FUNDAMENTAL ALGEBRA & ANALYSIS

Contents

Chapter 1

Differential Calculus

1.1 Landau symbol

In this section, we fix a complete valued field $(K, |\cdot|)$ and a normed vector space $(V, \|\cdot\|)$ over K.

Definition 1.1.1 Let X be a set, $f : X \longrightarrow V$, $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. Let $Y \subseteq X$ be a subset. We use the expression

$$f(x) = \mathcal{O}(g(x))$$

to denote the statement:

$$\exists C > 0, \forall x \in Y, \|f(x)\| \leq C \cdot g(x).$$

Let $\mathcal{F}$ be a filter on X, we use the expression

$$f(x) = \mathcal{O}\left(g(x)\right) \text{ along } \mathcal{F}$$

to denote the statement:

$$\exists C > 0, \exists A \in \mathcal{F}, \|f(x)\| \leq C \cdot g(x), \forall x \in A.$$

We use the expression

$$f(x) = o\left(g(x)\right) \text{ along } \mathcal{F}$$

to denote the statement:

$$\exists \varepsilon : X \longrightarrow \mathbb{R}_{\geq 0}, \exists A \in \mathcal{F}, \lim_{\mathcal{F}} \varepsilon = 0 \text{ and } \forall x \in A, \|f(x)\| \leq \varepsilon(x) g(x).$$

Proposition 1.1.2 Let X be a set and $\mathcal{F}$ be a filter on X.
(1) Let $f : X \longrightarrow V$, $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. If $f(x) = o(g(x))$ along $\mathcal{F}$, then $f(x) = \mathcal{O}(g(x))$ along $\mathcal{F}$.
(2)

1. Let $f_1 : X \longrightarrow V$, $f_2 : X \longrightarrow V$ and $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. If $f_1(x) = \mathcal{O}(g(x))$ and $f_2(x) = \mathcal{O}(g(x))$ along $\mathcal{F}$, then $f_1(x) + f_2(x) = \mathcal{O}(g(x))$ along $\mathcal{F}$.
2. Let $f_1 : X \longrightarrow V$, $f_2 : X \longrightarrow V$ and $g : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. If $f_1(x) = o(g(x))$ and $f_2(x) = o(g(x))$ along $\mathcal{F}$, then $f_1(x)+f_2(x) = o(g(x))$ along $\mathcal{F}$.

(3) Let $\lambda : X \longrightarrow K$, $f : X \longrightarrow V$, $g : X \longrightarrow \mathbb{R}_{\geq 0}$, $h : X \longrightarrow \mathbb{R}_{\geq 0}$ be mappings.

1. If $\lambda(x) = \mathcal{O}\left(g(x)\right)$ along $\mathcal{F}$, $f(x) = \mathcal{O}\left(h(x)\right)$ along $\mathcal{F}$, then
$$(\lambda f)(x) = \lambda(x) f(x) = \mathcal{O}\left(g(x)h(x)\right).$$
2. If $\lambda(x) = \mathcal{O}\left(g(x)\right)$ along $\mathcal{F}$, $f(x) = o\left(h(x)\right)$ along $\mathcal{F}$, or if $\lambda(x) = o\left(g(x)\right)$ along $\mathcal{F}$, $f(x) = \mathcal{O}\left(h(x)\right)$ along $\mathcal{F}$, then
$$\lambda(x) f(x) = o\left(g(x)h(x)\right).$$

Proof
(1) We have $\varepsilon : X \longrightarrow \mathbb{R}_{\geq 0}$, $A \in \mathcal{F}$ such that $\lim_{\mathcal{F}} \varepsilon = 0$ and $\forall x \in A$, $\|f(x)\| \leq \varepsilon(x)g(x)$. Since $\lim_{\mathcal{F}} \varepsilon = 0$, there exists $B \in \mathscr{T}$ such that $\forall x \in B$, $|\varepsilon(x)| < 1$, hence $\forall x \in A \cap B$, $\|f(x)\| \geq (x)$.
(2)

