Mathematics Notes

Kai YU

Contents

Chapter 1

Law and Order

$\exists x \in A,$ not $\mathbb{P}(x)$ and $\forall x \in A, \mathbb{P}(x)$ have the opposite truth value.
Let I be a set and $(A_i)_{i\in I}$ be a family of sets.

- $x \in \bigcup_{i\in I} A_i$ iff(if and only if) $\exists i \in I, x \in A_i$, the set $\bigcup_{i\in I} A_i$ is called the *union* of $(A_i)_{i\in I}$.
- $x \in \bigcap_{i\in I} A_i$ iff $\forall i \in I, x \in A_i$, the set $\bigcap_{i\in I} A_i$ is called the *intersection* of $(A_i)_{i\in I}$. $\bigcap_{i\in\emptyset} A_i$ is not defined.

Let A be a set, we denote by $\mathscr{P}(A)$ the set of all subsets of A, called the *power set* of A.
Equivalently, it is the set TV^S of all functions from S to the set TV of truth values. This is often written 2^S, since there are (at least in classical logic) exactly 2 truth values.

1.1 Sets and Correspondences

1.1.1 Def

f=$(\mathscr{D}_f, \mathscr{A}_f, \Gamma f)$ is a correspondence, where $\mathscr{D}_f$ is the *departure set* and $\mathscr{A}_f$ is the *arrival set*, and Γ_f is the subset of $\{(x,y) \mid x \in \mathscr{D}_f, y \in \mathscr{A}_f\}$, called the *graph* of f. f^{-1} called the *inverse correspondence*, $f^{-1} = (\mathscr{A}_f, \mathscr{D}_f, \Gamma_f^{-1})$, $\Gamma_{f^{-1}} = \{(y,x) \mid (x,y) \in \Gamma_f\}$.

1.1.1.1 Remark

$(f^{-1})^{-1} = f$.

1.1.1.2 Notation

We call $\Delta x := \{(x,x) \mid x \in X\}$ the *diagonal subset*, $Id_x := (X, X, \Delta x)$ is the *identity correspondence* of X, $(X, X, \emptyset)$ is the *empty correspondence*.

1.1.2 Image and Inverse Image

1.1.2.1 Def

We denote f(A)=$\{y \in \mathscr{A}_f \mid \exists x \in A, (x,y) \in \Gamma_f\}$ the image of A by f. We denote by Im(f) the set $f(\mathcal{D}_f)$, called the *range* of f. We denote by Dom(f) the set $f^{-1}(\mathscr{A}_f)$, called the *domain* of f. The *inverse image* of f means the image of f^{-1}.

1.1.2.2 Prop

$f(\bigcup_{i\in I} A_i) = \bigcup_{i\in I} f(A_i)$. If $I \neq \emptyset$, $f(\bigcap_{i\in I} A_i) \subseteq \bigcap_{i\in I} f(A_i)$.

1.1.3 Composition

$g \circ f = (\mathcal{D}_f, \mathscr{A}_g, \Gamma_{g\circ f})$, $\Gamma_{g\circ f} = \{(x,z) \mid \exists y,\ (x,y) \in \Gamma_f, (y,z) \in \Gamma_g\}$.
The correspondence $g \circ f$ is called the composite of f and g.

1.1.3.1 Prop

$(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, $h \circ (g \circ f) = (h \circ g) \circ f$.

1.1.4 Defs in the Correspondence

1.1.4.1 Surjective

$\mathscr{A}_f$=Im(f).

1.1.4.2 Multivalued Mapping

$\mathscr{D}_f$=Dom(f).

1.1.4.3 Function

$\forall x \in \mathscr{D}_f$ has at most an image of x by f.
Remark We denote by f(x) the unique image of x by f. We can also use $x \mapsto f(x)$ to denote.

1.1.4.4 Injective

$\forall x \in \mathscr{A}_f$ has at most an image of x by f^{-1}.

1.1.4.5 Mapping

If f is a function and a multivalued mapping, then we say that f is a mapping. We use $f : X \to Y$ or $X \xrightarrow{f} Y$ to denote "f is a mapping from X to Y".

f is a mapping + f is injective/surjective/injective and surjective$\to$ f is a/an injective/surjective/bijective mapping, also called f is a/an injection/surjection/bijection(one-to-one correspondence).

1.1.4.6 The Product

Let I be a set, $(A_i)_{i\in I}$ be a family of sets, and A=$\bigcup_{i\in I} A_i$. We denote by $\Pi_{i\in I}A_i$ the set of all mappings $\mu : I \to A$ s.t. $\mu(i) \in A_i$. $\forall i \in I$, if $\mu \in \Pi_{i\in I}A_i$, we often write μ in the form of a family $(\mu(i))_{i\in I}$. $\Pi_{i\in I}A_i$ is called the product of $(A_i)_{i\in I}$.

In particular, x^n denotes $x^{\{1,\ldots,n\}}$. $X^{\mathbb{N}}$ is called the *sequence* in X parametrized by $\mathbb{N}$.

1.1.4.7 Restriction and Extension

Let f and g be correspondences. If $\Gamma_f \subseteq \Gamma_g$, then we say that f is a restriction of g, g is an extension of f.

1.1.4.8 Prop

- If $g_1 \circ f = Id_X$, $f \circ g_2 = Id_Y$, then f is a bijection, $g_1 = g_2 = f^{-1}$.
- If $f \circ g$ and $g \circ f$ are bijections, then f and g are bijections.

1.2 Binary Relation

A binary relation on X, we refer to a correspondence from X to X.

1.2.1 Remark

xRy, xЯy denotes (x,y)$\in \Gamma_R$, x$\not R$y denotes not xRy.

1.2.2 Concept

- If $\forall x \in X$, xRx holds, then R is *reflexive.*
- If $\forall x \in x$, x$\not R$x holds, then R is *irreflexive.*
- If $\forall (x, y) \in X \times X$, xRy implies yRx, then R is *symmetric.*
- If $\forall (x, y) \in X \times X$, xRy and yRx implies x=y, then R is *antisymmetric.*
- If $\forall (x, y) \in X \times X$, xRy and yRx can't hold at the same time, then R is *asymmetric.*
- If $\forall (x, y, z) \in X \times X \times X$, xRy and yRz implies xRz, then R is *transitive.*
- If R is transitive, reflexive and antisymmetric, then it's an *order relation.*
- If R is transitive and asymmetric, then it's a *strict order relation.*
- If R is reflexive, symmetric and transitive, then we say that R is an *equivalence relation.*

In general, we use an underlined notation to denote an order relation.

1.2.3 Partially Ordered Sets

Let X be a set and $\underline{R}$ be an ordered relation on X, the pair (X,$\underline{R}$) is called a *partially ordered set.*

In addition, if $\forall x, y \in X$, either x$\underline{R}$y and y$\underline{R}$x, then we say that $\underline{R}$ is a *total order* and (X,$\underline{R}$) is a *total ordered set.*

1.2.4 Upperbounds and Lowerbounds

Let (X,$\underline{R}$) be a partially ordered set, $A \subseteq X$.

Let $M \in X$. If $\forall a \in A, a\underline{R}M$, we say that M is an upperbound of A, relatively to $\underline{R}$. If $\forall a \in A, a\underline{\text{Я}}M$, we say that M is a lowerbound of A, relative to $\underline{R}$.

1.2.5 Well-ordered Set

If any non-empty subset of X has a least element, we say that (X,$\underline{R}$) is a *well-ordered set.*

1.2.6 Induction Principle

Let (X,$\leq$) be a well-ordered set. Let $\mathbb{P}(.)$ be a statement depending on a parameter in X. Assume that $\forall x \in X$, $\forall y \in X_{<x}$, $\mathbb{P}(y)$ implies $\mathbb{P}(x)$, then $\forall x \in X, \mathbb{P}(x)$ holds.

1.3 Composition Laws

We mean a mapping from $X \times X \to X$.

1.3.1 Closed

Y is a subset of X. Y is closed under * means that $\forall a, b \in Y$, $a * b \in Y$.

1.3.2 The Restriction of *

Y is a subset of X. If Y is closed under *, which can also be presented as $\forall x, y \in Y, x * y \in Y$. This composition law is called the restriction of * on Y.

1.3.3 Semigroup, Monoid and Group

- If $\forall (x, y, z) \in X \times X \times X, (x * y) * z = x * (y * z)$, we say that $*$ is *associative* and that $(X, *)$ is a *semigroup*.

  Prop If * is associative, then $x_1 * x_2 * \cdots * x_n \cdots = x_1 * x_2 * x_3 * \cdots * (x_{n-1} * x_n) \cdots$.

- If $\forall (x, y) \in X^2, x * y = y * x$, then we say that $*$ is *commutative*.

- Let $(M, *)$ be a semigroup. If $\exists e \in M, \forall x \in M, e * x = x * e = x$, we say that e is a neutral element of $(M, *)$, and we say that $(M, *)$ is a *monoid*.

- If x*y=e, then we say that x is right invertible, y is a right inverse of x; y in left invertible, x is a left inverse of y.

1.3.3.1 Group We denote by $M^\times$ the set of all invertible elements of M, If $M^\times = M$, then we say that M is a group.

1.3.3.2 Abelian Group A commutative group.

1.3.3.3 Submonoids and Subgroups If M, N are monoids/groups and N is a subset of M which is closed and exists a neutral element, then N is a submonoid/subgroup.

1.3.3.4 Trivial Subgroup A subgroup of G which is G or $\{1_G\}$.

1.3.3.5 Prop $\{\mathbb{Z}/p\mathbb{Z}\}$ doesn't have a non-trivial group when p is a prime.

1.3.3.6 Finite Group A group which has finite elements.

1.3.3.7 Lagrange's Theorem Let G be a finite group, and let H be a subgroup of G. Then the order of H (that is, the number of elements in H) divides the order of G.

1.3.3.8 Prop

- In a group, x has a unique left/right inverse, which are equal.

  Proof: Let G be a group and $x \in G$. Suppose y is a left inverse of x, meaning $yx = e$, and z is a right inverse of x, meaning $xz = e$, where e is the identity element of G.

  We aim to show that $y = z$. Starting with $yx = e$, multiplying both sides on the right by z gives $(yx)z = ez = z$. Since $yx = e$, this simplifies to $z = z$. Now, consider $xz = e$, and multiply both sides on the left by y, giving $y(xz) = ye = y$. We conclude $y = z$.

  Thus, the left and right inverses of x are equal. Furthermore, if x had another distinct left or right inverse, it would contradict this result, so the inverse is unique.

  So we denote by x^{-1} its left and right inverse. If the composition law is denoted as +, the inverse of an invertible element x is usually denoted as -x, y+(-x) is denoted as y-x.

- $x \in M^\times \to x^{-1} \in M^\times$, $(x^{-1})^{-1} = x$;

- $(x, y) \in M^\times \cdot M^\times \to xy \in M^\times$, $(xy)^{-1} = y^{-1}x^{-1}$.

  Proof: $(y^{-1}x^{-1})(xy) = e$.

1.4 Ring Theory

1.4.1 Rings

1.4.1.1 Def If

- $(A,+)$ forms a commutative group,
- $(A,\cdot)$ forms a monoid,
- For all $(a,b,c)\in A^3$, $(a+b)c=ac+bc,\ c(a+b)=ca+cb$,

then we say that $(A,+,\cdot)$ is a unitary ring.

By convention, the neutral element of $(A,+)/(A,\cdot)$ is denoted as 0/1, called the *zero element/unity* of A.

If $A^\times = A\backslash\{0\}$, then A is a *division ring* or *skewfield*. If A is also commutative, then A is a *field.*

1.4.1.2 Prop If A is a field, $f: A\to B$ bijection s.t. $\forall a,b\in A,\ f(a+b)=f(a)+f(b),\ f(ab)=f(a)f(b)$, then B is a field.

1.4.1.3 Proof Firstly, we want to show that $(B,+)$ is a commutative group: $\forall a,b\in B$, we can find its inverse image in A, called a_0,b_0. $a+b=f(a_0+b_0)$, so it's closed. Since $\forall a,b,c\in B$, we can find their inverse image $a_0,b_0,c_0,\ a+(b+c)=f(a_0+(b_0+c_0))=f((a_0+b_0)+c_0)=(a+b)+c,\ a+b=f(a_0+b_0)=f(b_0+a_0)=b+a,\ \exists e=f(0)\in B,\ e+a=f(0+a_0)=f(a_0+0)=a+e=a$, so $(B,+)$ is a commutative monoid. For the same reason, $(B,\times)$ is also a commutative monoid. Since $\forall a\in B$, we can find its inverse $f(-a_0),\ a+f(-a_0)=f(-a_0+a_0)=f(0)=e$, so $(B,+)$ is an Abelian group.

Secondly, $\forall a,b,c\in B,\ a(b+c)=f(a_0(b_0+c_0))=f(a_0b_0+a_0c_0)=ab+ac$, so it's a ring. Since $B^\times=B\backslash\{f(0)\}$ (which is because $\forall a,b\in B\backslash\{f(0)\},\ f(a_0b_0)$ always exists), it's a field.

1.4.1.4 Prop If A is a unitary ring, then for all $a\in A$, $a\cdot 0=0\cdot a=0$; for all $(a,b)\in A$, $(-a)\cdot b=a\cdot(-b)=-a\cdot b$.

1.4.2 Morphisms and Isomorphisms

1.4.2.1 Def Let (M,*) and (N,+) be monoids, we call morphism of monoids from (M,*) to (N,+) any mapping $f: M\to N$ that satisfies:

- $f(e_M)=e_N$.
- For all $(x,y)\in M^2$, $f(x*y)=f(x)+f(y)$.

If $(M,*)$ and $(N,*)$ are groups, it's also a morphism of groups. If f is bijection, then it's an isomorphism.

1.4.2.2 Prop

- If $M_1\subseteq M$ is a submonoid, then $f(M_1)\subseteq N$ is also a submonoid.
- If $a\in M^\times$, then $f(a)\in N^\times$, $f(a)^{-1}=f(a^{-1})$.

1.4.2.3 Prop Let M be a monoid. $\forall x\in M^\times$, $\exists$ a unique morphism of monoids from $(\mathbb{Z},+)$ to (M,*) that sends $1\in\mathbb{Z}$ to x.

1.4.2.4 Morphism of unitary rings Let A and B be unitary rings. We call morphism of unitary rings from A to B only when mapping $A\to B$ is a morphism of group from (A,+) to (B,+), and a morphism of monoid from $(A,\cdot)\ to\ (B,\cdot)$.

1.4.2.5 Prop Let R be a unitary ring. There is a unique morphism from $\mathbb{Z}\ to\ R.$

1.4.3 K-Algebra

K is a commutative ring. We call the K-algebra any pair(R,f), when R is a unitary ring, and $f: K\to R$ is a morphism of unitary rings s.t. $\forall(b,x)\in K\times R, f(b)x=xf(b)$.

1.4.3.1 Example For any unitary ring R, the unique morphism of unitary rings $\mathbb{Z}\to R$ define a structure of $\mathbb{Z}$-algebra on R (extra: $\mathbb{Z}$ is commutative despite R isn't guaranteed)

1.4.3.2 Notation Let K be a commutative unitary ring, (A,f) be a K-algebra. If there is no ambiguity on f, for any $(\lambda,a)\in k\times A$, we denote $f(\lambda)a$ as λa.

1.5 Formal Power Series

1.5.1 Def

Let K be a commutative unitary ring, T be a formal symbol. We denote $K^{\mathbb{N}}$ as $K[\![T]\!]$. If $(a_n)_{n\in\mathbb{N}}$ is an element of $K^{\mathbb{N}}$, when we denote $K^{\mathbb{N}}$ as K[\![T]\!], this element is denoted as $\sum\limits_{n\in\mathbb{N}} a_n T^n$. Such element is called a formal power series over K and a_n is called the coefficient of T^n of this formal power series.

1.5.1.1 Remark $n \in \mathbb{N}$ is possible infinite, so $\sum\limits_{n\in\mathbb{N}} a_n$ couldn't be executed directly.

1.5.1.2 Notation

- omit terms with coefficient 0.
- write T^1 as T.
- omit coefficient those are 1.
- omit T^0.

1.5.1.3 Example $1T^0 + 2T^1 + 1T^2 + 0T^3 + ... + 0T^n + ...$ is written as $1 + 2T + T^2$.

1.5.2 Operations

Remind that K[\![T]\!]=$\{\sum\limits_{n\in\mathbb{N}} a_n T^n \mid (a_n)_{n\in\mathbb{N}} \in k^{\mathbb{N}}\}$, define two composition laws on K[\![T]\!]:
$\forall F(T) = a_0 + a_1 T + ...,\ G(T) = b_0 + b_1 T + ...$, let $F + G = (a_0 + b_0) + (a_1 + b_1)T + ...,\ FG = \sum\limits_{n\in\mathbb{N}} \sum\limits_{i+j=n} (a_i b_j) T^n$.

1.5.2.1 Prop

- $(FG)H = \left(\sum\limits_{n\in\mathbb{N}} \sum\limits_{i+j=n} (a_i b_j) T^n \right) (\sum\limits_{n\in\mathbb{N}} c_n T^n) = \sum\limits_{n\in\mathbb{N}} \left(\sum\limits_{p,q,l=n} a_p b_q c_l \right) T^n = F(GH)$.
- $0T^0 + \cdots + 0T^n + \ldots$ is the neutral element of (K[\![T]\!],+).
- $1T^0 + \cdots + 0T^n + \ldots$ is the neutral element of (K[\![T]\!],×).
- $(K[\![T]\!], +, \cdot)$ forms a commutative unitary ring.
- $K \to K[\![T]\!] \quad \lambda \mapsto \lambda T$ is a morphism.

1.5.3 Derivative

Let $F(T) \in K[\![T]\!]$. We denote by $F'(T)$ *or* $\mathcal{D}(F(T))$ the derivative of formal power series $\sum\limits_{n\in\mathbb{N}} (n+1) a_{n+1} T^n$.

1.5.3.1 Prop

- $\mathcal{D}(K[\![T]\!], +) \to (K[\![T]\!], +)$ is a morphism of groups.
- **Leibniz Rule** $\mathcal{D}(FG) = F'G + FG'$.

1.5.3.2 Prop a_0 is invertible in k, iff F(T) is invertible in K[\![T]\!].

1.5.3.3 exp We denote $\exp(T) \in K[\![T]\!]$ as $\sum\limits_{n\in\mathbb{N}} \frac{1}{n!} T^n$, called the *experimental series*, which fulfills the differential equation $\mathcal{D}(\exp(T)) = \exp(T)$.

1.5.3.4 Sin and Cos We define the solution of $z^2 = -1$ is $i \in K$. We define $\cos(T) := \frac{\exp(iT)+\exp(-iT)}{2}$, $\sin(T) := \frac{\exp(iT)-\exp(-iT)}{2i}$.

1.5.3.5 Ord $\operatorname{ord}(A)$ means the infimum of $i \in I$ s.t. $A_i \neq 0$. If the infimum not exist, then $\operatorname{ord}(A) = +\infty$.

1.5.3.6 ln We define another form of power series $\ln(1+T)$: $\ln(1+T) = \sum\limits_{n\in\mathbb{N}_{\geq 1}} \frac{(-1)^{n-1}}{n} T^n \quad (\operatorname{ord}(\ln(1+T)) = 1)$

1.6 Topology of Formal Power Series

1.6.1 Cauchy Sequence

1.6.1.1 Idea A Cauchy sequence is an infinite sequence, which ought to converge in the sense that successive terms get arbitrarily close together, as they would if they were getting arbitrarily close to a limit. Among sequences, only Cauchy sequences will converge; in a complete space, all Cauchy sequence converge.

$(F_i(T))_{i\in\mathbb{N}}$ be a sequence of elements in K⟦T⟧, and $F(T) \in K[\![T]\!]$. We say that $(F_i(T))_{i\in\mathbb{N}}$ is a Cauchy sequence if $\forall l \in \mathbb{N}$, there exists $N(l) \in \mathbb{N}$ s.t. $\forall (i,j) \in \mathbb{N}^2_{\geq N(l)}$, $\operatorname{ord}(F_i(T) - F_j(T)) \geq l$.

We will discuss it later.

1.7 Universal Property of the Formal Power Series

1.7.1 K-Algebra Morphism

Let A and B be two K-algebras. We say that ϕ is a K-algebra morphism from A to B if $\phi : A \to B$ satisfies $\forall \lambda \in K$, $\forall a \in A$, $\phi(\lambda a) = \lambda\phi(\alpha)$.

For $f : K \to A$ and $g : K \to B$, if ϕ is a ring morphism from A to B, then it is a K-algebra morphism iff $\phi \circ f = g$.

1.7.1.1 Theorem Let A be a K-algebra, and let $x \in A$. Then there exists a unique K-algebra morphism:

$$K[T] = \{\Sigma_{n\in\mathbb{N}} a_n T^n \mid \exists d \in \mathbb{N}, \forall n > d,\ a_n = 0\} \to A \quad \Sigma_{i=0}^{d} a_i T^i \mapsto \Sigma_{i=0}^{d} a_i x^i$$

1.7.1.2 Proof To prove that the morphism exists, we construct $f : K[\![T]\!] \to A \quad \Sigma_{i=0}^{d} a_i T^i \mapsto \Sigma_{i=0}^{d} a_i x^i$. Easy to prove that it's a K-algebra morphism.

To prove the morphism is unique: Obviously.

1.7.2 Polynomials

And the formal power series that belongs to $K[T]$ are called *polynomials*.

Monic polynomial refers to a polynomial whose leading coefficient is 1.

1.7.2.1 Def Let $F(T) \in K[T]$, $F(T) = \sum_{n\in\mathbb{N}} a_n T^n$. If $\{n \in \mathbb{N} \mid a_n \neq 0\} \neq \emptyset$, its greatest element is denoted as $\deg(F(T))$, called degree of $F(T)$.

1.7.3 Ord and Deg Explanation

For any $f \in K[T] \setminus \{0\}$, it can be written as $f(T) = C \cdot P_1(T)^{d_1} \cdot P_2(T)^{d_2} \cdots P_n(T)^{d_n}$, where $p = P_i \in K[T]$ are monic and irreducible, and $d_i =: \operatorname{ord}_p(f) \in \mathbb{Z}\backslash\{0\}$. Let $|f|_p := e^{-\operatorname{ord}_p(f)}$.

Properties of ord_p:

1. For $f, g \in K[T]$, $\operatorname{ord}_p(fg) = \operatorname{ord}_p(f) + \operatorname{ord}_p(g)$; $|fg|_p = |f|_p \times |g|_p$.
2. For $f, g \in K[T]$, $\operatorname{ord}_p(f+g) \geq \min\{\operatorname{ord}_p(f), \operatorname{ord}_p(g)\}$; $|f+g|_p \leq \max(|f|_p, |g|_p) \leq |f|_p + |g|_p$.
3. $|f|_p = 0 \Leftrightarrow \operatorname{ord}_p(f) = +\infty$, f=0 by convention.
4. If $f \neq 0$, then $\operatorname{ord}_p(f) > 0$.

The degree of a polynomial $f \in K[T]$, denoted as $\deg(f)$, is defined as the highest power of T with a non-zero coefficient. $|f|_\infty := e^{\deg(f)}$

Properties of deg:

1. For $f, g \in K[T]$, $\deg(fg) = \deg(f) + \deg(g)$; $|fg|_\infty = |f|_\infty \times |g|_\infty$.
2. For $f, g \in K[T]$, $\deg(f+g) \leq \max\{\deg(f), \deg(g)\}$; $|f+g|_\infty \leq \max\{|f|_\infty, |g|_\infty\}$.

The ord we used before is ord_T here.

Chapter 2

Sequences

2.1 Supremum and Infimum

2.1.1 Def

Let $(X, \leq)$ be a partially ordered set, and let A and Y be subsets of X, s.t. $A \subseteq Y$.

- If the set $\{y \in Y \mid \forall a \in A, a \leq y\}$ has a least element, then we say that A has a supremum in Y with respect to $\leq$, denoted by $\sup_{(Y,\leq)} A$ this least element, and we call it the supremum of A in Y with respect to $\leq$.
- If the set $\{y \in Y \mid \forall a \in A, y \leq a\}$ has a greatest element, then we say that A has a infimum in Y with respect to $\leq$, denoted by $\inf_{(Y,\leq)} A$ this greatest element, and we call it the infimum of A in Y with respect to $\leq$.
- **Observation** $\inf_{(Y,\leq)} A = \sup_{(Y,\geq)} A$.

2.1.1.1 Notation Let $(X, \leq)$ be a partially ordered set, and let I be a set.

- If f is a function from I to X, sup f denotes the supremum of f(I) in X. inf f takes the same.
- If $(x_i)_{i\in I}$ is a family of elements in X, then $\sup\limits_{i\in I} x_i$ denotes $\sup\{x_i \mid i \in I\}$ in X. $\inf\limits_{i\in I} x_i$ takes the same.

If moreover, $\mathbb{P}(\cdot)$ denotes a statement depending on a parameter in I, then $\sup\limits_{i\in I, \mathbb{P}(i)} x_i$ denotes $\sup\{x_i \mid i \in I, \mathbb{P}(i) \text{ holds}\}$.

2.1.1.2 Example Let $A = x \in R \mid 0 \leq x < 1 \subseteq \mathbb{R}$. We equip $\mathbb{R}$ with the usual order relation: $\{y \in \mathbb{R} \mid \forall x \in A, x \leq y\} = \{y \in \mathbb{R} \mid y \geq 1\}$, so $\sup A = 1$. $\{y \in \mathbb{R} \mid \forall x \in A, y \leq x\} = \{y \in \mathbb{R} \mid y \geq 0\}$, so $\inf A = 0$.

2.1.1.3 Example For $n \in \mathbb{N}$, let $x_n = (-1)^n \in R$, $\sup\limits_{n\in\mathbb{N}} \inf\limits_{k\in\mathbb{N}, k\geq n} x_k = -1$, $\inf\limits_{n\in\mathbb{N}} \sup\limits_{k\in\mathbb{N}, k\geq n} x_k = 1$.

2.1.1.4 Prop Let $(X, \leq)$ be a partially ordered set, A,Y,Z be subset of X, s.t. $A \subseteq Z \subseteq Y$:

- If max A exists, then is is also equal to $\sup_{(Y,\leq)} A$.
- If $\sup_{(Y,\leq)} A$ exists and belongs to Z, then it is equal to $\sup_{(Z,\leq)} A$.

inf takes the same.
Let $(X, \leq)$ be a partially ordered set, and A,B,Y be subsets of X s.t. $A \subseteq B \subseteq Y$:

- If $\sup_{(Y,\leq)} A$ and $\sup_{(Y,\leq)} B$ exist, then $\sup_{(Y,\leq)} A \leq \sup_{(Y,\leq)} B$.
- If $\inf_{(Y,\leq)} A$ and $\inf_{(Y,\leq)} B$ exist, then $\inf_{(Y,\leq)} A \geq \inf_{(Y,\leq)} B$.

Let $(X, \leq)$ be a partially ordered set, I be a set and $f, g : I \to X$ be mappings s.t. $\forall t \in I, f(t) \leq g(t)$:

- If $\inf f$ and $\inf g$ exist, then $\inf f \leq \inf g$.
- If $\sup f$ and $\sup g$ exist, then $\sup f \leq \sup g$.

2.2 Intervals

We fix a totally ordered set $(X, \leq)$.

2.2.1 Def

Let $I \subseteq X$. If $\forall (x, y) \in I \times I$ with $x \leq y$, one has $[x, y] := \{x \in X \mid a \leq x \leq b\} \subseteq I$, then we say that I is an interval in X.

2.2.1.1 Example Let $(a, b) \in X \times X$, s.t. $a \leq b$. Then the following sets are also intervals:

- $]a, b[:= \{x \in X \mid a < x < b\}$
- $[a, b[:= \{x \in X \mid a \leq x < b\}$
- $]a, b] := \{x \in X \mid a < x \leq b\}$

2.2.1.2 Prop Let Λ be a non-empty set and $(I_\lambda)_{\lambda\in\Lambda}$ be a family of intervals in X.

- $\bigcap\limits_{\lambda\in\Lambda} I_\lambda$ is a interval in X.
- If $\bigcap\limits_{\lambda\in\Lambda} I_\lambda \neq \varnothing, \bigcup\limits_{\lambda\in\Lambda} I_\lambda$ is a interval in X.

2.2.1.3 Def Let $(X, \leq)$ be a totally ordered set. I be a non-empty interval of X. If $\sup I$ exists in X, we call $\sup I$ the *right endpoint*.

inf takes the similar way.

2.2.1.4 Prop Let I be an interval in X.

- Suppose that $b = \sup I$ exists. Then $\forall x \in I, [x, b[\subseteq I$.
- Suppose that $a = \inf I$ exists. Then $\forall x \in I,]a, x] \subseteq I$.

2.2.1.5 Prop Let I be an interval in X. Suppose that I has a supremum b and an infimum a in X. Then I is equal to one of the following sets $[a, b],\ [a, b[,\ ]a, b],\ ]a, b[$.

2.2.2 Thick

Let $(X, \leq)$ be a totally ordered set. If $\forall (x, z) \in X \times X$, s.t. $x < z$, $\exists y \in X$ s.t. $x < y < z$, then we say that $(X, \leq)$ is thick.

2.2.2.1 Prop Let $(X, \leq)$ be a thick totally ordered set. $(a, b) \in X \times X,\ a < b$. If I is one of the following intervals $[a, b]; [a, b[;]a, b];]a, b[$. Then $\inf I = a,\ \sup I = b$.

2.2.2.2 Proof Since X is thick, there exists $x_0 \in]a, b[$. By Def, b is an upper bound of I. If b is not the supremum of I, there exists an upper bound M of I s.t. M<b. Since X is thick, there is $M' \in X$ s.t. $x_0 \leq M, M' < b$. Since $[x, b[\subseteq]a, b[\in I$. Hence M and M' belong to I, which conflicts with the uniqueness of supremum.

2.3 Enhanced Real Line

2.3.1 Def

Let $+\infty$ and $-\infty$ be two symbols that are different and don't belong to $\mathbb{R}$ We extend the usual total order $\leq$ on $\mathbb{R}$ to $\mathbb{R} \bigcup \{-\infty, +\infty\}$ s.t. $\forall x \in \mathbb{R}, -\infty < x < +\infty$.

Thus $\mathbb{R} \bigcup \{-\infty, +\infty\}$ become a totally ordered set, and $\mathbb{R} =] - \infty, +\infty[$. Obviously, this is a thick totally ordered set. We define:

- $\forall x \in] - \infty, +\infty] \quad x + (+\infty) := +\infty \quad (+\infty) + x := +\infty$
- $\forall x \in [-\infty, +\infty[\quad x + (-\infty) := -\infty \quad (-\infty) + x := -\infty$
- $\forall x \in]0, +\infty] \quad x(+\infty) := (+\infty)x := +\infty \quad x(-\infty) = (-\infty)x := -\infty$
- $\forall x \in [-\infty, 0[\quad x(+\infty) = (+\infty)x := -\infty \quad x(-\infty) = (-\infty)x := +\infty$
- $-(+\infty) := -\infty \quad -(-\infty) := +\infty \quad (\infty)^{-1} := 0$

$(+\infty) + (-\infty) \quad (-\infty) + (+\infty) \quad (+\infty)0 \quad 0(+\infty) \quad (-\infty)0 \quad 0(-\infty)$ **ARE NOT DEFINED**, they are *indeterminate forms*.

2.3.2 Order Complete

Let $(X, \leq)$ be a partially ordered set. If for any subset A of X, A has a supremum and an infimum in X, then we say that X is order complete.

2.3.2.1 Example Let Ω be a set $(\mathscr{P}(\Omega), \subseteq)$ is order complete. If $\mathcal{F}$ is a subset of $\mathscr{P}(\Omega)$, $\sup \mathcal{F} = \bigcup\limits_{A\in\mathcal{F}} A$. In particular, $\inf \varnothing = \Omega \quad \sup \varnothing = \varnothing$.

2.3.2.2 Axiom $(\mathbb{R}\bigcup\{-\infty,+\infty\},\leq)$ is order complete. In $\mathbb{R}\bigcup\{-\infty,+\infty\}$ $\sup\varnothing=-\infty,\ \inf\varnothing=+\infty$.

2.3.2.3 Notation

- For any $A\subseteq\mathbb{R}\bigcup\{-\infty,+\infty\}$ and $c\in\mathbb{R}$, we denote by $A+c$ the set $\{a+c\mid a\in A\}$.
- If $\lambda\in\mathbb{R}\backslash\{0\},\lambda A$ denotes $\{\lambda a\mid a\in A\}$.
- -A denotes (-1)A.

2.3.2.4 Prop For any $A\subseteq\mathbb{R}\bigcup\{-\infty,+\infty\},\ \sup(-A)=-\inf A,\inf(-A)=-\sup A$.

2.3.2.5 Def We denote by $(\mathbb{R},\leq)$ a field $\mathbb{R}$ equipped with a total order $\leq$, which satisfies the following condition:

- $\forall(a,b)\in\mathbb{R}\times\mathbb{R}$ s.t. $a<b$, one has $\forall c\in\mathbb{R},a+c<b+c$.
- $\forall(a,b)\in\mathbb{R}_{>0}\times\mathbb{R}_{>0},\ ab>0$.
- $\forall A\subseteq\mathbb{R}$, if A has an upper bound in $\mathbb{R}$, then it has a supremum in $\mathbb{R}$.

2.3.2.6 Prop Let $A\subseteq[-\infty,+\infty]$:

- $\forall c\in\mathbb{R}\quad\sup(A+c)=(\sup A)+c$.
- $\forall\lambda\in\mathbb{R}_{\geq 0}\quad\sup(\lambda A)=\lambda\sup(A)$.
- $\forall\lambda\in\mathbb{R}_{\leq 0}\quad\sup(\lambda A)=\lambda\inf(A)$.

inf takes the same.

2.3.2.7 Theorem Let I and J be non-empty sets.
$f:I\to[-\infty,+\infty],g:J\to[-\infty,+\infty]$.
$a=\sup\limits_{x\in I}f(x)\quad b=\sup\limits_{y\in J}g(y)\quad c=\sup\limits_{(x,y)\in I\times J,\{f(x),g(y)\}\neq\{+\infty,-\infty\}}(f(x)+g(y))$.
If $\{a,b\}\neq\{+\infty,-\infty\}$, then $c=a+b$.
inf takes the same.

2.3.2.8 Corollary Let I be a non-empty set, $f:I\to[-\infty,+\infty],g:J\to[-\infty,+\infty]$. Then

$$\sup_{x\in I,\{f(x),g(x)\}\neq\{+\infty,-\infty\}}(f(x)+g(x))\leq(\sup_{x\in I}f(x))+(\sup_{x\in I}g(x))$$

inf takes the similar($\leq\to\geq$) (provided when the sum are defined).

2.4 Vector Spaces

In this section, K denotes a unitary ring. Let 0 be zero element of K, 1 be the unity of K.

2.4.1 K-module

2.4.1.1 Def Let $(V,+)$ be a commutative group. We call *left/right K-module structure* any mapping $\Phi:K\times V\to V$:

- $\forall(a,b)\in K\times K,\forall x\in V\quad\Phi(ab,x)=\Phi(a,\Phi(b,x))/\Phi(b,\Phi(a,x))$;
- $\forall(a,b)\in K\times K,\forall x\in V\quad\Phi(a+b,x)=\Phi(a,x)+\Phi(b,x)$;
- $\forall a\in K,\forall(x,y)\in V\times V\quad\Phi(a,x+y)=\Phi(a,x)+\Phi(a,y)$;
- $\forall x\in V,\Phi(1,x)=x$.

A commutative group (V,+) equipped with a left/right K-module structure is called a *left/right K-module.*

2.4.1.2 Opposite Ring Let K^{op} be the set K equipped with the following composition laws:

- $K\times K\to K,\ (a,b)\mapsto a+b$
- $K\times K\to K,\ (a,b)\mapsto ba$

Then K^{op} forms a unitary ring.
Any left K^{op}-module is a right K-module;
Any right K^{op}-module is a left K-module.
$(K^{op})^{op}=K$.

2.4.1.3 Notation When we talk about a left/right K-module $(V, +)$, we often write its left K-module structure as $K \times V \to V \quad (a, x) \mapsto ax$

The defs become:

$$\begin{aligned} &\forall (a,b) \in K \times K, \forall x \in V \quad (ab)x = a(bx)/b(ax) \\ &\forall (a,b) \in K \times K, \forall x \in V \quad (a+b)x = ax + bx \\ &\forall a \in K, \forall (x,y) \in V \times V \quad a(x+y) = ax + ay \\ &\forall x \in V \quad 1x = x \end{aligned}$$

2.4.2 K-vector Space

If K is commutative, then $K^{op} = K$, so left K-module and right K-module structure are the same. We simply call them K-module structure. A commutative group equipped with a K-module structure is called a K-module. If K is a field, a K-module is also called a K-vector space, and the elements of K are called *scalars*, the elements of V are called *vectors*.

2.4.2.1 Remark Let $\Phi : K \times V \to V$ be a left or right K-module structure $\forall x \in V, \Phi(\cdot, x) : K \to V \quad (a \in K) \mapsto \Phi(a, x)$ is a morphism of commutative groups. Hence $\Phi(0, x) = 0, \Phi(-a, x) = -\Phi(a, x)$.

$\forall a \in K, \Phi(a, \cdot) : V \to V$ is a morphism of groups. Hence $\Phi(a, 0) = 0, \Phi(a, -x) = -\Phi(a, x) (\cdot \textit{ is a var})$.

2.4.2.2 Associativity $\forall x \in K$, $(f(f+g)+h)(x) = (f+g)(x)+h(x) = f(x)+g(x)+h(x) = f(x)+((g+h)(x)) = (f+(g+h))(x)$.

The mapping $K \times K^I \to K^I : (a, f : x \mapsto f(x)) \mapsto af : x \mapsto af(x)$ is a left K-module structure;

The mapping $K \times K^I \to K^I : (a, f : x \mapsto f(x)) \mapsto af : x \mapsto f(x)a$ is a right K-module structure.

2.4.2.3 Remark We can also write an element $\mu \in K^I$ in the form of a family $(\mu_i)_{i \in I}$ of elements in K, where μ_i is the image of $i \in I$ by μ. Then $(\mu_i)_{i\in I} + (\nu_i)_{i\in I} := (\mu_i + \nu_i)_{i\in I}$, $a(\mu_i)_{i\in I} := (a\mu_i)_{i\in I}$, $(\mu_i)_{i\in I}a = (\mu_i a)_{i\in I}$.

2.4.3 Sub K-module

Let V be a left/right K-module. If W is a subgroup of V s.t. $\forall a \in K, \forall x \in W \quad ax/xa \in W$, then we say that W is left/right sub K-module of V.

2.4.3.1 Example of Direct Sum Let I be a set. Let $K^{\oplus I}$ be the subset of K^I composed of mappings $f : I \to K$ s.t. $I_f = \{x \in I \mid f(x) \neq 0\}$ is finite. It is a left and right sub K-module of K^I.

In fact, $\forall (f, g) \in K^{\oplus I} \times K^I \quad I_{f-g} = \{x \in I \mid f(x) - g(x) \neq 0\} \subseteq I_f \bigcup I_g$, which is finite.

Hence $f - g \in K^{\oplus I}$. So $K^{\oplus I}$ is a subgroup of K^I.

$\forall a \in K, \forall f \in K^{\oplus I}$, $I_{af} \subseteq I_f, I_{(x \mapsto f(x)a)} \subseteq I_f$, so the proof is finished.

2.4.4 Morphisms of K-modules

2.4.4.1 Def Let V and W be left K-module, A morphism of groups $\phi : V \to W$ is called a morphism of left K-modules iff $\forall (a, x) \in K \times V, \phi(ax) = a\phi(x)$.

2.4.4.2 K-linear Mapping If K is commutative, a morphism of K-modules is also called a K-linear mapping. We denote by $\hom_{K-Mod}(V, W)$ the set of all morphism of left K-module from V to W. This is a subgroup of W^V. That is because if f and g are elements of $\hom_{K-Mod}(E, F)$, then $f - g$ is also a morphism of left K-module.

2.4.4.3 Proof

- $(f-g)(x+y) = f(x+y)-g(x+y) = (f(x)+f(y))-(g(x)+g(y)) = (f(x)-g(x))+(f(y)-g(y)) = (f-g)(x)+(f-g)(y)$
- $(f-g)(ax) = f(ax) - g(ax) = af(x) - ag(x) = a(f(x) - g(x)) = a(f-g)(x)$

If K isn't commutative, $\hom_{\text{K-Mod}}(E, F)$ isn't a left K-module, since $\lambda f(ax) = \lambda a f(x) \neq a\lambda f(x)$. Otherwise, it is.

2.4.4.4 Theorem (Usually used in k-linear mappings, but stronger than that Prop.)

Let V be a left K-module. Let I be a set.

The mapping $\hom_{K-Mod}(K^{\oplus I}, V) \to V^I : \ \phi \mapsto (\phi(e_i))_{i\in I}$ is a bijection where $e_i : I \to K : j \mapsto \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$

2.4.4.5 Proof *Injectivity* Suppose $\phi_1, \phi_2 \in \hom_{K-Mod}(K^{\oplus I}, V)$ and $(\phi_1(e_i))_{i\in I} = (\phi_2(e_i))_{i\in I}$. Then $\phi_1(X) = \phi_2(X)$ for all $X \in K^{\oplus I}$. Since any element of $K^{\oplus I}$ is a finite linear combination of e_i's and both ϕ_1 and ϕ_2 are K-module morphisms, we conclude that $\phi_1 = \phi_2$. Therefore, the map is injective.

Surjectivity Given $(v_i)_{i\in I} \in V^I$, define a morphism $\phi : K^{\oplus I} \to V$ by setting $\phi(e_i) = v_i$ for each $i \in I$. This def uniquely determines ϕ, since any element of $K^{\oplus I}$ is a finite sum of the basis elements $\{e_i\}_{i\in I}$. Thus, every element of V^I corresponds to a morphism, proving surjectivity.

2.4.5 Column

For any $(x_1, ..., x_n) \in V^n$, by the theorem, there exists a unique morphism of left K-modules $\phi : K^n \to V$ s.t. $\forall i \in 1, ..., n, \phi(e_i) = x_i$. We write this ϕ as a column $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. It sends $(a_1, ..., a_n) \in K^n$ to $a_1x_1 + ... + a_nx_n$.

2.4.6 Kernel

Let G and H be groups and $f : G \to H$ be a morphism of groups, $\ker(f) = \{x \in G \mid f(x) = e_H\}$, called the kernel of f.

2.4.6.1 Theorem f is injective iff $\ker(f) = \{e_G\}$.

2.4.6.2 Proof Suppose f is injective, then $\mathrm{Ker}(f) = e_G$.
Suppose $\mathrm{Ker}(f) = e_G$, let $(a, b) \in G^2$ s.t. $f(a) = f(b)$. $f(ab^{-1}) = f(a)f(b^{-1}) = f(b)f(b^{-1}) = e_H$. Hence $ab^{-1} = e_G$, therefore $a = b$.

2.4.6.3 A New Left K-module Let (V,+) be a commutative group, I be a set. We have defined the composition law + on V^I. Then V^I forms a commutative group. Equipping with K, it's a left K-module.

2.4.6.4 Theorem Let V be a left K-module, I be a set The mapping $f : \hom_{K-Mod}(K^{\oplus I}, V) \to V^I$: $\phi \mapsto (\phi(e_i))_i \in I$ is an isomorphism of groups, where $e_i : I \to K : j \mapsto \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$

2.4.6.5 Proof One has $(\phi + \psi)(e_i) = \phi(e_i) + \psi_i(e_i)$, $\forall (\phi, \psi) \in \hom_{K-Mod}(K^{\oplus I}, V)^2$. Hence $f(\phi + \psi) = (\phi(e_i) + \psi(e_i))_{i \in I} = f(\phi) + f(\psi)$. So f is a morphism of groups.
bijection We have proved this previously.

2.4.6.6 Remark Suppose that K' is a unitary ring and V is also equipped with a right K'-module structure, Then $\hom_{K-Mod}(K^{\oplus I}, V) \subseteq V^{K^{\oplus I}}$ is a right sub K'-module, and ψ_i in the theorem is a right K'-module isomorphism.

2.5 Monotone Mappings

2.5.0.1 Def Let I and X be partially ordered sets, $f : I \to X$ be a mapping.

- If $\forall (a, b) \in I \times I$ s.t. $a < b$. One has $f(a) \leq f(b)/f(a) < f(b)$, then we say that f is *increasing/strictly increasing*. *Decreasing* takes similar way.
- If f is (strictly) increasing or decreasing, we say that f is *(strictly) monotone*.

2.5.0.2 Prop Let X,Y,Z be partially ordered sets. $f : X \to Y, g : Y \to Z$ be mappings:

- If f and g have the same monotonicity, then $g \circ f$ is increasing.
- If f and g have different monotonicities, then $g \circ f$ is decreasing.

Strict monotonicities take the same.

2.5.1 Def on Functions

Let f be a function from a partially ordered set I to another partially ordered set. If $f \mid_{Dom(f)} \to X$ is (strictly) increasing/decreasing then we say that f is (strictly) increasing/decreasing.

2.5.1.1 Prop Let I and X be partially ordered sets. f be function from I to X.

- If f is increasing/decreasing and f is injection, then f is strictly increasing/decreasing.
- Assume that I is totally ordered and f is strictly monotone, then f is injection.

2.5.1.2 Prop Let A be totally ordered set, B be a partially ordered set, f be an injective function from A to B. If f is increasing/decreasing, then so is f^{-1}.

2.5.2 Def of Morphism on Partially Ordered Sets

Let X and Y be partially ordered sets. $f : X \to Y$ be a bijection. If both f and f^{-1} are increasing, then we say that f is an isomorphism of partially ordered sets. (If X is totally ordered, then a mapping $f : X \to Y$ is an isomorphism of partially ordered sets iff f is a bijection and f is increasing)

2.5.2.1 Prop Let I be a subset of $\mathbb{N}$ which is infinite. Then there is a unique increasing bijection $\lambda_I : \mathbb{N} \to I$.

2.5.2.2 Proof *bijection* We construct $f : \mathbb{N} \to I$ by induction as follows. Let $f(0) = \min I$, suppose that $f(0), ..., f(n)$ are constructed, then we take $f(n+1) := \min(I\backslash\{f(0), ..., f(n)\})$. Since $I\backslash\{f(0), ..., f(n-1)\} \supseteq I\backslash\{f(0), ..., f(n)\}$., therefore $f(n) \leq f(n+1)$. Since $f(n+1) \notin \{f(0), ..., f(n)\}$, we have $f(n) < f(n+1)$. Hence f is strictly increasing and this is injective.

If f is not surjective, then $I\backslash Im(f)$ has a element N. Let $m = \min\{n \in \mathbb{N} \mid N \leq f(n)\}$. Since $N \notin Im(f), N < f(m)$. So $m \neq 0$. Hence $f(m-1) < N < f(m) = \min(I\backslash\{f(0), ..., f(m-1)\})$. By Def, $N \in I\backslash Im(f) \subseteq I\backslash\{f(0), ..., f(m-1)\}$. Hence $f(m) \leq N$, causing contradiction.

uniqueness Easy to prove.

2.6 Sequence and Series

Let $I \subseteq \mathbb{N}$ be a infinite subset.

2.6.0.1 Remark Let X be a set. We call sequence in X parametrized by I a mapping from I to X.

2.6.0.2 Remark If K is a unitary ring and E is a left K-module then the set of sequence E^I admits a left K-module structure. If $x = (x_n)_{n\in I}$ is a sequence in E, we define a sequence $\sum(x) := (\sum\limits_{i\in I, i\leq n} x_i)_{n\in\mathbb{N}}$, called the series associated with the sequence x.

2.6.0.3 Prop $\sum : E^I \to E^{\mathbb{N}}$ is a morphism of left K-module.

2.6.0.4 Proof Let $x = (x_i)_{i\in I}$ and $y = (y_i)_{i\in I}$ be elements of E^I
$\sum\limits_{i\in I, i\leq n} (x_i + y_i) = (\sum\limits_{i\in I, i\leq n} x_i) + (\sum\limits_{i\in I, i\leq n} y_i), \lambda \sum\limits_{i\in I, i\leq n} x_i = \sum\limits_{i\in I, i\leq n} \lambda x_i$

2.6.0.5 Prop Let I be a totally ordered set. X be a partially ordered set, $f : I \to X$ be a mapping, $J \subseteq I$. Assume that J does not have any upper bound in I.(Equals to "J is infinite")

- If f is increasing, then $f(I)$ and $f(J)$ have the same upper bounds in X.
- If f is decreasing, then $f(I)$ and $f(J)$ have the same lower bounds in X.

2.6.0.6 Proof Since $f(J) \subset f(I)$, and upper bound of $f(I)$ is an upper bound of $f(J)$. Let M be an upper bound of $f(J)$. Let $x \in I$. Since J does not have any upper bound in I, $\exists y \in J, y > x$. Hence $f(x) < f(y)$. Hence $f(x) \leq f(y) \leq M$. Hence M is an upper bound of $f(I)$.

2.6.1 Limit

2.6.1.1 Def Let $I \subseteq \mathbb{N}$ be an infinite subset. $\forall (x_i)_{n\in I} \in [-\infty, +\infty]^I$ where $[-\infty, +\infty]$ denotes $\mathbb{R} \bigcup \{-\infty, +\infty\}$, we define:

$$\limsup_{n\in I, n\to+\infty} x_n := \inf_{n\in I}(\sup_{i\in I, i\geq n} x_i), \quad \liminf_{n\in I, n\to+\infty} x_n := \sup_{n\in I}(\inf_{i\in I, i\geq n} x_i)$$

If $\limsup\limits_{n\in I, n\to+\infty} x_n = \liminf\limits_{n\in I, n\to+\infty} x_n = l$, we then say that $(x_n)_{n\in I}$ tends to l and that l is the limit of $(x_n)_{n\in I}$. If in addition $(x_n)_{n\in I} \in \mathbb{R}^I$ and $l \in \mathbb{R}$, we say that $(x_n)_{n\in I}$ converges to l.

2.6.1.2 Remark If $J \subseteq \mathbb{N}$ is an infinite subset, then:

$$\limsup_{n\in I, n\to+\infty} = \inf_{n\in J}(\sup_{i\in I, i\geq n} x_i), \quad \liminf_{n\in I, n\to+\infty} x_n = \sup_{n\in J}(\inf_{i\in I, i\geq n} x_i)$$

Therefore, if we change the values of finitely many terms in $(x_i)_{i\in I}$ the limit superior and the limit inferior do not change. In fact, if we take $J = \mathbb{N}\backslash\{0, ..., m\}$, then $\inf\limits_{n\in J}(...)$ and $\sup\limits_{n\in J}(...)$ only depend on the values of $x_i, i \in I, i \geq m$.

2.6.1.3 Prop $\forall (x_n)_{n\in I} \in [-\infty, +\infty]^I$, $\liminf\limits_{n\in I, n\to+\infty} x_n \leq \limsup\limits_{n\in I, n\to+\infty} x_n$

2.6.1.4 Prop Let $(x_n)_{n\in I} \in [-\infty, +\infty]^I$

$$\forall c \in \mathbb{R} \qquad \limsup_{n\in I, n\to+\infty} (x_n + c) = (\limsup_{n\in I, n\to+\infty} x_n) + c, \ \liminf_{n\in I, n\to+\infty} (x_n + c) = (\liminf_{n\in I, n\to+\infty} x_n) + c$$

$$\forall c \in \mathbb{R}_{>0} \qquad \limsup_{n\in I, n\to+\infty} (\lambda x_n) = \lambda \limsup_{n\in I, n\to+\infty} x_n, \ \liminf_{n\in I, n\to+\infty} (\lambda x_n) = \lambda \liminf_{n\in I, n\to+\infty} x_n$$

$$\forall c \in \mathbb{R}_{<0} \qquad \limsup_{n\in I, n\to+\infty} (\lambda x_n) = \lambda \liminf_{n\in I, n\to+\infty} x_n, \ \liminf_{n\in I, n\to+\infty} (\lambda x_n) = \lambda \limsup_{n\in I, n\to+\infty} x_n$$

2.6.1.5 Prop Let $(x_n)_{n\in I}$ be elements in $[-\infty, +\infty]^I$. Suppose that there exists $N_0 \in \mathbb{N}$ s.t. $\forall n \in I, n \geq N_0$, one has $x_n \leq y_n$, then

$$\limsup_{n\in I, n\to+\infty} (x_n) \leq \limsup_{n\in I, n\to+\infty} y_n, \quad \liminf_{n\in I, n\to+\infty} (x_n) \leq \liminf_{n\in I, n\to+\infty} y_n$$

2.6.1.6 Theorem Let $(x_n)_{n\in I}, (y_n)_{n\in I}, (z_n)_{n\in I}$ be elements of $[-\infty, +\infty]^I$.

Suppose that

- $\exists N_0 \in \mathbb{N}, \forall n \in I, n \geq N_0$ one has $x_n \leq y_n \leq z_n$
- $(x_n)_{n\in I}$ and $(z_n)_{n\in I}$ tend to the same limit l

Then $(y_n)_{n\in I}$ tends to l.

2.6.1.7 Subsequence Let I be an infinite subset of $\mathbb{N}$, and $(x_n)_{n\in I}$ be a sequence in some set X. We call subsequence of $(x_n)_{n\in I}$ a sequence of the form $(x_n)_{n\in J}$, where J is an infinite subset of I.

2.6.1.8 Prop Let I and J be infinite subset of $\mathbb{N}$ s.t. $J \subseteq I$. $\forall (x_n)_{n\in I} \in [-\infty, +\infty]^I$, one has

$$\liminf_{n\in I, n\to+\infty} (x_n) \leq \liminf_{n\in J, n\to+\infty} x_n, \quad \limsup_{n\in I, n\to+\infty} (x_n) \geq \limsup_{n\in J, n\to+\infty} x_n$$

In particular, if $(x_n)_{n\in I}$ tends to $l \in [-\infty, +\infty]$, then $(x_n)_{n\in J}$ tends to l.

2.6.2 Monotone Bounded Principle

2.6.2.1 Theorem Let $I \subseteq \mathbb{N}$ be an infinite subset and $(x_N)_{n\in I}$ be a sequence in $[-\infty, +\infty]$

- If the mapping $(n \in I) \mapsto x_n$ is increasing, then $(x_N)_{i\in I}$ tends to $\sup_{n\in I} x_n$.
- If the mapping $(n \in I) \mapsto x_n$ is decreasing, then $(x_N)_{i\in I}$ tends to $\inf_{n\in I} x_n$.

2.6.2.2 Proof $\forall n \in I$, $\inf_{i\in I, i\geq n} x_i = x_n$, so liminf x_n=sup x_n. $\forall n, m \in I$, $\sup_{i\in I, i\geq n} x_i = \sup_{i\in I, i\geq m} x_i$, so $\limsup s_n = \sup x_n$.

2.6.2.3 Notation If a sequence $(x_N)_{n\in I} \in [-\infty, +\infty]$ tends to some $l \in [-\infty, +\infty]$ the expression $\lim_{n\in I, n\to+\infty} x_n$ denotes this limit l.

2.6.2.4 Corollary Let $(x_n)_{n\in I}$ be a sequence in $\mathbb{R}_{\geq 0}$ Then the series $\sum_{n\in I} x_n$(the sequence $(\sum_{i\in I, i\leq n} x_n)_{n\in\mathbb{R}}$) tends to an element in $\mathbb{R}_{\geq 0} \bigcup \{+\infty\}$. It converges in $\mathbb{R}$ iff it is bounded from above (namely, it has an upper bound in $\mathbb{R}$).

2.6.2.5 Notation If a series $\sum_{n\in I} x_n$ in $[-\infty, +\infty]$ tends to some limit, we use the expression $\sum_{n\in I} x_n$ to denote the limit.

2.6.2.6 Prop In the set of real numbers $\mathbb{R}$, every nonempty subset that is bounded above (or bounded below) has a least upper bound (supremum) (or greatest lower bound, infimum).

2.6.2.7 Prop Let $\alpha \in \mathbb{R}_{>1}$ and $b \in \mathbb{R}_{>0}$. The series $\Sigma_{n\in\mathbb{N}_{\geq 1}} \frac{1}{\alpha^{n^b}}$ converges.

2.6.3 Theorem: Bolzano-Weierstrass

(Stronger than sequentially compact in this topological space.)

Let $(x_n)_{n\in I}$ be a sequence in $[-\infty, +\infty]$. There exists a subsequence of $(x_n)_{n\in I}$ that tends to $\limsup_{n\in I, n\to+\infty} x_n$. There exists a subsequence of $(x_n)_{n\in I}$ that tends to $\liminf_{n\in I, n\to+\infty} x_n$.

2.6.3.1 Proof Let $J = \{n \in I \mid \forall m \in I, \text{ if } m \geq n \text{ then } x_m \leq x_n\}$. If J is infinite, the sequence $(x_N)_{n\in J}$ is decreasing, so it tends to $\liminf_{n\in J, n\to\infty} x_n = \inf_{n\in J} x_n$. $\forall n \in J$, by Def, $x_n = \sup_{i\in I, i\geq n} x_i$. So $\limsup_{n\in I, n\to+\infty} x_n = \inf_{n\in J} \sup_{i\in I, i\geq n} x_i = \inf_{n\in J} x_n = \lim_{n\in J, n\to+\infty} x_n$.

Assume that J is finite. Let $n_0 \in I$ s.t. $\forall n \in J, n < n_0$. Denote by $l = \sup_{n\in I, n\geq n_0} x_n$. Let $N \in \mathbb{N}$ s.t. $N \geq n_0$. By Def, $\sup_{i\in I, i\geq N} x_i \leq l$. If the strict inequality $\sup_{i\in I, i\geq N} x_i < l$ holds, then $\sup_{i\in I, i\geq N} x_i$ is NOT an upper bound of $\{x_n \mid n \in I, n_0 \leq n < N\}$. So there exists $n \in I$ s.t. $n_0 \leq n < N$ s.t. $x_n > \sup_{i\in I, i\geq N} x_i$.

We may also assume that n is the largest among the elements of $I \cap [n_0, N[$ that satisfy this inequality. Then $\forall m \in I$ if $m \geq n$ then $x_m \leq x_n$ Thus $n \in J$ contradicts the maximality of n_0. Therefore, $l = \sup_{i\in I, i\geq N} x_i$, which leads to $\limsup_{n\in I, n\to+\infty} x_n = l$.

Moreover, if $m \in I, m \geq n_0$, then $m \notin J$, so $x_m < l$ (since otherwise $x_m = \sup_{i\in I, i\geq m} x_i$, and hence $m \in J$). Hence, $\forall$ finite subset I' of $\{m \in I \mid m \geq n_0\}$, $\max_{i\in I} x_i < l$ and hence $\exists n \in I$, s.t. $n > \max I'$, and $\max_{i\in I'} x_i < x_n$.

We construct by induction an increasing sequence $(n_j)_{j\in\mathbb{N}}$ in I. Let n_0 be as above. Let $f : \mathbb{N} \to I_{\geq n_0}$ be a surjective mapping. If n_j is chosen, we choose $n_{j+1} \in I$ s.t. $n_{j+1} > n_j, x_{n_{j+1}} > \max\{x_{f(j)}, x_{n_j}\}$.

