Lecture Notes

Geometrical Anatomy of Theoretical Physics

Dr. Fredric P Schuller*

Piyush Kumar Singh †

December 2024
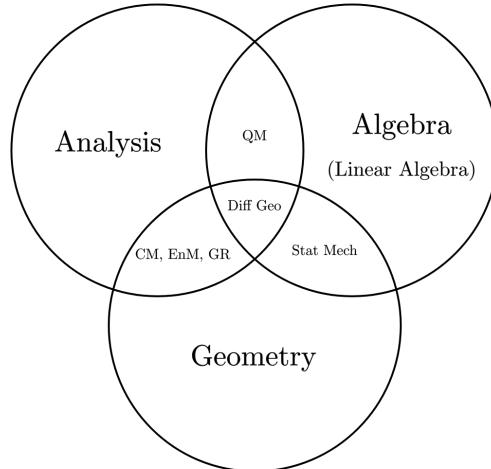
Structure of this Course

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 Axiomatic Set Theory

§1.1 Propositional Logic

> **Definition 1.1 (Proposition):**
> *A proposition p is a variable that can take the values "true" or "false". No other values are allowed.*

It is not the task of 'propositional logic' to determine whether a proposition is true or false. It is only concerned with the logical relationship between propositions.

Note: We can build new propositions from existing ones with the help of logical operators.

§1.1.1 Logical Operators

(a) <u>Unary Operators:</u> These operators operate on a single proposition. There are four unary operators:

p	$\neg p$ Negation	$\text{id}\, p$ Identity	$\top p$ Tautology	$\bot p$ Contradiction
T	F	T	T	F
F	T	F	T	F

Table 1.1: Unary Operators

(b) <u>Binary Operators:</u> These operators operate on two propositions. There are sixteen binary operators. Some important ones are:

p	q	$p \wedge q$ Conjunction (AND)	$p \vee q$ Disjunction (OR)	$p \underline{\vee} q$ Exclusive Or	$p \Rightarrow q$ Implication	$p \Leftrightarrow q$ Equivalence
T	T	T	T	F	T	T
T	F	F	T	T	F	F
F	T	F	T	T	T	F
F	F	F	F	F	T	T

Table 1.2: Binary Operators

> Remark 1.2 (*ex falso quodlibet*). The definition of implication has two not so obvious cases. The first one is when the hypothesis is false, and the conclusion is true. The second one is when both the hypothesis and the conclusion are false. In both cases, the implication is true.
>
> In other words, this says that we can conclude anything from a false assumption.

> **Theorem 1.3**
> $(p \Rightarrow q) \Leftrightarrow ((\neg q) \Rightarrow (\neg p))$

Proof:

We can prove this by truth table.

p	q	$\neg p$	$\neg q$	$p \Rightarrow q$	$(\neg q) \Rightarrow (\neg p)$
T	T	F	F	T	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

Since the columns for $p \Rightarrow q$ and $(\neg q) \Rightarrow (\neg p)$ are identical, we have shown that $(p \Rightarrow q) \Leftrightarrow ((\neg q) \Rightarrow (\neg p))$.

<div align="right">Q.E.D.</div>

Corollary 1.4

We can prove assertions by way of contradiction.

Remark 1.5 (Binding Order). We agree on the decreasing binding strength of the logical operators as follows:

$$\neg, \ \wedge, \ \vee, \ \Rightarrow, \ \Leftrightarrow$$

Remark 1.6 (Higher Order Operators). All higher order operators $(e.g. \ \heartsuit(p_1, p_2, \ldots, p_N))$ can be constructed from one single binary operator $i.e.$ NAND ($\uparrow$).

p	q	$p \uparrow q$
T	T	F
T	F	T
F	T	T
F	F	T

Table 1.3: NAND Operator

§1.2 Predicate Logic

Definition 1.7 (Predicate):
A predicate is a proposition-valued function of some variable(s).

Example 1.8

At this point, we don't know how to construct a predicate. But we can only talk about its value for a given value of the variable. In general, predicates are denoted as follows:

1. $P(x)$: true or false depending on the value of x.

2. $Q(x, y)$: true or false depending on the values of x and y.

We can construct new predicates from existing ones.

(a) Let $P(x)$ and $R(y, z)$ are two given predicates then we can define a third predicate $Q(x, y, z) :\Leftrightarrow P(x) \wedge R(y, z)$[1].

(b) Convert predicate P of one variable into a proposition:

---

[1]The symbol $(:\Leftrightarrow)$ means that the left-hand side is defined to be equivalent to the right-hand side.

> **Definition 1.9 ('Universal' quantifier):**
>
> $$\boxed{\forall x : P(x)} \quad \textit{reads as "for all } x, \ P(x) \textit{ is true."} \tag{1.1}$$
>
> <u>*defined*</u> *to be true, if $P(x)$ is true independently of x.*

Using for all quantifier, we can define a new proposition from a given predicate *i.e.* 'existence' quantifier.

> **Definition 1.10 ('Existential' quantifier):**
>
> $$\boxed{\exists x : P(x)} \quad \textit{reads as "there exists an } x \textit{ such that } P(x) \textit{ is true."} \tag{1.2}$$
>
> <u>*defined*</u> *as $\exists x : P(x) :\Leftrightarrow \neg(\forall x : \neg P(x))$*

> **Corollary 1.11**
>
> $$\forall x : \neg P(x) \Leftrightarrow \neg(\exists x : P(x)) \tag{1.3}$$

(c) Quantification for predicates of more than one variable:

$$Q(y) :\Leftrightarrow \forall x : P(x, y)$$

here, x is a *bound variable* and y is a *free variable*.

> Remark 1.12 (Order of Quantifiers). The order of quantifiers is important. For example,
>
> $$\underbrace{\forall x : \exists y : P(x, y)}_{\text{used for definition of inverse}} \quad \text{generically different proposition than} \quad \underbrace{\exists y : \forall x : P(x, y)}_{\text{used for definition of identity}}$$

§1.3 Axiomatic Systems and Theory of Proofs

> **Definition 1.13 (Axiomatic System):**
> *An axiomatic system is a finite sequence of propositions $a_1, a_2, \ldots, a_N$ called axioms.*

> **Definition 1.14 (Proof):**
> *A proof of a proposition p is within an axiomatic system $a_1, a_2, \ldots, a_N$ is a finite sequence of propositions $q_1, q_2, \ldots, (q_M = p)$ such that for any $1 \leq i \leq M$ in the sequence, either*
>
> *(A) q_i is a proposition from the list of axioms, or*
>
> *(T) q_i is a tautology, or*
>
> *(M) "modus ponens"*
> $$\exists \, 1 \leq m, n < i : (q_m \wedge q_n \Rightarrow q_i) \textit{ is true.}$$

This definition allows to easily recognize a proof by checking the sequence of propositions.

An altogether different matter is to actually find a proof.

> Remark 1.15. If proposition p can be proven from an axiomatic system $a_1, a_2, \ldots, a_N$, we often write this as
>
> $$a_1, a_2, \ldots, a_N \vdash p$$
>
> and say that axiomatic system proves proposition p.

> **Remark 1.16 (Redundant Axioms).** Any tautology, should it occur in the axioms, can be removed from the list of axioms without impairing the power of the axiomatic system.

Extreme case of this is: axiomatic system for propositional login is 'empty sequence.'

> **Definition 1.17 (Consisitency):**
> *An axiomatic system is* consistent *if there exists a proposition q which cannot be proven from the axioms.*
> $$\exists q : \neg(a_1, a_2, \ldots, a_N \vdash q) \tag{1.4}$$

<u>Idea behind consistency:</u> Consider an axiomatic system containing contradicting propositions:

$$a_1, a_2, \ldots, s, \ldots, \neg s, \ldots, a_N$$

Then by *modus ponens* (M), we can prove any proposition p as

$$s \wedge \neg s \Rightarrow p \text{ is a tautology}$$

This can be used as a marker for inconsistency *i.e.* if an axiomatic system can prove every proposition, then it is inconsistent.

> **Theorem 1.18**
> Propositional logic is consistent.

Proof:
Suffices to show that there exists a proposition which cannot be proven within propositional logic.

Propositional logic has an empty sequence of axioms. Only (T) and (M) must carry any proof $\Rightarrow$ only tautologies can be proven; *i.e.* for a proposition p, we can't prove $p \wedge \neg p$. Q.E.D.

> **Theorem 1.19 (Gödel's Incompleteness Theorem)**
> Any axiomatic system that is powerful enough to encode the elementary arithmetic of natural numbers is either inconsistent or contains a proposition that can neither be proven nor disproven.

Proof:
The proof of this theorem is complicated; but the basic idea is as follows:

1) assign a number to each (meta-)mathematical statement, now called *Gödel number*.
2) Use a "The barber shaves all man in his village who do not shave themselves"- type of argument to identify a proposition that is neither provable nor disprovable.

Q.E.D.

────────────────── $\boxed{\text{Lecture } 2}$ ──────────────────

§1.4 The $\in$-Relation

Set theory built on the postulate that there is a fundamental relation[2] called $\in$. There will be no definition of what $\in$ is, or what a set is.

Instead of defining what a set is, we have 9 axioms that speak of $\in$ and sets. Overviews of these axioms are as follows:

---
[2]A predicate of two variables

E	basic existence axiom
E	
P	
U	
R	Construction axiom
P	
I	further existence and construction
C	
F	Non-existence axiom

Table 1.4: Mnemonic for Axioms of Set Theory

Using the $\in$-relation, we can immediately define following relations:

Not an element: $x \notin y :\Leftrightarrow \neg(x \in y)$ $\hspace{5cm}$ (1.5)

Subset: $x \subseteq y :\Leftrightarrow \forall z : (z \in x \Rightarrow z \in y)$ $\hspace{4cm}$ (1.6)

Equality: $x = y :\Leftrightarrow x \subseteq y \wedge y \subseteq x$ $\hspace{4.5cm}$ (1.7)

§1.5 Zermelo-Fraenkel Axioms of Set Theory

Axiom 1.20 (Axiom on $\in$-relation)**:** *$x \in y$ is a proposition if and only if x and y are both sets.*

$$\forall x : \forall y : (x \in y) \veebar \neg(x \in y) \tag{1.8}$$

Counter example (Russell's Paradox): Assume there is some u that contains all sets that do not contain themselves as an element. To be precise,

$$\exists u : \forall z : (z \in u \Leftrightarrow z \notin z) \tag{1.9}$$

Problem: Is u a set?

If u was a set, then one must be able to determine whether $u \in u$ is true or false. As axiom 1.20 states, $u \in u$ is a proposition if and only if u is a set.

Case 1: Assume $u \in u$ is true: $\underset{\substack{\text{def of} \\ u}}{\Longrightarrow} u \notin u$. $\lightning$

Case 2: Assume $u \in u$ is false: $\underset{\substack{\text{def} \\ \notin}}{\Longleftrightarrow} u \notin u \underset{\substack{\text{def} \\ u}}{\Longrightarrow} u \in u$. $\lightning$

Conclusion: u is not a set. $\hspace{9cm}$ $\square$

Axiom 1.21 (Existence of an empty set)**:** *There exists a set that contains no elements.*

$$\exists x : \forall y : y \notin x \tag{1.10}$$

Theorem 1.22 (Uniqueness of the empty set)
There is only one empty set. And it is denoted by $\emptyset$.

Proof (Standard textbook style):

Assume x and x' are both empty sets. But then

$$\forall y : \overbrace{\underbrace{(y \notin x)}_{\text{always false}} \Rightarrow (y \in x')}^{\text{since hypothesis is always false the implication is always true}}$$

But this just means that $x \subseteq x'$ from eq. (1.6). Conversely,

$$\forall y : (y \notin x') \Rightarrow (y \in x)$$

But thus $x' \subseteq x$. Hence, $x = x'$.   Q.E.D.

The above proof is a bit informal. Now we will prove it using the definition of proof in an axiomatic system.

Proof (Formal Version):

Suffices to show that if x and x' are both empty sets, then $x = x'$.

$$\left.\begin{array}{l} a_1 \Leftrightarrow \forall y : y \notin x \\ a_2 \Leftrightarrow \forall y : y \notin x' \end{array}\right\} \text{assumptions or axioms}$$

$$q_1 \underset{\text{(T)}}{\Longleftrightarrow} \forall y : y \notin x \Rightarrow \forall y : (y \in x \Rightarrow y \in x')$$

$$q_2 \underset{\text{(A)1}}{\Longleftrightarrow} \forall y : y \notin x$$

$$q_3 \underset{\text{(M)1,2}}{\Longleftrightarrow} \forall y : (y \in x \Rightarrow y \in x') \boxed{\Leftrightarrow x \subseteq x'}$$

$$q_4 \underset{\text{(T)}}{\Longleftrightarrow} \forall y : y \notin x' \Rightarrow \forall y : (y \in x' \Rightarrow y \in x)$$

$$q_5 \underset{\text{(A)2}}{\Longleftrightarrow} \forall y : y \notin x'$$

$$q_6 \underset{\text{(M)4,5}}{\Longleftrightarrow} \forall y : (y \in x' \Rightarrow y \in x) \boxed{\Leftrightarrow x' \subseteq x}$$

$$q_7 \underset{\text{(M)3,6}}{\Longleftrightarrow} x = x'$$

          Q.E.D.

Axiom 1.23 (Axiom on Pair Sets): *Let x and y be sets. Then there exists a set that contains as its elements precisely the sets x and y.*

$$\forall x : \forall y : \exists m : \forall u : (u \in m \Leftrightarrow u = x \lor u = y) \tag{1.11}$$

Notation: denote this set m by $\{x, y\}$.

> **Theorem 1.24 (Ordering in Pair Sets)**
> Let x and y be sets. Then $\{x, y\} = \{y, x\}$.

Proof:

$$a \in \{x, y\} \Rightarrow a \in \{y, x\} \Rightarrow \{x, y\} \subseteq \{y, x\}$$
$$a \in \{y, x\} \Rightarrow a \in \{x, y\} \Rightarrow \{y, x\} \subseteq \{x, y\}$$

Thus, $\{x, y\} = \{y, x\}$.   Q.E.D.

> **Definition 1.25 (Set with one element):**
> *Let x be a set. Then there exists a set that contains as its element precisely the set x.*
>
> $$\{x\} := \{x, x\} \tag{1.12}$$

Axiom 1.26 (Axiom on Union Sets): *Let x be a set. Then there exists a set whose elements are precisely the elements of the elements of x.*

$$\forall x : \exists u : \forall y : (y \in u \Leftrightarrow \exists z : (z \in x \land y \in z)) \tag{1.13}$$

Notation: denote this set u by $\bigcup x$.

> **Example 1.27**
> Let a, b be sets. Then by axiom 1.23, $\{a\}$ and $\{b\}$ are sets, and hence $x := \{\{a\}, \{b\}\}$ is a set. Then $\bigcup x = \{a, b\}$.

Observe that, since a and b are sets, then by axiom 1.23, $\{a, b\}$ is a set. So it gives us a false impression that axiom 1.26 is redundant. But it is not. Consider the following example:

> **Example 1.28**
> Let a, b, c be sets. Then by axiom 1.23, $\{a, b\}$ and $\{c\}$ are sets, and hence $x := \{\{a, b\}, \{c\}\}$ is a set. Then
> $$\bigcup x =: \{a, b, c\}. \tag{1.14}$$

With this example, we can generalize the definition of a finite set.

> **Definition 1.29 (Finite Set):**
> *Let $a_1, a_2, \ldots, a_N$ be sets. Define, recursively for all $N \geq 3$,*
>
> $$\{a_1, a_2, \ldots, a_N\} := \bigcup \{\{a_1, a_2, \ldots, a_{N-1}\}, \{a_N\}\} \tag{1.15}$$

Axiom 1.30 (Axiom of Replacement): *Let R be a functional relation. Let m be a set. Then the image of m under R $[\mathrm{im}_R(m)]$ is a set.*

For this axiom to make sense, we need to define what a functional relation is.

> **Definition 1.31 (Functional Relation):**
> *A relation R is functional if and only if*
>
> $$\forall x : \exists! \, y : R(x, y)^{3} \tag{1.16}$$

We define the image of a set under a functional relation keeping in mind that axiom 1.30 guarantees this object to be a set.

> **Definition 1.32 (Image of a Set under a Functional Relation):**
> *Let R be a functional relation. Let m be a set. Then the image of m under R consists of all those elements y for which there is an element $x \in m$ such that $R(x, y)$.*

The axiom of replacement is a very powerful axiom. It implies, but is not implied by, the "*principle of restricted comprehension*".

---

[3]The symbol $\exists!$ is 'Uniqueness' quantifier and read as "there exists a unique" and is a shorthand for $\exists y : R(x, y) \land \forall y' : (R(x, y') \Rightarrow y = y')$.

> **Theorem 1.33 (Principle of Restricted Comprehension (PRC))**
> Let P be a predicate of one variable and let m be a set. Then those elements $y \in m$ for which $P(y)$ is true constitute a set.
>
> <u>Notation:</u> this set is denoted by $\{y \in m \mid P(y)\}$.

PRC is not to be confused with the inconsistent "*principle of unrestricted comprehension*" (PUC) which allows for paradoxes like 'Self-Reference Paradox.'

Consider the predicate $P(x)$ is true if x is a set that doesn't contain themselves as an element. Thus, from PUC we have $\{y \mid P(y)\}$ is a set (sick), but this give rise to Russell's Paradox in '*naive set theory*.'

Intuitively PRC makes sure that the image set doesn't grow bigger than a set itself.

Proof:
We can prove PRC using the axiom of replacement in following cases:

Case 1: $\neg \exists y \in m : P(y)$[4] in this case,
$$\{y \in m \mid P(y)\} := \emptyset$$

Case 2: $\exists \hat{y} \in m : P(\hat{y})$, then define a relation R as follows:
$$R(x, y) := (P(x) \wedge x = y) \vee (\neg P(x) \wedge \hat{y} = y)$$

> Claim: R is a functional relation.

Pf: Let x be a set. If $P(x)$ is true, then $R(x, y)$ is true if and only if $y = x$. If $P(x)$ is false, then $R(x, y)$ is true if and only if $y = \hat{y}$.
Thus, $\forall x : \exists! y : R(x, y)$. Hence, R is a functional relation. $\square$

Then by axiom 1.30, $\mathrm{im}_R(m)$ is a set. And by definition of R, $\forall y \in \mathrm{im}_R(m) : P(y)$[5], thus we can define $\{y \in m \mid P(y)\} := \mathrm{im}_R(m)$.

<div align="right">Q.E.D.</div>

> **Definition 1.34 (Set Difference):**
> *Let u and m be sets such that $u \subseteq m$. Then the set difference of m and u is defined as*
> $$m \setminus u := \{x \in m \mid x \notin u\}. \tag{1.17}$$

The object $m \setminus u$ is a set due to PRC *i.e.* ultimately due to the axiom of replacement.

> **Definition 1.35 (Intersection of Sets):**
> *Let x be a set. Then the intersection of x is defined as*
> $$\bigcap x := \left\{ y \in \bigcup x \ \middle| \ \forall z \in x : y \in z \right\} \tag{1.18}$$

Historically, in naive set theory, PUC was thought to be needed in order to define, for any set m,
$$\mathcal{P}(m) \underset{\text{inconsistently}}{:=} \{u \mid u \subseteq m\}$$

This definition is circular, as we need to know a priori, from which bigger set the elements of $\mathcal{P}(m)$ are taken.

---

[4]The quantifier $\exists y \in m : P(y)$ is logically equivalent to $\neg(\forall y \in m : \neg P(y))$.
[5]The quantifier $\forall y \in m : P(y)$ is logically equivalent to $\forall y : (y \in m \Rightarrow P(y))$.

Axiom 1.36 (Axiom on existence of power sets)**:** *Let m be a set. Then there exists a set, denoted by $\mathcal{P}(m)$, whose elements are precisely the subsets of m.*

$$\forall m : \exists p : \forall u : (u \in p \Leftrightarrow u \subseteq m) \tag{1.19}$$

Example 1.37

Let $m = \{a, b\}$, then $\mathcal{P}(m) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.

At this point, we can define what an ordered pair is and what a Cartesian product is.

Definition 1.38 (Ordered Pair):
Let X and Y be non-empty sets. For any $x \in X$ and $y \in Y$, define the ordered pair

$$(x, y) := \{\{x\}, \{x, y\}\} \tag{1.20}$$

From this definition, we can prove the following theorem which justifies the name 'ordered pair.'

Theorem 1.39 (Ordering in Ordered Pairs)
Let X and Y be non-empty sets. Then $(x, y) = (x', y')$ if and only if $x = x'$ and $y = y'$.

Proof:
$[\Rightarrow]$
We will prove this theorem in two cases:
Case 1: Assume $x = y$:

$$(x, y) = \{\{x\}, \{x, y\}\} = \{\{x\}, \{x, x\}\} = \{\{x\}, \{x\}\} = \{\{x\}\}$$
$$\{x', y'\} \in \{\{x'\}, \{x', y'\}\} = (x', y') = (x, y) = \{\{x\}\} \Rightarrow \{x', y'\} = \{x\}$$

Thus by axiom 1.23, $x' = y' = x$.
Case 2: Assume $x \neq y$:
Observe that $x' \neq y'$ if not then $\{x, y\} \in (x, y) = (x', y') = \{\{x'\}\} \Rightarrow x = y = x'$. ⨎
As $\{x'\} \in (x', y') = (x, y) = \{\{x\}, \{x, y\}\}$, we have $\{x'\} = \{x\}$ as $\{x'\} \neq \{x, y\}$ since $x \neq y$. Hence, $x' = x$.
Now, $\{x', y'\} \in (x', y') = (x, y) = \{\{x\}, \{x, y\}\}$, then $\{x', y'\} = \{x\}$ or $\{x', y'\} = \{x, y\}$. Observe $\{x', y'\} \neq \{x\}$ if not then $x' = y' = x$. ⨎
Hence $\{x', y'\} = \{x, y\}$ and thus $y' = y$ as $y' \neq x$ as $x' \neq y'$. Thus, $x' = x$ and $y' = y$.

