are invertible, we have

$$A^i_m B^m_j = B^i_m A^m_j = \delta^i_j \quad (1.54)$$

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

1. Covectors:

Let $\omega = \omega_j \epsilon^j$ be a covector, where $\omega_j = \mathbf{e}_j(\omega)$. Here \mathbf{e}_j is an element of $T_0^1 V$. Then we have

$$\omega_j = \mathbf{e}_j(\omega) = \omega(\mathbf{e}_j) = \omega(B^i_j \tilde{\mathbf{e}}_i) = B^i_j \omega(\tilde{\mathbf{e}}_i) = B^i_j \tilde{\mathbf{e}}_i(\omega) = B^i_j \tilde{\omega}_i. \quad (1.55)$$

2. Vectors:

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

$$v^i = \epsilon^i(\mathbf{v}) = \epsilon^i(\tilde{v}^j \tilde{\mathbf{e}}_j) = \tilde{v}^j \epsilon^i(\tilde{\mathbf{e}}_j) = \tilde{v}^j \epsilon^i(A^k_j \mathbf{e}_k) = \tilde{v}^j A^k_j \epsilon^i(\mathbf{e}_k) = \tilde{v}^j A^k_j \delta^i_k = A^i_j \tilde{v}^j. \quad (1.56)$$

Summarizing the above two equations, we have

$$v^i = A_j^i \tilde{v}^j \quad \omega_j = B_j^i \tilde{\omega}_i \quad (1.57)$$

$$\tilde{v}^i = B_j^i v^j \quad \tilde{\omega}_j = A_j^i \omega_i. \quad (1.58)$$

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

$$\epsilon^i(\mathbf{e}_j) = \delta_j^i \quad \tilde{\epsilon}^i(\tilde{\mathbf{e}}_j) = \delta_j^i. \quad (1.59)$$

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

$$\tilde{\epsilon}^i(\tilde{\mathbf{e}}_j) = C_k^i \epsilon^k(A_l^j \mathbf{e}_l) = C_k^i A_l^j \epsilon^k(\mathbf{e}_l) = C_k^i A_l^j \delta_l^k = C_k^i A^k_j = \delta_j^i. \quad (1.60)$$

Thus, using unique existence of the inverse, we have

$$C_j^i = B_j^i. \quad (1.61)$$

So we have the following relations for the dual basis:

$$\tilde{\epsilon}^i = B_j^i \epsilon^j \quad \epsilon^j = A^j_i \tilde{\epsilon}^i. \quad (1.62)$$

3. Tensors:

Let $T \in \mathbb{T}_q^p V$ be a tensor, then the components of T with respect to the basis \mathcal{B} and \mathcal{B}^* are given by

$$T^{i_1 \dots i_p}_{j_1 \dots j_q} = T(\epsilon^{i_1}, \dots, \epsilon^{i_p}, \mathbf{e}_{j_1}, \dots, \mathbf{e}_{j_q}). \quad (1.63)$$

Now, we can express the components of T with respect to the basis $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{B}}^*$ as follows:

$$\tilde{T}^{i_1 \dots i_p}_{j_1 \dots j_q} = T(\tilde{\epsilon}^{i_1}, \dots, \tilde{\epsilon}^{i_p}, \tilde{\mathbf{e}}_{j_1}, \dots, \tilde{\mathbf{e}}_{j_q}) \quad (1.64)$$

$$= T(B^{i_1}_{a_1} \epsilon^{a_1}, \dots, B^{i_p}_{a_p} \epsilon^{a_p}, A^{b_1}_{j_1} \mathbf{e}_{b_1}, \dots, A^{b_q}_{j_q} \mathbf{e}_{b_q}) \quad (1.65)$$

$$= B^{i_1}_{a_1} \cdots B^{i_p}_{a_p} \cdot A^{b_1}_{j_1} \cdots A^{b_q}_{j_q} \cdot T(\epsilon^{a_1}, \dots, \epsilon^{a_p}, \mathbf{e}_{b_1}, \dots, \mathbf{e}_{b_q}) \quad (1.66)$$

$$= B^{i_1}_{a_1} \cdots B^{i_p}_{a_p} \cdot A^{b_1}_{j_1} \cdots A^{b_q}_{j_q} \cdot T^{a_1 \dots a_p}_{b_1 \dots b_q}. \quad (1.67)$$

Thus, the reversal of the basis is given by

$$T^{i_1 \dots i_p}_{j_1 \dots j_q} = A^{i_1}_{a_1} \cdots A^{i_p}_{a_p} \cdot B^{b_1}_{j_1} \cdots B^{b_q}_{j_q} \cdot \tilde{T}^{a_1 \dots a_p}_{b_1 \dots b_q}. \quad (1.68)$$

§1.1.4 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 purely due to the fact that we are using a witchcraft called *matrices* to represent linear maps.