Hence, the sequence $(x_{n_j})_{j\in\mathbb{N}}$ is increasing, and $\sup_{j\in\mathbb{N}} x_{n_j} \leq \sup_{j\in\mathbb{N}} x_{f(j)} = \sup_{n\in I, n\geq n_0} x_n = l$.

So $(x_{n_j})_{j\in\mathbb{N}}$ tends to l, $\limsup_{n\in I, n\to+\infty} x_n$

2.7 Cauchy Sequence Again

Let $(x_n)_{n\in I}$ be a sequence in $\mathbb{R}$. If $\inf_{N\in\mathbb{N}} \sup_{(n,m)\in I\times I,\ n,m\geq N} |x_n - x_m| = \lim_{N\to+\infty} \sup_{(n,m)\in I\times I,\ n,m\geq N} |x_n - x_m| = 0$, then we say that $(x_n)_{n\in I}$ is a Cauchy sequence.

2.7.1 Prop

- If $(x_i)_{i\in I} \in \mathbb{R}^I$ converges to some $l \in \mathbb{R}$, then it is a Cauchy sequence.
- If $(x_i)_{i\in I}$ is a Cauchy sequence converges to 0, there exists $M > 0$ s.t. $\forall n \in I \quad |x_n| \leq M$.
- If $(x_n)_{n\in I}$ is a Cauchy sequence, then $\forall J \subseteq I$ infinite, $(x_n)_{n\in J}$ is a Cauchy sequence.
- If $(x_n)_{n\in I}$ is a Cauchy sequence, then $\forall J \subseteq I$ infinite and $l \in \mathbb{R}$ s.t. $(x_n)_{n\in I}$ converges to l, then $(x_n)_{n\in J}$ converges to l too.

2.7.2 Theorem: Completeness of Real Number

If $(x_n)_{n\in I} \in \mathbb{R}^I$ is a Cauchy sequence, then it converges in $\mathbb{R}$.

2.7.2.1 Proof Since $(x_n)_{n\in I}$ is a Cauchy sequence, $\exists M \in \mathbb{R}_{>0}$ s.t. $-M \leq x_n \leq M \quad \forall n \in I$. So $\limsup_{n\in I, n\to+\infty} x_n \in \mathbb{R}$. By Bolzano-Weierstrass theorem. $\exists J \subseteq I$ infinite s.t. $(x_n)_{n\in I}$ converges to $\limsup_{n\in I, n\to+\infty} x_n \in \mathbb{R}$. Therefore, $(x_n)_{n\in I}$ converges to the same limit.

2.7.3 Absolutely Converge

We say that a series $\sum_{n\in I} x_n \in \mathbb{R}$ converges absolutely if $\sum_{n\in I} |x_n| < +\infty$.

2.7.3.1 Prop If a series $\sum_{n\in I} x_n$ converges absolutely, then it converges in $\mathbb{R}$.

2.7.3.2 Remark $(x_n)_{n\in I}$ converges to 0 can't imply $\sum_{n\in I} x_n$ converges absolutely.

2.8 Comparison and Technics of Computation

2.8.1 Def of O(),o()

Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be sequences in $\mathbb{R}$.

- If there exists $M \in \mathbb{R}_{>0}$ and $N \in \mathbb{N}$ s.t. $\forall n \in I_{\geq N}, |x_n| \leq M|y_n|$ then we write $x_n = O(y_n), n \in I, n \to +\infty$.
- If there exists $(\varepsilon_n)_{n\in I} \in \mathbb{R}^I$ and $N \in \mathbb{N}$ s.t. $\lim_{n\in I, n\to+\infty} \varepsilon_n = 0$ and $\forall n \in I_{\geq N}, |x_n| \leq |\varepsilon_n y_n|$, then we write $x_n = o(y_n) \quad n \in I, n \to +\infty$.

2.8.1.1 Prop Let I and X be partially ordered sets and $f : I \to X$ be an increasing/decreasing mapping. Let J be a subset of I. Assume that any elements of I has an upper bound in J. Then $f(I)$ and $f(J)$ have the same upper/lower bounds in X.

2.8.1.2 Theorem Let I be a totally ordered set, $f : I \to [-\infty, +\infty]$ and $g : I \to [-\infty, +\infty]$ be two mappings that are both increasing/decreasing. Then the following equalities holds, provided that the sum on the right hand side of the equality is well defined.

$$\sup_{x\in I,\{f(x),g(x)\}\neq\{-\infty,+\infty\}}(f(x)+g(x)) = (\sup_{x\in I} f(x))+(\sup_{y\in I} g(y)),\quad \inf_{x\in I,\{f(x),g(x)\}\neq\{-\infty,+\infty\}}(f(x)+g(x)) = (\inf_{x\in I} f(x))+(\inf_{y\in I} g(y))$$

2.8.1.3 Proof We can assume that f and g increasing. Let $a = \sup f(I)$,$b = \sup g(I)$, $A = \{(x, y) \in I\times I \mid \{f(x), g(y)\} \neq \{-\infty, +\infty\}\}$.

We equip A with the following order relation: $(x, y) \leq (x', y')$ iff $x \leq x', y \leq y'$

Let $B = A \cap \Delta_I = \{(x, y) \in A \mid x = y\}$. Consider $h : A \to [-\infty, +\infty]\quad h(x, y) = f(x) + g(y)$, which is increasing.

Let $(x, y) \in A$. Assume that $x \leq y$. If $\{f(y), g(y)\} \neq \{-\infty, +\infty\}$, then $(y, y) \in B$ and $(x, y) \leq (y, y)$. If $\{f(y), g(y)\} = \{-\infty, +\infty\}$ and for $(x, y) \in A \to f(y) = +\infty, g(y) = -\infty$. So $a = +\infty$. Hence $b > -\infty$. So $\exists z \in I$ s.t. $g(z) > -\infty$. We should have $y \leq z$ Hence $f(z) + g(z)$ is well defined, $(z, z) \in B$ and $(x, y) \leq (z, z)$.

Similarly, if $x \geq y$, (x, y) has also an upper bound in B. Therefore: $\sup h(A) = \sup h(B)$.

inf takes the same.

2.8.1.4 Prop Let $I \subseteq \mathbb{N}$ be an infinite subset. Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be elements of $[-\infty, +\infty]^I$ s.t., $\forall n \in I \quad \{x_n, y_n\} \neq \{-\infty, +\infty\}$. Then the following inequalities holds, provided that the sum on the right hand side is well defined.

$$\limsup_{n\in I,n\to+\infty}(x_n + y_n) \leq (\limsup_{n\in I,n\to+\infty} x_n) + (\limsup_{n\in I,n\to+\infty} y_n),\quad \liminf_{n\in I,n\to+\infty}(x_n + y_n) \geq (\liminf_{n\in I,n\to+\infty} x_n) + (\liminf_{n\in I,n\to+\infty} y_n)$$

2.8.1.5 Proof $\forall n \in \mathbb{N}$, let $A_N = \sup_{n\in I,n\geq N} x_n \quad B_N = \sup_{n\in I,n\geq N} y_n$. $(A_N)_{N\in\mathbb{N}}$ and $(B_N)_{N\in\mathbb{N}}$ are decreasing, and $\limsup_{n\in I,n\to+\infty} x_n = \inf_{N\in\mathbb{N}} A_N \quad \limsup_{n\in I,n\to+\infty} y_n = \inf_{N\in\mathbb{N}} B_N$. By theorem: $\inf_{N\in\mathbb{N}} A_N + \inf_{N\in\mathbb{N}} B_N = \inf_{N\in\mathbb{N},\{A_N,B_N\}\neq\{-\infty,+\infty\}}(A_N + B_N)$ Let $C_N = \sup_{n\in I,n\geq N}(x_n + y_n) \leq A_N + B_N$ if $A_N + B_N$ is defined.

Therefore, $\inf_{N\in\mathbb{N}} C_N \leq \inf_{N\in\mathbb{N},\{A_N,B_N\}\neq\{-\infty,+\infty\}}(A_N + B_N) = \inf_{N\in\mathbb{N}} A_N + \inf_{N\in\mathbb{N}} B_N$.

2.8.1.6 Prop Let $I \subseteq \mathbb{N}$ be an infinite subset. Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be elements of $[-\infty, +\infty]^I$ s.t. $\forall n \in I \quad \{x_n, y_n\} \neq \{-\infty, +\infty\}$. Then the following inequalities holds, provided that the sum on the right hand side is well defined.

$$\limsup_{n\in I,n\to+\infty}(x_n + y_n) \geq (\limsup_{n\in I,n\to+\infty} x_n) + (\liminf_{n\in I,n\to+\infty} y_n),\quad \liminf_{n\in I,n\to+\infty}(x_n + y_n) \geq (\liminf_{n\in I,n\to+\infty} x_n) + (\limsup_{n\in I,n\to+\infty} y_n)$$

2.8.1.7 Proof Fake proof: $\limsup_{n\in I,n\to+\infty} x_n = \limsup_{n\in I,n\to+\infty}(x_n + y_n - y_n) \leq \limsup_{n\in I,n\to+\infty}(x_n + y_n) - \liminf_{n\in I,n\to+\infty} y_n$.

To have a true proof, we only need to discuss conditions with ∞.

2.8.1.8 Theorem Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be elements of $[-\infty, +\infty]^I$. Assume that $\forall n \in I, y_n \in \mathbb{R}$ and $(y_n)_{n\in I}$ converge to some $l \in \mathbb{R}$. Then: $\limsup_{n\in I,n\to+\infty}(x_n + y_n) = (\limsup_{n\in I,n\to+\infty} x_n) + l$, $\liminf_{n\in I,n\to+\infty}(x_n + y_n) = (\liminf_{n\in I,n\to+\infty} x_n) + l$.

2.8.1.9 Prop Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be elements of $[-\infty, +\infty]^I$. Then:

$$\liminf_{n\in I,n\to+\infty} \max\{x_n, y_n\} = \max\{\liminf_{n\in I,n\to+\infty} x_n, \liminf_{n\in I,n\to+\infty} y_n\},$$

$$\liminf_{n\in I,n\to+\infty} \min\{x_n, y_n\} = \min\{\liminf_{n\in I,n\to+\infty} x_n, \liminf_{n\in I,n\to+\infty} y_n\}.$$

2.8.1.10 Proof $\max\{x_n, y_n\} \geq x_n,\ y_n$, by the theorem of Bolzano-Weierstrass theorem, there exists an infinite subset J of I s.t. $\lim_{n\in J,n\to+\infty} \max\{x_n, y_n\} = \limsup_{n\in J,n\to+\infty} \max\{x_n, y_n\}$. Let $J_1 = \{n \in J \mid x_n \geq y_n\} \quad J_2 = \{n \in J \mid x_n \leq y_n\}$, $J_1 \bigcup J_2 = J$, so either J_1 or J_2 is infinite.

Suppose that J_1 is infinite, then $\lim_{n\in J,n\to+\infty} \max\{x_n, y_n\} = \lim_{n\in J_1,n\to+\infty} \max\{x_n, y_n\} = \lim_{n\in J,n\to+\infty} x_n \leq \limsup_{n\in I,n\to+\infty} x_n$. If J_2 is infinite, $\lim_{n\in J,n\to+\infty} \max\{x_n, y_n\} = \lim_{n\in J_2,n\to+\infty} \max\{x_n, y_n\} = \lim_{n\in J,n\to+\infty} x_n \leq \limsup_{n\in I,n\to+\infty} x_n$.

2.8.1.11 Theorem Let $(a_N)_{n\in I} \in \mathbb{R}^I$, $l \in \mathbb{R}$. The following statements are equivalent:

1. $(a_N)_{n\in I}$ converges to l.
2. $\limsup_{n\in I,n\to+\infty} |a_n - l| = 0$.

2.8.1.12 Proof

$$|a_n - l| = \max\{a_n - l, l - a_n\}$$

$$\limsup_{n\in I, n\to+\infty} |a_n - l| = \max\{(\limsup_{n\in I, n\to+\infty} a_n) - l, l - (\liminf_{n\in I, n\to+\infty} a_n)\}$$

(1) → (2): If $(a_n)_{n\in I}$ converges to l, then $\limsup_{n\in I, n\to+\infty} a_n = \liminf_{n\in I, n\to+\infty} a_n = l$.

(2) → (1): If $\limsup_{n\in I, n\to+\infty} |a_n - l| = 0$, then $\limsup_{n\in I, n\to+\infty} a_n \leq l \leq \liminf_{n\in I, n\to+\infty} a_n$.

Therefore: $\limsup_{n\in I, n\to+\infty} a_n = \liminf_{n\in I, n\to+\infty} a_n = l$.

2.8.1.13 Remark $(a_N)_{n\in I}$ converges to 0 is equivalent to $\limsup_{n\in I, n\to+\infty} |a_n| = 0$.

2.8.1.14 Remark Let $(a_n)_{n\in I}$ be a sequence in $\mathbb{R}$, $l \in \mathbb{R}$.
The sequence $(a_n)_{n\in I}$ converges to l iff $a_n - l = o(1) \quad n \in I, n \to +\infty$.

2.8.2 Calculates on O(),o()

2.8.2.1 Plus Let $(a_n)_{n\in I}$ $(a'_n)_{n\in I}$ and $(b_n)_{n\in I}$ be elements in $\mathbb{R}^I$:

- If $a_n = O(b_n), a'_n = O(b_n), n \in I, n \to +\infty$, then $\forall(\lambda, \mu) \in \mathbb{R}^2,\ \lambda a_n + \mu a'_n = O(b_n), n \in I, n \to +\infty$.
- If $a_n = o(b_n), a'_n = o(b_n), n \in I, n \to +\infty$, then $\forall(\lambda, \mu) \in \mathbb{R}^2,\ \lambda a_n + \mu a'_n = o(b_n), n \in I, n \to +\infty$.

2.8.2.2 Transform Let $(a_n)_{n\in I}$ and $(b_n)_{n\in I}$ be two sequence in $\mathbb{R}$ If $a_n = o(b_n), n \in I, n \to +\infty$, then $a_n = O(b_n), n \in I, n \to +\infty$.

2.8.2.3 Transition Let $(a_n)_{n\in I}, (b_n)_{n\in I}$ and $(c_n)_{n\in I}$ be elements in $\mathbb{R}^I$:

- If $a_n = O(b_n)$ and $b_n = O(c_n), n \in I, n \to +\infty$, then $a_n = O(c_n), n \in I, n \to +\infty$.
- If $a_n = O(b_n)$ and $b_n = o(c_n), n \in I, n \to +\infty$, then $a_n = o(c_n), n \in I, n \to +\infty$.
- If $a_n = o(b_n)$ and $b_n = O(c_n), n \in I, n \to +\infty$, then $a_n = o(c_n), n \in I, n \to +\infty$.

2.8.2.4 Times Let $(a_n)_{n\in I}, (b_n)_{n\in I}, (c_n)_{n\in I}, (d_n)_{n\in I}$ be sequences in $\mathbb{R}$:

- If $a - N = O(b_n), c_n = O(d_n), n \in I, n \to +\infty$, then $a_n c_n = O(b_n d_n), n \in I, n \to +\infty$.
- If $a - N = o(b_n), c_n = O(d_n), n \in I, n \to +\infty$, then $a_n c_n = o(b_n d_n), n \in I, n \to +\infty$.

2.8.3 On the Limit

Let $(a_n)_{n\in I}, (b_n)_{n\in I}$ be elements of $\mathbb{R}^I$ that converges to $l \in \mathbb{R}$ and $l' \in \mathbb{R}$ respectively. Then:

- $(a_n + b_n)_{n\in I}$ converges to $l + l'$.
- $(a_n b_n)_{n\in I}$ converges to ll'.

2.8.3.1 Prop Let $a \in \mathbb{R}$. Then $a^n = o(n!) \quad n \to +\infty$.

2.8.3.2 Proof Let $N \in \mathbb{N}$ s.t. $|a| < N$. For $n \in \mathbb{N}$ s.t. $n \geq N$

$$0 \leq \frac{|a^n|}{n!} = \frac{|a^N|}{N!} \cdot \frac{|a^{n-N}|N!}{n!} \leq \frac{|a^N|}{N!}(\frac{|a|}{N})^{n-N}$$

And $0 < \frac{|a|}{N} < 1 \to \lim_{n\to+\infty} (\frac{|a|}{N})^n = 0$. Therefore: $\lim_{n\to+\infty} \frac{|a^n|}{n!} = 0$, namely $a^n = o(n!)$.

2.8.3.3 Prop $n! = o(n^n) \quad n \to +\infty$

2.8.3.4 Proof Let $N \in \mathbb{N}_{\geq 1}$, $0 \leq \frac{n!}{n^n} \leq \frac{1}{n} \to \lim_{n\to+\infty} \frac{n!}{n^n} = 0$.

2.8.3.5 Prop Let $(a_n)_{n\in I}, (b_n)_{n\in I}$ be the elements of $\mathbb{R}^I$. If the series $\sum_{n\in I} b_n$ converges absolutely and if $a_n = O(b_n) \quad n \to +\infty$. Then $\sum_{n\in I} a_n$ converges absolutely.

2.8.3.6 Proof By Def, $\sum_{n\in I}|b_N| < +\infty$. If $|a_N| \leq M|b_N|$ for $n \in I, n \geq N$ where $N \in \mathbb{N}$, then:

$$\sum_{n\in I}|a_n| = \sum_{n\in I, n<N}|a_n| + \sum_{n\in I, n\geq N}|a_n| \leq \sum_{n\in I, n<N}|a_n| + \sum_{n\in I, n\geq N}|b_n| < +\infty.$$

2.8.4 Theorem: d'Alembort Ratio Test

Let $(a_N)_{n\in\mathbb{N}} \in (\mathbb{R}\backslash\{0\})^{\mathbb{N}}$

- If $\limsup_{n\to+\infty}|\frac{a_{n+1}}{a_n}| < 1$, then $\sum_{n\in\mathbb{N}} a_n$ converges absolutely.
- If $\liminf_{n\to+\infty}|\frac{a_{n+1}}{a_n}| > 1$, then $\sum_{n\in\mathbb{N}} a_n$ does not converge (diverges).

2.8.4.1 Proof

1. Let $\alpha \in \mathbb{R}$ s.t. $\limsup_{n\to+\infty}|\frac{a_{n+1}}{a_n}| < \alpha < 1$, α isn't a lower bound of $(\sup_{n\geq N}|\frac{a_{n+1}}{a_n}|)_{N\in\mathbb{N}}$.

   So $\exists N \in \mathbb{N}$ s.t. $\sup_{n\geq N}|\frac{a_{n+1}}{a_n}| < \alpha$. Hence for $n \geq N$ $|a_n| \leq \alpha^{n-N}|a_N|$, since $\frac{a_n}{a_N} = \frac{a_{N+1}}{a_N}\frac{a_{N+2}}{a_{N+1}}...\frac{a_n}{a_{n-1}}$. Therefore $a_n = O(\alpha^n)$ since $\sum_{n\in\mathbb{N}} = \frac{1}{1-\alpha} < +\infty$, $\sum_{n\in\mathbb{N}} a_n$ converge absolutely.

   Lemma If a series $\sum_{n\in\mathbb{N}} a_n \in \mathbb{R}$ converges absolutely, then $\lim_{n\to+\infty} a_n = 0$.

   Proof If $(\sum_{i=0}^{n} a_i)_{n\in\mathbb{N}}$ converges to some $l \in \mathbb{R}$, then $(\sum_{i=0}^{n-1} a_i)_{n\in\mathbb{N},n\geq 1}$ converges to l, too. Hence $\left(a_n = \left(\sum_{i=0}^{n} a_i\right) - \left(\sum_{i=0}^{n-1} a_i\right)\right)_n$ converges to l-l=0.

2. Let $\beta \in \mathbb{R}$ s.t. $1 < \beta < \liminf_{n\to+\infty}|\frac{a_{n+1}}{a_n}| = \sup_{N\in\mathbb{N}}\inf_{n\geq N}|\frac{a_{n+1}}{a_n}|$.

   So there exists $N \in \mathbb{N}$ s.t. $\beta < \inf_{n\geq N}|\frac{a_{n+1}}{a_n}|$.

   $\forall n \in \mathbb{N}, n \geq N \quad |\frac{a_{n+1}}{a_n}| \geq \beta$.

   Hence $(|a_n|)_{n\in\mathbb{N}}$ is not bounded since $|a_n| \geq \beta^{n-N}|a_n|$.

   By the lemma: $\sum_{n\in\mathbb{N}} a_n$ diverges.

2.8.4.2 Prop Let $a \in \mathbb{R}, a > 1$. Then $n = o(a^n), n \to +\infty$.

2.8.4.3 Proof Let $\varepsilon > 0$ s.t. $a = (1+\varepsilon)^2$. $a^n = (1+\varepsilon)^{2n} = (1+\varepsilon)^n(1+\varepsilon)^n \geq (1+n\varepsilon)(1+n\varepsilon) \geq \varepsilon^2 n^2$. Hence $n \leq \frac{a^n}{\varepsilon^2 n} = o(a^n)$.

2.8.4.4 Corollary Let $a > 1, t \in \mathbb{R}_{\geq 0}$ Then $n^t = o(a^n), n \to +\infty$

2.8.4.5 Proof Let $d \in \mathbb{N}_{\geq 1}$ s.t. $t \leq d$Then $n^{t-d} \leq 1$. So $n^t = n^d n^{t-d} = O(n^d)$. Let $b = \sqrt[d]{a} > 1$, $n^d = o((b^n)^d) = o(a^n)$. Hence $n^t = o(a^n)$.

2.8.4.6 Corollary There exists $M \geq 1$ s.t. $\forall x \in \mathbb{R}, x \geq M, \ln(x) \leq x$.

2.8.4.7 Proof Let $a \in \mathbb{R}$ s.t. $1 < a < e$.

2.8.5 Theorem: Cauchy Root Test

Let $(a_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}$. Let $\alpha = \limsup_{n\to+\infty}|a_n|^{\frac{1}{n}}$:

- If $\alpha < 1$, then $\sum_{n\in\mathbb{N}} a_n$ converges absolutely.
- If $a > 1$ then $\sum_{n\in\mathbb{N}} a_n$ diverges.

2.8.5.1 Proof

1. Let $\beta \in \mathbb{R}, \alpha < \beta < 1$. There exists $N \in \mathbb{N}$ s.t. $|a_N|^{\frac{1}{n}} \leq \beta$ for $n \geq N$. That means $|a_n| = O(\beta^n)$ since $0 < \beta < 1$, $\sum_{n\in\mathbb{N}} a_n$ converges absolutely.
2. If $\alpha > 1$ then $\forall N \in \mathbb{N} \quad \exists n \geq N$ s.t. $|a_n|^{\frac{1}{n}} \geq 1$, since otherwise $\exists N \in \mathbb{N}\ \forall n \geq N, |a_n|^{\frac{1}{n}} < 1$ contradiction. Hence $(|a_n|)_{n\in\mathbb{N}}$ cannot converge to 0.

2.8.6 Leibniz's Criterion

Let $\sum_{n=1}^{\infty}(-1)^{n+1}a_n$ be an alternating series where $a_n > 0$. The series converges if:

1. a_n is monotonic decreasing, i.e., $a_{n+1} \leq a_n$ for all n,

2. $\lim_{n\to\infty} a_n = 0$.

If these conditions are satisfied, the series converges.

Chapter 3

Topology

3.1 Absolute Value

3.1.0.1 Def Let K be a field. By absolute value on K, we mean a mapping $|\cdot| : K \to \mathbb{R}_{\geq 0}$ that satisfies:

1. $\forall a \in K \quad |a| = 0$ iff $a = 0$.
2. $\forall (a, b) \in K^2 \quad |ab| = |a| \cdot |b|$.
3. $\forall (a, b) \in K^2 \quad |a + b| \leq |a| + |b|$(triangle inequality). Or we can say $||x| - |y|| \leq |x - y|$.

3.1.1 Bernoulli's Inequality

Bernoulli's inequality states that for all $x \geq -1$ and $n \in \mathbb{N}_+$,

$$(1 + x)^n \geq 1 + nx.$$

3.1.1.1 Proof We will prove this inequality using mathematical induction on n.
For $n = 1$, we have

$$(1 + x)^1 = 1 + x.$$

This is exactly $1 + x \geq 1 + 1 \cdot x$, so the base case holds.
Assume that the inequality holds for some positive integer $n = k$, i.e.,

$$(1 + x)^k \geq 1 + kx.$$

We need to show that the inequality holds for $n = k + 1$, i.e.,

$$(1 + x)^{k+1} \geq 1 + (k + 1)x.$$

To prove this, start with the left-hand side:

$$(1 + x)^{k+1} = (1 + x)^k \cdot (1 + x).$$

Using the induction hypothesis, we know that $(1 + x)^k \geq 1 + kx$. Thus,

$$(1 + x)^{k+1} = (1 + x)^k \cdot (1 + x) \geq (1 + kx)(1 + x).$$

Expanding the right-hand side, we get

$$(1 + kx)(1 + x) = 1 + x + kx + kx^2 = 1 + (k + 1)x + kx^2.$$

Since $x^2 \geq 0$ for $x \geq -1$, we have $kx^2 \geq 0$, and therefore,

$$1 + (k + 1)x + kx^2 \geq 1 + (k + 1)x.$$

Thus,

$$(1 + x)^{k+1} \geq 1 + (k + 1)x.$$

This completes the inductive step.

3.1.2 Arithmetic-Geometric Mean Inequality (On real number's field)

The Arithmetic-Geometric Mean Inequality states that for any non-negative real numbers $a_1, a_2, \ldots, a_n$,

$$\frac{\sum_{i=1}^{n} a_i}{n} \geq \sqrt[n]{\prod_{i=1}^{n} a_i}.$$

Equality holds if and only if $a_1 = a_2 = \cdots = a_n$.

3.1.2.1 Example Trivial absolute value:

$$|a|_0 = \begin{cases} 0, & if\ a = 0 \\ 1, & if\ a \neq 0 \end{cases}$$

3.1.2.2 Notation Take a prime number p. $\forall \alpha \in \mathbb{Q}\backslash\{0\}$ there exists an integer s.t. $\frac{\alpha}{ord_p(\alpha)} = \frac{a}{b}$, where $\begin{cases} a \in \mathbb{Z}\backslash\{0\} \\ b \in \mathbb{N}\backslash\{0\} \end{cases}$ $, p\|ida, p\|idb,$ which is unique.

3.1.2.3 Prop $|\cdot| : \begin{array}{l} \mathbb{Q} \to \mathbb{R}_{\geq 0} \\ \alpha \mapsto \begin{cases} p^{-ord_p(\alpha)} & if\ \alpha \neq 0 \\ 0 & if\ \alpha = 0 \end{cases} \end{array}$ is a absolute value on $\mathbb{Q}$.

3.1.2.4 Proof

(1) Obviously.

(2) If $\alpha = p^{ord_p(\alpha)}\frac{a}{b}, \beta = p^{ord_p(\beta)}\frac{c}{d}$ $p\|idabcd.\alpha\beta = p^{ord_p(\alpha)+ord_p(\beta)}\frac{ac}{bd}$ $p\|idac, p\|idbd$

(3) $\alpha + \beta = p^{ord_p(\alpha)}\frac{a}{b} + p^{ord_p(\beta)}\frac{c}{d}$. Assume $ord_p(\alpha) \geq ord_p(\beta)$, $\alpha + \beta = p^{ord_p(\beta)}\left(p^{ord_p(\alpha)-ord_p(\beta)}\frac{a}{b} + \frac{c}{d}\right)$
$= p^{ord_p(\beta)}\frac{p^{ord_p(\alpha)-ord_p(\beta)}ad+bc}{bd}$ $p\|idbd$. So $ord_p(\alpha + \beta) \geq ord(\beta)$.
Hence $ord_p(\alpha+\beta) \geq \min\{ord_p(\alpha), ord_p(\beta)\}$. So $|\alpha+\beta|_p = p^{-ord_p(\alpha+\beta)} \leq \max\{p^{-ord_p(\alpha)}, p^{-ord_p(\beta)}\} = \max\{|\alpha|_p, |\alpha|_p\} \leq |\alpha|_p, |\alpha|_p$.

3.2 Quotient Structure

3.2.1 Quotient Set

Let X be a set and $\sim$ be an equivalence relation on X.

For any $x \in X$, we denote by $[x]$ the set $\{y \in X \mid y \sim x\}$ and call it the *equivalence class* of x. Let $X/\sim$ be the set $\{[x] \mid x \in X\}$, the name of it is *quotient set.*

3.2.1.1 Prop Let X be a set and $\sim$ be an equivalence relation on X.

1. For any $x \in X$ and any $y \in [x]$, one has $[x] = [y]$.
2. If α and β are elements of $X/\sim$ s.t. $\alpha \neq \beta$, then $\alpha \bigcap \beta = \emptyset$.
3. X=$\bigcup_{\alpha \in X/\sim} \alpha$.

3.2.1.2 Proof

1. Let $z \in [y]$. Then $y \sim z$. Since $y \in [x]$, one has $x \sim y$. Therefore, $x \sim z$ namely $z \in [x]$. This proves $[y] \subseteq [x]$.
   Moreover, since $x \sim y$, one has $x \in [y]$. For the same reason, $[x] \subseteq [y]$. Thus we obtain $[x] = [y]$.
2. Suppose that $\alpha \bigcap \beta \neq \emptyset$ and $y \in \alpha \bigcap \beta$. By (1), $\alpha = \beta = [y]$. This leads to a contradiction.
3. For any $x \in X$, $x \in [x]$. Hence $x \in \bigcup_{\alpha \in X/\sim} \alpha$. Hence $X \subseteq \bigcup_{\alpha \in X/\sim} \alpha$. Conversely, for any $\alpha \in X/\sim$, α is a subset of X. Hence $\bigcup_{\alpha \in X/\sim} \alpha \subseteq X$.

3.2.2 Left/Right Action

Let G be a group and X be a set.

We call left action of G on X any mapping $G \times X \to X \quad (g, x) \mapsto gx$ that satisfies $\forall x \in X, 1x = x$, $\forall (g, h) \in G^2, x \in X, g(hx) = (gh)x$

We call right action of G on X any mapping $G \times X \to X \quad (g, x) \mapsto xg$ that satisfies $\forall x \in X, x1 = x$, $\forall (g, h) \in G^2, x \in X, x(gh) = (xg)h$.

3.2.2.1 Remark If we denote by G^{op} the set G equipped with the composition law $G \times G \to G \quad (g, h) \mapsto hg$, then a right action of G on X is just a left action of G^{op} on X.

3.2.2.2 Prop Let G be a group and X be a set. Assume given a left action of G on X, Then the binary relation $\sim$ on X defined as $x \sim y$ iff $\exists g \in G$, y=gx is an equivalence relation.

For any $x \in X$, the equivalence class of x is denoted as Gx or $orb_G(x)$, called the *orbit* of x under the action of G.

3.2.2.3 Proof

1. $\forall x \in X$, x=1x, so $x \sim x$.
2. $\forall (x, y) \in X^2$, if y=gx for some $g \in G$, then $g^{-1}y = g^{-1}(gx) = (g^{-1}g)x = 1x = x$.
3. $\forall (x, y, z) \in X^3$, if $\exists (g, h) \in G^2$ s.t. y=gx and z=hy, then z=h(gx)=(hg)x. So $x \sim z$.

3.2.2.4 Notation We denote by $G\backslash X$ the set $X/\sim$.

3.2.3 Projection Mapping

Let X be a set and $\sim$ be an equivalence relation, the mapping $X \to X/\sim$, $(x \in X) \mapsto [x]$ is called the projection mapping.

3.2.3.1 Example Let G be a group and H be a subgroup of G(non-empty subset s.t. $\forall (x, y) \in H^2,\ xy^{-1} \in H$).
Then the mapping $H \times G \to G$ is a left action of H on G. Thus we obtain two quotient sets $H\backslash G$ and G/H.

3.2.4 Normal Subgroup

Let G be a group and H be a subgroup of G. If $\forall g \in G,\ h \in H,\ ghg^{-1} \in H$, then we say that H is a normal subgroup of G.

3.2.4.1 Remark

1. For any $g \in G$, gH⊆Hg provided that H is a normal subgroup of G.
   In fact, $\forall h \in H$, $\exists h' \in H$ s.t. $ghg^{-1} = h'$. Hence gh=h'g, this shows $gH \subseteq Hg$.
   $\exists h''$ s.t. $g^{-1}hg = h''$. Hence hg=gh". This shows $Hg \subseteq gH$.
2. If G is commutative, any subgroup of G is normal.

3.2.4.2 Theorem Let G be a group and H be a normal subgroup of G.
Then the mapping $G/H \times G/H \to G/H$ is well defined and determines a structure of group on the quotient set G/H.
Moreover, the projection mapping $\pi : G \to G/H \quad x \mapsto xH$ is a morphism of groups.

3.2.4.3 Proof

1. If xH=x'H and yH=y'H, then $\exists h_1 \in H$ and $h_2 \in H$ s.t. x'=xh_1, y'=yh_2. Hence x'y'=xh_1yh_2=(xy)$(y^{-1}h_1y)h_2$. Therefore $(x'y')H = (xy)H$. The mapping is well defined.
2. $\forall (x, y, z) \in G^3$, $(xH)(yH \times zH) = xH((yz)H) = (x(yz))H = (xH \times yH)zH$.
   $\forall x \in G$, $1H \times xH = xH \times 1H = xH,\ x^{-1}H \times xH = 1H$.
3. $\pi(xy) = (xy)H = xH \times yH = \pi(x)\pi(y)$.

3.2.4.4 Prop Let K be a unitary ring, E be a left K-module and F be a sub K-module. Then the mapping $K \times (E/F) \to E/F$ is well defined, and defines a left K-module $(a, [x]) \mapsto [ax]$ structure on E/F. Moreover, the projection mapping $\pi : E \to E/F$ is a morphism of left K-modules.

3.2.4.5 Proof Let x and x' be elements of E s.t. [x]=[x'], that means $x' - x \in F$. Hence $a(x' - x) = ax' - ax \in F$. So [ax]=[ax'].
Let we check that E/F forms a left K-module:

- $a([x] + [y]) = a([x + y]) = [a(x + y)] = [ax + ay] = [ax] + [ay]$.
- $(a + b)[x] = [(a + b)x] = [ax + bx] = [ax] + [bx]$.
- 1[x]=[1x]=[x].
- a(b[x])=a[bx]=[a(bx)]=[(ab)x]=(ab)[x].

By the previous Prop, π is a morphism of groups. Moreover, $\forall x \in E,\ a \in K,\ \pi(ax) = a\pi(x)$.

3.2.5 Two-sided Ideal

Let A be a unitary ring. We call two-sided ideal any subgroup I of (A,+) that satisfies the following condition: $\forall x \in I,\ a \in A,\ \{ax, xa\} \subseteq I$. (I is a left and right sub K-module of A)

3.2.5.1 Theorem Let A be a unitary ring and I be a two sided ideal of A. The mapping $(A/I)\times(A/I) \to A/I \quad ([a],[b]) \mapsto [ab]$ is well defined. Moreover, A/I becomes a unitary ring under the addition and this composition law, and the projection mapping $A \overset{\pi}{\to} A/I$ is a morphism of unitary ring.

3.2.5.2 Proof If $a' \sim a$, $b' \sim b$, it means $a'-a \in I$, $b'-b \in I$, then $a'b'-ab = a'b'-a'b+a'b-ab = a'(b'-b)+(a'-a)b \in I$, Therefore $a'b' \sim ab$. $\pi(ab) = \pi(a)\pi(b)$, $\pi(1)$ is the unity.

3.2.5.3 Example Let $d \in \mathbb{Z}$ and $d\mathbb{Z} = \{n \in \mathbb{Z} \mid \exists m \in \mathbb{Z}, n = dm\}$, $d\mathbb{Z}$ is a two sided ideal of $\mathbb{Z}$.

If $m \in \mathbb{Z}$, for any $a \in \mathbb{Z}$, adm=dma$\in d\mathbb{Z}$. Denote by $\mathbb{Z}/d\mathbb{Z}$ the quotient ring. The class of $n \in \mathbb{Z}$ in $\mathbb{Z}/d\mathbb{Z}$ is called the residu class of n modulo d.

If A is a commutative unitary ring, a two sided ideal of A is simply called an *ideal* of A.

3.2.5.4 Theorem Let $f : G \to H$ be a morphism of groups.

1. Im(f) is a subgroup of H.
2. Ker(f) is a normal subgroup of G.
3. The mapping $\widetilde{f} : G/Ker(f) \to Im(f) \quad [x] \to f(x)$ is well defined and is an isomorphism of groups.

3.2.5.5 Proof

1. Let α and β be elements of Im(f). Let (x,y)$\in G^2$ s.t. $\alpha = f(x), \beta = f(y)$. Then $\alpha\beta^{-1} \in Im(f)$. So Im($f$) is a subgroup.
2. Let x and y be elements of Ker(f). One has $f(xy^{-1}) = f(x)f(y^{-1}) = 1_H$, so $xy^{-1} \in Ker(f)$. Hence Ker(f) is a subgroup of G. Let $x \in Ker(f)$ and $y \in G$. One has $f(yxy^{-1}) = f(y)f(x)f(y)^{-1} = 1_H$. Hence $yxy^{-1} \in Ker(f)$. So Ker(f) is a normal subgroup.
3. If $x \sim y$ then $\exists z \in$ Ker(f) s.t. y=xz. Hence $f(y) = f(x)f(z) = f(x)$. So $\widetilde{f}$ is well defined. Moreover $\widetilde{f}([x][y]) = \widetilde{f}([xy]) = f(xy) = f(x)f(y) = \widetilde{f}([x])\widetilde{f}([y])$. Hence $\widetilde{f}$ is a morphism of groups.

   By Def, Im($\widetilde{f}$) = Im(f). If x and y are elements of G s.t. f(x)=f(y), then $f(xy^{-1}) = 1_H$. Hence $xy^{-1} \in Ker(f)$. Since $x = (xy^{-1})y$, $x \sim y$, that means $[x] = [y]$. Therefore $\widetilde{f}$ is injective.

3.2.5.6 Theorem Let K be a unitary ring and $f : E \to F$ be a morphism of left-K-modules.

1. Im(f) is a left sub-K-module of F.
2. Ker(f) is a left sub-K-module of E.
3. $\tilde{f} : E/\text{Ker}(f) \to \text{Im}(f)$, $[x] \mapsto f(x)$ is an isomorphism of left K-module.

3.2.5.7 Proof

1. Let $a \in K$, $x \in E$, $af(x) = f(ax) \in \text{Im}(f)$.
2. Let $a \in K$, $x \in$Ker(f), $f(ax) = af(x) = a0 = 0$.
3. Let $a \in K$, $x \in E$, $\tilde{f}(a[x]) = \tilde{f}([ax]) = f(ax) = af(x) = a\tilde{f}([x])$.

3.3 Topology

3.3.1 Topological Space

Here I skipped many things mentioned in the class "Geometry and Topology".

3.3.1.1 Remark While the union of the collection of topologies may not form a topology, the intersection of any collection of topologies do forms a topology.

3.3.2 Metric Space

We define (X, d) a metric space, s.t.

1. d(x,y)=0 iff x=y.
2. d(x,y)+d(y,z)$\geq$d(x,z).
3. d(x,y)=d(y,x).

3.3.2.1 Isometry isometry=distance preserving + bijective.

3.3.2.2 Prop Strongly equivalent metrics are topologically equivalent, and conversely, topologically equivalent metrics are strongly equivalent.

3.3.2.3 Example Let (X_i, d_i), $i \in \{1, \dots, n\}$ be a family of metric spaces. Let X be the product set $X_1 \times \cdots \times X_n$. We call $d(x, y) := \max_{i\in\{1,\dots,n\}} d_i(x_i, y_i)$ the *product metric* of $d_1, \dots, d_n$.

3.3.3 Axiom of Choice

Given a family of non-empty sets $\{A_i\}_{i\in I}$, where I is an index set, there exists a function f (called a *choice function*) s.t. for each $i \in I$, we have $f(i) \in A_i$, meaning that an element is chosen from each set A_i.

3.3.4 Theorem: Zorn's Lemma

If $\forall A \subseteq X$ that is totally ordered with respect to $\leq$, there exists a maximal element x_0 of X.

3.3.4.1 Remark Let (X,$\leq$) be a well ordered set, $y \notin X$. We extend $\leq$ to $X \bigcup \{y\}$ s.t. $\forall x \in X, x < y$. Then ($X \bigcup \{y\}$,$\leq$) is well ordered.

3.3.5 Initial Segment

Let (X,$\leq$) be a well ordered set, $S \subseteq X$, if $\forall s \in S$ and $x \in X$, $x < s$ implies $x \in S$, then we say that S is an initial segment of X.

If S is an initial segment s.t. $S \neq X$, then we say that S is a *proper initial segment.*

3.3.5.1 Prop Let $(X, \leq)$ be a well ordered set, if $(S_i)_{i\in I}$ is a family of initial segments of X, then $\bigcup_{i\in I} S_i$ is an initial segment of X.

3.3.5.2 Proof $\forall s \in \bigcup_{i\in I} S_i$, $\exists i \in I$ s.t. $s \in S_i$. Therefore $X_{<s} \subseteq S_i \subseteq \bigcup_{j\in I} S_j$.

3.3.5.3 Prop Let (X,$\leq$) be a well ordered set.

1. Let S be a proper initial segment of X, x=min($X\backslash S$), then $S = X_{<x}$.
2. $X \to \mathscr{P}(X) \quad x \mapsto X_{<x}$ is strictly increasing.
3. The set of all initial segments of X forms a well ordered subset of $(\mathscr{P}(X), \subseteq)$.

3.3.5.4 Proof of the Third Prop Let $\mathcal{F} \subseteq \mathscr{P}(X)$ be a set of initial segments s.t. $\mathcal{F} \neq \emptyset$. Then there exists $A \subseteq X$ s.t. $\mathcal{F}\backslash\{x\} = \{X_{<x} \mid x \in A\}$. If $A = \emptyset$ then $\mathcal{F} = \{X\}$, and X is the least element of $\mathcal{F}$. Otherwise $A \neq \emptyset$ and also have a least element a. Then by (2) $X_{<a}$ is the least element of $\mathcal{F}$.

3.3.5.5 Lemma Let $(X, \leq)$ be a well ordered set, $f : X \to X$ be a strictly increasing mapping. Then $\forall x \in X$, $x \leq f(x)$.

3.3.5.6 Prop Let $(X, \leq)$ be a well ordered set, S and T be two initial segments of X. If $f : S \to T$ is a bijection that is strictly increasing, then S=T and $f = Id_S$.

3.3.5.7 Proof We may assume $T \subseteq S$. Let $\iota : T \to S \quad t \mapsto t$ be the inclusion mapping and g=$\iota \circ f$, $S \to S$.

Since g is strictly increasing, by the lemma, $\forall s \in S, s \leq g(s) = f(s) \in T$. Since T is an initial segment, $s \in T$. Hence S=T.

Apply the lemma to f^{-1}, we get $\forall s \in S$, $s \leq f^{-1}(s)$. Hence $f(s) \leq s$, therefore f(s)=s.

3.3.6 Isomophic

Let (X,$\leq$), (Y,$\leq$) be partially ordered sets.

If there exists $f : X \to Y$ that is both increasing and a bijection, we say that (X,$\leq$), (Y,$\leq$) are isomophic.

3.3.6.1 Def of $\preceq$ / $\prec$ Let (X,$\leq$), (Y,$\leq$) be well ordered sets. If (X,$\leq$) is isomorphic to an initial segment of (Y,$\leq$), we denote $X \preceq Y$.

If X is isomorphic to Y, we note $X \sim Y$. If $x \preceq y$ but $X \nsim Y$, we note $X \prec Y$.

3.3.6.2 Prop Let X,Y be well ordered sets. Among the following conditions, one and only one holds: $X \prec Y$, $X \sim Y$, $Y \prec X$.

3.3.6.3 Proof

1. We construct a correspondence $f : X \to Y$ s.t. $(x, y) \in \Gamma_f$ iff $X_{<x} \sim Y_{<y}$. By the last prop of Oct. 11, f is a function.

   If $(a, b) \in \text{Dom}(f) \times \text{Dom}(f)$, $a < b$, then $X_{<a} \subset X_{<b}$. By Def, $Y_{<f(b)} \sim X_{<b}$, $Y_{<f(a)} \sim X_{<a}$. Hence $Y_{<f(a)}$ is isomorphic to a proper initial segment of $Y_{<f(b)}$. Therefore $Y_{<f(a)}$ is a proper initial segment of $Y_{<f(b)}$. We then get $f(a) < f(b)$. Thus f is strictly increasing.

2. Let $a \in \text{Dom}(f)$. Let $x \in X$, $x < a$. Then $X_{<x}$ is a proper initial segment of $X_{<a} \sim Y_{<f(a)}$. Hence $\exists y \in Y_{<f(a)}$, $X_{<x} \sim Y_{<y}$. This shows that $x \in$ Dom. Hence $\text{Dom}(f)$ is an initial segment of X. Applying this to f^{-1}, we get: $\text{Im}(f) = \text{Dom}(f^{-1})$ is an initial segment of Y.

3. To prove that either Dom(f)=X, or Im(f)=Y. We assume that $x \in X \backslash \text{Dom}(f)$, $y \in Y \backslash \text{Im}(f)$ are respectively the least elements of $X \backslash Dom(f)$ and $Y \backslash \text{Im}(f)$. Then we get $\text{Dom}(f) = X_{<x}$, $\text{Im}(f) = Y_{<y}$. We obtain $X_{<x} \sim Y_{<y}$, $(x, y) \in \Gamma_f$. Contradiction!

Cases

1. $\text{Dom}(f) = X$, $\text{Im}(f) \subset Y$, $X \prec Y$.

2. $\text{Dom}(f) \subset X$, $\text{Im}(f) = Y$, $Y \prec X$.

3. Dom(f)=X, Im(f)=Y, $X \sim Y$.

3.3.6.4 Lemma

Let $(X, \leq)$ be a partially ordered set. $\mathscr{S} \subseteq \mathscr{P}(X)$. Assume that:

1. $\forall A \in \mathscr{S}$, $(A, \leq)$ is a well ordered set.

2. $\forall (A, B) \in \mathscr{S} \times \mathscr{S}$, either A is an initial segment of B, or B is an initial segment of A.

Let $Y = \bigcup_{A \in \mathscr{S}} A$. Then $(Y, \leq)$ is a well ordered set, and, $\forall A \in \mathscr{S}$, A is an initial segment of Y.

3.3.6.5 Proof

1. Let $A \in \mathscr{S}$, $x \in A$, $y \in Y$, y<x. Since $Y = \bigcup_{B \in \mathscr{S}} B$, $\exists B \in \mathscr{S}$, s.t. $y \in B$. If $y \notin A$, then $B \nsubseteq A$. Hence A is an initial segment of B. Hence $y \in A$. Contradiction!

2. Let $Z \subseteq Y$, s.t. $Z \neq \emptyset$. Then $\exists A \in \mathscr{S}$, $A \bigcap Z \neq \emptyset$. Let m be the least element of $A \bigcap Z$. Let $z \in Z$ be the least element of Z. Let $B \in \mathscr{S}$ s.t. $z \in B$. If $z \in A$, then $m \leq z$. If $z \notin A$, then A is an initial segment of B.

   Since B is well ordered, $z < m$. Since $m \in A$, we get $z \in A$. Contradiction!

   Therefore, m is the least element of Z.

3.3.7 Proof of Zorn's Lemma

Suppose that X doesn't have any maximal element. Let $\omega = \{\text{well ordered subsets of } X\}$. Let $f : \omega \to X$ s.t. f(A) is an upper bound of $A \in \omega$. If $A \in \omega$ satisfies $\forall a \in A$, a=f($A_{<a}$), we say that A is a f-set.

Let $\mathscr{S} = \{\text{f-sets}\}$. Note that $\emptyset \in \mathscr{S}$. If $A \in \mathscr{S}$, $A \bigcup \{f(A)\} \in \mathscr{S}$. In fact, if $a \in A$, then $A_{<a} = (A \bigcup \{f(A)\})_{<a}$. If $a = f(A) \notin A$, then $(A \bigcup \{f(A)\})_{<a} = A$.

Let A and B be elements of $\mathscr{S}$. Let I be the union of all common initial segments of A and B. This is also a common initial segment of A and B.

If $I \neq A$ and $I \neq B$, then $\exists (a, b) \in A \times B$, $I = A_{<a} = B_{<b}$. $f(I) = f(A_{<a}) = f(B_{<b})$. Hence a=b. Then $I \bigcup \{a\}$ is also a common initial segment of A and B, contradiction!

By the lemma, $Y := \bigcup_{A \in \mathscr{S}} A$ is well ordered, and any $A \in \mathscr{S}$ is an initial segment of Y.

$\forall a \in Y$, $\exists A \in \mathscr{S}$, $a \in A$. Since A is an initial segment of Y, $A_{<a} = Y_{<a}$. Hence $f(Y_{<a}) = f(A_{<a}) = a$. Hence $Y \in \mathscr{S}$. Thus Y is the greatest element of $(\mathscr{S}, \subseteq)$. However, $Y \bigcup \{f(Y)\} \in \mathscr{S}$. Hence f(Y)∈Y. Contradiction!

Errata Suppose that X doesn't have any maximal element. $\forall A \in \omega$, $\exists f(A)$ s.t. $\forall a \in A$, $a < f(A)$.

3.4 Filter

3.4.0.1 Def

Let X be a set. We call *filter* of X any $\mathcal{F} \subseteq \mathscr{P}(X)$ that satisfies

1. $\mathcal{F} \neq \emptyset$, $\emptyset \notin \mathcal{F}$.

2. $\forall A \in \mathcal{F}$, $\forall B \in \mathscr{P}(X)$, if $A \subseteq B$, then $B \in \mathcal{F}$.

3. $\forall (A, B) \in \mathcal{F} \times \mathcal{F}$, $A \bigcap B \in \mathcal{F}$.

3.4.0.2 Example

1. Let $Y \subseteq X$, $Y \neq \emptyset$. $\mathcal{F}_Y := \{A \in \mathscr{P}(X) \mid Y \subseteq A\}$ is a filter, called the *principal filter* of Y.

2. Let X be an infinite set. $\mathcal{F}_{Fr}(X) := \{A \in \mathscr{P}(X) \mid X \backslash A \ is\ finite\}$ is a filter, called the *Frechet filter* of X.

3. Let (X, τ) be a topological space, $x \in X$. We call neighborhood of x any $V \in \mathscr{P}(X)$ s.t.$\exists U \in \tau$, satisfying $x \in U \subseteq V$. Then $\mathcal{V}_x = \{\text{neighborhoods of } x\}$ is a filter.

3.4.1 Filter Basis

Let X be a set, $\mathscr{B} \subseteq \mathscr{P}(X)$. If $\emptyset \notin \mathscr{B}$ and $\forall (B_1, B_2) \in \mathscr{B}^2$, $\exists B \in \mathscr{B}$, s.t. $B \subseteq B_1 \bigcap B_2$, we say that $\mathscr{B}$ is a filter basis.

3.4.2 Generating Filter

If $\mathscr{B}$ is a filter basis, then $\mathcal{F}(\mathscr{B}) := \{A \subseteq X \mid \exists B \in \mathscr{B},\ B \subseteq A\}$ is a filter.

3.4.2.1 Proof

$\emptyset \notin \mathcal{F}(\mathscr{B})$, $\mathcal{F}(\mathscr{B}) \neq \emptyset$ since $\emptyset \neq B \subseteq \mathcal{F}(\mathscr{B})$. If $A \in \mathcal{F}(\mathscr{B})$, $A' \in \mathscr{P}(x)$ s.t. $A \subseteq A'$, then $\exists B \in \mathscr{B}$ s.t. $B \subseteq A \subseteq A'$, hence $A' \in \mathcal{F}(\mathscr{B})$.

If A_1 and A_2 are elements of $\mathcal{F}(\mathscr{B})$, then $\exists (B_1, B_2) \in \mathscr{B}^2$, s.t. $B_1 \subseteq A_1,\ B_2 \subseteq A_2$. Since $\mathscr{B}$ is a filter basis, $\exists B \in \mathscr{B}$ s.t. $B \subseteq B_1 \bigcap B_2 \subseteq A_1 \bigcap A_2$. Hence $A_1 \bigcap A_2$. Hence $A_1 \bigcap A_2 \in \mathcal{F}(\mathscr{B})$.

3.4.3 Neighborhood

$V \in \mathscr{P}(X)$ is called a neighborhood of x if $\exists U \in \mathcal{J}$ s.t. $x \in U \subseteq V$.

3.4.4 Neighborhood Basis

$\mathscr{B}_x$ is a neighborhood basis of x iff

1. $\mathscr{B}_x \subseteq \mathcal{V}_x$.

2. $\forall V \in \mathcal{V}_x$, $\exists U \in \mathscr{B}_x$ s.t. $U \subseteq V$.

3. Let (X,d) be a metric space, $x \in X$. $\forall \varepsilon > 0$, let $\begin{cases} B(x, \varepsilon) = \{y \in X \mid d(x, y) < \varepsilon\} \\ \overline{B}(x, \varepsilon) = \{y \in X \mid d(x, y) \leq \varepsilon\} \end{cases}$.

   Then $\{B(x, \varepsilon) \mid \varepsilon > 0\}$, $\{B(x, \frac{1}{n}) \mid n \in \mathbb{N}_{\geq 1}\}$, $\{\overline{B}(x, \varepsilon) \mid \varepsilon > 0\}$, $\{\overline{B}(x, \frac{1}{n}) \mid n \in \mathbb{N}_{\geq 1}\}$ is a neighborhood basis of x.

For example, $\mathcal{V}_x \bigcap \tau$ is a neighborhood basis of x.

1. Let $Y \subseteq X,\ Y \neq \emptyset$. $\mathscr{B} = \{Y\}$ is a filter basis.

2. Let (X, τ) be a topological space, $x \in X$. If $\mathscr{B}_x$ is a filter basis s.t. $\mathcal{F}(\mathscr{B}_x) = \mathcal{V}_x = \{neighborhood\ of\ x\}$. Then we say that $\mathscr{B}_x$ is a neighborhood basis of x.

3.4.4.1 Remark of a Generating Method

Let (X, τ) be a topological space, $x \in X$ and $\mathscr{B}_x$ a neighborhood basis of x. Suppose that $\mathscr{B}_x$ is countable. We choose a surjective mapping $(B_n)_{n \in \mathbb{N}}$ from $\mathbb{N}$ to $\mathscr{B}_x$.

For any $n \in \mathbb{N}$, let $A_n = B_0 \bigcap B_1 \bigcap \cdots \bigcap B_n \in \mathcal{V}_x$. The sequence $(A_n)_{n \in \mathbb{N}}$ is decreasing and $\{A_n \mid n \in \mathbb{N}\}$ is a neighborhood basis of x.

3.4.4.2 Prop

Let Y and E be sets, $g : Y \to E$ be a mapping.

1. If $\mathcal{F}$ is a filter of Y, then
$$g_*(\mathcal{F}) := \{A \in \mathscr{P}(E) : g^{-1}(A) \in \mathcal{F}\}$$
   is a filter on E.

2. If $\mathscr{B}$ is a filter basis of Y, then
$$g(\mathscr{B}) := \{g(B) : B \in \mathscr{B}\}$$
   is a filter basis of E

3. $$\mathcal{F}(g(\mathscr{B})) = g_*(\mathcal{F}(\mathscr{B})) = \{A \subset E : \exists B \in \mathscr{B}, g^{-1}(A) \supset B\}.$$

3.4.4.3 Proof

1. $E \in g_*(\mathcal{F})$ since $g^{-1}(E) = Y$. $\varnothing \notin g_*(\mathcal{F})$ since $g^{-1}(\varnothing) = \varnothing$.
   If $A \in g_*(\mathcal{F})$ and $A' \supset A$, then $g^{-1}(A') \supset g^{-1}(A) \in \mathcal{F}$, so $g^{-1}(A') \in \mathcal{F}$. Hence $A' \in g_*(\mathcal{F})$.
   If A_1 and A_2 are element of $g_8(\mathcal{F})$. Then $g^{-1}(A_1) \in \mathcal{F}, g^{-1}(A_2) \in \mathcal{F}$. Hence $g^{-1}(A_1 \cap A_2) = g^{-1}(A_1) \cap g^{-1}(A_2) \in \mathcal{F}$. So $A_1 \cap A_2 \in g_*(\mathcal{F})$.

2. Since g is a mapping, and $\varnothing \notin \mathscr{B}$, we get $\varnothing \notin g(\mathscr{B})$. Since $\mathscr{B} \neq \varnothing$, $g(\mathscr{B}) \neq \varnothing$.
   Let $\mathscr{B}_1$ and $\mathscr{B}_2$ be elements of $\mathscr{B}$. There exists $C \in \mathscr{B}$ s.t. $C \subset \mathscr{B}_1 \cap \mathscr{B}_2$. Hence $g(C) \subset g(\mathscr{B}_1) \cap g(\mathscr{B}_2)$, namely $g(\mathscr{B}$ is a filter basis.

3. $g(B) \subseteq A$ iff $B \subseteq g^{-1}(A)$.

3.5 Limit Point and Accumulation Point

We fix a topological space (X, τ).

3.5.1 Def

Let $\mathcal{F}$ be a filter of X and $x \in X$.

1. If $\mathcal{V}_x \subseteq \mathcal{F}$, then we say that x is a *limit point* of $\mathcal{F}$.

2. If $\forall (A, V) \in \mathcal{F} \times \mathcal{V}_x$, $A \bigcap V \neq \emptyset$, we say that x is an *accumulation point* of $\mathcal{F}$.

3.5.1.1 Prop Let B be a filter basis of X, $x \in X$, $\mathscr{B}_x$ a neighborhood basis of x. Then x is an accumulation point of $\mathcal{F}(\mathscr{B})$ iff $\forall (B, U) \in \mathscr{B} \times \mathscr{B}_x$, $B \bigcup U \neq \emptyset$.

3.5.1.2 Proof Since $\mathscr{B} \subseteq \mathcal{F}(\mathscr{B})$, $\mathscr{B}_x \subseteq \mathcal{V}_x$, the *necessity* is true.
Sufficiency: Let (A,V)$\in \mathcal{F}(\mathscr{B}) \times \mathcal{V}_x$. There exist $B \in \mathscr{B}$, $U \in \mathscr{B}_x$, s.t. $B \subseteq A$, $U \subseteq V$. Hence $\emptyset \neq B \bigcap U \subseteq A \bigcup V$.