$[\Leftarrow]$
Trivial.   Q.E.D.

Is collection of all ordered pairs a set? Answer to this question leads us to the definition of Cartesian product.

Theorem 1.40 (Existence of Cartesian Product)
Let X and Y be non-empty sets. Then the collection of all ordered pairs of elements of X and Y is a set *i.e.*
$$X \times Y := \{(x, y) \mid x \in X \wedge y \in Y\} \tag{1.21}$$
is a set.

Proof:
Suffice to show that for any $x \in X$ and $y \in Y$, (x, y) exists in "bigger" set that we know exists. Then by PRC, $X \times Y$ is a set.

Fix $x \in X$ and $y \in Y$. Then by axiom 1.23, $\{x\}$ and $\{y\}$ are sets. Then by axiom 1.23, $\{x, y\}$ is a set. So by eq. (1.6), $\{x\} \subseteq X \Rightarrow \{x\} \subseteq \bigcup\{X, Y\}$ and $\{x, y\} \subseteq \bigcup\{X, Y\}$. So by axiom 1.36 we know $\mathcal{P}(\bigcup\{X, Y\})$ exists. Thus,

$$\{x\} \in \mathcal{P}\left(\bigcup\{X, Y\}\right) \quad \text{and} \quad \{x, y\} \in \mathcal{P}\left(\bigcup\{X, Y\}\right) \tag{1.22}$$

With this we can write $(x, y) = \{\{x\}, \{x, y\}\} \subseteq \mathcal{P}(\bigcup\{X, Y\})$. Hence, $(x, y) \in \mathcal{P}(\mathcal{P}(\bigcup\{X, Y\}))$. Thus, by PRC,

$$X \times Y = \left\{ u \in \mathcal{P}\left(\mathcal{P}\left(\bigcup\{X, Y\}\right)\right) \; \middle| \; \exists x \in X : \exists y \in Y : u = (x, y) \right\} \tag{1.23}$$

is a set.   Q.E.D.

Axiom 1.41 (Axiom of Infinity): *There exists a set that contains the empty set as an element and with every of its elements y it also contains the set $\{y\}$ as an element.*

$$\exists x : \emptyset \in x \wedge \forall y : (y \in x \Rightarrow \{y\} \in x). \tag{1.24}$$

Remark 1.42 (Smallest Infinite Set). Let x be the set guaranteed by axiom 1.41. Then $\emptyset \in x \Rightarrow \{\emptyset\} \in x \Rightarrow \cdots$,

$$x = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\{\emptyset\}\}\} \ldots\} \tag{1.25}$$

Notationally, writing down the set x is cumbersome as there are too many braces. So we denote each element of x by a natural number as follows:

$$0 := \emptyset, \quad 1 := \{\emptyset\}, \quad 2 := \{\{\emptyset\}\}, \quad 3 := \{\{\{\emptyset\}\}\}, \quad \ldots \tag{1.26}$$

Thus, the set x can be written as

$$x = \{0, 1, 2, 3, \ldots\} \tag{1.27}$$

Corollary 1.43 (Existence of Natural Numbers)
$\mathbb{N}$ is a set.

Remark 1.44 (Set of Real Numbers). As a set, $\mathbb{R} = \mathcal{P}(\mathbb{N})$.

Axiom 1.45 (Axiom of Choice): *Let x be a set whose elements are non-empty and mutually disjoint sets, then there exists a set y which contains precisely one element from each element of x.*

$$\forall x : P(x) \Rightarrow \exists y : \forall w \in x : \exists! z \in w : z \in y \tag{1.28}$$

where, $P(x) :\Leftrightarrow (\exists u : u \in x) \wedge (\forall u \in x : \forall v \in x : u \neq v \Rightarrow \bigcap\{u, v\} = \emptyset)$.

The set y is called "*dark*" set. Say we have a set x that contains pair of shoes, then we can algorithmically choose the left shoe from each pair of shoes to form the set y. But if we have a set x that contains a pair of socks, then we can't algorithmically choose one sock from each pair of socks to form the set y. This is where the axiom of choice comes into play.

The axiom of choice is independent of the other 8 axioms, which means that one could have set theory with or without the axiom of choice. However, standard mathematics uses the axiom of choice and hence so will we.

There is a number of theorems that can only be proved by using the axiom of choice. Amongst these we have:

- Proof that every vector space has a basis needs the axiom of choice.

- Proof that there exists a complete system of representatives for the equivalence classes of a set under an equivalence relation needs the axiom of choice.

Axiom 1.46 (Axiom of foundation)**:** *Every non-empty set x contains an element y that has none of its elements in common with x.*

$$\forall x : (\exists y : y \in x) \Rightarrow \exists y \in x : \bigcap \{x, y\} = \emptyset \tag{1.29}$$

Corollary 1.47

There is no set that contains itself as an element.

$$x \in x \quad \text{is false for all sets } x \tag{1.30}$$

The totality of these 9 axioms is called ZFC (Zermelo-Fraenkel with the axiom of choice) set theory.

───────────────── $\boxed{\text{LECTURE } 3}$ ─────────────────

§1.6 Classifications of Sets

A recurrent theme in mathematics is the study/classification of spaces by means of structure-preserving maps between those spaces.

A space is usually meant to be some set equipped with some additional structure. In this context, we are interested in the classification of sets which is a space without any additional structure.

Definition 1.48 (Map):
A map $\phi : A \to B$ is a relation such that for every $a \in A$, there exists exactly one $b \in B$ such that $\phi(a, b)$.

Notation: $\phi : A \to B$ *or* $A \xrightarrow{\phi} B$. *Since there is a unique b for every a, we have a notational abuse as $a \mapsto b =: \phi(a)$,*

Some basic terminologies:

- A is called the *domain* of ϕ.

- B is called the *co-domain* of ϕ.

- The set $\phi(A) \equiv \mathrm{im}_\phi(A) := \{\phi(a) \in B \mid a \in A\}$ is called the *image* of A under ϕ.

Definition 1.49:
Let A and B be sets. A map $\phi : A \to B$ is called

- Surjective *(or* onto*) if and only if* $\mathrm{im}_\phi(A) = B$.

- Injective *(or* one-to-one*) if and only if for all* $a_1, a_2 \in A$, $\phi(a_1) = \phi(a_2) \Rightarrow a_1 = a_2$.

- Bijective *if and only if it is both surjective and injective.*

Definition 1.50 (Iso-morphism):
Let A and B be sets. We say that A and B are (set-theoretic) isomorphic *if there exists a bijection $\phi : A \to B$. In this case, we write $A \cong_{set} B$.*

For two sets to be isomorphic, we only need to prove the existence of a bijection between them. We do not need to construct the bijection explicitly.

Remark 1.51 (Number of isomorphisms). If there is any bijection between two sets, then generically there are many bijections between them.

Intuition: pair the elements of the two sets in any way you like.

In case of set theory, the structure-preserving maps are bijections.

> **Definition 1.52:**
> *Let A be a set. The set A is*
>
> - Infinite *if there exists a proper subset $B \subsetneq A$ such that $B \cong_{set} A$.*
>
>   * *A is called* countable infinite *if and only if $A \cong_{set} \mathbb{N}$.*
>
>   * *A is called* uncountable infinite *if and only if it is infinite but not countably infinite.*
>
> - Finite *if it is not infinite.*
>
>   *In this case, we have $A \cong_{set} \{1, 2, 3, \ldots, N\}$ for some $N \in \mathbb{N}$. Then we write $|A| = N$ and call N the* cardinality *of A.*

<u>Composition of maps:</u> Given two maps $\phi : A \to B$ and $\psi : B \to C$, we can construct a new map know as the composition of ϕ and ψ, denoted by $\psi \circ \phi$, defined as

$$
\begin{aligned}
\psi \circ \phi : A &\to C \\
a &\mapsto \psi(\phi(a))
\end{aligned}
$$

Diagrammatically, we can represent the composition of maps as follows:

$$
\begin{array}{ccc}
 & B & \\
{}^{\phi}\nearrow & & \searrow{}^{\psi} \qquad \text{[Commutes]} \\
A \xrightarrow[\psi \circ \phi]{} & & C
\end{array}
$$

And the composition of maps is associative, i.e., $\xi \circ (\psi \circ \phi) = (\xi \circ \psi) \circ \phi$.

<u>Identity map:</u> For any set A, there exists a unique map $\mathrm{id}_A : A \to A$ such that

$$
\forall a \in A : \quad a \xrightarrow{\ \mathrm{id}_A\ } a
$$

> **Definition 1.53 (Inverse of a map):**
> *Let $\phi : A \to B$ be a bijection. Then the inverse of ϕ, is the map $\phi^{-1} : B \to A$ defined uniquely by*
>
> $$
> \mathrm{id}_A \circlearrowleft A \underset{\phi^{-1}}{\overset{\phi}{\rightleftarrows}} B \circlearrowright \mathrm{id}_B
> \qquad\qquad
> \begin{aligned}
> \phi^{-1} \circ \phi &= \mathrm{id}_A \\
> \phi \circ \phi^{-1} &= \mathrm{id}_B
> \end{aligned}
> \tag{1.31}
> $$

> **Definition 1.54 (Pre-image):**
> *Let $\phi : A \to B$ be a map and let $B' \subseteq B$. Then define the set*
>
> $$
> \mathrm{preim}_\phi(B') := \{a \in A \mid \phi(a) \in B'\}.
> \tag{1.32}
> $$
>
> *$\mathrm{preim}_\phi(B')$ is called the* pre-image *of B' under ϕ.*

§1.7 Equivalence Relations

Definition 1.55 (Equivalence relation):
Let M be a set and let $\sim$ be a relation such that:

(i) Reflexivity: $\forall m \in M : m \sim m$.

(ii) Symmetry: $\forall m, n \in M : m \sim n \Rightarrow n \sim m$.

(iii) Transitivity: $\forall m, n, p \in M : m \sim n \land n \sim p \Rightarrow m \sim p$.

Then $\sim$ is called an equivalence relation *on M.*

Example 1.56

Consider the following wordy examples.

(a) $p \sim q :\Leftrightarrow p$ is of the same opinion as q. This relation is reflexive, symmetric and transitive. Hence, it is an equivalence relation.

(b) $p \sim q :\Leftrightarrow p$ is a sibling of q. This relation is symmetric and transitive but not reflexive and hence, it is not an equivalence relation.

(c) $p \sim q :\Leftrightarrow p$ is taller q. This relation is transitive, but neither reflexive nor symmetric and hence, it is not an equivalence relation.

(d) $p \sim q :\Leftrightarrow p$ is in love with q. This relation is generally not reflexive. People don't like themselves very much. It is certainly not normally symmetric, which is the basis of much drama in literature. It is also not transitive, except in some French films.

Definition 1.57 (Equivalence class):
Let M be a set and let $\sim$ be an equivalence relation on M. For any $m \in M$, the set

$$[m] := \{n \in M \mid n \sim m\} \tag{1.33}$$

is called the equivalence class *of m under $\sim$.*

Two key properties of equivalence classes are:

Proposition 1.58

Let M be a set and let $\sim$ be an equivalence relation on M. Then:

(a) $a \in [m] \Rightarrow [a] = [m]$.

In other words, this means that "any element of an equivalence class can act as a representative of the equivalence class".

(b) Let $m, n \in M$, then either $[m] = [n]$ or $\bigcap \{[m], [n]\} = \emptyset$.

Proof: (a) Let $a \in [m]$. Then by definition, $a \sim m$. Let $b \in [a]$. Then $b \sim a$. By transitivity, $b \sim m$. Hence, $b \in [m]$. Therefore, $[a] \subseteq [m]$. Now, let $b \in [m]$. Then $b \sim m$. By transitivity, $b \sim a$. Hence, $b \in [a]$. Therefore, $[m] \subseteq [a]$. Hence, $[a] = [m]$.

(b) For two elements, either $m \sim n$ or $m \nsim n$. If $m \sim n$, then $n \in [m]$ and hence, $[m] = [n]$. If $m \nsim n$, then $\forall p \in M : (p \in [m]) \veebar (p \in [n])$. Hence, $\bigcap \{[m], [n]\} = \emptyset$.

Q.E.D.

Definition 1.59 (Quotient set):
Let M be a set and let $\sim$ be an equivalence relation on M. Then define the quotient set

$$M\big/_{\sim} := \{[m] \in \mathcal{P}(M) \mid m \in M\} \tag{1.34}$$

Notation: $M\big/_{\sim}$ *is read as "M modulo $\sim$". The elements of $M\big/_{\sim}$ are the equivalence classes of M under $\sim$.*

<u>Intuition</u>: The quotient set is the set of all equivalence classes of M under $\sim$.

> **Remark 1.60** (Set of representatives). Due to the Axiom of Choice, there exists a complete system of representatives for $\sim$ *i.e.* a set R such that $\forall m \in M : \exists! r \in R : m \sim r$. Then we can write
>
> $$R \cong_{\text{set}} M\big/_{\sim}. \tag{1.35}$$

This essentially means that we can choose a representative from each equivalence class and put them in a set namely R. Then this provides us a natural bijection between the set of representatives R and the quotient set $M\big/_{\sim}$.

> **Remark 1.61.** Care must be taken while defining maps whose domain or co-domain is a quotient set and if one uses representatives to define the map. The map must be well-defined, *i.e.* the map must not depend on the choice of representatives.

Example 1.62

Let $M = \mathbb{Z}$ and let $\sim$ be a relation defined by $m \sim n :\Leftrightarrow m - n \in 2\mathbb{Z}$. Then $\sim$ is an equivalence relation. The equivalence classes are

$$[0] = [2] = [4] = \ldots = [-2] = [-4] = \ldots = 2\mathbb{Z},$$
$$[1] = [3] = [5] = \ldots = [-1] = [-3] = \ldots = 2\mathbb{Z} + 1.$$

Therefore, the quotient set is
$$\mathbb{Z}\big/_{\sim} = \{[0], [1]\}.$$

<u>Idea</u>: on $\mathbb{Z}$ we have addition $+ : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$. We wish to inherit an addition on $\mathbb{Z}\big/_{\sim}$ from the addition on $\mathbb{Z}$ *i.e.* $\oplus : \mathbb{Z}\big/_{\sim} \times \mathbb{Z}\big/_{\sim} \to \mathbb{Z}\big/_{\sim}$. We can define this map as follows:

$$[a] \oplus [b] \mapsto [a + b]$$

Care needs to be taken precisely because this could be inconsistent (not well-defined). Check whether choice of representatives matters:
Let $a', b' \in \mathbb{Z}$ such that $[a] = [a']$ and $[b] = [b']$. Then we need to check whether $[a'] \oplus [b'] = [a] \oplus [b]$. So by assumptions we have $a \sim a' \Rightarrow a - a' = 2n \in 2\mathbb{Z}$ and $b \sim b' \Rightarrow b - b' = 2m \in 2\mathbb{Z}$. Then

$$[a'] \oplus [b'] \overset{\text{def}}{\underset{\oplus}{=}} [a' + b'] = [a - 2n + b - 2m]$$
$$= [a + b - 2(n + m)]$$
$$= [a + b] \underset{\overset{\text{def}}{\oplus}}{=} [a] \oplus [b].$$

Therefore, the addition is well-defined.

---

§1.8 Construction of $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$

We are only going through the outline of the construction of these sets.

§1.8.1 Natural Numbers $\mathbb{N}$

Recall, from the Axiom of Infinity, that there exists a set $\mathbb{N}$ such that

$$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$$
$$0 := \emptyset, \qquad 1 := \{\emptyset\}, \qquad 2 := \{\{\emptyset\}\}, \qquad 3 := \{\{\{\emptyset\}\}\}, \qquad \ldots \tag{1.36}$$

We wish to establish addition on $\mathbb{N}$. For that we need to define the *successor map* as

$$
\begin{aligned}
S : \mathbb{N} &\to \mathbb{N} \\
n &\mapsto \{n\}
\end{aligned}
\tag{1.37}
$$

This works as, $S(2) = S(\{\{\emptyset\}\}) = \{\{\{\emptyset\}\}\} = 3$.
We need to define the *predecessor map* as

$$
\begin{aligned}
P : \mathbb{N}^* &\to \mathbb{N} \\
n &\mapsto m \quad \text{such that} \quad m \in n
\end{aligned}
\tag{1.38}
$$

Here, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$. In this case, $P(2) = P(\{\{\emptyset\}\}) = \{\emptyset\} = 1$ as there is only one element in 2.
Now, with this we can define $(n \in \mathbb{N})^{\text{th}}$ power of S:

$$
\begin{aligned}
S^n &= S \circ S^{P(n)}, \quad \text{if } n \in \mathbb{N}^* \\
S^0 &= \text{id}_{\mathbb{N}}
\end{aligned}
\tag{1.39}
$$

At this point, we can define addition on $\mathbb{N}$ as

> **Definition 1.63 (Addition on $\mathbb{N}$):**
>
> $$
> \begin{aligned}
> + : \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\
> (m, n) &\mapsto m + n := S^n(m)
> \end{aligned}
> \tag{1.40}
> $$

> **Example 1.64 ($2 + 1 = 3$)**
> With this definition, we can verify our intuition of addition on $\mathbb{N}$. For example, $2 + 1 = 3$.
>
> $$
> 2 + 1 = S^1(2) = S(S^0(2)) = S(\text{id}_{\mathbb{N}}(2)) = S(2) = 3
> $$
>
> And with this, we should be able to check whether the addition is commutative or not.
>
> $$
> 1 + 2 = S^2(1) = S(S^1(1)) = S(S(S^0(1))) = S(S(1)) = S(2) = 3
> $$

This same calculation generalizes to any two natural numbers and hence, the addition is commutative.

> **Remark 1.65 (Identity element of addition).** We can see that 0 is the identity element of addition on $\mathbb{N}$. This is because
>
> $$
> \begin{aligned}
> \forall n \in \mathbb{N} &: 0 + n = S^n(0) = \cdots = n, \\
> \forall n \in \mathbb{N} &: n + 0 = S^0(n) = n.
> \end{aligned}
> $$

What about the inverse of addition? We can see that there is no inverse of addition on $\mathbb{N}$. This is because, for any $n \in \mathbb{N}$, there is no $m \in \mathbb{N}$ such that $n + m = 0$. In language of group theory, $\mathbb{N}$ is not a commutative group under addition.
Now this hurdle motivates us to define the set of integers $\mathbb{Z}$.

§1.8.2 Integers $\mathbb{Z}$

Let's start with defining an equivalence relation $\sim_{\mathbb{Z}}$ on $\mathbb{N} \times \mathbb{N}$ as

$$
(m, n) \sim_{\mathbb{Z}} (p, q) :\Leftrightarrow m + q = n + p
\tag{1.41}
$$

Check that this is an equivalence relation.

1. *Reflexivity:* $(m, n) \sim_{\mathbb{Z}} (m, n) \Leftrightarrow m + n = n + m$.

2. *Symmetry:* $(m, n) \sim_{\mathbb{Z}} (p, q) \Leftrightarrow m + q = n + p \Leftrightarrow n + p = m + q \Leftrightarrow (p, q) \sim_{\mathbb{Z}} (m, n)$.

3. *Transitivity:* $(m, n) \sim_{\mathbb{Z}} (p, q) \wedge (p, q) \sim_{\mathbb{Z}} (r, s) \Rightarrow m + q = n + p \wedge p + s = q + r \Rightarrow m + q + p + s = n + p + q + r \Rightarrow m + s = n + r \Rightarrow (m, n) \sim_{\mathbb{Z}} (r, s)$.

Thus, $\sim_{\mathbb{Z}}$ is an equivalence relation on $\mathbb{N} \times \mathbb{N}$.

Idea: in this case (m, n) corresponds to $m - n$. So, we can define the set of integers as

$$\mathbb{Z} := {}^{(\mathbb{N} \times \mathbb{N})} \big/ {}_{\sim_{\mathbb{Z}}} \tag{1.42}$$

Intuitively, we want $\mathbb{N} \subseteq \mathbb{Z}$ but at this stage this is nonsense as both the sets have different structures. We can resolve this by embedding $\mathbb{N}$ into $\mathbb{Z}$ with the help of an *inclusion map*:

$$\begin{aligned} \iota : \mathbb{N} &\hookrightarrow \mathbb{Z} \\ n &\mapsto [(n, 0)] \end{aligned} \tag{1.43}$$

With this new definition of $\mathbb{N}$ in $\mathbb{Z}$, we can say that $\mathbb{N} \subseteq \mathbb{Z}$. Now, in similar fashion we can define negative integers as well.

> **Definition 1.66 (Negative integers):**
> *Let $n \in \mathbb{N}$, then we define the negative integer as*
>
> $$-n := [(0, n)] \tag{1.44}$$

With this, let's define addition on $\mathbb{Z}$.

