Lecture Notes

Abyss

June 8, 2025

庾信平生最萧瑟，暮年诗赋动江关。

Contents

本笔记中“Morphism” 指的均是“Homomorphism”

Chapter 1

Mathematics Advanced Placement

本章内容大量参考了林嘉翊的笔记。因为一些历史原因, 本章的命题大多缺少证明。

1.1 Sets and correspondence

1.1.1 Construction of sets

1.1.1.1 Def

A collection of distinct elements is called a set.

Let A and B be two sets. If $\forall x \in A$, $x \in B$, we say that A is a subset of B, denoted it as $A \subseteq B$. If $A \subseteq B$ and $A \neq B$, then we say A is the proper subset of B, denoted it as $A \subsetneq B$.

$\varnothing$ is a set that doesn't have any element.

Notation

"$\forall/\exists x \in A$, $\mathbb{P}(x)$" is the notation of proposition.

Remark

$A = B$ iff $A \subseteq B$ and $B \subseteq A$.

"$\forall x \in A, \mathbb{P}(x)$" have the opposite meaning of "$\exists x \in A, \text{not } \mathbb{P}(x)$". Similarly, "$\forall x \in A, \text{not } \mathbb{P}(x)$" is opposite to "$\exists x \in A, \mathbb{P}(x)$".

"$\forall x \in \varnothing, \mathbb{P}(x)$" is always true. "$\exists x \in \varnothing$, $\mathbb{P}(x)$" is always false.

1.1.1.2 Def

Let I be a set, $(A_i)_{i\in I}$ is called a family of sets.

$\bigcup\limits_{i\in I} A_i$ is called the union of $(A_i)_{i\in I}$. $x \in \bigcup\limits_{i\in I} A_i$ iff $\exists i \in I$, s.t. $x \in A_i$.

If $I \neq \varnothing$, $\bigcap\limits_{i\in I} A_i$ is called the intersection of $(A_i)_{i\in I}$. $x \in \bigcap\limits_{i\in I} A_i$ iff $\forall i \in I$, $x \in A_i$.

Let A and B be two sets. $\mathscr{P}(A)$ is called power set, which is the set of all subsets of A. $A \times B$ is called the Cartesian product of A and B, which is also a set. $A \times B = \{(x, y)|x \in A, y \in B\}$.

$\prod\limits_{i\in I} A_i = \{(x_1, x_2, \cdots, x_n)|x_1 \in A_1, \cdots, x_n \in A_n\}$.

If all $A_i = A$, one can use A^I to denote $\prod\limits_{i\in I} A_i =$ and if $I = \{1, \cdots, n\}$. We use A^n to denote the set $A^{\{1,\cdots,n\}}$.

1.1.2 Correspondences

1.1.2.1 Def

$f = (\mathrm{D}_f, \mathrm{A}_f, \Gamma_f)$. D_f is departure set, A_f is arrival set and Γ_f is a subset of $\mathrm{D}_f \times \mathrm{A}_f$. Then we say f is a correspondence from D_f to A_f and Γ_f is the graph of f.

$f^{-1} = (\mathrm{A}_f, \mathrm{D}_f, \Gamma_{f^{-1}})$. f^{-1} is the inverse correspondence of f. $\Gamma_{f^{-1}} = \{(y,x)|(x,y) \in \Gamma_f\}$.

$\mathrm{Id}_X = (X, X, \triangle_X)$ is called the identity correpondence. $\triangle_X \subseteq \{(x,x)|x \in X\}$ is called the diagonal subset of X.

$(X, Y, \varnothing)$ is called the empty correpondence.

1.1.3 Image and inverse image

In this subsection, we fix a correspondence f from $\mathrm{D} = X$ to $\mathrm{A} = Y$.

1.1.3.1 Def

Let A be a subset of D_f. $f(A) = \{y \in \mathrm{A}_f | \exists x \in A, (x,y) \in \Gamma_f\}$ is called the image of A by f.

Let B be a subset of A_f. $f^{-1}(B) = \{x \in \mathrm{D}_f | \exists y \in B, (y,x) \in \Gamma_{f^{-1}}\}$ is called the inverse image of B by f.

1.1.3.2 Prop

Let A and A' be two sets, s.t. $A \subseteq A' \subseteq \mathrm{D}_f$, then $f(A) \subseteq f(A')$.

Proof

$\forall y \in f(A)$, $\exists x \in A \subseteq \mathrm{D}_f$, s.t. $(x,y) \in \Gamma_f$, since $A \subseteq A'$, $y \in f(A')$. □

Remark

Let B and B' be two sets, s.t. $B \subseteq B' \subseteq \mathrm{A}_f$, then $f^{-1}(B) \subseteq f^{-1}(B')$.

1.1.3.3 Def

$\mathrm{Dom}(f) = f^{-1}(\mathrm{A}_f)$ is called the domain of definition of f.

$\mathrm{Im}(f) = f(\mathrm{D}_f)$ is called the range of f.

1.1.3.4 Prop

$\mathrm{Im}(f) = f(\mathrm{Dom}(f))$, $\mathrm{Dom}(f) = f^{-1}(\mathrm{Im}(f))$.

Proof

Just need to prove $\mathrm{Im}(f) = f(\mathrm{Dom}(f))$.

$f(\mathrm{Dom}(f)) \subseteq \mathrm{Im}(f)$ is trivial.

$\forall y \in \mathrm{Im}(f)$, $\exists x \in \mathrm{D}_f$, s.t. $(x, y) \in \Gamma_f$, which means $x \in f^{-1}(\mathrm{A}_f) = \mathrm{Dom}(f)$, thus $y \in f(\mathrm{Dom}(f))$. $\square$

1.1.3.5 Prop

Let A be a set, an element $y \in Y$ belongs to $f(A)$ iff $f^{-1}(y) \bigcap A \neq \varnothing$.

Proof

Suppose $\exists x \in (f^{-1}(\{y\}) \bigcap A)$, then $(x, y) \in \Gamma_f$ and $x \in A$, which means $y \in f(A)$.

Suppose $y \in f(A)$, thus $\exists x \in A$, s.t. $(x, y) \in \Gamma_f$, which means $x \in f^{-1}(\{y\}) \bigcap A \neq \varnothing$. $\square$

Remark

Let B be a set, an element $x \in X$ belongs to $f^{-1}(B)$ iff $f(\{x\}) \bigcap B \neq \varnothing$.

1.1.3.6 Prop

Let I be a set, $(A_i)_{i \in I}$ be a family of sets.

(1) $f(\bigcup\limits_{i \in I} A_i) = \bigcup\limits_{i \in I} f(A_i)$

(2) $f(\bigcap\limits_{i \in I} A_i) \subseteq \bigcap\limits_{i \in I} f(A_i)$

Proof

(1) $f(\bigcup\limits_{i \in I} A_i) = \{y \in Y | f^{-1}(\{y\}) \bigcap (\bigcup\limits_{i \in I} A_i) \neq \varnothing\} = \{y \in Y | \exists i \in I, f^{-1}(\{y\}) \bigcap A_i \neq \varnothing\} = \{y \in Y | \exists i \in I, y \in f(A_i)\} = \bigcup\limits_{i \in I} f(A_i)$. $\square$

(2) **Own:** $\forall y \in f(\bigcap\limits_{i \in I} A_i)$, $\forall i \in I$, $\exists x_i \in A_i$, s.t. $(x_i, y) \in \Gamma_f$, thus $\forall i \in I$, $y \in f(A_i)$, hence $y \in \bigcap\limits_{i \in I} f(A_i)$. $\square$

CHEN Huayi: $f(\bigcap\limits_{i \in I} A_i) = \{y \in Y | f^{-1}(\{y\}) \bigcap (\bigcap\limits_{i \in I} A_i) \neq \varnothing\} \subseteq \{y \in Y | \forall i \in I, f^{-1}(\{y\}) \bigcap A_i \neq \varnothing\} = \{y \in Y | \forall i \in I, y \in f(A_i)\} = \bigcap\limits_{i \in I} f(A_i)$. $\square$

1.1.4 Composition

In this subsection, we fix two correspondences f and g from X to Y and from Y to Z, respectively.

1.1.4.1 Def

We define $g \circ f = (\mathrm{D}_{g\circ f}, \mathrm{A}_{g\circ f}, \Gamma_{g\circ f})$, $\Gamma_{g\circ f} = \{(x,z) | \exists y \in Y, \text{ s.t. } (x,y) \in \Gamma_f \text{ and } (y,z) \in \Gamma_g\}$ as the composition of f and g.

Remark

$\mathrm{D}_{g\circ f} \subseteq \mathrm{D}_f$, $\mathrm{A}_{g\circ f} \subseteq \mathrm{A}_g$.

1.1.4.2 Prop

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

Proof

$(z,x) \in \Gamma_{(g\circ f)^{-1}}$ is equivalent to $(x,z) \in \Gamma_{g\circ f}$ is equivalent to $\exists y \in Y$, $(x,y) \in \Gamma_f$, $(y,z) \in \Gamma_g$ is equivalent to $\exists y \in Y$, $(z,y) \in \Gamma_{g^{-1}}$, $(y,x) \in \Gamma_{f^{-1}}$ is equivalent to $(z,x) \in \Gamma_{f^{-1}\circ g^{-1}}$. $\square$

1.1.4.3 Prop

Let h be a correspondence from Z to H, then $h \circ (g \circ f) = (h \circ g) \circ f$.

Proof

It suffices to check that $\Gamma_{h\circ(g\circ f)} = \Gamma_{(h\circ g)\circ f}$.

$\forall (x,m) \in \Gamma_{h\circ(g\circ f)}$, iff $\exists z \in Z$, s.t. $(x,z) \in \Gamma_{g\circ f}$, $(z,m) \in \Gamma_h$, iff $\exists y \in Y$, s.t. $(x,y) \in \Gamma_f$, $(y,z) \in \Gamma_g$, $(z,m) \in \Gamma_h$, iff $(x,y) \in \Gamma_f$, $(y,m) \in \Gamma_{h\circ g}$, iff $(x,m) \in \Gamma_{(h\circ g)\circ f}$. $\square$

1.1.4.4 Prop

$f \circ \mathrm{Id}_X = f = \mathrm{Id}_Y \circ f$.

Proof

$(x,y) \in \Gamma_{f\circ \mathrm{Id}_X}$ iff $\exists x \in X$, $(x,x) \in \mathrm{Id}_X$, $(x,y) \in \Gamma_f$.

$\mathrm{Id}_Y \circ f = ((\mathrm{Id}_Y \circ f)^{-1})^{-1} = (f^{-1} \circ \mathrm{Id}_Y)^{-1} = (f^{-1})^{-1} = f$. $\square$

1.1.4.5 Prop

Let A be a subset of X, then $(g \circ f)(A) = g(f(A))$.

In particular, $\mathrm{Im}(g \circ f) \subseteq \mathrm{Im}(g)$.

Moreover if $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, then $\mathrm{Im}(g \circ f) = \mathrm{Im}(g)$.

Proof

$(g \circ f)(A) = \{z \in \mathrm{A}_g | \exists x \in A, (x, z) \in \Gamma_{g \circ f}\} = \{z \in \mathrm{A}_g | \exists x \in A, \exists y \in Y, (x, y) \in \Gamma_f, (y, z) \in \Gamma_g\} = \{z \in \mathrm{A}_g | \exists y \in f(A), (y, z) \in \Gamma_g\} = g(f(A))$.

$\mathrm{Im}(g \circ f) = (g \circ f)(\mathrm{D}_{g \circ f}) = g(f(\mathrm{D}_{g \circ f})) \subseteq g(\mathrm{D}_g) = \mathrm{Im}(g)$

$\mathrm{Im}(g) = g(\mathrm{Dom}(g)) \subseteq g(\mathrm{Im}(f)) = \mathrm{Im}(g \circ f)$. □

Remark

$\mathrm{Dom}(g \circ f) \subseteq \mathrm{Dom}(f)$, $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$

If $\mathrm{Im}(f) \subseteq \mathrm{Dom}(g)$, $\mathrm{Dom}(g \circ f) = \mathrm{Dom}(f)$.

1.1.5 Surjectivity

In this subsection, we fix a correspondence f from $\mathrm{D}_f = X$ to $\mathrm{A}_f = Y$.

1.1.5.1 Def

If $\mathrm{A}_f = \mathrm{Im}(f)$, then we call it is surjective.

If $\mathrm{D}_f = \mathrm{Dom}(f)$, then we call it is multivalued mapping.

1.1.5.2 Prop

(1) If f is surjective, $\forall B \subseteq \mathrm{A}_f$, $B \subseteq f(f^{-1}(B))$

(2) If f is multivalued mapping, $\forall A \subseteq \mathrm{D}_f$, $A \subseteq f^{-1}(f(A))$

Proof

(1) Let $y \in B$, since $B \subseteq \mathrm{A}_f = \mathrm{Im}(f)$ (since f is surjective). Then $\exists\ x \in \mathrm{D}_g$, such that $(x, y) \in \Gamma_f$. By definition, $x \in f^{-1}(B)$ □

1.1.5.3 Prop

Let f, g be correspondences.

(1) If $g \circ f$ is surjective, then g is surjective.

(2) If $g \circ f$ is multivalued, then f is multivalued.

Proof

Let $z \in \mathrm{A}_g = \mathrm{A}_{g\circ f}$. Since $g\circ f$ is surjective. $\exists\, x$ such that $(x, z) \in \Gamma_{g\circ f}$. That means $\exists\, x, \exists\, y$ such that $(x, y) \in \Gamma_f, (y, z) \in \Gamma_g$, this implies $z \in \mathrm{Im}(g)$

1.1.5.4 Prop

Let f, g be correspondences.

(1) Suppose that f, g are surjective and $\mathrm{Dom}(g) \subseteq \mathrm{A}_f$, then $g \circ f$ is also surjective.

(2) Suppose that f, g are multivalued and $\mathrm{Im}(f) \subseteq \mathrm{D}_g$, then $g \circ f$ is also multivalued.

Proof

(1) Let $z \in \mathrm{A}_{g\circ f} = \mathrm{A}_f$. Since g is surjective, $\exists\, y \in \mathrm{D}(g) \subseteq \mathrm{A}_f$ such that $(y, z) \in \Gamma_g$. Since f is surjective, $\exists\, x$ such that $(x, y) \in \Gamma_f$. We then get $(x, z) \in \Gamma_{g\circ f}$. Hence $z \in \mathrm{Im}(g \circ f)$, so $g \circ f$ is surjective. □

1.1.6 Injectivity

1.1.6.1 Def

Let f be a correspondence.

If $(x, y) \in \Gamma_f$, we say that y is an image of x by f and x is an inverse image.

If any element $x \in \mathrm{D}_f$ has at most an image, then f is functional.

If any element $y \in \mathrm{A}_f$ has at most an image, then f is injective.

Notation

If f is a function and $x \in \mathrm{Dom}(f)$, we denote by $f(x)$ the unique image of x by f.

We alse use the expression $x \mapsto f(x)$ to denote $(x, f(x)) \in \Gamma_f$ in the case $f(x)$ is functional.

1.1.6.2 Def

Let f be a correspondence.

If f is a function and $\mathrm{Dom}(f) = \mathrm{D}_f$, then we say that f is a mapping.

Notation

If f is a mapping from a set X to another, we use the expression: $f : X \to Y$ or $X \xrightarrow{f} Y$ to denote f is a mapping from X to Y.

1.1.6.3 Def

Let I be a set, $(A_i)_{i\in I}$ be a family of sets and $A = \bigcup\limits_{i\in I} A_i$. We denote $\prod\limits_{i\in I} A_i$ the set of all mappings such as $\mu : I \to A$ such that $\mu(j) \in A_j$ for any $j \in I$. $\prod\limits_{i\in I} A_i$ is called the product of $(A_i)_{i\in I}$.

If $\mu \in \prod\limits_{i\in I} A_i$, we often write μ in the form of a family $(\mu(i))_{i\in I}$.

In the case of all A_i equals to X, we denote by X^I the set $\prod\limits_{i\in I} X$. It is just the set of mappings from I to X.

If $n \in \mathbb{N}$, x^n denotes $X^{\{1,2,3,\ldots,n\}}$.

In particular if $x_1, x_2, x_3, \cdots, x_n$ are elements of X, then $(x_1, x_2, x_3, \cdots, x_n)$ denotes the mapping from $\{1, 2, 3, \ldots, n\}$ to X, sending $i \in \{1, 2, 3, \ldots, n\}$ to x_i.

$X^{\mathbb{N}} = \{\text{mappings from } \mathbb{N}\ \textit{to}\ X\}$. $(x_n)_{n\in\mathbb{N}}$ denotes the mapping sending $n \in \mathbb{N}$ to $x \in X$. Such mapping is called a sequence in X parameterized by $\mathbb{N}$.

1.1.6.4 Prop

Let f be a correspondence.

If f is injective, then for any set A, $f^{-1}(f(A)) \subseteq A$.

If f is functional, then for any set B, $f(f^{-1}(B)) \subseteq B$.

1.1.6.5 Prop

Let f, g be a correspondences.

If f, g are functional, then $g \circ f$ is alse functional.

Moreover, for any $x \in \mathrm{Dom}(g \circ f)$, $(g \circ f)(x) = g(f(x))$

If f, g are injective, then $g \circ f$ is alse injective. Moreover, for any $y \in \mathrm{Im}(g \circ f)$, $(g \circ f)^{-1}(y) = f^{-1}(g^{-1}(y))$

Proof

Take $z_1 = f(x_1)$ and $z_2 = f(x_2)$, just to prove that when $x_1 = x_2$,

$z_1 = z_2$. □

1.1.6.6 Prop

Let f, g be a correspondences.

If $g \circ f$ is injective and $\mathrm{Im}(f) \subseteq \mathrm{Dom}(g)$, then f is injective.

If $g \circ f$ is functional and $\mathrm{Dom}(g) \subseteq \mathrm{Im}(f)$, then g is functional.

1.1.6.7 Prop

Let f be a correspondence, I be a non-empty set.

(1) If f is functional, for any family set $(B_i)_{i \in I}$, $f^{-1}\left(\bigcap\limits_{i \in I} B_i\right) = \bigcap\limits_{i \in I} f^{-1}(B_i)$.

(2) If f is injective, for any family set $(A_i)_{i \in I}$, $f\left(\bigcap\limits_{i \in I} A_i\right) = \bigcap\limits_{i \in I} f^{-1}(A_i)$.

Proof

(1) We have already proved $f^{-1}\left(\bigcap\limits_{i \in I} B_i\right) \subseteq \bigcap\limits_{i \in I} f^{-1}(B_i)$. Let $x \in f^{-1}\left(\bigcap\limits_{i \in I} B_i\right)$, for any $i \in I$, $f(x) \in B_i$. So $f(x) \in \bigcap\limits_{i \in I} B_i$ and hence $x \in \bigcap\limits_{i \in I} f^{-1}(B_i)$. We get $f^{-1}\left(\bigcap\limits_{i \in I} B_i\right) \supseteq \bigcap\limits_{i \in I} f^{-1}(B_i)$ □

1.1.6.8 Def

Let f be a correspondence.

If f is functional and multivalued or if f^{-1} is injective and surjective then f is a mapping.

1.1.6.9 Prop

Let $f : X \to Y$ be a mapping.

(1) If $(A_i)_{i \in I}$ is a family of sets, $f\left(\bigcup\limits_{i \in I} A_i\right) = \bigcup\limits_{i \in I} f(A)$.

If $(B_i)_{i \in I}$ is a family of sets, $f^{-1}\left(\bigcup\limits_{i \in I} B_i\right) = \bigcup\limits_{i \in I} f^{-1}(B)$

(2) If $I \neq \varnothing$, $f\left(\bigcap\limits_{i \in I} A_i\right) \subseteq \bigcap\limits_{i \in I} f(A)$, $f^{-1}\left(\bigcup\limits_{i \in I} A_i\right) = \bigcup\limits_{i \in I} f^{-1}(A)$

(3) $\forall\ B \subseteq Y$, $f(f^{-1}(B)) \subseteq B$ while $\forall\ A \subseteq X$, $f^{-1}(f(A)) \supseteq A$

(4) Let $f : X \to Y$ and $g : Y \to Z$ be mappings, $g \circ f$ is also a mapping.

If f and g are surjective/injective, $g \circ f$ is alse surjective/injective.

If $g \circ f$ is surjective/injective, then g/f is surjective/injective.

1.1.7 Bijections

1.1.7.1 Def

Let f be a correspondence.

If f and f^{-1} are injective and surjective, then we say that f is a bijection or ono-to-one correspondence.

Example

Id_X is a bijection.

1.1.7.2 Prop

Let X, Y be sets, f be a correspondence $f = (X, Y, \Gamma_f)$. If f is a bijection, then $f^{-1} \circ f = \mathrm{Id}_X$, $f \circ f^{-1} = \mathrm{Id}_X$

1.1.7.3 Prop

Let X, Y be sets, f be a correspondence $f = (X, Y, \Gamma_f)$.

Assume that there exist correspondences g_1 and g_2 from Y to X, such that $g_1 \circ f = \mathrm{Id}_X$, $f \circ g_2 = \mathrm{Id}_Y$.

Then f is a bijection and $g_1 = g_2 = f^{-1}$.

1.1.7.4 Prop

Let $f : X \to Y$ and $g : Y \to Z$ be mappings.

If f and g are bijections, then $g \circ f$ is a bijection.

1.1.7.5 Prop

Let X, Y be sets, f be a correspondence from X to Y and g be a correspondence from Y to X.

If $g \circ f$ and $f \circ g$ are bijections, then f and g are both bijections.

Proof

$f \circ (g \circ (f \circ g)^{-1}) = (f \circ g) \circ (f \circ g)^{-1} = \mathrm{Id}_Y$, $((g \circ f)^{-1} \circ g) \circ f = (g \circ f)^{-1} \circ (g \circ f) = \mathrm{Id}_X$. □

1.1.7.6 Def

Let f, g be correspondences.

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

Notation

Let $f : X \to Y$ be a mapping and $A \subseteq X$, we denote by $f_{|A} : A \to X$ the mapping such that $\forall\ a \in A, f_{|A}(a) = f(a)$

Example

Let X be a set, $A \subseteq X$, we denote by $i_A : A \to X$ the mapping Id_A, called the inclusion mapping.

1.2 Binary Relation

1.2.1 Concept

1.2.1.1 Def

Let X be a set.

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

Notation

If R is a binary relation on X, $\forall\,(x, y) \in X^2$, the expression xRy denotes the statement $(x, y) \in \Gamma_R$.

We use the mirror symmetric expression to denote the inverse of a binary relation. For example, $x \ni y$ denotes $y \in x$.

We use the slashed expression to denote the negation of binary relation, that means $x \not R y$ denotes NOT xRy.

1.2.1.2 Def

Let X be a set, R be an binary relation on X.

- If $\forall\ x \in X$, xRx holds, we say that it is reflexive.
- If $\forall\ x \in X$, xRx does not hold, we say that it is irreflexive.
- If $\forall\ (x, y) \in X^2$, xRy implies yRx, we say that it is symmetric.
- If $\forall\ (x, y) \in X^2$, xRy and yRx implies $x = y$, we say that it is antisymmetric.
- If $\forall\ (x, y) \in X^2$, xRy and yRx cannot hold at the same time, we say that it is asymmetric.
- If $\forall\ (x, y, z) \in X^3$ such that xRy and yRz, xRz holds, we say that it is transitive.

1.2.2 Partially ordered relation

1.2.2.1 Def

If a binary relation is reflexive, antisymmetric and transitive, we say that it is an order relation.

Example

- If A is a set, then $\subseteq$ is an order relation on $\mathscr{P}(A)$.
- $\forall$ set X, the relation $=$ is an order relation on X.
- The relation of divisionability is an order relation on $\mathbb{N}$.

1.2.2.2 Def

If a binary relation is asymmetric and transitive, we say that it's a strict order relation.

Notation

In general, we use an underlined notaion to denote an order relation $\underline{R}$.

1.2.2.3 Prop

Let X be a set, $\underline{R}$ be an order relation on X.

Then the binary relation R on X defined as xRy iff $x\underline{R}y$ and $x \neq y$ is a strict order relation.

1.2.2.4 Def

Let $\underline{R}$ be an order relation, then the binary relation defined as xRy is called the strict order relation of $\underline{R}$.

1.2.2.5 Def

Let X be a set and $\leq$ be an order relation on X. The pair $(X, \leq)$ is called a partially ordered set.

If $\forall$ $(x, y) \in X^2$, either $x \leq y$ or $y \leq x$, then we say that $\leq$ is a total order

Example

$(\mathbb{R}, \leq)$ is a totally ordered set.

$(\mathbb{N}, \ |)$ is not a totally ordered set.

1.2.2.6 Def

Let $(X, \leq)$ be a partially ordered set, $A \subseteq X$.

- Let $M \in X$. If $\forall\ a \in A$, $a \leq M$, we say that M is an upperbound of A relatively to $\leq$.
- Let M be an upperbound of A relatively to $\leq$. If $M \in A$, then M is the greatest element of A denoted as $\max A$.
- Let $m \in X$. If $\forall\ a \in A$, $m \leq a$, we say that M is an lowerbound of A relatively to $\leq$.
- Let m be an lowerbound of A relatively to $\leq$. If $m \in A$, then m is the least element of A denoted as $\min A$.

1.2.2.7 Def

Let $(X, \leq)$ be a partially ordered set.

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

1.2.2.8 Prop

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

- $\forall\ Y \subseteq X$, if we equip Y with the binary relation $\leq_Y$, whose graph is $\Gamma_{\leq_Y} \subseteq \Gamma_{\leq_X}$, then $(Y, \leq_Y)$ is a well-order set.
- $(X, \leq)$ is a totally ordered set.

1.2.2.9 Theorem

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

Let $\mathbb{P}(\cdot)$ be a statement depending on a parameter in X. (For any $x \in X$, $\mathbb{P}(x)$ is a mathmetic statement)

Assume that $\forall\ x \in X$ ($\forall\ y \in X_{<x}, \mathbb{P}(y)$ holds) implies $\mathbb{P}(x)$, then $\forall\ x \in X\ \mathbb{P}(x)$ holds.

Proof

We argue it by contradiction.

Suppose that $A = \{x \in X | \mathbb{P}(x) \text{ doesn't hold}\}$ is non-empty. Since $(X, \leq)$ is a well-ordered set, it has a least element β. For any $x \in X$ such that $x < \beta$, one has $x \notin A$ and hence $\mathbb{P}(x)$ holds. By the hypothesis, $\mathbb{P}(\beta)$ holds. This leads to a contradiction. $\square$

1.3 Composition Laws

1.3.1 Definition

1.3.1.1 Def

Let X be a set. By composition law on X, we mean a mapping from (X, X) to X.

If $*$ is a composition law on X, $\forall\ (x, y) \in X^2$, we denote $x * y$ the image of (x, y) by $*$.

1.3.1.2 Def

Let X be a set and Y be a subset of X such that $\forall (x, y) \in Y^2$, $x*y \in Y$, we say that Y is closed under $*$. The restriction from $*$ defines a mapping $Y^2 \to Y$. The composition law is called the restriction of $*$ on Y. (Usually we still use $*$)

Example

$(\mathbb{Z}, +)$ and $(\mathbb{N}, +)$

1.3.1.3 Def

Let X be a set and $*$ be a composition law on X.

- If $\forall\ (x, y, z) \in X^3$, $(x * y) * z = x * (y * z)$, We say that $*$ is associative and that associative $(X, *)$ is a semi-group.
- If $\forall\ (x, y) \in X^2$, $x * y = y * x$, we say that X is commutative.

Example

Let X be a set and $\mathfrak{C}_X$ be the set of all the mappings from X to X.

Consider the composition law

$$\circ : \mathfrak{C}_X^2 \to \mathfrak{C}_X$$
$$(g, f) \mapsto g \circ f$$

$\forall\ (f, g, h) \in \mathfrak{C}_X^3$, $h \circ (g \circ f) = (h \circ g) \circ f$.

Therefore $(\mathfrak{C}_X, \circ)$ is a semi-group.

Notation

Let $(X, *)$ be a semi-group and $n \in \mathbb{N}_{\geq 2}$.

If $(x_1, \cdots, x_n) \in X^n$, we define $x_1 * \cdots * x_n$ in a recursive way as $x_1 * \cdots * x_n := (x_1 * \cdots * x_{n-1}) * x_n = ((x_1 * x_2) * x_3) * \cdots$

1.3.1.4 Prop

Let $(X, *)$ be a semi-group, $n \in \mathbb{N}_{>0}$ and $(x_1, \cdots, x_n) \in X^n$. $\forall\ i \in \{1, \ldots, n-1\}$, the following equality holds: $x_1 * \cdots * x_n = x_1 * \cdots * x_{i-1} * (x_i * x_{i+1}) * \cdots * x_n$

Note

结合律就是括号放在哪都没关系。

1.3.2 Monoids and groups

1.3.2.1 Def

Let $(M, *)$ be a semi-group. If an element $e \in M$ satisfies $\forall\ x \in M, e * x = x * e = x$, we say that e is a neutral element of $(M, *)$.

If $(M, *)$ has a neutral element, we say that it is a $Monoid$.

1.3.2.2 Prop

If a semi-group $(M, *)$ has a neutral element, then its neutral element is unique.

Proof

$e_1 = e_1 * e_2 = e_2$ $\square$

1.3.2.3 Def

Let $(M, *)$ be a monoid and e be its neutral element.

Let $x, y \in M$ such that $x * y = e$. In that case, we say that x is right invertical and y is a right inverse. We also say that y is left invertical and x is a left inverse.

1.3.2.4 Prop

Let $(M, *)$ be a monoid and $x \in M$. If x is left and right invertical, then x has a unique left inverse and right inverse, which are equal.

Proof

Let e be the neutral element of $(M, *)$, a be a left inverse of x, b be a right inverse of x.

$a = a * e = a * (x * b) = e * b = b$ □

1.3.2.5 Def

Let $(M, *)$ be a monoid.

We denote by $M^\times$ the set of all invertible elements of M.

If $M^\times = M$, we say that M is a group.

1.3.2.6 Def

a commutative group is called a abelian group.

Notation

Let X be a set. $\mathfrak{S}_X$ is the set of all bijections from X to X.

1.3.2.7 Prop

Let $(M, *)$ be a monoid.

- If $x \in M^\times$, then $x^{-1} \in M^\times$, and $((x^{-1})^{-1}) = x$.
- If $(x, y) \in (M^\times)^2$, then $xy \in M^\times$ and $(xy)^{-1} = y^{-1}x^{-1}$.

1.3.3 Submonoids and subgroups

1.3.3.1 Def

Let M be a monoid and $N \subset M$.

If the neutral element of M belongs to N and N is closed under the composition law of M, then we say that N is a *submonoid* of M. (Note that N equipped with this restriction composition law forms a monoid.)

1.3.3.2 Prop

Let M be a monoid, $H \subseteq M^\times$, $H \neq \varnothing$.

If $\forall\ (x, y) \in H^2, xy^{-1} \in H$, then H is a submonoid of M and H equipped with the restricted composition law forms a group.

Proof

Let e be the neutral element of H, $a \in H$.

$e = aa^{-1} \in H$.

$\forall\ y \in H, y^{-1} = ey^{-1} \in H$.

$\forall\ (x, y) \in H^2, xy = x(y^{-1})^{-1} \in H$.

Hence H is a submonoid.

$\forall\ y \in H, y^{-1} \in H$, so H forms a group. □

1.3.3.3 Def

Let G be a group.

We call subgroup of G any submonoid H of G such that $\forall\ x \in H, x^{-1} \in H$.

1.3.3.4 Prop

let M be a monoid.

- $M^\times$ is a submonoid of M, and $M^\times$ equipped with the restricted composition law forms a group.
- Let H be a non-empty subset of $M^\times$. If $\forall\ (x, y) \in H^2$, $xy^{-1} \in H$, then H is a subgroup of $M^\times$.

1.3.4 Rings

1.3.4.1 Def

Let A be a set, $+$ and $\times$ be composition laws on A.

- $(A, +)$ forms a commutative group.
- $(A, \times)$ forms a monoid.
- $\times$ is distributive with respect to $+$, namely $\forall\ (a, b, c) \in A^3$, $(a+b)c = ac + bc$ and $c(a + b) = ca + cb$.

We say the $(A, +, \times)$ satisfies the condition above is a unitary ring.

Notation

If we omit the composition law, while talking about a unitary ring, then by convention, the composition laws are denoted as $+$ and $\times$.

Also, the neutral element of $(A,+)$ is denoted as 0, called the zero element of A.

The neutral element of $(A,\times)$ is 1, which is called the *unity* of A.

The set $A^\times$ denotes the set of all invertible elements of $(A,\times)$.

Example

$(\mathbb{Z},+,\times)$ forms a commutative unitary ring.

So do $(\mathbb{Q},+,\times)$ and $(\mathbb{R},+,\times)$.

1.3.4.2 Def

Let A be a unitary ring.

If $A^\times = A\setminus\{0\}$, then we say A is a division ring.

If in addition $(A,\times)$ is commutative, then we say that A is a field.

1.3.4.3 Prop

Let A be a unitary ring.

- $\forall\ a\in A$, $a0 = 0a = 0$.
- $\forall\ (a,b)\in A^2, (-a)b = a(-b) = -ab$

Proof

$a0 = a(0+0) = a0 + a0$, $a0 = 0a = 0$

$ab + (-a)b = 0b = 0$, $ab + a(-b) = a0 = 0$ $\square$

1.3.5 Homomorphisms

1.3.5.1 Def

Let $(M,*), (N,\star)$ be monoids and e_M be the neutral element of M while e_N be the neutral element of N.

We call the morphism of monoids from $(M,*)$ to $(N,\star)$ any mapping $f: M\to N$ that satisfies

- $f(e_M) = e_N$
- $\forall\ (x,y)\in M^2, f(x*y) = f(x)\star f(y)$. (namely f preserves the neutral element and the composition law)

Remark

In the case where $(M, *), (N, \star)$ are groups, a morphism of monoid from $(M, *)$ to $(N, \star)$ is also called a morphism of groups.

Remark

If f is a bijection, we say that it's a isomorphism.

Example

$f : M \to N$, $m \mapsto e_N$, is a morphism of monoid.

1.3.5.2 Prop

Let $f : M \to N$ be a morphism of monoid. If $M_1 \subseteq M$ is a submonoid, $f(M_1) \subseteq N$ is also a submonoid.

1.3.5.3 Prop

Let $f : M \to N$ be a morphism of monoid.

$\forall\ a \in M^\times$, one has $f(a) \in N^\times$, and $(f(a))^{-1} = f(a^{-1})$.

Moreover, if G is a subgroup of $M^\times$, then $f(G)$ is also a subgroup of $N^\times$.

1.3.5.4 Prop

Let M be a monoid and G be a group. If a mapping $f : M \to G$ satisfies $\forall\ (a, b) \in M^2$, $f(ab) = f(a)f(b)$, then f is a morphism of monoids.

Proof

$f(e_M) = f(e_M e_M) = f(e_M)f(e_M)$, thus $f(e_M) = e_G$. $\square$

1.3.6 Universal morphism

1.3.6.1 Prop

Let $(M, *)$ be a monoid.

For any $x \in M$, there exists a unique morphism of monoid $f : (\mathbb{N}, +) \to (M, *)$, such that $f(1) = x$

Proof

Construct f in a recursive way.

Let $f(0) = e_M$, $f(1) = x$, $f(n+1) = f(n) * x$.

By induction on n, we can prove that $\forall\ (m,n) \in \mathbb{N}^2$, $f(m+n) = f(m) * f(n)$. The uniqueness is trivial. $\square$

1.3.6.2 Prop

Let M be a monoid.

- If $(x,y) \in M^2$ such that $xy = yx$, $\forall\ (m,n) \in \mathbb{N}^2$, $x^m y^n = y^n x^m$ and $(xy)^n = x^n y^n$
- If $x \in M^\times$, $\forall\ n \in \mathbb{N}$, $(x^n)^{-1} = (x^{-1})^n$

1.3.6.3 Prop

Let M be a monoid.

For any $x \in M^\times$, there exists a unique morphism of monoid from $(\mathbb{Z}, +)$ to M that sends $1 \in \mathbb{Z}$ to x. (The image of $n \in \mathbb{Z}$ by this is denoted as x^{*n})

Proof

Assume that such morphism $f : (\mathbb{Z}, +)$ exists.

The restriction of f to $\mathbb{N}$ is also a morphism of monoid from $\mathbb{N}$ to M that sends 1 to x.

Therefore, for any natural number, $f(n) = x^{*n}$. Since $(\mathbb{Z}, +)$ is a group, $f(-n) = (x^{*n})^{-1}$. Hence f is unique.

To prove the existence, check the mapping $f : \mathbb{Z} \to M$

$$f(n) = \begin{cases} x^{*n}, & \text{if } n \in \mathbb{N} \\ (x^{*n})^{-1}, & \text{if } n \in \mathbb{Z}_{<0} \end{cases}$$

$\square$

Remark

Let $(m,n) \in \mathbb{N}$, $x^{*m+n} = x^{*m} * x^{*n}$

Notation

Under the situation, if $*$ is written as $\times$, x^{*n} is denoted as x^n. If $*$ is written as $+$, x^{*n} is denoted as nx.

1.3.7 Morphism of rings

1.3.7.1 Def

Let A and B be unitary rings.

We call the morphism of unitary ring from A to B any mapping $f : A \to B$ that is a morphism of groups from $(A, +)$ to $(B, +)$ and a morphism of monoid from $(A, \times)$ to $(B, \times)$.

1.3.7.2 Prop

There is a unique morphism of unitary ring from $\mathbb{Z}$ to R.

Proof

Let 0_R be the zero element of R, 1_R be the unity of R.

There exists a unique morphism of groups from $(\mathbb{Z}, +)$ to $(R, +)$ that send $1 \in \mathbb{Z}$ to 1_R, $n \mapsto n1_R$

It means to check that this mapping preserves the multiplication laws. Namely, $\forall\ (n, m) \in \mathbb{Z}^2$, $(n1_R)(m1_R) = mn1_R$.

If $k \in \mathbb{N}$, then $k1_R = \underbrace{1_R + \cdots + 1_R}_{k \text{ copies}}$, $f(k) = f(\underbrace{1 + \cdots + 1}_{k \text{ copies}})$ □

1.3.7.3 Def

Let K be a commutative unitary ring.

We call K-algebra any unitary ring R equipped with a morphism of ring.

$f : K \to R$ such that $\forall\ b \in K$, $\forall\ x \in R$, $f(a)x = xf(a)$.

We often say (R, f) is a K-algebra.

1.3.8 Formal power series

Fix K as a commutative unitary ring.

1.3.8.1 Def

Let T only be a formal symbol. We denotes $K^{\mathbb{N}}$ as $K[|T|]$.

If $(a_n)_{n\in\mathbb{N}}$ is an element of $K^{\mathbb{N}}$, 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 coefficent of T^n of this formal power series.

Notation

We often:

- omit terms with coefficent 0.
- write T^1 as T.
- omit coefficent there are 1.
- omit T^0.

Notation

Let K be a commutative unitary ring and

(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.

1.3.8.2 Def

Let K be a commutative unitary ring. $K[|T|] = \{a_0 + a_1T + \cdots + a_nT^n + \cdots | (a_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}\}$.

Define the composition law on the set $\forall\ F(T) = a_0 + a_1T^+ \cdots + a_nT^n + \cdots$ *and* $G(T) = a_0 + a_1T^+ \cdots + a_nT^n + \cdots$, we let $F(T) + G(T) = \sum_{i=0} (a_i + b_i)T^i$, $F(T)G(T) = \cdots$ as you can imagine.

1.3.8.3 Prop

- $(K[|T|], +, \times)$ forms a commutative unitary ring.
- $K \to K[|T|], \lambda \mapsto \lambda T^0$ is a morphism of ring

1.3.8.4 Def

Let $F(T) = \sum_{n\in\mathbb{N}} a_nT^n \ \in K[|T|]$

We denote by $F'(T)$ or by $D(F(T))$ the differential of formal power series $\sum_{n\in\mathbb{N}} na_nT^{n-1}$.

1.3.8.5 Prop

the mapping of differential $D : (K[|T|], +) \to (K[|T|], +)$ is a morphism of groups.

1.3.8.6 Prop

If $F(T), G(T) \in K[|T|]$, then $D(F(T)G(T)) = F'(T)G(T) + F(T)G'(T)$

1.3.8.7 Corollary

$\forall\ (a, F(T)) \in K \times K[|T|]$, $D(aF(T)) = aF'(T)$.

Notation

We denote $\exp(T)$ as $\sum\limits_{n\in\mathbb{N}} \frac{1}{n!}T^n = \sum\limits_{n\in\mathbb{N}} (n!)^{-1}T^n$ which called the exponential series.

Remark

$\exp'(T) = \exp(T)$

1.3.8.8 Prop

Assume that any positive integer is invertable in K.

If $F(T) \in K[|T|]$ satisfies $F'(T) = 0$ then $F(T)$ is of the form aT^0 for any $a \in K$.

1.3.8.9 Theorem

Let $F(T) \in K[|T|]$. Then $F(T) \in K[|T|]^\times$ iff $a_0 \in K^\times$

Proof

Let $F(t) = \sum\limits_{n\in\mathbb{N}} a_n T^n$.

If $F(T)$ is invertable, there exists $G(t) = \sum\limits_{n\in\mathbb{N}} b_n T^n \in K[|T|]$ such that $F(T)G(T) = 1T^0$. So $a_0 b_0 = 1$ hence $a_0 \in K^\times$.

We suppose that $a_0 \in k^\times$, we Construct a formal power series $\sum\limits_{n\in\mathbb{N}} b_n T^n$ in a recurrsive way as $b_0 = a_0^{-1}$, $b_n = -\left(\sum\limits_{l=1}^{n} a_l b_{n-l}\right) a_0^{-1}$ Then $F(T)G(T) = 1$, hence $F(T) \in K[|T|]^\times$. $\square$

1.3.9 Topology of formal power series

1.3.9.1 Def

Let $F(T) \in K[|T|]$, $F(T) = \sum_{n \in \mathbb{N}} a_n T^n$.

If $\{n \in \mathbb{N} | a_n \neq 0\} \neq \varnothing$, its least element is denoted as $\mathrm{ord}(F(T))$ called the order of $F(T)$.

If $\{n \in \mathbb{N} | a_n \neq 0\} = \varnothing$, then by convention, $\mathrm{ord}(F(t))$ denote as $+\infty$.

If $\mathrm{ord}(F(T)) = l \in \mathbb{N}$, then $F(T)$ is of the form $a_l T^l + a_{l+1} T^{l+1} + \cdots$. Hence there exist $G(T) \in K[|T|]$ such that $F(T) = T^l G(T)$.

In general, for any $n \in \mathbb{N}$, $\mathrm{ord}(F(T)) \geq n$ iff $F(T)$ is of the form $T^n G(T)$.

Example

$\mathrm{ord}(T^l) = l$

1.3.9.2 Prop

For elements $F(t)$ and $G(T)$ of $K[|T|]$.

$\mathrm{ord}(F(T)+G(T)) \geq \min\{\mathrm{ord}(F(T)), \mathrm{ord}(G(T))\}$ and $\mathrm{ord}(F(T)G(T)) \geq \mathrm{ord}(F(T)) + \mathrm{ord}(G(T))$

1.3.9.3 Def

Let $(F_i(T))_{i \in \mathbb{N}}$ 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 exist $N_l \in \mathbb{N}$, such that $\forall\ (i,j) \in \mathbb{N}^2_{\geq N_l}$, $\mathrm{ord}(F_i(T) - F_j(T)) \geq l$.
- We say that $(F_i(T))_{i \in \mathbb{N}}$ converges to $F(T)$ if $\forall\ N \in \mathbb{N}$,$\exists\ i \in \mathbb{N}_{\geq N}$, $\mathrm{ord}(F_i(T) - F(T)) \geq l$

1.3.9.4 Theorem(Completeness of formal power series)

Let $(F_i(T))_{i \in \mathbb{N}}$ be a sequence in $K[|T|]$.

It is a Cauchy sequence iff then it converges to some $F(T) \in K[|T|]$. Moreover if it converges, then its limit if unique.

Proof

Suppose that $(F_i(T))_{i\in\mathbb{N}}$ converges to $F(T)$ and $G(T)$. $\forall\, l \in \mathbb{N}, \exists\, N_1, N_2 \in \mathbb{N}$ such that

- $\forall\; i \in \mathbb{N}_{\geq N_1}$, $\mathrm{ord}(F_i(T) - F(T)) \geq l$
- $\forall\; i \in \mathbb{N}_{\geq N_2}$, $\mathrm{ord}(F_i(T) - G(T)) \geq l$

So if we take $N = \max\{N_1, N_2\}$, then $\forall\; i \in \mathbb{N}_{>N}$, $\mathrm{ord}(F(T) - G(T)) = \mathrm{ord}((F(T)-F_i(T))+(F_i(T)-G(T))) \geq \min\{\mathrm{ord}(F(T)-F_i(T)), \mathrm{ord}(F_i(T)-G(T))\} \geq 1$. Hence $\forall\; l \in \mathbb{N}$, $\mathrm{ord}(F(T) - G(T)) \geq l$, this leads to $F(T) = G(T)$

$\forall\; (i,j) \in \mathbb{N}^2_{>N}$, $\mathrm{ord}(F_i(T) - F_j(T)) = \mathrm{ord}((F_i(T) - F(T)) + (F(T) - F_j(T))) \geq \min\{\mathrm{ord}(F(T) - F_i(T)), \mathrm{ord}(F_j(T) - G(T))\} \geq 1$. Therefore $(F_i(T))_{i\in\mathbb{N}}$ is a Cauchy sequence

We let $(F_i(T))_{i\in\mathbb{N}}$ be a Cauchy sequence. $\forall\; l \in \mathbb{N}, \exists\; N_l \in \mathbb{N}$ such that $\forall\; (i,j) \in \mathbb{N}^2_{>N}$, $\mathrm{ord}((F_i(T) - F(T)) + (F(T) - F_j(T))) \geq l+1$. In particular, $F_i(T)(i \geq N_l)$ have the same coefficent of T^l. We denote by a_l this common coefficent. Let $F(T) = \sum\limits_{n\in\mathbb{N}} a_n T^n$, $\forall\; i \in \mathbb{N}$ such that $i \geq \max\{N_0, \ldots, N_l\}$, $Fi(T)$ and $F(T)$ have the same coefficent of $T^0, T^1, \cdots, T^l$. Hence, $\mathrm{ord}(F_i(T) - F(T)) \geq l+1$ $\square$

1.3.10 Universal proporty of $K[T]$ and $K[|T|]$

1.3.10.1 Def

Let A and B be K-algebras. We call morphism of K-algebra from A to B any morphism of unitary ring $\varphi : A \to B$ such that $\forall\; \lambda \in K, \forall\; a \in A$ and $\varphi(\lambda a) = \lambda\varphi(a)$

1.3.10.2 Prop

Let A be a K-algebra and x be an element of A.

There exist a unique morphism of K-algebra from $K[T] := \{\sum\limits_{n\in\mathbb{N}} a_n T^n | \exists d \in \mathbb{N}, \text{s.t. } \forall n \geq d, a_n = 0\}$ to A that sends T to x.

Proof

If such morphism $f : k[T] \to A$ exists, then

$$\begin{aligned} f(a_0 + a_1T + \cdots + a_dT^d) &= \sum_{i=0}^{d} a_i f(T)^i \\ &= \sum_{i=0}^{d} a_i x^i \end{aligned}$$

Exercise

Check that $f : K[T] \to A$ defined as $f(\sum_{n\in\mathbb{N}} a_nT^n) := \sum_{n\in\mathbb{N}} a_nX^n$ is a morphism of K-algebra.

1.3.10.3 Corollary

Let $P(T) \in K[|T|]$ such that $\mathrm{ord}(P(T)) \geq 1$. There exists a unique morphism of K-algebra from $K[|T|]$ to $K[|T|]$ that send T to $P(T)$.

1.3.10.4 Def

Let $F(T) \in k[T]$, $F(T) = \sum_{n\in\mathbb{N}} a_nT^n$.

If $\{n \in \mathbb{N} | a_n \neq 0\} \neq \varnothing$, its greatest element is denoted as $\deg(F(T))$ called the degree of $F(T)$.

还有一些暑假的命题没有补全, 择日补全。

Chapter 2

Sequence

由于一些历史原因，本章前三节中的部分命题没有给出证明。

2.1 Supremum and infimum

2.1.1 Def

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

- If the set $\{y \in Y | \forall a \in A, a \leq y\}$ has a least element, then we say that A has a supremum in Y with respect to $\leq$, denoted by $\sup\limits_{(Y,\leq)} A$ this least element and called it the supremum of A in Y(with respect to $\leq$).
- If the set $\{y \in Y | \forall a \in A, y \leq a\}$ has a greatest element, we say that A has an infimum in Y with respect to $\leq$. We denote by $\inf\limits_{(Y,\leq)} A$ this greatest element and call it the infimum of A in Y.

Observation

$\inf\limits_{(Y,\leq)} A = \sup\limits_{(Y,\geq)} A$

Notation

Let $(X, \leq)$ be a partially ordered set, 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 $\sup\{x_i | i \in I\}$ (in X).

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 $\sup\{x_i | i \in I, \mathbb{P}(i) \text{ holds}\}$.

Example

Let $A = \{x \in \mathbb{R} | 0 \leq x < 1\} \subseteq \mathbb{R}$. We equip $\mathbb{R}$ with the usual order relation.

$\{y \in \mathbb{R} | \forall x \in A, x \leq y\} = \{y \in \mathbb{R} | y \geq 1\}$, so $\sup A = 1$.

$\{y \in \mathbb{R} | \forall x \in A, y \leq x\} = \{y \in \mathbb{R} | y \geq 0\}$, hence $\inf A = 0$.

Example

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

2.1.2 Prop

Let $(X, \leq)$ be a partially ordered set, A, Y, Z be subsets of X, such that $A \subseteq Z \subseteq Y$

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

The inf case is similar to sup.

2.1.3 Prop

Let $(X, \leq)$ be a partially ordered set, A, B, Y be subsets of X such that $A \subseteq B \subseteq Y$

- If $\sup\limits_{(Y,\leq)} A$ and $\sup\limits_{(Y,\leq)} B$ exists, then $\sup\limits_{(Y,\leq)} A \leq \sup\limits_{(Y,\leq)} B$
- If $\inf\limits_{(Y,\leq)} A$ and $\inf\limits_{(Y,\leq)} B$ exists, then $\inf\limits_{(Y,\leq)} A \geq \inf\limits_{(Y,\leq)} B$

2.1.4 Prop

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

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

2.2 Intervals

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

Notation

If $(a, b) \in X \times X$ such that $a \leq b$, $[a, b]$ denotes $\{x \in X | a \leq x \leq b\}$

2.2.1 Def

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

Note

区间具有“介值性”。

Example

Let $(a, b) \in X \times X$, such that $a \leq b$. Then the following sets are intervals

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

2.2.2 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

Proof

(1) Trivial. $\square$

(2) By induction, we only need to check that $\forall a \in I_\lambda$, $\forall b \in I_\mu$, $[a, b] \subseteq I_\lambda \cup I_\mu$

Let x be an element of $I_\lambda \bigcup I_\mu$.

- If $b \leq x$, $[a,b] \subseteq [a,x] \subseteq I_\lambda$ because $\{a,x\} \subseteq I_\lambda$
- If $x \leq a$, $[a,b] \subseteq [x,b] \subseteq I_\mu$ because $\{b,x\} \subseteq I_\mu$
- If $a < x < b$, then $[a,b] = [a,x] \cup [x,b] \subseteq I_\lambda \bigcup I_\mu$

□

2.2.3 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.

2.2.4 Prop

Let I be an interval in X.

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

Proof

(1) Let $y \in [x, b[$. Since $y < b$, y is not an upper bound of I, thus $\exists z \in I$, $y < z$. Therefore, $y \in [x, z[\subseteq [x, z] \subseteq I$, which means $[x, b[\subseteq I$.

(2) Similar to (1). □

2.2.5 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 $[a, b]$, $[a, b[$, $]a, b]$, $]a, b[$.

2.2.6 Def

Let $(X, \leq)$ be a totally ordered set.

If $\forall (x, z) \in X \times X$, such that $x < z$, $\exists y \in X$ such that $x < y < z$, then we say that $(X, \leq)$ is thick.

2.2.7 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 $\inf I = a$ and

$\sup I = b$ (For it's thick, I can not be empty set.)

Proof

Since X is thick, there exists $x_0 \in]a, b[$. By definition, b is an upper bound of I. If b is not the supremum of I, there exists an upper bound M of I such that $x_0 \leq M < b$. Since X is thick, there exists $M' \in X$ such that $M < M' < b$. Since $[x_0, b[\subseteq I$. Hence M and M' belong to I, which conflicts with the uniqueness of maximum. □

2.3 Enhanced real line

2.3.1 Def

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

Thus $\mathbb{R}\cup\{-\infty,+\infty\}$ becomes a totally ordered set, and $\mathbb{R}=]-\infty,+\infty[$ Obviously, this is a thick totally ordered set.

We define:

- $\forall x\in]-\infty,+\infty] \quad x+(+\infty):=+\infty \quad (+\infty)+x:=+\infty$
- $\forall x\in[-\infty,+\infty[\quad x+(-\infty):=-\infty \quad (-\infty)+x:=-\infty$
- $\forall x\in]0,+\infty] \quad x(+\infty):=(+\infty)x:=+\infty \quad x(-\infty):=(-\infty)x:=-\infty$
- $\forall x\in[-\infty,0[\quad x(+\infty):=(+\infty)x:=-\infty \quad x(-\infty):=(-\infty)x:=+\infty$
- $-(+\infty):=-\infty \quad -(-\infty):=+\infty \quad (\infty)^{-1}:=0$
- THE FOLLOWING LINE IS NOT DEFINED

  $(+\infty)+(-\infty) \quad (-\infty)+(+\infty) \quad (+\infty)0 \quad 0(+\infty) \quad (-\infty)0 \quad 0(-\infty)$

2.3.2 Def

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.

Example

Let Ω be a set. $(\mathscr{P}(\Omega),\subseteq)$ is order complete. If $\mathscr{F}$ is a subset of $\mathscr{P}(\Omega)$, $\sup\mathscr{F}=\bigcup\limits_{A\in\mathscr{F}}A$

Remark, $\inf\varnothing=\Omega \quad \sup\varnothing=\varnothing$

Remark

$(\mathbb{R}\bigcup\{-\infty,+\infty\},\leq)$ is order complete

In $\mathbb{R}\bigcup\{-\infty,+\infty\} \quad \sup\varnothing=-\infty \quad \inf\varnothing=+\infty$

Notation

- $[-\infty, +\infty]$ denotes $\mathbb{R} \bigcup \{-\infty, +\infty\}$
- 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 $(-1)A$

2.3.3 Prop

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

2.3.4 Def

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

- $\forall (a, b) \in \mathbb{R} \times \mathbb{R}$ such that $a < b$, 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 $\mathbb{R}$

Note

本笔记不会涉及实数构造的具体细节，在此只给出实数的公理化定义。

2.3.5 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.

2.3.6 Theorem

Let I and J be non-empty sets. $f : I \to [-\infty, +\infty]$, $g : J \to [-\infty, +\infty]$. $a = \sup_{x \in I} f(x), b = \sup_{y \in J} g(y), c = \sup_{(x,y) \in I \times J, \{f(x), g(y)\} \neq \{+\infty, -\infty\}} (f(x) + g(y))$

If $\{a, b\} \neq \{+\infty, -\infty\}$ then $c = a + b$

inf takes the same if $(-\infty) + (+\infty)$ doesn't happen.

Proof

$\forall (x, y) \in I \times J$, s.t. $f(x) + g(y)$ is well defined, $f(x) + g(y) \leq a + b$, thus $c \leq a + b$.

Suppose $-\infty \in \{a, b\}$, thus $c \leq a + b = -\infty$.

Suppose $-\infty \notin \{a, b\}$, thus $\exists (x_0, y_0) \in I \times J$, s.t. $f(x_0) > -\infty$ and $g(y_0) > -\infty$. Suppose $+\infty \in \{a, b\}$, thus $c = a + b = +\infty$. Suppsoe $+\infty \notin \{a, b\}$, thus $\forall y_0 \in J$, $g(y_0) \notin \{-\infty, +\infty\}$. Fix $y_0 \in J$, $c \geq \sup_{x \in I}(f(x) + g(y_0)) = a + g(y_0)$, take $\sup_{y_0 \in J}$, we get $c \geq a + b$. $\square$

Note

$+\infty$ 和 $-\infty$ 使得整个证明变得冗长且复杂。

2.3.7 Corollary

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

Then $\sup_{x \in I, \{f(x), g(x)\} \neq \{+\infty, -\infty\}} (f(x) + g(x)) \leq \sup_{x \in I} f(x) + \sup_{x \in I} g(x)$ (provided when the sum are defined)

inf takes the similar (provided when the sum are defined).

2.4 Vector space

K denotes a unitary ring.

Let 0 be the zero element of K, 1 be the unity of K.

2.4.1 Def

Let $(V, +)$ be a commutative group.

We call left K-module structure or right K-module structure any mapping $\phi : K \times V \to V$ that satisfies:

- $\forall\ (a, b) \in K \times K,\ x \in V,$

  Left：$\phi(ab, x) = \phi(a, \phi(b, x))$

  Right：$\phi(ab, x) = \phi(b, \phi(a, x))$

- $\forall\ (a, b) \in K \times K,\ x \in V,\ \phi(a + b, x) = \phi(a, x) + \phi(b, x)$

- $\forall\ a \in K,\ (x, y) \in V \times V,\ \phi(a, x + y) = \phi(a, x) + \phi(a, y)$

- $\forall\ x \in V,\ \phi(1, x) = x$

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

Note

注意到，左 K 模和右 K 模的实质差别只在于结合律时 K 中元素“作用”在 V 中元素上的先后顺序不同。

Remark

Let K^{op} be the set K equipped with the following 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 just a right K-module.

Any right K^{op}-module is just a left K-module.

Also, $(K^{\text{op}})^{\text{op}} = K$

Notation

When we talk about a left K-module $(V, +)$, we often write its left K-module structure as:

$K \times V \to V$

$(a, x) \mapsto ax$

Then,the axioms become:

- $\forall\ (a, b) \in K \times K,\ x \in V,\ (ab)x = a(bx)$
- $\forall\ (a, b) \in K \times K,\ x \in V,\ (a + b)x = ax + bx$
- $\forall\ a \in K,\ (x, y) \in V \times V,\ a(x + y) = ax + ay$
- $\forall\ x \in V,\ 1x = x$

Also, when we talk about a right K-module $(V, +)$, we often write its right K-module structure as:

$K \times V \to V$

$(a, x) \mapsto xa$

Then, the axioms become

- $\forall\ (a, b) \in K \times K,\ x \in V,\ x(ab) = (xa)b$
- $\forall\ (a, b) \in K \times K,\ x \in V,\ x(a + b) = xa + xb$
- $\forall\ a \in K,\ (x, y) \in V \times V,\ (x + y)a = xa + ya$
- $\forall\ x \in V,\ x1 = x$

Remark

If K is commutative, then $K^{\text{op}} = K$, so left K-module structure and right K-module structure are the same. We call them K-module structure. A commutative group equipped with a K-module sturcture is called a K-module. If K is a field, a K-module is also called a K-vector space.

Remark

Let ϕ :$K \times V \to V$ be a left K-module structure.

$\forall\ x \in V$, $\phi(\cdot, x) : K \to V$, $(a \in K) \mapsto ax$, is a morphism of groups.

Hence $0x = 0$, $(-a)x = -(ax)$

$\forall a \in K$, $\phi(a, \cdot) \to V$, $(x \in V) \mapsto V$ is also a morphism of groups.

Thus $a0 = 0$, $a(-x) = -(ax)$.

Example 1

Let $(V, +)$,be a commutative group. For any $x \in V$, $n \in \mathbb{Z}$, let nx be the image of n by the unique morphism of groups from $\mathbb{Z}$ to V sending 1 to x. Thus we obtain a mapping:

$\mathbb{Z} \times V \to V$

$(n, x) \mapsto nx$.

By definition, this mapping satisfies the five axiom a K-module needed.

Therefore,this morphism is a $\mathbb{Z}$-module sturcture.

Example 2

Let I be a set, $K^I = \{\text{mapping from } I \text{ to } K\}$

If f and g are elements of K^I, we let $f + g := I \to K$, $x \mapsto f(x) + g(x)$, because K is a unitary ring, it is easy to prove $(K^I, +)$ forms a commutative group.

The mapping $K \times K^I \to K^I$, $(a, f) \mapsto af$,$(af)(x) = af(x)$ is a left K-module structure, since K is a unitary ring.

Notation

We can also write an element μ of K^I in the form of a family $(u_i)_{i \in I}$ to describe a mapping from I to K. ((u_i) is the image of $i \in I$ by μ.)

Then we rewrite **Example 2** in the latest form of mappings.

Left K-module structure: $a(\mu_i)_{i\in I} = (a\mu_i)_{i\in I}$

Right K-module structure: $(\mu_i)_{i\in I}a = (\mu_i a)_{i\in I}$.

2.4.2 Def

Let V be a left K-module. If W is a subgroup of V, such that for any $a \in K$, for any $x \in W$, $ax \in W$, then we say that W is a left sub-K-module of V.

Example

Let I be a set, let $K^{\oplus I}$ be the subset of K^I composed of mappings $f : I \to K$ such that $\mathrm{Im}_f = \{x \in I | f(x) \neq 0\}$ is finite.

It can be a left or right sub-K-module of K^I.

2.4.3 Def

Let V and W be left K-modules. A morphism of groups φ: $V \to W$ is called a morphism of left K-module if it satisfies that $\forall (a, x) \in K \times V$, $\varphi(ax) = a\varphi(x)$

Moreover, if K is a commutative unitary ring, a morphism of K-module is also called a K-linear mapping.

We denote by $\mathrm{Hom}_{K-\mathrm{Mod}}(V, W)$ the set of all morphisms of left K-modules from V to W.

Notice that $\mathrm{Hom}_{K-\mathrm{Mod}}(V, W)$ is a subgroup of W^V.

2.4.4 Prop

Let G and H be groups, and $f : G \to H$ be a morphism of groups

(1)$\mathrm{Im}(f) \subseteq H$ is a subgroup of H

(2)$\mathrm{Ker}(f) = \{x \in G | f(x)$ is the neutral element of H$\}$ is a subgroup of G

(3)f is injective iff $\mathrm{Ker}(f) = \{$the neutral element of $G\}$

Proof

(1)$\forall (x, y) \in G^2$, $f(x)f^{-1}(y) = f(x)f(y^{-1}) = f(xy^{-1}) \in \mathrm{Im}(f)$ □

(2)$\forall (x, y) \in \mathrm{Ker}(f)$, $f(xy^{-1}) = f(x)f^{-1}(y) = e_H$ □

(3)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 $a = b$ □

Notation

$\mathrm{Ker}(f)$ is called the kernel of f.

2.4.5 Theorem

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

$\forall i \in I$, $e_i \in K^I$, s.t. $e_i = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{if } j \neq i \end{cases}$

The mapping $\Psi : \mathrm{Hom}_{K-\mathrm{Mod}}(K^{\oplus I}, V) \to V^I$, $\varphi \mapsto (\varphi(e_i))_{i \in I}$ is a group isomorphism.

Proof

$(\varphi + \psi)(e_i) = \varphi(e_i) + \psi(e_i)$, $\forall i \in I$, $\forall (\varphi, \psi) \in \mathrm{Hom}^2_{K-\mathrm{Mod}}(K^{\oplus I}, V)$. Thus $\Psi(\varphi + \psi) = \Psi(\varphi) + \Psi(\varphi)$, which means Ψ is a group morphism.

Let $\varphi \in \mathrm{Hom}_{K-\mathrm{Mod}}(K^{\oplus I}, V)$, s.t. $\forall i \in I$, $\varphi(e_i) = 0 \in V$. Let $a = (a_i)_{i \in I} \in K^{\oplus I}$, then $a = \sum\limits_{i \in I, a_i \neq 0} a_i e_i$. $\varphi(a) = \sum\limits_{i \in I, a_i \neq 0} a_i \varphi(e_i) = 0 \in V$. Therefore, Ψ is injective.

Let $x = (x_i)_{i \in I} \in V^I$. Define $\varphi_x \in \mathrm{Hom}_{K-\mathrm{Mod}}(K^{\oplus I}, V)$, s.t. $\forall a = (a_i)_{i \in I} \in K^{\oplus I}$, $\varphi_x(a) = \sum\limits_{i \in I, a_i \neq 0} a_i x_i$. Then $\forall i \in I$, $\varphi_x(e_i) = \sum\limits_{j \in I, e_i \neq 0} e_i x_j = x_i$. Therefore, $(\varphi_x(e_i))_{i \in I} = (x_i)_{i \in I}$. $\square$

Note 1

可以将 $e = (e_i)_{i \in I}$ 视作 $K^{\oplus I}$ 的一组"基"。

又注意到 $K^{\oplus I}$ 可以代表所有的自由左 K 模, 所以一个左 K 模同态完全由其在 "基" 上的表现所决定。这个证明与基是否有限无关。

Note 2

下面这个 theorem 2 是刚才 theorem 1 的一个简单推论

Let R be a commutative unitary ring, let I be a set and M be a R-module. Then any mapping f from I to M can be "extended" uniquely to a R-module morphism from $R^{\oplus I}$ to M.

$$\begin{tikzcd} I \arrow[r, "f"] \arrow[d, "i"'] & M \\ R^{\oplus I} \arrow[ur, dashed, "\exists! \varphi"] & \end{tikzcd}$$

使用这个 theorem 2 可以给出张量积的定义。

$$
\begin{tikzcd}
M\times N \arrow[r, "f\ \text{bilinear}"] \arrow[d, "i"'] \arrow[dd, bend right=60, "t\ R\text{-bilinear}"'] & P \\
R^{\oplus(M\times N)} \arrow[d, "\pi"'] \arrow[ur, dashed, "\exists!\varphi\ R\text{-linear}"] & \\
R^{\oplus(M\times N)}/F \arrow[uur, dashed, "\exists!\Psi\ R\text{-linear}"'] &
\end{tikzcd}
$$

Note 3

$\mathrm{Hom}(M^{\oplus I}, V) \cong \mathrm{Hom}(M, V)^I$, $\varphi \mapsto (\varphi \circ \iota_i)_{i\in I}$

单射和满射的证明思路与上文类似。

对于其几何意义暂无想法。

Remark

In the case where $I = \{1, 2, 3, ..., n\}$, V^I is denoted as V^n, K^I is denoted as K^n.

For any $(x_1, x_2, ..., x_n) \in V^n$, by the theorem, there exist a unique morphism of left K-module $\varphi : K^n \to V$ s.t. $\varphi(e_i) = x_i$ for any $i \in \{1, 2, 3, ...n\}$

We write this φ as a column $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$. It sends $(a_1, a_2, ..., a_n) \in K^n$ to $a_1x_1 + ... + a_nx_n$.

2.4.6 Def

Let K and K' be two unitary rings. We call (K, K')-bimodule any commutative group $(V, +)$ equipped with a left K-module structure $K\times V \to V$, $\quad (a, x) \longmapsto ax$ and a right K'-module structure $K'\times V \to V$, $\quad (b', x) \longmapsto xb'$, that satisfies the following condition $\forall (a, x, a') \in K \times V \times K'$, $\quad (ax)a' = a(xa')$.

Example

Let K be a unitary ring. The multiplicative law $K\times K \to K$, $\quad (a,b) \longmapsto ab$ defines a structure of (K,K)-bimodule on K because of the associativity.

Remark

Let K and K' be unitary rings, and V be a K, K'-bimodule. For any set I, the commutative group $(V^I,+)$ equipped with the left K-module structure $a(x_i)_{i\in I} := (ax_i)_{i\in I}$ and the right K'-module structure $(x_i)_{i\in I}b' := (x_ib')_{i\in I}$ forms a (K,K')-bimodule.

2.4.7 Prop

Let K and K' be unitary rings, E be a left K-module and F be a (K,K')-bimodule. Then the set $\mathrm{Hom}_{K-\mathrm{Mod}}(E,F)$ forms a right sub-K'-module of F^E.

Proof

We have seen that $\mathrm{Hom}_{K-\mathrm{Mod}}(E,F)$ forms a subgroup of F^E. To show that it is a right sub-K'-module of F^E, it suffices to show that, for any morphism of left K-module $\varphi : E \to F$ and any $b' \in K'$, the mapping $(x\in E) \mapsto \varphi(x)b'$ is also a morphism of left K-module. For $(x,y)\in E\times E$, one has $\varphi(x+y)b' = (\varphi(x)+\varphi(y))b' = \varphi(x)b' + \varphi(y)b'$; for $(a,x)\in K\times E$, one has (since F is a (K,K')-bimodule) $\varphi(ax)b' = (a\varphi(x))b' = a(\varphi(x)b')$. Therefore $(x\in E)\mapsto \varphi(x)b'$ is a morphism of left K-module.

2.4.8 Theorem

Let K and K' be unitary rings. Let I be a set and V be a (K,K')-bimodule. Then the mapping $\Psi : \mathrm{Hom}_{K-\mathrm{Mod}}(K^{\oplus I}, V) \to V^I, \varphi \longmapsto (\varphi(e_i))_{i\in I}$ is an isomorphism of right K'-modules.

Proof

We have seen that Ψ is an isomorphism of groups. It suffices to check that it is actually a morphism of right K'-modules. Let $\varphi : K^{\oplus I}\to V$ be a morphism of left K-modules, and $b'\in K'$, one has $\Psi(\varphi b') = (\varphi(e_i)b')_{i\in I} = (\varphi(e_i))_{i\in I}b'$. Hence Ψ is a morphism of right K'-modules.

2.5 Monotone mappings

2.5.1 Def

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

(1) If $\forall (a, b) \in I \times I$ such that $a < b$, one has $f(a) \leq f(b)$, then we say f is increasing.

If $\forall (a, b) \in I \times I$ such that $a < b$, one has $f(a) < f(b)$, then we say f is strictly increasing.

If $\forall (a, b) \in I \times I$ such that $a < b$, one has $f(a) \geq f(b)$, then we say f is decreasing.

If $\forall (a, b) \in I \times I$ such that $a < b$, one has $f(a) > f(b)$, then we say f is strictly decreasing.

(2) If f is increasing or decreasing, we say f is monotone, if f is strictly increasing or strictly decreasing, we say f is strictly monotone.

Example

Let X be a partially ordered set, $Id_x : X \to X$ is strictly increasing.

Moreover, $\forall A \subset X$, if we equipped A with the order relation, then the restriction of the mapping $\imath : A \to X$, $x \mapsto x$ is strictly increasing.

2.5.2 Prop

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

(1) If f and g are both (strictly) increasing/decreasing, then $g \circ f$ is (strictly) increasing/decreasing.

(2) If one of f and g is (strictly) increasing, the other is (strictly)decreasing, then $g \circ f$ is (strictly) decreasing.

Proof

(1) Suppose f and g are both increasing, let $(a, b) \in X \times X$ s.t. $a < b$, then $f(a) \leq f(b)$, if $f(a) = f(b)$, then $g(f(a)) = g(f(b)) \leq g(f(b))$, if $f(a) \leq f(b)$, then $g(f(a)) \leq g(f(b))$, hence $g \circ f$ is increasing.

Suppose f and g are both strictly increasing, let $(a, b) \in X \times X$ s.t. $a < b$, then $f(a) < f(b)$, therefore $g(f(a)) < g(f(b))$, hence $g \circ f$ is strictly increasing.

Similarly one can prove the rest situations. □

(2) Similar to (1). □

2.5.3 Def

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

2.5.4 Def

Let I and X be partially ordered set. f be a function from I to X.

(1) If f is increasing/decreasing and f is injective, then f is strictly increasing/decreasing.

(2) Assume that I is totally ordered and f is strictly monotone, then f is injective.

Proof

(1) Suppose that f is increasing. Let $(a, b) \in \mathrm{Dom}(f) \times \mathrm{Dom}(f)$, s.t. $a < b$, then $f(a) \leq f(b)$. If $f(a) = f(b)$, because f is injective, thus $a = b$, contradiction! Therefore, f is strictly increasing.

Similarly, one can prove the situation when f is decreasing. □

(2) Suppose that f is strictly increasing, let $(a, b) \in \mathrm{Dom}(f) \times \mathrm{Dom}(f)$ s.t. $f(a) = f(b)$. If $a \neq b$, because I is totally ordered, $a > b$ or $a < b$. If $a > b$, thus $f(a) > f(b)$, contradiction! Similarly, $a < b$ also has contradiction. Therefore, $a = b$, so f is injective.

Similarly, one can prove the case when f is strictly decreasing. □

2.5.5 Prop

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

Proof

Not loss generality, suppose f is increasing, let $(a, b) \in \mathrm{Dom}(f) \times \mathrm{Dom}(f)$, s.t. $f(a) < f(b)$. Because f is injective, so $a = f^{-1}(f(a))$ and

$b = f^{-1}(f(b))$. Since A is totally ordered, hence there are three situations: $a < b$, $a = b$ or $a > b$. Only $a < b$(which means $f^{-1}(f(a)) < f^{-1}(f(a)))$ satisfies $f(a) < f(b)$, so f^{-1} is increasing too. □

2.5.6 Def

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.

2.5.7 Prop

Let I be a subset of $\mathbb{N}$ which is infinite (infinite means I is not bounded). Then there is a unique increasing bijection $\lambda_I : \mathbb{N} \to I$.

Proof

Lemma

If $A \subseteq B$, then $\min A \geq \min B$

Proof

Since $\min A \in A \subseteq B$, so $\min A \in B$. By definition, $\min B \leq x(x \in B)$, therefore $\min B \leq \min A$. □

Existence:

We construct a function f by induction as follows.

Let $f(0) = \min I$.

Suppose that $f(0), f(1), f(2), ..., f(n)$ have constructed. Then we talk about $f(n+1) = \min(I \setminus \{f(0), ..., f(n)\})$. Since $I \setminus \{f(0), ..., f(n)\} \subseteq I \setminus \{f(0), ..., f(n-1)\}$, so $f(n+1) = \min(I \setminus \{f(0), ..., f(n)\}) \geq \min(I \setminus \{f(0), ..., f(n-1)\}) = f(n)$.

Because $f(n+1) \notin \{f(0), ..., f(n)\}$, we have $f(n+1) > f(n)$, therefore f is strictly increasing, and thus injective.

If f is not surjective, then $I \setminus \mathrm{Im}(f)$ has a least element N. Let $m = \min\{n \in \mathbb{N} | N \leq f(n)\}$.(Since f is injective. If its image is bounded from above, then $\mathbb{N}$ is finite, which leads to a contradiction.) Because $N \notin \mathrm{Im}(f)$, so $N < f(m)$, thus $m \neq 0$. Hence $f(m-1) < N < f(m) = \min(I \setminus \{f(0), ..., f(m-1)\})$. By definition, $N \in I \setminus \mathrm{Im}(f) \subseteq I \setminus \{f(0), ..., f(m-1)\}$, so $N \geq f(m)$, contradiction!

Uniqueness:

Let g be a increasing bijection that satisfies the same condition and is different from f. Hence the set $\{n \in \mathbb{N} | f(n) \neq g(n)\}$ has a least element N_0. Because g is surjective, there exists $N \in \mathbb{N}$, s.t. $g(N) = f(N_0) \neq f(N)$, so $N > N_0$, owing to g is increasing, thus $g(N) \geq g(N_0)$. But because g is injective, so there exist $M \in \mathbb{N}$ s.t. $f(M) = g(N_0) \neq g(M)$, therefore $M > N_0$. Since $M > N_0$, $f(M) > f(N_0)$, which is equivalent to $g(N_0) > g(N)$, contradiction! □

Hint of Uniqueness

Prove that $\mathrm{Id}_{\mathbb{N}} : \mathbb{N} \to \mathbb{N}$ is the only isomorphism of partially ordered sets from $\mathbb{N}$ to $\mathbb{N}$

Proof

Since $\mathrm{Id}_{\mathbb{N}}$ is a bijection and both $\mathrm{Id}_{\mathbb{N}}$ and $(\mathrm{Id}_{\mathbb{N}})^{-1}$ are increasing, so $\mathrm{Id}_{\mathbb{N}}$ is an isomorphism of partially ordered sets.

Assume that there exists g which satisfies the same condition and $g \neq \mathrm{Id}_{\mathbb{N}}$. Hence the set $\{n \in \mathbb{N} | g(n) \neq \mathrm{Id}_{\mathbb{N}}\}$ has a least element N_0, whose image by g is $N \in \mathbb{N}$. Because g is injective, so $g(N) \neq N$, which means $N_0 < N$. Since g is surjective, there exists $N_1 \in \mathbb{N}$, whose image is N_0. By the same reason, $N_0 < N_1$, because g is increasing, thus $g(N_0) \leq g(N_1)$, which is equivalent to $N \leq N_0$, contradiction! □

2.6 Sequences and series

In this section, we fix $I \subset \mathbb{N}$ be an infinite subset.

2.6.1 Def

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

Example 1

$X = \mathbb{R}$, $(n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ parametrized by $\mathbb{N}$.

Example 2

$(\frac{1}{n})_{n\in\mathbb{N}_{\geq 1}}$ is a sequence in $\mathbb{R}$ parametrized by $\mathbb{N}_{\geq 1}$.

Remark

If K is a unitary ring and E is a left K-module, then the set of sequences E^I admits a left K-module structure. If $x = (x_n)_{n\in I}$ is a sequence in E^I, 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.

2.6.2 Prop

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

Proof

Let $(x, y) \in E^I \times E^I$ and $\lambda \in K$

$\sum(x + y) = (\sum\limits_{i\in I, i\leq n} (x_i + y_i))_{n\in\mathbb{N}}$ when n is fixed, then it equals to $(\sum\limits_{i\in I, i\leq n} x_i)_{n\in\mathbb{N}} + (\sum\limits_{i\in I, i\leq n} y_i)_{n\in\mathbb{N}} = \sum(x) + \sum(y)$

$\sum(\lambda x) = (\sum\limits_{i\in I, i\leq n} \lambda x_i)_{n\in\mathbb{N}} = \lambda(\sum\limits_{i\in I, i\leq n} x_i)_{n\in\mathbb{N}} = \lambda \sum(x)$ □

2.6.3 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 doesn't have any upper bound in I.

(1) If f is increasing, then $f(I)$ and $f(J)$ have the same upper bound in X.

(2) If f is decreasing, then $f(I)$ and $f(J)$ have the same lower bound in X.

Proof

(1) Since $f(J) \subseteq f(I)$, any upper bound of $f(I)$ is an upper bound of $f(J)$.

Let M be an upper bound of $f(J)$, $\forall x \in I$, $\exists y_x \in J$, $M \geq f(y_x) \geq f(x)$, therefore, an upper bound of $f(J)$ is also an upper bound of $f(I)$. $\square$

(2) Similarly, one can prove this proposition. $\square$

2.6.4 Def

For any $(x_n)_{n\in I} \in [-\infty, +\infty]^I$. We define $\limsup\limits_{n\in I, n\to+\infty} x_n = \inf\limits_{n\in\mathbb{N}} (\sup\limits_{\substack{i\in I\\ i\geq n}} x_i)$ called the limit of superior, $\liminf\limits_{n\in I, n\to+\infty} x_n = \sup\limits_{n\in\mathbb{N}} (\inf\limits_{\substack{i\in I\\ i\geq n}} x_i)$ called the limit of inferior. If $\limsup\limits_{n\in I, n\to+\infty} x_n = \liminf\limits_{n\in I, n\to+\infty} x_n = l$, then we say that $(x_n)_{n\in I}$ tends to l, and l is the limit of $(x_n)_{n\in I}$. If in addition, $(x_n)_{n\in I} \in [-\infty, +\infty]^I$ and $l \in \mathbb{R}$, then we say $(x_n)_{n\in\mathbb{N}}$ converges to l.

Remark

Let $J \subseteq \mathbb{N}$ be an infinite set, then $\limsup\limits_{n\in I, n\to+\infty} x_n = \inf\limits_{n\in\mathbb{N}} (\sup\limits_{\substack{i\in I\\ i\geq n}} x_i) = \inf\limits_{n\in J} (\sup\limits_{\substack{i\in I\\ i\geq n}} x_i)$. Therefore, if you remove the first finite term of a sequence, the superior limit and the inferior limit don't change.

Note

注意到，收敛数列的所有子列都是收敛的。

2.6.5 Prop

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

Proof

For any $(n, m) \in \mathbb{N} \times \mathbb{N}$, $\inf\limits_{i\in I, i\geq n} (x_i) \leq x_{\max\{n,m\}} \leq \sup\limits_{i\in I, i\geq m} (x_i)$, thus $\liminf\limits_{n\in I, n\to+\infty} x_n = \inf\limits_{n\in\mathbb{N}} (\sup\limits_{\substack{i\in I\\ i\geq n}} x_i) = \inf\limits_{m\in\mathbb{N}} (\sup\limits_{\substack{i\in I\\ i\geq m}} x_i) \leq \limsup\limits_{n\in I, n\to+\infty} x_n$ $\square$

2.6.6 Prop

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

(1) $\forall c \in \mathbb{R}$, $\liminf\limits_{n\in I, n\to+\infty}(x_n + c) = (\liminf\limits_{n\in I, n\to+\infty} x_n) + c$, $\limsup\limits_{n\in I, n\to+\infty}(x_n + c) = (\limsup\limits_{n\in I, n\to+\infty} x_n) + c$

(2) $\forall \lambda \in \mathbb{R}_{\geq 0}$, $\liminf\limits_{n\in I, n\to+\infty}(\lambda x_n) = \lambda \liminf\limits_{n\in I, n\to+\infty} x_n$, $\limsup\limits_{n\in I, n\to+\infty}(\lambda x_n) = \lambda \limsup\limits_{n\in I, n\to+\infty} x_n$

(2) $\forall \lambda \in \mathbb{R}_{<0}$, $\liminf\limits_{n\in I, n\to+\infty}(\lambda x_n) = \lambda \limsup\limits_{n\in I, n\to+\infty} x_n$, $\limsup\limits_{n\in I, n\to+\infty}(\lambda x_n) = \lambda \liminf\limits_{n\in I, n\to+\infty} x_n$

Proof

In **1.3 Enhanced Real Line**, we have proved that

- For any $c \in \mathbb{R}$, any $A \in [-\infty, +\infty]$, $\inf(A+c) = \inf(A)+c$, $\sup(A+c) = \sup(A) + c$
- For any $\lambda \in \mathbb{R}_{\geq 0}$, any $A \in [-\infty, +\infty]$, $\inf(\lambda A) = \lambda \inf(A), \sup(\lambda A) = \lambda \sup(A)$
- For any $\lambda \in \mathbb{R}_{\leq 0}$, any $A \in [-\infty, +\infty]$, $\inf(\lambda A) = \lambda \sup(A), \sup(\lambda A) = \lambda \inf(A)$

Just using this conclusion, one can easily prove the proposition. □

2.6.7 Prop

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

Then

$\liminf\limits_{n\in I, n\to+\infty} x_n \leq \liminf\limits_{n\in I, n\to+\infty} y_n$

$\limsup\limits_{n\in I, n\to+\infty} x_n \leq \limsup\limits_{n\in I, n\to+\infty} y_n$

Proof

For any $n \in \mathbb{N}_{\geq N_0}$, since $x_n \leq y_n$, so $\sup\limits_{\substack{i\in I\\ i\geq n}} y_i$ is an upper bound of $(x_i)_{i\in I_{\geq n}}$, so $\sup\limits_{\substack{i\in I\\ i\geq n}} x_i \leq \sup\limits_{\substack{i\in I\\ i\geq n}} y_i$.

Repeat this step, one can obtain $\inf_{n\in\mathbb{N}_{\geq N_0}} (\sup_{\substack{i\in I\\ i\geq n}} x_i) \leq \inf_{n\in\mathbb{N}_{\geq N_0}} (\sup_{\substack{i\in I\\ i\geq n}} y_i)$

$\limsup_{n\in I, n\to+\infty} x_n = \inf_{n\in\mathbb{N}} (\sup_{\substack{i\in I\\ i\geq n}} x_i) = \inf_{n\in\mathbb{N}_{\geq N_0}} (\sup_{\substack{i\in I\\ i\geq n}} x_i) \leq \inf_{n\in\mathbb{N}_{\geq N_0}} (\sup_{\substack{i\in I\\ i\geq n}} y_i) = \inf_{n\in\mathbb{N}} (\sup_{\substack{i\in I\\ i\geq n}} y_i) = \limsup_{n\in I, n\to+\infty} y_n$

Similarly one can prove $\liminf_{n\in I, n\to+\infty} x_n \leq \liminf_{n\in I, n\to+\infty} y_n$ □

2.6.8 Theorem

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

Suppose that

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

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

Proof

$l = \limsup_{n\in I, n\to+\infty} z_n \geq \limsup_{n\in I, n\to+\infty} y_n \geq \liminf_{n\in I, n\to+\infty} y_n \geq \liminf_{n\in I, n\to+\infty} z_n = l$ □

2.6.9 Def

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.

2.6.10 Prop

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

$$\liminf_{n\in J, n\to+\infty} x_n \geq \liminf_{n\in I, n\to+\infty} x_n$$

$$\limsup_{n\in J, n\to+\infty} x_n \leq \limsup_{n\in I, n\to+\infty} x_n$$

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

Proof

In **1.1 Supremum and Infimum** we have proved that for any $n \in \mathbb{N}$, if $J \subseteq I$, then $\sup\limits_{\substack{i \in J \\ i \geq n}} x_i \leq \sup\limits_{\substack{i \in I \\ i \geq n}} x_i$, so $\limsup\limits_{n \in J, n \to +\infty} x_n \leq \limsup\limits_{n \in I, n \to +\infty} x_n$, similarly one can prove the other inequality. $\square$

2.6.11 Theorem

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

(1) If the mapping $(x_n)_{n \in I}$ is increasing, then $(x_n)_{n \in I}$ tends to $\sup\limits_{n \in I} x_n$

(1) If the mapping $(x_n)_{n \in I}$ is decreasing, then $(x_n)_{n \in I}$ tends to $\inf\limits_{n \in I} x_n$

Proof

(1)For any $n \in \mathbb{N}$, since $(x_n)_{n \in I}$ is increasing , so $\sup\limits_{\substack{i \in I \\ i \geq n}} x_i = \sup\limits_{i \in I} x_i$. Hence, $\limsup\limits_{n \in I, n \to +\infty} x_n = \inf\limits_{n \in \mathbb{N}} (\sup\limits_{i \in I} x_i) = \sup\limits_{i \in I} x_i$

Since it is increasing, $\inf\limits_{\substack{i \in I \\ i \geq n}} x_i = x_j$ (j is the minimum element in I that larger than n). Thus $\sup\limits_{n \in \mathbb{N}} (\inf\limits_{\substack{i \in I \\ i \geq n}} x_i) = \sup\limits_{n \in I} x_n$. $\square$

(2) It is similar to (1). $\square$

Note

这个 theorem 说明了单调有界的数列必定是收敛的。

Remark

Let $(x_n)_{n \in I}$ be a sequence in $\mathbb{R}_{\geq 0}$, then the series $\sum\limits_{n \in I} x_n$ (The sequence $(\sum\limits_{i \in I, i \leq n} x_i)_{n \in \mathbb{N}}$) tend to an element in $\mathbb{R} \bigcup \{+\infty\}$, it converges in $\mathbb{R}$ iff it is bounded.

Notation

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

2.6.12 Theorem(Bolzano-Weierstrass)

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 $\liminf\limits_{n \in I, n \to +\infty} x_n$ and $\limsup\limits_{n \in I, n \to +\infty} x_n$.