3.5.2 Closure

3.5.2.1 Def Let $Y \subseteq X$, $Y \neq \emptyset$. We call accumulation point of Y any accumulation point of the principal filter $\mathcal{F}_Y = \{A \subseteq X \mid Y \subseteq A\}$. We denote by $\overline{Y} = \{accumulation\ points\ of\ Y\}$. Note that $x \in \overline{Y}$ iff $\forall U \in \mathscr{B}_x$, $Y \bigcap U \neq \emptyset$. By convention $\overline{\emptyset} := \emptyset$. And we call it *closure.*

3.5.2.2 Prop Let $Y \subseteq X$. Then $\overline{Y}$ is the smallest closed subset of X containing Y.

3.5.2.3 Proof $\forall x \in X \backslash \overline{Y}$,there exists $U_x \in \mathcal{V}_x \bigcap \tau$ s.t. $Y \bigcap U_x = \emptyset$.
Moreover, $\forall y \in U_x$, $U_x \in \mathcal{V}_y \bigcap \tau$. This shows that $\forall y \in U_x$, $y \notin \overline{Y}$. Therefore $X \backslash \overline{Y} = \bigcup_{x \in X \backslash \overline{Y}} U_x \in \tau$.
Let $Z \subseteq X$ be a closed subset that contains Y. Supposed that $\exists y \in \overline{Y} \backslash Z$.
Then U=$X \backslash Z \in \mathcal{V}_y \bigcap \tau$ and $U \bigcap Y \subseteq U \bigcap Z = \emptyset$. So $y \notin \overline{Y}$. Contradiction! Hence $\overline{Y} \subseteq Z$.

3.6 Limits of Mappings

3.6.1 Limit of Filter

Let (E,τ_E) be a topological space, $f : Y \to E$ be a mapping, and $\mathcal{F}$ be a filter of Y. If $a \in E$ is a limit point of $f_*(F)$, namely, $\forall$ neighborhood V of a, $f^{-1}(V) \in \mathcal{F}$, then we say that a is a limit of the filter $\mathcal{F}$ by f.

3.6.1.1 Remark Let (E, τ_E) be a topological space. $f : Y \to E$ be a mapping, and $\mathcal{F}$ be a filter of Y. If $a \in E$, is a limit point of $f_*(\mathcal{F}) = \{A \in E | f^{-1}(A) \in \mathcal{F}\}$, namely, for any neighborhood V of a, $f^{-1}(V) \in \mathcal{F}$, then we say that a is a limit of the filter $\mathcal{F}$ by f.

3.6.1.2 Remark Let $\mathscr{B}_a$ be a neighborhood basis of a. Then $V_a \subseteq f_*(\mathcal{F})$ iff $B_a \subseteq f_*(\mathcal{F})$.
Therefore, a is a limit of $\mathcal{F}$ iff for any $B \in \mathscr{B}_a$, $f^{-1}(B) \in \mathcal{F}$.

3.6.2 Convergent of a Mapping

Let (E, τ) be a topological space. $I \subseteq \mathbb{N}$ be an infinite subset, $x = (x_n)_{n \in I} \in E^I$. If the Frechet filter $\mathcal{F}_{Fr}(I)$ has a limit $a \in E$ by the mapping $x : I \to E$. We say that $(x_n)_{n \in I}$ converges to a, denote as $a = \lim_{n \in I, n \to +\infty} x_n$.

3.6.2.1 Remark $a = \lim_{n\in I, n\to+\infty} x_n$ iff for any $B \in \mathscr{B}_a$, there exists $N \in \mathbb{N}$, s.t. $x_n \in B$ for any $n \in \mathbb{N}_{\geq N}$.
Suppose that τ_E is induced a metric d.
$\{B(x,\varepsilon)|\varepsilon > 0\}$, $\{B(x, \frac{1}{n})|n \in \mathbb{N}_{>0}\}$, $\{\overline{B}(x,\varepsilon)|\varepsilon > 0\}$ and $\{\overline{B}(x, \frac{1}{n})|n \in \mathbb{N}_{>0}\}$ are all neighborhood basis of a.
Therefore, the following are equivalent:

- $a = \lim_{n\in I, n\to+\infty} x_n$
- $\forall\varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall n \in I_{\geq N}$, $d(x_n, a) < \varepsilon$
- $\forall\varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall n \in I_{\geq N}$, $d(x_n, a) \leq \varepsilon$
- $\forall k \in \mathbb{N}_{\geq 1}$, $\exists N \in \mathbb{N}$, $\forall n \in \mathbb{N}_{\geq N}$, $d(x_n, a) < \frac{1}{k}$
- $\forall k \in \mathbb{N}_{\geq 1}$, $\exists N \in \mathbb{N}$, $\forall n \in \mathbb{N}_{\geq N}$, $d(x_n, a) \leq \frac{1}{k}$

3.6.2.2 Remark We consider the metric d on $\mathbb{R}$ defined as $\forall(x, y) \in \mathbb{R}\times\mathbb{R}$, $d(x, y) := |x - y|$.
The topology of $\mathbb{R}$ defined by this metric is called the *usual topology* on $\mathbb{R}$.

3.6.2.3 Theorem Let $(x_n)_{n\in I} \in \mathbb{R}^I$, when $I \subseteq \mathbb{N}$ is an infinite subset. Let $l \in \mathbb{R}$, the following statements are equivalent.

1. The sequence $(x_n)_{n\in I}$ converges to l in the topological space $\mathbb{R}$.
2. $\limsup_{n\in I, n\to+\infty} x_n = \liminf_{n\in I, n\to+\infty} x_n = l$.
3. $\limsup_{n\in I, n\to+\infty} |x_n - l| = 0$.

3.6.2.4 A More Powerful Version Let (X, d) be a metric space, let $I \subseteq \mathbb{N}$ be an infinite subset and $(x_n)_{n\in I} \in X_I$. Let $l \in X$, the following statements are equivalent.

1. $\lim_{n\in I, n\to+\infty} x_n = l$.
2. $\limsup_{n\in I, n\to+\infty} d(x_n, l) = 0$.

3.6.2.5 Proof (1) $\to$ (2). The condition (1) is equivalent to for any $\varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall n \in I_{\geq N}$, $d(x_n, l) < \varepsilon$, hence $\sup_{n\in I_{\geq N}} d(x_n, l) \leq \varepsilon$, therefore, $\limsup_{n\in I, n\to+\infty} d(x_n, l) \leq \varepsilon$, we then obtain $\limsup_{n\in I, n\to+\infty} d(x_n, l) = 0$.
(2) $\to$ (1). If $\inf_{N\in\mathbb{N}} \sup_{n\in I_{\geq N}} d(x_n, l) = 0$, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$, s.t. $\sup_{i\in I_{\geq N}} d(x_n, l) < \varepsilon$. Hence $\forall n \in I_{\geq N}$, $d(x_n, l) < \varepsilon$, since ε is arbitrary, (1) is proved.

3.6.2.6 Prop Let (X, τ) be a topological space. $Y \subseteq X$, $p \in \overline{Y}\backslash Y$. Then $V_{p,Y} = \{V \bigcap Y|V \in V_p\}$ is a filter of Y.

3.6.2.7 Proof Y is not empty, since otherwise $\overline{Y} = \varnothing$.
$X \bigcap Y = Y \in V_{p,Y}$ and $\emptyset \notin V_{p,Y}$ since $p \in \overline{Y}$.
Let $V \in V_p$ and $W \subseteq Y$, s.t. $V \bigcap Y \subseteq W$. Let $U = V \bigcup (W\backslash(V \bigcap Y)) \in V_p$, and $U \bigcap Y = W$, hence $W \in V_{p,Y}$.
Let $(V_1, V_2) \in V_p^2$, $(V_1 \bigcap Y) \bigcap (V_2 \bigcap Y) = (V_1 \bigcap V_2) \bigcap Y \in V_{p,Y}$.

3.6.3 Limit of Mappings

Let (X, τ_x) and (E, τ_E) be topological spaces, $Y \subseteq X$, $p \in \overline{Y}\backslash Y$, and $f : Y \to E$ be a mapping. If a is a limit point of $f_*(V_{p,Y})$, then we say that a is a limit of f, when the variable $y \in Y$ tends to p, denoted as $a = \lim_{y\in Y, y\to p} f(p)$.

3.6.3.1 Remark If $\mathscr{B}_a$ is a neighborhood basis of a, then $a = \lim_{y\in Y, y\to p} f(y)$ is equivalent to for any $U \in \mathscr{B}_a$, there exists $V \in V_p$, s.t. $V \bigcap Y \subseteq f^{-1}(U)$ ($f(V \bigcap Y) \subseteq U$).

3.6.3.2 Theorem Let (X, τ_X) and (E, τ_E) be topological spaces. $Y \subseteq X$, $p \in \overline{Y}\backslash Y$, $a \in E$. We consider the following conditions.
(i) $a = \lim_{y\in Y, y\to p} f(y)$.
(ii) For any $(y_n)_{n\in\mathbb{N}} \in Y^{\mathbb{N}}$, if $\lim_{n\to+\infty} y_n = p$, then $\lim_{n\to+\infty} f(y_n) = a$.
The following statements are true:

- If (i) holds, then (ii) also holds.
- If p has a countable neighborhood basis, then (i) and (ii) are equivalent.

3.6.3.3 Proof

1. For any $U \in V_p$, $\exists N \in \mathbb{N}$, s.t. $\forall n \in \mathbb{N}_{\geq N}$, $y_n \in U \bigcap Y$. Therefore, $V_{p,Y} \subseteq y_*(\mathcal{F}_{Fr}(\mathbb{N}))$. We then get $f_*(V_{p,Y}) \subseteq f_*(y_*(\mathcal{F}_{Fr}(\mathbb{N}))) = (f \circ y)_*(\mathcal{F}_{Fr}(\mathbb{N}))$. (In fact, $f_*(y_*(\mathcal{F}_{Fr}(\mathbb{N}))) = \{F \subseteq E | f^{-1}(F) \subseteq y_*(\mathcal{F}_{Fr}(\mathbb{N}))\} = \{F \subseteq E | y^{-1}(f^{-1}(F)) \subseteq \mathcal{F}_{Fr}(\mathbb{N})\} = \{F \subseteq E | (f \circ y)^{-1}(F)) \subseteq \mathcal{F}_{Fr}(\mathbb{N})\} = (f \circ y)_*(\mathcal{F}_{Fr}(\mathbb{N})))$

   Condition (i) leads $V_a \subseteq f_*(V_{p,Y}) \subseteq (f \circ y)_*(\mathcal{F}_{Fr}(\mathbb{N}))$.

2. Assume that p has a countable neighborhood basis, there exists a decreasing sequence $(V_n)_{n \in \mathbb{N}} \in V_p^{\mathbb{N}}$, s.t. $\{V_n | n \in \mathbb{N}\}$ forms a neighborhood basis of p.

3.7 Continuity

3.7.0.1 Def Let (X, τ_X) and (Y, τ_Y) be topological spaces, f be a function from X to Y, $x \in \mathrm{Dom}(f)$. If, for any neighborhood U of $f(x)$, there exists a neighborhood V of x s.t. $f(V) \subset U$, then we say that f is continuous at x. If f is continuous at any $x \in \mathrm{Dom}(f)$, then we say that f is continuous.

3.7.0.2 Prop If f(x) and g(x) are continuous at x_0, then $f(x) \pm g(x)$, $f(x)g(x)$ are also continuous at x_0.

3.7.0.3 Remark

- Let $\mathscr{B}_{f(x)}$ be a neighborhood basis of $f(x)$. If, $\forall U \in \mathscr{B}_{f(x)}$ there exists $V \in \mathcal{V}_x$ s.t. $f(V) \subset U$. Then we say that f is continuous at x.
- We couldn't write $V \subset f^{-1}(U)$ or "$f^{-1}(U)$ is open" since f may not be a mapping.
- Suppose that X and Y are metric spaces. Then f is continuous at x iff

$$\forall \varepsilon > 0, \exists \delta > 0, \text{ s.t. } \forall y \in \mathrm{Dom}(f), d(y, x) < \delta \text{ implies } d(f(y), f(x)) < \varepsilon.$$

3.7.0.4 Prop of Transitivity Let $(X, \tau_X), (Y, \tau_Y)$ and (Z, τ_Z) be topological spaces, f be a function from X to Y, g be a function from Y to Z. Let $x \in \mathrm{Dom}(g \circ f)$. If f is continuous at x and g is continuous at $f(x)$, then $g \circ f$ is continuous at x.

3.7.0.5 Proof Let $U \in \mathcal{V}_{g(f(x))}$. Since g is continuous at $f(x)$, $\exists W \in \mathcal{V}_{f(x)}, g(W) \subset U$. Since f is continuous at x, $\exists V \in \mathcal{V}_x, f(V) \subset W$. Therefore, $g(f(V)) \subset f(W) \subset U$. Hence $g \circ f$ is continuous at x.

3.7.0.6 Theorem Let $(X, \tau_x), (Y, \tau_Y)$ be topological spaces, f be a function from X to Y. Consider the following conditions

1. f is continuous at x.
2. $\forall (x_n)_{n \in \mathbb{N}} \in \mathrm{Dom}(f)^{\mathbb{N}}$, if $\lim_{n \to +\infty} x_n = x$, then $\lim_{n \to +\infty} f(x_n) = f(x)$.

Then (i) implies (ii). Moreover, if x has a countable neighborhood basis, then (i) and (ii) are equivalent.

3.7.0.7 Proof (i)→(ii). Let $(x_n)_{n \in \mathbb{N}}$ converges to x. For any $U \in \mathcal{V}_{f(x)}, \exists V \in \mathcal{V}_x, f(V) \subset U$. Since $\lim_{n \to +\infty} x_n = x$, there exists $N \in \mathbb{N}$ s.t. $\forall n \in \mathbb{N}_{\geq N}, x_n \in V$. Hence $\forall n \in \mathbb{N}_{\geq N}, f(x_n) \in f(V) \subset U$. Thus $\lim_{n \to +\infty} f(x_n) = f(x)$.
(ii)→(i). Under the hypothesis that x has countable neighbourhood basis. Actually we will prove NOT(i) → NOT(ii).
Let $(V_n)_{n \in \mathbb{N}}$ be a decreasing sequence in $\mathcal{V}_x$ s.t. $\{V_n : n \in \mathbb{N}\}$ forms a neighborhood basis of x. If (i) does not hold, then $\exists U \in \mathcal{V}_{f(x)}, \forall n \in \mathbb{N}, f(V_n) \not\subset U$. Pick $x_n \in V_n$ s.t. $f(x_n) \notin U$. $\forall N \in \mathbb{N}, \forall n \in \mathbb{N}_{\geq N}, x_n \in V_N$. Hence $(x_n)_{n \in \mathbb{N}}$ converges to x. However, $f(x_n) \notin U$ for any n. So $(f(x_n))_{n \in \mathbb{N}}$ does not converges to $f(x)$. Therefore, (ii) does not hold.

3.7.1 Topology Basis

Let (X, τ) be a topological space, $\mathscr{B} \subset \tau$. If any element of τ can br written as the union of a family of sets in $\mathscr{B}$, we say that $\mathscr{B}$ is a topological basis of τ.

3.7.1.1 Prop Let (X, τ) be topological spaces, $\mathscr{B} \subset \tau$. $\mathscr{B}$ is a topological basis iff $\forall x \in X, \mathscr{B}_x := \{V \in \mathscr{B} : x \in V\}$ is a neighborhood basis of x.

3.7.1.2 Proof "→". $\forall x \in X, \mathscr{B}_x \subset \mathcal{V}_x$. Moreover, $\forall U \in \mathcal{V}_x, \exists V \in \tau, x \in V \subset U$. Since $\mathscr{B}$ is a topological basis of $\tau, \exists W \in \mathscr{B}, x \in W \subset V \subset U$. Hence $\mathcal{V}_x$ is generated by $\mathscr{B}_x$.
"⇐". Let $U \in \tau$. $\forall x \in U, U \in \mathcal{V}_x$. So $\exists V_x \in \mathscr{B}_x, x \in V_x \subset U$. Hence $U \subset \bigcup_{x \in U} V_x \subset U$. Hence $U = \bigcup_{x \in U} V_x$.

3.7.1.3 Lemma Let (X, τ) be a topological space, $V \in \mathscr{P}(X)$. Then $V \in \tau$ iff $\forall x \in V$, V is a neighborhood of x.

3.7.1.4 Proof "$\rightarrow$" follows by Def.
"$\Leftarrow$". $\forall x \in V, \exists E_x \in \tau, x \in W_x \subset V$. Hence $V = \bigcup_{x\in V} W_x \in \tau$.

3.7.1.5 Prop Let (X, τ_X) and (Y, τ_Y) be topological spaces. $\mathscr{B}_Y$ be a topological basis of τ_Y, $f : X \to Y$ be a mapping. The following conditions are equivalent:

1. f is continuous.
2. $\forall U \in \tau_Y, f^{-1}(U) \in \tau_X$.
3. $\forall U \in \mathscr{B}_Y, f^{-1}(U) \in \tau_X$.

3.7.1.6 Proof

1. (i)$\rightarrow$(ii). Let $U \in \tau_Y$. $\forall x \in f^{-1}(U), f(x) \in U$. Hence $U \in \mathcal{V}_{f(x)}$. Hence, there exists an open neighborhood W of x s.t. $f(W) \subset U$. Since f is a mapping, $W \subset f^{-1}(U)$. Therefore, $f^{-1}(U) \in \mathcal{V}_x$. Since x is arbitrary, $f^{-1}(U) \in \tau_X$. (By the lemma.)
2. (3)$\rightarrow$(1). For any $U \in \mathscr{B}_Y$ s.t. $f(x) \in U$, $f^{-1}(U)$ is an open neighborhood of x, and $f(f^{-1}(U)) \subset U$.

3.7.1.7 Def Let X be a set, and let $((Y_i, \tau_i))_{i\in I}$ be a family of topological spaces. For any $i \in I$, let $f_i : X \to Y_i$ be a mapping. We call initial topology of $(f_i)_{i\in I}$ on X the smallest topology on X making all f_i continuous.

3.7.1.8 Remark If τ is the initial topology of $(f_i)_{i\in I}$. $\forall i \in I\ \forall U_i \in \tau_i$, $f_i^{-1}(U_i) \in \tau$.
If $\tau \subseteq I$ is a finite subset, $(U_j)_{j\in J} \in \Pi_{j\in J}\tau_j$, then $\bigcap_{j\in J} f_j^{-1}(U_j) \in J$.

3.7.1.9 Prop

$$\mathscr{B} := \left\{ \bigcap_{j\in J} f_j^{-1}(U_j) : J \subset I \text{ finite}, (U_j)_{j\in J} \in \prod_{j\in J} \tau_j \right\}$$

is a topology basis of the initial topology τ.

3.7.1.10 Proof First, $\mathscr{B} \subset \tau$.
Let $\tau\{$subset V of X that can be written as the union of a family of sets in $\mathscr{B}\}$.

- $\phi \in \tau'$. $X \in \mathscr{B} \subset \tau'$.
- τ' is stable by taking the union of any family of elements in τ'.
- If V_1 and V_2 are elements of τ'. Then $V_1 \cap V_2 \in \tau'$. In fact, V_1 and V_2 are of the form of the union of some sets of $\mathscr{B}$. The intersection of two elements of $\mathscr{B}$ is still an element of $\mathscr{B}$

$$\begin{aligned}\left(\bigcap_{j\in J} f_j^{-1}(U_j)\right) \cap \left(\bigcap_{j\in J'} f_j^{-1}(U_j')\right) &= \bigcup_{j\in J\bigcup J} f_j^{-1}(W_j) \\ &= \left(\bigcap_{j\in J\setminus J'} f_j^{-1}(U_j)\right) \cap \left(\bigcap_{j\in J\cap J'} f_j^{-1}(U_j) \cap f_j^{-1}(U_j')\right) \cap \left(\bigcap_{j\in J'\setminus J} f_j^{-1}(U_j')\right)\end{aligned}$$

where

$$W_j = \begin{cases} U_j, & j \in J\setminus J' \\ U_j', & j \in J'\setminus J \\ U_j \cap U_j', & j \in J \cap J'. \end{cases}.$$

So τ' is a topology making all f_i continuous. Hence

$$\tau \subset \tau' \subset \tau \ \rightarrow\ \tau' = \tau.$$

3.7.1.11 Example Let $((Y_i, \tau_i))_{i\in I}$ be topological spaces, $Y = \prod_{i\in I} Y_i$ and $\pi_i : Y \to Y_i$ be the projection mapping. The product topology on Y is by Def the initial topology of $(\pi_i)_{i\in I}$.

3.7.1.12 Theorem Let X be a set, $((Y_i, \tau_i))_{i\in I}$ be a family of topological spaces, $((f_i : X \to Y_i))_{i\in I}$ be a family of mappings, and we equip X with the initial topology τ_X of $(f_i)_{i\in I}$. Let (Z, τ_Z) be a topological space and $h : Z \to X$ be a mapping. Then h is continuous iff $\forall i \in I$, $f_i \circ h$ is continuous.

3.7.1.13 Proof "$\to$". If h is continuous. Since each f_i is continuous, $f_i \circ h$ is also continuous.
"$\Leftarrow$". Suppose that $\forall i \in I$, $f_i \circ h$ is continuous. Hence, $\forall U_i \in \tau_i, (f_i \circ h)^{-1}(U_i) = h^{-1}(f_i^{-1}(U_i)) \in \tau_Z$. Let

$$\mathscr{B} = \left\{ \bigcap_{j\in J} f_j^{-1}(U_j) : J \subset I \text{ finite}, (U_j)_{j\in J} \in \prod_{j\in J} \tau_j \right\}.$$

$\forall U := \bigcap_{j\in J} f_j^{-1}(U_j) \in \mathscr{B}$, $h^{-1}(U) = \bigcap_{j\in J} h^{-1}(f_j^{-1}(J_j)) \in \tau_Z$. Therefore, h is continuous.

3.7.1.14 Remark We keep the notation of the Def of initial topology. If $\forall i \in I, \mathscr{B}_i$ is a topological basis of τ_i, then

$$\mathscr{B} := \left\{ \bigcap_{j\in J} f_j^{-1}(U_j) : J \subset I \text{ finite}, (U_j)_{j\in J} \in \prod_{j\in J} \mathscr{B}_j \right\}$$

is also a topological basis of the initial topology.

3.7.1.15 Example Let $((X_i, d_i))_1^n$ be a family of metric spaces, $X = \prod_{i=1}^n X_i$. We define a mapping

$$\begin{aligned} d : X \times X &\longrightarrow \mathbb{R}_{\geq 0} \\ \left((x_i)_1^n, (y_i)_1^n\right) &\longmapsto \max_{i\in[n]} d_i(x_i, y_i). \end{aligned}$$

d is a metric on X: If $x = (x_i)_1^n, y = (y_i)_1^n, z = (z_i)_1^n$, then

$$\begin{aligned} d(x, z = \max_{i\in[n]} d_i(x_i, z_i) &\leq \max_{i\in[n]}(d_i(x_i, y_i) + d(y_i, z_i)) \\ &\leq d(x, y) + d(y, z). \end{aligned}$$

Each $\begin{matrix} \pi_i : X \to X_i \\ (x_j)_1^n \mapsto x_i \end{matrix}$ is continuous. Hence the product topology τ is contained in τ_d.
Let $x = (x_i)_1^n \in X, \varepsilon > 0$.

$$\begin{aligned} B(x, \varepsilon) &= \left\{ y = (y_i)_1^n : \max_{i\in[n]} d_i(x_i, y_i) < \varepsilon \right\} \\ &= \prod_{i=1}^n B(x_i, \varepsilon) = \bigcap_{i=1}^n \pi_i^{-1}(B(x_i, \varepsilon)) \in \tau. \end{aligned}$$

Hence $\tau_d = \tau$.

3.8 Uniform Continuity and Convergence

3.8.1 Diameter

Let (X, d) be a metric space. $\forall A \subset X$ not empty, we define $\operatorname{diam}(A) := \sup_{(x,y)\in A^2} d(x, y)$, called the diameter of A. If $A = \phi$, by convention $\operatorname{diam}(A) := 0$. If $\operatorname{diam}(A) < +\infty$, we say that A is bounded.

3.8.1.1 Example $\bar{d}(x, y) = \min\{|x - y|, 1\}$ and $\bar{d}(x, y) = \frac{|x-y|}{1+|x-y|}$ are both bounded metrics.

3.8.1.2 Remark

- If A is a finite, then it's bounded.
- If $A \subset B$, then $\operatorname{diam}(A) \leq \operatorname{diam}(B)$.

3.8.1.3 Prop Let (X, d) be a metric space, $A \subset X, B \subset X, (x_0, y_0) \in A \times B$. Then $\operatorname{diam}(A \bigcup B) \leq \operatorname{diam}(A) + d(x_0, y_0) + \operatorname{diam}(B)$. In particular, if A and B are bounded, then $A \bigcup B$ is bounded.

3.8.1.4 Proof Let $(x, y) \in (A \bigcup B)^2$. If $\{x, y\} \subset A$, then $d(x, y) \leq \operatorname{diam}(A)$. If $\{x, y\} \subset B$, then $\operatorname{diam}(B) \geq d(x, y)$.
If $x \in A, y \in B$, then $d(x, y) \leq d(x, x_0) + d(x_0, y_0) + d(y_0, y) \leq \operatorname{diam}(A) + d(x_0, y_0) + \operatorname{diam}(B)$. Similarly when $x \in B, y \in A$.

3.8.1.5 Example $\operatorname{diam}(\overline{B}(x, r)) \leq 2r$. If $(y, z) \in \overline{B}(x, r)$, then $d(y, z) \leq d(y, x) + d(x, z) \leq r + r = 2r$.

3.8.2 Cauchy Sequence

Let (X, d) be a metric space, $I \subset N$ be an infinite subset, $(x_n)_{n\in I} \in X^I$. If

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \operatorname{diam}\big(\{x_n : n \in I_{\geq N}\} \big) \leq \varepsilon,$$

then we say that $(x_n)_{n\in I}$ is a Cauchy sequence.

3.8.2.1 Prop

1. If $(x_n)_{n\in I}$ converges, then it is a Cauchy sequence.
2. If $(x_n)_{n\in I}$ is a Cauchy sequences, $\{x_n : n \in I\}$ is bounded.
3. Suppose that $(x_n)_{n\in I}$ is a Cauchy sequence. If there exists an infinite subset J of I s.t. $(x_n)_{n\in J}$ converges to some $x \in X$, then $(x_n)_{n\in I}$ converges to x.

3.8.2.2 Proof

1. Let l be the limit of $(x_n)_{n\in I}$. $\forall\varepsilon > 0, \exists N \in \mathbb{N}$ s.t. $\{x_n : n \in I_{\geq N}\} \subset \overline{B}(l, \varepsilon/2)$. Hence $\operatorname{diam}(\{x_n : n \in I_{\geq N}\}) \leq \varepsilon$.
2. $\exists N \in \mathbb{N}$ s.t. $\operatorname{diam}(\{x_n : n \in I_{\geq N}\}) \leq 1$. Hence $\operatorname{diam}(\{x_n : n \in I\})$ is finite, since $\{x_n : n \in I\} = \{x_n : n \in I_{<N}\}(\text{finite}) \bigcup \{x_n : n \in I_{\geq N}\}(\text{bounded})$.
3. Let $\varepsilon > 0, \exists N \in \mathbb{N}, \operatorname{diam}(\{x_n : n \in I_{\geq N}\}) \leq \varepsilon/2$.
   Take $n_0 \in J_{\geq N \subset I_{\geq N}}$.
   $$\forall n \in I_{\geq N}, d(x_n, x) \leq d(x_n, x_{n_0}) + d(x_{n_0}, x) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$
   Hence $(x_n)_{n\in I}$ converges to x.

3.8.3 Uniformly Continuous

Let (X, d_X) and (Y, d_Y) be metric spaces, f be a function from X to Y. If $\forall\varepsilon > 0, \exists\delta > 0$ s.t.

$$\forall(x, y) \in (\operatorname{Dom}(f))^2, d(x, y) \leq \delta \text{ implies } d(f(x), f(y)) \leq \varepsilon,$$

namely

$$\inf_{\delta>0} \sup_{d(x,y)\leq\delta} d(f(x), f(y)) = 0,$$

we say f is uniformly continuous.

3.8.3.1 Prop

Let (X, d_X) and (Y, d_Y) be metric spaces, f be a function from X to Y which is uniformly continuous.

1. If $I \subset \mathbb{N}$ is infinite, and $(x_n)_{n\in I}$ is a Cauchy sequence in $\operatorname{Dom}(f)^I$. Then $(f(x_n))_{n\in I}$ is a Cauchy sequence.
2. f is continuous.

3.8.3.2 Proof

1. $\forall\varepsilon > 0, \exists\delta > 0$ s.t. $\forall(x, y) \in \operatorname{Dom}(f)^2, d_X(x, y) \leq \delta \to d_Y(f(x), f(y)) \leq \varepsilon$. Since $(x_n)_{n\in I}$ is a Cauchy sequence, $\exists N \in \mathbb{N}$ s.t. $\forall(n, m) \in I^2_{\geq N}, d_X(x_n, x_m) \leq \delta$. Hence $d_Y(f(x_n), f(x_m)) \leq \varepsilon$. Therefore $(f(x_n))_{n\in I}$ is a Cauchy sequence.
2. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $\operatorname{Dom}(f)^{\mathbb{N}}$ that converges to $x \in \operatorname{Dom}(f)$. We define $(y_n)_{n\in N}$ as
   $$y_n = \begin{cases} x, & 2\|idn \\ x_{n/2}, & 2 \mid n. \end{cases}$$
   Then $(y_n)_{n\mathbb{N}}$ converges to x.
   Hence $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since f is uniformly continuous, $(f(y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. $(f(y_n))_{n\in\mathbb{N},n \text{ is odd}} = (f(x))_{n\in\mathbb{N},n \text{ is odd}}$ converges to $f(x)$.
   This leads to $\lim_{n\to+\infty} f(x_n) = f(x)$. Hence f is continuous at x.

3.8.4 Convergence

Let X be a set, $Z \subset X$, (Y, d) be a metric space, $I \subset \mathbb{N}$ infinite, $(f_n)_{n\in I}$ and f be functions from X to Y, having Z as their common domain of Def.

1. If $\forall x \in Z$, $(f_n(x))_{n\in I}$ converges to $f(x)$, we say that $(f_n)_{n\in I}$ converges pointwisely to f.
2. If $\lim_{n\to+\infty} \sup_{x\in Z} d(f_n(x), f(x)) = 0$, we say that $(f_n)_{n\in I}$ converges uniformly to f.

3.8.4.1 Theorem

Let X and Y be metric spaces, $Z \subset X$, $I \subset \mathbb{N}$ infinite. $(f_n)_{n\in I}$ and f be function from X to Y, having Z as domain of Def. Suppose that

1. $(f_n)_{n\in I}$ converges uniformly to f.
2. each f_n ins uniformly continuous.

Then f is uniformly continuous.

3.8.4.2 Notation $f_n = x^n, f = \mathbb{1}_{\{1\}}$.

3.8.4.3 Proof For $n \in I$, let $A_n = \sup_{x\in Z} d(f_n(x), f(x)), \lim A_n = 0$. $\forall (x,y) \in Z^2, \forall n \in I$,

$$d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) + d(f_n(y), d(y)) \quad \leq 2A_n + d(f_n(x), f_n(y)),$$

$$\inf_{\delta>0} \sup_{d(x,y)\leq\delta} d(f(x), f(y)) \leq 2A_n + \inf_{\delta>0} \sup_{d(x,y)\leq\delta} d(f_n(x), f_n(y)) = 0. \text{ (because } f_n \text{ is uniformly continuous.)}$$

Hence $0 \leq \inf_{\delta>0} \sup_{d(x,y)\leq\delta} d(f_n(x), f_n(y)) \leq 2A_n$. Take $\lim_{n\to+\infty}$, by *Squeeze Theorem*, we get $\inf_{\delta>0} \sup_{d(x,y)\leq\delta} d(f(x), f(y)) = 0$.

3.8.4.4 Theorem Let X be a topological space, Y be a metric space, $Z \subset X, P \in Z$. Let $(f_n)_{n\in I}, f$ functions form X to Y, having Z as domain of Def. Suppose that

1. $(f_n)_{n\in I}$ converges uniformly to f.
2. each f_n is continuous at P.

Then f is continuous at P.

3.8.4.5 Proof $\forall n \in I$, let $A_n = \sup_{x\in Z} d(f_n(x), f(x))$. $\forall \varepsilon, \exists n \in I, A_n \leq \frac{\varepsilon}{3}$. Since f_n is continuous, $\exists U \in \mathcal{V}_p, f_n(U) \subset \overline{B}(f_n(P), \frac{\varepsilon}{3})$.

$$\forall x \in U \cap Z, f(U) \subset \overline{B}(f(P), \varepsilon), \quad d(f(x), f(P)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(P)) + d(f_n(P), f(P)) \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

3.8.5 Epsilon-Lipschitzian

In the study of metric spaces, an isometry is a function $f : X \to Y$ between metric spaces X and Y that preserves distances exactly, i.e., for all $x, y \in X$, we have $d_Y(f(x), f(y)) = d_X(x, y)$.

Isometries play a crucial role in understanding the structure of metric spaces, as they identify spaces that are "essentially the same" in terms of their metric properties.

When an isometry f is also bijective, it is called an *isometric isomorphism*, and the metric spaces X and Y are said to be isometrically isomorphic. Such maps allow for perfect preservation of the metric structure.

However, in many contexts, exact preservation of distances is too restrictive. Instead, we may allow the distances between points to be distorted by a controlled factor. This leads to the concept of an ε-Lipschitzian function.

Let X and Y be metric spaces, f be a function from X to Y, $\varepsilon > 0$. If $\forall x, y, d(f(x), f(y)) \leq \varepsilon d(x, y)$, then we say that f is ε-Lipschitzian.

If $\exists \varepsilon > 0$ s.t. f is ε-Lipschitzian, we say that f is Lipschitzian.

3.8.5.1 Remark If f is Lipschitzian, then it is uniformly continuous.

3.8.5.2 Example

1. Let $((X_i, d_i))_{i\in I}$ be metric spaces, $X \prod_i X_i$ where $i \in I$ finite. $d : X \times X \to \mathbb{R}_{\geq 0}, d((x_i)_{i\in I}, (y_i)_{i\in I}) = \max_{i\in I} d_i(x_i, y_i))$. Then $\pi_i : X \to X_i$ is Lipschitzian. $d_i(x_i, y_i) = d_i(\pi_i(x), \pi_i(y)) \leq d(x, y)$.
2. Let (X, d) be a metric space, $d : X \times X \to \mathbb{R}_{\geq 0}$ is Lipschitzian.

$$|d(x, y) - d(x', y')| \leq 2 \max\{d(x, x'), d(y, y')\}.$$

Chapter 4

Normed Vector Space

4.1 Linear Algebra

Let K be a unitary ring.

4.1.1 Induced Morphism

Let M be a left K-module, and let $x = (x_i)_{i\in I}$ be a family of elements of M.
We define a morphism of left K-module as follows:

$$\varphi_x : K^{\oplus I} \to M,\ (a_i)_{i\in I} \mapsto \sum_{i\in I} a_i x_i$$

Let M be a left K-module , I be a set, $x = (x_i)_{i\in I} \in M^I$.
If φ_x is injective, then we say that $(x_i)_{i\in I}$ is *K-linear independent*.
If this mapping is surjective, then we say that $(x_i)_{i\in I}$ is a *system of generators*.
If this mapping is a bijection, then we say that $(x_i)_{i\in I}$ is a *basis* of M.

4.1.1.1 Example Let e_i be the element $(\delta_{i,j})_{j\in I}$ where $\delta_{i,j} = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$, then the family $e = (e_i)_{i\in I} \in (K^{\oplus I})^I$ is a basis of $K^{\oplus I}$. (In this example $M = K^{\oplus I}$).

4.1.1.2 Proof $\varphi_e((a_i)_{i\in I}) = \sum\limits_{i\in I} a_i e_i = (a_i)_{i\in I}$, so $\varphi_e = \mathrm{Id}_{K^{\oplus I}}$ is a bijection.
Here e_i should be considered as a mapping from I to K that sends $j \in I$ to $\delta_{i,j}$.

4.1.1.3 Remark Let $I = \{1, 2, ..., n\}$, $x = (x_i)_{i\in I} \in M^n$, where $n \in \mathbb{N}$.

- x is linearly independent iff $\forall (a_1, ..., a_n) \in K^n$, $a_1x_1 + ... + a_nx_n = 0$ implies $a_1 = a_2 = ... = a_n = 0$
- x is a system of generators iff any elements of M can be written in the form $b_1x_1 + ... + b_nx_n$, $(b_1, ..., b_n) \in K^n$. Such expression is called a *K-linear combination* of $(x_1, ..., x_n)$.

4.1.2 Free K-Module and Finite Type

Let M be a left K-module.
If M has a basis, we say that M is a free K-module.
If M has a finite system of generators, then we say that M is of finite type.

4.1.2.1 Notation $(x_i)_{i=1}^n$ denotes $(x_i)_{i\in\{1,...,n\}}$.

4.1.3 Supplemented Submodule Theorem

Let K be a unitary ring and V be a left K-module. W be a left sub K-module of V. Let $(x_i)_{i=1}^n$ be an element of W^n. $(\alpha_j)_{j=1}^l \in (V/W)^l$, where $(n, l) \in \mathbb{N}^2$. For any $j \in \{1, \ldots, l\}$, let x_{n+j} be an element in the equivalence class α_j.

If both $(x_i)_{i=1}^n$ and $(\alpha_j)_{j=1}^l$ are linearly independent/system of generators/basis, then $(x_i)_{i=1}^{n+j}$ is also linearly independent/system of generators/basis.

4.1.3.1 Proof

1. **Linearly independent**: Let $\pi : V \to V/W \quad x \mapsto [x]$ be the projection morphism. Suppose that $(b_i)_{i=1}^{n+l}$ s.t. $\sum_{i=1}^{n+l} b_i x_i = 0$. Then $0 = \pi(\sum_{i=1}^{n+l} b_i x_i) = \sum_{i=1}^{n+l} b_i \pi(x_i) = \sum_{j=1}^{l} b_{n+j}\alpha_j$.

   Since $(\alpha_j)_{j=1}^{l}$ is linearly independent, $b_{n+1} = \cdots = b_{n+l} = 0$; since $(x_i)_{i=1}^{n}$ is linearly independent. $b_1 = \cdots = b_n = 0$.

2. **System of generators**: Let $y \in W$, then $\pi(y) \in V/W$. So there exists $(c_{n+1}, \dots, c_{n=l}) \in K^l$, s.t. $[y] = c_{n+1}\alpha_{n+1} + .. + c_{n+1}\alpha_{n+l} = \pi(c_{n+1}x_{n+1} + .. + c_{n+l}x_{n+l})$.

   Hence $y - (c_{n+1}x_{n+1} + .. + c_{n+l}x_{n+l}) \in W$, which means $\exists (c_1, \dots, c_n) \in K^n$, s.t. $y - (c_{n+1}x_{n+1} + .. + c_{n+l}x_{n+l}) = c_1x_1 + .. + c_nx_n$. Therefore, $y = \sum_{i=1}^{n+l} c_i x_i$.

3. **Basis**: Combine the two proofs above.

4.1.4 Steinitz Exchange Theorem

Let K be a division ring (any element of $K\backslash\{0\}$ is invertible). Let V be a left K-module of finite type, and $(x_i)_{i\in\{1,\dots,n\}}$ be a system of generators of V. Then there exists a subset I of $\{1, \dots, n\}$ s.t. $(x_i)_{i\in I}$ forms a basis of V.

4.1.4.1 Proof (By induction on n) If $n = 0$, then $V = \{0\}$. In this case $\emptyset$ is a basis of V. $\varphi_\emptyset : K^\emptyset \to V$.

Induction hypothesis: True for a system of generators of $n-1$ elements.

Let $(x_i)_{i\in\{1,\dots,n\}}$ be a system of generators of V. If $(x_i)_{i\in\{1,\dots,n\}}$ is linearly independent, then it is a basis.

Otherwise, $\exists (a_1, \dots, a_n) \in K^n$, s.t. $(a_1, \dots, a_n) \neq (0, \dots, 0)$ and $a_1x_1 + ... + a_nx_n = 0$. Not loss generality, we suppose $a_n \neq 0$, then $x_n = -a_n^{-1}(a_1x_1 + .. + a_{n-1}x_{n-1})$.

Since $(x_i)_{i\in\{1,\dots,n\}}$ is a system of generator, any elements of V can be written as $b_1x_1 + .. + b_nx_n = b_1x_1 + .. + b_{n-1}x_{n-1} - b_na_n^{-1}(a_1x_1 + .. + a_{n-1}x_{n-1}) = (b_1 - b_na_n^{-1}a_1)x_1 + .. + (b_{n-1} - b_na_n^{-1}a_{n-1})x_{n-1}$

Thus $(x_i)_{i\in\{1,\dots,n-1\}}$ forms a system of generators. By the induction hypothesis, there exists $I \subseteq \{1, \dots, n-1\}$ s.t. $(x_i)_{i\in I}$ forms a basis of V.

4.1.4.2 Corollary Let K be a division ring and V be a left K-module of finite type. If $(x_i)_{i=1}^{n}$ is a linearly independent family of elements of V ($n \in \mathbb{N}$), then there exists $l \in \mathbb{N}$ and $(x_{n+j})_{j=1}^{l} \in V^l$, s.t. $(x_i)_{i=1}^{n+l}$ forms a basis of V.

4.1.4.3 Proof Let W be the image of $\varphi_{(x_i)_{i=1}^n} : K^n \to V$, $(a_i)_{i=1}^{n} \mapsto \sum_{i=1}^{n} a_i x_i$. It is a left sub K-module of V.

Note that $(x_i)_{i=1}^{n}$ forms a basis of W.

Moreover, since V is of finite type, there exists $d \in \mathbb{N}$ and a surjective morphism of left K-modules. $\psi : K^d \to V$. Since the projection morphism $\pi : V \to V/W$ is surjective, the composite morphism $K^d \to V \to V/W$ is surjective. Thus V/W is of finite type. There exists a basis $(\alpha_j)_{j=1}^{l}$ of V/W.

4.1.5 Rank/Dimension of a Left K-module

Let K be a division ring and V be a left K-module of finite type. We call rank of V the minimal number of elements of its basis, denoted as $\mathrm{rk}_K(V)$ or simply $\mathrm{rk}(V)$.

If K is a field(commutative division ring), $\mathrm{rk}(V)$ is also denoted as $\dim(V)$ or $\dim_K(V)$, called the dimension of V.

4.1.5.1 Theorem Let K be a division ring and V be a left K-module of finite type. Let W be a left sub K-module of V.

1. W and V/W are both of finite type, and $\mathrm{rk}(V) = \mathrm{rk}(W) + \mathrm{rk}(V/W)$.
2. Any basis of V has exactly $\mathrm{rk}(V)$ elements.

4.1.5.2 Proof Let $(x_i)_{i=1}^{n}$ be a basis of V, let π be the projection mapping.

1. In $(\pi(x_i))_{i=1}^{n}$ we extract a basis of V/W, say $(\pi(x_i))_{i=1}^{l}$. For $j \in \{l+1, \dots, n\}$, $\exists (b_{j,1}, \dots, b_{j,l}) \in K^l$ s.t. $\pi(x_j) = \sum_{i=1}^{l} b_{j,i}\pi(x_i)$.

   Let $y_j = x_j - \sum_{i=1}^{l} b_{j,i}x_i \in W$, $\pi(y_j) = 0$. For any $x \in W$, $\exists (a_i)_{i=1}^{n} \in K^n$, $x = \sum_{i=1}^{n} a_i x_i = \sum_{i=1}^{l} a_i x_i + \sum_{j=l+1}^{n} a_j(y_j + \sum_{i=1}^{l} b_{j,i}x_i) = \sum_{j=l+1}^{n} a_j y_j + \sum_{i=1}^{l}(a_i + \sum_{j=l+1}^{n} a_j b_{j,i})x_i$.

   Since $\pi(x) = \sum_{i=1}^{l}(a_i + \sum_{j=l+1}^{n} a_j b_{j,i})\pi(x_i) = 0$. So $\forall i \in \{1, \dots, n\}$, $a_i + \sum_{j=l+1}^{n} a_j b_{j,i} = 0$. Hence $x = \sum_{i=l+1}^{n} a_j y_j$. Hence W is of finite type, and $\mathrm{rk}(V) \geq \mathrm{rk}(W) + \mathrm{rk}(V/W)$. Moreover, the previous theorem shows that $\mathrm{rk}(V) \leq \mathrm{rk}(W) + \mathrm{rk}(V/W)$.

2. We reason by induction on $\mathrm{rk}(V)$.

   If $\mathrm{rk}(V) = 0$, in this case $V = \{0\}$. Since $\{0\}$ is not a system of generators, the only basis of V is $\varnothing$. So the statement holds.

   Suppose that there exists $e \in V\backslash\{0\}$ s.t. $V = \{\lambda e | \lambda \in K\}$.

   Then any basis of V is of the form $(ae)_{i\in\{1\}}$ where $a \in K\backslash\{0\}$.

   For all $\lambda \in V$, there exists $(\lambda a^{-1}) \in K$ s.t. $(\lambda a^{-1})ae = \lambda e$. Hence, $\mathrm{rk}(V) = 1$.

   Let $(e_i)_{i=1}^m$ be a basis of V. We reason by induction on m to prove that $m = \mathrm{rk}(V)$. The cases where $m = 0$ or $m = 1$ have been proved respectively.

   Induction hypothesis: True for a basis of fewer than m elements.

   Let $W = \{\lambda e_1 | \lambda \in K\}$. Let π be the projection mapping. Then $(\pi(e_i))_{i=2}^m$ forms a system of generators of V/W.

   If $(a_i)_{i=2}^m \in K^{m-1}$ s.t. $\sum_{i=2}^m a_i\pi(e_i) = 0$, then $\sum_{i=2}^m a_ie_i \in W$. Hence there exists $a_i \in K$ s.t. $\sum_{i=1}^m a_ie_i - a_1e_1 = 0$. Therefore, $a_1 = \cdots = a_m = 0$. Thus, $(\pi(e_i))_{i=2}^m$ is a basis of V/W.

   By the induction hypothesis, $\mathrm{rk}(V) = m - 1 + 1 = m$.

4.1.5.3 Prop If U,V are two left K-modules, $f : U \to V$ is an isomorphism. Then rk(U)=rk(V).

4.1.5.4 Proof Let n=rk(U), m=rk(V). Since there exists $f : K^n \to U$ and $g : K^m \to V$ which are isomorphism, $U \cong V$, so $K^n \cong K^m$, let $f : K^n \to K^m$ be the isomorphism. Since $\dim K^n = \dim K^m + \dim Ker(f) = \dim K^m$, so n=m.

4.1.5.5 Prop Let K be a unitary ring and $f : E \to F$ be a morphism of left K-modules. Let I be a set, and let $x = (x_i)_{i\in I} \in E^I$.

1. If $(x_i)_{i\in I}$ is linearly independent and f is injective, then $(f(e_i))_{i\in I}$ is linearly independent.
2. If $(x_i)_{i\in I}$ is a system of generators and f is surjective, then $(f(e_i))_{i\in I}$ is a system of generators.
3. If $(x_i)_{i\in I}$ is a basis and f is a bijection, then $(f(e_i))_{i\in I}$ is a basis.

4.1.5.6 Proof $\varphi_{(f(e_i))_{i\in I}} = f \circ \varphi_{(e_i)_{i\in I}}$.

4.2 Matrix

We fix a unitary ring K.

4.2.1 Column

Let $n \in \mathbb{N}$ and V be a left K-module. For any $(x_i)_{i=1}^n \in V^n$, we denote by $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ the morphism $\varphi_{(x_i)_{i=1}^n} : K^n \to V$.

4.2.1.1 Example Suppose that $V = K^p (p \in \mathbb{N})$. Then each $x_i \in K^p$ is of the form $(x_{i,1}, x_{i,2}, \ldots, x_{i,p})$. Hence $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ can be written $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{np} \end{pmatrix}$.

4.2.2 Matrix

Let $(n, p) \in \mathbb{N}^2$. We call n by p matrix of coefficient in K any morphism of left K-module form K^n to K^p.

4.2.2.1 Example Denote by I_n the identity mapping $K^n \to K^n$. Then $(e_i)_{i=1}^n$ is a basis of K^n called the *canonical basis* of K^n.

$$\varphi_{(e_i)_{i=1}^n} = \mathrm{Id}_{K^n} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Here, let e_i follows the previous Def.

4.2.2.2 Example Let $(x_1, \ldots, x_n) \in K^n$. Denote by $\mathrm{diag}(x_1, \ldots, x_n) : K^n \to K^n$, $(a_1, \ldots, a_n) \mapsto (a_1 x_1, \ldots, a_n x_n)$.

It looks like: $\mathrm{diag}(x_1, \ldots, x_n) = \begin{pmatrix} x_1 e_1 \\ x_2 e_2 \\ \vdots \\ x_n e_n \end{pmatrix} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{pmatrix}$

4.2.3 Composition Laws on the Matrices

We denote by $M_{n,p}(K)$ the set of all n by p matrices of coefficient in K. For $(n, p, r) \in \mathbb{N}^3$, We define $M_{n,p}(K) \times M_{p,r}(K) \to M_{n,r}(K)$, $(A, B) \mapsto AB := B \circ A$.

4.2.3.1 Example Let V be a left K-module Let $n \in \mathbb{N}$ and $x = (x_1, \ldots, x_n) \in V^n$.

Consider a matrix $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,n} \end{pmatrix} \in M_{p,n}(K)$.

A is a morphism of left K-module from K^p to K^n.

$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is defined as $\varphi_x \circ A : K^p \to K^n \to V$.

Let $(b_1, ..., b_p) \in K^p$, $A((b_1, ..., b_p)) = \sum_{i=1}^{p} b_i(a_{i,1}, \ldots, a_{i,n})$.

$\varphi_x(A(b_1, ..., b_p)) = \varphi_x(\sum_{i=1}^{p} b_i(a_{i,1}, \ldots, a_{i,n})) = \sum_{i=1}^{p} b_i \varphi_x((a_{i,1}, ..., a_{i,n})) = \sum_{i=1}^{p} b_i \sum_{j=1}^{n} a_{i,j} x_j$.

So $A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1,j} x_j \\ \vdots \\ \sum_{j=1}^{n} a_{p,j} x_j \end{pmatrix}$.

Let $B = \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,r} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,r} \end{pmatrix} : K^n \to K^r$;

$$AB = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{p,1} & \cdots & a_{p,n} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{1,2} & \cdots & b_{1,r} \\ b_{2,1} & b_{2,2} & \cdots & b_{2,r} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n,1} & b_{n,2} & \cdots & b_{n,r} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1,i}(b_{i,1}, \ldots, b_{i,r}) \\ \vdots \\ \sum_{i=1}^{n} a_{p,i}(b_{i,1}, \ldots, b_{i,r}) \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} a_{1,i} b_{i,1} & \sum_{i=1}^{n} a_{1,i} b_{i,2} & \cdots & \sum_{i=1}^{n} a_{1,i} b_{i,r} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_{p,i} b_{i,1} & \sum_{i=1}^{n} a_{p,i} b_{i,2} & \cdots & \sum_{i=1}^{n} a_{p,i} b_{i,r} \end{pmatrix} \in M_{p,r}(K)$$

The coefficient at j^{th} line and the k^{th} column of AB is given by $\sum_{i=1}^{n} a_{j,i} b_{i,k}$.

A matrix is in $\mathscr{L}(K^n, K^m)$, and we can write

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_m}{\partial x_1}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_n}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix}$$

to denote the differential of f at p.

4.3 Block Matrix

Matrix computation is a relatively complex operation. To simplify these operations, we introduce block matrices and their operations. Please note that block matrices and their operations are not new types of operations, but a simplified form of matrix operations.

What is a block matrix? Simply put, a block matrix is one where a matrix is divided into smaller matrices along rows and columns. A matrix formed this way is called a "block matrix." For example, let:

$$A = \begin{pmatrix} 1 & 2 & -1 & 0 \\ 2 & 5 & 0 & -2 \\ 3 & -1 & 1 & 3 \end{pmatrix}$$

be a block matrix. If we write

$$A_{11} = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, A_{12} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}, A_{21} = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}, A_{22} = \begin{pmatrix} -1 & 3 \end{pmatrix},$$

then matrix A can be written as:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}.$$

This is a block matrix with 4 submatrices.

In general, for an $m \times n$ matrix A, if we first divide it into r blocks along rows and then into s blocks along columns, we obtain a block matrix of size $r \times s$. It is represented as:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix}.$$

Note that A_{ij} represents the submatrix at the (i, j) position of A. A is also referred to as a block matrix $A = (A_{ij})$, with the condition that these are block matrices.

4.3.1 Composition Laws

The addition, and scalar multiplication of block matrices are trivial.

4.3.1.1 Block Matrix Multiplication Block matrix multiplication is similar to standard matrix multiplication but with additional considerations when dealing with blocks between matrices. For block matrices, the operation must satisfy certain conditions to ensure validity. Let $A = (A_{ij})_{r\times s}$ and $B = (B_{ij})_{s\times t}$ be two block matrices (note: the number of columns of A equals the number of rows of B). We then define:

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{r1} & A_{r2} & \cdots & A_{rs} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{st} \end{pmatrix}.$$

The above block matrices satisfy the following conditions: For matrix A, the block at position (1,1), denoted A_{11}, has the row size m_1 and column size n_1, the block at position (1,2), denoted A_{12}, has row size m_2 and column size n_2, and so on. Similarly, for matrix B, the block at position (i,j), denoted B_{ij}, has row size n_i and column size l_j. This type of block matrix multiplication ensures the validity of the product of block matrices A and B.

If the product $C = A \cdot B$ is:

$$C = \begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1t} \\ C_{21} & C_{22} & \cdots & C_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ C_{r1} & C_{r2} & \cdots & C_{rt} \end{pmatrix},$$

then each block C_{ij} is a matrix of size $m_i \times l_j$ and is computed as:

$$C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{is}B_{sj}.$$

4.3.2 Determinant

4.3.2.1 Lemma Let A and C be matrices of dimensions $m \times n$. Then, the block matrix determinant for the following block structure is:

$$|G| = \begin{vmatrix} A & B \\ O & C \end{vmatrix} = |A| \cdot |C|, \quad H = \begin{vmatrix} A & O \\ B & C \end{vmatrix} = |A| \cdot |C|.$$

4.3.2.2 Block Elementary Transformations Block elementary transformations are similar to ordinary elementary transformations and include three types:

1. Interchange two block rows or two block columns of a block matrix.
2. Multiply a block row of a block matrix by an invertible matrix on the left, or multiply a block column by an invertible matrix on the right.
3. Multiply a block row of a block matrix by a matrix on the left and add it to another block row, or multiply a block column by a matrix on the right and add it to another block column.

4.3.2.3 Theorem Let A be an $m \times m$ invertible matrix, D be an $n \times n$ matrix, B be an $m \times n$ matrix, and C be an $n \times m$ matrix, then

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D - CA^{-1}B|.$$

If D is invertible (at this time A is not necessarily invertible), then

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D||A - BD^{-1}C|.$$

4.3.2.4 Proof Using the third type of block elementary transformation, multiply the first block row by $-CA^{-1}$ and add it to the second block row to obtain

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \to \begin{pmatrix} A & B \\ O & D - CA^{-1}B \end{pmatrix}.$$

The third type of block elementary transformation does not change the value of the determinant, from which the conclusion can be drawn. Another conclusion can be similarly proved.

4.3.2.5 Remark When both A and D are invertible matrices, we get the equation

$$|D||A - BD^{-1}C| = |A||D - CA^{-1}B|.$$

This equation is called the determinant reduction formula. Because when the orders of D and A are not equal, it can be used to reduce the calculation of high-order determinants to the calculation of low-order determinants.

4.4 Transpose

We fix a unitary ring K.

4.4.1 Dual of a Left K-Module

Let E be a left K-module. Denote by $E^\vee := \{\text{morphisms of left K-modules from E to K}\}$.

$\forall (f, g) \in E^\vee$, let $f + g : E \to K$, $x \mapsto f(x) + g(x)$. $(E^\vee, +)$ forms a commutative group.

We define $K \times E^\vee \to E^\vee \quad (a, f) \mapsto f_a \circ f$, where f_a is defined as: $E^\vee \mapsto E^\vee \quad x \mapsto xa$. $\forall x \in E, \lambda \in K$, $(fa)(\lambda x) = f(\lambda x)a = \lambda f(x)a = \lambda(fa)(x)$. This mapping defines a structure of right K-module on $E^\vee$.

4.4.2 Dual of a Morphism of Left K-Modules

Let E and F be two left K-modules, $\varphi : E \to F$ be a morphism of left K-modules. We denote by $\varphi^\vee : F^\vee \to E^\vee$ the morphism of right K-modules sending $g \in F^\vee$ to $g \circ \varphi \in E^\vee$. Actually $\forall a \in K$, $g \in F^\vee$, $f_a(\varphi^\vee(g)) = f_a \circ g \circ \varphi = (f_a \circ g) \circ \varphi = \varphi^\vee(ga)$.

4.4.2.1 Example Suppose that $E = K^n$, $F = K^p$.

$\varphi = \begin{pmatrix} b_{1,1} & \cdots & b_{1,p} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \cdots & b_{n,p} \end{pmatrix}$. Which sends $(a_1, ..., a_n) \in K^n$ to $(\sum_{i=1}^{n} a_i b_{i,1}, ..., \sum_{i=1}^{n} a_i b_{i,p})$.

Let $g \in F^\vee$, thus $g : K^p \to K$. g is of the form $\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}$, $y_i \in K$. $g \circ \varphi$ sends $(a_1, ..., a_n)$ to $\sum_{j=1}^{p}(\sum_{i=1}^{n} a_i b_{i,j}) y_j$.

$$g \circ \varphi = \varphi g = \begin{pmatrix} \sum_{j=1}^{p} b_{1,j} y_j \\ \sum_{j=1}^{p} b_{2,j} y_j \\ \vdots \\ \sum_{j=1}^{p} b_{n,j} y_j \end{pmatrix}.$$

4.4.3 Transpose

Assume that K is commutative. We denote by $\iota_p : (K^p)^\vee \to K^p$, $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \mapsto (x_1, ..., x_p)$. This is an isomorphism of K-module.

For any morphism of K-modules, $\varphi : K^n \to K^p$, we denote by φ^τ the morphism of K-modules $K^p \to K^n$ given by $\iota_n \circ \varphi^\vee \circ \iota_p^{-1}$.

$$\begin{array}{ccc} (K^p)^\vee & \xrightarrow{\varphi^\vee} & (K^n)^\vee \\ \iota_p \downarrow & & \downarrow \iota_n \\ K^p & \xrightarrow[\phi^\tau]{} & K^n \end{array}$$

φ^τ is called the transpose of φ.

Let $(y_1, ..., y_p) \in K^p$, $\iota_p^{-1}((y_1, ..., y_p)) = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$.

$$\varphi^\vee(\begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}) = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \circ \varphi = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^p b_{1,i} y_i \\ \vdots \\ \sum_{i=1}^p b_{n,i} y_i \end{pmatrix}$$

$$\iota_n(\begin{pmatrix} \sum_{i=1}^p b_{1,i} y_i \\ \vdots \\ \sum_{i=1}^p b_{n,i} y_i \end{pmatrix}) = (\sum_{i=1}^p b_{1,i} y_i, ..., \sum_{i=1}^p b_{n,i} y_i) = y_1(b_{1,1}, b_{2,1}, ..., b_{n,1}) + y_2(b_{1,2}, ..., b_{n,2}) + ... + y_p(b_{1,p}, ..., b_{n,p}).$$

Therefore, $\varphi^\tau = \begin{pmatrix} (b_{1,1}, \ldots, b_{n,1}) \\ (b_{1,2}, \ldots, b_{n,2}) \\ \vdots \\ (b_{1,p}, \ldots, b_{n,p}) \end{pmatrix} = \begin{pmatrix} b_{11} & \cdots & b_{n1} \\ \vdots & & \vdots \\ b_{1p} & \cdots & b_{np} \end{pmatrix} \in M_{p,n}(K)$.