> **Definition 1.67 (Addition on $\mathbb{Z}$):**
>
> $$\begin{aligned} +_{\mathbb{Z}} : \mathbb{Z} \times \mathbb{Z} &\to \mathbb{Z} \\ [(m, n)] +_{\mathbb{Z}} [(p, q)] &:= [(m + p, n + q)] \end{aligned} \tag{1.45}$$

It is easy to see that this addition is well-defined.

Example 1.68 (2 + (-3) = -1)
Using our definition of $2 = [(2, 0)]$ and $-3 = [(0, 3)]$, we can calculate $2 +_{\mathbb{Z}} (-3)$ as:

$$2 +_{\mathbb{Z}} (-3) = [(2, 0)] +_{\mathbb{Z}} [(0, 3)] = [(2 + 0, 0 + 3)] = [(2, 3)] = [(0, 1)] = -1$$

We will now assume that we have constructed the multiplication on $\mathbb{Z}$[6], and then we can define the set of rational numbers $\mathbb{Q}$.

§1.8.3 Rational Numbers $\mathbb{Q}$

With similar construction as above, we define an equivalence relation $\sim_{\mathbb{Q}}$ on $\mathbb{Z} \times \mathbb{Z}^*$[7] as

$$(m, n) \sim_{\mathbb{Q}} (p, q) :\Leftrightarrow mq = np \tag{1.46}$$

Check that this is an equivalence relation.

1. *Reflexivity:* $(m, n) \sim_{\mathbb{Q}} (m, n) \Leftrightarrow mn = nm$.

2. *Symmetry:* $(m, n) \sim_{\mathbb{Q}} (p, q) \Leftrightarrow mq = np \Leftrightarrow np = mq \Leftrightarrow (p, q) \sim_{\mathbb{Q}} (m, n)$.

3. *Transitivity:* $(m, n) \sim_{\mathbb{Q}} (p, q) \wedge (p, q) \sim_{\mathbb{Q}} (r, s) \Rightarrow mq = np \wedge ps = qr \Rightarrow mqps = npqr \Rightarrow ms = nr \Rightarrow (m, n) \sim_{\mathbb{Q}} (r, s)$.

---

[6]Multiplication $\cdot_{\mathbb{Z}} : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$, defined as $[(m, n)] \cdot_{\mathbb{Z}} [(p, q)] = [(mp + nq, mq + np)]$.
[7]$\mathbb{Z}^* := \mathbb{Z} \setminus \{0\}$

Thus, $\sim_{\mathbb{Q}}$ is an equivalence relation on $\mathbb{Z} \times \mathbb{Z}^*$.

<u>Idea:</u> in this case (m, n) corresponds to $\frac{m}{n}$. So, we can define the set of rational numbers as

$$\mathbb{Q} := \left.(\mathbb{Z} \times \mathbb{Z}^*)\middle/\sim_{\mathbb{Q}}\right. \tag{1.47}$$

Intuitively, we want $\mathbb{Z} \subseteq \mathbb{Q}$ but at this stage this is nonsense as both the sets have different structures. We can resolve this by embedding $\mathbb{Z}$ into $\mathbb{Q}$ with the help of an *inclusion map*:

$$\begin{aligned} \iota : \mathbb{Z} &\hookrightarrow \mathbb{Q} \\ n &\mapsto [(n, 1)] \end{aligned} \tag{1.48}$$

With this new definition of $\mathbb{Z}$ in $\mathbb{Q}$, we can say that $\mathbb{Z} \subseteq \mathbb{Q}$.

> **Definition 1.69 (Addition on $\mathbb{Q}$):**
>
> $$\begin{aligned} +_{\mathbb{Q}} : \mathbb{Q} &\times \mathbb{Q} \to \mathbb{Q} \\ [(m,n)] +_{\mathbb{Q}} [(p,q)] &:= [(mq + np, nq)] \end{aligned} \tag{1.49}$$

Similarly, we can define the multiplication on $\mathbb{Q}$

> **Definition 1.70 (Multiplication on $\mathbb{Q}$):**
>
> $$\begin{aligned} \cdot_{\mathbb{Q}} : \mathbb{Q} &\times \mathbb{Q} \to \mathbb{Q} \\ [(m,n)] \cdot_{\mathbb{Q}} [(p,q)] &:= [(mp, nq)] \end{aligned} \tag{1.50}$$

§1.8.4 Real Numbers $\mathbb{R}$

There are many ways to construct the reals from the rationals. One is to define a set $\mathscr{A}$ of *almost homomorphism* on $\mathbb{Z}$ and hence define:

$$\mathbb{R} := \left.\mathscr{A}\middle/\sim\right., \tag{1.51}$$

where $\sim$ is a "suitable" equivalence relation on $\mathscr{A}$.

We will not go into the details of this construction. We will just assume that the reals are constructed, and we have all the usual operations defined on them. Now define the set of positive reals as

$$\mathbb{R}^+ := \{r \in \mathbb{R} \mid r > 0\} \tag{1.52}$$

With this construction, we can define $\mathbb{R}^d$ for any $d \in \mathbb{N}$ as

$$\mathbb{R}^d := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{d \text{ times}} \tag{1.53}$$

For $x \in \mathbb{R}^d$, we write $x = (x^1, x^2, \ldots, x^d)$ where $x^i \in \mathbb{R}$. For further convenience, we need to define the *norm* on $\mathbb{R}^d$.

> **Definition 1.71 (norm):**
> Let $x = (x^1, x^2, \ldots, x^d) \in \mathbb{R}^d$. *Then the norm of x is defined as*
>
> $$\begin{aligned} \|\cdot\| : \mathbb{R}^d &\to \mathbb{R}^+ \\ x &\mapsto \sqrt{\sum_{i=1}^{d} (x^i)^2} \end{aligned} \tag{1.54}$$

We can extend this definition of norm as

$$x \mapsto \sqrt[2n]{\sum_{i=1}^{d} (x^i)^{2n}} \quad \text{for any } n \in \mathbb{N} \tag{1.55}$$

In order to avoid confusion, we will denote the norm of x as $\|x\|_2$ for $n = 1$ and in general as $\|x\|_{2n}$ for $n \in \mathbb{N}$. For convenience, we will use $n = 1$ most of the time.

With this, one can define a ball in $\mathbb{R}^d$ as

> **Definition 1.72 (Ball in $\mathbb{R}^d$):**
> *Let $x \in \mathbb{R}^d$ and $r \in \mathbb{R}^+$. Then define the set $B_r(x)$ as*
>
> $$B_r(x) := \left\{ y \in \mathbb{R}^d \;\middle|\; \sqrt{\sum_{i=1}^{d} (x^i - y^i)^2} < r \right\} \tag{1.56}$$
>
> *is called the* ball *of radius r centered at x.*

Later on, we will see that we denote these sets as "open balls".

2 Topological Spaces

We will now discuss topological spaces based on our previous development of set theory. As we will see, a topology on a set provides the weakest structure in order to define the two very important notions of convergence of sequences to points in a set, and of continuity of maps between two sets.

§2.1 Topological Spaces

> **Definition 2.1 (Topology & Topological Space):**
> *Let M be a set. Then a choice of subsets $\mathcal{O} \subseteq \mathcal{P}(M)$ is called a* topology *on M if*
>
> *(i) $\emptyset \in \mathcal{O}$ and $M \in \mathcal{O}$,*
>
> *(ii) $U, V \in \mathcal{O} \Rightarrow \bigcap \{U, V\} \in \mathcal{O}$,*
>
> *(iii) $C \subseteq \mathcal{O} \Rightarrow \bigcup C \in \mathcal{O}$.*
>
> *The pair $(M, \mathcal{O})$ is then called a* topological space.

> **Remark 2.2 (Open & Closed Sets).** Let $(M, \mathcal{O})$ be a topological space. Then
>
> (a) A set $U \in \mathcal{O}$ is called an *open set* in $(M, \mathcal{O})$.
>
> (b) A set $C \subseteq M$ is called a *closed set* in $(M, \mathcal{O})$ if $M \setminus C \in \mathcal{O}$ i.e. $M \setminus C$ is open in $(M, \mathcal{O})$.

> **Remark 2.3 (choice of topology).** Unless $|M| = 1$, there are different topologies $\mathcal{O}$ one can choose on one and the same set M. Based on the choice of topology $\mathcal{O}$, the notion of convergence of sequences and continuity of maps will change.

For finite set, we can count the number of topologies on it.

$|M|$	Number of Topologies
1	1
2	4
3	29
4	355
5	6,942
6	209,527
7	9,535,241
⋮	⋮

> **Example 2.4 (Basic Topologies)**
> Let M be any set, then we can define following "trivial" topologies on M:
>
> (a) $\mathcal{O} = \{\emptyset, M\}$. It is easy to see that, $\mathcal{O}$ is a topology on M. This is called the *chaotic topology*.
>
> (b) $\mathcal{O} = \mathcal{P}(M)$. It is easy to see that, $\mathcal{O}$ is a topology on M. This is called the *discrete topology*.

For a less abstract example, consider $M = \{1, 2, 3\}$, then we can define 29 different topologies on M. For example, $\mathcal{O} = \{\emptyset, \{1\}, \{2\}, \{1, 2\}, \{1, 2, 3\}\}$ is a topology on M.

Example 2.5 (Standard Topology on $\mathbb{R}^d$)

This is the most important and heavily used topology in mathematics and physics.
Consider $M = \mathbb{R}^d$, where $d \in \mathbb{N}$. Then the *standard topology* on $\mathcal{O}_{\text{std.}}$ is constructed as follows:

$$U \in \mathcal{O}_{\text{std.}} :\Leftrightarrow \forall p \in U : \exists r \in \mathbb{R}^+ : B_r(p) \subseteq U$$

where $B_r(p) = \left\{ x \in \mathbb{R}^d \ \middle| \ \|x - p\| < r \right\}$ is the open ball of radius r around p. In more involved terms, we can write the definition of the standard topology on $\mathbb{R}^d$ as

$$\mathcal{O}_{\text{std.}} = \left\{ U \in \mathcal{P}(\mathbb{R}^d) \ \middle| \ \forall p \in U : \exists r \in \mathbb{R}^+ : B_r(p) \subseteq U \right\}$$

Then by theorem 1.33, $\mathcal{O}_{\text{std.}}$ is a set. Now we have to prove that $\mathcal{O}_{\text{std.}}$ is indeed a topology on $\mathbb{R}^d$.

(i) First, we need to check whether $\emptyset \in \mathcal{O}_{\text{std.}}$, *i.e.* whether:

$$\forall p \in \emptyset : \exists r \in \mathbb{R}^+ : B_r(p) \subseteq \emptyset$$

is true. This proposition is of the form $\forall p \in \emptyset : Q(p)$, which was defined as being equivalent to:

$$\forall p : p \in \emptyset \Rightarrow Q(p).$$

However, since $p \in \emptyset$ is false, the implication is true independent of p. Hence, the initial proposition is true and thus $\emptyset \in \mathcal{O}_{\text{std.}}$.

Second, by definition, we have $B_r(x) \subseteq \mathbb{R}^d$ independent of x and r, hence:

$$\forall p \in \mathbb{R}^d : \exists r \in \mathbb{R}^+ : B_r(p) \subseteq \mathbb{R}^d$$

is true and thus $\mathbb{R}^d \in \mathcal{O}_{\text{std.}}$.

(ii) Let $U, V \in \mathcal{O}_{\text{std.}}$, then we have to show that $U \cap V \in \mathcal{O}_{\text{std.}}$. Let $p \in U \cap V$, then $p \in U$ and $p \in V$. Since $U, V \in \mathcal{O}_{\text{std.}}$, we have:

$$\exists r_1 \in \mathbb{R}^+ : B_{r_1}(p) \subseteq U,$$
$$\exists r_2 \in \mathbb{R}^+ : B_{r_2}(p) \subseteq V.$$

Now, let $r = \min \{r_1, r_2\}$, then we have $B_r(p) \subseteq B_{r_1}(p) \subseteq U$ and $B_r(p) \subseteq B_{r_2}(p) \subseteq V$. Hence, $B_r(p) \subseteq U \cap V$ and thus $U \cap V \in \mathcal{O}_{\text{std.}}$.

(iii) Let $C \subseteq \mathcal{O}_{\text{std.}}$, then we have to show that $\bigcup C \in \mathcal{O}_{\text{std.}}$. Let $p \in \bigcup C$, then $p \in U$ for some $U \in C$. Since $U \in \mathcal{O}_{\text{std.}}$, we have:

$$\exists r \in \mathbb{R}^+ : B_r(p) \subseteq U \subseteq \bigcup C.$$

Hence, $\bigcup C \in \mathcal{O}_{\text{std.}}$.

Thus, $\mathcal{O}_{\text{std.}}$ is a topology on $\mathbb{R}^d$.

Definition 2.6 (Comparison of Topologies):
Let M be a set and $\mathcal{O}_1, \mathcal{O}_2$ be two topologies on M. If $\mathcal{O}_1 \subseteq \mathcal{O}_2$, then we say that $\mathcal{O}_2$ is finer *(or* stronger*) topology than $\mathcal{O}_1$, and $\mathcal{O}_1$ is* coarser *(or* weaker*) topology than $\mathcal{O}_2$.*

From this definition, we can see that the chaotic topology is the coarsest topology on a set, and the discrete topology is the finest topology on a set.

§2.2 Construction of new Topologies from given topologies

§2.2.1 Induced Topology

> **Theorem 2.7 (Induced Topology)**
> Let $(M, \mathcal{O})$ be a topological space and $N \subset M$[1]. Then the following set
>
> $$\mathcal{O}\Big|_N = \{U \cap N \mid U \in \mathcal{O}\} \subseteq \mathcal{P}(N) \tag{2.1}$$
>
> is a topology on N, called the *induced topology* on N.

Proof:

We have to show that $\mathcal{O}|_N$ is a topology on N.

  (i) First, as $\emptyset \in \mathcal{O}$, we have $\emptyset \cap N = \emptyset \in \mathcal{O}|_N$.

      Second, since $M \in \mathcal{O}$, we have $M \cap N = N \in \mathcal{O}|_N$.

  (ii) $U, V \in \mathcal{O}|_N \overset{?}{\Rightarrow} U \cap V \in \mathcal{O}|_N$. From (2.1), $\exists U_1, V_1 \in \mathcal{O} : U = U_1 \cap N \wedge V = V_1 \cap N$. Then we have:

$$U \cap V = (U_1 \cap N) \cap (V_1 \cap N) = (U_1 \cap V_1) \cap N \in \mathcal{O}\Big|_N$$

     as $U_1 \cap V_1 \in \mathcal{O}$.

  (iii) $C \subseteq \mathcal{O}|_N \overset{?}{\Rightarrow} \bigcup C \in \mathcal{O}|_N$. Since $C \subseteq \mathcal{O}|_N$, we have $C = \{U_i \cap N \mid U_i \in \mathcal{O}\}$. Then we have:

$$\bigcup C = \bigcup \{U_i \cap N \mid U_i \in \mathcal{O}\} = \Big\{\bigcup U_i\Big\} \cap N \in \mathcal{O}\Big|_N .$$

Thus, $\mathcal{O}|_N$ is a topology on N.   Q.E.D.

> **Example 2.8**
> Consider $(\mathbb{R}, \mathcal{O}_{\text{std.}})$ and $N = [0,1] := \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$. Then from theorem 2.7, $(N, \mathcal{O}_{\text{std.}}|_N)$ is a topological space.
>
> It is easy to see that $(0,1] \notin \mathcal{O}_{\text{std.}}$ so $(0,1]$ is not open in $(\mathbb{R}, \mathcal{O}_{\text{std.}})$, but $(0,1] = (0,2) \cap N \in \mathcal{O}_{\text{std.}}|_N$. Hence, $(0,1]$ is an open set in $(N, \mathcal{O}_{\text{std.}}|_N)$.

At this point we can talk about openness and closedness of sets in a topological space.

> Remark 2.9. Let $(M, \mathcal{O})$ be a topological space and $N \subset M$. Then N is either
>
> - open, or
> - closed, or
> - open and closed, or
> - open and not closed, or
> - not open and closed, or
> - not open and not closed.

For this very unintuitive idea of a set to be open and closed at the same time, see the following example.

> **Example 2.10**
> For any topological space $(M, \mathcal{O})$, we know by definition of topology that $\emptyset, M \in \mathcal{O}$. Hence, $\emptyset$ and M are open in $(M, \mathcal{O})$.
>
> - Now, let $N = M$, then $M \setminus N = \emptyset \in \mathcal{O}$. Hence, M is closed in $(M, \mathcal{O})$.

---

[1] We will use either $N \subset M$ or $N \subsetneq M$ or $N \subsetneqq M$ to denote that N is a proper subset of M.

- Let $N = \emptyset$, then $M \setminus N = M \in \mathcal{O}$. Hence, $\emptyset$ is closed in $(M, \mathcal{O})$.

Thus, $\emptyset$ and M are both open and closed in $(M, \mathcal{O})$.

§2.2.2 Quotient Topology

Definition 2.11 (Quotient Topology):
Let $(M, \mathcal{O})$ be a topological space and $\sim$ be an equivalence relation on M. Then the quotient set $M/_\sim$ can be equipped with the quotient topology $\mathcal{O}_{M/_\sim}$ *defined as follows:*

$$\mathcal{O}_{M/_\sim} = \left\{ U \subseteq M/_\sim \ \middle| \ \bigcup U = \bigcup_{[u] \in U} [u] \in \mathcal{O} \right\} \tag{2.2}$$
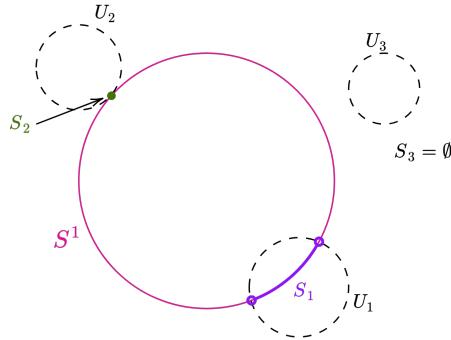
An equivalent definition of the quotient topology is as follows: Let $\pi : M \to M/_\sim$ be the map defined as $\pi(m) = [m]$. Then the quotient topology is defined as

$$\mathcal{O}_{M/_\sim} = \left\{ U \subseteq M/_\sim \ \middle| \ \mathrm{preim}_\pi(U) \in \mathcal{O} \right\} \tag{2.3}$$

Example 2.12 (Topology on 1-sphere)
Consider $M = \mathbb{R}^2$. The *circle* (or *1-sphere*) is defined as a set $S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$. We have two ways to construct a new topological space:

(a) <u>Induced Topology:</u> As $S^1 \subset \mathbb{R}^2$, we can construct a topology $\mathcal{O}$ on S^1 by theorem 2.7 using $(\mathbb{R}^2, \mathcal{O}_{\text{std.}})$ topological space. This is the *standard topology* on S^1.


Figure 2.1: 1-sphere induced topology

From fig. 2.1, we can see that U_i's are open in $\mathbb{R}^2$ and $U_i \cap S^1$ are open in S^1. See, $U_3 \in \mathcal{O}_{\text{std.}} \Rightarrow \emptyset \in \mathcal{O}$ and $U_1 \in \mathcal{O}_{\text{std.}} \Rightarrow S_1 \in \mathcal{O}$. Also, the set S_2 is not open in S^1 as $\forall U \in \mathcal{P}(\mathbb{R}) : U \cap S^1 \neq S_2$. Hence, $S_2 \notin \mathcal{O}$.

(b) <u>Quotient Topology:</u> Let $\sim$ be the equivalence relation on $\mathbb{R}$ defined by:

$$x \sim y :\Leftrightarrow \exists n \in \mathbb{Z} : x = y + 2\pi n.$$

Then the circle can be defined as the set $S^1 := \mathbb{R}/_\sim$ equipped with the quotient topology.

§2.2.3 Product Topology

Definition 2.13 (Product Topology):
Let $(A, \mathcal{O}_A)$ and $(B, \mathcal{O}_B)$ be two topological spaces. Then the set $\mathcal{O}_{A \times B}$ defined implicitly as
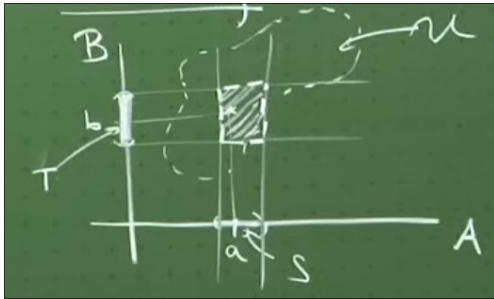
$$U \in \mathcal{O}_{A \times B} :\Leftrightarrow \forall (a, b) \in U : \exists U_A \in \mathcal{O}_A, U_B \in \mathcal{O}_B : a \in U_A \land b \in U_B \land U_A \times U_B \subseteq U \tag{2.4}$$

is a topology on $A \times B$, called the product topology *on $A \times B$.*

Here we used the existence of U_A and U_B such that it contains a and b respectively. And we will use this kind of argument throughout our discussion.

Let $(M, \mathcal{O})$ be a topological space. Let $a \in M$ be a point. Then we say that a set $U \in \mathcal{O}$ is an *open neighbourhood* of a if $a \in U$.