Proof

Not loss generality, let's prove the lim sup case. Let $J = \{n \in I|$ For any $m \in I$, if $m > n$, then $x_m \leq x_n\}$

If J is infinite, then $(x_n)_{n\in J}$ is the subsequence we need. $\limsup\limits_{n\in I, n\to+\infty} x_n = \inf\limits_{n\in\mathbb{N}}(\sup\limits_{\substack{i\in I\\ i\geq n}} x_i) = \inf\limits_{n\in J}(\sup\limits_{\substack{i\in I\\ i\geq n}} x_i) = \inf\limits_{n\in J} x_n$. And since $(x_n)_{n\in J}$ is decreasing, it tends to $\inf\limits_{n\in J} x_n$, so $(x_n)_{n\in J}$ tends to the limit of superior of $(x_n)_{n\in I}$.

If J is finite, then we try to make use of J to find a increasing infinite set, since there doesn't exist a "supremum"s decreasing infinite set. Let $n_0 \in I$, s.t. for any $n \in J$, $n < n_0$. Denote $\sup\limits_{n\in I, n\geq n_0} x_n$ by l.

Let $N \in \mathbb{N}$, s.t. $N \geq n_0$. By definition, $\sup\limits_{i\in I, l\geq N} x_i \leq l$. If the strict equality $\sup\limits_{i\in I, i\geq N} x_i < l$, holds. Then $\sup\limits_{i\in I, i\geq N} x_i$ is not an upper bound of $\{x_n | n_0 \leq n < N\}$. So there exists $m \in I$, s.t. $n_0 \leq m < N$ and $x_m > \sup\limits_{i\in I, i\geq N} x_i$. We may assume that m is the largest element that satisfies this condition. So $m \in J$, contradiction!

Hence, $\sup\limits_{i\in I, i\geq N} x_i = l$, which leads to $\limsup\limits_{n\in I, n\to+\infty} x_n = \inf\limits_{N\in\mathbb{N}, N\geq n_0}(\sup\limits_{\substack{i\in I\\ i\geq N}} x_i) = \inf\limits_{N\in\mathbb{N}, N\geq n_0} l = l$.

Then we try to construct an increasing infinite set who tends to l. Moreover, if $m \in I$ and $m \geq n_0$, then $x_m < l$. Because if $x_m = l$, then $x_m \in J$, contradiction! So for any $m \in I_{\geq n_0}$, there exists an element $t \in I$ which is greater than m and $x_t > x_m$. Along with this we can have our construction of an infinite increasing set.

We construct by induction an increasing sequence $(n_j)_{j\in\mathbb{N}}$ in I. n_0 be as above. Let $f : \mathbb{N} \to I_{\geq n_0}$ be a surgective mapping. If n_j is chosen, we choose $n_{j+1} \in I$, s.t. $n_{j+1} > n_j$ and $x_{n_{j+1}} > \max\{x_{f(j+1)}, x_{n_j}\}$. Therefore this sequence $(x_{n_j})_{j\in\mathbb{N}}$ is increasing, and $l = \sup\limits_{n\in I, n\geq n_0} x_n \geq \sup\limits_{j\in\mathbb{N}} x_{n_j} \geq \sup\limits_{n\in I, n\geq n_0} x_n = l$. □

Note

给出了比有界数列必有收敛子列强的多的结论。

推论是上极限就是最大的子列极限, 下极限就是最小的子列极限。

2.7 Cauchy sequences

2.7.1 Def

Let $(x_n)_{n\in I}$ be a sequence in $\mathbb{R}$.

If $\lim\limits_{N\to+\infty}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|)=\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|)=0$, then we say $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.

2.7.2 Prop

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

(2) If $(x_n)_{n\in I}$ is a Cauchy sequence, there exists $M\geq 0$, s.t. for any $n\in I$, $|x_n|\leq M$.

(3) If $(x_n)_{n\in I}$ is a Cauchy sequence, then for any infinite set $J\subseteq I$, $(x_n)_{n\in J}$ is also a Cauchy sequence.

(4) If $(x_n)_{n\in I}$ is a Cauchy sequence, and if there exists infinite set $J\subseteq I$ and $l\in\mathbb{R}$ s.t. $(x_n)_{n\in J}$ converges to l, then $(x_n)_{n\in I}$ converges to l, too.

Proof

(1) $0\leq\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|)=\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-l+l-x_m|)=\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-l|+|x_m-l|)=2\inf\limits_{n\in\mathbb{N}}(\sup\limits_{\substack{i\in I\\ i\geq n}}|x_i-l|)=0$ □

(2) Since $\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|)=0$, there exists $N\in I$, such that $\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|=l\in\mathbb{R}$, so for any $n\in I_{\geq N}$, $|x_n|=|x_n-x_N+x_N|\leq|x_n-x_N|+|x_N|=l+|x_N|$.

Let $M=\max\{\sup\limits_{n\in I_{<N}}x_n, l+|x_N|\}$ □

(3) $0\leq\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in J\times J\\ n\geq N,m\geq N}}|x_n-x_m|)\leq\inf\limits_{N\in\mathbb{N}}(\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|)=0$ □

(4) For any $N\in\mathbb{N}$, let $B_N=\sup\limits_{\substack{(n,m)\in I\times I\\ n\geq N,m\geq N}}|x_n-x_m|$, $C_N=\sup\limits_{\substack{n\in J\\ n\geq N}}|x_n-l|$

For any $n \in I$, $n \geq N$, for any $m \in J$, $m \geq N$, $|x_n - l| = |x_n - x_m + x_m - l| \leq |x_n - x_m| + |x_m - l| \leq B_N + C_N$, so $B_N + C_N$ is an upper bound of $|x_n - l|$, so $\sup\limits_{n\in I, n\geq N} |x_n - l| \leq B_N + C_N$.

Since $\lim\limits_{N\to+\infty} B_N = \lim\limits_{N\to+\infty} C_N = 0$, thus $\limsup\limits_{n\in I, n\to+\infty} |x_n - l| = 0$ □

2.7.3 Def

If any Cauchy sequence in X converges to a value in X, then we say X is complete.

2.7.4 Theorem(Completeness of $\mathbb{R}$)

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

Proof

Since $(x_n)_{n\in I}$ is a Cauchy sequence, there exists $M \in \mathbb{R}_{>0}$, s.t. $-M \leq (x_n)_{n\in I} \leq M$, so $\limsup\limits_{n\in I, n\to+\infty} x_n \in \mathbb{R}$.

By Bolzano-Weierstrass theorem, there exists an infinite set $J \subseteq I$, s.t. $(x_n)_{n\in J}$ converges to $\limsup\limits_{n\in I, n\to+\infty} x_n \in \mathbb{R}$, by previous conclusion $(x_n)_{n\in I}$ also converges to the same limit. □

2.7.5 Def

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

2.7.6 Prop

If a series $\sum\limits_{n\in I} x_n$ converges absolutely, then it converges in $\mathbb{R}$.

Proof

Since $\sum\limits_{n\in I} |x_n| < +\infty$, so $(\sum\limits_{i\in I, i\leq n} |x_i|)_{n\in\mathbb{N}}$ converges in $\mathbb{R}$, hence it is a Cauchy sequence.

For any $N \in \mathbb{N}$, for any $(n,m) \in I \times I$, $N \leq n \leq m$, $|\sum\limits_{i\in I, i\leq m} |x_i| - \sum\limits_{i\in I, i\leq n} |x_i|| = \sum\limits_{i\in I, n<i\leq m} |x_i| \geq |\sum\limits_{i\in I, n<i\leq m} x_i| = |\sum\limits_{i\in I, i\leq m} x_i - \sum\limits_{i\in I, i\leq n} x_i| \geq 0$

Hence $\sup\limits_{\substack{(n,m)\in I\times I \\ m\geq n\geq N}} |\sum\limits_{i\in I, i\leq m} x_i - \sum\limits_{i\in I, i\leq n} x_i| \leq \sup\limits_{\substack{(n,m)\in I\times I \\ m\geq n\geq N}} |\sum\limits_{i\in I, i\leq m} |x_i| - \sum\limits_{i\in I, i\leq n} |x_i||$

Taking $\lim\limits_{N \to +\infty}$, we obtain $(\sum\limits_{i \in I, i \leq n} x_i)_{n \in \mathbb{N}}$ is a Cauchy sequence. $\square$

2.8 Comparison and technics of computation

2.8.1 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 element of I has an upper bound in J. Then $f(I)$ and $f(J)$ have the same upper/lower bounds in X.

Proof

Since $J \subseteq I$, hence $f(J) \subseteq f(I)$, any upper bound of $f(I)$ is an upper bound of $f(J)$.

For any upper bound A of J, let $x \in I$, there exists an $y \in J$, s.t. $x \leq y$, hence $f(x) \leq f(y) \leq A$, so any upper bound of $f(J)$ is an upper bound of $f(I)$. $\square$

2.8.2 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 or both decreasing. Then the following qualities holds, provided that the sum on the RHS equality is well defined.

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

or

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

Proof

Let $a = \sup_{x\in I} f(x)$ and $b = \sup_{x\in I} g(x)$. Assume that f and g are both increasing.

Let $A = \{(x, y) \in I \times I | \{f(x), g(y)\} \neq \{-\infty, +\infty\}\}$, equip A with an order relation if $a \leq a'$ and $b \leq b'$, then $(a, b) \leq (a', b')$. Let $h : A \to [-\infty, +\infty]$, s.t. $h((x, y)) = f(x) + g(y)$, thus $h(x, y)$ is well-defined. And it is easy to check this mapping is increasing.

Let $B = A \bigcap \Delta I$, $B \subseteq A$. For any $(x, y) \in A$, not loss generality, let $x \leq y$, then if $\{f(y), g(y)\} \neq \{-\infty, +\infty\}$, there exists an upper bound $(y, y) \in B$.

If $\{f(y), g(y)\} = \{-\infty, +\infty\}$, since $x \leq y$ and f is increasing, so $f(y) = +\infty$, $g(y) = -\infty$, thus $a = +\infty$, because of we have guaranteed that RHS is well defined, so there exists $z \in I$, s.t. $g(z) > -\infty$. Because g is increasing, so $z \geq y$, and thus there exists $(z, z) \in B$, which is an upper bound of (x, y). $\square$

2.8.3 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. for any $n \in I$, $\{x_n, y_n\} \neq \{-\infty, +\infty\}$. Then the following inequality holds, provided that the sum on RHS is well defined.

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

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

Proof

Let $A_N = \sup_{n\in I, n\geq N} x_n$ and $B_N = \sup_{n\in I, n\geq N} y_n$, since $(A_N)_{N\in\mathbb{N}}$ and $(B_N)_{N\in\mathbb{N}}$ are both decreasing, by the theorem, $\inf_{N\in\mathbb{N}, \{A(N), B(N)\}\neq\{-\infty,+\infty\}} (A_N + B_N) = \inf_{N\in\mathbb{N}} A_N + \inf_{N\in\mathbb{N}} B_N$, if $\{\inf_{N\in\mathbb{N}} A_N, \inf_{N\in\mathbb{N}} B_N\} \neq \{-\infty, +\infty\}$.

For any $N \in \mathbb{N}$, and for any $n \in I$, s.t. $n \geq N$, $x_n \leq A_N$ and $y_n \leq B_N$, so $A_N + B_N$ is an upper bound of $(x_n + y_n)_{n\in I_{\geq N}}$, so $\sup_{n\in I, n\geq N} (x_n + y_n) \leq A_N + B_N$, if $A_N + B_N$ is defined. Hence $\inf_{N\in\mathbb{N}} (\sup_{n\in I, n\geq N} (x_n + y_n)) \leq \inf_{N\in\mathbb{N}, \{A(N), B(N)\}\neq\{-\infty,+\infty\}} (A_N + B_N) = \inf_{N\in\mathbb{N}} A_N + \inf_{N\in\mathbb{N}} B_N$. $\square$

Remark

If $\lim_{n\to+\infty} a_n = l \in \mathbb{R}$ and $\lim_{n\to+\infty} b_n = l' \in \mathbb{R}$, then $\lim_{n\to+\infty} (a_n + b_n)$ exists and equals to $l + l'$.

2.8.4 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. for any $n \in I$, $\{x_n, y_n\} \neq \{-\infty, +\infty\}$. Then the following inequality holds, provided that the sum on RHS is well defined.

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

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

Proof

Not loss generality, only need to prove one case.

If $\{\limsup_{n\in I, n\to+\infty}(x_n+y_n), \limsup_{n\in I, n\to+\infty}(-y_n)\} \neq \{-\infty, +\infty\}$, then, $\limsup_{n\in I, n\to+\infty} x_n = \limsup_{n\in I, n\to+\infty}(x_n+y_n-y_n) \leq \limsup_{n\in I, n\to+\infty}(x_n+y_n) + \limsup_{n\in I, n\to+\infty}(-y_n) = \limsup_{n\in I, n\to+\infty}(x_n+y_n) - \liminf_{n\in I, n\to+\infty}(y_n)$

If $\{\limsup_{n\in I, n\to+\infty}(x_n+y_n), \limsup_{n\in I, n\to+\infty}(-y_n)\} = \{-\infty, +\infty\}$, which means $\limsup_{n\in I, n\to+\infty}(x_n+y_n) = \liminf_{n\in I, n\to+\infty}(y_n) = +\infty$ or $-\infty$, since RHS is always well defined the inequality becomes $+\infty \geq +\infty$ or $-\infty \geq -\infty$, which always holds. □

2.8.5 Theorem

Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be elements of $[-\infty, +\infty]^I$. If for any $n \in I$, $y_n \in \mathbb{R}$ and $(y_n)_{n\in I}$ converges to some $l \in \mathbb{R}$, then $\limsup_{n\in I, n\to+\infty}(x_n+y_n) = \limsup_{n\in I, n\to+\infty} x_n + l$.

Proof

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

□

2.8.6 Prop

Let $(x_n)_{n\in I}$ and $(y_n)_{n\in I}$ be elements in $[-\infty, +\infty]^I$, then

$$\limsup_{n\in I, n\to+\infty}(\max\{x_n, y_n\}) = \max\{\limsup_{n\in I, n\to+\infty} x_n, \limsup_{n\in I, n\to+\infty} y_n\}$$

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

Proof

Since $\max\{x_n, y_n\} \geq x_n$ and $\max\{x_n, y_n\} \geq y_n$, so $\limsup_{n\in I, n\to+\infty}(\max\{x_n, y_n\}) \geq \limsup_{n\in I, n\to+\infty} x_n$ and $\limsup_{n\in I, n\to+\infty}(\max\{x_n, y_n\}) \geq y_n$, therefore, $\limsup_{n\in I, n\to+\infty}(\max\{x_n, y_n\}) \geq$

$\max\{\limsup\limits_{n\in I, n\to+\infty} x_n, \limsup\limits_{n\in I, n\to+\infty} y_n\}$

Because of **Bolzano-Weierstrass Theorem**, one can find an infinite set $J \subseteq I$, s.t. $\lim\limits_{n\in J, n\to+\infty}(\max\{x_n, y_n\}) = \limsup\limits_{n\in I, n\to+\infty}(\max\{x_n, y_n\})$.

Let $J_1 = \{n \in J | x_n \geq y_n\}$, $J_2 = \{n \in J | x_n < y_n\}$, either J_1 is infinite or J_2 is infinite.

Not loss generality, suppose that J_1 is infinite, hence

$\limsup\limits_{n\in I, n\to+\infty}(\max\{x_n, y_n\}) = \lim\limits_{n\in J, n\to+\infty}(\max\{x_n, y_n\}) = \lim\limits_{n\in J_1, n\to+\infty} x_n$
$\leq \limsup\limits_{n\in I, n\to+\infty} x_n \leq \max\{\limsup\limits_{n\in I, n\to+\infty} x_n, \limsup\limits_{n\in I, n\to+\infty} y_n\}$ □

2.8.7 Theorem

Let $(a_n)_{n\in I} \in \mathbb{R}^I$ and $l \in \mathbb{R}$. The following statements are equivalent.

- $\lim\limits_{n\in I, n\to+\infty} a_n = l$
- $\limsup\limits_{n\in I, n\to+\infty} |a_n - l| = 0$

Proof

Suppose $\lim\limits_{n\in I, n\to+\infty} a_n = l$, then $\limsup\limits_{n\in I, n\to+\infty} |a_n - l| = \limsup\limits_{n\in I, n\to+\infty} \max\{a_n - l, l - a_n\} = \max\{\limsup\limits_{n\in I, n\to+\infty}(a_n - l), \limsup\limits_{n\in I, n\to+\infty}(l - a_n)\} = \max\{0, 0\} = 0$

Suppose $\limsup\limits_{n\in I, n\to+\infty} |a_n - l| = 0$, then $\limsup\limits_{n\in I, n\to+\infty} a_n \leq l \leq \liminf\limits_{n\in I, n\to+\infty} a_n$. □

2.8.8 Def

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

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

Remark

A sequence $(a_n)_{n\in I}$ in $\mathbb{R}$ converges to l iff $a_n - l = o(1)$.

Suppose $(a_n)_{n\in I}$ converges to l, thus $\lim\limits_{n\in I, n\to+\infty} |a_n - l| = 0$, so there exists $(|a_n - l|)_{n\in I}$ which converges to 0 and $|a_n - l| \leq |a_n - l|$, for any $N_0 \in \mathbb{N}$, $n \in \mathbb{N}_{\geq N_0}$.

Suppose $a_n - l = o(1)$, which means $0 \leq \limsup\limits_{n\in I, n\to+\infty} |a_n - l| = \limsup\limits_{n\in I_{\geq N_0}, n\to+\infty} |a_n - l| \leq \limsup\limits_{n\in I_{\geq N_0}, n\to+\infty} |\varepsilon_n| = 0$ □

2.8.9 Prop

Let $(x_n)_{n\in I}$, $(a'_n)_{n\in I}$ and $(b_n)_{n\in I}$ be elements of $\mathbb{R}^I$.

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

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

Proof

(1)There exists $(M, M') \in \mathbb{N}_{>0} \times \mathbb{N}_{>0}$, for any n is sufficient large, $|a_n| \leq M|b_n|$ and $|a'_n| \leq M'|b_n|$, then $|\lambda a_n + \mu a'_n| \leq |\lambda a_n| + |\mu a'_n| \leq (|\lambda|M + |\mu|M')|b_n|$. □

(2) Similar to (1), just turn M and M' into $(\epsilon_n)_{n\in I}$ and $(\epsilon'_n)_{n\in I}$. □

2.8.10 Prop

Let $(a_n)_{n\in I}$ and $(b_n)_{n\in I}$ be two sequences in $\mathbb{R}$.

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

Proof

Since there exists a $(\epsilon_n)_{n\in I}$ which converges to 0 and when n sufficient large $|a_n| \leq |\epsilon_n||b_n|$, thus $(\epsilon_n)_{n\in I}$ is a Cauchy sequence , thus it is bounded, so there exists $M \in \mathbb{R}_{>0}$, s.t. for any $n \in I$, $|\epsilon_n| < M$.

When n is sufficient large, $|a_n| \leq |\epsilon_n||b_n| \leq M|b_n|$, thus $a_n = O(b_n)$. □

2.8.11 Prop

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

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

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

Proof

trivial. $\square$

2.8.12 Prop

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

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

Proof

trivial. $\square$

2.8.13 Prop

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

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

Proof

$a_n - l = o(1)$ and $b_n - l' = o(1)$, hence $a_n + b_n - l - l' = o(1) + o(1) = o(1)$

$a_n b_n - ll' = (a_n - l)b_n + (b_n - l')l = o(1)b_n + o(1)l = o(1)$. $\square$

2.8.14 Prop

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

Proof

When n is sufficient large,

Let $a = (1 + \epsilon)^2$

$a^n = (1 + \epsilon)^n \times (1 + \epsilon)^n \geq (1 + n\epsilon) \times (1 + n\epsilon) \geq n^2\epsilon^2$

Hence $n \le \frac{1}{n\epsilon^2} \times a^n$ $\square$

2.8.15 Corollary

Let $a > 1$ and $t \in \mathbb{R}_{\ge 0}$, then $n^t = o(a^n)$, $n \to +\infty$"

Proof

When n is sufficient large,

Let $d \in \mathbb{N}$, s.t. $d \ge t$, then $n^t = o(n^d)$.

Let $b = \sqrt[d]{a} > 1$, then $n^d = o(b^n \cdots b^n) = o(a^n)$

Therefore, $n^t = o(a^n)$. $\square$

2.8.16 Prop

Let $a \in \mathbb{R}$, then $a^n = o(n!)$, $n \to +\infty$

Proof

Let $N \in \mathbb{N}$, s.t. $N > a$, for any $n \in \mathbb{N}_{\ge N}$.

$0 \le |\frac{a^n}{n!}| = |\frac{a^N}{N!}||\frac{a^{n-N}}{(N+1)...(n)}| \le |\frac{a^N}{N!}||(\frac{a}{N})^{n-N}|$

Since $\frac{a}{N} < 1$, so when $n \to +\infty$, $\lim\limits_{n\to+\infty} |\frac{a^n}{n!}| = 0$, which means $a^n = o(n!)$. $\square$

2.8.17 Prop

$n! = o(n^n), n \to +\infty$

Proof

$0 \le \lim\limits_{n\to+\infty} \frac{n!}{n^n} \le \lim\limits_{n\to+\infty} \frac{1}{n} = 0$ $\square$

2.8.18 Prop

Let $(a_n)_{n\in I}$ and $(b_n)_{n\in I}$ be elements of $\mathbb{R}^I$.

If the series $\sum\limits_{n\in I} b_n$ converges absolutely and $a_n = O(b_n)$, $n \to +\infty$, then $\sum\limits_{n\in I} a_n$ converges absolutely.

Proof

Let N_0 be the element belongs to $\mathbb{N}$, s.t. for any $n \in \mathbb{N}$, $|a_n| \le |M||b_n|$.

$\sum\limits_{n\in I} |a_n| = \sum\limits_{i=0}^{N_0} |a_i| + \sum\limits_{n\in N_{>N_0}} |a_n| < \sum\limits_{i=0}^{N_0} |a_i| + M \sum\limits_{n\in N_{>N_0}} |b_n| < +\infty$ $\square$

2.8.19 Theorem(d'Alembert ratio test)

Let $(a_n)_{n\in\mathbb{N}} \in (\mathbb{R}-\{0\})^{\mathbb{N}}$

(1)If $\limsup\limits_{n\to+\infty} |\frac{a_{n+1}}{a_n}| < 1$, then $\sum\limits_{n\in\mathbb{N}} a_n$ converges absolutely.

(2)If $\liminf\limits_{n\to+\infty} |\frac{a_{n+1}}{a_n}| > 1$, then $\sum\limits_{n\in\mathbb{N}} a_n$ diverges.

Proof

(1) Let $\alpha \in \mathbb{R}$, s.t. $\limsup\limits_{n\to+\infty} |\frac{a_{n+1}}{a_n}| < \alpha < 1$. Hence α is not a lower bound of $(\sup\limits_{n\in\mathbb{N}_{\geq N}} |\frac{a_{n+1}}{a_n}|)_{N\in\mathbb{N}}$, so there exists $N \in \mathbb{N}$, s.t. for any $n \in \mathbb{N}_{\geq N}$, $\sup |\frac{a_{n+1}}{a_n}| < \alpha$.

Therefore, $a_n = O(\alpha^n)$, $n \to +\infty$. Because $\sum\limits_{n\in\mathbb{N}} \alpha^n = \frac{1}{1-\alpha} < +\infty$, $\sum\limits_{n\in\mathbb{N}} a_n$ converges absolutely □

Lemma

If a series $\sum\limits_{n\in\mathbb{N}} a_n$ in $\mathbb{R}$ converges, then $\lim\limits_{n\to+\infty} a_n = 0$

Proof

If $(\sum\limits_{n=0}^{N} a_n)_{N\in\mathbb{N}}$ converges to $l \in \mathbb{R}$, then $(\sum\limits_{n=0}^{N-1} a_n)_{N\in\mathbb{N}_{\geq 1}}$ also have the same limit.

$(a_N)_{N\in\mathbb{N}} = (\sum\limits_{n=0}^{N} a_n - \sum\limits_{n=0}^{N-1} a_n)_{N\in\mathbb{N}}$ converges to $l - l = 0$ □

(2) Let $\beta \in \mathbb{R}$, s.t. $\liminf\limits_{n\to+\infty} |\frac{a_{n+1}}{a_n}| > \beta > 1$. Hence β is not an upper bound of $(\inf\limits_{n\in\mathbb{N}_{\geq N}} |\frac{a_{n+1}}{a_n}|)_{N\in\mathbb{N}}$. So there exists $N \in \mathbb{N}$, s.t. for any $n \in \mathbb{N}_{\geq N}$, $\inf\limits_{n\in\mathbb{N}_{\geq N}} |\frac{a_{n+1}}{a_n}| > \beta$.

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

Therefore, $(a_n)_{n\in\mathbb{N}}$ not converges to 0. □

2.8.20 Theorem(Cauchy root test)

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

(1) If $\alpha < 1$, then $\sum\limits_{n\in\mathbb{N}} a_n$ converges absolutely.

(2) If $\alpha > 1$, then $\sum\limits_{n\in\mathbb{N}} a_n$ diverges.

Proof

(1) Let $\beta \in \mathbb{R}$, s.t. $\alpha < \beta < 1$. Hence β is not a lower bound of $(\sup_{n \geq \mathbb{N}_{\geq N}} |a_n|^{\frac{1}{n}})_{N \in \mathbb{N}}$. Thus there exists $N \in \mathbb{N}$, s.t. for any $n \in \mathbb{N}_{\geq N}$, $\sup |a_n|^{\frac{1}{n}} < \beta$. So $a_n = O(\beta^n)$, and $\beta < 1$, which means $\sum_{n \in \mathbb{N}} a_n$ converges absolutely. □

(2) Because $\limsup_{n \to +\infty} |a_n|^{\frac{1}{n}} > 1$, thus for any $N \in \mathbb{N}$, there exists $n \in \mathbb{N}_{\geq N}$, $|a_n|^{\frac{1}{n}} > 1$, since otherwise there exists $N \in \mathbb{N}$, for any $n \in \mathbb{N}_{\geq N}$, $|a_n|^{\frac{1}{n}} \leq 1$, contradiction! Therefore, $(|a_n|)_{n \in \mathbb{N}}$ can not converge at 0. □

Example

Let $t \in \mathbb{R}$. $\limsup_{n \in \mathbb{N}, n \to +\infty} \frac{\frac{t^{n+1}}{(n+1)!}}{\frac{t^n}{n!}} = \limsup_{n \in \mathbb{N}, n \to +\infty} \frac{t}{n+1} = 0$.

By **d'Alembert ratio test**, $\sum_{n \in \mathbb{N}} \frac{t^n}{n!}$ converges absolutely. We denote by $\exp(t)$ its limit. Sometimes $\exp(t)$ is also written as e^t.

Chapter 3

Topology

3.1 Absolute values

2.8.1 Def

Let K be a field. By absolute value on K, we mean a mapping.

$|.| : K \to \mathbb{R}_{\geq 0}, a \mapsto |a|$ that satisfies:

T1 For any $a \in K$, $|a| = 0$ iff $a = 0$

T2 For any $(a, b) \in K \times K$, $|a||b| = |ab|$

T3 For any $(a, b) \in K \times K$, $|a + b| \leq |a| + |b|$(triangle inequality)

Example 1

$K = \mathbb{R}$, $|a| = \max\{a, -a\}$

Example 2

$K = \mathbb{C}$, $|a + bi| = \sqrt{a^2 + b^2}$

Example 3

K any field. $|\cdot| : K \to \mathbb{R}_{\geq 0}$

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

Example 4

Let p be a prime number. For any number α belongs to $\mathbb{Q} - \{0\}$, there exists an integer $\mathrm{ord}_p(\alpha)$ s.t. α is of the form $p^{\mathrm{ord}_p(\alpha)}\frac{a}{b}$, where $a \in \mathbb{Z} - \{0\}$ and $b \in \mathbb{Z} - \{0\}$ and $p \nmid a$, $p \nmid b$.

$$|\cdot|_p : \mathbb{Q} \to \mathbb{R}_{\geq 0},\ \alpha \mapsto \begin{cases} 0 & \text{if } \alpha = 0, \\ p^{-\mathrm{ord}_p(\alpha)} & \text{if } \alpha \neq 0. \end{cases}$$

T2: Let $x = p^{\mathrm{ord}_p(x)}\frac{a}{b}$, $y = p^{\mathrm{ord}_p(y)}\frac{c}{d}$ and $p \nmid abcd$. $xy = p^{\mathrm{ord}_p(x)+\mathrm{ord}_p(y)}\frac{ac}{bd}$ because $p \nmid ac$, $p \nmid bd$, which means $\mathrm{ord}_p(ab) = \mathrm{ord}_p(a) + \mathrm{ord}_p(b)$.

T3: Let $A = \min\{\mathrm{ord}_p(x), \mathrm{ord}_p(y)\}$

$|x + y|_p = |\frac{p^{\mathrm{ord}_p(x)}ad + p^{\mathrm{ord}_p(y)}bc}{bd}|_p = |p^A \frac{p^{\mathrm{ord}_p(x)-A}ad + p^{\mathrm{ord}_p(y)-A}bc}{bd}|_p \leq p^{-A} = p^{\max\{-\mathrm{ord}_p(x), -\mathrm{ord}_p(y)\}} \leq p^{-\mathrm{ord}_p(x)} + p^{-\mathrm{ord}_p(y)} \leq |x|_p + |y|_p$

3.2 Quotient structure

3.2.1 Prop

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

For any $x \in X$ and any $y \in X$, $x \sim y$ iff $[x] = [y]$.

Proof

For any $z \in [x]$, by definition, $z \sim x$, since $x \sim y$, hence $z \sim y$, which means $[x] \subseteq [y]$. Because $y \sim x$, thus $[y] \subseteq [x]$.

Let $z \in [x]$, then $x \sim z$ since $[x] = [y]$, thus $z \sim y$, hence $x \sim y$ $\square$

3.2.2 Def

Let G be a group and X be a set.

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

Also, right action of G on X is any mapping $G \times X \to X$, $(g, x) \mapsto xg$ that satisfies for any $x \in X$, $x1 = x$ and for any $(g, h) \in G^2$, $x(gh) = (xg)h$.

Remark

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

Note

与 K 模类似，左作用与右作用的区别只在于结合律时作用的先后。放在左侧亦或是右侧只是形式。

3.2.3 Prop

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

Proof

$x = 1x$.

If $x \sim y$, which means $x = gy$ and $g \in G$, then $g^{-1}x = g^{-1}(gy) =$

$(g^{-1}g)y = 1y = y$.

If $x \sim y$ and $y \sim z$, then $x = g_1 y = g_1(g_2 z) = (g_1 g_2) z$. □

Notation

We denote by $G \setminus X$ the set $X/\sim$.

For any $x \in X$, the equivalence class of x is denoted as Gx or $\mathrm{orb}_G(x)$, called the orbit of x under the action of G.

3.2.4 Def

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

Example

Let G be a group and H be a subgroup of G. Then the mapping $H \times G \to G$, $(h, g) \mapsto hg$ is a left action of H on G, $(h, g) \mapsto gh$ is a right action of H on G. Then we obtain two quotient set $H \setminus G$ and G/H, respectively.

3.2.5 Def

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

Remark

For any $g \in G$, $gH = Hg$, provided that H is a normal subgroup of G.

If G is commutative, any subgroup H is normal.

3.2.6 Theorem

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

Proof

If $x'H = xH$, $y'H = yH$, then there exists $(h_1, h_2) \in H^2$, s.t. $x' = xh_1$ and $y' = yh_2$. For any $h \in H$, $(x'y')h = (xh_1yh_2)h = xh_1yh_2h =$

$x(yy^{-1})h_1yh_2h = xy(y^{-1}h_1y)(h_2h)$, since both $y^{-1}h_1y$ and h_2h belongs to H, thus $xyH = x'y'H$.

Associative: For any $(x, y, z) \in G^3$, $xHyHzH = (xy)HzH = (xyz)H = (x(yz))H = xH(yz)H = xH(yHzH)$.

Neutral element: $1HxH = xH1H = xH$.

Reversible: $xHx^{-1}H = x^{-1}HxH = 1H$

$\pi(xy) = (xy)H = xHyH = \pi(x)\pi(y)$. □

Notation

G/H is called the quotient group of G by H.

3.2.7 Def

Let K be a unitary ring and E be a left K-module. We say that a subgroup F of $(E, +)$ is a left sub-K-module of E if for any $x \in F$, $a \in K$, $ax \in F$.

3.2.8 Prop

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

Proof

Let $(x, x') \in E^2$, s.t. $[x] = [x']$. thus $(x - x') \in F$, hence $a(x - x') \in F$ and $a(x - x') = ax - ax'$, therefore, $[ax] = [ax']$.

Let $x \in E$, $1[x] = [1x] = [x]$.

Let $(x, y) \in E^2$, $a \in K$, $a[x+y] = [a(x+y)] = [ax+ay] = [ax]+[ay] = a[x] + a[y]$.

Let $(a, b) \in K^2$, $x \in E$, $(ab)[x] = [(ab)x] = [a(bx)] = a[bx] = a(b[x])$. $(a + b)[x] = [(a + b)x] = [ax + bx] = [ax] + [bx]$.

By the previous theorem, π is a morphism of groups (Since E is a commutative group thus F is a normal subgroup, which shows that E/F is a quotient group, therefore it is a morphism of groups.). Let $(a, x) \in K \times E$, $a\pi(x) = a[x] = [ax] = \pi(ax)$. $\square$

3.2.9 Def

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

3.2.10 Theorem

Let A be a unitary ring and I be a two sided ideal of A. The mapping $(A/I) \times (A/I) \to A/I$, $([a], [b]) \mapsto [ab]$ is well-defined. Moreover, A/I becomes a unitary ring under the addition and thus composition law, and the projection mapping $\pi : A \to A/I$ is a morphism of unitary rings.

Proof

Let $(x, x', y, y') \in A^4$, s.t. $x \sim x'$ and $y \sim y'$, thus $(x - x', y - y') \in I^2$. $xy - x'y' = xy - xy' + xy' - x'y' = x(y - y') + y'(x - x') \in I$.

Let $(a, b) \in A^2$, $\pi(a + b) = \pi(a) + \pi(b)$ and $\pi(ab) = \pi(a)\pi(b)$. $\pi(1) = [1]$. $\square$

Example

Let $d \in \mathbb{Z}$ and $d\mathbb{Z} = \{m | \exists n \in \mathbb{Z}, m = nd\}$. $d\mathbb{Z}$ is a two sided ideal of $\mathbb{Z}$.

Denote by $\mathbb{Z}/d\mathbb{Z}$ the quotient ring. The equivalence class of $n \in \mathbb{Z}$ is called the residua class of n modulo d.

Notation

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

3.2.11 Theorem

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

(1) $\mathrm{Im}(f)$ is a subgroup of H.

(2) $\mathrm{Ker}(f) = \{x \in G | f(x) = 1_H\}$ is a normal subgroup of G.

(3) The mapping $\tilde{f} : G/\mathrm{Ker}(f) \to \mathrm{Im}(f)$, $[x] \mapsto f(x)$ is well-defined and is an isomorphism of groups.

(4) f is injective iff Ker(f)=$\{1_G\}$

Proof

(1) Let $(x, y) \in G^2$, $f(x)f^{-1}(y) = f(xy^{-1}) \in \mathrm{Im}(f)$. □

(2) Let $(x, y) \in \mathrm{Ker}(f)$, $f(xy^{-1}) = f(x)f^{-1}(y) = 1_H$, let $(g, h) \in G{\times}\mathrm{Ker}(f)$, $f(ghg^{-1}) = f(g)f(h)f(g^{-1}) = f(g)f(g^{-1}) = 1_H$. □

(3) Let $[x] = [x']$, thus there exists $h \in \mathrm{Ker}(f)$, s.t. $x' = hx$. $\tilde{f}([x']) = f(hx) = f(h)f(x) = \tilde{f}([x])$.

$\tilde{f}([x][y]) = \tilde{f}([xy]) = f(xy) = f(x)f(y) = \tilde{f}(x)\tilde{f}(y)$

Injective Let $[x]$ and $[y]$ have the same image, which means $f(x) = f(y)$, thus $xy^{-1} \in \mathrm{Ker}(f)$, since $x = (xy^{-1})y$, $[x] = [y]$.

Surjective By definition, Im($\tilde{(f)}$)=Im(f). □

(4) Proved in **2.4 Vector Space** □

3.2.12 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) $\tilde{f} : E/\mathrm{Ker}(f) \to \mathrm{Im}(f)$, $[x] \mapsto f(x)$ is an isomorphism of left-K-module.

Proof

(1) Let $a \in K$, $x \in E$, $af(x) = f(ax) \in \mathrm{Im}(f)$. □

(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 Def

If there exists d a metric on X, s.t. $J_d = J$, then we say J is metrizable.

Example

Let X be a set. The discrete topology on X is metrizable, since there exists a metric

$$d(x,y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y. \end{cases}$$

, which shows that for any $x \in X$, $\{x\} = B(x,1) \in J_d$.

3.4 Axiom of choice

3.4.1 Axiom

For any set I and any family $(A_i)_{i\in I}$ of non-empty sets, there exists a mapping $f : I \to \bigcup_{i\in I} A_i$ s.t. $\forall i \in I$, $f(i) \in A_i$.

3.4.2 Theorem

For any set X, there exists an order relation $\leq$ on X, s.t. $(X, \leq)$ forms a well ordered set.

3.4.3 Theorem(Zorn's Lemma)

Let $(X, \leq)$ be a partially ordered set. If for any $A \subseteq X$, which is well ordered with respect to $\leq$, there exists an upper bound of A inside X. Then there exists a maximum element x_0 of X.

3.4.4 Prop

Let $(X, \leq)$ be a well ordered set and $y \notin X$. We extends $\leq$ to $X \bigcup y$ s.t. for any $x \in X$, $x < y$, then $(X \bigcup \{y\}, \leq)$ is well ordered.

Proof

Let A be a non-empty subset of $X \bigcup \{y\}$. If $A = \{y\}$, then y is its least element. If $A \neq \{y\}$, then $A \setminus \{y\}$ is not empty, let $x_0 \in X$ be $A \setminus \{y\}$'s least element, since $x_0 < y$, hence x_0 is still the least element of A. $\square$

3.4.5 Def

Let $(X, \leq)$ be a well ordered set and S be a subset of X. If for any $s \in S$ and $x \in X$, $x < s$ implies $x \in S$ ($X_{<s} \subseteq S$). Then we say that S is an initial segment of X.

If S is an initial segment of X and $S \neq X$, then we say that S is a proper initial segment of X.

Example

For any $x \in X$, $X_{<x} = \{s \in X | s < x\}$ is an initial segment of X.

3.4.6 Prop

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

Proof

For any $s \in \bigcup_{i\in I} S_i$, there exists $j \in I$, s.t. $s \in S_j$. Thus $X_{<s} \subseteq S_j \subseteq \bigcup_{i\in I} S_i$. $\square$

3.4.7 Prop

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

(1) Let S be a proper initial segment of X, $x = \min(X \setminus S)$, then $S = X_{<x}$.

(2) The mapping $X \to \mathscr{P}(X)$, $x \mapsto X_{<x}$ is a strictly increasing mapping.

(3) The set of all initial segments of X forms a well ordered subset of $(\mathscr{P}(X), \subseteq)$.

Proof

(1) For any $s \in S$, if $x < s$, $x \in S$, contradiction! If $x = s$, also contradiction! Hence $x > s$, $s \in X_{<x}$.

For any $y \in X_<x$, since x is the least element of $X \setminus S$ and $y < x$, thus $y \in S$. $\square$

(2) Let a and b be two elements of X, s.t. $a < b$, by definition, $X_{<a} \subseteq X_{<b}$ and since $a \in X_{<b} \setminus X_{<a}$, $X_{<a} \subsetneq X_{<b}$. $\square$

(3) Let J be the set of all the initial segments and $\mathcal{F}$ be a non-empty subset of J. Let $(I_i)_{i\in I}$ be a sequence of all the proper initial segments in $\mathcal{F}$, because of (1), we know every proper initial segment can be rewrite as the form $X_{<x}$, let $I_i = X_{<x_i}$ for each I_i belongs to $\mathcal{F}$. Let $\mathcal{Q} = \mathcal{F} - \{X\}$. If $\mathcal{Q} = \varnothing$, then $\mathcal{F} = X$, X is the least element of $\mathcal{F}$. If $\mathcal{Q} \neq \varnothing$, which means there exists a sequence $(x_i)_{i\in I}$, since the set $\{x_i | i \in I\}$ is a subset of X, it has a least element, denoted it as m. Using the conclusion of (2), $X_{<m}$ is the least element of $\mathcal{Q}$, which also means it is the least element of $\mathcal{F}$. $\square$

3.4.8 Lemma

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

Proof

Let $A = \{x \in X | x > f(x)\}$. Suppose A is not an empty set, thus there exists a least element $m \in X$, $f(m) < m$. Since f is strictly increasing, thus $f(f(m)) < f(m)$, contradiction! Hence A is empty. □

Note

不平凡的构造。

3.4.9 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 increasing, then $S = T$ and $f = \mathrm{Id}_S$.

Proof

Not loss generality, suppose $T \subseteq S$. (if $S \subseteq T$, use f^{-1} to do the following steps.) Let $\tau : T \to S$ be the inclusion mapping and $g = \tau \circ f$ be a strictly increasing mapping. For any $s \in S$, because of the lemma $s \leq g(s) = f(s) \in T$, hence $s \in T$, $T = S$.

f is strictly increasing and X is well ordered, thus f^{-1} is strictly increasing. Because of the lemma, $s \leq f^{-1}(s)$, and f is bijection, hence $f(s) \leq s$, $s = f(s)$. □

3.4.10 Def

Let $(X, \leq)$ and $(Y, \leq)$ be partially ordered set, we call X is isomorphic to Y if there exists f from X to Y, which is a isomorphism of partially ordered sets.

3.4.11 Def

Let $(X, \leq)$ and $(Y, \leq)$ be two well ordered sets, if $(X, \leq)$ is isomorphic to an initial segment of Y, we note $X \preceq Y$, if X is isomorphic to Y we note $X \sim Y$, if $X \preceq Y$ and $X \nsim Y$, we note $X \prec Y$ or $Y \succ X$.

3.4.12 Prop

Let X and Y be two well ordered sets. Among the following conditions, one and only one holds:

$$X \succ Y,\ X \sim Y,\ X \prec Y$$

Proof

Create a correspondence from X to Y, s.t. $(x, y) \in \Gamma_f$ iff $X_{<x} \sim Y_{<y}$. Let $(a, b) \in \mathrm{Dom}^2(f)$, s.t. $a < b$, then $X_{<a} \subsetneq X_{<b}$, $Y_{<f(b)} \sim X_{<b}$ and $Y_{<f(a)} \sim X_{<a}$, thus $Y_{<f(a)} \subseteq Y_{<f(b)}$(In fact, there exists an increasing bijection f s.t. $f(Y_{<f(a)})$ is a proper initial segement of $Y_{f(b)}$, from **3.4.9** we know $f(Y_{<f(a)}) = Y_{<f(a)}$, thus $Y_{<f(a)} \subsetneq Y_{<f(b)}$), which means $f(a) < f(b)$ since $f(a) \neq f(b)$ and if $f(a) > f(b)$, $f(b) \in Y_{<f(a)}$ and $f(b) \notin Y_{<f(b)}$, contradiction! Therefore, this function is strictly increasing.

Let $a \in \mathrm{Dom}(f)$. Let $x \in X$, $x < a$, then $X_{<x}$ is an initial segment of $X_{<a} \sim Y_{<f(a)}$, thus there exists $y < f(a)$, s.t. $X_{<x} \sim Y_{<y}$, this shows that $x \in \mathrm{Dom}(f)$, which means $\mathrm{Dom}(f)$ is an initial segment of X. Applying this to f^{-1}, we get $\mathrm{Im}(f)$=$\mathrm{Dom}(f^{-1})$ is an initial segment of Y.

Either $\mathrm{Dom}(f)$=X or $\mathrm{Im}(f)$=Y. Assume that $x \in X\backslash\mathrm{Dom}(f)$, $y \in Y\backslash\mathrm{Im}(f)$ are the least elements of $X\backslash\mathrm{Dom}(f)$ and $Y\backslash\mathrm{Im}(f)$, respectively. Then we get $\mathrm{Dom}(f)$=$X_{<x}$ and $\mathrm{Im}(f)$=$Y_{<y}$. f is an increasing bijection from $\mathrm{Dom}(f)$ to $\mathrm{Im}(f)$, hence $x \in$$\mathrm{Dom}(f)$, contradiction!

case1 $\mathrm{Dom}(f) = X$ $\mathrm{Im}(f)$$\neq Y$, $X \preceq Y$.

case2 $\mathrm{Dom}(f)$$\neq X$ $\mathrm{Im}(f)$$= Y$, $Y \preceq X$.

case3 $\mathrm{Dom}(f) =$ and $\mathrm{Im}(f) = Y$, $X \sim Y$. □

3.4.13 Lemma

Let $(X, \leq)$ be a partially ordered set. $\mathscr{S}$ be a subset of $\mathscr{P}(X)$.

Assume for any $A \in \mathscr{S}$, $(A, \leq)$ is a well ordered set. And for any $(A, B) \in \mathscr{S}^2$, either A is an initial segment of B or B is an initial segment of A.

Let $Y = \bigcup\limits_{A \in \mathscr{S}} A$, then $(Y, \leq)$ is well ordered and for any $A \in \mathscr{S}$, A is an initial segment of Y.

Proof

Let $A \in \mathscr{S}$, $x \in A$, $y \in Y$, s.t. $y < x$. Since $Y = \bigcup\limits_{B \in \mathscr{S}} B$, there exists $B \in \mathscr{S}$, s.t. $y \in B$. If $y \notin A$, then $B \nsubseteq A$, which means A is an initial segment of B, thus $y \in A$, contradiction!

Let $Z \subseteq Y$, $Z \neq \varnothing$, then there exists $A \in \mathscr{S}$, $Z \bigcap A \neq \varnothing$. Let m be the least element of $Z \bigcap A$. Let $z \in Z$, Let $B \in \mathscr{S}$, s.t. $z \in B$. If $z \in A$, then $m \leq z$, if $z \notin A$, which means A is an initial segment of B, thus $m < z$. Therefore, m is the least element of Z. $\square$

3.4.14 Theorem (Zorn's Lemma)

Let $(X, \leq)$ be a partially ordered set. If for any well ordered subset of X have an upper bound in X, then X has a maximum element. (Maximum means there doesn't exist an element larger than it but it may not larger than any other element.)

Proof

Suppose that X doesn't have any maximum element. Let W ={well ordered subsets of X}, let $f : W \to X$, s.t. for any $A \in W$, $f(A)$ is an upper bound of A(**Axiom of choice**) and $\forall x \in A$, $x < f(A)$.

If $A \in W$ satisfies that for any $a \in A$, $a = f(A_{<a})$, we say that A is a f-set.

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

Let A and B be elements of $\mathscr{S}$. Let I be the union of all common intial segments of A and B, which is also a common initial segment. If $I \neq A$ and $I \neq B$, then ther 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, $Y = \bigcup\limits_{A \in \mathscr{S}} A$ is well ordered and for any $A \in \mathscr{S}$, A is an initial segment of Y.

For any $a \in Y$, there 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, $Y \bigcup f(Y) \in \mathscr{S}$, thus $f(Y) \in Y$, contradiction! $\square$

3.5 Filter

3.5.1 Def

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

(1) $\varnothing \neq \mathscr{F}$ and $\varnothing \notin \mathscr{F}$.

(2) For any $A \in \mathscr{F}$, for any $B \in \mathscr{P}(X)$, if $A \subseteq B$, then $B \in \mathscr{F}$.

(3) For any $(A, B) \in \mathscr{F}$, $A \bigcap B \in \mathscr{F}$.

Example

(1) Let Y be a non-empty subset of X, $\mathscr{F}_Y = \{A \in \mathscr{P}(X) | Y \subseteq A\}$ is a filter, called the principal filter of Y.

(2) Let X be an infinite set.

$\mathscr{F}_{Fr}(X) = \{A \in \mathscr{P}(X) | X \setminus A \text{ is finite}\}$ is a filter called the Frechet filter of X.

(3) Let (X, τ) be a topological space and $x \in X$. We call neighborhood of x any $V \subseteq X$, s.t. there exists an open set U, $x \in U \subseteq V$.

$V_x = \{\text{neighborhoods of } x\ \}$ is a filter.

3.5.2 Def

Let X be a set. $\mathscr{B}$ is a non-empty subset of $\mathscr{P}(X)$. If $\varnothing \notin \mathscr{B}$ and for any $(B_1, B_2) \in \mathscr{B}^2$, there exists $B \in \mathscr{B}$, s.t. $B \subseteq B_1 \bigcap B_2$. We say that $\mathscr{B}$ is a filter basis.

Remark

If $\mathscr{B}$ is a filter basis, then $\mathscr{F}(\mathscr{B}) = \{A \subseteq X | \exists B \in \mathscr{B}, B \subseteq A\}$, is a filter, called the filter generated by $\mathscr{B}$.

In fact, $\varnothing \notin \mathscr{F}(\mathscr{B})$ and $\varnothing \neq \mathscr{F}(\mathscr{B})$ since $\varnothing \neq \mathscr{B} \subseteq \mathscr{F}(\mathscr{B})$.

For any $A \in \mathscr{F}(\mathscr{B})$, there exists $B \in \mathscr{B}$, s.t. $B \subseteq A$, for any subset C of X, s.t.$C \supseteq A \supseteq B$, hence $C \in \mathscr{F}(\mathscr{B})$.

For any $(A, B) \in \mathscr{F}^2(\mathscr{B})$, there exists $(C, D) \in \mathscr{B}^2$, s.t. $C \subseteq A$, $D \subseteq B$, there exists $E \in \mathscr{B}$, s.t. $E \subseteq (C \bigcap D) \subseteq (A \bigcap B)$, hence $(A \bigcap B) \in \mathscr{F}(\mathscr{B})$.

Example

(1) Let $Y \subseteq X$, $Y \neq \varnothing$, $\mathscr{B} = \{Y\}$ is a filter basis. $\mathscr{F}(\mathscr{B}) = \{A \in \mathscr{P}(X) | Y \subseteq A\}$

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

(3) Let (X, d) be a metric space, $x \in X$.

For any $\varepsilon > 0$, let $B(x, \varepsilon) = \{y \in X | d(x, y) < \varepsilon\}$ and $\overline{B}(x, \varepsilon) = \{y \in X | d(x, y) \leq \varepsilon\}$.

Then $\{B(x, \varepsilon) | \varepsilon > 0\}$ is a neighborhood basis of x.

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

$\{\overline{B}(x, \varepsilon) | \varepsilon > 0\}$ is a neighborhood basis of x.

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

Remark

$\mathscr{B}_x$ is a neighborhood basis if and only if $\mathscr{B}_x \subseteq V_x$ and for any $V \in V_x$ there exists $U \in \mathscr{B}_x$, s.t. $U \subseteq V$.

For example, $V_x \bigcap \tau$ is a neighborhood basis of x.

Remark

Let (X, τ) be a topological space, $x \in X$ and B_x is a neighborhood basis of x. Suppose that B_x is countable, we choose a surjective mapping $(B_n)_{n \in \mathbb{N}}$ from $\mathbb{N}$ to B_x. For any $n \in \mathbb{N}$, $A_n = B_0 \bigcap B_1 \cdots B_{n-1} \bigcap B_n \in V_x$. The sequence $(A_n)_{n \in \mathbb{N}}$ is decreasing and $\{A_n | n \in \mathbb{N}\}$ is a neighborhood basis of x.

3.5.3 Prop

$\mathscr{P}(\mathbb{N})$ is not countable.

Proof

Suppose there exists an surjective mapping f from $\mathbb{N}$ to $\mathscr{P}(\mathbb{N})$. Since it is surjective, there exists $n \in \mathbb{N}$, s.t. $f(n) = \{x \in \mathbb{N} | x \notin f(x)\}$, if $n \in f(n)$, then $n \notin f(n)$, if $n \notin f(n)$, then $n \in f(n)$, contradiction!

3.5.4 Prop

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

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

Proof

(1) $E \in g_*(\mathscr{F})$ and $\varnothing \notin g_*(\mathscr{F})$ since $g^{-1}(E) = Y \in \mathscr{F}$ and $g^{-1}(\varnothing) = \varnothing \notin \mathscr{F}$.

Let $A \in g_*(\mathscr{F})$ and $B \subseteq E$, s.t. $A \subseteq B$, then $g^{-1}(A) \subseteq g^{-1}(B)$, thus $B \in g_*(\mathscr{F})$.

Let $(A, B) \in g_*^2(\mathscr{F})$, since g is a mapping $g^{-1}(A \bigcap B) = (g^{-1}(A) \bigcap g^{-1}(B)) \in \mathscr{F}$.

(2) Since g is a mapping and $\varnothing \notin \mathscr{B}$, hence $\varnothing \notin g(\mathscr{B})$ and since $\mathscr{B} \neq \varnothing$, we get $g(\mathscr{B}) \neq \varnothing$.

For any $(A, B) \in \mathscr{B}^2$, there exists $C \in \mathscr{B}$, s.t. $C \subseteq (A \bigcap B)$, hence $g(C) \subseteq g(A \bigcap B) \subseteq g(A) \bigcap g(B)$.

$g(B) \subseteq A$ if and only if $B \subseteq g^{-1}(A)$, therefore, $\mathscr{F}(g(\mathscr{B})) = \{A \subseteq E|\exists B \in \mathscr{B}, g(B) \subseteq A\} = \{A \subseteq E|\exists B \in \mathscr{B}, B \subseteq g^{-1}(A)\}$. $\square$

3.6 Limit point and accumulation point

We fix a topological space (X, τ).

3.6.1 Def

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

(1) If $V_x \subseteq \mathscr{F}$, then we say x is a limit point of $\mathscr{F}$.

(2) If for any $(A, V) \in \mathscr{F} \times V_x$, $A \bigcap V \neq \varnothing$, then we say x is an accumulation point of $\mathscr{F}$.

Remark

Any limit point of $\mathscr{F}$ is an accumulation point of $\mathscr{F}$.

3.6.2 Prop

Let $\mathscr{B}$ be a filter basis of X, $x \in X$, $\mathscr{B}_x$ is a neighborhood basis of X. Then x is an accumulation point of $\mathscr{F}(\mathscr{B})$ iff for any $(B, U) \in \mathscr{B} \times \mathscr{B}_x$, $B \bigcap U \neq \varnothing$.

Proof

Since $\mathscr{B} \subseteq \mathscr{F}(\mathscr{B})$ and $\mathscr{B}_x \subseteq V_x$, the **Necessity** is true.

Sufficiency: Let $(A, V) \in \mathscr{F}(B) \times V_x$, there exists $(B, U) \in \mathscr{B} \times \mathscr{B}_x$, s.t. $B \subseteq A$ and $U \subseteq V$, hence $\varnothing \neq B \bigcap U \subseteq A \bigcap V$. □

3.6.3 Def

Let $Y \subseteq X$, $Y \neq \varnothing$. We call accumulation point of Y any accumulation point of the principal filter $\mathscr{F}_Y$. We denote by $\overline{Y} = \{$accumulation points of $Y\}$. Note that $x \in \overline{Y}$ iff for any $B \in \mathscr{B}_x$, $B \bigcap Y \neq \varnothing$.

By convention, $\overline{\varnothing} = \varnothing$.

3.6.4 Prop

Let $Y \subseteq X$, then $\overline{Y}$ is the smallest closed subset of X containing Y.

Proof

For any $x \in X \setminus \overline{Y}$, there exists $U_x \in V_x \bigcap \tau$, s.t. $U_x \bigcap Y = \varnothing$.

Moreover, for any $y \in U_x$, $U_x \in V_y \bigcap \tau$. This shows that for any $y \in U_x$,

$y \notin \overline{Y}$. Therefore, $X \setminus \overline{Y} = \bigcup_{x \in X \setminus \overline{Y}} U_x \in \tau$.

(**Own**)Let $Y \subseteq A$ be an closed set. For any $x \in X \setminus A$, $(X \setminus A) \bigcap Y = \varnothing$, thus $x \notin \overline{Y}$

(**Chen Huayi**) Let $Z \subseteq X$ be a closed set contains Y. Suppose that there exists $y \in \overline{Y} \setminus Z$, then $U = (X \setminus Z) \in V_y \bigcap \tau$ and $U \bigcap Y = \varnothing$, hence $y \notin \overline{Y}$, contradiction! $\square$

3.7 Limits of mappings

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

Remark

Let $\mathscr{B}_a$ be a neighborhood basis of a. Then $V_a \subseteq f_*(\mathscr{F})$ iff $B_a \subseteq f_*(\mathscr{F})$.

Therefore, a is a limit of $\mathscr{F}$ iff for any $B \in \mathscr{B}_a$, $f^{-1}(B) \in \mathscr{F}$.

Example

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 $\mathscr{F}_{Fr}(I)$ 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 $a = \lim\limits_{n \in I, n \to +\infty} x_n$.

Remark

$a = \lim\limits_{n \in I, n \to +\infty} x_n$ 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.

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

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 of $\mathbb{R}$.

3.7.1 Prop

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) $\limsup\limits_{n\in I, n\to+\infty} x_n = \liminf\limits_{n\in I, n\to+\infty} x_n = l.$

(3) $\limsup\limits_{n\in I, n\to+\infty} |x_n - l| = 0$

Proof

To prove this proposition, we introduce a more powerful proposition.

3.7.2 Theorem

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) $\lim\limits_{n\in I, n\to+\infty} x_n = l$

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

Proof

(1) $\Rightarrow$ (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 $\sup\limits_{n\in I_{\geq N}} d(x_n, l) \leq \varepsilon$, therefore, $\limsup\limits_{n\in I, n\to+\infty} d(x_n, l) \leq \varepsilon$, we then obtain $\limsup\limits_{n\in I, n\to+\infty} d(x_n, l) = 0$.

(2) $\Rightarrow$ (1). If $\inf\limits_{N\in\mathbb{N}} \sup\limits_{n\in I_{\geq N}} d(x_n, l) = 0$, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$, s.t. $\sup\limits_{i\in I_{\geq N}} d(x_n, l) < \varepsilon$. Hence $\forall n \in I_{\geq N}$, $d(x_n, l) < \varepsilon$, since ε is arbitrary, (1) is proved. $\square$

3.7.3 Prop

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

Proof

Y is not empty, since otherwise $\overline{Y} = \varnothing$.

$X \bigcap Y = Y \in \mathcal{V}_{p,Y}$ and $\varnothing \notin \mathcal{V}_{p,Y}$ since $p \in \overline{Y}$.

Let $V \in \mathcal{V}_p$ and $W \subseteq Y$, s.t. $V \bigcap Y \subseteq W$. Let $U = V \bigcup (W \backslash (V \bigcap Y)) \in \mathcal{V}_p$, and $U \bigcap Y = W$, hence $W \in \mathcal{V}_{p,Y}$.

Let $(V_1, V_2) \in \mathcal{V}_p^2$, $(V_1 \bigcap Y) \bigcap (V_2 \bigcap Y) = (V_1 \bigcap V_2) \bigcap Y \in \mathcal{V}_{p,Y}$. $\square$

3.7.4 Def

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

Remark

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

3.7.5 Theorem(Heine)

Let (X, τ_X) and (E, τ_E) be topological spaces. $Y \subseteq X$, $p \in \overline{Y} \setminus Y$, $a \in E$. We consider the following conditions.

(i) $a = \lim\limits_{y \in Y, y \to p} f(y)$

(ii) For any $(y_n)_{n \in \mathbb{N}} \in Y^{\mathbb{N}}$, if $\lim\limits_{n \to +\infty} y_n = p$, then $\lim\limits_{n \to +\infty} f(y_n) = a$

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.

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, $V_{p,Y} \subseteq y_*(\mathscr{F}_{Fr}(\mathbb{N}))$. We then get $f_*(V_{p,Y}) \subseteq f_*(y_*(\mathscr{F}_{Fr}(\mathbb{N}))) = (f \circ y)_*(\mathscr{F}_{Fr}(\mathbb{N}))$. (In fact, $f_*(y_*(\mathscr{F}_{Fr}(\mathbb{N}))) = \{F \subseteq E | f^{-1}(F) \subseteq y_*(\mathscr{F}_{Fr}(\mathbb{N}))\} = \{F \subseteq E | y^{-1}(f^{-1}(F)) \subseteq \mathscr{F}_{Fr}(\mathbb{N})\} = \{F \subseteq E | (f \circ y)^{-1}(F)) \subseteq \mathscr{F}_{Fr}(\mathbb{N})\} = (f \circ y)_*(\mathscr{F}_{Fr}(\mathbb{N})))$

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

(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.

Assume that (1) doesn't hold. Then there exists $U \in V_a$, s.t. for any $n \in \mathbb{N}$, $V_n \bigcap Y \nsubseteq f^{-1}(U)$. Take an arbitrary $y_n \in (V_n \bigcap Y) \setminus f^{-1}(U)$. Therefore, $\lim\limits_{n \to +\infty} y_n = p$. In fact, for any $W \in V_p$, there exists $N \in \mathbb{N}$, s.t. $V_N \subseteq W$. Hence $y_n \in W$ for any $\mathbb{N}_{\geq N}$. However $f(y_n) \notin U$ for any $n \in \mathbb{N}$, so $(f(y_n))_{n \in \mathbb{N}}$ can not converges to a. □

3.8 Continuity

3.8.1 Def

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

Remark

Let $\mathscr{B}_{f(x)}$ be a neighborhood basis of $f(x)$, then if any $B \in \mathscr{B}_{f(x)}$, there exists $V \in V_x$, s.t. $f(V) \subseteq B$, then f is continuous.

Suppose X and Y be metric spaces. Then f is continuous at x iff $\forall \varepsilon > 0$, $\exists \delta > 0$, s.t. if $d(x, y) < \delta$, $\rho(f(x), f(y)) < \varepsilon$.

3.8.2 Prop

Let (X, τ_X), (Y, τ_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.

Proof

Proof followed by definition. □

3.8.3 Theorem

Let (X, τ_X) and (Y, τ_Y) be two topological spaces. f be a function from X to Y, $x \in \mathrm{Dom}(f)$. Consider the following conditions.

- (i) f is continuous at x.
- (ii) $\forall (x_n) \in$Dom(f)$^{\mathbb{N}}$, if $\lim\limits_{n \to +\infty} x_n = x$, then $\lim\limits_{n \to +\infty} f(x_n) = f(x)$

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

Proof

For any $U \in V_{f(x)}$, there exists $V \in V_x$, s.t. $f(V) \subseteq U$, there also exists $N \in \mathbb{N}$, for any $n \in \mathbb{N}_{\geq N}$, $x_n \in V$, hence $f(x_n) \in f(V) \subseteq U$.

One can construct a neighborhood basis $\{V_n|n \in \mathbb{N}, V_{n+1} \subseteq V_n\}$. Suppose f is not continuous at x, then there exists a neighborhood $U \in V_{f(x)}$, for any $n \in \mathbb{N}$, $f(V_n) \nsubseteq U$. Fetch $y_n \in V_n$ and $f(y) \notin U$, then y_n converges to x, but $f(y_n)$ not converges to $f(x)$, since there exists $U \in V_{f(x)}$, s.t. $f(y_n) \notin U$ for any $n \in \mathbb{N}$. □

3.8.4 Def

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

3.8.5 Prop

Let (X, τ) be a topological space, $\mathscr{B} \subseteq \tau$. $\mathscr{B}$ 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.

Proof

Let U be an open set. For any $x \in U$, $\mathscr{B}_x$ is a neighborhood basis of x, thus there exists $W_x \subseteq U$, by definition, $\bigcup\limits_{x \in U} W_x = U$.

Own: For any $x \in X$, let V be a neighborhood of x, hence there exists an open set $U \ni x$ which is a subset of V. By definition, there exists some members of $\mathscr{B}$, s.t. the union of these members equal to U, therefore, there exists $B \in \mathscr{B}$, s.t. $x \in B \subseteq U$, because $\mathscr{B} \subseteq \tau$, hence $\mathscr{B}_x \subseteq V_x$. Therefore $\mathscr{B}_x$ is a neighborhood basis of x.

CHEN Huayi: $\mathscr{B} \neq \varnothing$, since $\exists V \in \mathscr{B}$, $x \in V$, $\varnothing \notin \mathscr{B}$, since $\forall V \in \mathscr{B}$, $x \in V$.

$\forall V_1$ and V_2 in $\mathscr{B}$, $V_1 \bigcap V_2$ is a element of τ containing x, so $\exists V_3 \in \mathscr{B}$, $x \in V_3 \subseteq V_1 \bigcap V_2$.

$\mathscr{F}(\mathscr{B}_x) = \{U \in \mathscr{P}(X)|\exists V \in \mathscr{B}_x, V \subseteq U\}$, hence $V_x \subseteq \mathscr{F}(\mathscr{B}_x)$, moreover, $\mathscr{B}_x \subseteq V_x$, hence $\mathscr{F}(\mathscr{B}_x) \subseteq V_x$, so $\mathscr{F}(\mathscr{B}_x) = V_x$ □

3.8.6 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:

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

Proof

When f is a mapping $f(U) \subseteq V$ is equivalent to $U \subseteq f^{-1}(V)$.

Own: Let $x \in X$ and $U \in V_{f(x)} \bigcap \tau_X$, since f is continuous, there exists $W_x \in V_x \bigcap \tau_X$, s.t. $W_x \subseteq f^{-1}(U)$, for any $y \in f^{-1}(U)$, there exists $W_y \subseteq f^{-1}(U)$, hence $\bigcup\limits_{x \in f^{-1}(U)} W_x = f^{-1}(U)$. If an open set in Y denoted as D doesn't have any inverse image, then $f^{-1}(D) = \varnothing \in \tau_X$. Therefore, for any $U \in \tau_Y$, $f^{-1}(U) \in \tau_X$.

CHEN Huayi: Let $U \in \tau_Y$. $\forall x \in f^{-1}(U)$, $f(x) \in U$. Hence $U \in V_{f(x)}$, so there exists an open neighborhood W of x, s.t. $f(W) \subseteq U$. Since f is a mapping, $W \subseteq f^{-1}(U)$, therefore, $f^{-1}(U) \in V_x$. Since x is arbitrary $f^{-1}(U) \in \tau_X$. (In fact, if $V \in \mathscr{P}(X)$ then $V \in \tau$ iff $\forall x \in V$, V is a neighborhood of x.)

Let $x \in X$ and $V \in V_{f(x)}$, then there exists $U \in B_Y$, s.t. $U \subseteq V$, by definition $f^{-1}(U) \subseteq f^{-1}(V)$. □

3.8.7 Def

Let X be a set, $((Y_i, \tau_i))_{i \in I}$ 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.

Remark

If τ is the initial topology of $(f_i)_{i \in I}$, then $\forall i \in I$, $\forall U_i \in \tau_i$, $f_i^{-1}(U_i) \in \tau$.

If $J \subseteq I$ is a finite subset, $(U_j)_{j \in J} \in \prod\limits_{j \in J} \tau_j$, then $\bigcap\limits_{j \in J} f_j^{-1}(U_j) \in \tau$.

3.8.8 Prop

$\mathscr{B} = \{\bigcap\limits_{j \in J} f_j^{-1}(U_j) | J \subseteq I$ finite and $(U_j)_{j \in J} \in \prod\limits_{j \in J} \tau_j\}$ is a topological basis of the initial topology τ.

Proof

Let $\tau' = \{$ unions of members in $\mathscr{B}\}$. What we need to prove is that τ' is a topology and $\tau' = \tau$.

$\varnothing = \bigcup\limits_{i \in \varnothing, U_i \in \mathscr{B}} U_i \in \tau'$ and $X \in \tau'$. The union of "unions of members in $\mathscr{B}$" is still unions of members in $\mathscr{B}$.

Let $\bigcap\limits_{j \in J} f_j^{-1}(U_j)$ and $\bigcap\limits_{k \in K} f_j^{-1}(U_k)$ be two elements of τ'.

Then $(\bigcap\limits_{j \in J} f_j^{-1}(U_j)) \bigcap (\bigcap\limits_{k \in K} f_j^{-1}(U_k))$
$= (\bigcap\limits_{j \in J \setminus K} f_j^{-1}(U_j)) \bigcup (\bigcap\limits_{j \in J \bigcap K} f_j^{-1}(U_j \bigcap U_k)) \bigcup (\bigcap\limits_{k \in K \setminus J} f_k^{-1}(U_k))$

Notice that $\mathscr{B} \subseteq \tau$, thus $\tau' \subseteq \tau$. Since $\forall i \in I$, τ' can let f_i be a continuous mapping, thus $\tau \subseteq \tau'$. □

Example

Let $((Y_i, \tau_i))_{i \in I}$ be topological spaces, $Y = \prod\limits_{i \in I} Y_i$ and $\pi_i : Y \to Y_i$, $(y_j)_{j \in I} \mapsto y_i$ be the projection mapping.

The product topology on Y is by definition the initial topology of $(\pi_i)_{i \in I}$.

3.8.9 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.

Proof

$\forall i \in I$, $\forall U \in \tau_i$, $(f_i \circ h)^{-1}(U) \in \tau_Z$. Therefore, $\forall i \in I$, $\forall U \in \tau_i$, $h^{-1}(f^{-1}(U)) \in \tau_Z$. □

Remark

We keep the notation of the definition of initial topology. If $\forall i \in I$, $\mathscr{B}_i$ is a topological basis of τ_i, then $\mathscr{B}' = \{\bigcap\limits_{j \in J} f_j^{-1}(U_j) | J \subseteq I$ finite and

$(U_j)_{j\in J} \in \prod_{j\in J} \mathscr{B}_j\}$ is also a topological basis of the initial topology.

Example

Let $I = \{1, 2, 3, ..., n\}$. Let $((X_i, d_i))_{i\in I}$ be a family of metric spaces. $X = \prod_{n\in I} X_i$. The mapping defined as $d; X \times X \to \mathbb{R}_{\geq 0}$, $(x_i)_{i\in I} \times (y_i)_{i\in I} \mapsto \max_{i\in I} d_i(x_i, y_i)$ is a metric on X.

Each $\pi_j : X \to X_i$, $(x_i)_{i\in I} \mapsto x_j$ is continuous(**Will be proved in 3.9**). Hence the product topology τ is contained in τ_d.

For any $(x_i)_{i\in I} \in X$, for any $\varepsilon > 0$, $B(x, \varepsilon) = \{(y_i)_{i\in I} | \forall i \in I, \max d(x_i, y_i) < \varepsilon\} = \prod_{i\in I} B_i(x_i, \varepsilon) = \bigcap_{i\in I} \pi^{-1}(B(x_i, \varepsilon)) \in \tau$, therefore, $\tau_d = \tau$.

3.9 Uniform continuity and convergence

3.9.1 Def

Let (X, d) be a metric space. For any non-empty set A, s.t. $A \subseteq X$, we define $\operatorname{diam}(A) = \sup\limits_{(x,y)\in A\times A} d(x, y)$, called the diameter of A, if $A = \varnothing$, by convention $\operatorname{diam}(A) = 0$. If $\operatorname{diam}(A) < +\infty$, we say A is bounded.

Remark

If A is finite, then A is bounded.

If $A \subseteq B$, then $\operatorname{diam}(A) \leq \operatorname{diam}(B)$.

3.9.2 Prop

Let (X, d) be a metric space. $A \subseteq X$, $B \subseteq X$, $(x_0, y_0) \in A \times B$, 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.

Proof

Let $(x, y) \in (A \bigcup B)^2$.

If $(x, y) \in A^2$, then $d(x, y) \leq \operatorname{diam}(A)$.

If $(x, y) \in B^2$, then $d(x, y) \leq \operatorname{diam}(B)$.

If $x \in A$ and $y \in B$, then $d(x, y) \leq d(x, x_0) + d(x_0, y_0) + d(y_0, y) \leq \operatorname{diam}(A) + d(x_0, y_0) + d(y_0, y)$.

Similarly, if $x \in B$ and $y \in A$, then $d(x, y) \leq \operatorname{diam}(A) + d(x_0, y_0) + d(y_0, y)$. $\square$

Example

$\operatorname{diam}(\overline{B}(x, r)) \leq 2r$.

If $(y, z) \in \overline{B}(x, r)$, $d(y, z) \leq d(y, x) + d(x, z) \leq 2r$.

3.9.3 Def

Let (X, d) be a metric space, $I \subseteq \mathbb{N}$ be an infinite subset, $(x_n)_{n\in I} \in X^I$. If $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, $\operatorname{diam}\{x_n | n \geq N\} \leq \varepsilon$, then we say that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.

3.9.4 Prop

(1) If $(x_n)_{n\in I}$ converges to x, then it is a Cauchy sequence.

(2) If $(x_n)_{n\in I}$ is a Cauchy sequence, then it is bounded.

(3) Suppose that $(x_n)_{n\in I}$ is a Cauchy sequence, if there exists $J\subseteq I$, s.t. $(x_n)_{n\in J}$ converges to $l\in X$, then $(x_n)_{n\in I}$ converges to l.

Proof

(1) $\forall\varepsilon>0$, $\exists N\in\mathbb{N}$, s.t. $\forall n\in I_{\geq N}$, $d(x_n,x)<\frac{\varepsilon}{2}$, let $(m,n)\in I_{\geq N}\times I_{\geq N}$, then $d(x_m,x_n)\leq d(x_m,x)+d(x,x_n)<\varepsilon$. $\square$

(2) $\exists N\in\mathbb{N}$, s.t. $\operatorname{diam}\{x_n|n\in I_{\geq N}\}\leq 1$, because $\{x_n|n\in I_{<n}\}$ is finite and $\{x_n|n\in I_{\geq N}\}$ is bounded, hence their union is bounded. $\square$

(3) $\forall\varepsilon>0$, $\exists N_1\in\mathbb{N}$, $\{x_n|n\in J_{\geq N_1}\}\subseteq B(l,\frac{\varepsilon}{2})$, $\exists N_2\in\mathbb{N}$, $\operatorname{diam}\{x_n|n\in I_{\geq N_2}\}<\frac{\varepsilon}{2}$. Let $N=\max\{N_1,N_2\}$, then for any $(n,m)\in I_{\geq N}$, $\exists p\in J_{\geq N}$, s.t. $d(x_m,l)\leq d(x_m,x_p)+d(x_p,x_n)<\varepsilon$. $\square$

3.9.5 Def

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. $\forall(x,y)\in\operatorname{Dom}(f)\times\operatorname{Dom}(f)$, $d_X(x,y)\leq\delta$ implies $d_Y(f(x),f(y))\leq\varepsilon$, namely $\inf\limits_{\delta>0}\sup\limits_{\substack{(x,y)\in\operatorname{Dom}(f)^2\\ d(x,y)\leq\delta}}d(f(x),f(y))=0$, we say that f is uniformly continuous.

3.9.6 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 $I\subseteq\mathbb{N}$ 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.

Proof

(1) $\forall\varepsilon>0$, $\exists\delta>0$, $d_X(x_m,x_n)<\delta$ implies $d_Y(f(x_m),f(x_n))<\varepsilon$, since $(x_n)_{n\in I}$ is a Cauchy sequence, there exists $N\in\mathbb{N}$, for any $(n,m)\in I_{\geq N}\times I_{\geq N}$, $d_X(x_m,x_n)<\delta$. $\square$

(2) Let $(x_n)_{n\in I}$ be a sequence in $\operatorname{Dom}(f)^I$ that converges to $x\in\operatorname{Dom}(f)$.

We define $(y_n)_{n\in\mathbb{N}}$ as if n is odd then $y_n = x$, if n is even, then $y_n = x_{\frac{n}{2}}$. Then $(y_n)_{n\in\mathbb{N}}$ converges to x. Hence $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Since f is uniformly continuous, thus $(f(y_n))_{n\in\mathbb{N}}$ is also a Cauchy sequence in Y. $(f(y_n))_{n \text{ is odd}} = (f(x))_{n \text{ is odd}}$ converges to $f(x)$, which means $(f(y_n))_{n\in\mathbb{N}}$ converges to $f(x)$.

This leads to $\lim\limits_{n\to+\infty} f(x_n) = f(x)$. Hence f is continuous at x.

3.9.7 Def

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

(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 $\lim\limits_{n\in I, n\to+\infty} \sup\limits_{x\in Z} d(f_n(x), f(x)) = 0$, we say that $(f_n)_{n\in I}$ converges uniformly to f.

3.9.8 Theorem

Let X and Y be metric spaces, $Z \subseteq X$, $I \subseteq \mathbb{N}$ infinite, $(f_n)_{n\in I}$ and f be function from X to Y, having Z as domain of definition. Suppose that $(f_n)_{n\in I}$ converges uniformly to f and each f_n is uniformly continuous.

Then, f is uniformly continuous.

Proof

Own: Fix $\varepsilon > 0$

For any $t \in I$, $\exists \delta > 0$, s.t. $\forall (m, n) \in Z^2$, $d_Z(m, n) < \delta$ implies $d_Y(f_t(m), f_t(n)) < \frac{\varepsilon}{3}$.

$\exists N \in \mathbb{N}$, $\forall n \in I_{\geq N}$, $d_Y(f_n(x), f(x)) < \frac{\varepsilon}{3}$.

Let $(x, y) \in Z^2$, s.t. $d_Z(x, y) < \delta$, then $d_Y(f(x), f(y)) \leq d_Y(f(x), f_k(x)) + d_Y(f_k(x), f_k(y)) + d_Y(f_k(y), f(y)) < \varepsilon$.

CHEN Huayi: Using d to denote the metric of Y.

For $n \in I$, let $A_n = \sup\limits_{x\in Z} d(f_n(x), f(x))$. $\lim\limits_{n\to+\infty} A_n = 0$.

$\forall (x, y) \in Z \times Z, \forall n \in I$, $d(f(x), f(y)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(y)) +$

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

$\inf_{\delta>0} \sup_{\substack{(x,y)\in Z^2 \\ d(x,y)\leq\delta}} d(f(x), f(y)) \leq 2A_n$ (f_n is uniformly continuous.)

Hence $0 \leq \inf_{\delta>0} \sup_{\substack{(x,y)\in Z^2 \\ d(x,y)\leq\delta}} d(f(x), f(y)) \leq 2A_n$.

Take $\lim_{n\to+\infty}$, by squeeze theorem, we get $\inf_{\delta>0} \sup_{\substack{(x,y)\in Z^2 \\ d(x,y)\leq\delta}} d(f(x), f(y)) = 0$. □

Example

(To clarify why do we need to have uniformly convergence but not just convergence)

$f_n = x^n$ on $[0, 1]$

$f = \mathbb{1}_{\{1\}}$ on $[0, 1]$

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A \end{cases}$$

3.9.9 Theorem

Let X be a topological space, Y be a metric space. $Z \subseteq X$, $p \in Z$, $I \subseteq \mathbb{N}$ infinite. $(f_n)_{n\in I}$, and f be a function from X to Y, having Z as domain of definition. Suppose that $(f_n)_{n\in I}$ converges uniformly to f and each f_n is continuous at p. Then f is continuous at p.

Proof

Let d be the metric of Y.

$\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, s.t. $\forall n \in I_{\geq N}$, $d(f_n(p), f(p)) < \frac{\varepsilon}{3}$ .Since f_N is continuous at p, hence there exists $V \in \mathcal{V}_p$ and $x \in V \bigcap Z$, s.t. $d(f(x), f(p)) < \frac{\varepsilon}{3}$ and $d(f_N(x), f(x)) < \frac{\varepsilon}{3}$

Therefore, $d(f(x), f(p)) \leq d(f(x), f_N(x)) + d(f_N(x), f_N(p)) + d(f_N(p), f(p)) \leq \varepsilon$, which shows $f(V \bigcap Z) \subseteq \overline{B}(f(p), \varepsilon)$ □

Example

Let X and Y be metric spaces, f be a function from X to Y. $\varepsilon > 0$. If $\forall (x, y) \in \mathrm{Dom}(f)^2$, $d(f(x), f(y)) \leq \varepsilon d(x, y)$, then we say that f is a

ε-Lipschitzian.

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

Remark

If f is Lipschiztian, then it is uniformly continuous.

Example

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

Example

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

We fix a unitary ring K.

4.1.1 Def

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.

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

$(K^{\oplus I} := \{(a_i)_{i\in I} \in K^I | \exists J \subseteq I \text{ finite, s.t. } a_i = 0 \text{ for } i \in I \setminus J\})$

In fact, $\varphi_x((a_i)_{i\in I} + (b_i)_{i\in I}) = \varphi_x((a_i)_{i\in I}) + \varphi_x((b_i)_{i\in I})$ is equivalent to $\sum\limits_{i\in I}(a_i + b_i)x_i = \sum\limits_{i\in I} a_i x_i + \sum\limits_{i\in I} b_i x_i$. And for any $\lambda \in K$, $\varphi_x(\lambda(a_i)_{i\in I}) = \sum\limits_{i\in I}(\lambda a_i)x_i = \lambda \sum\limits_{i\in I} a_i x_i = \lambda\varphi_x((a_i)_{i\in I})$.

4.1.2 Def

Let M be a left K-module , I be a set, $x = (x_i)_{i\in I} \in M^I$.

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

Example 1

Let $e_i \in K^{\oplus I}$ be the element $(\delta_{i,j})_{j\in I}$ with

$$\delta_{i,j} = \begin{cases} 1, & \text{if } j = i \\ 0, & \text{if } j \neq i \end{cases}$$

Then the family $e = (e_i)_{i\in I} \in (K^{\oplus I})^I$ is a basis of $K^{\oplus I}$.

In fact, e_i should be considered as a mapping from I to K that sends $j \in I$ to $\delta_{i,j}$. $\sum\limits_{i\in I} a_i e_i$ sends $j \in I$ to $\sum\limits_{i\in I} a_i \delta_{i,j} = a_i$, $(a_i)_{i\in I}$ also sends $i \in I$ to a_i.

Thus $\varphi_e((a_i)_{i\in I}) = \sum\limits_{i\in I} a_i e_i = (a_i)_{i\in I}$, $\varphi_e = \mathrm{Id}_{K^{\oplus I}}$ is a bijection.

Example 2

For the case of K^n $(n \in \mathbb{N})$

$e_i = (0, ..., 0, 1, 0, ..., 0)$

$\forall (a_1, ..., a_n) \in K^n$, $(a_1, ..., a_n) = \sum\limits_{i=1}^{n} (0, ..., a_i, ..., 0) = \sum\limits_{i=1}^{n} a_i(0, ..., 1, ..., 0) = \sum\limits_{i=1}^{n} a_i e_i$.

4.1.3 Def

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.

Remark

Let R be a unitary ring. One can notice that any free $R-$module is isomorphism to $R^{\oplus I} := \{(a_i)_{i \in I} | a_i \in R, a_i \neq 0 \text{ for finitely many } i\}$ for some I.

In fact, $\forall ((a_i)_{i\in I}, (b_i)_{i\in I}) \in (R^{\oplus I})^2$, $(a_i)_{i\in I} + (b_i)_{i\in I} := (a_i + b_i)_{i\in I}$, $\lambda \in R$, $\lambda (a_i)_{i\in I} := (\lambda a_i)_{i\in I}$

Notation

By convention, one can write $(a_i)_{i\in I} \in R^{\oplus I}$ as the form $\sum\limits_{i\in I} a_i i$

Example

$j = (a_i)_{i\in I} \in R^{\oplus I}$, s.t. $a_i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$. Then one can seen I as a subset of $R^{\oplus I}$

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 $\forall (a_1, ..., a_n) \in K^n$, $a_1 x_1 + ... + a_n x_n = 0$ 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 $b_1 x_1 + ... + b_n x_n$, $(b_1, ..., b_n) \in K^n$. Such expression is called a

$K-$linear combination of $x_1, ..., x_n$.

4.1.4 Theorem

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

Proof(By induction on n)

If $n = 0$, then $V = \{0\}$.

In this case $\varnothing$ is a basis of V. $\varphi_\varnothing : K^\varnothing \to V$.

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

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

Otherwise, $\exists (a_1, .., a_n) \in K^n$, s.t. $(a_1, .., a_n) \neq (0, .., 0)$ 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 $(x_i)_{i \in \{1,..,n\}}$ 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 $(x_i)_{i \in \{1,..,n-1\}}$ forms a system of generators. By the induction hypothesis, there exists $I \subseteq \{1, .., n-1\}$ s.t. $(x_i)_{i \in I}$ forms a basis of V. $\square$

Notation

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

4.1.5 Theorem

Let K be a unitary ring and V be a left K-module . W be a left K-submodule 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 $j \in \{1, .., l\}$, 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.

Proof

(1) Let $\pi : V \to V/W$ be the projection morphism. ($\pi(x) = [x]$).

Suppose that $\exists (b_i)_{i=1}^{n+l} \in K^{n+1}$ 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, $b_{n+1} = .. = b_{n+l} = 0$. Since $(x_i)_{i=1}^{n}$ is linearly independent. $b_1 = .. = b_n = 0$ □

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

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

Therefore, $y = \sum_{i=1}^{n+l} c_i x_i$. □

4.1.6 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.

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 K-submodule 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.

Hence 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. □

4.1.7 Def

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), $\text{rk}(V)$ is also denoted as $\dim(V)$ or $\dim_K(V)$, called the dimension of V.

4.1.8 Theorem

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

(1) W and V/W are both of finite type, and $\text{rk}(V) = \text{rk}(W) + \text{rk}(V/W)$.

(2) Any basis of V has exactly $\text{rk}(V)$ elements.

Proof

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

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

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

Since $\pi(x) = \sum\limits_{i=1}^{l}(a_i + \sum\limits_{j=l+1}^{n} a_j b_{j,i})\pi(x_i) = 0$. Hence $x = \sum\limits_{j=l+1}^{n} a_j y_j$.

Hence W is of finite type, and $\text{rk}(V) \geq \text{rk}(W) + \text{rk}(V/W)$

Moreover, the previous theorem shows that $\text{rk}(V) \leq \text{rk}(W) + \text{rk}(V/W)$.

$\square$

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

If $\text{rk}(V) = 0$, in this case $V = \{0\}$. Since $\{0\}$ is not linear independent thus the only basis of V is $\varnothing$. So the statement holds.

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

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

$\forall \lambda \in K$, $\exists (\lambda a^{-1}) \in K$, s.t. $(\lambda a^{-1})ae = \lambda e$. $\text{rk}(V) = 1$.

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

Induction hypothesis: True for a basis of $< m$ 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 $\sum_{i=2}^m a_i e_i \in W$. Hence $\exists a_i \in K$, s.t. $\sum_{i=1}^m a_i e_i - a_1 e_1 = 0$. Hence $a_1 = ... = a_m = 0$. Therefore, $(\pi(e_i))_{i=2}^m$ is a basis of V/W.

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

4.1.9 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 $x = (x_i)_{i \in I} \in E^I$.

(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.

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 Def

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 $\varphi_{(x_i)_{i=1}^n}$: $K^n \to V$

Example

Suppose that $V = K^p (p \in \mathbb{N})$. Then each $x_i \in K^p$ is of the form $(x_{i,1}, x_{i,2}, .., x_{i,p})$. 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 Def

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

Example

Denote by $\boldsymbol{I_n}$ the identity mapping $K^n \to K^n$. Let $e_i = (0, .., 1, .., 0)$. Then $(e_i)_{i=1}^n$ is a basis of K^n called tha canonical basis of K^n.

$\varphi_{(e_i)_{i=1}^n} = \mathrm{Id}_{K^n}$

$$\boldsymbol{I_n} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

Example

Let $(x_1, .., x_n) \in K^n$. Denote by $\mathbf{diag}(\boldsymbol{x_1, .., x_n}) : K^n \to K^n, (a_1, .., a_n) \mapsto (a_1x_1, .., a_nx_n)$

$$\mathbf{diag}(\boldsymbol{x_1},..,\boldsymbol{x_n}) = \begin{pmatrix} x_1e_1 \\ x_2e_2 \\ \vdots \\ x_ne_n \end{pmatrix} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{pmatrix}$$

4.2.3 Def

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 $M_{n,p}(K) \times M_{p,r}(K) \to M_{n,r}(K)$, $(\boldsymbol{A},\boldsymbol{B}) \mapsto \boldsymbol{AB} := \boldsymbol{B} \circ \boldsymbol{A}$.