4.4.3.1 Prop Let E, F and G be left K-modules. $\varphi : E \to F$ and $\psi : F \to G$ be morphisms of left K-modules. Then $(\psi \circ \varphi)^\vee = \varphi^\vee \circ \psi^\vee$.

4.4.3.2 Proof $\forall f \in G^\vee$, $(\varphi^\vee \circ \psi^\vee)(f) = \varphi^\vee(\psi^\vee(f)) = \varphi^\vee(f \circ \psi) = f \circ \psi \circ \varphi = f \circ (\psi \circ \varphi) = (\psi \circ \varphi)^\vee(f)$.

4.4.3.3 Corollary Assume K is commutative. Let n, p, q be natural numbers. $A \in M_{n,p}(K)$, $B \in M_{p,q}(K)$. Then $(AB)^\tau = B^\tau A^\tau$.

4.4.3.4 Proof $B^\tau A^\tau = A^\tau \circ B^\tau = \iota_n \circ A^\vee \circ \iota_p^{-1} \circ \iota_p \circ B^\vee \circ \iota_q^{-1} = \iota_n \circ (B \circ A)^\vee \circ \iota_q^{-1} = \iota_n \circ (AB)^\vee \circ \iota_q^{-1} = (AB)^\tau$.

4.4.3.5 Remark

1. For $A \in M_{n,p}(K)$, one has $(A^\tau)^\tau = A$.

2. We have a mapping $E \to (E^\vee)^\vee$, $x \mapsto ((f \in E^\vee) \mapsto f(x))$. This is a K-linear mapping.

   If K is a field and E is of finite dimension, this is an isomorphism of K-modules.

   In fact, if $e = (e_i)_{i=1}^n$ is a basis of E over K. For $i \in \{1, ..., n\}$, let $e_i^\vee : E \to K$, $\lambda_1 e_1 + ... + \lambda_n e_n \mapsto \lambda_i$. $e^\vee = (e_i^\vee)_{i=1}^n$ is called the dual basis of e. $(e^\vee)^\vee$ gives a basis of $(E^\vee)^\vee$. Hence $E \to E^\vee$ is an isomorphism.

4.5 Linear Equations

We fix a unitary ring K.

4.5.1 (Reduced) Row Echelon

For $a = (a_1, \dots, a_n) \in K^n \backslash \{(0, \dots, 0)\}$. Denote by $j(a)$ the first index $j \in \{1, ..., n\}$ s.t. $a_j \neq 0$.

Let $(n,p) \in \mathbb{N}^2$, $A \in M_{n,p}(K)$, We write A as a column $A = \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(n)} \end{pmatrix}$, $a^{(i)} = (a_1^{(i)}, \dots, a_p^{(i)}) \in K^p$.

We say that A is of row echelon form if, $\forall i \in \{1, \dots, n-1\}$, one of the following conditions is satisfied:

- $a^{(i+1)} = (0, \dots, 0)$;
- $a^{(i)}$ and $a^{(i+1)}$ are both non-zero, and $j(a^{(i)}) < j(a^{(i+1)})$.

If in addition the following condition is satisfied: $\forall i \in \{1, \dots, n\}$, if $a^{(i)} \neq (0, \dots, 0)$, then $a^{(i)}_{j(a^{(i)})} = 1$ and $\forall s \in \{1, \dots, i-1\}$ $a^{(s)}_{j(a^i)} = 0$, we say that A is of reduced row echelon form.

4.5.1.1 Prop

Suppose that $A = \begin{pmatrix} a^{(1)} \\ \vdots \\ a^{(n)} \end{pmatrix} \in M_{n,p}(K)$ is of row echelon form. Then $\{i \in \{1, \dots, n\} | a^{(i)} \neq (0, \dots, 0)\}$ is of cardinal $\leq p$.

4.5.1.2 Proof

Let $k = \text{Card}\{i \in \{1, \dots, n\} | a^{(i)} \neq (0, \dots, 0)\}$, $a^{(k+1)} = .. = a^{(n)} = (0, \dots, 0)$ and $j(a^{(1)}) < j(a^{(2)}) < .. < j(a^{(k)})$. Hence $\{1, \dots, k\} \to \{1, \dots, p\}$ is injective, so $k \leq p$.

4.5.1.3 Def

Let $A = \begin{pmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{np} \end{pmatrix} \in M_{n,p}(K)$.

Let V be a left K-module and $(b_1, \dots, b_n) \in V^n$.

We consider the equation $A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ (*), which can also be written as $\begin{cases} a_{1,1}x_1 + .. + a_{1,p}x_p = b_1 \\ \cdots \\ a_{n,1}x_1 + .. + a_{n,p}x_p = b_n \end{cases}$

The set of $(x_1, \dots, x_p) \in V^p$ that satisfies (*) is called the solution set of (*).

4.5.1.4 Prop

Suppose that A is of reduced row echelon form.

Let $I(A) = \{i \in \{1, \dots, n\} | (a_{i,1}, \dots, a_{i,p}) \neq (0, \dots, 0)\}$, $J_0(A) = \{1, \dots, p\} \backslash \{j((a_{i,1}, \dots, a_{i,p})) | i \in I(A)\}$.

1. If $\exists i \in \{1, \dots, n\} \backslash I(A)$ s.t. $b_i \neq 0$ then $(*)$ does not have any solution in V^p.
2. Suppose that $\forall i \in \{1, \dots, n\} \backslash I(A)$, $b_i = 0$. Then $(*)$ has at least one solution.

   Moreover, $V^{J_0(A)} \to V^p$, $(z_k)_{k \in J_0(A)} \mapsto (x_1, \dots, x_p)$ with $x_j = \begin{cases} z_j & j \in J_0(A) \\ b_i - \sum\limits_{l \in J_0(A)} a_{i,l} z_l, & j \in j((a_{i,1}, \dots, a_{i,p})), i \in I(A) \end{cases}$

is an injective mapping, whose image is equal to the set of solution of $(*)$.

4.5.1.5 Proof

1. Trivial.
2. Tedious and complex.

4.5.1.6 Prop

Let $m \in \mathbb{N}$ and $S \in M_{m,n}(K)$.

If $(x_1, \dots, x_p) \in V^p$ is a solution of $(*)$: $A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$,

then $(x_1, \dots, x_p)$ is a solution of $(*)_S$: $(SA) \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = S \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

In the case where S is left invertible, namely there exists $R \in M_{n,m}(K)$ s.t. $RS = I_n \in M_{n,n}(K)$. Then $(*)$ abd $(*)_S$ has the same solution set.

4.5.1.7 Example 1 Let $\sigma : \{1,\dots,n\} \to \{1,\dots,n\}$ be a bijection. Let $P_\sigma \in M_{n,n}(K)$, $P_\sigma : K^n \to K^n$, $(\lambda_1,\dots,\lambda_n) \mapsto (\lambda_{\sigma^{-1}(1)},\dots,\lambda_{\sigma^{-1}(n)})$.

$P_{\sigma^{-1}}P_\sigma = P_\sigma \circ P_{\sigma^{-1}} = I_n$.

Let W be a left K-module, $(y_1,\dots,y_n) \in W^n$.

$P_\sigma \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \circ P_\sigma : K^n \to K^n \to W$, $(\lambda_1,\dots,\lambda_n) \mapsto (\lambda_{\sigma^{-1}(1)},\dots,\lambda_{\sigma^{-1}(n)}) \mapsto \lambda_{\sigma^{-1}(1)}y_1 + .. + \lambda_{\sigma^{-1}(n)}y_n = \lambda_1 y_{\sigma(1)} + .. + \lambda_n y_{\sigma(n)}$.

$P_\sigma \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_{\sigma(1)} \\ \vdots \\ y_{\sigma(n)} \end{pmatrix}$, $P_\sigma = P_\sigma I_n = P_\sigma \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$.

4.5.1.8 Example 2 Let $(r_1,\dots,r_n) \in K^n$, suppose that each r_i is left invertible and $s_i \in K$, s.t. $s_i r_i = 1$.

$\mathrm{diag}(r_1,\dots,r_n) : K^n \to K^n$, $(\lambda_1,\dots,\lambda_n) \mapsto (\lambda_1 r_1,\dots,\lambda_n r_n)$, $\mathrm{diag}(s_1,\dots,s_n)\mathrm{diag}(r_1,\dots,r_n) : (\lambda_1,\dots,\lambda_n) \mapsto (\lambda_1 s_1,\dots,\lambda_n s$ $(\lambda_1 s_1 r_1,\dots,\lambda_n s_n r_n) = (\lambda_1,\dots,\lambda_n)$.

$\mathrm{diag}(r_1,\dots,r_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \circ \mathrm{diag}(r_1,\dots,r_n) : K^n \to W$, $(\lambda_1,\dots,\lambda_n) \mapsto (\lambda_1 r_1,\dots,\lambda_n r_n) \mapsto \lambda_1 r_1 y_1 + .. + \lambda_n r_n y_n$.

$\mathrm{diag}(r_1,\dots,r_n) \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} r_1 y_1 \\ \vdots \\ r_n y_n \end{pmatrix}$, $\mathrm{diag}(r_1,\dots,r_n)I_n = \begin{pmatrix} r_1 e_1 \\ \vdots \\ r_n e_n \end{pmatrix}$.

4.5.1.9 Example 3 Let $i \in \{1,\dots,n\}$, $c = (c_1,\dots,c_n) \in K^n$, $c_i = 0$, $S_{i,c} \in M_{n,n}(K)$.

$S_{i,c} : K^n \to K^n$, $(\lambda_1,\dots,\lambda_n) \mapsto (\lambda_1,\dots,\lambda_{i-1},\lambda_i + \sum_{j=1}^n \lambda_j c_j,\lambda_{i+1},\dots,\lambda_n)$.

Since $c_i = 0$, $S_{i,-c}S_{i,c} = I_n$, $S_{i,c} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \circ S_{i,c} : K^n \to W$.

$(\lambda_1,\dots,\lambda_n) \mapsto (\lambda_1,\dots,\lambda_i + \sum_{j=1}^n \lambda_j c_j,\dots,\lambda_n) \mapsto \lambda_1 y_1 + ... + (\lambda_i + \sum_{j=1}^n \lambda_j c_j)y_i + ... + \lambda_n y_n = \lambda_1(y_1 + c_1 y_i) + \lambda_2(y_2 + c_2 y_i) + ... + y_i + ... + \lambda_n(y_n + c_n y_i)$.

$$S_{i,c} \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1 + c_1 y_i \\ \vdots \\ y_i \\ \vdots \\ y_n + c_n y_i \end{pmatrix}.$$

$$S_{i,c}I_n = \begin{pmatrix} 1 & 0 & \cdots & c_1 & \cdots & 0 \\ 0 & 1 & \cdots & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n & \cdots & 1 \end{pmatrix}.$$

4.5.1.10 Def Let $G_n(K)$ be the set of $S \in M_{n,n}(K)$ that can be written as $U_1...U_N$ (by convention, $S = I_n$ when $N = 0$), where each U_i is of one of the following forms.

1. P_σ where $\sigma \in \mathfrak{S}_n = \{\text{bijections from } n \text{ to } n\}$.

2. $\mathrm{diag}(r_1,\dots,r_n)$ where each $r_i \in K$ is left K-module.

3. $S_{i,c}$ with $i \in \{1,\dots,n\}$, $c = (c_1,\dots,c_n) \in K^n$, $c_i = 0$.

Let $p \in \mathbb{N}$, we say that $A \in M_{n,p}(K)$ is reducible by Gauss elimination if $\exists S \in G_N(K)$, s.t. SA is of reduced row echelon form.

4.5.1.11 Theorem Assume that K is a division ring. $\forall (n,p) \in \mathbb{N}^2$, any $A \in M_{n,p}(K)$ is reduced by Gauss elimination.

4.5.1.12 Proof (by induction)

The case when $n = 0$ or $p = 0$ is trivial.

We assume $n \geq 1$, $p \geq 1$. We write A as $\begin{pmatrix} \lambda_1 & \\ \vdots & B \\ \lambda_n & \end{pmatrix}$, $\lambda_i \in K$, $B \in M_{n,p-1}(K)$.

If $\lambda_1 = ... = \lambda_n = 0$, applying the induction hypothesis to B.

(For $S \in G_n(K)$, $SA = \begin{pmatrix} S\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} & SB \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} & SB \end{pmatrix}$).

Suppose that $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$, by permuting the rows, we may assume $\lambda_1 \neq 0$. As K is a division ring, by multiplying the first low λ_1^{-1}, we may assume $\lambda_1 = 1$.

We add $(-\lambda_i)$ times the first row to the i^{th} row, to reduce A to the form $\begin{pmatrix} 1 & \mu_2 & \cdots & \mu_p \\ 0 & & & \\ \vdots & & \boldsymbol{C} & \\ 0 & & & \end{pmatrix}$. $C \in M_{n-1,p-1}(K)$, $(\mu_2, \mu_3, \dots, \mu_p) \in K^{p-1}$.

Applying the induction hypothesis to C, we may assume that C is of reduced row echelon form $\begin{pmatrix} c_2 \\ \vdots \\ c_n \end{pmatrix}$.

For $i \in \{2, \dots, k\}$, we add $-\mu_{j(c_i)}$ times the i^{th} row of A to the first line to obtain a matrix of reduced row echelon.

4.6 Normed Vector Space

4.6.1 Cauchy Sequence

Let (X, d) be a metric space. If $(x_n)_{n\in\mathbb{N}}$ is an element of $X^{\mathbb{N}}$ s.t. $\lim\limits_{N\to+\infty} \sup\limits_{(n,m)\in\mathbb{N}^2_{\geq N}} d(x_n, x_m) = 0$, we say the $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. If any Cauchy sequence in X converges, then we say (X, d) is *complete*.

Let $\mathrm{Cau}(X, d)$ be the set of all Cauchy sequence in X. We define an equivalence relation $\sim$ on $\mathrm{Cau}(X, d)$ as $(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$ iff $\lim\limits_{n\to+\infty} d(x_n, y_n) = 0$. Actually, $\lim\limits_{n\to+\infty} d(x_n, x_n) = 0$; $d(x_n, y_n) = d(y_n, x_n)$; If $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ be elements of $\mathrm{Cau}(X, d)$, $0 \leq d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n) \leq 0$ (Here I omit some "lim" in convenience).

4.6.2 Completion

The completion of (X, d) is defined as $\widehat{X} := \mathrm{Cau}(X, d)/\sim$.

4.6.2.1 Example $k[\![T]\!] = (\widehat{K[T], |\cdot|_T})$.

4.6.2.2 Theorem The mapping $\hat{d} : \hat{X} \times \hat{X} \to \mathbb{R}_{\geq 0}$, $([x], [y]) \mapsto \lim\limits_{n\to+\infty} d(x_n, y_n)$ is well defined, and it is a metric on $\hat{X}$.

4.6.2.3 Proof To check that d is well defined, it suffices to prove that $\forall([x], [y]) \in \hat{X} \times \hat{X}$, $(d(x_n, y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence (since $\mathbb{R}$ is complete) and its limit does not depend on the choice of x and y. For $N \in \mathbb{N}$ and $(n, m) \in \mathbb{N}^2_{\geq N}$, one has $d(x_n, y_n) \leq d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n)$, thus $d(x_n, y_n) - d(x_m, y_m) \leq d(x_n, x_m) + d(y_m, y_n)$, similarly, $d(x_m, y_m) - d(x_n, y_n) \leq d(x_n, x_m) + d(y_m, y_n)$. Therefore, $0 \leq \sup\limits_{(n,m)\in\mathbb{N}^2_{\geq N}} |d(x_n, y_n) - d(x_m, y_m)| \leq \sup\limits_{(n,m)\in\mathbb{N}^2_{\geq N}} d(x_n, x_m) + \sup\limits_{(n,m)\in\mathbb{N}^2_{\geq N}} d(y_n, y_m)$. Taking $\lim\limits_{N\to+\infty}$, we obtain that $(d(x_n, y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence and hence converges in $\mathbb{R}$.

If $x' = (x'_n)_{n\in\mathbb{N}} \in [x]$, $y' \in [y]$, then $\lim\limits_{n\to+\infty} d(x_n, x'_n) = \lim\limits_{n\to+\infty} d(y_n, y'_n) = 0$, $|d(x_n, y_n) - d(x'_n, y'_n)| \leq d(x_n, x'_n) + d(y_n, y'_n)$, taking $\lim\limits_{n\to+\infty}$, we get $\lim\limits_{n\to+\infty} |d(x_n, y_n) - d(x'_n, y'_n)| = 0$. So $\lim\limits_{n\to+\infty} d(x_n, y_n) = \lim\limits_{n\to+\infty} d(x'_n, y'_n)$.

In the following, we check that $\hat{d}$ is a metric.

$\hat{d}([x], [y]) = 0$ iff $[x] = [y]$. If $[x] = [y]$ it is trivial, if $\hat{d}([x], [y]) = 0$, which actually means $x \sim y$, thus $[x] = [y]$. Symmetry is trivial. If $[x]$ $[y]$ and $[z]$ are elements of $\hat{X}$. $d([x], [z]) = \lim\limits_{n\to+\infty} d(x_n, z_n) \leq \lim\limits_{n\to+\infty} d(x_n, y_n) + \lim\limits_{n\to+\infty} d(y_n, z_n) = \hat{d}([x], [y]) + \hat{d}([y], [z])$.

4.6.2.4 Remark $i_X : X \to \hat{X}$, $a \mapsto [(a, a, a, ...)]$. $\hat{d}(i_X(a), i_X(b)) = d(a, b)$. In particular, i_X is injective, if $i_X(a) = i_X(b)$, then $d(a, b) = 0$, hence $a = b$.

4.6.2.5 Prop i_X is dense in $\hat{X}$. (The closure of $i_X(X)$ in $\hat{X}$ is equal to $\hat{X}$).

4.6.2.6 Proof Let $[x]$ be an element of $\hat{X}$. We claim that $[x] = \lim_{n\to+\infty} i_X(x_n)$.

For any $N \in \mathbb{N}$, $0 \le \hat{d}(i_X(x_N), [x]) = \lim_{n\to+\infty} d(x_N, x_n) \le \sup_{(n,m)\in\mathbb{N}^2_{\ge N}} d(x_n, x_m)$, taking $\lim_{N\to+\infty}$, we get $\lim_{N\to+\infty} \hat{d}(i_X(x_N), [x]) = 0$.

4.6.2.7 Theorem $(\hat{X}, \hat{d})$ is a complete metric space.

4.6.2.8 Proof Let $([x^{(N)}])_{N\in\mathbb{N}}$ be a Cauchy sequence in $\hat{X}$, where, for any $N \in \mathbb{N}$, $x^{(N)} = (x_n^{(N)})_{n\in\mathbb{N}}$ is a Cauchy sequence in X.

$\forall \varepsilon > 0$, $\exists N_0 \in \mathbb{N}$, s.t. $\forall (k,l) \in \mathbb{N}^2_{\ge N_0}$, $\hat{d}([x^{(k)}], [x^{(l)}]) = \lim_{n\to+\infty} d(x_n^{(k)}, x_n^{(l)}) \le \varepsilon$.

$\forall N \in \mathbb{N}$, $\exists \alpha(N) \in \mathbb{N}$, $d(x_\mu^{(N)}, x_v^{(N)}) \le \frac{1}{N+1}$ for any $(\mu, v) \in \mathbb{N}^2_{\ge\alpha(N)}$.

Let $y_N = x_{\alpha(N)}^{(N)}$ for any $N \in \mathbb{N}$. Without loss of generality, we assume that $\alpha(0) < \alpha(1) < \ldots$.

Let $\varepsilon > 0$. Take $N_0 \in \mathbb{N}$, s.t. $\hat{d}([x^{(k)}], [x^{(l)}]) < \frac{\varepsilon}{3}$, for any $(k,l) \in \mathbb{N}^2_{\ge N_0}$ and $\frac{1}{N_0+1} < \frac{\varepsilon}{3}$.

Let $(k,l) \in \mathbb{N}^2_{\ge N_0}$, $d(y_k, y_l) = d(x_{\alpha(k)}^{(k)}, x_{\alpha(l)}^{(l)})$. $\alpha(k) \ge N_0$.

$\forall n \in \mathbb{N}_{\ge N_0}$, $d(y_k, y_l) \le d(x_{\alpha(k)}^{(k)}, x_n^{(k)}) + d(x_n^{(k)}, x_n^{(l)}) + d(x_n^{(l)}, x_{\alpha(l)}^{(l)}) \le \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + d(x_n^{(k)}, x_n^{(l)})$

Taking $\lim_{n\to+\infty}$, get $d(y_k, y_l) \le \varepsilon$. So $y = (y_N)_{N\in\mathbb{N}}$ is a Cauchy sequence.

Now we check that $\lim_{N\to+\infty} \hat{d}([x^{(N)}], [y]) = 0$:

$0 \le d(x_n^{(N)}, y_n) \le d(x_n^{(N)}, y_N) + d(y_n, y_N)$, $0 \le \limsup_{N\to+\infty} \lim_{n\to+\infty} d(x_n^{(N)}, y_n) \le \limsup_{N\to+\infty}(\frac{1}{N+1} + \lim_{n\to+\infty} d(y_n, y_N))$
$= \limsup_{N\to+\infty} \lim_{n\to+\infty} d(y_n, y_N) = 0$.

So $\lim_{N\to+\infty} \lim_{n\to+\infty} d(x_n^{(N)}, y_n) = 0$.

4.6.2.9 Example Let $(K, |\cdot|)$ be a valued field. It has an absolute value. There is a metric space with $d(a,b) := |a-b|$, then $\mathrm{Cau}(K)$ forms a commutative unitary ring:

- $(a_n)_{n\in\mathbb{N}} \sim (b_n)_{n\in\mathbb{N}}$ iff $\lim_{n\to+\infty}(a_n - b_n) = 0$.
- $(a_n - b_n)_{n\in\mathbb{N}} \in \mathrm{Cau}_0(K) = \{\text{Cauchy sequences that converges to } 0\}$, which is an ideal of $\mathrm{Cau}(K)$.

Hence $\hat{K} = \mathrm{Cau}(K)/\mathrm{Cau}_0(K)$ is a quotient ring of $\mathrm{Cau}(K)$. Absolute value extends to $\hat{K} : |[(a_n)_{n\in\mathbb{N}}]| = \lim_{n\to+\infty} |a_n|$ that forms an absolute value.

4.7 Norms

Previously, we have defined absolute value $|\cdot|$ on field K.

4.7.1 Def of Semi-norms and Norms

4.7.1.1 Def Let V be a vector space over K. We call a semi-norm on V any mapping $\|\cdot\| : V \to \mathbb{R}_{\ge 0} \quad s \mapsto \|s\|$ s.t.

1. $\forall (a,s) \in K \times V$, $\|as\| = |a| \cdot \|s\|$.
2. $\forall (s,t) \in V \times V$, $\|s+t\| \le \|s\| + \|t\|$.

If in addition $\|\cdot\|$ satisfies $\forall s \in V$, $\|s\| = 0$ iff s=0, we say that $\|\cdot\|$ is a norm and $(V, \|\cdot\|)$ is a normed vector space over K.

4.7.1.2 Example

1. If we consider K as a vector space over K, then $(K, |\cdot|)$ forms a normed vector spaces over K.
2. Let $(V_1, \|\cdot\|_1), \ldots, (V_n, \|\cdot\|_n)$ be vector spaces equipped with semi-norms. Let $V = V_1 \oplus \cdots \oplus V_n = V_1 \times \cdots \times V_n$.
   $\|\cdot\|_{l^\infty} : V \to \mathbb{R}_{\ge 0} \quad (x_1, \ldots, x_n) \mapsto \max_{i\in\{1,\ldots,n\}} \|x_i\|_i$, $\|\cdot\|_{l^p} : V \to \mathbb{R}_{\ge 0} \quad (x_1, \ldots, x_n) \mapsto (\|x_1\|_1^p + \cdots + \|x_n\|_n^p)^{\frac{1}{p}}$.
   There are semi-norms. They are norms if $\|\cdot\|_1, \ldots, \|\cdot\|_n$ are all norms.

4.7.1.3 Def Let $(V, \|\cdot\|)$ be a vector space over K equipped with a semi-norm, and W be a vector subspace of V.

1. The restriction of $\|\cdot\| : V \to \mathbb{R}_{\ge 0}$ to W forms a semi-norm on W. It is a norm if $\|\cdot\|$ is a norm. $\|\cdot\| : W \to \mathbb{R}_{\ge 0}$.
2. The mapping $\|\cdot\|_{V/W} : V/W \to \mathbb{R}_{\ge 0} \quad \alpha \mapsto \inf_{s\in\alpha} \|s\|$.
   Remark $\|[s]\|_{V/W} = \inf_{w\in W} \|s+w\|$ is a semi-norm on V/W.
   Warning Even if $\|\cdot\|$ is a norm, $\|\cdot\|_{V/W}$ might only be a semi-norm.

4.7.1.4 Prop Let $(V, \|\cdot\|)$ be a vector space over K, equipped with a semi-norm. Then N=$\{s \in V \mid \|s\| = 0\}$ forms a vector subspace of V. Moreover, $\|\cdot\|_{V/N}$ is a norm.

4.7.1.5 Proof If $(a,s) \in K \times N$, then $\|as\| = |a| \times \|s\| = 0$, so $as \in N$. If $(s_1, s_2) \in N \times N$, then $0 \le \|s_1 + s_2\| \le \|s_1\| + \|s_2\| = 0$, so $s_1 + s_2 \in N$. So we proved that N forms a vector subspace of V.

$\lambda \in K,\ \alpha \in V/N,\ \|\lambda\alpha\|_{V/N} = \inf_{S\in\alpha} \|\lambda s\| = \inf_{s\in\alpha} |\lambda| \times \|s\| = |\lambda| \times \|\alpha\|_{V/N}$.

$\|\alpha + \beta\|_{V/N} = \inf_{S\in\alpha+\beta} \|s\| = \inf_{(x,y)\in\alpha\times\beta} \|x + y\| \le \inf_{(x,y)\in\alpha\times\beta}(\|x\| + \|y\|) = \|\alpha\|_{V/N} + \|\beta\|_{V/N}$.

Let $\alpha \in V/N$, s.t. $\|\alpha\|_{V/N} = 0$ and we want to prove that $\alpha = [0]$.

Let $s \in \alpha$. $\forall t \in N$, $\|s+t\| \le \|s\|+\|t\| = \|s\| = \|(s+t)+(-t)\| \le \|s+t\|+\|-t\| = \|s+t\|$. $\|\alpha\|_{V/N} = \inf_{t\in N} \|s+t\| = \|s\|$.

Hence $\|\alpha\|_{V/N} = \|s\| = 0$. We obtain that $\alpha = N = [0]$.

4.7.1.6 Def Let $(V, \|\cdot\|)$ be a vector space over K, equipped with a semi-norm. For any $x \in V$, and $r \ge 0$, we denote by $B(x, r)$ the set $\{y \in V | \|y - x\| < r\}$, $\overline{B}(x, r)$ the set $\{y \in V \mid \|y - x\| \le r\}$.

4.7.1.7 Remark If $N = \{s \in V | \|s\| = 0\}$, then $x + N \subseteq \overline{B}(x, r)$, $x + N \subseteq B(x, r)$, when $r > 0$.

$\forall s \in N$, $\|(x + s) - x\| = 0 \le r$, $< r$ when $r > 0$.

We equip V with the topology s.t. $\forall U \subseteq V$, U is open iff $\forall x \in U$, $\exists r_x > 0$, $B(x, r_x) \subseteq U$.

4.7.1.8 Prop Let $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ be vector spaces over K, equipped with semi-norms. Let $f : V_1 \to V_2$ be a K-linear mapping.

1. If f is continuous, $\forall s \in V_1$, if $\|s\|_1 = 0$, then $\|f(s)\|_2 = 0$.

2. If there exists $C > 0$ s.t. $\forall x \in V_1$, $\|f(x)\|_2 \le C\|x\|_1$, then f is continuous.

   The inverse is true when (a) $|\cdot|$ is nontrivial or (b) $V_2/\{y \in V_2 | \|y\|_2 = 0\}$ is of finite type.

4.7.1.9 Proof

1. *Lemma*: If $(V, \|\cdot\|)$ is a vector space over K equipped with a semi-norm, then $N_{\|\cdot\|} := \{s \in V | \|s\| = 0\}$ is closed.

   Proof: Let $s \in V\backslash N_{\|\cdot\|}$. Then $\|s\| > 0$. Let $\varepsilon = \frac{\|s\|}{2}$, $\forall x \in B(s, \varepsilon)$, $\|x\| \ge |\|s\| - \|s - x\|| \ge \|s\| - \varepsilon = \varepsilon > 0$, so $B(s, \varepsilon) \subseteq V\backslash N_{\|\cdot\|}$.

   $f^{-1}(N_{\|\cdot\|_2})$ is closed. Note that $0 \in f^{-1}(N_{\|\cdot\|_2})$, $\overline{\{0\}} \subseteq f^{-1}(N_{\|\cdot\|_2})$.

   $x \in N_{\|\cdot\|_1}$, s.t. $\forall\varepsilon > 0$, $x + N_{\|\cdot\|_1} \subseteq B(x, \varepsilon)$, therefore, $x \in \overline{\{0\}}$.

2. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of V_1 that converges to some $x \in V_1$ (this means $\lim\limits_{n\to+\infty} \|x_n - x\|_1 = 0$).

   Hence $\limsup\limits_{n\to+\infty} \|f(x_n) - f(x)\|_2 = \limsup\limits_{n\to+\infty} \|f(x_n - x)\|_2 \le \limsup\limits_{n\to+\infty} C\|x_n - x\|_1 = C \limsup\limits_{n\to+\infty} \|x_n - x\|_1 = 0$.

   So $(f(x_n))_{n\in\mathbb{N}}$ converges to $f(x)$. Hence f is continuous at x.

   Assume that $|\cdot|$ is non-trivial and f is continuous. Then $f^{-1}(\{y \in V_2 | \|y\|_2 < 1\})$ is an open subset of V_1 containing $0 \in V_1$.

   So there exists $\varepsilon > 0$, s.t. $\{x \in V_1 | \|x\|_1 < \varepsilon\} \subseteq f^{-1}(\{y \in V_2 | \|y\|_2 < 1\})$, namely, $\forall x \in V_1$, if $\|x\|_1 < \varepsilon$, then $\|f(x)\|_2 < 1$

   Since $|\cdot|$ is nontrivial, $\exists a \in K$, ,$0 < |a| < 1$. We prove that $\forall x \in V_1$, $\|f(x)\|_2 \le \frac{1}{\varepsilon|a|}\|x\|_1$.

   If $\|x\|_1 = 0$, by (1) we obtain $\|f(x)\|_2 = 0$, ok.

   Suppose that $\|x\|_1 > 0$, then $\exists n \in \mathbb{Z}$, s.t. $\|a^n x\|_1 = |a|^n\|x\|_1 < \varepsilon \le \|a^{n-1}x\|_1 = |a|^{n-1}\|x\|_1$.

   Thus $|a^n|\|f(x)\|_2 = \|f(a^n x)\|_2 < 1$. Hence $\|f(x)\|_2 < \frac{1}{|a|^n} = \frac{1}{|a|^{n-1}} \times \frac{1}{a} \le \frac{1}{\varepsilon}\|x\|_1 \frac{1}{|a|} = \frac{\|x\|_1}{\varepsilon|a|}$.

4.7.1.10 Def Let $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ be vector spaces over K equipped with semi-norm. We say that a K-linear mapping is bounded if there exists $C > 0$, s.t, $\forall x \in V_1$, $\|f(x)\|_2 \le C\|x\|_1$.

For a general K-linear mapping $f : V_1 \to V_2$ we define

$$\|f\| := \begin{cases} \sup\limits_{x\in V_1 \|x\|>0} \left(\frac{\|f(x)\|_2}{\|x\|_1}\right) & \text{if } f(N_{\|\cdot\|_1}) \subseteq N_{\|\cdot\|_2} \\ +\infty & \text{if } f(N_{\|\cdot\|_1}) \nsubseteq N_{\|\cdot\|_2} \end{cases}$$

f is bounded iff $\|f\| < +\infty$. $\|f\|$ is called the *operator semi-norm* of f.

We denote by $\mathscr{L}(V_1, V_2)$ the set of all bounded K-linear mappings from V_1 to V_2.

4.7.1.11 Prop $\mathscr{L}(V_1, V_2)$ is a vector subspace of $\mathrm{Hom}_K(V_1, V_2)$. Moreover, $\|\cdot\|$ is a semi-norm on $\mathscr{L}(V_1, V_2)$.

4.7.1.12 Proof Let f and g be elements of $\mathscr{L}(V_1, V_2)$.

$\|f+g\| = \sup\limits_{x\in V_1, \|x\|_1\neq 0} \frac{\|f(x)+g(x)\|_2}{\|x\|_1} \leq \sup\limits_{x\in V_1, \|x\|_1\neq 0} \frac{\|f(x)\|_2+\|g(x)\|_2}{\|x\|_1} \leq \sup\limits_{x\in V_1, \|x\|_1\neq 0} \frac{\|f(x)\|_2}{\|x\|_1} + \sup\limits_{x\in V_1, \|x\|_1\neq 0} \frac{\|g(x)\|_2}{\|x\|_1} = \|f\| + \|g\| < +\infty$.

Hence $f + g \in \mathscr{L}(V_1, V_2)$.

Let $\lambda \in K$, $\lambda f : x \mapsto \lambda f(x)$, $\|\lambda f\| = \sup\limits_{x\in V_1, \|x\|_1>0} \frac{\|\lambda f(x)\|_2}{\|x\|_1} = |\lambda|\|f(x)\| < +\infty$.

Here I used the same method as in **4.6.1.5 Proof**.

4.7.1.13 Remark If $\|\cdot\|_2$ is a norm, then $\|\cdot\|$ is a norm.

In fact, let $f \in \mathscr{L}(V_1, V_2)$. Suppose that $\exists x \in V_1$ s.t. $f(x) \neq 0$. Since $f(x) \notin N_{\|\cdot\|_2} = \{0\}$, we obtain $\|x\|_1 \neq 0$ (**4.6.1.8 Prop(2)**).

Thus $\|f\| \geq \frac{\|f(x)\|_2}{\|x\|_1} > 0$. Therefore, $\|\cdot\|$ is a norm.

4.7.2 Banach Space

4.7.2.1 Idea A Banach space $\mathscr{B}$ is both a vector space (over a normed field such as $\mathbb{R}$) and a complete metric space, in a compatible way. Hence a complete normed vector space.

A source of simple Banach spaces comes from considering a Cartesian space $\mathbb{R}^n$ (or K^n where K is the normed field) with the norm:

$$\|(x_1, \ldots, x_n)\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

as we mentioned before.

However, the theory of these spaces is not much more complicated than that of finite-dimensional vector spaces because they all have the same underlying topology. When we look at infinite-dimensional examples, however, things become trickier. Common examples are Lebesgue spaces, Hilbert spaces, and sequence spaces.

In the literature, one most often sees Banach spaces over the field $\mathbb{R}$ of real numbers; Banach spaces over the field $\mathbb{C}$ of complex numbers are not much different, since they are also over $\mathbb{R}$. But people do study them over p-adic numbers too.

4.7.2.2 Def Let $(V, \|\cdot\|)$ be a normed vector space. If V is complete with respect to the metric $V \times V \to \mathbb{R}_{\geq 0}$, $(x, y) \mapsto \|x - y\|$, then we say that $(V, \|\cdot\|)$ is a Banach space.

4.7.2.3 Theorem Let $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ be vector spaces over K, equipped with semi-norms. If $(V_2, \|\cdot\|_2)$ is a Banach space, then $(\mathscr{L}(V_1, V_2), \|\cdot\|)$ is a Banach space.

4.7.2.4 Proof Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathscr{L}(V_1, V_2)$ converges to g. $\forall x \in V_1$, the mapping $(f \in \mathscr{L}(V_1, V_2)) \mapsto f(x)$ is $\|x\|_1$-Lipschitz mapping. $\|f(x) - g(x)\|_2 = \|(f - g)(x)\|_2 \leq \|f - g\|\|x\|_1$

So $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence, that converges to some $g(x) \in V_2$. Thus we obtain a mapping $g : V_1 \to V_2$.

We want to prove that g is an element of $\mathscr{L}(V_1, V_2)$.

$\forall (x, y) \in V_1 \times V_2$, $g(x + y) = \lim\limits_{n\to+\infty} f_n(x + y) = \lim\limits_{n\to+\infty} f_n(x) + \lim\limits_{n\to+\infty} f_n(y) = g(x) + g(y)$.

$\|f_n(x) + f_n(y) - g(x) - g(y)\| \leq \|f_n(x) - g(x)\| + \|f_n(y) - g(y)\| = o(1) + o(1) = o(1)$, $(n \to +\infty)$

$\forall x \in V_1$, $\forall \lambda \in K$, $g(\lambda x) = \lim\limits_{n\to+\infty} f_n(\lambda x) = \lim\limits_{n\to+\infty} \lambda f_n(x)$, $\|\lambda f_n(x) - \lambda g(x)\| = \lambda\|f_n(x) - g(x)\| = o(1)$, $(n \to +\infty)$

So $g(\lambda x) = \lambda g(x)$

$\forall x \in V_1$, $\|g(x)\| = \lim\limits_{n\to+\infty} \|f_n(x)\| \leq \lim\limits_{n\to+\infty} \|f_n\|\|x\|$ (This is because $\forall (a, b) \in V_2^2$, $|\|a\| - \|b\|| \leq \|a - b\|$)

Hence $|\|f_n(x)\| - \|g(x)\|| \leq \|f_n(x) - g(x)\| = o(1)$, $g \in \mathscr{L}(V_1, V_2)$

$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall (n, m) \in \mathbb{N}_{\geq N}, \|f_n - f_m\| \leq \varepsilon, \forall x \in V_1, \|(f_N - f_m)(x)\| \leq \varepsilon\|x\|$, take $\lim\limits_{m\to+\infty}$, we get $\|(f_n - g)(x)\| \leq \varepsilon\|x\|$, so $\|f_n - g\| \leq \varepsilon$, $\forall n \in \mathbb{N}$, $n \geq N$.

In fact, $\|f_n(x) - g(x)\| \leq \|f_n(x) - f_m(x)\| + \|f_m(x) - g(x)\| \leq \varepsilon\|x\| + \|f_m(x) - g(x)\|$.

4.8 Differentiability

In this section, we fix a complete valued field $(K, |\cdot|)$. The absolute value is nontrivial.

4.8.1 Defs

4.8.1.1 Def Let X be a topological space and $p \in X$.

Let K be a complete value field and $(E, \|\cdot\|)$ be a normed vector space over K.

Let $f : X \to E$ be a mapping and $g : X \to \mathbb{R}_{\geq 0}$ be a non-negative mapping.

We say that $f(x) = O(g(x))$, $x \to p$, if there are a neighborhood V of p in X and a constant $C > 0$, s.t. $\|f(x)\| \leq Cg(x)$ for any $x \in V$; $f(x) = o(g(x))$, $x \to p$ if there exist a neighborhood V of p in X and a mapping $\varepsilon : V \to \mathbb{R}_{\geq 0}$, s.t. $\lim\limits_{x\in V, x\to p} \varepsilon(x) = 0$ and $\forall x \in V$, $\|f(x)\| \leq \varepsilon(x)g(x)$.

$\lim\limits_{x\in V, x\to p} \varepsilon(x)$ means $\forall \delta > 0$, $\exists$ open neighborhood U of p, $U \subseteq V$, and $\forall x \in U$, $0 \leq \varepsilon(x) \leq \delta$.

4.8.1.2 Def Let E and F be normed vector space over K, $U \subseteq E$ be an open subset, $f : U \to F$ be a mapping and $p \in U$. If there exists $\varphi \in \mathscr{L}(E, F)$, s.t. $f(x) = f(p) + \varphi(x - p) + o(\|x - p\|)$, $x \to p$. We say that f is differentiable at p, and φ is the differential of f at p.

$f(x) = f(p) + \varphi(x - p) + o(\|x - p\|)$, $x \to p$ means that there exists an open neighborhood V of p with $V \subseteq U$, and a mapping $\varepsilon : V \to \mathbb{R}_{\geq 0}$, s.t. $\lim_{x\to p} \varepsilon(x) = 0$ and that $\|f(x) - f(p) - \varphi(x - p)\| \leq \varepsilon(x)\|x - p\|$, $\forall x \in V$.

4.8.1.3 Prop If f is differentiable at p, then its differential at p is unique.

4.8.1.4 Proof Suppose that there exists φ and ψ in $\mathscr{L}(E, F)$, s.t. $f(x) = f(p) + \varphi(x - p) + o(\|x - p\|)$ and $f(x) = f(p) + \psi(x - p) + o(\|x - p\|)$

$(\varphi - \psi)(x - p) = o(\|x - p\|)$. $\exists \varepsilon : V \to \mathbb{R}_{\geq 0}$. V neighborhood of p, $V \subseteq U$.

$\|\varphi - \psi\| = \sup_{y\in E\setminus\{0\}} \frac{\|(\varphi-\psi)(y)\|}{\|y\|} \leq \sup_{y\in E\setminus\{0\},\|y\|\leq\delta} \frac{\|(\varphi-\psi)(y)\|}{\|y\|}$

$\lambda \neq 0$, $\frac{\|(\varphi-\psi)(\lambda y)\|}{\|\lambda y\|} = \frac{\|(\varphi-\psi)(y)\|}{\|y\|}$

Therefore, $\|\varphi - \psi\| = \inf_{\delta>0} \sup_{y\in E,0<\|y-p\|\leq\delta} \frac{\|(\varphi-\psi)(y-p)\|}{\|y-p\|} \leq \inf_{\delta>0} \sup_{y\in E,0<\|y-p\|\leq\delta} \varepsilon(y) = \limsup_{y\to p} \varepsilon(y) = 0$

Hence $\varphi = \psi$.

4.8.1.5 Def Suppose that f is differentiable at p. We denote by d_pf the differential of f at p.

4.8.2 Zero Mapping

$f : U \to F$, $f(x) = y_0$. $\forall x \in U$, $\forall p \in U$, $f(x) - f(p) = 0 = 0 + o(\|x - p\|)$. Hence $d_pf(x) = 0$, $\forall x \in E$.

4.8.3 Linear Mapping

Let $f \in \mathscr{L}(E, F)$. $f(x) - f(p) = f(x - p)$, hence $d_pf = f$.

4.8.4 Addition Mapping

$A : E \times E \to E$, $(x, y) \mapsto x + y$. $\|(x, y)\|_{l^1} = \|x\| + \|y\|$. $\|x + y\| \leq \|x\| + \|y\| = \|(x, y)\|_{l^1} \leq 2\|(x, y)\|_{l^\infty}$. $\forall (p, q) \in E^2$, $d_{(p,q)}A = A$, $d_{(p,q)}A(x, y) = A(x, y) = x + y$.

4.8.5 Scalar Multiplication Mapping

$m : K \times E \to E$, $(\lambda, x) \mapsto \lambda x$. Let $(a, p) \in K \times E$, $\lambda x - ap = \lambda x - ax + ax - ap = (\lambda - a)x + a(x - p) = (\lambda - a)p + a(x - p) + (\lambda - a)(x - p)$. $\|(\lambda - a)(x - p)\| = |\lambda - a|\|x - p\| \leq \max\{|\lambda - a|, \|x - p\|\}^2 = o(\max\{|\lambda - a|, \|x - p\|\})$. When $(\lambda, x) \to (a, p)$, $\|(\lambda - a, x - p)\|_{l^\infty} \to 0$, the mapping $((\mu, y) \in K \times E) \mapsto \mu p + ay \in E$ is a bounded K-linear mapping, because $(\mu_1 + \mu_2)p + a(y_1 + y_2) = (\mu_1 p + ay_1) + (\mu_2 p + ay_2)$, $b\mu p + a(by) = b(\mu p + ay)$, $\|\mu p + ay\| \leq |\mu|\|p\| + |a|\|y\| \leq \max\{|\mu|, \|y\|\}(|a| + \|p\|) = \|(\mu, y)\|_{l^\infty}(|a| + \|p\|)$. Thus this mapping is an element of $\mathscr{L}(K \times E, E)$. Hence m is differentiable, and $d_{(a,p)}m(\mu, y) = \mu p + ay$, $\forall (\mu, y) \in K \times E$, $d_{(a,p)}(m)(\lambda - a, x - p) = (\lambda - a)p + a(x - p)$.

4.8.6 Theorem(Chain Rule Differentials)

Let E, F and G be three normed vector space, $U \subseteq E$, $V \subseteq F$ be open subsets.

Let $f : U \to F$ and $g : V \to G$ be mappings s.t. $f(U) \subseteq V$. Let $p \in U$. Assume that f is differentiable at p and g is differentiable at $f(p)$. Then $g \circ f$ is differentiable at p and $d_p(g \circ f) = d_{f(p)}g \circ d_pf$.

4.8.6.1 Proof Let $x \in U$. By Def, $f(x) = f(p) + d_pf(x-p) + o(\|x-p\|)$, $f(x) - f(p) = O(\|x-p\|)$. $(g \circ f)(x) = g(f(x)) = g(f(p)) + d_{f(p)}g(f(x) - f(p)) + o(\|f(x) - f(p)\|) = g(f(p)) + d_{f(p)}g(f(x) - f(p)) + o(\|x - p\|) = g(f(p)) + d_{f(p)}g(d_pf(x - p) + o(\|x - p\|)) + o(\|x - p\|) = g(f(p)) + d_{f(p)}g(d_pf(x - p)) + o(\|x - p\|)$

So $g \circ f$ is differentiable at p, and $d(g \circ f) = d_{f(p)}g \circ d_pf$.

4.8.6.2 Prop Let n be a positive integer, $E, F_1, \cdots, F_n$ be normed vector spaces over K. $U \subseteq E$ an open subset, $p \in U$. For any $i \in \{1, \cdots, n\}$, let $f_i : U \to F_i$ be a mapping. Let $f : U \to F = F_1 \times \cdots \times F_n$ be the mapping that sends $x \in U$ to $(f_1(x), \cdots, f_n(x))$. We equip F with the norm $\|\cdot\|$ defined as follows: $\|(y_1, \cdots, y_n)\| = \max_{i\in\{1,\cdots,n\}} \|y_i\|$.

Then f is differentiable at p iff each f_i is differentiable at p. Moreover, when this happens, one has $\forall x \in E$, $d_pf(x) = (d_pf_1(x), \cdots, d_pf_n(x))$.

4.8.6.3 Proof Suppose $\forall i \in \{1, \cdots, n\}$, f_i is differentiable at p. $f(x) - f(p) = (f_1(x) - f_1(p), \cdots, f_n(x) - f_n(p)) = (d_pf_1(x - p), \cdots, d_pf_n(x - p)) + o(\|x - p\|)$

Therefore, f is differentiable at p and $d_pf(\cdot) = (d_pf_1(\cdot), \cdots, d_pf_n(\cdot))$.

$\pi_i : F \to F_i$, $(x_1, \cdots, x_n) \mapsto x_i$ is a bounded linear mapping. One has $\|\pi_i\| \leq 1$.

π_i is then differentiable at $f(p)$. Hence $\pi_i \circ f = f_i$ is differentiable at p.

4.8.6.4 Def Let U be an open subset of K, and $(F, \|cd)$ be a normed vector space over K. If $f : U \to F$ is a mapping that differentiable at some $p \in U$. We denote by $f'(p)$ the element $d_p f(1) \in F$, called the derivative of f at p.

4.8.6.5 Corollary Let U and V be open subsets of K, $(F, \|cd)$ be a normed vector space, $f : U \to K$ and $g : V \to F$ be mappings s.t. $f(U) \subseteq V$. Let $p \in U$. If f is differentiable at p and g is differentiable at $f(p)$, then $(g \circ f)'(p) = f'(p)g'(f(p))$.

4.8.6.6 Proof By Def,
$d_p(g \circ f)(1) = d_{f(p)}g(d_p(f)(1)) = d_{f(p)}g(f'(p)) = d_{f(p)}g(f'(p)1) = f'(p)d_{f(p)}g(1) = f'(p)g(1) = f'(p)g'(f(p))$.

4.8.7 Corollary(Leibniz rule)

Let E and F be normed vector spaces over K, $U \subseteq E$ an open subset, $f : U \to K$ and $g : U \to F$ be mappings, and $p \in U$. If both f and g are differentiable at p, then $fg : U \to F$, $x \mapsto f(x)g(x)$ is also differentiable at p, and $\forall l \in E$, $d_p(fg)(l) = d_p f(l)g(p) + f(p)d_p g(l)$.

4.8.7.1 Proof Consider $m : K \times F \to F$, $(a, y) \mapsto ay$. We have shown that m is differentiable, and $d_{(a,y)}m(b, z) = by + az$
fg is the following composite, $U \to K \times F \to F$, $x \mapsto (f(x), g(x)) \mapsto f(x)g(x)$.
$d_p(fg)(l) = d_p(m \circ h)(l) = d_{h(p)}m(d_p h(l)) = d_{(f(p),g(p))}m(d_p f(l), d_p g(l)) = f(p)d_p g(l) + d_p f(l)g(p)$.

4.8.7.2 Corollary Let U be an open subset of K, f and g be mappings from U to K, and to a normed vector space F respectively. If f and g are differentiable at $p \in U$, then
$(fg)'(p) = f'(p)g(p) + f(p)g'(p)$

4.8.7.3 Proof $(fg)'(p) = d_p(fg)(1) = d_p f(1)g(p) + f(p)d_p g(1) = f'(p)g(p) + f(p)g'(p)$.

4.8.7.4 Example $f_n : K \to K$, $x \mapsto x^n$ is differentiable at any $x \in K$. $f_n'(x) = nx^{n-1}$.

4.8.7.5 Proof $f_1 : K \to K$ is differentiable, $\forall x \in K$, $d_x f_1 = f_1$.
$f_1'(x) = d_x f_1(1) = f_1(1) = 1$, $\forall x \in K$.
If $f_n'(x) = nx^{n-1}$, then $f_{n+1}'(x) = (f_n f_1)'(x) = f_n(x)f_1'(x) + f_n'(x)f_1(x) = x^n + xf_n'(x) = (n+1)x^n$.

4.8.7.6 Prop Let E, F and G be normed vector space. $U \subseteq E$ be an open subset, $\varphi \in \mathscr{L}(F, G)$, $p \in U$. If $f : U \to F$ is differentiable at p, then so is $\varphi \circ f$. Moreover, $d_p(\varphi \circ f) = \varphi \circ d_p f$.

4.8.7.7 Proof φ is differentiable at $f(p)$, and $d_{f(p)}\varphi = \varphi$.

4.8.7.8 Corollary Let E and F be normed vector spaces, $U \subseteq E$ be an open subset, $p \in U$. Let $f : U \to F$ and $g : U \to F$ be mappings that are differentiable at p, $(a, b) \in K \times K$.
Then $af + bg$ is differentiable at p, and $d_p(af + bg) = ad_p f + bd_p g$.

4.8.7.9 Proof $af + bg$ is the composite $U \to F \times F \to F$, $x \mapsto (f(x), g(x)) \mapsto af(x) + bg(x)$.
$\|ay + bz\| \le |a|\|y\| + |b|\|z\| \le (|a| + |b|)\max\{\|y\|, \|z\|\}$.

4.8.7.10 Def Let E be a vector space over K, and $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on E. We say that $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exists constants $C_1 > 0$ and $C_2 > 0$, s.t. $\forall s \in E$, $C_1\|s\|_1 \le \|s\|_2 \le C_2\|s\|_1$.

4.8.7.11 Prop If $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, then $\mathrm{Id}_E : (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)$ and $\mathrm{Id}_E : (E, \|\cdot\|_2) \to (E, \|\cdot\|_1)$ are bounded linear mappings. Moreover, $\|\cdot\|_1$ and $\|\cdot\|_2$ defines the same topology on E.

4.8.7.12 Proof $\|s\|_2 \le C_2\|s\|_1 \le C_1^{-1}\|s\|_2$. So these linear mappings are bounded. Hence $\mathrm{Id}_E : (E, \|\cdot\|_1) \to (E, \|\cdot\|_2)$ and $\mathrm{Id}_E : (E, \|\cdot\|_2) \to (E, \|\cdot\|_1)$ are continuous. So $\forall$ open subset of $(E, \|\cdot\|_2)$, $\mathrm{Id}_E^{-1}(U) = U$ is open in $(E, \|\cdot\|_1)$. Conversely if V is open in $(E, \|\cdot\|_1)$ then $V = \mathrm{Id}_E^{-1}(V)$ is open in $(E, \|\cdot\|_2)$.

4.8.7.13 Remark If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on E that define the same topology on E, then they are equivalent. (Under the assumptions that $|\cdot|$ is not trivial.)

4.8.7.14 Prop Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces, $\|\cdot\|_E'$ and $\|\cdot\|_F'$ be norms on E and F that are equivalent to $\|\cdot\|_E$ and $\|\cdot\|_F$, respectively. Let $U \subseteq E$ be an open subset and $f : U \to F$ be a mapping. Let $p \in U$. Then f is differentiable at p with respect to $\|\cdot\|_E$ and $\|\cdot\|_F$ iff it is differentiable with respect to $\|\cdot\|_E'$ and $\|\cdot\|_F'$. Moreover the differential of f at p is not changed in the change of norms from $(\|\cdot\|_E, \|\cdot\|_F)$ to $(\|\cdot\|_E', \|\cdot\|_F')$.

4.8.7.15 Proof $f = \mathrm{Id}_F \circ f \circ \mathrm{Id}_U$.
$d_p f = d_{f(p)}\mathrm{Id}_F \circ d_p f \circ d_p\mathrm{Id}_U = \mathrm{Id}_F \circ d_p f \circ \mathrm{Id}_E = d_p f$.

4.8.7.16 Theorem Let V be a finite dimensional vector space over K. Then all norms on V are equivalent. Moreover, V is complete with respect to any norms on V.

4.8.7.17 Proof Let $(e_i)_{i=1}^n$ be a basis of V. Then the mapping $V \to \mathbb{R}_{\geq 0}$, $a_1e_1 + \cdots + a_ne_n \mapsto \max\{|a_1|, \cdots, |a_n|\}$ is a norm on V.

Let $\|\cdot\|$ be another norm on V. One has $\|a_1e_1+\cdots+a_ne_n\| \leq |a_1|\|e_1\|+\cdots+|a_n|\|e_n\| \leq (\|e_1\|+\cdots+\|e_n\|)\max\{|a_1|+\cdots+|a_n|\}$

We reason by induction that there exists $C > 0$ s.t. $\max\{|a_1|, \cdots, |a_n|\} \leq C\|a_1e_1 + \cdots + a_ne_n\|$

The case where $n = 0$ is trivial(We just have a single norm on $\{0\}$).

Case where $n = 1$, $\|a_1e_1\| = |a_1|\|e_1\|$. $|a_1| = \|e_1\|^{-1}\|a_1e_1\|$.(It is complete since K is complete.)

Induction hypothesis: true for vector spaces of dimension $< n$.

Let $W = \{a_1e_1+\cdots+a_{n-1}e_{n-1}|(a_1, \cdots, a_{n-1}) \in K^{n-1}\}$ equipped with the restriction of $\|\cdot\|$. The induction shows that W is complete. Hence it is closed in V.

Let $Q = V/W$ and $\|\cdot\|_Q$ be the quotient norm on Q, that is defined as $\forall x \in Q$, $\|\alpha\|_Q = \inf_{s\in\alpha} \|s\|$.

If $s \in V \setminus W$, $\exists \varepsilon > 0$, s.t. $\overline{B}(s, \varepsilon) \bigcap W = \emptyset$. $\forall t \in W$, $s + t \notin \overline{B}(0, \varepsilon)$. Since otherwise $-t \in W \bigcap \overline{B}(s, \varepsilon)$.

Therefore, $\|[s]\|_Q = \inf_{t\in W} \|s + t\| \geq \varepsilon > 0$

Applying the induction hypothesis to W, we obtain the existence of some $A > 0$ s.t. $\max\{|a_1|, \cdots, |a_{n-1}|\} \leq A\|a_1e_1 + \cdots + a_{n-1}en - 1\|$, for any $(a_1, \cdots, a_{n-1}) \in K^{n-1}$.

Take $s = a_1e_1 + \cdots + a_{n-1}e_{n-1} + a_ne_n \in V$. Let $\alpha = [s] = a_n[e_n] \in Q$

$A^{-1}\max\{|a_1|, \cdots, |a_{n-1}|\} \leq \|a_1e_1 + \cdots + a_{n-1}e_{n-1}\| = \|s - a_ne_n\| \leq \|s\| + |a_n|\|e_n\|$

$\|\alpha\|_Q = |a_n|\|[e_n]\|_Q = |a_n| \inf_{t\in W} \|e_n + t\|$

Take $e'_n \in V$ s.t. $[e'_n] = [e_n]$ and $\|e'_n\| \leq \|[e_n]\|_Q + \varepsilon$.

$\|e_n\| \geq \|[e_n]\|$. Because of induction hypothesis, $\|[e_n]\|_Q > 0$, thus $\|e_n\| \leq B\|[e_n]\|$

$s = a_ne_n + t \in V$ with $t \in W$.

$\|s\| \geq \|[a_ne_n]\|_Q = |a_n|\|[e_n]\|_Q \geq B^{-1}|a_n|\|e_n\|$

If $\|a_ne_n\| < \frac{1}{2}\|t\|$, $\|s\| \geq \|t\| - \|a_ne_n\| > \frac{1}{2}\|t\| \geq \frac{1}{2}A\max\{|a_1|, \cdots, |a_{n-1}|\}$.

If $\|a_ne_n\| \geq \frac{1}{2}\|t\|$, $\|s\| \geq B^{-1}|a_n|\|e_n\| \geq \frac{B^{-1}}{2}\|t\| \geq \frac{B^{-1}A}{2}\max\{|a_1|, \cdots, |a_{n+1}|\}$.

We take $C = \min\{B^{-1}\|e_n\|, \frac{A}{2}, \frac{B^{-1}A}{2}\}$

Then $\|s\| \geq C\max\{|a_1|, \cdots, |a_n|\}$.

4.8.8 Theorem: Differentiability Implies Continuity

Let E and F be normed vector space over a complete valued field, $U \subseteq E$ be an open subset, and $f : U \to F$ be a mapping. If f is differentiable at $p \in U$, then f is continuous at p.

4.8.9 Proof

$f(x) = f(p) + d_pf(x - p) + o(\|x - p\|) = O(\|x - p\|) + f(p) = f(p) + o(1) \quad x \to p$. So $\lim_{x\to p} f(x) = f(p)$.

4.9 Compactness

4.9.1 Quasi-Compact/Compact

Let X be a topological space, $Y \subseteq X$. We call open cover of Y any family $(U_i)_{i\in I}$ s.t. $Y \subseteq \bigcup_{i\in I} U_i$. If I is a finite set, we say that $(U_i)_{i\in I}$ is a finite open cover. If $J \subseteq I$ s.t. $Y \subseteq \bigcup_{j\in J} U_j$ is a subcover of $(U_i)_{i\in I}$.