Intuitively, the definition can be understood with the help of following diagram:

> **Remark 2.14 (Extending product topology to finite cartesian product).** Let $A_1, A_2, \ldots, A_N$ be topological spaces. Then in same spirit, we can define the product topology
>
> $$\mathcal{O}_{A_1 \times A_2 \times \cdots \times A_N}$$

With this, we can check that the standard topology on $\mathbb{R}^d$ is a product topology with open ball defined using $\lVert \cdot \rVert_\infty$

$$\mathcal{O}_{\text{std. } \mathbb{R}^d} = \mathcal{O}_{\underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{d \text{ times}}}$$

Till this point, we have seen topology and topological spaces from geometric view-point. But one can use these things beyond the geometric (or manifold) applications. For example, a guy named Furstenberg introduced a simple topology on the set of integers to prove the infinitude of primes.

§2.3 Convergence

> **Definition 2.15 (Sequence):**
> *A sequence (of points) in a set M is a map from $\mathbb{N}$ onto M, i.e. $q : \mathbb{N} \to M$.*

> **Definition 2.16 (Convergence and Limit Point):**
> *Let $(M, \mathcal{O})$ be a topological space. A sequence q in M is said to* converge *against a* limit point $a \in M$ *if*
> $$\forall U \in \mathcal{O} : a \in U \Rightarrow \exists N \in \mathbb{N} : \forall n > N : q(n) \in U \tag{2.5}$$

> **Example 2.17 (Justifying different topologies)**
>
> With the notion of convergence, we can justify the different topologies on a set.
>
> (a) <u>Chaotic Topology:</u> Consider the chaotic topology $\mathcal{O} = \{\emptyset, M\}$ on M. Then any sequence in M converges against any point in M. Let $a \in M$ and q be a sequence in M, then we have only two choice for $U \in \mathcal{O}$, i.e. $U = \emptyset$ or $U = M$. In case $U = \emptyset$, the definition of convergence is vacuously true. In case $U = M$, the definition of convergence is true as $q(n) \in M$ for all $n \in \mathbb{N}$. Hence, (arbitrary) sequence q converges against (arbitrary) point a in $(M, \mathcal{O})$.
>
> This weird behavior is the reason for the name "chaotic topology".
>
> (b) <u>Discrete Topology:</u> Consider the discrete topology $\mathcal{O} = \mathcal{P}(M)$ on M. Then only almost constant sequences converge against a (unique) point in M. Let $a \in M$ and q be a sequence in M, then

we have $U = \{a\} \in \mathcal{O}$. Then the definition of convergence is true if and only if $q(n) = a$ for all $n > N$. Hence, only almost constant sequences converge against a point in $(M, \mathcal{O})$.

Here we can see that every sequence is convergent in chaotic topology (the coarsest topology) and only a specific kind of sequence is convergent in discrete topology (the finest topology). So we deduce the following remark.

Remark 2.18. Let M be a set and $\mathcal{O}_1, \mathcal{O}_2$ be two topologies on M such that $\mathcal{O}_1 \subseteq \mathcal{O}_2$ and let q be a sequence in M. Then if q converges in $(M, \mathcal{O}_2)$, then q converges in $(M, \mathcal{O}_1)$ but not necessarily vice-versa.

Theorem 2.19 (Convergence in standard topology on $\mathbb{R}^d$)

Let q be a sequence in $\mathbb{R}^d$ and $a \in \mathbb{R}^d$. Then q converges against a in $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ if and only if

$$\forall \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n > N : \|q(n) - a\| < \varepsilon \tag{2.6}$$

Proof:

Let q be a sequence in $\mathbb{R}^d$ and $a \in \mathbb{R}^d$. Then q converges against a in $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ if and only if

$$\forall U \in \mathcal{O}_{\text{std.}} : a \in U \Rightarrow \exists N \in \mathbb{N} : \forall n > N : q(n) \in U \tag{2.7}$$

From the definition of standard topology on $\mathbb{R}^d$, we have $U \in \mathcal{O}_{\text{std.}} \Rightarrow \exists r \in \mathbb{R}^+ : B_r(a) \subseteq U$. Hence, (2.7) is equivalent to

$$\forall r > 0 : \exists N \in \mathbb{N} : \forall n > N : q(n) \in B_r(a) \tag{2.8}$$

which is equivalent to (2.6). <div style="text-align:right">Q.E.D.</div>

Example 2.20 (Convergence in $\mathbb{R}$)

Let $q(n) = 1 - \frac{1}{n+1}$ be a sequence in $\mathbb{R}$. Then q converges against 1 in $(\mathbb{R}, \mathcal{O}_{\text{std.}})$ as

$$\forall \varepsilon \in \mathbb{R}^+ : \exists N \in \mathbb{N} : \forall n > N : |q(n) - 1| = \left| 1 - \frac{1}{n+1} - 1 \right| = \frac{1}{n+1} < \varepsilon$$

But observe q is not almost constant sequence and hence does not converge in $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$.

---

§2.4 Continuity

Definition 2.21 (Continuity):

Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two topological spaces and $\phi : M \to N$ be a map. Then ϕ is said to be continuous if

$$\forall V \in \mathcal{O}_N : \operatorname{preim}_\phi(V) \in \mathcal{O}_M \tag{2.9}$$

In other words, ϕ is continuous iff the pre-image of every open set in N is open in M.

Example 2.22 ('Trivial' Continuity)

Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two topological spaces and $\phi : M \to N$ be a map. Then

(a) Let $\mathcal{O}_M = \mathcal{P}(M)$ and $\mathcal{O}_N$ be any topology on N. Then every map is continuous.

   Let $V \in \mathcal{O}_N$ be any open set in N. Then $\operatorname{preim}_\phi(V) \subseteq M$ and hence $\operatorname{preim}_\phi(V) \in \mathcal{O}_M$. Hence, ϕ is continuous.

(b) Let $\mathcal{O}_M$ be any topology on M and $\mathcal{O}_N = \{\emptyset, N\}$. Then every map is continuous.

   Here, either $V = \emptyset$ or $V = N$. In case $V = \emptyset$, we have $\operatorname{preim}_\phi(\emptyset) = \emptyset \in \mathcal{O}_M$. In case $V = N$, we have $\operatorname{preim}_\phi(N) = M \in \mathcal{O}_M$. Hence, ϕ is continuous.

Recovering the definition of $\varepsilon - \delta$ definition of continuity on 'Euclidean' spaces from the definition of continuity in topological spaces.

Theorem 2.23 (Continuity in Euclidean spaces)
Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a map. Then f is continuous if and only if

$$\forall a \in \mathbb{R}^m : \forall \varepsilon > 0 : \exists \delta > 0 : \|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \varepsilon \tag{2.10}$$

Proof:
Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a map. Then f is continuous if and only if

$$\forall U \in \mathcal{O}_{\mathbb{R}^n} : \operatorname{preim}_f(U) \in \mathcal{O}_{\mathbb{R}^m} \tag{2.11}$$

From the definition of standard topology on $\mathbb{R}^n$, we have $U \in \mathcal{O}_{\mathbb{R}^n} \Rightarrow \exists \varepsilon > 0 : B_\varepsilon(f(a)) \subseteq U$. Hence, (2.11) is equivalent to

$$\forall \varepsilon > 0 : \exists \delta > 0 : \operatorname{preim}_f(B_\varepsilon(f(a))) \in \mathcal{O}_{\mathbb{R}^m} \tag{2.12}$$

which is equivalent to (2.10).   Q.E.D.

§2.4.1 Homeomorphism
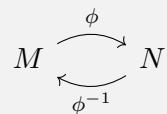
Definition 2.24 (Homeomorphism):
Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two topological spaces. Let $\phi : M \to N$ be a bijection. Then ϕ is said to be a homeomorphism *if*

(a) ϕ is continuous, and

(b) ϕ^{-1} is continuous.

Remark 2.25 (Structure preserving maps). Homeomorphisms are structure preserving maps in topology.

Remark 2.26 (One-to-one pairing of open sets). Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two topological spaces and let $\phi : M \to N$ be a homeomorphism.

$$M \underset{\phi^{-1}}{\overset{\phi}{\rightleftarrows}} N$$

Then ϕ provides a *one-to-one* pairing of the open sets of M with those of N.

Definition 2.27 (Homeomorphic):
Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two topological spaces. Then M and N are said to be homeomorphic *if there exists a homeomorphism between them. And we write*

$$M \cong_{top.} N \tag{2.13}$$

Remark 2.28 (Homeomorphic & Isomorphic). Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two homeomorphic topological spaces. Then they are isomorphic as sets. But the converse is not true.

$$M \cong_{\text{top.}} N \Rightarrow M \cong_{\text{set}} N \tag{2.14}$$

$$\boxed{\text{L{\small ECTURE} 5}}$$

In case of set theory, we had a notion of 'cardinality' of a set. With the help of this notion, we can say that if two sets are isomorphic (set theoretically) then they have the same cardinality. But in case of topological spaces, we don't have any such quantity *i.e.* we don't have any property which can tell us that two topological spaces are 'homeomorphic' or not.

Building on this idea, we had classified sets into different types based on their cardinality like finite, countable, uncountable, etc. But in case of topology, classification of topological spaces is an open problem.

§2.5 Topological Properties I: Separation Axioms

> **Definition 2.29 (T1 Space):**
> *A topological space $(M, \mathcal{O})$ is called* T1 *if for any two distinct points $p, q \in M$ (i.e. $p \neq q$), then*
> $$\exists U_p \in \mathcal{O} : q \notin U_p \tag{2.15}$$
> *here U_p is an open neighbourhood of p.*
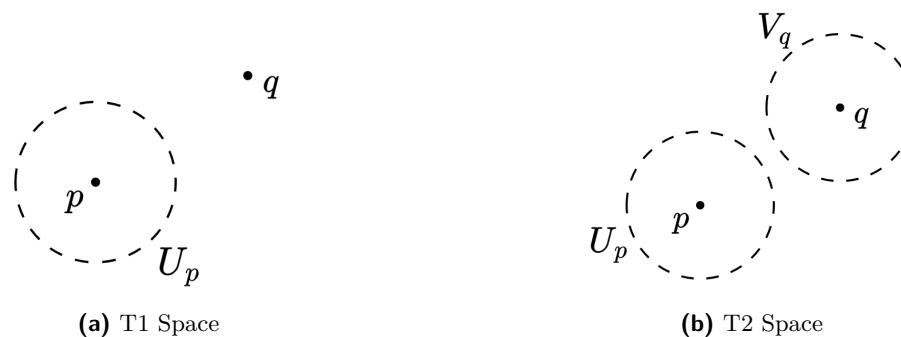
> **Definition 2.30 (T2 Space):**
> *A topological space $(M, \mathcal{O})$ is called* T2 *or* Hausdorff *if for any two distinct points $p, q \in M$ (i.e. $p \neq q$), then*
> $$\exists U_p, V_q \in \mathcal{O} : U_p \cap V_q = \emptyset \tag{2.16}$$
> *here U_p and V_q are open neighbourhoods of p and q respectively.*

Intuitively, we can see why definition 2.29 and definition 2.30 are called 'separation axioms'. Consider standard topology on $\mathbb{R}^2$, see fig. 2.2.

**(a)** T1 Space   **(b)** T2 Space

Figure 2.2: Separation Axioms

> Remark 2.31 (T1 vs T2). Every T2 space is a T1 space, but the converse is not true.

> **Example 2.32**
> Some examples of topological spaces based on separation axioms are:
>
> (a) $\mathbb{R}^d$ with standard topology is a T2 space.
>
> (b) Zariski topology (algebraic geometry) is T1 but not T2.
>
> (c) Chaotic topology is neither T1 nor T2.
>
> (d) Discrete topology is always T1 and T2.

There are many other "T" properties, like T3, T4, etc. All of them have use cases in some area of mathematics or physics.

Remark 2.33 (T2$\frac{1}{2}$ Space). A topological space $(M, \mathcal{O})$ is called $T2\frac{1}{2}$ if for any two distinct points $p, q \in M$ (*i.e.* $p \neq q$), then there exists a closed neighbourhood U_p of p such that $q \notin U_p$.

> **Theorem 2.34 (Uniqueness of Limit Point)**
> Let $(M, \mathcal{O})$ be a Hausdorff space. Then every convergent sequence in M has a unique limit point.

§2.6 Compactness and Paracompactness

Before we define compactness, we need to understand the concept of 'open cover'.

> **Definition 2.35 (Cover and Open Cover):**
> *Let $(M, \mathcal{O})$ be a topological space. A set $C \subseteq \mathcal{P}(M)$ is called a* cover *of M if*
>
> $$\bigcup C = M. \tag{2.17}$$
>
> *If $C \subseteq \mathcal{O}$ then it is called an* open cover*.*

> **Definition 2.36 (Subcover and Finite Subcover):**
> *Let C be a cover of M. A subset $\tilde{C} \subseteq C$ is called a* subcover *of C if*
>
> $$\bigcup \tilde{C} = M. \tag{2.18}$$
>
> *If $\tilde{C}$ is finite then it is called a* finite subcover*.*

§2.6.1 Compactness

> **Definition 2.37 (Compact Space):**
> *A topological space $(M, \mathcal{O})$ is called* compact *if every open cover of M has a finite subcover.*

It is clear from the definition that compactness is a topological property. So we have,

Remark 2.38. Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two homeomorphic topological spaces. Then M is compact if and only if N is compact.

> **Definition 2.39 (Compact Subset):**
> *Let $(M, \mathcal{O})$ be a topological space. A subset $N \subseteq M$ is called* compact *if the induced topology $(N, \mathcal{O}|_N)$ is compact.*

Determining whether a topological space is compact or not, is not an easy task. With the help of a suitable homeomorphism we can simplify this task.

> **Theorem 2.40 (Heine-Borel Theorem)**
> Let $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ be a topological space. Then a subset $K \subseteq \mathbb{R}^d$ is compact if and only if it is closed and bounded.

Remark 2.41 (Bounded Set in $\mathbb{R}^d$). A set $K \subseteq \mathbb{R}^d$ is called *bounded* if

$$\exists r > 0 : K \subseteq B_r(0) \tag{2.19}$$

We can extend Heine-Borel Theorem to any metric space. But first we need to define the concept of 'metric space'.

> **Definition 2.42 (Metric Space):**
> *Let M be a set. A function $d : M \times M \to \mathbb{R}$ is called a* metric *on M if it satisfies the following properties for all $p, q, r \in M$:*
>
> - Symmetry*: $d(p, q) = d(q, p)$.*
>
> - Positivity*: $d(p, q) \geq 0$ and $d(p, q) = 0 \Leftrightarrow p = q$.*
>
> - Triangle Inequality*: $d(p, q) + d(q, r) \geq d(p, r)$.*
>
> *A set M equipped with a metric d is called a* metric space*.*

We can define 'metric-induced topology' on a metric space as we did for $\mathbb{R}^d$.

Remark 2.43 (Metric-Induced Topology). Let (M, d) be a metric space. Then the set $B_r(p)$ is an open ball of radius $r \in \mathbb{R}^+$ centered at $p \in M$ defined as

$$B_r(p) = \{q \in M : d(p, q) < r\}. \tag{2.20}$$

With this, we can define the metric-induced topology $\mathcal{O}_d$ on M as

$$U \in \mathcal{O}_d :\Leftrightarrow \forall p \in U, \exists r \in \mathbb{R}^+ : B_r(p) \subseteq U. \tag{2.21}$$

Theorem 2.44 (Generalized Heine-Borel Theorem)
Let (M, d) be a metric space with metric-induced topology $\mathcal{O}_d$. Then a subset $K \subseteq M$ is compact if and only if it is *complete* and *totally bounded*.

Example 2.45
Example of compact subsets in $\mathbb{R}$ are:

(a) $[0, 1]$ is compact in $(\mathbb{R}, \mathcal{O}_{\text{std.}})$.

(b) $\mathbb{R}$ is not compact in $(\mathbb{R}, \mathcal{O}_{\text{std.}})$. It suffices to show that there exists an open cover of $\mathbb{R}$ which does not have a finite subcover. Consider the open cover

$$C := \{(n, n + 1) \mid n \in \mathbb{Z}\} \cup \{(n + 1/2 \,, n + 3/2) \mid n \in \mathbb{Z}\}. \tag{2.22}$$

We can prove that the product of two compact spaces is compact for product topology.

Theorem 2.46
Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be two compact topological spaces. Then the product space $(M \times N, \mathcal{O}_{M \times N})$ is compact.

Now extending this to finite product spaces is easy.

Compactness is a very strong property, and sometimes it doesn't hold. So we need a weaker property which can be used in such cases. This is where the concept of 'paracompactness' comes in.

§2.6.2 **Paracompactness**

Definition 2.47 (Refinement of a Cover):
Let C be a cover of M. A cover R of M is called a refinement *of C if*

$$\forall \tilde{U} \in R : \exists U \in C : \tilde{U} \subseteq U. \tag{2.23}$$

A refinement R is said to be:

- open *if $R \subseteq \mathcal{O}$.*

- Locally finite *if for every $p \in M$, there exists an open neighbourhood U_p of p such that*

$$\left| \left\{ \tilde{U} \in R \;\middle|\; \tilde{U} \cap U_p \neq \emptyset \right\} \right| < \infty. \tag{2.24}$$

Remark 2.48. Let C be a cover of M and let $\tilde{C}$ be a subcover of C. Then $\tilde{C}$ is a refinement of C. But the converse is not true.

Definition 2.49 (Paracompact Space):
A topological space $(M, \mathcal{O})$ is called paracompact *if every open cover of M has a locally finite open refinement.*

Corollary 2.50
Every compact space is paracompact.

Definition 2.51 (Metrisable Space):
A topological space $(M, \mathcal{O})$ is called metrizable *if there exists a metric d on M such that the metric-induced topology $\mathcal{O}_d$ is same as $\mathcal{O}$.*

Theorem 2.52 (Stone's Theorem)
Every metrizable space is paracompact.

Example 2.53
Let's see some topological spaces which are paracompact or not:

(a) The topological space $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ is metrizable since $\mathcal{O}_{\text{std.}} = \mathcal{O}_d$ where $d = \|\cdot\|_2$. Hence, $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ is paracompact.

(b) Alexandroff Long Line is not paracompact. To construct it, we first observe that we could "build" $\mathbb{R}$ by taking the interval $[0,1)$ and stacking countably many copies of it one after the other. Hence, in a sense, $\mathbb{R}$ is equivalent to $\mathbb{Z} \times [0,1)$. The long line L is defined analogously as $L : \omega_1 \times [0,1)$, where ω_1 is an uncountable infinite set. The resulting space L is not paracompact.

Theorem 2.54
Let $(M, \mathcal{O})$ be a paracompact space and $(N, \mathcal{O}_N)$ be a compact space. Then the product space $(M \times N, \mathcal{O}_{M \times N})$ is paracompact.

Corollary 2.55
Let $(M, \mathcal{O})$ be a paracompact space and $(N_i, \mathcal{O}_{N_i})$ be compact spaces for all $i \in \{1, 2, \ldots, n\}$. Then the product space $M \times N_1 \times N_2 \times \ldots \times N_n$ is paracompact.

At this point, we can discuss the sufficient and necessary criteria for paracompactness which are used as definitions in some books. For that we need to define the concept of '*partition of unity.*'

Definition 2.56 (Partition of Unity):
Let $(M, \mathcal{O})$ be a topological space. A set $\mathcal{F}$ of continuous functions $\phi : M \to [0,1]$ is called a partition
of unity if for each $p \in M$,

   (a) There exists an open neighbourhood U_p such that $\phi(p) \neq 0$ for finitely many $\phi \in \mathcal{F}$.

   (b) $\sum\limits_{\phi \in \mathcal{F}} \phi(p) = 1$.

Definition 2.57 (Partition of Unity Subordinate to a Cover):
Let $(M, \mathcal{O})$ be a topological space and let C be a cover of M. A partition of unity $\mathcal{F}$ is said to be
subordinate to C if:
$$\forall \phi \in \mathcal{F} : \exists U \in C : f(p) \neq 0 \Rightarrow p \in U. \tag{2.25}$$

Theorem 2.58
Let $(M, \mathcal{O})$ be a Hausdorff space. Then M is paracompact if and only if every open cover C admits
a partition of unity subordinate to C.

Example 2.59
Let $(\mathbb{R}, \mathcal{O}_{\text{std.}})$ be a topological space. By theorem 2.52, $(\mathbb{R}, \mathcal{O}_{\text{std.}})$ is paracompact. Consider the
open cover $C := \{U_1 = (-\infty, 1), U_2 = (0, \infty)\}$. We can construct a partition of unity $\mathcal{F} = \{\phi_1, \phi_2\}$
subordinate to C as follows:

$$\phi_1(x) = \begin{cases} 1 & \text{if } x \leq 0 \\ 1 - x^2 & \text{if } 0 \leq x \leq 1 \\ 0 & \text{if } x \geq 1 \end{cases} \quad \text{and} \quad \phi_2(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x^2 & \text{if } 0 \leq x \leq 1 \\ 1 & \text{if } x \geq 1 \end{cases}$$

For ϕ_1, we have $U_1 = (-\infty, 1)$ for which $\phi_1(x) \neq 0$ and for ϕ_2, we have $U_2 = (0, \infty)$ for which
$\phi_2(x) \neq 0$. Also, $\phi_1(x) + \phi_2(x) = 1$ for all $x \in \mathbb{R}$.