Example

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

Consider a matrix $\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \in M_{p,n}(K)$.

$\boldsymbol{A}$ is a morphism of left K-module from K^p to K^n.

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

Let $(b_1,...,b_p) \in K^p$, $\boldsymbol{A}((b_1,...,b_p)) = \sum_{i=1}^{p} b_i(a_{i1},...,a_{in})$.

$$\varphi_x(\boldsymbol{A}(b_1,...,b_p)) = \varphi_x(\sum_{i=1}^{p} b_i(a_{i1},...,a_{in})) = \sum_{i=1}^{p} b_i\varphi_x((a_{i1},...,a_{in})) = \sum_{i=1}^{p} b_i \sum_{j=1}^{n} a_{ij}x_j$$

$$\boldsymbol{A}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{pj}x_j \end{pmatrix}$$

Let $\boldsymbol{B} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nr} \end{pmatrix} : K^n \to K^r$

$$\boldsymbol{AB} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots \\ a_{p1} & \cdots & a_{pn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1r} \\ b_{21} & b_{22} & \cdots & b_{2r} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nr} \end{pmatrix} = \begin{pmatrix} \sum\limits_{i=1}^{n} a_{1,i}(b_{i,1}, ..., b_{i,r}) \\ \vdots \\ \sum\limits_{i=1}^{n} a_{p,i}(b_{i,1}, ..., b_{i,r}) \end{pmatrix}$$

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

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

4.3 Transpose

We fix a unitary ring K.

4.3.1 Def

Let E be a left K-module .

Denote by $E^\vee := \{$morphisms of left K-module 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 $K \times E^\vee \to E^\vee$, $(a,f) \mapsto fa : x \in E, f(x)a$. $\forall \lambda \in K$, $(fa)(\lambda x) = f(\lambda x)a = \lambda f(x)a = \lambda(fa)(x)$.

This mapping defines a structure of right K-module on $E^\vee$.

4.3.2 Def

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 $\forall a \in K$, $g \in F^\vee$, $\varphi^\vee(g)a = (g \circ \varphi)a = g(\varphi(\cdot))a = (ga) \circ \varphi(\cdot) = \varphi^\vee(ga)$.

Example

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

$\varphi = \begin{pmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{np} \end{pmatrix}$.

φ sends $(a_1, \cdots, a_n) \in K^n$ to $(\sum\limits_{i=1}^n a_i b_{i1}, \cdots, \sum\limits_{i=1}^n a_i b_{ip}) \in K^p$.

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 $(a_1, \cdots, a_n)$ to $\sum\limits_{j=1}^p (\sum\limits_{i=1}^n a_i b_{ij}) y_j$.

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

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 $\iota_n \circ \varphi^\vee \circ \iota_p^{-1}$.

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

φ^τ is called the transpose of φ.

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

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

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

$y_2(b_{1,2},\cdots,b_{n,2})+\cdots+y_p(b_{1,p},\cdots,b_{n,p})$.

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

4.3.3 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 $(\psi\circ\varphi)^\vee = \varphi^\vee \circ \psi^\vee$.

Proof

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

4.3.4 Corollary

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

Proof

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

Remark

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

(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 $E\to(E^\vee)^\vee$ is an isomorphism.

$$
\begin{tikzcd}
K^n \arrow[d, "\varphi_e", "\cong"'] \arrow[dr, "\varphi_{e^\vee}"] & (K^n)^\vee \arrow[l, "\cong"', "\iota_n"] \\
E & E^\vee \arrow[l, "\cong"] \arrow[u, "\varphi_e^\vee", "\cong"']
\end{tikzcd}
$$

Note

这一节将几何和代数两种视角融合的很好, 这也是线性代数的核心。用几何观点避开繁琐的代数验证, 又用代数将抽象的映射具体起来。

4.4 Linear equations

We fix a unitary ring K.

4.4.1 Def

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

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

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

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

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

4.4.2 Prop

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

Proof

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

Note

这段证明比较诡异，感觉直接 by definition 说明是一个单射就行了()

□

4.4.3 Def

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

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

We consider the equation $\boldsymbol{A} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ (*)

$$\begin{cases} a_{1,1}x_1 + ... + a_{1,p}x_p = b_1 \\ \cdots \\ a_{n,1}x_1 + ... + a_{n,p}x_p = b_n \end{cases}$$

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

4.4.4 Prop

Suppose that A is of reduced row echelon form.

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

(1) If $\exists i \in \{1, ..., n\} \setminus I(A)$ s.t. $b_i \neq 0$ then $(*)$ does not have any solution in V^p.

(2) Suppose that $\forall i \in \{1, ..., n\} \setminus I(A)$, $b_i = 0$. Then $(*)$ has at least one solution.

Moreover, $V^{J_0(A)} \to V^p$, $(z_k)_{k \in J_0(A)} \mapsto (x_1, ..., x_p)$

with $x_j = \begin{cases} z_j & j \in J_0(A) \\ b_i - \sum\limits_{l \in J_0(A)} a_{i,l} z_l, & j \in j((a_{i,1}, ..., a_{i,p})), i \in I(A) \end{cases}$

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

Proof

(1) Trivial

(2) Tedious and complex.

4.4.5 Prop

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

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

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

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

Example 1

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

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

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

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

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

Example 2

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

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

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

$$\mathbf{diag}(\boldsymbol{r_1,...,r_n})\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} r_1y_1 \\ \vdots \\ r_ny_n \end{pmatrix}, \ \mathbf{diag}(\boldsymbol{r_1,...,r_n})\boldsymbol{I_n} = \begin{pmatrix} r_1e_1 \\ \vdots \\ r_ne_n \end{pmatrix}$$

Example 3

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

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

Since $c_i = 0$, $\boldsymbol{S_{i,-c}S_{i,c} = I_n}$

$$\boldsymbol{S_{i,c}}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \circ \boldsymbol{S_{i,c}} : K^n \to W.$$

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

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

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

4.4.6 Def

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

(1) $\boldsymbol{P_\sigma}$ where $\sigma \in \mathfrak{S}_n = \{\text{bijections from } n \text{ to } n\}$.

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

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

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

4.4.7 Theorem

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

Proof (by induction)

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

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

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

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

Suppose that $(\lambda_1, ..., \lambda_n) \neq (0, ..., 0)$, 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 $\boldsymbol{A}$ to the form $\begin{pmatrix} 1 & \mu_2 & \cdots & \mu_p \\ 0 & & & \\ \vdots & & \boldsymbol{C} & \\ 0 & & & \end{pmatrix}$. $\boldsymbol{C} \in M_{n-1,p-1}(K)$, $(\mu_2, \mu_3, ..., \mu_p) \in K^{p-1}$.

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

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

4.5 Normed vector space

4.5.1 Def

Let (X,d) be a metric space. If $(x_n)_{n\in\mathbb{N}}$ is an element of $X^{\mathbb{N}}$ s.t. $\lim_{N\to+\infty} \sup_{(n,m)\in\mathbb{N}^2_{\geq N}} d(x_n,x_m) = 0$, 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 a binary relation $\sim$ on $\mathrm{Cau}(X,d)$ as

$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$ iff $\lim_{n\to+\infty} d(x_n,y_n) = 0$

4.5.2 Prop

$\sim$ is an equivalence relation.

Proof

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

If $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ be elements of $\mathrm{Cau}(X,d)$.
$0 \leq d(x_n,z_n) \leq d(x_n,y_n) + d(y_n,z_n) \leq 0$. □ □

4.5.3 Def

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

4.5.4 Theorem

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

Proof

To check that d is well defined, it suffices to prove that $\forall([x],[y]) \in \widehat{X}\times\widehat{X}$, $(d(x_n,y_n))_{n\in\mathbb{N}}$ is a Cauchy sequence (Since $\mathbb{R}$ is complete) and its limit doesn't depends 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, $0 \le \sup_{(n,m)\in\mathbb{N}^2_{\ge N}} |d(x_n,y_n)-d(x_m,y_m)| \le \sup_{(n,m)\in\mathbb{N}^2_{\ge N}} d(x_n,x_m)+ \sup_{(n,m)\in\mathbb{N}^2_{\ge N}} d(y_n,y_m)$

Taking $\lim_{N\to+\infty}$, 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 $\lim_{n\to+\infty} d(x_n,x'_n) = \lim_{n\to+\infty} d(y_n,y'_n) = 0$

$0 \le |d(x_n,y_n) - d(x'_n,y'_n)| \le d(x_n,x'_n) + d(y_n,y'_n)$, taking $\lim_{n\to+\infty}$, we get $\lim_{n\to+\infty} |d(x_n,y_n) - d(x'_n,y'_n)| = 0$

So $\lim_{n\to+\infty} d(x_n,y_n) = \lim_{n\to+\infty} d(x'_n,y'_n)$

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

$\widehat{d}([x],[y]) = 0$ iff $[x] = [y]$. if $[x] = [y]$ it is trivial, if $\widehat{d}([x],[y]) = 0$, which actually means $x \sim y$, thus $[x] = [y]$.

Symmetry is trivial.

If $[x]$ $[y]$ and $[z]$ are elements of $\widehat{X}$.

$d([x],[z]) = \lim_{n\to+\infty} d(x_n,z_n) \le \lim_{n\to+\infty} d(x_n,y_n) + \lim_{n\to+\infty} d(y_n,z_n) = \widehat{d}([x],[y]) + \widehat{d}([y],[z])$.

Remark

$i_X : X \to \widehat{X}$, $a \mapsto [(a,a,a,...)]$

$\widehat{d}(i_X(a), i_X(b)) = d(a,b)$. In particular, i_X is injective, if $i_X(a) = i_X(b)$, then $d(a,b) = 0$, hence $a = b$.

4.5.5 Prop

$i_X(X)$ is dense in $\widehat{X}$.(The closure of $i_X(X)$ is $\widehat{X}$).

Proof

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

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

4.5.6 Theorem

$(\widehat{X}, \widehat{d})$ is a complete metric space.

Proof

Let $([x^{(N)}])_{N\in\mathbb{N}}$ be a Cauchy sequence in $\widehat{X}$, 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. $\forall (k,l) \in \mathbb{N}^2_{\geq N_0}$, $\widehat{d}([x^{(k)}], [x^{(l)}]) = \lim\limits_{n\to+\infty} d(x_n^{(k)}, x_n^{(l)}) \leq \varepsilon$

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

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

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

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

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

Taking $\lim\limits_{n\to+\infty}$, get $d(y_k, y_l) \leq \varepsilon$.

So $y = (y_N)_{N\in\mathbb{N}}$ is a Cauchy sequence. We check that $\lim\limits_{N\to+\infty} \widehat{d}([x^{(N)}], [y]) = 0$

$d(x_n^{(N)}, y_n) \leq d(x_n^{(N)}, y_N) + d(y_n, y_N)$

$0 \leq \limsup\limits_{N\to+\infty} \lim\limits_{n\to+\infty} d(x_n^{(N)}, y_n) \leq \limsup\limits_{N\to+\infty} (\frac{1}{N+1} + \lim\limits_{n\to+\infty} d(y_n, y_N))$

$= \limsup\limits_{N\to+\infty} \lim\limits_{n\to+\infty} d(y_n, y_N) = 0.$ □

Example

Let $(K, |\cdot|)$ be a valued field. It has a absolute value. This is a metric space with $d(a,b) := |a-b|$.

$\mathrm{Cau}(K)$ forms a commutative unitary ring.

$(a_n)_{n\in\mathbb{N}} \sim (b_n)_{n\in\mathbb{N}}$ iff $\lim\limits_{n\to+\infty}(a_n - b_n) = 0$

$(a_n - b_n)_{n\in\mathbb{N}} \in \mathrm{Cau}_0(K) = \{\text{Cauchy sequences that converges to } 0\}$.

This is an ideal of $\mathrm{Cau}(K)$

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

4.6 Norms

In this section, we fix a field K and an absolute value $|\cdot|$ on K.

Example: trivial absolute value $|x|_0 = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \neq 0. \end{cases}$

We assume that $(K, |\cdot|)$ forms a complete metric space with respect to the metric $K \times K \to \mathbb{R}_{\geq 0}$, $(a, b) \mapsto |a - b|$.

4.6.1 Def

Let V be a vector space over K (K-module), we call seminorm on V any mapping $||\cdot|| : V \to \mathbb{R}_{\geq 0}$, $s \mapsto \|s\|$, s.t.

- $\forall (a, s) \in K \times V$, $\|as\| = |a| \cdot \|s\|$
- $\forall (s, t) \in V \times V$, $\|s + t\| \leq \|s\| + \|t\|$

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.

Remark

If $\|\cdot\|$ is a norm on V, then $V \times V \to \mathbb{R}_{\geq 0}$, $(s, t) \mapsto \|s - t\|$ is a metric on V.

Example 1

If we consider K as a vector space over K, then $(K, |\cdot|)$ forms a normed vector space over K.

Example 2

Let $(V_1, \|\cdot\|_1)$, ..., $(V_n, \|\cdot\|_n)$ be vector spaces equipped with seminorms. Let $V = V_1 \oplus ... \oplus V_n = V_1 \times ... \times V_n$, $\|\cdot\|_{l^\infty} : V \to \mathbb{R}_{\geq 0}$, $(x_1, ..., x_n) \mapsto \max\limits_{i \in \{1,...,n\}} \|x_i\|_i$, $\|\cdot\|_{l^p} : V \to \mathbb{R}_{\geq 0}$, $(x_1, ..., x_n) \mapsto (\|x_1\|_1^p + ... + \|x_n\|_n^p)^{\frac{1}{p}}$ ($p \in \mathbb{R}_{\geq 1}$). There are seminorms, they are norms if $\|\cdot\|_1, \cdots, \|\cdot\|_n$ are all norms.

Note

Below is the proof of $\|\cdot\|_{l^p}$ satisfies triangular inequality.

In this note, we default that x, y, a, b, p, q are all positive numbers.

Let $(p,q) \in \mathbb{R}^2$, s.t. $\frac{1}{p} + \frac{1}{q} = 1$, then $\forall (a,b) \in \mathbb{R}$, one can claim that $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$. In fact, since **Jensen's inequality** shows that

$\ln(a_1x_1 + a_2x_2) \geq a_1\ln(x_1) + a_2\ln(x_2)$ if $a_1 + a_2 = 1$, let $\frac{1}{p} = a_1$, $\frac{1}{q} = a_2$, then one can obtain the previous inequality, whose name is **Young inequality**.

With that inequality, one can get another inequality immediately, $\sum\limits_{i=1}^{n} x_iy_i \leq \frac{\sum\limits_{i=1}^{n} x_i^p}{p} + \frac{\sum\limits_{i=1}^{n} y_i^q}{q}$, let $x_i = \frac{a_i}{(\sum\limits_{i=1}^{n} a_i^p)^{\frac{1}{p}}}$ and $y_i = \frac{b_i}{(\sum\limits_{i=1}^{n} b_i^q)^{\frac{1}{q}}}$, then one can prove another well-known inequality **Hölder inequality**: $\sum\limits_{i=1}^{n} a_ib_i \leq (\sum\limits_{i=1}^{n} a_i^p)^{\frac{1}{p}}(\sum\limits_{i=1}^{n} b_i^q)^{\frac{1}{q}}$ if $\frac{1}{p} + \frac{1}{q} = 1$.

Next is the last inequality **Minkowski inequality**. Let $p > 1$, $\sum\limits_{i=1}^{n}(x_i + y_i)^p = \sum\limits_{i=1}^{n} x_i(x_i+y_i)^{p-1} + \sum\limits_{i=1}^{n} y_i(x_i+y_i)^{p-1} \leq (\sum\limits_{i=1}^{n} x_i^p)^{\frac{1}{p}}(\sum\limits_{i=1}^{n}(x_i+y_i)^{q(p-1)})^{\frac{1}{q}} + (\sum\limits_{i=1}^{n} y_i^p)^{\frac{1}{p}}(\sum\limits_{i=1}^{n}(x_i+y_i)^{q(p-1)})^{\frac{1}{q}} = ((\sum\limits_{i=1}^{n} x_i^p)^{\frac{1}{p}} + (\sum\limits_{i=1}^{n} y_i^p)^{\frac{1}{p}})(\sum\limits_{i=1}^{n}(x_i+y_i)^p)^{\frac{1}{p}\frac{p}{q}}$, hence $(\sum\limits_{i=1}^{n}(x_i + y_i)^p)^{\frac{1}{p}(p-\frac{p}{q})} = (\sum\limits_{i=1}^{n}(x_i + y_i)^p)^{\frac{1}{p}} \leq (\sum\limits_{i=1}^{n} x_i^p)^{\frac{1}{p}} + (\sum\limits_{i=1}^{n} y_i^p)^{\frac{1}{p}}$

Therefore, $\|x + y\|_{l^p} = (\sum\limits_{i=1}^{n} \|x_i + y_i\|_i^p)^{\frac{1}{p}} \leq (\sum\limits_{i=1}^{n}(\|x_i\|_i + \|y_i\|_i)^p)^{\frac{1}{p}} \leq (\sum\limits_{i=1}^{n} \|x_i\|_i^p)^{\frac{1}{p}} + (\sum\limits_{i=1}^{n} \|y_i\|_i^q)^{\frac{1}{q}} = \|x\| + \|y\|$

4.6.2 Def

Let $(V, \|\cdot\|)$ be a vector space over K equipped with a seminorm, and W be a vector subspace of V (sub-K-module).

- The restriction of $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ to W forms a seminorm on W. It is a norm if $\|\cdot\|$ is a norm.

  $\|\cdot\|_W : W \to \mathbb{R}_{\geq 0}$, $x \mapsto \|x\|$.

  $\|\cdot\|_W$ is called the restricted seminorm of $\|\cdot\|$ on W.

- The mapping $\|\cdot\|_{V/W} : V/W \to \mathbb{R}_{\geq 0}$, $\alpha \mapsto \inf\limits_{s\in\alpha} \|s\|$. $\|[s]\|_{V/W} = \inf\limits_{w\in W} \|s + w\|$ is a seminorm on V/W.

  Even if $\|\cdot\|$ is a norm, $\|\cdot\|_{V/W}$ might only be a seminorm.

$\|\cdot\|_{V/W}$ is called the quotient seminorm of $\|\cdot\|$.

$\lambda \in K$, $\alpha \in V/W$, $\|\lambda\alpha\|_{V/W} = \inf_{s\in\alpha} \|\lambda s\| = \inf_{s\in\alpha} |\lambda|\|s\| = |\lambda| \cdot \|\alpha\|_{V/W}$

$\|\alpha + \beta\|_{V/W} = \inf_{s\in\alpha+\beta} \|s\| = \inf_{(x,y)\in\alpha\times\beta} \|x + y\| \leq \|\alpha\|_{V/W} + \|\beta\|_{V/W}$

Note

商了之后变成半范数的例子 $\mathbb{R}/\mathbb{Q}$。

4.6.3 Prop

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

Proof

If $(a, s) \in K \times N$, then $\|as\| = |a|\|s\| = 0$, so $as \in N$.

If $(s_1, s_2) \in N\times N$, then $0 \leq \|s_1+s_2\| \leq \|s_1\|+\|s_2\| = 0$, so $s_1+s_2 \in N$.

Let $\alpha \in V/N$, s.t. $\|\alpha\|_{V/N} = 0$.

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

Hence $\|\alpha\|_{V/N} = \|s\| = 0$, we obatain that $\alpha = N = [0]$. □

4.6.4 Def

Let $(V, \|\cdot\|)$ be a vector space over K, equipped with a seminorm. For any $x \in V$, and $r \geq 0$, we denote by $B(x, r)$ the set $\{y \in V | \|y - x\| < r\}$, $\overline{B}(x, r)$ the set $\{y \in V | \|y - x\| \leq r\}$.

Let $N = \{s \in V | \|s\| = 0\}$. $\forall x \in V$, $\forall r \geq 0$, $x+N \subseteq \overline{B}(x, r)$. Moreover, if $r > 0$, $x + N \subseteq B(x, r)$. We equip V with the topology s.t. $\forall U \subseteq V$, U is open iff $\forall x \in U$, $\exists r_x > 0$, $B(x, r_x) \subseteq U$.

4.6.5 Prop

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

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

- If there exists $C > 0$ s.t. $\forall x \in V_1$, $\|f(x)\|_2 \leq C\|x\|_1$, 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.

Proof

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

Proof: Let $s \in V \setminus N_{\|\cdot\|}$. Then $\|s\| > 0$.

Let $\varepsilon = \frac{\|s\|}{2}$, $\forall x \in B(s,\varepsilon)$, $\|x\| \geq |\|s\| - \|s-x\|| \geq \|s\| - \varepsilon = \varepsilon > 0$, so $B(s,\varepsilon) \subseteq V \setminus N_{\|\cdot\|}$.

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

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

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

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

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 $\forall x \in V_1$, $\|f(x)\|_2 \leq \frac{1}{\varepsilon|a|}\|x\|_1$.

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. $\|a^n x\| = |a|^n\|x\|_1 < \varepsilon \leq \|a^{n-1}x\|_1 = |a|^{n-1}\|x\|_1$.

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

4.6.6 Def

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

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

$$\|f\| := \begin{cases} \sup\limits_{x \in V_1 \|x\|_1 > 0} (\frac{\|f(x)\|_2}{\|x\|_1}) & \text{if } f(N_{\|\cdot\|_1}) \subseteq N_{\|\cdot\|_2} \\ +\infty & \text{if } f(N_{\|\cdot\|_1}) \not\subseteq N_{\|\cdot\|_2} \end{cases}$$

f is bounded iff $\|f\| < +\infty$. $\|f\|$ is called the operator seminorm 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.

4.6.7 Prop

$\mathscr{L}(V_1, V_2)$ is a vector subspace (sub-K-module) of $\mathrm{Hom}_K(V_1, V_2)$. Moreover, $\|\cdot\|$ is a seminorm on $\mathscr{L}(V_1, V_2)$.

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 $\lambda \in K$, $\lambda f : x \mapsto \lambda f(x)$, $\|\lambda f\| = \sup\limits_{x \in V_1, \|x\|_1 > 0} \frac{\|\lambda f(x)\|_2}{\|x\|_1} = |\lambda| \|f(x)\| < +\infty$

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 $\|x\|_1 \neq 0$

Thus $\|f\| \geq \frac{\|f(x)\|_2}{\|x\|_1} > 0$. Therefore, $\|\cdot\|$ is a norm.

4.6.8 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.

4.6.9 Theorem

Let $(V_1, \|\cdot\|_1)$ and $(V_2, \|\cdot\|_2)$ be vector spaces over K, equipped with seminorms. If $(V_2, \|\cdot\|_2)$ is a Banach space, then $(\mathscr{L}(V_1, V_2), \|\cdot\|)$ is a Banach space.

Proof

Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathscr{L}(V_1, V_2)$. $\forall x \in V_1$, the mapping $(f \in \mathscr{L}(V_1, V_2)) \mapsto f(x)$ is $\|x\|_1$-Lipschitz mapping. $\|f(x) - g(x)\| = \|(f - g)(x)\| \leq \|f - g\|\|x\|$

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 proved that g is an element of $\mathscr{L}(V_1, V_2)$.

$\forall (x, y) \in V_1 \times V_2$, $g(x + y) = \lim\limits_{n\to+\infty} f_n(x + y) = \lim\limits_{n\to+\infty} (f_n(x) + f_n(y))$

$\|f_n(x) + f_n(y) - g(x) - g(y)\| \leq \|f_n(x) - g(x)\| + \|f_n(y) - g(y)\| = o(1) + o(1) = o(1)$, $(n \to +\infty)$

$\forall x \in V_1$, $\forall \lambda \in K$, $g(\lambda x) = \lim\limits_{n\to+\infty} f_n(\lambda x) = \lim\limits_{n\to+\infty} \lambda f_n(x)$, $\|\lambda f_n(x) - \lambda g(x)\| = |\lambda|\|f_n(x) - g(x)\| = o(1)$, $(n \to +\infty)$

So $g(\lambda x) = \lambda g(x)$

$\forall x \in V_1$, $\|g(x)\| = \lim\limits_{n\to+\infty} \|f_n(x)\| \leq \lim\limits_{n\to+\infty} \|f_n\|\|x\|$.

In fact, $\forall (a, b) \in V_2^2$, $|\|a\| - \|b\|| \leq \|a - b\|$. $|\|f_n(x)\| - \|g(x)\|| \leq \|f_n(x) - g(x)\| = o(1)$, $g \in \mathscr{L}(V_1, V_2)$.

$\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall (n, m) \in \mathbb{N}_{\geq N}$, $\|f_n - f_m\| \leq \varepsilon$, $\forall x \in V_1$, $\|(f_n - f_m)(x)\| \leq \varepsilon\|x\|$, 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, $\|f_n(x) - g(x)\| \leq \|f_n(x) - f_m(x)\| + \|f_m(x) - g(x)\| \leq \varepsilon\|x\| + \|f_m(x) - g(x)\|$. $\square$

4.7 Differentiability

In this section, we fix a complete valued field $(K, |\cdot|)$. The absolute value is nontrivial.

4.7.1 Def

Let X be a topological space and $p \in X$ 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 $x \in V$.

We say that $f(x) = o(g(x))$, $x \to p$ 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 $\forall x \in V$, $\|f(x)\| \leq \varepsilon(x)g(x)$. $\lim\limits_{x \in V, x \to p} \varepsilon(x) = 0$ means $\forall \delta > 0$, $\exists$ open neighborhood U of p, $U \subseteq V$, and $\forall x \in U$, $0 \leq \varepsilon(x) \leq \delta$.

4.7.2 Def

Let E and F be normed vector spaces 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. $f(x) = f(p) + \varphi(x - p) + o(\|x - p\|)$, $x \to p$, then 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. $\lim\limits_{x \to p} \varepsilon(x) = 0$ and that $\|f(x) - f(p) - \varphi(x - p)\| \leq \varepsilon(x)\|x - p\|$, $\forall x \in V$.

4.7.3 Prop

If f is differentiable at p, then its differential at p is unique.

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 $f(x) = f(p) + \psi(x - p) + o(\|x - p\|)$

$(\varphi - \psi)(x - p) = o(\|x - p\|)$. $\exists \varepsilon : V \to \mathbb{R}_{\geq 0}$ s.t. V is a neighborhood of

p, $V \subseteq U$.

$\forall \delta > 0$, $\|\varphi - \psi\| = \sup\limits_{y \in E\setminus\{0\}} \frac{\|(\varphi-\psi)(y)\|}{\|y\|} = \sup\limits_{y\in E\setminus\{0\}, \|y\|\leq\delta} \frac{\|(\varphi-\psi)(y)\|}{\|y\|}$. In fact, $\lambda \neq 0$, $\frac{\|(\varphi-\psi)(\lambda y)\|}{\|\lambda y\|} = \frac{\|(\varphi-\psi)(y)\|}{\|y\|}$.

Therefore, $\|\varphi-\psi\| = \inf\limits_{\delta>0} \sup\limits_{y\in E, 0<\|y-p\|\leq\delta} \frac{\|(\varphi-\psi)(y-p)\|}{\|y-p\|} \leq \inf\limits_{\delta>0} \sup\limits_{y\in E,0<\|y-p\|\leq\delta} \varepsilon(y) = \limsup\limits_{y\to p} \varepsilon(y) = 0$. Hence $\varphi = \psi$. □

4.7.4 Def

Suppose that f is differentiable at p. We denote by d_pf the differential of f at p.

Example 1

$f : U \to F$, $f(x) = y_0$. $\forall x \in U, \forall p \in U$, $f(x)-f(p) = 0 = 0+o(\|x-p\|)$. Hence $d_pf(x) = 0$, $\forall x \in E$.

Example 2

Let $f \in \mathscr{L}(E, F)$. $f(x) - f(p) = f(x - p)$, hence $d_pf = f$.

Example 3

$A : E \times E \to E$, $(x, y) \mapsto x + y$.

$\|(x, y)\|_{l^1} = \|x\| + \|y\|$. $\|x + y\| \leq \|x\| + \|y\| = \|(x, y)\|_{l^1} \leq 2\|(x, y)\|_{l^\infty}$. And A is a K-linear mapping.

$\forall (p, q) \in E^2$, $d_{(p,q)}A = A$, $d_{(p,q)}A(x, y) = A(x, y) = x + y$

Example 4

$m : K \times E \to E$, $(\lambda, x) \mapsto \lambda x$.

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 K-linear mapping.

$(\mu_1 + \mu_2)p + a(y_1 + y_2) = (\mu_1p + ay_1) + (\mu_2p + ay_2)$

$b\mu p + a(by) = b(\mu p + ay)$

$\|\mu p+ay\| \leq |\mu|\|p\|+|a|\|y\| \leq \max\{|\mu|, \|y\|\}(|a|+\|p\|) = \|(\mu, y)\|_{l^\infty}(|a|+\|p\|)$

Thus this mapping is an element of $\mathscr{L}(K \times E, E)$

Hence m is differentiable, and $d_{(a,p)}m(\mu, y) = \mu p + ay$, $\forall(\mu, y) \in K \times E$.

$d_{(a,p)}(m)(\lambda - a, x - p) = (\lambda - a)p + a(x - p)$.

4.7.5 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 $d_p(g \circ f) = d_{f(p)}g \circ d_p f$.

Proof

Let $x \in U$. By definition, $f(x) = f(p) + d_p f(x - p) + o(\|x - p\|)$.

$f(x) - f(p) = O(\|x - p\|)$

$(g\circ f)(x) = g(f(x)) = g(f(p))+d_{f(p)}g(f(x)-f(p))+o(\|f(x)-f(p)\|) = g(f(p)) + d_{f(p)}g(f(x) - f(p)) + o(\|x - p\|) = g(f(p)) + d_{f(p)}g(d_p f(x - p) + o(\|x - p\|)) + o(\|x - p\|) = g(f(p)) + d_{f(p)}g(d_p f(x - p)) + o(\|x - p\|)$

So $g \circ f$ is differentiable at p, and $d(g \circ f) = d_{f(p)}g \circ d_p f$. □

4.7.6 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 $\forall x \in E$, $d_p f(x) = (d_p f_1(x), \cdots, d_p f_n(x))$.

Proof

Suppose $\forall i \in \{1, \cdots, n\}$, f_i is differentiable at p. $f(x)-f(p) = (f_1(x)-f_1(p), \cdots, f_n(x) - f_n(p)) = (d_p f_1(x - p), \cdots, d_p f_n(x - p)) + o(\|x - p\|)$

Therefore, f is differentiable at p and $d_p f(\cdot) = (d_p f_1(\cdot), \cdots, d_p f_n(\cdot))$

$\pi_i : F \to F_i$, $(x_1, \cdots, x_n) \mapsto x_i$ 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. □

4.7.7 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, $d_p(\varphi \circ f) = \varphi \circ d_p f$.

Proof

φ is differentiable at $f(p)$, and $d_{f(p)}\varphi = \varphi$. □

4.7.8 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 $d_p(af + bg) = ad_p f + bd_p g$.

Proof

$af + bg$ is the composite $U \to F \times F \to F$, $x \mapsto (f(x), g(x)) \mapsto af(x) + bg(x)$.

$\|ay + bz\| \leq |a|\|y\| + |b|\|z\| \leq (|a| + |b|) \max\{\|y\|, \|z\|\}$. □

4.7.9 Def

Let U be an open subset of K, and $(F, \|\cdot\|)$ 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.

4.7.10 Corollary

Let U and V be open subsets of K, $(F, \|\cdot\|)$ 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 $(g \circ f)'(p) = f'(p)g'(f(p))$.

Proof

By definition,

$d_p(g \circ f)(1) = d_{f(p)}g(d_p f(1))$

$= d_{f(p)}g(f'(p)) = d_{f(p)}g(f'(p)1)$

$= f'(p)d_{f(p)}g(1) = f'(p)g(1) = f'(p)g'(f(p))$ □

4.7.11 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

$\forall l \in E$, $d_p(fg)(l) = d_p f(l)g(p) + f(p)d_p g(l)$.

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, $U \to K \times F \to F$, $x \mapsto (f(x), g(x)) \mapsto f(x)g(x)$.

$d_p(fg)(l) = d_p(m \circ h)(l) = d_{h(p)}m(d_p h(l)) = d_{(f(p),g(p))}m(d_p f(l), d_p g(l)) = f(p)d_p g(l) + d_p f(l)g(p)$. □

4.7.12 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

$(fg)'(p) = f'(p)g(p) + f(p)g'(p)$

Proof

$(fg)'(p) = d_p(fg)(1) = d_p f(1)g(p) + f(p)d_p g(1) = f'(p)g(p) + f(p)g'(p)$. □

Example

$f_n : K \to K$, $x \mapsto x^n$ is differentiable at any $x \in K$. $f_n'(x) = nx^{n-1}$.

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 $f_{n+1}'(x) = (f_n f_1)'(x) = f_n(x) f_1'(x) + f_n'(x) f_1(x) = x^n + x f_n'(x) = (n+1)x^n$. $\square$

4.7.13 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. $\forall s \in E$, $C_1\|s\|_1 \le \|s\|_2 \le C_2\|s\|_1$.

4.7.14 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.

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 $(E, \|\cdot\|_2)$. $\square$

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.)

4.7.15 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 tio $\|\cdot\|_E'$ and $\|\cdot\|_F'$. Moreover the differential of f at p is not changed in the change of norms from $(\|\cdot\|_E, \|\cdot\|_F)$ to $(\|\cdot\|_E', \|\cdot\|_F')$.

Proof

$f = \mathrm{Id}_F \circ f \circ \mathrm{Id}_U$.

$d_p f = d_{f(p)}\mathrm{Id}_F \circ d_p f \circ d_p \mathrm{Id}_U = \mathrm{Id}_F \circ d_p f \circ \mathrm{Id}_E = d_p f$. □

4.7.16 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.

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.

$C \max\{|a_1|, \cdots, |a_n|\} \leq \|a_1e_1 + \cdots + a_ne_n\|$

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. $\overline{B}(s, \varepsilon) \bigcap W = \varnothing$. $\forall t \in W$, $s + t \notin \overline{B}(0, \varepsilon)$. 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. $\max\{|a_1|, \cdots, |a_{n-1}|\} \leq A\|a_1e_1 + \cdots + a_{n-1}e_{n-1}\|$, 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 $\alpha = [s] = a_n[e_n] \in Q$

$A^{-1} \max\{|a_1|, \cdots, |a_{n-1}|\} \leq \|a_1e_1 + \cdots + a_{n-1}e_{n-1}\| = \|s - a_ne_n\| \leq \|s\| + |a_n|\|e_n\|$

$\|\alpha\|_Q = |a_n|\|[e_n]\|_Q = |a_n| \inf_{t \in W} \|e_n + t\|$.

$\|e_n\| \geq \|[e_n]\|_Q$. Because of induction hypothesis, $\|[e_n]\|_Q > 0$, thus $\|e_n\| \leq B\|[e_n]\|_Q$

$s = a_n e_n + t \in V$ with $t \in W$.

$\|s\| \geq \|[a_n e_n]\|_Q = |a_n|\|[e_n]\|_Q \geq B^{-1}|a_n|\|e_n\|$

If $\|a_n e_n\| < \frac{1}{2}\|t\|$, $\|s\| \geq \|t\| - \|a_n e_n\| > \frac{1}{2}\|t\| \geq \frac{1}{2}A^{-1}\max\{|a_1|, \cdots, |a_{n-1}|\}$.

If $\|a_n e_n\| \geq \frac{1}{2}\|t\|$, $\|s\| \geq B^{-1}|a_n|\|e_n\| \geq \frac{B^{-1}}{2}\|t\| \geq \frac{B^{-1}A^{-1}}{2}\max\{|a_1|, \cdots, |a_{n-1}|\}$.

We take $C = \min\{B^{-1}\|e_n\|, \frac{A^{-1}}{2}, \frac{B^{-1}A^{-1}}{2}\}$

Then $\|s\| \geq C\max\{|a_1|, \cdots, |a_n|\}$.

Under the norm $\max\{|a_1|, \cdots, |a_n|\}$, a sequence $(a_1^{(k)}e_1 + \cdots + a_n^{(k)}e_n)_{k \in \mathbb{N}}$ is a Cauchy sequence iff $\forall i \in \{1, \cdots, n\}$, $(a_i^{(k)})_{k \in \mathbb{N}}$ is a Cauchy sequence. Since K is complete, each $(a_i^{(k)})_{k \in \mathbb{N}}$ converges to some $a_i \in K$, hence any Cauchy sequence is convergent. $\square$

4.8 Compactness

4.8.1 Def

Let X be a topological space, $Y \subseteq X$. We call open cover of Y any family $(U_i)_{i \in I}$ s.t. $Y \subseteq \bigcup\limits_{i \in I} U_i$. 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. $Y \subseteq \bigcup\limits_{j \in J} U_j$ 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.

4.8.2 Def

Let X be a set and $\mathscr{F}$ be a filter on X.

If there doesn't exist any filter $\mathscr{F}'$ of X s.t. $\mathscr{F} \subsetneq \mathscr{F}'$, then we say that $\mathscr{F}$ is an ultrafilter.

Zorn's Lemma implies that for any filter $\mathscr{F}_0$ of X, there exists an ultrafilter $\mathscr{F}$ of X containing $\mathscr{F}_0$.

4.8.3 Prop

Let $\mathscr{F}$ be a filter on a set X. The following statements are equivalent.

(1) $\mathscr{F}$ is an ultrafilter.

(2) $\forall A \subseteq X$, either $A \in \mathscr{F}$ or $X \setminus A \in \mathscr{F}$.

(3) $\forall (A, B) \in \mathscr{P}(X)^2$, if $A \bigcup B \in \mathscr{F}$, then $A \in \mathscr{F}$ or $B \in \mathscr{F}$.

Proof

From (1) to (2): Suppose thagt $A \in \mathscr{P}(X)$ s.t. $A \notin \mathscr{F}$ and $X \setminus A \notin \mathscr{F}$.

For any $B \in \mathscr{F}$. One has $B \bigcap A \neq \varnothing$, since otherwise $B \subseteq X \setminus A$ and hence $X \setminus A \in \mathscr{F}$, contradiction!

Hence $\mathscr{F} \bigcup \{A\}$ generates some filter strictly larger than $\mathscr{F}$, $\mathscr{F}$ cannot be an ultrafilter.

From (2) to (3): Suppose $B \notin \mathscr{F}$, then $X \setminus B \in \mathscr{F}$

$(A \bigcup B) \bigcap (X \setminus B) = A \setminus B \in \mathscr{F}$. So $A \in \mathscr{F}$.

From (3) to (1): Suppose $\mathscr{F}'$ is a filter s.t. $\mathscr{F} \subsetneq \mathscr{F}'$. Take $A \in \mathscr{F}' \setminus \mathscr{F}$. Then by $X = A \bigcup (X \setminus A) \in \mathscr{F}$. Hence $X \setminus A \in \mathscr{F} \subseteq \mathscr{F}'$. $\varnothing = A \bigcap (X \setminus A) \in \mathscr{F}'$ is impossible. $\square$

4.8.4 Theorem

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.

Proof

From (1) to (2): Assume that a filter $\mathscr{F}$ of X doesn't have any accumulation point.

$\forall x \in X$, $\exists A_x \in \mathscr{F}$, $\exists$ open neighborhood V_x of x s.t. $A_x \bigcap V_x = \varnothing$. Since $X = \bigcup\limits_{x \in X} V_x$, there is $\{x_1, \cdots, x_n\} \subseteq X$, s.t. $X = \bigcup\limits_{i=1}^{n} V_{x_i}$. Take $B = \bigcap\limits_{i=1}^{n} A_{x_i} \in \mathscr{F}$, $B \bigcap X = B = \varnothing$. Since $B \bigcap V_{x_i} = \varnothing$, $\forall i \in \{1, \cdots, n\}$.

From (2) to (3): Let $\mathscr{F}$ be an ultrafilter of X, by (2), there is $x \in X$ s.t. $\mathscr{F} \bigcup V_x$ generates a filter $\mathscr{F}'$. Since $\mathscr{F}$ is an ultrafilter $\mathscr{F} = \mathscr{F}'$ and hence $V_x \subseteq \mathscr{F}$.

From (3) to (1): Let $(U_i)_{i \in I}$ be an open cover of X. Suppose that this cover doesn't have any finite subcover. For any $i \in I$, let $F_i = X \setminus U_i$. For any $J \subseteq I$ finite, $F_J = \bigcap\limits_{j \in J} F_j = X \setminus \bigcup\limits_{j \in J} U_j \neq \varnothing$.

Let $\mathscr{F}$ be the smallest filter on X that contains $\{F_J | J \subseteq I \text{ finite}\}$.

Let $\mathscr{F}$ be an ultrafilter containing $\mathscr{F}$. 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 $V_x \subseteq \mathscr{F}'$, we get $U_i \in \mathscr{F}'$. This is impossible since $F_i \in \mathscr{F}'$. $\square$

4.8.5 Theorem

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. $X = \bigcup\limits_{x \in X_\varepsilon} B(x, \varepsilon)$

(2) X is compact.

Proof

From (1) to (2): Let $\mathscr{F}$ be an ultrafilter.

Let $\varepsilon > 0$ and $\{x_1, \cdots, x_n\} \subseteq X$ s.t. $X = \bigcup\limits_{i=1}^{n} B(x, \varepsilon)$.

There exists some $i \in \{1, \cdots, n\}$ s.t. $B(x_i, \varepsilon) \in \mathscr{F}$.

Therefore, $\mathscr{F}$ is a Cauchy filter. (For any $\delta > 0$, $\exists A \in \mathscr{F}$, whose diameter $\leq \delta$)

Since X is complete, $\mathscr{F}$ has a limit point. So $\mathscr{F}$ is compact.

From (2) to (1): Let $\varepsilon > 0$, one has $X = \bigcup\limits_{x \in X} B(x, \varepsilon)$. Since X is compact, $\exists X_\varepsilon \subseteq X$ finite s,t, $X = \bigcup\limits_{x \in X_\varepsilon} B(x, \varepsilon)$

Let $\mathscr{F}$ be a Cauchy filter. Let $x \in X$ be an accumulation point of $\mathscr{F}$. For any $\varepsilon > 0$, $\exists A \in \mathscr{F}$ with diameter $\leq \frac{\varepsilon}{2}$. Note that $A \bigcap B(x, \frac{\varepsilon}{2}) \neq \varnothing$. Take $y \in A \bigcap B(x, \frac{\varepsilon}{2})$, $\forall z \in A$, $d(x, z) \leq d(x, y) + d(y, z) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Thus $z \in B(x, \varepsilon)$.

Therefore, $A \subseteq B(x, \varepsilon)$. So $B(x, \varepsilon) \in \mathscr{F}$. This implies $V_x \subseteq \mathscr{F}$. □

4.8.6 Lemma

Let (X, d) be a metric space, X is complete iff any Cauchy filter of X has a limit point.

Proof

Suppose that X is complete. Let $\mathscr{F}$ be a Cauchy filter. $\forall n \in \mathbb{N}_{\geq 1}$, let $A_n \in \mathscr{F}$ s.t. $\operatorname{diam}(A_n) \leq \frac{1}{n}$. Take $x_n \in \bigcap\limits_{k=1}^{n} A_k \in \mathscr{F}$. Then $(x_n)_{n \in \mathbb{N}_{\geq 1}}$ is a Cauchy sequence since $\forall \varepsilon > 0$, if we take $N \in \mathbb{N}$ with $\frac{1}{N} \leq \varepsilon$, 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 $(x_n)_{n \in \mathbb{N}_{\geq 1}}$ converges to some $x \in X$. Note that x is a limit point of $\mathscr{F}$. In fact, $\forall \varepsilon > 0$, $\exists n \in \mathbb{N}$ with $A_n \subseteq B(x, \varepsilon)$, which implies $B(x, \varepsilon) \in \mathscr{F}$.

Let $(x_n)_{n \in \mathbb{N}}$ be a Cauchy sequence in X. Let $\mathscr{F} = \{A \subseteq X | \exists N \in \mathbb{N}, \{x_N, x_{N+1}, \cdots\} \subseteq A\}$. This is a Cauchy filter on X since $\lim\limits_{N \to +\infty} \operatorname{diam}\{x_N, x_{N+1}, \cdots\} = 0$. Hence $\mathscr{F}$ has a limit point $x \in X$.

By definition, $\forall U \in V_x$, $\exists N \in \mathbb{N}$, $\{x_N, x_{N+1}, \cdots\} \subseteq U$. So $x = \lim\limits_{n \to +\infty} x_n$. □

4.8.7 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.

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. So $\exists J \subseteq I$ s.t. $A \subseteq \bigcup\limits_{j\in J} f^{-1}(V_j)$.

This implies $f(A) \subseteq \bigcup\limits_{j\in J} V_j$. So $f(A)$ is quasi-compact.

4.8.8 Prop

Let X be a topological space and $A \subseteq X$ be a quasi-compact subset. For any closed subset F of X, $A\bigcap F$ is quasi-compact.

Proof

Let $(U_i)_{i\in I}$ be an open cover of $A\bigcap F$. Then $A \subseteq (\bigcup\limits_{i\in I} U_i)\bigcup(X \setminus F)$. Since A is quasi-compact, there exists $J \subseteq I$ s.t. $A \subseteq (\bigcup\limits_{j\in J} U_j)\bigcup(X \setminus F)$. Hence $A\bigcap F \subseteq \bigcup\limits_{j\in J} U_j$.

4.8.9 Prop

Let X be a Hausdorff space. Any compact subset A of X is closed.

Proof

Let $x \in X \setminus A$. $\forall y \in A$, $\exists$ open subsets U_y and V_y s.t. $y \in U_y$, $x \in V_y$ and $U_y \bigcap V_y = \varnothing$.

Since $A \subseteq \bigcup\limits_{y\in A} U_y$, $\exists\{y_1, \cdots, y_n\} \subseteq A$, s.t. $A \subseteq \bigcup\limits_{i=1}^{n} U_{y_i}$.

Let $U = \bigcup\limits_{i=1}^{n} U_{y_i}$ $V = \bigcap\limits_{i=1}^{n} V_{y_i}$. These are open subsets. Moreover $A \subseteq U$, $x \in V$, and $U \bigcap V = \bigcup\limits_{i=1}^{n}(U_{y_i} \bigcap V) = \varnothing$. In particular, $x \in V \subseteq X \setminus A$.

So $X \setminus A$ is open. $\square$

4.8.10 Prop

Let X be a Hausdorff topological space and A and B be compact subsets of X s.t. $A\bigcap B = \varnothing$. Then there exist open subsets U and V s.t. $A \subseteq U$, $B \subseteq V$ and $U \bigcap V = \varnothing$.

Proof

We have seen in the proof of the previous proposition that $\forall x \in B$, $\exists U_x$ and V_x open s.t. $A \subseteq U_x$, $x \in V_x$ and $U_x \bigcap V_x = \varnothing$. Since $B \subseteq \bigcup\limits_{x\in B} V_x$.

$\exists\{x_1, \cdots, x_m\} \subseteq B$ s.t. $B \subseteq \bigcup\limits_{i=1}^{m} V_{x_i}$.

We take $U = \bigcap\limits_{i=1}^{m} U_{x_i}$, $V = \bigcup\limits_{i=1}^{m} V_{x_i}$. One has $A \subseteq U$, $B \subseteq V$, $U \bigcap V = \varnothing$. $\square$

4.8.11 Theorem

Let (X, τ) be a Hausdorff topological space. If $(A_n)_{n\in\mathbb{N}}$ is a sequence of non-empty compact subsets of X s.t. $A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots$, then $\bigcap\limits_{n\in\mathbb{N}} A_n \neq \varnothing$.

Proof

Suppose that $\bigcap\limits_{n\in\mathbb{N}} A_n = \varnothing$, then $A_0 \subseteq \bigcup\limits_{n\in\mathbb{N}} (X \setminus A_n)$. Since A_0 is compact, $\exists N \in \mathbb{N}$, s.t. $A_0 \subseteq \bigcup\limits_{n=0}^{N} (X \setminus A_n) = X \setminus \bigcap\limits_{n=0}^{N} A_n = X \setminus A_N$

So $A_N = \varnothing$.(Since $A_N \subseteq A_0 \subseteq X \setminus A_N$). $\square$

The rest part of this section, we fix a complete valued field $(K, |\cdot|)$.

4.8.12 Prop

Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K. Assume that E is finite dimensional. Then any K-linear mapping $\varphi : E \to F$ is bounded and continuous.

Proof

Let $(e_i)_{i=1}^n$ be a basis of E. $\forall (a_1, \cdots, a_n) \in K^n$, $\|a_1e_1 + \cdots + a_ne_n\|_E = \max\{|a_1|, \cdots, |a_n|\}$.

$\forall s = a_1e_1 + \cdots + a_ne_n$. $\|\varphi(s)\|_F = \|\varphi(a_1e_1 + \cdots + a_ne_n)\|_F \leq |a_1|\|\varphi(e_1)\|_F + \cdots + |a_n|\|\varphi(e_n)\|_F \leq (\sum\limits_{i=1}^{n} \|\varphi(e_i)\|_F)\|s\|_E$. $\square$

4.8.13 Def

Let (X, τ) be a topological space. If any sequence in X has a convergent subsequence, we say that X is sequentially compact.

4.8.14 Def

Let (X, τ) be a topological space and $p \notin X$.

$X^* = X \bigcup \{p\}$ and the $\tau^* = \tau \bigcup \{\{p\} \cup (X \backslash K) | K$ is compact and closed$\}$ with this topology X^* is quasi-compact and we call this "one-point-compactification". Moreover, if X is hausdorff, then X^* is compact.

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.

4.8.15 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.

Proof

From (1) to (2): Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X. Assume $(x_n)_{n \in \mathbb{N}}$ doesn't have any convergent subsequence in X. For any $p \in X$, there exists $\varepsilon_p > 0$, s.t. $\{n \in \mathbb{N} | d(p, x_n) < \varepsilon_p\}$ is finite.

Since (X, d) is compact and $X = \bigcup_{p \in X} B(p, \varepsilon_p)$, then $X = \bigcup_{i=1}^{n} B(p_i, \varepsilon_{p_i})$, which is absurd.

From (2) to (1): Let $(x_n)_{n \in \mathbb{N}}$ be a Cauchy sequence. By (2) it contains a convergent subsequence, hence $(x_n)_{n \in \mathbb{N}}$ converges to the same limit. So (X, d) is complete.

Let $\varepsilon > 0$. If X is not covered by finitely many balls of radius ε, then we can construct a sequence $(x_n)_{n \in \mathbb{N}}$, s.t. $x_{n+1} \in X \setminus \bigcup_{k=0}^{n} B(x_k, \varepsilon)$. Any subsequence of this one is not a Cauchy sequence thus it is not convergent, contradiction!

By previous theorem, X is compact. $\square$

4.8.16 Def

Let X be a Hausdorff topological space. If for any $x \in X$, there exists a compact neighborhood C_x, we say X is locally compact.

Example

$\mathbb{R}$ is locally compact. $C_x = [x-1, x+1]$.

4.8.17 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 iff it is closed and bounded.

Proof

Let $Y \subseteq E$ be compact. Then Y is closed.

Moreover, $Y \subseteq \bigcup\limits_{n \in \mathbb{N}_{\geq 1}} B(0, n)$. We can find finitely many positive integers $n_1 < n_2 < \cdots < n_k$, s.t. $Y \subseteq B(0, n_1) \bigcup \cdots \bigcup B(0, n_k) \subseteq B(0, n_k)$, so Y is bounded.(This implication holds in any Hausdorff space.)

Let $(e_i)_{i=1}^d$ be a basis of E. Again we assume $\|a_1 e_1 + \cdots + a_d e_d\| = \max\{|a_1|, \cdots, |a_d|\}$.

Let $(a_1^{(n)} e_1 + \cdots + a_d^{(n)} e_d)_{n \in \mathbb{N}}$ be a sequence in Y. Since Y is bounded, for any $i \in \{1, \cdots, d\}$, the sequence $(a_i^{(n)})_{n \in \mathbb{N}}$ is bounded. In particular, we find $M > 0$ s.t. $|a_i^{(n)}| < M (\forall i \in \{1, \cdots, d\})$.

Since $(K, |\cdot|)$ is locally compact there is a compact set $C = C_0$ on K that is a neighborhood of 0. Let $\varepsilon > 0$, $\overline{B}(0, \varepsilon) \subseteq C$. Since the absolute value is not trivial, there exists $a \in K$, s.t. $|a| \geq \frac{M}{\varepsilon}$. Then $\overline{B}(0, M) \subseteq aC$.

$C \subseteq K$ is compact. We have a K-linear mapping $K \to K$, $y \mapsto ay$, this map is bounded, hence it is a continuous map. Hence aC is compact. $\overline{B}(0, M) \subseteq aC$ is a closed subspace of a compact, so it is compact, hence it is sequentially compact.

Therefore, we can find a $I_1, \cdots, I_d$ that an infinite subsets of $\mathbb{N}$, with $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_d$, s.t. $(a_i)_{n \in I_i}^{(n)}$ converges.

They converge to some $a_i \in K$. It follows that our original sequence has a convergent subsequence converges to $a_1 e_1 + \cdots + a_d e_d$. Y is sequentially compact. $\square$

4.8.18 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. $\forall x \in Y$, $f(a) \leq f(x) \leq f(b)$. Namely the restriction of f to Y attains its maximum and minimum.

Proof

$f(Y) \subseteq \mathbb{R}$ is a compact subset since Y is quasi-compact and $\mathbb{R}$ is hausdorff. Moreover, since $\mathbb{R}$ is locally compact. $f(Y)$ is bounded and closed.

Note there exists sequences $(\alpha_n)_{n\in\mathbb{N}}$ and $(\beta_n)_{n\in\mathbb{N}}$ in $f(Y)$ that tend to $\sup f(Y)$ and $\inf f(Y)$ respectively.(In fact, Let $M = \sup f(Y) \in \mathbb{R}$, $\forall n \in \mathbb{N}$, $n \neq 0$, $M - \frac{1}{n}$ is not an upper bound of $f(Y)$, $\exists \alpha_n \in f(Y)$ s.t. $M - \frac{1}{n} < \alpha_n \leq M$) Since $f(Y)$ is closed, $\sup f(Y)$ and $\inf f(Y)$ belongs to $f(Y)$. So $f(Y)$ has a greatest and a least element. $\square$

4.9 Mean value theorems

4.9.1 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 $\exists t \in]a, b[$, s.t. $f'(t) = 0$.

Proof

Since $[a, b]$ is compact, f attains its maximum and minimum. Let $M = \max f([a, b])$, $m = \min f([a, b])$, $l = f(a) = f(b)$.

If $M \neq l$, $\exists t \in]a, b[$, s.t. $f(t) = M$.

$f(t + x) = f(t) + f'(t)x + o(|x|)$

$f(t - x) = f(t) - f'(t)x + o(|x|)$

$0 \leq (f(t + x) - f(t))(f(t - x) - f(t)) = -f'(t)^2x^2 + o(|x|^2)$

$0 \leq -f^{'}(t)^2 + o(1)$, $x \to 0$, taking the limit when $x \to 0$, we get $f'(t) = 0$

If $m \neq l$, then any $t \in]a, b[$, s.t. $f(t) = m$ verifies $f'(t) = 0$

If $m = l = M$, f is constant, so $\forall t \in]a, b[$, $f^{'}(t) = 0$. □

Note

The reason why $0 \leq -f^{'}(t)^2 + o(1)$, $x \to 0$, may have a more precise explanation.

Not loss generality, let $x \geq 0$. Since $|f(t+x) - f(t) - f'(t)x| \leq \varepsilon_1(x)x$, thus $-\varepsilon_1(x)x \leq f(t + x) - f(t) - f'(t)x \leq \varepsilon_1(x)x$. Let $f(t + x) - f(t) = f'(t)x + \varepsilon_3(x)x$, $\lim\limits_{x\to 0} \varepsilon_3(x) = 0$. Similarly, one can obtain $f(t - x) - f(t) = -f'(t)x + \varepsilon_4(x)x$. $0 \leq (f(t + x) - f(t))(f(t - x) - f(t)) = x^2(-f'(t)^2 - \varepsilon_3(x)f'(t) + \varepsilon_4(x)f'(t) + \varepsilon_3(x)\varepsilon_4(x))$, take $\lim\limits_{x\to 0}$, one can obtain $f'(t) = 0$

4.9.2 Theorem(Mean value theorem, Lagrange)

Let a and b be two different real numbers, $f : [a, b] \to \mathbb{R}$ be a continuous mapping, that is differentiable on $]a, b[$. Then there exists $t \in]a, b[$ s.t. $f(b) - f(a) = f'(t)(b - a)$.

Proof

Let $g : [a, b] \to \mathbb{R}$ be defined as $g(x) = f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$

$g(a) = g(b) = f(a)$

$g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$, by Rolle's theorem, $\exists t \in]a, b[$, s.t. $g'(t) = 0$, that is $f'(t) = \frac{f(b)-f(a)}{b-a}$. $\square$

4.9.3 Theorem

Let E and F be normed vector spaces 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.

Proof

$f(x) = f(p)+d_pf(x-p)+o(\|x-p\|) = f(p)+O(\|x-p\|) = f(p)+o(1)$, $x \to p$. $\square$

4.9.4 Theorem (Mean value inequality)

Let a and b be real numbers s.t. $a < b$. $(E, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$. $f : [a, b] \to E$ be a continuous mapping s.t. f is differentiable on $]a, b[$. Then $\|f(b) - f(a)\| \leq \sup\limits_{x\in]a,b[} \|f'(x)\|(b - a)$.

Proof

Suppose that $\sup\limits_{x\in]a,b[} \|f'(x)\| < +\infty$.

Let $M \in \mathbb{R}$, s.t. $M > \sup\limits_{x\in]a,b[} \|f'(x)\|$

Let $J = \{x \in [a, b] | \forall y \in [a, x], \|f(y) - f(a)\| \leq M(y - a)\}$

By definition, 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 $\|f(c) - f(a)\| = \lim\limits_{n\to+\infty} \|f(c_n) - f(a)\| \leq \lim\limits_{n\to+\infty} M(c_n - a) = M(c - a)$. Hence $c \in J$, namely $J = [a, c]$

(**Unstrictly**) We assume that $c > a$.

We argue $c = b$ by contradiction. Suppose that $c < b$, $\forall h \in]0, b - c[$, $\|f(c + h) - f(c)\| = \|hf'(c) + o(h)\| \leq \|f'(c)\|h + o(h)$, $h \to 0$

Since $M > \|f'(c)\|$, $\exists h_0 > 0$, s.t. $\forall 0 < h < h_0$, $\|f(c+h) - f(c)\| \leq Mh$. Hence $\|f(c + h) - f(a)\| \leq \|f(c + h) - f(c)\| + \|f(c) - f(a)\| \leq M(c + h -$

$c + c - a) = M(c + h - a)$, so $c + h_0 \in J$, contradiction!

Since M is arbitrary, the expected inequality holds. □

Remark

If f is defined on an open neighborhood of a and is differentiable at a, then the same arguments hold without the assumption $c > a$.

In general we apply the particular case (f is differentiable mapping at a to $[\frac{a+b}{2}, b]$ and $[a, \frac{a+b}{2}]$) to get $\|f(b) - f(\frac{a+b}{2})\| \leq C\frac{b-a}{2}$, $\|f(\frac{a+b}{2}) - f(a)\| \leq C\frac{b-a}{2}$, with $C = \sup_{x\in]a,b[} \|f'(x)\|$.

4.9.5 Theorem

Let I be an interval in $\mathbb{R}$ and $f : I \to \mathbb{R}$ be a continuous mapping, then $f(I)$ is an interval.

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 loss of generality, we suppose $a < b$. Let $z \in \mathbb{R}$ s.t. $(z - x)(z - y) \leq 0$.

We construct 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) $a_0 = a$, $b_0 = b$, $c_0 = \frac{a+b}{2}$

(2) If a_n, b_n and c_n are constructed, satisfying $c_n = \frac{1}{2}(a_n + b_n)$, $(z - f(a_n))(z - f(b_n)) \leq 0$.

$(a_{n+1}, b_{n+1}) = (a_n, c_n)$ if $(z - f(a_n))(z - f(c_n)) \leq 0$

$(a_{n+1}, b_{n+1}) = (c_n, b_n)$ if $(z - f(a_n))(z - f(c_n)) > 0$

$c_{n+1} = \frac{a_{n+1}+b_{n+1}}{2}$

The sequence $(a_n)_{n\in\mathbb{N}}$ is increasing and bounded, $(b_n)_{n\in\mathbb{N}}$ is decreasing and bounded. Hence converges to some $l \in [a, b]$, $,m \in [a, b]$.

Note that $|b_n - a_n| = \frac{1}{2^n}|b - a| \to 0$ $(n \to +\infty)$ so $l = m$.

By $(z - f(a_n))(z - f(b_n)) \leq 0$, we obtain by letting $n \to +\infty$, $(z - f(l))^2 \leq 0$, so $z = f(l)$. □

4.9.6 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.

Proof

Let $(a,b) \in I \times I$ s.t. $a < b$. Consider the following mappings. $g : [a,b] \to \mathbb{R}$, $g(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \neq a \\ f'(a) & \text{if } x = a \end{cases}$, $h : [a,b] \to \mathbb{R}$, $h(x) = \begin{cases} \frac{f(b)-f(x)}{b-x} & \text{if } x \neq b \\ f'(b) & \text{if } x = b \end{cases}$. g and h are continuous.($\frac{f(x)-f(a)}{x-a} = f'(a) + o(1)$, $x \to a$)

So $g([a,b])$ and $h([a,b])$ are intervals. By the mean value theorem, $g([a,b]) \subseteq f'(I)$, $h([a,b]) \subseteq f'(I)$. Moreover, $\{f'(a), f'(b)\} \subseteq g([a,b]) \bigcup h([a,b]) \subseteq f'(I)$.

Since $g(b) = h(a)$, $g([a,b]) \bigcup h([a,b])$ is an interval. Hence $f'(I)$ is an interval. $\square$

4.10 Fixed point theorem

4.10.1 Def

Let X be a set and $T : X \to X$ be a mapping. If $x \in X$ satisfies $T(x) = x$, we say that x is a fixed point of T.

4.10.2 Def

Let (X, d) be a metric space, and $T : X \to X$ be a mapping. If $\exists \varepsilon \in [0, 1[$ s.t. T is ε-Lipschitzian.($d(T(x), T(y)) \leq \varepsilon d(x, y)$) then we say that T is a contraction.

4.10.3 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, $\forall x_0 \in X$, if we let $x_n = (T \circ \cdots \circ T)(x)$, then $(x_n)_{n \in \mathbb{N}}$ converges to the fixed point.

Proof

If p and q are two fixed point of T, then $d(p, q) = d(T(p), T(q)) \leq \varepsilon d(p, q)$, so $d(p, q) = 0$

Let $x_0 \in X$, $x_n = T \circ \cdots \circ T(x_0)$, $x_{n+1} = T(x_n)$

$\forall n \in \mathbb{N}$, $d(x_n, x_{n+1}) \leq \varepsilon^n d(x_0, x_1)$

For any $N \in \mathbb{N}$, $\forall (n, m) \in \mathbb{N}_{\geq N}$, $n < m$. $d(x_n, x_m) \leq \sum\limits_{k=n}^{m-1} d(x_k, x_{k+1}) \leq \sum\limits_{k=n}^{m-1} \varepsilon^k d(x_0, x_1) \leq \frac{\varepsilon^n}{1-\varepsilon} d(x_0, x_1) \leq \frac{\varepsilon^N}{1-\varepsilon} d(x_0, x_1)$

So $\lim\limits_{N \to +\infty} \sup\limits_{(n,m) \in \mathbb{N}_{\geq N}} d(x_n, x_m) = 0$, $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence, hence converges to some $p \in X$. $d(T(p), p) = \lim\limits_{n \to +\infty} d(T(x_{n-1}), x_n) = 0$, since d is continuous. $\square$

Chapter 5

Higher differentials

5.1 Multilinear mapping

Let K be a commutative unitary ring.

5.1.1 Def

Let $n \in \mathbb{N}$, $V_1, \cdots, V_n$, W be K-module. 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. $\forall i \in \{1, \cdots, n\}$, $\forall (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \in V_1 \times V_{i-1} \times V_{i+1} \times \cdots V_n$ the mapping $f(x_1, \cdots, x_{i-1}, \cdot, x_{i+1}, \cdots, x_n) : V_i \to W$, $x_i \mapsto f(x_1, \cdots, x_n)$ is a morphism of K-module.

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.

Example

$K \times K \to K$, $(a, b) \mapsto ab$ is a 2-linear mapping(bilinear mapping).

Remark

When $n = 0$, $\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W)$ is considered as W by convention.

$\mathrm{Hom}^{(1)}(V_1, W) = \mathrm{Hom}(V_1, W) = \{$morphisms of K-module from V_1 to $W\}$.

5.1.2 Prop

Suppose that $n \geq 2$. For any $i \in \{1, \cdots, n-1\}$. $\Phi : \mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W) \to \mathrm{Hom}^{(i)}(V_1 \times \cdots \times V_i, \mathrm{Hom}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W))$, $f \mapsto ((x_1, \cdots, x_i) \mapsto ((x_{i+1}, \cdots, x_n) \mapsto f(x_1, \cdots, x_n)))$ is a bijection.

Proof

The inverse of Φ is given by $g \in \mathrm{Hom}^{(i)}(V_1 \times \cdots \times V_i, \mathrm{Hom}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W)) \mapsto (((x_1, \cdots, x_n) \in V_1 \times \cdots \times V_n) \mapsto g(x_1, \cdots, x_i)(x_{i+1}, \cdots, x_n))$.

$\square$

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 $K-$module.

5.2 Operator norm of multilinear mappings

Let $(K, |\cdot|)$ be a complete valued field.

5.2.1 Def

Let $V_1, \cdot, V_n$ and W be normed vector spaces over K. We define $\|\cdot\| :$ $\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W) \mapsto [0, +\infty]$ as $\|f\| := \sup_{(x_1,\cdots,x_n) \in V_1 \times \cdots \times V_n \setminus \{0\}} \frac{\|f(x_1,\cdots,x_n)\|}{\|x_1\| \cdots \|x_n\|}$

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.

5.2.2 Theorem

For any $i \in \{1, \cdots, n-1\}$, $\forall f \in \mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W)$.$\forall (x_1, \cdots, x_i) \in V_1 \times \cdots \times V_i$ the $(n-i)$-linear mapping $f(x_1, \cdots, x_i, \cdot) : V_{i+1} \times \cdots \times V_n \to W$, $(x_{i+1}, \cdots, x_n) \mapsto f(x_1, \cdots, x_n)$ belongs to $\mathscr{L}^{n-i}(V_{i+1} \times \cdots \times V_n, W)$.

Moreover $\|f\| = \sup_{(x_1,\cdots,x_i) \in V_1 \times \cdots \times V_i \setminus \{0\}} \frac{\|f(x_1,\cdots,x_i,\cdot)\|}{\|x_1\| \cdots \|x_i\|}$.

Proof

$\forall (x_{i+1}, \cdots, x_n) \in V_{i+1} \times \cdots \times V_n$, $\|f(x_1, \cdots, x_n)\| \leq \|f\| \|x_1\| \cdots \|x_n\| = (\|f\| \|x_1\| \cdots \|x_i\|) \|x_{i+1}\| \cdots \|x_n\|$. So $\|f(x_1, \cdots, x_i, \cdot)\| \leq \|f\| \|x_1\| \cdots \|x_i\|$. If we define $\|f\|' = \sup_{(x_1,\cdots,x_i) \in V_1 \times \cdots \times V_i \setminus \{0\}} \frac{\|f(x_1,\cdots,x_i,\cdot)\|}{\|x_1\| \cdots \|x_i\|}$, then $\|f\|' \leq \|f\|$.

Conversely, $\forall (x_1, \cdots, x_n) \in V_1 \times \cdots \times V_n$, $\|f(x_1, \cdots, x_n)\| \leq \|f(x_1, \cdots, x_n, \cdot)\| \cdots \|x_{i+1}\| \|x_n\|$

Hence $\frac{\|f(x_1,\cdots,x_n)\|}{\|x_1\| \cdots \|x_n\|} \leq \frac{\|f(x_1,\cdots,x_i,\cdot)\|}{\|x_1\| \cdots \|x_i\|} \leq \|f\|'$

Taking sup, we get $\|f\| \leq \|f\|'$. $\square$

5.2.3 Corollary

(1) $\mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W)$ is a vector subspace of $\mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W)$

(2) $\|\cdot\|$ is a norm on $\mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W)$

(3) $\forall i \in \{1, \cdots, n\}$, $\mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W) \to \mathscr{L}^{(i)}(V_1 \times \cdots \times V_i, \mathscr{L}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W))$ is a K-linear isomorphism that preserves operator norms.$\|f\| = \|\Phi(f)\|$.

Proof

We reason by induction on n.

$n = 1$, $\mathscr{L}^{(1)}(V_1, W) = \mathscr{L}(V_1, W)$.

Suppose that the corollary is true for m-linear mappings with $m < n$. Take $i \in \{1, \cdots, n-1\}$. We consider the following diagram of mapping.

$$\begin{array}{ccc} \mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W) & \xrightarrow{\Phi} & \mathscr{L}^{(i)}(V_1 \times \cdots \times V_i, \mathscr{L}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W)) \\ \cap\!\!| & & \cap\!\!| \\ \mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, W) & \xrightarrow{\Phi} & \mathrm{Hom}^{(i)}(V_1 \times \cdots \times V_i, \mathrm{Hom}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W)) \end{array}$$

To show that $\mathscr{L}^{(n)}(V_1 \times \cdots \times V_n, W)$ is a vector subspace, it suffices to check that $\forall g \in \mathscr{L}^{(i)}(V_1 \times \cdots \times V_i, \mathscr{L}^{(n-i)}(V_{i+1} \times \cdots \times V_n, W))$, one has $\|\Phi^{-1}(g)\| = \|g\| < +\infty$.

For any $(x_1, \cdots, x_n) \in V_1 \times \cdots \times V_n$.

$\|\Phi^{-1}(g)(x_1, \cdots, x_n)\| \|g(x_1, \cdots, x_i)(x_{i+1} \cdots, x_n)\| \leq \|g(x_1, \cdots, x_i)\| \|x_{i+1}\| \cdots \|x_n\| \leq \|g\| \|x_1\| \cdots \|x_i\| \|x_{i+1}\| \cdots \|x_n\|$

Therefore, $\|\Phi^{-1}(g)\| \leq \|g\| \leq \|\Phi^{-1}(g)\|$. $\square$

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) We denote by $D^0 f$ any mapping $f : U \to F$. If f is continuous, we say that f is of class C^0. Any mapping $f : U \to F$ is considered as 0-time differential.

(2) By abuse of notation we use $D^1 f$ to denote df. 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 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 $d(D^n f)(x)$ by the K-linear bijection $\mathscr{L}(E, \mathscr{L}^{(n)}(E^n, F)) \to \mathscr{L}^{(n+1)}(E^{n+1}, F)$. $df : U \to \mathscr{L}(E, F)$, $d(D^n f)) : U \to \mathscr{L}(E, \mathscr{L}^{(n)}(E^n, F)) \cong \mathscr{L}^{n+1}(E^{n+1}, F)$. If $D^{n+1} f$ is continuous, we say that f is of class C^{n+1} $(n \geq 0)$.

Remark

If f is n-times differentiable, $\forall i \in \{1, \cdots, n-1\}$. $\forall p \in U, \forall (h_1, \cdots, h_n) \in E^n$, one has $D^i(D^{n-i} f)(p)(h_1, \cdots, h_{n-i})(h_{n-i+1}, \cdots, h_n) = D^n f(p)(h_1, \cdots, h_n)$. $D^{n-i} f : U \to \mathscr{L}^{n-i}(E^{n-i}, F)$, $D^i(D^{n-i} f) : U \to \mathscr{L}^{(i)}(E^i, \mathscr{L}^{(n-i)}(E^{n-i}, F))$.

$D^{n-i} f : U \to \mathscr{L}^{(n-i)}(E^{n-i}, F),\ D^i(D^{n-i} f) \quad :$

$$\begin{tikzcd} U \arrow[r] \arrow[dr, "D^n f"'] & \mathscr{L}^{(i)}(E^i, \mathscr{L}^{(n-i)}(E^{n-i}, F)) \arrow[d, phantom, "\cong" rotate=-90] \\ & \mathscr{L}^{(n)}(E^n, F) \end{tikzcd}$$

5.3.2 Prop(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 $\|f(b)-f(a)\| \leq g(b)-g(a)$.

Proof

Let $c \in]a,b[$. Let $\varepsilon > 0$. Let $J = \{t \in [c,b] | \forall s \in [c,t], \|f(s)-f(c)\| \leq g(s)-g(c)+\varepsilon(s-c)\}$. By definition, J is an interval.

Since f and g are continuous, J is a closed interval, hence J is of the form $[c,t]$. If $t<b$, then for $h>0$ sufficiently small $f(t+h)-f(t) = hf'(t)+o(h)$, $g(t+h)-g(t) = hg'(t)+o(h)$.

$\exists \delta > 0$, $\forall h \in [0,\delta]$, $\|f(t+h)-f(t)\| \leq \|f'(t)\|h + \frac{\varepsilon}{2}h$, $g(t+h)-g(t) \geq g'(t)h - \frac{\varepsilon}{2}h$.

So $\|f(t+h)-f(t)\| \leq g(t+h)-g(t)+\varepsilon h$. Moreover, $\|f(t)-f(c)\| \leq g(t)-g(c)+\varepsilon(t-c)$.

$\|f(t+h)-f(c)\| \leq g(t+h)-g(c)+\varepsilon(t+h-c)$. So $J \supseteq [c,t+\delta]$, contradiction! For the same reason, $\|f(c)-f(a)\| \leq g(c)-g(a)+\varepsilon(c-a)$, Hence $\|f(b)-f(a)\| \leq g(b)-g(a)+\varepsilon(b-a)$. Since $\varepsilon > 0$ is arbitrary, $\|f(b)-f(a)\| \leq g(b)-g(a)$. $\square$

Note

这个命题的价值在于没有使用积分便完成了证明。

实际上, 当 F 是 $\mathbb{R}$ 时, 上述命题是平凡的, 只需两侧同时积分即可, 以下是在 $\mathbb{R}$ 上一个更实用版本的 Gronwall 不等式。

f, g 和 h 是 $[a,b]$ 上的连续函数, $g(x)>0$, 且 $f(x) \leq h(x)+\int_a^x f(s)g(s)\mathrm{d}s$, 则 $f(x) \leq h(x)+e^{\int_a^x g(t)\mathrm{d}t}\int_a^x e^{-\int_a^s g(t)\mathrm{d}t}g(s)h(s)\mathrm{d}s$。特别的, 当 $h(x)$ 单调递增时 $f(x) \leq h(x)e^{\int_a^x g(t)\mathrm{d}t}$

设 $u(x) = \int_a^x f(s)g(s)\mathrm{d}s$, 则 $f(x) \leq h(x)+u(x)$, $u'(x) = f(x)g(x) \leq (h(x)+u(x))g(x) = h(x)g(x)+u(x)g(x)$。

两侧乘上积分因子 $\mu(x) = \int_a^x e^{-\int_a^x g(s)\mathrm{d}s}$, $(u(x)\mu(x))' = h(x)g(x)\mu(x)$, 两侧同时积分 $u(x)\mu(x) = \int_a^x h(x)g(x)\mu(x)$, 整理得到 $f(x) \leq h(x)+u(x) \leq h(x)+e^{\int_a^x g(s)\mathrm{d}s}\int_a^x h(s)g(s)e^{-\int_a^s g(t)\mathrm{d}t}\mathrm{d}s$

特别的, 当 $h(x)$ 单调递增时, $f(x) \leq h(x)(1+e^{\int_a^x g(s)\mathrm{d}s}\int_a^x g(s)e^{-\int_a^s g(t)\mathrm{d}t}\mathrm{d}s) =$

$h(x)e^{\int_a^x g(t)\mathrm{d}t}$

5.3.3 Theorem(Taylor-Langrange formula)

Let $n \in \mathbb{N}$, E and F be normed vector spaces over $\mathbb{R}$ and $U \subseteq E$ be a open set. 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. $p + th \in U$, $\forall t \in [0,1]$. Let $M = \sup\limits_{t\in]0,1[} \|D^{n+1}f(p+th)\|$. Then $\|f(p+h) - \sum\limits_{k=1}^{n} \frac{1}{k!}D^k f(p)(h,\cdots,h)\| \leq \frac{M}{(n+1)!}\|h\|^{n+1}$.

Proof

In this proof, we use 0^0 to denote 1.

Consider $\phi : [0,1] \to F$. $\phi(t) = \sum\limits_{k=0}^{n} \frac{(1-t)^k}{k!} D^k f(p+th)(h,\cdots,h)$.

$\phi(1) = f(p+h)$, $\phi(0) = \sum\limits_{k=0}^{n} \frac{1}{k!} D^k f(p)(h,\cdots,h)$

$\phi'(t) = \sum\limits_{k=0}^{n} \frac{(1-t)^k}{k!} D^{k+1} f(p+th)(h,\cdots,h) - \sum\limits_{k=1}^{n} \frac{(1-t)^{k-1}}{(k-1)!} D^k f(p+th)(h,\cdots,h) = \frac{(1-t)^n}{n!} D^{n+1} f(p+th)(h,\cdots,h)$

$\|\phi'(t)\| \leq M\frac{(1-t)^n}{n!}\|h\|^{n+1} = (-M\frac{(1-t)^{n+1}\|h\|^{n+1}}{(n+1)!})'$

By Gronwall inequality, $\|\phi(1) - \phi(0)\| \leq \frac{M}{(n+1)!}\|h\|^{n+1}$. □

5.3.4 Def

Let $n \in \mathbb{N}_{\geq 1}$, $E_1, \cdots, E_n$ 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 subset, $p = (p_1, \cdots, p_n) \in U$, $i \in \{1, \cdots, n\}$, $f : U \to F$.

If there exists an open neighborhood U_i of p_i in E_i s.t. $U_i \to F$, $x_i \mapsto f(p_1, \cdots, p_{i-1}, x_i, p_{i+1}, \cdots, p_n)$ is well defined and is differentiable at p_i. We denote by $\frac{\partial f}{\partial x_i}(p)$ the differential of this mapping $U_i \to F$ at p_i and say that f admits the i^{th} partial differentiable at p.

5.3.5 Prop

Let $K = \mathbb{R}$, suppose that f has all partial differentials on U and $\frac{\partial f}{\partial x_i} : U \to \mathscr{L}(E_i, F)$ is continuous for any $i \in \{1, \cdots, n\}$.

Then f is of class C^1, and $\forall h = (h_1, \cdots, h_n) \in E_1 \times \cdots \times E_n$. $\forall p \in U$, $d_p f(h) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)(h_i)$.

Proof

By induction, it suffices to treat the case where $n = 2$.

Let $p = (a, b) \in U$. $\forall \varepsilon > 0$, $\exists \delta > 0$, $\forall (h, k) \in E_1 \times E_2$, $\max\{\|h\|, \|k\|\} \leq \delta$, one has $\|\frac{\partial f}{\partial x_2}(a+h, b+k) - \frac{\partial f}{\partial x_2}(a, b)\| \leq \varepsilon$. (By continuity of $\frac{\partial f}{\partial x_2}$).

Consider the mapping $\phi : [0, 1] \to F$. $\phi(t) = f(a+h, b+tk) - f(a+h, b) - t\frac{\partial f}{\partial x_2}(a+h, b)(k)$.

$\phi'(t) = \frac{\partial f}{\partial x_2}(a+h, b+tk)(k) - \frac{\partial f}{\partial x_2}(a+h, b)(k)$, $\|\phi'(t)\| \leq 2\|\varepsilon\|k$.

Thus $\|\phi(1) - \phi(0)\| \leq 2\varepsilon\|k\|$. So $\|f(a+h, b+k) - f(a+h, b) - \frac{\partial f}{\partial x_2}(a+h, b)(k)\| \leq 2\varepsilon\|k\| = o(\max\{\|h\|, \|k\|\})$. $(\max\{\|h\|, \|k\|\} \to 0)$.

f has 1^{st} partial differential. $\|f(a+h, b) - f(a, b) - \frac{\partial f}{\partial x_1}(a, b)(h)\| \leq 2\varepsilon\|k\| = o(\max\{\|h\|, \|k\|\})$. $(\max\{\|h\|, \|k\|\} \to 0)$.

Continuity of $\frac{\partial f}{\partial x_2}$, $\|\frac{\partial f}{\partial x_2}(a+h, b)(k) - \frac{\partial f}{\partial x_2}(a, b)(k)\| = o(\max\{\|h\|, \|k\|\})$. $(\max\{\|h\|, \|k\|\} \to 0)$.

Taking the sum, we get $\|f(a+h, b+k) - f(a, b) - \frac{\partial f}{\partial x_1}(a, b)(h) - \frac{\partial f}{\partial x_2}(a, b)(k)\| = o(\max\{\|h\|, \|k\|\})$. □

5.3.6 Theorem

Let E and F be normed vector spaces over $\mathbb{R}$, $U \subseteq E$ open. $(f_n)_{n\in\mathbb{N}}$ 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 $f : U \to F$.

Then f is differentiable and $df = g$.

Proof

Let $p \in U$, $\forall (m, n) \in \mathbb{N}^2$.

$\|f_n(x) - f_m(x) - (f_n(p) - f_m(p))\| \leq (\sup_{\xi\in U} \|d_\xi f_m - d_\xi f_n\|)\|x - p\|$.(mean value inequality)

Taking $\lim_{m\to+\infty}$, we get $\|(f_n(x) - f(x)) - (f_n(p) - f(p))\| \leq \varepsilon_n\|x - p\|$, where $\varepsilon_n = \sup_{\xi\in U} \|d_\xi f_n - g(\xi)\|$.

So $\|f(x) - f(p) - g(p)(x-p)\| \leq \|(f(x) - f_n(x)) - (f(p) - f_n(p))\| + \|f_n(x) - f_n(p) - d_p f_n(x-p)\| + \|d_p f_n(x-p) - g(p)(x-p)\| \leq \varepsilon_n\|x - p\| + \|f_n(x) - f_n(p) - d_p f_n(x-p)\| + \varepsilon_n\|x - p\|$

$\limsup_{x\to p} \frac{\|f(x)-f(p)-g(p)(x-p)\|}{\|x-p\|} \leq 2\varepsilon_n$

Take $\lim_{n\to+\infty}$, get $\limsup_{x\to p} \frac{\|f(x)-f(p)-g(p)(x-p)\|}{\|x-p\|} = 0$。 □

5.4 Symmetric group

5.4.1 Def

Let X be a set, we denote with $\mathfrak{S}_X$ the set of all bijections from X to itself. The elements of $\mathfrak{S}_X$ are called permutations if the set X is finite. If $x_1, \cdots, x_n \in X$ are distinct element then $(x_1, \cdots, x_n) \in \mathfrak{S}_X$ s.t. $x_i \mapsto x_{i+1}$, $i = 1, \cdots, n-1$, $x_n \mapsto x_1$ this is called a n-cycle. A 2-cycle is called a transposition.

5.4.2 Def

We denote by $\operatorname{orb}_\sigma(x) = \{\sigma \circ \cdots \circ \sigma(x)(\text{n times}), n \in \mathbb{N}\}, x \in X, \sigma \in \mathfrak{S}_X$.

5.4.3 Prop

If $\operatorname{orb}_\sigma(x)$ is a finite set of d elements, then one has $\sigma^d(x) = x$, $\operatorname{orb}_\sigma(x) = \{x, \sigma(x), \sigma^2(x), \cdots, \sigma^{d-1}(x)\}$, moreover $\sigma^{-1}(x) \in \operatorname{orb}_\sigma(x)$.

Proof

The set $\{(n,m) \in \mathbb{N}^2 | n < m, \sigma^n(x) = \sigma^m(x)\}$ is not empty. Let $d' = \min\{m - n | (n,m) \in \mathbb{N}^2, n < m, \sigma^n(x) = \sigma^m(x)\}$. Therefore, $x, \sigma(x), \cdots, \sigma^{d-1}(x)$ are all distinct.

$\{x, \sigma(x), \sigma^2(x) \cdots, \sigma^{d-1}(x)\} \subseteq \operatorname{orb}_\sigma(x)$, $d' \leq d$. Now use the Euclidean division, $h = qd' + \tau (0 \leq \tau \leq d')$, $\sigma^h(x) = \sigma^\tau(x)$.

$\sigma^{-1}(x) = \sigma^{d-1}(x) \in \operatorname{orb}_\sigma(x)$. □

Remark

Let $Y \subseteq X$, then we have a morphism of groups, $\mathfrak{S}_Y \to \mathfrak{S}_X$, $\sigma \mapsto (x \mapsto \begin{cases} \sigma(x) & \text{if } x \in Y, \\ x & \text{if } x \notin Y \end{cases})$.

If Y and Z are subsets of X, $Y \bigcap Z = \varnothing$, $\sigma \in \mathfrak{S}_Y$, $\tau \in \mathfrak{S}_Z$, then $\sigma \circ \tau = \tau \circ \sigma$.

5.4.4 Theorem

Let X be a finite set and let $\sigma \in \mathfrak{S}_X$. There exists $d \in \mathbb{N}$ and $(n_1, \cdots, n_d) \in \mathbb{N}^d_{\geq 2}$ and pairwise disjoint subset $X_1, \cdots, X_d$ of X of cardinalities $n_1, \cdots, n_d$, together with n_i-cycles 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.

Proof

By induction on the cardinality of X. The case $\sigma = \mathrm{Id}_X$ is trivial(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 $Y = X \setminus X_1$, $\forall y \in Y$, we have that $\sigma(y) \in Y$ because if $\sigma(y) \in X_1$, by the previous proposition, $y \in X_1$.

Let $\tau = \sigma|_Y \in \mathfrak{S}_Y$, use induction hypothesis, we get $X_2, \cdots, X_d$ of cardinaties $n_2, \cdots, n_d$, 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. $\square$

Remark

The previous theorem says that the group of permutations is generated by cycles.

5.4.5 Corollary

Let X be a finite set. Then $\mathfrak{S}_X$ is generated by transposition.

Proof

Thanks to the theorem before, it is enough to decompose cycles in terms of transpositions.

$(x_1\ x_2 \cdots x_n) = (x_1\ x_2) \circ (x_2 \cdots x_n) = (x_1\ x_2) \circ (x_2\ x_3) \circ \cdots \circ (x_{n-1}\ x_n)$.

$\square$

Remark

Be careful, the decomposition of transpositions is not unique.

$|X| = n$, $\mathfrak{S}_X = S_n = \mathfrak{S}_n$.

5.4.6 Def

Consider $\mathfrak{S}_n$, a transposition is called adjacent if is of the form $(j\ j{+}1)$, for $j = 1, \cdots, n-1$.

5.4.7 Corollary

$\mathfrak{S}_n$ is generated by adjacent transpositions.

Proof

It is enough to decompose transpositions, because of the previous Corollary.

$i < j$, $(i\ j) = (i\ i{+}1)\circ(i{+}1\ i{+}2)\circ\cdots\circ(j{-}1\ j)\circ(j{-}2\ j{-}1)\circ\cdots(i\ i{+}1)$.

5.4.8 Cayley Theorem

Any finite group can be embedded(injective morphism) in a $\mathfrak{S}_n$ for some $n \in \mathbb{N}$.

Proof

Let $|G| = n$, $\varphi : G \to \mathfrak{S}_G \cong \mathfrak{S}_n$, $g \mapsto lg$, $lg(x) = gx$.

5.4.9 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 put $\mathrm{sgn}(\sigma) = (-1)^d$

This is well defined function and moreover $\mathrm{sgn} : \mathfrak{S}_X \to \{1, -1\}$ is a morphism of groups.

Proof

Let's define the map $\phi : \mathfrak{S}_X \to \mathbb{Q} \setminus \{0\}$, $\phi(\sigma) = \prod_{(i,j)\in\{1,\cdots,n\}^2} \frac{\sigma(i)-\sigma(j)}{i-j}$.