If any open cover of Y has a finite subcover, we say that Y is quasi-compact. If in addition X is Hausdorff, we say that Y is compact.

4.9.1.1 Ultrafilter Let X be a set and $\mathcal{F}$ be a filter on X.

If there doesn't exist any filter $\mathcal{F}'$ of X s.t. $\mathcal{F} \subsetneq \mathcal{F}'$, then we say that $\mathcal{F}$ is an ultrafilter.

Zorn's Lemma implies that for any filter $\mathcal{F}_0$ of X, there exists an ultrafilter $\mathcal{F}$ of X containing $\mathcal{F}_0$.

4.9.1.2 Prop Let $\mathcal{F}$ be a filter on a set X. The following statements are equivalent.

(1) $\mathcal{F}$ is an ultrafilter.

(2) $\forall A \subseteq X$, either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$.

(3) $\forall (A, B) \in \mathscr{P}(X)^2$, if $A \bigcup B \in \mathcal{F}$, then $A \in \mathcal{F}$ or $B \in \mathcal{F}$.

4.9.1.3 Proof From (1) to (2): Suppose that $A \in \mathscr{P}(X)$ s.t. $A \notin \mathcal{F}$ and $X \setminus A \notin \mathcal{F}$.
For any $B \in \mathcal{F}$. One has $B \bigcap A \notin \emptyset$, since otherwise $B \subseteq X \setminus A$ and hence $X \setminus A \in \mathcal{F}$, contradiction!
Hence $\mathcal{F} \bigcup \{A\}$ generates some filter strictly larger than $\mathcal{F}$, $\mathcal{F}$ cannot be an ultrafilter.
From (2) to (3): Suppose $B \notin \mathcal{F}$, then $X \setminus B \in \mathcal{F}$
$(A \bigcup B) \bigcap (X \setminus B) = A \setminus B \in \mathcal{F}$. So $A \in \mathcal{F}$.
From (3) to (1): Suppose $\mathcal{F}'$ is a filter s.t. $\mathcal{F} \subsetneq \mathcal{F}'$. Take $A \in \mathcal{F}' \setminus \mathcal{F}$. Then by $X - A \bigcup (X \setminus A) \in \mathcal{F}$. Hence $X \setminus A \in \mathcal{F} \subseteq \mathcal{F}'$. $\emptyset = A \bigcap (X \setminus A) \in \mathcal{F}'$ is impossible.

4.9.2 Equivalent Conditions for Quasi-Compactness

Let (X, τ) be a topological space. The following are equivalent:

1. X is quasi-compact.
2. Any filter of X has an accumulation point.
3. Any ultrafilter of X is convergent.

4.9.2.1 Proof From (1) to (2): Assume that a filter $\mathcal{F}$ of X doesn't have any accumulation point.
$\forall x \in X$, $\exists A_x \in \mathcal{F}$, $\exists$ open neighborhood V_x of x s.t. $A_x \bigcap V_x = \emptyset$. Since $X = \bigcup\limits_{x \in X} V_x$, there is $\{x_1, \cdots, x_n\} \subseteq X$, s.t. $X = \bigcup\limits_{i=1}^{n} U_{x_i}$. Take $B = \bigcap\limits_{i=1}^{n} A_{x_i} \in \mathcal{F}$, $B \bigcap X = B = \emptyset$. Since $B \bigcap V_{x_i} = \emptyset$, $\forall i \in \{1, \cdots, n\}$.
From (2) to (3): Let $\mathcal{F}$ be an ultrafilter of X, by (2), there is $x \in X$ s.t. $\mathcal{F} \bigcup V_x$ generates a filter $\mathcal{F}'$. Since $\mathcal{F}$ is an ultrafilter $\mathcal{F} = \mathcal{F}'$ and hence $V_x \subseteq \mathcal{F}$.
From (3) to (1): Let $(U_i)_{i \in I}$ be an open cover of X. We suppose that $(U_i)_{i \in I}$ does not have any finite subcover. For any $i \in I$, let $F_i = X \setminus U_i$. For any $J \subseteq I$ finite, $F_J = \bigcap_{j \in J} F_j = X \setminus \bigcup_{j \in J} U_j \neq \emptyset$. Let $\mathcal{F}$ be the smallest filter on X that contains $\{F_J \mid J \in I \text{finite}\}$. Let $\mathcal{F}'$ be an ultrafilter containing $\mathcal{F}$. It has a limit point x. There exists $i \in I$ s.t. $x \in U_i$. Since U_i is a neighborhood of x, and $\mathcal{V} \subseteq \mathcal{F}'$, we get $U_i \in \mathcal{F}'$. This is impossible since $F_i \in \mathcal{F}'$.

4.9.2.2 Def A filter $\mathcal{F}$ on X is called a Cauchy filter if $\forall \delta > 0$, $\exists A \in \mathcal{F}$ s.t. $\text{diam}(A) \leq \delta$.

4.9.3 Heine-Borel Theorem in Metric Spaces

Let (X, d) be a metric space. The following statements are equivalent:

1. X is complete, and $\forall \varepsilon > 0, \exists X_\varepsilon \subseteq X$ finite s.t. $X = \bigcup_{x \in X_\varepsilon} B(x, \varepsilon)$, where $B(x, \varepsilon)$ denotes the open ball centered at x with radius ε.
2. X is compact.

4.9.3.1 Lemma Let (X, d) be a metric space.

1. Let $\mathcal{F}$ be a Cauchy filter on X. Any accumulation point of $\mathcal{F}$ is a limit point of $\mathcal{F}$.
2. X is complete if and only if any Cauchy filter of X has a limit point.

4.9.3.2 Proof of the Lemma

1. Let $\mathcal{F}$ be a Cauchy filter. Let $x \in X$ be an accumulation point of $\mathcal{F}$. For any $\varepsilon > 0$, $\exists A \in \mathcal{F}$ with diameter $\leq \frac{\varepsilon}{2}$. Note that $A \cap B(x, \frac{\varepsilon}{2}) \neq \varnothing$. Take $y \in A \cap B(x, \frac{\varepsilon}{2})$. Then, $\forall z \in A$, we have $d(x, z) \leq d(x, y) + d(y, z)$. Since $y \in B(x, \frac{\varepsilon}{2})$, it follows that $d(x, y) < \frac{\varepsilon}{2}$. Also, since A has diameter $\leq \frac{\varepsilon}{2}$, we have $d(y, z) \leq \frac{\varepsilon}{2}$. Therefore, $d(x, z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. This implies $z \in B(x, \varepsilon)$. Hence, $A \subseteq B(x, \varepsilon)$. So $B(x, \varepsilon) \in \mathcal{F}$. This implies $x \in \mathcal{F}$.
2. " $\Longleftarrow$ " Let $(x_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in X. Let $\mathcal{F} = \{A \subseteq X \mid \exists N \in \mathbb{N}, \{x_N, x_{N+1}, \cdots\} \subseteq A\}$. This is a Cauchy filter on X since $\lim_{n \to \infty} \text{diam}\{x_n, x_{n+1}, \cdots\} = 0$. Hence $\mathcal{F}$ has a limit point $x \in X$. By Def, $\forall U \in \mathcal{V}_x, \exists N \in \mathbb{N}, \{x_N, x_{N+1}, \cdots\} \subseteq U$. So $x = \lim_{n \to \infty} x_n$.
   " $\Longrightarrow$ " Suppose that X is complete. Let $\mathcal{F}$ be a Cauchy filter. $\forall n \in \mathbb{N}_{\geq 1}$, let $A_n \in \mathcal{F}$ s.t. $\text{diam}(A_n) \leq \frac{1}{n}$. Take $x_n \in \bigcap_{k=1}^{n} A_k \in \mathcal{F}$. Then $(x_n)_{n \in \mathbb{N}_{\geq 1}}$ is a Cauchy sequence since $\forall \epsilon > 0$, if we take $N \in \mathbb{N}$ with $\frac{1}{N} \leq \epsilon$, then $\forall (n, m) \in \mathbb{N}_{\geq N}$, $d(x_n, x_m) \leq \frac{1}{N}$ since $\{x_n, x_m\} \subseteq A_N$. Hence $(x_n)_{n \in \mathbb{N}}$ converges to some $x \in X$. Note that x is a limit point of $\mathcal{F}$.

4.9.3.3 Proof of the Theorem *From (1) to (2):* Let $\mathcal{F}$ be an ultrafilter. Then $X = \bigcup_{x \in X} B(x, \varepsilon)$ for some $\varepsilon > 0$. Let $\{x_1, x_2, \cdots, x_n\} \subseteq X$ s.t. $B(x_i, \varepsilon) \in \mathcal{F}$ for all $i \in \{1, \cdots, n\}$. There exists some $i \in \{1, \cdots, n\}$ s.t. $B(x_i, \varepsilon) \in \mathcal{F}$ (By induction). Thus, $\mathcal{F}$ is a Cauchy filter (for any $\delta > 0$, $\exists A \in \mathcal{F}$ of diameter $\leq \delta$). Since X is complete, $\mathcal{F}$ has a limit point. So $\mathcal{F}$ is compact.
From (2) to (1): Let $\varepsilon > 0$. One has $X = \bigcup_{x \in X} B(x, \varepsilon)$. Since X is compact, $\exists X_\varepsilon \subseteq X$ finite s.t. $X = \bigcup_{x \in X_\varepsilon} B(x, \varepsilon)$. Compact $\to$ completeness is trivial.

4.9.3.4 Prop Let $f : X \to Y$ be a continuous mapping of topological spaces. If $A \subseteq X$ is quasi-compact, then $f(A) \subseteq Y$ is also quasi-compact.

4.9.3.5 Proof Let $(V_i)_{i\in I}$ be an open cover of $f(A)$. Then $(f^{-1}(V_i))_{i\in I}$ is an open cover of A. Since A is quasi-compact, there exists a finite subcover $(V_{i_j})_{j\in J}$ s.t. $A \subseteq \bigcup_{j\in J} f^{-1}(V_{i_j})$. This implies $f(A) \subseteq \bigcup_{j\in J} V_{i_j}$. So $f(A)$ is quasi-compact.

4.9.3.6 Prop Let X be a topological space and $A \subseteq X$ be a compact subset. For any closed subset F of X, $A \cap F$ is quasi-compact.

4.9.3.7 Proof Let $(U_i)_{i\in I}$ be an open cover of $A \cap F$. Then $(U_i)_{i\in I} \bigcup (X \setminus F)$ is an open cover of A. Since A is compact, there exists a finite subcover $(U_j)_{j\in J}$ s.t. $A \subseteq \bigcup_{j\in J}(U_j) \bigcup (X \setminus F)$. This implies $A \cap F \subseteq \bigcup_{j\in J} U_{i_j}$. So $A \cap F$ is quasi-compact.

4.9.3.8 Prop Let X be a Hausdorff topological space. Any compact subset A of X is closed.

4.9.3.9 Proof Let $x \in X \setminus A$. For every $y \in A$, there exist open subsets U_y and V_y s.t. $y \in U_y$, $x \in V_y$, and $U_y \cap V_y = \varnothing$.
Since $A \subseteq \bigcup_{y\in A} U_y$, there exists a finite subcover $\{U_{y_1}, U_{y_2}, \cdots, U_{y_n}\} \subseteq A$ s.t. $A \subseteq \bigcup_{k=1}^n U_{y_k}$.
Let $U = \bigcup_{i=1}^n U_{y_i}$ and $V = \bigcap_{i=1}^n V_{y_i}$. These are open subsets.
Moreover, $A \subseteq U$, $x \in V$, and $U \cap V = \bigcup_{i=1}^n (U_{y_i} \cap V) = \varnothing$.
In particular, $x \in V \subseteq X \setminus A$. Consequently, $X \setminus A$ is open.

4.9.3.10 Prop Let X be a Hausdorff topological space, and A and B be compact subsets of X, $A \bigcap B = \emptyset$. Then there exist open subsets U and V of X s.t. $A \subseteq U$, $B \subseteq V$, and $U \cap V = \varnothing$.

4.9.3.11 Proof We have seen in the proof of the previous Prop that for every $x \in B$, there exist U_x and V_x open s.t. $A \subseteq U_x$, $x \in V_x$, and $U_x \cap V_x = \varnothing$. Since $B \subseteq \bigcup_{x\in B} V_x$, there exists a finite subset $\{x_1, \cdots, x_m\} \subseteq B$ s.t. $B \subseteq \bigcup_{i=1}^m V_{x_i}$. We take $U = \bigcap_{i=1}^m U_{x_i}$ and $V = \bigcup_{i=1}^m V_{x_i}$. Clearly, $A \subseteq U$, $B \subseteq V$, and $U \cap V = \varnothing$.

4.9.4 Cantor's Intersection Theorem

Let (X, τ) be a Hausdorff topological space. If $(A_n)_{n\in\mathbb{N}}$ is a sequence of compact subsets of X s.t. $A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots$, then $\bigcap_{n\in\mathbb{N}} A_n \neq \varnothing$.

4.9.4.1 Proof Suppose that $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$. Then $A_0 \subseteq \bigcup_{n\in\mathbb{N}}(X \setminus A_n)$.
Since A_0 is compact, there exists $N \in \mathbb{N}$ s.t. $A_0 \subseteq \bigcup_{n=0}^N (X \setminus A_n) = X \setminus \bigcap_{n=0}^N A_n = X \setminus A_N$.
So $A_N = \varnothing$. (Since $A_N \subseteq A_0 \subseteq X \setminus A_N$).

4.9.4.2 Prop We fix a complete value field $(K, |\cdot|)$. Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be normed vector spaces over K. Assume that E is finite dimensional. Then any K-linear map $\Psi : E \to F$ is bounded.

4.9.4.3 Proof Let $(e_i)_{i=1}^n$ be a basis of E over K. Use **4.7.7.16 Theorem**: norms on E are equivalent.

$$\forall (a_1, \ldots, a_n) \in K, \|a_1e_1 + \ldots + a_ne_n\|_E := \max\{|a_1|, \ldots, |a_n|\}.$$

Then $b_j = a_1e_1 + \ldots + a_ne_n$.

$$\|\Psi(b_j)\|_F = \|a_1\Psi(e_1) + \ldots + a_n\Psi(e_n)\|_F \leq \sum_{i=1}^n |a_i|\|\Psi(e_i)\|_F \leq \left(\sum_{i=1}^n \|\Psi(e_i)\|_F\right) \|b_j\|_E.$$

4.9.5 Sequentially Compact

Let (X, τ) be a topological space. If any sequence in X has a convergent subsequence, we say that X is sequentially compact.

4.9.5.1 Example By Bolzano-Weierstrass, any bounded sequence in $\mathbb{R}$ has a convergent subsequence. So any bounded and closed subset of $\mathbb{R}$ is sequentially compact.

4.9.5.2 Theorem Let (X, d) be a metric space. Then the following statements are equivalent:

1. (X, d) is compact.
2. (X, d) is sequentially compact.

4.9.5.3 Proof $1 \to 2$ Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. Assume that no subsequence of (x_n) converges in X. $\forall p \in X$, there exists $\epsilon_p > 0$ s.t. $\{n \in \mathbb{N} \mid d(p, x_n) < \epsilon_p\}$ is finite.

This is true, otherwise we can construct a strictly increasing sequence $\{n_k\}_{k\in\mathbb{N}}$ s.t.

$$\forall k \in \mathbb{N}, \exists x_n \text{ s.t. } d(p, x_n) < \frac{1}{k}, \text{ which implies } x_n \to p.$$

Since X is compact, then $\mathbb{N} = \bigcup_{i=1}^{d}\{n \in \mathbb{N},\ d(p_i, x_n) < \epsilon p_i\}$, which is finite, leading to contradiction.

$2 \to 1$ We show that (X, d) is complete. Let (x_n) be a Cauchy sequence. By (2), it contains a convergent subsequence. Therefore, by a fact proved before (**2.7.1(4) Prop**), (x_n) must converge to the same limit. Hence, (X, d) is complete. By implying **4.8.3 Heine-Borel Theorem in Metric Spaces**, we only need to prove that it's covered by finitely many balls of radius ϵ.

If X is not covered by finitely many balls of radius ϵ, then we can construct a sequence (x_n) s.t. $x_{n+1} \notin \bigcup B(x_k, \xi)$ for any subsequence of this is not Cauchy (not convergent), which leads to contradiction.

4.9.6 Locally Compact

Let X be a Hausdorff topological space. If for any $x \in X$ there exists a compact neighborhood C_x of x in X, we say that X is locally compact.

4.9.6.1 Remark Locally compact $\nRightarrow$ compact; compact $\to$ locally compact.

4.9.6.2 Example $\mathbb{R}$ is locally compact, since $C_x = [x-1, x+1]$ is a compact neighborhood of x.

4.9.6.3 Prop Assume that $(K, |\cdot|)$ is a locally compact non-trivial valued field. Let $(E, \|\cdot\|)$ be a finite dimensional normed K-vector space. A subset $Y \subseteq E$ is compact if and only if it is closed and bounded.

4.9.6.4 Proof $(\to)$ Let $Y \subseteq X$ be compact. Then Y is closed (**4.8.5.2 Theorem**). Moreover, $Y \subseteq \bigcup_{n\in\mathbb{N}_{\geq 1}} B(0, n)$. We can find finitely many positive integers $n_1, \ldots, n_k$ s.t. $Y \subseteq B(0, n_1) \bigcup \ldots \bigcup B(0, n_k) \subseteq B(0, n_k)$, hence Y is bounded.

$(\Leftarrow)$ Let (e_i) be a basis of E. Again, we assume $\|a_1e_1 + \ldots + a_de_d\| = \max\{\|a_1\|, \ldots, \|a_d\|\}$. We only need to prove sequential compactness by the theorem proved before.

Let $(x_n) = (a_1^{(n)}e_1 + \ldots + a_d^{(n)}e_d)$. Since Y is bounded, for $m \in \{1, \ldots, d\}$, the sequence $(a_i^{(n)})$ is bounded. In particular, we find $M > 0$ s.t. $|a_i^{(n)}| < M$ for all $i \in \{1, \ldots, d\}$. Since $(K, |\cdot|)$ is locally compact, there is a compact set C_0 of K that is a neighborhood of 0. Let $\epsilon > 0$, $\overline{B}(0, \epsilon) \subseteq C$. Since K is not trivially valued, there exists $a \in K$ s.t. $|a| \geq \frac{M}{\epsilon}$.

Then $\overline{B}(0, M) \subseteq aC$. $C \subseteq K$ is compact. We have a K-linear mapping $K \to K \quad y \mapsto ay$, this mapping is bounded, hence continuous, hence aC is compact (the image of C through a continuous mapping). $\overline{B}(0, M) \subseteq aC$ is a closed subspace of a compact.

So it is compact $\to$ it is sequentially compact.

Therefore, we can find $I_1, \ldots, I_d$ that are infinite subsets of $\mathbb{N}$, with $I_1 \supset I_2 \supset \ldots \supset I_d$, s.t. $(a_i^{(n)})_{n\in I_i}$ converges. They converge to some $a_i \in K$. It follows that our original sequence has a convergent subsequence converging to $a_1e_1 + \ldots + a_de_d$. So Y is sequentially compact.

4.9.7 Extreme Value Theorem

Let X be a topological space and $f : X \to \mathbb{R}$ be a continuous mapping. If $Y \subseteq X$ is a non-empty quasi-compact subset, then there exists $a \in Y$ and $b \in Y$ s.t. for all $x \in Y$, $f(a) \leq f(x) \leq f(b)$. Namely, the restriction of f to Y attains its maximum and minimum.

4.9.7.1 Proof $f(Y) \subseteq \mathbb{R}$ is a compact subset since Y is quasi-compact and $\mathbb{R}$ is Hausdorff (**4.8.3.4 Prop**). Moreover, since $\mathbb{R}$ is locally compact, $f(Y)$ is bounded and closed. Note that there exists sequences $\{d_n\}_{n\in\mathbb{N}}$ and $\{e_n\}_{n\in\mathbb{N}}$ in $f(Y)$ that tend to $\sup f(Y)$ and $\inf f(Y)$ respectively. Since $f(Y)$ is closed, $\sup f(Y)$ and $\inf f(Y)$ belong to $f(Y)$. So $f(Y)$ has a greatest and a least element.

4.10 Mean Value Theorems

4.10.1 Theorem (Rolle)

Let a and b be real numbers s.t. $a < b$. Let $f : [a, b] \to \mathbb{R}$ be a continuous mapping that is differentiable on $]a, b[$. If $f(a) = f(b)$, then there exists $c \in]a, b[$ s.t. $f'(c) = 0$.

4.10.1.1 Proof Since $[a, b]$ is compact, f attains its maximum and minimum. Let $M = \max f([a, b])$, $m = \min f([a, b])$. Let $l = f(a) = f(b)$.

If $M \neq l$, there exists $a < t < b$ s.t. $f(t) = M$. Then, $f(t+x) = f(t) + f'(t)x + o(|x|)$, $\quad f(t-x) = f(t) - f'(t)x + o(|x|)$. So $0 \leq (f(t+x) - f(t))(f(t-x) - f(t)) = f'(t)^2x^2 + o(|x|^2)$ Taking the limit when $x \to 0$, we get $f'(t) = 0$.

If $m \neq l$, then any $t \in]a, b[$ s.t. $f(t) = m$ satisfies $f'(t) = 0$.

If $m = l = M$, f is constant, so $\forall t \in]a, b[$, $f'(t) = 0$.

4.10.2 Theorem (Mean value theorem, Lagrange)

Let a and b be two real numbers, $a < b$. Let $f : [a, b] \to \mathbb{R}$ be a continuous mapping that is differentiable on $]a, b[$. Then there exists $t \in]a, b[$ s.t. $f'(t) = \frac{f(b)-f(a)}{b-a}$.

4.10.2.1 Proof Let $g : [a, b] \to \mathbb{R}$ be defined as $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$. Then,

$$g(a) = f(a), \quad g(b) = f(b) - (f(b) - f(a)) = f(a)$$

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}$$

By Rolle's theorem, there exists $t \in]a, b[$ s.t. $g'(t) = 0$, that is,

$$f'(t) = \frac{f(b) - f(a)}{b - a}$$

4.10.3 Theorem (Mean value inequality)

Let a and b be two real numbers, s.t. $a < b$. Let $(\mathbb{E}, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$. Let $f : [a, b] \to \mathbb{E}$ be a continuous mapping s.t. f is differentiable on $]a, b[$. Then,

$$\|f(b) - f(a)\| \leq \sup_{x \in]a,b[} \|f'(x)\|(b - a)$$

4.10.3.1 Proof Suppose that $\sup_{x\in]a,b[} \|f'(x)\| < +\infty$. Let $M \in \mathbb{R}$ s.t. $M > \sup_{x\in]a,b[} \|f'(x)\|$. Let $J = \{x \in [a, b] \mid \forall y \in [a, x], \|f(y) - f(a)\| \leq M(y - a)\}$. By Def, J is an interval containing a, so J is of the form $[a, c]$ or $[a, c[$. Since f is continuous, by taking a sequence $(c_n)_{n\in\mathbb{N}}$ in $[a, c[$ that converges to c, we obtain

$$\|f(c) - f(a)\| = \lim_{n\to+\infty} \|f(c_n) - f(a)\| \leq \lim_{n\to+\infty} M(c_n - a)$$

Hence, $c \in J$, namely $J = [a, c]$.

We assume that $c > a$. We will prove that $c = b$ by contradiction.

Suppose that $c < b$. (Attention: This is a bit imprecise, but the correct proof is to do something similar to this twice, which is a bit lengthy.) $\forall b \in]0, b - c[$, $\|f(c + b) - f(c)\| = \|h \times f'(c) + o(h)\| \leq \|f'(c)\| \times h + o(h) \quad h \to 0$. Since $M > \|f'(c)\|$, $\exists h_0 > 0 (h_0 < b - c)$ s.t. $\forall 0 < h < h_0$, $\|f(c + h) - f(c)\| \leq Mh$ since $\|o(h)\| \leq (M - \|f'(c)\|) \times h$. Hence

$$\|f(c + h) - f(a)\| \leq \|f)c + h) - f(c)\| + \|f(c) - f(a)\| \leq M(c + h - a)$$

So, $c + h_0$ is in J, a contradiction.

4.10.4 Intermediate Value Theorem

Let I be an open interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ be a differentiable mapping, then $f(I)$ is an interval.

4.10.4.1 Proof Let x and y be two elements of $f(I)$ with $x \neq y$. Let a and b be elements of I s.t. $x = f(a)$, $y = f(b)$. Without generality, we suppose $a < b$. Let $z \in \mathbb{R}$ s.t. $(z - x)(z - y) \leq 0$.

We conduct by induction three sequences $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$, and $(c_n)_{n\in\mathbb{N}}$ s.t.:

1. $a = a_0$, $b_0 = b$, $c_0 = \frac{a+b}{2}$ at t_0.
2. If a_n, b_n, c_n are constructed, satisfying $c_n = \frac{1}{2}(a_n + b_n)$, $(z - a_n)(z - b_n) \leq 0$, we let $(a_{n+1}, b_{n+1}) = (a_n, c_n)$ if $(z - f(a_n))(z - f(c_n)) \leq 0$, $(a_{n+1}, b_{n+1}) = (c_n, b_n)$ if $(z - f(a_n))(z - f(c_n)) > 0$ (in this case $(z - f(c_n))(z - f(b_n)) \leq 0$).

The sequence $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ are strictly increasing/decreasing and bounded, hence converges to some $l \in]a, b[$, $m \in]a, b[$. Note that $|b_n - a_n| = \frac{1}{2^n}|b - a| \to 0$ (as $n \to +\infty$), so $l = m$.

By $(z - f(a_n))(z - f(b_n)) \leq 0$, we obtain by letting $n \to +\infty$ that $(z - f(l))(z - f(l)) \leq 0$, which implies $z \in f(l)$. (Using dichotomy.)

4.10.5 Theorem (Darboux)

Let I be an open interval of $\mathbb{R}$ and $f : I \to \mathbb{R}$ be a differentiable mapping. Then $f'(I)$ is an interval.

4.10.5.1 Proof Consider the following mappings:

$$g : [a, b] \to \mathbb{R}, \quad g(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & x \neq a \\ f'(a) & x = a \end{cases}$$

$$h : [a, b] \to \mathbb{R}, \quad h(x) = \begin{cases} \frac{f(b)-f(x)}{b-x} & x \neq b \\ f'(b) & x = b \end{cases}$$

g and h are continuous, and $\frac{f(x)-f(a)}{x-a} = f'(a) + o(1)$ as $x \to a$. So $g([a, b])$ and $h([a, b])$ are intervals.

Moreover, by the mean value theorem, $g([a, b]) \subseteq f'(I)$ and $h([a, b]) \subseteq f'(I)$. So $\{f'(a) - f'(b)\} \leq g([a, b]) \bigcup h([a, b]) \subseteq f'(I)$.

Note that $g(b) = h(a)$, so $g([a, b]) \bigcup h([a, b])$ is an interval. Hence $f'(I)$ is an interval.

4.11 Fixed Point Theorem

4.11.1 Def of Fixed Point and Contraction

Let X be a set and $T : X \to X$ be a mapping. If there exists $x \in X$ s.t. $T(x) = x$, we say that x is a *fixed point* of T.

Let (X, d) be a metric space, and $T : X \to X$ be a mapping. If there exists $\varepsilon \in]0, 1[$ s.t. T is ε-Lipschitzian, $cd(T(x), T(y)) \leq \epsilon d(x, y))$, then we say that T is a *contraction*.

4.11.2 Theorem (Fixed Point Theorem)

Let (X, d) be a complete non-empty metric space, and $T : X \to X$ be a contraction. Then T has a unique fixed point. Moreover, for any $x_0 \in X$, the sequence $(x_n)_{n\in\mathbb{N}}$ defined by $x_{n+1} = T(x_n)$ converges to the fixed point.

4.11.2.1 Proof If p and q are two fixed points of T, then

$$d(p, q) = d(T(p), T(q)) \leq \varepsilon d(p, q)$$

for some $0 < \varepsilon < 1$, hence $d(p, q) = 0$.

Let $x_0 \in X$, $x_n = (T \circ \cdots \circ T)(x_0)$ (with n applications of T), $x_{n+1} = T(x_n)$.

For $n \in \mathbb{N}$, $d(x_n, x_{n+1}) \leq \varepsilon^n d(x_0, x_1)$.

For any $N \in \mathbb{N}$, $\forall (n, m) \in \mathbb{N}^2$, $n < m$,

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} d(x_k, x_{k+1}) \leq \sum_{k=n}^{m-1} \varepsilon^k d(x_0, x_1) \leq \frac{\varepsilon^n}{1-\varepsilon} d(x_0, x_1)$$

So $\lim_{N\to\infty} \sup_{n,m\in\mathbb{N}, n\leq m\leq N} d(x_n, x_m) = 0$.

$(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, hence it converges to some $p \in X$. $d(T(p), p) = \lim_{n\to\infty} d(T(x_n), x_n) = 0$ since $d : X \times X \to \mathbb{R}_{\geq 0}$ is continuous.

Chapter 5

Higher differentials

5.1 Multi-linear mapping

Let K be a commutative unitary ring.

5.1.0.1 Def Let $n \in \mathbb{N}$, $V_1, \ldots, V_n, W$ be K-modules.

We call n-linear mapping from $V_1 \times \cdots \times V_n$ to W any mapping $f : V_1 \times \cdots \times V_n \to W$ s.t. $\forall i \in \{1, \ldots, n\}$, $\forall (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in V_1 \times \cdots \times V_{i-1} \times V_{i+1} \times \cdots \times V_n$, the mapping $f(x_1, \ldots, x_{i-1}, \cdot, x_{i+1}, \ldots, x_n) : V_i \to W$ is a morphism of K-modules.

We denote by $\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W)$ the set of all n-linear mappings from $V_1 \times \cdots \times V_n$ to W.

5.1.0.2 Example $K \times K \to K$, $(a, b) \mapsto ab$ is a 2-linear mapping (bilinear mapping).

5.1.0.3 Remark When $n = 0$, $\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W)$ is considered as W by convention.

$$\mathrm{Hom}^{(1)}(V_1, W) = \mathrm{Hom}(V_1, W) = \{\text{morphisms of } K\text{-modules from } V_1 \text{ to } W\}$$

5.1.0.4 Prop Suppose that $n \geq 2$. For any $i \in \{1, \ldots, n-1\}$,

$$\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W) \xrightarrow{\Phi} \mathrm{Hom}^{(i)}(V_1 \times \cdots \times V_i, \mathrm{Hom}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W))$$
$$f \mapsto ((x_1, \ldots, x_i) \mapsto ((x_{i+1}, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)))$$

5.1.0.5 Proof The inverse of Φ is given by

$$g \in \mathrm{Hom}^{(i)}(V_1 \times \cdots \times V_i, \mathrm{Hom}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W)) \mapsto$$
$$(x_1, \ldots, x_n) \in V_1 \times \cdots \times V_n \mapsto g(x_1, \ldots, x_i)(x_{i+1}, \ldots, x_n)$$

5.1.0.6 Remark $\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W)$ is a sub-K-module of $W^{V_1 \times \cdots \times V_n}$ and Φ is an isomorphism of K-modules.

5.2 Operator norm of multi-linear mappings

5.2.1 Bounded n-linear Mappings

5.2.1.1 Def Let $(K, |\cdot|)$ be a complete valued field. Let $V_1, \ldots, V_n$ and W be normed vector spaces over K.

We define $\|\cdot\| : \mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W) \to [0, +\infty]$ as

$$\|f\| := \sup_{\substack{(x_1, \ldots, x_n) \in V_1 \times \cdots \times V_n \\ x_i \neq 0, \forall i}} \frac{\|f(x_1, \ldots, x_n)\|}{\|x_1\| \cdots \|x_n\|}$$

If $\|f\| < +\infty$, we say that f is bounded. We denote by $\mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W)$ the set of bounded n-linear mappings from $V_1 \times \cdots \times V_n$ to W.

5.2.1.2 Theorem For any $i \in \{1, \ldots, n-1\}$, if $f \in \mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W)$, then the (n-i)-linear mapping

$$f_i(x_1, \ldots, x_{i-1}, x_i, \cdot) : (x_{i+1}, \ldots, x_n) \mapsto f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

belongs to $\mathscr{L}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W)$.

Moreover,

$$\|f\| = \sup_{\substack{(x_1, \ldots, x_n) \in V_1 \times \cdots \times V_n \\ x_i \neq 0, \forall i}} \frac{\|f(x_1, \ldots, x_i, \cdot)\|}{\|x_1\| \cdots \|x_i\|}$$

5.2.1.3 Proof Let $(x_{i+1}, \dots, x_n) \in V_{i+1} \times \dots \times V_n$, $\|f(x_1, \dots, x_n)\| \le \|f\| \cdot \|x_1\| \cdot \dots \cdot \|x_n\| = (\|f\| \cdot \|x_1\| \cdot \dots \cdot \|x_i\|) \cdot \|x_{i+1}\| \cdot \dots \cdot \|x_n\|$. So $\|f(x_1, \dots, x_i, \cdot)\| \le \|f\| \cdot \|x_1\| \cdot \dots \cdot \|x_i\|$.

If we define $\|f\|' := \sup_{\substack{(x_1,\dots,x_i)\in V_1\times\dots\times V_i \\ x_j \neq 0,\, \forall j \le i}} \frac{\|f(x_1,\dots,x_i,\cdot)\|}{\|x_1\|\cdots\|x_i\|}$, then $\|f\|' \le \|f\|$.

Conversely, for any $(x_1, \dots, x_n) \in V_1 \times \dots \times V_n$, s.t. $x_i \neq 0$.

$\|f(x_1, \dots, x_n)\| \le \|f(x_1, \dots, x_i, \cdots)\| \cdot \|x_{i+1}\| \cdot \dots \cdot \|x_n\|$.

Hence, $\|f\| = \frac{\|f(x_1,\dots,x_n)\|}{\|x_1\|\cdot\dots\cdot\|x_n\|} \le \frac{\|f(x_1,\dots,x_i,\cdot)\|}{\|x_1\|\cdot\dots\cdot\|x_i\|} \le \|f\|'$.

Taking the supremum, we get $\|f\| = \|f\|'$.

5.2.1.4 Corollary

1. $\mathscr{L}^{(n)}(V_1 \times \dots \times V_n, W)$ is a vector subspace of $\mathrm{Hom}^{(n)}(V_1 \times \dots \times V_n, W)$.
2. $\|\cdot\|$ is a norm on $\mathscr{L}^{(n)}(V_1 \times \dots \times V_n, W)$.
3. For all $i \in \{1, \dots, n\}$, the mapping $\mathscr{L}^{(n)}(V_1 \times \dots \times V_n, W) \xrightarrow{\Phi} \mathscr{L}^{(i)}(V_1 \times \dots \times V_i, \mathscr{L}^{(n-i)}(V_{i+1} \times \dots \times V_n, W))$ is a K-linear isomorphism that preserves operator norms.

Here W is arbitrary.

5.2.1.5 Proof We reason by induction on n.

- For $n = 1$, $\mathscr{L}^{(1)}(V_1, W) = \mathscr{L}(V_1, W)$, which is trivially true.
- Suppose that the corollary is true for m-linear mappings with $m < n$. Take $i \in \{1, \dots, n-1\}$. We consider the following diagram of mappings:

$$\begin{array}{ccc} \mathscr{L}^{(n)}(V_1 \times \dots \times V_n, W) & \xrightarrow{\quad\Phi\quad} & \mathscr{L}^{(i)}(V_1 \times \dots \times V_i, \mathscr{L}^{(n-i)}(V_{i+1} \times \dots \times V_n, W)) \\ \cap\!\!\!| & & \cap\!\!\!| \\ \mathrm{Hom}^{(n)}(V_1 \times \dots \times V_n, W) & \xrightarrow{\quad\Phi\quad} & \mathrm{Hom}^{(i)}(V_1 \times \dots \times V_i, \mathrm{Hom}^{(n-i)}(V_{i+1} \times \dots \times V_n, W)) \end{array}$$

  WTS $\mathscr{L}^{(n)}(V_1 \times \dots \times V_n, W)$ is a vector subspace, we only need to show $\forall g \in \mathscr{L}^{(i)}(V_1 \times \dots \times V_i, \mathscr{L}^{(n-i)}(V_{i+1} \times \dots \times V_n, W))$, one has $\|\Phi^{-1}(g)\| = \|g\| < +\infty$ (**5.1.0.6 Remark**).

  For any $(x_1, \dots, x_n) \in V_1 \times \dots \times V_n$,

$$\begin{aligned} \|\Phi^{-1}(g)(x_1, \dots, x_n)\| \le \|g(x_1, \dots, x_i, x_{i+1}, \dots, x_n)\| &\le \|g\| \cdot \|x_1\| \cdots \|x_n\| \\ &\le \|g\| \cdot \|x_i\| \cdot \|x_{i+1}\| \cdots \|x_n\| \end{aligned}$$

  Therefore, $\|\Phi^{-1}(g)\| \le \|g\| \le \|\Phi^{-1}(g)\|$. ($\|\Phi(g)\| = \|g\| \forall g \in \mathscr{L}^{(n)}(V_1 \times \dots \times V_n, W)$).

$\|g\| = \sup\{\frac{\|g(x_1,\dots,x_n)\|}{\|x_1\|\cdot\dots\cdot\|x_n\|}\}$.

5.3 Higher Differentials

We fix a complete non-trivially valued field $(K, |\cdot|)$ and normed K-vector spaces E and F.

5.3.1 Def

Let $U \subseteq E$ be an open subset and $f : U \to F$ be a mapping.

1. If f is continuous, we say that f is of class C^0; any mapping from U to F is 0-time differentiable and we denote by $D^0 f$ the mapping $f : U \to F$.
2. If f is differentiable on an open neighborhood $V \subseteq U$ of some point $p \in U$, and $df : V \to \mathscr{L}(E, F)$, $x \mapsto d_x f$, is n-times differentiable at p, then we say that f is $(n+1)$-times differentiable at p. If f is $(n+1)$-times differentiable at any point $p \in U$, we denote by $D^{n+1} f : U \to \mathscr{L}^{(n+1)}(E^{n+1}, F)$ the mapping that sends $x \in U$ to the image of $D^n(df)(x)$ by the K-linear bijection $\mathscr{L}^{(n)}(E^n, \mathscr{L}(E, F)) \to \mathscr{L}^{(n+1)}(E^{n+1}, F)$.

If $D^n f$ is continuous, we say that f is of class $C^n (n \ge 0)$.

5.3.1.1 Remark If f is n-times differentiable, $\forall i \in \{1, \dots, n-1\}$, $\forall p \in U$, $\forall (h_1, \dots, h_n) \in E^n$, one has

$$D^i(D^{n-i} f)(p)(h_1, \dots, h_i)(h_{i+1}, \dots, h_n) = D^n f(p)(h_1, \dots, h_n)$$

$$D^{n-i} f : U \to \mathscr{L}^{(n-i)}(E^{n-i}, F), \ D^i(D^{n-i} f) \quad : \quad \begin{array}{ccc} U & \longrightarrow & \mathscr{L}^{(i)}(E^i, \mathscr{L}^{(n-i)}(E^{n-i}, F)) \\ & \searrow^{D^n f} & \wr\!\| \\ & & \mathscr{L}^{(n)}(E^n, F) \end{array}$$

5.3.1.2 Theorem Assume that $(K, |\cdot|) = (\mathbb{R}, |\cdot|)$. Let $f : U \to F$ be a mapping that is $(n+1)$-times differentiable on U. Let $p \in U$ and $h \in E$ s.t. $p + th \in U$ $\forall t \in I \subseteq \mathbb{R}$.

Then,

$$\|f(p+h) - f(p) - \sum_{k=1}^{n} \frac{1}{k!} D^k f(p)(h, \ldots, h)\| \leq \left(\sup_{t \in I \setminus \{0,1\}} \frac{(1-t)^n}{n!} \|D^{(n+1)} f(p+th)\| \right) \|h\|^{n+1}$$

(Taylor-Lagrange formula)

5.3.2 Gronwall Inequality

Let F be a normed vector space over $\mathbb{R}$, $(a, b) \in \mathbb{R}^2$, $a < b$. Let $f : [a, b] \to F$ and $g : [a, b] \to \mathbb{R}$ be continuous mappings that are differentiable on $]a, b[$. Suppose that $\forall t \in]a, b[$, $\|f'(t)\| \leq g'(t)$. Then $\|f(b) - f(a)\| \leq g(b) - g(a)$.

5.3.2.1 Proof Let $c \in]a, b[$, $\varepsilon > 0$. Let $J = \{t \in [c, b] \mid \forall s \in [c, t], \|f(s) - f(c)\| \leq g(s) - g(c) + \varepsilon(s - c)\}$. By Def, J is an interval.

Since f and g are continuous, J is a closed interval, hence J is of the form $[c, t]$. (**4.9.3.1 Proof**) If $t < b$, then for $h > 0$, sufficiently small, we have:

$$f(t+h) - f(t) = hf'(t) + o(h), \; g(t+h) - g(t) = hg'(t) + o(h).$$

There exists $\delta > 0$, for $h \in [0, \delta]$:

$$\|f(t+h) - f(t)\| \leq \|f'(t)\| \cdot h + \frac{3}{2}\varepsilon h, \; g(t+h) - g(t) \geq g'(t)h - \frac{1}{2}\varepsilon h.$$

So, $\|f(t+h) - f(t)\| \leq g(t+h) - g(t) + \varepsilon h$. Moreover, $\|f(t) - f(c)\| \leq g(t) - g(c) + \varepsilon(t - c)$.

Thus, $\|f(t+h) - f(c)\| \leq g(t+h) - g(c) + \varepsilon(t + h - c)$. This implies $[c, t+\delta] \subseteq J$, which is a contradiction! Hence, $\|f(b) - f(c)\| \leq g(b) - g(c) + \varepsilon(b - c)$. For the same reason, $\|f(c) - f(a)\| \leq g(c) - g(a) + \varepsilon(c - a)$.

Therefore, $\|f(b) - f(a)\| \leq g(b) - g(a) + \varepsilon(b - a)$. Since ε is arbitrary, we conclude $\|f(b) - f(a)\| \leq g(b) - g(a)$.

5.3.3 Taylor-Lagrange Formula

Let $n \in \mathbb{N}$, E and F be normed vector spaces over $\mathbb{R}$. Let $U \subseteq E$ be open and $f : U \to F$ be a mapping that is $(n+1)$-times differentiable. Let $p \in U$ and $h \in E$ s.t. $p + th \in U$ for $t \in]0, 1[$. Let $M = \sup_{t \in [0,1]} \|D^n f(p+th)\|$. Then:

$$\|f(p+th) - \sum_{k=0}^{n} \frac{1}{k!} D^k f(p)(\underbrace{h, \ldots, h}_{k \text{ copies}})\| \leq \frac{M}{(n+1)!} \|h\|^{n+1}.$$

5.3.3.1 Proof Consider $\phi : [0, 1] \to F$, defined by:

$$\phi(t) = \sum_{k=0}^{n} \frac{(1-t)^k}{k!} D^k f(p+th)(h, \ldots, h).$$

We have $\phi(1) = f(p+th)$ and $\phi(0) = \sum_{k=0}^{n} \frac{1}{k!} D^k f(p)(h, \ldots, h)$. By the properties of the derivative, we can show that:

$$\phi'(t) = \Sigma_{k=0}^{n} \frac{(1-t)^k}{k!} D^{k+1} f(p+th)(\underbrace{h, \ldots, h}_{k+1 \text{ copies}}) - \Sigma_{k=1}^{n} \frac{(1-t)^{k-1}}{(k-1)!} D^k f(p+th)(\underbrace{h, \ldots, h}_{k \text{ copies}})$$

$$= \Sigma_{k=0}^{n} \frac{(1-t)^k}{k!} D^{k+1} f(p+th)(\underbrace{h, \ldots, h}_{k+1 \text{ copies}}) - \Sigma_{k=0}^{n-1} \frac{(1-t)^k}{k!} D^{k+1} f(p+th)(\underbrace{h, \ldots, h}_{k+1 \text{ copies}})$$

$$\|\phi'(t)\| = \frac{(1-t)^n}{n!} D^{n+1} f(p+th)(h, \ldots, h) \leq M \frac{(1-t)^n}{n!} \|h\|^{n+1} = (-M \frac{(1-t)^{n+1} \|h\|^{n+1}}{(n+1)!})'.$$

By **Gronwall's inequality**, we get:

$$\|\phi(1) - \phi(0)\| \leq \frac{M}{(n+1)!} \|h\|^{n+1}.$$

5.3.3.2 Partial Differentials Let $n \in \mathbb{N}_{\geq 1}$, $E_1, \ldots, E_n$, and F be normed vector spaces over a complete valued field $(K, |\cdot|)$. Let $U \subseteq E_1 \times \cdots \times E_n$ be an open set, $p = (p_1, \ldots, p_n) \in U$, $i \in \{1, \ldots, n\}$. $f : U \to F$. If there exists an open neighborhood U_i of p_i in E_i s.t. the mapping $U_i \to F$, $x_i \mapsto f(p_1, \ldots, p_{i-1}, x_i, p_{i+1}, \ldots, p_n)$ is well defined and differentiable at p_i, we denote by $\frac{\partial f}{\partial x_i}(p)$ the differential at p of this mapping $U_i \to F$ and say that f admits the i-th partial differential at p.

5.3.3.3 Prop Suppose that $(K, |\cdot|) = \mathbb{R}$. Suppose that f has all partial differentials on U as $\frac{\partial f}{\partial x_i} : U \to \mathscr{L}(E_i, F)$ is continuous for any $i \in \{1, \dots, n\}$. Then f is of class C^1, and for any $h = (h_1, \dots, h_n) \in E1 \times \cdots \times E_n$, $\forall p \in U$, $d_p f(h) = \Sigma_{i=1}^n \frac{\partial f}{\partial x_i}(p)(h_i)$.

5.3.3.4 Proof By induction, it suffices to treat the case where $n = 2$. Let $p = (a, b) \in E_1 \times E_2$, $\forall \varepsilon \in \mathbb{R}_+$, $\exists \delta > 0$, $\forall (h, k) \in E_1 \times E_2$, $\max\{|h|, |k|\} \le \delta$, one has:

$$\left\| \frac{\partial f}{\partial x_2}(a+h, b+k) - \frac{\partial f}{\partial x_2}(a, b) \right\| \le \varepsilon$$

(by the continuity of $\frac{\partial f}{\partial x_2}$).

Consider the mapping $\phi : [0, 1] \to F$, $\phi(t) = f(a+h, b+tk) - f(a+h, b) - t\frac{\partial f}{\partial x_2}(a+h, b)(k)$.

$$\phi'(t) = \frac{\partial f}{\partial x_2}(a+h, b+tk)(k) - \frac{\partial}{\partial x_2}(a+h, b)(k)$$

$$\|\phi(1) - \phi(0)\| \le 2\varepsilon \|k\|$$

(by the **Gronwall's inequality**)

$$\|f(a+h, b+k) - f(a+h, b) - \frac{\partial f}{\partial x_2}(a+h, b)(k)\| \le 2\varepsilon \|k\|$$

So, $\|f(a+h, b+k) - f(a+h, b) - \frac{\partial}{\partial x_2}(a+h, b)(k)\| = o(\max\{\|h\|, \|k\|\})$ (since f has second partial derivatives).
(f has a 1-st partial differential:)

$$\left\| \left(f(a+h, b) - f(a, b) - \frac{\partial f}{\partial x_1}(a, b)(h) \right) \right\| = o(\max\{\|h\|, \|k\|\})$$

(Continuity of $\frac{\partial}{\partial x_2}$:)

$$\left\| \frac{\partial}{\partial x_2}(a+h, b+k) - \frac{\partial}{\partial x_2}(a, b)(k) \right\| = o(\max\{\|h\|, \|k\|\})$$

Taking the sum, we get:

$$\|f(a+h, b+k) - f(a, b) - \left(\frac{\partial f}{\partial x_1}(a, b)(h) + \frac{\partial f}{\partial x_2}(a, b)(k) \right) \| = o(\max\{\|h\|, \|k\|\})$$

5.3.3.5 Theorem Let E and F be normed vector spaces over $\mathbb{R}$, $U \subseteq E$ open, and $(f_n)_{n \in \mathbb{N}}$ be a sequence of differentiable mappings from U to F. Let $g : U \to \mathscr{L}(E, F)$, suppose that:

1. $(df_n)_{n \in \mathbb{N}}$ converges uniformly to g,
2. $(f_n)_{n \in \mathbb{N}}$ converges pointwisely to some mapping $f : U \to F$ ($\forall p \in U$, $(f_n(p))_{n \in \mathbb{N}}$ converges to $f(p)$).

Then f is differentiable and $df = g$.

5.3.3.6 Proof Let $p \in U$, $\forall (m, n) \in \mathbb{N}^2$, we have:

$$\|f_n(x) - f_m(x) - (f_n(p) - f_m(p))\| \le \left(\sup_{\xi \in U} \|d_\xi f_m - d_\xi f_n\| \right) \|x - p\|$$

(mean value inequality).

Taking $\lim_{m \to +\infty}$, we get:

$$\|(f_n(x) - f(x)) - (f_n(p) - f(p))\| \le \varepsilon_n \|x - p\|$$

where

$$\varepsilon_n = \sup_{\xi \in U} \|d_\xi f_n - g\|.$$

So,

$$\|f(x) - f(p) - g(p)(x-p)\| \le \|(f(x) - f_n(x)) - (f(p) - f_n(p))\| + \|f_n(x) - f_n(p) - df_n(x-p)\| \le 2\varepsilon_n \|x-p\| + \|f_n(x) - f_n(p) - df_n(x-p)\|.$$

$$\limsup_{x \to p} \frac{\|f(x) - f(p) - g(p)(x-p)\|}{\|x - p\|} \le 2\varepsilon_n$$

Taking $\lim_{n \to +\infty}$, we get

$$\limsup_{x \to p} \frac{\|f(x) - f(p) - g(p)(x-p)\|}{\|x - p\|} = 0.$$

5.4 Symmetric Group

5.4.1 Permutations, n-cycle and Transposition

Let X be a set. We denote by $\mathfrak{S}_X$ the set of all bijections from X to itself. The elements of $\mathfrak{S}_X$ we call *permutations* if the set X is finite. If $x_1, \dots, x_n \in X$ are distinct elements, then $(x_1, \dots, x_n) \in \mathfrak{S}_X$ s.t. $x_i \mapsto x_{i+1} \quad i = 1, \dots, n-1; x_n \mapsto x_1$ is called an *n-cycle*. A 2-cycle is called a *transposition*.

5.4.1.1 Example $X = \{1, \dots, 7\}$.
$(2\ 3) \circ (4\ 2\ 1) = (1\ 4\ 3\ 2)$.

5.4.1.2 Def We denote with $\mathrm{orb}_\sigma(x) = \{\underbrace{\sigma \circ \cdots \circ \sigma}_{n \text{ times}}(x),\ n \in \mathbb{N}\}$, $x \in X, \sigma \in \mathfrak{S}_X$.

5.4.1.3 Prop If $\mathrm{orb}_\sigma(x)$ is a finite set of d elements, then one has $\sigma^d(x) = x$, $\mathrm{orb}_\sigma(x) = \{x, \sigma(x), \dots, \sigma^{d-1}(x)\}$, moreover $\sigma^{-1}(x) \in \mathrm{orb}_\sigma(x)$.

5.4.1.4 Proof The set $\{(n,m) \in \mathbb{N}^2,\ n < m,\ \sigma^n(x) = \sigma^m(x)\}$ is not empty. Let $d' := \min\{m - n,\ (n,m) \in \mathbb{N}^2,\ n < m,\ \sigma^n(x) = \sigma^m(x)\}$, where $\sigma^n(x) = \sigma^m(x) \Leftrightarrow x = \sigma^{m-n}(x)$.

Therefore, $x, \sigma(x), \dots, \sigma^{d'-1}(x)$ are all distinct.

$\{x, \sigma(x), \dots, \sigma^{d-1}(x)\} \subseteq \mathrm{orb}_\sigma(x)$. Now use the euclidean division, $h = qd' + r \quad r < d'$. $\sigma^h(x) = \sigma^r(x),\ \sigma^{qd'}(x) = x$. So we have $d' = d$.

Moreover, $\sigma^{-1}(x) = \sigma^{d-1}(x) \in \mathrm{orb}_\sigma(x)$.

5.4.1.5 Remark Let $Y \subseteq X$, then we have a morphism of groups $\mathfrak{S}_Y \to \mathfrak{S}_X \quad \sigma \mapsto \left(x \mapsto \begin{cases} \sigma(x) & \text{if } x \in Y \\ x & \text{on } X \backslash Y \end{cases} \right)$.

5.4.1.6 Theorem Let X be a finite set and let $\alpha \in \mathfrak{S}_X$. Then exist $d \in \mathbb{N}$ and $(n_1, \dots, n_d) \in \mathbb{N}^d_{\geq 2}$, and pairwise disjoint subsets $X_1, \dots, X_d$ of X of coordinates $n_1, \dots, n_d$, together with n_i-circles τ_i of X_i s.t. $\sigma = \tau_1 \circ \cdots \circ \tau_d$.

In other words, any permutation can be decomposed in the composition of finitely many cycles on disjoint subsets.

5.4.1.7 Proof By induction on the cardinality N of X. The case $\sigma = \mathrm{id}_X$ is trival (with d=0). So the case when $N = 0, 1$ is clear.

Assume $N \geq 2$. Take $x \in X$ s.t. $\sigma(x) \neq x$ and let $X_1 = \mathrm{orb}_\sigma(x)$. Let $Y = X \backslash X_1$, $\forall y \in Y$ we have that $\sigma(y) \in Y$. (because if $\sigma(y) \in X$, by the previous prop $\sigma^{-1}(\sigma(y)) \in X \Rightarrow y \in X$). Let $\tau = \sigma\ |_y \in \mathfrak{S}_Y$. Use the induction hypothesis, we get $X_2, \dots, X_d$ of cardinalities $n_2, \dots, n_d$, and n_i-cycles τ_i s.t. $\tau = \tau_2 \circ \cdots \circ \tau_d$. Consider $\tau_1 = \sigma\ |_{X_1}$, then τ_1 is a n_1-cycle of X_1.

$X_1 = \{x, \sigma x, \dots, \sigma^{m-1}(x)\} \Rightarrow \tau = \tau_1 \circ \cdots \circ \tau_d$.

This theorem says that the group of permutations is generated by cycles.

5.4.1.8 Corollary Let X be a finite set. Then $\mathfrak{S}_X$ is generated by transposition.

5.4.1.9 Proof Thanks to the theorem before, it is enough to decompose cycles in terms of transpositions.
$(x_1, x_2, \dots, x_n) = (x_1, x_2) \circ (x_2, \dots, x_n) = \cdots = (x_1, x_2) \circ \cdots \circ (x_{n-1}, x_n)$.

5.4.1.10 Remark The decomposition in terms of transposition is not unique.

5.4.2 Adjacent

Consider $\mathfrak{S}_n$ a transposition is called adjacent if is of the form $(j, j+1)$, for $j = 1, \dots, n-1$.

5.4.2.1 Corollary $\mathfrak{S}_n$ can be generated by adjacent transpositions.

5.4.2.2 Proof It's enough to decomprise transpositions, because of the previous **Cor**.
$i < j$: $(i, j) = (i, i+1) \circ (i+1, i+2) \circ \cdots \circ (j-1, j) \circ (j-2, j-1) \circ \cdots \circ (i, i+1)$.

5.4.3 Some Extra Information on $\mathfrak{S}_n$

5.4.3.1 Cayley Theorem Any finite group can be embedded (injection morphism) in a $\mathfrak{S}_n$ for some $n \in \mathbb{N}$.

5.4.3.2 Proof Let $|G| = n$, $\varphi : G \to \mathfrak{S}_G \cong \mathfrak{S}_n$, $g \mapsto lg$, $lg(x) = gx$.

5.4.3.3 Theorem Assume that X is finite, and $\sigma \in \mathfrak{S}_X$ can be written as $\sigma = \tau_1 \circ \cdots \circ \tau_d$ where each τ_i is a transposition. We define $\text{sgn}(\sigma) = (-1)^d$.

This is a well-defined function and, moreover, $\text{sgn} : \mathfrak{S}_X \to \{1, -1\}$ is a morphism of groups.

5.4.3.4 Proof Let us define the map $\phi : \mathfrak{S}_X \to \mathbb{Q} \setminus \{0\}$, $\phi(\sigma) = \prod_{(i,j)\in\{1,\ldots,n\}^2} \frac{\sigma(i)-\sigma(j)}{i-j}$.

Let $\theta = \{(i,j) \in \{1, \ldots, n\}^2 | i < j\}$.

$\phi(\sigma \circ \tau) = \prod_{(i,j)\in\theta} \frac{\sigma(\tau(i))-\sigma(\tau(j))}{i-j} = \left(\prod_{(i,j)\in\theta} \frac{\sigma(\tau(i))-\sigma(\tau(j))}{\tau(i)-\tau(j)}\right) \left(\prod_{(i,j)\in\theta} \frac{\tau(i)-\tau(j)}{i-j}\right) = \phi(\sigma)\phi(\tau)$

$\phi(\sigma) = -1$ if σ is a transposition. Therefore, since $\sigma = \tau_1 \circ \cdots \circ \tau_d$, $\phi(\sigma) = (-1)^d$.

5.5 Symmetry of Multilinear Maps

In this section, we fix a commutative unitary ring K, and two $K-$modules E, F.

5.5.0.1 Def Let $n \in \mathbb{N}$ and $f \in \text{Hom}^{(n)}(E^n, F)$ if for any $\sigma \in \mathfrak{S}_n$ one has $f(x_1, \cdots, x_n) = f(x_{\sigma(1)}, \cdots, x_{\sigma(n)})$. $\forall(x_1, \cdots, x_n) \in E^n$. We say that f is symmetric.

If for any $(i,j) \in \{1, \cdots, n\}^2$ and any $(x_1, \cdots, x_n) \in E^n$ s.t. $x_i = x_j$ one has that $f(x_1, \cdots, x_n) = 0$, we say that f is alternating.

5.5.0.2 Prop Suppose that $f \in \text{Hom}^{(n)}(E^n, F)$ is alternating, then $f(x_1, \cdots, x_n) = \text{sgn}(\sigma) f(x_{\sigma(1)}, \cdots, x_{\sigma(n)})$, $\forall(x_1, \cdots, x_n)$ E^n, $\forall \sigma \in \mathfrak{S}_n$.

5.5.0.3 Proof By what we proved on permutations, it is enough to prove the Prop for adjacent transpositions. Let $i \in \{1, \cdots, n-1\}$, then $0 = f(x_1, \cdots, x_{i-1}, x_i + x_{i+1}, x_i + x_{i+1}, x_{i+2}, \cdots, x_n) = f(x_1, \cdots, x_n) + f(x_1, \cdots, x_{i-1}, x_{i+1}, x_i, \cdots, x_n)$.