1. $A_1, A_2 \in \mathcal{F}, C_1, C_2 > 0, \forall x \in A_1, \|f_1(x)\| \leq C_1 g(x), \forall x \in A_2, \|f_2(x)\| \leq C_2 g(x)$. So $f_1(x) + f_2(x) = \mathcal{O}(g(x))$
2. Let $\varepsilon_1 : X \longrightarrow \mathbb{R}_{\geq 0}$, $\varepsilon_2 : X \longrightarrow \mathbb{R}_{\geq 0}$, $A \in \mathcal{F}$, $\lim_{\mathcal{F}} \varepsilon_1 = \lim_{\mathcal{F}} \varepsilon_2 = 0$. $\forall x \in A_1, \|f_1(x)\| \leq \varepsilon_1(x) \cdot g(x), \forall x \in A_2, \|f_2(x)\| \leq \varepsilon_2(x) g(x)$. So $\lim_{\mathcal{F}} \varepsilon_1 + \varepsilon_2 = 0$.
$$\forall x \in A_1 \cap A_2,\ \|f_1(x) + f_2(x)\| \leq \|f_1(x)\| + \|f_2(x)\| \leq \left(\varepsilon_1(x) + \varepsilon_2(x)\right) g(x).$$

(3)

1. There exists $(C_1, C_2) \in \mathbb{R}^2_{>0}$ and $(A_1, A_2) \in \mathcal{F}^2$ such that

$$\forall x \in A_1,\ |\lambda(x)| \leq C_1 g(x),\ \forall x \in A_2,\ \|f(x)\| \leq C_2 h(x).$$

   Hence,

$$\forall x \in A_1 \cap A_2,\ \|(\lambda(x)f(x))\| \leq |\lambda(x)| \cdot \|f(x)\| \leq C_1 C_2 g(x) h(x).$$

2. We assume that

$$\lambda(x) = \mathcal{O}\left(g(x)\right) \text{ along } \mathcal{F},\ f(x) = o\left(h(x)\right) \text{ along } \mathcal{F}.$$

   There exists $(A_1, A_2) \in \mathcal{F} \times \mathcal{F}, C \in \mathbb{R}_{\geq 0}$ and a mapping $\varepsilon : X \longrightarrow \mathbb{R}_{\geq 0}$ such that

$$\forall x \in A_1,\ |\lambda(x)| \leq C \cdot g(x),\ \forall x \in A_2, \|f(x)\| \leq \varepsilon(x) h(x).$$

   Then one has

$$\lim_{\mathcal{F}} C\varepsilon(x) = 0$$

   and

$$\forall x \in A_1 \cap A_2,\ \|(\lambda(x)f(x))\| \leq |\lambda(x)| \cdot \|f(x)\| \leq C \cdot g(x) \cdot \varepsilon(x) h(x)$$

   As required.

□

Example 1.1.3
(1) Let $I \subseteq \mathbb{N}$ infinite. Let $(V, \|\cdot\|)$ be a normed vector space over complete valued field $(K, |\cdot|)$. Let $\mathcal{F}$ be the filter on I. Let $(x_n)_{n\in I} \in V^I, (b_n)_{n\in I} \in \mathbb{R}^I_{\geq 0}$. We denote by

$$x_n = \mathcal{O}\left(b_n\right),\ n \in I,\ n \to +\infty$$

the statement $x_n = \mathcal{O}\left(b_n\right)$ along $\mathcal{F}$. Namely,

$$\exists N \in \mathbb{N},\ \exists C > 0,\ \forall n \in I_{\geq N},\ \|x_n\| \leq C \cdot b_n.$$

$$x_n = o\left(b_n\right),\ n \in I,\ n \to +\infty$$

denotes the statement $x_n = o\left(b_n\right)$ along $\mathcal{F}$. Namely,

$$\exists (\varepsilon_n)_{n\in I} \text{ such that } \lim_{n\to+\infty} \varepsilon_n = 0,\ \exists N \in \mathbb{N},\ \forall n \in I_{\geq N},\ \|x_n\| \leq \varepsilon_n \cdot b_n.$$