Remark 1.21. Let $\phi \in \mathbb{T}_1^1 V$ be an endomorphism and $g \in \mathbb{T}_2^0 V$ be a bilinear form. We have

$$\phi = \phi^i_j \mathbf{e}_i \otimes \epsilon^j \quad \text{and} \quad g = g_{ij} \epsilon^i \otimes \epsilon^j. \quad (1.69)$$

We can arrange the components of ϕ and g in a matrix form as follows:

$$\Phi = \begin{pmatrix} \phi^1_1 & \cdots & \phi^1_n \\ \vdots & \ddots & \vdots \\ \phi^n_1 & \cdots & \phi^n_n \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \cdots & g_{nn} \end{pmatrix}. \quad (1.70)$$

But say, we change the basis of V to $\tilde{\mathcal{B}}$, then the components of ϕ and g change as follows:

$$\tilde{\phi}^i_j = B_k^i \cdot \phi^k_l \cdot A_l^j \quad \text{and} \quad \tilde{g}_{ij} = A_i^k \cdot g_{kl} \cdot A_l^j. \quad (1.71)$$

Recall $(A^{-1})^i_j = B_j^i$, so we can write the above equations as (remember that $A^i_j = A_{ij}$, which give rise to

transpose in the expression)

$$\tilde{\Phi} = A^{-1} \cdot \Phi \cdot A \quad \text{and} \quad \tilde{G} = A^\top \cdot G \cdot A. \quad (1.72)$$

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 is the case, we have to look at basis free definition of determinants.

§1.1.4.1 Permutation Group

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

Definition 1.22 (Permutation and Permutation Group):

A permutation of a non-empty set M is a bijective function $\sigma : M \rightarrow M$.

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:

$$\operatorname{sgn}(\sigma) := \begin{cases} +1 & \text{if } \sigma \text{ is the product of an even number of transpositions} \\ -1 & \text{if } \sigma \text{ is the product of an odd number of transpositions} \end{cases} \quad (1.73)$$

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 1.23 (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 $\omega \in T_n^0 V$ 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

$$\omega(\mathbf{v}_1, \dots, \mathbf{v}_n) = \operatorname{sgn}(\sigma) \cdot \omega(\mathbf{v}_{\sigma(1)}, \dots, \mathbf{v}_{\sigma(n)}). \quad (1.74)$$

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 1.24

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 1.25 (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

$$\dim \Lambda^n V = \begin{cases} \binom{d}{n} & \text{if } 0 \leq n \leq d \\ 0 & \text{if } n > d \end{cases} \quad (1.75)$$

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 1.26 (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

$$\text{vol}(\mathbf{v}_1, \dots, \mathbf{v}_d) := \omega(\mathbf{v}_1, \dots, \mathbf{v}_d). \quad (1.76)$$

Due to the antisymmetry of the volume form, we have $\text{vol}(\mathbf{v}_1, \dots, \mathbf{v}_d) = 0$ 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 $\phi \in \text{End}(V)$ as follows:

Definition 1.27 (Determinant):

Let V be a d -dimensional vector space. The determinant of a linear map $\phi \in \text{End}(V)$ is defined as

$$\det(\phi) := \frac{\omega(\phi(\mathbf{e}_1), \dots, \phi(\mathbf{e}_d))}{\omega(\mathbf{e}_1, \dots, \mathbf{e}_d)} \quad (1.77)$$

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

$$\tilde{\omega} = \lambda \cdot \omega \Rightarrow \det(\phi) = \frac{\tilde{\omega}(\phi(\mathbf{e}_1), \dots, \phi(\mathbf{e}_d))}{\tilde{\omega}(\mathbf{e}_1, \dots, \mathbf{e}_d)} = \frac{\lambda \cdot \omega(\phi(\mathbf{e}_1), \dots, \phi(\mathbf{e}_d))}{\lambda \cdot \omega(\mathbf{e}_1, \dots, \mathbf{e}_d)}. \quad (1.78)$$