5.5.0.4 Def We denote by $\text{Hom}_S^{(n)}(E^n, F)$ and $\text{Hom}_a^{(n)}(E^n, F)$ the set of symmetric and alternating $n-$linear maps from E^n to F. These are sub-K-modules of $\text{Hom}^{(n)}(E, F)$ and when $n = 1$, $\text{Hom}_S^{(1)}(E, F) = \text{Hom}_a^{(1)}(E, F) = \text{Hom}(E, F)$.

E, F are two normed vector spaces over $\mathbb{R}$. $f : E \to F$, is differentiable (twice): $df : E \to \mathscr{L}(E, F)$, $D^2 f : E \to \mathscr{L}(E, F)) \cong \mathscr{L}^2(E^2, F)$.

5.5.0.5 Theorem (Schweiz) $U \subseteq E$ is an open set, $f : U \to F$ is a function of class C^n. Then for any $p \in U$ $D^n f(p) \in \mathscr{L}^n(E^n, F)$ is symmetric.

5.5.0.6 Proof By induction and by the fact that permutations are decomposed in transpositions, we can induce to prove only the case $n = 2$.

$d_{p+u} f - d_p f = D^2 f(p)(u, \cdot) + o(u)$, $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ s.t. $0 < \|u\| < \delta$, then take $h, k \in E$, $0 < \|h\| < \frac{\delta}{4}, x \in B(p, \frac{\delta}{2})$, $0 < \|k\| \frac{delta}{4}$.

for any $x \in B(p, \frac{\delta}{2})$, let's introduce the following function: $\phi(x) = f(x+k) - f(x) - D^2 f(p)(k, x)$.

We use the "mean value inequality on ϕ":

$$\|\phi(p+h) - \phi(p)\| = \|f(x+h+k) - f(p+h) - D^2 f(p)(K, p+h) - f(p+k) + f(p) + D^2 f(p)(k, p)\|\|$$

$$= \|f(x+h+k) - f(p+h) - f(p+k) - D^2 f(p)(k, h) \leq (\sup_{t\in[0,1]} \|d_{p+th}(\phi)\|)\|h\|)$$

$$\|d_{p+th}(\phi)\| = \|d_{p+th+k} f - d_{p+th} f - D^2 f(p)(k, \cdot)\|$$

Add and subtract $d_p f$, $D^2 f(p)(th, \cdot)$, by applying triangle inequality, we can get:

$$\leq \|d_{p+th+k} f - d_p f - D^2 f(p)(k+th, \cdot)\| + \|d_{p+th} f - d_p f - D^2 f(p)(th, \cdot)\| \leq \|d_{p+u} f - d_p f - D^2 f(p)(u, \cdot)\| < \varepsilon\|u\|$$

$$\varepsilon\|th + k\| + \varepsilon\|th\| \leq 2\varepsilon(\|h\| + \|k\|)$$

$$f(p+h+k) - f(p+k) - f(p+h) - D^2 f(p)(k, h) + f(p) = o(\max\{\|h\|, \|k\|\}^2)$$

Exchange the role of h,k, then we get:

$$\|f(p+h+k) - f(p+k) - f(p+h) - D^2 f(p)(h, k) + f\| = o(\max\{\|h\|, \|k\|\|^2)$$

$$D^2 f(p)(k, h) - D^2 f(p)(h, k)\| = o(\max\{\|h\|, \|k\|\}^2)$$

This implies that the LHS is 0.

5.5.0.7 Def Let E, F be normal vector spaces over a complete valued field $(K, |\cdot|)$. Let $U \subseteq E$, $V \subseteq F$ be open subsets and $f : U \to V$ is a bijection.

1. If f and f^{-1} are both continuous, we say that f is a homeomorphism.
2. If f and F^{-1} are both of class e^n, we say that f is a C^n-diffeomorphism. If (2) is true for any $n \in \mathbb{N}$, we say that f is a C^∞-diffeomorphism.

5.5.0.8 Prop Let E and F be Bannach space. Let $I(E, F) \subseteq \mathscr{L}(E, F)$ be the set of linear, continuous, invertible maps s.t. $\|\phi^{-1}\| < +\infty$. Then $I(E, F)$ is open in $\mathscr{L}(E, F)^V$. Moreover, the map $i : I(E, F) \to I(F, E) \quad \phi \mapsto \phi^{-1}$ is a C^1-diffeomorphism.

5.5.0.9 Proof Let $\phi \in I(E, F)$, we want to show that $\phi - \psi \in I(E, F)$ for $\psi \in \mathscr{L}(E, F)$ s.t. $\|\psi\| < \frac{1}{\|\psi^{-1}\|}$.

Notice that $\phi - \psi = \phi \circ (\mathrm{Id}_E - \phi^{-1} \circ \psi)$. Since $\|\phi^{-1} \circ \psi\| \leq \|\phi^{-1}\| \|\psi\| < 1$, it means that the series $\sum\limits_{n \in \mathbb{N}} (\varphi^{-1} \circ \psi)^{\circ n}$ is absolutely convergent in $\mathscr{L}(E, E)$. This series is the inverse of $(Id_E - \varphi^{-1}\psi)$. $(Id_e - \varphi^{-1}\psi) \circ \sum\limits_{n=0}^{N-1} (\varphi^{-1} \circ \psi) = Id_E - (\varphi^{-1} \circ \psi)^{\circ N}$. Take $\lim\limits_{N \to +\infty}$, then $(\varphi - \psi)^{-1} = \sum\limits_{n \in \mathbb{N}} (\varphi^{-1} \circ \psi)^{\circ n} \circ \varphi^{-1}$ and $(\varphi - \psi)^{-1} = \varphi^{-1} + \varphi^{-1} \circ \psi \circ \varphi^{-1} + o(\|\psi\|)$. Replace the inverse with i: $i(\varphi - \psi) - i(\varphi) = \varphi^{-1} + \varphi^{-1} \circ \psi \circ \varphi^{-1} + o(\|\psi\|)$. Then $d_\varphi i(\psi) = i(\varphi) \circ (-\psi) \circ i(\varphi)$. So i is differentiable. Moreover, i and i^{-1} are continuous.

5.5.0.10 Remark By induction, we can show that i is a $C^{+\infty}$-diffeomorphism.

5.5.0.11 Prop Let $n \in \mathbb{N} \bigcup \{\infty\}$ Let E, F, G be normed vector spaces over a complete valued field $(K, |\cdot|)$ $U \subseteq E, V \subseteq F$ be open sets. $f : U \to V$ $g : V \to G$ be mappings of class C^n, then $g \circ f$ also of class C^n.

5.5.0.12 Proof The case where $n = 0$ is known.

Denote by $\Phi : \mathscr{L}(E, F) \times E \to F \quad (\beta, \alpha) \mapsto \beta \circ \alpha$. Φ is a bounded bilinear mapping: $\|\Phi(\beta, \alpha)\| \leq \|\beta\| \cdot \|\alpha\|$. Suppose that $n \geq 1$ and the statement is true for mappings of class C^{n-1} $g \circ f$ is differentiable. $\forall p \in U \quad d_p(g \circ f) = d_{f(p)} g \circ d_p f$, $D^1(g \circ f) : U \to \mathscr{L}(E, G)$, $D^1 = \Phi \circ (D^1 g \circ f, D^1 f)$, $(D^1 g \circ f, D^1 f) : U \to \mathscr{L}(F, G) \times \mathscr{L}(E, F) \quad p \mapsto (d_{f(p)} g, d_p f)$. $d_{\beta_0, \alpha_0} \Phi(\beta, \alpha) = \beta_0 \circ \alpha + \beta_0 \circ \alpha$, $D^1 \Phi : \mathscr{L}(F, G) \times \mathscr{L}(E, F) \to \mathscr{L}(\mathscr{L}(F, G) \times \mathscr{L}(E, F), \mathscr{L}(E, G)) \quad (\alpha_0, \beta_0) \mapsto ((\alpha, \beta) \mapsto \beta_0 = \alpha + \beta_0 \alpha_0)$.

Since g, f are of class C^n $D^1 f, D^1 g$ are of class C^{n-1} Thus, by induction hypothesis,

$$(D^1 g \circ f, D^1 f)$$

is of class C^{n-1}. Since Φ is of class C^∞, we obtain that $D^1(g \circ f)$ is of class C^{n-1} then $g \circ f$ is of class C^n.

5.5.0.13 Prop Let E and F be Banach space over a complete valued field $(K, |\cdot|)$. U and V be open subsets of E and F respectively.$n \in \mathbb{N} \bigcup \{\infty\}$ and $f : U \to V$ be a bijection. If f is of class C^n, then f^{-1} is differentiable, then f^{-1} is of class C^n.

5.5.0.14 Proof $f \circ f^{-1} = Id_V$. $\forall y \in V$, $d_y(f \circ f^{-1}) = d_{f^{-1}(p)} f \circ d_y f^{-1} = Id_F$. For $x \in U, y = f(x)$, $d_y(f \circ f^{-1}) = d_x f \circ d_y f^{-1} = Id_F$, $d_x(f^{-1} \circ f) = d_y f \circ d_x f^{-1} = Id_E$. So $d_y f^{-1} - (d_x f)^{-1}$, that is $D^1 f^{-1} = \iota \circ (D^1 f \circ f^{-1})$ where $\iota : I(E, F) \to I(F, E) \quad \phi \mapsto \phi^{-1}$. Suppose that f^{-1} is of class C^{n-1} then $D^1 f^{-1} = \iota D^1 f \circ f^{-1}$ is of class C^{n-1}.

5.5.0.15 Local Inversion Theorem Let E and F be Banach space over $\mathbb{R}$, $U \in E$ open, $f : U \to F$ be a mapping of class C^n and $a \in U$. Suppose that $d_a f \in I(E, F)$(which means $d_a f$ is invertible and of bounded inverse), then there exists open neighborhoods V and W of a and $f(a)$ respectively, s.t.:

- $V \subseteq U$ and $f(V) \subseteq W$
- The restriction of f to V defines a bijection from V to W
- $(f |_V)^{-1} W \to V$ is of class C^n

5.5.0.16 Proof For $y \in F$ consider the mapping:

$$\phi_y : U \to F$$
$$x \mapsto x - (d_a f)^{-1}(f(x) - y)$$

$f(x) = y$ iff $\phi_y(x) = x$ i.e. x is a fix point of ϕ_y ϕ_y is of class C^1 and $d_x \phi_y(v) = v - d_a f^{-1}(d_x f(v))$. $\forall v$, $d_a \phi y^{(v)} = 0$. By the continuity of $D^1 f$, there exists $r > 0$ s.t. $\overline{\mathcal{B}}(a, r) \subseteq U$ and $\forall y \in F, \forall x \in \overline{\mathcal{B}}(a, r)$, $\|d_x \phi_y\| \leq \frac{1}{2}$.

By the mean value inequality. $\forall (x_1, x_2) \in \overline{\mathcal{B}}(a, r)$, $\|\phi_y(x_1) - \phi_{=y(x_2)}\| \leq \frac{1}{2} \|x_1 - x_2\|$. Hence ϕ_y is contraction.

By the boundedness of $(d_af)^{-1}$, $\exists\delta > 0$ s.t. $\forall y \in \overline{\mathcal{B}}(f(a),\delta) \quad \|(d_af)^{-1}(f(a)-y)\| \leq \frac{r}{2}$. Then $\forall x \in \overline{\mathcal{B}}(a,r)$ $y \in \overline{\mathcal{B}}(f(a),\delta)$ $\|\phi_y(x) - a\| \leq \|\phi_y(x) - \phi_y(a)\| + \|\phi_y(a) - a\| \leq \frac{1}{2}\|x - a\| + \frac{r}{2} \leq \frac{r}{2} + \frac{r}{2} = r$, $\phi_y(\overline{\dashv,\nabla}) \in \overline{\mathcal{B}}(a,r)$. By the fixed point theorem, $\exists g : \overline{\mathcal{B}}(f(a),\delta) \to \overline{\mathcal{B}}(a,r)$ sending y to the fixed point of ϕ_y. Let $W = \mathcal{B}(f(a),g)$,then $g\ |_W \colon W \to V$ is the inverse of $f\ |_V \colon V \to W$. Hence $f^{-1}(W) = V$ is open.

In the following, we prove that g is of class C^n on an open neighborhood of $f(a)$. By reducing V and W, we may assume that $\forall x \in V$, $d_xf \in I(E,F)$. Let $x_0 \in V$ $y_0 = f(x_0) \quad x_0 = g(y_0)$, $y - y_0 = f(g(y)) - f(g(y_0)) = d_{x_0}f(g(y) - g(y_0)) + o(\|g(y) - g(y_0)\|)$. So g$(y) - g(y_0) = (d_xf)^{-1}(y - y_0) + o(\|g(y) - g(y_0)\|)$. Thus leads to $g(y) - g(y_0) = O(\|y - y_0\|)$ $(\exists\epsilon > 0 \quad (1-\epsilon)\|g(y) - g(y_0)\| \leq \|d_{x_0}f\|^{-1}$ when $\|y - y_0\|$ is sufficiently small).

So $d_{y_0}g = (d_xf)^{-1}$. By the previous Prop, g is of class C^n.

Chapter 6

Integration

6.1 Integral Operators

6.1.1 Riesz Space

Let Ω be a non-empty set and S be a vector subspace of $\mathbb{R}^{\Omega}$. If for all $(f,g) \in S^2$, $f \wedge g : \Omega \to \mathbb{R}, \quad \omega \mapsto \min\{f(\omega), g(\omega)\} \in S$, we say that S is a Riesz space.

6.1.1.1 Prop

1. For all $(f,g) \in S^2$, $f \vee g : \Omega \to \mathbb{R}, \quad \omega \mapsto \max\{f(\omega), g(\omega)\}$ and $f \vee g \in S$.
2. For all $f \in S$, $|f| : \Omega \to \mathbb{R}, \quad \omega \mapsto |f(\omega)|$ and $|f| \in S$.

6.1.1.2 Proof

1. $f \vee g = f + g - f \wedge g$
2. $|f| = f \vee (-f)$

6.1.2 Integral Operators

We call an integral operator on S any $\mathbb{R}$-linear mapping $I : S \to \mathbb{R}$ that satisfies the following conditions:

(1) If $f \in S$ is s.t. for all $\omega \in \Omega, f(\omega) \geq 0$, then $I(f) \geq 0$.

(2) If $(f_n)_{n\in\mathbb{N}}$ is a decreasing sequence of elements in S s.t. for all $\omega \in \Omega, \lim_{n\to+\infty} f_n(\omega) = 0$, then $\lim_{n\to+\infty} I(f_n) = 0$.

6.1.2.1 Prop Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence in S that converges pointwisely to some $f \in S$. Then $\lim_{n\to+\infty} I(f_n) = I(f)$.

6.1.2.2 Proof Let $g_n = f - f_n \in S$, and $(g_n)_{n\in\mathbb{N}}$ is decreasing and converges pointwise to 0. Then $\lim_{n\to+\infty} I(g_n) = 0$, so $\lim_{n\to+\infty} I(f_n) = I(f)$.

6.1.2.3 Prop Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence in S and $f \in S$. If $f \leq \lim_{n\to+\infty} f_n$, then $I(f) \leq \lim_{n\to+\infty} I(f_n)$.

6.1.2.4 Proof We have $f = \lim_{n\to+\infty} f \wedge f_n$, so $I(f) = \lim_{n\to+\infty} I(f \wedge f_n) \leq \lim_{n\to+\infty} I(f_n)$.

6.1.2.5 Example

(1) Let $\Omega = \mathbb{R}$ and S be the vector subspace of $\mathbb{R}^{\mathbb{R}}$ generated by mappings of the form $\mathbb{1}_{]a,b]}$, where $(a,b) \in \mathbb{R}^2, a < b$.

$$\mathbb{1}_{]a,b]} = \begin{cases} 1, & x \in]a,b], \\ 0, & \text{else.} \end{cases}$$

Any element of S is of the form $\sum_{i=1}^{n} \lambda_i \mathbb{1}_{]a_i,b_i]}$. Define $I : S \to \mathbb{R}$ as $I\left(\sum_{i=1}^{n} \lambda_i \mathbb{1}_{]a_i,b_i]}\right) = \sum_{i=1}^{n} \lambda_i (b_i - a_i)$. More generally, if $\varphi : \mathbb{R} \to \mathbb{R}$ is increasing and right-continuous ($\forall x \in \mathbb{R}, \lim_{\epsilon\to 0^+} \varphi(x+\epsilon) = \varphi(x)$), we define $I_\varphi : S \to \mathbb{R}$, with $I_\varphi\left(\sum_{i=1}^{n} \lambda_i \mathbb{1}_{]a_i,b_i]}\right) = \sum_{i=1}^{n} \lambda_i (\varphi(b_i) - \varphi(a_i))$.

(2) (*Radon measure*) Let Ω be a quasi-compact topological space, and $S = C^0(\Omega) := \{f : \Omega \to \mathbb{R} \mid f \text{ is continuous}\}$. Let $I : S \to \mathbb{R}$ be $\mathbb{R}$-linear, s.t. for all $f \in S, f \geq 0$, one has $I(f) \geq 0$.

6.1.3 Dini's Theorem

Let $(f_n)_{n\in\mathbb{N}}$ be a decreasing sequence in $C^0(\Omega)$, that converges pointwise to some $f \in C^0(\Omega)$. Then $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f.

6.1.3.1 Proof Let $g_n = f_n - f \geq 0$. Fix $\epsilon > 0$. For all $n \in \mathbb{N}$, let $U_n = \{\omega \in \Omega \mid g_n(\omega) < \epsilon\}$ which is open. Moreover, $\bigcup_{n\in\mathbb{N}} U_n = \Omega \quad (U_0 \subseteq U_1 \subseteq \cdots)$. Since Ω is quasi-compact, there exists $N \in \mathbb{N}$ s.t. $\Omega = U_N$. Therefore, for all $n \in \mathbb{N}$, $n \geq N$, and for all $\omega \in \Omega$, we have $g_n(\omega) < \epsilon$.

6.1.3.2 Corollary If $(f_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ is decreasing and converges point-wise to 0, then $\|f_n\|_{\sup} := \sup_{\omega\in\Omega} |f_n(\omega)|$ converges to 0 as $n \to +\infty$. For all $n \in \mathbb{N}$, we have $f_n \leq \|f_n\|_{\sup} \cdot \mathbb{1}_\Omega$, so $0 \leq I(f_n) \leq \|f_n\|_{\sup} I(\mathbb{1}_\Omega) \to 0 \quad (n \to +\infty)$. (If $f \leq g$, then $g - f \geq 0$, so $I(g-f) = I(g) - I(f) \geq 0$, and thus $I(g) \geq I(f)$).

6.1.4 σ-Algebra

We call a σ-algebra any subset $\mathscr{A}$ of $\mathscr{P}(\Omega)$ that satisfies the following conditions:

- $\varnothing \in \mathscr{A}$,
- If $A \in \mathscr{A}$, then $\Omega \setminus A \in \mathscr{A}$,
- If $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

6.1.5 Measure

Given a σ-algebra $\mathscr{A}$ on Ω, we mean by a measure on $(\Omega, \mathscr{A})$ any mapping $\mu : \mathscr{A} \to [0, +\infty]$ s.t.

- $\mu(\varnothing) = 0$,
- If $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$ are pairwise disjoint, then $mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

6.2 Riemann Integral

In this section, we fix a Riesz space S and an integral operator $I : S \to \mathbb{R}$.

6.2.1 I-Riemann Integrable

For any $f : \Omega \to \mathbb{R}$, define $I^*(f) := \inf_{\mu\in S, \mu\geq f} I(\mu), \quad I_*(f) := \sup_{l\in S, l\leq f} I(l)$. If $I^*(f) = I_*(f)$, then we say that f is I-Riemann integrable and denote the value by $I(f)$ (or $I_*(f)$).

6.2.2 Riemann Integrable Mappings Linear Extension Theorem

The set $\mathscr{R}$ of all I-Riemann integrable mappings forms a vector space of $\mathbb{R}^\Omega$ that contains S. Moreover, $I : \mathscr{R} \to \mathbb{R}$ is a $\mathbb{R}$ linear mapping that extends $I : S \to \mathbb{R}$.

6.2.2.1 Proof For all $h \in S$, we have $I^*(h) = I_*(h) = I(h)$, so $h \in \mathscr{R}$. Let $(f_1, f_2) \in \mathscr{R}$. If $(\mu_1, \mu_2) \in S^2, \mu_1 \geq f_1, \mu_2 \geq f_2$, then $\mu_1 + \mu_2 \in S$ and $\mu_1 + \mu_2 \geq f_1 + f_2$, so $I(\mu_1) + I(\mu_2) \geq I^*(f_1 + f_2)$. Taking the infimum with respect to (μ_1, μ_2), we get $I^*(f_1) + I^*(f_2) \geq I^*(f_1 + f_2)$. Similarly, $I_*(f_1) + I_*(f_2) \leq I_*(f_1 + f_2)$. Hence, $I^*(f_1 + f_2) = I_*(f_1 + f_2) = I(f_1) + I(f_2)$.

Let $f : \Omega \to \mathbb{R}$ be a mapping, and $\lambda \in \mathbb{R}_{>0}$. Then $I^*(\lambda f) = \inf_{\mu\in S, \mu\geq\lambda f} I(\mu) = \lambda I^*(f)$, and similarly $I_*(\lambda f) = \lambda I_*(f)$. Hence, if $f \in \mathscr{R}$, then $\lambda f \in \mathscr{R}$ and $I(\lambda f) = \lambda I(f)$.

Finally, $I^*(-f) = -I_*(f)$ and $I_*(-f) = -I^*(f)$, so if $f \in \mathscr{R}$, then $-f \in \mathscr{R}$ and $I(-f) = -I(f)$.

6.3 Daniell Integral

We fix an integral operator $I : S \to \mathbb{R}$.

6.3.1 $S^\uparrow$

Let $S^\uparrow = \left\{f : \Omega \to \mathbb{R} \bigcup \{+\infty\} \mid \exists (f_n)_{n\in\mathbb{N}} \subseteq S^{\mathbb{N}},\ f_n \text{ increasing, and } f = \lim_{n\to\infty} f_n \text{ pointwise}\right\}$.

6.3.1.1 Prop Let f, g be elements of $S^\uparrow$ s.t. $f \leq g$. Let $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ be increasing sequences in S s.t. $f = \lim\limits_{n\to+\infty} f_n,\ g = \lim\limits_{n\to+\infty} g_n$. Then $\lim\limits_{n\to+\infty} I(f_n) \leq \lim\limits_{n\to+\infty} I(g_n)$.

6.3.1.2 Proof For any $m \in \mathbb{N}$, $f_m \leq f \leq g$ Hence $I(f_m) \leq \lim\limits_{n\to+\infty} I(g_n)$. Taking $\lim\limits_{m\to+\infty}$ we get $\lim\limits_{m\to+\infty} I(f_m) \leq \lim\limits_{n\to+\infty} I(g_n)$.

6.3.1.3 Corollary Let $f \in S^{\uparrow}$. If $(f_n)_{n\in\mathbb{N}}$ and $(\widetilde{f}_n)_{n\in\mathbb{N}}$ are both increasing sequences in S s.t. $f = \lim\limits_{n\to+\infty} f_n = \lim\limits_{n\to+\infty} \widetilde{f}_n$ then $\lim\limits_{n\to+\infty} I(f_n) = \lim\limits_{n\to+\infty} I(\widetilde{f}_n)$ So we denote by $I(f)$ the limit $\lim\limits_{n\to+\infty} I(f_n)$, which is well-defined. Thus we obtain a mapping $I : S^{\uparrow} \to \mathbb{R}\bigcup\{+\infty\}$ s.t.

- If $(f_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ is increasing then $I(\lim\limits_{n\to+\infty} f_n) = \lim\limits_{n\to+\infty} I(f_n)$,
- If $(f,g) \in {S^{\uparrow}}^2$, $f \le g$ then $I(f) \le I(g)$,
- If $(f,g) \in {S^{\uparrow}}^2$ then $f+g \in S^{\uparrow}$ and $I(f+g) = I(f) + I(g)$,
- If $f \in S^{\uparrow}, \lambda \ge 0$, then $\lambda f \in S^{\uparrow}$ and $I(\lambda f) = \lambda I(f)$.

6.3.1.4 Prop Let $(f_n)_{n\in\mathbb{N}} \in (S^{\uparrow})^{\mathbb{N}}$ be an increasing sequence and $f = \lim\limits_{n\to+\infty} f_n$. Then $f \in S^{\uparrow}$ and $I(f) = \lim\limits_{n\to+\infty} I(f_n)$

6.3.1.5 Proof For $k \in \mathbb{N}$ let $(g_{k,m})_{m\in\mathbb{N}} \in S^{\mathbb{N}}$ be an increasing sequence s.t. $f_k = \lim\limits_{m\to+\infty} g_{k,m}$. For $n \in \mathbb{N}$, let $h_n = g_{0,n} \vee \cdots \vee g_{n,n} \in S$. The sequence $(h_n)_{n\in\mathbb{N}}$ is increasing. Moreover, $f_n \ge f_k \ge g_{k,n} \quad (k \le n)$. Hence $f_n \ge h_n$. Taking $\lim\limits_{n\to+\infty}$, we get $\forall k \in \mathbb{N}$, $f = \lim\limits_{n\to+\infty} f_n \ge \lim\limits_{n\to+\infty} h_n \ge \lim\limits_{n\to+\infty} g_{k,n} = f_k$. Taking $\lim\limits_{k\to+\infty}$, we get $f = \lim\limits_{n\to+\infty} h_n$. Hence $f \in S^{\uparrow}$ and $I(f) = \lim\limits_{n\to+\infty} I(h_n) \le \lim\limits_{n\to+\infty} I(f_n)$. Conversely, $\forall n \in \mathbb{N}, f \ge f_n$. Hence $I(f) \ge \lim\limits_{n\to+\infty} I(f_n)$.

6.3.2 S$^{\downarrow}$

Let $S^{\downarrow} = \{-f \mid f \in S^{\uparrow}\}$. We extend I to $I : S^{\downarrow} \to \mathbb{R}\bigcup\{-\infty\}$. By letting $I(-f) := -I(f)$ for $f \in S^{\uparrow}$.

6.3.2.1 Prop Let $(f,g) \in (S^{\uparrow}\bigcup S^{\downarrow})^2$ If $f \le g$, then $I(f) \le I(g)$.

6.3.2.2 Proof It suffices to treat the cases where $(f,g) \in S^{\uparrow} \times S^{\downarrow}$ and $(f,g) \in S^{\uparrow} \times S^{\downarrow}$.
If $(f,g) \in S^{\uparrow} \times S^{\downarrow}$, then $-f \in S^{\downarrow}$ and hence $g-f \in S^{\uparrow}, g-f \ge 0$ In both cases, $0 \le I(g-f) = I(g)+I(-f) = I(g)-I(f)$.

6.3.3 I-Integrable

Let $f : \Omega \to \mathbb{R}$ be a mapping. We define $\overline{I}(f) := \inf\limits_{\mu\in S^{\uparrow},\mu\ge f} I(\mu) \le \inf\limits_{\mu\in S,\mu\ge f} I(\mu) = I^*(f)$, $\underline{I}(f) := \sup\limits_{\mu\in S^{\downarrow},\mu\le f} I(\mu) \ge \sup\limits_{\mu\in S,\mu\le f} I(\mu) = I_*(f)$. If $\overline{I}(f) = \underline{I}(f)$ then we say that f is I-integrable (in the sense of Daniell).

6.3.3.1 Remark If f is I-integrable in the sense of Riemann, then it is I-integrable in the sense of Daniell.

6.3.4 Daniell Theorem

The set $\mathrm{L}^1(I)$ of all I-integrable mappings forms a vector subspace of $\mathbb{R}$. Moreover,

- $\forall (f,g) \in L^1(I)\ f \wedge g \in L^1(I)$;
- $I : L^1(I) \to \mathbb{R}$ is an integral operator extending $I : S \to \mathbb{R}$.

6.3.4.1 Proof Let $(f_1,f_2) \in L^1(I)^2$, $(l_1,l_2) \in {S^{\downarrow}}^2, l_1 \le f_1, l_2 \le f_1$. Let $(\mu_1,\mu_2) \in {S^{\uparrow}}^2, f_1 \le \mu_1, f_2 \le \mu_2$. We have $l_1+l_2 \le f_1+f_2 \le \mu_1+\mu_2$. Taking the supremum with respect to (l_1,l_2), we get $I(f_1)+I(f_2) = \underline{I}(f_1)+\underline{I}(f_2) \le \underline{I}(f_1+f_2)$. Taking the infimum with respect to (l_1,l_2), we get $\overline{I}(f_1+f_2) \le I(f_1)+I(f_2)$. Then $\overline{I}(f_1+f_2) = \underline{I}(f_1+f_2)$. So $f_1+f_2 \in L^1(I)$ and $I(f_1+f_2) = I(f_1)+I(f_2)$.

Similarly, if $f \in L^1(I), \lambda \ge 0$, then $\underline{I}(\lambda f) = \sup\limits_{l\le\lambda f, l\in S^{\downarrow}} I(l) = \sup\limits_{l\le f, l\in S^{\downarrow}} I(\lambda l) = \lambda\underline{I}(f) = \lambda I(f)$, $\overline{I}(\lambda f) = \lambda\overline{I}(f) = \lambda I(f)$. So $\lambda f \in L^1(I)$ and $I(\lambda f) = \lambda I(f)$. Moreover, if $f \in L^1(I), \mu \in S^{\uparrow}, l \in S^{\downarrow}, l \le f \le \mu$ then $-\mu \in S^{\downarrow}, -l \in S^{\uparrow}, -\mu \le -f \le -l$. Hence $\overline{I}(-f) = -\underline{I}(f) = -I(f) \quad \underline{I}(-f) = -\overline{I}(f) = -I(f)$. So $-f \in L^1(I)$ and $I(-f) = -I(f)$.

Want to prove that $\forall (f_1,f_2) \in L^1(I)^2,\ f_1 \wedge f_2 \in L^1(I)$.

Let $(f_1,f_2) \in L^1(I)^2$. For any $\varepsilon > 0$, $\exists (l_1,l_2) \in S^{\uparrow} \times S^{\uparrow}$, $(\mu_1,\mu_2) \in S^{\downarrow} \times S^{\downarrow}$ s.t. $l_1 \le f_1 \le \mu_1,\quad l_2 \le f_2 \le \mu_2$ and $I(\mu_1 - l_1) \le \frac{\varepsilon}{2},\quad I(\mu_2 - l_2) \le \frac{\varepsilon}{2}$. One has $l_1 \wedge l_2 \le f_1 \wedge f_2 \le \mu_1 \wedge \mu_2$, $\mu_1 \wedge \mu_2 - l_1 \wedge l_2 \le (\mu_1 - l_1) + (\mu_2 - l_2)$.

Hence $\overline{I}(f_1 \wedge f_2) - I(f_1 \wedge f_2) \le \varepsilon$.

6.3.5 Beppo Levi Theorem

Let $(f_n)_{n\in\mathbb{N}}$ be a monotone sequence of elements of $L_1(I)$, which converges pointwisely to some $f : \Omega \to \mathbb{R}$. If $(I(f_n))_{n\in\mathbb{N}}$ converges to a real number α Then $f \in L^1(I)$ and $I(f) = \alpha$.

6.3.5.1 Proof Assume that $(f_n)_{n\in\mathbb{N}}$ is increasing. Moreover, by replacing f_n by $f_n - f_0$ we may assume that $f_0 = 0$. Let $\epsilon > 0$ $\forall n \in \mathbb{N}$ let $\mu_n \in S^{\uparrow}$ s.t. $f_n - f_{n-1} \leq \mu_n$ and $I(f_n - f_{n-1}) \geq I(\mu_n) - \frac{\epsilon}{2^n}$.

the existence: $I(f_n - f_{n-1}) = \inf\limits_{\mu\in S^{\uparrow}, \mu\geq f_n - f_{n-1}} I(\mu)$. If $\forall \mu \in S^{\uparrow}, \mu \geq f_n - f_{n-1}$ one has $I(\mu) > I(f_n - f_{n-1}) + \frac{\epsilon}{2^n}$, then $I(f_n - f_{n-1}) + \frac{\epsilon}{2^n} \leq I(f_n - f_{n-1})$ contraction.

Thus $f_n = \sum\limits_{k=1}^{n}(f_k - f_{k-1}) \leq \mu_1 + \cdots + \mu_n$ and $I(f_n) \geq \sum\limits_{k=1}^{n}(I(\mu_k) - \frac{\epsilon}{2^k}) \geq I(\mu_1) + \cdots + I(\mu_n) - \epsilon$ Let $\mu = \mu_1 + \cdots + \mu_n + \cdots \in S^{\uparrow}$, $I(\mu) = \sum\limits_{n\in\mathbb{N}} I(\mu_n)$ One has $\mu \geq f$. $\lim\limits_{n\to+\infty} I(f_n) \geq I(\mu) - \epsilon \geq \overline{I}(f) - \epsilon$. Similarly, one can choose $l_n \in S^{\downarrow}, l_n \leq f_n, I(l_n) \geq I(f_n) - \epsilon$, $\liminf\limits_{n\to+\infty} I(l_n) \geq \alpha - \epsilon$. Note that $l_n \leq f_n \leq f$, so $\alpha - \epsilon \leq \liminf\limits_{n\to+\infty} I(l_n) \leq \underline{I}(f)$. Thus $\alpha - \epsilon \leq \underline{I}(f) \leq \overline{I}(f) \leq \alpha + \epsilon$. Let $\epsilon \to 0$ we get $\overline{I}(f) = \underline{I}(f) = \alpha$.

6.3.6 Fatou's Lemma

Let $(f_n)_{n\in\mathbb{N}} \in L^1(I)^{\mathbb{N}}$ Assume that there is $g \in L^1(I)$ s.t. $\forall n \in \mathbb{N} \quad f_n \geq g$. If $\liminf\limits_{n\to+\infty} f_n$ is a mapping from Ω to $\mathbb{R}$ and $\liminf\limits_{n\to+\infty} I(f_n) < +\infty$, then $\liminf\limits_{n\to+\infty} f_n \in L^1(I)$ and $I(\liminf\limits_{n\to+\infty} f_n) \leq \liminf\limits_{n\to+\infty} I(f_n)$.

6.3.6.1 Proof For any $n \in \mathbb{N}$, let $g_n = \lim\limits_{k\to+\infty}(f_n \wedge f_{n+1} \wedge \cdots \wedge f_{n+k})$. Then $\liminf\limits_{n\to+\infty} f_n = \lim\limits_{n\to+\infty} g_n$. For any k, one has $f_n \wedge \cdots \wedge f_{n+k} \geq g$. Hence $I(f_n) \geq \lim\limits_{k\to+\infty} I(f_n \wedge \cdots \wedge f_{n+k}) \geq I(g)$. By the theorem of Beppo Levi, $g_n \in L^1(I)$ and $I(g_n) = \lim\limits_{n\to+\infty} I(f_n \wedge \cdots \wedge f_{n+k}) \leq I(f_n)$. Note that $(g_n)_{n\in\mathbb{N}}$ is increasing and $\liminf\limits_{n\to+\infty} I(f_n) < +\infty$ Hence $\lim\limits_{n\to+\infty} I(g_n) = \liminf\limits_{n\to+\infty} I(g_n) \leq \liminf\limits_{n\to+\infty} I(f_n) < +\infty$. By the theorem of Beppo Levi, $\lim\limits_{n\to+\infty} g_n \in L^1(I)$ and $I(\liminf\limits_{n\to+\infty} f_n) = I(\lim\limits_{n\to+\infty} g_n) = \lim\limits_{n\to+\infty} I(g_n) \leq \liminf\limits_{n\to+\infty} I(f_n)$

6.3.7 Lebesgue Dominated Convergence Theorem

Let $(f_n)_{n\in\mathbb{N}}$ be a sequence in $L^1(I)$ that converges pointwisely to some $f : \Omega \to \mathbb{R}$ Assume that there exists $g \in L^1(I)$ s.t. $\forall n \in \mathbb{N}, |f_n| \leq g$ Then $f \in L^1(I)$ and $I(f) = \lim\limits_{n\to+\infty} I(f_n)$.

6.3.7.1 Proof Apply Fatou's lemma to $(f_n)_{n\in\mathbb{N}}$ and $(-f_n)_{n\in\mathbb{N}}$ to get $I(f) \geq \limsup I(f_n) \geq \liminf I(f_n) \geq I(f)$.

6.3.7.2 Notation Let $\varphi : \mathbb{R} \to \mathbb{R}$ be an increasing and right continuous mapping. Let S be the vector subspace of $\mathbb{R}^{\mathbb{R}}$ generated by $\mathbb{1}_{]a,b]}$ with $(a,b) \in \mathbb{R}^2, a < b$. For any $f \in L^1(I_\varphi)$. $I_\varphi(f)$ is denoted as $\int_{\mathbb{R}} f(x)d\varphi(x)$. For any subset A of $\mathbb{R}$, if $\mathbb{1}_A f \in L^1(I)$, then $\int_A f(x)d\varphi(x)$ denotes $\int_{\mathbb{R}} \mathbb{1}_A(x)f(x)d\varphi(x) = I(\mathbb{1}_A f)$. If $(a,b) \in \mathbb{R}^2, a < b$, $\int_a^b f(x)d\varphi(x)$ denotes $\int_{]a,b]} f(x)d\varphi(x)$; $\int_b^a f(x)d\varphi(x)$ denotes $-\int_{]a,b]} f(x)d\varphi(x)$. If $\varphi(x) = x$ for any $x \in \mathbb{R}$, we replace $d\varphi(x)$ by dx. $\forall \epsilon > 0, \exists \delta$, $m(E) < \delta$, $\int_E f d\mu < \epsilon$.

6.4 Semi-Algebra

6.4.1 Disjoint Union

Let A be sets, then notation $A = \bigsqcup\limits_{i\in I} A_i$ denotes:

1. $(A_i)_{i\in I}$ is a pairwisely disjoint family of sets.
2. $A = \bigcup\limits_{i\in I} A_i$.

6.4.2 Def

Let Ω be a set. We call a *semi-algebra* on Ω any $\mathcal{C} \subseteq \mathscr{P}(\Omega)$ satisfies:

1. $\emptyset \in \mathcal{C}$
2. $\forall\ (A,B) \in \mathcal{C}^2$, $A \cap B \in \mathcal{C}$
3. $\forall\ (A,B) \in \mathcal{C}^2$, $\exists\ (C_i)_{i=1}^n$ a finite family of elements in $\mathcal{C}$ s.t. $B \backslash A = \bigsqcup\limits_{i=1}^{n} C_i$

6.4.2.1 Algebra Let $\mathcal{C}$ be a semi-algebra on Ω. The set

$$\left\{A \in \mathscr{P}(\Omega) \middle| \exists\ n \in \mathbb{N},\ \exists\ (A_i)_{i=1}^n \in \mathcal{C}^n, \text{ s.t. } A = \bigsqcup_{i=1}^{n} A_i\right\}$$

called the *algebra* generated by $\mathcal{C}$.

6.4.2.2 Prop Let $\mathcal{C}$ be a semi-algebra on Ω, $\mathcal{A}$ be the algebra generated by $\mathcal{C}$. Then:

1. $\emptyset \in \mathcal{A}$;
2. $\forall\ (A,B) \in \mathcal{A}^2$, $A \cap B \in \mathcal{A}$, $B \backslash A \in \mathcal{A}$, $A \bigcup B \in \mathcal{A}$.

6.4.2.3 Proof By Def, $\emptyset \in \mathcal{A}$, $\mathcal{C} \subseteq \mathcal{A}$.
Moreover, if A and B be elements of $\mathcal{A}$ s.t. $A \cap B = \emptyset$, then $A \bigcup B \in \mathcal{A}$.
Let $A = \bigsqcup_{i=1}^{n} A_i$, $B = \bigsqcup_{j=1}^{m} B_j$, then $A \bigcup B = \bigsqcup_{(i,j)\in\{1,\cdots,n\}\times\{1,\cdots,m\}} A_i \cap B_j$. Hence $A \cap B \in \mathcal{A}$.
Since $B\backslash A = ((B\backslash A_1)\backslash A_2 \cdots)\backslash A_n$, by induction, it suffices to treat the case where $A \in \mathcal{C}$.
Then $B\backslash A = \bigsqcup_{B_j\backslash A} \in \mathcal{A}$. Finally, $A \bigcup B = (A\cap B) \sqcup (A\backslash B) \sqcup (B\backslash A)$.

6.4.2.4 Prop Let $\mathcal{C}$ be a semi-algebra on Ω, $\mathcal{A}$ be a algebra generated by $\mathcal{C}$.
Let S be the vector subspace of $\mathbb{R}^\Omega$ over $\mathbb{R}$ generated by mappings of the form:

$$\mathbb{1}_A, \quad A \in \mathcal{C} \quad (\text{ then } \forall\ B \in \mathcal{A},\ \mathbb{1}_B \in S)$$

$I : S \to \mathbb{R}$ be a $\mathbb{R}$-linear mapping. Suppose that: $\forall\ (f,g) \in S \times S$ s.t. $f \le g$, one has $I(f) \le I(g)$. Then: I is an integral-operator iff for any decreasing ($\subseteq$) sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{A}^{\mathbb{N}}$ s.t. $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, one has $\lim_{n\to+\infty} I(\mathbb{1}_{A_n}) = 0$.

6.4.2.5 Lemma $\forall\ (f,g) \in S^2$, $f \wedge g \in S$.

6.4.2.6 Proof $\forall\ A \in \mathcal{A}$, $\exists\ (A_i)_{i=1}^n \in \mathcal{C}^n$, $A = \bigsqcup_{i=1}^{n} A_i$, so $\mathbb{1}_A = \sum_{i=1}^{n} \mathbb{1}_{A_i} \in S$.
$\Rightarrow$: Suppose that I is an integral operator, $(\mathbb{1}_{A_n})_{n\in\mathbb{N}}$ is a decreasing sequence in S and $\lim_{n\to+\infty} \mathbb{1}_{A_n}(\omega) = 0$, $\forall\ \omega \in \Omega$. Hence, $\lim_{n\to+\infty} I(\mathbb{1}_{A_n}) = 0$.

$\Leftarrow$: Let $(F_n)_{n\in\mathbb{N}}$ be a decreasing sequence in S that converges pointwisely to 0.
Let $B_0 = \{\omega \in \Omega | f_0(\omega) > 0\} \in \mathcal{A}$, $M = \max f_0(\Omega)$.

Lemma: $\forall f \in S$, $\exists (A_i)_{i=1}^n \in \mathcal{C}^n$ pairwise disjoint, and $(\lambda_i)_{i=1}^n \in \mathbb{R}^n$, $f = \sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}$.

Proof: f is of the form $\sum_{j=1}^{m} a_j \mathbb{1}_{B_j}$, where $B_j \in \mathcal{C}$. For any subset $I \subseteq \{1,\ldots,m\}$, $I \ne \varnothing$, $I \ne \{1,\ldots,m\}$, let $B_I = \left(\bigcap_{i\in I} B_i\right) \cap \left(\bigcap_{j\in\{1,\ldots,m\}\backslash I} (\Omega \setminus B_j)\right)$, if $I = \{1,\ldots,m\}$, let $B_I = \bigcap_{i\in I} B_i$. Then $(B_I)_{I\subseteq\{1,\ldots,m\}}$ are pairwise disjoint. Moreover, since $I \ne \varnothing$, $B_I \in \mathcal{C}$. We have $B_j = \bigsqcup_{I\subseteq\{1,\ldots,m\}, j\in I} B_I$. Hence $f = \sum_{\varnothing\ne I\subseteq\{1,\ldots,m\}} \left(\sum_{j\in I} a_j\right) \mathbb{1}_{B_I}$.

Corollary:

1. If $f \in S$, then $f \wedge 0 \in S$.
2. If $(f,g) \in S^2$, then $f \wedge g = (f-g) \wedge 0 + g \in S$.

So we proved the **6.4.2.5 Lemma** mentioned before.

For any $\varepsilon > 0$, $A_n^\varepsilon = \{\omega \in \Omega | f_n(\omega) \ge \varepsilon\} \in \mathcal{A}$. Moreover, since $\lim_{n\to+\infty} f_n = 0$, $\bigcap_{n\in\mathbb{N}} A_n^\varepsilon = \emptyset$. Note that $0 \le f_n \le \varepsilon \mathbb{1}_B + M\mathbb{1}_{A_n^\varepsilon}$, so $0 \le I(f_n) \le \varepsilon I(\mathbb{1}_B) + MI(\mathbb{1}_{A_n^\varepsilon})$, which leads to $\limsup_{n\to+\infty} I(f_n) \le \varepsilon I(\mathbb{1}_B)$, $\forall\ \varepsilon > 0$. So $\lim_{n\to+\infty} I(f_n) = 0$.

6.4.2.7 Example $\Omega = \mathbb{R}$, $\mathcal{C} = \{]a,b] | (a,b) \in \mathbb{R}^2,\ a \le b\}$, $\mathcal{A} =$ algebra generated by $\mathcal{C}$, $\varphi : \mathbb{R} \to \mathbb{R}$ be an increasing right continuous mapping.
$I_\varphi : S \to \mathbb{R}$, $I_\varphi(\mathbb{1}_{]a,b]}) = \varphi(b) - \varphi(a)$

In the following, we will prove that I_φ is a well-defined integral operator.

6.4.2.8 Lemma $\forall\ \varepsilon > 0$, $\forall\ A \in \mathcal{A}$, $A \ne \emptyset$, $\exists B \in \mathcal{A}$ s.t. $\emptyset \ne \bar{B} \subseteq A$ and $I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) \le \varepsilon$.

6.4.2.9 Proof We first consider the case where $A \in \mathcal{C}$, $A =]a,b]$, $a < b$. By the right continuity of φ, $\exists\ a' \in]a,b[$ s.t. $\varphi(a') - \varphi(a) \le \varepsilon$. Let $B =]a',b[$, $\bar{B} = [a',b] \subseteq]a,b]$, $I_\varphi(\mathbb{1}_B) = \varphi(b) - \varphi(a)$,
$I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) = \varphi(a') - \varphi(a) \le \varepsilon$. In general, $A = \bigsqcup_{i=1}^{n} A_i$ with $A_i \in \mathcal{C}$, $\forall\ i \in \{i,\cdots,n\}$, $\exists\ B_i \in \mathcal{C}$, $\emptyset \ne \bar{B}_i \subseteq A_i$, $I(\mathbb{1}_{A_i}) - I(\mathbb{1}_{B_i}) \le \frac{\varepsilon}{n}$, $B = \bigsqcup_{j=1}^{m} B_j$, then $I(\mathbb{1}_A) - I(\mathbb{1}_B) \le \varepsilon$.

6.4.2.10 Theorem I_φ is an integral operator.

6.4.2.11 Proof Let $(A_n)_{n\in\mathbb{N}}$ be a decreasing sequence in $\mathcal{A}$ s.t. $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$. Let $\varepsilon > 0$, for any $n \in \mathbb{N}$, let $B_n \in \mathcal{A}$ s.t. $\emptyset \neq \bar{B}_n \subseteq A_n$ and $I_\varphi(\mathbb{1}_{A_n}) - I_\varphi(\mathbb{1}_{B_n}) \leq \frac{\varepsilon}{2^n}$. Note that $\bar{B}$ is compact. For any $n \in \mathbb{N}$, let $C_n = B_0 \cap \cdots \cap B_n \subseteq \bar{B}_0 \cap \cdots \cap \bar{B}_n$. Since $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, $\bigcap_{n\in\mathbb{N}} \bar{B}_n = \emptyset$. Hence $\exists\ N \in \mathbb{N}$, $\bigcap_{n=0}^{N} \bar{B}_n = \emptyset$. Moreover,

$$B_n \backslash C_n = B_n \backslash (B_n \cap C_{n-1}) = B_n \backslash C_{n-1} \subseteq A_n \backslash C_{n-1} \subseteq A_{n-1} \backslash C_{n-1}$$

Hence $I_\varphi(\mathbb{1}_{A_n \backslash C_n}) = I_\varphi(\mathbb{1}_{B_n \backslash C_n}) + I_\varphi(\mathbb{1}_{A_n \backslash B_n}) \leq I_\varphi(\mathbb{1}_{A_{n-1} \backslash C_{n-1}}) + \frac{\varepsilon}{2^n}$. So $I_\varphi(\mathbb{1}_{A_n}) \leq \varepsilon$, $\forall\ n \geq N$. Thus $\lim_{n\to+\infty} I_\varphi(\mathbb{1}_{A_n})$.

6.4.2.12 Prop Let $\mu : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a mapping s.t. for any $(A, B) \in \mathcal{C}^2$, $A \subseteq B$, and any $(C_i)_{i=1}^n \in \mathcal{C}^n$ s.t. $B \setminus A = \coprod_{i=1}^{n} C_i$, one has $\mu(B) = \mu(A) + \sum_{i=1}^{n} \mu(C_i)$.

Then there is a unique $\mathbb{R}$-linear mapping $I_\mu : S \to \mathbb{R}$ s.t. $I_\mu(\mathbb{1}_A) = \mu(A)$ for all $A \in \mathcal{C}$.

6.4.2.13 Proof We intend to define $I_\mu\left(\sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}\right)$ as $\sum_{i=1}^{n} \lambda_i \mu(A_i)$ for $A_i \in \mathcal{C}$.

We need to check that if $f \in S$ is written as $f = \sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i} = \sum_{j=1}^{m} \xi_j \mathbb{1}_{B_j}$, then $\sum_{i=1}^{n} \lambda_i \mu(A_i) = \sum_{j=1}^{m} \xi_j \mu(B_j)$.

We have $0 = \sum_{i=1}^{n} \lambda_i \mu(A_i) - \sum_{j=1}^{m} \xi_j \mu(B_j)$.

It suffices to prove that if $\sum_{i=1}^{n} a_i \mathbb{1}_{A_i} = 0$ for $a_i \in \mathbb{R}$ and $A_i \in \mathcal{C}$, then $\sum_{i=1}^{n} a_i \mu(A_i) = 0$.

For $I \subseteq \{1, \ldots, n\}$, let $A_I = \{\omega \in \Omega \mid \forall i \in I, \omega \in A_i, \forall i \in \{1, \ldots, n\} \setminus I, \omega \in \Omega \setminus A_i\} \in \mathscr{A}$, when $I \neq \varnothing$.

Lemma: Let $B \in \mathscr{A}$. If $B = \coprod_{i=1}^{n} B_i = \coprod_{j=1}^{m} C_j$ with $B_i \in \mathcal{C}$ and $C_j \in \mathcal{C}$, then $\sum_{i=1}^{n} \mu(B_i) = \sum_{j=1}^{m} \mu(C_j)$. In particular, we can extend $\mu : \mathcal{C} \to \mathbb{R}_{\geq 0}$ to $\mu : \mathscr{A} \to \mathbb{R}_{\geq 0}$ s.t. for any $D_1, \ldots, D_n$ in $\mathscr{A}$ disjoint, $\mu(D_1 \coprod \cdots \coprod D_n) = \sum_{i=1}^{n} \mu(D_i)$.

Proof: We have $B_i = \coprod_{j=1}^{m} (B_i \cap C_j)$, so $\mu(B_i) = \sum_{j=1}^{m} \mu(B_i \cap C_j)$.

Therefore, $\sum_{i=1}^{n} \mu(B_i) = \sum_{i=1}^{n} \sum_{j=1}^{m} \mu(B_i \cap C_j) = \sum_{j=1}^{m} \sum_{i=1}^{n} \mu(C_j \cap B_i) = \sum_{j=1}^{m} \mu(C_j)$.

Returning to the proof of the proposition, we have $0 = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} = \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} \left(\sum_{i\in I} a_i\right) \mathbb{1}_{A_I}$.

Hence, when $A_I \neq \varnothing$, we have $\sum_{i\in I} a_i = 0$.

Thus, $\sum_{i=1}^{n} a_i \mu(A_i) = \sum_{i=1}^{n} a_i \sum_{I \subseteq \{1,\ldots,n\}} \mu(A_I) = \sum_{\varnothing \neq I \subseteq \{1,\ldots,n\}} \mu(A_I) \sum_{i\in I} a_i = 0$.

6.5 Integrable Functions

In this section, let Ω be a set, $s \subseteq \mathbb{R}^\Omega$ be a vector space over $\mathbb{R}$. $\forall\ (f, g) \in S^2$, $f \wedge g \in S$, $I : S \to \mathbb{R}$ be a integral operator.

6.5.1 Prop

Suppose that $\mathbb{1}_\Omega \in L^1(I)^\uparrow$, the set $\mathcal{G} = \{A \subseteq \Omega | \mathbb{1}_A \in L^1(I)^\uparrow\}$ is a σ-algebra on Ω. Moreover, if we denote by $\mu : \mathcal{G} \to \mathbb{R}_{\geq 0}$, then mapping defined as $\mu(A) := I(\mathbb{1}_A)$, then μ satisfies: $\forall\ (A_n)_{n\in\mathbb{N}} \in \mathcal{G}^\mathbb{N}$ that is pairwisely disjoint, then $\mu(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n\in\mathbb{N}} \mu(A_n)$.

6.5.2 Proof

1. $\emptyset \in \mathcal{G}$, since $0 = \mathbb{1}_\emptyset \in L^1(I)^\uparrow$. $\Omega \in \mathcal{G}$, since $\mathbb{1}_\Omega \in L^1(I)^\uparrow$.
2. If A and B are elements of $\mathcal{G}$, $A \subseteq B$, then $\mathbb{1}_A \leq \mathbb{1}_B$, so $\mathbb{1}_B - \mathbb{1}_A = \mathbb{1}_{B\setminus A} \in L^1(I)^\uparrow$
3. If $(A, B) \in \mathcal{G}^2$, $\mathbb{1}_{A \bigcup B} = \mathbb{1}_A \wedge \mathbb{1}_B \in L^1(I)^\uparrow$, so $A \bigcup B \in \mathcal{G}$. If $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}^n$, $A = \bigcup_{n\in\mathbb{N}} A_n$, then $\mathbb{1}_A = \lim_{n\to+\infty} \mathbb{1}_{A_0 \bigcup \cdots \bigcup A_n} \in L^1(I)^\uparrow \implies A \in \mathcal{G}$. $\mathbb{1}_{\bigcup_{n\in\mathbb{N}} A_n} = \lim_{n\to+\infty} \mathbb{1}_{A_0 \bigcup \cdots \bigcup A_n} = \lim_{n\to+\infty} \sum_{i=0}^{n} \mathbb{1}_{A_i}$

$$\mu(\bigcup_{n\in\mathbb{N}} A_n) = I(\mathbb{1}_{\bigcup_{n\in\mathbb{N}} A_n}) = \lim_{n\to+\infty} I(\sum_{i=1}^{n} \mathbb{1}_{A_i}) = \lim_{n\to+\infty} (\mu(A_0) + \cdots + \mu(A_n)) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

(Using **Beppo Levi Theorem**).

6.6 Limit and Differential of Integrals with Parameters

Let Ω be a set, $S \subseteq \mathbb{R}$ be an $\mathbb{R}$-vector subspace s.t. $\forall\ (f,g) \in S^2,\ f \wedge g \in S$ Let $I : S \to \mathbb{R}$ be an integral operator.

6.6.0.1 Theorem Let X be a topological space, $p \in X$, $f : \Omega \times X \to \mathbb{R}$ be a mapping, $g \in L^1(I)$. Suppose that:

1. $\forall\ \omega \in \Omega,\ f(\omega,\cdot) : X \to \mathbb{R},\quad x \mapsto f(\omega,x)$ is continuous at p.
2. $\forall\ n \in X,\ f(\cdot,x) : X \to \mathbb{R},\quad x \mapsto f(\omega,x)$ belongs to $L^1(I)$ and $\forall\ \omega \in \Omega,\ |f(\omega,x)| \leq g(\omega)$.
3. p has a countable neighborhood basis in X.

Then $(x \in X) \mapsto I(f(\cdot,x))$ is continuous at p

6.6.0.2 Proof Take $\omega \in \Omega$, Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X that converges to p. Then for any n, $f_n : \Omega \to \mathbb{R}$, $f_n(\omega) := f(\omega,x)$, one has $|f_n| \leq g$. Moreover, $\forall\ \omega \in \Omega$, $\lim_{n\to+\infty} f_n(\omega) = \lim_{n\to+\infty} f(\omega,x) = f(\omega,p)$. Then, by **Dominated Convergence Theorem**, $\lim_{n\to+\infty} I(f_n) = I(f(\cdot,p))$.

6.6.0.3 Theorem Let J be an open interval in $\mathbb{R}$, $f : \Omega \times J \to \mathbb{R}$ be a mapping, $g \in L^1(I)$.
Assume that:

1. $\forall\ \omega \in \Omega,\ f(\omega,\cdot) : J \to \mathbb{R},\quad f \mapsto f(\omega,t)$ is differentiable. (We denote by $\frac{\partial f}{\partial t}(\omega,t)$ the derivation at t) $\forall\ t \in J$, $\left|\frac{\partial f}{\partial t}(\omega,t)\right| \leq g(\omega)$
2. $\forall\ t \in J,,\ f(\cdot,t) : \Omega \to \mathbb{R},\quad \omega \mapsto f(\omega,t)$ belongs to $L^1(I)$

Then, $\varphi : J \to \mathbb{R},\quad t \mapsto I(f(\cdot,t))$ is differentiable, and $\varphi'(t) = I(\frac{\partial f}{\partial t}(\cdot,t))$

6.6.0.4 Proof Let $a \in J$ and $(t_n)_{n\in\mathbb{N}}$ be a sequence in $J\backslash\{a\}$ s.t. $\lim_{n\to+\infty} t_n = a$. Then

$$\frac{\varphi(t_n)-\varphi(a)}{t_n-a} = I\left(\frac{f(\cdot,t_n)-f(\cdot,a)}{t_n-a}\right)$$

$\forall\ \omega \in \Omega$, $\left|\frac{f(\omega,t_n)-f(\omega,a)}{t_n-a}\right|$, and its limit is $\frac{\partial f}{\partial t}(\omega,t)$. Then $\lim_{n\to+\infty} \frac{\varphi(t_n)-\varphi(a)}{t_n-a} = I\left(\frac{\partial f}{\partial t}(\cdot,t)\right)$

6.7 Measure Theory

6.7.1 Measurable Space

We call *measurable space* any pair $(E,\mathcal{E})$ where E is a set any $\mathcal{E}$ is a σ-algebra on E.
Let (X,Σ_X) and (Y,Σ_Y) be two measurable spaces. $\Sigma_X \otimes_R \Sigma_Y := \sigma(\{S_1 \times S_2 : S_1 \in \Sigma_X, S_2 \in \Sigma_Y\})$.

6.7.1.1 Notation Take $A \subseteq X \times Y$.
For $x \in X$, $A_x := \{y \in Y : (x,y) \in A\} = \pi_2((\{x\} \times Y) \cap A)$, called *vertical section* of A.
For $y \in Y$, $A_y := \pi_1((X \times \{y\}) \cap A)$, called *horizontal section* of A.

6.7.1.2 Def Let X be a set, then $\mathcal{D} \subseteq \mathcal{P}(X)$ is a *Dynkin system* if:

1. $X \in \mathcal{D}$.
2. $\forall D \in \mathcal{D},\ X\backslash D \in \mathcal{D}$.
3. If $\{D_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathcal{D}$ of pairwise disjoint sets, then $\bigcup_n D)n \in \mathcal{D}$.