(2) Let $(X, \mathscr{T})$ be a topological space, $Y \subseteq X$, $y_0 \in \overline{Y}$. Let $f : Y \longrightarrow V$ and $g : Y \longrightarrow \mathbb{R}_{\geq 0}$ be mappings.

$$\mathcal{F} = \mathcal{V}_{y_0}(\mathscr{T})|_Y := \{U \cap Y \mid U \text{ is a neighborhood of } y_0\}$$

$f(y)\mathcal{O}(g(y))$, $y \in Y$, $y \to y_0$ denotes $f(y) = \mathcal{O}(g(y))$ along $\mathcal{F}$. Namely,

$$\exists C > 0,\ \exists U \in \mathcal{V}_{y_0}(\mathscr{T}),\ \forall y \in U \cap Y,\ \|f(y)\| \leq C \cdot g(y).$$

$$f(y) = o(g(y)),\ y \in Y,\ y \to y_0$$

denotes $f(y) = o(g(y))$ along $\mathcal{F}$. Namely,

$$\exists \varepsilon : Y \longrightarrow \mathbb{R}_{\geq 0},\ \lim_{y \in Y, y \to y_0} \varepsilon(y) = 0,\ \exists U \in \mathcal{V}_{y_0}(\mathscr{T}),$$

$$\forall y \in U \cap Y,\ \|f(y)\| \leq \varepsilon(y) g(y).$$

(3) Let $\mathcal{F}$ be a filter on $\mathbb{R}$ generated by subsets of the form $[a, +\infty[$. Let $Y \subseteq \mathbb{R}$ not bounded from above. Let $f : Y \longrightarrow V$ and $g : Y \longrightarrow \mathbb{R}_{\geq 0}$ be mappings. Then

$$f(y) = \mathcal{O}(g(y)),\ y \in Y,\ y \to +\infty$$

denotes $f(y) = \mathcal{O}(g(y))$ along $\mathcal{F}|_Y$. Namely,

$$\exists C > 0,\ \exists a \in \mathbb{R},\ \forall y \in Y_{\geq a},\ \|f(y)\| \leq C \cdot g(y).$$

$$f(y) = o(g(y)),\ y \in Y,\ y \to +\infty$$

denotes $f(y) = o(g(y))$ along $\mathcal{F}|_Y$. Namely,

$$\exists \varepsilon : Y \longrightarrow \mathbb{R}_{\geq 0},\ \lim_{y \to +\infty} \varepsilon(y) = 0,\ \exists a \in \mathbb{R},\ \forall y \in Y_{\geq a},\ \|f(y)\| \leq \varepsilon(y) g(y).$$

1.2 Differentiability

We fix a complete valued field $(K, |\cdot|)$. We suppose that there exists $a \in K^{\times}$, such that $|a| < 1$. Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K.

$$\mathscr{L}(E, F) := \{\varphi \in \mathrm{Hom}_K(E, F) \mid \|\varphi\| < +\infty\}.$$

$(\mathscr{L}(E, F), \|\cdot\|)$ is a normed vector space over K.

Definition 1.2.1 Let $U \subseteq E$ be subset and $p \in U^\circ$. We say that a mapping $f : U \longrightarrow F$ is **differentiable** at p if there exists $\varphi \in \mathscr{L}(E, F)$ such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

If $U = U^\circ$ and f is differentiable at every point of U, we say that f is **differentiable** on U.