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

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

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

$$= \frac{A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\phi(\mathbf{e}_{\sigma_{i_1 \dots i_d}(1)}), \dots, \phi(\mathbf{e}_{\sigma_{i_1 \dots i_d}(d)}))}{A^{i_1}{}_1 \cdots A^{i_d}{}_d \cdot \omega(\mathbf{e}_{\sigma_{i_1 \dots i_d}(1)}, \dots, \mathbf{e}_{\sigma_{i_1 \dots i_d}(d)})}$$

$$= \frac{\operatorname{sgn}(\sigma_{i_1 \dots i_d}) \cdot A^{i_1}_1 \cdots A^{i_d}_d \cdot \omega(\phi(\mathbf{e}_1), \dots, \phi(\mathbf{e}_d))}{\operatorname{sgn}(\sigma_{i_1 \dots i_d}) \cdot A^{i_1}_1 \cdots A^{i_d}_d \cdot \omega(\mathbf{e}_1, \dots, \mathbf{e}_d)}$$

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

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

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

$$\det(\tilde{\Phi}) = \det(A^{-1} \cdot \Phi \cdot A) = \det(A^{-1}) \cdot \det(\Phi) \cdot \det(A) = \det(\Phi). \quad (1.79)$$

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

$$g \rightarrow A^\top g A. \quad (1.80)$$

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

$$\det(A^\top g A) = \det(A^\top) \cdot \det(g) \cdot \det(A) = \det(A)^2 \cdot \det(g). \quad (1.81)$$

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:

$$X \rightarrow \frac{1}{\det(A)^2} X. \quad (1.82)$$

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

$$\det(g) \cdot X \rightarrow \det(A)^2 \cdot \det(g) \cdot \frac{1}{\det(A)^2} X = \det(g) \cdot X. \quad (1.83)$$

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

§1.2 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 following vector space over \mathbb{R} :

$$(\mathcal{C}^\infty(M), +, \cdot) \quad (1.84)$$

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

$$(f + g)(p) := f(p) + g(p) \quad (1.85)$$

$$(\lambda \cdot f)(p) := \lambda \cdot f(p). \quad (1.86)$$

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

Definition 1.28 (Directional Derivative):

Let $\gamma : \mathbb{R} \rightarrow M$ be a smooth curve¹ through a point $p \in M$, and WLOG let $\gamma(0) = p$. Then the directional derivative operator along γ at p is a map

$$X_{\gamma,p} : \mathcal{C}^\infty(M) \rightarrow \mathbb{R} \quad (1.87)$$

defined as

$$\mathcal{C}^\infty(M) \ni f \mapsto X_{\gamma,p}(f) := (f \circ \gamma)'(0) \in \mathbb{R}. \quad (1.88)$$

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 $x \circ \gamma : \mathbb{R} \rightarrow \mathbb{R}^d$ 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 1.29

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 $\lambda, \mu \in \mathbb{R}$,

$$X_{\gamma,p}(\lambda f + \mu g) = \lambda X_{\gamma,p}(f) + \mu X_{\gamma,p}(g). \quad (1.89)$$

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 $\gamma_1, \gamma_2 : \mathbb{R} \rightarrow M$ 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

$$X_{\gamma_1,p}(f) = (f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0) \cdot 2 = 2(f \circ \gamma_2)'(0) = 2X_{\gamma_2,p}(f). \quad (1.90)$$

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 1.30 (Tangent Vector Space):

Let M be a smooth manifold and $p \in M$. The tangent vector space to M at p , denoted by $\mathcal{T}_p M$, is defined as

$$\mathcal{T}_p M := \{X_{\gamma,p} \mid \gamma : \mathbb{R} \rightarrow M \text{ is a smooth curve with } \gamma(0) = p\}. \quad (1.91)$$

Equipped with following operations:

$$\begin{aligned} \oplus : \mathcal{T}_p M \times \mathcal{T}_p M &\rightarrow \mathcal{T}_p M, \\ \odot : \mathbb{R} \times \mathcal{T}_p M &\rightarrow \mathcal{T}_p M, \end{aligned}$$

defined pointwise as

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

$(\mathcal{T}_p M, \oplus, \odot)$ 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 $\mathcal{T}_p M$. So, we have to prove the following proposition.