A σ-algebra is a Dynkin system.

6.7.1.3 Def Let $\mathcal{G} \subseteq \mathcal{P}(X)$, then $\delta(\mathcal{G}) \subseteq \mathcal{P}(X)$ is called the Dynkin system generated by $\mathcal{G}$:

1. $\mathcal{G} \subseteq \delta(\mathcal{G})$;
2. If $\mathcal{D}$ is a Dynkin system containing $\mathcal{G}$, then $\delta(\mathcal{G}) \subseteq \mathcal{D}$.

This is the smallest Dynkin system containing $\mathcal{G}$.

6.7.1.4 Exercise $\delta(\mathcal{G})$ exists and it's unique.

6.7.1.5 Prop If $\mathcal{D}$ is a Dynkin system closed under the finite intersection, then it is a σ-algebra.

6.7.1.6 Proof We have to show that $\mathcal{D}$ is closed under any countable union.

Let $\{\mathcal{D}_n\}$ be any sequence in $\mathcal{D}$, define: $E_n := \mathcal{D}_n \cap (X \cup E_n)$, $E_0 = \mathcal{D}_0$

Since $\mathcal{D}$ is closed under intersection, $E_n \in \mathcal{D}$. Now the E_n are all disjoint, and by induction we can prove that $\bigcup_{k=0}^n E_k = \bigcup_{k=0}^n \mathcal{D}_k$. Thus we have shown that $\bigcup_n \mathcal{D}_n = \bigcup_n E_n \in \mathcal{D}$.

6.7.1.7 Prop Let X be a set, and let $G \subseteq P(X)$. Assume that G is closed under finite intersection, then $\delta(G) = \sigma(G)$.

6.7.1.8 Proof We know $\delta(\mathcal{G}) \subseteq \sigma(\mathcal{G})$. We need to show that $\delta(\mathcal{G})$ is a σ-algebra, so we must demonstrate it is closed under intersection.

Let $D \in \delta(\mathcal{G})$, and define $\delta_D := \{E \subseteq X \mid E \cap D \in \delta(\mathcal{G})\}$. One can claim that δ_D is a Dynkin system.

1. $X \cup D = D$, so $X \in \delta_D$.
2. (Paolo) For all $E \in \delta_D$, $(X \setminus E) \cap D = ((X \setminus E) \cap D) \cup ((X \setminus D) \cap D) = ((X \setminus E) \cup (X \setminus D)) \cap D = (X \setminus (E \cap D)) \cap D$. Further, $(X \setminus (E \cap D)) \cap D = X \setminus ((E \cap D) \cup (X \setminus D))$, and $(E \cap D) \cap (X \setminus D) = \emptyset$.

   Since $E \cap D \in \delta(\mathcal{G})$ (because $E \in \delta_D$) and $X \setminus D \in \delta(\mathcal{G})$ (because $D \in \delta(\mathcal{G})$), it follows that $(E \cap D) \cup (X \setminus D) \in \delta(\mathcal{G})$. Therefore, $(X \setminus E) \cap D \in \delta(\mathcal{G})$, and thus $X \setminus E \in \delta_D$.
3. (Own) For $A \in \delta_D$, the goal is to show $X \setminus A \in \delta_D$: $A \cap D \in \delta(\mathcal{G})$ and $D \in \delta(\mathcal{G})$. Then, $(A \cap D) \sqcup (X \setminus D) \in \delta(\mathcal{G})$. The complement is: $X \setminus ((X \setminus (A \cap D)) \cap D) \to (X \setminus (A \cap D)) \cap D = (X \setminus A) \bigcap D$. These things are all in $\delta(\mathcal{G})$.
4. Let $\{E_n\}_{n \in \mathbb{N}}$ be pairwise disjoint elements in δ_D. Then,

$$\left(\bigsqcup_n E_n\right) \cap D = \bigsqcup_n (E_n \cap D).$$

   Since $E_n \cap D \in \delta(\mathcal{G})$ and the disjoint union belongs to $\delta(\mathcal{G})$, it follows that $\bigsqcup_n E_n \in \delta_D$.

   Note that for all $G \in \mathcal{G}$, $\mathcal{G} \subseteq \delta_G$, hence $\delta(\mathcal{G}) \subseteq \delta_G$. For all $(A, B) \in \delta(\mathcal{G})^2$, since $\delta(\mathcal{G}) \subseteq \delta_B$, $A \in \delta_B$, and thus $A \cap B \in \delta(\mathcal{G})$.

6.7.1.9 Theorem Let (X, Σ_X, μ) and (Y, Σ_Y, ν) be two σ-finite measure spaces. Then for any $E \in \Sigma_X \otimes \Sigma_Y$, the functions $f_E : X \to \mathbb{R} \cup \{\infty\}, \quad x \mapsto \nu(E_x)$ and $g_E : Y \to \mathbb{R} \cup \{\infty\}, \quad y \mapsto \mu(E^y)$ are Σ_X-measurable and Σ_Y-measurable, respectively.

6.7.1.10 Proof We prove the result only for f_E.

Assume ν is a finite measure ($\nu(Y) < +\infty$). Define $\mathcal{F} = \{E \in \Sigma_X \otimes \Sigma_Y \mid f_E \text{ is measurable}\}$. We aim to show $\mathcal{F} = \Sigma_X \otimes \Sigma_Y$.

- Let $S_1 \in \Sigma_X$ and $S_2 \in \Sigma_Y$. Then,

$$(S_1 \times S_2)_x = \begin{cases} S_2 & \text{if } x \in S_1, \\ \emptyset & \text{otherwise.} \end{cases}$$

  Thus,

$$f_{S_1 \times S_2}(x) = \nu((S_1 \times S_2)_x) = \nu(S_2)\mathbb{I}_{S_1}(x).$$

  Since $S_1 \in \Sigma_X$, $\mathbb{I}_{S_1}$ is measurable, and $\nu(S_2)$ is a constant, $f_{S_1 \times S_2}$ is measurable. Hence, $S_1 \times S_2 \in \mathcal{F}$.

- Now, show $\mathcal{F}$ is a Dynkin system:
  1. $X \times Y \in \mathcal{F}$ (take $S_1 = X$, $S_2 = Y$).
  2. Let $D \in \mathcal{F}$. Then,

$$((X \times Y) \setminus D)_x = Y \setminus D_x.$$

     Since ν is finite,

$$f_{(X \times Y) \setminus D}(x) = \nu(Y \setminus D_x) = \nu(Y) - \nu(D_x) = \nu(Y) - f_D(x).$$

     Thus, $f_{(X \times Y) \setminus D}$ is measurable, and $(X \times Y) \setminus D \in \mathcal{F}$.
  3. Let $\{D_n\}$ be a sequence of disjoint sets in $\mathcal{F}$. Then,

$$f_{\bigsqcup_n D_n}(x) = \nu\left(\left(\bigsqcup_n D_n\right)_x\right) = \nu\left(\bigsqcup_n (D_n)_x\right) = \sum_n \nu((D_n)_x) = \sum_n f_{D_n}(x),$$

     so $f_{\bigsqcup_n D_n}$ is measurable.

- Let $\mathcal{G} = \{S_1 \times S_2 \mid S_1 \in \Sigma_X, S_2 \in \Sigma_Y\}$. Then $\mathcal{G} \subseteq \mathcal{F}$, and $\mathcal{G}$ is closed under intersection:

$$(S_1 \times T_1) \cap (S_2 \times T_2) = (S_1 \cap S_2) \times (T_1 \cap T_2).$$

  Hence, $\delta(\mathcal{G}) = \sigma(\mathcal{G}) = \Sigma_X \otimes \Sigma_Y \subseteq \mathcal{F}$.

- For the σ-finite case, let $Y = \bigcup_n Y_n$ with $\nu(Y_n) < +\infty$. Define disjoint sets $F_n = Y_n \cap \left(X \setminus \bigcup_{k<n} Y_k\right)$, so $Y = \bigsqcup_n F_n$ and $\nu(F_n) < +\infty$. Define finite measures $\nu^{(n)}(E) := \nu(E \cap F_n)$. Then,

$$f_E(x) = \nu(E_x) = \sum_n \nu^{(n)}(E_x) = \sum_n f_E^{(n)}(x),$$

where each $f_E^{(n)}$ is measurable. Thus, f_E is measurable.

6.7.1.11 Prop Let Ω be a set and $(\mathcal{G}_i)_{i\in I}$ be a family of σ-algebra on Ω. Then $\mathcal{G} = \bigcap_{i\in I} \mathcal{G}_i$ is a σ-algebra.

6.7.1.12 Proof

1. $\emptyset \in \mathcal{G}, \Omega \in \mathcal{G}$
2. If $A \in \mathcal{G}$, then $\Omega \backslash A \in \mathcal{G}$
3. Let $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$. For any $i \in I$, $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}_i^{\mathbb{N}}$, so $\bigcup_{n\in\mathbb{N}} \in \mathcal{G}_i$, hence $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{G}$

6.7.1.13 Def Let $\mathcal{C} \subseteq \mathscr{P}(\Omega)$, we denote by $\sigma(\mathcal{C})$ the intersection of all σ-algebras on Ω containing $\mathcal{C}$. It is the smallest σ-algebra containing $\mathcal{C}$

6.7.1.14 Example

1. Let (X, τ) be a topological space, $\sigma(\tau)$ is called the *Borel σ-algebra* of X.
2. On $[-\infty, +\infty]$, the following σ-algebra are the same:
   - $\mathcal{G}_1 = \sigma(\{[a, +\infty]|a \in \mathbb{R}\})$
   - $\mathcal{G}_2 = \sigma(\{]a, +\infty]|a \in \mathbb{R}\})$
   - $\mathcal{G}_3 = \sigma(\{[-\infty, a]|a \in \mathbb{R}\})$
   - $\mathcal{G}_4 = \sigma(\{[-\infty, a[|a \in \mathbb{R}\})$

   Moreover, $\mathcal{B} = \{A \subseteq |A \in \mathcal{G}_1\}$ is equal to the Borel σ-algebra of $\mathbb{R}$

6.7.1.15 Proof $[a, +\infty[= \bigcap_{n\in\mathbb{N}_+}]a - \frac{1}{n}, +\infty] \in \mathcal{G}_2$

6.7.1.16 Exercise Borel σ-algebra of $\mathbb{R} = \sigma(\{]-\infty, a]|a \in \mathbb{R}\})$.

6.7.1.17 Def Let $f : X \to Y$ be a mapping of sets. For any $\mathcal{C}_Y \subseteq \mathscr{P}(Y)$, we denoted by $f^{-1}(\mathcal{C}_Y) := \{f^{-1}(B)|B \in \mathcal{C}_Y\}$. For any $\mathcal{C}_X \subseteq \mathscr{P}(X)$, we denoted by $f_*(\mathcal{C}_X) := \{B \subseteq Y|f^{-1}(B) \in \mathcal{C}_X\}$.

6.7.1.18 Prop Let $f : X \to Y$ be a mapping. If $\mathcal{G}_Y$ is a σ-algebra on Y, $f^{-1}(\mathcal{G}_Y)$ is a σ-algebra on X. If $\mathcal{G}_X$ is a σ-algebra on X, $f_*(\mathcal{G}_X)$ is a σ-algebra on Y.

6.7.1.19 Proof

1. $\emptyset = f^{-1}(\emptyset) \in f^{-1}(\mathcal{G}_Y)$, $\forall\, B \in \mathcal{G}_Y$, $X\backslash f^{-1}(B) = f^{-1}(Y\backslash B)$. If $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}_Y^{\mathbb{N}}$, $A = \bigcup_{n\in\mathbb{N}} A_n$, then $\bigcup_{n\in\mathbb{N}} f^{-1}(A_n) = f^{-1}(A) \in f^{-1}(\mathcal{G}_Y)$
2. $f^{-1}(\emptyset) = \emptyset \in \mathcal{G}_X$, so $\emptyset \in f^*(\mathcal{G}_X)$. $\forall\, B \in f_*(\mathcal{G}_X)$, $f^{-1}(Y\backslash B) = X\backslash f^{-1}(B) \in \mathcal{G}_Y$, so $Y\backslash B \in f_*(\mathcal{G}_X)$. $\forall\, (B_n)_{n\in\mathbb{N}} \in f_*(\mathcal{G}_X)^{\mathbb{N}}$, $B = \bigcup_{n\in\mathbb{N}} B_n$, $f^{-1}(B) = \bigcup_{n\in\mathbb{N}} f^{-1}(B_n)$ so $B \in f_*(\mathcal{G}_X)$

6.7.1.20 measurable Let $(X, \mathcal{G}_X)$ and $(Y, \mathcal{G}_Y)$ be measurable spaces, $f : X \to Y$ be a mapping. If $f^{-1}(\mathcal{G}_Y) \subseteq \mathcal{G}_X$ or equivalently $\mathcal{G}_Y \subseteq f_*(\mathcal{G}_X)$, then we say that f is *measurable*.

6.7.1.21 Prop Let $(X, \mathcal{G}_X)$, $(Y, \mathcal{G}_Y)$ and $(Z, \mathcal{G}_Z)$ be measurable spaces, $f : X \to Y$ and $g : Y \to Z$ be measurable mappings. Then $g \circ f$ is measurable.

6.7.1.22 Proof $\forall\, B \in \mathcal{G}_Z$, $(g \circ f)^{-1}(B) = f^{-1}(g^{-1}(B))$, $g^{-1}(B) \in \mathcal{G}_Y$, so $f^{-1}(g^{-1}(B)) \in \mathcal{G}_Y$.

6.7.1.23 Def Let Ω be a set, $((E_i, \mathcal{E}_i))_{i\in I}$ be a family of measurable spaces. $f : (f_i)_{i\in I}$ where $f_i : \Omega \to E_i$ is a mapping. We denote by $\sigma(f)$ the σ-algebra on Ω $\sigma\left(\bigcup_{i\in I} f_i^{-1}(\mathcal{E}_i)\right)$. It is the smallest σ-algebra on Ω making all f_i measurable.

6.7.1.24 Prop We keep the notation of the above Def. For any $i \in I$, let $\mathcal{C}_i \subseteq \mathscr{P}(E_i)$ s.t. $\sigma(\mathcal{C}_i) = \mathcal{E}_i$. Then $\sigma(f) = \sigma(\bigcup_{i\in I} f_i^{-1}(\mathcal{C}_i))$.

6.7.1.25 Proof Let $g = \sigma(\bigcup_{i\in I} f_i^{-1}(\mathcal{C}_i))$, by definition, $\mathcal{G} \subseteq \sigma(f)$. For any $i \in I$, $f_{i,*}(\sigma(f_i^{-1}(\mathcal{C}_i)))$ is a σ-algebra on E_i containing $\mathcal{C}_i$. So $\mathcal{E}_i \subseteq f_{i,*}(\sigma(f_i^{-1}(\mathcal{C}_i)))$. which leads to $f_i^{-1}(\mathcal{E}_i) \subseteq \sigma(f_i^{-1}(\mathcal{C}_i)) \subseteq \mathcal{G}$. Hence $\bigcup_{i\in I} f_i^{-1}(\mathcal{E}_i) \subseteq \mathcal{G} \Rightarrow \sigma(f) \subseteq \mathcal{G}$.

6.7.1.26 Corollary Let $(X, \mathcal{G}_X)$ and $(Y, \mathcal{G}_Y)$ be measurable spaces, $f : X \to Y$ be a mapping. $\mathcal{C}_Y \subseteq \mathcal{G}_Y$ s.t. $\sigma(\mathcal{C}_Y) = \mathcal{G}_Y$. Then f is measurable iff $\forall\ B \in \mathcal{C}_Y$, $f^{-1}(B) \in \mathcal{G}_X$.

6.7.1.27 Proof $\sigma(f) = \sigma(f^{-1}(\mathcal{C}_Y))$ is measurable iff $\sigma(f) \subseteq \mathcal{G}_X$

6.7.1.28 Example Let $((E_i, \mathcal{E}_i))_{i\in I}$ be a family of measurable spaces, $E = \prod_{i\in I} E_i$. $\forall\ i \in I$, $\pi_i : E \to E_i$, $\quad (x_j)_{j\in I} \mapsto x_i$. We denote by $\bigotimes_{i\in I} \mathcal{E}_i$ the σ-algebra $\sigma((\pi_i)_{i\in I})$.

6.7.1.29 Prop Let X be a set, $((E_i, \mathcal{E}_i))_{i\in I}$ be measurable spaces, $(\Omega, \mathcal{G})$ be a measurable space. $f = (f_i : X \to E_i)_{i\in I}$ be mappings. Then $\varphi : (\Omega, \mathcal{G}) \to (X, \sigma(f))$ is measurable iff $\forall\ i \in I$, $f_i \circ \varphi : (\Omega, \mathcal{G}) \to (E, \mathcal{E}_i)$ is measurable.

6.7.1.30 Proof $\Rightarrow$: If φ is measurable since each f_i is measurable, one has $f_i \circ \varphi$ is measurable.
$\Leftarrow$: If $f_i \circ \varphi$ is measurable, $\forall\ B \in \mathcal{E}_i$, $(f_i \circ \varphi)^{-1}(B) = \varphi^{-1}(f^{-1}(B)) \in \mathcal{G}$. Hence $\varphi^{-1}\left(\bigcup_{i\in I} f^{-1}(\mathcal{E}_i)\right) \subseteq \mathcal{G}$, since $\sigma(f) = \sigma\left(\bigcup_{i\in I} f^{-1}(\mathcal{E}_i)\right)$, φ is measurable.

6.7.1.31 example Let $(\Omega, \mathcal{G})$ be a measurable space.

1. $\forall\ A \in \mathcal{G}$, $\mathbb{1}_A : \Omega \to \mathbb{R}$ is measurable. For any $U \subseteq \mathbb{R}$, $\mathbb{1}_A^{-1}(U) = A$ or $\Omega \backslash A$ or Ω or $\emptyset$.
2. Let X ,Y be topological spaces. If $f : X \to Y$ is a continuous mapping, then f is measurable with respect to Borel σ-algebra.
3. Let $(\Omega, \mathcal{G})$ be a measurable space. If $f, g : \Omega \to \mathbb{R}$ are measurable, then
$$(f + g) \quad (f \times g) \quad (f \wedge g) \quad (f \vee g)$$
are all measurable.
4. Let $(f_n)_{n\in\mathbb{N}}$ be a family of measurable mappings from Ω to $[-\infty, \infty]$ $f = \sup_{n\in\mathbb{N}} f_n$ (It's necessary to be countable f_n). Then f is measurable. (Similarly, $\inf_{n\in\mathbb{N}}$ is measurable). In fact, for any $q \in \mathbb{R}$, $\{\omega \in \Omega | f(\omega) > a\} = \bigcup_{n\in\mathbb{N}} \omega \in \Omega | f_n(\omega) > a \in \mathcal{G}$

6.8 Measure

6.8.1 σ-Additive

Let Ω be a set, $\mathcal{C}$ be a semi-algebra on Ω. $\mu : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a mapping. If $\forall\ n \in \mathbb{N}$, $\forall\ (A_i)_{i=1}^n \in \mathcal{C}^n$ pairwisely disjoint with $A = \bigsqcup_{i=1}^n A_i \in \mathcal{C}$, one has $\mu(A) = \mu(A_1) + \cdots + \mu(A_n)$, we say that μ is *additive*.
Let $S = \mathbb{R}$-vector subspace of $\mathbb{R}^\Omega$ generated by $(\mathbb{1}_A)_{A\in\mathcal{C}}$, then $I_\mu : S \to \mathbb{R}$, $\quad \sum_{i=1}^n \lambda_i \mathbb{1}_{A_i} \mapsto \sum_{i=1}^n \lambda_i \mu(A_i)$ is well-defined. If I_μ is an integral operator, we say that μ is *σ-additive*.

6.8.2 Measure Space

Let $(\Omega, \mathcal{G})$ be a measurable space, $\mu : \mathcal{G} \to [0, +\infty]$ be a mapping. If $\mu(\emptyset) = 0$ and if for any $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$ pairwisely disjoint, $\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$, $(\Omega, \mathcal{G}, \mu)$ is called a *measure space*.

6.8.2.1 Def If $\exists\ (A_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$ s.t. $\Omega = \bigcup_{n\in\mathbb{N}} A_n$ and $\mu(A_n) < +\infty$, then μ is said to be *σ-finite*

6.8.2.2 Example $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ is σ-finite because $\mathbb{R} = \bigcup_n [-n, n]$.

6.8.3 Caratheodory

Let Ω be a set, $\mathcal{C}$ be a semi-algebra on Ω, $\mu : \Omega \to \mathbb{R}_{\geq 0}$ be a σ-additive mapping. Assume that there is a sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{C}^{\mathbb{N}}$ s.t. $\Omega = \bigcup_{n\in\mathbb{N}} A_n$. The μ extends to a σ-finite measure on $\sigma(\mathcal{C})$

6.8.3.1 Proof Let $S \subseteq \mathbb{R}$ be the vector subspace generated by $\mathbb{1}_A$, $A \in \mathcal{C}$. Let $\mathcal{G} = \{A \subseteq \Omega | \mathbb{1}_A \in L^1(I_\mu)^\uparrow\}$, then $\mathcal{G}$ is a σ-algebra containing $\mathcal{C}$. Hence, $\sigma(\mathcal{C}) \subseteq \mathcal{G}$.
Moreover, $(A \in \mathcal{G}) \mapsto I_\mu(\mathbb{1}_A)$ is a measure on $\mathcal{G}$, which is σ-algebra.

6.8.3.2 Example $\Omega = \mathbb{R}$, $\mathcal{C} = \{]a,b] | (a,b) \in \mathbb{R}^2, a < b\}$, $\sigma(\mathcal{C}) =$ Borel σ-algebra, $\varphi : \mathbb{R} \to \mathbb{R}$ is increasing right continuous. $\mu_\varphi : \mathcal{C} \to \mathbb{R}_{\geq 0}$, $\quad]a,b] \mapsto \varphi(b) - \varphi(a)$ is σ-additive.
Hence, μ_φ extends to a measure $\sigma(\mathcal{C}) \to [0, +\infty]$ called the *Stieltjes measure*. In particular case where $\mu(x) = x$, $\forall\, x \in \mathbb{R}$, μ_φ is called a *Lebesgue measure*.

6.8.3.3 Def Let $(\Omega, \mathcal{G}, \mu)$ be a σ-finite measure space, where $\varphi = \{A \in \mathcal{G} | \mu(A) < +\infty\}$ is a semi-algebra and $\sigma(\mathcal{C}) = \mathcal{G}$, and $\mu_{|\mathcal{C}}$ is additive. We denoted by $L^1(\Omega, \mathcal{G}, \mu)$ the set of measurable mappings $f : \Omega \to \mathbb{R}$, that belongs to $L^1(I_\mu)$. For $f \in L^1(\Omega, \mathcal{G}, \mu)^\uparrow$, $I_\mu(f)$ is denoted as:

$$\int_\Omega f(\omega)\mu(d\omega)$$

Particular case:
If $\Omega = \mathbb{R}$, $\mu = \mu_\varphi$ Stieltjes measure, $\int_\mathbb{R} f(x)\mu_\varphi(dx)$ is denoted as $\int_\mathbb{R} f(x)d\varphi(x)$.

6.8.3.4 Prop Let $(\Omega, \mathcal{G}, \mu)$ be a σ-finite measure space, $f : \Omega \to \mathbb{R}$ be a measurable mapping. If $\exists\, g \in L^1(\Omega, \mathcal{G}, \mu)$, $g \leq f$, then $f \in L^1(\Omega, \mathcal{G}, \mu)^\uparrow$.

6.8.3.5 Proof By replacing f by $f - g$, we may assume that $g = 0$. Consider first the case where $f = \mathbb{1}_B$, $B \in \mathcal{G}$. Let $(A_n)_{n\in\mathbb{N}}$ be an increasing sequence in $\mathcal{G}$, $\mu(A_n) < +\infty$,, $\bigcup_{n\in\mathbb{N}} A_n = \Omega$. Then $\mathbb{1}_B = \lim_{n\to+\infty} \mathbb{1}_{B\cap A} \in L^1(\Omega, \mathcal{G}, \mu)^\uparrow$. In general, $f = \lim_{n\to+\infty} f_n$,

$$f_n = \sum_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{\{\omega\in\Omega | \frac{k}{2^n} \leq f(\omega) < \frac{k+1}{2^n}\}} + n\mathbb{1}_{\{\omega\in\Omega | f(\omega) \geq n\}}$$

6.8.3.6 Corollary Let $f : \Omega \to \mathbb{R}$ be a measurable mapping. Then $f \in L^1(\Omega, \mathcal{G}, \mu)$ iff $\int_\Omega |f(\omega)|\mu(d\omega) < +\infty$.

6.8.3.7 Proof $\Rightarrow$:
One has $f \in L^1(I_\mu)$, hence $|f| \in L^1(I_\mu)$, so $I_\mu < +\infty$ $\Leftarrow$:
Suppose that $\int_\Omega |f(\omega)|\mu(d\omega) < +\infty$, since $f \vee 0$ and $-(f \wedge 0)$ belong to $L^1(\Omega, \mathcal{G}, \mu)^\uparrow$, and $f \vee 0 \leq |f|$, $-(f \wedge 0) \leq |f|$, so $f \vee 0$ and $-(f \wedge 0)$ belong to $L^1(\Omega, \mathcal{G}, \mu)$. Hence, $f = f \vee 0 - f \wedge 0 \in L^1(\Omega, \mathcal{G}, \mu)$.

6.9 Fundamental Theorem of Calculus

6.9.0.1 Theorem Let J be an open interval over $\mathbb{R}$, $x_0 \in J$, $f : J \to \mathbb{R}$ be a continuous mapping.

1. $\forall\, (a,b) \in J^2$, $a < b$,

$$\begin{aligned} \mathbb{R} \to & \quad \mathbb{R} \\ \mathbb{1}_{]a,b]} f : x \mapsto & \quad f(x) \quad x \in]a,b] \\ & \quad 0 \quad x \notin]a,b] \end{aligned}$$

   belongs to $L^1(\mathbb{R}, \mathcal{B}, \mu)$, where $\mathcal{B}$ is Borel σ-algebra, μ is Lebesgue measure.

2. Let $F : J \to \mathbb{R}$, $F(x) := \int_{x_0}^x f(t)dt$. Then F is differentiable on J with $F'(x) = f(x)$

6.9.0.2 Corollary If $G : J \to \mathbb{R}$ is a mapping s.t. $G' = f$, then $\forall\, (a,b) \in J^2$, $a < b$, $G(b) - G(a) = \int_a^b f(t)dt$

6.9.0.3 Proof

1. f is bounded by $[a,b]$, hence $\int_\mathbb{R} \mathbb{1}_{]a,b]} |f| dx < +\infty$.
2. Let $x \in J$, $h > 0$ s.t. $[x, x+h] \subseteq J$, f is uniformly continuous on $[x, x+h]$. For $0 < t \leq h$, $\inf f_{|[x,x+h]} \leq \frac{F(x+t)-F(x)}{t} = \frac{1}{t}\int_x^{x+t} f(s)ds \leq \sup f_{|[x,x+h]}$, since f is continuous, $\lim_{t\to 0} \inf f_{|[x,x+h]} = \lim_{t\to 0} \sup f_{|[x,x+h]} = f(x)$, so $\lim_{t>0,\ t\to 0} \frac{F(x+t)-F(x)}{t} = f(x)h$. Similarly, $\lim_{t>0,\ t\to 0} \frac{F(x)-F(x-t)}{t} = f(x)h$, hence $F'(x) = f(x)$.

6.9.0.4 Remark

1. Let F and G be two mappings of class C^1 from J to $\mathbb{R}$, then $F \times G$ is of class C^1 and $(F \times G)' = F' \times G + F \times G'$. Let $f = F'$, $g = G'$, then $\forall\ (a,b) \in J^2$, $a < b$, $\int_a^b f(t)G(t)dt = F(b)G(b) - F(a)G(a) - \int_a^b F(t)g(t)dt$.

2. Let $\varphi : I \to J$ be a mapping of class C^1, where I is an interval. Let $F : J \to \mathbb{R}$ be a mapping of class C^1, $(f \circ \varphi)'(x) = F'(\varphi(x))\varphi'(x)$. Then $\forall\ (\alpha, \beta) \in I^2$, $\alpha < \beta$, $\int_\alpha^\beta F'(\varphi(x))\varphi'(x)dx = (F \circ \varphi)(\beta) - (F \circ \varphi)(\alpha)$.

6.10 L^p space

6.10.0.1 Def We fix a measure space $(\Omega, \mathcal{G}, \mu)$. Let $p \in \mathbb{R}_{\geq 1}$, we denote by $L^p(\Omega, \mathcal{G}, \mu)$ the set of measurable mappings $f : \Omega \to \mathbb{R}$, s.t. $\|f\|_{L^p} := \left(\int_\Omega |f(x)|^p \mu(dx)\right)^{\frac{1}{p}}$.

6.10.0.2 Lemma Let $(p,q) \in \mathbb{R}^2_{\geq 1}$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$. For any $(a,b) \in \mathbb{R}^2$, $\frac{a^p}{p} + \frac{b^q}{q} \geq ab$.

6.10.0.3 Proof Since *exp* is convex, $\frac{1}{p} \exp p \ln a + \frac{1}{q} \exp q \ln b \geq \exp \ln a + \ln b = ab$.

6.10.1 Holder's Inequality

Let $f : \Omega \to \mathbb{R}$ and $g : \Omega \to \mathbb{R}$ be measurable mappings. One has $\|f \times g\|_{L^1} \leq \|f\|_{L^p} + \|g\|_{L^p}$.

6.10.1.1 Proof Take $\varphi = \frac{f}{\|f\|_{L^p}}$ and $\psi = \frac{g}{\|g\|_{L^p}}$, $|\varphi(x)\psi x| \leq \frac{|\varphi(x)|^p}{p} + \frac{|\psi(x)|^q}{q}$. $\frac{\int_\Omega |\varphi(x)\psi(x)|\mu(dx)}{\|f\|_{L^p}\|g\|_{L^p}} \leq \frac{\int_\Omega |f(x)|^p \mu(dx)}{p\|f\|^p_{L^p}} + \frac{\int_\Omega |g(x)|^q \mu(dx)}{q\|g\|^q_{L^p}} = \frac{1}{p} + \frac{1}{q} = 1$.

6.10.1.2 Corollary Let $p \geq 1$, $\forall\ (f,g) \in L^p(\Omega, \mathcal{G}, \mu)$, $\|f+g\|_{L^p} \leq \|f\|^{l^p} + \|g\|^{L^p}$

6.10.1.3 Proof Apply Holder's inequality to show: $f \times (f+g)^{p-1}$ and $g \times (f+g)^{p-1}$.

Chapter 7

Multilinear Algebra

7.1 Tensor Products of Linear Spaces

7.1.1 Goal

Study in a systemmetical way the relation between linear mappings and multilinear mappings.

7.1.2 Tensor Products

Let R be a commutative ring with unity.

7.1.2.1 Theorem Let M and N be two R-modules. There exists an R-module denoted by $M \otimes_R N$ and a bilinear mapping $t : M \times N \to M \otimes_R N$ having the following properties:

1. $\forall$ R-module P and any bilinear mapping $s : M \times N \to P$, $\exists!$ linear mapping $f_s : M \otimes_R N \to P$ s.t. $s = f_s \circ t$. (Universal property)

2. If T,t' is another couple satisfying (1) s.t. $s = g_s \circ t'$, then there exists an unique isomorphism $T \cong M \otimes_R N$.

7.1.2.2 Proof

1. Let $\mathcal{F}$ be the free R-module generated by $M \times N$, that is, $\mathcal{F} = \{\Sigma_{finite(i,j)} a_{i,j}(m_i, n_i) \mid a_{i,j} \in R, m_i \in M, n_j \in N\}$.

   Let $\mathcal{G}$ be the free R-submodule generated by the elements of the following shape (written in the following way):

$$\begin{cases} (m + m', n) - (m, n) - (m', n) \\ (m, n + n') - (m, n) - (m, n') \\ (rm, n) - r(m, n) \\ (m, rn) - r(m, n) \end{cases}$$

   $\forall (m, m') \in M^2,\ (n, n') \in N^2,\ r \in R.$

   Then we can define $M \otimes_R N := \mathcal{F}/\mathcal{G}$, so $t(m + m', n) = t(m, n) + t(m', n),\ t(m, n + n') = t(m, n) + t(m, n')$, so t is bilinear. So $f_s(\mathcal{G} + (m, n)) = s(m, n)$, which is linear and unique.

2. $f_{t'} \circ g_t \circ t' = f_{t'} \circ t = t'$, so $f_{t'} \circ g_t = \mathrm{Id}_T$. (Using the first property.)

7.1.2.3 Def The R-module $M \otimes_R N$ constructed above is called the tensor product of M and N. An element of $M \otimes_R N$ is called tensor. We denote $m \otimes_R n$ as $t(m, n)$, and any element of the form is called pure tensor.

7.1.2.4 Remark Pure tensors generate $M \otimes_R N$. In particular, any tensor can be written of pure tensors.

7.1.2.5 Corollary The mapping $s \mapsto f_s$ defined above gives an isomorphism: $\mathscr{L}(M, N; P) \cong \mathscr{L}(M \otimes_R N; P)$ for any R-module P.

7.1.2.6 Proof

- surj: Take $\phi \in \mathscr{L}(M \otimes_R N; P)$, then $t \circ \phi$ is bilinear, which is in $\mathscr{L}(M, N; P)$.

- inj: $\forall s \in \mathscr{L}(m, n; P)$, which is bilinear. By **7.1.2.1 Theorem**, we can infer that f_s exists, hence injective.

7.1.2.7 Prop

1. $M \otimes_R N \cong N \otimes_R M$.
2. $M \otimes_R (N \otimes_R P) = (M \otimes_R N) \otimes_R P$.

   So that we can remove the parenthesis and write $M_1 \otimes_R M_2 \otimes_R \cdots \otimes_R M_n$. We call it *n-fold tensor product.*
3. Show that $M_1 \otimes_R M_2 \otimes_R \cdots \otimes_R M_n$ factorizes the multi-linear mappings and $\mathscr{L}(M_1, \ldots, M_n; P) \cong \mathscr{L}(M_1 \otimes_R \ldots \otimes_R M_n; P)$. We can solve this by induction.

We have the general definition of tensors product for R-modules, but we are interested in the case where R is a field. And we will omit R this time.

7.1.2.8 Lemma Let $V_1, \ldots, V_n$ be vector spaces of finite dimension d. Let $\{e_{i_1}, e_{i_1} \ldots, e_{i_{d_i}}\}$ be a basis of V_i. Let's define the following functions: $\phi_{i_1,\ldots,i_n} : V_1 \times V_2 \times \cdots \times V_n \to K \quad (v_1, \ldots, v_n) \mapsto \Pi_{j=1}^n e_{j,i_j}^\vee(v_j)$. Then the set $\phi_{i_1,\ldots,i_n}$ is a basis of $\mathscr{L}(V_1, \ldots, V_n; K)$.

7.1.2.9 Proof We do the proof for $n = 2$, since $n = 1$ is obvious. Then the generated case follows by induction.

$V_1 = \langle e_1, \ldots, e_{d_1} \rangle$, $V_2 = \langle \omega_1, \ldots, \omega_{d_2} \rangle$. ϕ denoted by $\xi_{i,j}(x, y) = e_i^\vee(x)\omega_j^\vee(y)(x \in V_1, y \in V_2)$.

To show that $\{\xi_{i,j}\}$ is a generated set: Take $\Phi \in \mathscr{L}(V_1, V_2; K)$, let $A_{i,j} = \Phi(e_i, \omega_j) \in K$. $\Phi(x, y) = \Phi(\Sigma_i \alpha e_i, \Sigma_j \beta_j \omega_j) = \Sigma_{i,j}\alpha_i\beta_j\Phi(e_i, \omega_j) = \Sigma_{i,j}\alpha_i\beta_j A_{i,j}$, Let $f = \Sigma_{i,j} A_{i,j}\xi_{i,j}$, $f(x, y) = f(\Sigma_i \alpha e_i, \Sigma_j \beta_j \omega_j) = \Sigma_{i,j}\alpha_i\beta_j f(e_i, \omega_j) = \Sigma_{i,j}\alpha_i\beta_j A_{i,j}$. So we have proved that $\xi_{i,j}$ are system of generators.

Now we prove that $\xi_{i,j}$ are linearly independent: Assume that $\Sigma_{i,j} A_{i,j}\xi_{i,j}(x, y) = 0, \forall(x, y) \in V_1 \times V_2$. Put in $(x, y) = (e_i, \omega_j) \Rightarrow A_{i,j} = 0 \quad \forall i, j$.

7.1.2.10 Prop Assume that $V_1, \ldots, V_n$ are vector spaces and V_i has basis given by $\{e_{i_1}, \ldots, e_{i_{d_i}}\}$. Then $\mathcal{B} = \{e_{1i_1} \otimes \ldots \otimes e_{ni_n} : 1 \le i_j \le d_j\}$ is a basis for $V_1 \otimes \ldots \otimes V_n$. In particular, $V_1 \otimes \ldots \otimes V_n$ has dimension $\Pi_i d_i$.

7.1.2.11 Proof Again we assume $n = 2$. $V_1 = \langle e_1, \ldots, e_m \rangle$, $V_2 = \langle \omega_1, \ldots, \omega_n \rangle$.

We knew that $\mathscr{L}(V_1, V_2; K) \cong (V_1 \otimes V_2)^\vee \quad s \mapsto f_s$. Let's check what happens when $s = \xi_{i,j}$.

Recall that $f_{\xi_{i,j}}(x \otimes y) = \xi_{i,j}(x, y) = e_i^\vee(x)e_j^\vee(y)$. So $f_{\xi_{i,j}}(e_k \otimes w_l) = \begin{cases} 1 & if\ (i,j) = (k,l) \\ 0 & otherwise. \end{cases}$. It follows that $\{e_k \otimes w_l\}_{k,l}$ is a basis of $V_1 \otimes V_2$.

7.1.3 Tensor Product and Duality

7.1.3.1 Prop Let $V_1, \ldots, V_n$ be vector spaces as above. Then $(V_1^\vee \otimes \ldots \otimes V_n^\vee) \cong (V_1 \otimes \ldots \otimes V_n)^\vee$.

7.1.3.2 Proof Define $V_1^\vee \times \cdots \times V_n^\vee \to \mathscr{L}(V_1, \ldots, V_n; K) \cong (V_1 \otimes \ldots \otimes V_n)^\vee \quad (\phi_1, \ldots, \phi_n) \mapsto [(v_1, \ldots, v_n) \mapsto \Pi_i\phi_i(v_i)]$. This map is multi-linear. It descends by the property of tensor product to a map $F : V_1^\vee \otimes \ldots \otimes V_n^\vee \to \mathscr{L}(V_1, \ldots, V_n; K) \cong (V_1 \otimes \ldots \otimes V_n)^\vee \quad \phi_1 \otimes \ldots \otimes \phi_n \mapsto [(v_1, \ldots, v_n) \mapsto \Pi_i\phi_i(v_i)]$.

By **7.1.2.10 Prop**, these two spaces have the same dim Πd_i. It is enough to show that the map is surjective. Let's do it for $n = 2$. (Keep the same notation as above.) Take $\xi_{i,j}$, $\xi_{i,j}(x, y) = e_i^\vee(x)e_j^\vee(y) = F(e_i^\vee \otimes e_j^\vee)$.

7.1.3.3 Prop Let V and W be two vector spaces (finite dim). Then $\mathscr{L}(V, W) \cong V^\vee \otimes W$.

7.1.3.4 Proof $s : V^\vee \times W \to \mathscr{L}(V, W) \quad (\phi, w) \mapsto [v \mapsto \phi(v)w]$. Let's check that s is bilinear: $((\phi + \psi)(v))w = (\phi(v) + \psi(v))w = \phi(v)w + \psi(v)w$, $\phi(v)(w + w') = w\phi(v) + w'\phi(v)$, so bilinear.

So it indices $f_s : V^\vee \otimes W \to \mathscr{L}(V, W)$. We have to show that this is the required isomorphism.

Let $\{v_1^\vee, \ldots, v_m^\vee\}$ be a basis of $V^\vee$ and let $\{w_1, \ldots, w_n\}$ be a basis of W. Let's see what happens to $f_s(v_i^\vee \otimes w_j) = [v_k \mapsto v_i^\vee(v_k)w_j = \delta_{i,k}w_j]$.

Let's see what happens to $f_s(v_i^\vee \otimes w_j) = [v_k \mapsto v_i^\vee(v_k)w_j = S_{i,k}w_j]$.

Consider the matrix associated to f_s with respect to the fixed bases. Call this matrix $M_{a,b} = \begin{cases} 1 & if\ (a,b) = (j,i) \\ 0 & otherwise. \end{cases}$.

The matrices of this form are a basis of $\mathscr{L}(K^m, K^n) \cong \mathscr{L}(V, W)$. The proof is over.

An important case of this prop is when V=W, $\mathscr{L}(V, V) \cong V \otimes V$. More general, $\mathscr{L}(V, W) \to V^\vee \otimes W \quad f \mapsto \Sigma a_{i,j}v_i^\vee \otimes w_j$. $(a_{i,j})$ is exactly the matrix associated to f with respect to the bases $\{v_i\}$ $\{w_i\}$. For instance $V = W$, $\mathrm{Id}_V \in \mathscr{L}(V, V) \mapsto \Sigma_i v_i^\vee \otimes v_i$.

7.1.3.5 Prop Let M,N,P be R-modules. $\mathscr{L}(M \otimes_R N; P) \cong \mathscr{L}(M, \mathscr{L}(N, P))$.

7.1.3.6 Proof Use **7.1.3.3 Prop**.

7.1.3.7 Def Let M_1, M_2, N_1, N_2 be R-modules and let $f_i : M_i \to N_i$ be linear maps. Then we define $f_1 \otimes f_2 : M_1 \otimes M_2 \to N_1 \otimes N_2 \quad m_1 \otimes m_2 \mapsto f_1(m_1) \otimes f_2(m_2)$. This is a linear map.

7.1.4 Extension of Scalars

Let $\phi : R \to S$ be a ring homomorphism (commutative, with unity). Let M be a R-module. Our goal is to give to M also a structure of S-module "carried by ϕ".

Notice that S has a structure of R-module: $s \in S, r \in R$; $rs := \phi(r)s$. Now take the tensor product $M \otimes_R S$. By def, it's an R-module.

Now we give a structure of S-module to $M \otimes_R S$. Take $s \in S$, $m \otimes s' \in M \otimes_R S$, $s(m \otimes s') := m \otimes ss'$ (ss' is the multiplication in S). Note that $M \otimes_R S$ is a S-module.

Note that we have a map: $i : M \to M \otimes_R S \quad m \mapsto m \otimes 1$ (map of R-modules). Be careful: In general, the map i is not injective. e.g. $R = \mathbb{Z}$, $S = \mathbb{Z}/2\mathbb{Z}$, $\alpha : \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, $M = \mathbb{Z}[t]$. $i(2x) = 2x \otimes 1 = 2(X \otimes 1) = X \otimes \alpha(2) \times 1 = X \otimes 0 = 0$.

7.1.4.1 Prop Let $K \subseteq L$ be a field extension and let V be a K-vector space. Moreover, let's denote $V_L := V \otimes_K L$. If $\{e_i\}_{i=1}^n$ is a basis of V, then $\{e_i \otimes 1\}$ is a L-basis of V_L (V_L has the same dim of V).

7.1.4.2 Proof The set $\{e_i \otimes 1\}$ generates V_L. In fact, $v \otimes l = (\Sigma \alpha_i e_i \otimes l) = \Sigma l \alpha_i (e_i \otimes 1)$. We have to show that the elements are linearly independent. $0 = \Sigma \alpha_i (e_i \otimes 1) = \Sigma e_i \otimes \alpha_i; \alpha_i \in L$. Define the map $b_i : V \times L \to L \quad (\Sigma \lambda_i e_i, \beta) \mapsto \lambda_i \beta$. This map is bilinear.

It induces a map $f_i = f_{b_i} : (\Sigma \lambda_i e_i) \otimes \beta \mapsto \lambda_i \beta_i$. Note that $f_i(e_j \otimes \beta) = \delta_{i,j} \beta$. Go back to the expression: $f_i(\Sigma_j e_j \otimes \alpha_j) = \alpha_i$. But $0 = f_i(0) = f_i(\Sigma_j e_i \otimes \alpha_j) = \alpha_i$, $\forall i$.

7.1.4.3 Remark As a consequence, we have that the map $i : V \to V_L$ (map of K-vector spaces) is injective.

7.1.4.4 Prop $V \otimes_K K \cong V \quad v \otimes a \mapsto av$.

7.1.5 Exactness of the Tensor Product

Fix a R-module N, and consider $\_ \otimes N : M \mapsto M \otimes_R N$ for any R-module M.

Moreover, for any linear map $f : M \to P$, we induce a map $f \otimes \mathrm{Id}_N : M \otimes_R N \to P \otimes_R N$. This association sends Id_M to $\mathrm{Id}_{M \otimes_R N}$ and moreover, it's well behaved with respect to the composition: $f \circ \mapsto (f \circ g) \otimes \mathrm{Id}_N = (f \otimes \mathrm{id}_N) \circ (g \otimes \mathrm{id}_N)$, $g : M \to P, f : P \to E$.

7.1.5.1 Def A sequence of R-modules (chain complex of R-modules) is a diagram of the following form: $M_1 \xrightarrow{d^1} M_2 \xrightarrow{d^2} M_3 \xrightarrow{d^3} \cdots$. d^i are morphisms.

7.1.5.2 Remark $d^{i+1} \circ d^i = 0 \Leftrightarrow \mathrm{Im}(d^i) \subseteq d^{i+1}$.

Why? Some information about morphisms of R-modules can be nicely expressed in this language.

Here $\forall i$ M_i is an R-module, d^i is a linear mapping, s.t. $\ker(d^{i+1}) \subseteq \mathrm{Im}(d^i)$. The sequence is *exact* when $\ker(d^{i+1}) = \mathrm{Im}(d^i)$, $\forall i$. $0 \to M \xrightarrow{f} N \to 0$: f is an isomorphism, iff the complex is exact.

Take a morphism $f : M \to N$, then f is injective iff $0 \to M \xrightarrow{f} N$ is exact, f is surjective iff $M \xrightarrow{f} N \to 0$ is exact.

The first theorem of homomorphism (namely, $\mathrm{Im}(f) = f(M/\ker(f))$) can be written as an exact sequence: $0 \to \ker(f) \xrightarrow{i} M \xrightarrow{f} \mathrm{Im}(f) \to 0$.

More in general, sequences in the form: $0 \to M_1 \to M_2 \to M_3 \to 0$ are called short exact sequence.

N,P is an R-module; we have morphisms of R-modules: $\forall M$, $M \mapsto M \otimes_R N$. Let $f : M \to P$ be R-linear. $f \otimes \mathrm{id}_N : M \otimes_R N \to P \otimes_R N \quad m \otimes n \mapsto f(m) \otimes n$.

Assume that we have a short exacct sequence of R-modules.

$0 \to M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \to 0$.

Let us transform this by adding all $- \otimes_R N$, it becomes:

$0 \to M_1 \otimes_R N \xrightarrow{f \otimes_R N} M_2 \otimes_R N \xrightarrow{g \otimes_R N} M_3 \otimes_R N \to 0$.

Example: Use the same construction before and the first mapping is zero mapping, which is not injective.

7.1.5.3 Exercise(Important) If R=K, then $- \otimes_K N$ (where N is a finite dim vector space) is exact. **Hint:** use the bases.

7.1.5.4 Answer Let $(e_i)_{i=1}^n$ be a basis of N. $\forall v \otimes n \in M_1 \otimes N$ can be written as the form $\Sigma_{i=1}^n a_i \otimes e_i$. $f \otimes \mathrm{id}_N(\Sigma_{i=1}^n a_i \otimes e_i) = \Sigma_{i=1}^n f(a_i) \otimes e_i$. Since $\{e_i\}$ is a basis, one can construct a projection mapping to show $f(a_i) = 0$, $\forall i$. The concrete construction is let $s_i : M_2 \times N \to M_2 \quad (m, n) \mapsto e_i^\vee(n)m$, which is bilinear, which induces $f_{s_i} : M_2 \otimes N \to M_2$

7.1.6 Tensor Algebra

Fix a vector space V over K of finite dimension.
Let us denote: $T_p^q(V) := (V^\vee)^{\otimes p} \otimes V^{\otimes q} \quad p, q \in \mathbb{N}$.
An element of $T_p^q(V)$ is called a tensor of type (p, q) or a mixed tensor, which is p-covariant and q-contravariant.
Let's denote: $T(V) := \bigoplus_{q=0}^{\infty} T_0^q(V)$, then:

- $T_0^0(V) := K$.
- $T_1^0(V) = V^\vee$.
- $T_0^1(V) = V$.
- $T_1^1(V) = V^\vee \otimes V \cong \mathscr{L}(V; V)$.
- $T_2^0(V) = V^\vee \otimes V^\vee \cong (V \otimes V)^\vee \cong \mathscr{L}(V, V; K)$.

In $T(V)$ we have the following operation:
$T_0^l(V) \times T_0^q(V) \to T_0^{l+q}(V) \quad ((x_1 \otimes x_2 \otimes \ldots \otimes x_l), (y_1 \otimes y_2 \otimes \ldots \otimes y_q)) \mapsto (x_1 \otimes x_2 \otimes \ldots \otimes x_l) \otimes (y_1 \otimes y_2 \otimes \ldots \otimes y_q)$ and we can extend it by linearity.
With this operation, T(V) becomes a K-algebra. It's called the tensor algebra associated to V.

7.2 Exterior Product

7.2.1 Def

Let W be the two-sided ideal of $T(V)$ generated by the elements of type $x \otimes x$.
$W = \{\Sigma_{finite\ i}(y_i \otimes \ldots \otimes y_{m_i}) \otimes (x_i \otimes x_i) \otimes (z_i \otimes \ldots \otimes z_{n_i}) : \ x_i, y_i, z_i \in V, \ n_i, m_i \in \mathbb{N}\}$.
The quotient algebra $\bigwedge(V) := T(V)/W$ is called the *exterior algebra* of V, which is a K-algebra.
$\pi : T(V) \to \bigwedge(V) \quad x_1 \otimes \ldots \otimes x_n \mapsto \pi(x_1 \otimes \ldots \otimes x_n) =: x_1 \wedge \cdots \wedge x_n$.
$\bigwedge^n(V) := T_0^n(V)/(W \bigcup T_0^n(V))$ and we call it *n-fold wedge product*, and we can use it to define $\bigwedge(V) = \bigoplus_{n=0}^{\infty} \bigwedge^n(V)$, and we call it *extension product*.

7.2.1.1 Prop Let $\sigma \in S_n$, then $x_1 \wedge \cdots \wedge x_n = \text{sgn}(\sigma) x_{\sigma(1)} \wedge \cdots \wedge x_{\sigma(n)}$.

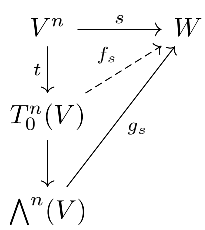
7.2.1.2 Proof Since any permutation can be written as the product of adjacent transformations, it is enough to do the proof for $\sigma = (i, i+1)$.
$0 = (x_i + x_{i+1}) \wedge (x_i + x_{i+1}) = (x_i \wedge x_{i+1}) + (x_{i+1} \wedge x_i)$.

7.2.1.3 Question Can the exterior algebra be defined in a similar way?

7.2.2 Alternating

V is a K-vector space, a multi-linear map $\phi : V \times V \times \cdots \times V \to W$ is called *skew-symmetric* (or *alternating*) if $\phi(x_1, \ldots, x_n) = 0$ when $\exists i \neq j$ s.t. $x_i = x_j$.

7.2.2.1 Prop Fix a vector space V. $\forall$ alternating multi-linear map $s : V^n \to W$ when W is another vector space, there exists a unique linear map $g_s : \bigwedge^n V \to W$ s.t. the following diagram commutes.



7.2.2.2 Proof $g_s(v_1 \wedge \cdots \wedge v_n) := s(v_1, \ldots, v_n)$.

7.2.2.3 Remark/Exercise The couple $\bigwedge^n V$ with $V \times \cdots \times V \to \bigwedge^n(V)$ that satisfies **7.1.7.5 Prop**, is unique up to unique isomorphism.

7.2.2.4 Notation *automorphism* is an isomorphism $f : V \to V$.

7.2.2.5 Prop Let V be a vector space of dimension n with a basis $\{e_1, \dots, e_n\}$. Then $\bigwedge^k(V)$ is a vector space with a basis given by $\mathcal{B} = \{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \leq i_1 < \cdots < i_k \leq n\}$.

In particular, $\bigwedge^k(V)$ has dimension $\binom{n}{k}$.

7.2.2.6 Proof $\mathcal{B}$ is clearly a generating set. The different part is to show that $\mathcal{B}$ is made up of linearly independent elements.

$I = (i_1, \dots, i_k)$ with $1 \leq i_1 < \cdots < i_k \leq n$, define: $\phi_I : V \times \cdots \times V \to K \quad (e_{j_1}, \dots, e_{j_k}) \mapsto \begin{cases} sgn(\tau) & if\ \exists \tau \in S_I, \tau(j_m) = i_m \\ 0 & otherwise. \end{cases}$.

Φ_I is multi-linear and alternating, hence it induces a linear map.

$\overline{\phi_I} := g_{\phi_I} : \bigwedge^k(V) \to K \quad defined\ as\ above.$

Assume that $0 = v \in \bigwedge^k(V) = \Sigma_{1 \leq j_1 < \cdots < j_k \leq n} \lambda_{j_1,\dots,j_k} e_{j_1} \wedge \cdots \wedge e_{j_k}$.

By linearity, $0 = \overline{\phi_I}(v) = \pm\lambda_I$. Do it for any possible I. This shows that any $\lambda_{j_1,\dots,j_k} = 0$.

7.3 Determinant

7.3.1 Def

Let V be a vector space of dimension n, then $\det(V) := \bigwedge^n(V)$ is called the determinant of V. It's a vector space of dimension $1 = \binom{n}{n}$ and a basis is given by $e_1 \wedge \cdots \wedge e_n$ when $\{e_1, \dots, e_n\}$ is a basis of V.

Let $f \in \mathscr{L}(V;V)$, then consider $\tilde{f} : V^k \to \bigwedge^k(V) \quad (v_1, \dots, v_k) \mapsto f(v_1) \wedge \cdots \wedge f(v_k)$, ans this is multi-linear and alternating. Therefore, it induces a map $g_{\tilde{f}}$. Since $\det(f)$ has dim 1, $\det(f) : v_1 \wedge \cdots \wedge v_n \mapsto \det_f \times (v_1 \wedge \cdots \wedge v_n) = f(v_1) \wedge \cdots \wedge f(v_n)$.

By abuse of notation, we identify $\det(f) = \det_f$.

7.3.1.1 Prop $f \in \mathscr{L}(V, V)$ is invertible iff $det(f) \neq 0$.

7.3.1.2 Proof f is not invertible iff $\{f(e_1), \dots, f(e_n)\}$ is not a basis iff there is a non-trivial linear combination: $\Sigma_i \lambda_i f(e_i) = 0$. After relabeling the e_i, we can assume $f(e_1) = \Sigma_{i \geq 2} \mu_i f(e_i)$.

So $\det(f)(e_1 \wedge \cdots \wedge e_n) = \det_f \times (e_1 \wedge \cdots \wedge e_n) = (\Sigma_{i \geq 2} \mu_i f(e_i)) \wedge f(e_2) \wedge \cdots \wedge f(e_n) = \Sigma_{i \geq 2} \mu_i (f(e_i) \wedge f(e_2) \wedge \cdots \wedge f(e_n)) = 0$.

7.3.2 Binet Theorem

$\det(f \circ g) = \det(f) \circ \det(g)$.

7.3.2.1 Proof $\det(f \circ g)(e_1 \wedge \cdots \wedge e_n) = (f \circ g)(e_1) \wedge \cdots \wedge (f \circ g)(e_n) = f(g(e_1)) \wedge \cdots \wedge f(g(e_n)) = (\det f)(g(e_1) \wedge \cdots \wedge g(e_n)) = \det(f) \circ \det(g)(e_1 \wedge \cdots \wedge e_n)$.

7.3.3 Determinant

$f : V \to W$ is a linear map, $k \in \mathbb{N}$. $\bigwedge^k f : \bigwedge^k(V) \to \bigwedge^k(W) \quad v_1 \wedge \cdots \wedge v_k \mapsto f(v_1) \wedge \cdots \wedge f(v_k)$. When V=W, $n = \dim(V)$, $\bigwedge^n(V) = \det(V)$; $\bigwedge^n f = \det(f)$.

In this case we identify $\det(f)$ with an element of K.

7.3.3.1 Prop The determinant of f is equals to the determinant of any matrix that represents f with respect to a fixed basis. This doesn't depend on the choice of the basis.

Fix $f \in \mathscr{L}(V;V)$, let $\{v_1, \dots, v_n\}$ be a basis of V, $b : e_i \mapsto v_i \quad e_i = \{0, 0, \dots, i, \dots, 0\}$.

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ {\scriptstyle b}\uparrow & & \uparrow{\scriptstyle b} \\ K^n & \xrightarrow{A_f} & K^n \end{array}$$

Take $\bigwedge^n(\cdot)$:

$$\begin{array}{ccc} \det(V) & \xrightarrow{\det(f)} & \det(V) \\ {\scriptstyle \bigwedge^n b}\uparrow & & \uparrow{\scriptstyle \bigwedge^n b} \\ \det(K^n) & \xrightarrow{\det(A_f)} & \det(K^n) \end{array}$$

$A_f = b^{-1} \circ f \circ b \Rightarrow \det(A_f) = \bigwedge^n(b)^{-1} \times \det(f) \times \bigwedge^n(b) \in K$.

7.3.3.2 Upper Triangular Matrix Let A be an upper triangular matrix. $\det(A)(e_1 \wedge \cdots \wedge e_n) = A(e_1) \wedge \cdots \wedge A(e_n) = (a_{1,1}e_1) \wedge (a_{1,2}e_1 + a_{2,2}e_2) \wedge \cdots \wedge (a_{1,n}e_1 + \cdots + a_{n,n}e_n) = (\Pi_1^n a_{i,i})e^1 \wedge \cdots \wedge e_n$. So $\det(A) = \Pi_{i=1}^n a_{i,i}$.

1. $A = (a_{i,j})$. If one column of A can be expressed as a linear combination of other columns of A, then $\det(A) = 0$.

   The columns are images of $\{e_1, \ldots, e_n\}$, means that $A(e_1), \ldots, A(e_n)$ are linearly dependent, which infers to A is not an asomorphism, $\det(A) = 0$.

2. If we exchange two columns of A, then $\det(A)$ changes sign.

7.3.4 The First Method to Compute the Determinant

Let $(a_{i,j})$ be a matrix of dimension $n \times n$. Then $\det(A) = \Sigma_{\sigma \in S_n} \mathrm{sgn}(\sigma) \Pi_{i=1}^n a_{\sigma(i)i}$.