§2.7 Connectedness and Path-Connectedness

§2.7.1 Connectedness

Definition 2.60 (Connected Space):
A topological space $(M, \mathcal{O})$ is called connected *unless there exists two non-empty open sets A, B such*
that
$$A \cup B = M \quad \text{and} \quad A \cap B = \emptyset. \tag{2.26}$$

Example 2.61
Let $\left(M := \mathbb{R} \setminus \{0\}, \mathcal{O} := \mathcal{O}_{\text{std.}}|_{\mathbb{R} \setminus \{0\}} \right)$ be a topological space. Then M is not connected since we can
write $M = A \cup B$ where $A = (-\infty, 0)$ and $B = (0, \infty)$ with $A \cap B = \emptyset$.

Theorem 2.62
The closed interval $[0, 1]$ with induced topology from $(\mathbb{R}, \mathcal{O}_{\text{std.}})$ is connected.

Theorem 2.63
Let $(M, \mathcal{O})$ be a topological space. Then M is connected if and only if $\emptyset$ and M are the only subsets
of M that are both open and closed.

Proof:

[$\Rightarrow$]:

Suppose not there exists a set $U \subseteq M$ that is both open and closed and $U \neq \emptyset, M$. Then $M = U \,\dot{\cup}\, (M \backslash U)$. Since U is open and $M \backslash U$ is open (as U is closed). Hence, M is not connected. $\lightning$

[$\Leftarrow$]:

Suppose not M is not connected. Then there exists two non-empty open sets A, B such that $A \cup B = M$ and $A \cap B = \emptyset$. Clearly, $A \neq M$, if not then $B = \emptyset$. Since A is open, $M \backslash A$ is closed. But $M \backslash A = B$ is also open. Hence, $B \neq \emptyset, M$ is open and closed. $\lightning$ <div style="text-align:right">Q.E.D.</div>

§2.7.2 Path-Connectedness

> **Definition 2.64 (Path-Connected Space):**
> *A topological space $(M, \mathcal{O})$ is called* path-connected *if for every pair of points $p, q \in M$, there exists a continuous curve $\gamma : [0, 1] \to M$ such that $\gamma(0) = p$ and $\gamma(1) = q$.*
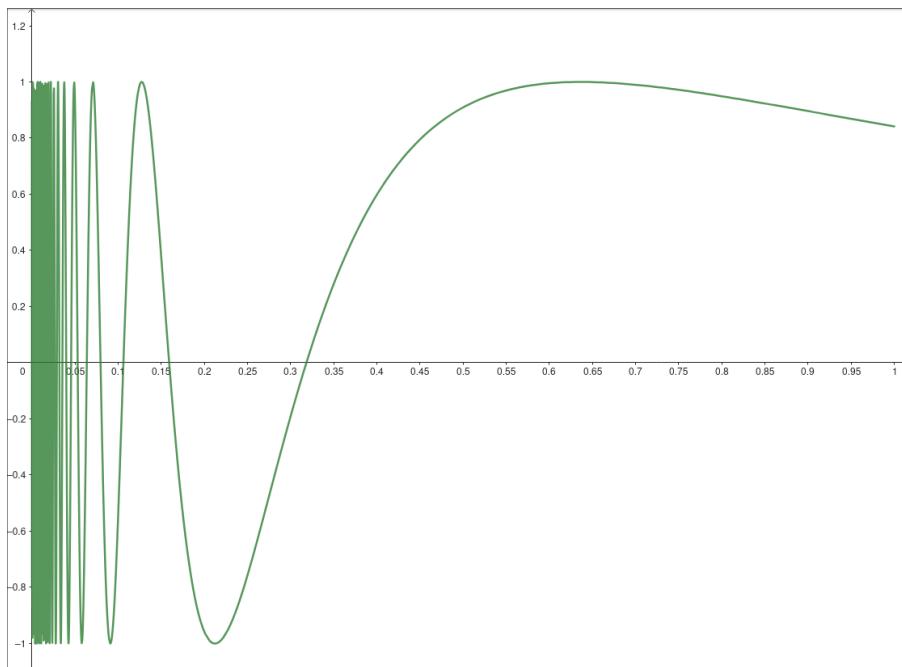
> **Example 2.65**
> Let $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ be a topological space. Then $(\mathbb{R}^d, \mathcal{O}_{\text{std.}})$ is path-connected.

Proof:

Let $p, q \in \mathbb{R}^d$. Then the curve $\gamma(\lambda) := (1 - \lambda)p + \lambda q$ is continuous and $\gamma(0) = p$ and $\gamma(1) = q$. <div style="text-align:right">Q.E.D.</div>

> **Example 2.66**
> Let $S := \left\{ \left(x, \sin\left(\frac{1}{x}\right) \right) \,\middle|\, x \in (0, 1] \right\} \cup \{(0, 0)\} \subseteq \mathbb{R}^2$. Then $(S, \mathcal{O}_{\text{std.}}|_S)$ is not path-connected but connected.
>
>

> **Theorem 2.67**
> Every path-connected space is connected.

Proof:

Let $(M, \mathcal{O})$ be a path-connected space. Suppose not M is not connected. Then there exists two non-empty open sets A, B such that $A \,\dot{\cup}\, B = M$. Since, A and B are non-empty, we can choose $a \in A$ and $b \in B$. Since M is path-connected, there exists a continuous curve $\gamma : [0, 1] \to M$ such that $\gamma(0) = a$

and $\gamma(1) = b$. Moreover, γ is continuous and A, B are open, disjoint sets, $\mathrm{preim}_\gamma(A)$ and $\mathrm{preim}_\gamma(B)$ are open, disjoint sets. We know

$$[0,1] = \mathrm{preim}_\gamma(M) = \mathrm{preim}_\gamma(A \,\dot\cup\, B) = \mathrm{preim}_\gamma(A) \,\dot\cup\, \mathrm{preim}_\gamma(B).$$

Thus, existence of $\mathrm{preim}_\gamma(A)$ and $\mathrm{preim}_\gamma(B)$ implies that $[0,1]$ is not connected. $\lightning$ Q.E.D.

§2.8 Homotopic Curves and Fundamental Group

> **Definition 2.68 (Homotopic Curves):**
> Let $(M, \mathcal{O})$ be a topological space. Two continuous curves $\gamma_1, \gamma_2 : [0,1] \to M$ such that
>
> $$\gamma_1(0) = \gamma_2(0) \quad and \quad \gamma_1(1) = \gamma_2(1)$$
>
> are called homotopic if there exists a continuous function $h : [0,1] \times [0,1] \to M$ such that for all $\lambda \in [0,1]$,
> $$h(0, \lambda) = \gamma_1(\lambda) \quad and \quad h(1, \lambda) = \gamma_2(\lambda). \tag{2.27}$$

Intuitively, two curves are homotopic if one can be continuously deformed into the other. Pictorially, see fig. 2.3.
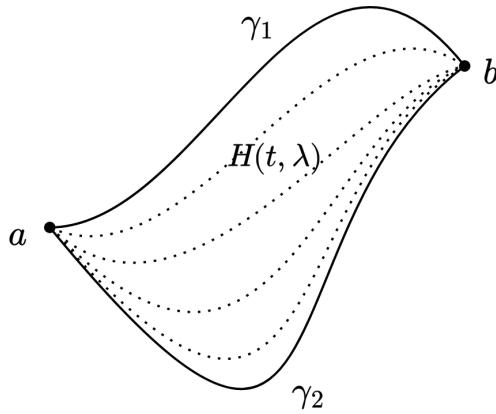
Figure 2.3: Homotopic Curves

> **Proposition 2.69 (Homotopy is an Equivalence Relation)**
> Let $(M, \mathcal{O})$ be a topological space. Define a relation $\sim$ on the set of continuous curves $\{\gamma : [0,1] \to M\}$ by
>
> $$\gamma_1 \sim \gamma_2 :\Leftrightarrow \gamma_1 \text{ and } \gamma_2 \text{ are homotopic.} \tag{2.28}$$
>
> Then $\sim$ is an equivalence relation.

Proof:
Let $\gamma_1, \gamma_2, \gamma_3 : [0,1] \to M$ be continuous curves.

(a) <u>Reflexivity:</u> Define a continuous function $h : [0,1] \times [0,1] \to M$ as $h(t, \lambda) := \gamma_1(\lambda)$. Then $h(0, \lambda) = \gamma_1(\lambda)$ and $h(1, \lambda) = \gamma_1(\lambda)$ for all $\lambda \in [0,1]$. Hence, $\gamma_1 \sim \gamma_1$.

(b) <u>Symmetry:</u> Let $\gamma_1 \sim \gamma_2$. Then there exists a continuous function $h : [0,1] \times [0,1] \to M$ such that $h(0, \lambda) = \gamma_1(\lambda)$ and $h(1, \lambda) = \gamma_2(\lambda)$ for all $\lambda \in [0,1]$. Define a continuous function $h' : [0,1] \times [0,1] \to M$ as $h'(t, \lambda) := h(1 - t, \lambda)$. Then $h'(0, \lambda) = \gamma_2(\lambda)$ and $h'(1, \lambda) = \gamma_1(\lambda)$ for all $\lambda \in [0,1]$. Hence, $\gamma_2 \sim \gamma_1$.

(c) <u>Transitivity:</u> Let $\gamma_1 \sim \gamma_2$ and $\gamma_2 \sim \gamma_3$. Then there exists continuous functions $h_1, h_2 : [0,1] \times [0,1] \to M$ such that $h_1(0, \lambda) = \gamma_1(\lambda)$, $h_1(1, \lambda) = \gamma_2(\lambda)$ and $h_2(0, \lambda) = \gamma_2(\lambda)$, $h_2(1, \lambda) = \gamma_3(\lambda)$ for all $\lambda \in [0,1]$. Define a continuous function $h : [0,1] \times [0,1] \to M$ as

$$h(t, \lambda) := \begin{cases} h_1(2t, \lambda) & \text{if } 0 \leq t \leq \frac{1}{2} \\ h_2(2t - 1, \lambda) & \text{if } \frac{1}{2} \leq t \leq 1 \end{cases} \tag{2.29}$$

Then $h(0, \lambda) = \gamma_1(\lambda)$ and $h(1, \lambda) = \gamma_3(\lambda)$ for all $\lambda \in [0,1]$. Hence, $\gamma_1 \sim \gamma_3$.

Q.E.D.

Definition 2.70 (Space of Loops):
Let $(M, \mathcal{O})$ be a topological space. Then, for every point $p \in M$, the set $\mathscr{L}_p$ defined as

$$\mathscr{L}_p := \{\gamma : [0,1] \to M \mid \gamma \text{ is continuous and } \gamma(0) = \gamma(1) = p\} \tag{2.30}$$

is called the space of loops *at p.*

Definition 2.71 (Concatenation of Loops):
Let $(M, \mathcal{O})$ be a topological space. Fix a point $p \in M$. We define the concatenation *operation $*_p : \mathscr{L}_p \times \mathscr{L}_p \to \mathscr{L}_p$ as*

$$(\gamma_1 *_p \gamma_2)(\lambda) := \begin{cases} \gamma_1(2\lambda) & \text{if } 0 \leq \lambda < \frac{1}{2} \\ \gamma_2(2\lambda - 1) & \text{if } \frac{1}{2} \leq \lambda \leq 1 \end{cases} \tag{2.31}$$

§2.8.1 Fundamental Group

We will first recall the definition of a group.

Definition 2.72 (Group):
A set G equipped with a binary operation $* : G \times G \to G$ is called a group *if it satisfies the following properties:*

(a) <u>Associativity</u>: $\forall a, b, c \in G : (a * b) * c = a * (b * c)$.

(b) <u>Identity Element</u>: $\exists e \in G : \forall a \in G : a * e = e * a = a$.

(c) <u>Inverse Element</u>: $\forall a \in G : \exists a^{-1} \in G : a * a^{-1} = a^{-1} * a = e$.

A group is called abelian *(or* commutative*) if $\forall a, b \in G : a * b = b * a$.*

For future discussion, we need the notion of isomorphism between groups.

Definition 2.73 (Group Isomorphism):
Let $(G, *)$ and $(H, \star)$ be two groups. A function $\varphi : G \to H$ is called a group isomorphism *if it is bijective and satisfies*

$$\forall a, b \in G : \varphi(a * b) = \varphi(a) \star \varphi(b). \tag{2.32}$$

<u>Notation:</u> If there exists an isomorphism between groups G and H, we say that G and H are isomorphic and write $G \cong_{grp} H$.

Now we can define the fundamental group.

> **Definition 2.74 (Fundamental Group):**
> Let $(M, \mathcal{O})$ be a topological space and let $p \in M$. Define the set $\pi_1(p)$ as
>
> $$\pi_1(p) := \mathscr{L}_p/{\sim} = \{[\gamma] \mid \gamma \in \mathscr{L}_p\} \tag{2.33}$$
>
> where $\sim$ is the equivalence relation defined using homotopy. Define a binary operation $\bullet_p : \pi_1(p) \times \pi_1(p) \to \pi_1(p)$ as
>
> $$[\gamma_1] \bullet_p [\gamma_2] := [\gamma_1 *_p \gamma_2]. \tag{2.34}$$
>
> Then $\pi_1(p)$ equipped with the operation $\bullet_p$ is called the fundamental group of M at p.

To prove that $(\pi_1(p), \bullet_p)$ is a group, we need to show following properties:

1. $\bullet_p$ is well-defined.
2. $\bullet_p$ is associative.
3. There exists an identity element.
4. There exists an inverse element for every element.

Proof:

Using the steps mentioned above, we can prove that $(\pi_1(p), \bullet_p)$ is a group.

(a) Well-Defined:

Let $\gamma_1, \gamma_1', \gamma_2, \gamma_2' \in \mathscr{L}_p$ such that $\gamma_1 \sim \gamma_1'$ and $\gamma_2 \sim \gamma_2'$. Then there exists continuous functions $h_1, h_2 : [0,1] \times [0,1] \to M$ such that $h_1(0, \lambda) = \gamma_1(\lambda)$, $h_1(1, \lambda) = \gamma_1'(\lambda)$ and $h_2(0, \lambda) = \gamma_2(\lambda)$, $h_2(1, \lambda) = \gamma_2'(\lambda)$ for all $\lambda \in [0,1]$. Define a continuous function $h : [0,1] \times [0,1] \to M$ as

$$h(t, \lambda) := \begin{cases} h_1(2t, \lambda) & \text{if } 0 \leq t \leq \frac{1}{2} \\ h_2(2t-1, \lambda) & \text{if } \frac{1}{2} \leq t \leq 1 \end{cases} \tag{2.35}$$

Then $h(0, \lambda) = \gamma_1(\lambda) *_p \gamma_2(\lambda)$ and $h(1, \lambda) = \gamma_1'(\lambda) *_p \gamma_2'(\lambda)$. Hence, $\gamma_1 *_p \gamma_2 \sim \gamma_1' *_p \gamma_2'$.

(b) Associativity:

Let $[\gamma_1], [\gamma_2], [\gamma_3] \in \pi_1(p)$. We need to show that $([\gamma_1] \bullet_p [\gamma_2]) \bullet_p [\gamma_3] = [\gamma_1] \bullet_p ([\gamma_2] \bullet_p [\gamma_3])$ which is equivalent to showing $(\gamma_1 *_p \gamma_2) *_p \gamma_3$ and $\gamma_1 *_p (\gamma_2 *_p \gamma_3)$ are homotopic. From definition of concatenation,

$$(\gamma_1 *_p \gamma_2) *_p \gamma_3(\lambda) = \begin{cases} \gamma_1(4\lambda) & \text{if } 0 \leq \lambda \leq \frac{1}{4} \\ \gamma_2(4\lambda - 2) & \text{if } \frac{1}{4} \leq \lambda \leq \frac{1}{2} \\ \gamma_3(2\lambda - 1) & \text{if } \frac{1}{2} \leq \lambda \leq 1 \end{cases} ; \quad \gamma_1 *_p (\gamma_2 *_p \gamma_3)(\lambda) = \begin{cases} \gamma_1(2\lambda) & \text{if } 0 \leq \lambda \leq \frac{1}{2} \\ \gamma_2(4\lambda - 2) & \text{if } \frac{1}{2} \leq \lambda \leq \frac{3}{4} \\ \gamma_3(4\lambda - 3) & \text{if } \frac{3}{4} \leq \lambda \leq 1 \end{cases}$$

Now define $h : [0,1] \times [0,1] \to [0,1]$ as

$$h(t, \lambda) := \begin{cases} \gamma_1((2\lambda)t + (4\lambda)(1-t)) & \text{if } 0 \leq \lambda \leq \frac{t}{2} + \frac{1-t}{4} \\ \gamma_2(4\lambda - 2) & \text{if } \frac{t}{2} + \frac{1-t}{4} \leq \lambda \leq \frac{3t}{4} + \frac{1-t}{2} \\ \gamma_3((4\lambda - 3)t + (2\lambda - 1)(1-t)) & \text{if } \frac{3t}{4} + \frac{1-t}{2} \leq \lambda \leq 1 \end{cases}$$

Then $h(0, \lambda) = (\gamma_1 *_p \gamma_2) *_p \gamma_3(\lambda)$ and $h(1, \lambda) = \gamma_1 *_p (\gamma_2 *_p \gamma_3)(\lambda)$. Hence, $(\gamma_1 *_p \gamma_2) *_p \gamma_3 \sim \gamma_1 *_p (\gamma_2 *_p \gamma_3)$.

(c) Identity Element:

Observe that the constant curve $\gamma_{e,p} : [0,1] \to M$ defined as $\gamma_{e,p}(\lambda) = p$ for all $\lambda \in [0,1]$ is the identity element.

(d) Inverse Element:

Let $[\gamma] \in \pi_1(p)$. Then the curve $\gamma^{-1} : [0,1] \to M$ defined as $\gamma^{-1}(\lambda) = \gamma(1 - \lambda)$ is the inverse element.

This completes the proof.   Q.E.D.

> **Remark 2.75 (Notion of Topological Invariance).** A "topological property" (*i.e.* a property that only depends upon the topological space) is said to be *topologically invariant* if it is shared between homeomorphic spaces.
>
> For example, connectedness, compactness, path-connectedness, etc., are topologically invariant properties.
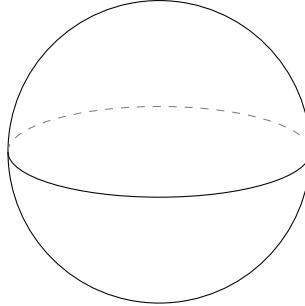
Example 2.76 (2-sphere S^2)

The 2-sphere S^2 is defined as

$$S^2 := \Big\{ (x, y, z) \in \mathbb{R}^3 \ \Big| \ x^2 + y^2 + z^2 = 1 \Big\}. \tag{2.36}$$

We define the topological space $(S^2, \mathcal{O})$ where $\mathcal{O}$ is the induced topology from $(\mathbb{R}^3, \mathcal{O}_{\text{std.}})$.


Figure 2.4: A homeomorphic representation of the 2-sphere

The sphere has the property that all the loops at any point are homotopic, hence the fundamental group (at every point) of the sphere is the trivial group:
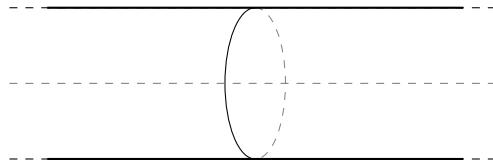
$$\forall p \in S^2 : \pi_1(p) = \{[\gamma_{e,p}]\}. \tag{2.37}$$

Example 2.77 (Cylinder $S^1 \times \mathbb{R}$)

The cylinder C is defined as

$$C := \mathbb{R} \times S^1 \tag{2.38}$$

equipped with product topology.

Figure 2.5: A homeomorphic representation of the cylinder

A loop in C can either go around the cylinder (*i.e.* around its central axis) or not. If it does not, then it can be continuously deformed to a point (the identity loop). If it does, then it cannot be deformed to the identity loop (intuitively because the cylinder is infinitely long) and hence it is a homotopically different loop. The number of times a loop winds around the cylinder is called the *winding number*. Loops with different winding numbers are not homotopic. Moreover, loops with different oreintations are also not homotopic. Hence, the fundamental group of the cylinder at any point is isomorphic to the additive group of integers:
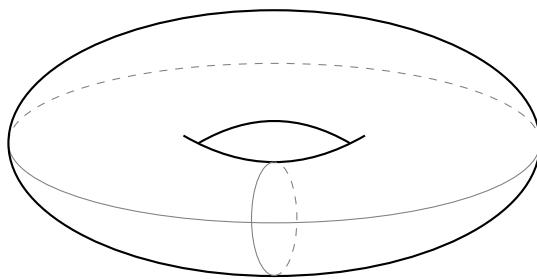
$$\forall p \in C : (\pi_1(p), \bullet_p) \cong_{\text{grp}} (\mathbb{Z}, +). \tag{2.39}$$

Example 2.78 (2-Torus $S^1 \times S^1$)

The 2-torus T^2 is defined as

$$T^2 := S^1 \times S^1 \tag{2.40}$$

equipped with product topology.

Figure 2.6: A homeomorphic representation of the 2-torus

For the 2-torus, the fundamental group at any point is isomorphic to the additive group of integers squared:

$$\forall p \in T^2 : (\pi_1(p), \bullet_p) \cong_{\mathrm{grp}} (\mathbb{Z} \times \mathbb{Z}, +). \tag{2.41}$$

Here the group operation is component-wise addition.