Let $\theta = \{(i, j) \in \{1, \cdots, n\}^2 | i < j\}$.

$\phi(\sigma\circ\tau) = \prod_{(i,j)\in\theta} \frac{\sigma(\tau(i))-\sigma(\tau(j))}{i-j} = (\prod_{(i,j)\in\theta} \frac{\sigma(\tau(i))-\sigma(\tau(j))}{\tau(i)-\tau(j)})(\prod_{(i,j)\in\theta} \frac{\tau(i)-\tau(j)}{i-j}) = \phi(\sigma)\phi(\tau)$

$\phi(\sigma) = -1$ if σ is a transposition. Therefore since $\sigma = \tau_1 \circ \cdots \circ \tau_d$, $\phi(\sigma) = (-1)^d$. $\square$

5.5 Symmetry of multilinear maps

In this section, we fix a commutative unitary ring K, and two $K-$modules E, F.

5.5.1 Def

Let $n \in \mathbb{N}$ and $f \in \mathrm{Hom}^{(n)}(E^n, F)$ if for any $\sigma \in \mathfrak{S}_n$ one has that $f(x_1, \cdots, x_n) = f(x_{\sigma(1)}, \cdots, x_{\sigma(n)})$. $\forall (x_1, \cdots, x_n) \in E^n$. We say tha 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.

5.5.2 rop

Suppose that $f \in \mathrm{Hom}^{(n)}(E^n, F)$ is alternating, then $f(x_1, \cdots, x_n) = \mathrm{sgn}(\sigma) f(x_{\sigma(1)}, \cdots, x_{\sigma(n)})$, $\forall (x_1, \cdots, x_n) \in E^n$, $\forall \sigma \in \mathfrak{S}_n$.

Proof

By what we proved on permutations it is enough to prove the proposition for adjacent transpositions. Let $i \in \{1, \cdots, n-1\}$, then $0 = f(x_1, \cdots, x_{i-1}, x_i + x_{i+1}, x_i + x_{i+1}, x_{i+2}, \cdots, x_n) = f(x_1, \cdots, x_n) + f(x_1, \cdots, x_{i-1}, x_{i+1}, x_i, \cdots, x_n)$.

$\square$

5.5.3 Def

We denote with $\mathrm{Hom}_S^{(n)}(E^n, F)$ and $\mathrm{Hom}_a^{(n)}(E^n, F)$ the set of symmetric and alternating $n-$linear maps from E^n to F. These are K-submodules of $\mathrm{Hom}^{(n)}(E^n, F)$ and when $n = 1$, $\mathrm{Hom}_S^{(1)}(E, F) = \mathrm{Hom}_a^{(1)}(E, F) = \mathrm{Hom}(E, F)$.

5.5.4 Theorem(Schwarz)

Let E and F be two vector spaces over $\mathbb{R}$. Let $U \subseteq E$ be 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.

Proof

By induction and by the fact that permutations are decomposed in transpositions, we can reduce to prove when the $n = 2$.

Since f is of class C^2, $d_{p+u}f - d_pf = D^2f(p)(u,\cdot) + o(u)$. $\forall\varepsilon > 0$, $\exists\delta > 0$, s.t. if $0 < \|u\| < \delta$, then $\|d_{p+u}f - d_pf - D^2f(p)(u,\cdot)\| < \varepsilon\|u\|$.

Take $(h,k) \in E^2$, $0 < \|h\| < \frac{\delta}{4}$, $0 < \|k\| < \frac{\delta}{4}$, $x \in B(p, \frac{\delta}{2})$.

For any $x \in B(p, \frac{\delta}{2})$. Let's introduce the following function. $\varphi(x) = f(x+k) - f(x) - D^2f(p)(k,x)$. We use the mean inequality on φ.

$\|\varphi(p+h) - \varphi(p)\| = \|f(p+h+k) - f(p+h) - D^2f(p)(k,p+h) - f(p+k) + f(p) + D^2f(p)(k,p)\| = \|f(p+h+k) - f(p+h) - f(p+k) + f(p) - D^2f(p)(k,h)\| \leq \sup\limits_{t\in]0,1[} \|d_{p+th}(\varphi)\|\|h\|$.

$\|d_{p+th}(\varphi)\| = \|d_{p+th+k}f - d_{p+th}f - D^2f(p)(k,\cdot)\| \leq \|d_{p+th+k}f - d_pf - D^2f(p)(k+th,\cdot)\| + \|d_{p+th}f - d_pf - D^2f(p)(th,\cdot)\| \leq \varepsilon\|th+k\| + \varepsilon\|th\| \leq 2\varepsilon(\|h\| + \|k\|)$.

$\|f(p+h+k) - f(p+k) - f(p+h) + f(p) - D^2f(p)(k,h)\| = o(\max\{\|h\|,\|k\|\}^2)$.

Exchange the role of h, k, then we get $\|f(p+h+k) - f(p+h) - f(p+k) + f(p) - D^2f(p)(h,k)\| = o(\max\{\|h\|,\|k\|\}^2)$.

$\|D^2f(p)(k,h) - D^2f(p)(h,k)\| = o(\max\{\|h\|,\|k\|\}^2)$. □

5.5.5 Def

Let E and F be normed 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. If f and f^{-1} are both continuous we say that f is a homeomorphism.

If f and f^{-1} are both of class C^n, we say that f is a C^n-diffeomorphism. If $\forall n \in \mathbb{N}$, it holds, then f is a C^∞-diffeomorphism.

5.5.6 Prop

Let E and F be Banach spaces. Let $I(E,F) \subseteq \mathscr{L}(E,F)$ be the set of linear, continuous, invertible mappings s.t. $\|\varphi^{-1}\| < +\infty$. Then $I(E,F)$ is open in $\mathscr{L}(E,F)$. Moreover, the mapping $i : I(E,F) \to I(F,E)$, $\varphi \mapsto \varphi^{-1}$, is a C^1-diffeomorphism.

Proof

Let $\varphi \in I(E,F)$, we want to show $\varphi - \psi \in I(E,F)$ for $\psi \in \mathscr{L}(E,F)$ and $\|\psi\| < \frac{1}{\|\varphi^{-1}\|}$

Notice that $\varphi - \psi = \varphi \circ (\mathrm{Id}_E - \varphi^{-1} \circ \psi)$. Since $\|\varphi^{-1} \circ \psi\| \leq \|\varphi^{-1}\|\|\psi\| < 1$

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 $(\mathrm{Id}_E-\varphi^{-1}\circ\psi)$. $(\mathrm{Id}_E-\varphi^{-1}\circ\psi)\circ\sum\limits_{n=1}^{N-1}(\varphi\circ\psi)^{\circ n}=Id_E-(\varphi^{-1}\circ\psi)^{\circ N}$, $N\to+\infty$.

$(\varphi-\psi)^{-1}=\sum\limits_{n\in\mathbb{N}}(\varphi\circ\psi)^{\circ n}\circ\varphi^{-1}$.

Hence $i(\varphi-\psi)=(\varphi-\psi)^{-1}=\varphi^{-1}+\varphi^{-1}\circ\psi\circ\varphi^{-1}+o(\|\psi\|)$, so $i(\varphi-\psi)-i(\varphi)=\varphi^{-1}\circ\psi\circ\varphi^{-1}+o(\|\psi\|)$.

$d_\varphi i(\psi)=i(\varphi)\circ(-\psi)\circ i(\varphi)$, therefore, i is differentiable. Moreover, $I(E,F)\to I(F,E)$ is continuous. Do the same with i^{-1}. □

Remark

By induction, you can show that i is a $C^\infty-$diffeomorphism.

5.5.7 Prop

Let $n\in\mathbb{N}\bigcup\{\infty\}$. Let E, F and G be normed vector spaces over a complete valued field $(K,|\cdot|)$. $U\subseteq E$, $V\subseteq F$ be open subsets. $f:U\to V$ and $g:V\to G$ be mappings of class C^n, then $g\circ f$ is class of C^n.

Proof

The case when $n=0$ is known.

Denote by $\Phi:\mathscr{L}(F,G)\times\mathscr{L}(E,F)\to\mathscr{L}(E,G)$, $(\beta,\alpha)\mapsto\beta\circ\alpha$. Φ is a bounded bilinear mapping.

$\|\Phi(\beta,\alpha)\|\le\|\beta\|\|\alpha\|$.

Suppose that $n\ge 1$ and the statement is true for mappings of class C^{n-1}.

Since $n\ge1$, $g\circ f$ is differentiable. $\forall p\in U$, $d_p(g\circ f)=d_{f(p)}g\circ d_pf$

$\Phi\circ(D^1g\circ f,D^1f)=D^1(g\circ f):U\to\mathscr{L}(E,G)$.

Since g and f are of C^n, D^1g and D^1f are of class C^{n-1}. Thus by induction hypothesis, $(D^1g\circ f,D^1f)$ 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}, $g\circ f$ is of class C^n. □

5.5.8 Prop

Let E and F be Banach spaces over a complete valued field $(K,|\cdot|)$, U

and V be open subsets of E and F, respectively. $n \in \mathbb{N}\bigcup\{\infty\}$, $n \geq 1$, and $f : U \to V$ be a bijection.

If f is of class C^n, and f^{-1} is differentiable, then f^{-1} is of class C^n.

Proof

$f \circ f^{-1} = \mathrm{Id}_V$. For any $y \in V$, $d_y(f \circ f^{-1}) = d_{f^{-1}(y)} f \circ d_y f^{-1} = \mathrm{Id}_F$

For $x \in U$, with $y = f(x)$, $d_y(f \circ f^{-1}) = d_x f \circ d_y f^{-1} = \mathrm{Id}_F$,

$d_x(f^{-1} \circ f) = d_y f^{-1} \circ d_x f = \mathrm{Id}_E$.

So $d_y f^{-1} = (d_x f)^{-1}$, that is $D^1 f^{-1} = i \circ D^1 f \circ f^{-1}$, $i : I(E, F) \to I(F, E)$, $\varphi \mapsto \varphi^{-1}$.

Suppose that f^{-1} is of class C^{n-1}, then, $D^1 f^{-1} = i \circ D^1 f \circ f^{-1}$ is of class C^{n-1}. Hence f^{-1} is of class C^n.

5.5.9 Theorem(Local inversion)

Let E and F be Banach spaces over $\mathbb{R}$, $U \subseteq E$ open, $f : U \to F$ be a mapping of class C^n ($n \in \mathbb{N}_{\geq 1}\bigcup\{\infty\}$), and $a \in U$. Suppose that $d_a f \in I(E, F)$.($d_a f$ is invertible and of bounded inverse.)

Then there exists open neighborhood V and W of a and $f(a)$, respectively, s.t. (1) $V \subseteq U$ and $f(V) \subseteq W$.

(2) The restriction of f to V defines a bijection from V to W.

(3) $(f_{|V})^{-1} : W \to V$ is of class C^n.

Proof

For $y \in F$, consider the mapping, $\phi_y : U \to E$, $\phi_y(x) = x - (d_a f)^{-1}(f(x) - y)$, $f(x) = y$ iff $\phi_y(x) = x$.

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 $D^1 f$, $\forall y \in F$, there exists $r > 0$, s.t. $\overline{B}(a, r) \subseteq U$, $\forall x \in \overline{B}(a, r)$, $\|d_x\phi_y\| \leq \frac{1}{2}$.

By the mean value inequality, $\forall (x_1, x_2) \in \overline{B}(a, r)$, $\|\phi_y(x_1) - \phi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|$. Hence ϕ_y is a contraction.

By the boundedness of $(d_a f)^{-1}$, $\exists \delta > 0$ s.t. $\forall y \in \overline{B}(f(a), \delta)$, $\|(d_a f)^{-1}(f(a) - y)\| \leq \frac{r}{2}$.

Then $\forall x \in \overline{B}(a, r)$, $y \in \overline{B}(f(a), \delta)$. $\|\phi_y(x) - a\| \leq \|\phi_y(x) - \phi_y(a)\| + \|\phi_y(a) - a\| \leq \frac{1}{2}\|x - a\| + \frac{r}{2} \leq r$.

$\phi_y(\overline{B}(a,r)) \subseteq \overline{B}(a,r)$. By the fixed point theorem, $\exists g : \overline{B}(f(a),\delta) \to \overline{B}(a,r)$, sending y to the fixed point of ϕ_y. Let $W = B(f(a),\delta)$, $V = g(W)$.

$g_{|W} : W \to V$ is the inverse of $f_{|V} : V \to W$. 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, $\forall x \in V$, $d_x f \in I(E,F)$.

Let $x_0 \in V$, $y_0 = f(x_0)$, $x_0 = g(y_0)$.

$y - y_0 = f(g(y)) - f(g(y_0)) = d_{x_0} f(g(y) - g(y_0)) + o(\|g(y) - g(y_0)\|)$

So $g(y) - g(y_0) = (d_{x_0} f)^{-1}(y - y_0) + o(\|g(y) - g(y_0)\|)$. This leads to $g(y) - g(y_0) = O(\|y - y_0\|)$.($\exists \varepsilon > 0$, $(1-\varepsilon)\|g(y) - g(y_0)\| \leq \|d_{x_0} f\|^{-1}\|y - y_0\|$ when $\|y - y_0\|$ is sufficient small.)

So $d_{y_0} g = (d_{x_0} f)^{-1}$. By the previous proposition, g is of class C^n. $\square$

5.5.10 Theorem

Let E_1, E_2, F be Banach spaces. U and V be open subsets of E_1 and E_2, respectively. $f : U \times V \to F$ be a mapping of class C^1, $(a,b) \in U \times V$ s.t. $f(a,b) = 0$ and $\frac{\partial f}{\partial y}(a,b) : E_2 \to F$ is invertible of bounded inverse.

Then $\exists$ open neigbborhood U_a and V_b of a and b, and a mapping $\varphi : U_a \to V_b$ of class C^1, s.t.

(1) $\varphi(a) = b$

(2) $\forall (x,y) \in U_a \times V_b$, $f(a,b) = 0$ iff $y = \varphi(x)$

Moreover, if f is of class C^n, then φ is also of class C^n.

Proof

We define $g : U \times V \to U \times F$, $(x,y) \mapsto (x, f(x,y))$.

$Dg(a,b)(h,k) = (h, \frac{\partial f}{\partial x}(a,b)(h) + \frac{\partial f}{\partial y}(a,b)(k)) = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x}(a,b) & \frac{\partial f}{\partial y}(a,b) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$

$Dg(a,b)^{-1}(h',k') = (h', (\frac{\partial f}{\partial y}(a,b))^{-1}(k' - \frac{\partial f}{\partial x}(a,b)))$

By local inversion theorem, $\exists W'$ neighborhoods (a,b) and $W \subseteq U \times F$ neighborhoods of $(a,0)$ s.t. $g_{|W'} : W' \to W$ bijection.

$(g_{|W'})^{-1} : W \to W'$, $g(x,y) = (x, f(x,y))$. $g^{-1}(x,z)$ should be of the form $g^{-1}(x,z) = (x, \phi(x,z))$

Take $\varphi(x) = \phi(x,0)$. Then $f(x,y) = 0$ iff $y = \varphi(x)$. $\square$

5.6 Power series

First, we consider two fundamental mappings: $\lambda\mathbb{1}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$, $x \mapsto \lambda$ and $\mathrm{Id}_{\mathbb{R}} : \mathbb{R} \to \mathbb{R}$, $x \mapsto x$. By applying addition and multiplication to these mappings, we obtain polynomials. The set of all polynomials is denoted by $\mathbb{R}[T]$.

Intuitively, $\mathbb{R}^{\oplus\mathbb{N}} \cong \mathbb{R}[T]$, $(a_n)_{n\in\mathbb{N}} \mapsto a_0 + a_1T + a_2T^2 + \cdots$

5.6.1 Prop

Let $(K, |\cdot|)$ be a complete valued field. $(a_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$ be a sequence in K. $R \in \mathbb{R}$ is a positive number.

Suppose $\sum\limits_{n\in\mathbb{N}} |a_n|R^n < \infty$. Then, for any $z \in \overline{B}(0, R)$, the series $\sum\limits_{n\in\mathbb{N}} a_n z^n$ converges.

Proof

Note that $\sum\limits_{n\in\mathbb{N}} |a_n|R^n$ converges. Hence, $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, $\forall (n, m) \in \mathbb{N}^2_{\geq N}$, $|\sum\limits_{i=0}^{n} |a_i|R^i - \sum\limits_{j=0}^{m} |a_j|R^j| < \varepsilon$.

Not loss generality, let $n \leq m$, $|\sum\limits_{i=0}^{n} a_i z^i - \sum\limits_{j=0}^{m} a_j z^j| = |\sum\limits_{i=n+1}^{m} a_i z^i| \leq \sum\limits_{i=n+1}^{m} |a_i||z^i| \leq \sum\limits_{i=n+1}^{m} |a_i|R^i = |\sum\limits_{i=0}^{n} |a_i|R^i - \sum\limits_{j=0}^{m} |a_j|R^j| < \varepsilon$. The absolute value inequality shows that $\sum\limits_{n\in\mathbb{N}} a_n z^n$ is a Cauchy sequence in K, which converges due to completeness. □

5.6.2 Def

Let X be a set, K be a field, $(E, \|\cdot\|)$ be a Banach space over K. $\forall n \in \mathbb{N}$, $f_n : X \to E$ is a mapping.

A series $\sum\limits_{n\in\mathbb{N}} f_n$ is said to converge normally on X if $\sum\limits_{n\in\mathbb{N}} \sup\limits_{x\in X} \|f_n(x)\| < +\infty$.

Remark

Note that normal convergence implies uniform convergence. The proof parallels proposition 1.

Example(Geometric series)

We formally define the geometric series $z \mapsto \sum_{n\in\mathbb{N}} z^n$, $z \in \mathbb{C}$, $|z| < 1$.

Using proposition 1, for any $R = |z| < 1$, $\sum_{n\in\mathbb{N}} R^n = \frac{1}{1-R} < +\infty$. Hence, the geometric series actually defines a function for $z \in B(0,1)$.

Example(Exponential series)

Define the exponential series $\exp(T) = \sum_{n\in\mathbb{N}} \frac{T^n}{n!} \in \mathbb{C}[|T|]$.

We verify convergence for $R \in \mathbb{R}_{>0}$, $z \in \overline{B}(0,R)$ via estimation. Choose integer $N > R$. For $n \geq N$: $\sum_{i=0}^{n} \frac{R^i}{i!} \leq \sum_{i=0}^{N-1} \frac{R^i}{i!} + \frac{R^{N-1}}{(N-1)!} \sum_{i=N}^{\infty} \left(\frac{R}{N}\right)^{i-N+1} < +\infty$, since geometric series is convergent.

Thus, $\exp(z) = \sum_{n\in\mathbb{N}} \frac{z^n}{n!}$ converges normally on $\overline{B}(0,R)$, $\forall R > 0$. We denote $\exp(z)$ as e^z.

Restricting $\exp : \mathbb{R} \to \mathbb{R}$, for any $R > 0$, the partial sums $(\sum_{i=1}^{n} \frac{x^i}{i!})_{n\in\mathbb{N}}$ converge uniformly to $\exp(x)$ on $[-R,R]$ (by normal convergence). Since partial sums are continuous, exp is continuous on $\mathbb{R}$(by theorem 3.9.9).

5.6.3 Theorem

Let E and F be normed vector spaces over $\mathbb{R}$, $(f_n)_{n\in\mathbb{N}}$ a sequence of maps from open set $U \subseteq E$ to F. If $(D^1 f_n)_{n\in\mathbb{N}}$ converges uniformly to $g : U \to \mathscr{L}(E,F)$ and $(f_n)_{n\in\mathbb{N}}$ converges pointwise to $f : U \to F$, then f is differentiable on U with $D^1 f = g$.

Proof

See Theorem 5.3.6 □

5.6.4 Corollary

Let $U \subseteq \mathbb{R}$ be an open set, $f_n : U \to \mathbb{R}$ differentiable. If (f_n') converges uniformly to g and $(f_n)_{n\in\mathbb{N}}$ converges pointwise to f, then f is differentiable with $f' = g$.

Proof

$\forall p \in U$, $\forall (n,m) \in \mathbb{N}^2$, $|f_n(x) - f_m(x) - (f_n(p) - f_m(p))| \leq \sup_{\xi\in U} |f_n'(\xi) -$

$f'_m(\xi)||x-p|$. Taking $\lim\limits_{m\to+\infty} |f_n(x)-f(x)-(f_n(p)-f(p))| \leq \sup\limits_{\xi\in U}|f'_n(\xi)-g'(\xi)||x-p|$. Let $\varepsilon_n = \sup\limits_{\xi\in U}|f'_n(\xi)-g'(\xi)|$. $|f(x)-f(p)-(x-p)g(p)| \leq 2\varepsilon_n|x-p| + |f_n(x)-f_n(p)-(x-p)f'_n(p)|$. Hence $\limsup\limits_{x\to p} \frac{|f(x)-f(p)-(x-p)g(p)|}{|x-p|} \leq 2\varepsilon_n$, proving $f'(p) = g(p)$. $\square$

5.6.5 Prop

If $\sum\limits_{n\in\mathbb{N}} |a_n|r^n < +\infty \ \forall r \in [0, R[$, then $\sum\limits_{n\in\mathbb{N}} (n+1)|a_{n+1}|r^n < +\infty \ \forall r \in [0, R[$.

Proof

Take $s \in]r, R[$. Then $\sum\limits_{n\in\mathbb{N}} (n+1)|a_{n+1}|r^n = \frac{1}{s}\sum\limits_{n\in\mathbb{N}} |a_{n+1}|s^{n+1}\frac{(n+1)r^n}{s^n} < +\infty$. $\square$

Remark

This proposition implies that if a series converges on $]-R, R[$, its term-wise derivative also converges there.

5.6.7 Corollary

If $\sum\limits_{n\in\mathbb{N}} a_n r^n$ converges on $]-R, R[$, then $\sum\limits_{n\in\mathbb{N}} a_n t^n$ is differentiable in $]-R, R[$ with derivative $\sum\limits_{n\in\mathbb{N}} (n+1)a_{n+1}t^n$.

Example

Differentiating $\exp(t) = \sum\limits_{n\in\mathbb{N}} \frac{t^n}{n!}$ yields $\exp'(t) = \exp(t)$, proving smoothness.

Example(Trigonometric functions)

Define trigonometric functions: $\cos(t) = \frac{\exp(it)+\exp(-it)}{2} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}t^{2n}$,, $\sin(t) = \frac{\exp(it)-\exp(-it)}{2i} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!}t^{2n-1}$.

Derivatives satisfy $\sin'(z) = \cos(z)$, $\cos'(z) = -\sin(z)$, both smooth.

5.6.8 Theorem(Taylor's Formula)

Let $f : U \to \mathbb{R}$ be n-times differentiable near $a \in \mathbb{R}$. Then $f(x) = \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o(|x-a|^n)$.

If f is $(n+1)$-times differentiable, $\exists \varepsilon_x \in]0,1[$ s.t. $f(x) - \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k = \frac{f^{(n+1)}(a+\varepsilon_x(x-a))}{(n+1)!}(x-a)^{n+1}$.

Proof

See Taylor-Lagrange formula in 5.3.3. $\square$

Example

Taylor series applications to power series: Let $(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ satisfy $\sum\limits_{n\in\mathbb{N}} |a_n| r^n < +\infty \ \forall r \in [0, R[$. Then $f(x) = \sum\limits_{n\in\mathbb{N}} a_n t^n$ satisfies: $f^{(k)}(t) = \sum\limits_{n\in\mathbb{N}} a_{n+k} \frac{(n+k)!}{n!} t^n, \quad f^{(k)}(0) = k! a_k$.

Substituting into Taylor's formula: $f(x) = \sum\limits_{k=0}^{n} a_k x^k + o(|x|^n)$. Taylor series perfectly approximates f near a.

Example

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

All derivatives at 0 vanish: $f^{(k)}(0) = 0$. Taylor series fails to approximate f near 0.

5.6.9 Prop

Let $f \in C^\infty(U)$ satisfy $\lim\limits_{n\to+\infty} \frac{r^n}{n!} \sup\limits_{t\in[a-r,a+r]} |f^{(n)}(t)| = 0$. Then its Taylor series converges uniformly to f on $[a-r, a+r]$.

Proof

$|f(x) - \sum\limits_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k| \leq \frac{r^{k+1}}{(k+1)!} \sup\limits_{t\in[a-r,a+r]} |f^{(k+1)}(t)|$. $\square$

Example(Logarithm)

Fix exp on $\mathbb{R}$ to make it be a mapping from $\mathbb{R}$ to $\mathbb{R}$. By the binomial theorem and the absolutely convergence of exp, we obtain $\exp(x+y) = \exp(x)\exp(y)$. Thus, $\exp(\cdot)$ is a group homomorphism with $\exp(0) = 1$.

For $x > 0$, $\exp(x) > 1$ is obvious. For $x < 0$, $\exp(x) = \frac{1}{\exp(-x)} > 0$.

Additionally, $\exp(\cdot)$ is monotonically increasing because for $x > y$, $\exp(x) - \exp(y) = \exp(y)(\exp(x-y) - 1) > 0$.

Since it is increasing, we define its inverse function as $\ln(\cdot) : \mathbb{R}_{>0} \to \mathbb{R}$. By the inverse function theorem, $\ln(\cdot)$ is smooth.

Using the chain rule: $(\mathrm{Id}_{\mathbb{R}_{>0}})' = (\exp(\ln(x)))' = \ln'(x)\exp'(\ln(x)) = 1$. Hence, $\ln'(x) = \frac{1}{x}$.

$$\ln^{(n)}(x) = (-1)^{n-1}\frac{(n-1)!}{x^n}$$

The Taylor expansion of $\ln(x)$ at $x = 1$ is:

$$\ln(x) = \sum_{n\in\mathbb{N}} \frac{f^{(n)}(1)}{n!}(x-1)^n = \sum_{n\in\mathbb{N}} (-1)^{n-1}\frac{(x-1)^n}{n!}$$

This series converges normally on $[1-r, 1+r]$ for $r \in [0,1]$, but does not necessarily converge uniformly to $\ln(x)$.

To determine uniform convergence, compute this

$$\lim_{n\to+\infty} \frac{r^n}{n!} \sup_{t\in[1-r,1+r]} \left|\frac{(n-1)!}{t^n}\right| = 0$$

Thus, the series converges uniformly to $\ln(x)$ on $[\frac{1}{2}, 2]$

$$\sum_{n\in\mathbb{N}} \frac{(-1)^{k-1}}{k}(x-1)^k = \ln(x) \quad \text{for } x \in \left[\frac{1}{2}, 2\right]$$

Example(Power Functions)

For $x > 0$ and $\alpha \in \mathbb{R}$, define $x^\alpha = \exp(\alpha \ln(x))$. Since this is a composition of group homomorphisms, $x \mapsto x^\alpha$ is also a group homomorphism, so $(xy)^\alpha = x^\alpha y^\alpha$.

$$(x^\alpha)' = \alpha\frac{1}{x}\exp'(\alpha\ln(x)) = \alpha x^{\alpha-1}$$

Define $f_\alpha : \mathbb{R}_{>0} \to \mathbb{R}$ by $x \mapsto x^\alpha$. Its n-th derivative is:

$$f_\alpha^{(n)}(x) = \alpha(\alpha-1)\cdots(\alpha-(n-1))x^{\alpha-n}$$

The Taylor expansion at $x = 1$ is:

$$\sum_{n\in\mathbb{N}} \frac{\alpha(\alpha-1)\cdots(\alpha-(n-1))}{n!}(x-1)^n$$

This series converges normally on $[1,2]$.

$$\lim_{n\to+\infty}\frac{r^n}{n!}\sup_{[1,1+r]}|\alpha(\alpha-1)\cdots(\alpha-(n-1))|=0$$

Hence, the series converges uniformly to x^α on $[1,2]$.

Example(Inverse Trigonometric Functions)

Note that $\exp(it)\exp(-it)=1$, implying $|\exp(it)|=1$. Define $q:\mathbb{R}\to\mathbb{C}$ by $t\mapsto\exp(it)$, a group homomorphism from $(\mathbb{R},+)$ to $(\mathbb{C}^\times,\times)$. The kernel of q contains a smallest positive element, defined as 2π.

Lemma

$\ker(q)$ contains a smallest positive number.

Proof

Assume no smallest positive element exists. For any $x\in\ker(q)\setminus\{0\}$, $nx\in\ker(q)$ for all $n\in\mathbb{Z}$, making $\ker(q)$ unbounded. For intervals (a,b), there exists $x\in\ker(q)$ with $x>b$. Let $\varepsilon=b-a$, if $y\in\ker(q)$ satisfies $0<y<\epsilon$, then $\exists N\in\mathbb{Z}$ s.t. $x-Ny\in\ker(q)\cap(a,b)$. Otherwise, positive elements in $\ker(q)$ are spaced by at least ϵ, leading to a contradiction. Thus, $\ker(q)$ is dense. However, continuity of q implies $\ker(q)$ is closed, so $\ker(q)=\mathbb{R}$, contradicting $\exp(1)>1$. Therefore, $\ker(q)$ has a smallest positive element.

$\square$

$\sin(x)$ is monotonically increasing on $[-\frac{\pi}{2},\frac{\pi}{2}]$ since (To be continued...).

Define $\arcsin(x):[-1,1]\to[-\frac{\pi}{2},\frac{\pi}{2}]$ as the inverse of $\sin(x)$. Its derivative is

$$\arcsin'(x)=\frac{1}{\sqrt{1-x^2}}$$

The Taylor series for $\arcsin(x)$ is derived via its derivative

$$(1-x^2)^{-1/2}=\sum_{n\in\mathbb{N}}\frac{(2n-1)!!}{2^n n!}x^{2n},\quad |x|<1$$

Integrating term-by-term

$$\arcsin(x)=\sum_{n\in\mathbb{N}}\frac{(2n)!}{(2n+1)4^n(n!)^2}x^{2n+1},\quad |x|<1$$

Since the derivative of $\arcsin(x)$ is too complicated, the fitting effect of Taylor expansion on it will not be discussed.

Since $\arccos(x) = \frac{\pi}{2} - \arcsin(x)$, $\arccos(x)$ will no longer be discussed separately.

Define $\tan(x) = \frac{\sin(x)}{\cos(x)}, x \notin \{n\pi + \frac{\pi}{2} | n \in \mathbb{Z}\}$. $\tan'(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x)$, $\arctan'(x) = \frac{1}{\tan'(\arctan(x))} = \frac{1}{1+x^2}$.

Example

$$\lim_{n\to+\infty} \left(1 + \frac{1}{n}\right)^n = \lim_{n\to+\infty} \exp\left(n \ln\left(1 + \frac{1}{n}\right)\right) = \lim_{n\to+\infty} \exp\left(1 + o(1)\right) = e$$

5.6.11 Theorem(Fundamental theorem of algebra)

Every non constant polynomial with coefficients in $\mathbb{C}$ admits a root.

Proof

$f(z) = a_0 + a_1 z + \cdots + a_n z^n \in \mathbb{C}[z]$ and $a_n \neq 0$.

If $|z| \to +\infty$, then $|f(z)| \to +\infty$. In fact, $|f(z)| = |z|^n |a_n + \cdots + a_0 z^{-n}| \geq |z|^n (|a_n| - \frac{|a_{n-1}|}{|z|} - \cdots - \frac{|a_0|}{|z|^n})$.

Then we want to prove that $f(z)$ admits a global minimum. Let $M = |f(0)|$, $\exists N \in \mathbb{R}$, $\forall z$ with $|z| > N$, we have $|f(z)| > M$. Then consider the complex disk $\overline{B}(0, N) = \{z \in \mathbb{C} | |z| \leq N\}$. $\overline{B}(0, N)$ is compact, f admits a minimum in $\overline{B}(0, N)$. Call $z_0 \in \overline{B}(0, N)$ the point of minimum $|f(z_0)| \leq |f(0)| = M$, thus z_0 is a global point of minimum.

We show that $f(z_0) = 0$. Assume that $f(z_0) = a \neq 0$. Write the taylor expansion of f around z.

$f(z) = a + c_k (z - z_0)^k + \cdots + c_n (z - z_0)^n$, $k \geq 1$. Take $q(z) = a + c_k (z - z_0)^k$.

$f(z) - q(z) = O((z - z_0)^{k+1})$ for $z \to z_0$, which is equivalent to $|\frac{f(z)-q(z)}{(z-z_0)^{k+1}}| \leq M'$ for z in a neighborhood of z_0.

Fix $\theta_0 = \frac{\arg(a)+\pi-\arg(c_k)}{k}$. $|f(z)| \leq |q(z)| + r^{k+1} |\frac{f(z)-q(z)}{r^{k+1}}|$. By taking $z = z_0 + re^{i\theta_0}$, $|f(z)| \leq |a + (-1) c_k r^k e^{i(\arg(a)-\arg(c_k))}| + M' r^{k+1} \leq ||a| e^{i \arg(a)} + (-1) c_k r^k e^{i \arg(a)} e^{-i \arg(c_k)}| + M' r^{k+1} = ||a| - |c_k| r^k| + M' r^{k+1} = |a| - |c_k| r^k + M' r^{k+1} < |a|$ $\square$

Chapter 6

Integration

6.1 Integral operators

We fix a set Ω and a vector subspace S of $\mathbb{R}^{\Omega}$.

We suppose that $\forall (f,g) \in S^2$, $f \wedge g : \Omega \to \mathbb{R}$, $\omega \mapsto \min\{f(\omega), g(\omega)\}$ belongs to S.

6.1.1 Prop

$\forall (f,g) \in S^2$, $f \vee g : \Omega \to \mathbb{R}$, $\omega \to \max\{f(\omega), g(\omega)\}$ belongs to S.

Proof

$f \vee g = f + g - f \wedge g$. □

6.1.2 Prop

$\forall f \in S$, $|f| : \Omega \to \mathbb{R}$, $\omega \mapsto |f(\omega)|$ belongs to S

Proof

$|f| = f \vee (-f)$.

6.1.3 Def

We call integral operator on S any $\mathbb{R}-$linear mapping $I : S \to \mathbb{R}$ that satisfies the following conditions.

(1) If $f \in S$ satisfies the condition that $\forall \omega \in \Omega$, $f(\omega) \geq 0$, then $I(f) \geq 0$.

(2) If $(f_n)_{n \in \mathbb{N}}$ is a decreasing sequence of elements in S s.t $\forall \omega \in \Omega$, $\lim\limits_{n \to +\infty} f_n(\omega) = 0$, then $\lim\limits_{n \to +\infty} I(f_n) = 0$.

Example 1

$\Omega = \mathbb{R}$, S=Vector subspaces of $\mathbb{R}^{\mathbb{R}}$ generated by mappings of the form $\mathbb{1}_A$, $A =]a,b]$, $(a,b) \in \mathbb{R}^2$, $a < b$.

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{else} \end{cases}$$

Any elements of S is of the form $\sum\limits_{i=1}^{n} \lambda_i \mathbb{1}_{]a_i,b_i]}$, $\lambda_i \in \mathbb{R}$, $a_i < b_i$.

Our $I : S \to \mathbb{R}$ is defined as $I(\sum\limits_{i=1}^{n} \lambda_i \mathbb{1}_{]a_i,b_i]}) = \sum\limits_{i=1}^{n} \lambda_i (b_i - a_i)$.

More generally if $\varphi : \mathbb{R} \to \mathbb{R}$ is increasing and right continuous. ($\forall x \in \mathbb{R}$, $\lim\limits_{\varepsilon>0,\varepsilon\to 0} \varphi(x+\varepsilon) = \varphi(x)$)

We define $I_\varphi : S \to \mathbb{R}$, $I_\varphi(\sum\limits_{i=1}^{n} \lambda_i \mathbb{1}_{]a_i,b_i]}) = \sum\limits_{i=0}^{n} \lambda_i(\varphi(b_i) - \varphi(a_i))$

Example 2(Radon measure)

Let Ω be a quasi-compact topological space. $S = C^0(\Omega) := \{f : \Omega \to \mathbb{R}$ continuous$\}$

Let $I : S \to \mathbb{R}$, $R-$linear s.t. $\forall f \in S$, $f \geq 0$, one has $I(f) \geq 0$.

6.1.4 Prop(Dini's theorem)

Let $(f_n)_{n\in\mathbb{N}}$ be a decreasing sequence in $C^0(\Omega)$, that converges pointwisely to some $f \in C^0(\Omega)$. Then $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f.

Proof

Let $g_n = f_n - f \geq 0$.

Fix $\varepsilon > 0$, $\forall n \in \mathbb{N}$, let $U_n = \{\omega \in \Omega | g_n(\omega) < \varepsilon\}$ is open.

Moreover, $\bigcup\limits_{n\in\mathbb{N}} U_n = \Omega$.

Since Ω is quasi-compact, $\exists N \in \mathbb{N}$, $\Omega = U_N$.

Therefore, $\forall n \in \mathbb{N}$, $n \geq N$, $\forall \omega \in \Omega$, $g_n(\omega) < \varepsilon$. $\square$

6.1.5 Consequence

If $(f_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ is decreasing and converges pointwisely to 0, then $\|f_n\|_{\sup} := \sup\limits_{\omega\in\Omega} |f_n(\omega)|$ converges to 0, when $n \to +\infty$.

$\forall n \in \mathbb{N}$, $f_n \leq \|f_n\|_{\sup}\mathbb{1}_\Omega$. So $0 \leq I(f_n) \leq \|f_n\|_{\sup} I(\mathbb{1}_\Omega) \to 0$. ($n \to +\infty$). $\square$

Remark

If $f \leq g$, then $f - g \geq 0$, so $I(g-f) = I(g) - I(f) \geq 0$.

Example 3(Measures on σ-algebras)

6.1.6 Def

We call σ-algebra any subset $\mathscr{A}$ of $\mathscr{P}(\Omega)$ that satisfies the following conditions:

(1) $\varnothing \in \mathscr{A}$

(2) If $A \in \mathscr{A}$, then $\Omega \setminus A \in \mathscr{A}$

(3) If $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$, then $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

6.1.7 Prop

(1) $\Omega \in \mathscr{A}$

(2) If $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$, then $\bigcap\limits_{n\in\mathbb{N}} A_n \in \mathscr{A}$.

Proof

Trivial. $\square$

Given a σ-algebra $\mathscr{A}$ on Ω, we mean by measure on $(\Omega, \mathscr{A})$ any mapping $\mu : \mathscr{A} \to [0, +\infty]$ s.t.

(1) $\mu(\varnothing) = 0$

(2) If $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$ s.t. $A_0, \cdots, A_n, \cdots$ are pairwisely disjoint, then $\mu(\bigcup\limits_{n\in\mathbb{N}} A_n) = \sum\limits_{n\in\mathbb{N}} \mu(A_n)$

We let $S =$ vector subspace of $\mathbb{R}^{\Omega}$ generated by mapping of the form $\mathbb{1}_A$, where $A \in \mathscr{A}$, $\mu(A) < +\infty$.

Any element of S is of the form $\sum\limits_{i=1}^{n} \lambda \mathbb{1}_{A_i}$.

Let $I_\mu : S \to \mathbb{R}$. $I_\mu(\sum\limits_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}) = \sum\limits_{i=1}^{n} \lambda_i \mu(A_i)$.

6.2 Riemann integral

6.2.1 Def

Let Ω be a non-empty set and S be a vector subspace of $\mathbb{R}^\Omega$. If $\forall (f,g) \in S^2$. $f \wedge g \in S$, we say that S is a Riesz space.

In this section, we fix a Riesz space and an integral operator $I : S \to \mathbb{R}$.

6.2.2 Def

For any $f : \Omega \to \mathbb{R}$, let $I^*(f) := \inf\limits_{\mu \in S, \mu \geq f} I(\mu)$, $I_*(f) := \sup\limits_{l \in S, l \leq f} I(l)$.

If $I^*(f) = I_*(f) \in \mathbb{R}$, we say that f is $I-$Riemann integrable, and denote by $I(f)$ the value $I^*(f)(= I_*(f))$.

6.2.3 Theorem

The set V of all I-Riemann integrable mappings forms a vector subspace of $\mathbb{R}^\Omega$ that contains S. Moreover, $I : V \to \mathbb{R}$ is an $\mathbb{R}$-linear mapping extending $I : S \to \mathbb{R}$.

Proof

$\forall h \in S$, $I^*(h) = I_*(h) = I(h)$, so $h \in V$.

Let $(f_1, f_2) \in V^2$. If $(\mu_1, \mu_2) \in S^2$, $\mu_1 \geq f_1$, $\mu_2 \geq f_2$, then $\mu_1 + \mu_2 \in S$, $\mu_1 + \mu_2 \geq f_1 + f_2$. Hence $I(\mu_1) + I(\mu_2) \geq I^*(f_1 + f_2)$.

Take 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, $\lambda \in \mathbb{R}_{>0}$.

$I^*(\lambda f) = \inf\limits_{\mu \in S, \mu \geq \lambda f} I(\mu) = \inf\limits_{v \in S, v \geq f} I(\lambda v) = \lambda I^*(f)$.

$I_*(\lambda f) = \lambda I_*(f)$. Hence if $f \in V$, then $\lambda f \in V$, and $I(\lambda f) = \lambda I(f)$.

$I^*(-f) = \inf\limits_{\mu \in S, \mu \geq -f} I(\mu) = \inf\limits_{l \in S, l \leq f} I(-l) = -\sup\limits_{l \in S, l \leq f} I(l) = -I_*(f)$.

$I_*(-f) = -f^*(f)$. Hence if $f \in V$, then $-f \in V$ and $I(-f) = -I(f)$.

$\square$

6.3 Daniell integral

We fix an integral operator $I : S \to \mathbb{R}$.

6.3.1 Prop

Let $(f_n)_{n\in\mathbb{N}}$ be an increasing sequence in S that converges pointwisely to some $f \in S$. Then $\lim\limits_{n\to+\infty} I(f_n) = I(f)$.

Proof

Let $g_n = f - f_n \in S$. $(g_n)_{n\in\mathbb{N}}$ is decreasing and converges pointwisely to 0. So $\lim\limits_{n\to+\infty} I(g_n) = 0$. Hence $\lim\limits_{n\to+\infty} I(f_n) = I(f)$. $\square$

6.3.2 Prop

Let $(f_n)_{n\in\mathbb{N}}$ be a increasing sequence in S, $f \in S$. If $f \leq \lim\limits_{n\to+\infty} f_n$, then $I(f) \leq \lim\limits_{n\to+\infty} I(f_n)$.

Proof

$f = \lim\limits_{n\to+\infty} f \wedge f_n$.(In fact, $\lim\limits_{n\to+\infty} f \wedge f_n = \lim\limits_{n\to+\infty} \min\{f_n, f\} = \min\{f, \lim\limits_{n\to+\infty} f_n\} = f$) So $I(f) = \lim\limits_{n\to+\infty} I(f \wedge f_n) \leq \lim\limits_{n\to+\infty} I(f_n)$. $\square$

6.3.3 Def

Let $S^{\uparrow} = \{f : \Omega \to \mathbb{R} \bigcup \{+\infty\} | \exists (f_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ increasing s.t. $f = \lim\limits_{n\to+\infty} f_n$ pointwisely$\}$.

6.3.4 Prop

Let f and 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 $\lim\limits_{n\to+\infty} I(f_n) \leq \lim\limits_{n\to+\infty} I(g_n)$.

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 $\lim\limits_{m\to+\infty} I(f_m) \leq \lim\limits_{n\to+\infty} I(g_n)$. $\square$

6.3.5 Corollary

Let $f \in S^{\uparrow}$. If $(f_n)_{n\in\mathbb{N}}$ and $(\tilde{f}_n)_{n\in\mathbb{N}}$ be increasing sequences in S s.t.

$f = \lim_{n\to+\infty} f_n = \lim_{n\to+\infty} \tilde{f}_n$, then $\lim_{n\to+\infty} I(f_n) = \lim_{n\to+\infty} I(\tilde{f}_n)$.

We denote by $I(f)$ the limit $\lim_{n\to+\infty} I(f_n)$.

Thus we obtain a mapping $I : S^\uparrow \to \mathbb{R}\bigcup\{+\infty\}$ s.t.

(1) If $(f_n)_{n\in\mathbb{N}} \in S^{\mathbb{N}}$ is increasing, then $I(\lim_{n\to+\infty} f_n) = \lim_{n\to+\infty} I(f_n)$.

(2) If $(f, g) \in S^\uparrow \times S^\uparrow$, $f \le g$, then $I(f) \le I(g)$.

(3) If $(f, g) \in S^\uparrow \times S^\uparrow$, then $f + g \in S^\uparrow$, and $I(f + g) = I(f) + I(g)$.

(4) If $f \in S^\uparrow$, $\lambda \ge 0$, then $\lambda f \in S^\uparrow$ and $I(\lambda f) = \lambda I(f)$.

6.3.6 Prop

Let $(f_n)_{n\in\mathbb{N}} \in (S^\uparrow)^{\mathbb{N}}$ be an increasing sequence and $f = \lim_{n\to+\infty} f_n$. Then $f \in S^\uparrow$ and $I(f) = \lim_{n\to+\infty} I(f_n)$.

Proof

For $k \in \mathbb{N}$, let $(g_{k,m})_{m\in\mathbb{N}} \in S^{\mathbb{N}}$ be an increasing sequence s.t. $f_k = \lim_{m\to+\infty} g_{k,m}$. 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, $f_n \ge f_k \ge g_{k,n}$. $(k \le n)$. Hence $f_n \ge h_n$.

Taking $\lim_{n\to+\infty}$, we get $f = \lim_{n\to+\infty} f_n \ge \lim_{n\to+\infty} h_n \ge \lim_{n\to+\infty} g_{k,n} = f_k$.

Taking $\lim_{k\to+\infty}$, we get $f = \lim_{n\to+\infty} h_n$. Hence $f \in S^\uparrow$, and $I(f) = \lim_{n\to+\infty} I(h_n) \le \lim_{n\to+\infty} I(f_n)$.

Conversely, $\forall n \in \mathbb{N}$, $f \ge f_n$. Hence $I(f) \ge \lim_{n\to+\infty} I(f_n)$. $\square$

6.3.7 Def

Let $S^\downarrow = \{-f | f \in S^\uparrow\}$. We extend I to $I : S^\downarrow \to \mathbb{R}\bigcup\{-\infty\}$ by letting $I(-f) := -I(f)$ for $f \in S^\uparrow$.

6.3.8 Prop

Let $(f, g) \in (S^\uparrow \bigcap S^\downarrow)^2$. If $f \le g$, then $I(f) \le I(g)$.

Proof

It suffices to treat the cases where $(f, g) \in S^\uparrow \times S^\downarrow$ and $(f, g) \in S^\downarrow \times S^\uparrow$.

If $(f, g) \in S^\uparrow \times S^\downarrow$, then $-f \in S^\downarrow$ and hence $g - f \in S^\downarrow$, $g - f \ge 0$.

If $(f, g) \in S^\downarrow \times S^\uparrow$, then $-f \in S^\uparrow$ and hence $g - f \in S^\uparrow$, $g - f \ge 0$.

In both cases, $0 \leq I(g-f) = I(g) + I(-f) = I(g) - I(f)$. $\square$

Remark

We extend I to a mapping $S^{\uparrow} \bigcup S^{\downarrow} \to [-\infty, +\infty]$.

6.3.9 Def

Let $f : \Omega \to \mathbb{R}$ be a mapping. We define

$\overline{I}(f) := \inf_{\mu \in S^{\uparrow}, u \geq f} I(\mu) \leq \inf_{\mu \in S, \mu \geq f} I(\mu) = I^*(f)$.

$\underline{I}(f) := \sup_{l \in S^{\downarrow}, l \leq f} I(l) \geq \sup_{l \in S, l \leq f} I(l) = I_*(f)$

If $\overline{I}(f) = \underline{I}(f) \in \mathbb{R}$, then we say that f is I-integrable (in the sense of Daniell). And we denote by $I(f)$ the real number $\overline{I}(f) (= \underline{I}(f))$.

Remark

If f is I-integrable in the sense of Riemann, then it is I-integrable in the sense of Daniell.

6.3.10 Theorem(Daniell)

The set $L^1(I)$ of all I-integrable mappings forms a vector subspace of $\mathbb{R}^{\Omega}$.

Moreover, (1) $\forall (f,g) \in L^1(I)$, $f \wedge g \in L^1(I)$

(2) $I : L^1(I) \to \mathbb{R}$ is an integral operator extending $I : S \to \mathbb{R}$.

Proof

Let $(f_1, f_2) \in L^1(I)^2$. Let $(l_1, l_2) \in S^{\downarrow} \times S^{\downarrow}$, $l_1 \leq f_1$, $l_2 \leq f_2$. Let $(\mu_1, \mu_2) \in S^{\uparrow} \times S^{\uparrow}$, $f_1 \leq \mu_1$, $f_2 \leq \mu_2$.

We have $l_1 + l_2 \leq f_1 + f_2 \leq \mu_1 + \mu_2$.

Taking the supremum with respect to (l_1, l_2), we get $I(f_1) + I(f_2) \leq \underline{I}(f_1 + f_2)$

Taking the infimum with respect to (μ_1, μ_2), we get $\overline{I}(f_1 + f_2) \leq I(f_1) + I(f_2)$

So $f_1 + f_2 \in L^1(I)$ and $I(f_1 + f_2) = I(f_1) + I(f_2)$.

Similarly, if $f \in L^1(I)$, $\lambda \geq 0$, then $\underline{I}(\lambda f) = \sup_{l \leq \lambda f, l \in S^{\downarrow}} I(l) = \sup_{l \leq f, l \in S^{\downarrow}} I(\lambda l) = \lambda \underline{I}(f) = \lambda I(f)$

$\overline{I}(\lambda f) = \lambda \overline{I}(I)(f) = \lambda I(f)$, 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 $\overline{I}(-f) = -\underline{I}(f) = -I(f)$, $\underline{I}(-f) = -\overline{I}(f) = -I(f)$.

So $-f \in L^1(I)$ and $I(-f) = -I(f)$.

We prove that $L^1(I)$ is stable by $\wedge$.

Let $(f_1, f_2) \in L^1(I)^2$. For any $\varepsilon > 0$, $\exists (l_1, l_2) \in S^{\downarrow} \times S^{\downarrow}$, $(\mu_1, \mu_2) \in S^{\uparrow} \times S^{\uparrow}$ s.t. $l_1 \le f_1 \le \mu_1$, $l_2 \le f_2 \le \mu_2$ and $I(\mu_1 - l_1) \le \frac{\varepsilon}{2}$, $I(\mu_2 - l_2) \le \frac{\varepsilon}{2}$.

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 $\overline{I}(f_1 \wedge f_2) - \underline{I}(f_1 \wedge f_2) \le \varepsilon$. □

6.3.11 Theorem(BeppoLevi)

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 $I(f) = \alpha$.

Proof

We may 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 $\varepsilon > 0$. For any $n \in \mathbb{N}$, let $\mu_n \in S^{\uparrow}$ s.t. $f_n - f_{n-1} \le \mu_n$ and $I(f_n - f_{n-1}) \ge I(\mu_n) - \frac{\varepsilon}{2^n}$

Hence $f_n = \sum\limits_{k=1}^{n} (f_k - f_{k-1}) \le \mu_0 + \cdots + \mu_n$ and $I(f_n) \ge \sum\limits_{k=1}^{n} (I(\mu_k) - \frac{\varepsilon}{2^k}) \ge I(\mu_1) + \cdots + I(\mu_n) - \varepsilon$.

Let $\mu = \mu_1 + \cdots + \mu_n + \cdots \in S^{\uparrow}$. $I(\mu) = \sum\limits_{n\in\mathbb{N}} I(\mu_n)$.

One has $\mu \ge f$, $\lim\limits_{n\to+\infty} I(f_n) \ge I(\mu) - \varepsilon \ge \overline{I}(f) - \varepsilon$.

Similarly, one can choose $l_n \in S^{\downarrow}$, $l_n \le f_n$, $I(l_n) \ge I(f_n) - \varepsilon$. $\liminf\limits_{n\to+\infty} I(l_n) \ge \alpha - \varepsilon$.

Note that $l_n \le f_n \le f$, so $\alpha - \varepsilon \le \liminf\limits_{n\to+\infty} I(l_n) \le \underline{I}(f)$.

Thus $\alpha - \varepsilon \le \underline{I}(f) \le \overline{I}(f) \le \alpha + \varepsilon$. Let $\varepsilon \to 0$, we get $\overline{I}(f) = \underline{I}(f) = \alpha$. □

6.3.12 Lemma(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. $f_n \ge g$ for any $n \in \mathbb{N}$. If $\liminf_{n\to+\infty} f_n$ is a mapping from Ω to $\mathbb{R}$ and $\liminf_{n\to+\infty} I(f_n) < +\infty$, then $\liminf_{n\to+\infty} f_n \in L^1(I)$, and $I(\liminf_{n\to+\infty} f_n) \leq \liminf_{n\to+\infty} I(f_n)$.

Proof

For any $n \in \mathbb{N}$, let $g_n = \lim_{k\to+\infty}(f_n \wedge f_{n+1} \wedge \cdots \wedge f_{n+k})$. Then $\liminf_{n\to+\infty} f_n = \lim_{n\to+\infty} g_n$.($\liminf_{n\to+\infty} f_n : \Omega \to \mathbb{R}$, $\omega \mapsto \liminf_{n\to\infty} f_n(w)$)

For any k, one has $f_n \wedge \cdots \wedge f_{n+k} \geq g$. Hence $I(f_n) \geq \lim_{k\to+\infty} I(f_n \wedge \cdots \wedge f_{n+k}) \geq I(g)$.

By the theorem of BeppoLevi, $g_n \in L^1(I)$ and $I(g_n) = \lim_{k\to+\infty} I(f_n \wedge \cdots \wedge f_{n+k}) \leq I(f_n)$

Note that $(g_n)_{n\in\mathbb{N}}$ is increasing and $\liminf_{n\to+\infty} I(f_n) < +\infty$.

Hence $\lim_{n\to+\infty} I(g_n) = \liminf_{n\to+\infty} I(g_n) \leq \liminf_{n\to+\infty} I(f_n) < +\infty$

By the theorem of Beppolevi, $\lim_{n\to+\infty} g_n \in L^1(I)$ and $I(\liminf_{n\to+\infty} f_n) = I(\lim_{n\to+\infty} g_n) = \lim_{n\to+\infty} I(g_n) \leq \liminf_{n\to+\infty} I(f_n)$. □

6.3.13 Theorem(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 $I(f) = \lim_{n\to+\infty} I(f_n)$.

Proof

Apply Fatou's lemma to $(f_n)_{n\in\mathbb{N}}$ and $(-f_n)_{n\in\mathbb{N}}$ to get

$I(f) \leq \liminf_{n\to+\infty} I(f_n)$ and $I(-f) = -I(f) \leq \liminf_{n\to+\infty} I(-f_n) = -\limsup_{n\to+\infty} I(f_n)$.

Hence $I(f) \leq \liminf_{n\to+\infty} I(f_n) \leq \limsup_{n\to+\infty} I(f_n) \leq I(f)$ □

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 $\int_{\mathbb{R}} f(x)d\varphi(x)$. ($I_\varphi : S \to \mathbb{R}$, $\mathbb{1}_{]a,b]} \mapsto \varphi(b) - \varphi(a)$)

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 $(a,b) \in \mathbb{R}^2$, $a < b$. $\displaystyle\int_a^b f(x)d\varphi(x)$ denotes $\displaystyle\int_{]a,b]} f(x)d\varphi(x)$ and $\int_b^a f(x)d\varphi(x)$ denotes $-\displaystyle\int_{]a,b]} f(x)d\varphi(x)$.

If $\varphi(x) = x$ for any $x \in \mathbb{R}$, we replace $d\varphi(x)$ by dx.

Notation

Let A, $(A_i)_{i\in I}$ be sets. The 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 Semialgebra

6.4.1 Def

Let Ω be a set. We call semialgebra on Ω any $\mathcal{C} \subseteq \mathscr{P}(\Omega)$ that verifies:

(1) $\varnothing \in \mathcal{C}$

(2) $\forall (A,B) \in \mathcal{C}^2$, $A \bigcap B \in \mathcal{C}$

(3) $\forall (A,B) \in \mathcal{C}^2$, $\exists (C_i)_{i=1}^n$ a finite family of elements in $\mathcal{C}$, s.t. $B \setminus A = \bigsqcup_{i=1}^n C_i$

Example

$\Omega = \mathbb{R}$, $\mathcal{C} = \{]a,b] | (a,b) \in \mathbb{R}^2, a \leq b\}$.

6.4.2 Def

Let $\mathcal{C}$ be a semialgebra on Ω. The set $\{A \in \mathscr{P}(\Omega) | \exists n \in \mathbb{N}, \exists (A_i)_{i=1}^n \in \mathcal{C}^n, A = \bigsqcup_{i=1}^n A_i\}$ is called the algebra generated by $\mathcal{C}$.

6.4.3 Prop

Let $\mathcal{C}$ be a semialgebra on Ω, $\mathscr{A}$ be the algebra generated by $\mathcal{C}$.

Then (1) $\varnothing \in \mathscr{A}$

(2) $\forall (A,B) \in \mathscr{A}^2$, $A \bigcap B \in \mathscr{A}$, $B \setminus A \in \mathscr{A}$, $A \bigcup B \in \mathscr{A}$.

Proof

By definition, $\varnothing \in \mathscr{A}$.

Let $A = \bigsqcup_{i=1}^n A_i$, $B = \bigsqcup_{j=1}^m B_j$ be elements of $\mathscr{A}$, then $A \bigcap B = \bigsqcup_{(i,j) \in \{1,\cdots,n\} \times \{1,\cdots,m\}} A_i \bigcap B_j$. Hence $A \bigcap B \in \mathscr{A}$.

Since $B \setminus A = ((B \setminus A_1) \setminus A_2 \cdots) \setminus A_n$, by induction it suffices to treat the case where $A \in \mathcal{C}$.

In this case $B \setminus A = \bigsqcup_{j=1}^m (B_j \setminus A) \in \mathscr{A}$.

Finally $A \bigcup B = (A \bigcap B) \bigsqcup (A \setminus B) \bigsqcup (B \setminus A) \in \mathscr{A}$. $\square$

6.4.4 Prop

Let $\mathcal{C}$ be a semialgebra on Ω. $\mathscr{A}$ be the algebra generated by $\mathcal{C}$. Let S be the $\mathbb{R}$-vector subspace of $\mathbb{R}^\Omega$ generated by mappings of the form $\mathbb{1}_A$,

$A \in \mathcal{C}$, $f \in S$ is of the form $f = \sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}$, $\lambda_i \in \mathbb{R}$, $A_i \in \mathcal{C}$. $I : S \to \mathbb{R}$ be an $\mathbb{R}$-linear mapping.

Assume that $\forall (f,g) \in S \times S$, $f \leq g$, one has $I(f) \leq I(g)$.

Then I is an integral operator iff $\forall$ decreasing sequence $(A_n)_{n\in\mathbb{N}}$ in $\mathscr{A}^{\mathbb{N}}$ s.t. $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$, one has $\lim_{n\to+\infty} I(\mathbb{1}_{A_n}) = 0$

Proof

$\forall A \in \mathscr{A}$, $\exists (A_i)_{i=1}^n \in \mathcal{C}^n$, $A = \bigsqcup_{i=1}^{n} A_i$. So $\mathbb{1}_A = \sum_{i=1}^{n} \mathbb{1}_{A_i} \in S$.

Lemma: If $f = \sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}$, $g = \sum_{j=1}^{m} \mu_j \mathbb{1}_{B_j} \in S$, $A_i \in \mathcal{C}$, $B_j \in \mathcal{C}$, then $f \wedge g \in S$.

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 $\lim_{n\to+\infty} I(\mathbb{1}_{A_n}) = 0$.

Let $(f_n)_{n\in\mathbb{N}}$ be a decreasing sequence in S that converges pointwisely to 0.

Let $B = \{\omega \in \Omega | f_0(\omega) > 0\} \in \mathscr{A}$, $M = \max\{f_0(\omega) | \omega \in \Omega\}$

Lemma: $\forall f \in S$, $\exists (A_i)_{i=1}^n \in \mathcal{C}^n$ pairwisely disjoint, and $(\lambda_i)_{i=1}^n \in \mathbb{R}^n$, $f = \sum_{i=1}^{n} \lambda_i \mathbb{1}_{A_i}$.

For any $\varepsilon > 0$, let $A_n^\varepsilon = \{\omega \in \Omega | f_n(\omega) \geq \varepsilon\} \in \mathscr{A}$. Moreover, since $\lim_{n\to+\infty} f_n = 0$, $\bigcap_{n\in\mathbb{N}} A_n^\varepsilon = \varnothing$.

Note that $0 \leq f_n \leq \varepsilon \mathbb{1}_B + M \mathbb{1}_{A_n^\varepsilon}$. So $0 \leq I(f_n) \leq \varepsilon I(\mathbb{1}_B) + M I(\mathbb{1}_{A_n^\varepsilon})$, which leads to $\limsup_{n\to+\infty} I(f_n) \leq \varepsilon I(\mathbb{1}_B)$, $\forall \varepsilon > 0$.

Thus $\limsup_{n\to+\infty} I(f_n) = 0$. □

Example

$\Omega = \mathbb{R}$, $\mathcal{C} = \{]a,b] | (a,b) \in \mathbb{R}^2, a \leq b\}$, $\mathscr{A} =$ algebra generated by $\mathcal{C}$. $\varphi : \mathbb{R} \to \mathbb{R}$ increasing, right continuous. $S = \mathbb{R}$-vector subspace generated by $\mathbb{1}_{]a,b]}$, $(a,b) \in \mathbb{R}^2$, $a \leq b$.

$I_\varphi : S \to \mathbb{R}$. $I_\varphi(\mathbb{1}_{]a,b]}) = \varphi(b) - \varphi(a)$.

6.4.5 Lemma

$\forall \varepsilon > 0$, $\forall A \in \mathscr{A}$, $A \neq \varnothing$, $\exists B \in \mathscr{A}$, $\varnothing \neq \overline{B} \subseteq A$ and $I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) \leq \varepsilon$.

Proof

We first consider the case where $A \in \mathcal{C}$. $A =]a, b]$, $a < b$.

By the right continuous of φ, $\exists a' \in]a, b[$ s.t. $\varphi(a') - \varphi(a) \leq \varepsilon$.

Let $B =]a', b]$. $\overline{B} = [a', b] \subseteq]a, b]$.

$I_\varphi(\mathbb{1}_B) = \varphi(b) - \varphi(a')$, $I_\varphi(\mathbb{1}_A) = \varphi(b) - \varphi(a)$.

$I_\varphi(\mathbb{1}_A) - I_\varphi(\mathbb{1}_B) = \varphi(a') - \varphi(a) \leq \varepsilon$.

In general $A = \bigsqcup_{i=1}^{n} A_i$ with $A_i \in \mathcal{C}$。 $\forall i \in \{1, \cdots, n\}$, $\exists B_i \in \mathcal{C}$.

$\varnothing \neq \overline{B_i} \subseteq A_i$, $I(\mathbb{1}_{A_i}) - I(\mathbb{1}_{B_i}) \leq \frac{\varepsilon}{n}$.

Let $B = \bigsqcup_{i=1}^{n} B_i$. Then $I(\mathbb{1}_A) - I(\mathbb{1}_B) = \sum_{i=1}^{n} (I(\mathbb{1}_{A_i}) - I(\mathbb{1}_{B_i})) \leq \varepsilon$. □

6.4.6 Theorem

I_φ is an integral operator.

Proof

Let $(A_n)_{n\in\mathbb{N}}$ be a decreasing sequence in $\mathscr{A}$ s.t. $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$. Let $\varepsilon > 0$. For any $n \in \mathbb{N}$, let $B_n \in \mathscr{A}$ s.t. $\overline{B_n} \subseteq A_n$ and $I_\varphi(\mathbb{1}_{A_n}) - I_\varphi(\mathbb{1}_{B_n}) \leq \frac{\varepsilon}{2^n}$。

Note that $\overline{B_n}$ is compact. For any $n \in \mathbb{N}$, let $C_n = B_0 \bigcap \cdots \bigcap B_n \subseteq \overline{B_0} \bigcap \cdots \bigcap \overline{B_n}$

Since $\bigcap_{n\in\mathbb{N}} A_n = \varnothing$, $\bigcap_{n\in\mathbb{N}} \overline{B_n} = \varnothing$.

Hence $\exists N \in \mathbb{N}$, $\bigcap_{n=0}^{N} \overline{B_n} = \varnothing$. Moreover, $B_n \setminus C_n = B_n \setminus (B_n \bigcap C_{n-1}) = B_n \setminus C_{n-1} \subseteq A_n \setminus C_{n-1} \subseteq A_{n-1} \setminus C_{n-1}$

Hence $I_\varphi(\mathbb{1}_{A_n \setminus C_n}) = I_\varphi(\mathbb{1}_{B_n \setminus C_n}) + I_\varphi(\mathbb{1}_{A_n \setminus B_n}) \leq I_\varphi(\mathbb{1}_{A_{n-1} \setminus C_{n-1}}) + \frac{\varepsilon}{2^n}$.

So $I_\varphi(\mathbb{1}_{A_n}) \leq \frac{\varepsilon}{2^n} + \frac{\varepsilon}{2^{n-1}} + \cdots + \frac{\varepsilon}{2} \leq \varepsilon$, $\forall n \geq N$.

Thus $\lim_{n\to+\infty} I_\varphi(\mathbb{1}_{A_n}) = 0$. □

Let Ω be a set, $\mathcal{C}$ be a semialgebra on Ω, $\mathscr{A}$ be the algebra generated by $\mathcal{C}$, S be the vector subspace of $\mathbb{R}^\Omega$ generated by mappings of the form $\mathbb{1}_A$ with $A \in \mathcal{C}$.

6.4.7 Prop

For any $f \in S$, $\exists (A_i)_{i=1}^n \in \mathcal{C}^n$ pairwisely disjoint, and $(\lambda_i)_{i=1}^n \in \mathbb{R}^n$, $f = \sum\limits_{i=1}^n \lambda_i \mathbb{1}_{A_i}$.

Proof

f is of the form $\sum\limits_{j=1}^m a_j \mathbb{1}_{B_j}$. $B_j \in \mathcal{C}$.

For any $I \subseteq \{1, \cdots, m\}$, $I \neq \varnothing$, let $B_I = (\bigcap\limits_{i \in I} B_i) \bigcap (\bigcap\limits_{j \in \{1,\cdots,m\} \setminus I} (\Omega \setminus B_j))$

Then $(B_I)_{I \subseteq \{1,\cdots,m\}}$ are pairwisely disjoint. Moreover, if $I \neq \varnothing$, $B_I \in \mathcal{C}$.

$B_i = \bigsqcup\limits_{I \subseteq \{1,\cdots,m\}, i \in I} B_I$. Hence $f = \sum\limits_{\varnothing \neq I \subseteq \{1,\cdots,m\}} (\sum\limits_{j \in I} a_j) \mathbb{1}_{B_I}$. $\square$

6.4.8 Corollary

(1) If $f \in S$, then $f \wedge 0 \in S$.

(2) If $(f, g) \in S^2$, then $f \wedge g = (f - g) \wedge 0 + g \in S$

6.4.9 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. $B \setminus A = \bigsqcup\limits_{i=1}^n C_i$, one has $\mu(B) = \mu(A) + \sum\limits_{i=1}^n \mu(C_i)$.

Then there is a unique $\mathbb{R}$-linear mapping $I_\mu : S \to \mathbb{R}$ s.t. $I_\mu(\mathbb{1}_A) = \mu(A)$, $\forall A \in \mathcal{C}$.

Proof

We intend to define $I_\mu(\sum\limits_{i=1}^n \lambda_i \mathbb{1}_{A_i})$ as $\sum\limits_{i=1}^n \lambda_i \mu(A_i)$ for $A_i \in \mathcal{C}$.

We need to check that if $f \in S$ is written as $f = \sum\limits_{i=1}^n \lambda_i \mathbb{1}_{A_i} = \sum\limits_{j=1}^m \xi_j \mathbb{1}_{B_j}$, then $\sum\limits_{i=1}^n \lambda_i \mu(A_i) = \sum\limits_{j=1}^m \xi_j \mu(B_j)$.

$0 = \sum\limits_{i=1}^n \lambda_i \mu(A_i) - \sum\limits_{j=1}^m \xi_j \mu(B_j)$

It suffices to prove that, if $\sum\limits_{i=1}^n a_i \mathbb{1}_{A_i} = 0$, $a_i \in \mathbb{R}$, $A_i \in \mathcal{C}$, then $\sum\limits_{i=1}^n a_i \mu(A_i) = 0$.

For $I \subseteq \{1, \cdots, n\}$, let $A_I = \{\omega \in \Omega | \forall i \in I, \omega \in A_i, \forall i \in \{1, \cdots, n\} \setminus I, \omega \in \Omega \setminus A_i\} \in \mathscr{A}$, when $I \neq \varnothing$.

Lemma

Let $B \in \mathscr{A}$. If $B = \bigsqcup\limits_{i=1}^{n} B_i = \bigsqcup\limits_{j=1}^{m} C_j$ with $B_i \in \mathcal{C}$, $C_j \in \mathcal{C}$, then $\sum\limits_{i=1}^{n} \mu(B_i) = \sum\limits_{j=1}^{m} \mu(C_j)$. 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 $D_1, \cdots, D_n$ in $\mathscr{A}$ disjoint, $\mu(D_1 \bigsqcup \cdots \bigsqcup D_n) = \sum\limits_{i=1}^{n} \mu(D_i)$.

Proof

$B_i = \bigsqcup\limits_{j=1}^{m} (B_i \bigcap C_j)$, $\mu(B_i) = \sum\limits_{j=1}^{m} \mu(B_i \bigcap C_j)$

$\sum\limits_{i=1}^{n} \mu(B_i) = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{m} \mu(B_i \bigcap C_j) = \sum\limits_{j=1}^{m} \sum\limits_{i=1}^{n} \mu(C_j \bigcap B_i) = \sum\limits_{j=1}^{m} \mu(C_j)$. $\square$

Back to the proof of the proposition, $0 = \sum\limits_{i=1}^{n} a_i \mathbb{1}_{A_i} = \sum\limits_{\varnothing \neq I \subseteq \{1,\cdots,n\}} (\sum\limits_{i \in I} a_i) \mathbb{1}_{A_I}$

Hence when $A_I \neq \varnothing$, $\sum\limits_{i \in I} a_i = 0$

$\sum\limits_{i=1}^{n} a_i \mu(A_i) = \sum\limits_{i=1}^{n} a_i \sum\limits_{i \in I \subseteq \{1,\cdots,n\}} \mu(A_I) = \sum\limits_{\varnothing \neq I \subseteq \{1,\cdots,n\}} \mu(A_I) \sum\limits_{i \in I} a_i = 0.$

$\square$

6.5 Integrable functions

Let Ω be a set, $S \subseteq \mathbb{R}^{\Omega}$ be a vector subspace, $\forall (f,g) \in S^2$, $f \wedge g \in S$. $I : S \to \mathbb{R}$ is an 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 $\mu : \mathcal{G} \to \mathbb{R}_{\geq 0} \bigcup \{+\infty\}$ the mapping defined as $\mu(A) := I(\mathbb{1}_A)$, then μ satisfies: for any $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$ that is pairwisely disjoint, then $\mu(\bigsqcup\limits_{n\in\mathbb{N}} A_n) = \sum\limits_{n\in\mathbb{N}} \mu(A_n)$.

Proof

(1) $\varnothing \in \mathcal{G}$ since $0 = \mathbb{1}_{\varnothing} \in L^1(I)$, $\Omega \in \mathcal{G}$ 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 $\mathbb{1}_B - \mathbb{1}_A = \mathbb{1}_{B\setminus A} \in L^1(I)^{\uparrow}$, $B \setminus A \in \mathcal{G}$.

(In fact, let f and g be two mappings belong to $L^1(I)^{\uparrow}$ and $0 \leq f \leq g$. Let $(g_n)_{n\in\mathbb{N}}$ be an increasing sequence that converges pointwisely to g, then $\forall n \in \mathbb{N}$, $g_n \wedge f \in L^1(I)^{\uparrow}$, since $|g_n \wedge f| \leq g_n$, hence $g_n \wedge f \in L^1(I)$, thus $g_n - g_n \wedge f \in L^1(I)$, $\lim\limits_{n\to+\infty} g_n - g_n \wedge f = g - f$ and $g_n - g_n \wedge f = g_n \vee f - f$ is increaing. Therefore, $g - f \in L^1(I)$)

(3) If $(A,B) \in \mathcal{G}^2$, $\mathbb{1}_{A\bigcup B} = \mathbb{1}_A \vee \mathbb{1}_B \in L^1(I)^{\uparrow}$, so $A \bigcup B \in \mathcal{G}$

If $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$, $A = \bigcup\limits_{n\in\mathbb{N}} A_n$, then $\mathbb{1}_A = \lim\limits_{n\to+\infty} \mathbb{1}_{A_0\bigcup\cdots\bigcup A_n} \in L^1(I)^{\uparrow}$, thus $A \in \mathcal{G}$.

$\mathbb{1}_{\bigsqcup\limits_{n\in\mathbb{N}} A_n} = \lim\limits_{n\to+\infty} \mathbb{1}_{A_0\bigsqcup\cdots\bigsqcup A_n} = \lim\limits_{n\to+\infty} \sum\limits_{i=0}^{n} \mathbb{1}_{A_i}$. Then $\mu(\bigsqcup\limits_{n\in\mathbb{N}} A_n) = I(\mathbb{1}_{\bigsqcup\limits_{n\in\mathbb{N}} A_n}) = \lim\limits_{n\to+\infty} I(\sum\limits_{i=0}^{n} \mathbb{1}_{A_i}) = \lim\limits_{n\to+\infty} \sum\limits_{i=1}^{n} \mu(A_i) = \sum\limits_{n\in\mathbb{N}} \mu(A_n)$. $\square$

6.6 Limit and differential of integrals with parameters

Let Ω be a set, $S \subseteq \mathbb{R}^{\Omega}$ be an vector subspace over $\mathbb{R}$ s.t. $\forall (f,g) \in S^2$, $f \wedge g \in S$.

Let $I : S \to \mathbb{R}$ be an integral operator.

6.6.1 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) $\forall \omega \in \Omega$, $f(\omega, \cdot) : X \to \mathbb{R}$, $x \mapsto f(\omega, x)$ is continuous at p.

(2) $\forall x \in X$, $f(\cdot, x) : \Omega \to \mathbb{R}$, $\omega \mapsto f(\omega, x)$ 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.

Proof

Let $(x_n)_{n \in \mathbb{N}}$ be a sequence in X that converges to p.

For any $n \in \mathbb{N}$, let $f_n : \Omega \to \mathbb{R}$, $f_n(\omega) := f(\omega, x_n)$. One has $|f_n| \leq g$.

Moreover, $\forall \omega \in \Omega$, $\lim\limits_{n \to +\infty} f_n(\omega) = \lim\limits_{n \to +\infty} f(\omega, x_n) = f(\omega, p)$.

Hence, by dominated convergence theorem, $\lim\limits_{n \to +\infty} I(f_n) = I(f(\cdot, p))$. $\square$

Note

在 ode "解对初值和参数的连续依赖性"中有用。

6.6.2 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) $\forall \omega \in \Omega$, $f(\omega, \cdot) : J \to \mathbb{R}$, $t \mapsto f(\omega, t)$ is differentiable, (we denote by $\frac{\partial f}{\partial t}(\omega, t)$ its derivative at t) and $\forall t \in J$, $\frac{\partial f}{\partial t}(\cdot, t) \in L^1(I)$, $|\frac{\partial f}{\partial t}(\omega, t)| \leq g(\omega)$.

(2) $\forall t \in J$, $f(\cdot, t) : \Omega \to \mathbb{R}$, $\omega \mapsto f(\omega, t)$ belongs to $L^1(I)$.

Then, $\varphi : J \to \mathbb{R}$, $t \mapsto I(f(\cdot, t))$ is differentiable and $\varphi'(t) = I(\frac{\partial f}{\partial t}(\cdot, t))$.

Proof

Let $a \in J$ and $(t_n)_{n\in\mathbb{N}}$ be a sequence in $J\setminus\{a\}$ s.t. $\lim\limits_{n\to+\infty} t_n = a$. Then $\frac{\varphi(t_n)-\varphi(a)}{t_n-a} = I(\frac{f(\cdot,t_n)-f(\cdot,a)}{t_n-a})$

$\forall\omega\in\Omega$, $|\frac{f(\omega,t_n)-f(\omega,a)}{t_n-a}| \le g(\omega)$(by mean value theorem) and $\lim\limits_{n\to+\infty} \frac{f(\omega,t_n)-f(\omega,a)}{t_n-a} = \frac{\partial f}{\partial t}(\omega,a)$

Hence $\lim\limits_{n\to+\infty} \frac{\varphi(t_n)-\varphi(a)}{t_n-a} = I(\frac{\partial f}{\partial t}(\cdot,a))$ $\square$

Note

在庞加莱引理和 ode "解对初值和参数的连续可微性" 中有用。

6.7 Measure theory

6.7.1 Def

We call measurable space any pair (E, ε), where E is a set and ε is a σ-algebra on E.

6.7.2 Prop

Let Ω be a set and $(\mathcal{G}_i)_{i\in I}$ be a family of σ-algebras on Ω. Then $\bigcap\limits_{i\in I} \mathcal{G}_i$ is a σ-algebra.

Proof

$\forall i \in I$, $\varnothing \in \mathcal{G}_i$, hence $\varnothing \in \mathcal{G}_i$

If $A \in \bigcap\limits_{i\in I} \mathcal{G}_i$, then $\forall i \in I$, $A \in \mathcal{G}_i$. Hence $\forall i \in I$, $\Omega \setminus A \in \mathcal{G}_i$, so $\Omega \setminus A \in \bigcap\limits_{i\in I} \mathcal{G}_i$.

Let $(A_n)_{n\in\mathbb{N}} \in (\bigcap\limits_{i\in I} \mathcal{G}_i)^{\mathbb{N}}$. For any $i \in I$, $(A_n)_{n\in\mathbb{N}} \in (\mathcal{G}_i)^{\mathbb{N}}$. So $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{G}_i$ so $\bigcup\limits_{n\in\mathbb{N}} A_n \in \bigcap\limits_{i\in I} \mathcal{G}_i$. $\square$

6.7.3 Def

Let $\mathcal{C} \subseteq \mathscr{P}(\Omega)$. We denote by $\sigma(\mathcal{C})$ the intersection of all $\sigma-$algebras on Ω containg $\mathcal{C}$. It is the smallest σ-algebra containing $\mathcal{C}$.

Example 1

Let (X, τ) be a topological space. $\sigma(\tau)$ is called the Borel $\sigma-$algebra of X.

Example 2

On $[-\infty, +\infty]$, the following $\sigma-$algebras 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 $\mathscr{B} = \{A \subseteq \mathbb{R} | A \in \mathcal{G}_1\}$ is equal to the Borel $\sigma-$algebra of $\mathbb{R}$.

$\forall a \in \mathbb{R}$, $[a, +\infty] = \bigcap\limits_{n\in\mathbb{N}_{\geq 1}}]a - \frac{1}{n}, +\infty] \in \mathcal{G}_2$, $\mathcal{G}_1 \subseteq \mathcal{G}_2$.

$]a, +\infty] = [-\infty, +\infty] \setminus [-\infty, a] \in \mathcal{G}_3$, $\mathcal{G}_2 \subseteq \mathcal{G}_3$

$[-\infty, a] = \bigcap\limits_{n\in\mathbb{N}_{\geq 1}} [-\infty, a+\frac{1}{n}[\in \mathcal{G}_4$, $\mathcal{G}_3 \subseteq \mathcal{G}_4$

$[-\infty, a[= [-\infty, +\infty] \setminus [a, +\infty] \in \mathcal{G}_1$, $\mathcal{G}_4 \subseteq \mathcal{G}_1$.

Borel $\sigma-$algebra of $\mathbb{R}$=$\sigma(\{]a, b[|a < b, (a, b) \in \mathbb{R}^2\})$.

But $]a, b[=]a, +\infty] \bigcap [-\infty, b[$, hence Borel $\sigma-$algebra of $\mathbb{R} \subseteq \{A \subseteq \mathbb{R}|A \in \mathcal{G}_1\}$.

Then we need to show that $\{A \subseteq \mathbb{R}|A \in \mathcal{G}_1\} \subseteq$ Borel $\sigma-$algebra of $\mathbb{R}$. Let $G = \{[a, +\infty]|a \in \mathbb{R}\}$, Only needs to prove $\{A \subseteq \mathbb{R}|A \in \sigma(G)\} = \sigma(\{A \bigcap \mathbb{R}|A \in G\})$. Notice that $\{A \subseteq \mathbb{R}|A \in \sigma(G)\}$ forms a σ-algebra in $\mathbb{R}$. In fact, $\forall A \in \overline{\mathbb{R}}$, $\mathbb{R} \setminus (A \bigcap \mathbb{R}) = \mathbb{R} \setminus A = (\overline{\mathbb{R}} \setminus A) \setminus (\overline{\mathbb{R}} \setminus \mathbb{R}) = (\overline{\mathbb{R}} \setminus A) \bigcap (\overline{\mathbb{R}} \setminus (\overline{\mathbb{R}} \setminus \mathbb{R})) = (\overline{\mathbb{R}} \setminus A) \bigcap \mathbb{R}$, thus $\mathbb{R} \setminus (A \bigcap \mathbb{R}) \in \{A \subseteq \mathbb{R}|A \in \sigma(G)\}$ and $\forall \{E_n\} \subseteq \overline{\mathbb{R}}$, $\bigcup\limits_n (E_n \bigcap \mathbb{R}) = (\bigcup\limits_n E_n) \bigcap \mathbb{R}$.

6.7.4 Def

Let $f : X \to Y$ be a mapping of sets.

For any $\mathcal{C}_Y \subseteq \mathscr{P}(Y)$, we denote by $f^{-1}(\mathcal{C}_Y) := \{f^{-1}(B)|B \in \mathcal{C}_Y\}$.

For any $\mathcal{C}_X \subseteq \mathscr{P}(X)$, we denote by $f_*(\mathcal{C}_X) := \{B \subseteq Y|f^{-1}(B) \in \mathcal{C}_X\}$

6.7.5 Prop

Let $f : X \to Y$ be a mapping. If $\mathcal{G}_Y$ is a $\sigma-$algebra on Y, $f^{-1}(\mathcal{G}_Y)$ is a $\sigma-$algebra on X.

If $\mathcal{G}_X$ is a $\sigma-$algebra on X, $f_*(\mathcal{G}_X)$ is a $\sigma-$algebra on Y.

Proof

(1) $\varnothing = f^{-1}(\varnothing) \in f^{-1}(\mathcal{G}_Y)$

$\forall B \in \mathcal{G}_Y$, $X \setminus f^{-1}(B) = f^{-1}(Y \setminus B)$

If $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}_Y^{\mathbb{N}}$, $A = \bigcup\limits_{n\in\mathbb{N}} A_n \in \mathcal{G}_Y$, then $\bigcup\limits_{n\in\mathbb{N}} f^{-1}(A_n) = f^{-1}(A) \in f^{-1}(\mathcal{G}_Y)$

(2) $f^{-1}(\varnothing) = \varnothing \in \mathcal{G}_X$, so $\varnothing \in f_*(\mathcal{G}_X)$

$\forall B \in f_*(\mathcal{G}_X)$, $f^{-1}(Y \setminus B) = X \setminus f^{-1}(Y) \in \mathcal{G}_X$, so $Y \setminus B \in f_*(\mathcal{G}_X)$

$\forall (B_n)_{n\in\mathbb{N}} \in f_*(\mathcal{G}_X)^{\mathbb{N}}$, $B = \bigcup\limits_{n\in\mathbb{N}} B_n$, $f^{-1}(B) = \bigcup\limits_{n\in\mathbb{N}} f^{-1}(B_n)$, so $B \in f_*(\mathcal{G}_X)$. $\square$

6.7.6 Def

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 $\mathcal{G}_Y \subseteq f_*(\mathcal{G}_X)$ (or $\forall B \in \mathcal{G}_Y$, $f^{-1}(B) \in \mathcal{G}_X$), then we say that f is measurable.

6.7.7 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.

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 $f^{-1}(g^{-1}(B)) \in \mathcal{G}_X$. $\square$

6.7.8 Def

Let Ω be a set, $((E_i, \varepsilon_i))_{i \in I}$ 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 $\sigma-$algebra. $\sigma(\bigcup\limits_{i \in I} f_i^{-1}(\varepsilon_i))$. It is the smallest $\sigma-$algebra on Ω making all f_i measurable.

6.7.9 Prop

We keep the notation of the above definition. For any $i \in I$, let $\mathcal{C}_i \subseteq \mathscr{P}(E_i)$ s.t. $\sigma(\mathcal{C}_i) = \varepsilon_i$. Then $\sigma(f) = \sigma(\bigcup\limits_{i \in I} f_i^{-1}(\mathcal{C}_i))$

Proof

Let $\mathcal{G} = \sigma(\bigcup\limits_{i \in I} f_i^{-1}(\mathcal{C}_i))$. By definition $\mathcal{G} \subseteq \sigma(f)$.

For any $i \in I$, $f_{i,*}(\sigma(f_i^{-1}(\mathcal{C}_i)))$ is a $\sigma-$algebra on E_i containing $\mathcal{C}_i$, so $\varepsilon_i \subseteq f_{i,*}(\sigma(f_i^{-1}(\mathcal{C}_i)))$, which leads to $f_i^{-1}(\varepsilon_i) \subseteq \sigma(f_i^{-1}(\mathcal{C}_i)) \subseteq \mathcal{G}$. Hence $\bigcup\limits_{i \in I} f_i^{-1}(\varepsilon_i) \subseteq \mathcal{G}$, $\sigma(f) \subseteq \mathcal{G}$. $\square$

6.7.10 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. $\mathcal{G}_Y = \sigma(\mathcal{C}_Y)$. Then f is measurable iff $\forall B \in \mathcal{C}_Y$, $f^{-1}(B) \in \mathcal{G}_X$.

Proof

$\sigma(f) = \sigma(f^{-1}(\mathcal{C}_Y))$. f is measurable iff $\sigma(f) \subseteq \mathcal{G}_X$. $\square$

Example

Let $((E_i, \varepsilon_i))_{i\in I}$ be a family of measurable spaces. $E = \prod\limits_{i\in I} E_i$.

$\forall i \in I$, $\pi_i : E \to E_i$, $(x_j)_{j\in I} \mapsto x_i$. We denote by $\bigotimes\limits_{i\in I} \varepsilon_i$ the σ-algebra $\sigma((\pi_i)_{i\in I})$.

6.7.11 Prop

Let X be a set, $((E_i, \varepsilon_i))_{i\in I}$ be measurable spaces. $(\Omega, \mathcal{G})$ be a measurable space. $f = (f_i : X \to E_i)_{i\in I}$ be mappings, $\varphi : \Omega \to X$ be a mapping. Then $\varphi : (\Omega, \mathcal{G}) \to (X, \sigma(f))$ is measurable iff $\forall i \in I$, $f_i \circ \varphi : (\Omega, \mathcal{G}) \to (E, \varepsilon_i)$ is measurable.

Proof

If φ is measurable, since each f_i is measurable, one has $f_i \circ \varphi$ is measurable.

If $f_i \circ \varphi$ is measurable, $\forall B \in \varepsilon_i$, $(f_i \circ \varphi)^{-1}(B) = \varphi^{-1}(f_i^{-1}(B)) \in \mathcal{G}$. Hence $\varphi^{-1}(\bigcup\limits_{i\in I} f_i^{-1}(\varepsilon_i)) \subseteq \mathcal{G}$. Since $\sigma(f) = \sigma(\bigcup\limits_{i\in I} f_i^{-1}(\varepsilon_i))$, φ is measurable. $\square$

Example

(1) Let $(\Omega, \mathcal{G})$ be a measurable space.