Proposition 1.31

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

$$X_{\gamma_1,p} \oplus X_{\gamma_2,p} \in \mathcal{T}_p M, \quad (1.92)$$

$$\lambda \odot X_{\gamma,p} \in \mathcal{T}_p M. \quad (1.93)$$

So we need to show that there exist smooth curves $\gamma_3, \gamma_4 : \mathbb{R} \rightarrow M$ 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 $x : U \rightarrow x(U) \subseteq \mathbb{R}^d$ 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 $\gamma_3 : I \rightarrow M$ using the chart (U, x) as follows:

$$\gamma_3(t) := x^{-1} \circ (x \circ \gamma_1(t) + x \circ \gamma_2(t) - x(p)), \quad \forall t \in I. \quad (1.94)$$

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

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

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

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

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:

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

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

$$X_{\gamma_3,p} = X_{\gamma_1,p} \oplus X_{\gamma_2,p}. \quad (1.95)$$

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

$$\gamma_4(t) := x^{-1} \circ (x \circ \gamma(\lambda t)), \quad \forall t \in I. \quad (1.96)$$

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

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

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

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

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

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

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

$$X_{\gamma_4,p} = \lambda \odot X_{\gamma,p}. \quad (1.97)$$

Q.E.D.

Remark 1.32 (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 $y \circ x^{-1} : x(U \cap V) \rightarrow y(U \cap V)$ 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 $X_{\gamma_4,p}$.

§1.2.1 Algebras and Derivations

Definition 1.33 (Algebra over a Field):

Let $(V, +, \cdot)$ be a vector space over a field K equipped with a “product” operation,

$$\bullet : V \times V \rightarrow V, \quad (1.98)$$

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 1.34 (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:

$$\begin{aligned}\bullet : \mathcal{C}^\infty(M) \times \mathcal{C}^\infty(M) &\rightarrow \mathcal{C}^\infty(M), \\ (f, g) &\mapsto f \bullet g\end{aligned}\tag{1.99}$$

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 1.35:

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

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

$$(u \bullet v) \bullet w = u \bullet (v \bullet w).\tag{1.100}$$

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

$$u \bullet v = v \bullet u.\tag{1.101}$$

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

$$\mathbf{1} \bullet v = v \bullet \mathbf{1} = v, \quad \forall v \in V.\tag{1.102}$$

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

Definition 1.36 (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 $u, v \in V$,

$$[u, v] = -[v, u].\tag{1.103}$$

- Jacobi Identity: for all $u, v, w \in V$,

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0.\tag{1.104}$$

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 1.37 (Derivation):

Let $(V, +, \cdot, \bullet)$ be an algebra over a field K . A derivation on V is a linear map $D : V \rightarrow V$ such that it satisfies the Leibniz rule:

$$D(u \bullet v) = D(u) \bullet v + u \bullet D(v), \quad \forall u, v \in V.\tag{1.105}$$

Example 1.38 (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 $X_{\gamma,p} \in T_p M$. We know from the definition that $X_{\gamma,p} : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}$ is a linear map. Now let's check if it satisfies the Leibniz rule, for all $f, g \in \mathcal{C}^\infty(M)$,

$$X_{\gamma,p}(f \bullet g) = ((f \bullet g) \circ \gamma)'(0)$$

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

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 $(C^\infty(M), +, \cdot, \bullet)$, usually called a derivation at point p .

Now define a map

$$\begin{aligned}
D : C^\infty(M) &\rightarrow C^\infty(M), \\
f &\mapsto D(f)
\end{aligned} \tag{1.106}$$

where $D(f)(p) := X_{\gamma,p}(f)$ for all $p \in M$. Then D is a derivation on the algebra $(C^\infty(M), +, \cdot, \bullet)$.