But still, the question remains: does a complete list of topological invariants exist *i.e.* can we conclude that two spaces are homeomorphic if and only if they have the same topological invariants?

This problem is still open in topology and is known as the *classification of topological spaces.*

3 Topological Manifolds and Bundles

Roughly speaking, a topological manifold is a topological space that locally looks like $\mathbb{R}^d$ for some fixed $d \in \mathbb{N}$. For example, 2-sphere is a topological manifold as well as a 2-torus and a pretzel with $d = 2$ in each case.

§3.1 Definition and Construction of Topological Manifolds

> **Definition 3.1 (Topological Manifold):**
> *A paracompact Hausdorff topological space $(M, \mathcal{O})$ is called a d-dimensional (topological) manifold if for every $p \in M$ there exists an open neighborhood $U_p \in \mathcal{O}$ of p and a homeomorphism $x : U_p \to x(U_p) \subseteq \mathbb{R}^d$. And we write $\dim M = d$.*

> **Example 3.2**
> Let's go through some known topological spaces to see if they are topological manifolds.
>
> 1. Trivially, $\mathbb{R}^d$ is a d-dimensional topological manifold for all $d \geq 1$.
>
> 2. The 1-sphere S^1 is a 1-dimensional topological manifold.
>
> 3. The topological spaces S^2, C and T^2 are 2-dimensional topological manifolds.

> **Remark 3.3 (Local Homeomorphisms vs. Homeomorphisms).** The definition of a topological manifold requires only local homeomorphisms. It is not necessary for the whole space to be homeomorphic to $\mathbb{R}^d$.
> For example, the 1-sphere S^1 is a 1-dimensional topological manifold, but it is not topologically isomorphic to $\mathbb{R}$.

§3.1.1 Submanifolds
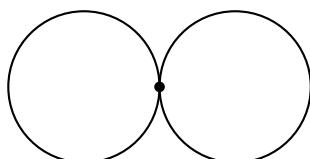
> **Definition 3.4 (Submanifold):**
> *Let $(M, \mathcal{O})$ be a topological manifold and $N \subseteq M$ be a subset. Then $(N, \mathcal{O}|_N)$ is called a submanifold of $(M, \mathcal{O})$ if it is a topological manifold in its own right.*
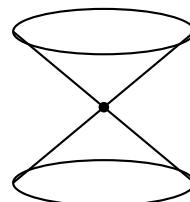
> **Example 3.5**
> The manifold S^1 is a submanifold of $\mathbb{R}^2$ while the manifolds S^2, C and T^2 are submanifolds of $\mathbb{R}^3$.

Counterexample: The figure-eight curve is not a submanifold of $\mathbb{R}^2$. As at the intersection point, we can't define dimension.

(a) Figure-Eight in $\mathbb{R}^2$ **(b)** Light-Cone in Minkowski Spacetime

Figure 3.1: Non-Submanifolds

We can see this in case of light-cone in Minkowski spacetime as well. The light-cone is also not a submanifold of Minkowski spacetime.

§3.1.2 Product Manifolds

> **Definition 3.6 (Product Manifold):**
> *Let $(M, \mathcal{O}_M)$ and $(N, \mathcal{O}_N)$ be topological manifolds. Then $(M \times N, \mathcal{O}_{M \times N})$ is a topological manifold called the* product manifold *with* $\dim(M \times N) = \dim M + \dim N$.
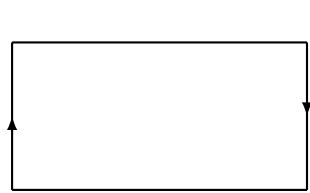
> **Example 3.7**
> This shows that $T^2 = S^1 \times S^1$ is a topological manifold of dimension 2. And we can generalize this to define *n-torus* as
> $$T^n := \underbrace{S^1 \times \cdots \times S^1}_{n \text{ times}} \tag{3.1}$$
> which is a topological manifold of dimension n.
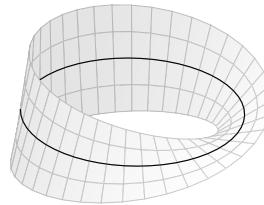
Products are very useful. Very often in physics one intuitively thinks of the product of two manifolds as attaching a copy of the second manifold to each point of the first.

Counterexample: Möbius strip is not a product manifold.

(a) Möbius Strip (easy to draw)   **(b)** Möbius Strip (hard to draw)

Figure 3.2: Non-Product Manifolds

§3.2 Bundles

> **Definition 3.8 (Bundle):**
> *A* bundle *(of a topological space) is a triple (E, π, M) where E and M are topological spaces called the* total space *and the* base space *respectively, and $\pi : E \to M$ is a continuous surjective map called the* projection map.
> *We often denote the bundle as*
> $$E \xrightarrow{\ \pi\ } M \tag{3.2}$$

> **Definition 3.9 (Fiber):**
> *Let (E, π, M) be a bundle and $p \in M$. Then the set $F_p := \operatorname{preim}_\pi(\{p\})$ is called the* fiber *over p.*

> **Example 3.10 (Product Bundle)**
> Let M and F be topological spaces. Define the product space $E := M \times F$ and the projection map $\pi : E \to M$ as $\pi(m, f) = m$. Then (E, π, M) is a bundle called the *product bundle*.
> Define $\pi' : E \to F$ as $\pi'(m, f) = f$. Then (E, π', F) is also a product bundle.

> **Example 3.11 (Möbius Strip as a Bundle)**
> Consider the Möbius strip as a bundle. Define the total space E as the Möbius strip and the projection map $\pi : E \to S^1$ as the projection of the Möbius strip onto the circle. Then (E, π, S^1) is a bundle.
>
> It is easy to see that the fiber over any point $p \in S^1$ is a line segment. See fig. 3.3.
> $$\forall p \in S^1 : F_p = [-1, 1] \tag{3.3}$$
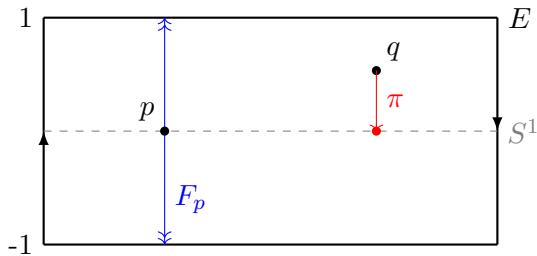
Figure 3.3: Möbius Strip as a Bundle

In both the examples, the fiber over any point in the base space is the same. But this is not necessary. The fiber can be different over different points in the base space.

Example 3.12

Consider the base space $M = \mathbb{R}$ and the total space as defined in fig. 3.4.
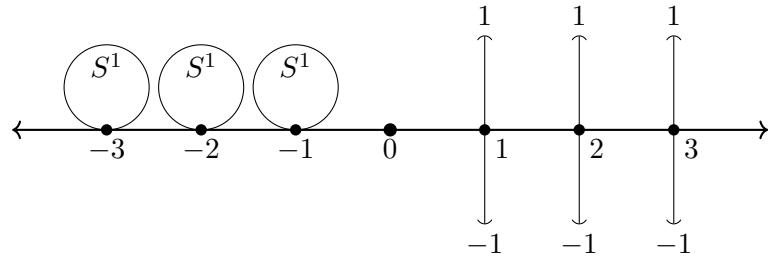
Figure 3.4: Total Space E

Here, the fiber over every point in the base space is not the same.

$$\forall p \in \mathbb{R} : F_p \cong_{\text{top.}} \begin{cases} S^1 & \text{if } p < 0 \\ \{0\} & \text{if } p = 0 \\ (1, -1) & \text{if } p > 0 \end{cases} \tag{3.4}$$

§3.2.1 Fiber Bundles

Definition 3.13 (Fiber Bundle):
Let $E \xrightarrow{\pi} M$ be a bundle. Then (E, π, M) is called a fiber bundle *if*

$$\exists F (= \text{topological space}) : \forall p \in M : F_p \cong_{top.} F \tag{3.5}$$

We can think of a fiber bundle as an attachment of a copy of the fiber F to each point of the base space M. Notationally, we write

$$\begin{array}{ccc} F & \longrightarrow & E \\ & & \downarrow \pi \\ & & M \end{array} \tag{3.6}$$

It is easy to see that the product bundle is a fiber bundle. But the converse is not true.

Example 3.14 (Möbius Strip as a Fiber Bundle)
Let Möbius strip be our total space and 1-sphere S^1 be our base space. Then the Möbius strip is a fiber bundle over the 1-sphere with the typical fiber $[-1, 1]$. See fig. 3.3.
Keep in mind that $E \neq S^1 \times [-1, 1]$ *i.e.* Möbius strip is not a product bundle.

Example 3.15 ($\mathbb{C}^1$-line bundle)

A $\mathbb{C}$-line bundle over M is the fiber bundle (E, π, M) with fiber $\mathbb{C}$.

Note that $M \times \mathbb{C} \xrightarrow{\pi} M$ is a $\mathbb{C}$-line bundle over M, but a $\mathbb{C}$-line bundle over M need not be a product bundle.

Definition 3.16 (Section):

Let $E \xrightarrow{\pi} M$ be a bundle. A map $\sigma : M \to E$ is called a (cross-)section of the bundle if $\pi \circ \sigma = \mathrm{id}_M$.

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

Example 3.17 (Special Case for Product Bundles)

Let $M \times F \xrightarrow{\pi} M$ be a product bundle. Let $s : M \to F$ be a map. Then the map

$$\begin{aligned} \sigma : M &\to M \times F \\ p &\mapsto (p, s(p)) \end{aligned} \tag{3.7}$$

is a section of the product bundle.

Thus, for product bundles, sections are in one-to-one correspondence with maps $M \to F$ *i.e.* by choosing a map $s : M \to F$, we get a section $\sigma : M \to M \times F$.

In case of product bundles, we can analyze 'section' of the bundle by looking at the corresponding map $s : M \to F$.

With this we can order the bundles as follows:

$$\text{Product Bundles} \subset \text{Fiber Bundles} \subset \text{Bundles} \tag{3.8}$$

At this point, we can have a physics example (from quantum mechanics) but to understand it fully, we require more maturity.

Example 3.18 (Wave functions)

Consider our base space M to be our physical space (for case $\mathbb{R}^3$). The mathematical structure, we are analyzing in quantum mechanics is $\mathbb{C}^1$-line bundle over physical space,

$$\text{wave function} := \text{section of } (E, \pi, M). \tag{3.9}$$

In case, E is a product manifold of M and $\mathbb{C}$, then wave function is a function.

---

§3.2.2 Constructing Bundles

Definition 3.19 (Sub-Bundle):

Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be bundles. Then (E', π', M') is called a sub-bundle of (E, π, M) if

1. *$E' \subseteq E$ be a submanifold,*

2. *$M' \subseteq M$ be a submanifold, and*

3. *$\pi' = \pi|_{E'}$.*

Definition 3.20 (Restricted Bundle):

Let $E \xrightarrow{\pi} M$ be a bundle and $N \subseteq M$ be a submanifold. Then the bundle

$$\mathrm{preim}_\pi(N) \xrightarrow{\pi|_{\mathrm{preim}_\pi(N)}} N \tag{3.10}$$

is called the restricted bundle of (E, π, M) to N.

§3.3 Bundle Morphisms

Definition 3.21 (Bundle Morphism):
Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be bundles and $u : E \to E'$ and $v : M \to M'$ be continuous maps. Then (u, v) is called a bundle morphism if

$$
\begin{array}{ccc}
E & \xrightarrow{\;u\;} & E' \\
\pi \downarrow & & \downarrow \pi' \\
M & \xrightarrow{\;v\;} & M'
\end{array}
\tag{3.11}
$$

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

Definition 3.22 (Bundle Isomorphism):
Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be bundles. They are called isomorphic (as bundles) if there exist bundle morphisms (u, v) such that

$$
\begin{array}{ccc}
E & \underset{u^{-1}}{\overset{u}{\rightleftarrows}} & E' \\
\pi \downarrow & & \downarrow \pi' \\
M & \underset{v^{-1}}{\overset{v}{\rightleftarrows}} & M'
\end{array}
\tag{3.12}
$$

is a commutative diagram. Such a bundle morphism (u, v) is called a bundle isomorphism. And we write $E \xrightarrow{\pi} M \cong_{bdl} E' \xrightarrow{\pi'} M'$.

Bundle isomorphisms are structure-preserving maps between bundles.

Definition 3.23 (Local Isomorphism):
Let $E \xrightarrow{\pi} M$ and $E' \xrightarrow{\pi'} M'$ be bundles. Then $E \xrightarrow{\pi} M$ is called locally isomorphic to $E' \xrightarrow{\pi'} M'$ if

$$
\forall p \in M : \exists U_p \in \mathcal{O}_M : \mathrm{preim}_\pi(U_p) \xrightarrow{\;\pi|_{\mathrm{preim}_\pi(U_p)}\;} U_p \cong_{bdl} E' \xrightarrow{\pi'} M'.
\tag{3.13}
$$

Some useful terminologies:

Definition 3.24:
A bundle $E \xrightarrow{\pi} M$ is called

(a) trivial if it is isomorphic to a product bundle $M \times F \xrightarrow{\pi_1} M$,

(b) locally trivial if it is locally isomorphic to some product bundle.

It is obvious that every trivial bundle is locally trivial. But the converse is not true.

Example 3.25
Let's see some insightful examples:

1. A cylinder C is trivial $\Rightarrow$ locally trivial.

2. Möbius Strip is not trivial, but it is locally trivial.

3. A very specific construction, such that the bundle is not even locally trivial.
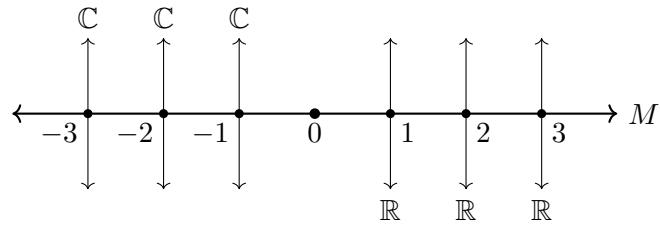
Figure 3.5: Non Locally Trivial bundle

From now on, we will be focusing on locally trivial bundles unless mentioned otherwise.

Remark 3.26. Since we have restricted ourselves to locally trivial bundles. Thus, any section can be thought of as a map $M \to F$ locally.

Example 3.27 (Wave functions (revisited))

Assume that our $\mathbb{C}^1$-line bundle over $M(= \mathbb{R}^3)$ is locally trivial. Then the wave function is a function $M \to \mathbb{C}$ locally.

Definition 3.28 (Pull-Back Bundle):

Let $E \xrightarrow{\pi} M$ be a bundle and $f : M' \to M$ be a map from another manifold M' to M. Define the manifold E' as

$$E' := \{(m', e) \in M' \times E \mid f(m') = \pi(e)\}. \tag{3.14}$$

Then the bundle (E', π', M') with $\pi'(m', e) = m'$ is called the pull-back bundle of $E \xrightarrow{\pi} M$ induced by $f : M' \to M$.

Remark 3.29 (Bundle Morphism on Pull-Back Bundle). Let $E' \xrightarrow{\pi'} M'$ be a pull-back bundle of $E \xrightarrow{\pi} M$ induced by $v : M' \to M$. Define the map

$$\begin{aligned} u : E' &\to E \\ (m', e) &\mapsto e. \end{aligned} \tag{3.15}$$

Then (u, v) is a bundle morphism. This corresponds to the commutative diagram

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

Remark 3.30 (Pull-back of Sections). Let $E' \xrightarrow{\pi'} M'$ be a pull-back bundle of $E \xrightarrow{\pi} M$ induced by $f : M' \to M$. Let σ be a section of $E \xrightarrow{\pi} M$.

Observe that $\sigma \circ f$ is a map $M' \to E$. Using the fact that σ is a section, we have,

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

This can be interpreted as $\pi((\sigma \circ f)(m')) = f(m')$. Thus, $(m', (\sigma \circ f)(m')) \in E'$. Define the map σ' as

$$\begin{aligned} \sigma' : M' &\to E' \\ m' &\mapsto (m', (\sigma \circ f)(m')). \end{aligned} \tag{3.16}$$

Furthermore, $\pi' \circ \sigma' = \mathrm{id}_{M'}$ as $\pi'(m', (\sigma \circ f)(m')) = m'$. Thus, σ' is a section of the pull-back bundle.

§3.4 Viewing Manifolds from Atlases

Definition 3.31 (Chart):
Let $(M, \mathcal{O})$ be a d-dimensional manifold. Then the pair (U, x), where $U \in \mathcal{O}$ and $x : U \to x(U) \subseteq \mathbb{R}^d$ is a homeomorphism, is called a chart *on M.*

Remark 3.32 (Component Functions). Let (U, x) be a chart on M. Then the *component functions* of x are the functions

$$x^i : U \to \mathbb{R}$$
$$p \mapsto \text{proj}^i(x(p)) \tag{3.17}$$

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

Definition 3.33 (Atlas):
Let $(M, \mathcal{O})$ be a d-dimensional manifold. Then the collection of all charts $\mathscr{A} := \{(U_\alpha, x_\alpha) \mid \alpha \in I\}$ is called an atlas *of the manifold M if,*

$$\bigcup_{\alpha \in I} U_\alpha = M, \tag{3.18}$$

for some index set I.

This is obvious to see that this collection is a set, as from the definition of manifolds, we have an open set U_p at each point with an appropriate homeomorphism x_p. Thus, this provides a trivial atlas of M,
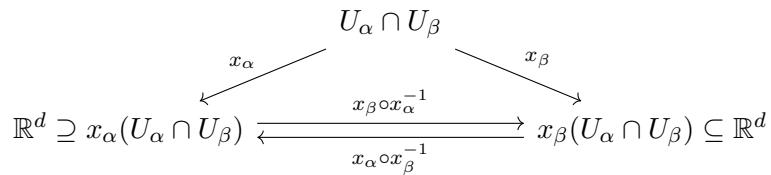
$$\mathscr{A} = \{(U_p, x_p) \mid p \in M\} \tag{3.19}$$

Definition 3.34 ($\mathcal{C}^0$-Compatibility):
Let (U_α, x_α) and (U_β, x_β) be two charts of a manifold M. They are called $\mathcal{C}^0$-compatible either

(a) $U_\alpha \cap U_\beta = \emptyset$, *or*

(b) *the map $x_\beta \circ x_\alpha^{-1} : x_\alpha(U_\alpha \cap U_\beta) \to x_\beta(U_\alpha \cap U_\beta)$ is continuous.*

$x_\beta \circ x_\alpha^{-1}$ (and its inverse $x_\alpha \circ x_\beta^{-1}$) is a map from a subset of $\mathbb{R}^d$ to another subset of $\mathbb{R}^d$.

$$U_\alpha \cap U_\beta$$

$$\overset{x_\alpha}{\swarrow} \qquad \overset{x_\beta}{\searrow}$$

$$\mathbb{R}^d \supseteq x_\alpha(U_\alpha \cap U_\beta) \underset{x_\alpha \circ x_\beta^{-1}}{\overset{x_\beta \circ x_\alpha^{-1}}{\rightleftarrows}} x_\beta(U_\alpha \cap U_\beta) \subseteq \mathbb{R}^d$$

Here it may seem like a redundant definition, since x_α and x_β are homeomorphisms, the composition map $x_\beta \circ x_\alpha^{-1}$ (or its inverse $x_\alpha \circ x_\beta^{-1}$) is also a homeomorphism thus continuous. Therefore, any two charts on a topological manifold are $\mathcal{C}^0$-compatible.

This notion will be useful, when we want to talk about differentiable manifolds. In that case, we require the map $x_\beta \circ x_\alpha^{-1}$ to be continuously differentiable or k-smooth for $\mathcal{C}^1$ or $\mathcal{C}^k$-compatible respectively using the notion of differentiability on $\mathbb{R}^d$.

> **Remark 3.35 (Co-ordinate Change).** The map $\tau_{\alpha,\beta} := x_\beta \circ x_\alpha^{-1}$ (and its inverse $\tau_{\beta,\alpha} := x_\alpha \circ x_\beta^{-1}$) is called the *co-ordinate change map* or *chart transition map*.

This also express that, all the physics that we have done so far lives in chart. For example, the motion of a particle is a subset of our physical manifold, and we use different co-ordinates (or charts) for our understanding.

> **Definition 3.36 ($\mathcal{C}^0$-Atlas):**
> *An atlas $\mathscr{A}$ is a $\mathcal{C}^0$-atlas if all the charts in $\mathscr{A}$ are pairwise $\mathcal{C}^0$-compatible.*

Any atlas is a $\mathcal{C}^0$-atlas.