$\forall A \in \mathcal{G}$, $\mathbb{1}_A : \Omega \to \mathbb{R}$ is measurable.

For any $U \subseteq \mathbb{R}$, $\mathbb{1}_A^{-1}(U) = A$ or $\Omega \setminus A$ or Ω or $\varnothing$.

(2) If X and Y be topological spaces.

$f : X \to Y$ is a continuous mapping, then f is measurable with respect to Borel σ-algebras.

In fact, $\forall V \subseteq Y$ open, $f^{-1}(V) \subseteq X$ open.

(3) Let $(\Omega, \mathcal{G})$ be a measurable space. If $f : \Omega \to \mathbb{R}$, $g : \Omega \to \mathbb{R}$ are measurable, then $f + g$, fg, $f \wedge g$, $f \vee g$, $|f|$ are measurable.

$\Omega \to \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $\omega \mapsto (f(\omega), g(\omega)) \mapsto f(\omega) + g(\omega)$(Notice that we use the fact that $\sigma(\mathcal{O}_{\mathbb{R}^2}) = \sigma(\mathcal{O}_{\mathbb{R}}) \otimes \sigma(\mathcal{O}_{\mathbb{R}})$ implicitly).

In fact,

$$\begin{tikzcd}
 & & (\mathbb{R}, \mathcal{O}_{\mathbb{R}}) \\
(\Omega, \mathcal{G}) \arrow[r, "\Psi"] & (\mathbb{R}^2, \mathcal{O}_{\mathbb{R}} \otimes \mathcal{O}_{\mathbb{R}} = \sigma(g)) \arrow[ur, "\pi_1"] \arrow[dr, "\pi_2"'] & \\
 & & (\mathbb{R}, \mathcal{O}_{\mathbb{R}})
\end{tikzcd}$$

Since $f = \pi_1 \circ \Psi$ and $g = \pi_2 \circ \Psi$ are measurable, Ψ is measurable.

$$\begin{tikzcd}
(\Omega, \mathcal{G}) \arrow[r, "\Psi"] & (\mathbb{R}^2, \mathcal{O}_{\mathbb{R}^2}) \arrow[r, "\Phi"] & (\mathbb{R}, \mathcal{O}_{\mathbb{R}}) \\
x \arrow[r, mapsto] & (f(x), g(x)) \arrow[r, mapsto] & f(x) + g(x)
\end{tikzcd}$$

Since $\mathcal{O}_{\mathbb{R}^2} = \mathcal{O}_{\mathbb{R}} \otimes \mathcal{O}_{\mathbb{R}}$, Ψ is still measurable and Φ is a continuous mapping, hence $\Phi \circ \Psi = f + g$ is measurable.

(4) Let $(f_n)_{n \in \mathbb{N}}$ be a family of measurable mappings from Ω to $[-\infty, +\infty]$.

$f = \sup\limits_{n \in \mathbb{N}} f_n$, $f : \Omega \to [-\infty, +\infty]$, $f(\omega) = \sup\limits_{n \in \mathbb{N}} f_n(\omega)$

Then f is measurable. (Similarly, $\inf\limits_{n \in \mathbb{N}} f_n$ is measurable)

In fact, for any $a \in \mathbb{R}$, $\{\omega \in \Omega | f(\omega) > a\} = \bigcup\limits_{n \in \mathbb{N}} \{\omega \in \Omega | f_n(\omega) > a\} = \bigcup\limits_{n \in \mathbb{N}} f_n^{-1}(]a, +\infty])$

6.8 Measure

6.8.1 Def

Let Ω be a set. $\mathcal{C}$ be a semialgebra 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 $A = A_1 \bigsqcup \cdots \bigsqcup A_n \in \mathcal{C}$, one has $\mu(A) = \mu(A_1) + \cdots + \mu(A_n)$, we say that μ is additive.

Moreover, let $S =$ vector subspace of $\mathbb{R}^\Omega$ generated by $\mathbb{1}_A$, $A \in \mathcal{C}$. Then $I_\mu : S \to \mathbb{R}$, $\sum\limits_{i=1}^n \lambda_i \mathbb{1}_{A_i} \mapsto \sum\limits_{i=1}^n \lambda_i \mu(A_i)$ is well defined. If I_μ is an integral operator, we say that μ is $\sigma-$ additive.

6.8.2 Def

Let $(\Omega, \mathcal{G})$ be a measurable space. $\mu : \mathcal{G} \to [0, +\infty]$ be a mapping. If $\mu(\varnothing) = 0$ and if for any $(A_n)_{n\in\mathbb{N}} \in \mathcal{G}^{\mathbb{N}}$ pairwisely disjoint, $\mu(\bigsqcup\limits_{n\in\mathbb{N}} A_n) = \sum\limits_{n\in\mathbb{N}} \mu(A_n)$, we say that μ is a measure.

Notation

$(\Omega, \mathcal{G}, \mu)$ is called a measure space.

6.8.3 Theorem(Caratheodory)

Let Ω be a set, $\mathcal{C}$ be a semialgebra on Ω, $\mu : \mathcal{C} \to \mathbb{R}_{\geq 0}$ be a σ-additive mapping. Assume that there is a sequence $(A_n)_{n\in\mathbb{N}} \in \mathcal{C}^{\mathbb{N}}$ s.t. $\Omega = \bigcup\limits_{n\in\mathbb{N}} A_n$. Then μ extends to a σ-finite measure on $\sigma(\mathcal{C})$. (σ-finite measure means if $\exists (A_n)_{n\in\mathbb{N}} \in \sigma(\mathcal{C})^{\mathbb{N}}$, s.t. $\Omega = \bigcup\limits_{n\in\mathbb{N}} A_n$ and $\forall n \in \mathbb{N}$, $\mu(A_n) < +\infty$, then μ is said to be σ-finite.)

Proof

Let $S \subseteq \mathbb{R}^\Omega$ 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 containging $\mathcal{C}$. 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 $\sigma-$finite. $\square$

Example

$\Omega = \mathbb{R}$, $\mathcal{C} = \{]a, b] | (a, b) \in \mathbb{R}^2, a < b\}$. $\sigma(\mathcal{C}) =$ Borel σ-algebra of $\mathbb{R}$. $\varphi : \mathbb{R} \to \mathbb{R}$ increasing right continuous, $\mu_\varphi : \mathcal{C} \to \mathbb{R}_{\geq 0}$, $]a, b] \mapsto \varphi(b) - \varphi(a)$

is $\sigma-$additive.

Hence μ_φ extends to a measure $\sigma(\mathcal{C}) \to [0, +\infty]$ called the Stieltjes measure. In the particular case where $\varphi(x) = x$, $\forall x \in \mathbb{R}$, μ_φ is called the Lebesgue measure.

6.8.4 Def

Let $(\Omega, \mathcal{G}, \mu)$ be a σ-finite measure space. Then $\mathcal{C} = \{A \in \mathcal{G} | \mu(A) < +\infty\}$ is a semialgebra, $\sigma(\mathcal{C}) = \mathcal{G}$ (Actually, $\exists \bigcup\limits_{n\in\mathbb{N}} A_n = \mathcal{G}$ s.t. $\mu(A_n) < +\infty$, $\forall n$, hence $\forall E \in \mathcal{G}$, $E = \bigcup\limits_{n\in\mathbb{N}} (E \bigcap A_n)$, $\mu(A_n \bigcap E) < +\infty$, thus $E \in \sigma(\mathcal{C})$) and $\mu|_{\mathcal{C}}$ is σ-additive.

In fact, if $A_0 \supseteq A_1 \supseteq \cdots \supseteq \cdots$, $\bigcap\limits_{n\in\mathbb{N}} A_n = \varnothing$. Let $B_n = A_n \setminus A_{n+1}$, then $\bigcup\limits_{n\in\mathbb{N}} B_n = A_0$, $\mu(A_0) = \sum\limits_{n\in\mathbb{N}} \mu(B_n) < +\infty$, $\sum\limits_{k\geq n} \mu(B_k) = \mu(A_n) \to 0$, $n \to +\infty$

We denote 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 denote as $\displaystyle\int_\Omega f(\omega)\mu(d\omega)$.

Particular case, if $\Omega = \mathbb{R}$. $\mu = \mu_\varphi$ Stieltjes measure, $\displaystyle\int_{\mathbb{R}} f(x)\mu_\varphi(dx)$ is denoted as $\displaystyle\int_{\mathbb{R}} f(x)d\varphi(x)$.

6.8.5 Prop

Let $(\Omega, \mathcal{G}, \mu)$ be a $\sigma-$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 $f \in L^1(\Omega, \mathcal{G}, \mu)^\uparrow$

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 $\mathcal{G}$. $\forall n \in \mathbb{N}$, $\mu(A_n) < +\infty$, $\bigcup\limits_{n\in\mathbb{N}} A_n = \Omega$. Then $\mathbb{1}_B = \lim\limits_{n\to+\infty} \mathbb{1}_{B\bigcap A_n} \in L^1(\Omega, \mathcal{G}, \mu)^\uparrow$

For general $f \geq 0$, $f = \lim\limits_{n\to+\infty} f_n$

$f_n = \sum\limits_{k=0}^{n2^n-1} \frac{k}{2^n} \mathbb{1}_{\{\omega\in\Omega | \frac{k}{2^n} \leq f(\omega) < \frac{k+1}{2^n}\}} + n\mathbb{1}_{\{\omega\in\Omega | f(\omega)\geq n\}} \in L^1(\Omega, \mathcal{G}, \mu)^\uparrow$ $\square$

6.8.6 Corollary

Let $f : \Omega \to \mathbb{R}$ be a measurable mapping. Then $f \in L^1(\Omega, \mathcal{G}, \mu)$ iff

$\int_\Omega |f(\omega)|\mu(dw) < +\infty$.

Proof

One has $f \in L^1(\Omega, \mathcal{G}, \mu)$. Hence $|f| \in L^1(\Omega, \mathcal{G}, \mu)$. So $\int_\Omega |f(\omega)|\mu(dw) = I_\mu(|f|) < +\infty$.

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 $f = f \vee 0 + f \wedge 0 \in L^1(\Omega, \mathcal{G}, \mu)$. $\square$

6.9 Fundamental theorem of calculus

6.9.1 Theorem

Let J be an open interval in $\mathbb{R}$, $x_0 \in J$, $f : J \to \mathbb{R}$ be a continuous mapping.

(1) $\forall (a,b) \in J \times J$, $a < b$.

$\mathbb{1}_{]a,b]} f(x) = \begin{cases} f(x) & \text{if } x \in]a,b] \\ 0 & \text{if else} \end{cases}$

This mapping belongs to $L^1(\mathbb{R}, \mathcal{B}(\text{Borel } \sigma\text{-algebra of } \mathbb{R}), \mu(\text{Lebesgue measure}))$

(2) Let $F : J \to \mathbb{R}$. $F(x) = \displaystyle\int_{x_0}^{x} f(t)dt$. Then F is differentiable on J with $F'(x) = f(x)$, $\forall x \in J$

Proof

(1) f is bounded on $[a,b]$. Hence $\displaystyle\int_{\mathbb{R}} \mathbb{1}_{]a,b]}(x)|f(x)|dx < +\infty$

(2) Let $x \in J$, $h > 0$, s.t. $[x, x+h] \subseteq J$. f is uniformly continuous on $[x, x+h]$.

For $0 < t \leq h$, $\inf f|_{[x,x+t]} \leq \frac{F(x+t)-F(x)}{t} = \frac{1}{t}\displaystyle\int_{x}^{x+t} f(s)ds \leq \sup f|_{[x,x+t]}$

Since f is uniformly continuous $\lim\limits_{t\to 0} \inf f|_{[x,x+t]} = \lim\limits_{t\to 0} \sup f|_{[x,x+t]} = f(x)$.

So $\lim\limits_{t>0, t\to 0} \frac{F(x+t)-F(x)}{t} = f(x)$, similarly $\lim\limits_{t>0, t\to 0} \frac{F(x)-F(x-t)}{t} = f(x)$.

Hence $F'(x) = f(x)$. □

6.9.2 Corollary

If $G : J \to \mathbb{R}$ is a mapping s.t. $G' = f$, then $\forall (a,b) \in J \times J$, $a < b$, $G(b) - G(a) = \displaystyle\int_a^b f(t)dt$.

Application

(1) Let F and G be two mappings of class C^1 from J to $\mathbb{R}$. Then FG is of class C^1, and $(FG)' = F'G + FG'$. Let $f = F'$, $g = G'$.

Then $\forall (a,b) \in J \times J$, $a < b$

$\displaystyle\int_a^b f(t)G(t)dt = F(b)G(b) - F(a)G(a) - \int_a^b F(t)g(t)dt = \int_a^b (FG)'(t)dt - \int_a^b F(t)g(t)dt$

(2) Let $\varphi : I \to J$ be a mapping of class C^1, where I is an open interval. Let $F : J \to \mathbb{R}$ be a mapping of class C^1.

$(F\circ\varphi)'(x) = F'(\varphi(x))\varphi'(x)$. Hence, $\forall(\alpha, \beta) \in I\times I$, $\alpha < \beta$, $\displaystyle\int_\alpha^\beta F'(\varphi(x))\varphi'(x)dx = F(\varphi(\beta)) - F(\varphi(\alpha))$.

6.10 L^p space

6.10.1 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. $\|f\|_{L^p} := (\int_\Omega |f(x)|^p \mu(dx))^{\frac{1}{p}} < +\infty$

6.10.2 Lemma

Let $(p, q) \in \mathbb{R}^2_{\geq 1}$ s.t. $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

For any $(a, b) \in \mathbb{R}^2_{\geq 0}$, $\dfrac{a^p}{p} + \dfrac{b^q}{q} \geq ab$.

Proof

We may assume $(a, b) \in \mathbb{R}^2_{>0}$, $\dfrac{a^p}{p} + \dfrac{b^q}{q} = \dfrac{1}{p}\exp(p\ln(a)) + \dfrac{1}{q}\exp(q\ln(b)) \geq \exp(\ln(a) + \ln(b)) = ab$. $\square$

6.10.3 Theorem(Hölder inequality)

Let $(p, q) \in \mathbb{R}^2_{\geq 1}$ s.t. $\dfrac{1}{p} + \dfrac{1}{q} = 1$, let $f : \Omega \to \mathbb{R}$ and $g : \Omega \to \mathbb{R}$ be measurable mappings that belongs to $L^p(\Omega, \mathcal{G}, \mu)$ and $L^q(\Omega, \mathcal{G}, \mu)$, respectively. Then $\|fg\|_{L^1} \leq \|f\|_{L^p}\|g\|_{L^q}$, which means $fg \in L^1(\Omega, \mathcal{G}, \mu)$.

Proof

Take $\varphi = \dfrac{f}{\|f\|_{L^p}}$, $\psi = \dfrac{g}{\|g\|_{L^q}}$.

$|\varphi(x)\psi(x)| \leq \dfrac{|\varphi(x)|^p}{p} + \dfrac{|\psi(x)|^q}{q}$.

$\int_\Omega |\varphi(x)\psi(x)|\mu(dx) \leq \dfrac{\int_\Omega |f(x)|^p \mu(dx)}{p\|f\|^p_{L^p}} + \dfrac{\int_\Omega |g(x)|^q \mu(dx)}{q\|g\|^q_{L^q}} = \dfrac{1}{p} + \dfrac{1}{q} = 1$ $\square$

6.10.4 Corollary

Let $p \geq 1$.

$\forall (f, g) \in L^p(\Omega, \mathcal{G}, \mu)^2$, $\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^p}$

Proof

Apply Hölder inequality to $f(f+g)^{p-1}$ and $g(f+g)^{p-1}$. $\square$

Chapter 7

Compliment of linear algebra

7.1 Tensor product

Let R be a commutative unitary ring.

7.1.1 Theorem

Let M and N be two R-modules. There exists an R-module denote by $M \otimes_R N$ and a bilinearmap $t : M \times N \to M \otimes_R N$ having the following properties :

(1) For any R-module P and any bilinear map $s : M \times N \to P$ there exists a unique linear map $f_s : M \otimes_R N \to P$ s.t. $s = f_s \circ t$

(2) If T, t' is another couple that satisfies (1) with $s \mapsto g_s$ then there exists a unique isomorphism $T \cong M \otimes_R N$

Proof

Uniqueness :

$$\begin{tikzcd}
M \times N \arrow[rr, "t"] \arrow[dr, "t'"'] & & M \otimes_R N \arrow[dl, shift left, "f_{t'}"] \\
& T \arrow[ur, shift left, "g_t"] &
\end{tikzcd}$$

This commutative diagram shows that $f_{t'} \circ g_t \circ t' = f_{t'} \circ t = t'$

$$\begin{tikzcd}
M \times N \arrow[rr, "t"] \arrow[dr, "t"'] & & T \arrow[dl, "f_{t'} \circ g_t"] \\
& T &
\end{tikzcd}$$

Since $f_{t'} \circ g_t$ is unique, thus $f_{t'} \circ g_t = \mathrm{Id}_T$. Similarly, $g_t \circ f_{t'} = \mathrm{Id}_{M \otimes_R N}$.

Existence : Let $\mathcal{F}$ be the free R-module generated by $M \times N$.

Let $\mathcal{G}$ be the R-submodule generated by the elements of the following shape :

$(m + m', n) - (m, n) - (m'n)$

$(m, n + n') - (m, n) - (m, n')$

$(rm, n) - r(m, n)$

$(m, rn) - r(m, n)$

Define $M \otimes_R N := \mathcal{F}/\mathcal{G}$, $t : M \times N \to M \otimes_R N$, $(m, n) \mapsto \mathcal{G} + (m, n) =: t(m, n)$

$t(m + m', n) = \mathcal{G} + (m + m', n) = \mathcal{G} + (m, n) + (m', n) = (\mathcal{G} + (m, n)) + (\mathcal{G} + (m', n)) = t(m, n) + t(m', n)$. Hence t is bilinear.

We define $f_s(\mathcal{G} + (m, n)) := s(m, n)$. Extend this map by linarity. This makes this diagram commutative it is clear the unique map.

7.1.2 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 $t(m, n) =: m \otimes_R n$ and any element of this form is called pure tensor.

Remark

Pure tensors generate $M \otimes_R N$. In particular any tensor can be written as sum of pure tensors. But it is not a basis of the tensor, since it's not linearly independent.

7.1.3 Corollary

The map $s \mapsto f_s$ defined above gives an isomorphism $\mathrm{Hom}^{(2)}(M \times N, P) \cong \mathrm{Hom}(M \otimes_R N, P)$ for any $R-$module P.

Proof

Surjective: Take $\varphi \in \mathrm{Hom}(M \otimes_R N, P)$, the $\varphi \circ t$ is clearly bilinear.

Injective : If $0 \neq s = f_s \circ t$, then $f_s \neq 0$ hence injective.

Exercises(Use only the theorem 1)

(1) Show that $M \otimes_R N \cong N \otimes_R M$

(2) Show that $(M \otimes_R N) \otimes_R P \cong M \otimes_R (N \otimes_R P)$

(3) Show that $M_1 \otimes_R M_2 \otimes_R M_3 \otimes_R \cdots \otimes_R M_n$ factorizes the multilinear maps and $\mathrm{Hom}^{(n)}(M_1 \times \cdots \times M_n, P) \cong \mathrm{Hom}(M_1 \otimes_R \cdots \otimes_R M_n, P)$(Hint: Use induction)

Answer

(1) Define $t' : M \times N \to N \otimes_R M$, $(m, n) \mapsto n \otimes_R m$, it satisfies theorem (1), thus $N \otimes_R M \cong M \otimes_R N$. $\square$

(2) Fix $p \in P$. Define $\phi_p : M \times N \to M \otimes_R (N \otimes_R P)$, $(m, n) \mapsto m \otimes_R (n \otimes_R p)$, ϕ_p is a bilinear mapping.

$$\begin{tikzcd}
M \times N \arrow[rr] \arrow[dr, "\phi_p"'] & & M \otimes_R N \arrow[dl, "\phi'_p"] \\
& M \otimes_R (N \otimes_R P) &
\end{tikzcd}$$

Thus $\exists! \phi'_p : M \otimes_R N \to M \otimes_R (N \otimes_R P)$, $m \otimes_R n \mapsto m \otimes_R (n \otimes_R p)$. Define $\psi : (M \otimes_R N) \times P \to M \otimes_R (N \otimes_R P)$ s.t. $\psi(x, p) = \phi'_p(x)$, it is a bilinear mapping.

$$\begin{tikzcd}
(M \otimes_R N) \times P \arrow[rr] \arrow[dr, "\psi"'] & & (M \otimes_R N) \otimes_R P \arrow[dl, "\psi'"] \\
& M \otimes_R (N \otimes_R P) &
\end{tikzcd}$$

Hence $\exists! \psi' : (M \otimes_R N) \otimes_R P \to M \otimes_R (N \otimes_R P)$, $(m \otimes_R n) \otimes_R p \mapsto m \otimes_R (n \otimes_R p)$. Similarly, one can prove it is a unique isomorphism.

(3) Proof it by induction.

The case when $n = 1$ is trivial, the case when $n = 2$ has proved in corollary.

Suppose when $n = k$ it holds, then when $n = k+1$, let $f \in \mathrm{Hom}^{(k+1)}(M_1 \times \cdots \times M_{k+1}, P)$, fix $p \in M_{k+1}$, then $f(\cdot, p) : M_1 \times \cdots \times M_k \to P$ is a multi-linear mapping. Then $\exists! \phi_p : M_1 \otimes_R \cdots \otimes_R M_k \to P$ s.t. $f_p(m_1, \cdots, m_k) = \phi_p(m_1 \otimes_R \cdots \otimes_R m_k)$.

Define $g : (M_1 \otimes_R \cdots \otimes_R M_k) \times M_{k+1} \to P$ s.t. $g(\cdot, p) = \phi_p(\cdot)$, hence $\exists! \psi : (M_1 \otimes_R \cdots \otimes_R M_k) \otimes_R M_{k+1} = M_1 \otimes_R \cdots \otimes_R M_{k+1} \to P$ s.t. $\psi(m_1 \otimes_R \cdots \otimes_R m_{k+1}) = f(m_1, \cdots, m_{k+1})$. $\square$

Change R into a field K.

Remark

$\mathrm{Hom}^{(n)}(V_1\times\cdots\times V_n, K)\cong(V_1\otimes\cdots\otimes V_n)^\vee$, this is the previsous exercise for $P=K$.

Remark

$e_i\otimes e_j\neq e_j\otimes e_i$

7.1.4 Lemma

Let $V_1,\cdots,V_n$ be vector spaces of finite dimension $d_i>0$. Let $\{e_{i1},\cdots,e_{id_i}\}$ be a basis of V_i. Let's define the following functions.

$\varphi_{i_1,\cdots,i_n}: V_1\times\cdots\times V_n\to K$, $(v_1,\cdots,v_n)\mapsto\prod\limits_{i=1}^{n}e_{ii_1}^\vee(v_i)$

Then the set $\{\varphi_{i_1,\cdots,i_n}\}$ is a basis of $\mathrm{Hom}^{(n)}(V_1\times\cdots\times V_n, K)$

Proof

We do the proof for $n=2$, then the general case follows by induction.

$V_1=<e_1,\cdots,e_{d_1}>$, $V_2=<w_1,\cdots,w_{d_2}>$.

Let $\psi\in\mathrm{Hom}^{(2)}(V_1\times V_2,K)$, $(a,b)\in V_1\times V_2$ s.t. $a=\sum\limits_{i=1}^{d_1}a_ie_i$, $b=\sum\limits_{j=1}^{d_2}b_jw_j$.

Surjective: $\psi(a,b)=\psi(\sum\limits_{i=1}^{d_1}a_ie_i,\sum\limits_{j=1}^{d_2}b_jw_j)=\sum\limits_{i=1}^{d_1}\sum\limits_{j=1}^{d_2}a_ib_j\psi(e_i,w_j)=\sum\limits_{i=1}^{d_1}\sum\limits_{j=1}^{d_2}\psi(e_i,w_j)e_{1i}^\vee(a)e_{2j}^\vee(b)=\sum\limits_{i=1}^{d_1}\sum\limits_{j=1}^{d_2}\psi(e_i,w_j)\varphi_{i,j}(a,b)$

Injective: If $\psi\equiv 0$, then $\psi(e_i,w_j)=0$, $\forall(i,j)$. $\square$

7.1.5 Prop

Assume that $V_1,\cdots,V_n$ are vector spaces and V_i has basis given by $\{e_{i1},\cdots,e_{id_i}\}$, then $\mathscr{B}=\{e_{1i_1}\otimes\cdots\otimes e_{ni_n}, 1\le i_j\le d_j\}$ is a basis for $V_1\otimes\cdots\otimes V_n$. In particular, $V_1\otimes\cdots\otimes V_n$ has dimension $\prod\limits_{i=1}^{n}d_i$.

Proof

Again we assume $n=2$.

$V_1 =< e_1, \cdots, e_m >$, $V_2 =< w_1, \cdots, w_n >$.

We know $\mathrm{Hom}^{(2)}(V_1 \times V_2, K) \cong (V_1 \otimes V_2)^\vee$, $s \mapsto f_s$. Take $s = e_i^\vee(x) w_j^\vee(y)$, then we get $f_{ij}(v_1 \otimes v_2) = e_i^\vee(v_1) w_j^\vee(v_2)$, which forms a basis of $(V_1 \otimes V_2)^\vee$. It follows that $\{e_i \otimes w_j\}$ forms a basis of $V_1 \otimes V_2$. $\square$

7.1.6 Prop

Let $V_1, \cdots, V_n$ be vector spaces as above. Then $V_1^\vee \otimes \cdots \otimes V_n^\vee \cong (V_1 \otimes \cdots \otimes V_n)^\vee$.

Proof

Define $V_1^\vee \times \cdots \times V_n^\vee \to \mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, K) \cong (V_1 \otimes \cdots \otimes V_n)^\vee$, $(\varphi_1, \cdots, \varphi_n) \mapsto ((v_1, \cdots, v_n) \to \prod_{i=1}^n \varphi_i(v_i))$.

This mapping is multilinear. It descents by the property of tensor product to a map $F : V_1^\vee \otimes \cdots \otimes V_n^\vee \to \mathrm{Hom}^{(n)}(V_1 \times \cdots \times V_n, K) \cong (V_1 \otimes \cdots \otimes V_n)^\vee$, $\varphi_1 \otimes \cdots \otimes \varphi_n \mapsto ((v_1, \cdots, v_n) \mapsto \prod_{i=1}^n \varphi_i(v_i))$. These two spaces have the same dimension $\prod_{i=1}^n d_i$. It is enough to show that the map is surjective.

Let's do it for $n = 2$. $F(e_i^\vee \otimes w_j^\vee) = e_i^\vee w_j^\vee$ and $\{e_i^\vee w_j^\vee\}_{ij}$ forms a basis of $\mathrm{Hom}^{(n)}(V_1 \times V_2, K)$. $\square$

7.1.7 Prop

Let V and W be two vector spaces (finite dimension). Then $\mathrm{Hom}(V, W) \cong V^\vee \otimes W$.

Proof

$s : V^\vee \times W \to \mathrm{Hom}(V, W)$, $(\varphi, w) \mapsto (\sigma \mapsto \varphi(\sigma) w)$. It is bilinear, thus it induces $f_s : V^\vee \otimes W \to \mathrm{Hom}(V, W)$, we have to show that this is the required isomorphism.

Let $\{v_1^\vee, \cdots, v_m^\vee\}$ be a basis of $V^\vee$ and let $\{w_1, \cdots, w_n\}$ be a basis of W. Let's see what happens to $f_s(v_i^\vee \otimes w_j) = (v_k \mapsto v_i^\vee(v_k) w_j = \delta_{ik} w_j)$

Consider that matrix associated to f_s with respect to the fixed basis. Call this matrix M_{ji}.

$$M_{ji} = \begin{cases} 1 & \text{if } (a,b) = (j,i) \\ 0 & \text{elsewhere} \end{cases}$$

This matrix of this form are a basis of $\mathrm{Hom}^{(m)}(K^m, K^n) \cong \mathrm{Hom}(V, W)$.

□

Remark

An important case of the last proposition is when $W = V$, $V^\vee \otimes V \cong \mathrm{Hom}(V, V)$. $\sum v_i^\vee \otimes v_i \mapsto \mathrm{Id}_V \in \mathrm{Hom}(V, V)$.

Exercise

Let M, N, P be R-modules.

Show that $\mathrm{Hom}(M \otimes_R N, P) \cong \mathrm{Hom}(M, \mathrm{Hom}(N, P))$.

Answer

$\mathrm{Hom}(M \otimes_R N, P) \cong \mathrm{Hom}^{(2)}(M \times N, P) \cong \mathrm{Hom}(M, \mathrm{Hom}(N, P))$ □

Let $\varphi : R \to S$ be a ring homomorphism. Let M be a R-module.

Notice that S has a structure of R-module. $s \in S$, $r \in R$, $rs := \varphi(r)s$. Now take the tensor product $M \otimes_R S$.

Now we give a structure of $S-$module to $M \otimes_R S$. Take $s \in S$, $s(m \otimes_R s') := m \otimes_R ss'$. $M \otimes_R S$ is a $S-$module.

Note that we have a map $i : M \to M \otimes_R S$, $m \mapsto m \otimes 1$. Be careful, in general the map i isn't injective.

Example

$R = \mathbb{Z}$, $S = \mathbb{Z}/2\mathbb{Z}$, $\alpha : \mathbb{Z} \to \mathbb{Z}/2\mathbb{Z}$, $M = \mathbb{Z}[X]$.

$i(2X) = 2X \otimes 1 = 2(X \otimes 1) = X \otimes \alpha(2)1 = X \otimes 0 = 0$

7.1.8 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 $\{e_i \otimes_K 1\}$ is a $L-$basis of V_L.

Proof

The set $\{e_i \otimes_K 1\}$ generates V_L since $v \otimes_K l = (\sum \alpha_i e_i \otimes_K l) = \sum l\alpha_i(e_i \otimes_K 1)$. We have to show that the elements are linearly independent.

Suppose $0 = \sum \alpha_i(e_i \otimes_K 1) = \sum e_i \otimes_K \alpha_i$, $\alpha_i \in L$. Define the map $b_i : V \times L \to L$, $(\sum \lambda_i e_i, \beta) \mapsto \lambda_i \beta$. This map is bilinear. It induces a map $f_{b_i} : (\sum \lambda_i e_i \otimes_K \beta) \mapsto \lambda_i \beta$. Note that $f_i(e_j \otimes_K \beta) = \delta_{ij}\beta$.

$f_i(\sum_j e_j \otimes_K \alpha_j) = \alpha_i$. But $0 = f_i(0) = f_i(\sum_j e_j \otimes_K \alpha_j) = \alpha_i$, $\forall i$. $\square$

Remark

As a consequence we have that the map $i : V \to V_L$ (map of $K-$vector space) is injective.

Exercise

Show that $V \otimes_K K \cong V$, $v \otimes a \mapsto av$.

Answer

Since $s : V \times K \to V$, $(v, a) \mapsto av$ is bilinear, it induces a mosphism of $K-$module, $f_s : V \otimes_K K \to V$. Surjective is trivial. If $av = 0$, either $a = 0$ or $v = 0$, which means it is injective. $\square$

Remark

In fact, $M \otimes_R R \cong M$, $m \otimes_R r \mapsto rm$. It is injective since if $r_1 m_1 = r_2 m_2$, then $m_1 \otimes_R r_1 = r_1 m_1 \otimes_R 1 = r_2 m_2 \otimes_R 1 = m_2 \otimes_R r_2$

Exercise

Let I be an index set and M be a R-module, then $M \otimes_R R^{\oplus I} \cong M^{\oplus I}$.

Answer

$f : M \otimes_R R^{\oplus I} \to M^{\oplus I}$, $v \otimes (a_i)_{i \in I} \mapsto (a_i v)_{i \in I}$.

$g : M^{\oplus I} \to M \otimes_R R^{\oplus I}$, $(v_i)_{i \in I} \mapsto \sum_{i \in I} v_i \otimes_R e_i$. $\square$

7.2 Exact sequence

7.2.1 Def

Let M_1, M_2, N_1, N_2 be R-modules and let $f_i : M_i \to N_i$ be linear mappings. Then we define $f_1 \otimes f_2 : M_1 \otimes M_2 \to N_1 \otimes N_2$, $m_1 \otimes m_2 \mapsto f(m_1) \otimes f(m_2)$. This is a linear map.

$$\begin{tikzcd} M_1 \times M_2 \arrow[rr, "(m_1,m_2)\mapsto f(m_1)\otimes f(m_2)"] \arrow[dr] & & N_1 \otimes N_2 \\ & M_1 \otimes M_2 \arrow[ur, dashed] & \end{tikzcd}$$

Fix a R-module N, and consider $\_ \otimes N : M \mapsto M \otimes_R N$ for any R-module M. And for any linear map $(f : M \to P) \mapsto (f \otimes \mathrm{Id}_N : M \otimes_R N \to P \otimes_R N)$.

Notice that this association sends Id_M to $\mathrm{Id}_{M \otimes_R N}$ and moreover is well behaved with respect to the composition $f \circ g \mapsto (f \circ g) \otimes \mathrm{Id}_N = (f \otimes \mathrm{Id}_N) \circ (g \otimes \mathrm{Id}_N)$, $g : M \to P$, $f : P \to E$, above properties show that it forms a category.

7.2.2 Def

A sequence of R-modules (complex chain of $R-$modules) is a diagram of the following form.

$$M_1 \xrightarrow{d^1} M_2 \xrightarrow{d^2} M_3 \xrightarrow{d^3} M_4 \longrightarrow \cdots$$

M_i is a R-module, d^i is a linear map, $\mathrm{Ker}(d^{i+1}) \supseteq \mathrm{Im}(d^i)$, which means $d^{i+1} \circ d^i = 0$.

The sequence is exact when $\mathrm{Ker}(d^{i+1}) = \mathrm{Im}(d^i)$

Example 1

Take a morphism $f : M \to N$.

f is injective iff

$$0 \longrightarrow M \xrightarrow{f} N$$

is exact.

f is surjective iff

$$M \xrightarrow{f} N \longrightarrow 0$$

is exact.

The first theorem of homomorphism can be written as an exact sequence.

$$0 \longrightarrow \mathrm{Ker}(f) \xrightarrow{i} M \xrightarrow{f} \mathrm{Im}(f) \longrightarrow 0$$

More in general sequence like

$$0 \longrightarrow M_1 \longrightarrow M_2 \longrightarrow M_3 \longrightarrow 0$$

are called short exact sequences.

Example 2

Let N be an $R-$module, $\forall M,\ M \mapsto M \otimes_R N$

$f : M \to P$, f is R-linear, $f \otimes_R \mathrm{Id}_N : M \otimes_R N \to P \otimes_R N$

Assume that we have a short exact sequence of R-modules.

$$0 \longrightarrow M_1 \xrightarrow{f} M_2 \xrightarrow{g} M_3 \longrightarrow 0$$

Then apply $\_ \otimes_R N$ to it.

$$0 \longrightarrow M_1 \otimes_R N \xrightarrow{f\otimes \mathrm{Id}_N} M_2 \otimes_R N \xrightarrow{g\otimes \mathrm{Id}_N} M_3 \otimes_R N \longrightarrow 0$$

This is still a chain complex. But an important question is that : Is this chain complex also exact?

Example 3

$$0 \longrightarrow \mathbb{Z} \xrightarrow{\mu : x \mapsto 2x} \mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/2\mathbb{Z} \longrightarrow 0$$

This sequence is exact.

Now apply $\_ \otimes_{\mathbb{Z}} (\mathbb{Z}/2\mathbb{Z})$

$$0 \longrightarrow \mathbb{Z} \otimes (\mathbb{Z}/2\mathbb{Z}) \xrightarrow{\mu \otimes \mathrm{Id}} \mathbb{Z} \otimes (\mathbb{Z}/2\mathbb{Z}) \xrightarrow{\pi \otimes \mathrm{Id}} (\mathbb{Z}/2\mathbb{Z}) \otimes (\mathbb{Z}/2\mathbb{Z}) \longrightarrow 0$$

This sequence is not exact, since $\mu \otimes \mathrm{Id} = 0$.

Exercise(Important)

If $R = K$, then $\_ \otimes_K N$ (where N is a finite dimension vector space) is exact.(Hint : Use the basis)

Answer

$$\begin{tikzcd} 0 \arrow[r] & M_1 \arrow[r, "f"] & M_2 \arrow[r, "g"] \arrow[d, squiggly, "\_ \bigotimes N"] & M_3 \arrow[r] & 0 \\ 0 \arrow[r] & M_1 \otimes N \arrow[r, "f \otimes \mathrm{Id}_N"] & M_2 \otimes N \arrow[r, "g \otimes \mathrm{Id}_N"] & M_3 \otimes N \arrow[r] & 0 \end{tikzcd}$$

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 $\sum\limits_{i=1}^n a_i \otimes e_i$. $f \otimes \mathrm{Id}_N(\sum\limits_{i=1}^n a_i \otimes e_i) = \sum\limits_{i=1}^n f(a_i) \otimes e_i$. 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 $s_i : M_2 \times N \to M_2$, $(m, n) \mapsto e_i^\vee(n)m$, which is bilinear, thus it induces $f_{s_i} : M_2 \otimes N \to M_2$. $f(a_i) = f_{s_i}(\sum\limits_{i=1}^n f(a_i) \otimes e_i) = f_{s_i}(0) = 0$. Since f is injective, thus $a_i = 0$, $\forall i$. Therefore, $f \otimes \mathrm{Id}_N$ is injective.

$\mathrm{Im}(f \otimes \mathrm{Id}_N) \subseteq \mathrm{Ker}(g \otimes \mathrm{Id}_N)$ is trivial. $\forall u \otimes m \in \mathrm{Ker}(g \otimes \mathrm{Id}_N)$, similar to above $g(a_i) = 0$, $\forall i$. Thus there exists $\sum\limits_{i=1}^n v_i \otimes e_i \in M_1 \otimes N$, s.t. $f \otimes \mathrm{Id}_N(\sum\limits_{i=1}^n v_i \otimes e_i) = \sum\limits_{i=1}^n a_i \otimes e_i = u \otimes m$.

Surjective of $g \otimes \mathrm{Id}_N$ is trivial. □

7.2.3 Def

Let R be a commutative unitary ring and N be a R-module. We say that N is a flat R-module if apply $\_ \otimes_R N$ to any short exact sequence, the sequence is still exact.

7.2.4 Theorem

Any free R-module is flat.

Proof

See above exercise. □

Exercise(Harder)

Prove the right exactness of the tensor product functor.(To simplify this question, let R be a commutative unitary ring.)

Answer

$$\begin{tikzcd}
M_1 \arrow[r, "f"] & M_2 \arrow[r, "g"] \arrow[d, phantom, "_ \bigotimes N"] & M_3 \arrow[r] & 0 \\
M_1 \otimes N \arrow[r, "f\otimes \mathrm{Id}_N"'] & M_2 \otimes N \arrow[r, "g\otimes \mathrm{Id}_N"'] & M_3 \otimes N \arrow[r] & 0
\end{tikzcd}$$

Surjective of g is trivial.

To prove $\mathrm{Ker}(g \otimes \mathrm{Id}_N) = \mathrm{Im}(f \otimes \mathrm{Id}_N)$, one need to back to the construction of tensor product.

$$\begin{tikzcd}
R(M_2 \times N) \arrow[r, "\pi_1"] \arrow[d, dashed, "g'\otimes \mathrm{Id}_N"'] & M_2 \otimes N \arrow[d, "g\otimes \mathrm{Id}_N"] \\
R(M_3 \times N) \arrow[r, "\pi_2"'] & M_3 \otimes N
\end{tikzcd}$$

To make this diagram commutative, $g' \otimes \mathrm{Id}_N$ is defined that $r(u, v) \mapsto r(g(u), v)$.

Let $m \otimes n \in \mathrm{Ker}(g \otimes \mathrm{Id}_N)$. By definition $m \otimes n = [\sum_{i=1}^{N_1} (a_i, n_i)] =$

$\sum_{i=1}^{N_1}[(a_i, n_i)]$. Then $g \otimes \mathrm{Id}_N(m \otimes n) = \sum_{i=1}^{N_1}[(g(a_i), n_i)]$

Since it equal to zero, the internal structure of $g \otimes \mathrm{Id}_N(m \otimes n)$ must be of the form $[\sum_{j=1}^{N_2}((b_j, n_j + n_j') - (b_j, n_j) - (b_j, n_j'))]$.

But b_j is a image of an element inside M_2, since g is surjective, thus $g' \otimes \mathrm{Id}_N(\pi_2(\sum_{j=1}^{N_2}((a_j+k_1, n_j+n_j') - (a_j+k_2, n_j) - (a_j+k_3, n_j')))) = [\sum_{j=1}^{N_2}((b_j, n_j + n_j') - (b_j, n_j) - (b_j, n_j'))]$

Therefore, $\pi_1(\sum_{j=1}^{N_2}((a_j + k_1, n_j + n_j') - (a_j + k_2, n_j) - (a_j + k_3, n_j'))) = \sum_{j=1}^{N_2}((a_j \otimes (n_j + n_j') - a_j \otimes n_j - a_j \otimes n_j') + k_1 \otimes (n_j + n_j') - k_2 \otimes n_j - k_3 \otimes n_j') = \sum_{j=1}^{N_2}(k_1 \otimes (n_j + n_j') - k_2 \otimes n_j - k_3 \otimes n_j') \in \mathrm{Im}(f \otimes N)$. $\square$

7.3 Exterior product

Fix a vector space V over K of finite dimension.

7.3.1 Def

$T_p^q(V) := (V^\vee)^{\otimes p} \otimes_K V^{\otimes q}$, $p, q \in \mathbb{N}$

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.

Example

$T_0^0(V) := K$, $T_1^0(V) = V^\vee$, $T_0^1(V) = V$, $T_1^1(V) = V^\vee \otimes V \cong \mathrm{Hom}(V, V)$, $T_2^0(V) = V^\vee \otimes V^\vee \cong (V \otimes V)^\vee \cong \mathrm{Hom}^{(2)}(V \times V, K)$.

7.3.2 Def

$T(V) := \bigoplus\limits_{q=0}^{\infty} T_0^q(V)$.

In $T(V)$ we define the multiplication. $T(V) \times T(V) \to T(V)$, $(\sum\limits_{n\in\mathbb{N}} \alpha_n, \sum\limits_{n\in\mathbb{N}} \beta_n) \mapsto \sum\limits_{i\in\mathbb{N}} \sum\limits_{n+m=i} \alpha_n \otimes \beta_m$ and $T_0^l(V) \times T_0^q(V) \to T_0^{l+q}(V)$, $(\alpha_l, \beta_q) = (x_1 \otimes \cdots \otimes x_l, y_1 \otimes \cdots \otimes y_q) \mapsto x_1 \otimes \cdots \otimes x_l \otimes y_1 \otimes \cdots \otimes y_q := \alpha_l \otimes \beta_q$, extend it by linear. When $k \in T_0^0(V)$, we also define $k \otimes \alpha_n := k\alpha_n$

With the operation $T(V)$ becomes a $K-$algebra. $T(V)$ is the tensor algebra associated to V.

7.3.3 Def

Let W be the two sided ideal of $T(V)$ generated by the elements of the type $x \otimes x$. $W = \{ \sum\limits_{i(\text{finite})} (y_1 \otimes \cdots \otimes y_{m_i}) \otimes (x_i \otimes x_i) \otimes (z_1 \otimes \cdots \otimes z_{n_i}), x_j, y_j, z_j \in V, n_j, m_j \in \mathbb{N}\}$

The quotient algebra $\Lambda(V) := T(V)/W$ is called the exterior algebra of V. $\pi : T(V) \to \Lambda(V)$, $x_1 \otimes \cdots \otimes x_n \mapsto \pi(x_1 \otimes \cdots \otimes x_n) =: x_1 \wedge \cdots \wedge x_n$ extended by $K-$linear. Notice that $\Lambda(V)$ is a K-algebra.

$\Lambda^n(V) := T_0^n(V)/(W \bigcap T_0^n(V))$ (n-fold wedge product or n-fold exterior product), thus $\Lambda(V) = \bigoplus\limits_{n=0}^{\infty} \Lambda^n(V)$.

7.3.4 Prop

Let $\sigma \in \mathfrak{S}_n$, then $x_1 \wedge \cdots \wedge x_n = \operatorname{sgn}(\sigma) x_{\sigma(1)} \wedge \cdots \wedge x_{\sigma(n)}$.

Proof

Since any permutation can be written as product of adjactent transpositions. It is enough to do the proof for $\sigma = (i\ i+1)$.

$0 = (x_i + x_{i+1}) \wedge (x_i + x_{i+1}) = (x_i \wedge x_i) + (x_i \wedge x_{i+1}) + (x_{i+1} \wedge x_i) + (x_{i+1} \wedge x_{i+1}) = (x_i \wedge x_{i+1}) + (x_{i+1} \wedge x_i)$. $\square$

7.3.5 Corollary

If $\exists x_i = x_j$, $i < j$, then $x_1 \wedge \cdots \wedge x_i \wedge \cdots \wedge x_j \wedge \cdots \wedge x_n = 0$

7.3.6 Def (4.5 Symmetry of multilinear maps)

V is a vector space. A multilinear mapping.

$$V \times \cdots \times V \xrightarrow{\quad\varphi\quad} W$$

is called alternating symmetric(skew-symmetric) if $\varphi(x_1, \cdots, x_n) = 0$ when $\exists i \neq j$, st. $x_i = x_j$.

7.3.7 Prop

Fix a vector space V. For any alternating multilinear mapping $s : V \times \cdots \times V \to W$, where W is another vector space. There exists a unique linear mapping $g_s : \bigwedge^n(V) \to W$ s.t. the following diagram commutes.

$$\begin{tikzcd}
V \times \cdots \times V \arrow[rr, "s"] \arrow[d] & & W \\
T_0^n(V) \arrow[urr, "f_s"] \arrow[d] & & \\
\bigwedge^n(V) \arrow[uurr, "g_s"'] & &
\end{tikzcd}$$

Proof

$g_s(v_1 \wedge \cdots \wedge v_n) := s(v_1, \cdots, v_n)$ $\square$

Remark

The couple $\bigwedge^n(V)$ with $V \times \cdots \times V \to \bigwedge^n(V)$ that satisfies the above proposition is unique up to isomorphism.

7.3.8 Prop

Let V be a vector space of dimension n with a basis $\{e_1, \cdots, e_n\}$. Then $\bigwedge^k(V)$ is a vector space with a basis given by $\mathscr{B} = \{e_{i_1} \wedge \cdots \wedge e_{i_k}, 1 \leq i_i < \cdots < i_k \leq n\}$. In particular $\bigwedge^k(V)$ has dimension $\begin{pmatrix} n \\ k \end{pmatrix}$.

Proof

$\mathscr{B}$ is clearly a generating set. The difficult part is to show that $\mathscr{B}$ is made of linearly independent elements.

$I = \{i_1, \cdots, i_k\}$ with $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, define $\varphi_I : V \times \cdots \times V \to K$, $(e_{j_1}, \cdots, e_{j_k}) \mapsto \begin{cases} \mathrm{sgn}(\tau) & \text{if } \exists \tau \in \mathfrak{S}_I,\ \tau(j_m) = i_m \\ 0 & \text{otherwise} \end{cases}$

φ_I is multilinear and alternating, hence it induces a linear mapping. $g_{\varphi_I} = \overline{\varphi_I} : \bigwedge^k(V) \to K$. $\square$

7.3.9 Prop

Let R be a commutative unitary ring, E be a R-module, $\{e_i\}_{i=1}^n \in E^n$.

(1) Assume that $\forall \lambda \in R \setminus \{0\}$, $\lambda e_1 \wedge \cdots \wedge e_n \neq 0$, then $\{e_i\}_{i=1}^n$ is R-linearly independent. ($R^n \to E$, $(a_1, \cdots, a_n) \mapsto a_1 e_1 + \cdots + a_n e_n$ is injective)

(2) Assume that $\{e_i\}_{i=1}^n$ is a system of generators of E over R. $\forall d \in \mathbb{N}$, $\{e_{i_1} \wedge \cdots \wedge e_{i_d}\}_{1 \leq i_1 < \cdots < i_d \leq n}$ is a system of generators of $\bigwedge^d E$ over R.(When $d > n$, $\bigwedge^d E = 0$).

(3) Assume that E is a free R-module and $\{e_i\}_{i=1}^n$ is a basis of E. $\forall d \in \mathbb{N}$, $\{e_{i_1} \wedge \cdots \wedge e_{i_d}\}_{1 \leq i_1 < \cdots < i_d \leq n}$ forms a basis of $\bigwedge^d E$.

7.4 Determinant

7.4.1 Def

Let V be a n dimensional vector space, then $\det(V) := \bigwedge^n(V)$ is called the determinant of V. It is a vector space of dimension $\binom{n}{n} = 1$ and a basis of $\det(V)$ is given by $e_1 \wedge \cdots \wedge e_n$ where $\{e_1, \cdots, e_n\}$ is a basis of V.

Let W also be a n dimensional vector space, let $(v_i)_{i=1}^n$ and $(w_j)_{j=1}^n$ be a basis of V and W respectively and let $f \in \mathrm{Hom}(V, W)$ then consider $f' : V \times \cdots \times V \to \bigwedge^k(W)$, $(a_1, \cdots, a_k) \mapsto f(a_1) \wedge \cdots \wedge f(a_k)$ this mapping is multilinear and alternating thus it induces a mapping $g_{f'} : \bigwedge^k(V) \to \bigwedge^k(W)$. When $k = n$ and $\dim(V) = \dim(W) = n$, $\bigwedge^n f : \det(V) \to \det(W)$ is called the determinant of f in the fixed bases $\{v_i\}_{i=1}^n$ and $\{w_j\}_{j=1}^n$ and the notation of $\bigwedge^n f$ is $\det(f)$.

Since $\det(f)$ is a linear mapping betweent two one dimensional vector spaces, the influence of $\det(f)$ can be seen as a scalar. Then define $\det_f$ as the scalar put on the basis of $\bigwedge^n(V)$, $\det(f)(v_1 \wedge \cdots \wedge v_n) =: \det_f w_1 \wedge \cdots \wedge w_n$. By abuse of notation we identify $\det(f) = \det_f$. When $f \in \mathrm{Hom}(V, V)$, by convention, we use the same basis in V, and then the value of $\det(f)$ is independent with the same choice of basis, so we can not mension the basis when f is an endomorphism.

7.4.2 Prop

$\det(f \circ g) = \det(f) \det(g)$

Proof

$\det(f \circ g)(e_1 \wedge \cdots \wedge e_n) = (f \circ g)(e_1) \wedge \cdots \wedge (f \circ g)(e_n) = f(g(e_1)) \wedge \cdots \wedge f(g(e_n)) = (\det(f))(g(e_1) \wedge \cdots \wedge g(e_n)) = \det(f) \det(g)(e_1 \wedge \cdots \wedge e_n)$. □

7.4.3 Prop

Let V and W be two vector spaces with the same dimension. $f \in \mathrm{Hom}(V, W)$ is invertible iff $\det(f) \neq 0$.

Proof

(**CHEN Huayi**) Suppose f is invertible. $1 = \det(\mathrm{Id}_V) = \det(f^{-1} \circ f) =$

$\det(f^{-1})\det(f)$

Suppose f isn't invertible, which means $\dim(f(V)) < n$, thus $\bigwedge^n(f(V)) = \{0\}$, so $f(e_1) \wedge \cdots \wedge f(e_n) \equiv 0$ $\square$

(**Paolo**)f is not invertible iff $\{f(e_1), \cdots, f(e_n)\}$ is not a basis iff there is a non-trivial linear combination $\sum\limits_i \lambda_i f(e_i) = 0$ not loss generality $f(e_1) = \sum\limits_{i\geq 2} \mu_i f(e_i)$. $\det(f)(e_1 \wedge \cdots \wedge e_n) = (\sum\limits_{i\geq 2} \mu_i f(e_i)) \wedge f(e_2) \wedge \cdots \wedge f(e_n) = 0$ iff $\det(f) = 0$ $\square$

7.4.4 Prop

Let V and W be two vector spaces with the same dimension, let $f \in \mathrm{Hom}(V, W)$. The determinant of f is equal to the determinant of any matrix that represents f with respect fixed bases. This doesn't depend on the choice of basis.

Proof

Fix a basis $\{v_1, \cdots, v_n\}$ of V.

$$\begin{tikzcd}
V \arrow[r, "f"] & W \\
K^n \arrow[u, "b"] \arrow[r, "\mathbf{A}"] & K^n \arrow[u, "b"']
\end{tikzcd}$$

$$\begin{tikzcd}
\det(V) \arrow[r, "\det(f)"] & \det(W) \\
\det(K^n) \arrow[u, "\bigwedge^n b"] \arrow[r, "\det(\mathbf{A})"] & \det(K^n) \arrow[u, "\bigwedge^n b"']
\end{tikzcd}$$

$\det(\mathbf{A}) = \det(b^{-1} \circ f \circ b) = \det(b)^{-1}\det(f)\det(b) = \det(f)$ $\square$

Let $\{e_1, \cdots, e_n\}$ be the standard basis(column) of K^n.

$$K^n \xrightarrow{\mathbf{A}} K^n$$

$\mathbf{A}$ is a matrix $n \times n$, $\mathbf{A} = (\mathbf{A}e_1, \cdots, \mathbf{A}e_n)$

(1)

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

This is called a upper triangular matrix.

$\mathbf{A}e_1 = a_{11}e_1$, $\mathbf{A}e_2 = a_{12}e_1 + a_{22}e_2$, $\mathbf{A}e_i = a_{1i}e_1 + \cdots + a_{ii}e_i$

$\det(\mathbf{A})(e_1 \wedge \cdots \wedge e_n) = \mathbf{A}e_1 \wedge \cdots \wedge \mathbf{A}e_n = \prod_{i=1}^{n} a_{ii}(e_1 \wedge \cdots \wedge e_n)$. For upper triangular matrices the determinant is the product of the elements on the diagonal.

(2) $\mathbf{A} = (a_{ij})_{n \times n}$. If one column of $\mathbf{A}$ can be expressed as a linear combination of other columns of $\mathbf{A}$, then $\det(\mathbf{A}) = 0$.

(3) $\det(\boldsymbol{P_\sigma})(e_1 \wedge \cdots \wedge e_n) = e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)} = \operatorname{sgn}(\sigma)(e_1 \wedge \cdots \wedge e_n)$.

(4) $\det(\boldsymbol{S_{i,c}})(e_1 \wedge \cdots \wedge e_n) = (e_1 + c_1 e_i) \wedge \cdots \wedge e_i \wedge \cdots \wedge (e_n + c_n e_i) = e_1 \wedge \cdots \wedge e_n$

7.4.5 Prop

Let (a_{ij}) be a matrix of dimension $n \times n$. Then $\det(\mathbf{A}) = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i)i}$

Proof

$\det(\mathbf{A})(e_1 \wedge \cdots \wedge e_n) = (\sum_i a_{i1}e_i) \wedge (\sum_i a_{i2}e_i) \wedge \cdots \wedge (\sum_i a_{in}e_i) = \sum_{\sigma \in \mathfrak{S}_n} \prod_i a_{\sigma(i)i}(e_{\sigma(1)} \wedge \cdots \wedge e_{\sigma(n)}) = \sum_i \operatorname{sgn}(\sigma) \prod_i a_{\sigma(i)i}(e_1 \wedge \cdots \wedge e_n)$ □

7.4.6 Prop

$\det(\mathbf{A}) = \det(\boldsymbol{A^\tau})$

Proof

$\boldsymbol{A}^{\tau} = (\alpha_{ij})$, $\mathbf{A} = (a_{ij})$, $\forall i, j$, $a_{ij} = \alpha_{ji}$

$\det(\boldsymbol{A}^{\tau}) = \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{i} \alpha_{\sigma(i)i} = \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{i} a_{i\sigma(i)} = \sum_{\sigma} \operatorname{sgn}(\sigma^{-1}) \prod_{i} a_{\sigma^{-1}(\sigma(i))\sigma(i)} = \sum_{\sigma} \operatorname{sgn}(\sigma^{-1}) \prod_{i} a_{\sigma^{-1}(i)i} = \det(\mathbf{A})$ □

Reminder

$f : V \to W$, $f^{\vee} : W^{\vee} \to V^{\vee}$, $\varphi \mapsto \varphi \circ f$

If you fix some bases on V and W, then $\mathbf{A}$ is the matrix asssociated to f. Take the dual bases of $W^{\vee}$ and $V^{\vee}$, then the matrix associated to $f^{\vee}$ is $\mathbf{A}^{\tau}$.

7.4.7 Corollary

Let V and W be two vector spaces with the same dimension. $f \in \operatorname{Hom}(V, W)$, $\det(f) = \det(f^{\vee})$

Remark

Fix $\mathbf{A}$ of dimension $n \times n$, apply Gauss Reduction, $\mathbf{B}$ is upper triangular.

By the property listed above $|\det(\mathbf{A})| = |\det(\mathbf{B})|$.

Notation

Fix $\mathbf{A} = (a_{ij})$, denote with $\boldsymbol{A}_{[ij]}$ the $(n-1) \times (n-1)$ matrix obtained after removing the $i-$th row and $j-$th column from $\mathbf{A}$.

7.4.8 Prop

Let $\mathbf{A} = (a_{ij})$ then $\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\boldsymbol{A}_{[ij]})$, $\forall i \in \{1, \cdots, n\}$

Proof

(**Paolo**)

$$\begin{tikzcd} K^n \arrow[r, "\mathbf{A}"] & K^n \arrow[d, "p_i"] \\ K^{n-1} \arrow[u, "\tau_j"] \arrow[r, "\boldsymbol{A}_{[ij]}"'] & K^{n-1} \end{tikzcd}$$

$\{e'_1, \cdots, e'_n\}$ is a standard basis of K^n. $\{e_1, \cdots, e_{n-1}\}$ is a standard basis of K^{n-1}.

$$\tau_j(e_i) = \begin{cases} e'_i & \text{if } i < j \\ e'_{i+1} & \text{if } i \geq j \end{cases}$$

$p_i =$"map that forgets about the i-th coordinates", $(x_1, \cdots, x_i, \cdots, x_n) \mapsto (x-1, \cdots, \hat{x_i}, \cdots, x_n) := (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$

You can check that the above diagram is commutative.

$$\begin{tikzcd} \wedge^{n-1}K^n \arrow[r, "\wedge^{n-1}\mathbf{A}"] & \wedge^{n-1}K^n \arrow[d, "\wedge^{n-1}p_i"] \\ \det(K^{n-1}) \arrow[u, "\wedge^{n-1}\tau_j"] \arrow[r, "\det(\boldsymbol{A}_{[ij]})"'] & \det(K^{n-1}) \end{tikzcd}$$

$\det(\mathbf{A})(e'_1 \wedge \cdots \wedge e'_n) = (-1)^{i-1} \det(\mathbf{A})(e'_i \wedge e'_1 \wedge \cdots \wedge \hat{e_i} \wedge \cdots \wedge e'_n) = (-1)^{i-1}\mathbf{A}e'_i \wedge \mathbf{A}e_1 \wedge \cdots \wedge \mathbf{A}\hat{e'_i} \wedge \cdots \wedge \mathbf{A}e'_n = (-1)^{i-1}\mathbf{A}e'_i \wedge \bigwedge^{n-1} \mathbf{A}(e'_1 \wedge \cdots \wedge \hat{e_i} \wedge \cdots \wedge e'_n)(*)$

Let $\pi_j : K^n \to K^n$, $(x_1, \cdots, x_n) \mapsto (0, \cdots, 0, x_j, 0, \cdots, 0)$, then $\mathbf{A} = \sum\limits_j (\pi_j \circ \mathbf{A})$, it means that $(*) = (-1)^{i-1}\mathbf{A}e_i \wedge \sum\limits_j \bigwedge^{n-1}(\pi_j \circ \mathbf{A})(e'_1 \wedge \cdots \wedge \hat{e'_i} \wedge \cdots \wedge e'_n) = (-1)^{i-1}\mathbf{A}e'_i \wedge \sum\limits_j \bigwedge^{n-1}(\pi_j \circ \mathbf{A} \circ \pi_i)(e_1 \wedge \cdots \wedge e_{n-1}) = \sum\limits_{k,j}((-1)^{i-1}a_{ki}e'_k \wedge \bigwedge^{n-1}(\pi_j \circ \mathbf{A} \circ \tau_i)(e_1 \wedge \cdots \wedge e_{n-1}))$

But $\pi_j(*)$ is always collinear of e_j, so when $k = j$, the element in the sum is 0. We can remove the terms $k = j$.

$\rho_k := \tau \circ p_k : K^n \to K^n$, $(x_1, \cdots, x_n) \mapsto (x_1, \cdots, 0, \cdots, x_n)$, $\pi_k = \mathrm{Id}_{K^n} - \rho_k$ and $\sum\limits_{j \neq k} \pi_j = \rho_k$

$(**) = \sum\limits_k (-1)^{i-1}a_{ki}e'_k \wedge \bigwedge^{n-1} \tau_k \circ \bigwedge(\rho_k \circ \mathbf{A} \circ \tau_i)(e_1 \wedge \cdots \wedge e_{n-1})$

But, by the diagram $\wedge^{n-1}(p_k \circ \mathbf{A} \circ \tau) = \det(\boldsymbol{A}_{[\boldsymbol{ki}]})$, $\wedge^{n-1}\tau_k(e_1 \wedge \cdots \wedge e_{n-1}) = e'_1 \wedge \cdots \wedge \hat{e_k} \wedge \cdots \wedge e'_n$

$(**) = \sum\limits_k (-1)^{i-1}a_{ki} \det(\boldsymbol{A}_{[\boldsymbol{ki}]})(e'_k \wedge e'_1 \wedge \cdots \wedge \hat{e_k} \wedge \cdots \wedge e'_n) = \sum\limits_k (-1)^{i+k}a_{ki} \det(\mathbf{A}_{[ki]})e'_1 \wedge \cdots \wedge e'_n$ $\square$

(Own) $\det(\boldsymbol{A}) = \sum\limits_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) \prod\limits_{i=1}^n a_{\sigma(i)i} = \sum\limits_{j=1}^n a_{ij} \sum\limits_{\sigma \in \mathfrak{S}_n, \sigma(j)=i} \mathrm{sgn}(\sigma) \prod\limits_{k=1, k\neq j}^n a_{\sigma(k)k} = \sum\limits_{j=1}^n (-1)^{i+j}a_{ij} \sum\limits_{\tau \in \mathfrak{S}_{n-1}} \mathrm{sgn}(\tau) \prod\limits_{k=1}^{j-1} a_{\tau(k)k} \prod\limits_{k=j+1}^n a_{\tau(k-1)k} = \sum\limits_{j=1}^n (-1)^{i+j}a_{ij}\boldsymbol{A}_{[\boldsymbol{ij}]}$ $\square$

Exercise

Let R be a commutative unitary ring, let F be a free $K-$module with finite type. Prove $\bigwedge^r(F^\vee) \cong \bigwedge^r(F)^\vee$, $r \in \mathbb{N}_{\geq 1}$.

Answer

Define $\delta : \underbrace{F \times \cdots \times F}_{r \text{ copies}} \times \underbrace{F^\vee \times \cdots \times F^\vee}_{r \text{ copies}} \to R$, $(v_1, \cdots, v_r, w_1, \cdots, w_r) \mapsto \det(w_i(v_j))$.

It is easy to check δ is multilinear and alternating in the first r and in the last r entries thus δ induces a bilinear mapping $\delta' : \bigwedge^r(F) \times \bigwedge^r(F^\vee) \to R$, $(v_1 \wedge \cdots \wedge v_r, w_1 \wedge \cdots \wedge w_r) \mapsto \det(w_i(v_j))$.

Define the isomorphism mapping as $\phi : \bigwedge^r(F^\vee) \to (\bigwedge^r(F))^\vee$, $\eta \mapsto \delta'(\cdot, \eta)$. Since they have the same dimension, only needs to check it is injective.

Let $(e_i)_{i=1}^n$ be a basis of F, then $(e_i^\vee)_{i=1}^n$ forms a basis of $F^\vee$. The case when $r > n$ is trivial. Suppose $\delta'(\cdot, \eta) = 0$, then $\sum\limits_{1 \leq k_1 < \cdots < k_r \leq n} c_{k_1, \cdots, k_r} \delta'(\cdot, e_{k_1}^\vee \wedge \cdots \wedge e_{k_r}^\vee) = 0$. Take $\xi = e_{k_1} \wedge \cdots \wedge e_{k_r}$, thus $0 = \phi(\xi, \eta) = \det(e_{k_i}^\vee(e_{k_j})) = c_{k_1, \cdots, k_n}$. $\square$

7.5 Eigenvalues and eigenvectors

7.5.1 Theorem

Let $f : V \to W$ be a linear mapping bewteen vector spaces of the finite dimension. Then

(1) There exists decompositions $V = V_0 \oplus V_1$, $W = W_1 \oplus W_2$ s.t. $V_0 = \mathrm{Ker}(f)$ and f induces an isomorphism between V_1 and W_1.

(2) There exist bases in V and W s.t. the associated matrix $\boldsymbol{A_f} = (a_{ij})$ satisfies $a_{ii} = 1$ for $1 \leq i \leq \tau$, $(\tau \leq n)$ and $a_{ij} = 0$ elsewhere. The number τ is unique.

(3) Let $\mathbf{A}$ be a $m \times n$ matrix. Then there exist two square matrices (with $\det \neq 0$) $\mathbf{B}$ and $\mathbf{C}$ of dimension $m \times m$ and $n \times n$ respectively and a number $\tau \leq \min\{m, n\}$ s.t. $\mathbf{BAC}$ has the form describd in 2. The number τ is unique and $\tau = \mathrm{rk}(\mathbf{A})$

Proof

(1) $V_0 = \mathrm{Ker}(f)$. Let V_1 be the direct complement of V_0. We put $W_1 = \mathrm{Im}(f)$ and W_2 is the direct complement of W_1. Now consider the restriction $f_{|V_1}$

This restriction is injective $V_1 \bigcap \mathrm{Ker}(f) = \{0\}$. Moreover $f_{|V_1}$ is surjective, in fact $w = f(\sigma) \in W_1$, then $f(\sigma) = f(v_0 + v_1) = f(v_0) + f(v_1) = f(v_1)$. $\square$

(2) Put $\tau = \dim V_1 = \dim W_1$. Take a basis of V given by $\{v_1, \cdots, v_\tau, v_{\tau+1}, \cdots, v_n\}$ where $\{v_1, \cdots, v_\tau\}$ is a basis of V_1.

The vectors $f(v_i)$, $1 \leq i \leq \tau$ are a basis of W_1. Extend such basis to a basis $\{w_1, \cdots, w_m\}$ of W.

Then we have the following situation.

$$\begin{tikzcd}
V \arrow[r, "f"] & W \\
K^n \arrow[u, "b"] \arrow[r, "\boldsymbol{A_f}"] & K^m \arrow[u, "b'"']
\end{tikzcd}$$

$b(e_i) = v_i$, $b'(e'_i) = w_i$

$$\boldsymbol{A_f}(e_i) = (b^{-1}) \circ f \circ b(e_i) = (b')^{-1} \circ f(v_i) = \begin{cases} e_i & \text{if } 1 \le i \le \tau \\ 0 & \text{otherwise} \end{cases}$$

This shows that $\boldsymbol{A_f}$ has the required form. In particular $\tau = \dim(\mathrm{Im}(f))$

$\square$

(3) The same as 2 with $\mathbf{A} = f$, $\mathrm{rk}(\mathbf{A}) = \dim(\mathrm{Im}(\mathbf{A}))$. $\mathbf{A} : K^n \to K^m$. $\square$

Remark

Let $f : V \to W$ be an isomorphism of vector spaces I can always find two bases in V and W s.t. $\boldsymbol{A_f} = \mathrm{Id}_n$. It follows from the theorem above.

7.5.2 Def

Let $f : V \to V$ be a linear mapping. A subspace $V_0 \subseteq V$ is said to be an invariant subspace of f if $f(V_0) \subseteq V_0$.

7.5.3 Def

A linear mapping $f : V \to V$ is diagonalizable if one of the following equivalent conditions are satisfied

(1) V dicomposed as a direct sum of one dimensional invariant subspaces of f.

(2) There exists a basis of V in which the matrix $\boldsymbol{A_f}$ is diagonal.

Let's show the equivalence between (1) and (2).

From (2) to (1) : Assume that in the base $\{v_1, \cdots, v_n\}$.

$$\boldsymbol{A_f} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

$f(v_i) = b \circ \boldsymbol{A_f}(e_i) = \lambda_i v_i \subseteq < v_i >$

So $V =< v_1 > \oplus \cdots \oplus < v_n >$.

From (1) to (2) : Assume $V =< v_1 > \oplus \cdots \oplus < v_n >$. Then $\{v_1, \cdots, v_n\}$

forms a basis of V. Consider the previous diagram.

$\mathbf{A}(e_i) = b^{-1} \circ f \circ b(e_i) = b^{-1}(f(v_i)) = b^{-1}(\lambda_i v_i) = \lambda_i e_i$

Example

$\mathbf{A} : \mathbb{R}^2 \to \mathbb{R}^2$, $\mathbf{A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $0 \leq \theta \leq \frac{\pi}{2}$

$\mathbf{A}$ isn't diagonalizable.

7.5.4 Def

Let L be a one dimensional invariant subspace for $f : V \to V$, then $f_{|L}$ is a multiplication by a scalar $\lambda \in K$. Such λ is called an eigenvalue of f.

A non-zero vector $v \in V$ is said to be an eigenvector of V, if $< v >$ is an invariant subspace of f.

Remark

If v is an eigenvector, then μv is an eigenvector for $\mu \in K^\times$.

Let V be a vector space over K. $\dim(V) = n$, $f \in \mathrm{End}(V)$. Let $\boldsymbol{A_f}$ be an associated matrix(in any basis).

The mapping $P(t) : K \to K$, $t \mapsto \det(t\boldsymbol{I_n} - \boldsymbol{A_f})$. This is a polynomial in $K[t]$.

7.5.5 Lemma

$P(t)$ is a monic polynomial of degree n.

Proof

$P(t) = \det(t\boldsymbol{I_n} - \boldsymbol{A_f}) = \sum\limits_{\sigma} \mathrm{sgn}(\sigma) \prod\limits_{i=1}^{n} (t\delta_{\sigma(i)i} - \boldsymbol{A_{\sigma(i)i}})$.
The only term giving t^n is when $\sigma = \mathrm{Id}$. □

7.5.6 Theorem

Use the notations introduced before.

(1) $P(t)$ doesn't depend on $\boldsymbol{A_f}$(If you change bases $P(t)$ doesn't change)

(2) Any eigenvalue of f is a root of $P(t)$. Conversely any $K-$root of $P(t)$ is an eigenvalue.

Proof

(1) Let $\boldsymbol{A'}$ be another matrix representation of f. Then $\boldsymbol{A'} = \boldsymbol{B^{-1}AB}$ where $\boldsymbol{B}$ is an invertible $n \times n$ matrix.

$\det(t\boldsymbol{I_n} - \boldsymbol{A'}) = \det(t\boldsymbol{I_n} - \boldsymbol{A})$ □

(2) Let $\lambda \in K$ be a k-root of $P(t)$, then $\det(\lambda\boldsymbol{I}_n - \boldsymbol{A_f}) = 0$. $\lambda\boldsymbol{I_n} - \boldsymbol{A_f}$ is not invertible, $\exists v \neq 0 \in \ker(\lambda\boldsymbol{I_n} - \boldsymbol{A_f})$, $\lambda v = \boldsymbol{A_f}v$ □

7.5.7 Corollary

Let $f \in \mathrm{End}(V)$ be diagonalizable, $\dim(V) = n$ and $\boldsymbol{A_f}$ is a diagonal matrix that presents f. Then $\boldsymbol{A_f}$ is unique up to permutation of the elements in the diagonal.

7.5.8 Def

The polynomial $P(t)$ would be denoted by $P_f(t)$. It is called the characteristic polynomial of f.

7.5.9 Corollary

If $P_f(t)$ has n different roots in K, then f is diagnolizable.

7.5.10 Def

A matrix of the form $\boldsymbol{J_r(\lambda)} = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ 0 & 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & \cdots & \lambda \end{pmatrix} \in M_{r\times r}(K)$, $r \geq 1$

is called a Jordan block.

A Jordan matrix is a matrix of the type $\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J_{r_1}(\lambda_1)} & 0 & \cdots & 0 \\ 0 & \boldsymbol{J_{r_2}(\lambda_2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{J_{r_k}(\lambda_k)} \end{pmatrix}$

Example

$V_n(\lambda) := \{F(x)|F(x) = e^{\lambda x}f(x)$ where $\lambda \in \mathbb{C}, f \in \mathbb{C}[x]_{\leq n-1}\}$. It is a n dimension vector space over $\mathbb{C}$.

Notice that $\frac{\mathrm{d}}{\mathrm{d}x} \in \mathrm{End}(V_n(\lambda))$. Consider $v_{i+1} = \frac{x^i}{i!}e^{\lambda x}$, notice that $\{v_0, \cdots, v_{n-1}\}$ form a basis of $V_n(\lambda)$.

$\frac{\mathrm{d}}{\mathrm{d}x}(v_{i+1}) = \lambda v_{i+1} + v_i$, $\boldsymbol{A}_{\frac{\mathrm{d}}{\mathrm{d}x}} = \boldsymbol{J_n(\lambda)}$.

7.6 Cayley-Hamilton theorem

7.6.1 Def

Let $a_0 + a_1 t + \cdots + a_n t^n = Q(t) \in K[t]$, then for $f \in \mathrm{End}(V)$, we define $Q(f) := a_0 \mathrm{Id}_V + a_1 f + \cdots + a_n f^{\circ n}$. From now we define $f^{\circ n} =: f^n$.

We say that Q is a polynomial annihilates f if $Q(f) = 0$.

7.6.2 Prop

Let $f \in \mathrm{End}(V)$. There exists a polynomial $Q \in K[t] \setminus \{0\}$ that annihilates f.

Proof

$\dim(\mathrm{End}(V)) = n^2$. Hence $\mathrm{Id}_V,\ f, \cdots, f^{n^2} \in \mathrm{End}(V)$ are linearly dependent. There exists a non-trivial linear combination s.t. $\lambda_0 \mathrm{Id} + \lambda_1 f + \cdots + \lambda_{n^2} f^{n^2} = 0$. So take $Q(t) = \lambda_0 + \lambda_1 t + \cdots + \lambda_{n^2} t^{n^2}$. This shows that $Q \neq 0$, and $Q(f) = 0$. □

7.6.3 Def

Let $m(t) \in K[t] \setminus \{0\}$ be a monic polynomial of minimal degree that annihilates $f \in \mathrm{End}(V)$. Then $m(t)$ is called the minimal polynomial of f.

7.6.4 Prop

If $m(t)$ is a minimal polynomial of f, then it is unique.

Proof

Suppose $m_1(t)$ is another minimal polynomial ($\deg(m_1) = \deg(m)$). Then $m - m_1 \in K[t]$, $(m - m_1)(f) = 0$. Now m and m_1 are both monic, so $\deg(m - m_1) < \deg(m) = \deg(m_1)$. This means $m - m_1 = 0 \in K[t]$. □

Notation

We denote the minimal polynomial of f with m_f.

7.6.5 Prop

Let $Q \in K[t] \setminus \{0\}$ be a polynomial that annihilates f. Then $m_f | Q$.

Proof

By the Euclidean division between polynomials. $Q(t) = m_f(t)s(t) + \tau(t)$ s.t. $\deg(\tau) < \deg(m_f)$. So $\tau(t) = 0$. $\square$

7.6.6 Def

Let $\boldsymbol{A}$ be a matrix of dimension $n \times n$, $M_{ij} = (-1)^{i+j} \det(\boldsymbol{A_{[i,j]}}) \in K$. $\det(\boldsymbol{A_{[i,j]}})$ is called the (i,j)-minor of $\boldsymbol{A}$. Then we define $\mathrm{Adj}(\boldsymbol{A})$ the adjugate matrix of $\boldsymbol{A} = (M_{ij})^{\tau}_{i,j}$.

7.6.7 Lemma

$\boldsymbol{A}\mathrm{Adj}(\boldsymbol{A}) = \mathrm{Adj}(\boldsymbol{A})\boldsymbol{A} = \det(\boldsymbol{A})\boldsymbol{I}_n$

Proof

One needs to prove that when $i \neq j$, $\sum\limits_{k=1}^{n} a_{ik}M_{jk} = 0$

Suppose that $\forall 1 \leq k \leq n$, $a_{ik} = a_{jk}$, then $\sum\limits_{k=1}^{n} a_{ik}M_{jk} = \sum\limits_{k=1}^{n} a_{jk}M_{jk} = \det(\boldsymbol{A}) = 0$ $\square$

7.6.8 Theorem (Cayley-Hamilton)

The characteristic polynomial P_f annihilates f.

Proof

Let $\boldsymbol{A} = \boldsymbol{A_f}$ any matrix that represents f. Consider $\boldsymbol{B} = \mathrm{Adj}(t\boldsymbol{I_n} - \boldsymbol{A})$, $\boldsymbol{B}$ is a matrix with coefficients in $K[t]$. Then $(t\boldsymbol{I_n} - \boldsymbol{A})\boldsymbol{B} = \det(t\boldsymbol{I_n} - \boldsymbol{A})\boldsymbol{I_n} = P_f\boldsymbol{I_n}$

We can decompose $\boldsymbol{B}$ in the following way. $\boldsymbol{B} = \sum\limits_{i=0}^{n-1} t^i\boldsymbol{B}_i$, $\boldsymbol{B}_i \in M_{n\times n}(K)$.

$P_f(t)\boldsymbol{I_n} = (t\boldsymbol{I_n} - \boldsymbol{A})\boldsymbol{B} = (t\boldsymbol{I_n} - \boldsymbol{A})(\sum\limits_{i=0}^{n-1} t^i\boldsymbol{B_i}) = (\sum\limits_{i=0}^{n-1} t^{i+1}\boldsymbol{B_i}) - (\sum\limits_{i=0}^{n-1} t^i\boldsymbol{A}\boldsymbol{B}_i) = t^n\boldsymbol{B_{n-1}} + (\sum\limits_{i=1}^{n-1} t^i(\boldsymbol{B_{i-1}} - \boldsymbol{A}\boldsymbol{B_i})) - \boldsymbol{A}\boldsymbol{B_0}$

Recall that $P_f(t)\boldsymbol{I_n} = t^n\boldsymbol{I_n} + c_{n-1}t^{n-1}\boldsymbol{I_n} + \cdots + c_1t\boldsymbol{I_n} + c_0\boldsymbol{I_n}$

$t^n\boldsymbol{I_n} + c_{n-1}t^{n-1}\boldsymbol{I_n} + \cdots + c_1t\boldsymbol{I_n} + c_0\boldsymbol{I_n} = t^n\boldsymbol{B_{n-1}} + (\sum\limits_{i=1}^{n-1} t^i(\boldsymbol{B_{i-1}} - \boldsymbol{A}\boldsymbol{B_i})) - \boldsymbol{A}\boldsymbol{B_0}$

We can compose the cofficients $\boldsymbol{B_{n-1}} = \boldsymbol{I_n}$, $\boldsymbol{B_{i-1}} - \boldsymbol{A}\boldsymbol{B_i} = c_i\boldsymbol{I_n}$,

$-\boldsymbol{AB_0} = c_0\boldsymbol{I_n}$

Multiply by $\boldsymbol{A}^i$, $0 \leq i \leq n$.

$\boldsymbol{A^n B_{n-1}} + \sum_{i=1}^{n-1}(\boldsymbol{A^i B_{i-1} - A^{i+1} B_i}) - \boldsymbol{AB_0} = \boldsymbol{A^n} + c_{n-1}\boldsymbol{A^{n-1}} + \cdots + c_1\boldsymbol{A} + c_0\boldsymbol{I_n}$

$0 = P_f(\boldsymbol{A})$ □

7.6.9 Corollary

$m_f | P_f$

Example

(a) m_f and P_f are in general different. $f = \mathrm{Id}_V$, $\dim(V) = n$, $P_f(t) = (t-1)^n$, $m_f(t) = t - 1$

(b) Assume $f : V \to V$, $\dim(V) = r$, and assume that $\boldsymbol{A_f} = \boldsymbol{J_r(\lambda)}$. Then $P_f(t) = (t-\lambda)^r$. Moreover, $\boldsymbol{J_r(\lambda)} = \lambda\boldsymbol{I_n} + \boldsymbol{J_r(0)}$.

Notice that if $k \geq r$, $\boldsymbol{J_r(0)}^k = \boldsymbol{0}$

$(\boldsymbol{J_r(\lambda) - \lambda I_r})^k = \boldsymbol{J_r(0)}^k \neq 0$ when $0 \leq k \leq r-1$. We know that $m_f | (t-\lambda)^r$, m_f must be of the type $m_f = (t-\lambda)^k$. But the only possible is $k = r$. $m_f = P_f$.

For Jordan blocks the minimal polynomial is equal to the characteristic polynomial.

7.7 Jordan canonical form

7.7.1 Def

A field K is algebraic closed if any non-zero polynomial has a root in K.

Given a field K, we call algebraic closure of K the smallest algebraic closed field that contains K denoted as $\overline{K}$.

7.7.2 Theorem(Jordan canonical form)

Let $f \in \mathrm{End}(V)$ where V is a vector space of dimension n over an algebraic closed field. Then

(1) f can be represented by a Jordan matrix.

(2) The above matrix is unique up to premutation of Jordan blocks.

7.7.3 Def

Let $f \in \mathrm{End}(V)$ and let $\lambda \in K$. A vector $w \in V \setminus \{0\}$ is called a root vector of f corrsponding to λ, if there exists $r \in \mathbb{N}$, $(f - \lambda \mathrm{Id}_V)^r(w) = 0$

Remark

Eigenvectors are root vectors(corresponding to the eigenvalue) take $r = 1$.

Remark

Let $f = \boldsymbol{J_r(\lambda)} \in \mathrm{End}(V)$ be a Jordan block. Then any $v \in V$ is a root vector if f corresponding to λ.

7.7.4 Prop

Let $V(\lambda) = \{\text{root vectors of } f \text{ corresponding to } \lambda\} \bigcup \{0\}$. Then $V(\lambda)$ is a vector subspace of V. Moreover $V(\lambda) \neq \{0\}$ iff λ is an eigenvalue.

Proof

Take v_1, v_2 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 $(f - \lambda \mathrm{Id})^r(v_1 + v_2) = 0$.

$\alpha \in K$, $(f - \lambda \mathrm{Id})^r(v) = 0$, then $(f - \lambda \mathrm{Id})^r(\alpha v) = 0$.

Assume λ is an eigenvalue, we have an eigenvector corresponding to it,

$v \neq 0$ s.t. $(f - \lambda\mathrm{Id})(v) = 0$.

Vice versa, take $0 \neq v \in V(\lambda)$. Take the smallest r s.t. $(f - \lambda\mathrm{Id})^r(v) = 0$. Since $v' = (f - \lambda\mathrm{Id})^{r-1}(v) \neq 0$ and $(f - \lambda\mathrm{Id})(v') = 0$. This means v' is an eigenvector of λ, which means λ is an eiganvalue. $\square$

7.7.5 Prop

Let K be an algebraic closed field. Let $\lambda_1, \cdots, \lambda_k$ be the set of all distinct eigenvalues of f $(k \geq 1)$. Then $V = \bigoplus_{i=1}^{k} V(\lambda_i)$

Proof

Since K is algebraic closed, $P_f(t) = \prod_{i=1}^{k}(t - \lambda_i)^{r_i} \in K[t]$. Consider $F_i(t) = P_f(t)(t - \lambda_i)^{-r_i} \in K[t]$. Then define $f_i := F_i(f) \in \mathrm{End}(V)$, $V_i = \mathrm{Im}(f_i)$.

I want to prove that $(f - \lambda_i\mathrm{Id})^{r_i}(V_i) = 0$, which means $V_i \subseteq V(\lambda_i)$. $(f - \lambda_i\mathrm{Id})^{r_i} \circ f_i = P_f(f) = 0$.

Then I want to prove $V = V_1 + \cdots + V_n$. The polynomials $G_i(t)$ are coprime. $\exists G_i(t) \in K[t]$ s.t. $1 = \gcd(F_1(t), \cdots, F_n(t)) = F_1(t)G_1(t) + \cdots + F_n(t)G_n(t)$.

$\sum_{i=1}^{k} F_i(f) \circ G_i(f) = \mathrm{Id}$. Take $v \in V$, $v = \sum_{i=1}^{k} F_i(f) \circ (G_i(f))(v) = \sum_{i=1}^{k} f_i(G_i(f))(v) \in V_1 + \cdots + V_k \subseteq V(\lambda_1) + \cdots + V(\lambda_k)$.

Take $1 \leq i \leq k$, I want to prove that $V_i \bigcap (\sum_{i \neq j} V_j) = \{0\}$. Let v be a vector in this intersection. $(f - \lambda\mathrm{Id})^{r_i}(v) = 0$ and $F_i(f)(v) = \prod_{j \neq i}(f - \lambda_j\mathrm{Id})^{r_j}(v) = 0$. Notice that $(t - \lambda_i)^{r_i}$ and $F_i(t)$ are coprime. Thus there exists $G_1(t)$ and $G_2(t)$ s.t. $G_1(t)(t - \lambda_i)^{r_i} + G_2(t)F_i(t) = 1$. Substitute f instead of t, thus $v = 0$.

Finally, we want to prove that $V_i = V(\lambda_i)$. Take $v \in V(\lambda_i)$, write it as $v = v' + v'' \in V_i \oplus \bigoplus_{j \neq i} V_j$. Then $v'' = v - v' \in V(\lambda_i)$. It exists some $r \in \mathbb{N}$ s.t. $(f - \lambda_i\mathrm{Id})^r(v'') = 0$. Like in the previous step, $(t - \lambda_i)^r$ and $F_i(t)$ are coprime, $\exists H_1(t), H_2(t)$, s.t. $(t - \lambda_i)^r H_1(t) + F_i(t)H_2(t) = 1$. Hence $v'' = 0$. $v = v' \in V(\lambda)$ $\square$

7.7.6 Def

Let $f \in \mathrm{End}(V)$. Then f is said to be nilpotent if there exists $r \in \mathbb{N}$ s.t. $f^r = 0$.

7.7.7 Lemma 1

Let f be an nilpotent mapping, then $\mathrm{Ker}(f) = \{\text{set of eigenvectors of } f\} \bigcup \{0\}$

Proof

Let $0 \neq v \in \mathrm{Ker}(f)$, then v is an eigenvector.

Let v be an eigenvector, then $0 = f^m(v) = f^{m-1}(f(v)) = f^{m-1}(\lambda v) = \lambda^m v = 0$, thus $\lambda^m = 0$, which means $\lambda = 0$ □

7.7.8 Lemma 2

Let f be a nilpotent mapping, then $\mathrm{Ker}(f) \neq \{0\}$

Proof

Let r be the minimal integer s.t. $f^r = 0$, $f^{r-1}(V) \subseteq \mathrm{Ker}(f)$, but $f^{r-1}(V) \neq \{0\}$ because of the minimal of r.

Remark

Another way to prove the above lemma is to notice that $m_f = t^{r'}$, $1 \leq r' \leq r$. By Cayley-Hamilton theorem, 0 is an eigenvalue thus $f(x) = 0$ for some $x \neq 0$.

7.7.9 Theorem(Jordan form for nilpotent mappings)

Let f be a nilpotent mapping, then there exists a Jordan basis for f which means this basis gives a Jordan matrix made of blocks of the type $\boldsymbol{J_r(0)}$ for f.

Proof

Proof by induction on the dimension of the vector space V. If $\dim(V) = 1$, then $f = 0$ and $0 = \boldsymbol{J_1(0)}$.