Example 1.39

Let V be the vector space over \mathbb{R} , define $A := \text{End}(V)$, we know that $(A, +, \cdot)$ is a vector space over \mathbb{R} . Now define a product operation on A as follows:

$$\begin{aligned}
[\star, \star] : A \times A &\rightarrow A, \\
(\phi, \psi) &\mapsto [\phi, \psi] := \phi \circ \psi - \psi \circ \phi,
\end{aligned} \tag{1.107}$$

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 $\phi, \psi, \rho \in A$,

$$[\phi, [\psi, \rho]] + [\psi, [\rho, \phi]] + [\rho, [\phi, \psi]] = 0, \tag{1.108}$$

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

$$\begin{aligned}
D_H : A &\rightarrow A, \\
\phi &\mapsto D_H(\phi) := [H, \phi].
\end{aligned} \tag{1.109}$$

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 1.40 (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 1.41 (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.

§1.2.2 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 $\mathcal{T}_p M$ 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 $\mathcal{T}_p M$ is equal to the dimension of the manifold M .

Theorem 1.42 (Dimension of Tangent Space)

Let M be a smooth manifold of dimension d , then for all $p \in M$, the tangent vector space $\mathcal{T}_p M$ is a vector space over \mathbb{R} of dimension d .

$$\dim(\mathcal{T}_p M) = \dim(M) = d. \quad (1.110)$$

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 $\mathcal{T}_p M$ consisting of d tangent vectors.

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

$$\gamma_j(0) = p; \quad x^i \circ \gamma_j(t) = \delta_j^i t, \quad \forall t \in \mathbb{R}, \quad (1.111)$$

where x^i is the i -th coordinate function of chart x , and δ_j^i 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

$$\mathbf{e}_j := X_{\gamma_j, p}, \quad j = 1, \dots, d. \quad (1.112)$$

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

$$\mathbf{e}_j(f) = (f \circ \gamma_j)'(0) = (f \circ \text{id}_U \circ \gamma_j)'(0)$$

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

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

Define a formal symbol as

$$\left(\frac{\partial}{\partial x^j} \right)_p (f) := \partial_j(f \circ x^{-1})(x(p)), \quad \forall f \in \mathcal{C}^\infty(M). \quad (1.113)$$

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

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

So we have

$$\mathbf{e}_j = \left(\frac{\partial}{\partial x^j} \right)_p, \quad j = 1, \dots, d. \quad (1.114)$$

Define the set

$$\mathcal{B} := \{\mathbf{e}_1, \dots, \mathbf{e}_d\} = \left\{ \left(\frac{\partial}{\partial x^1} \right)_p, \dots, \left(\frac{\partial}{\partial x^d} \right)_p \right\}. \quad (1.115)$$

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

$$X = X^j \mathbf{e}_j = X^j \left(\frac{\partial}{\partial x^j} \right)_p. \quad (\text{sum over } j \text{ from 1 to } d \text{ is implied}) \quad (1.116)$$

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

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

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

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

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

$$X^i := (x^i \circ \gamma)'(0) \in \mathbb{R}, \quad i = 1, \dots, d. \quad (1.117)$$

So we have

$$X(f) = X^i \left(\frac{\partial}{\partial x^i} \right)_p (f) = X^i \mathbf{e}_i(f), \quad \forall f \in \mathcal{C}^\infty(M). \quad (1.118)$$

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

$$X = X^i \mathbf{e}_i. \quad (1.119)$$

Thus, $\mathcal{T}_p M = \text{span}(\mathcal{B})$.

Remark 1.43 (Smoothness of Chart map and co-ordinate functions). In general to talk about smooth of any function $f : M \rightarrow \mathbb{R}$, we have used charts such that f is smooth if and only if $f \circ x^{-1} : x(U) \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$.

So by this definition, the chart map $x : U \rightarrow x(U) \subseteq \mathbb{R}^d$ is trivially smooth, since $x \circ x^{-1} = \text{id}_{x(U)}$ is smooth. Similarly, the coordinate functions $x^i : U \rightarrow \mathbb{R}$ are also smooth, since $x^i \circ x^{-1} : x(U) \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ 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

$$X^j \mathbf{e}_j = 0. \quad (1.120)$$

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

$$\begin{aligned} 0 &= X^j \mathbf{e}_j(x^i) = X^j \left(\frac{\partial}{\partial x^j} \right)_p (x^i) = X^j \partial_j(x^i \circ x^{-1})(x(p)) \\ &= X^j \partial_j(\text{proj}^i)(x(p)) = X^j \delta_j^i = X^i. \end{aligned}$$

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

$$X = X^j \left(\frac{\partial}{\partial x^j} \right)_p, \quad (1.121)$$

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) .