> **Definition 3.37 (Maximal Atlas):**
> *A $\mathcal{C}^0$-atlas $\mathscr{A}$ is called a* maximal *if it contains every chart that is $\mathcal{C}^0$-compatible with every chart in $\mathscr{A}$.*
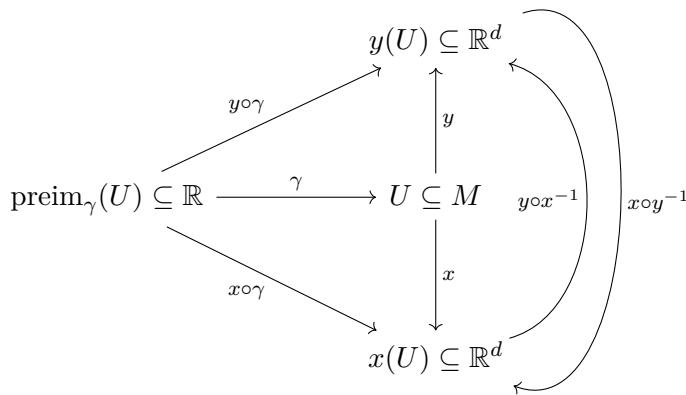
> **Example 3.38**
> Let $(\mathbb{R}, \mathcal{O}_{\text{std.}})$ be our topological manifold. Then we have a trivial atlas $\mathscr{A} = \{(\mathbb{R}, \text{id}_{\mathbb{R}})\}$, but it is not maximal. Since chart $((0,1), \text{id}_{(0,1)})$ is $\mathcal{C}^0$-compatible with the chart $(\mathbb{R}, \text{id}_{\mathbb{R}})$, but it is not in the atlas $\mathscr{A}$.

We can look at "objects on" topological manifolds from two point of view. Consider a curve γ in a d-dimensional manifold, *i.e.* $\gamma : \mathbb{R} \to M$. From our previous knowledge, we know this curve has to be continuous if it describes the trajectory of a classical particle.

One way to verify this, let an open set in M and check whether its pre-image is open or not (topological definition).

Another way is through the concept of charts. Let (U, x) be a chart, rather than studying γ directly, we will focus on $x \circ \gamma : \mathbb{R} \supseteq \text{preim}_\gamma(U) \to x(U) \subseteq \mathbb{R}^d$. Checking continuity of this composition map is easier as well as guarantees the continuity of γ.

Let the problem is very complicated in this co-ordinate system, let (U, y) be another chart. And this change of co-ordinate is facilitated by transition maps. We can summarize all this in the following commutative diagram:

**Figure 3.6:** Looking at Continuity of a curve in a manifold using charts

Intuitively, this description allows us to forget about the inner structure (*i.e.* U and the maps γ, x and y) which, in a sense, is the real world, and only consider $\text{preim}_\gamma(U) \subseteq \mathbb{R}$ and $x(U), y(U) \subseteq \mathbb{R}^d$ together with the maps between them, which is our representation of the real world.

4 Differentiable Manifolds

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

> **Definition 4.1 (❀-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 $\mathbb{R}^d$.

$$
\begin{array}{ccc}
 & U \cap V & \\
{\scriptstyle x}\swarrow & & \searrow{\scriptstyle y} \\
\mathbb{R}^d \supseteq x(U \cap V) & \underset{x \circ y^{-1}}{\overset{y \circ x^{-1}}{\rightleftarrows}} & y(U \cap V) \subseteq \mathbb{R}^d
\end{array}
$$

Now, we can define the placeholder symbol ❀ as:

- ❀ $= \mathcal{C}^0$: see definition 3.36.

- ❀ $= \mathcal{C}^k$: 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

$$
\begin{aligned}
f : \mathbb{C} &\to \mathbb{C} \\
x + \mathrm{i}y &\mapsto u(x, y) + \mathrm{i}v(x, y)
\end{aligned}
\tag{4.1}
$$

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:

$$
\frac{\partial u}{\partial x}(x_0, y_0) = \frac{\partial v}{\partial y}(x_0, y_0) \quad \text{and} \quad \frac{\partial u}{\partial y}(x_0, y_0) = -\frac{\partial v}{\partial x}(x_0, y_0).
\tag{4.2}
$$

> **Theorem 4.2 (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 4.3 ($\mathcal{C}^k$-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 4.4 (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 4.5. A given topological manifold $(M, \mathcal{O})$ can have different incompatible atlases.

A simple example of incompatible atlases,

> **Example 4.6**
> 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 4.7 (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 $\mathbb{R}^n$.*
>
> $$
> \begin{array}{ccc}
> M \supseteq U & \overset{f}{\longrightarrow} & N \subseteq V \\
> \downarrow{\scriptstyle x} & & \downarrow{\scriptstyle y} \\
> \mathbb{R}^m \supseteq x(U) & \overset{y \circ f \circ x^{-1}}{\longrightarrow} & y(V) \subseteq \mathbb{R}^n
> \end{array}
> $$
>
> *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 4.8**
> 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:

$$
\begin{array}{ccc}
\mathbb{R}^m \supseteq \tilde{x}(U \cap \tilde{U}) & \xrightarrow{\ \tilde{y} \circ f \circ \tilde{x}^{-1}\ } & \tilde{y}(V \cap \tilde{V}) \subseteq \mathbb{R}^n \\
\Big\uparrow \tilde{x} & & \Big\uparrow \tilde{y} \\
M \supseteq U \cap \tilde{U} \ni p & \xrightarrow{\quad f \quad} & f(p) \in V \cap \tilde{V} \subseteq N \\
\Big\downarrow x & & \Big\downarrow y \\
\mathbb{R}^m \supseteq x(U \cap \tilde{U}) & \xrightarrow{\ y \circ f \circ x^{-1}\ } & y(V \cap \tilde{V}) \subseteq \mathbb{R}^n
\end{array}
$$

with outer arrows $x \circ \tilde{x}^{-1}$ (left) and $\tilde{y} \circ y^{-1}$ (right).

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

$$\tilde{y} \circ f \circ \tilde{x}^{-1} = (\tilde{y} \circ y^{-1}) \circ (y \circ f \circ x^{-1}) \circ (x \circ \tilde{x}^{-1}) \tag{4.3}$$

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 $(\tilde{V}, \tilde{y})$. $\quad$ Q.E.D.

> **Definition 4.9 (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 4.10 (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*
>
> $$M \cong_{diff} N \quad or \quad (M, \mathcal{O}_M, \mathscr{A}_M) \cong_{diff} (N, \mathcal{O}_N, \mathscr{A}_N). \tag{4.4}$$

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

$$(\mathbb{R}, \mathcal{O}_{\text{std.}}, \mathscr{A}_{1,\text{max}}) \cong_{\text{diff}} (\mathbb{R}, \mathcal{O}_{\text{std.}}, \mathscr{A}_{2,\text{max}}) \tag{4.5}$$

where $\mathscr{A}_{1,\text{max}}$ and $\mathscr{A}_{2,\text{max}}$ 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 4.11 (Radon-Moise)**
> For $d \leq 3$, 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 $d \leq 3$. Then
>
> $$(M, \mathcal{O}_M, \mathscr{A}_M) \cong_{\text{diff}} (N, \mathcal{O}_N, \mathscr{A}_N) \Leftrightarrow (M, \mathcal{O}_M) \cong_{\text{top.}} (N, \mathcal{O}_N). \tag{4.6}$$
>
> 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 4.12 (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 4.13 (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.

$$\boxed{\text{L\textsc{ecture} 8}}$$

§4.2 Review of Vector Spaces

Definition 4.14 ((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 $a, b, c \in K$,*

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad and \quad (a + b) \cdot c = a \cdot c + b \cdot c. \tag{4.7}$$

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 4.15 (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 $a, b, c \in R$,*

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad and \quad (a + b) \cdot c = a \cdot c + b \cdot c. \tag{4.8}$$

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 4.16 (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 4.17 ((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.*
>
> $$\forall \mathbf{u}, \mathbf{v} \in U : \mathbf{u} \oplus \mathbf{v} \in U \quad and \quad \forall a \in K : \forall \mathbf{u} \in U : a \odot \mathbf{u} \in U. \tag{4.9}$$

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 4.18 (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:

$$\forall \mathbf{v}_1, \mathbf{v}_2 \in V, \forall a, b \in K : L(a \odot \mathbf{v}_1 \oplus b \odot \mathbf{v}_2) = a \boxdot L(\mathbf{v}_1) \boxplus b \boxdot L(\mathbf{v}_2). \tag{4.10}$$

Remark 4.19 (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. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

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

$$\mathrm{Hom}(V, W) := \{L : V \to W \mid L \text{ is a linear map}\} \equiv \left\{L : V \xrightarrow{\sim} W\right\} \tag{4.11}$$

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

> **Proposition 4.20**
> $\mathrm{Hom}(V, W)$ 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
>
> $$(L_1, L_2) \xmapsto{\oplus} L_1 \oplus L_2 \tag{4.12}$$

where

$$L_1 \oplus L_2 : V \xrightarrow{\sim} W, \quad \mathbf{v} \mapsto (L_1 \oplus L_2)(\mathbf{v}) := L_1(\mathbf{v}) \boxplus L_2(\mathbf{v}). \tag{4.13}$$

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

$$(a, L) \xmapsto{\diamondsuit} a \diamond L \tag{4.14}$$

where

$$a \diamond L : V \xrightarrow{\sim} W, \quad \mathbf{v} \mapsto (a \diamond L)(\mathbf{v}) := a \boxdot L(\mathbf{v}). \tag{4.15}$$

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 $\oplus$ and $\diamond$.

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

$$(L_1 \oplus L_2)(\mathbf{v}) = L_1(\mathbf{v}) \boxplus L_2(\mathbf{v}). \tag{4.16}$$

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

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

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

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

$$(a \diamond L)(\mathbf{v}) = a \boxdot L(\mathbf{v}). \tag{4.17}$$

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

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

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

We have shown that $\mathrm{Hom}(V, W)$ is closed under both operations $\oplus$ and $\diamond$. 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 4.21 (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.

$$\mathrm{End}(V) := \mathrm{Hom}(V, V) = \left\{ L : V \xrightarrow{\sim} V \right\} \tag{4.18}$$

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

$$\mathrm{Aut}(V) := \left\{ L : V \xrightarrow{\sim} V \ \middle|\ L \text{ is a linear isomorphism} \right\}. \tag{4.19}$$

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:

$$V^* := \mathrm{Hom}(V, K) \tag{4.20}$$

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

§4.2.1 Tensors

Definition 4.22 (Tensor):
A type (p,q)-tensor is a multilinear map (linear in each argument) of the form

$$T : \underbrace{V^* \times \cdots \times V^*}_{p \ times} \times \underbrace{V \times \cdots \times V}_{q \ times} \to K \tag{4.21}$$

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

$$\mathsf{T}^p_q V := \underbrace{V \otimes \cdots \otimes V}_{p \ times} \otimes \underbrace{V^* \otimes \cdots \otimes V^*}_{q \ times} = \{T \mid T \text{ is type } (p,q)\text{-tensor}\} \tag{4.22}$$

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:
$$\oplus : \mathsf{T}^p_q V \times \mathsf{T}^p_q V \to \mathsf{T}^p_q V \tag{4.23}$$
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

$$(T_1 \oplus T_2)(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) = T_1(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) + T_2(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) \tag{4.24}$$

2. Scalar multiplication of tensors:
$$\odot : K \times \mathsf{T}^p_q V \to \mathsf{T}^p_q V \tag{4.25}$$
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

$$(a \odot T)(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) = a \cdot T(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) \tag{4.26}$$

Proposition 4.23
$\mathsf{T}^p_q V$ 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 4.24 (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

$$(T_1 \otimes T_2)(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_{p+r}, \mathbf{v}_1, \ldots, \mathbf{v}_{q+s}) =$$
$$T_1(\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_p, \mathbf{v}_1, \ldots, \mathbf{v}_q) \cdot T_2(\boldsymbol{\omega}_{p+1}, \ldots, \boldsymbol{\omega}_{p+r}, \mathbf{v}_{q+1}, \ldots, \mathbf{v}_{q+s}). \tag{4.27}$$

Example 4.25

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

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

$$\forall \mathbf{v} \in V : L_{\boldsymbol{\omega}}(\mathbf{v}) = T(\boldsymbol{\omega}, \mathbf{v}). \tag{4.28}$$

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

$$L(\boldsymbol{\omega}) = L_{\boldsymbol{\omega}} \qquad \Rightarrow \qquad \boldsymbol{\omega} \mapsto \underbrace{(\mathbf{v} \mapsto T(\boldsymbol{\omega}, \mathbf{v}))}_{L_{\boldsymbol{\omega}}}. \tag{4.29}$$

Now, let $L \in \mathrm{End}(V^*)$ 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

$$T(\boldsymbol{\omega}, \mathbf{v}) = L(\boldsymbol{\omega})(\mathbf{v}) \tag{4.30}$$

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

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

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

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

$$(V^*)^* \not\cong_{\mathrm{vec}} V. \tag{4.31}$$

This is true only for finite-dimensional vector spaces.

§4.2.2 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 4.26 (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*

$$\sum_{i=1}^{N} \lambda^i \cdot \mathbf{b}_i = \mathbf{0} \quad \text{for } \lambda^i \in K \tag{4.32}$$

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.*

$$\forall \mathbf{v} \in V : \exists M \in \mathbb{N} : \exists v^1, \ldots, v^M \in K : \exists \mathbf{b}_1, \ldots, \mathbf{b}_M \in \mathcal{B} : \mathbf{v} = \sum_{i=1}^{M} v^i \cdot \mathbf{b}_i. \tag{4.33}$$

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 4.27

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

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

Definition 4.28 (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 $\dim V$.*

$$\dim V := |\mathcal{B}| \quad \text{for any Hamel basis } \mathcal{B} \subseteq V. \tag{4.34}$$

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 4.29

Let V be a finite-dimensional vector space. Then

$$(V^*)^* \cong_{\text{vec}} V. \tag{4.35}$$

Proof:

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

$$\begin{aligned} L : V &\to (V^*)^* \\ \mathbf{v} &\mapsto L_{\mathbf{v}} \end{aligned} \tag{4.36}$$

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

$$\begin{aligned} L_{\mathbf{v}} : V^* &\to K \\ \boldsymbol{\omega} &\mapsto L_{\mathbf{v}}(\boldsymbol{\omega}) := \boldsymbol{\omega}(\mathbf{v}) \end{aligned} \tag{4.37}$$

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 $\mathrm{span}(\mathbf{v}) \subseteq V$ 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

$$
\mathbf{u} = b \cdot \mathbf{v} + \sum_{i=2}^{n} b^i \cdot \mathbf{e}_i \tag{4.38}
$$

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

$$
\mathbf{u} \mapsto \boldsymbol{\omega}(\mathbf{u}) := b. \tag{4.39}
$$

Thus, we have

$$
L(\mathbf{v})(\boldsymbol{\omega}) = \boldsymbol{\omega}(\mathbf{v}) = 1 \neq 0. \tag{4.40}
$$

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
>
> $$
> \forall i, j \in \{1, \ldots, n\} : \mathbf{f}^i(\mathbf{e}_j) = \delta^i_j := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{4.41}
> $$

(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:

$$
\mathbf{v} = \sum_{i=1}^{n} F(\mathbf{f}^i) \cdot \mathbf{e}_i \tag{4.42}
$$

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.

<div align="right">Q.E.D.</div>

> **Corollary 4.30**
> 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 4.31 (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:*
>
> $$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} := T(\boldsymbol{\epsilon}^{i_1}, \ldots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q}) \quad \text{for } i_k, j_l \in \{1, \ldots, n\}. \tag{4.43}$$

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

$$T = \underbrace{\sum_{i_1=1}^{n} \cdots \sum_{j_q=1}^{n}}_{p+q \text{ sums}} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \cdot \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_q}. \tag{4.44}$$

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

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

$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \cdot \mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_p} \otimes \boldsymbol{\epsilon}^{j_1} \otimes \cdots \otimes \boldsymbol{\epsilon}^{j_q}. \tag{4.45}$$

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.

$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} := T(\boldsymbol{\epsilon}^{i_1}, \ldots, \boldsymbol{\epsilon}^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q}). \tag{4.46}$$

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

$$\mathbf{v} = v^i \cdot \mathbf{e}_i \quad \text{where } v^i = \boldsymbol{\epsilon}^i(\mathbf{v}) \tag{4.47}$$

$$\boldsymbol{\omega} = \omega_j \cdot \boldsymbol{\epsilon}^j \quad \text{where } \omega_j = \mathbf{e}_j(\boldsymbol{\omega}). \tag{4.48}$$

Remark 4.32 (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*.

§4.2.3 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)

$$\tilde{\mathbf{e}}_i = A^j{}_i \mathbf{e}_j \tag{4.49}$$

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

$$\mathbf{e}_i = B^j{}_i \tilde{\mathbf{e}}_j \tag{4.50}$$

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

$$A^{-1} = B \quad \text{and} \quad B^{-1} = A. \tag{4.51}$$

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 4.33 (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

$$
\mathbf{v} = v^i \mathbf{e}_i \quad \leftrightsquigarrow \quad \mathbf{v} \,\hat{=}\, \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \tag{4.52}
$$

and

$$
\boldsymbol{\omega} = \omega_j \boldsymbol{\epsilon}^j \quad \leftrightsquigarrow \quad \boldsymbol{\omega}^\top \,\hat{=}\, \begin{pmatrix} \omega_1 & \cdots & \omega_n \end{pmatrix}. \tag{4.53}
$$

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

$$
\phi = \phi^i{}_j \mathbf{e}_i \otimes \boldsymbol{\epsilon}^j \quad \leftrightsquigarrow \quad \Phi \,\hat{=}\, \begin{pmatrix} \phi^1{}_1 & \cdots & \phi^1{}_n \\ \vdots & \ddots & \vdots \\ \phi^n{}_1 & \cdots & \phi^n{}_n \end{pmatrix}. \tag{4.54}
$$

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 $\text{End}(V) \cong_{\text{vec}} \mathsf{T}^1_1 V$. Let $\phi \in \text{End}(V)$, so we can rethink ϕ as a type $(1,1)$-tensor in following sense:

$$\phi(\boldsymbol{\omega}, \mathbf{v}) := \boldsymbol{\omega}(\phi(\mathbf{v})). \tag{4.55}$$

So the components of tensor ϕ is defined as,

$$\phi^i_{\ j} := \phi(\boldsymbol{\epsilon}^i, \mathbf{e}_j) = \boldsymbol{\epsilon}^i(\phi(\mathbf{e}_j)). \tag{4.56}$$

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

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

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

$$\Phi \cdot \Psi \,\hat{=}\, \left(\phi^i_{\ a} \cdot \psi^a_{\ j}\right)^n_{i,j=1} \tag{4.57}$$

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,

$$\phi(\boldsymbol{\omega}, \mathbf{v}) = \omega_i \phi^i_{\ j} v^j \quad \longleftrightarrow \quad \boldsymbol{\omega}^\top \cdot \Phi \cdot \mathbf{v} \tag{4.58}$$

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, $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\tilde{\mathcal{B}} = \{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_n\}$ of a vector space V. Then the change of basis for vectors is given by

$$\tilde{\mathbf{e}}_i = A^j_{\ i}\mathbf{e}_j \qquad\qquad \mathbf{e}_j = B^i_{\ j}\tilde{\mathbf{e}}_i. \tag{4.59}$$

Since these relations are invertible, we have

$$A^i_{\ m}B^m_{\ j} = B^i_{\ m}A^m_{\ j} = \delta^i_{\ j} \tag{4.60}$$

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

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

$$\omega_j = \mathbf{e}_j(\boldsymbol{\omega}) = \boldsymbol{\omega}(\mathbf{e}_j) = \boldsymbol{\omega}(B^i_{\ j}\tilde{\mathbf{e}}_i) = B^i_{\ j}\boldsymbol{\omega}(\tilde{\mathbf{e}}_i) = B^i_{\ j}\tilde{\mathbf{e}}_i(\boldsymbol{\omega}) = B^i_{\ j}\tilde{\omega}_i. \tag{4.61}$$

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

$$v^i = \boldsymbol{\epsilon}^i(\mathbf{v}) = \boldsymbol{\epsilon}^i(\tilde{v}^j\tilde{\mathbf{e}}_j) = \tilde{v}^j\boldsymbol{\epsilon}^i(\tilde{\mathbf{e}}_j) = \tilde{v}^j\boldsymbol{\epsilon}^i(A^k_{\ j}\mathbf{e}_k) = \tilde{v}^jA^k_{\ j}\boldsymbol{\epsilon}^i(\mathbf{e}_k) = \tilde{v}^jA^k_{\ j}\delta^i_{\ k} = A^i_{\ j}\tilde{v}^j. \tag{4.62}$$

Summarizing the above two equations, we have

$$v^i = A^i{}_j \tilde{v}^j \qquad\qquad \omega_j = B^i{}_j \tilde{\omega}_i \qquad\qquad (4.63)$$

$$\tilde{v}^i = B^i{}_j v^j \qquad\qquad \tilde{\omega}_j = A^i{}_j \omega_i. \qquad\qquad (4.64)$$

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

$$\epsilon^i(\mathbf{e}_j) = \delta^i_j \qquad\qquad \tilde{\epsilon}^i(\tilde{\mathbf{e}}_j) = \delta^i_j. \qquad\qquad (4.65)$$

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

$$\tilde{\epsilon}^i(\tilde{\mathbf{e}}_j) = C^i{}_k \epsilon^k(A^l{}_j \mathbf{e}_l) = C^i{}_k A^l{}_j \epsilon^k(\mathbf{e}_l) = C^i{}_k A^l{}_j \delta^k_l = C^i{}_k A^k{}_j = \delta^i_j. \qquad (4.66)$$

Thus, using unique existence of the inverse, we have

$$C^i{}_j = B^i{}_j. \qquad\qquad (4.67)$$

So we have the following relations for the dual basis:

$$\tilde{\epsilon}^i = B^i{}_j \epsilon^j \qquad\qquad \epsilon^j = A^j{}_i \tilde{\epsilon}^i. \qquad\qquad (4.68)$$

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

$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} = T(\epsilon^{i_1}, \ldots, \epsilon^{i_p}, \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q}). \qquad (4.69)$$

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

$$\tilde{T}^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} = T(\tilde{\epsilon}^{i_1}, \ldots, \tilde{\epsilon}^{i_p}, \tilde{\mathbf{e}}_{j_1}, \ldots, \tilde{\mathbf{e}}_{j_q}) \qquad (4.70)$$

$$= T(B^{i_1}{}_{a_1} \epsilon^{a_1}, \ldots, B^{i_p}{}_{a_p} \epsilon^{a_p}, A^{b_1}{}_{j_1} \mathbf{e}_{b_1}, \ldots, A^{b_q}{}_{j_q} \mathbf{e}_{b_q}) \qquad (4.71)$$

$$= B^{i_1}{}_{a_1} \cdots B^{i_p}{}_{a_p} \cdot A^{b_1}{}_{j_1} \cdots A^{b_q}{}_{j_q} \cdot T(\epsilon^{a_1}, \ldots, \epsilon^{a_p}, \mathbf{e}_{b_1}, \ldots, \mathbf{e}_{b_q}) \qquad (4.72)$$

$$= B^{i_1}{}_{a_1} \cdots B^{i_p}{}_{a_p} \cdot A^{b_1}{}_{j_1} \cdots A^{b_q}{}_{j_q} \cdot T^{a_1 \ldots a_p}{}_{b_1 \ldots b_q}. \qquad (4.73)$$

Thus, the reversal of the basis is given by

$$T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} = A^{i_1}{}_{a_1} \cdots A^{i_p}{}_{a_p} \cdot B^{b_1}{}_{j_1} \cdots B^{b_q}{}_{j_q} \cdot \tilde{T}^{a_1 \ldots a_p}{}_{b_1 \ldots b_q}. \qquad (4.74)$$

§4.2.4 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 4.34. 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

$$\phi = \phi^i{}_j \mathbf{e}_i \otimes \epsilon^j \quad \text{and} \quad g = g_{ij} \epsilon^i \otimes \epsilon^j. \qquad (4.75)$$

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

$$\Phi = \begin{pmatrix} \phi^1{}_1 & \cdots & \phi^1{}_n \\ \vdots & \ddots & \vdots \\ \phi^n{}_1 & \cdots & \phi^n{}_n \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} g_{11} & \cdots & g_{1n} \\ \vdots & \ddots & \vdots \\ g_{n1} & \cdots & g_{nn} \end{pmatrix}. \qquad (4.76)$$

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

$$\tilde{\phi}^i{}_j = B^i{}_k \cdot \phi^k{}_l \cdot A^l{}_j \quad \text{and} \quad \tilde{g}_{ij} = A^k{}_i \cdot g_{kl} \cdot A^l{}_j. \qquad (4.77)$$

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)

$$\tilde{\Phi} = A^{-1} \cdot \Phi \cdot A \quad \text{and} \quad \tilde{G} = A^{\top} \cdot G \cdot A. \tag{4.78}$$

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.