7.3.4.1 Proof Let $\{v_1, \ldots, v_n\}$ be columns of A, $v_i = A(e_i)$.
$\det(A)(e_1 \wedge \cdots \wedge e_n) = v_1 \wedge \cdots \wedge v_n = (\Sigma_i a_{i,1}e_i) \wedge \cdots \wedge (\Sigma_i a_{i,n}e_i) = \Sigma_{\sigma \in S_n} \Pi_i a_{\sigma i,i} e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = (\Sigma_\sigma \mathrm{sgn}(\sigma) \Pi_i a_{\sigma(i),i}) e_1 \wedge \cdots \wedge e_n$.

7.3.4.2 Corollary $\det(A) = \det(A^\tau)$.

7.3.4.3 Proof $A^\tau = (\alpha_{i,j}), A = (a_{i,j}), \forall i,j \quad a_{i,j} = \alpha_{j,i}$.
$\det(A^\tau) = \Sigma_\sigma \mathrm{sgn}(\sigma) \Pi_i \alpha_{\sigma(i),i} = \Sigma_\sigma \mathrm{sgn}(\sigma) \Pi_i a_{i,\sigma(i)} = \Sigma_{\sigma^{-1}} \mathrm{sgn}(\sigma^{-1}) \Pi_j a_{\sigma^{-1}(j)j} = \det(A)$. It's easy to prove that $\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma)$.

7.3.5 The Second Method to Compute the Determinant

Fix A of dimension $n \times n$. Apply Gauss Reduction, A' is it's upper triangular matrix. By the properties listed above, $|\det(A)| = |\det(A')|$, but on A', the determinant is just the product of elements on the diagonal. So this is the second method to compute the determinant.

Fix $A = (a_{i,j})$. Denote with $A_{[i,j]}$ the $(n-1) \times (n-1)$ matrix obtained after removing the i-th and j-th column from A.

7.3.6 The Third Method to Compute the Determinant(Laplace Expansion of the Determinant)

Let $A = (a_{i,j})$, then $\det(A) = \Sigma_{j=1}^n (-1)^{i+j} a_{i,j} \det(A_{[i,j]}) \quad \forall i \in \{1, \ldots, n\} = \Sigma_{i=1}^n (-1)^{i+j} a_{i,j} \det(A_{[i,j]}) \quad \forall j \in \{1, \ldots, n\}$.

7.3.6.1 Proof

$$\begin{array}{ccc} K^n & \xrightarrow{A} & K^n \\ {\scriptstyle \tau_j}\uparrow & & \downarrow{\scriptstyle p_i} \\ K^{n-1} & \xrightarrow{A_{[i,j]}} & K^{n-1} \end{array}$$

$\{e'_1, \ldots, e'_n\}$ is a standard basis of K^n.
$\{e_1, \ldots, e_{n-1}\}$ is a standard basis of K^{n-1}.
$\tau_j(e_i) = \begin{cases} e'_i & if\ i < j \\ e'_{i+1} & if\ i \geq j \end{cases}$. Now take $\bigwedge^{n-1}$ of the diagram:

$$\begin{array}{ccc} \bigwedge^{n-1} K^n & \xrightarrow{\bigwedge^{n-1} A} & \bigwedge^{n-1} K^n \\ {\scriptstyle \bigwedge^{n-1}\tau_j}\uparrow & & \downarrow{\scriptstyle \bigwedge^{n-1} p_i} \\ \det(K^{n-1}) & \xrightarrow{\det(A_{[i,j]})} & \det(K^{n-1}) \end{array}$$

$$\det(A)(e'_1 \wedge \cdots \wedge e'_n)$$

$$= (-1)^{i-1} \det(A)(e'_i \wedge e'_1 \wedge \cdots \wedge e'_{i-1} \wedge e'_{i+1} \wedge \cdots \wedge e'_n)$$

$$= (-1)^{i-1} A(e'_i) \wedge A(e'_1) \wedge \cdots \wedge A(e'_{i-1}) \wedge A(e'_{i+1}) \wedge \cdots \wedge A(e'_n)$$

$$(-1)^{i-1} A(e'_i) \wedge \bigwedge^{n-1} A(e'_1 \wedge \cdots \wedge e'_{i-1} \wedge e'_{i+1} \wedge \cdots \wedge e'_n) = (*)$$

.

Let $\pi_j : K^n \to K^n \quad (x_i) \mapsto (0, \ldots, x_j, \ldots, 0)$, then $A = \Sigma_j(\pi_j \circ A)$, it means that:

$$(*) = (-1)^{i-1}A(e'_i) \wedge \Sigma_j \bigwedge^{n-1}(\pi_j \circ A)(e'_1 \wedge \cdots \wedge e'_{i-1} \wedge e'_{i+1} \wedge \cdots \wedge e'_n)$$

$$= (-1)^{i-1}A(e'_i) \wedge \Sigma_j \bigwedge^{n-1}(\pi_j \circ A \circ \tau_i)(e_1 \wedge \cdots \wedge e_{n-1})$$

$$= \Sigma_{k,j}((-1)^{i-1}a_{k,i}e'_k \wedge \bigwedge^{n-1}(\pi_j \circ A \circ \tau_i)(e_1 \wedge \cdots \wedge e_{n-1}))$$

$$= \Sigma_k((-1)^{i-1}a_{k,i}e'_k \wedge \bigwedge^{n-1}(\rho_k \circ A \circ \tau_i)(e_1 \wedge \cdots \wedge e_{n-1})) = (**)$$

Here $\rho_k := \pi_k - \mathrm{id}_{K^n}$.

$$(**) = \Sigma_k(-1)^{i-1}a_{k,i}e'_k \wedge \bigwedge^{n-1}\tau_k \circ \bigwedge^{n-1}(p_k \circ A \circ \tau_i)(e_1 \wedge \cdots \wedge e_{n-1})$$

By the diagram, $\bigwedge^{n-1}(p_k \circ A \circ \tau_i) = \det A_{[k,i]}$, so:

$$(**) = \Sigma_k(-1)^{i-1}a_{k,i}\det(A_{[k,i]})(e'_k \wedge e'_1 \wedge \cdots \wedge e'_{k-1} \wedge e'_{k+1} \wedge \cdots \wedge e'_n) = \Sigma_k(-1)^{k-1}a_{k,i}\det(A_{[k,i]})(e'_1 \wedge \cdots \wedge e'_n)$$

7.4 The Structure of Linear Mappings

7.4.1 A Set of Theorems

Let $f : V \to W$ be a linear mapping between vector spaces of finite and same dim. Then:

1. There exists a decomposition $V = V_0 \oplus V_1$ and $W = W_1 \oplus W_2$ s.t. $V_0 = \ker f$ and f includes an isomorphism between V_1 and W_1(namely $f\mid_{V_1}$).
2. There exists a basis in V and W s.t. the associated matrix $A_f = a_{ij}$ satisfies $\forall 1 \leq i \leq r, \exists r \leq n$ have $a_{ii} = 1$ and have $a_{ij} = 0$ elsewhere.
3. Let A be a $m \times n$ matrix. Then there exists two square matrices (with $\det \neq 0$) B and C of dim $m \times m$ and $n \times n$ and a number $r \leq min(m, n)$ s.t. BAC has the form in (2). Moreover, the number r is unique $r = rank(A)$.

7.4.1.1 Invariant Subspace

Let $F : V \to V$ be a linear mapping. A subspace $V_0 \subseteq V$ is said to be an invariant subspace of F is $F(V_0) \subseteq V_0$.

7.4.2 Diagonalizable

A linear mapping $f : V \to V$ (finite dim) is diagonalizable if the following equivalent conditions are satisfied:

1. V decomposes as a direct sum of a one-dimensional invariant subspace of f.
2. There exists a basis of V, in which the matrix A_f is diagonal.

7.4.2.1 Proof

$2 \Rightarrow 1$ Assume that in the base $\{v_1, \cdots, v_n\}$, we have $A_f = \begin{pmatrix} \lambda & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ by the familiar diagram

$$\begin{array}{ccc} V & \xrightarrow{f} & V \\ \uparrow b & & \uparrow b \\ K^n & \xrightarrow{A_f} & K^n \end{array}$$

$$f(v_i) = b \circ A_f(e_i) = b(\lambda_i e_i) = \lambda_i v_i \in \langle v_i \rangle$$

So

$$V = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$$

$1 \Rightarrow 2$ Assume that $V = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$, where $f(\langle v_i \rangle) \subseteq \langle v_i \rangle$, then $\{v_1, \cdots, v_n\}$ forms a basis of V.

Consider the previous diagram

$$A(e_1) = b^1 \circ f \circ b(e_i) = b^{-1}(f(v_i)) = b^{-1}(\lambda_i v_i) = \lambda_i e_i$$

7.4.2.2 Example Take

$$A : \mathbb{R}^2 \to \mathbb{R}^2 \quad A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

A is not diagonalizable.

7.4.2.3 Def Let L be a one-dimensional invariant subspace of $f : V \to V$. Then $F\mid_L$ is a multiplication by a scalar $\lambda \in K$. Such λ is called *eigenvalue* of f. A non-zero vector $v \in V$ is called an *eigenvector* of V if $\langle v\rangle$ is an invariant subspace of f.

7.4.2.4 Direct Sum(Recall) $\bigoplus_{i\in\mathbb{N}} V_i = \{(x_i)_{i\in\mathbb{N}} \quad \text{all but finite many } x_i = 0\}$.
$T_0^p(V) \times T_0^q(V) \to T_0^{p+q}(V) \quad (x = x_1 \otimes \times \otimes x_p, y = y_1 \otimes \times \otimes y_p \mapsto x \otimes y.$

7.4.2.5 Remark/Exercise Assume that f is diagonalizable and A_f is the diagonal matrix that represents f. Then A_f is unique up to permutation of the elements in the diagonal.

7.4.2.6 Hint $V =< v_1 > \oplus \cdots \oplus < v_n >=< v_{\sigma(1)} > \oplus \cdots \oplus < v_{\sigma(n)} >$, $\sigma \in S_n$.
Let V be a vector space over K. $\dim(V) = n$, $f \in \mathscr{L}(V, V)$. Let A_f be an associated matrix in any basis. The map: $K \to K \quad t \mapsto \det(tI_n - A_f)$.
This is a polynomial in $K[T]$.

7.4.2.7 Lemma $P(t)$ is a monic polynomial of degree n.

7.4.2.8 Proof $P(t) = \det(tI_n - A_f) = \Sigma_\sigma \mathrm{sgn}(\sigma)\Pi_{i=1}^n(t\delta_{\sigma(i),i} - A_{\sigma(i),i})$.
The only term that gives "t^n" is when $\sigma = \mathrm{id}$.

7.4.2.9 Theorem Use the notation introduced previously.

1. $P(t)$ doesn't depend on A_f (if you change basis, $P(t)$ doesn't change).
2. Any eigenvalue of f is a root of $P(t)$. Conversely, any K-root of $P(t)$ is an eigenvalue.

7.4.2.10 Proof

1. Put $A = A_f$, and let A' be another matrix presentation of f. Then $A' = B^{-1}AB$ where B is invertible $n \times n$ matrix.

   $\det(tI_n - A') = \det(tI_n - B^{-1}AB) = \det(B^{-1}(tI_n)B - B^{-1}AB = \det(B^{-1}(tI_n - A)B) = \det(tI_n - A)$. (Because $(\lambda I)A = A(\lambda I)$).

2. Let $\lambda \in K$ be a K-root of $P(t)$, then: $\det(\lambda I_n - A_f) = 0 = P(\lambda)$. $\lambda I_n - A_f$ is not invertible, so $\exists v \neq 0 \in \ker(\lambda I_n - A_f)$, $A_f(v) = \lambda v$.

   Viceversa (the other implication) if, $\sigma \neq 0$, $f(v) = \lambda v$, $v \in \ker(\lambda I_n - A_f)$, $\det(\lambda I_n - A_f) = 0 = P(\lambda)$.

7.4.3 Characteristic Polynomial

The polynomial $P(t)$ will be denoted by $P_f(t)$. It is called the characteristic polynomial of f.

7.4.3.1 Corollary If $P_f(t)$ splits with no respected roots, then f is diagonalizable.

7.4.3.2 Def A matrix of the form $J_r(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$ is called a *Jordan block.*

A *Jordan matrix* is a matrix of the type $J = \begin{pmatrix} J_{r_1}(\lambda_1) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & J_{r_2}(\lambda_2) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & J_{r_3}(\lambda_3) \end{pmatrix}$.

7.4.3.3 Example Let $V_n(\lambda)$ be the vector space of complex functions: $\{F : \mathbb{C} \to \mathbb{C} : F(x) = e^{\lambda x} f(x)$, where $\lambda \in \mathbb{C} \quad f \in \mathbb{C}[x]_{\leq n-1}\} = V_n(\lambda)$.

Verify that $V_n(\lambda)$ is a vector space of dim n

$$\begin{aligned}\frac{d}{dx}(e^{\lambda} x f(x)) &= \lambda e^{\lambda x} f(x) + e^{\lambda x} f'(x) \\ &= e^{\lambda x}(\lambda f(x) + f'(x))\end{aligned}$$

$\frac{d}{dx} \in \mathscr{L}(V_n(\lambda); V_n(\lambda))$. Consider $v_{i+1} = \frac{x^i}{i!} e^{\lambda x}$. Show that $\{v_0, \cdots, v_{n-1}\}$ forms a basis of $V_n(\lambda)$

$$\begin{aligned}\frac{d}{dx} v_{i+1} &= \lambda v_{i+1} + \frac{x^{i-1}}{(i-1)!} e^{\lambda x} \\ &= \lambda \sigma_{i+1} + \sigma_i\end{aligned}$$

Then

$$A_{\frac{d}{dx}} = \begin{pmatrix} \lambda & 0 & \dots & 0 \\ 1 & \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 1 & \lambda \end{pmatrix} = (J_n(\lambda))^T$$

7.4.4 Annihilate

Let $a_0 + a_1 t + \cdots + a_n t^n = Q(t) \in K[T]$, then for $f \in \mathscr{L}(V, V)$, we define $Q(f) := a_0 \mathrm{Id}_v + a_1 f + \cdots + a_n f^{\circ f}$. From now on, we denote $f^k := f^{\circ k}$.

We say that Q annihilates f is $Q(f) = 0$.

7.4.4.1 Prop Let $f \in \mathscr{L}(V, V)$. There exists a polynomial $Q \in K[T] \backslash \{0\}$ that annihilates f (i.e. Q(f)=0).

7.4.4.2 Proof $\dim(\mathscr{L}(V, V)) = n^2$. Hence the maps $Id_V, f, f^2, \ldots, f^{n^2} \in \mathscr{L}(V, V)$ are linearly dependent. So there exists a non-trivial linear combination: $\lambda_0 Id_V + \lambda_1 f + \cdots + \lambda_{n^2} f^{n^2} = 0$, so take $Q(t) = \lambda_0 + \lambda_1 t + \cdots + \lambda_{n^2} t^{n^2}$.

This shows that $Q \neq 0$, and $Q(f) = 0$.

7.4.4.3 Remark The proof of the Prop also gives the degree of a polynomial that annihilates is $\leq n^2$.

7.4.4.4 Minimal Polynomial Let $m(t) \in K[T] \backslash \{0\}$ be a monic polynomial of minimal degree that annihilates $f \in \mathscr{L}(V, V)$. Then $m(t)$ is called minimal polynomial of f.

7.4.4.5 Prop If $m(t)$ is a minimal polynomial of f, then it's unique.

7.4.4.6 Proof Assume that $m_1(t)$ is another minimal polynomial, then $m - m_1 \in K[T]$; $(m - m_1)(f) = m(f) - m_1(f) = 0$. Since m and m_1 are both monic, so $\deg(m - m_1) < \deg(m)$ is smaller, contradiction!

7.4.4.7 Notation From now on we denote the minimal polynomial of f with m_f.

7.4.4.8 Prop Let $Q \in K[T] \backslash \{0\}$ a polynomial that annihilates f. Then $m_f \mid Q$.

7.4.4.9 Proof By the division between polynomials: $Q(t) = m_f(t)s(t) + r(t)$ s.t. $\deg(r) < \deg(m_f)$. So $0 = Q(f) = m_f(f)s(f) + r(f) = r(f)$. So $r(f)$ must be zero.

7.4.4.10 Adjugate Matrix Let A be a matrix of dim $n \times n$, $M_{i,j} := (-1)^{i+j} \det(A_{[i,j]}) \in K$. Here $\det(A_{[i,j]})$ is called *(i,j)-minor of A*.

Then we define $\mathrm{Adj}(A) \in M_{n \times n}(K) := (M_{i,j})^T_{i,j}$, called the adjugate matrix.

7.4.4.11 Lemma $A\mathrm{Adj}(A) = \mathrm{Adj}(A)A = \det(A) I_n$.

7.4.4.12 Proof Exercise.

7.4.5 Cayley-Hamilton Theorem

The characteristic polynomial p_f annihilates f. Consequence, $m_f \mid p_f$.

7.4.5.1 Proof Let $A = A_f$ any matrix that represents f. Consider $B := \mathrm{Adj}(tI_n - A)$, then $(tI_n - A)B = \det(tI_n - A)I_n = p_f(t)I_n$. We can decompose B in the following way: $B = \Sigma_{i=0}^{n-1} t^i B_i,\ B_i \in M_{n\times n}(K)$.

So $P_f(t)I_n = (\Sigma_{i=1}^{n-1} tI_n t^i B_i) - (\Sigma_{i=0}^{n-1} AT^i B_i) = \Sigma_{i=0}^{n-1} t^{i+1} B_i - \Sigma_{i=0}^{n-1} At^i B_i = t^n B_{n-1} + \Sigma_{i=1}^{n-1} t^i (B_{i-1} - AB_i) - AB_0$.

We can compare the coefficients, and both times A^i, the equation became $A^n B_{n-1} + \Sigma_{i=1}^{n-1}(A^i B_{i-1} - A^{i+1}B_i) - AB_0 = 0 = A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n = P_f(A)$, so $p_f(f) = 0$.

Here $p_f(f) = p_f(C^{-1}AC) = \det(tI_n - C^{-1}AC) = \det(C^{-1}(tI_n)C - C^{-1}AC) = \det(C^{-1}(tI_n - A)C) = \det(C^{-1})P_f(A)\det(C) = p_f(A)$.

7.4.5.2 Examples

1. m_f and p_f are in general different. $f = \mathrm{id}_V$, then $p_f(t) = (t-1)^n, m_f(t) = t - 1$.

2. Assume that $f : V \to V$ and dim V=r, $A_f = J_r(\lambda)$. Then $p_f(t) = (t0\lambda)^r$. Moreover, $J_r(\lambda) = \lambda I_r + J_r(0)$, and if $k \geq r$, $(J_r(0))^k = 0$. $(J_r(\lambda) - \lambda I_n)^k = (J_r(0))^k \neq 0$ if $0 \leq k \leq r-1$.

   We know that $m_f \mid (t-\lambda)^r$, m_f must be of the type $m_f = (t-\lambda)^k$, but the only possibility is $k = r$. So $m_f = p_f$.

   For Jordan blocks, the minimal polynomial is equal to the characteristic polynomial.

7.4.5.3 Algebraically Closed Field A field K is algebraically closed if any nonzero polynomial has a root in K.

7.4.6 Theorem

Let $f \in \mathscr{L}(V, V)$, where V is a vector space of dim n,over an algebraically closed field. Then:

1. f can be represented by a Jordan matrix.

2. The above matrix is unique up to permutation of Jordan blocks.

7.4.6.1 Root Vector Let $f \in \mathscr{L}(V, V)$ and let $\lambda \in K$. A vector $w \in V\backslash\{0\}$ is called a root vector of f corresponding to λ, if there exists $m \in \mathbb{N}$, $(f - \lambda\mathrm{id}_V)^m(w) = 0$.

7.4.6.2 Remark Eigenvectors are root vectors that take $r = 1$.

7.4.6.3 Remark Let $f = J_r(\lambda)$ be a Jordan block. Then any $v \in V$ is a root vector of f corresponding to λ. In fact: $\forall v,\ (J_r(\lambda) - \lambda I_r)^m(v) = 0 \quad m \geq r$.

7.4.6.4 Prop Let $V(\lambda)$ be the set of root vectors of f corresponding to λ. Then $V(\lambda)$ is a vector subspace of V. Moreover, $V(\lambda) \neq \{0\}$ if and only if λ is an eigenvalue.

7.4.6.5 Proof Take v_1, v_2 to be λ-root vectors. Assume that

$$(f - \lambda\mathrm{Id})^{r_1}(v_1) = (f - \lambda\mathrm{Id})^{r_2}(v_2) = 0.$$

Take $r = \max\{r_1, r_2\}$, then

$$(f - \lambda\mathrm{Id})^r(v_1 + v_2) = 0.$$

For any $\alpha \in K$, if $(f - \lambda\mathrm{Id})^r(v) = 0$, then

$$(f - \lambda\mathrm{Id})^r(\alpha v) = 0.$$

Assume λ is an eigenvalue. Then there exists an eigenvector $v \neq 0$ s.t.

$$(f - \lambda\mathrm{Id})(v) = 0.$$

Conversely, take $0 \neq v \in V(\lambda)$. Let r be the smallest integer s.t. $(f - \lambda\mathrm{Id})^r(v) = 0$. Since $v' = (f - \lambda\mathrm{Id})^{r-1}(v) \neq 0$ and $(f - \lambda\mathrm{Id})(v') = 0$, v' is an eigenvector of λ, which means λ is an eigenvalue.

7.4.6.6 Prop Let K be an algebraically closed field. Let $\lambda_1, \cdots, \lambda_k$ be the set of all distinct eigenvalues of f $(k \geq 1)$. Then

$$V = \bigoplus_{i=1}^{k} V(\lambda_i).$$

7.4.6.7 Proof Since K is algebraically closed, the characteristic polynomial $P_f(t)$ can be factored as

$$P_f(t) = \prod_{i=1}^{k} (t - \lambda_i)^{r_i} \in K[T].$$

Consider

$$F_i(t) = P_f(t)(t - \lambda_i)^{-r_i} \in K[T].$$

Define $f_i := F_i(f) \in \mathrm{End}(V)$ and $V_i = \mathrm{Im}(f_i)$.

We want to prove that $(f - \lambda_i \mathrm{Id})^{r_i}(V_i) = 0$, which means $V_i \subseteq V(\lambda_i)$. Note that

$$(f - \lambda_i \mathrm{Id})^{r_i} \circ f_i = P_f(f) = 0.$$

Next, we prove that $V = V_1 + \cdots + V_k$. The polynomials $F_i(t)$ are coprime. There exist polynomials $G_i(t) \in K[T]$ s.t.

$$1 = \gcd(F_1(t), \cdots, F_k(t)) = F_1(t)G_1(t) + \cdots + F_k(t)G_k(t).$$

Applying f to both sides, we get

$$\sum_{i=1}^{k} F_i(f) \circ G_i(f) = \mathrm{Id}.$$

For any $v \in V$,

$$v = \sum_{i=1}^{k} F_i(f)(G_i(v)) = \sum_{i=1}^{k} f_i(G_i(v)) \in V_1 + \cdots + V_k \subseteq V(\lambda_1) + \cdots + V(\lambda_k).$$

To show that the sum is direct, take $1 \le i \le k$ and suppose v is in the intersection $V_i \cap \left(\sum_{j \neq i} V_j\right)$. Then

$$(f - \lambda_i \mathrm{Id})^{r_i}(v) = 0 \quad \text{and} \quad F_i(f)(v) = \prod_{j \neq i} (f - \lambda_j \mathrm{Id})^{r_j}(v) = 0.$$

Since $(t - \lambda_i)^{r_i}$ and $F_i(t)$ are coprime, there exist polynomials $G_1(t)$ and $G_2(t)$ s.t.

$$G_1(t)(t - \lambda_i)^{r_i} + G_2(t)F_i(t) = 1.$$

Substituting f for t, we get $v = 0$.

Finally, we prove that $V_i = V(\lambda_i)$. Take $v \in V(\lambda_i)$ and write it as

$$v = v' + v'' \in V_i \oplus \bigoplus_{j \neq i} V_j.$$

Then $v'' = v - v' \in V(\lambda_i)$. There exists some $r \in \mathbb{N}$ s.t. $(f - \lambda_i \mathrm{Id})^r(v'') = 0$. Since $(t - \lambda_i)^r$ and $F_i(t)$ are coprime, there exist polynomials $H_1(t)$ and $H_2(t)$ s.t.

$$(t - \lambda_i)^r H_1(t) + F_i(t)H_2(t) = 1.$$

Hence, $v'' = 0$ and $v = v' \in V_i$.

7.4.7 Nilpotent

Let $f \in Z(V, V)$. Then f is said to be *nilpotent* if there exists $z \in \mathbb{N}$ s.t. $f^z = 0$.

7.4.7.1 Lemma Let f be a nilpotent map. Then $\ker(f)f = \{\text{set of eigenvectors of f}\} \cup \{0\}$.

7.4.7.2 Proof Let $v \in \ker(f)$, then v is an eigenvector with eigenvalue=0.

Let v be an eigenvector of f with eigenvalue λ. Then

$$f(v) = \lambda v.$$

Applying f repeatedly, we get

$$f^m(v) = f^{m-1}(f(v)) = f^{m-1}(\lambda v) = \lambda f^{m-1}(v) = \cdots = \lambda^m v.$$

Since f is nilpotent, there exists $m \in \mathbb{N}$ s.t. $f^m = 0$. Therefore,

$$0 = f^m(v) = \lambda^m v.$$

Since $v \neq 0$, it follows that $\lambda^m = 0$, which implies $\lambda = 0$.

7.5 Jordan Normal Form

7.5.0.1 Lemma Let f be a nilpotent mapping, then $\mathrm{Ker}(f) \neq \{0\}$.

7.5.0.2 Proof Let r be the minimal integer s.t. $f^r = 0$. Then $f^{r-1}(V) \subseteq \mathrm{Ker}(f)$, but $f^{r-1}(V) \neq \{0\}$ due to the minimality of r.

7.5.0.3 Remark Another way to prove the above lemma is to notice that $m_f = t^{r'}, 1 \leq r' \leq r$. By the Cayley-Hamilton theorem, 0 is an eigenvalue, thus $f(x) = 0$ for some $x \neq 0$.

7.5.1 Theorem (Jordan form for nilpotent mappings)

Let f be a nilpotent mapping. Then there exists a Jordan basis for f, which means that this basis gives a Jordan matrix made of blocks of the type $J_r(0)$ for f.

7.5.1.1 Proof Proof by induction on the dimension of the vector space V. If $\dim(V) = 1$, then $f = 0$ and $0 = J_1(0)$.

Assume that the induction is true for $\dim(V) < n$, then we prove it on dimension n. Let $V_0 = \mathrm{Ker}(f) = \{$the set of eigenvalues $\{0\}$. Since f is nilpotent, $\dim(V_0) \geq 1$. Therefore, $\dim(V/V_0) < n$. Define the following mapping $\overline{f} \in \mathrm{End}(V/V_0)$, $\overline{v} \mapsto \overline{f}(v)$. It is well defined and nilpotent. We use the induction hypothesis, so we have a Jordan basis for $\overline{f}$, and thus we have elements $\overline{v_0}, \cdots, \overline{v_m} \in V/V_0$.

Now lift $\overline{v_i}$ to some elements $v_i \in V$. Applying f to these elements v_i, b_i is the first integer s.t. $\overline{f^{b_i}}(\overline{v_i}) = 0$. This means that $f^{b_i}(v_i) \in V_0$, hence $f^{b_i}(v_i)$ is an eigenvector.

Consider now the vector subspace generated by $f^{b_1}(v_1), \cdots, f^{b_m}(v_m)$:

$$\langle f^{b_1}(v_1), \cdots, f^{b_m}(v_m) \rangle \subseteq V_0$$

Extract a basis and complete it to a basis of V_0. The new vectors are denoted by $u_1, \cdots, u_t$. Then we want to prove that $f^k(v_i), u_j$ forms a basis of V.

Let $v \in V$, $\pi(v) = \overline{v} = \sum_{i=1}^m \sum_{j=0}^{b_i-1} a_{ij} f^j(\overline{v_i}) = \sum_{i=1}^m \sum_{j=0}^{b_i-1} a_{ij} f^j(v_i)$.

Hence $v - \sum_{i=1}^m \sum_{j=0}^{b_i-1} a_{ij} f^j(v_i) \in V_0$, this finishes because V_0 is generated by $f^{b_1}(v_1), \cdots, f^{b_m}(v_m), u_1, \cdots, u_t$.

We show that $f^{b_i}(v_i)$ and u_j are linearly independent:

$$\sum_{i=1}^m a_i f^{b_i}(v_i) + \sum_{i=1}^t b_i u_t = 0.$$

The first observation is that $b_i = 0, \forall i$. So

$$0 = \sum_{i=1}^m a_i f^{b_i}(v_i) = f\left(\sum_{i=1}^m a_i f^{b_i-1}(v_i)\right).$$

So $\sum_{i=1}^m a_i f^{b_i-1}(v_i) \in V_0$.

This means that $\sum_{i=1}^m a_i f^{b_i-1}(\overline{v_i}) = 0$. Hence $a_i = 0, \forall i$.

By applying f to $\sum_{i=1}^m \sum_{j=0}^{b_i-1} a_{ij} f^j(v_i) + \sum_{i=1}^t b_i u_i = 0$ many times, one can prove that $\{f^k(b_i), \cdots, u_t\}$ forms a Jordan basis of V.

7.5.1.2 Prop The Jordan matrix that represents a nilpotent mapping $f \in \mathrm{End}(V)$ is unique up to permutations of the blocks.

7.5.1.3 Proof $f^{b_i}(v_i), u_j$ forms a basis of $\mathrm{Ker}(f)$, thus these elements have exactly $\dim(V_0)$ elements, which is independent of the choice of basis.

Let's look at the elements like $f^{b_i-1}(v_i)$. If we work by induction, then the proof is finished.

7.5.2 Theorem

Let K be an algebraically closed field and $f \in \mathrm{End}(V)$. Then f admits a Jordan basis. Moreover, the Jordan matrix is unique up to permutation of blocks.

7.5.2.1 Proof Since K is algebraically closed, $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_k)$, where $\lambda_1, \cdots, \lambda_k$ are the distinct eigenvalues of f.

Now consider $f|_{V(\lambda_i)} = g$, $\lambda_i = \lambda$. If we prove the theorem for g, then we are done.

$g - \lambda\mathrm{Id} \in \mathrm{End}(V(\lambda))$, this function is nilpotent on $V(\lambda)$. In fact, you choose a basis of $V(\lambda)$, then pick the largest p of them.

Apply the theorem for nilpotent mappings to $g - \lambda\mathrm{Id}$. Then we have $J_g - \lambda\mathrm{Id}$ made of blocks of the type $J_r(0)$.

Take the matrix $J_g - \lambda\mathrm{Id}$. Restrict to a block $J_r(0)$. $g - \lambda\mathrm{Id} = P^{-1}J_r(0)P$. I want to show that with the same P, $J_r(0) + \lambda I_r = J_r(\lambda)$. This is a Jordan block for g. In fact, $J_r(\lambda) = P^{-1}(J_r(0) + \lambda I_r)P = g - \lambda\mathrm{Id} + \lambda\mathrm{Id} = g$.

The uniqueness follows from the uniqueness of the $J_r(0)$.

7.5.3 Geometric Multiplicity and Algebraic Multiplicity

Let λ be an eigenvalue of $f \in \mathrm{End}(V)$.

$$E(\lambda) = \mathrm{Ker}(f - \lambda \mathrm{Id}) = \{\text{all eigenvectors of } f \text{ with } \lambda\} \cup \{0\} \subseteq V(\lambda).$$

This is called the eigenspace of λ. $\mathrm{mult}(\lambda)_{\mathrm{geo}} = \dim(E(\lambda))$ is the geometric multiplicity of λ. Moreover, $\mathrm{mult}(\lambda)_{\mathrm{alg}} = \max\{k \in \mathbb{N}, (t-\lambda)^k | P_f(t)\}$ is the algebraic multiplicity of λ.

7.5.3.1 Prop Let K be an algebraically closed field. Then $\mathrm{mult}(\lambda)_{\mathrm{geo}} \leq \mathrm{mult}(\lambda)_{\mathrm{alg}}, \forall$ eigenvalue of f.

7.5.3.2 Proof $V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_n)$. Take $\lambda = \lambda_i$. Let J_f be the Jordan matrix of f. Then $\det(J_f) = \det(f)$. So $P_f(t) = \prod (t - \lambda_i) \dim(V(\lambda_i))$. $\dim(V(\lambda_i)) = \mathrm{mult}(\lambda)_{\mathrm{alg}}$.

Since $E(\lambda) \subseteq V(\lambda)$, then $\dim(E(\lambda)) \leq \dim(V(\lambda))$.

7.5.3.3 Corollary Let K be an algebraically closed field, let $f \in \mathrm{End}(V)$. f is diagonalizable iff $\mathrm{mult}(\lambda)_{\mathrm{geo}} = \mathrm{mult}(\lambda)_{\mathrm{alg}}, \forall \lambda$.

7.5.3.4 Proof Work on a single $V(\lambda)$. $r = \dim(V(\lambda)) = \mathrm{mult}(\lambda)_{\mathrm{alg}}$.

We get a diagonal Jordan matrix iff we have exactly r blocks of length 1. It means that $r = \dim(\mathrm{Ker}(f - \lambda \mathrm{Id})) = \mathrm{mult}(\lambda)_{\mathrm{geo}}$.

Chapter 8

Inner Product Space

8.1 Inner Product

8.1.1 Bilinear Form

Let V be an n-dimensional vector space over K (where $K = \mathbb{R}$ or $K = \mathbb{C}$).

A bilinear form $g \in \mathscr{L}(V \times V, K)$ is called a bilinear form. Choose a basis $\{v_1, \cdots, v_n\}$ of V. The matrix $G = (g(v_i, v_j))_{ij} \in M_{n\times n}(K)$ is called the Gram matrix of g with respect to $\{v_1, \cdots, v_n\}$. By linearity, G determines g uniquely.

8.1.1.1 Remark For all $(x, y) \in V$,

$$g(x, y) = \sum_{ij} x_i y_j g(v_i, v_j) = \underline{x}^\tau G \underline{y}.$$

On the other hand, given a basis $\{v_1, \cdots, v_n\}$ and $G \in M_{n\times n}(K)$, the mapping $V \times V \to K$, $(x, y) \mapsto \underline{x}^\tau G \underline{y}$, is a bilinear form, and the associated Gram matrix is exactly G.

Fixed a pair $(V, \{v_1, \cdots, v_n\})$, we have defined a bijection:

$$\mathscr{L}(V \times V, K) \cong M_{n\times n}(K), \quad g \mapsto G.$$

8.1.2 Conjugate

Two matrices G, $G' \in M_{n\times n}(K)$ are said to be conjugate if there exists $A \in \mathrm{GL}(n; K)$ s.t. $G' = A^\tau G A$. Notice that conjugacy is an equivalence relation.

8.1.2.1 Notation The Gram matrices of an inner product under different bases are conjugate.

8.1.2.2 Remark

$$\mathscr{L}(V \times V, K) \cong \mathscr{L}(T_0^2(V), K) \cong \mathscr{L}(V, V^\vee), \quad g \mapsto g_s := (g_s(x \otimes y) := g(x, y)) \mapsto (x \mapsto g_s(x \otimes \cdot)) =: \widetilde{g}$$

8.1.2.3 Def Given $g \in \mathscr{L}(V \times V, K)$, define $g_p(x, y) := g(y, x)$, $\overline{g_p}(x, y) := \overline{g_p(x, y)} = \overline{g(y, x)}$. Notice that if $K = \mathbb{R}$, then $g_p = \overline{g_p}$.

A bilinear form g is said to be:

- Symmetric if $g = g_p$.
- Symplectic or skew-symmetric if $g = -g_p$.
- Hermitian if $g = \overline{g_p}$.

8.1.2.4 Example

- $K^n \times K^n \to K$, $(x, y) \mapsto x^\tau y$ is symmetric.
- $\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, $(x, y) \mapsto x^\tau \overline{y}$ is hermitian.
- $K^2 \times K^2 \to K$, $(v_1, v_2) \mapsto \det(v_1|v_2)$ is skew-symmetric.

8.1.2.5 Def $g \in \mathscr{L}(V \times V, K)$ is an inner product of V if g is either symmetric, symplectic, or Hermitian.

8.1.3 Orthogonal

Let (V, g) be an inner product space. Two vectors $v_1, v_2 \in V$ are orthogonal (with respect to g) if $g(v_1, v_2) = 0$. Two vector subspaces $V_1, V_2 \subseteq V$ are orthogonal if $g(v_1, v_2) = 0$ for all $(v_1, v_2) \in V_1 \times V_2$. Notice that:

- If g is symmetric, then $A = A^\tau$.
- If g is symplectic, then $A = -A^\tau$.
- If g is Hermitian, then $A = \overline{A^\tau}$.

Here A is the matrix presenting g.

8.1.3.1 Def Let (V, g) be an inner product space. The kernel of g is defined as: $\mathrm{Ker}(g) = \{v \in V \mid g(v, w) = 0, \forall w \in V\}$. Moreover, g is said to be non-degenerate if $\mathrm{Ker}(g) = \{0\}$.

8.1.3.2 Remark Note that $\mathrm{Ker}(g) = \mathrm{Ker}(\widetilde{g})$ when $\widetilde{g} \in \mathscr{L}(V, V^\vee)$.

$x \in \mathrm{Ker}(\widetilde{g})$ means $\widetilde{g_x} = 0$, which is equivalent to $g(x, y) = 0$ for all $y \in V$. This implies that $\mathrm{Ker}(g)$ is a linear subspace of V.

Chapter 9

Differential Forms in $\mathbb{R}^n$

9.1 Differential Forms

9.1.1 Def

Let $p \in \mathbb{R}^n$ be a fixed point. $\mathbb{R}^n_p = \{p\} \times \mathbb{R}^n$, $(p,a) \in \mathbb{R}^n_p$, $a \in \mathbb{R}^n$. Define the operations: $(p,a)+(p,b) := (p, a+b)$, $\alpha(p,a) := (p, \alpha a)$, $\forall \alpha \in \mathbb{R}$.

With these operations, $\mathbb{R}^n_p$ is a vector space over $\mathbb{R}$. $a|_p$ denotes (p,a), and a basis of $\mathbb{R}^n_p$ is denoted by $\{e_1|_p, \cdots, e_n|_p\}$.

$\mathbb{R}^n_p$ is called the *tangent space* of $\mathbb{R}^n$ at p. The dual space $(\mathbb{R}^n_p)^\vee \cong \{p\} \times (\mathbb{R}^n)^\vee$. The dual basis is denoted by $\{dx_1|_p, \cdots, dx_n|_p\} := \{(e_1|_p)^\vee, \cdots, (e_n|_p)^\vee\}$.

$\mathbb{R}^n \times \mathbb{R}^n \cong \bigsqcup_{p \in \mathbb{R}^n} \mathbb{R}^n_p$. $\bigsqcup_p \mathbb{R}^n_p$ is called the *tangent bundle* of $\mathbb{R}^n$. We have a projection mapping: $\pi : \bigsqcup_{p\in\mathbb{R}^n} \mathbb{R}^n_p \to \mathbb{R}^n$, $a|_p \mapsto p$, and $\pi^{-1}(p) = \mathbb{R}^n_p$.

9.1.1.1 Remark

$$\frac{\partial x_i}{\partial x_j} := dx_i|_p(e_j|_p) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

9.1.1.2 Def

$$(\mathbb{R}^n_p)^\vee := \bigoplus_{k\in\mathbb{N}} \bigwedge^k (\mathbb{R}^n_p)^\vee.$$

Consider $\bigwedge^k (\mathbb{R}^n_p)^\vee \cong (\bigwedge^k \mathbb{R}^n_p)^\vee$. The basis of this vector space is:

$$\{dx_{i_1}|_p \wedge \cdots \wedge dx_{i_k}|_p \mid 1 \le i_1 < \cdots < i_k \le n\}.$$

$$\dim \left(\bigwedge^k (\mathbb{R}^n_p)^\vee \right) = \binom{n}{k}.$$

$$\bigsqcup_p \left(\bigwedge^k (\mathbb{R}^n_p)^\vee \right) \to \mathbb{R}^n, \quad f|_p \mapsto p$$

is induced by the natural projection map.

9.1.2 Exterior k-form

An exterior k-form in $\mathbb{R}^n$ is a mapping:

$$\omega : \mathbb{R}^n \to \bigsqcup_p \left(\bigwedge^k (\mathbb{R}^n_p)^\vee \right), \quad p \mapsto \omega(p),$$

that is a section of the projection π. $\pi \circ \omega = \mathrm{Id}_{\mathbb{R}^n}$ means $\omega(p) \in \bigwedge^k (\mathbb{R}^n_p)^\vee$. For all $p \in \mathbb{R}^n$,

$$\omega(p) = \sum_{1 \le i_1 < \cdots < i_k \le n} a_{i_1, \cdots, i_k}(p) dx_{i_1}|_p \wedge \cdots \wedge dx_{i_k}|_p.$$

Fix ω, if all $a_{i_1,\cdots,i_k}$ are differentiable of class C^m, then ω is called a C^m-differential k-form. If $m = +\infty$, then ω is a smooth k-form.

9.1.2.1 Notation

$\omega = \sum_I a_I dx_I$, $I = (i_1, \cdots, i_k)$.

9.1.2.2 Example Take $n = 4$.

- 1-form: $\omega = a_1 dx_1 + a_2 dx_2 + a_3 dx_3 + a_4 dx_4$.

$$\omega(p) = a_1(p)dx_1|_p + a_2(p)dx_2|_p + a_3(p)dx_3|_p + a_4(p)dx_4|_p \in \bigwedge^1(\mathbb{R}^4_p)^\vee.$$

- 2-form:

$$\omega = a_{12}dx_1 \wedge dx_2 + a_{13}dx_1 \wedge dx_3 + a_{14}dx_1 \wedge dx_4 + a_{23}dx_2 \wedge dx_3 + a_{24}dx_2 \wedge dx_4 + a_{34}dx_3 \wedge dx_4.$$

- 4-form:

$$\omega(p) = a_{1234}(p)dx_1|_p \wedge dx_2|_p \wedge dx_3|_p \wedge dx_4|_p \in \bigwedge^4(\mathbb{R}^4_p)^\vee.$$

9.1.2.3 Remark When $k = 0$, a C^m-differential 0-form is a mapping $f : \mathbb{R}^n \to \mathbb{R}$ of class C^m.

9.1.2.4 Def $\Omega^k_{(m)}(\mathbb{R}^n) := \{\text{set of } C^m\text{-differential } k\text{-forms}\}$.

9.1.2.5 Prop $\Omega^k_{(m)}(\mathbb{R}^n)$ is a module over $\Omega^0_{(m)}(\mathbb{R}^n)$.

9.1.2.6 Proof $\omega, \eta \in \Omega^k(\mathbb{R}^n)$, $(\omega + \eta)(p) := \omega(p) + \eta(p)$. $f \in \Omega^0(\mathbb{R}^n)$, $\omega \in \Omega^k(\mathbb{R}^n)$, $f\omega \in \Omega^k(\mathbb{R}^n)$, $(f\omega)(p) := f(p)\omega(p)$.

9.1.2.7 Notation We use $df|_p$ to denote the differential of f at p.

9.1.2.8 Remark Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable mapping, then its differential $df|_p : \mathbb{R}^n_p \to \mathbb{R}_{f(p)} \cong \mathbb{R}$, thus $df|_p \in (\mathbb{R}^n_p)^\vee$. Since $df|_p \in (\mathbb{R}^n_p)^\vee$, hence $df|_p = \sum f_i(p)dx_i|_p$. By definition, f_i are the partial derivatives of f. This means df is a differential 1-form.

Moreover, $F : \mathbb{R}^n \to \mathbb{R}^m$ differentiable then $F = (F_1, \cdots, F_m)$, where $F_i : \mathbb{R}^n \to \mathbb{R}$ is differentiable $dF|_p : \mathbb{R}^n_p \to \mathbb{R}^m_{F(p)} \cdot dF_i|_p = dx_i|_{F(p)}(dF|_p) = d(x_i \circ F)|_p$.

9.1.2.9 Notation Using the definition, we can derive:

$$\frac{\partial f}{\partial x_i}(p) = df|_p(e_i), \quad \text{thus} \quad df|_p(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(p)h_i.$$

Therefore, f_i is the i-th partial derivative of f at point p.

9.1.2.10 Def $\Omega_{(m)}(\mathbb{R}^n) := \bigoplus_k \Omega^k_{(m)}(\mathbb{R}^n)$.

We define a wedge product on $\Omega(\mathbb{R}^n)$:

$$\Omega^k(\mathbb{R}^n) \times \Omega^l(\mathbb{R}^n) \to \Omega^{k+l}(\mathbb{R}^n), \quad (\omega, \eta) \mapsto \omega \wedge \eta.$$

Take $\omega = \sum_I a_I dx_I$, $\eta = \sum_J b_J x_J$. Then:

$$\omega \wedge \eta := \sum_{IJ} a_I b_J dx_{IJ}, \quad I = (i_1, \cdots, i_k), \quad J = (j_1, \cdots, j_l), \quad IJ := (i_1, \cdots, i_k, j_1, \cdots, j_l).$$

9.1.2.11 Exercise Take $\omega \in \Omega^k(\mathbb{R}^n)$, $\eta \in \Omega^l(\mathbb{R}^n)$, $\varphi \in \Omega^s(\mathbb{R}^n)$. Then:

1. $(\omega \wedge \eta) \wedge \varphi = \omega \wedge (\eta \wedge \varphi)$
2. $(\omega \wedge \eta) = (-1)^{kl}(\eta \wedge \omega)$
3. Take $\theta \in \Omega^s(\mathbb{R}^n)$, $\omega \wedge (\varphi + \theta) = \omega \wedge \varphi + \omega \wedge \theta$

9.2 Pullback of Forms

9.2.1 Def

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a mapping of class C^{r+1}. It induces a mapping:

$$f^* : \Omega^k_{(r)}(\mathbb{R}^m) \to \Omega^k_{(r)}(\mathbb{R}^n), \quad \omega \mapsto f^*\omega.$$

$$\Lambda^k(V^\vee) \cong (\Lambda^k V)^\vee \cong \{\psi : V \times \cdots \times V \to K \mid \psi \text{ is multilinear and alternating}\},$$

hence by abuse of notation:

$$(f^*\omega)(p)(v_1, \cdots, v_k) := \omega(f(p))(\mathrm{d}f|_p(v_1), \cdots, \mathrm{d}f|_p(v_k)) := \omega(f(p))(\mathrm{d}f|_p(v_1) \wedge \cdots \wedge \mathrm{d}f|_p(v_k)).$$

9.2.1.1 Prop
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable mapping. For $\omega, \eta \in \Omega^k(\mathbb{R}^m)$ and $g : \mathbb{R}^m \to \mathbb{R}$ a differentiable mapping ($g \in \Omega^0(\mathbb{R}^m)$), the following hold:

1. $f^*(\omega + \eta) = f^*\omega + f^*\eta$
2. $f^*(g\omega) = f^*g f^*\omega$ where $f^*g := g \circ f$
3. If $\omega_1, \cdots, \omega_k$ are 1-forms in $\mathbb{R}^m$, then $f^*(\omega_1 \wedge \cdots \wedge \omega_k) = f^*\omega_1 \wedge \cdots \wedge f^*\omega_k$

9.2.1.2 Proof

1. $f^*(\omega + \eta)(p)(v_1, \cdots, v_k) = (\omega + \eta)(f(p))(df|_p(v_1), \cdots, df|_p(v_k)) =$
 $\omega(f(p))(df|_p(v_1), \cdots, df|_p(v_k)) + \eta(f(p))(df|_p(v_1), \cdots, df|_p(v_k)) = (f^*\omega)(p)(v_1, \cdots, v_k) + (f^*\eta)(p)(v_1, \cdots, v_k)$ □
2. $f^*(g\omega)(p)(v_1, \cdots, v_k) = g\omega(f(p))(df|_p(v_1), \cdots, df|_p(v_n)) = (g \circ f)(p)(f^*\omega)(p)(v_1, \cdots, v_k)$ □
3. $f^*(\omega_1 \wedge \cdots \wedge \omega_k)(p)(v_1, \cdots, v_k) = (\omega_1 \wedge \cdots \wedge \omega_k)(f(p))(df|_p(v_1), \cdots, df|_p(v_k)) = \det(\omega_i(df|_p(v_j)))_{ij} = \det(f^*\omega_i(p)(v_j))_{ij} =$
 $(f^*\omega_1 \wedge \cdots \wedge f^*\omega_k)(p)(v_1, \cdots, v_k)$ □

9.2.1.3 Remark
Let $f : \mathbb{R}^n \to \mathbb{R}^m$, and let $\{y_1, \cdots, y_m\}$ be a basis of $\mathbb{R}^m$. Then:

$$(x_1, \cdots, x_n)^T \mapsto (f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n))^T.$$

Let $\omega = \sum_I a_I dy_I \in \Omega(\mathbb{R}^m)$. Then:

$$f^*\omega = \sum_I f^*(a_I)(f^*dy_{i1}) \wedge \cdots \wedge (f^*dy_{ik}).$$

Note that:

$$(f^*dy_i)(v) = dy_i(df(v)) = d(y_i \circ f)(v) = (df_i)(v).$$

Thus:

$$f^*\omega = \sum_I a_I(f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n))df_{i1} \wedge \cdots \wedge df_{ik}.$$

9.2.1.4 Remark
Let U be an open set of $\mathbb{R}^n$. Then consider $\Omega^k(U) \subseteq \Omega^k(\mathbb{R}^n)$.

9.2.1.5 Example

$$\omega = -\frac{y}{x^2 + y^2}dx + \frac{x}{x^2 + y^2}dy \in \Omega^1(\mathbb{R}^2 \setminus \{(0,0)\}).$$

$$V = \{(r, \theta) \in \mathbb{R}^2 \mid r > 0, 0 \le \theta < 2\pi\}, \quad f : V \to U, \quad (r, \theta)^T \mapsto (r\cos\theta, r\sin\theta)^T.$$

Let's compute $f^*\omega : \bigsqcup_{r>0, 0\le\theta<2\pi} \mathbb{R}^2_{(r,\theta)} \to V$. Then:

$$df_1 = \cos\theta dr - r\sin\theta d\theta, \quad df_2 = \sin\theta dr + r\cos\theta d\theta.$$

$$f^*\omega = -\frac{r\sin\theta}{r^2}(\cos\theta dr - r\sin\theta d\theta) + \frac{r\cos\theta}{r^2}(\sin\theta dr + r\cos\theta d\theta) = d\theta.$$

9.2.1.6 Prop
Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable mapping. Then:

1. $f^*(\omega \wedge \eta) = (f^*\omega) \wedge (f^*\eta)$ for any two forms on $\mathbb{R}^m$.
2. $(f \circ g)^*\omega = g^*(f^*\omega)$, where $g : \mathbb{R}^p \to \mathbb{R}^n$ is differentiable.

9.2.1.7 Proof

1. Let $(y_1, \cdots, y_m) = (f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_n)) \in \mathbb{R}^m$, $(x_1, \cdots, x_n) \in \mathbb{R}^n$. Let $\omega = \sum_I a_I dy_I$, $\eta = \sum_J b_J dy_J$. Then:

$$f^*(\omega \wedge \eta) = f^*\left(\sum_I a_I b_J dy_I \wedge dy_J\right) = \sum_{IJ} a_I(f_1, \cdots, f_m)b_J(f_1, \cdots, f_m)df_I \wedge df_J =$$

$$\left(\sum_I a_I(f_1, \cdots, f_m)df_I\right) \wedge \left(\sum_J b_J(f_1, \cdots, f_m)df_J\right) = (f^*\omega) \wedge (f^*\eta).$$

2. $[f \circ g]^*(\omega) = \sum_I a_I((f \circ g)_1, \cdots, (f \circ g)_m)d(f \circ g)_I = \sum_I a_I(f_1(g_1, \cdots, g_n), \cdots, f_m(g_1, \cdots, g_n))df_I(dg_1, \cdots, dg_n) =$
 $g^*(f^*\omega)$.

9.2.2 Exterior Differential

The differential of a mapping is a 1-form, $f \to df$, from 0-form to 1-form. We want to generalize this process for any k-form.

$$d : \Omega^k_{(m)}(U) \to \Omega^{k+1}_{(m-1)}(U), \quad \omega \mapsto d\omega, \quad \omega = \sum_I a_I dx_I, \quad d\omega := \sum_I da_I \wedge dx_I,$$

$$da_I = \sum_I \frac{\partial a_{I,i}}{\partial x_i} dx_i.$$

9.2.2.1 Example

$$\omega = xyzdx + yzdy + (x+z)dz.$$

$$d\omega = d(xyz) \wedge dx + d(yz) \wedge dy + d(x+z) \wedge dz = -xzdx \wedge dy + (1-xy)dx \wedge dz - ydy \wedge dz.$$

9.2.2.2 Prop

1. $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2, \quad \forall \omega_1, \omega_2 \in \Omega^k(U), \quad \forall \omega_1, \omega_2 \in \Omega(U)$
2. $d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^k \omega \wedge d\eta, \quad \omega \in \Omega^k(U), \quad \eta \in \Omega^l(U)$
3. $d(d\omega) = 0 \quad (d^2\omega = 0) \quad \omega \in \Omega^k(U)$
4. $d(f^*\omega) = f^*(d\omega), \quad \omega \in \Omega^k(U), \quad f : U \to V$ is a differentiable function.

(1) Obvious. □

(2) Let $\omega = \sum_I a_I dx_I$ and $\eta = \sum_J b_J dx_J$ then:

$$\begin{aligned} d(\omega \wedge \eta) &= \sum_{IJ} d(a_I b_J) \wedge dx_I \wedge dx_J \\ &= \sum_{IJ} b_J da_I \wedge dx_I \wedge dx_J + \sum_{IJ} a_I db_J \wedge dx_I \wedge dx_J \\ &= d\omega \wedge \eta + (-1)^k \sum_{IJ} a_I dx_I \wedge db_J \wedge dx_J \\ &= d\omega \wedge \eta + (-1)^k \omega \wedge d\eta \quad \square \end{aligned}$$

(3) First assume $\omega = f \in \Omega^0(U)$:

$$\begin{aligned} d(df) &= d\left(\sum_{j=1}^n \frac{\partial f}{\partial x_j} dx_j\right) = \sum_{j=1}^n d\left(\frac{\partial f}{\partial x_j}\right) \wedge dx_j \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} dx_i \wedge dx_j\right) \end{aligned}$$

Since $dx_i \wedge dx_i = 0$ and $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$, we have $d(df) = 0$.

For general $\omega = a_I dx_I$, note that $d^2 x_I = d(dx_I) = d(1) \wedge dx_I = 0$. Hence:

$$d^2\omega = d(d\omega) = d(da_I \wedge dx_I) = d^2 a_I \wedge dx_I - da_I \wedge d^2 x_I = 0 \quad \square$$

(4) For $\omega = g \in \Omega^0(U)$, $g \circ f : \mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}$:

$$(x_1, \cdots, x_n) \mapsto (f_1(x), \cdots, f_m(x)) \mapsto g(f_1(x), \cdots, f_m(x))$$

□

9.2.2.3 Remark

This forms a complex chain:

$$0 \to \Omega^0(U) \to \Omega^1(U) \to \cdots \to 0$$

9.2.2.4 Chain Rule Computation

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m \xrightarrow{g} \mathbb{R}$$

Previously proved that $d(g \circ f)|_p = dg|_{f(p)} \circ df_p$, thus:

$$\begin{aligned} d(g \circ f)|_p(\cdot) &= dg|_{f(p)}(df|_p(\cdot)) = \nabla g(f(p)) df|_p(\cdot) \\ &= \nabla g(f(p)) J_f(p)(\cdot) \\ &= \left(\frac{\partial g}{\partial y_1}(f(p)) \cdots \frac{\partial g}{\partial y_m}(f(p)) \right) \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix} (\cdot) \\ &= \sum_{ij} \frac{\partial g}{\partial y_i}(f(p)) \frac{\partial f_i}{\partial x_j}(p)(\cdot) \end{aligned}$$

Moreover:

$$d(g \circ f) = \sum_j \frac{\partial (g \circ f)}{\partial x_j} dx_j \Rightarrow \frac{\partial (g \circ f)}{\partial x_j} = \sum_i \frac{\partial g}{\partial y_i}(f(\cdot)) \frac{\partial f_i}{\partial x_j}$$

9.3 Line Integrals

Fix $\omega = \sum_i a_i dx_i \in \Omega^1_{(m)}(\mathbb{R}^n)$, $m \geq 1$.

9.3.1 Def

Let $U \subseteq \mathbb{R}^n$ be an open set, $\gamma : [a, b] \to \mathbb{R}^n$. $\exists t_0 = a < t_1 < \cdots < t_k < t_{k+1} = b$ such that $\gamma|_{[t_j, t_{j+1}]} =: \gamma_j$ is of class C^1. Then we define γ as a parametric curve and $t \in [a, b]$ as parameters.

9.3.1.1 Example $f : \mathbb{R}^n \to \mathbb{R}$ of class C^1, then $\gamma : t \mapsto (t, f(t))$ is a parametric curve.

$\gamma : t \mapsto (\cos t, \sin t)$ is a parametric curve.

For $\gamma_j :]t_k, t_{k+1}[\to \mathbb{R}^n$, we can define $\gamma_j^* \omega$. This is a one form in $\Omega^1([t_k, t_{k+1}])$. If $\gamma_j(t) = (x_1(t), \cdots, x_n(t))$ then

$$\gamma_j^* \omega = \sum_i a_i(x_1(t), \cdots, x_n(t)) dx_i = \sum_i a_i(x_1(t), \cdots, x_n(t)) \frac{dx_i}{dt} dt$$

9.3.2 Line Integral

Let ω and γ be as above. Define:

$$\int_\gamma \omega := \sum_j \int_{t_j}^{t_{j+1}} \gamma_j^* \omega$$

Remark

Fix $\gamma(t)$, $\gamma'(t) = \left(\frac{dx_1}{dt}(t), \cdots, \frac{dx_n}{dt}(t)\right) = \{$the tangent vector of γ in $\gamma(t)\}$.

$$\int_{t_j}^{t_{j+1}} \gamma_j^* \omega = \int_{t_j}^{t_{j+1}} \langle a \circ \gamma, \gamma_j' \rangle dt$$

since $a \circ \gamma = (a_1(x_1(t), \cdots, x_n(t)), \cdots, a_n(x_1(t), \cdots, x_n(t)))$ and $\gamma_j' = \left(\frac{dx_1}{dt}, \cdots, \frac{dx_n}{dt}\right)$.

9.3.2.1 Example Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ be a field. γ is the path of a particle under the action of F.

Fix t, $\Delta\gamma = \gamma(t + h) - \gamma(t)$, $\Delta t = h$. $\Delta\gamma \sim \gamma'(t)\Delta t$ and $\lim_{\Delta t \to 0} \frac{\Delta\gamma}{\Delta t} = \gamma'(t)$.

$$\langle F(\gamma(t)), \Delta\gamma \rangle \sim \langle F(\gamma(t)), \gamma'(t) \rangle \Delta t$$

Index