§1.2.3 Change of Basis and Coordinate Transformation

In the construction of the basis \mathcal{B} of $\mathcal{T}_p M$, we have used a chart (U, x) around p . Now suppose we have another chart (V', y) around p , we can similarly construct another basis of $\mathcal{T}_p M$ as

$$\tilde{\mathcal{B}} := \{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_d\} = \left\{ \left(\frac{\partial}{\partial y^1} \right)_p, \dots, \left(\frac{\partial}{\partial y^d} \right)_p \right\}. \quad (1.122)$$

For simplicity, define $V := U \cap V'$, which is also a chart neighborhood of p , now we have two charts (V, x) and (V, y) around p .

These two charts induce two different bases \mathcal{B} and $\tilde{\mathcal{B}}$ of the same vector space $\mathcal{T}_p M$, so there must exist a change of basis map between them $A, B \in \text{End}(\mathcal{T}_p M)$ such that

$$\tilde{\mathbf{e}}_j = A^i{}_j \mathbf{e}_i, \quad \mathbf{e}_j = B^i{}_j \tilde{\mathbf{e}}_i, \quad (1.123)$$

where $A^i{}_j, B^i{}_j \in \mathbb{R}$ are the components of the linear maps A and B with respect to the basis \mathcal{B} . Since change of basis maps are invertible, we have the following relation

$$A^i{}_j B^k{}_i = \delta^k{}_j = B^i{}_j A^k{}_i. \quad (1.124)$$

Now let's find the explicit form of the components $A^i{}_j$ and $B^i{}_j$. Take x^k , the k -th coordinate function of chart x , and apply $\tilde{\mathbf{e}}_j$ on it, we have

$$\begin{aligned} \tilde{\mathbf{e}}_j(x^k) &= A^i{}_j \mathbf{e}_i(x^k) \\ \left(\frac{\partial}{\partial y^j} \right)_p(x^k) &= A^i{}_j \left(\frac{\partial}{\partial x^i} \right)_p(x^k) \\ \partial_j(x^k \circ y^{-1})(y(p)) &= A^i{}_j \partial_i(x^k \circ x^{-1})(x(p)) \\ \partial_j(x^k \circ y^{-1})(y(p)) &= A^i{}_j \delta_i^k \\ \partial_j(x^k \circ y^{-1})(y(p)) &= A^k{}_j. \end{aligned}$$

So we have

$$A^k{}_j = \partial_j(x^k \circ y^{-1})(y(p)). \quad (1.125)$$

Similarly, take y^k , the k -th coordinate function of chart y , and apply \mathbf{e}_j on it, we have

$$B^k{}_j = \partial_j(y^k \circ x^{-1})(x(p)). \quad (1.126)$$

With this let's look at the transformation of components of a tangent vector $X \in \mathcal{T}_p M$ under change of basis. We have

$$X = X^j \mathbf{e}_j = \tilde{X}^j \tilde{\mathbf{e}}_j, \quad (1.127)$$

where $X^j = X(x^j)$ and $\tilde{X}^j = X(y^j)$ are the components of X with respect to the bases \mathcal{B} and $\tilde{\mathcal{B}}$ respectively. Using the change of basis relation (1.123), we have

$$X = X^j \mathbf{e}_j = X^j B^i{}_j \tilde{\mathbf{e}}_i = \tilde{X}^i \tilde{\mathbf{e}}_i.$$

Since $\tilde{\mathcal{B}}$ is a basis, we have

$$\tilde{X}^i = X^j B^i{}_j = X^j \partial_j(y^i \circ x^{-1})(x(p)). \quad (1.128)$$

Now, let's write this component of change of basis in a more familiar form, *i.e.*, the Jacobian matrix of the coordinate transformation. Note that the transition map between the two charts (V, x) and (V, y) is given by

$$y^i \circ x^{-1} : x(V) \subseteq \mathbb{R}^d \rightarrow \mathbb{R}. \quad (1.129)$$

Abusing the notation a bit, for now call $y^i \circ x^{-1}$ as y^i , so y^i is a real-valued smooth function of d real variables. Since we are only interested in a single point $p \in V$, we can think of x^j as independent variables for the function y^i . So we have

$$y^i = y^i(x^1, \dots, x^d). \quad (1.130)$$

With this new interpretation, we have

$$\partial_j(y^i \circ x^{-1})(x(p)) = \frac{\partial y^i}{\partial x^j}(x(p)). \quad (1.131)$$