Assume that the induction is true for $\dim(V) < n$, then we prove it on dimension n. Let $V_0 = \mathrm{Ker}(f) = \{\text{the set of eigenvectors}\} \bigcup \{0\}$. 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, we have a Jordan basis for $\overline{f}$, so 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 $f^{b_1}(v_1), \cdots, f^{b_m}(v_m)$. $< f^{b_1}(v_1), \cdots, f^{b_m}(v_m) > \subseteq V_0$

Extract a basis and complete 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 $v \in V$, $\pi(v) = \overline{v} = \sum\limits_{i=1}^{m} \sum\limits_{j=0}^{b_i-1} a_{ij} \overline{f}^j(\overline{v_i}) = \overline{\sum\limits_{i=1}^{m} \sum\limits_{j=0}^{b_i-1} a_{ij} f^j(v_i)}$.

Hence $v - \sum\limits_{i=1}^{m} \sum\limits_{j=0}^{b_i-1} a_{ij} f^j(v_i) \in V_0$, this finishes because I know V_0 is generated by $f^{b_1}(v_1), \cdots, f^{b_m}(v_m), u_1, \cdots, u_t$.

We show that the $f^{b_i}(v_i)$ and u_j are linearly independent.

$\sum\limits_{i=1}^{m} a_i f^{b_i}(v_i) + \sum\limits_{i=1}^{t} b_i u_t = 0$. The first observation is that $b_i = 0$, $\forall i$. So $0 = \sum\limits_{i=1}^{m} a_i f^{b_i}(v_i) = f(\sum\limits_{i=1}^{m} a_i f^{b_i - 1}(v_i))$. So $\sum\limits_{i=1}^{m} a_i f^{b_i-1}(v_i) \in V_0$.

It means that $\sum\limits_{i=1}^{m} a_i \overline{f}^{b_i-1}(\overline{v_i}) = 0$. Hence $a_i = 0$, $\forall i$.

By applying f to $\sum\limits_{i=1}^{m} \sum\limits_{j=0}^{b_i-1} a_{ij} f^j(v_i) + \sum\limits_{i=1}^{t} b_i u_i = 0$ many times, one can prove that $\{f^k(b_i), \cdots, u_t\}$ forms a Jordan basis of V. □

Note

实际上, 最后一部分只需证明 $\{f^{b_1}(v_1), \cdots, f^{b_m}(v_m)\}$ 线性无关即可, 利用定理 4.1.5 无需证明单射和满射。

7.7.10 Prop

The Jordan matrix that represents a nilpotent mapping $f \in \mathrm{End}(V)$ is unique up to permutations of the blocks.

Proof

$f^{b_i}(v_i), u_j$ forms a basis of $\mathrm{Ker}(f)$, thus these elements has exactly $\dim(V_0)$ elements, which is independent with the choice of basis.

Let's look at the elements looks like $f^{b_i-1}(v_i)$, if we work by induction then the proof is finished. $\square$

7.7.11 Theorem

Let K be an algebraic 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.

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 eigenvalue 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 $\boldsymbol{J_{g-\lambda\mathrm{Id}}}$ made of blocks of the type $\boldsymbol{J_r(0)}$.

Take the matrix $\boldsymbol{J_{g-\lambda\mathrm{Id}}}$. Restrict to a block $\boldsymbol{J_r(0)}$. $g-\lambda\mathrm{Id} = \boldsymbol{P^{-1}J_r(0)P}$. I want to show that with the same $\boldsymbol{P}$, $\boldsymbol{J_r(0)}+\lambda\boldsymbol{I_r} = \boldsymbol{J_r(\lambda)}$. This is a Jordan block for g. In fact, $\boldsymbol{J_r(\lambda)} = \boldsymbol{P^{-1}}(\boldsymbol{J_r(0)} + \lambda\boldsymbol{I_r})\boldsymbol{P} = g - \lambda\mathrm{Id} + \lambda\mathrm{Id} = g$.

Uniqueness from the uniqueness of the $\boldsymbol{J_r(0)}$. $\square$

7.7.12 Def

Let λ be an eigenvalue of $f \in \mathrm{End}(V)$.

$E(\lambda) = \mathrm{Ker}(f-\lambda\mathrm{Id}) = \{\text{all eigenvectors of } f \text{ with } \lambda\} \bigcup \{0\} \subseteq V(\lambda)$.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.

Note

几何重数等于 Jordan 块的个数。

代数重数等于 Jordan 块的阶数和。

7.7.13 Prop

Let K be an algebraic closed field. Then $\text{mult}(\lambda)_{\text{geo}} \leq \text{mult}(\lambda)_{\text{alg}}$, $\forall$ eigenvalue of f.

Proof

$V = V(\lambda_1) \oplus \cdots \oplus V(\lambda_n)$. Take $\lambda = \lambda_i$. Let $\boldsymbol{J_f}$ be the Jordan matrix of f. Then $\det(\boldsymbol{J_f}) = \det(f)$. So $P_f(t) = \prod(t - \lambda_i)^{\dim(V(\lambda_i))}$. $\dim(V(\lambda_i)) = \text{mult}(\lambda)_{\text{alg}}$.

Since $E(\lambda) \subseteq V(\lambda)$, then $\dim(E(\lambda)) \leq \dim(V(\lambda))$. $\square$

7.7.14 Corollary

Let K be an algebraic closed field, let $f \in \text{End}(V)$. f is diagonalizable iff $\text{mult}(\lambda)_{\text{geo}} = \text{mult}(\lambda)_{\text{alg}}$, $\forall\lambda$.

Proof

Work on a single $V(\lambda)$. $r = \dim(V(\lambda)) = \text{mult}(\lambda)_{\text{alg}}$.

We get a diagonal Jordan matrix iff we have exactly r blocks of length 1. It means that $r = \dim(\text{Ker}(f - \lambda\text{Id})) = \text{mult}(\lambda)_{\text{geo}}$. $\square$

Chapter 8

Inner product space

In this chapter, when we use K, we always mean $\mathbb{R}$ or $\mathbb{C}$.

8.1 Inner product

8.1.1 Def

Let V be a n dimensional vector space over K.

A bilinear form g on V means an element of $\mathrm{Hom}^{(2)}(V \times V, K)$. Choose a basis $\{v_1, \cdots, v_n\}$ of V. The matrix $\boldsymbol{G} = (g(v_i, v_j))_{ij} \in M_{n\times n}(K)$ is called the Gram matrix of g with respect to $\{v_1, \cdots, v_n\}$. Notice that $\boldsymbol{G}$ determines g uniquely.

8.1.2 Def

Let V be a vector space over $\mathbb{C}$.

The complex conjugate $\overline{V}$ is the same set of V, the sum on $\overline{V}$ is the same of V, but we define $\alpha * v := \overline{\alpha} v$, where $\alpha \in \mathbb{C}$.

If V and W are two complex vector spaces, then a semilinear mapping is a mapping $f : V \to W$ s.t. $f(v_1 + v_2) = f(v_1) + f(v_2)$ and $f(\alpha v) = \alpha * f(v)$, so a semilinear mapping is a linear mapping $f : V \to \overline{W}$. A sesquilinear form is a bilinear mapping, $g : V \times \overline{V} \to \mathbb{C}$, $g(av, bw) := a\bar{b}g(v, w)$.

Notation

$K^n \to V$, $e_i \mapsto v_i$. It's inverse is denoted as $x \mapsto \underline{x}$.(And by convention we always use $\underline{x}$ to denote the column vector only.)

Remark

Let g be a bilinear mapping and $\boldsymbol{G}$ be its Gram matrix.

If $g \in \mathrm{Hom}^{(2)}(V \times V, K)$, then $\forall (x, y) \in V^2$, $g(x, y) = \sum\limits_{ij} x_i y_j g(v_i, v_j) = \underline{x}^\tau \boldsymbol{G} \underline{y}$. If $g \in \mathrm{Hom}^{(2)}(V \times \overline{V}, K)$, then $\forall (x, y) \in V^2$, $g(x, y) = \underline{x}^\tau \boldsymbol{G} \overline{\underline{y}}$

Remark

Fix a basis $\{v_1, \cdots, v_n\}$ and $\boldsymbol{G} \in M_{n\times n}(K)$, the mapping $V \times V \to K$, $(x, y) \mapsto \underline{x}^\tau \boldsymbol{G} \underline{y}$, this is a bilinear form and the associated Gram matrix is exactly $\boldsymbol{G}$. Therefore, fixed a couple $(V, \{v_1, \cdots, v_n\})$, we have defined a bijection. $\mathrm{Hom}^{(2)}(V \times V, K) \cong M_{n\times n}(K)$, $g \mapsto \boldsymbol{G}$.

Similarly, $\mathrm{Hom}^{(2)}(V \times \overline{V}, K) \cong M_{n\times n}(K)$.

Remark

In the previous remark, we figure out that if you fix a basis and have a $n \times n$ matrix, then it would determine a unique bilinear mapping, also a blinear mapping would be a $n \times n$ matrix if you fix a basis.

Then we discuss what would happen to the Gram matrix of a fixed bilinear mapping if you change the basis.

Suppose we have a bilinear mapping g and its Gram matrix under the basis $\{w_j\}_{j=1}^n$ is $\boldsymbol{G}$. Then we change the basis from $\{v_i\}$ to $\{w_j\}$.

$$\begin{tikzcd} V \arrow[r, "\mathrm{Id}"] & V \\ K^n \arrow[u, "b_v"] \arrow[r, "\boldsymbol{P}"'] & K^n \arrow[u, "b_w"'] \end{tikzcd}$$

$\forall (x, y) \in V^2$, $g(x, y) = b_w^{-1}(x)^\tau \boldsymbol{G} b_w^{-1}(y) = (\boldsymbol{P} b_v^{-1}(x))^\tau \boldsymbol{G}\boldsymbol{P} b_v^{-1}(y) = b_v^{-1}(x)^\tau \boldsymbol{P}^\tau \boldsymbol{G}\boldsymbol{P} b_v^{-1}(y)$, thus the Gram matrix of g under the basis $\{v_i\}$ is $\boldsymbol{P}^\tau \boldsymbol{G}\boldsymbol{P}$, where $\boldsymbol{P} \in \mathrm{GL}(n; K)$, since Id, b_v and b_w are invertible.

$$\begin{tikzcd} K^n \arrow[r, "\boldsymbol{B}"] \arrow[d, "b_w"'] & K^n \arrow[d, "b_w"'] \\ V \arrow[r, "\varphi"] & V \\ K^n \arrow[u, "b_v"] \arrow[r, "\boldsymbol{A}"'] & K^n \arrow[u, "b_v"'] \end{tikzcd}$$

$\varphi \in \mathrm{End}(V)$, above commutative diagram shows that the two matrices of two different basis would satisfy the following property $\boldsymbol{B} = b_w \circ b_v^{-1} \circ \boldsymbol{A} \circ b_v \circ b_w^{-1}$. Notice that $b_v \circ b_w^{-1} \in \mathrm{GL}(n, K)$, thus it equals to some matrix $\boldsymbol{P}$, therefore, $\boldsymbol{B} = \boldsymbol{P^{-1}AP}$.

We say $\boldsymbol{B}$ and $\boldsymbol{A}$ are similar if $\exists P \in \mathrm{GL}(n; K)$ s.t. $\boldsymbol{B} = \boldsymbol{P^{-1}AP}$. Two different representative matrix of the same endomorphism are similar and we want to give a definition to show the relation between two Gram matrix of the same bilinear mapping.

8.1.3 Def

Two matrices $\boldsymbol{G}$, $\boldsymbol{G'} \in M_{n\times n}(K)$ are said to be congruent if $\exists \boldsymbol{A} \in \mathrm{GL}(n;K)$ s.t. $\boldsymbol{G'} = \boldsymbol{A}^\tau \boldsymbol{G}\boldsymbol{A}$.

Two matrices $\boldsymbol{G}$, $\boldsymbol{G'} \in M_{n\times n}(\mathbb{C})$ are said to be complex congruent if $\exists \boldsymbol{A} \in \mathrm{GL}(n;\mathbb{C})$ s.t. $\boldsymbol{G'} = \boldsymbol{A}^\tau \boldsymbol{G}\overline{\boldsymbol{A}}$.

Congruent and complex congruent are equivalence relations.

Remark

$\mathrm{Hom}^{(2)}(V\times V, K) \cong \mathrm{Hom}(T_0^2(V), K) \cong \mathrm{Hom}(V, V^\vee)$, $g \mapsto (g_s(x\otimes y) := g(x,y)) \mapsto (\tilde{g}(x) := g(x,\cdot))$

$\mathrm{Hom}^{(2)}(V\times \overline{V}, \mathbb{C}) \cong \mathrm{Hom}(V\otimes \overline{V}, K) \cong \mathrm{Hom}(V, \overline{V}^\vee)$, $g \mapsto (g_s(x\otimes y) := g(x,y)) \mapsto (\tilde{g}(x) := g(x,\cdot))$

8.1.4 Def

A bilinear mapping form $g : V\times V \to K$ is

- Symmetric if $g(x,y) = g(y,x)$, $\forall x, y$
- Sympletic or skew-symmetric if $g(x,y) = -g(y,x)$, $\forall x, y$

A sesquilinear form $g : V\times \overline{V} \to \mathbb{C}$ is

- Hermitian if $g(x,y) = \overline{g(y,x)}$

An inner product is any of (1) (2) (3). We say inner product space over K any pair (V, g), where V is a vector space over K and g is an inner product on V.

Example

$K^n \times K^n \to K$, $(x,y) \mapsto x^\tau y$ is symmetric.

$K^2 \times K^2 \to K$, $(v_1, v_2) \mapsto \det(v_1|v_2)$ is sympletic.

$\mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, $(x,y) \mapsto x^\tau \overline{y}$ is Hermitian.

8.1.5 Def

Let (V, g) be an inner product space over K. We say two vectors $v_1, v_2 \in V$ are orthogonal(with respect to g) if $g(v_1, v_2) = 0$. Notice that $g(v_1, v_2) = 0$ is equivalent to $g(v_2, v_1) = 0$ thus this definition satisfies our intuition.

We say two vector subspaces $V_1, V_2 \subseteq V$ are orthogonal if $g(v_1, v_2) = 0$, $\forall (v_1, v_2) \in V_1 \times V_2$.

Exercise

Fix a basis and use $\boldsymbol{A}$ to denote the Gram matrix of g.

- g is symmetric iff $\boldsymbol{A} = \boldsymbol{A}^{\tau}$
- g is symplectic iff $\boldsymbol{A} = -\boldsymbol{A}^{\tau}$
- g is Hermitian iff $\boldsymbol{A} = \overline{\boldsymbol{A}^{\tau}}$

Answer

Trivial. $\square$

Example

Let $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ be an open set.

$D^2 f : U \to \mathscr{L}(\mathbb{R}^n, \mathscr{L}(\mathbb{R}^n, \mathbb{R})) \cong \mathscr{L}^{(2)}(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$, $p \mapsto D^2 f(p)$. $D^2 f(p)$ is a bilinear form and its Gram matrix $\boldsymbol{G} = (\frac{\partial^2 f}{\partial x_j \partial x_i}(p))_{ij}$.

Since we have proved if $D^2 f(p)$ is continuous then $\forall p$, $D^2 f(p)$ is a symmetric blinear form which showed in Schwarz theorem before, thus its Gram matrix satisfies $\mathbf{G} = \boldsymbol{G}^{\tau}$, which means $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$.

Let's do some calculations to show the Gram matrix of $D^2 f(p)$ is $(\frac{\partial^2 f}{\partial x_j \partial x_i}(p))_{ij}$ which is equivaltent to prove that $\forall f : \mathbb{R}^n \to \mathbb{R}$, $D^2 f(p)(e_j, e_i) = \frac{\partial^2 f}{\partial x_i \partial x_j}(p)$, where e_i is the canonical basis of $\mathbb{R}^n$.

By definition, $f(p + x) - f(p) = \mathrm{d}_p f(x) + o(\|x\|)$, take $x = te_i$, then $\mathrm{d}_p f(e_i) = \lim\limits_{t \to 0} \frac{f(p + te_i) - f(p)}{t} =: \frac{\partial f}{\partial x_i}(p)$. Hence $\mathrm{d}f = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i} \mathrm{d}x_i$.

$\mathrm{d}f(p + x)(h) - \mathrm{d}f(p)(h) = D^2(p)(x)(h) + o(\|x\|)$. Write it more clear, $\mathrm{d}_{x+p} f(h) - \mathrm{d}_p f(h) = D^2(p)(x, h) + o(\|x\|)$. Take $x = te_j$, $h = e_i$, then $\frac{\partial f}{\partial x_i}(p + te_j) - \frac{\partial f}{\partial x_i}(p) = D^2(p)(te_j, e_i) + o(t)$, hence $D^2(p)(e_j, e_i) = \frac{\partial^2 f}{\partial x_i \partial x_j}(p)$.

8.1.6 Def

Let (V, g) be an inner product space over K. $\mathrm{Ker}(g) := \{v \in V | g(v, w) = 0, \forall w \in V \text{ or } \overline{V}\}$. Moreover, g is said non-degenerate if $\mathrm{Ker}(g) = \{0\}$.

Remark

Note that $\mathrm{Ker}(g) = \mathrm{Ker}(\tilde{g})$ when $\tilde{g} \in \mathrm{Hom}(V, V^\vee)$ or $\mathrm{Hom}(V, \overline{V}^\vee)$. In fact, $x \in \mathrm{Ker}(\tilde{g})$ means $\tilde{g_x} = 0$, which is equivalent to $g(x, y) = 0$, $\forall y \in V$.

This implies that $\mathrm{Ker}(g)$ is a linear subspace of V.

8.1.7 Prop

Let $g \in \mathrm{Hom}^{(2)}(V \times V, K)$ and let $\boldsymbol{G} = (G_{ij})_{n\times n}$ be the Gram matrix of g with respect to the basis $\{v_1, \cdots, v_n\}$. Then the matrix of $\tilde{g}$ w.r.t the basis $\{v_1, \cdots, v_n\}$ and $\{v_1^\vee, \cdots, v_n^\vee\}$ is $\boldsymbol{G^\tau}$.

Proof

Notation is in the below graph. $((b^\vee)^{-1} \circ \tilde{g} \circ b)(e_i) = (b^\vee)^{-1} \circ \tilde{g}(v_i) = (b^\vee)^{-1} \circ g(v_i, \cdot) = (b^\vee)^{-1} \circ g_{v_i}(\cdot)$.

$g_{v_i}(x) = g(v_i, \sum\limits_j x_j v_j) = \sum\limits_j x_j G_{ij} = \sum\limits_j G_{ij} v_j^\vee(x)$, $g_{v_i}(\cdot) = \sum\limits_j G_{ij} v_j^\vee$.

Thus $((b^\vee)^{-1} \circ \tilde{g} \circ b)(e_i) = (b^\vee)^{-1}(\sum\limits_j G_{ij} v_j^\vee) = \sum\limits_j G_{ij} e_j$ the i-row of $\boldsymbol{G}$.

$$\begin{tikzcd}
V \arrow[r, "\tilde{g}"] & V^\vee \\
K^n \arrow[u, "b"] \arrow[r, "\boldsymbol{G^\tau}"] & K^n \arrow[u, "b^\vee"']
\end{tikzcd}$$

$\square$

Remark

For $g \in \mathrm{Hom}^{(2)}(V \times \overline{V}, \mathbb{C})$, the matrix of $\tilde{g}$ w.r.t the basis $\{v_1, \cdots, v_n\}$ and $\{\overline{v_1}^\vee, \cdots, \overline{v_n}^\vee\}$ is $\boldsymbol{G^\tau}$.

8.1.8 Prop

Let g be an inner product on V. Then g is non-degenerate iff $\det(\boldsymbol{G}) \neq 0$.

Proof

Let $\boldsymbol{G}$ and $\boldsymbol{G'}$ be two Gram matrix of g. $\boldsymbol{G'} = \boldsymbol{A^\tau G A}$ for $\boldsymbol{A}$ invertible. $\det(\boldsymbol{G'}) = \det(\boldsymbol{A})^2 \det(\boldsymbol{G})$, hence $\det(\boldsymbol{G'}) \neq 0$ is equivalent to $\det(\boldsymbol{G}) \neq 0$.

Fix a basis $\{v_1, \cdots, v_n\}$, $\mathrm{Ker}(g) = \mathrm{Ker}(\tilde{g}) \cong \mathrm{Ker}(\boldsymbol{G}^{\boldsymbol{\tau}})$. $\mathrm{Ker}(g) = \{0\}$ iff $\mathrm{Ker}(\boldsymbol{G}^{\boldsymbol{\tau}}) = \{0\}$ iff $\boldsymbol{G}^{\boldsymbol{\tau}}$ is invertible iff $\det(\boldsymbol{G}^{\boldsymbol{\tau}}) \neq 0$ iff $\det(\boldsymbol{G}) \neq 0$. $\square$

8.2 Classification of low-dimensional spaces

8.2.1 Def

Let (V, g) and (W, g') be two inner product spaces s.t. $V \cong W$.

An isomorphism $f : (V, g) \to (W, g')$ of vector spaces is an isometry if $g(x, y) = g'(f(x), f(y))$. We say (V, g) and (W, g') are isometric if there are two isometrics $f : (V, g) \to (W, g')$, $g : (W, g') \to (V, g)$ s.t. $f \circ g = \mathrm{Id}_W$ and $g \circ f = \mathrm{Id}_V$.

Example 1

Let (V, g) be a vector spaces over K. $\dim V = 1$, g is symmetric. If $g(v, v) = 0$, $\forall v \in V \setminus \{0\}$, then g is degenerate and $g = 0$

If g is non-degenerate, let $v \in V \setminus \{0\}$, s.t. $g(v, v) \neq 0$. Let $c = g(v, v)$, $\forall u \neq 0 \in V$, $\exists x \neq 0 \in K$, $g(u, u) = x^2 c$. Therefore, each v induces a unique set $C(v) = \{x^2 c | x \in K^\times\}$, which is an element of $K^\times/(K^\times)^2$.

For $K = \mathbb{R}$, $\mathbb{R}^\times/(\mathbb{R}^\times)^2 \cong \{\pm 1\}$. For $K = \mathbb{C}$, $\mathbb{C}^\times \cong (\mathbb{C}^\times)^2$, $\mathbb{C}^\times/(\mathbb{C}^\times)^2 = 1$.

8.2.2 Prop

Let (V_1, g_1), (V_2, g_2) be two one dimensional vector spaces s.t. g_1 and g_2 are non-degenerate. Then (V_1, g_1) and (V_2, g_2) are isometric iff $\exists (v_1, v_2) \in V_1 \times V_2 \setminus \{(0, 0)\}$ s.t. $C_{g_1}(v_1) = C_{g_2}(v_2)$

Proof

From left to right : Let $f : V_1 \to V_2$ be an isometry, let $v_1 \in V_1 \setminus \{0\}$, then $v_2 = f(v_1) \neq 0$. $g_2(v_2, v_2) = g_2(f(v_1), f(v_1)) = g_1(v_1, v_1)$ thus $C_{g_1}(v_1) = C_{g_2}(v_2)$.

From right to left : Assume $\exists v_1 \neq 0, v_2 \neq 0$ s.t. $C_{g_1}(v_1) = C_{g_2}(v_2)$. Thus $g_1(v_1, v_1) = g_2(v_2, v_2)x^2$, $x \neq 0$. Then define $f : V_1 \to V_2$, $v_1 \mapsto xv_2$, $g_2(xv_2, xv_2) = x^2 g(v_2, v_2) = g(v_1, v_1)$. $\square$

8.2.3 Theorem

Let (V, g) be an inner product space. $\dim(V) = 1$ and g is symmetric. Then (V, g) is isometric to one of the following

- $K=\mathbb{R}$, $(\mathbb{R},\ g(x,y)=xy)$, $(\mathbb{R},\ g(x,y)=-xy)$, $(\mathbb{R}, g(x,y)=0))$
- $K=\mathbb{C}$, $(\mathbb{C}, g(x,y)=xy)$, $(\mathbb{C}, g(x,y)=0)$

Proof

Only prove the $\mathbb{R}$ case, if g is degenerate, then $g=0$. If g is non-degenerate, then $\exists v\neq 0$ s.t. $g(v,v)\neq 0$. Hence g corresponds to one element of $K^\times/(K^\times)^2$, which means (V,g) isometric to $(\mathbb{R}, g(x,y)=xy)$ or $(\mathbb{R}, g(x,y)=-xy)$. □

Example 2

Let (V,g) be a vector spaces over $\mathbb{C}$. $\dim V=1$, g is Hermitian. Choose $v\in V\setminus\{0\}$. If g is degenerate then $g=0$.

We use the same reasoning as above. $v\in V\setminus\{0\}$, $g(v,v)=a\neq 0$. $\forall w\neq 0\in V$, $\exists k\in\mathbb{C}^\times$, $g(w,w)=|k|^2 g(v,v)$. So any vector induces a coset in $\mathbb{C}^\times/\mathbb{R}_{>0}$.

Inside $\mathbb{C}^\times$, $\mathbb{R}_{>0}$ is a multiplication normal subgroup. Any $z\in\mathbb{C}^\times$ can write uniquely as $z=re^{i\theta}$, hence $C^\times\cong\mathbb{R}_{>0}\times S^1\to S^1$, $z\mapsto(r,e^{i\theta})\mapsto e^{i\theta}$. The kernel of $z\mapsto e^{i\theta}$ is $\mathbb{R}_{>0}$, therefore, $\mathbb{C}^\times/\mathbb{R}_{>0}\cong S^1$. But g is Hermitian $g(v,v)=\overline{g(v,v)}\in\mathbb{R}$.

It follows that the coset $\{|a|^2 g(v,v)|a\in\mathbb{C}^\times\}\in(\mathbb{C}^\times/\mathbb{R}_{>0})\bigcap\mathbb{R}=\{\pm 1\}$.

8.2.4 Theorem

Let (V,g) be an inner product space over $\mathbb{C}$. $\dim(V)=1$ and g is Hermitian. Then (V,g) is isometric to one of the following:

$(\mathbb{C}, g(x,y)=x\overline{y})$, $(\mathbb{C}, g(x,y)=-x\overline{y})$, $(\mathbb{C}, g(x,y)=0)$.

Example 3

Let (V,g) be a vector spaces over K. $\dim V=1$, g is symplectic. $g(v,v)=-g(v,v)$, thus $g(v,v)=0$, $\forall v\in V$.

8.2.5 Theorem

Let (V,g) be an inner product space. $\dim(V)=1$ and g is symplectic. Then (V,g) is isometric to $(K,0)$.

Example 4

Let (V, g) be a vector spaces over K. $\dim V = 2$, g is symplectic.

Assume that g is degenerate, and let $g(x, y) = 0$, $\forall y \in V$, for a fixed $x \in V \setminus \{0\}$. Extend x to a basis $\{x, x'\}$ of V, $g(ax + a'x', bx + bx') = 0$.

Take g non-degenerate, $\exists v_1, v_2$ s.t. $g(v_1, v_2) = a \neq 0$. We can assume $a = 1$, since $g(a^{-1}v_1, v_2) = 1$. Notice that v_1 and v_2 are linearly independent, then $\{v_1, v_2\}$ forms a basis of V.

$$g(\alpha_1 v_1 + \alpha_2 v_2, \beta_1 v_1 + \beta_2 v_2) = \begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

8.2.6 Theorem

Let (V, g) be an inner product space. $\dim(V) = 2$ and g is symplectic. Then (V, g) is isometric to the following

$(K^2, g(x, y) = 0)$, $(K^2, g(x, y) = x_1 y_2 - x_2 y_1)$.

8.3 Orthogonal structure

8.3.1 Def

Let (V, g) be an inner product space. Let $V_0 \subseteq V$ be a subspace. We say that V_0 is non-degenerate if $g|_{V_0}$ is non-degenerate. Moreover, we say V_0 is isotropic if $g|_{V_0} = 0$.

Example

$(\mathbb{R}^2, g)$, $g(x, y) = x_1 y_1 - x_2 y_2$ is symmetric and non-degenerate.
$\boldsymbol{G} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $V_0 =< \begin{pmatrix} 1 \\ 1 \end{pmatrix} >$, V_0 is isotropic $g|_{V_0} = 0$.

$(\mathbb{R}^2, g)$, $g(x, y) = x_1 y_2 - x_2 y_1$ is symplectic and non-degenerate.
$\boldsymbol{G} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, V_0 is isotropic $g|_{V_0} = 0$.

8.3.2 Def

Let (V, g) be an inner product space. Let $V_0 \subseteq V$ be a subspace. $V_0^\perp := \{v \in V | g(v_0, v) = 0, \forall v_0 \in V_0\}$, we say $V_0^\perp$ is the orthogonal compliment of V_0.

Exercise

$V_0^\perp$ is a vector subspace of V.

Answer

Trivial $\square$

8.3.3 Prop

Let (V, g) be an inner product space ($\dim(V) < +\infty$). If V_0 is non-degenerate, then $V = V_0 \oplus V_0^{\perp}$

Proof

Let $\{v_1, \cdots, v_r\}$ be a basis of V_0. Extend this basis to a basis $\{v_1, \cdots, v_r, v_{r+1}, \cdots, v_n\}$ of V. Let $\boldsymbol{G}$ be the Gram matrix of g w.r.t such basis.

$$\boldsymbol{G} = \begin{pmatrix} g(v_1, v_1) & \cdots & g(v_1, v_r) & \cdots & g(v_1, v_n) \\ \vdots & & & & \\ g(v_r, v_1) & \cdots & g(v_r, v_r) & \cdots & g(v_r, v_n) \\ \vdots & & & & \\ g(v_n, v_1) & \cdots & g(v_n, v_r) & \cdots & g(v_n, v_n) \end{pmatrix}$$

Let $\boldsymbol{G_0} := (g(v_i, v_j))_{r\times n} \in M_{r\times n}(K)$, $\boldsymbol{G_1} := (g(v_i, v_j))_{r\times r} \in M_{r\times r}(K)$

$V_0^{\perp} \cong \mathrm{Ker}(\boldsymbol{G_0})$, $\dim(V_0^{\perp}) = \dim(\mathrm{Ker}(\boldsymbol{G_0})) = n - \mathrm{rk}(\boldsymbol{G_0})$. Since $g|_{V_0}$ is non-degenerate, which means $\mathrm{rk}(\boldsymbol{G_1}) = r$, hence $\mathrm{rk}(\boldsymbol{G_0}) = \mathrm{rk}(\boldsymbol{G_1}) = r$. Thus $\dim(V_0^{\perp}) + \dim(V_0) = n$.

$\forall x \in V_0 \bigcap V_0^{\perp}$, $g(x, V_0) = 0$, which means $x = 0$. Therefore, $V_0 \oplus V_0^{\perp} = V$. □

8.3.4 Theorem

Let (V, g) be an inner product space. V_0 be a subspace of V. If both V_0 and $V_0^{\perp}$ are non-degenerate, then $(V_0^{\perp})^{\perp} = V_0$.

Proof

$\forall v_0 \in V_0$, $g(v_0, V_0^{\perp}) = 0$, thus $V_0 \subseteq (V_0^{\perp})^{\perp}$.

Since both V_0 and $V_0^{\perp}$ are non-degenerate, then $\dim(V_0) = \dim((V_0^{\perp})^{\perp})$, hence $V_0 = (V_0^{\perp})^{\perp}$. □

Remark

Notice that $V_0 \subseteq (V_0^{\perp})^{\perp}$ is also true in the infinite dimensional case.

8.3.5 Def

Let (V, g) be an inner product space. A basis $\{v_1, \cdots, v_n\}$ is said orthogonal if $g(v_i, v_j) = 0$, $\forall i \neq j$, which is equivalent to $V = V_1 \oplus \cdots \oplus V_n$

s.t. $g(V_i, V_j) = 0$ and $\dim(V_i) = 1$, $\forall i$.

8.3.6 Theorem

Let (V, g) be an inner product space, then there exists a decomposition $V = V_1 \oplus \cdots \oplus V_k$ s.t. V_i are pairwise orthogonal and

(a) They are one dimensional if g is symmetric or Hermitian

(b) They are one dimensional degenerate or 2-dimensional non-degenerate if g is symplectic.

Proof

We work by induction on $n = \dim(V)$.

If $\dim(V) = 1$, it is true.

Assume $\dim(V) = k$, if $g = 0$ there is nothing to prove. (Any decomposition into a direct sum of two one dimensional spaces is ok.)

- g symplectic : As we proved last section, since $g \neq 0$, $\exists x, x'$ s.t. $g(x, x') \neq 0$ and use V_0 to denote $< x, x' >$.

  Then $V = V_0 \oplus V_0^{\perp}$, since V_0 is non-degenerate(By classification of low dimensional inner product space). Then apply induction on $V^{\perp}$.

- g symmetric : It is enough to show that there exists $V_0 \subseteq V$ of dimension 1, which is non-degenerate. If such V_0 exists, then $V = V_0 \oplus V_0^{\perp}$ and apply induction.

  Assume that such V_0 doesn't exist, then $g(x, x) = 0$, $\forall x$. $0 = g(x + y, x + y) = 2g(x, y)$, $\forall x, y$, then $g = 0$.

- g Hermitian : Similarly, assume V_0 doesn't exist, then $0 = g(x+y, x+y) = 2\Re g(x, y)$, which means $g(x, y) = i\alpha$, $\alpha \in \mathbb{R}$. But if $\alpha \neq 0$, $0 = \Re g((i\alpha)^{-1}x, y) = \Re 1 = 1$, contradiction!

$\square$

8.3.7 Def

(V, g) is an inner product space with $\dim(V) = 1$. Assume that g is real symmetric or Hermitian. Then we say that (V, g) is positive if (V, g) is isometric to $(\mathbb{R}, g(x, y) = xy)$ or to $(\mathbb{C}, g(x, y) = x\overline{y})$.

We say that (V, g) is negative if (V, g) is isometric to $(\mathbb{R}, g(x, y) = -xy)$ or to $(\mathbb{C}, g(x, y) = -x\overline{y})$

8.3.8 Def

Let (V, g) be an inner product space, $\dim(V) = n$, $\dim(\mathrm{Ker}(g)) = \tau_0$, τ_+ is the number of positive subspaces in the orthogonal decomposition of V, τ_- is similar.

(1) If g is real symmetric or Hermitian, then (τ_0, τ_+, τ_-) is the signature of V.

(2) Otherwise τ_0 is the signature of V.

8.3.9 Theorem

Let (V, g) and (V', g') be two same dimensional inner product spaces with g, g' that are either (both) symplectic or complex symmetric. Then (V, g) and (V', g') are isometric iff $\tau_0 = \tau_0'$.

Proof

One direction is trivial.

- g symplectic : From the orthogonal decomposition we have a decomposition $V = \bigoplus_{i=1}^{k_1} V_i$. Notice that τ_0 is the number of the 1-dimensional degenerate subspaces of V. Similarly, let $V' = \bigoplus_{j=1}^{k_2} V_j'$, since $\tau_0 = \tau_0'$ and $\dim(V) = \dim(V')$, we know $k_1 = k_2$.

  Not loss generality, let $V_1, \cdots, V_{\tau_0}$ and $V_1', \cdots, V_{\tau_0'}'$ $(i \leq \tau_0)$ be one dimensional subspaces of V and V', respectively.

  By the classification theorem we have an isometry $f : V_i \to V_i'$.

  (1) $i \leq \tau_0$, they are one dimensional vector spaces with trivial inner product, so they are isometric.

  (2) $i > \tau_0$, $V_i \cong V_i' \cong (K, g(x, y) = x_1y_2 - x_2y_1)$.

- g complex symmetric : Similarly we have $V = \bigoplus_{i=1}^{n} V_i(*)$ and we want to show that τ_0 is the number of degenerate one dimensional subspaces.

Let $V_0 \subseteq V$ be the sum of the degenerate subspaces in $(*)$. $V_0 = V_1 \oplus \cdots \oplus V_k$. Take $v_0 \in V_0$, $v_0 = v_1' + \cdots + v_k'$, $\forall v \in V$, $v = v_1 + \cdots + v_k + \cdots + v_n$. $g(v_0, v) = \sum_{i \leq k} g(v_i', v_i) + \sum_{i \neq j} g(v_i', v_j) = 0$. Hence $V_0 \subseteq \mathrm{Ker}(g)$.

Viceversa if $v = \sum_{i=1}^{n} v_i$ and $\exists h > k$ s.t. $v_h \neq 0$. Then $g(v, v_h) = g(v_h, v_h) \neq 0$, so $v \notin \mathrm{Ker}(g)$. Therefore, $V_0 = \mathrm{Ker}(g)$.

Similar to the symplectic case, we construct the isometry in this way.

(1) $i \leq \tau_0$, Trivial.

(2) $i > \tau_0$, $V_i \cong V_i \cong (\mathbb{C}, g(x,y) = xy)$

8.3.10 Theorem(Sylvester's Law of Inertia)

Let (V, g) and (V', g') be two inner product spaces with g and g' real symmetric or Hermitian. Then (V, g) and (V', g') are isometric iff $(\tau_0, \tau_+, \tau_-) = (\tau_0', \tau_+', \tau_-)$.

Proof

If the spaces are isometric, then $n = n'$ and $\tau_0 = \tau_0'$, hence $\tau_+ + \tau_- = \tau_+' + \tau_-'$. $V = V_0 \oplus V_+ \oplus V_-$, $V' = V_0' \oplus V_+' \oplus V_-'$.

We have to prove that $\tau_+ = \tau_+'$. Assume by contradiction that $\tau_+ = \dim(V_+) > \dim(V_+') = \tau_+'$. Take $v \in V_+$, $f : V \to V'$ is the isometry, then $f(v) = f(v)_0 + f(v)_+ + f(v)_-$. We can consider the linear mapping $\pi_+ : V_+ \to V_+'$, $v \mapsto f(v)_+$, but by hypothesis $\exists v_0 \in V_+ \setminus \{0\}$ s.t. $\pi_+(v_0) = 0$, which means $f(v_0) = f(v_0)_0 + f(v_0)_-$. Since $v_0 \in V_+$, so $g(v_0, v_0) > 0$, since f is an isometry, $g'(f(v_0), f(v_0)) = g(v_0, v_0) > 0$. But $g'(f(v_0), f(v_0)) = g'(f(v_0)_0 + f(v_0)_+, f(v_0)_0 + f(v_0)_+) \leq 0$.

Assume that the signature are the same. $V = \bigoplus_{i=1}^{n} V_i$, $V' = \bigoplus_{i=1}^{n} V_i'$.

This is the orthogonal decomposition in one dimensional spaces. We have the same number of positive or negative vector spaces. By the classification we have isometries $f_i : V_i \to V_i'$, so we define $f = \sum_{i=1}^{n} f_i$. □

Remark

The dimension is a complete invariant for vector spaces (when considering their isomorphism).

The signature is a complete invariant for inner product spaces (when considering their isometry).

Reminder

Let (V, g) be an inner product space. A basis $\{v_1, \cdots, v_n\}$ of V is said orthogonal if $g(v_i, v_j) = 0$, $\forall i \neq j$.

8.3.11 Def

An orthogonal basis $\{v_1, \cdots, v_n\}$ is called orthonormal if $g(v_i, v_i) \in \{0, \pm 1\}$, $\forall i$.

Remark

If g is Hermitian or symmetric one can always find an orthogonal basis.

From an orthogonal basis, we can always construct an orthonormal basis.

Remark

If g is non-degenerate($\tau_0 = 0$), one can determine elements by the inner product.

We fix a orthogonal basis $\{v_1, \cdots, v_n\}$. Let $v \in V$, $v = \sum \alpha_i v_i$. $g(v, v_i) = \alpha_i g(v_i, v_i)$, thus $\alpha_i = \dfrac{g(v, v_i)}{g(v_i, v_i)}$, hence $v = \sum \dfrac{g(v, v_i)}{g(v_i, v_i)} v_i$. Moreover, if $\{v_1, \cdots, v_n\}$ is an orthonormal basis $v = \sum \pm g(v, v_i) v_i$.

Remark

Our theorems on the signature say that in the Hermitian or symmetric case, the number of elements in a orthonormal basis for which we have $g(v_i, v_i) = \begin{cases} 0 \\ +1 \\ -1 \end{cases}$ doen't depend on the basis.(This is essentially the signature.) It means that the Gram matrix of g w.r.t. any orthonormal basis (up to differs the order of the elements on the diagonal)

Real symmetric or Hermitian $\boldsymbol{G} = \begin{pmatrix} \boldsymbol{I}_{\tau_+} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I}_{\tau_-} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$

Complex symmetric $\boldsymbol{G} = \begin{pmatrix} \boldsymbol{I}_{\tau_+} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$

Remark

Let (V, g) be an inner product space and g symplectic.

$V = V_1 \oplus \cdots \oplus V_\tau \oplus V_{2\tau+1} \oplus \cdots \oplus V_n$. Consider a basis of V in the following way $\{v_1, \cdots, v_{2\tau}, v_{2\tau+1}, \cdots, v_n\}$ s.t. $g(v_i, v_{i+1}) = -g(v_{i+1}, v_i) = 1$, where $\{v_{2i-1}, v_{2i}\}$ is a basis for V_i, $i \leq \tau$. And $\{v_j\}$ is a basis for V_j where $j \geq 2\tau + 1$. Except $\{v_{2i-1}, v_{2i}\}(i \leq \tau)$, $g(v_k, v_j) = 0$.

A basis like this is called a symplectic basis for V. With respect to such basis the Gram matrix of g is like $\boldsymbol{G} = \begin{pmatrix} \boldsymbol{O} & \boldsymbol{I}_\tau & \boldsymbol{O} \\ -\boldsymbol{I}_\tau & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$

Conclusion

(1) Let $\boldsymbol{G}$ be a square symmetric matrix on $\mathbb{R}$ or a Hermitian matrix (over $\mathbb{C}$), then $\boldsymbol{G}$ can be reduced to diagonal form $\boldsymbol{D}$ with the transformation($\boldsymbol{A} \in \mathrm{GL}(n; K)$)

$\boldsymbol{G} \mapsto \boldsymbol{A}^\tau \boldsymbol{G} \boldsymbol{A} = \boldsymbol{D}$(Real case)

$\boldsymbol{G} \mapsto \boldsymbol{A}^\tau \boldsymbol{G} \overline{\boldsymbol{A}} = \boldsymbol{D}$ (Complex case)

where $\boldsymbol{D} = \begin{pmatrix} \boldsymbol{I}_{d_1} & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I}_{d_2} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$

(2) Let $\boldsymbol{G}$ be a square symmetric matrix over $\mathbb{C}$ then $\boldsymbol{G}$ can be reduced to a diagonal form $\boldsymbol{D}$ with the same transformation where $\boldsymbol{D} = \begin{pmatrix} \boldsymbol{I}_d & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} \end{pmatrix}$

8.4 Quadratic form

8.4.1 Def

Let V be a finite dimensional vector space over K and $\mathrm{char}(K) \neq 2$.

A quadratic form on V is a mapping $q : V \to K$ s.t.

- $q(\alpha v) = \alpha^2 q(v)$, $\forall v \in V$, $\forall \alpha \in K$
- $V \times V \to K$, $(u, v) \mapsto q(u+v) - q(u) - q(v)$ is bilinear

Remark

Given a quadratic form $q : V \to K$ we define a symmetric bilinear form $h_q(u, v) := \frac{1}{2}(q(u+v) - q(u) - q(v))$. And any symmetric bilinear form $h : V \times V \to K$ induces a quadratic form $q_h(v) := h(v, v)$.

These mappings are one the inverse of the other.(Since field's character must be a prime, we know $4 \neq 0$.) Hence we have a bijective correspondence between quadratic forms and bilinear symmetric forms.

Remark

Fix a basis $\{v_1, \cdots, v_n\}$, then any quadratic form can be written uniquely as $q(v) = h(v, v) = \underline{v}^\tau \boldsymbol{G} \underline{v}$.

$$\begin{tikzcd} v_i & V \arrow[r, "q"] \arrow[d, "b^{-1}"'] & K \\ e_i \arrow[u, mapsto] & K^n \arrow[u, bend right, "b"'] \arrow[ur, "\overline{q}"'] & \end{tikzcd}$$

$\overline{q}(x_1, \cdots, x_n) = \sum\limits_{1 \le i,j \le n} a_{ij} x_i x_j$, $a_{ij} = a_{ji}$, these are the elements of the Gram matrix $\boldsymbol{G}$ of the h w.r.t. the basis $\{v_1, \cdots, v_n\}$ and $\boldsymbol{G} = (a_{ij})_{n \times n}$.

The previous discussions shows that after a linear change of variable(choose an orthonormal basis) any quadratic form in $\mathbb{R}$ or $\mathbb{C}$ can be written as the form $\overline{q}(x_1, \cdots, x_n) = \sum a_i x_i^2$ with $a_i = \begin{cases} 0 \\ \pm 1 \end{cases}$.

Example

$f(x, y) = ax^2 + by^2 + cxy$, then you can rewrite it as $f(x', y') = a'x'^2 + b'y'^2$, which has distinct geometric meaning.

This simplification corresponds to rotating the coordinate system such that the axes align with the principal axes of the quadratic curve, eliminating the cross term cxy which indicates misalignment with these axes. The transformed coefficients a', b' (related to the eigenvalues(Why they are eigenvalues would be explain in the following course) of the quadratic form's matrix) dictate the curve's geometric type—ellipse, hyperbola, or degenerate case—and its orientation, with the new coordinates (x', y') directly spanning the principal axes.

8.4.2 Theorem(Gram-Schmidt algorithm)

Let (V, g) be an inner product space with g symmetric or Hermitian. Let $\{v'_1, \cdots, v'_n\}$ be a basis of V s.t. $V_i =< v'_1, \cdots, v'_i >$, $\forall i = 1, \cdots, n$ is non-degenerate.

Then there exists an orthogonal basis $\{v_1, \cdots, v_n\}$ s.t. $V_i =< v'_1, \cdots, v'_i > =< v_1, \cdots, v_i >$, $\forall i = 1, \cdots, n$.

Moreover, each v_i is determined uniquely up to multiplication by scalar.

Proof

We construct the v_i by induction. $v_1 = v'_1$.

If $v_1, \cdots, v_{i-1}$ have been already constructed, then v_i must be of the form $v_i = v'_i - \sum\limits_{j=1}^{i-1} x_j v_j$, with $x_j \in K$. I have $< v_1, \cdots, v_{i-1} >=< v'_1, \cdots, v'_{i-1} >$ by induction hypothesis, then $< v_1, \cdots, v_{i-1}, v_i >=< v_1, \cdots, v_{i-1}, v'_i >=< v'_1, \cdots, v'_{i-1}, v'_i >$.

We have to show that v_i is orthogonal to $\{v_1, \cdots, v_{i-1}\}$. It is enough to check $g(v_i, v'_k) = 0$, $\forall k < i$, which is equivalent to $g(v'_i, v'_k) = \sum\limits_{j=1}^{i-1} x_j g(v'_j, v'_k)$, $\forall k = 1, \cdots, i-1$.

This is a system of $i-1$ equations in $i-1$ unknowns (the x_j). The matrix of the linear system is the Gram matrix of $g|_{V_i}$ w.r.t. $\{v_1, \cdots, v_i\}$.

But since the V_i are non-degenerate, then the matrix is invertible, this means that the system as a unqiue solution. $\square$

Remark

For computations there is an easier way $v_i = v'_i - \sum_{j=1}^{i-1} y_j v_j$ instead of v'_j.

Now the condition $g(v_i, v_k) = 0$, $\forall 1 \le k \le i-1$ means $0 = g(v'_i, v_k) - y_k g(v_k, v_k)$, thus $y_k = \dfrac{g(v'_i, v_k)}{g(v_k, v_k)}$, $k = 1, \cdots, i-1$.(We use the condition V_k is non-degenerate through $g(v_k, v_k) \neq 0$)

8.5 Euclidean and Unitary spaces

8.5.1 Def

A Euclidean vector space (E, g) is a finite dimensional inner product space over $\mathbb{R}$ with symmetric and positive definite($= g(x, x) > 0$, if $x \neq 0$) inner product. And we denote $g(x, y) =: \langle x, y\rangle$

Remark

Any non-zero subspace of $(E, \langle\ ,\ \rangle)$ is non-degenerate. In fact, take $V \subseteq E$, if $\exists x \in V \setminus \{0\}$, $\langle x, V\rangle = 0$ then $\langle x, x\rangle = 0$.

The signature of E is of the type $(0, \tau_+, 0)$ denoted by $(p, q = 0)$.

Remark

A Euclidean space is a normed vector space whose norm is defined by $\|x\| := \sqrt{\langle x, x\rangle}$(Hence we need $\langle\ ,\ \rangle$ positive definite).

A Euclidean space admits an orthonormal basis $\{v_1, \cdots, v_n\}$, $\langle v_i, v_i\rangle = 1$. So orthonormal means $\|v_i\| = 1$.

In a Euclidean space we have a distance $d(x, y) = \|x - y\|$ thus it is a topological space.

Exercise

1. $|\langle x, y\rangle| \leq \|x\|\|y\|$
2. $\|x + y\| \leq \|x\| + \|y\|$
3. (Pythagoras's theorem) If $x_1, \cdots, x_k$ are pairwise orthogonal, then $\|\sum\limits_{i=1}^{k} x_i\|^2 = \sum\limits_{i=1}^{n} \|x_i\|^2$.

Answer

1. Not loss generality, suppose $x \neq 0$. Let $v = y - \frac{\langle x, y\rangle}{\|x\|^2}x$.

   Then $0 \leq \langle v, v\rangle = \langle y - \frac{\langle x, y\rangle}{\|x\|^2}x, y - \frac{\langle x, y\rangle}{\|x\|^2}x\rangle = \|y\|^2 - \frac{2\langle x, y\rangle}{\|x\|^2}\langle x, y\rangle + \frac{\langle x, y\rangle^2}{\|x\|^2}$, which means $\|x\|^2\|y\|^2 \geq \langle x, y\rangle^2$, which is equivalent to $|\langle x, y\rangle| \leq \|x\|\|y\|$. $\square$

2. $\|x+y\|^2 = \langle x+y, x+y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x,y\rangle \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2$.qed

3. Trivial. $\square$

Remark

It is noted that the above proof can be used to illustrate that any positive definite symmetric bilinear form can define a norm

8.5.2 Prop

A Euclidean space $(E, \langle\ ,\ \rangle)$ of dimension n is isometric to $(\mathbb{R}^n, \langle\ ,\ \rangle)$.

Proof

Let $\{v_1, \cdots, v_n\}$ be an orthonormal basis of E. Consider the mapping $f : v_i \mapsto e_i$, extends to an isomorphism of vector spaces.

$x = \sum \alpha_i v_i$, $y = \sum \beta_i v_i$, $f(x) = \sum \alpha_i e_i$, $f(y) = \sum \beta_i e_i$. Then $\langle x, y\rangle = \sum_i \alpha_i \beta_i = \langle f(x), f(y)\rangle$. $\square$

8.5.3 Def

Let $U, V \subseteq E$ be two subsets, then $d(U,V) := \inf\{\|v-u\|, u \in U, v \in V\} \geq 0$.

8.5.4 Def

Let V be a vector subspace of E, $x \in E$, since $E = V \oplus V^{\perp}$, we can write(uniquely) $x = x_0 + x_0'(x_0 \in V,\ x_0' \in V^{\perp})$. Then x_0 is called the orthogonal projection of x on V, x_0' is the orthogonal projection of x on $V^{\perp}$.

8.5.5 Prop

Use the above notation, $d(x,V) = \|x_0'\|$.

Proof

$\forall y \in V$, $\|x-y\|^2 = \|x_0 + x_0' - y\|^2 = \|x_0 - y\|^2 + \|x_0'\|^2 \geq \|x_0\|^2$ $\square$

8.5.6 Prop

Use the previous notations. Assume that $m = \dim(V)$, $\{v_1, \cdots, v_m\}$ is an orthonormal basis of V, then $x_0 = \sum_{j=1}^{n} \langle x_i, v_i \rangle v_i$.

Proof

Consider $y = x_0 - \sum_{i=1}^{m} \langle x, v_i \rangle v_i$, $y \in V$.

Moreover, since $\forall j$, $\langle y, v_j \rangle = \langle x - \sum_{i=1}^{m} \langle x, v_i \rangle v_i, v_j \rangle = \langle x, v_j \rangle - \langle x, v_j \rangle = 0$, $y \in V^{\perp}$, thus $y = 0$. $\square$

Remark

On $\mathbb{R}^n$ we have the notion of the volumes $\mathrm{vol}(B) = \lambda^n(B)$. B is a borel set. A n dimensional parallelepiped is $P_n := \{t_1 v_1 + \cdots + t_n v_n | t_i \in [0,1], \forall i\}$. Consider a linear mapping $\boldsymbol{A}_{\boldsymbol{P_n}} = \boldsymbol{A} : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto \boldsymbol{A}x$ where $\boldsymbol{A} = (v_1 | \cdots | v_n) \in M_{n \times n}(\mathbb{R})$. $\boldsymbol{A}$ is invertible iff $\{v_1, \cdots, v_n\}$ is a basis. Let $\Pi_n = [0,1]^n$ and $A(\Pi_n) = P_n$.

$\mathrm{vol}(P_n) = \lambda^n(P_n) = \int_{A(\Pi_n)} \mathbb{1}_{P_n} \mathrm{d}\lambda^n = \int_{\Pi_n} |\det(\boldsymbol{A})| \mathbb{1}_{\Pi_n} \mathrm{d}\lambda^n = |\det(\boldsymbol{A})| = \sqrt{\det(\boldsymbol{A}^{\tau} \boldsymbol{A})}$

Consider the Gram matrix $\boldsymbol{G} = (\langle v_i, v_j \rangle)_{n \times n}$. One can claim that $\det(\boldsymbol{G}) \neq 0$ iff $\{v_1, \cdots, v_n\}$ forms a basis. $\boldsymbol{A} = (v_1 | \cdots | v_n)$, $\boldsymbol{G} = \boldsymbol{A}^{\tau} \boldsymbol{A}$, $\det(\boldsymbol{G}) = \det^2(\boldsymbol{A})$.

8.5.7 Prop

$\mathrm{vol}(P_n) = \sqrt{\det(\boldsymbol{G})}$

Proof

See above. $\square$

8.5.8 Def

A complex vector space (H, h) which is an inner product where h is Hermitian and positive definite is called unitary space.

As in the Euclidean case we have orthonormal bases and we can define a norm $\|x\| := \sqrt{h(x,x)}$. A n dimensional unitary space is isomertic to $(\mathbb{C}^n, \langle \ , \ \rangle)$. The proof is the same as Euclidean case.

Exercise

Prove the triangle inequality.

Answer

One can claim that $|\Re h(x,y)| \leq \|x\|\|y\|$. Then the following thing is trivial.

In fact, not loss generality, suppose $x \neq 0$. Let $v = y - \frac{\Re h(x,y)}{\|x\|^2}x$. Then $0 \leq h(v,v) = h(y - \frac{\Re h(x,y)}{\|x\|^2}x, y - \frac{\Re h(x,y)}{\|x\|^2}x) = \|y\|^2 + \frac{(\Re h(x,y))^2}{\|x\|^2} - \frac{\Re h(x,y)}{\|x\|^2}h(x,y) - \frac{\Re h(x,y)}{\|x\|^2}h(y,x) = \|y\|^2 + \frac{(\Re h(x,y))^2}{\|x\|^2} - (\frac{\Re h(x,y)}{\|x\|^2}h(x,y) + \overline{\frac{\Re h(x,y)}{\|x\|^2}h(x,y)}) = \|y\|^2 + \frac{(\Re h(x,y))^2}{\|x\|^2} - 2\Re\frac{\Re h(x,y)}{\|x\|^2}h(x,y) = \|y\|^2 - \frac{(\Re h(x,y))^2}{\|x\|^2}$. □

8.5.9 Prop(Cauchy-Schwarz inequality)

Let (H,h) be a unitary space over finite dimensional case then $|h(x,y)| \leq \|x\|\|y\|$

Proof

Remember that we have proved $|\Re h(x,y)| \leq \|x\|\|y\|$.

Now take $h(x,y) = |h(x,y)|e^{i\theta}$. Then $\Re h(e^{-i\theta}x,y) = |h(x,y)|$, hence $|h(x,y)| = \Re(h(e^{-i\theta}x,y)) \leq \|e^{-i\theta}x\|\|y\| = \|x\|\|y\|$ □

8.6 Complex structure

8.6.1 Def

Let V be a complex vector space and $\dim(V) = n$. Since V is a vector space over $\mathbb{C}$, we have a mapping $\mathbb{C}\times V \to V$. If we restrict the multiplication by scalar to $\mathbb{R} \times V \to V$ then we have a new vector space $V_{\mathbb{R}}$, it has the same vectors of V and the same sum. $V_{\mathbb{R}}$ is called the decomplexification of V.

Let W be another $\mathbb{C}$-vector space $f : V \to W$ then we have an induced mapping $f_{\mathbb{R}} : V_{\mathbb{R}} \to W_{\mathbb{R}}$, $(\mathrm{Id}_V)_{\mathbb{R}} = \mathrm{Id}_{V_{\mathbb{R}}}$, $(f \circ g)_{\mathbb{R}} = f_{\mathbb{R}} \circ g_{\mathbb{R}}$, $(af + bg)_{\mathbb{R}} = af_{\mathbb{R}} + bg_{\mathbb{R}}$.

8.6.2 Theorem

Let V be a complex vector space, $n = \dim(V)$.

(1) Let $\{v_1, \cdots, v_n\}$ be a basis of V($\mathbb{C}$ basis). Then $\{v_1, \cdots, v_n, iv_1, \cdots, iv_n\}$ is a basis of $V_{\mathbb{R}}$. In particular $\dim(V_{\mathbb{R}}) = 2n$.

(2) $f : V \to W$ is a linear mapping. Its matrix representation w.r.t. to the basis $\{v_1, \cdots, v_n\}$, $\{v'_1, \cdots, v'_n\}$ (of W) is $\boldsymbol{A} = \boldsymbol{B}+i\boldsymbol{C}$. Then the matrix representation of $f_{\mathbb{R}} : V_{\mathbb{R}} \to W_{\mathbb{R}}$ w.r.t. the bases $\{v_1, \cdots, v_n, iv_1, \cdots, iv_n\}$ and $\{v'_1, \cdots, v'_m, iv'_1, \cdots, iv'_m\}$ is $\begin{pmatrix} \boldsymbol{B} & -\boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$

Proof

(1) Trivial. □

(2)

$$\begin{array}{ccc} V & \xrightarrow{\quad f \quad} & W \\ \uparrow{\scriptstyle \mathscr{B}} & & \uparrow{\scriptstyle \mathscr{B}'} \\ \mathbb{C}^n & \xrightarrow{\boldsymbol{B}+i\boldsymbol{C}} & \mathbb{C}^m \end{array}$$

$\boldsymbol{B} = (b_{jk})_{j,k}$, $i\boldsymbol{C} = (ic_{jk})_{j,k}$, $\forall k \in \{1, \cdots, n\}$, $f(v_k) = \mathscr{B}'(\sum\limits_{j=1}^{m} b_{jk}e_j + \sum\limits_{j=1}^{m} c_{jk}(ie_j)) = \sum\limits_{j=1}^{m} b_{jk}v'_j + \sum\limits_{j=1}^{m} c_{jk}(iv'_j)$

$f(iv_k) = \sum\limits_{j=1}^{m} b_{jk}(iv'_j) - \sum\limits_{j=1}^{m} c_{jk}v'_j$. We have the value of $f_{\mathbb{R}}$ on $\{v_1, \cdots, v_n, iv_1,$

$\cdots, iv_n\}$ the matrix is $\begin{pmatrix} \boldsymbol{B} & -\boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$ □

8.6.3 Corollary

Let $f : V \to V$ be a $\mathbb{C}$-linear mapping then $\det(f_{\mathbb{R}}) = |\det(f)|^2$

Proof

$$\det(f_{\mathbb{R}}) = \begin{vmatrix} \boldsymbol{B} & -\boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{vmatrix} = \begin{vmatrix} \boldsymbol{B}+i\boldsymbol{C} & -\boldsymbol{C}+i\boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{B} \end{vmatrix} = \begin{vmatrix} \boldsymbol{B}+i\boldsymbol{C} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{B}-i\boldsymbol{C} \end{vmatrix} =$$
$\det(\boldsymbol{B}+i\boldsymbol{C})\det(\boldsymbol{B}-i\boldsymbol{C}) = \det(f)\det(\overline{f}) = \det(f)\overline{\det(f)} = |\det(f)|^2$ □

8.6.4 Def

Let W be a real vector space. Consider $J \in \mathrm{End}(W)$ s.t. $J^2 = -\mathrm{Id}$. Then J is called a complex structure on W. The couple (W, J) is said to be a vector space with a complex structure.

Example

$J : V_{\mathbb{R}} \to V_{\mathbb{R}}$, $v \mapsto iv$. $(V_{\mathbb{R}}, J)$ is a vector space with a complex structure.

8.6.5 Theorem

Let (W, J) be a real vector space with complex structure then on W we introduce the following complex scalar multiplication $\mathbb{C} \times W \to W$, $(a + bi)w := aw + bJ(w)$. From this mapping we obtain a complex vector space $\tilde{W}$ s.t. $\tilde{W}_{\mathbb{R}} = W$.

Proof

Trivial. □

8.6.6 Corollary

If (W, J) is a vector space with a complex structure then $\dim(W)$ is even. Assume it is $2n$, then it is possible to find a basis s.t. J is represented by $\begin{pmatrix} \boldsymbol{O} & -\boldsymbol{I}_n \\ \boldsymbol{I}_n & \boldsymbol{O} \end{pmatrix}$.

Proof

W is the decomplexfication of $\tilde{W}$, this operation duplicates the dimension ($W = \tilde{W}_{\mathbb{R}}$). Take an arbitrary basis $\{v_1, \cdots, v_n\}$ of $\tilde{W}$ then $J_i : \tilde{W} \to \tilde{W}$, $w \mapsto iw$ can be expressed as $\boldsymbol{A} = i\boldsymbol{I_n}$. Use the theorem $(J_i)_{\mathbb{R}} = J$ is written as $\begin{pmatrix} \boldsymbol{O} & -\boldsymbol{I_n} \\ \boldsymbol{I_n} & \boldsymbol{O} \end{pmatrix}$

Remark

Let (H, h) be a unitary space. Consider a basis $\{v_1, \cdots, v_n\}$ and let $\boldsymbol{G}$ be the Gram matrix of h w.r.t. $\{v_1, \cdots, v_n\}$.

$\boldsymbol{G_h} = \boldsymbol{B} + i\boldsymbol{C}$, and now consider $\boldsymbol{G}_{\mathbb{R}} = \begin{pmatrix} \boldsymbol{B} & -\boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$ and on $\boldsymbol{H}_{\mathbb{R}}$ consider $\{v_1, \cdots, v_n, iv_1, \cdots, iv_n\}$. Then $\boldsymbol{G}_{\mathbb{R}}$ defines a real symmetric inner product $g = \langle\ ,\ \rangle$, which induces same norm of vectors. Be careful, h and $\langle\ ,\ \rangle$ are not the same.

Exercise

Check $\boldsymbol{G}_{\mathbb{R}} = \begin{pmatrix} \boldsymbol{B} & -\boldsymbol{C} \\ \boldsymbol{C} & \boldsymbol{B} \end{pmatrix}$ induces the same norm of vectors.

Answer

Easy but nontrivial. □

8.6.7 Prop

Let h be a Hermitian inner product on V, let J be a complex structure on $V_{\mathbb{R}}$ and $h(x, y) = a(x, y) + ib(x, y)$, where $a, b : V \times V \to \mathbb{R}$. Then the following things hold.

- $a(x, y)$ and $b(x, y)$ are inner product on $V_{\mathbb{R}}$, with a symmetric and b symplectic.

  In addition, $a(ix, iy) = a(x, y)$, $b(ix, iy) = b(x, y)$.(In other words, a and b are invariant by the multiplication by i, invariant w.r.t. the complex structure J of $V_{\mathbb{R}}$)

- $a(x, y) = b(ix, y)$, $b(x, y) = -a(ix, y)$

- Any pair of J-invariant bilinear forms on $V_{\mathbb{R}}$, $a, b : V_{\mathbb{R}} \times V_{\mathbb{R}} \to \mathbb{R}$ that are symmetric and symplectic, respectively and satisfies condition (2) actually define an Hermitian inner product $h(x, y) = a(x, y) + ib(x, y)$ on V. Moreover, h is positive definite iff a is positive definite.

Proof

(1) $h(x, y) = \overline{h(y, x)}$ this means $a(x, y) + ib(x, y) = a(y, x) - ib(y, x)$, so $a(x, y) = a(y, x)$ and $b(x, y) = -b(y, x)$.

$h(ix, iy) = h(x, y)$ so $a(ix, iy) = a(x, y)$ and $b(ix, iy) = b(x, y)$. $\square$

(2) $h(ix, y) = ih(x, y) = a(ix, y) + ib(ix, y) = i(a(x, y) + ib(x, y))$ $\square$

(3) The only non obvious axiom is linearity w.r.t. the first component for non-real numbers. $ih(x, y) = ia(x, y) + iib(x, y) = ia(x, y) - b(x, y) = ib(ix, y) + a(ix, y) = a(ix, y) + ib(ix, y) = h(ix, y)$. We have linearity in the first variable.

Positive definite follows from $b(x, x) = 0$ $\square$

Remark

The angle in unitary space is from 0 to $\frac{\pi}{2}$, which is defined by $\cos\phi := \frac{|h(x, y)|}{\|x\|\|y\|}$.

8.7 Orthogonal operator and Unitary operator

8.7.1 Prop

Let (V, g) be an inner product space with a non-degenerate inner product Hermitian or real symmetric and let $f \in \mathrm{End}(V)$. TFAE:

- f is an isometry
- $g(f(x), f(x)) = g(x, x)$, $\forall x \in V$
- Let $\{v_1, \cdots, v_n\}$ be a basis for V and let $\boldsymbol{G}$ be a Gram matrix of g w.r.t. such basis. If $\boldsymbol{A}$ is the matrix of f w.r.t. $\{v_1, \cdots, v_n\}$ then $\boldsymbol{A}^\tau \boldsymbol{G} \boldsymbol{A} = \boldsymbol{G}$ or $\boldsymbol{A}^\tau \boldsymbol{G} \overline{\boldsymbol{A}} = \boldsymbol{G}$
- Let $\{e_1, \cdots, e_n\}$ be an orthonormal basis, then $g(f(e_i), f(e_j)) = g(e_i, e_j)$, $\forall i, j$.
- The matrix of f w.r.t. any fixed orthonormal basis $\{v_1, \cdots, v_p, v_{p+1}, \cdots, v_{p+q}\}$ is $\boldsymbol{A}$, where $\begin{cases} (v_i, v_i) = 1 \text{ if } i \leq p \\ (v_i, v_i) = -1 \text{ if } i > p \end{cases}$. Then $\mathbf{A}$ satisfies the following property
  $\boldsymbol{A}^\tau \begin{pmatrix} \boldsymbol{I_p} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I_q} \end{pmatrix} \boldsymbol{A} = \begin{pmatrix} \boldsymbol{I_p} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I_q} \end{pmatrix}$ or $\boldsymbol{A}^\tau \begin{pmatrix} \boldsymbol{I_p} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I_q} \end{pmatrix} \overline{\boldsymbol{A}} = \begin{pmatrix} \boldsymbol{I_p} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{I_q} \end{pmatrix}$

Proof

From (1) to (2) is trivial. From (2) to (1) needs to distinguish two case.

1. symmetric real case $g(x, y) = \dfrac{1}{2}(g(x+y, x+y) - g(x, x) - g(y, y))$.
2. $q(x) = g(x, x)$, $\dfrac{1}{2}(q(x+y) - q(x) - q(y)) = \Re g(x, y)$, $\Im g(x, y) = -\Re g(ix, y)$(From 8.5.15). So if f preserves the quadratic form, then f preserves g.

The equivalence between (1) and (3) is trivial. $\square$

Exercise

Complete the remaining proof.

Answer

From (1) to (4), from (4) to (2).

From (3) to (5), from (5) to (4). □

8.7.2 Def

Let $(E, \langle\ ,\ \rangle)$ be a Euclidean space then an isometry $f \in \mathrm{Aut}(V)$ is called an orthogonal operator.

Let (H, h) be a unitary space then an isometry $f \in \mathrm{Aut}(H)$ is said to be a unitary operator.

8.7.3 Corollary

Orthogonal and unitary operators have the following property. Fix any orthogonal basis, the matrix representation $\boldsymbol{U}$ satisfies $\boldsymbol{U}\boldsymbol{U}^{\tau} = \boldsymbol{I_n}$ or $\boldsymbol{U}\overline{\boldsymbol{U}}^{\tau} = \boldsymbol{I_n}$.

8.7.4 Def

$\mathcal{O}(n) := \{\boldsymbol{A} \in \mathrm{GL}(n;\mathbb{R}) | \boldsymbol{A}\boldsymbol{A}^{\tau} = \boldsymbol{I_n}\}$

$\mathcal{U}(n) := \{\boldsymbol{A} \in \mathrm{GL}(n;\mathbb{C}) | \boldsymbol{A}\overline{\boldsymbol{A}}^{\tau} = \boldsymbol{I_n}\}$

Remark

Notice that $\mathcal{O}(n)$ is the group of matrices representation of orthogonal operators of $(\mathbb{R}^n, \langle\ ,\ \rangle)$ and $\mathcal{U}(n)$ is the group of matrices representation of unitary operators of $(\mathbb{C}^n, \langle\ ,\ \rangle)$.

Example

$\mathcal{U}(1) = \{a \in \mathbb{C} | a\overline{a} = 1\} = S^1 \subseteq \mathbb{C}$

$\mathcal{O}(1) = \{\pm 1\}$

8.7.5 Def

Remember that $\det : \mathrm{GL}(n;K) \to K^{\times}$ is a group morphism, then we define $\mathrm{SL}(n;\mathbb{R}) := \mathrm{Ker}(\det)$ and define $\mathrm{SO}(n;\mathbb{R}) := \mathrm{SL}(n;\mathbb{R}) \bigcap \mathcal{O}(n)$.

Remark

$\det|_{\mathcal{O}(n)} : \mathcal{O}(n) \to \{\pm 1\}$, $\mathrm{SO}(n) := \mathrm{Ker}(\det|_{\mathcal{O}(n)})$. $\mathcal{O}(n)/\mathrm{SO}(n;\mathbb{R}) \cong \{\pm 1\}$. $\mathrm{SO}(n;\mathbb{R})$ is a normal subgroup of $\mathcal{O}(n)$ of index 2.

Remark

$$\begin{tikzcd}[column sep=small] & M_{n\times n}(\mathbb{R}) \arrow[d, no head] & \\ & \mathrm{GL}(n;\mathbb{R}) \arrow[dl, no head] \arrow[dr, no head] & \\ \mathrm{SL}(n;\mathbb{R}) \arrow[dr, no head] & & \mathcal{O}(n) \arrow[dl, no head] \\ & \mathrm{SO}(n;\mathbb{R}) & \end{tikzcd}$$

Example

Let $\boldsymbol{T} \in \mathcal{O}(2)$. We have two cases (1) $\boldsymbol{T} \in \mathrm{SO}(2;\mathbb{R})$ (2) $\boldsymbol{T} \in \mathcal{O}(2) \setminus \mathrm{SO}(2;\mathbb{R})$

1. $\boldsymbol{T} \in \mathrm{SO}(2;\mathbb{R})$.

   Let $\boldsymbol{T} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Since $\boldsymbol{T}\boldsymbol{T}^{\boldsymbol{\tau}} = \boldsymbol{I_2}$, we have $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Hence $\begin{cases} ad - bc = 1 \\ a^2 + b^2 = 1 \\ c^2 + d^2 = 1 \\ ac + bd = 0 \end{cases}$.

   Let $a = \cos\alpha$, $b = \sin\alpha$, $adc - bc^2 = c$, $-bd^2 - bc^2 = c$, $c = -b = -\sin\alpha$.

   If $b \neq 0$, then $a = d = \cos\alpha$, if $b = c = 0$ then $\begin{cases} ad = 1 \\ a^2 = d^2 = 1 \end{cases}$, which means $\boldsymbol{T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

   So $\mathrm{SO}(2;\mathbb{R}) = \{ \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} | \alpha \in [0, 2\pi[\}$

2. $T \in \mathcal{O}(2) \setminus \mathrm{SO}(2;\mathbb{R})$. $\boldsymbol{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ is the reflection w.r.t. the line

$y = 0$.

$\boldsymbol{TA} \in \mathrm{SO}(2;\mathbb{R})$ since $\det(\boldsymbol{TA}) = 1$ and $\boldsymbol{TA} \in \mathcal{O}(2)$. By the previous reasoning $\boldsymbol{TA} = \boldsymbol{B} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ for $\alpha \in [0, 2\pi[$.

Then $\boldsymbol{T} = \boldsymbol{BA^{-1}} = \boldsymbol{BA}$. $\boldsymbol{T}$ is the reflection w.r.t. the line l whose angle with real line is $\frac{\alpha}{2}$.

8.7.6 Def

l is a line in $\mathbb{R}^2$.

A reflection R_l w.r.t. l is the unique endormorphism that satisfies $R_l|_l = \mathrm{Id}_l$ and $R_l(l^\perp) = -\mathrm{Id}_{l^\perp}$

Remark

We have shown that $\mathcal{O}(2)\backslash\mathrm{SO}(2;\mathbb{R}) = \{\text{reflections}\}$, $\mathrm{SO}(2;\mathbb{R}) = \{\text{rotations}\}$. Then $\mathcal{O}(2) = \{\text{rotations}\} \bigsqcup \{\text{reflections}\}$

Remark

Rotations form a subgroup of $\mathcal{O}(2)$.

$U(1) \cong \mathrm{SO}(2;\mathbb{R}) < \mathcal{O}(2)$

8.7.7 Theorem

(1) Let (H, h) be a unitary space. $f \in \mathrm{End}(H)$ is unitary iff it is diagonalizable in an orthonormal basis and with eigenvalues in S^1.

(2) Let $(E, \langle\ ,\ \rangle)$ be a euclidean space. $f \in \mathrm{End}(V)$ is orthogonal iff in an orthonormal basis f is represented by a matrix $\begin{pmatrix} \boldsymbol{R(\phi_1)} & \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{R(\phi_2)} & \cdots & \boldsymbol{O} & \boldsymbol{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{I_{d_1}} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} & \boldsymbol{-I_{d_2}} \end{pmatrix}$, which $\boldsymbol{R(\phi_i)} = \begin{pmatrix} \cos\phi_i & -\sin\phi_i \\ \sin\phi_i & \cos\phi_i \end{pmatrix}$

(3) The eigenvectors of orthogonal/unitary operators corresponding to different eigenvalues are orthogonal.

Proof

(1) Assume that $f = \boldsymbol{A}\boldsymbol{D}\overline{\boldsymbol{A}^{\tau}}$, $\boldsymbol{A} \in \mathcal{O}(n)$, $\boldsymbol{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$ with $\lambda_i\overline{\lambda_i} = 1$. Then notice that $\boldsymbol{D}\overline{\boldsymbol{D}}^{\tau} = \boldsymbol{I_n}$. This means that f can be represented by a unitary matrix $\boldsymbol{D}$ in some basis, thus f is unitary.

Assume f is unitary. Notice that we only needs to prove there exists an orthonormal basis s.t. the matrix representation of f is a diagonal matrix.

Let $\lambda \in \mathbb{C}$ be an eigenvalue and let L_λ be the eigenspace of λ. We can then write $H = L_\lambda \oplus L_\lambda^{\perp}$, since it's in unitary space. L_λ is an invariant subspace for f. Any orthonormal basis of L_λ is a basis of eigenvectors for $f|_{L_\lambda}$. So if we show that $L_\lambda^{\perp}$ is f invariant we are done, because then we can prove the theorem for $f|_{L_\lambda^{\perp}}$ and proved by induction.

Take $v \in L_\lambda^{\perp}$, $\forall v_0 \in L_\lambda$, then $h(f(v), v_0) = h(f(v), f(\lambda^{-1}v_0)) = \overline{\lambda^{-1}}h(v, v_0) = 0$. □

(2) From right to left is a simple check.

From left to right : Cases where E has dimension 1,2 are ok. Because of the classification of $\mathcal{O}(1)$, $\mathcal{O}(2)$. Assume $\dim(E) \geq 3$.

(a) f has at least one real eigenvalue. $E = L_\lambda \oplus L_\lambda^{\perp}$ and the proof is the same as in (1).

(b) f doesn't have real eigenvalues. I have two complex conjugate eigenvalues λ and $\overline{\lambda}$ with v and $\overline{v}$ be two eigenvectors of v and $\overline{v}$, respectively.

One can claim that v cannot be an eigenvectors of $\overline{\lambda}$. It means that $< v, \overline{v} >= L_0$ is two dimensional which is clearly f invariant. f has a two dimensional invariant subspace then $E = L_0 \oplus L_0^{\perp}$. It is enough to show that $f(L_0^{\perp}) \subseteq L_0^{\perp}$. $v_0 \in L_0$, $v \in L_0^{\perp}$, $\langle v_0, f(v)\rangle = \langle f \circ f^{-1}(v_0), f(v)\rangle = \langle f^{-1}(v_0), v\rangle = 0$ □

(3) $\lambda_1 \neq \lambda_2$, $h(v_1, v_2) = h(f(v_1), f(v_2)) = \lambda_1\overline{\lambda_2}h(v_1, v_2)$, which means $\lambda_1\overline{\lambda_2} = 1$, contradiction. □

Chapter 9

Differential forms in $\mathbb{R}^n$

本章记号十分混乱, 有历史因素影响，也与教师和教师选取的参考教材有关。

9.1 Differential forms

9.1.1 Def

Let $p \in \mathbb{R}^n$ be a fixed point. $\mathbb{R}^n_p = \{p\} \times \mathbb{R}^n$, $(p, a) \in \mathbb{R}^n_p$, $a \in \mathbb{R}^n$. $(p, a) + (p, b) := (p, a + b)$, $\alpha(p, a) := (p, \alpha a)$, $\forall \alpha \in \mathbb{R}$. 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 $(\mathbb{R}^n_p)^\vee \cong \{p\} \times (\mathbb{R}^n)^\vee$. And the dual basis is denoted by $\{\mathrm{d}x_1|_p, \cdots, \mathrm{d}x_n|_p\} := \{(e_1|_p)^\vee, \cdots, (e_n|_p)^\vee\}$.

$\mathbb{R}^n \times \mathbb{R}^n \cong \bigsqcup\limits_{p \in \mathbb{R}^n} \mathbb{R}^n_p$. $\bigsqcup\limits_p \mathbb{R}^n_p$ is called the tangent bundle of $\mathbb{R}^n$. We have a projection mapping $\pi : \bigsqcup\limits_{p \in \mathbb{R}^n} \mathbb{R}^n_p \to \mathbb{R}^n$, $a|_p \mapsto p$ and $\pi^{-1}(p) = \mathbb{R}^n_p$.

Remark

$$\frac{\partial x_i}{\partial x_j} = \mathrm{d}x_i|_p(e_j|_p) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise} \end{cases}$$

9.1.2 Def

$\bigwedge(\mathbb{R}^n_p)^\vee := \bigoplus\limits_{k \in \mathbb{N}} \bigwedge^k(\mathbb{R}^n_p)^\vee$. Fix k and consider $\bigwedge^k(\mathbb{R}^n_p)^\vee$ which is isomorphic to $(\bigwedge^k \mathbb{R}^n_p)^\vee$. The basis of this vector space is $\{\mathrm{d}x_{i_1}|_p \wedge \cdots \wedge \mathrm{d}x_{i_k}|_p | 1 \leq i_1 < \cdots < i_k \leq n\}$. $\dim(\bigwedge^k(\mathbb{R}^n_p)^\vee) = \binom{n}{k}$.

$\pi : \bigsqcup\limits_p (\bigwedge^k(\mathbb{R}^n_p)^\vee) \to \mathbb{R}^n$, $f|_p \mapsto p$ is induced by the natural projection mapping. An exterior k-form in $\mathbb{R}^n$ is a mapping $\omega : \mathbb{R}^n \to \bigsqcup\limits_p (\bigwedge^k(\mathbb{R}^n_p)^\vee)$, $p \mapsto \omega(p)$, that is a section of the projection π. $\pi \circ \omega = \mathrm{Id}_{\mathbb{R}^n}$ which means $\omega(p) \in \bigwedge^k(\mathbb{R}^n_p)^\vee$.

$\forall p \in \mathbb{R}^n$, $\omega(p) = \sum\limits_{1 \leq i_1 < \cdots < i_k \leq n} a_{i_1, \cdots, i_k}(p) \mathrm{d}x_{i_1}|_p \wedge \cdots \wedge \mathrm{d}x_{i_k}|_p$. 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.

Notation

$\omega = \sum\limits_I a_I \mathrm{d}x_I$, $I = (i_1, \cdots, i_k)$.

Example

Take $n = 4$.

1-form $\omega = a_1 \mathrm{d}x_1 + a_2 \mathrm{d}x_2 + a_3 \mathrm{d}x_3 + a_4 \mathrm{d}x_4$. $\omega(p) = a_1(p)\mathrm{d}x_1|_p + a_2(p)\mathrm{d}x_2|_p + a_3(p)\mathrm{d}x_3|_p + a_4(p)\mathrm{d}x_4|_p \in \bigwedge^1(\mathbb{R}^4_p)^\vee$

2-form $\omega = a_{12}\mathrm{d}x_1 \wedge \mathrm{d}x_2 + a_{13}\mathrm{d}x_1 \wedge \mathrm{d}x_3 + a_{14}\mathrm{d}x_1 \wedge \mathrm{d}x_4 + a_{23}\mathrm{d}x_2 \wedge \mathrm{d}x_3 + a_{24}\mathrm{d}x_2 \wedge \mathrm{d}x_4 + a_{34}\mathrm{d}x_3 \wedge \mathrm{d}x_4$

...

4-form $\omega(p) = a_{1234}(p)\mathrm{d}x_1|_p \wedge \mathrm{d}x_2|_p \wedge \mathrm{d}x_3|_p \wedge \mathrm{d}x_3|_p \wedge \mathrm{d}x_4|_p \in \bigwedge^4(\mathbb{R}^4_p)^\vee$

Remark

When $k = 0$, a C^m-differential 0-form is a mapping $f : \mathbb{R}^n \to \mathbb{R}$ of class C^m.

9.1.3 Def

$\Omega^k_{(m)}(\mathbb{R}^n) := \{\text{set of } C^m\text{-differential } k\text{-forms}\}$.

9.1.4 Prop

$\Omega^k_{(m)}(\mathbb{R}^n)$ is a module over $\Omega^0_{(m)}(\mathbb{R}^n)$

Proof

$\omega, \eta \in \Omega^k(\mathbb{R}^n)$, $(\omega + \eta)(p) := \omega(p) + \eta(p)$

$f \in \Omega^0(\mathbb{R}^n)$, $\omega \in \Omega^k(\mathbb{R}^n)$, $f\omega \in \Omega^k(\mathbb{R}^n)$, $(f\omega)(p) := f(p)\omega(p)$. $\square$

Notation

We use $\mathrm{d}f|_p$ to denote the differential of f at p.(Not use $d_p f$ anymore.)

Remark

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable mapping, then its differential $\mathrm{d}f|_p : \mathbb{R}^n_p \to \mathbb{R}_{f(p)} \cong \mathbb{R}$, thus $\mathrm{d}f|_p \in (\mathbb{R}^n_p)^\vee$, hence $\mathrm{d}f|_p = \sum f_i(p)\mathrm{d}x_i|_p$. By definition f_i are the partial derivatives of f. This means $\mathrm{d}f$ is a differential 1-form.

Moreover, $F : \mathbb{R}^n \to \mathbb{R}^m$ differetiable then $F = (F_1, \cdots, F_m)$, where $F_i : \mathbb{R}^n \to \mathbb{R}$ is differentiable. $\mathrm{d}F|_p : \mathbb{R}^n_p \to \mathbb{R}^m_{F(p)}$ and $\mathrm{d}F_i|_p = \mathrm{d}x_i|_{F(p)} \circ \mathrm{d}F|_p = \mathrm{d}(\mathrm{d}x_i \circ F)|_p$.

Note

用定义可以推导出 $\frac{\partial f}{\partial x_i}(p)(h_i) = \mathrm{d}f|_p(h_i)$, 从而 $\mathrm{d}f|_p(h) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)(h_i) = \sum\limits_{i=1}^{n} \frac{\partial f}{\partial x_i}(p)h_i$, 故 f_i 就是 f 在 p 点的 i-th 偏导数。

9.1.5 Def

$\Omega_{(m)}(\mathbb{R}^n) := \bigoplus\limits_k \Omega^k_{(m)}(\mathbb{R}^n)$. We define a wedge product on $\Omega(\mathbb{R}^n)$. $\Omega^k(\mathbb{R}^n) \times \Omega^l(\mathbb{R}^n) \to \Omega^{k+l}(\mathbb{R}^n)$, $(\omega, \eta) \mapsto \omega \wedge \eta$. Take $\omega = \sum\limits_I a_I \mathrm{d}x_I$, $\eta = \sum\limits_J b_J \mathrm{d}x_J$. $\omega \wedge \eta := \sum\limits_{IJ} a_I b_J \mathrm{d}x_{IJ}$. $I = (i_1, \cdots, i_k)$, $J = (j_1, \cdots, j_l)$, $IJ := (i_1, \cdots, i_k, j_1, \cdots, j_l)$.

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 $\theta \in \Omega^s(\mathbb{R}^n)$, $\omega \wedge (\varphi + \theta) = \omega \wedge \varphi + \omega \wedge \theta$ □

Proof

(1) Let $\omega = \sum\limits_I a_I \mathrm{d}x_I$, $\eta = \sum\limits_J b_J \mathrm{d}x_J$, and $\varphi = \sum\limits_K c_K \mathrm{d}x_K$.

Then $(\omega \wedge \eta) \wedge \varphi = (\sum\limits_{IJ} a_I b_J \mathrm{d}x_I \wedge \mathrm{d}x_j) \wedge (\sum\limits_K c_K \mathrm{d}x_K) = \sum\limits_{IJK} a_I b_J c_K \mathrm{d}x_I \wedge \mathrm{d}x_J \wedge \mathrm{d}x_K = \omega \wedge (\eta \wedge \varphi)$ □

(2) (3) Similar to (1) some calculations. □

9.2 Pullback of forms

9.2.1 Def

Fix $\omega \in \Omega^k_{(r)}(\mathbb{R}^m)$. Let $p \in \mathbb{R}^m$, use g to denote $\omega(p)$. Since $g \in \bigwedge^k(\mathbb{R}^m_p)^\vee$ thus $g = \sum\limits_{1\le i_1<\cdots<i_k\le n} a_{i_1,\cdots,i_k}(p)\mathrm{d}x_{i_1}|_p \wedge \cdots \wedge \mathrm{d}x_{ik}|_p$ actually is not a mapping. But remind that we have proved an isomprhism which is $\Phi : \bigwedge^k(V)^\vee \cong (\bigwedge^k(V))^\vee$, $e^\vee_{i_1}\wedge\cdots\wedge e^\vee_{i_k} \mapsto (e_{i_1}\wedge\cdots\wedge e_{i_k})^\vee := (e_{i_1}\wedge\cdots\wedge e_{i_k} \mapsto \det(\mathrm{d}x_j(e_{i_k}))_{jk})$ hence we define $g(v_1,\cdots,v_m) := \Phi(g)(v_1\wedge\cdots\wedge v_m)$.