§4.2.4.1 Permutation Group

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

> **Definition 4.35 (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 $M = \{1, 2, \ldots, n\}$. 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:

$$\text{sgn}(\sigma) := \begin{cases} +1 & \text{if } \sigma \text{ is the product of an even number of transpositions} \\ -1 & \text{if } \sigma \text{ is the product of an odd number of transpositions} \end{cases} \tag{4.79}$$

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 4.36 (n-form):**
> *Let V be a finite-dimensional vector space of dimension d. An n-form on V (where $0 \le n \le d$) is a type $(0, n)$-tensor $\omega \in \mathsf{T}_n^0 V$ such that it is* totally antisymmetric, *i.e. for any n vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ and any permutation $\sigma \in S_n$, we have*
>
> $$\omega(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \text{sgn}(\sigma) \cdot \omega(\mathbf{v}_{\sigma(1)}, \ldots, \mathbf{v}_{\sigma(n)}). \tag{4.80}$$
>
> *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 4.37**
> A $(0, n)$-tensor ω is an n-form if and only if, $\omega(\mathbf{v}_1, \ldots, \mathbf{v}_n) = 0$ whenever any $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ 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 4.38 (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
> $$\dim \Lambda^n V = \begin{cases} \binom{d}{n} & \text{if } 0 \leq n \leq d \\ 0 & \text{if } n > d \end{cases} \tag{4.81}$$

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 4.39 (Volume Span):**
> *Let V be a d-dimensional vector space with a volume form ω. Given a set of d vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\} \subseteq V$, the* volume span *of these vectors is defined as*
> $$\mathrm{vol}(\mathbf{v}_1, \ldots, \mathbf{v}_d) := \omega(\mathbf{v}_1, \ldots, \mathbf{v}_d). \tag{4.82}$$

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 4.40 (Determinant):**
> *Let V be a d-dimensional vector space. The* determinant *of a linear map $\phi \in \mathrm{End}(V)$ is defined as*
> $$\det(\phi) := \frac{\omega(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\omega(\mathbf{e}_1, \ldots, \mathbf{e}_d)} \tag{4.83}$$
> *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

$$\tilde{\omega} = \lambda \cdot \omega \quad \Rightarrow \quad \det(\phi) = \frac{\tilde{\omega}(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\tilde{\omega}(\mathbf{e}_1, \ldots, \mathbf{e}_d)} = \frac{\lambda \cdot \omega(\phi(\mathbf{e}_1), \ldots, \phi(\mathbf{e}_d))}{\lambda \cdot \omega(\mathbf{e}_1, \ldots, \mathbf{e}_d)}. \tag{4.84}$$

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

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

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

call $\text{sgn}(\sigma_{i_1 \ldots i_d}) \cdot A^{i_1}{}_1 \cdots A^{i_d}{}_d = \lambda$, 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

$$\det\left(\tilde{\Phi}\right) = \det\left(A^{-1} \cdot \Phi \cdot A\right) = \det\left(A^{-1}\right) \cdot \det(\Phi) \cdot \det(A) = \det(\Phi). \tag{4.85}$$

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

$$g \to A^\top g A. \tag{4.86}$$

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

$$\det\left(A^\top g A\right) = \det\left(A^\top\right) \cdot \det(g) \cdot \det(A) = \det(A)^2 \cdot \det(g). \tag{4.87}$$

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:

$$X \to \frac{1}{\det(A)^2} X. \tag{4.88}$$

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

$$\det(g) \cdot X \to \det(A)^2 \cdot \det(g) \cdot \frac{1}{\det(A)^2} X = \det(g) \cdot X. \tag{4.89}$$

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.

$$\boxed{\text{L\textsc{ecture} 9}}$$

§4.3 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 following vector space over $\mathbb{R}$:

$$(\mathcal{C}^\infty(M), +, \cdot) \tag{4.90}$$

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

$$(f + g)(p) := f(p) + g(p) \tag{4.91}$$
$$(\lambda \cdot f)(p) := \lambda \cdot f(p). \tag{4.92}$$

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

Definition 4.41 (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

$$X_{\gamma,p} : \mathcal{C}^\infty(M) \to \mathbb{R} \tag{4.93}$$

defined as

$$\mathcal{C}^\infty(M) \ni f \mapsto X_{\gamma,p}(f) := (f \circ \gamma)'(0) \in \mathbb{R}. \tag{4.94}$$

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 4.42
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 $\lambda, \mu \in \mathbb{R}$,

$$X_{\gamma,p}(\lambda f + \mu g) = \lambda X_{\gamma,p}(f) + \mu X_{\gamma,p}(g). \tag{4.95}$$

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

$$X_{\gamma_1,p}(f) = (f \circ \gamma_1)'(0) = (f \circ \gamma_2)'(0) \cdot 2 = 2(f \circ \gamma_2)'(0) = 2X_{\gamma_2,p}(f). \tag{4.96}$$

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

$$\mathcal{T}_p M := \{X_{\gamma,p} \mid \gamma : \mathbb{R} \to M \text{ is a smooth curve with } \gamma(0) = p\}. \tag{4.97}$$

Equipped with following operations:

$$\oplus : \mathcal{T}_p M \times \mathcal{T}_p M \to \mathcal{T}_p M,$$
$$\odot : \mathbb{R} \times \mathcal{T}_p M \to \mathcal{T}_p M,$$

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 4.44

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

$$X_{\gamma_1,p} \oplus X_{\gamma_2,p} \in \mathcal{T}_p M, \tag{4.98}$$
$$\lambda \odot X_{\gamma,p} \in \mathcal{T}_p M. \tag{4.99}$$

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

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

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:

$$\gamma_3(t) := x^{-1} \circ (x \circ \gamma_1(t) + x \circ \gamma_2(t) - x(p)), \quad \forall t \in I. \tag{4.100}$$

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

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

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, $j = 1, \dots, d$:

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

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:

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

combining the term in the square brackets, we get:

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

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

$$X_{\gamma_3,p} = X_{\gamma_1,p} \oplus X_{\gamma_2,p}. \tag{4.101}$$

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

$$\gamma_4(t) := x^{-1} \circ (x \circ \gamma(\lambda t)), \quad \forall t \in I. \tag{4.102}$$

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

$$X_{\gamma_4, p}(f) = (f \circ \gamma_4)'(0)$$
$$= \left(f \circ x^{-1} \circ (x \circ \gamma(\lambda t)) \right)'(0)$$

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, $j = 1, \ldots, d$:

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

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

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

combining the term in the square brackets, we get:

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

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

$$X_{\gamma_4, p} = \lambda \odot X_{\gamma, p}. \tag{4.103}$$

<div align="right">Q.E.D.</div>

> **Remark 4.45 (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 $X_{\gamma_4, p}$.

§4.3.1 Algebras and Derivations

> **Definition 4.46 (Algebra over a Field):**
> Let $(V, +, \cdot)$ be a vector space over a field K equipped with a "product" operation,
>
> $$\bullet : V \times V \to V, \tag{4.104}$$
>
> 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 4.47 (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:

$$\bullet : \mathcal{C}^\infty(M) \times \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M),$$
$$(f, g) \mapsto f \bullet g \tag{4.105}$$

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 4.48:
Let $(V, +, \cdot, \bullet)$ be an algebra over a field K. The algebra is called:

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

$$(u \bullet v) \bullet w = u \bullet (v \bullet w). \tag{4.106}$$

- Commutative *if for all $u, v \in V$,*
$$u \bullet v = v \bullet u. \tag{4.107}$$

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

$$\mathbf{1} \bullet v = v \bullet \mathbf{1} = v, \quad \forall v \in V. \tag{4.108}$$

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

Definition 4.49 (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 $u, v \in V$,*
$$[u, v] = -[v, u]. \tag{4.109}$$

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

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = 0. \tag{4.110}$$

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 4.50 (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:*

$$D(u \bullet v) = D(u) \bullet v + u \bullet D(v), \quad \forall u, v \in V. \tag{4.111}$$

Example 4.51 (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 $f, g \in \mathcal{C}^\infty(M)$,

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

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

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

Now define a map

$$\begin{aligned}
D : \mathcal{C}^\infty(M) &\to \mathcal{C}^\infty(M), \\
f &\mapsto D(f)
\end{aligned} \tag{4.112}$$

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 4.52

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:

$$\begin{aligned}
[\star, \star] : A \times A &\to A, \\
(\phi, \psi) &\mapsto [\phi, \psi] := \phi \circ \psi - \psi \circ \phi,
\end{aligned} \tag{4.113}$$

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

$$[\phi, [\psi, \rho]] + [\psi, [\rho, \phi]] + [\rho, [\phi, \psi]] = 0, \tag{4.114}$$

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

$$\begin{aligned}
D_H : A &\to A, \\
\phi &\mapsto D_H(\phi) := [H, \phi].
\end{aligned} \tag{4.115}$$

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 4.53 (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 4.54 (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.

§4.3.2 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 4.55 (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.
>
> $$\dim(\mathcal{T}_pM) = \dim(M) = d. \tag{4.116}$$

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

$$\gamma_j(0) = p; \qquad\qquad x^i \circ \gamma_j(t) = \delta^i_j t, \quad \forall t \in \mathbb{R}, \tag{4.117}$$

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

$$\mathbf{e}_j := X_{\gamma_j, p}, \quad j = 1, \ldots, d. \tag{4.118}$$

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

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

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

$$= \left(f \circ x^{-1} \circ (x \circ \gamma_j) \right)'(0)$$
$$= \left[\partial_i (f \circ x^{-1})(x(p)) \right] \cdot \left(x^i \circ \gamma_j \right)'(0)$$
$$= \left[\partial_i (f \circ x^{-1})(x(p)) \right] \cdot \delta^i_j$$
$$= \partial_j (f \circ x^{-1})(x(p)).$$

Define a formal symbol as

$$\left(\frac{\partial}{\partial x^j} \right)_p (f) := \partial_j (f \circ x^{-1})(x(p)), \quad \forall f \in \mathcal{C}^\infty(M). \tag{4.119}$$

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

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

So we have

$$\mathbf{e}_j = \left(\frac{\partial}{\partial x^j}\right)_p, \quad j = 1, \ldots, d. \tag{4.120}$$

Define the set

$$\mathcal{B} := \{\mathbf{e}_1, \ldots, \mathbf{e}_d\} = \left\{\left(\frac{\partial}{\partial x^1}\right)_p, \ldots, \left(\frac{\partial}{\partial x^d}\right)_p\right\}. \tag{4.121}$$

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 $X^1, \ldots, X^d$ such that

$$X = X^j \mathbf{e}_j = X^j \left(\frac{\partial}{\partial x^j}\right)_p. \qquad \text{(sum over } j \text{ from 1 to } d \text{ is implied)} \tag{4.122}$$

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:

$$= \left(f \circ x^{-1} \circ (x \circ \gamma)\right)'(0)$$

$$= \left[\partial_i(f \circ x^{-1})(x(p))\right] \cdot \left(x^i \circ \gamma\right)'(0)$$

$$= (x^i \circ \gamma)'(0) \cdot \left(\frac{\partial}{\partial x^i}\right)_p (f).$$

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

$$X^i := (x^i \circ \gamma)'(0) \in \mathbb{R}, \quad i = 1, \ldots, d. \tag{4.123}$$

So we have

$$X(f) = X^i \left(\frac{\partial}{\partial x^i}\right)_p (f) = X^i \mathbf{e}_i(f), \quad \forall f \in \mathcal{C}^\infty(M). \tag{4.124}$$

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

$$X = X^i \mathbf{e}_i. \tag{4.125}$$

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

> **Remark 4.56 (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 $X^1, \ldots, X^d$, not all zero, such that

$$X^j \mathbf{e}_j = 0. \tag{4.126}$$

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

$$0 = X^j \mathbf{e}_j(x^i) = X^j \left(\frac{\partial}{\partial x^j}\right)_p (x^i) = X^j \partial_j(x^i \circ x^{-1})(x(p))$$

$$= X^j \partial_j(\mathrm{proj}^i)(x(p)) = X^j \delta^i_j = X^i.$$

Since this is true for all $i = 1, \ldots, d$, we have $X^1 = X^2 = \ldots = X^d = 0$, 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

$$X = X^j \left(\frac{\partial}{\partial x^j}\right)_p, \tag{4.127}$$

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).

§4.3.3 Change of Basis and Coordinate Transformation

In the construction of the basis $\mathcal{B}$ of $\mathcal{T}_pM$, 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 $\mathcal{T}_pM$ as

$$\tilde{\mathcal{B}} := \{\tilde{\mathbf{e}}_1, \ldots, \tilde{\mathbf{e}}_d\} = \left\{ \left(\frac{\partial}{\partial y^1} \right)_p, \ldots, \left(\frac{\partial}{\partial y^d} \right)_p \right\}. \tag{4.128}$$

For simiplicity, 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 $\mathcal{T}_pM$, so there must exist a change of basis map between them $A, B \in \mathrm{End}(\mathcal{T}_pM)$ such that

$$\tilde{\mathbf{e}}_j = A^i{}_j \mathbf{e}_i, \qquad\qquad\qquad \mathbf{e}_j = B^i{}_j \tilde{\mathbf{e}}_i, \tag{4.129}$$

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

$$A^i{}_j B^k{}_i = \delta^k{}_j = B^i{}_j A^k{}_i. \tag{4.130}$$

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

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

So we have

$$A^k{}_j = \partial_j(x^k \circ y^{-1})(y(p)). \tag{4.131}$$

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

$$B^k{}_j = \partial_j(y^k \circ x^{-1})(x(p)). \tag{4.132}$$

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

$$X = X^j \mathbf{e}_j = \tilde{X}^j \tilde{\mathbf{e}}_j, \tag{4.133}$$

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 (4.129), 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

$$\tilde{X}^i = X^j B^i{}_j = X^j \partial_j(y^i \circ x^{-1})(x(p)). \tag{4.134}$$

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

$$y^i \circ x^{-1} : x(V) \subseteq \mathbb{R}^d \to \mathbb{R}. \tag{4.135}$$

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

$$y^i = y^i(x^1, \ldots, x^d). \tag{4.136}$$

With this new interpretation, we have

$$\partial_j(y^i \circ x^{-1})(x(p)) = \frac{\partial y^i}{\partial x^j}(x(p)). \tag{4.137}$$

Similarly, we can also write

$$\partial_j(x^i \circ y^{-1})(y(p)) = \frac{\partial x^i}{\partial y^j}(y(p)). \tag{4.138}$$

So the change of basis relation (4.129) can be rewritten as

$$\tilde{\mathbf{e}}_j = \frac{\partial y^i}{\partial x^j}(x(p))\mathbf{e}_i, \qquad\qquad \mathbf{e}_j = \frac{\partial x^i}{\partial y^j}(y(p))\tilde{\mathbf{e}}_i, \tag{4.139}$$

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

$$\tilde{X}^i = X^j \frac{\partial y^i}{\partial x^j}(x(p)). \tag{4.140}$$

Remark 4.57 (Jacobian Matrix). Let $f : \mathbb{R}^m \to \mathbb{R}^n$ 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

$$\frac{\partial f^i}{\partial x^j}(a), \tag{4.141}$$

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 (4.140)) is being used to define a 'vector'. It goes like this: an array of d real numbers $(X^1, \ldots, X^d)$ is called a vector if under a change of coordinates $x \to y = y(x)$, the components transform as

$$\tilde{X}^i = X^j \frac{\partial y^i}{\partial x^j}. \tag{4.142}$$

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

§4.3.3.1 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 ―――――――――――――――――――――

§4.4 Cotangent Spaces and Gradient

Definition 4.58 (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 $\mathcal{T}_p M$:*

$$\mathcal{T}_p^* M := (\mathcal{T}_p M)^*. \tag{4.143}$$

Since the manifold is finite-dimensional[2], the tangent space $\mathcal{T}_pM$ 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

$$\mathcal{T}_pM \cong_{\text{vec}} \mathcal{T}_p^*M. \tag{4.144}$$

However, this isomorphism is not canonical, *i.e.*, there is no natural way to identify vectors in $\mathcal{T}_pM$ 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 $p \in M$,

$$\mathsf{T}_s^r(\mathcal{T}_pM) := \left\{ t : \underbrace{\mathcal{T}_p^*M \times \cdots \times \mathcal{T}_p^*M}_{r \text{ times}} \times \underbrace{\mathcal{T}_pM \times \cdots \times \mathcal{T}_pM}_{s \text{ times}} \to \mathbb{R} \;\middle|\; t \text{ is multilinear} \right\}. \tag{4.145}$$

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

Definition 4.59 (Gradient):
Let $f \in \mathcal{C}^\infty(M)$. Then at each point $p \in M$, we have a linear map

$$\begin{aligned} \mathrm{d}_p : \mathcal{C}^\infty(M) &\to \mathcal{T}_p^*M \\ f &\mapsto \mathrm{d}_pf, \end{aligned} \tag{4.146}$$

where d_pf is defined by

$$\mathrm{d}_pf(X) := X(f), \quad \forall X \in \mathcal{T}_pM. \tag{4.147}$$

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

Observe that the gradient d_pf 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 $X \in \mathcal{T}_pM$ be tangent to the level set[3] of f at the point p, then we have

$$\mathrm{d}_pf(X) = X(f) = 0,$$

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

§4.4.1 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 4.60 (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

$$\left\{ \mathrm{d}_px^1, \mathrm{d}_px^2, \ldots, \mathrm{d}_px^{\dim M} \right\} \tag{4.148}$$

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

---

[2]The 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.

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

sufficient to show that the set $\{d_p x^i\}_{i=1}^{\dim M}$ 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

$$\omega_i d_p x^i = 0. \qquad (\text{sum over } i \text{ from 1 to } \dim M \text{ is implied}).$$

So for any tangent vector $X \in \mathcal{T}_p M$, we have $\omega_i d_p x^i(X) = 0$. 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

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

Here, $\text{proj}^i : \mathbb{R}^{\dim M} \to \mathbb{R}$ is the projection map onto the i-th coordinate. Since j was arbitrary, we have $\omega_j = 0$ for all $j = 1, 2, \ldots, \dim M$, which contradicts our assumption that not all ω_i are zero. Thus, the set $\{d_p x^i\}_{i=1}^{\dim M}$ 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):

$$d_p x^i\left(\left(\frac{\partial}{\partial x^j}\right)_p\right) = \delta_j^i. \tag{4.149}$$

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

§4.5 Push-Forward and Pull-Back