Similarly, we can also write

$$\partial_j(x^i \circ y^{-1})(y(p)) = \frac{\partial x^i}{\partial y^j}(y(p)). \quad (1.132)$$

So the change of basis relation (1.123) can be rewritten as

$$\tilde{\mathbf{e}}_j = \frac{\partial y^i}{\partial x^j}(x(p))\mathbf{e}_i, \quad \mathbf{e}_j = \frac{\partial x^i}{\partial y^j}(y(p))\tilde{\mathbf{e}}_i, \quad (1.133)$$

and the transformation of components of a tangent vector $X \in \mathcal{T}_p M$ under change of basis can be rewritten as

$$\tilde{X}^i = X^j \frac{\partial y^i}{\partial x^j}(x(p)). \quad (1.134)$$

Remark 1.44 (Jacobian Matrix). Let $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a smooth map, then the Jacobian matrix of f at a point $a \in \mathbb{R}^m$ is defined as the $n \times m$ matrix whose (i, j) -th entry is given by

$$\frac{\partial f^i}{\partial x^j}(a), \quad (1.135)$$

where f^i is the i -th component function of f , and x^j is the j -th coordinate of \mathbb{R}^m .

In older literature, you may see that the transformation law (see (1.134)) is being used to define a ‘vector’. It goes like this: an array of d real numbers (X^1, \dots, X^d) is called a vector if under a change of coordinates $x \rightarrow y = y(x)$, the components transform as

$$\tilde{X}^i = X^j \frac{\partial y^i}{\partial x^j}. \quad (1.136)$$

This definition is not wrong, but it is not very useful, since it does not tell you what a vector actually is.

§1.2.3.1 Position Vector in Physics

In physics, we often talk about the position vector of a point in space. From all this discussion, only coordinate functions x^i looks like a good candidate for a position vector. However, this does not make sense, as we have seen that vectors have a specific transformation law under change of coordinates. Let’s consider a chart (U, y) such that y is highly non-linear, for example, polar coordinates in \mathbb{R}^2 . So the transformation of components of y^i under change of basis will not be linear, and hence y^i cannot be a vector.

This shows that the concept of position vector in physics is not very well-defined in the context of differential geometry. In fact, the position vector is not a vector in the sense of differential geometry, but rather an image of a point on a manifold under a chart map.

LECTURE 10

§1.3 Cotangent Spaces and Gradient

Definition 1.45 (Cotangent Space):

Let M be a smooth manifold and $p \in M$. The cotangent space at p , denoted $\mathcal{T}_p^* M$, is the dual space of the tangent space $\mathcal{T}_p M$:

$$\mathcal{T}_p^* M := (\mathcal{T}_p M)^*. \quad (1.137)$$

Since the manifold is finite-dimensional², the tangent space $\mathcal{T}_p M$ is a finite-dimensional vector space. Therefore, its dual space $\mathcal{T}_p^* M$ is also a finite-dimensional vector space of the same dimension, moreover, we have the isomorphism

$$\mathcal{T}_p M \cong_{\text{vec}} \mathcal{T}_p^* M. \quad (1.138)$$

However, this isomorphism is not canonical, *i.e.*, there is no natural way to identify vectors in $\mathcal{T}_p M$ with covectors in $\mathcal{T}_p^* M$ without introducing some additional structure on these tangent and cotangent spaces.

Now, using these vector and dual vector spaces, we can define tensor spaces at a point $p \in M$,

$$\mathcal{T}_s^r(\mathcal{T}_p M) := \left\{ t : \underbrace{\mathcal{T}_p^* M \times \cdots \times \mathcal{T}_p^* M}_{r \text{ times}} \times \underbrace{\mathcal{T}_p M \times \cdots \times \mathcal{T}_p M}_{s \text{ times}} \rightarrow \mathbb{R} \mid t \text{ is multilinear} \right\}. \quad (1.139)$$

This is set (even vector space) of all (r, s) -type tensors at the point p .