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a mapping of class C^r, then it induces a mapping $f^* : \Omega^k_{(r)}(\mathbb{R}^m) \to \Omega^k_{(r)}(\mathbb{R}^n)$, $\omega \mapsto f^*\omega$. $(f^*\omega)(p)(v_1,\cdots,v_k) := \omega(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f|_p(v_k))$

Remark

Remember that $\bigwedge^k(\mathbb{R}^n)^\vee \cong (\bigwedge^k(\mathbb{R}^n))^\vee$, $e^\vee_{i_1}\wedge\cdots\wedge e^\vee_{ik} \mapsto (e_{i_1}\wedge\cdots\wedge e_{i_k})^\vee = (e_{j_1}\wedge\cdots\wedge e_{j_k} \mapsto \det(e^\vee_{i_l}(e_{j_\eta}))_{l\eta})$, which is equivalent to $f_1\wedge\cdots\wedge f_k \mapsto (e_{i_1}\wedge\cdots\wedge e_{i_k} \mapsto \det(f_l(e_{i_\eta}))_{l\eta})$. So we define $f_1\wedge\cdots\wedge f_k(v_1,\cdots,v_k) := \det(f_l(v_\eta))_{l\eta}$

Also remember that we define a wedge product on $\Omega(\mathbb{R}^n) := \bigoplus\limits_i \Omega^i(\mathbb{R}^n)$. $\Omega^k(\mathbb{R}^n)\times\Omega^l(\mathbb{R}^n) \to \Omega^{k+l}(\mathbb{R}^n)$, $(\omega,\eta)\mapsto \omega\wedge\eta$. Take $\omega = \sum\limits_I a_I\mathrm{d}x_I$, $\eta = \sum\limits_J b_J\mathrm{d}x_J$. $\omega\wedge\eta := \sum\limits_{IJ} a_Ib_J\mathrm{d}x_{IJ}$. $I = (i_1,\cdots,i_k)$, $J = (j_1,\cdots,j_l)$, $IJ := (i_1,\cdots,i_k,j_1,\cdots,j_l)$.

Let $\omega_i \in \Omega^k(\mathbb{R}^m)$, $\forall i \in \{1,\cdots,n\}$, $p \in \mathbb{R}^m$, then $(\omega_1\wedge\cdots\wedge\omega_n)(p) = \omega_1(p)\wedge\cdots\wedge\omega_n(p) = ((v_1,\cdots,v_k) \mapsto \det(\omega_l(p)(v_\eta))_{l\eta}$. The verification of the first equal sign is trivial(But not definition!) and the conclusion is important.

9.2.2 Prop

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable mapping. $\omega,\eta \in \Omega^k(\mathbb{R}^m)$ and $g : \mathbb{R}^m \to \mathbb{R}$ be a differentiable mapping ($g \in \Omega^0(\mathbb{R}^m)$). Then

(1) $f^*(\omega+\eta) = f^*\omega + f^*\eta$

(2) $f^*(g\omega) = f^*gf^*\omega$ where $f^*g := g\circ f$

(3) If $\omega_1,\cdots,\omega_k$ are 1-forms in $\mathbb{R}^m$, then $f^*(\omega_1\wedge\cdots\wedge\omega_k) = f^*\omega_1\wedge\cdots\wedge f^*\omega_k$

Proof

(1) $f^*(\omega+\eta)(p)(v_1,\cdots,v_k) = (\omega+\eta)(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f|_p(v_k)) = \omega(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f|_p(v_k)) + \eta(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f|_p(v_k)) = (f^*\omega)(p)(v_1,\cdots,v_k) + (f^*\eta)(p)(v_1,\cdots,v_k)$ $\square$

(2) $f^*(g\omega)(p)(v_1,\cdots,v_k) = (g\omega)(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f|_p(v_n)) = g(f(p))\,\omega(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f_p(v_k)) = (f^*g)(p)(f^*\omega)(p)(v_1,\cdots,v_k) = (f^*gf^*\omega)(p)(v_1,\cdots,v_k)$ $\square$

(3) $f^*(\omega_1\wedge\cdots\wedge\omega_k)(p)(v_1,\cdots,v_k) = \omega_1\wedge\cdots\wedge\omega_k(f(p))(\mathrm{d}f|_p(v_1),\cdots,\mathrm{d}f|_p(v_k)) = \det(\omega_i(f(p))(\mathrm{d}f|_p(v_j)))_{ij} = \det(f^*\omega_i(p)(v_j))_{ij} = f^*\omega_1\wedge\cdots\wedge f^*\omega_k(p)(v_1,\cdots,v_k)$

$\square$

Remark

$f:\mathbb{R}^n\to\mathbb{R}^m$, $(x_1,\cdots,x_n)\mapsto(f_1(x_1,\cdots,x_n),\cdots,f_m(x_1,\cdots,x_n))$. Let $\{y_1,\cdots,y_m\}$ be a standard basis of $\mathbb{R}^m$.

Let $\omega=\sum\limits_I a_I\mathrm{d}y_I\in\Omega^k(\mathbb{R}^m)$. $f^*\omega=\sum\limits_I f^*a_I(f^*\mathrm{d}y_{i_1})\wedge\cdots\wedge(f^*\mathrm{d}y_{i_k})$. Note that $f^*\mathrm{d}y_i(p)(v)=\mathrm{d}y_i(f(p))(\mathrm{d}f(v))=\mathrm{d}y_i|_{f(p)}(\mathrm{d}f(v))=\mathrm{d}f_i|_{f(p)}(v)$. Thus $f^*\omega=\sum\limits_I a_I(f)\mathrm{d}f_{i_1}\wedge\cdots\wedge\mathrm{d}f_{i_k}$.

Remark

Let U be an open set of $\mathbb{R}^n$ then consider $\Omega^k(U)\subseteq\Omega^k(\mathbb{R}^n)$.

Example

Let $U=\mathbb{R}^2\setminus\{(0,0)\}$ and $V=\{(r,\theta)|r>0, 0\le\theta<2\pi\}$.

Let $\omega=-\dfrac{y}{x^2+y^2}\mathrm{d}x+\dfrac{x}{x^2+y^2}\mathrm{d}y\in\Omega^1(U)$ and $f:V\to U$, $(r,\theta)\mapsto(r\cos\theta,r\sin\theta)$.

Let's compute $f^*\omega$. $\mathrm{d}f_1=\cos\theta\mathrm{d}r-r\sin\theta\mathrm{d}\theta$, $\mathrm{d}f_2=\sin\theta\mathrm{d}r+r\cos\theta\mathrm{d}\theta$. Then $f^*\omega=-\dfrac{r\sin\theta}{r^2}(\cos\theta\mathrm{d}r-r\sin\theta\mathrm{d}\theta)+\dfrac{r\cos\theta}{r^2}(\sin\theta\mathrm{d}r+r\cos\theta\mathrm{d}\theta)=\mathrm{d}\theta$

9.2.3 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.

Proof

(1) Let $\omega = \sum_I a_I \mathrm{d}x_I$, $\eta = \sum_J b_J \mathrm{d}y_J$. $f^*(\omega\wedge\eta) = f^*(\sum_{IJ} a_I b_J \mathrm{d}x_I \wedge \mathrm{d}y_J) = \sum_{IJ} a_I(f) b_J(f) \mathrm{d}f_I \wedge \mathrm{d}f_J = (\sum_I a_I(f)\mathrm{d}f_I) \wedge (\sum_J b_J(f)\mathrm{d}f_J) = (f^*\omega) \wedge (f^*\eta)$ □

(2) Let $\omega = \sum_I a_I \mathrm{d}y_I$. $(f\circ g)^*(\omega) = \sum_I a_I(f\circ g)\mathrm{d}(f\circ g)_I = \sum_I a_I(f(g))\mathrm{d}(f_{i_1}\circ g) \wedge \cdots \wedge \mathrm{d}(f_{i_k} \circ g) = \sum_I a_I(f\circ g)\mathrm{d}(f\circ g)_I = \sum_I a_I(f(g))\mathrm{d}(f_{i_1}(\mathrm{d}g)) \wedge \cdots \wedge \mathrm{d}(f_{i_k}(\mathrm{d}g)) = g^*(\sum_I a_I(f)\mathrm{d}f_I) = g^*(f^*\omega)$ □

9.3 Exterior differential

The differential of a mappings is a 1-form. $f \to \mathrm{d}f$ is from 0-form to 1-form. This section wants to generalize this process for any k-form.

$\mathrm{d} : \Omega^k_{(m)}(U) \to \Omega^{k+1}_{(m-1)}(U)$, $\omega \mapsto \mathrm{d}\omega$, $\omega = \sum\limits_I a_I \mathrm{d}x_I$, $\mathrm{d}\omega := \sum\limits_I \mathrm{d}a_I \wedge \mathrm{d}x_I$, $\mathrm{d}a_I = \sum\limits_I \frac{\partial a_{I,i}}{\partial x_i}\mathrm{d}x_i$.

Example

$\omega = xyz\mathrm{d}x + yz\mathrm{d}y + (x+z)\mathrm{d}z$

$\mathrm{d}\omega = \mathrm{d}(xyz) \wedge \mathrm{d}x + \mathrm{d}(yz) \wedge \mathrm{d}y + d(x+z) \wedge \mathrm{d}z = -xz\mathrm{d}x \wedge \mathrm{d}y + (1 - xy)\mathrm{d}x \wedge \mathrm{d}z - y\mathrm{d}y \wedge \mathrm{d}z$

9.3.1 Prop

(1) $\mathrm{d}(\omega_1 + \omega_2) = \mathrm{d}\omega_1 + \mathrm{d}\omega_2$, $\forall \omega_1, \omega_2 \in \Omega^k(U)$

(2) $\mathrm{d}(\omega \wedge \eta) = \mathrm{d}\omega \wedge \eta + (-1)^k \omega \wedge \mathrm{d}\eta$, $\omega \in \Omega^k(U)$, $\eta \in \Omega^l(U)$

(3) $\mathrm{d}(\mathrm{d}\omega) = 0$ ($\mathrm{d}^2\omega = 0$) $\omega \in \Omega^k(U)$

(4) $\mathrm{d}(f^*\omega) = f^*(\mathrm{d}\omega)$, $\omega \in \Omega^k(U)$, $f : V \to U$ is a differentiable function.

Proof

(1) Obvious. $\square$

(2) Let $\omega = \sum\limits_I a_I \mathrm{d}x_I$ and $\eta = \sum\limits_J b_J \mathrm{d}x_J$ then $\omega \wedge \eta = \sum\limits_{IJ} a_I b_J \mathrm{d}x_I \wedge \mathrm{d}x_J$

$\mathrm{d}(\omega \wedge \eta) = \sum\limits_{IJ} \mathrm{d}(a_I b_J) \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J = \sum\limits_{IJ} b_J \mathrm{d}a_I \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J + \sum\limits_{IJ} a_I \mathrm{d}b_J \wedge \mathrm{d}x_I \wedge \mathrm{d}x_J = \mathrm{d}\omega \wedge \eta + (-1)^k \sum\limits_{IJ} a_I \mathrm{d}x_I \wedge \mathrm{d}b_J \wedge \mathrm{d}x_J = \mathrm{d}\omega \wedge \eta + (-1)^k \omega \wedge \mathrm{d}\eta$ $\square$

(3) First assume $\omega = f \in \Omega^0(U)$. $\mathrm{d}(\mathrm{d}f) = \mathrm{d}(\sum\limits_{j=1}^n \frac{\partial f}{\partial x_j}\mathrm{d}x_j) = \sum\limits_{j=1}^n \mathrm{d}(\frac{\partial f}{\partial x_j}) \wedge \mathrm{d}x_j = \sum\limits_{j=1}^n (\sum\limits_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}\mathrm{d}x_i \wedge \mathrm{d}x_j)$. Remind that $\mathrm{d}x_i \wedge \mathrm{d}x_i = 0$ and $\frac{\partial f}{\partial x_i \partial x_j} = \frac{\partial f}{\partial x_j \partial x_i}$, thus $\mathrm{d}(\mathrm{d}f) = 0$.

By (1) not loss generality let $\omega = a_I \mathrm{d}x_I$. Notice that $\mathrm{d}^2 x_I = \mathrm{d}(1\mathrm{d}x_I) = \mathrm{d}(1) \wedge \mathrm{d}x_I = 0$. Hence $\mathrm{d}^2\omega = \mathrm{d}(\mathrm{d}\omega) = \mathrm{d}(\mathrm{d}a_I \wedge \mathrm{d}x_I) = \mathrm{d}^2 a_I \wedge \mathrm{d}x_I - \mathrm{d}a_I \wedge \mathrm{d}^2 x_I = 0$ $\square$

(4) Let $\omega = g \in \Omega^0(U)$, $g \circ f : \mathbb{R}^n \to \mathbb{R}^m \to \mathbb{R}$, $(x_1, \cdots, x_n) \mapsto (f_1(x_1, \cdots, x_n), \cdots, f_m(x_1, \cdots, x_m)) := (y_1, \cdots, y_m) \mapsto g(y_1, \cdots, y_m)$

$f^*(\mathrm{d}g) = f^*(\sum\limits_i \frac{\partial g}{\partial y_i}\mathrm{d}y_i) = \sum\limits_{ij} \frac{\partial g}{\partial y_i}(f(\cdot))\frac{\partial f_i}{x_j}\mathrm{d}x_j = \sum\limits_j \frac{\partial(g\circ f)}{\partial x_j}\mathrm{d}x_j = \mathrm{d}(g \circ f) = \mathrm{d}(f^*g)$

Let $\omega \in \Omega^k(U)$, $\omega = \sum\limits_I a_I \mathrm{d}x_I$. $\mathrm{d}(f^*\omega) = \mathrm{d}(f^*(\sum\limits_I a_I \mathrm{d}x_I)) = d(\sum\limits_I f^*a_I \wedge f^*\mathrm{d}x_I) = \sum\limits_I \mathrm{d}(f^*a_I \wedge f^*\mathrm{d}x_I) = \sum\limits_I \mathrm{d}(f^*a_I) \wedge f^*\mathrm{d}x_I + f^*a_I \wedge \mathrm{d}(f^*\mathrm{d}x_I) = \sum\limits_I f^*(\mathrm{d}a_I) \wedge f^*\mathrm{d}x_I = f^*(\sum\limits_I \mathrm{d}a_I \wedge \mathrm{d}x_I) = f^*(\mathrm{d}\omega)$ □

Remark

It forms a complex chain.

$$0 \xrightarrow{\quad i \quad} \Omega^0(U) \xrightarrow{\quad \mathrm{d} \quad} \Omega^1(U) \xrightarrow{\quad \mathrm{d} \quad} \cdots \xrightarrow{\quad \mathrm{d} \quad} 0$$

Note

$$\mathbb{R}^n \xrightarrow{\quad f \quad} \mathbb{R}^m \xrightarrow{\quad g \quad} \mathbb{R}$$

先前证明过 $\mathrm{d}(g \circ f)|_p = \mathrm{d}g|_{f(p)} \circ \mathrm{d}f_p$, 从而

$\mathrm{d}(g \circ f)|_p(\cdot) = \mathrm{d}g|_{f(p)}(\mathrm{d}f|_p(\cdot)) = \nabla g(f(p))\mathrm{d}f|_p(\cdot) = \nabla g(f(p))\mathrm{d}f|_p(\cdot) =$

$$\nabla g(f(p))\boldsymbol{J_f(p)}(\cdot) = (\tfrac{\partial g}{\partial y_1}(f(p))\ \tfrac{\partial g}{\partial y_2}(f(p))\ \cdots\ \tfrac{\partial g}{\partial y_m}(f(p))) \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \cdots & \frac{\partial f_1}{\partial x_n}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) & \cdots & \frac{\partial f_2}{\partial x_n}(p) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \frac{\partial f_m}{\partial x_2}(p) & \cdots & \frac{\partial f_m}{\partial x_n}(p) \end{pmatrix} (\cdot)$$

$= \sum\limits_{ij} \frac{\partial g}{\partial y_i}(f(p))\frac{\partial f_i}{\partial x_j}(p)(\cdot)$

又有 $\mathrm{d}(g \circ f) = \sum\limits_j \frac{\partial(g\circ f)}{\partial x_j}\mathrm{d}x_j$, 从而 $\frac{\partial(g\circ f)}{\partial x_j} = \sum\limits_i \frac{\partial g}{\partial y_i}(f(\cdot))\frac{\partial f_i}{\partial x_j}$

9.4 Line integral of the second kind

Fix $\omega = \sum\limits_i a_i \mathrm{d}x_i \in \Omega^1_{(m)}(\mathbb{R}^n)$, $m \geq 1$.

9.4.1 Def

Let $U \subseteq \mathbb{R}^n$ be an open set, $\gamma : [a,b] \to \mathbb{R}^n$. $\exists t_0 = a < t_1 < \cdots < t_k < t_{k+1} = b$ s.t. $\gamma|_{]t_j,t_{j+1}[} =: \gamma_j$ is of class C^1. Then we define γ as a parametric curve and $t \in [a,b]$ are parameters.

Example

$f : \mathbb{R} \to \mathbb{R}$ of class C^1, then $\gamma : t \mapsto (t, f(t))$ is a parametric curve.

$\gamma : t \mapsto (\cos t, \sin t)$ this is a parametric curve.

$\gamma_j :]t_j, t_{j+1}[\to \mathbb{R}^n$, we can define $\gamma_j^* \omega$. This is a one form in $\Omega^1(]t_j, t_{j+1}[)$. If $\gamma_j(t) = (x_1(t), \cdots, x_n(t))$ then $\gamma_j^*\omega = \sum\limits_i a_i(x_1(t), \cdots, x_n(t))\mathrm{d}x_i = \sum\limits_i a_i(x_1(t), \cdots, x_n(t))\dfrac{\mathrm{d}x_i}{\mathrm{d}t}\mathrm{d}t$

Remark

Notice that the $\mathrm{d}x_i$ in $\gamma_j^*\omega$ is different from the $\mathrm{d}x_i$ in ω.

9.4.2 Def

Let ω and γ be as above $\displaystyle\int_\gamma \omega := \sum_j \int_{t_j}^{t_{j+1}} \gamma_j^* \omega$

Remark

Fix $\gamma(t)$, $\gamma'(t) = (\dfrac{\mathrm{d}x_1}{\mathrm{d}t}(t), \cdots, \dfrac{\mathrm{d}x_n}{\mathrm{d}t}(t))$ = the tangent vector of γ in $\gamma(t)$

$\displaystyle\int_{t_j}^{t_{j+1}} \gamma_j^*\omega = \int_{t_j}^{t_{j+1}} \langle a \circ \gamma, \gamma_j' \rangle \mathrm{d}t$, since $a \circ \gamma = (a_1(x_1(t), \cdots, x_n(t)), \cdots, a_n(x_1(t), \cdots, x_n(t)))$ and $\gamma_j' = (\dfrac{\mathrm{d}x_1}{\mathrm{d}t}, \cdots, \dfrac{\mathrm{d}x_n}{\mathrm{d}t})$.

Example(Physics)

Let $F : \mathbb{R}^3 \to \mathbb{R}^3$ be a field.

γ is the path of the particle under the action of F.

Fix t, $\Delta t = h$, $\Delta\gamma = \gamma(t+h) - \gamma(t)$. $\Delta\gamma \sim \gamma'(t)\Delta t$ and $\lim\limits_{\Delta t \to 0} \dfrac{\Delta\gamma}{\Delta t} = \gamma'(t)$. And then $\langle F(\gamma(t)), \Delta\gamma \rangle \sim \langle F(\gamma(t)), \gamma'(t) \rangle \Delta t$

Consider a "conservative force" if there exists a potential $\phi : \mathbb{R}^n \to \mathbb{R}$ s.t. $F = -\nabla\phi = -(\frac{\partial\phi}{\partial x_1}, \cdots, \frac{\partial\phi}{\partial x_n})$, $F : \mathbb{R}^n \to \mathbb{R}^n$

$W_{F,\gamma} = \int_\gamma -\mathrm{d}\phi$ and $\mathrm{d}\phi = \frac{\partial\phi}{\partial x_1}\mathrm{d}x_1 + \cdots + \frac{\partial\phi}{\partial x_n}\mathrm{d}x_n$

$G : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto yg$. g gravity constant and G is a potential. $-\nabla G = (0, -g, 0)$ this is the gravitational force field in $\mathbb{R}^3$.

9.5 Interlude : Complements of measure theory

This section wants to define a measure on $(X \times Y, \sum_X \otimes \sum_Y)$ from $(X, \sum_X, \mu)$ and $(Y, \sum_Y, \nu)$.

Reminder

Let $(X, \sum_X)$ and $(Y, \sum_Y)$ be two measurable spaces. ($\sum_X$ and $\sum_Y$ are two $\sigma-$algebras) The $\sigma-$algebra on $X \times Y$ is defined as $\sum_X \otimes \sum_Y := \sigma(\{S_1 \times S_2 | S_1 \in \sum_X, S_2 \in \sum_Y\})$.

Let $(X, \sum_X, \mu)$ be a measure space. We say that it is $\sigma-$finite if there exists a sequence $\{E_n\}_{n\in\mathbb{N}}$ of measurable sets ($E_n \in \sum_X$) s.t. $X = \bigcup\limits_n E_n$, $\mu(E_n) < +\infty$, $\forall n$.

Exercise

The two definitions of $\sum_X \otimes \sum_Y$ are equivalent.(The other definition of $\sum_X \otimes \sum_Y$ is $\sigma((\pi_1^{-1}(\sum_X) \bigcup \pi_2^{-1}(\sum_Y)))$, which is defined in 6.7)

Answer

Notice that the previous definition of $\sum_X \otimes \sum_Y = \sigma(\{A \times Y | A \in \sum_X\} \bigcup \{X \times B | B \in \sum_Y\})$. It's trivial to check that $\sigma(\{A \times Y | A \in \sum_X\} \bigcup \{X \times B | B \in \sum_Y\}) \subseteq \sigma(\{S_1 \times S_2 | S_1 \in \sum_X, S_2 \in \sum_Y\})$ and $\sigma(\{S_1 \times S_2 | S_1 \in \sum_X, S_2 \in \sum_Y\}) \subseteq \sigma(\{A \times Y | A \in \sum_X\} \bigcup \{X \times B | B \in \sum_Y\})$. $\square$

Example

$(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda(\text{lebesgue}))$ is σ-finite. $\mathbb{R} = \bigcup\limits_n [-n, n]$, $\lambda([-n, n]) = 2n < +\infty$

Notation

Take $A \subseteq X \times Y$.

For $x \in X$, $A_x := \{y \in Y | (x, y) \in A\} = \pi_2((\{x\} \times Y) \bigcap A)$. "vertical section" of A or x-fiber of A

For $y \in Y$, $A_y := \{x \in X | (x, y) \in A\} = \pi_1((X \times \{y\}) \bigcap A)$. "horizontal section" of A or y-fiber of A

9.5.1 Def

Let X be a set, then $\mathcal{D} \subseteq \mathscr{P}(X)$ is a Dynkin system iff

(1) $X \in \mathcal{D}$

(2) $\forall D \in \mathcal{D}$, $X \setminus D \in \mathcal{D}$

(3) If $\{D_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathcal{D}$ of pointwise disjoint sets then $\bigsqcup\limits_{n\in\mathbb{N}} D_n \in \mathcal{D}$

Example

A σ-algebra is a Dynkin system.

9.5.2 Def

Let $\mathcal{G} \subseteq \mathscr{P}(X)$, then $\delta(\mathcal{G}) \subseteq \mathscr{P}(X)$ 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 $\delta(\mathcal{G}) \subseteq \mathcal{D}$

Exercise

$\delta(\mathcal{G})$ exists and it is unique.

Answer

Notice that the intersection of Dynkin systems is still a Dynkin system, thus just take the intersection of all the Dynkin system containing $\mathcal{G}$, one can get $\delta(\mathcal{G})$, uniqueness is trivial. $\square$

9.5.3 Prop

If $\mathcal{D}$ is a Dynkin system closed under the finite intersection then it is a σ-algebra.

Proof

Let $\{D_n\}$ be a sequence in $\mathcal{D}$, define $E_n := D_n \bigcap (X \setminus \bigsqcup\limits_{m<n} E_m) = D_n \setminus \bigsqcup\limits_{m<n} E_m$ $\square$

9.5.4 Prop

Let X be a set, and let $\mathcal{G} \subseteq \mathscr{P}(X)$. Assume that $\mathcal{G}$ is closed under the finite intersection, then $\delta(\mathcal{G}) = \sigma(\mathcal{G})$.

Proof

We know $\delta(\mathcal{G}) \subseteq \sigma(\mathcal{G})$. We need to show that $\delta(\mathcal{G})$ is a σ-algebra so we need to show it is closed under intersection.

Let $D \in \delta(\mathcal{G})$, and define $\delta_D := \{E \subseteq X | E \bigcap D \in \delta(\mathcal{G})\}$. One can claim that δ_D is a Dynkin system.

(1) $X \bigcup D = D$, so $X \in \delta_D$.

(2) (**Paolo**) $\forall E \in \delta_D$, $(X \setminus E) \bigcap D = ((X \setminus E) \bigcap D) \bigcup ((X \setminus D) \bigcap D) = ((X \setminus E) \bigcup (X \setminus D)) \bigcap D = (X \setminus (E \bigcap D)) \bigcap D = (X \setminus (E \bigcap D)) \bigcap (X \setminus (X \setminus D)) = X \setminus ((E \bigcap D) \bigcup (X \setminus D))$ And $(E \bigcap D) \bigcap (X \setminus D) = \varnothing$.

$E \bigcap D \in \delta(\mathcal{G})$, since $E \in \delta_D$. $X \setminus D \in \delta(\mathcal{G})$ since $D \in \delta(\mathcal{G})$, which means $(E \bigcap D) \bigcup (X \setminus D) \in \delta(\mathcal{G})$. $(X \setminus E) \bigcap D \in \delta(\mathcal{G})$, hence $X \setminus E \in \delta_D$

(**Own**)

$$\begin{tikzcd}
A \in \delta_D \arrow[rr, "\text{Goal}"] \arrow[d] & & X \setminus A \in \delta_D & \\
A \bigcap D \in \delta(\mathcal{G}) \text{ and } D \in \delta(\mathcal{G}) \arrow[d] & & & (X \setminus A) \bigcap D \in \delta(\mathcal{G}) \arrow[ul] \\
(A \bigcap D) \bigsqcup (X \setminus D) \arrow[d, equal] & & & \\
X \setminus ((X \setminus (A \bigcap D)) \bigcap D) \arrow[rr] & & (X \setminus (A \bigcap D)) \bigcap D \arrow[uur, equal] &
\end{tikzcd}$$

(3) Let $\{E_n\}_{n \in \mathbb{N}}$ be elements in δ_D which are pairwisely disjoint. $(\bigsqcup_n E_n) \bigcap D = \bigsqcup_n (E_n \bigcap D)$. $E_n \bigcap D$ are in $\delta(\mathcal{G})$, the disjoint union belongs to $\delta(\mathcal{G})$, thus $(\bigsqcup_n E_n) \bigcap D \in \delta(\mathcal{G})$, so $\bigsqcup_n E_n \in \delta_D$.

Note that $\forall G \in \mathcal{G}$, $\mathcal{G} \subseteq \delta_G$, hence $\delta(\mathcal{G}) \subseteq \delta_G$, which actually means $\forall F \in \delta(\mathcal{G})$, $\mathcal{G} \subseteq \delta_F$, thus $\delta(\mathcal{G}) \subseteq \delta_F$. $\forall (A, B) \in \delta(\mathcal{G})^2$, since $\delta(\mathcal{G}) \subseteq \delta_B$, $A \in \delta_B$ thus $A \bigcap B \in \delta(\mathcal{G})$ $\square$

Remark

Let $(X, \sum_X)$ and $(Y, \sum_Y)$ be two measurable spaces. We want to show that $\forall E \in \sum_X \otimes \sum_Y$, $\forall (x,y) \in X \times Y$, $E_x \in \sum_Y$ and $E_y \in \sum_X$.

In fact, fix y, let $\mathcal{F} = \{E \subseteq X \times Y | E_y \in \sum_X\}$. We need to show that $\sum_X \times \sum_Y \subseteq \mathcal{F}$ and $\mathcal{F}$ is a σ-algebra.

Let $A \times B \in \mathcal{F}$, then $(X \times Y) \setminus (A \times B) = ((X \setminus A) \times Y) \bigcup (A \setminus (Y \setminus B))$. If $y \in B$, then $A_y \in \sum_X$, thus $((X \times Y) \setminus (A \times B))_y = X \setminus A \in \sum_X$, if $y \notin B$, then $((X \times Y) \setminus (A \times B))_y = X \in \sum_X$. Other conditions is easy to check.

9.5.5 Theorem

Let $(X, \sum_X, \mu)$ and $(Y, \sum_Y, \nu)$ be two $\sigma-$finite measure spaces. Then for any $E \in \sum_X \otimes \sum_Y$, the function $f_E : X \to \mathbb{R} \bigcup \{+\infty\}$, $x \mapsto \nu(E_x)$ and $g_E : Y \to \mathbb{R} \bigcup \{+\infty\}$, $y \mapsto \mu(E_y)$ are respectively $\sum_X$-measurable and $\sum_Y$-measurable.

Proof

We do the proof only for f_E.

Assume that ν is finite measure ($\nu(Y) < +\infty$). $\mathcal{F} = \{E \in \sum_X \otimes \sum_Y | f_E$ is measurable$\}$. We want $\mathcal{F} = \sum_X \otimes \sum_Y$.

Let $S_1 \in \sum_X$ and $S_2 \in \sum_Y$. $(S_1 \times S_2)_x = \begin{cases} S_2 & \text{if } x \in S_1 \\ \varnothing & \text{otherwise} \end{cases}$. $f_{S_1 \times S_2}(x) = \nu((S_1 \times S_2)_x) = \nu(S_2) \mathbb{1}_{S_1}(x)$. Since $S_1 \in \sum_X$, $\mathbb{1}_{S_1}$ is measurable, also $y \equiv \nu(S_2)$ is measurable then $f_{S_1 \times S_2}$ is measurable. $\forall S_1 \in \sum_X$, $\forall S_2 \in \sum_Y$, $S_1 \times S_2 \in \mathcal{F}$.

Now let's show that $\mathcal{F}$ is a Dynkin system.

(1) $X \times Y \in \mathcal{F}$, taking $S_1 = X$, $S_2 = Y$.

(2) Let $D \in \mathcal{F}$, we want to show that $(X \times Y) \setminus D \in \mathcal{F}$. Note that $((X \times Y) \setminus D)_x = Y \setminus D_x$(This equation isn't trivial.). Recall that ν is a finite measure. $f_{(X \times Y) \setminus D}(x) = \nu(((X \times Y) \setminus D)_x) = \nu(Y \setminus D_x) = \nu(Y) - \nu(D_x) = \nu(Y) - f_D(x)$.$f_{(X \times Y) \setminus D}$ is measurable, so $(X \times Y) \setminus D \in \mathcal{F}$

(3) $\{D_n\}$ is a sequence of disjoint sets s.t. $D_n \in \mathcal{F}$, $\forall n$. $D = \bigsqcup_n D_n$. $f_D(x) = \nu(D_x) = \nu((\bigsqcup_n D_n)_x) = \nu(\bigsqcup_n (D_n)_x) = \sum_n \nu((D_n)_x) = \sum_n f_{D_n}(x)$, hence f_D is measurable.

$\mathcal{G} = \{S_1 \times S_2 | S_1 \in \sum_X, S_2 \in \sum_Y\} \subseteq \mathcal{F}$. $\delta(\mathcal{G}) \subseteq \mathcal{F}$, moreover $\mathcal{G}$ is closed under the intersection. $(S_1 \times T_1) \bigcap (S_2 \times T_2) = (S_1 \bigcap S_2) \times (T_1 \bigcap T_2)$, $\forall (S_1, S_2) \in \sum_X^2, \forall (T_1, T_2) \in \sum_Y^2$. So $\delta(\mathcal{G}) = \sigma(\mathcal{G}) = \sum_X \otimes \sum_Y \subseteq \mathcal{F}$

Since ν is $\sigma-$finite. $Y = \bigcup\limits_n Y_n$, $\nu(Y_n) < +\infty$. Let $F_0 = Y_0$, $F_n = Y_n \bigcap (X \setminus \bigcup\limits_{k<n} Y_k)$, $F_n \subseteq Y_n$, $\nu(F_n) < \nu(Y_n) < +\infty$.

$\{F_n\}_{n\in\mathbb{N}}$ are disjoint, in the σ-algebra and of finite measures. $Y = \bigsqcup\limits_n F_n$. Define a measure $\nu^{(n)}$ on Y for any n. $\nu^{(n)}(E) := \nu(E \bigcap F_n)$. Notice that $\nu^{(n)}(Y) = \nu(Y \bigcap F_n) = \nu(F_n) < +\infty$. Therefore, for any $n \in \mathbb{N}$, $\nu^{(n)}$ is a finite measure. Hence we have $f_E^{(n)} : X \to \mathbb{R} \bigcup \{+\infty\}$, $x \mapsto \nu^{(n)}(E_x)$. $f_E^{(n)}$ is measurable $\forall E$, $\forall n$.

$f_E(x) = \nu(E_x) = \nu(E_x \bigcap Y) = \nu(E_x \bigcap (\bigsqcup\limits_n F_n)) = \nu(\bigsqcup\limits_n (E_x \bigcap F_n)) = \sum\limits_n \nu(E_x \bigcap F_n) = \sum\limits_n \nu^{(n)}(E_x) = \sum\limits_n f_E^{(n)}(x)$. $\square$

Notation

$\{E_n\} \uparrow E$ means $\bigcup\limits_n E_n = E$ and $E_n \subseteq E_{n+1}$, $\forall n$. $\{E_n\} \downarrow E$ means $\bigcap\limits_n E_n = E$ and $E_{n+1} \subseteq E_n$, $\forall n$

9.5.6 Def

Let $(X, \mathscr{A})$ be a measurable space. We say $\mu : \mathscr{A} \to [0, +\infty]$ has lower continuity if $\forall \{E_n\} \subseteq \mathscr{A}$ s.t. $\{E_n\} \uparrow E$, $\lim\limits_{n\to+\infty} \mu(E_n) = \mu(E)$. We say μ has upper continuity if $\forall \{E_n\} \subseteq \mathscr{A}$ s.t. $\{E_n\} \downarrow E$, $\lim\limits_{n\to+\infty} \mu(E_n) = \mu(E)$.

9.5.7 Prop

Let $(X, \mathscr{A})$ be a measurable space, then $\mu : \mathscr{A} \to [0, +\infty]$ is a measure iff

(1) $\forall \{E_n\}_{n=1}^N \subseteq \mathscr{A}$ s.t. pairwisely disjoint, $\mu(\bigsqcup\limits_{n=1}^N E_n) = \sum\limits_{n=1}^N \mu(E_n)$

(2) μ has lower continuity.

Moreover, if $\mu(X) < +\infty$, μ is a measure iff it satisfies (1) and it has upper continuity.

Proof

Suppose μ is a measure then μ satisfies (1). $\forall \{E_n\} \subseteq \mathscr{A}$ s.t. $\{E_n\} \uparrow E \in \mathscr{A}$, let $F_0 := E_0$ and $F_n := E_n \setminus (\bigcup\limits_{m<n} E_m)$, then $\mu(E) = \mu(\bigsqcup\limits_n F_n) =$

$\sum_{n} \mu(F_n) = \lim_{n\to+\infty} \sum_{i=1}^{n} \mu(F_i) = \lim_{n\to+\infty} \mu(E_i)$.

Suppose μ satisfies (1) and (2). Let $\{E_n\} \subseteq \mathscr{A}$ s.t. they are disjoint pairwisely let $F_n = \bigsqcup_{i=1}^{n} E_i$, thus $F_n \uparrow \bigcup_{n} E_n$ so $\mu(\bigsqcup_{n} E_n) = \lim_{n\to+\infty} \mu(F_n) = \lim_{n\to+\infty} \sum_{i=1}^{n} \mu(E_i) = \sum_{n} \mu(E_n)$.

If $\mu(X) < +\infty$. Suppose μ is a measure, then μ satisfies (1). Let $\{E_n\} \subseteq \mathscr{A}$ s.t. $\{E_n\} \downarrow E$ notice that $(E_1 \setminus E_n) \uparrow (E_1 \setminus E)$ hence $\lim_{n\to+\infty} \mu(E_1 \setminus E_n) = \mu(E_1 \setminus E)$, since $\mu(X) < +\infty$, $\lim_{n\to+\infty} \mu(E_n) = \mu(E)$.

Suppose μ has upper continuity and satisfies (1). Let $\{E_n\} \subseteq \mathscr{A}$ s.t. they are disjoint. $0 = \mu(\varnothing) = \lim_{n\to+\infty} \mu(\bigsqcup_{m\in\mathbb{N}} E_m \setminus \bigsqcup_{i=1}^{n} E_i) = \lim_{n\to+\infty} (\mu(\bigsqcup_{m\in\mathbb{N}} E_m) - \mu(\bigsqcup_{i=1}^{n} E_i)) = \mu(\bigsqcup_{m\in\mathbb{N}} E_m) - \lim_{n\to+\infty} (\sum_{i=1}^{n} \mu(E_i)) = \mu(\bigsqcup_{m\in\mathbb{N}} E_m) - (\sum_{n} \mu(E_i))$ □

9.5.8 Theorem(Monotone Convergence)

Let $(\Omega, \mathcal{G}, \mu)$ be a σ-finite measure space. Let $\{f_n\}_{n\in\mathbb{N}} \subseteq L^1(\Omega, \mathcal{G}, \mu)^{\uparrow}$ be an increasing sequence $\forall n \in \mathbb{N}$, $f_n : \Omega \to \mathbb{R}$ is non-negative and measurable. Use f to denote the pointwise limit of f_n, then $\displaystyle\int f \mathrm{d}\mu = \lim_{n\to+\infty} \int f_n \mathrm{d}\mu$.

Proof

$f = \lim_{n\to+\infty} f_n = \sup_{N\in\mathbb{N}} \inf_{n\geq N} f_n$ is measurable, thus $\displaystyle\int f \mathrm{d}\mu$ is well defined.

Not loss generality let $\displaystyle\int f \mathrm{d}\mu = +\infty$. In fact, if $\displaystyle\int f \mathrm{d}\mu < +\infty$ which means $f \in L^1(\Omega, \mathcal{G}, \mu)$, by Dominated convergence theorem(Use DCT to denote this theorem in the following), $\displaystyle\lim_{n\to+\infty} \int f_n \mathrm{d}\mu = \int \lim_{n\to+\infty} f_n \mathrm{d}\mu$.

We argue by contradiction. Suppose $\displaystyle\lim_{n\to+\infty} \int f_n = B < +\infty$ and μ is finite measure. Because μ is finite measure and the definition of $\displaystyle\int f \mathrm{d}\mu$, one can obtain a simple mapping $s = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i} \in L^1(\Omega, \mathcal{G}, \mu)$ s.t. $s \leq f$ and $\displaystyle\int s \mathrm{d}\mu > B$. Take $g_n := \min\{s, f_n\}$ notice that $\lim_{n\to+\infty} g_n = s$. By DCT, $\displaystyle\int s \mathrm{d}\mu = \int \lim_{n\to+\infty} g_n \mathrm{d}\mu = \lim_{n\to+\infty} \int g \mathrm{d}\mu \leq \lim_{n\to+\infty} \int f_n \mathrm{d}\mu \leq B$, contradiction!

Argue for the general case and still suppose $\displaystyle\lim_{n\to+\infty} \int f_n = B < +\infty$. Since μ is σ-finite, $\exists \{A_j\} \subseteq \mathcal{G}$ s.t. $\{A_j\} \uparrow \Omega$ and $\forall j$, $\mu(A_j) < +\infty$. For each j, $\displaystyle\int_{A_j} f\mathrm{d}\mu = \lim_{n\to+\infty}\int_{A_j} f_n\mathrm{d}\mu \le \lim_{n\to+\infty}\int f_n\mathrm{d}\mu = B$. Taking $j \to +\infty$, $\displaystyle\lim_{j\to+\infty}\int_{A_j} f\mathrm{d}\mu \le B$.

Similarly, one can obtain a simple mapping $s = \sum\limits_{k=1}^{m} a_k \mathbb{1}_{C_k} \le f$ s.t. $B < \displaystyle\int s\mathrm{d}\mu = \sum_{k=1}^{m} a_k\mu(C_k)$. We have $\displaystyle\int_{A_j} f\mathrm{d}\mu \ge \int_{A_j} s\mathrm{d}\mu = \sum_{k=1}^{m} a_k\mu(A_j \bigcap C_k)$.

However, $B \ge \displaystyle\lim_{j\to+\infty}\int_{A_j} f\mathrm{d}\mu \ge \lim_{j\to+\infty}\sum_{k=1}^{m}\mu(A_j \bigcap C_k) = \sum_{k=1}^{m}\lim_{j\to+\infty}$ $\mu(A_j \bigcap C_k) = \sum\limits_{k=1}^{m}\mu(C_k) > B$, contradiction! □

9.5.9 Corollary

Let $(\Omega, \mathcal{G}, \mu)$ be a σ-finite measure space. Let $\{f_n\}_{n\in\mathbb{N}} \subseteq L^1(\Omega, \mathcal{G}, \mu)^{\uparrow}$ be a sequence and $\forall n \in \mathbb{N}$, $f_n : \Omega \to \mathbb{R}$ is a non-negative measurable mapping, then $\displaystyle\int \sum_n f_n\mathrm{d}\mu = \sum_n \int f_n\mathrm{d}\mu$.

Proof

$$\int \sum_n f_n\mathrm{d}\mu = \int \lim_{n\to+\infty}\sum_{i=1}^{n} f_n\mathrm{d}\mu = \lim_{n\to+\infty}\int\sum_{i=1}^{n} f_n\mathrm{d}\mu = \lim_{n\to+\infty}\sum_{i=1}^{n}\int f_n\mathrm{d}\mu = \sum_n \int f_n\mathrm{d}\mu$$ □

9.5.10 Prop

Let $(X, \sum_X, \mu)$ and $(Y, \sum_Y, \nu)$ be two σ-finite measure spaces. Then functions $\rho_X(E) := \displaystyle\int_X f_E(x)\mathrm{d}\mu(x)$ and $\rho_Y(E) := \displaystyle\int_Y g_E(y)\mathrm{d}\nu(y)$, $\forall E \in \sum_X \otimes \sum_Y$ define two measures on $(X \times Y, \sum_X \otimes \sum_Y)$ s.t. $\rho_X(S_1 \times S_2) = \rho_Y(S_1 \times S_2) = \mu(S_1)\nu(S_2)$, $\forall S_1 \times S_2 \in \sum_X \times \sum_Y$

Proof

We already know that f_E and g_E are measurable, so the integral make sense. And Since $f_E \ge 0$ and $g_E \ge 0$, then $\rho_i(E) \ge 0$, $\forall i \in \{X, Y\}$, $\forall E \in \sum_X \otimes \sum_Y$.

$\rho_X(\varnothing) = \displaystyle\int_X \nu(\varnothing_x)\mathrm{d}\mu(x) = 0$, similarly for $\rho_Y(\varnothing)$.

Assume $\{E_n\}_{n\in\mathbb{N}}$ is a sequence in $\sum_X \otimes \sum_Y$ of disjoint subsets. $\rho_X(\bigsqcup\limits_n E_n) =$

$\int_X \nu((\bigsqcup_n E_n)_x)\mathrm{d}\mu(x) = \int_X \nu(\bigsqcup_n((E_n)_x))\ \mathrm{d}\mu(x) = \int_X \sum_n \nu((E_n)_x)\mathrm{d}\mu(x) = \sum_n \int_X \nu((E_n)_x)\mathrm{d}\mu(x) = \sum_n \rho_X(E_n)$, similarly for ρ_Y.

$\rho_X(S_1 \times S_2) = \int_X \nu((S_1 \times S_2)_x)\mathrm{d}\mu(x) = \int_X \nu(S_2)\mathbb{1}_{S_1}(x)\mathrm{d}\mu(x) = \nu(S_2) \int_X \mathbb{1}_{S_1}(x)\mathrm{d}\mu(x) = \nu(S_2)\mu(S_1)$. For $\rho(y)$ is the same calculation.

9.5.11 Prop

Let $(X, \sum_X, \mu)$ and $(Y, \sum_Y, \nu)$ be two σ-finite measure spaces. Any measure η on $(X \times Y, \sum_X \otimes \sum_Y)$ that satisfies $\eta(S_1 \times S_2) = \mu(S_1)\nu(S_2)$, $\forall S_1 \times S_2 \in \sum_X \times \sum_Y$ is σ-finite.

Proof

Since μ and ν are σ-finite, $X = \bigcup_n E_n$, $\mu(E_n) < +\infty$, $Y = \bigcup_n F_n$, $\nu(F_n) < +\infty$.

$X \times Y = \bigcup_n E_n \times \bigcup_n F_n = \bigcup_{m,n} (E_m \times F_n)$, $\eta(E_n \times F_n) < +\infty$. $\square$

9.5.12 Prop

Let $(X, \sum)$ be a measurable space and assume $\mathcal{G} \subseteq \mathscr{P}(X)$ s.t. $\sum = \sigma(\mathcal{G})$. Assume that $\mathcal{G}$ satisfies the following conditions

(1) It is closed under finite intersection.

(2) There exists a sequence $\{G_n\}_n$ in $\mathcal{G}$ s.t. $\{G_n\} \uparrow X$.

Let μ and ν be two measures on $(X, \sum)$ s.t.

(a) $\mu(G) = \nu(G)$, $\forall G \in \mathcal{G}$.

(b) $\mu(G_n) < +\infty$, $\forall n$.

Then $\mu = \nu$.

Proof

$\mathcal{D}_n = \{E \in \sum \,|\, \mu(G_n \bigcap E) = \nu(G_n \bigcap E)\} \subseteq \sum$. We show that $\mathcal{D}_n$ is a Dynkin system for any n.

(1) $G_n \bigcap X = G_n$

(2) Assume $D \in \mathcal{D}_n$, $\mu(G_n \bigcap (X \setminus D)) = \mu(G_n) - \mu(G_n \bigcap D) = \nu(G_n) - \nu(G_n \bigcap D) = \nu(G_n \bigcap (X \setminus D))$

(3) Take $\{D_m\}_{m\in\mathbb{N}}$ in $\mathcal{D}_n$, of pairwise disjoint sets. $\mu(G_n \bigcap (\bigsqcup_m D_m)) = \mu(\bigsqcup_m (G_n \bigcap D_m)) = \sum_m \mu(G_n \bigsqcup D_m) = \sum_m \nu(G_n \bigcap D_m) = \nu(G_n \bigcap (\bigsqcup_m D_m))$

Combining (1) and (a), $\mathcal{G} \subseteq \mathcal{D}_n$. Consider $\delta(\mathcal{G})$, $\delta(\mathcal{G})$ is a σ-algebra, $\delta(\mathcal{G}) = \sigma(\mathcal{G}) = \sum$, $\delta(\mathcal{G}) \subseteq \mathcal{D}_n$, thus $\sum = \mathcal{D}_n$, $\forall n$.

For any $n \in \mathbb{N}$, $\mu(G_n \bigcap E) = \nu(G_n \bigcap E)$, $\forall E \in \sum$. Define $E_n := G_n \bigcap E \in \mathcal{G}$, E_n is increasing and $\bigcup_n (G_n \bigcap E) = (\bigcup_n G_n) \bigcap E = E$. So $\{E_n\} \uparrow E$.

Take $E \in \sum$, $\mu(E) = \lim\limits_{n\to+\infty} \mu(E_n) = \lim\limits_{n\to+\infty} \mu(G_n \bigcap E) = \lim\limits_{n\to+\infty} \nu(G_n \bigcap E) = \nu(E)$ □

9.5.13 Theorem

Let $(X, \sum_X, \mu)$ and $(Y, \sum_Y, \nu)$ be two σ-finite measure spaces. There exists a unique σ-finite measure $\mu \times \nu$ on $(X \times Y, \sum_X \otimes \sum_Y)$ s.t. $(*) (\mu \times \nu)(S_1 \times S_2) = \mu(S_1)\nu(S_2)$, $\forall S_1 \times S_2 \in \sum_X \times \sum_Y$ and moreover we have that $(\mu \times \nu)(E) = \int_X f_E \mathrm{d}\mu = \int_Y g_E \mathrm{d}\nu$, $\forall E \in \sum_X \otimes \sum_Y$.

Proof

Assume that η and η' are two measures on the product satisfying the condition.

Let $\mathcal{G} = \{S_1 \times S_2 | (S_1, S_2) \in \sum_X \times \sum_Y\}$, $\sigma(\mathcal{G}) = \sum_X \otimes \sum_Y$.

$\mathcal{G}$ is stable under finite intersection.

Since μ and ν are σ-finite, we can construct $\{E_n\} \uparrow X$, $\{F_n\} \uparrow Y$. $X \times Y = \bigcup_{n,m} (E_n \times F_m) = \bigcup_{i_k} (E_{i_k} \times F_{i_k})$. $G_k := E_{i_k} \times F_{i_k}$, $\{G_k\} \uparrow X \times Y$.

By $(*)$, $\eta(G_k) = \eta'(G_k)$. Since we have constructed a measure satisfies the condition, by the uniqueness we obtain $(\mu \times \nu)(E) = \int_X f_E \mathrm{d}\mu = \int_Y g_E \mathrm{d}\nu$, $\forall E \in \sum_X \otimes \sum_Y$. □

9.5.14 Corollary

On $\mathbb{R}^n$ we can define a unique measure $\lambda^{(n)}$ as product of the lebesgue measure on $\mathbb{R}$. This is called the lebesgue measure on $\mathbb{R}^n$.

9.6 Integrals in $\mathbb{R}^n$

Notation

$\mathcal{O}^n = \{\text{set of open sets of } \mathbb{R}^n\}$

$\mathscr{C}^n = \{\text{set of closed sets of } \mathbb{R}^n\}$

$\mathcal{R}^n = \{\text{set of compact sets of } \mathbb{R}^n\}$

$\mathcal{I}^n_{ho} = \{\text{set of all half-open rectangles in } \mathbb{R}^n\}$

$\mathcal{I}^m_{ho,rat} = \{\text{set of all half-open rectangles of } \mathbb{R}^n, \text{ with rational end points}\}$

Exercise

$\mathscr{B}(\mathbb{R}^n) := \sigma(\mathcal{O}^n) = \sigma(\mathscr{C}^n) = \sigma(\mathcal{R}^n) = \sigma(\mathcal{I}^n_{ho}) = \sigma(\mathcal{I}^n_{ho,rat})$

Answer

$\sigma(\mathcal{O}^n) = \sigma(\mathscr{C}^n)$ and $\sigma(\mathcal{I}^n_{ho}) = \sigma(\mathcal{I}^n_{ho,rat})$ are trivial.

Then We prove $\mathcal{R}^n = \mathcal{O}^n$. $\mathcal{R}^n \subseteq \mathcal{O}^n$ is trivial. $\forall U \in \mathcal{R}^n$, define $K_m := \{x \in U | \inf\limits_{p \in \mathbb{R}^n \setminus U} d(x,p) \geq \frac{1}{m} \text{ and } \|x\| \leq m\}$. Suppose $\exists x \in U$ s.t. $\forall n \in \mathbb{N}$, $\inf\limits_{p \in \mathbb{R}^n \setminus U} d(x,p) < \frac{1}{n}$, which means $\forall r \in \mathbb{R}_{>0}$, $B(x,r) \not\subseteq U$, contradiction! Therefore, $\bigcup\limits_{m \in \mathbb{N}_{\geq 1}} K_m = U$. We also needs to show that K_m is compact, $\forall m$. Let $\{x_n\}_{n \in \mathbb{N}} \subseteq K_m$ be a sequence, notice that if $\lim\limits_{n \to +\infty} x_n = x$, then $x \in K_m$, thus K_m is closed also K_m is bounded, hence K_m is compact.

Next we want to prove $\mathcal{O}^n = \mathcal{I}^n_{ho,rat}$. Let $U \in \mathcal{R}^n$, take $x = (x_1, \cdots, x_n) \in U$, by definition $\exists r > 0$ s.t. $B(x,r) \subseteq U$, let $A_x = \prod\limits_{i=1}^{n} [x_i - \frac{1}{2}\sqrt[n]{r}, x_i + \frac{1}{2}\sqrt[n]{r}[\subseteq U$ and $A_x \in \mathcal{I}^n_{ho}$, hence $\bigcup\limits_{x \in U} A_x = U$. Take $\xi_{x_i} \in \mathbb{Q}$ that satisfies $x_i - \xi_{x_i} \in \mathbb{Q}$ and $x_i - \frac{1}{2}\sqrt[n]{r} < x_i - \xi_{x_i} < x_i$, then $B_x := \prod\limits_{i=1}^{n} [x_i - \xi_{x_i}, x_i + \xi_{x_i}[\subseteq U$. Notice that $\mathcal{I}^n_{ho,rat}$ is countable, take all $B_n \in \mathcal{I}^n_{ho,rat}$ that satisfies $B_n \subseteq U$, then $\bigcup\limits_{n \in \mathbb{N}} B_n = U$. Let $\prod\limits_{i=1}^{n} [a_i, b_i[\in \mathcal{I}^n$, then $\bigcap\limits_{n \in \mathbb{N}_{\geq 1}} \prod\limits_{i=1}^{n}]a_i + \frac{1}{n}, b_i[= \prod\limits_{i=1}^{n} [a_i, b_i[$, and it's easy to check $\forall n$, $\prod\limits_{i=1}^{n}]a_i + \frac{1}{n}, b_i[$ is an open set.(Another way to prove the compact set part is to notice that $\sigma(\mathcal{I}^n_{ho,rat}) = \sigma(\{\prod\limits_{i=1}^{n} [a_i, b_i] | a_i, b_i \in \mathbb{R}^2\}) \subseteq \sigma(\mathcal{R}^n) \subseteq \sigma(\mathscr{C}^n))$ $\square$

Reminder

Let $(X, \sum_X)$ and $(Y, \sum_Y)$ be two measurable spaces. Moreover assume $\sum_Y = \sigma(\mathcal{G})$, where $\mathcal{G} \subseteq \mathscr{P}(Y)$. A function $f : X \to Y$ is measurable iff $f^{-1}(S) \in \sum_X$, $\forall S \in \mathcal{G}$

Exercise

Try to prove it by your own.

Answer

Let $(\mathscr{A}_1, \mathscr{A}_2) \subseteq \mathscr{P}(X) \times \mathscr{P}(Y)$, define $f_*(\mathscr{A}_1) := \{B | f^{-1}(B) \in \mathscr{A}_1\}$ and $f^{-1}(\mathscr{A}_2) = \{f^{-1}(B) | B \in \mathscr{A}_2\}$.

f is measurable iff $\sigma(\mathcal{G}) = \sum_Y \subseteq f_*(\sum_X)$ iff $\mathcal{G} \subseteq f_*(\sum_X)$ iff $f^{-1}(\mathcal{G}) \subseteq \sum_X$ $\square$

Exercise

$f : \mathbb{R}^n \to \mathbb{R}^m$, if f is continuous then f is measurable with respect to the lebesgue measure.

Answer

A direct corollary of the previous exercise. $\square$

9.6.1 Def

Let $(X, \sum_X, \mu)$ be a measure space, and let $(Y, \sum_Y)$ be a measurable space. If $f : X \to Y$ is a measurable function, then define: $f_{*\mu}(E) := \mu(f^{-1}(E))$, $\forall E \in \sum_Y$.

This is a measure on Y, called the pushforward of μ through f.

Exercise

Check $f_{*\mu}$ is a measure.

Answer

$f_{*\mu}(\varnothing) = \mu(f^{-1}(\varnothing)) = 0$

Let $B_n \subseteq \sum_Y$ s.t. B_n disjoint pairwisely. $f_{*\mu}(\bigsqcup_n B_n) = \mu(f^{-1}(\bigsqcup_n B_n)) = \mu(\bigsqcup_n f^{-1}(B_n)) = \sum_n f_{*\mu}(B_n)$ $\square$

9.6.2 Prop

Let $p \in \mathbb{R}^n$ and let $E \in \mathscr{B}(\mathbb{R}^n)$, then $\lambda^n(E+p) = \lambda^n(E)$

Proof

$p = (p_1, \cdots, p_n)$. Consider the translation $\tau_p : \mathbb{R}^n \to \mathbb{R}^n$, $x \mapsto x - p$ this mapping is continuous, so measurable. We consider $\tau_{p_{*\lambda^n}} =: \lambda_p^n$

Let's show $\lambda_p^n = \lambda^n$. $\lambda_p^n(\prod_{i=1}^n [a_i, b_i[) = \lambda^n(\tau_p^{-1}(\prod_{i=1}^n [a_i, b_i[)) = \lambda^n(\prod_{i=1}^n [a_i + p_i, b_i + p_i[) = \prod_{i=1}^n (b_i - a_i)$.

By the uniqueness of the product measure, we have $\lambda_p^n = \lambda^n$. $\lambda^n(E + p) = \lambda^n(\tau_p^{-1}(E)) = \lambda_p^n(E) = \lambda(E)$ $\square$

9.6.3 Theorem(Fubini-Tonelli)

Let $(X, \sum_X, \mu)$ and $(Y, \sum_Y, \nu)$ be two σ-finite measure spaces. Let $(X \times Y, \sum_X \otimes \sum_Y, \mu \times \nu)$ be the product space. Let $f : X \times Y \to \mathbb{R}$ be a measurable function. Then

$$\int_{X\times Y} |f| \mathrm{d}(\mu\times\nu) = \int_Y (\int_X |f(x,y)| \mathrm{d}\mu(x)) \mathrm{d}\nu(x) = \int_X (\int_Y |f(x,y)| \mathrm{d}\nu(y)) \mathrm{d}\mu(x)$$

Proof

We can assume that $f \geq 0$.

$\forall (x,y) \in X \times Y$, define $f_x : Y \to \mathbb{R}$, $x \mapsto f(x,y)$ and $f_y : X \to \mathbb{R}$. Then we have to show $\int_{X\times Y} f \mathrm{d}(\mu \times \nu) = \int_Y (\int_X f_y \mathrm{d}\mu) \mathrm{d}\nu = \int_X (\int_Y f_x \mathrm{d}\nu) \mathrm{d}\mu$.

Step 1: f_x and f_y are measurable. Let's do it for f_y. We have to show that $D \in \mathscr{B}(\mathbb{R})$, then $f_y^{-1}(D) \in \sum_X$.

$f_y^{-1}(D) = \{x \in X | f(x,y) \in D\} = \{x \in X | (x,y) \in f^{-1}(D)\} = (f^{-1}(D))_y$. We have shown that if $E \in \sum_X \otimes \sum_Y$, then $E_y \in \sum_X$.

Step 2: Consider the function $G : Y \to \mathbb{R} \bigcup \{+\infty\}$, $y \mapsto \int_X f_y \mathrm{d}\mu$ and $F : X \to \mathbb{R} \bigcup \{+\infty\}$, $x \mapsto \int_Y f_x \mathrm{d}\nu$. They are both measurable.

Let's do the proof for G. Assume that $f = \mathbb{1}_E$ for $E \in \sum_X \otimes \sum_Y$. $(\mathbb{1}_E)_y(x) = \mathbb{1}_E(x,y) = 1$ iff $(x,y) \in E$ iff $x \in E_y$ iff $\mathbb{1}_{E_y}(x) = 1$. This chains of implications shows $(\mathbb{1}_E)_y = \mathbb{1}_{E_y}$. Hence $G(y) = \int_X (\mathbb{1}_E)_y \mathrm{d}\mu = \int_X \mathbb{1}_{E_y} \mathrm{d}\mu = \mu(E_y)$. And we have prove that such functions are measurable.

Now assume $f = \sum_{k=1}^n a_k \mathbb{1}_{E_k}$, $E_1, E_n \in \sum_X \otimes \sum_Y$, $a_1, \cdots, a_n \in \mathbb{R}_{\geq 0}$.

$f_y = \sum_{k=1}^{n} a_k \mathbb{1}_{(E_k)_y}$. $G(y) = \int_X f_y \mathrm{d}\mu = \sum_{k=1}^{n} a_k \int_X \mathbb{1}_{(E_k)_y} \mathrm{d}\mu = \sum_{k=1}^{n} a_k \mu((E_k)_y)$, G is measurable.

Now assume f measurable. $\exists f_n$ simple function s.t. $0 \le f_i \le f_j$ if $i < j$ and $\lim_{n\to+\infty} f_n = f$ pointwise. Since $\{f_n\}$ is increasing, also $(f_n)_y$ is increasing and $\lim_{n\to+\infty}(f_n)_y = f_y$ pointwise. By the monotone convergence theorem, $G(y) = \int_X f_y \mathrm{d}\mu = \int_X \lim_{n\to+\infty}(f_n)_y(x)\mathrm{d}\mu = \lim_{n\to+\infty}\int_X (f_n)_y \mathrm{d}\mu$.

Consider $g_n : Y \to \mathbb{R}$, $y \mapsto g_n(y) := \int_X (f_n)_y \mathrm{d}\mu$. Since f_n is simple, by the previous claim in step 2, we know that g_n is measurable. $G(y) = \lim_{n\to+\infty} g_n(y)$, this implies G is measurable.

step 3: We show the theorem for $f = \mathbb{1}_E$.

$\int_{X\times Y} f\mathrm{d}(\mu\times\nu) = (\mu\times\nu)(E)$, $f_y = \mathbb{1}_{E_y}$. $\int_X f_y \mathrm{d}\mu = \mu(E_y)$, $\int_X (\int_Y f_x \mathrm{d}\mu)\mathrm{d}\nu = \int_X \nu(E_x)\mathrm{d}\mu = \int_X f_E \mathrm{d}\mu = (\mu\times\nu)(E)$

step 4: We show the theorem for $f = \sum_{k=0}^{n} a_k \mathbb{1}_{E_k}$, $E_1, \cdots, E_n \in \sum_X \otimes \sum_Y$.

$\int_X f_y \mathrm{d}\mu = \sum_{k=1}^{n} a_j \mu((E_k)_y)$, $a_k \ge 0$. $\int_Y (\int_X f_y \mathrm{d}\mu)\mathrm{d}\nu = \int_Y (\sum_{k=1}^{n} a_k \mu((E_k)_y))\mathrm{d}\nu = \sum_{k=1}^{n} a_k \int_Y \mu((E_k)_y)\mathrm{d}\nu = \sum_{k=1}^{n} a_k (\mu\times\nu)(E_k)$

$\int_{X\times Y} f\mathrm{d}(\mu\times\nu) = \int_{X\times Y}(\sum_{k=1}^{n} a_k \mathbb{1}_{E_k})\mathrm{d}(\mu\times\nu) = \sum_{k=1}^{n} a_k \int_{X\times Y} \mathbb{1}_{E_k}\mathrm{d}(\mu\times\nu) = \sum_{k=1}^{n} a_k (\mu\times\nu)(E_k)$

step 5: We prove the theorem for f measurable.

$\exists\{f_n\}$ simple functions increasing and $\lim_{n\to+\infty} f_n = f$ pointwise. We induce a sequence $\{(f_n)_y\}$ non-negative, simple, increasing, $\lim_{n\to+\infty}(f_n)_y = f_y$ pointwise.

$g_n(y) = \int_X (f_n)_y \mathrm{d}\mu$, apply MCT, $\lim_{n\to+\infty} g_n(y) = \lim_{n\to+\infty}\int_X (f_n)_y \mathrm{d}\mu = \int_X \lim_{n\to+\infty}(f_n)_y \mathrm{d}\mu = \int_X f_y \mathrm{d}\mu$. $\{g_n\}$ is measurable, increasing and $\lim_{n\to+\infty} g_n(y) = G(y)$ pointwise.

$\int_{X\times Y} f_n \mathrm{d}(\mu\times\nu) = \int_Y (\int_X (f_n)_y \mathrm{d}\mu)\mathrm{d}\nu = \int_Y g_n \mathrm{d}\nu$. $\int_Y (\int_X f_y \mathrm{d}\mu)\mathrm{d}\nu =$

$\displaystyle\int_Y \lim_{n\to+\infty} g_n(y)\mathrm{d}\nu = \lim_{n\to+\infty}\int_Y g_n(y)\mathrm{d}\nu = \lim_{n\to+\infty}\int_{X\times Y} f_n\mathrm{d}(\mu\times\nu) = \int_{X\times Y} f\mathrm{d}(\mu\times\nu)$. □

9.6.4 Corollary

Fubini-Tonelli theorem holds for $f \in L^1(X\times Y, \sum_X \otimes \sum_Y, \mu\times\nu)$.

Proof

Remember that $f \in L^1(X\times Y, \sum_X\otimes\sum_Y, \mu\times\nu)$ iff $\displaystyle\int_{X\times Y}|f| < +\infty$ iff $\displaystyle\int_{X\times Y} f^+ < \infty$ and $\displaystyle\int_{X\times Y} f^- < +\infty$(Easy to check if f is measurable then $f^+ := f\vee 0$ and $f^- := -f\vee 0$ are measurable.). Apply Fubini-Tonelli to f^+ and f^-.

Computations of double Integrals

$f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$, $f \in L^1(U, \mathscr{B}(\mathbb{R}^2), \lambda^2)$, $E \in \mathscr{B}(\mathbb{R}^2)$ and $E \subseteq U$.

Let $E = [a,b]\times[c,d]$. $\displaystyle\int_E f(x,y)\mathrm{d}\lambda^2 = \int_E f(x,y)\mathrm{d}x\mathrm{d}y = \int_a^b(\int_c^d f(x,y)\mathrm{d}y)\mathrm{d}x.$

Let $E = \{(x,y) \in \mathbb{R}^2 | a \le x \le b, \alpha(x) \le y \le \beta(x)\}$. α and β are functions. Induce E in the rectangle $R = [a,b]\times I$.

Define $\tilde{f}(x,y) = \begin{cases} f(x,y) & \text{if } (x,y) \in E \\ 0 & \text{otherwise} \end{cases}$

$$\int_E f\mathrm{d}\lambda^2 = \int_R \tilde{f}\mathrm{d}\lambda^2 = \int_a^b \int_I \tilde{f}(x,y)\mathrm{d}y\mathrm{d}x = \int_a^b \int_{\alpha(x)}^{\beta(x)} f(x,y)\mathrm{d}y\mathrm{d}x$$

Exercise

In $\displaystyle\int_E f(x,y)\mathrm{d}\lambda^2 = \int_E f(x,y)\mathrm{d}x\mathrm{d}y$, we use the fact that $\sigma_E = \{U\bigcap E | U \in \mathscr{B}(\mathbb{R}^2)\} = \mathscr{B}(E)$, $\forall E \in \mathscr{B}(\mathbb{R}^2)$ implictly, prove it.

Answer

Easy to check σ_E is a σ-algebra containing $\mathcal{O}_E$, thus $\mathscr{B}(E) \subseteq \sigma_E$.

Define $\mathscr{A} = \{U \subseteq \mathbb{R}^2 | U\bigcap E \in \mathscr{B}(E)\}$. The goal is to prove $\sigma(\mathcal{O}^2) \subseteq \mathscr{A}$. Easy to prove $\mathscr{A}$ forms a σ-algebra and $\mathcal{O}^2 \subseteq \mathscr{A}$. □

Example

(1) $E = \{(x,y) \in \mathbb{R}^2 | 0 \le x \le \frac{\pi}{2}, 0 \le y \le x\}$.

$$\iint_E \sin(x+y)\mathrm{d}x\mathrm{d}y = \int_0^{\frac{\pi}{2}}\int_0^x \sin(x+y)\mathrm{d}y\mathrm{d}x = \int_0^{\frac{\pi}{2}}[-\cos(x+y)]_0^x\mathrm{d}x =$$
$$\int_0^{\frac{\pi}{2}}(-\cos(2x)+\cos(x))\mathrm{d}x = 1$$

(2) $E = \{(x,y)\in\mathbb{R}^2 | 1\le y\le 2, y\le x\le y^2\}$

$$\iint_E e^{\frac{x}{y}}\mathrm{d}x\mathrm{d}y = \int_1^2\int_y^{y^2} e^{\frac{x}{y}}\mathrm{d}x\mathrm{d}y = \int_1^2 [ye^{\frac{x}{y}}]_y^{y^2}\mathrm{d}y = \int_1^2 (ye^y - ye)\mathrm{d}y =$$
$$e^2 - \frac{3}{2}e.$$

Counter example

$$f(x,y) = \frac{x^2-y^2}{(x^2+y^2)^2},\ R = [0,1]^2.$$

$$\int_0^1(\int_0^1 f(x,y)\mathrm{d}y)\mathrm{d}x = \int_0^1 [\frac{y}{x^2+y^2}]_0^1\mathrm{d}x = \int_0^1 \frac{1}{1+x^2}\mathrm{d}x = \arctan x|_0^1 =$$
$$\frac{\pi}{4}$$

$$\int_0^1(\int_0^1 f(x,y)\mathrm{d}x)\mathrm{d}y = \int_0^1 [-\frac{x}{x^2+y^2}]_0^1\mathrm{d}y = -\int_0^1 \frac{1}{1+y^2}\mathrm{d}y = [-\arctan y]_0^1 =$$
$$-\frac{\pi}{4}$$

We now check that the problem is that $f \notin L^1(R, \mathscr{B}(R), \lambda^2)$.

$f(x,y) \ge 0$ if $1 \ge x^2 \ge y^2 \ge 0$ which is equivlant to $0 \le x \le 1$ and $0 \le y \le x$.

$$\int_R |f(x,y)|\mathrm{d}x\mathrm{d}y = \int_0^1\int_0^1 |f(x,y)|\mathrm{d}x\mathrm{d}y = 2\int_0^1\int_0^x f(x,y)\mathrm{d}y\mathrm{d}x = 2\int_0^1\int_0^x$$
$$\frac{x^2-y^2}{(x^2+y^2)^2}\mathrm{d}y\mathrm{d}x = 2\int_0^1 [\frac{y}{x^2+y^2}]_0^x\mathrm{d}x = \int_0^1 \frac{1}{x}\mathrm{d}x = +\infty!$$

Notation

Let $U \subseteq \mathbb{R}^2$ be an open set. $C_0^0(U)$ denotes the set of continuous functions $f : U \to \mathbb{R}^2$ that they have compact support. Compact support means the set $\mathrm{Supp}(f) = \{x \in U | f(x) \neq 0\}$ is compact.

Remark

Let U be an open set of $\mathbb{R}^n$. Functions of $C_0^0(U)$ are measurable, since they are continuous. Let $g : U \to \mathbb{R}^m$ be a differentiable function.

Then the Jacobian of g is the matrix $\boldsymbol{J_g} = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$ where $g(x_1, \cdots, x_n) = \begin{pmatrix} g_1(x_1, \cdots, x_n) \\ \vdots \\ g_m(x_1, \cdots, x_n) \end{pmatrix}$

The Jacobian is related to the differential of g. In fact, $\forall p \in U$, $\mathrm{d}g|_p : \mathbb{R}^n_p \to \mathbb{R}^m_{g(p)}$, $v \mapsto \boldsymbol{J_g|_p(v)}$

9.6.5 Theorem(Change of variables for the Lebesgue integral)

Let $V \subseteq \mathbb{R}^n$ be an open set, and let $\varphi : V \to \mathbb{R}^n$ be a C^1-diffeomorphism, then $\displaystyle\int_{\varphi(V)} f \mathrm{d}\lambda^n = \int_V (f \circ \varphi)|\det(\boldsymbol{J_\varphi})|\mathrm{d}\lambda^n$, $\forall f \in C^0_0(\varphi(V))$

Proof

Difficult. (At the end of the course.)

Remark

The theorem can be generalize to a bigger class of functions. In fact, it is possible to show that this theorem holds whenever one of the two integrals exists.

Example

$f(x,y) = \dfrac{1}{1+x^2+y^2}$, $A = \{(x,y) \in \mathbb{R}^2 | 0 < y < \sqrt{3}, 1 < x^2+y^2 < 4\}$

$\varphi(\rho, \theta) : [0, +\infty[\times [0, 2\pi[\to \mathbb{R}^2$, $(\rho, \theta) \mapsto (\rho\cos\theta, \rho\sin\theta)$ is C^1-diffeomorphism.

$\tilde{A} = \varphi^{-1}(A) = \{(\rho, \theta) | 1 < \rho < 2, 0 < \theta < \frac{\pi}{3}\}$

Use the theorem $\displaystyle\int_A f \mathrm{d}x\mathrm{d}y = \int_{\tilde{A}} (f \circ \varphi)|\det(\boldsymbol{J_\varphi})|\mathrm{d}\rho\mathrm{d}\theta = \int_1^2 \int_0^{\frac{\pi}{3}} \frac{\rho}{1+\rho^2}\mathrm{d}\rho\mathrm{d}\theta = \int_0^{\frac{\pi}{3}} [\frac{1}{2}\ln(1+\rho^2)]_1^2 \mathrm{d}\theta = \frac{\pi}{6}\ln(\frac{5}{2})$

9.7 Closed and exact forms

Reminder

Let U be an open set of $\mathbb{R}^n$, $\omega = \sum a_i \mathrm{d}x_i \in \Omega^1(U)$, $\gamma : [a,b] \to U$ and γ is of class C^1.

Then we define $\gamma(t) = (x_1(t), \cdots, x_n(t))$ and $\int_\gamma \omega$. The fact that γ is differentiable is crucial since we need $\gamma^*\omega$.

Let $\varphi : [c,d] \to [a,b]$, $\tau \mapsto t$ be a C^1-diffeomorphism that satisfies $\varphi(c) = a$ and $\varphi(d) = b$. $t = \varphi(\tau)$ and $\tau = \varphi^{-1}(t)$.

$$\int_\gamma \omega := \int_a^b (\sum_i a_i(\gamma(t)) \frac{\mathrm{d}x_i}{\mathrm{d}t}) \mathrm{d}t = \int_a^b (\sum_i a_i(\gamma(t)) \frac{\mathrm{d}x_i}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}t}) \mathrm{d}t = \int_c^d (\sum_i a_i(\gamma(\varphi(\tau))) \frac{\mathrm{d}t}{\mathrm{d}\tau}\frac{\mathrm{d}x_i}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}t}) \mathrm{d}\tau = \int_a^b (\sum_i a_i(\gamma(\varphi(\tau))) \frac{\mathrm{d}x_i}{\mathrm{d}\tau}) \mathrm{d}\tau = \int_{\gamma\circ\varphi} \omega.$$

We call $\gamma\circ\varphi$ a reparametrization of the curve γ, with the C^1-diffeomorphism φ. We say that φ preserves the orientation if $\varphi' > 0$, we say that φ reverses the orientation if $\varphi' < 0$. The φ in the above example preserves the orientation.

Notice that if φ preserves the orientation, then $\displaystyle\int_\gamma \omega = \int_{\gamma\circ\varphi} \omega$. If φ reverses the orientation, then $\displaystyle\int_\gamma \omega = -\int_{\gamma\circ\varphi} \omega$.

Consider a special $\varphi : [a,b] \to [a,b]$, $\tau \mapsto -\tau + a - \frac{b-a}{a-b} b$, $a < b$. $\varphi(a) = b$ and $\varphi(b) = a$, $\varphi' < 0$. $-\gamma = \gamma \circ \varphi$.

Notation

Let $U \subseteq \mathbb{R}^n$ s.t. U is open and connected, $\gamma : [a,b] \to U$.

We say that γ is a curve if γ is of class C^1 piecewisely (We can integrate 1-forms along curves), a curve is closed if $\gamma(a) = \gamma(b)$.

We say that γ is a path if γ is of class C^0. A closed path is called loop.

9.7.1 Def

Let $\omega = \sum_i a_i \mathrm{d}x_i \in \Omega^1(U)$. We say that ω is closed if $\mathrm{d}\omega = 0$. We say that ω is exact in $V \subseteq U$ if there exists a mapping $f : V \to \mathbb{R}$ s.t. $\omega = \mathrm{d}f$ in V.

9.7.2 Lemma

Let $\varnothing \neq U \subseteq \mathbb{R}^n$ be a connected open set. Then any two points of U can be joined by a piecewise C^1 curve.

Proof

Take $a \in U$. Let $H \subseteq U$ the set of points that can be joined to a with a piecewise C^1 curve. Take $K := U \setminus H$. Take $x \in H$, $\exists B(x, \varepsilon) \subseteq U$, since U is open. Any two points in $B(x, \varepsilon)$ can be joined with a segment. Take any $y \in B(x, \varepsilon)$ this y can be joined to a with a piecewise C^1 curve. This means that H is open. For the same reason K is open, hence $H = U$. $\square$

Notation

If γ is a curve, then we have defined $-\gamma$ using reparametrization.

Let $\gamma_1 : [a, b] \to U$, $\gamma_2 : [b, c] \to U$, assume $\gamma_1(b) = \gamma_2(b)$. Then we define $\gamma_1 \bigsqcup \gamma_2 : [a, c] \to U$, $\gamma_1 \bigsqcup \gamma_2(t) := \gamma_i(t)$ if t is in the domain of γ_i, $i \in \{1, 2\}$. By definition of the line integral $\displaystyle\int_{\gamma_1 \bigsqcup \gamma_2} \omega = \int_{\gamma_1} \omega + \int_{\gamma_2} \omega$.

9.7.3 Theorem

The following statements are equivalent

- ω is exact in a connected open set $V \subseteq U$.
- $\displaystyle\int_\gamma \omega$ depends only on the end points of γ for any γ.(Let $\gamma_1 : [a, b] \to V$, $\gamma_2 : [c, d] \to V$ s.t. $\gamma_1(a) = \gamma(c)$, $\gamma_1(b) = \gamma_2(d)$ then $\displaystyle\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$.)
- $\displaystyle\int_\gamma \omega = 0$ for all closed curves $\gamma \subseteq V$. ($\gamma : [a, b] \to V$ is closed, if $\gamma(a) = \gamma(b)$)

Proof

From (1) to (2) : $\omega = \mathrm{d}f$ in V. $\displaystyle\int_\gamma \omega = \int_\gamma \mathrm{d}f = \int_a^b \gamma^*(\mathrm{d}f) = f(\gamma(b)) - f(\gamma(a))$

From (2) to (3) : Take a closed curve, s.t. $\gamma(a) = \gamma(b) = p \in U$. Take $\gamma' \equiv p$. $\displaystyle\int_\gamma \omega = \int_{\gamma'} \omega = 0$

From (3) to (2) : Consider $\gamma : [a, b] \to V$, $\delta : [c, d] \to V$ s.t. $\gamma(a) = \delta(c)$ and $\gamma(b) = \delta(d)$.

Reparametrize δ, to obtain $\tilde{\delta} : [b, c'] \to V$ s.t. $\tilde{\delta}(b) = \delta(d) = \gamma(b)$, $\tilde{\delta}(c') = \gamma(a) = \delta(c)$. $\delta \bigsqcup \tilde{\delta}$ is a closed curve, $0 = \int_{\gamma \bigsqcup \tilde{\delta}} \omega = \int_{\gamma} \omega + \int_{\tilde{\delta}} \omega$, hence $\int_{\delta} \omega = \int_{\gamma} \omega$

From (2) to (1) : Fix $p \in V$. For any $x \in V$ be the lemma, there exists a curve γ_x piecewise C^1 that connects p and x. Define $f : V \to \mathbb{R}$, $x \mapsto \int_{\gamma_x} \omega$ this is well defined.

We want to show $\omega = \mathrm{d}f$. $\mathrm{d}f = \sum_i \frac{\partial f}{\partial x_i} \mathrm{d}x_i$, then we want to show $\frac{\partial f}{\partial x_i} = a_i$.

Define $\delta_i(t) = x + te_i$, $e_i = (0, \cdots, 0, 1, 0, \cdots, 0)$, $t \in]0, \varepsilon[$. Since V is open, if ε is small enough, the entire curve $\delta_i \subseteq V$.

$\frac{\partial f}{\partial x_i}(x) = \lim_{t \to 0} \frac{1}{t}(f(x+te_i) - f(x)) = \lim_{t \to 0} \frac{1}{t}(\int_{\gamma_x \bigsqcup \delta_i} \omega - \int_{\gamma_x} \omega) = \lim_{t \to 0} \frac{1}{t} \int_{\delta_i} \omega =$ $\lim_{t \to 0} \frac{1}{t} \int_0^t \delta_i^* \omega = \lim_{t \to 0} \frac{1}{t} \int_0^t a_i(\delta_i(s)) \mathrm{d}s = a_i(\delta_i(0)) = a_i(x)$ □

9.7.4 Theorem(Poincaré lemma)

Let U be an open set of $\mathbb{R}^n$, let $\omega = \sum_i a_i \mathrm{d}x_i \in \Omega^1(U)$. Then $\mathrm{d}\omega = 0$ iff $\forall p \in U$, there exists an open neighborhood $p \in V \subseteq U$ and a differentiable function $f : V \to \mathbb{R}$ s.t. $\mathrm{d}f = \omega$

Proof

Assume that Ω is locally exact. Then locally $\omega = \mathrm{d}f$, $\mathrm{d}\omega = \mathrm{d}(\mathrm{d}f) = \mathrm{d}^2 f = 0$.

Assume that $\mathrm{d}\omega = 0$. $0 = \mathrm{d}\omega = \sum_i \mathrm{d}a_i \wedge \mathrm{d}x_i = \sum_i (\sum_j \frac{\partial a_i}{\partial x_j} \mathrm{d}x_j) \wedge \mathrm{d}x_i = \sum_{i,j} \frac{\partial a_i}{\partial x_j} \mathrm{d}x_j \wedge \mathrm{d}x_i$, which means $\frac{\partial a_i}{\partial x_j} = \frac{\partial a_j}{\partial x_i}$. $\forall p \in U$, $p = (p_1, \cdots, p_n)$. Consider an open ball $B(p, \varepsilon) \subseteq U$.

$\forall x = (x_1, \cdots, x_n) \in B(p, \varepsilon)$. Consider $\beta(t) = p + t(x - p)$, $t \in [0, 1]$. $\beta'(t) = (x_1 - p_1, \cdots, x_n - p_n)$. Define $f(x) = \int_{\beta(t)} \omega = \int_0^1 \sum_i a_i(\beta(t))(x_i - p_i) \mathrm{d}t$, $f : B(p, \varepsilon) \to \mathbb{R}$. We want to show $\frac{\partial f}{\partial x_i} = a_i$.

Fix $x_i = x_1$. $\frac{\partial f}{\partial x_1}(x) = \int_0^1 \frac{\partial a_1}{\partial x_1} t(x_1 - p_1) + a_1(\beta(t)) + \sum_{i>1} \frac{\partial a_i}{\partial x_1} t(x_i -$

$p_i)\mathrm{d}t = \int_0^1 \frac{\partial a_1}{\partial x_1} t(x_1 - p_1) + a_1(\beta(t)) + \sum_{i>1} \frac{\partial a_1}{\partial x_i} t(x_i - p_i)\mathrm{d}t = \int_0^1 \sum_i \frac{\partial a_1}{\partial x_i} t(x_i - p_i) + a_1(\beta(t))\mathrm{d}t = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}[a_1(\beta(t))t]\mathrm{d}t = a_1(\beta(1)) = a_1(x)$ □

Example

A 1-form is closed iff it is locally exact and locally exact is different from exact.

$$\omega = a_1\mathrm{d}x + a_2\mathrm{d}y = -\frac{y}{x^2+y^2}\mathrm{d}x + \frac{x}{x^2+y^2}\mathrm{d}y \in \Omega^1(\mathbb{R}^2 \setminus \{(0,0)\})$$

$\mathrm{d}\omega = -\frac{(x^2+y^2)-2y^2}{(x^2+y^2)^2}\mathrm{d}y \wedge \mathrm{d}x + \frac{(x^2+y^2)-2x^2}{(x^2+y^2)^2}\mathrm{d}x \wedge \mathrm{d}y = 0$. ω is closed. By Poincaré lemma it is locally exact.

Consider $\gamma(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$. $\gamma'(t) = (-\sin t, \cos t)$. Define $F : \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto (a_1(x,y), a_2(x,y))$, $F(\gamma(t)) = (-\sin t, \cos t)$

$\int_\gamma \omega = \int_0^{2\pi} \langle(-\sin t, \cos t), (-\sin t, \cos t)\rangle \mathrm{d}t = \int_0^{2\pi} \mathrm{d}t = 2\pi$. Thus ω is not exact.

Remark

We can integrate forms along curves.

But it is also possible to integrate closed forms (locally exact) along paths. In fact, assume ω is closed, γ is a path. Since $[a,b]$ is compact, $\gamma([a,b])$ is compact. One can find a finite covering $\bigcup_{i=1}^n B_i \supseteq \gamma([a,b])$. ω is locally compact on each B_i.

Find a partition of $[a,b] : 0 = t_0 < t_1 < \cdots < t_{n+1} = b$ s.t. $\gamma|_{[t_i, t_{i+1}]} =: \gamma_i$ is contained in B_i.

In each B_i, there exists a function $f_i : B_i \to \mathbb{R}$ s.t. $\mathrm{d}f_i = \omega$. So we define $\int_\gamma \omega = \sum_{i=1}^n f_i(\gamma(t_{i+1})) - f_i(\gamma(t_i))$. Notice that if γ is a curve this definition coincides with the usual definition.

To make this definition wedll defined we have to show that $\int_\gamma \omega$ doesn't depend on the partition. Taking $\mathscr{P} = \{t_0, \cdots, t_{k+1}\}$. Let $\mathscr{P}'$ be a refinement of $\mathscr{P}(\mathscr{P} \subseteq \mathscr{P}'$ and $\mathscr{P} \neq \mathscr{P}')$. $\exists t' \in]t_i, t_{i+1}[$, $\gamma(t') \subseteq B_i$. Compute the integral with the partition $\mathscr{P}'$, in the summation we have $f_i(\gamma(t_{i+1})) - f_i(\gamma(t')) + f_i(\gamma(t')) - f_i(\gamma(t_i)) = f_i(\gamma(t_{i+1})) - f_i(\gamma(t_i))$. It means that if

$\mathscr{P}' \supseteq \mathscr{P}$, then $\displaystyle\int_\gamma^{(\mathscr{P})} \omega = \int_\gamma^{\mathscr{P}'} \omega$.

Take two different partition $\mathscr{P}_1$ and $\mathscr{P}_2$, then one can find a common refinement $\mathscr{P}_1 \bigcup \mathscr{P}_2$. $\displaystyle\int_\gamma^{(\mathscr{P}_1)} \omega = \int_\gamma^{(\mathscr{P}_1 \bigcup \mathscr{P}_2)} \omega = \int_\gamma^{\mathscr{P}_2} \omega$

9.7.5 Def

Let $\gamma_0, \gamma_1 : [a,b] \to U$ be two paths. A homotopy bewteen γ_0 and γ_1 is a continuous function $H : [a,b] \times [0,1] \to U$, $(s,t) \mapsto H(s,t) = H_t(s)$ s.t.

(1) $H_0(s) = \gamma_0(s)$, $H_1(s) = \gamma_1(s)$, $\forall s \in [a,b]$.

(2) $H_t(a) = \gamma_0(a) = \gamma_1(a)$, $H_t(b) = \gamma_0(b) = \gamma_1(b)$.

Two paths are said homotopy if there is a homotopy beween them.

9.7.6 Def

Let (X, τ) be a topological space induced by a metric d, and let $U = \{U_i\}$ be an open covering of X. The lebesgue number δ_U(of the open covering) is a non-negative real number that satisfies the following condition : if $Z \subseteq X$ is a subset with $\operatorname{diam}(Z) < \delta$, then $Z \subseteq U_j$ for some $U_j \in U$.

Notice that if δ is a lebesgue number, then any $\delta' < \delta$ is also a lebesgue number. 0 is a lebesgue number.

9.7.7 Lemma

If X is a compact metric space, then for any open covering there exists a positive lebesgue number.

Proof

Let $U = \{U_i\}$ be an open covering. Since X is compact, then $\bigcup\limits_{i=1}^{n} U_i = X$. If one of the U_i is equal to X, then any positive number is a lebesgue number. So we can assume that $\forall i \in \{1, \cdots, n\}$, the set $V_i := X \setminus U_i$ is not empty.

Define the following function $f : X \to \mathbb{R}$, $x \mapsto \dfrac{1}{n} \sum\limits_{i=1}^{n} d(x, V_i)$. f is continuous on a compact set (In fact, $|d(x,V) - d(z,V)| \leq |d(x,y) - d(z,y)| \leq d(x,z)$), hence it attains a minimum. This minimum is not 0, because $d(x, V_i) > 0$ for some i (In fact, $d(x,V) = 0$ means $x \in \overline{V}$), take $\delta = \min\limits_{x \in X} f(x) > 0$, we show that δ is a lebesgue number.

Let $Y \subseteq X$ s.t. $\operatorname{diam}(Y) < \delta$. Take $x_0 \in Y$, then $Y \subseteq B(x_0, \delta)$. Since $f(x_0) \geq \delta$, then there exists at least one i s.t. $d(x_0, V_i) \geq \delta$, this means $B(x_0, \delta) \subseteq U_i$, hence $Y \subseteq U_i$. $\square$

9.7.8 Theorem(Homotopy invariance of the integral)

Let ω be a closed 1-form on an open set U. Let γ_0, γ_1 be two homotopic paths in U, then $\displaystyle\int_{\gamma_0} \omega = \int_{\gamma_1} \omega$.

Proof

Since ω is closed, then it is locally exact. Let $H : [a, b] \times [0, 1] \to U$ be the homotopy between γ_0 and γ_1, and let $\{B_i\}_{i=1}^n = \mathscr{B}$ be an open covering of $H([a, b] \times [0, 1]) \subseteq U$ made of open balls where ω is locally exact.

Define $R := [a, b] \times [0, 1]$. We have covered R with $\{W_i := H^{-1}(B_i)\}_{i=1}^n$. Since R is compact we can choose a lebesgue number $\delta > 0$ for this covering.

Divide R into rectangles $\{R_{jk}\}$ having diameter $< \delta$.

$$\begin{tikzcd}
(s_R, t_{j+1}) \arrow[rr, "\alpha_{j,k+1}"] & & (s_{R+1}, t_{j+1}) \\
& R_{jk} & \\
(s_R, t_j) \arrow[rr, "\alpha_{j,k}"] \arrow[uu, "\beta_{j,k}"] & & (s_{R+1}, t_j) \arrow[uu, "\beta_{j+1,k}"']
\end{tikzcd}$$

$\partial R_{jk} = \alpha_{j,k} \bigsqcup \beta_{j,k+1} \bigsqcup (-\alpha_{k+1,k}) \bigsqcup (-\beta_{j,k})$. $H(\partial R_{jk}) = H(s, t_j) \bigsqcup H(s_{R+1}, t)$ $\bigsqcup (-H(s, t_{j+1})) \bigsqcup (-H(s_R, t))$ $s \in [s_R, s_{R+1}]$, $t \in [t_j, t_{j+1}]$, $s \in [s_R, s_{R+1}]$, $t \in [t_j, t_{j+1}]$.

$H(\partial R_{jk})$ is a closed curve contained in some B_i but in such balls ω is exact, then $\displaystyle\int_{H(\partial R_{jk})} \omega = 0$. Do it $\forall ij$

$$0 = \sum_{j,k} \int_{H(\partial R_{jk})} \omega = \sum_{j,k} \left(\int_{H(s,t_j)} \omega + \int_{H(s_{R+1},t)} \omega - \int_{H(s,t_{j+1})} \omega - \int_{J(s_R,t)} \omega \right) =$$
(*)

Moreover, $H(s, 0) = \gamma_0(s)$, $H(s, 1) = \gamma_1(s)$, $H(b, t) = \gamma_0(b) = \gamma_1(b)$, $H(a, t) = \gamma_0(a) = \gamma_1(a)$.

(*) $= \displaystyle\int_{\gamma_0} \omega + \int_{\gamma_0(b)} \omega - \int_{\gamma_1} \omega - \int_{\gamma_0(a)} \omega = \int_{\gamma_0} \omega - \int_{\gamma_1} \omega$. Thus $\displaystyle\int_{\gamma_0} \omega =$

$\int_{\gamma_1} \omega$ □

9.7.9 Def

Let $\gamma_0, \gamma_1 : [a,b] \to U$ be two loops. A free homotopy between γ_0 and γ_1 is a continuous mapping $H : [a,b] \times [0,1] \to U$, $(s,t) \mapsto H(s,t) = H_t(s)$ s.t.

(1) $H(s,0) = \gamma_0(s)$, $H(s,1) = \gamma_1(s)$

(2) For any fixed t_0, $H(s,t_0)$ is a loop.

Notice that free homotopy isn't homotopy at all. Two loops with a free homotopy between them are said freely homotopic.

9.7.10 Corollary

Let $\omega \in \Omega^1(U)$ be a closed form and let γ_0 and γ_1 be two freely homotopic loops, then $\displaystyle\int_{\gamma_0} \omega = \int_{\gamma_1} \omega$

Proof

Do the same proof. Only needs to change reason of the equation $\displaystyle\int_{\gamma_0} \omega + \int_{H(b,t)} \omega - \int_{\gamma_1} \omega - \int_{H(a,t)} \omega = \int_{\gamma_0} \omega - \int_{\gamma_1} \omega$. □

Exercise

Being homotopic and being freely homotopic are equivalence relations.(We use $\backsimeq$ to denote these equivalence relation.)

Answer

Trivial. □

9.7.11 Def

Let $U \subseteq \mathbb{R}^n$ be a path connected subspace. Then U is said simply connected if any loop is freely homotopic to a point.

9.7.12 Prop

Let ω be a closed 1-form on a simply connected open set $U \subseteq \mathbb{R}^n$. Then ω is exact.

Proof

By the corollary, if γ is a loop, then $\int_\gamma \omega = \int_{\{p\}} \omega = 0$ means the form is exact. $\square$

Application

$\mathbb{R}^2 \setminus \{0\}$ isn't simply connected.
$\omega = -\frac{y}{x^2+y^2}\mathrm{d}x + \frac{x}{x^2+y^2}\mathrm{d}y$ isn't exact in $\mathbb{R}^2 \setminus \{0\}$.

9.8 Winding numbers

9.8.1 Def

A path $\gamma : [a, b] \to \mathbb{R}^2$ is said simple if $\gamma|_{]a,b[}$ is injective.

A connected component means a closed and open set which is also connected.

9.8.2 Theorem(Jordan)

Let $\gamma : [a, b] \to \mathbb{R}^2$ be a simple loop , then $\mathbb{R}^2 \setminus \gamma([a, b])$ consists of two connected components. One of this is bounded(interior), the other is unbounded (exterior). Moreover, $\gamma([a, b])$ is the boundary of any of the two components.

Proof

Won't be covered in this course. $\square$

9.8.3 Def

Let $S^1 \subseteq \mathbb{R}^2$ be the unit circle and let $c : I \to S^1, t \mapsto (x(t), y(t))$ $(0 \in I)$ be a closed curve. Consider $\varphi_0 \in [0, 2\pi[$ s.t. $\cos \varphi_0 = x(0)$, $\sin \varphi_0 = y(0)$.

Define the function $\varphi : I \to \mathbb{R}$, $t \mapsto \varphi(t) := \varphi_0 + \displaystyle\int_0^t (xy' - yx')\mathrm{d}\tau$. φ is called an angle function of c since by calculation(Let $F(t) = (x(t) - \cos\varphi(t))^2 + (y(t) - \sin\varphi(t))^2$, then $F'(t) = 0$.) one can check φ is such a function that $\begin{cases} \cos\varphi(t) = x(t) \\ \sin\varphi(t) = y(t) \end{cases}$, $\forall t \in I$.

9.8.4 Def

Let $c : [0, b] \to S^1$ be a closed curve. Let φ be the angle function of c. We define the winding number of c as $n(c) := \dfrac{1}{2\pi}(\varphi(b) - \varphi(0))$. Since c is a closed curve, then $n(c) \in \mathbb{Z}$.

Example

$c : [0, 2\pi] \to S^1$, $t \mapsto (\cos kt, \sin kt)$, $k \in \mathbb{Z}$. $c'(t) = (-k \sin kt, k \cos kt)$.

$\varphi(t) = \displaystyle\int_0^t k\mathrm{d}\tau = kt$, $n(c) = k$.

9.8.5 Def

Let $\gamma : [0,b] \to \mathbb{R}^2 \setminus \{p\}$ be a closed curve, $\gamma(t) = p + \rho(t)c(t)$, where $c(t) \in S^1$. $r(t) = p + \rho(t)(\cos\theta(t), \sin\theta(t))$.

Then we define the winding number of γ at p, $n_{p(\gamma)} := n(c)$.

9.8.6 Prop

Let $\gamma(t) = p + \rho(t)c(t)$ be a closed curve.

$\gamma : [0,b] \to \mathbb{R}^2 \setminus \{p\}$. Then $n_{p(\gamma)} = \dfrac{1}{2\pi}\displaystyle\int_C \omega_0$ where $\omega_0 = \dfrac{-y}{x^2+y^2}\mathrm{d}x + \dfrac{x}{x^2+y^2}\mathrm{d}y$.

Proof

$$\frac{1}{2\pi}\int_C \omega_0 = \frac{1}{2\pi}\int_0^b (x(t)y'(t) - y(t)x'(t))\mathrm{d}t = \frac{\varphi(b)-\varphi(0)}{2\pi} = n(c) \qquad \square$$

9.8.7 Prop

Let $\gamma_0, \gamma_1 : [a,b] \to \mathbb{R}^2 \setminus \{p\}$ be two closed curves. Then γ_0 and γ_1 are freely homotopic iff $n_{p(\gamma_0)} = n_{p(\gamma_1)}$.

Proof

From left to right : Let $\hat{\gamma_1}(t) = p + c_0(t)$ and $\hat{\gamma_2}(t) = p + c_1(t)$. The freely homotopy between γ_i and $\hat{\gamma_i}$ is given by $H_i(s,t) = p + \dfrac{\rho_i(s)}{(1-t)+t\rho_i(s)}c_i(s)$. Hence $\hat{\gamma_0} \backsimeq \hat{\gamma_1}$ in $\mathbb{R}^2 - \{p\}$, which is equivalent to $c_0(t) \backsimeq c_1(t)$ in $\mathbb{R}^2 - \{0\}$. The rest of the proof follows from the invariance of the integral of ω_0 along free homotopic loops.

From right to left : Assume $\gamma_0(t) = p + \rho_0(t)c_0(t)$, $\gamma_1(t) = p + \rho_1(t)c_1(t)$. The angle functions of c_1 and c_2 are φ_0 and φ_1, respectively.

Consider $\varphi(s,t) = (1-t)\varphi_0(s) + t\varphi_1(s)$, $H(s,t) = p + (\cos\varphi(s,t), \sin\varphi(s,t))$, $(s,t) \in [0,b] \times [0,1]$. One can claim H is a free homotopy between $\hat{\gamma_1}(t)$ and $\hat{\gamma_2}(t)$.

We have to check that any curve in the homotopy is closed. $\varphi(b,t) - \varphi(0,t) = (1-t)\varphi_0(b) + t\varphi_1(b) - (1-t)\varphi_0(0) - t\varphi_1(0) = (1-t)(\varphi_0(b) - \varphi_0(0)) + t(\varphi_1(b) - \varphi_1(0)) = (1-t)2\pi n_{p(\gamma_0)} + t2\pi n_{p(\gamma_1)} = 2\pi n_{p(\gamma_0)}$. So we proved that $\hat{\gamma_0}$ and $\hat{\gamma_1}$ are freely homotopic. Therefore, $\gamma_0 \backsimeq \hat{\gamma_0} \backsimeq \hat{\gamma_1} \backsimeq \gamma_1$. $\square$

9.8.8 Def

Let $F : U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a differentiable mapping. We say that $p \in U$ is a zero of F if $F(p) = 0$. If there exists a neighborhood V of p s.t. V contains no zeros of F other than p then p is called isolated zero.

If p is a zero of F and $\mathrm{d}F|_p$ is invertible at p, then we say that p is a simple zero.

Remark

By the local inversion theorem F is a bijection in a neighborhood of a simple zero. Hence a simple zero is isolated.

9.8.9 Def

Let $F : U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ be a differentiable mapping. $F(x, y) = (f(x, y), g(x, y))$. $D \subseteq U$ is a closed disk, with boundary $\partial D = C$ ($C : [0, 2\pi] \to U$). Assume that C doesn't contain zeros of F.

Consider the form $\theta = \dfrac{f\mathrm{d}g - g\mathrm{d}f}{f^2 + g^2} \in \Omega^1(U \setminus \{(x, y)|F(x, y) = 0\})$.

The index of F in D is defined as $n(F, D) := \dfrac{1}{2\pi}\displaystyle\int_C \theta$. See that $\theta = F^*\omega_0$, $\omega_0 = \dfrac{-y}{x^2 + y^2}\mathrm{d}x + \dfrac{x}{x^2 + y^2}\mathrm{d}y$. $n(F, D) = \dfrac{1}{2\pi}\displaystyle\int_C \theta = \dfrac{1}{2\pi}\displaystyle\int_C F^*\omega_0 = \dfrac{1}{2\pi}\displaystyle\int_{F\circ C} \omega_0 =$ winding number of $F \circ C$ at the center of D.

9.8.10 Prop

If $n(F, D) \neq 0$, then $\exists$ a point $q \in D$ s.t. $F(q) = 0$.

Proof

Assume we don't have zeros, let p be the center of D.

Let $H(s, t) : [0, 2\pi] \times [0, 1] \to \mathbb{R}^2 \setminus \{(0, 0)\}$,$(s, t) \mapsto F((1 - t)C(s) + tp)$, this is a free homotopy in D between $F(p)$ and $F(C)$. $0 = \dfrac{1}{2\pi}\displaystyle\int_{F(p)} \omega_0 = \dfrac{1}{2\pi}\displaystyle\int_{F\circ C} \omega_0 = n(F, D)$ $\square$

Remark

Assume that $F_1, F_2 : U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ are two differentiable mappings with no zeros in $C = \partial D$. If there is a continuous mapping $H(s, t) : \partial D \times [0, 1] \to$

$\mathbb{R}^2\backslash\{(0,0)\}$ s.t. $\begin{cases} H(q,0) = F_1(q) \\ H(q,1) = F_2(q) \\ H(q,t) \neq (0,0), \forall (q,t) \end{cases}$ then the curves $F_1 \circ C$ and $F_2 \circ C$ are freely homotopic, $K : [0, 2\pi] \times [0,1] \to \mathbb{R}^2 - \{(0,0)\}$, $(s,t) \mapsto H(C(s),t)$, which means $n(F_1, D) = n(F_2, D)$.

9.8.11 Def

A simple zero p of F is said positive if $\det(\mathrm{d}F|_p) > 0$, otherwise it is said negative.

9.8.12 Def

We say $B \subseteq X$ is discrete if $\forall x \in B$, $\exists$ an open neighborhood $U \ni x$ s.t. $B \bigcap U = \{x\}$.

9.8.13 Lemma

Assume that F has a unique simple zero $p \in (D_1 \backslash C_1) \subseteq U$, then $\exists D_2 \subseteq D_1$ s.t. $n(F, D_2) = \pm 1$ occur along wih $\det(\mathrm{d}F|_p)$. Moreover, $n(F, D_1) = n(F, D_2)$.

Proof

Not loss generality assume $p = (0,0)$. By the taylor formula $F(q) = Tq + R(q)\|q\|$, $\lim\limits_{q \to 0} R(q) = 0$.

Consider the mapping $H : U \times [0,1] \to \mathbb{R}^2$, $(q,t) \mapsto Tq + (1-t)R(q)\|q\|$. If we show that when D_2 small enough, $H(q,t) \neq 0$, $\forall q \in D_2$, $\forall t \in [0,1]$, then $H(C(s),t)$ is a free homotopy between $F \circ C$ and $T \circ C$ thus $n(F, D) = n(T, D)$, which should prove that $n(T, D) = \pm 1$. We have two approaches.

(1) Direct computation : $T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $T \circ C = \begin{pmatrix} a\rho\cos\theta + b\rho\sin\theta \\ c\rho\cos\theta + d\rho\sin\theta \end{pmatrix}$ then compute $\dfrac{1}{2\pi}\displaystyle\int_{T \circ C} \omega_0$.

(2) $T(C)$ is an ellipse, which is homotopic to circles. $n(T, D) = n(T, C) = \pm 1$.

Define $c = \dfrac{1}{\|T^{-1}\|}$, then $\|q\| = \|T^{-1} \circ Tq\| \leq \|T^{-1}\|\|Tq\| = \dfrac{\|Tq\|}{c}$, hence $\|Tq\| \geq c\|q\|$. Take $\varepsilon > 0$, s.t. $\forall q \in D_\varepsilon$, $\|R(q)\| \leq \dfrac{c}{2}$.

Then if $q \in D_\varepsilon \setminus \{(0,0)\}$, $\|H(q,t)\| = \|Tq + (1-t)R(q)\|q\|\| \geq \|Tq\| - (1-t)\|R(q)\|\|q\| \geq c\|q\| - \frac{(1-t)c}{2}\|q\| \geq \frac{c}{2}\|q\| > 0$. We find the $D_2 = D_\varepsilon$ we need. □

9.8.14 Theorem(Kronecker index formula)

Assume that $F : U \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ has only simple zeros in a disk $D \subseteq U$, and none of them in ∂D, then $n(F,D) = P - N$, where P is the number of positive zeros and N is the number of negative zeros.

Proof

Since simple zeros are isolated, they form a closed discrete subset of the compact D, hence simple zeros in D are finitely many $p_1, \cdots, p_k$.

$$n(F,D) = \frac{1}{2\pi}\int_C \theta = \frac{1}{2\pi}\sum_{i=1}^{k}\int_{\partial S_i}\theta = \frac{1}{2\pi}\sum_{i=1}^{k}\int_{C_i}\theta = P - N$$ □

9.9 Length and regularity

Let $\gamma : [a,b] \to \mathbb{R}^n$ be a path. $\gamma(t) = (\gamma_1(t), \cdots, \gamma_n(t))$, where each $\gamma_i : [a,b] \to \mathbb{R}$ is continuous. Then $\int_a^b \gamma(t)\mathrm{d}t := (\int_a^b \gamma_1(t)\mathrm{d}t, \cdots, \int_a^b \gamma_n(t)\mathrm{d}t) \in \mathbb{R}^n$

9.9.1 Lemma

Let γ be as above.

(1) If $\gamma \in C^1([a,b])$, then $\int_a^b \gamma'(t)\mathrm{d}t = \gamma(b) - \gamma(a)$.

(2) For any $c \in \mathbb{R}^n$, $\langle c, \int_a^b \gamma(t)\mathrm{d}t\rangle = \int_a^b \langle c, \gamma(t)\rangle \mathrm{d}t$.

(3) $\| \int_a^b \gamma(t)\mathrm{d}t\| \leq \int_a^b \|\gamma(t)\|\mathrm{d}t$.

Proof

(1) A consequence of the fundamental theorem of calculus componentwise. □

(2) $c = (c_1, \cdots, c_n)$, $\langle c, \int_a^b \gamma(t)\mathrm{d}t\rangle = \sum_{i=1}^n c_i \int_a^b \gamma_i(t)\mathrm{d}t = \int_a^b \sum_{i=1}^n c_i\gamma_i(t)\mathrm{d}t = \int_a^b \langle c, \gamma_i(t)\rangle \mathrm{d}t$ □

(3) We apply (2) with $c = \int_a^b \gamma(t)\mathrm{d}t \in \mathbb{R}^n$.

$\| \int_a^b \gamma(t)\mathrm{d}t\|^2 = \langle \int_a^b \gamma(t)\mathrm{d}t, \int_a^b \gamma(t)\mathrm{d}t\rangle = \int_a^b \langle \int_a^b \gamma(t)\mathrm{d}t, \gamma(t)\rangle \mathrm{d}t \leq \int_a^b \| \int_a^b \gamma(t)\mathrm{d}t\| \|\gamma(t)\|\mathrm{d}t = \| \int_a^b \gamma(t)\mathrm{d}t\| \int_a^b \|\gamma(t)\|\mathrm{d}t$. The case when $\| \int_a^b \gamma(t)\mathrm{d}t\| = 0$ is trivial. □

9.9.2 Def

Let $\gamma : [a,b] \to \mathbb{R}^n$ be a path.

$\mathscr{P} = \{t_0, t_1, \cdots, t_n\}$, s.t. $a = t_0 < t_1 < \cdots < t_n = b$ or $a = t_n < t_{n-1} < \cdots < t_0 = a$, $P_i = \gamma(t_i)$, $l_{\mathscr{P}}(\gamma) = \sum_{i=1}^{n-1} \|P_{i+1} - P_i\|$

The length of γ is $l(\gamma) := \sup_{\mathscr{P}} \{l_{\mathscr{P}}(\gamma)\}$. If $l(\gamma) < +\infty$, then the path γ is said rectifiable.

Example

Consider $\gamma : [0,1] \to \mathbb{R}^2$, $\gamma_1(t) = t$, $\gamma_2(t) = \begin{cases} t\sin\dfrac{\pi}{2t}, & \text{if } t \neq 0 \\ 0, & \text{otherwise} \end{cases}$

Take $\mathscr{P}_n = \{1, \frac{1}{3}, \cdots, \frac{1}{2n-1}, \frac{1}{2n+1}, 0\}$. Recall $\sin(j+\frac{1}{2})\pi = (-1)^j$, then for $i \in \mathbb{N}_{\leq n}$, $P_i = \gamma(t_i) = (\frac{1}{2i+1}, \frac{1}{2i+1}(-1)^i)$

$\|P_j - P_{j-1}\| = \sqrt{(\frac{2}{4j^2-1})^2 + (\frac{4j}{4j^2-1})^2} = \sqrt{\frac{4(4j^2+1)}{(4j^2-1)^2}} \geq \sqrt{\frac{4(4j^2-1)}{(4j^2-1)^2}} = \sqrt{\frac{4}{4j^2-1}} \geq \frac{1}{j}$, then $l_{\mathscr{P}}(\gamma) \geq \sum\limits_{n\in\mathbb{N}} \frac{1}{n} = +\infty$.

9.9.3 Prop

Let $\gamma : [a,b] \to \mathbb{R}^n$ be of class C^1, then γ is rectifiable and $l(\gamma) = \int_a^b \|\gamma'(t)\|\mathrm{d}t$. Moreover, $l(\gamma)$ doesn't depend on the parametrization of the γ.

Proof

Let $\mathscr{P}$ be a partition of $[a,b]$, $\|\gamma(t_{j+1}) - \gamma(t_j)\| = \|\int_{t_j}^{t_{j+1}} \gamma'(t)\mathrm{d}t\| \leq \int_{t_j}^{t_{j+1}} \|\gamma'(t)\|\mathrm{d}t$. So $l_{\mathscr{P}}(\gamma) = \sum\limits_{i=0}^{n-1} \|\gamma(t_{j+1}) - \gamma(t_j)\| \leq \sum\limits_{i=0}^{n-1} \int_{t_j}^{t_{j+1}} \|\gamma'(t)\|\mathrm{d}t = \int_a^b \|\gamma'(t)\|\mathrm{d}t$. Therefore, $l(\gamma) \leq \int_a^b \|\gamma'(t)\|\mathrm{d}t$, which means $l(\gamma)$ is finite.

$\gamma'(t)$ is continuous on $[a,b]$, so it is uniformly continuous. $\forall\varepsilon > 0$, $\exists\delta$ s.t. if $t,s \in [a,b]$ s.t. $|t-s| < \delta$ then $\|\gamma'(t) - \gamma'(s)\| < \varepsilon$. Choose a partition $\mathscr{P}$ s.t. $|t_{j+1} - t_j| < \delta$, $\forall i$, then $\forall t \in [t_j, t_{j+1}]$, then $\|\gamma'(t) - \gamma'(t_j)\| < \varepsilon$.

Hence $\int_{t_j}^{t_{j+1}} \|\gamma'(t)\|\mathrm{d}t < \|\int_{t_j}^{t_{j+1}} \gamma'(t_j)\mathrm{d}t\| + (t_{j+1}-t_j)\varepsilon \leq \|\int_{t_j}^{t_{j+1}} \gamma'(t)\mathrm{d}t\| + \|\int_{t_j}^{t_{j+1}} (\gamma'(t) - \gamma'(t_j))\mathrm{d}t\| + (t_{j+1} - t_j)\varepsilon \leq \|\gamma(t_{j+1}) - \gamma(t_j)\| + 2\varepsilon(t_{j+1} - t_j)$

So $\int_a^b \|\gamma'(t)\|\mathrm{d}t = \sum\limits_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|\gamma'(t)\|\mathrm{d}t \leq \sum\limits_{j=0}^{n-1} \|\gamma(t_{j+1}) - \gamma(t_j)\| + 2\varepsilon(b-a) \leq l(\gamma) + 2\varepsilon(b-a)$. Taking $\varepsilon \to 0$.

Let $\varphi : [\alpha,\beta] \to [a,b]$, φ is of class C^1, $\varphi(\tau) \neq 0$, $\forall\tau \in [\alpha,\beta]$. $\tilde{\gamma} : [\alpha,\beta] \to \mathbb{R}^n$ where $\tilde{\gamma} = \gamma \circ \varphi$.

Not loss generality, let $\varphi'(\tau) > 0$, $\varphi(\alpha) = a$ and $\varphi(\beta) = b$. $l(\gamma) =$

$\displaystyle\int_a^b \|\gamma'(t)\|\mathrm{d}t = \int_\alpha^\beta \|\gamma'(\varphi(\tau))\|\varphi'(\tau)\mathrm{d}\tau = \int_\alpha^\beta \|\gamma'(\varphi(\tau))\varphi'(\tau)\| = \int_\alpha^\beta \|\tilde{\gamma}'(\tau)\|\mathrm{d}\tau =$ $l(\tilde{\gamma})$ □

Exercise

If γ is a curve, then γ is rectifiable and the length is the sum of the lengths of its C^1 pieces.

Answer

For those partitions that have not taken all the endpoints, by taking all the endpoints to refine them, the triangle inequality can be used. □

9.9.4 Def

A C^1-curve is regular if $\gamma'(t) \neq 0 \ \forall t \in [a, b]$.

A piecewise C^1-path is regular if all its C^1-pieces are regular.

9.9.5 Def

Let $\gamma : [a, b] \to \mathbb{R}^n$ be a curve of class C^1. $s(t) := \displaystyle\int_a^t \|\gamma'(u)\|\mathrm{d}u$, $s(t)$ is the length of $\gamma|_{[a,t]}$ and $s'(t) = \|\gamma'(t)\|$.

Now assume that γ is regular thus $s'(t) > 0$. Therefore, $s : [a, b] \to [0, l]$ is a C^1-diffeomorphism, the inverse is $t : [0, l] \to [a, b]$, $\dfrac{\mathrm{d}t}{\mathrm{d}s} = \dfrac{1}{\|\gamma'(t)\|}$.

We reparametrze γ with respect to t. We obtain $\tilde{\gamma}(s) = (\gamma \circ t)(s)$, $\tilde{\gamma} : [0, l] \to \mathbb{R}^n$ we say that $\tilde{\gamma}$ is the reparametrization of γ with respect to the curvilinear coordinate $s(t)$.

Example

$\gamma(t) = (R\cos t, R\sin t, ct)$, $t \in [0, 2\pi]$, $R, c > 0$.

$s(t) = \displaystyle\int_0^t \sqrt{R^2 + c^2}\mathrm{d}u = \sqrt{R^2 + c^2}t$, $t(s) = \dfrac{s}{\sqrt{R^2 + c^2}}$.

$\tilde{\gamma}(s) = (R\cos\dfrac{s}{R^2 + c^2}, R\sin\dfrac{s}{R^2 + c^2}, \dfrac{cs}{\sqrt{R^2 + c^2}})$, $l(\gamma) = s(2\pi) = 2\pi\sqrt{R^2 + c^2}$

9.9.6 Def

Let $\gamma : [a, b] \to \mathbb{R}^n$, $\tilde{\gamma} : [0, l] \to \mathbb{R}^n$. $\dfrac{\mathrm{d}\tilde{\gamma}}{\mathrm{d}s} = \dfrac{\mathrm{d}\gamma}{\mathrm{d}t}\dfrac{\mathrm{d}t}{\mathrm{d}s} = \dfrac{\gamma'(t)}{\|\gamma'(t)\|}$, $\|\dfrac{\mathrm{d}\tilde{\gamma}}{\mathrm{d}s}\| = 1$.

Define the tangent vector of norm 1 as $T(t) := \dfrac{\mathrm{d}\tilde{\gamma}}{\mathrm{d}s} = \dfrac{\gamma'(t)}{\|\gamma'(t)\|}$.

$\|T(t)\| \equiv 1$, thus $0 = \dfrac{\mathrm{d}}{\mathrm{d}t}\|T(t)\|^2 = \dfrac{\mathrm{d}}{\mathrm{d}t}\langle T(t), T(t)\rangle = 2\langle T'(t), T(t)\rangle$, which means $T'(t) \perp T(t)$. We define the normal vector as $N(t) = \dfrac{T'(t)}{\|T'(t)\|}$.

$\dfrac{\mathrm{d}^2\tilde{\gamma}}{\mathrm{d}s^2} = \dfrac{\mathrm{d}}{\mathrm{d}s}(\dfrac{\mathrm{d}\tilde{\gamma}}{\mathrm{d}s}) = \dfrac{\mathrm{d}}{\mathrm{d}s}(T(t)) = \dfrac{T'(t)}{\|\gamma'(t)\|} = N(t)\dfrac{\|T'(t)\|}{\|\gamma'(t)\|}$, which indicates $N(t) = \dfrac{\frac{\mathrm{d}^2\tilde{\gamma}}{\mathrm{d}s^2}}{\|\frac{\mathrm{d}^2\tilde{\gamma}}{\mathrm{d}s^2}\|}$

Along the curve we have a moving canonical basis of $\mathbb{R}^2_{\gamma(t)} = \mathrm{span}\{T(t), N(t)\}$.

9.9.7 Def(Line integral of the first kind)

Let $f : U \to \mathbb{R}^n \to \mathbb{R}$ be measurable. $\gamma : [a, b] \to \mathbb{R}^n$ be a regular curve of class C^1.

The curvilinear integral of f along γ is $\int_\gamma f\mathrm{d}s := \int_a^b f(\gamma(t))\|\gamma'(t)\|\mathrm{d}t$. In fact, it can be seen as an 'inverse' reparametrization of the curvilinear coordinate $\int_\gamma f\mathrm{d}s = \int_{s(a)}^{s(b)} f(\tilde{\gamma}(s))\mathrm{d}s = \int_a^b f(\gamma(t))\dfrac{\mathrm{d}s}{\mathrm{d}t}\mathrm{d}t = \int_a^b f(\gamma(t))s'(t)\mathrm{d}t = \int_a^b f(\gamma(t))\|\gamma'(t)\|\mathrm{d}t$.

Exercise

Let $f, g : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be two measurable functions γ, γ_1, γ_2 (piecewise) regular C^1 curves, then

(1) $\int_\gamma \alpha f + \beta g\mathrm{d}s = \alpha \int_\gamma f\mathrm{d}s + \beta \int_\gamma g\mathrm{d}s$, $\alpha, \beta \in \mathbb{R}$

(2) If $\gamma_1 \bigsqcup \gamma_2$ make sense, then $\int_{\gamma_1 \bigsqcup \gamma_2} f\mathrm{d}s = \int_{\gamma_1} f\mathrm{d}s + \int_{\gamma_2} f\mathrm{d}s$

(3) $\int_{\gamma_1} f\mathrm{d}s = \int_{-\gamma_1} f\mathrm{d}s$

(4) $\int_\gamma f\mathrm{d}s$ doesn't depend on the parametrization of γ.

Answer

(1) (2) Trivial. $\square$

(3) Not loss generality, let $\gamma : [a, b] \to \mathbb{R}^n$ be of class C^1.

$\int_{-\gamma} f\mathrm{d}s = \int_b^a f(-\gamma(t))\|-\gamma(t)\|\mathrm{d}t = \int_b^a f(\gamma(a+b-t))\|\gamma(a+b-t)\|\mathrm{d}t = \int_a^b f(\gamma(t'))\|\gamma(t')\|\mathrm{d}t' = \int_\gamma f\mathrm{d}s$ $\square$

(4) Let $\alpha : [c, d] \to [a, b]$, $\tau \mapsto t$ s.t. $\tilde{\gamma} = \gamma \circ \alpha$ is a reparametrization of γ.

$$\int_\gamma f \mathrm{d}s = \int_a^b f(\gamma(t))\|\gamma'(t)\|\mathrm{d}t = \int_c^d f(\gamma(\alpha(\tau)))\|\gamma'(\alpha(\tau))\|\frac{\mathrm{d}t}{\mathrm{d}\tau}\mathrm{d}\tau = \int_c^d f(\tilde{\gamma}(\tau))\|\tilde{\gamma}'(\tau)\|\mathrm{d}\tau = \int_{\tilde{\gamma}} f \mathrm{d}s$$

□

Chapter 10

Fourier analysis

In this chapter, we fix an inner product space V over $\mathbb{R}$ or $\mathbb{C}$
whose non-degenerate and positive definite inner product $\langle\ ,\ \rangle$ is either real symmetric or Hermitian.

Also the functions in L^p space are real or complex value.

10.1 Orthogonal expansions

Reminder

Let V be the inner product space mentioned in the beginning of the chapter. We reprove the Cauchy-Schwarz inequality in $\mathbb{R}$ (applicable to the infinite-dimensional case), and the generalization from $\mathbb{R}$ to $\mathbb{C}$ is the same as that in the finite-dimensional case, so it will not be elaborated here.

$f(t) = \langle x + ty, x + ty\rangle = \|x\|^2 + 2t\langle x, y\rangle + \|y\|^2 \geq 0$, then consider the discriminant of a quadratic equation.

10.1.1 Def

Let V be the vector space agreed upon at the beginning of this chapter. If V is complete with the norm given by the inner product then we say V is a Hilbert space.

A set of vectors $\{l_k | k \in I\}$ is said to be an orthogonal system if $\langle l_k, l_j\rangle = 0$ iff $k \neq j$. Moreover $\{l_k\}$ is an orthonormal system if $\langle l_k, l_j\rangle = \delta_{k,j}$. And in this chapter we always assume the index set I is up to countable.

Note

中文书里的内积空间一般就是指 $\mathbb{C}$ 或 $\mathbb{R}$ 上的线性空间带有一个正定的厄米特型或对称型。从而希尔伯特空间的定义可以很简洁的写成完备的内积空间。

10.1.2 Prop

Let $\{l_k\}$ be an orthogonal system, then l_k is a set of linearly independent vectors.

Proof

Take $c_1, \cdots, c_n$, $n \in \mathbb{N}$ s.t. $c_1 l_1 + \cdots + c_n l_n = 0$. We have to show $c_1 = \cdots = c_n = 0$. $0 = \langle 0, l_j\rangle = c_j\langle l_j, l_j\rangle$, since $\langle l_j, l_j\rangle > 0$ then $c_j = 0$. $\square$

10.1.3 Prop

(1) If $\lim\limits_{n\to+\infty} x_n = x$, then $\forall y \in V$, $\lim\limits_{n\to+\infty}\langle x_n, y\rangle = \langle x, y\rangle$ and $\lim\limits_{n\to+\infty}\langle y, x_n\rangle = \langle y, x\rangle$(The inner product $\langle\ ,\ \rangle$ is continuous).

(2) If $x = \sum_{k=1}^{\infty} v_k$, then $\forall y \in V$, $\langle x, y\rangle = \sum_{k=1}^{\infty} \langle v_k, y\rangle$.

(3) If $\{l_k\}$ is orthonormal and $x = \sum_{k=1}^{\infty} x_k l_k$, $y = \sum_{k=1}^{\infty} y_k l_k$ then $\langle x, y\rangle = \sum_{k=1}^{\infty} x_k \overline{y_k}$ or $\sum_{k=1}^{\infty} x_k y_k$.

Proof

(1) $|\langle x_n, y\rangle - \langle x, y\rangle| = |\langle x_n - x, y\rangle| \leq \|x_n - x\|\|y\|$ □

(2) This is an application of (1). □

(3) This is an application of (2). □

10.1.4 Corollary(Pythagoras)

(1) If $\{v_k\}$ is a system of orthogonal vectors and $V \ni v = \sum_{k=1}^{\infty} v_k$, then $\|v\|^2 = \sum_{k=1}^{\infty} \|v_k\|^2$

(2) If $\{l_k\}$ is a system of orthonormal vectors and $x = \sum_{k=1}^{\infty} x_k l_k$, then $\|x\|^2 = \sum_{k=1}^{\infty} |x_k|^2$

10.1.5 Def

Let $\{l_k\}$ be an orthogonal system of V. Then $x_k := \dfrac{\langle x, l_k\rangle}{\langle l_k, l_k\rangle}$ is called a Fourier coefficient of x in $\{l_k\}$.

Remark

Consider $x \in V$, then the Fourier series of x in $\{l_k\}$ is $\sum_{k=1}^{\infty} \dfrac{\langle x, l_k\rangle}{\langle l_k, l_k\rangle} l_k$. We don't know if $\sum_{k=1}^{\infty} \dfrac{\langle x, l_k\rangle}{\langle l_k, l_k\rangle} l_k$ converges and if it converges to x.

Example

Let $X \subseteq \mathbb{R}^n$ be a λ^n-measurable set. We define the integral on complex number by $\int_X f \mathrm{d}\mu := \int_X \Re(f)\mathrm{d}\mu + i\int_X \Im(f)\mathrm{d}\mu$ and let $L^2(X, \mathbb{C}) := \{f : X \to \mathbb{C} | \int_X |f|^2 \mathrm{d}\lambda^n < +\infty\}$.(Since it's in $\mathbb{R}^n$ we omit the familiar borel sigma algebra of $\mathbb{R}^n$ and lebesgue measure in $L^2(X, \mathbb{C})$.) The definition of $L^2(X, \mathbb{R})$ is similar.

We use K to denote both $\mathbb{R}$ and $\mathbb{C}$. We define an equivalence relation on $L^2(X,K)$ in the following way $f \sim g$ iff $\lambda^n(\{x \in X | f(x) \neq g(x)\}) = 0$. Then we define $V = L^2(X,K)/\sim$. We define an inner product on the vector space V. $\langle\ ,\ \rangle : V \times V \to K$, $(f,g) \mapsto \int_X f\overline{g}\mathrm{d}\lambda^n$. Since $f, g \in V$, $|\int_X f\overline{g}\mathrm{d}\lambda^n| \leq \int_X |f\overline{g}|\mathrm{d}\lambda^n \leq \frac{1}{2}\int_X |f|^2 + |g|^2\mathrm{d}\lambda^n < +\infty$. It is easy to show that $\langle\ ,\ \rangle$ is bilinear.(It is symmetric if $K = \mathbb{R}$ and Hermitian if $K = \mathbb{C}$.)

The only non-trivial thing is $0 = \langle f, f\rangle = \int_X |f|^2\mathrm{d}\lambda^n$ implies $f \equiv 0$ almost everywhere, $f \sim 0$. This means if we don't use $\sim$ that we cannot say that $\langle\ ,\ \rangle$ is positive definite.

Exercise

Check $L^2(X,K)$ forms a vector space.

Check $\sim$ gives an equivalence relation. And check plus, scalar multiply and $\langle\ ,\ \rangle$ is well defined w.r.t. $\sim$.(V is an inner product space.)

Check $|\int_X f\overline{g}\mathrm{d}\lambda^n| \leq \int_X |f\overline{g}|\mathrm{d}\lambda^n$.

Why $0 = \int_X |f|^2\mathrm{d}\lambda^n$ implies $f \equiv 0$ almost everywhere?

Answer

(1) See 6.10.14 thus $L^2(X,K)$ forms a vector space. □

(2) Reflexive and symmetric is trivial.

Let $f, g, h \in L^2(X,K)$. Notice that if $f(x) \neq h(x)$ then either $f(x) \neq g(x)$ or $g(x) \neq h(x)$. Hence $\{x \in X | f(x) \neq h(x)\} \subseteq \{x \in X | f(x) \neq g(x)\} \bigcup \{x \in X | g(x) \neq h(x)\}$. Thus $\lambda^n(\{x \in X | f(x) \neq h(x)\}) = 0$.

The reason why plus and scalar multiply is well defined is similar to above. And since $\int_X f\overline{g} - f'\overline{g}' = 0$, the inner product is well defined. □

(3) Let's prove $|\int_X f\mathrm{d}\mu| \leq \int_X |f|\mathrm{d}\mu$.

Use I to denote $\int_X f\mathrm{d}\mu$ and take $\theta = \arg(I)$. Then $|I| = |Ie^{-i\theta}| = |\int_X fe^{-i\theta}\mathrm{d}\mu| = |\Re(\int_X fe^{-i\theta}\mathrm{d}\mu)| = |\Re(\int_X \Re(fe^{-i\theta}) + i\int_X \Im(fe^{i\theta})\mathrm{d}\mu)| = |\int_X \Re(fe^{i\theta})| \leq |\int_X |fe^{-i\theta}|\mathrm{d}\mu| = \int_X |f|\mathrm{d}\mu$ □

(4) In fact, we can prove a stronger conclusion, the seminorm of $L^p(\Omega, \mathscr{A}, \mu)/\sim$ $\|f\|_{L^p} := (\int_\Omega |f(x)|^p\mathrm{d}\mu)^{\frac{1}{p}}$ is positive definite.(Forms a norm)

We needs to prove $\int_{\Omega} |f(x)|^p \mathrm{d}\mu = 0$ iff $f = 0$ almost everywhere.

Suppose $\int_{\Omega} |f(x)|^p \mathrm{d}\mu = 0$. Take $E_n = \{x \in \Omega | |f(x)| > \frac{1}{n}\}$ then $\bigcup_{n\in\mathbb{N}} E_n = \{x \in \Omega | f(x) \neq 0\}$. $\forall n \in \mathbb{N}$, $0 = \int_{\Omega} |f(x)|^p \mathrm{d}\mu \geq \int_{E_n} |f(x)|^p \mathrm{d}\mu = \frac{1}{n}\mu(E_n)$, hence $\mu(E_n) = 0, \forall n$. Therefore, $\mu(\{x \in \Omega | f(x) \neq 0\}) \leq \sum_{n\in\mathbb{N}} \mu(E_n) = 0$.

The other direction is trivial. $\square$

Example

Consider the following integrals, $m, n \in \mathbb{Z}$. $\int_{-\pi}^{\pi} e^{imx} e^{-inx} \mathrm{d}x = \begin{cases} 0 & \text{if } m \neq n \\ 2\pi & \text{if } m = n \end{cases}$.

Hence $\{e^{imx} | m \in \mathbb{Z}\}$ is an orthogonal system for $V = L^2([-\pi, \pi], \mathbb{C})/\sim$.

If you want to make it orthonormal, just consider $\{\frac{1}{\sqrt{2\pi}} e^{imx} | m \in \mathbb{Z}\}$. If you want to replace $[-\pi, \pi]$ by $[-a, a]$ then consider $\{e^{\frac{im\pi x}{a}} | m \in \mathbb{Z}\}$ and this is an orthonormal system for $V = L^2([-a, a], \mathbb{C})/\sim$.

Example

In the real case we consider the following integrals.

$$\int_{-\pi}^{\pi} \cos(mx)\cos(nx)\mathrm{d}x = \begin{cases} 0 & \text{if } m \neq n \\ \pi & \text{if } m = n \neq 0 \\ 2\pi & \text{if } m = n = 0 \end{cases}$$

$$\int_{-\pi}^{\pi} \cos(mx)\sin(nx)\mathrm{d}x = 0$$

$$\int_{-\pi}^{\pi} \sin(mx)\sin(nx)\mathrm{d}x = \begin{cases} 0 & \text{if } m \neq n \text{ or } m = n = 0 \\ \pi & \text{if } m = n \neq 0 \end{cases}$$

Then $\{1, \cos(nx), \sin(mx) | m, n \in \mathbb{N}_{>0}\}$ forms an orthogonal system for $V = L^2([-\pi, \pi], \mathbb{R})/\sim$. If you want to make it orthonormal then $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(mx) | m, n \in \mathbb{N}_{>0}\}$. From $[-\pi, \pi]$ to $[-a, a]$ is similar.

Example

Let $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(mx) | m, n \in \mathbb{N}_{>0}\}$ be an orthonormal system of $L^2([-\pi,\pi],\mathbb{R})$. Take $f \in L^2([-\pi,\pi],\mathbb{R})/\sim$, the fourier series of f is $\dfrac{a_0(f)}{2}+\sum\limits_{k=1}^{\infty} a_k(f)\cos(kx)+b_k(f)\sin(kx)$, where $a_k(f) = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\cos(kx)\mathrm{d}x$ and $b_k(f) = \dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\sin(kx)\mathrm{d}x$.

For instance if $f(x) = x$, then $\begin{cases} a_k = 0 \\ b_k = (-1)^{k+1}\dfrac{2}{k} \end{cases}$

If $f \in L^2([-\pi,\pi],\mathbb{C})/\sim$, the fourier series of f is $\sum c_k e^{ikx}$, where $c_k = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} f(x)e^{ikx}\mathrm{d}x = \begin{cases} \frac{1}{2}(a_k+ib_k) \text{ if } k \geq 0 \\ \frac{1}{2}(a_{-k}-ib_{-k}) \text{ if } k < 0 \end{cases}$.

Notice that the Fourier series of complex valued function or real valued function can be written in a united way $\sum\limits_{-\infty}^{+\infty} c_k e^{ikx}$.

10.1.6 Def

A first-countable space is a topological space (X,τ) for each point $x \in X$, there exists a countable collection $\{U_n\}_{n\in\mathbb{N}}$ of open sets containing x such that for any open set $V \in \tau$ containing x, there exists an n with $U_n \subseteq V$.

Example

Metric space is a first-countable space.

10.1.7 Prop

Let (X,τ) be a first-countable space and let A be a subset of X, then $\overline{A} = \{a \in X | \exists (a_n)_{n\in\mathbb{N}} \subseteq A^{\mathbb{N}} \text{ s.t. } \lim\limits_{n\to+\infty} a_n = a\}$

Proof

Trivial. □

10.1.8 Prop

Let $\{l_k\}$ be an orthogonal system on V. Take $x \in V$ and assume $\sum\limits_k \dfrac{\langle x, l_k\rangle}{\langle l_k, l_k\rangle} l_k \in V$. We use $\overline{x}$ to denote $\sum\limits_k \dfrac{\langle x, l_k\rangle}{\langle l_k, l_k\rangle} l_k$.

Then if we write $x = \overline{x} + h$ then h is orthogonal to $\overline{x}$. Moreover, h is orthogonal to $\overline{<\{l_k\}>}$.

Proof

Since $\langle\ ,\ \rangle$ is continuous in $V \times V$, it is enough to show that $\langle h, l_m \rangle = 0$, $\forall m$. $\langle h, l_m \rangle = \langle x, l_m \rangle - \langle \sum_k \frac{\langle x, l_k \rangle}{\langle l_k, l_k \rangle} l_k, l_m \rangle = \langle x, l_m \rangle - \langle x, l_m \rangle = 0$. $\square$

Remark

Assume $\overline{x}$ exists. By Pythagoras, since $x = \overline{x} + h$, $\|x\|^2 = \|\overline{x}\|^2 + \|h\|^2 \geq \|\overline{x}\|^2$. If we write that inequality w.r.t. the Fourier coefficent, we get Bessel's inequality $\|x\|^2 \geq \|\overline{x}\|^2 = \sum_k |\frac{\langle x, l_k \rangle}{\langle l_k, l_k \rangle}|^2 \langle l_k, l_k \rangle = \sum_k \frac{|\langle x, l_k \rangle|^2}{\langle l_k, l_k \rangle}$.

10.1.9 Corollary

When V is a Hilbert space, the Fourier series is convergent($\overline{x}$ exists).

Proof

In fact, we don't need the Fourier series converges to prove Bessel's inequality. $\overline{x}$ exists where $\{l_k\}$ is finite. In this case $\sum_{k=1}^{n} \frac{|\langle x, l_k \rangle|^2}{\langle l_k, l_k \rangle} \leq \|x\|^2$ is true(Consider $\langle x - \overline{x}, x - \overline{x} \rangle \geq 0$). So we can just pass to the limit.

Use x_k to denote $\frac{\langle x, l_k \rangle}{\langle l_k, l_k \rangle}$. From Bessel's inequality we know $\sum_k |x_k|^2 \langle l_k, l_k \rangle$ converges in $\mathbb{R}$, which means $\forall \varepsilon > 0$, $\exists N$, $\forall m, n \geq N$, $\|x_m l_m + \cdots + x_n l_n\| = \sqrt{|x_m|^2 \langle l_m, l_m \rangle + \cdots + |x_n|^2 \langle l_n, l_n \rangle} < \sqrt{\varepsilon}$. Hence $\sum_k x_k l_k$ is a Cauchy sequence in V, which is convergent. $\square$

10.1.10 Prop

Let $\{l_k\}$ be an orthogonal system. Take $x \in V$ and assume that $\overline{x} \in V$. Then for any $y = \sum_k \alpha_k l_k$, $\|x - \overline{x}\| \leq \|x - y\|$ and the equality is ture only if $\overline{x} = y$.

Proof

$\|x - y\|^2 = \|x - \overline{x}\|^2 + \|\overline{x} - y\|^2$ $\square$

10.2 Complete vector families and basis

10.2.1 Def

A family of vectors $\mathcal{F} = \{x_\alpha | \alpha \in A\}$ in a normed vector space V is said to be complete in a subset $E \subseteq V$ if every vector $x \in E$ can be approximated with arbitrary accuracy by a finite linear combination of elements of $\mathcal{F}$.

Example

$V = L^2([a,b], K)/\sim$ and the family $\mathcal{F} = \{x^k | k \in \mathbb{N}\}$ then $\mathcal{F}$ is complete in $C^0([a,b])$.

10.2.2 Theorem(Weierstrass approximation theorem)

Let $f \in C^0([a,b])$, $\forall \varepsilon > 0$ there exists a polynomial $P \in K[x]$ s.t. $\forall x \in [a,b]$ we have $|f(x) - P(x)| < \varepsilon$.

Proof

By a linear change of variables, it suffices to prove the theorem for $[0,1]$. Let $t = \dfrac{x-a}{b-a}$. Then $x = a + (b-a)t$ and $f(x) = f(a + (b-a)t)$. If there exists a polynomial $Q(t)$ approximating $f(a+(b-a)t)$ on $[0,1]$, then $P(x) = Q(\dfrac{x-a}{b-a})$ is a polynomial approximating $f(x)$ on $[a,b]$.

Define the n-th Bernstein polynomial of f as: $B_n(f,x) = \sum\limits_{k=0}^{n} f(\dfrac{k}{n}) \dbinom{n}{k}$ $x^k(1-x)^{n-k}$. Using the binomial theorem, $\sum\limits_{k=0}^{n} \dbinom{n}{k} x^k (1-x)^{n-k} = 1$. Thus $|f(x) - B_n(f,x)| = |\sum\limits_{k=0}^{n} (f(x) - f(\dfrac{k}{n})) \dbinom{n}{k} x^k (1-x)^{n-k}|$.

Since f is continuous on $[0,1]$, it is uniformly continuous. For $\varepsilon > 0$, there exists $\delta > 0$ such that $|x-y| < \delta$ implies $|f(x) - f(y)| < \dfrac{\varepsilon}{2}$. Split the sum into two parts

$$\sum_{|x-\frac{k}{n}|<\delta} \quad \text{and} \quad \sum_{|x-\frac{k}{n}|\geq\delta}.$$

For the first sum

$$\left| \sum_{|x-\frac{k}{n}|<\delta} (f(x) - f(\frac{k}{n})) \binom{n}{k} x^k (1-x)^{n-k} \right| \leq \frac{\varepsilon}{2} \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \frac{\varepsilon}{2}.$$

Let $M = \sup_{x\in[0,1]} |f(x)|$. And for the second sum, there are many calculations.

Notice that $na(a+b)^{n-1} = \sum_{k=1}^{n} k\binom{n}{k}a^k b^{n-k}$ and $na((a+b)^{n-1} + (n-1)(a+b)^{n-2}) = \sum_{k=1}^{n} k^2\binom{n}{k}a^k b^{n-k}$.

$$\begin{aligned}
&\sum_{k=0}^{n}\left(x-\frac{k}{n}\right)^2\binom{n}{k}x^k(1-x)^{n-k}\\
&=\sum_{k=0}^{n}\left(x^2-2x\frac{k}{n}+\frac{k^2}{n^2}\right)\binom{n}{k}x^k(1-x)^{n-k}\\
&=x^2\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}-\frac{2x}{n}\sum_{k=0}^{n}k\binom{n}{k}x^k(1-x)^{n-k}+\frac{1}{n^2}\sum_{k=0}^{n}k^2\binom{n}{k}x^k(1-x)^{n-k}\\
&=x^2\cdot 1-\frac{2x}{n}\cdot nx+\frac{1}{n^2}(n(n-1)x^2+nx)\\
&=\frac{x(1-x)}{n}\\
&\le\frac{1}{4n}
\end{aligned}$$

For $|x-\frac{k}{n}|\ge\delta$, we have

$$\sum_{|x-\frac{k}{n}|\ge\delta}\binom{n}{k}x^k(1-x)^{n-k}\le\frac{1}{\delta^2}\sum_{k=0}^{n}\left(x-\frac{k}{n}\right)^2\binom{n}{k}x^k(1-x)^{n-k}\le\frac{1}{4n\delta^2}.$$

Thus

$$\left|\sum_{|x-\frac{k}{n}|\ge\delta}\left[f(x)-f\left(\frac{k}{n}\right)\right]\binom{n}{k}x^k(1-x)^{n-k}\right|\le 2M\cdot\frac{1}{4n\delta^2}.$$

Choose N such that $n\ge N$ implies $\frac{M}{2n\delta^2}<\frac{\varepsilon}{2}$. Then

$$|f(x)-B_n(f,x)|<\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon\quad\text{for all}\quad x\in[0,1].$$

$\square$

Remark

$$\|f-P\|_{L^2}=\sqrt{\int_a^b(f(x)-P(x))^2\mathrm{d}\mu}\le\varepsilon\sqrt{b-a}$$

10.2.3 Prop

Let V be a Hilbert space over K and $\{l_k\}$ be an orthogonal system. TFAE:

(1) $\{l_k\}$ is complete in $E \subseteq V$.

(2) $\forall x \in E$, we have $x = \sum\limits_k \dfrac{\langle x, l_k \rangle}{\langle l_k, l_k \rangle} l_k$.

(3) $\forall x \in E$ satisfies $\|x\|^2 = \sum\limits_k \dfrac{|\langle x, l_k \rangle|^2}{\langle l_k, l_k \rangle}$.

Proof

From (1) to (2) : Since V is complete, $\overline{x} = \sum\limits_k \dfrac{\langle x_k, l_k \rangle}{\langle l_k, l_k \rangle} l_k \in V$.

$\forall \varepsilon > 0$, $\exists$ a finite combination $\alpha_1 l_1 + \cdots + \alpha_m l_m$ s.t. $\|x - (\alpha_1 l_1 + \cdots + \alpha_m l_m)\| < \varepsilon$. Since $\|x - \overline{x}\| \leq \|x - (\alpha_1 l_1 + \cdots + \alpha_m l_m)\| < \varepsilon$, $x = \overline{x}$.

From (2) to (3) : Pythagoras.

From (3) to (1) : $x - \sum\limits_{k=1}^{n} \dfrac{\langle x, l_k \rangle}{\langle l_k, l_k \rangle} l_k = h + \sum\limits_{k>n} \dfrac{\langle x, l_k \rangle}{\langle l_k, l_k \rangle} l_k$. Then consider $\langle x - \sum\limits_{i=1}^{n} x_k l_k, \sum\limits_{k=1}^{n} x_k l_k \rangle = \langle h + \sum\limits_{k>n} x_k l_k, \sum\limits_{k=1}^{n} x_k l_k \rangle = 0$.

So we use Pythagoras $\|x - \sum\limits_{k=1}^{n} x_k l_k\|^2 = \|x\|^2 - \|\sum\limits_{k=1}^{n} x_k l_k\|^2$. By (3) the RHS can be made arbitrary small. This proves the claim. □

10.2.4 Def

Let V be a normed vector space over K.

A family of vectors $\{b_k\}$ is a (Hamel) basis of V if $\forall v \in V$, $\exists$ a unique sequence $\{\alpha_k\}$ in K with $\alpha_k = 0$ for all but finitely many k s.t. $v = \sum\limits_k \alpha_k b_k$.

A countable family of vectors $\{b_k\}$ is a (Schauder) basis for V if $\forall v \in V$, $\exists$ a unique sequence $\{\alpha_k\}$ in K s.t. $v = \sum\limits_k \alpha_k b_k$(as convergent series).

Remark

A Schauder basis is a complete family of vectors in V.

Remark

Let $\{l_k\}$ be a complete orthogonal system in a Hilbert space V. Then any $x \in V$ can be written as $x = \sum\limits_k \alpha_k l_k$ where α_k are the Fourier coefficent.

But in general it is false that a complete family of vectors(not orthogonal) in a Hilbert space is a Schauder basis. Since in many situations $\{\sum_k a_k l_k\} \subsetneq \overline{<\{l_k\}>}$.

Reminder

Before we introduce a concrete counter example, we need to review some theorems about the convergence of power series. In fact, all the following contents in this reminder can be found in section 5.6.

The first theorem is **Cauchy criterion for uniform convergence**. A series $\sum_{n=1}^{\infty} u_n(x)$ converges uniformly on I iff $\forall \varepsilon > 0$, $\exists N_0 \in \mathbb{N}$ s.t. $\forall N > n > N_0$ and $\forall x \in I$, $\left|\sum_{k=n+1}^{N} u_k(x)\right| < \varepsilon$.. The only if case is trivial and the if case just take $N \to +\infty$.

Abel's Theorem is the second theorem, it indicates that if a power series $\sum_{n=0}^{\infty} a_n x^n$ converges at a point $x = R \neq 0$, then it converges absolutely for any x s.t. $|x| < R$ and uniformly on every closed interval $[-r, r]$ where $r < R$. The first half of the sentence can be illustrated by the convergence of geometric series and the the second half of the sentence is obvious after using the Cauchy convergence criterion.

From above theorem we can get an important insight that if a power series converges uniformly on $[-r, r]$, then its term-by-term differentiation also converges uniformly on $[-r, r]$.

The last theorem(**Weierstrass**) is about the derivative of power series. If the power series $\sum_{n=0}^{\infty} a_n x^n$ is convergent at $x = R > 0$, then within the interval $(-R, R)$, $S(x) = \sum_{n=0}^{\infty} a_n x^n$ is differentiable, and its derivative can be obtained by term-by-term differentiation.

Let $V(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}$. We already know $V(x)$ is well defined in $I = (-R, R)$. To prove $S'(x) = V(x)$, we need to show that for any $x \in I$,

$$\lim_{h \to 0} \frac{S(x+h) - S(x)}{h} = V(x).$$

Since $S(x)$ is the uniformly limit of $\sum\limits_{n=0}^{\infty} a_n x^n$, we have:

$$S(x+h) - S(x) = \int_x^{x+h} V(t)\, dt.$$

Therefore,

$$\frac{S(x+h) - S(x)}{h} - V(x) = \frac{1}{h} \int_x^{x+h} (V(t) - V(x))\, dt.$$

Since the series $\sum\limits_{n=1}^{\infty} u_n'(t)$ converges uniformly to $V(t)$ on I, for any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $t \in I$ and $n > N$,

$$\left| \sum_{k=n+1}^{\infty} u_k'(t) \right| < \frac{\varepsilon}{3}.$$

Define the partial sum $V_N(t) = \sum\limits_{k=1}^{N} u_k'(t)$. Then $V(t) = V_N(t) + \sum\limits_{k=N+1}^{\infty} u_k'(t)$, and thus

$$V(t) - V(x) = (V_N(t) - V_N(x)) + \left(\sum_{k=N+1}^{\infty} u_k'(t) - \sum_{k=N+1}^{\infty} u_k'(x) \right).$$

For the first term $V_N(t) - V_N(x)$, since $V_N(t)$ is a finite sum of differentiable functions, it is continuous and differentiable at x. Thus, as $h \to 0$, there exists $\delta_1 > 0$ such that for all $t \in [x, x+h]$ (or $[x+h, x]$),

$$|V_N(t) - V_N(x)| < \frac{\varepsilon}{3}.$$

For the second term, using the Cauchy criterion for uniform convergence, when $|h| < \delta_2$, the interval length is sufficiently small such that t is close to x. Combining this with the uniform convergence, there exists $\delta_2 > 0$ such that

$$\left| \sum_{k=N+1}^{\infty} u_k'(t) - \sum_{k=N+1}^{\infty} u_k'(x) \right| \leq \left| \sum_{k=N+1}^{\infty} u_k'(t) \right| + \left| \sum_{k=N+1}^{\infty} u_k'(x) \right| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3}.$$

Taking $\delta = \min\{\delta_1, \delta_2\}$, when $|h| < \delta$, we have

$$\left| \frac{1}{h} \int_x^{x+h} (V(t) - V(x))\, dt \right| < \varepsilon.$$

Therefore,

$$\lim_{h\to 0} \frac{S(x+h)-S(x)}{h} = V(x),$$

which implies $S'(x) = V(x) = \sum_{n=1}^{\infty} u_n'(x)$, proving term-by-term differentiation.

By induction, we can easily prove if $f(x)$ is the uniform limit of $\sum\limits_{n\in\mathbb{N}} a_n x^n$ in $(-R,R)$ then $f(x)$ is C^∞ in $(-R,R)$.

Counter example

$V = C^0([-1,1],\mathbb{R}) \subseteq L^2([-1,1],\mathbb{R})$, we induce on V the inner product from $L^2([-1,1],\mathbb{R})$, and At this point ,we acknowledge that $L^2([-1,1],\mathbb{R})$ is a Hilbert space.. We have proved $\{x^k | k \geq 0\}$ is a complete family in V and now we show that it is not a Schauder basis.

First we show that if $\sum a_k x^k$ converges in the L^2 norm then it converges pointwise in $]-1,1[$. In fact, $\|\alpha_k x^k\|_{L^2} = \sqrt{\int_{-1}^{1} (\alpha_k x^k)^2 \mathrm{d}\lambda} = |\alpha_k| \sqrt{\frac{2}{2k+1}}$. Since $\|\alpha_k x^k\|_{L^2} \to 0$, we know $\alpha_k^2 < 2k+1$ for k large enough. Take $x \in]-1,1[\setminus\{0\}$, consider $\sum\limits_{k\geq N} |\alpha_k x^k| \leq \sum\limits_{k\geq N} \sqrt{2k+1}|x^k|$. $\limsup\limits_{k\to+\infty} \frac{|x^{k+1}|\sqrt{2k+3}}{|x^k|\sqrt{2k+1}} = |x| < 1$. It means that the series converges pointwisely in $]-1,1[$.

Then from previous lectures we know $f = \sum a_k x^k$ would belong to $C^\infty(]-1,1[)$. Therefore it is enough to take $g \in V \setminus (C^\infty(]-1,1[))$. Such function cannot be a L^2-limit. $\{x^k | k \geq 0\}$ is not a Schauder basis. For instance $f(x) = |x|$ this function cannot be expressed by a limit of of power series.

Example

$V = L^2([-\pi,\pi],\mathbb{R})$, $\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\cos(nx), \frac{1}{\sqrt{\pi}}\sin(mx) | m,n \in \mathbb{N}_{>0}\}$ is an orthonormal system.

$\forall f \in V$ the Fourier series of f is $\frac{a_0(f)}{2} + \sum\limits_{k=1}^{\infty} a_k(f)\cos(kx) + b_k(f)\sin(kx)$, where $a_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\mathrm{d}x$ and $b_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\mathrm{d}x$.

From Bessel's inequality we know $\frac{a_0^2(f)}{2} + \sum\limits_{k=1}^{\infty} a_k^2(f) + b_k^2(f) \leq \frac{1}{\pi}\int_{-\pi}^{\pi} f^2 \mathrm{d}\lambda = \frac{1}{\pi}\|f\|^2$. Hence we know $\sum\limits_{k=1}^{\infty} \frac{\sin(kx)}{\sqrt{k}}$ won't by any Fourier series of $f \in V$.

10.3 Interlude : Complement of L^p space

Reminder

Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. We have proved that $L^p(\Omega, \mathscr{A}, \mu)$ is a vector space with a seminorm $\|f\|_{L^p} := (\int_\Omega |f(x)|^p \mathrm{d}\mu)^{\frac{1}{p}} (f : \Omega \to \mathbb{R})$. And in the first section of this chapter we have proved the quotient space $L^2(\Omega, \mathscr{A}, \mu)/\sim$ $(f : X \subseteq \mathbb{R}^n \to K)$ is a normed vector space with the equivalence relation mentioned before.

Without causing ambiguity, we use $L^p(X, \mu)$ to represent $L^p(X, \mu)/\sim$.

10.3.1 Theorem

Let $(\Omega, \mathscr{A}, \mu)$ be a measure space. If $p > 1$ then $L^p(\Omega, \mathscr{A}, \mu)$ forms a Banach space.

Proof

We have to show that any Cauchy sequence (w.r.t. $\|\cdot\|_{L^p}$) is convergent.

Consider $\{f_n\} \subseteq L^p(\Omega, \mathscr{A}, \mu)$ Cauchy sequence. Choose $n_1 < n_2 < \cdots < n_j$ s.t. $\|f_n - f_m\|_{L^p} \leq 2^{-2j}$ for $m, n \geq n_j$. Define $g_j := f_{n_j}$, $h_k := g_{k+1} - g_k$, $\forall k \geq 1$. Now we study $\{h_k\}_{k \in \mathbb{N}}$ and the goal is to find the limit of $\{g_j\}_{j \in \mathbb{N}}$. Notice that $\int_\Omega |h_k(x)|^p \mathrm{d}\mu = \|h_k\|_{L^p}^p = \|g_{k+1} - g_k\|_{L^p}^p \leq 2^{-2pk}$.

We show that $\sum\limits_{k=1}^\infty h_k(x)$ converges pointwise almost everywhere in Ω. Let $a_k(x) = h_k(x) 2^{\frac{k}{p}}$, $b_k = 2^{-\frac{k}{p}}$. Then $\sum\limits_{k=1}^\infty |h_k(x)| = \sum\limits_{k=1}^\infty |a_k(x) b_k(x)| \leq (\sum\limits_{k=1}^\infty |a_k(x)|^p)^{\frac{1}{p}} (\sum\limits_{k=1}^\infty |b_k(x)|^q)^{\frac{1}{q}} = (\sum\limits_{k=1}^\infty 2^k |h_k(x)|^p)^{\frac{1}{p}} (\sum\limits_{k=1}^\infty 2^{\frac{-kq}{p}})^{\frac{1}{q}} \leq C(\sum\limits_{k=1}^\infty 2^k |h_k(x)|^p)^{\frac{1}{p}}$.
Therefore, $\int_\Omega (\sum\limits_{k=1}^\infty |h_k(x)|)^p \mathrm{d}\mu \leq C^p \int_\Omega \sum\limits_{k=1}^\infty 2^k |h_k(x)|^p \mathrm{d}\mu = C^p \sum\limits_{k=1}^\infty \int_\Omega 2^k |h_k(x)|^p \mathrm{d}\mu \leq C^p \sum\limits_{k=1}^\infty 2^k 2^{-2pk} = C^p \sum\limits_{k=1}^\infty 2^{-k(2p-1)} < +\infty$

$\int_\Omega (\sum\limits_{k=1}^\infty |h_k(x)|)^p < +\infty$ implies $\sum\limits_{k=1}^\infty |h_k(x)| < +\infty$ almost everywhere. In fact, without loss of generality, let $f : \Omega \to \mathbb{R}$ measurable, if $\int_\Omega |f|^p \mathrm{d}\mu < +\infty$ then $|f|^p$ is bounded almost everywhere. $E_n = \{x \in X \,|\, |f(x)|^p > n\}$, $C \geq \int_\Omega |f(x)|^p \mathrm{d}\mu \geq \int_{E_n} |f(x)|^p \ \mathrm{d}\mu \geq n\mu(E_n)$, thus $\mu(E_n) < \dfrac{C}{n}$, thus

$\lim\limits_{n\to+\infty} \mu(E_n) = 0$. $E = \bigcap\limits_{n=1}^{\infty} E_n = \{x \in X | |f(x)|^p = +\infty\}$, $\mu(E) = \lim\limits_{n\to+\infty} \mu(E_n) = 0$. This means $\{\sum h_k(x)\}$ absolutely converges(almost everywhere). Since K is complete then $\{\sum h_k(x)\}$ converges almost everywhere.

Define $f(x) = \begin{cases} g_1(x) + \sum\limits_{k=1}^{\infty} h_k(x) & \text{when the series converges} \\ 0 & \text{otherwise} \end{cases}$

Fix an index k, $2^{-2kp} \geq \liminf\limits_{j\to+\infty} \int_\Omega |g_j(x)-g_k(x)|^p \mathrm{d}\mu \geq \int_\Omega \liminf\limits_{j\to+\infty} |g_j(x)-g_k(x)|^p \mathrm{d}\mu = \int_\Omega |f(x) - g_k(x)|^p \mathrm{d}\mu$, thus $\|f - g_k\|_{L^p} \leq 2^{-2k}$. Notice that for $k = 1$, $f - g_1 \in L^p(\Omega, \mathscr{A}, \mu)$ so $f \in L^p(\Omega, \mathscr{A}, \mu)$. □

Exercise

Prove the upper continuity of measure in the $\exists E_n$ s.t. $\mu(E_n) < +\infty$ case.

Answer

Without loss of generality, assume $\mu(E_1) < \infty$ (otherwise, consider the sequence starting from E_n). Define $F_n = E_1 \backslash E_n$. Then $\{F_n\}$ is an increasing sequence of sets, and

$$\bigcup_{n=1}^{\infty} F_n = E_1 \setminus \bigcap_{n=1}^{\infty} E_n.$$

By the lower continuity of measures,

$$\lim_{n\to\infty} \mu(F_n) = \mu\left(\bigcup_{n=1}^{\infty} F_n\right).$$

Since $\mu(E_1) < +\infty$, by the subtractivity of measures,

$$\mu(F_n) = \mu(E_1) - \mu(E_n),$$

$$\mu\left(\bigcup_{n=1}^{\infty} F_n\right) = \mu\left(E_1 \setminus \bigcap_{n=1}^{\infty} E_n\right) = \mu(E_1) - \mu\left(\bigcap_{n=1}^{\infty} E_n\right).$$

Combining these equalities,

$$\lim_{n\to\infty} [\mu(E_1) - \mu(E_n)] = \mu(E_1) - \mu\left(\bigcap_{n=1}^{\infty} E_n\right).$$

Subtracting $\mu(E_1)$ from both sides yields

$$\lim_{n\to\infty} \mu(E_n) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right).$$

$\square$

10.3.2 Corollary

$L^2([-\pi,\pi],K)$ is a Hilbert space.

10.4 Fourier series

10.4.1 Def

Define $T_n(x) := \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos(kx) + b_k \sin(kx)$. We say $T_n(x)$ is the trigonometric polynomial of order n, where $a_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\mathrm{d}x$ and $b_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\mathrm{d}x$.

Reminder

We can write $T_n(x)$ in a more elegant way $T_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}$, where

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx = \begin{cases} \frac{1}{2}(a_k - ib_k) & k > 0 \\ \frac{1}{2}a_0 & k = 0 \\ \frac{1}{2}(a_{-k} + ib_{-k}) & k < 0 \end{cases}.$$

Example(Dirichlet kernel)

We want to write the Fourier coefficent in a better way, so we needs to do more calculations.

Since $a_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(kx)\mathrm{d}x$ and $b_k(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(kx)\mathrm{d}x$. Then $T_n(x) = \sum_{k=-n}^{n}(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\mathrm{d}t)e^{ikx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)(\sum_{k=-n}^{n} e^{ik(x-t)})\mathrm{d}t$

Now define $D_n(u) = \sum_{k=-n}^{n} e^{iku} = e^{-inu}\frac{e^{iu(2n+1)} - 1}{e^{iu} - 1} = \frac{\sin((n+\frac{1}{2})u)}{\sin(\frac{1}{2}u)}$, which is called Dirichlet kernel, since $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(u)\mathrm{d}u = \frac{1}{\pi}\int_{0}^{\pi} D_n(u)\mathrm{d}u = 1$

Extend this function by periodicity $L = 2\pi$, $\tilde{f}(x) = f(x - kL)$, k is the unique integer s.t. $x = x_0 + kL$, $x_0 \in [-\pi, \pi]$(If the values at the original endpoints are not equal, the extended function will have jump discontinuity points, and it is still integrable). Let's go back to T_n, putting $u = x - t$, $T_n(x) = \frac{1}{2\pi}\int_{x-\pi}^{x+\pi} f(x-u)D_n(u)\mathrm{d}u = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-u)D_n(u)\mathrm{d}u$.

Now use that D_n is an even function $T_n(x) = \frac{1}{2\pi}\int_{0}^{\pi}(f(x-u) + f(x+u))D_n(u)\mathrm{d}u$

10.4.2 Lemma(Riemann-Lebesgue)

Let $f : [a, b] \to \mathbb{R}$ be an integrable function($\int_a^b |f| d\lambda < +\infty$), then $\lim_{\lambda \to +\infty} \int_a^b f(x)e^{i\lambda x} dx = 0$.

Proof

Suppose f is a simple function. $f(x) = \sum_{k=1}^{n} c_k \mathbb{1}_{[x_{k-1}, x_k[}(x)$. $a = x_0 < x_1 < \cdots < x_n = b$. $\int_a^b f(x)e^{i\lambda x} = \sum_{k=1}^{n} c_k \int_{x_{k-1}}^{x_k} e^{i\lambda x} dx = \sum_{k=1}^{n} c_k (\int_{x_{k-1}}^{x_k} \cos(\lambda x) dx + i \int_{x_{k-1}}^{x_k} \sin(\lambda x) dx) = \sum_{k=1}^{n} c_k (\frac{1}{\lambda} \sin(\lambda x)|_{x_{k-1}}^{x_k} - \frac{i}{\lambda} \cos(\lambda x)|_{x_{k-1}}^{x_k}) = \sum_{k=1}^{n} c_k \frac{e^{i\lambda x_k} - e^{i\lambda x_{k-1}}}{i\lambda}$ $\to 0$ when $\lambda \to +\infty$.

Since f is integrable, f and its integral can be approximated arbitrariely by simple functions. For any $\varepsilon > 0$, $\exists \phi : [a, b] \to \mathbb{R}$ simple s.t. $\int_a^b |f(x) - \phi(x)| dx < \varepsilon$.

$|\int_a^b f(x)e^{i\lambda x} dx| = |\int_a^b (f(x) - \phi(x))e^{i\lambda x} dx| + |\int_a^b \phi(x)e^{i\lambda x} dx| \leq \int_a^b |(f(x) - \phi(x))e^{i\lambda x}| dx + |\int_a^b \phi(x)e^{i\lambda x} dx| \leq \varepsilon + |\int_a^b \phi(x)e^{i\lambda x} dx|$. Then $\limsup_{\lambda \to +\infty} |\int_a^b f(x)e^{i\lambda x} dx| \leq \varepsilon$ for $\forall \varepsilon > 0$. □

10.4.3 Corollary

If $f : [a, b] \to \mathbb{R}$ is integrable, then $\lim_{\lambda \to +\infty} \int_a^b f(x) \cos(\lambda x) dx = 0$ and $\lim_{\lambda \to +\infty} \int_a^b f(x) \sin(\lambda x) dx = 0$.

10.4.4 Corollary

If $f : [a, b] \to \mathbb{C}$ is integrable($\int |\Re(f)| dx < +\infty$ and $\int |\Im(f)| dx < +\infty$, which is equivalent to $\int |f| < +\infty$), then $\lim_{\lambda \to +\infty} \int_a^b f(x) \cos(\lambda x) dx = 0$ and $\lim_{\lambda \to +\infty} \int_a^b f(x) \sin(\lambda x) dx = 0$.

10.4.5 Theorem(Localization principle)

Let $f, g \in L^2([-\pi, \pi], K)$. If f and g coincides in a neighborhood of $x_0 \in]-\pi, \pi[$. Then the Fourier series of f and g either both diverge at x_0 or both converges at x_0. Moreover if they converge at x_0 then the two limits are the same.

Proof

Since $\|f \cdot g\|_{L^1} \leq \|f\|_{L^p}\|g\|_{L^q}$, we know $\|f\|_{L^1} \leq \sqrt{2\pi}\|f\|_{L^2} < +\infty$, thus f is integrable.

The n-th partial sum of the Fourier series of f at x is given by:

$$T_{n,f}(x) = \frac{1}{2\pi}\int_0^{\pi} [f(x-t) + f(x+t)] \frac{\sin\left((n+\frac{1}{2})t\right)}{\sin\left(\frac{1}{2}t\right)}\, dt.$$

Fix $\delta > 0$ such that $f = g$ on $[x_0 - \delta, x_0 + \delta]$.

Define the difference function:

$$\phi(t) = [f(x_0 - t) + f(x_0 + t)] - [g(x_0 - t) + g(x_0 + t)].$$

Since $f = g$ on $[x_0 - \delta, x_0 + \delta]$, $\phi(t) = 0$ for $t \in [0, \delta]$. Therefore:

$$\int_0^{\delta} \phi(t)\frac{\sin\left((n+\frac{1}{2})t\right)}{\sin\left(\frac{1}{2}t\right)}\, dt = 0.$$

Thus,

$$T_{n,f}(x_0) - T_{n,g}(x_0) = \frac{1}{2\pi}\int_0^{\pi} \phi(t)\frac{\sin\left((n+\frac{1}{2})t\right)}{\sin\left(\frac{1}{2}t\right)}\, dt = \frac{1}{2\pi}\int_{\delta}^{\pi} \phi(t)\frac{\sin\left((n+\frac{1}{2})t\right)}{\sin\left(\frac{1}{2}t\right)}\, dt.$$

The function $\dfrac{\phi(t)}{2\pi\sin\left(\frac{1}{2}t\right)}$ is integrable on $[\delta, \pi]$(bounded), so by the Riemann-Lebesgue lemma:

$$\lim_{n\to+\infty} [T_{n,f}(x_0) - T_{n,g}(x_0)] = 0.$$

This implies that the Fourier series of f and g either both diverge or both converge at x_0, and if they converge, they converge to the same limit. □

Remark

The theorem is still true for $x_0 = -\pi, \pi$ by extending f, g by periodicity.

10.4.6 Def

Define $U_x := [-\delta, \delta] \subseteq \mathbb{R}$ and $U_x^0 := [-\delta, x[\bigcup]x, \delta[\subseteq U_x$. Let $f : U_x^0 \to K$. We say that f satisfies the Dini conditions at x if

(a) $f(x_-) := \lim\limits_{t\to 0^+} f(x-t)$ and $f(x_+) := \lim\limits_{t\to 0^+} f(x+t)$ both exists and finite.

(b) $\exists \varepsilon > 0$ s.t. $\displaystyle\int_0^{\varepsilon} |\frac{(f(x-t)-f(x_-))+(f(x+t)-f(x_+))}{t}|\mathrm{d}t < +\infty$

Example

Suppose $f : U_x \to K$ continuous and satisfies the Holder inequality. $|f(x+t)-f(x)| \leq M|t|^{\alpha}$, for $M > 0$ and $0 < \alpha \leq 1$. Therefore, $|\dfrac{f(x+t)-f(x)}{t}| \leq \dfrac{M}{|t|^{1-\alpha}}$ So condition (b) is satisfied.

Moreover, $f : U_x^0 \to K$ and $f(x_-)$ and $f(x_+)$ exists and we assume $|f(x+t)-f(x_+)| \leq Mt^{\alpha}$ and $|f(x-t)-f(x_-)| \leq Mt^{\alpha}$ for $t > 0$, $M > 0$ and $0 < \alpha \leq 1$. Also in this case Dini's condition is true.

10.4.7 Theorem(Pointwise convergence of Fourier series)

Let $f : \mathbb{R} \to K$ be a periodic function of period $T = 2\pi$, s.t. f is integrable in $[-\pi,\pi]$. If f satisfies the Dini's conditions at $x \in \mathbb{R}$, then its Fourier series converges at x and $\sum\limits_{-\infty}^{+\infty} c_k(f)e^{i\lambda x} = \dfrac{f(x_-)+f(x_+)}{2}$

Proof

Recall the n-th partial sum of the Fourier series:

$$T_n(x) = \sum_{\lambda=-n}^{n} c_{\lambda}(f)e^{i\lambda x} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\frac{\sin\left((n+\frac{1}{2})t\right)}{\sin\left(\frac{1}{2}t\right)}\,dt.$$

We aim to show that:

$$T_n(x) - \frac{f(x_-)+f(x_+)}{2} \to 0 \quad \text{as} \quad n \to +\infty.$$

Using the periodicity and symmetry, we rewrite the partial sum as:

$$T_n(x) = \frac{1}{\pi}\int_0^{\pi} [f(x-t)+f(x+t)]\,\frac{\sin\left((n+\frac{1}{2})t\right)}{2\sin\left(\frac{1}{2}t\right)}\,dt.$$

Thus, the difference becomes:

$$T_n(x) - \frac{f(x_-)+f(x_+)}{2} = \frac{1}{\pi}\int_0^{\pi} \frac{[f(x-t)-f(x_-)]+[f(x+t)-f(x_+)]}{2\sin\left(\frac{1}{2}t\right)}\sin((n+\frac{1}{2})t)\,dt.$$

Define the function:

$$\Phi(t) = \frac{[f(x-t)-f(x_-)]+[f(x+t)-f(x_+)]}{2\sin\left(\frac{1}{2}t\right)}.$$

. We needs to prove $\Phi(t)$ is integrable in $[0,\pi]$.

Since $\lim_{x\to 0}\frac{x}{\sin(x)}=1$, $\exists\delta>0$ s.t. $\forall x\in[-\delta,\delta]$, $|\frac{x}{2\sin(\frac{x}{2})}|\le 2$.

$$\int_0^\delta |\Phi(t)| = \int_0^\delta |\frac{(f(x-t)-f(x_-))+(f(x+t)-f(x_+))}{t}||\frac{t}{2\sin(\frac{1}{2}t)}|\mathrm{d}t<+\infty$$

Therefore, $\Phi(t)$ is integrable on $[0,\delta]$.

On the interval $[\delta,\pi]$, the function $\Phi(t)$ is bounded because f is integrable and $\sin\left(\frac{1}{2}t\right)\ge\sin\left(\frac{\delta}{2}\right)>0$. Hence, $\Phi(t)$ is integrable on $[\delta,\pi]$. $\square$

Example(Fejer kernel)

Consider a function $f:\mathbb{R}\to K$ of period 2π. Let $T_0(x),\cdots,T_n(x)$ be the trigonometric polynomials of f of order $0,1,\cdots,n$. We define $\sigma_{n,f}=\sigma_n(x):=\frac{T_0(x)+\cdots+T_n(x)}{n+1}$. Recall the expression of $T_i(x)$ in terms of the Dirichlet kernel ($T_n(x)=\int_{-\pi}^{\pi}f(t)D(x-t)\mathrm{d}t$), we get $\sigma_n(x)=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x-t)\mathcal{F}_n(t)\mathrm{d}t$, where $\mathcal{F}_n(x)=\frac{1}{n+1}(D_0(t)+\cdots+D_n(t))$ this is the average of the Dirichlet kernels and it is called Fejer's kernel.

One can claim that $\sum_{k=0}^{n}\sin(k+\frac{1}{2})t=\frac{\sin^2(\frac{n+1}{2})t}{\sin\frac{t}{2}}$. In fact, $2\sin(k+\frac{1}{2})t\sin\frac{t}{2}=\cos kt-\cos(k+1)t$. Hence $\sum_{k=0}^{n}\sin(k+\frac{1}{2})t=\frac{\sin^2(\frac{n+1}{2})t}{\sin\frac{t}{2}}$.

Remember that $D_n(t)=\frac{\sin(n+\frac{1}{2})t}{\sin\frac{1}{2}t}$. Thus we get the following expression of the Fejer kernel $\mathcal{F}_n(t)=\frac{\sin^2(\frac{n+1}{2}t)}{(n+1)\sin^2(\frac{t}{2})}$. Let's compute $\lim_{t\to 0}\mathcal{F}_n(t)$. We use the taylor expansion around zero. $\mathcal{F}_n(t)\sim\frac{(\frac{n+1}{2}t)^2}{(n+1)(\frac{t}{2})^2}=n+1$. So we can extend $\mathcal{F}_n(t)$ at all points $2k\pi$ by putting $\mathcal{F}_n(2k\pi)=n+1$. $\mathcal{F}_n(t)$ becomes a periodic continuous function.

10.4.8 Def

A family of functions $\{K_n\}_{n\ge 0}$ with $K_n:\mathbb{R}\to\mathbb{R}$ is called an approximate identity if

(1) $\int_{-\infty}^{\infty}K_n(t)\mathrm{d}t=1$ for all n.

(2) $K_n(t)\ge 0$, $\forall t\in\mathbb{R}$, $\forall n\ge 0$.

(3) For any $\delta > 0$, $\lim_{n\to+\infty} \int_{|t|>\delta} K_n(t)\mathrm{d}t = 0$.

10.4.9 Prop

Consider $\Delta_n(x) = \begin{cases} \frac{1}{2\pi}\mathcal{F}_n(x) & \text{if } |x| \leq \pi \\ 0 & \text{otherwise} \end{cases}$. Then Δ_n is an approximate identity.

Proof

Remember that $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t)\mathrm{d}t = 1$. $\int_{-\infty}^{\infty} \Delta_n(t)\mathrm{d}t = \int_{-\pi}^{\pi} \Delta_n(t)\mathrm{d}t = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathcal{F}_n(t)\mathrm{d}t = \frac{1}{2\pi(n+1)}\sum_{k=0}^{n}\int_{-\pi}^{\pi} D_k(t)\mathrm{d}t = 1$.

Take $\delta > 0$, $0 \leq \int_{-\infty}^{-\delta} \Delta_n(t)\mathrm{d}t = \int_{\delta}^{+\infty} \Delta_n(t)\mathrm{d}t = \int_{\delta}^{\pi} \Delta_n(t)\mathrm{d}t \leq \frac{1}{2\pi(n+1)}\int_{\delta}^{\pi} \frac{\mathrm{d}t}{\sin^2(\frac{1}{2}t)} \to 0$. □

10.4.10 Theorem(Fejer)

Let $f : \mathbb{R} \to K$ continuous, of period 2π, and integrable in $[-\pi, \pi]$. Then $\sigma_n \rightrightarrows f$.

Proof

We want to prove $\lim_{n\to+\infty} \sup_{x\in[-\pi,\pi]} |\sigma_n(x) - f(x)| = 0$.

Since f is continuous in $[-\pi, \pi]$, it is uniformly continuous on $[-\pi, \pi]$. $\forall\varepsilon > 0$, $\exists\delta > 0$ s.t. $|f(x - t) - f(x)| < \varepsilon$ for $\forall x$ and $\forall |t| < \delta$.

Now $\sigma_n(x) - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x - t)\mathcal{F}_n(t)\mathrm{d}t - f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi}(f(x - t) - f(x))\mathcal{F}_n(t)\mathrm{d}t$. Now split the integral in two parts.

$\frac{1}{2\pi}\int_{|t|\leq\delta}(f(x - t) - f(x))\mathcal{F}_n(t)\mathrm{d}t \leq \varepsilon$.

$\frac{1}{2\pi}\int_{\delta\leq|t|\leq\pi}(f(x - t) - f(x))\mathcal{F}_n(t)\mathrm{d}t \leq M\int_{|t|>\lambda} \mathcal{F}_n(t)\mathrm{d}t \to 0$. □

10.4.11 Corollary(Weierstrass)

Let $f : [-\pi, \pi] \to K$ be a continuous function s.t. $f(-\pi) = f(\pi)$ then such function can be approximated uniformly by σ_n with arbitrary precision.

10.4.12 Theorem

In $V = L^2([-\pi,\pi],K)$, $\{1,\cos(mx),\sin(nx)|m,n\in\mathbb{N}_{>0}\}$ is complete.

Proof

HARD. $\square$