Definition 1.46 (Gradient):

Let $f \in \mathcal{C}^\infty(M)$. Then at each point $p \in M$, we have a linear map

$$\begin{aligned} d_p : \mathcal{C}^\infty(M) &\rightarrow \mathcal{T}_p^* M \\ f &\mapsto d_p f, \end{aligned} \quad (1.140)$$

where $d_p f$ is defined by

$$d_p f(X) := X(f), \quad \forall X \in \mathcal{T}_p M. \quad (1.141)$$

The linear map d_p is called the gradient operator at the point p , and $d_p f$ is called the gradient of the function f at the point p .

Observe that the gradient $d_p f$ is a covector at the point p not a vector. Also, the gradient of a function f at a point p is “orthogonal” to the level set of f passing through the point p , *i.e.*, let a tangent vector $X \in \mathcal{T}_p M$ be tangent to the level set³ of f at the point p , then we have

$$d_p f(X) = X(f) = 0,$$

since the function f is constant on its level set. Thus, the gradient $d_p f$ annihilates all tangent vectors to the level set of f at the point p .

§1.3.1 Basis of Cotangent Space

We have defined the gradient operator at a point $p \in M$, using which we can construct a basis for the cotangent space $\mathcal{T}_p^* M$.

Theorem 1.47 (Basis of Cotangent Space)

Let M be a smooth manifold and $p \in M$. Let (U, x) be a chart around the point p . Then the set

$$\{d_p x^1, d_p x^2, \dots, d_p x^{\dim M}\} \quad (1.142)$$

where x^i is the i -th component function of the chart x , forms a basis of the cotangent space $\mathcal{T}_p^* M$.

Proof:

Since the dimension of the cotangent space $\mathcal{T}_p^* M$ is equal to the dimension of the manifold M , it is

²The way we have defined manifolds in this course, they are finite-dimensional by default as we always map them to \mathbb{R}^d . For infinite-dimensional manifolds, we need to more sophisticated tools like Banach and Hilbert spaces, which are beyond the scope of this course.

³The level set of a function $f : M \rightarrow \mathbb{R}$ corresponding to a value $c \in \mathbb{R}$ is defined as the set $\mathcal{L}_c(f) := \{p \in M \mid f(p) = c\}$.

sufficient to show that the set $\{\mathbf{d}_p x^i\}_{i=1}^{\dim M}$ is linearly independent. Suppose not the set is linearly dependent, then there exist real numbers $\omega_i \in \mathbb{R}$, not all zero, such that

$$\omega_i \mathbf{d}_p x^i = 0. \quad (\text{sum over } i \text{ from 1 to } \dim M \text{ is implied}).$$

So for any tangent vector $X \in \mathcal{T}_p M$, we have $\omega_i \mathbf{d}_p x^i(X) = 0$. In particular, let us choose $X = \left(\frac{\partial}{\partial x^j}\right)_p$, the j -th basis vector of the tangent space $\mathcal{T}_p M$ induced by the chart (U, x) . Then we have

$$\begin{aligned} 0 &= \omega_i \mathbf{d}_p x^i \left(\left(\frac{\partial}{\partial x^j} \right)_p \right) = \omega_i \left(\frac{\partial}{\partial x^j} \right)_p (x^i) = \omega_i \partial_j (x^i \circ x^{-1})(x(p)) \\ &= \omega_i \partial_j (\text{proj}^i)(x(p)) = \omega_i \delta_j^i = \omega_j. \end{aligned}$$

Here, $\text{proj}^i : \mathbb{R}^{\dim M} \rightarrow \mathbb{R}$ is the projection map onto the i -th coordinate. Since j was arbitrary, we have $\omega_j = 0$ for all $j = 1, 2, \dots, \dim M$, which contradicts our assumption that not all ω_i are zero. Thus, the set $\{\mathbf{d}_p x^i\}_{i=1}^{\dim M}$ is linearly independent and hence forms a basis of the cotangent space $\mathcal{T}_p^* M$. Q.E.D.

Moreover, we have the following duality relation between the basis of the tangent space and the basis of the cotangent space induced by the same chart (U, x) :

$$\mathbf{d}_p x^i \left(\left(\frac{\partial}{\partial x^j} \right)_p \right) = \delta_j^i. \quad (1.143)$$

Thus, the basis $\{\mathbf{d}_p x^i\}_{i=1}^{\dim M}$ is *dual basis* to the basis $\left\{ \left(\frac{\partial}{\partial x^i} \right)_p \right\}_{i=1}^{\dim M}$.

§1.4 Push-Forward and Pull-Back