Proposition 1.2.2 Assume that $f : U \longrightarrow F$ is differentiable at $p \in U^\circ$. There exists a unique $\varphi \in \mathscr{L}(E, F)$ such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

Lemma 1.2.3 $\forall \eta \in \mathscr{L}(E, F), \forall r > 0.$

$$\|\eta\| = \sup_{x \in E, 0 < \|x\|_E \leq r} \frac{\|\eta(x)\|_F}{\|x\|_E} = \sup_{\substack{x \in E \\ 0 < \|x\|_E < r}} \frac{\|\eta(x)\|_F}{\|x\|_E}.$$

Proof (of Lemma) $\|\eta\| \geq \sup_{x \in E, 0 < \|x\|_E < r} \frac{\|\eta(x)\|_F}{\|x\|_E}$. $\forall y \in E \backslash \{0\}$, $\|a^N y\|_E = |a|^N \|y\|_E < r$.

$$\frac{\|\eta(a^N y)\|_F}{\|a^N y\|_E} = \frac{|a|^N \cdot \|\eta(y)\|_F}{|a|^N \cdot \|y\|_E} = \frac{\|\eta(y)\|_F}{\|y\|_E} \leq \sup_{\substack{x \in E \\ 0 < \|x\|_E < r}} \frac{\|\eta(x)\|_F}{\|x\|_E}.$$

□

Proof (of Proposition) Suppose $\varphi, \psi \in \mathscr{L}(E, F)$ are such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right),\ h \to 0_E,$$

$$f(p+h) - f(p) - \psi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

Then

$$\varphi(h) - \psi(h) = o\left(\|h\|_E\right),\ h \to 0_E.$$

$$\exists r > 0, \exists \varepsilon : \overline{B}(0_E, r) \longrightarrow \mathbb{R}_{\geq 0} \text{ such that } \lim_{h \to 0_E} \varepsilon(h) = 0.$$

$$\forall h \in \overline{B}(0_E, r),\ \|(\varphi - \psi)(h)\|_F = \varepsilon(h)\|h\|_E.$$

$$\|\varphi - \psi\| = \sup_{\substack{x \in E \\ 0 < \|h\|_E < r'}} \frac{\|\varphi(h) - \psi(h)\|_F}{\|h\|_E} \leq \sup_{0 < \|h\|_E < r'} \varepsilon(h).$$

Taking the limit when $r' \to 0$, by $\limsup_{h \to 0_E} \varepsilon(h) = 0$. We get $\|\varphi - \psi\| = 0$, hence $\varphi = \psi$. □

Definition 1.2.4 Let $U \subseteq E$ and $f : U \longrightarrow F$ be a mapping that is differentiable at $p \in U^\circ$. The unique $\varphi \in \mathscr{L}(E, F)$ such that

$$f(p+h) - f(p) - \varphi(h) = o\left(\|h\|_E\right), \ h \to 0_E$$

is called the **differential** of f at p and is denoted as

$$\mathrm{D}(f(p)).$$

Example 1.2.5
(1) $f : U \longrightarrow F$, $f(x) \equiv c$, $c \in F$.

$$f(x+h) - f(x) = 0_E = o\left(\|h\|_E\right).$$

So f is differentiable at every point of U and $\mathrm{D}(f(x)) = 0_F$.
(2) $\varphi \in \mathscr{L}(E, F)$.

$$\varphi(p+h) - \varphi(p) - \varphi(h) = 0_F = o\left(\|h\|_E\right).$$

So φ is differentiable at every point of E and $\mathrm{D}(\varphi(p)) = \varphi$.
(3) Let $(F_i, \|\cdot\|_i)$ be normed vector spaces over K, $i \in \{1, \ldots, n\}$, $n \in \mathbb{N}$. Suppose that $F = F_1 \oplus \cdots \oplus F_n$ and

$$\|\left(s_1, \ldots, s_n\right)\|_F = \max\{\|s_1\|_1, \ldots, \|s_n\|_n\}.$$

Let $U \subseteq E$ be an open subset, $f_i : U \longrightarrow F_i$ be a mapping.

$$f : U \longrightarrow F, \ f(x) = \left(f_1(x), \ldots, f_n(x)\right).$$

$$f(p+h) - f(p) = \left(f_1(p+h) - f_1(p), \ldots, f_n(p+h) - f_n(p)\right).$$

Suppose that each f_i is differentiable

$$\begin{aligned} & f(p+h) - f(p) - \left(\mathrm{D}f_1(p)(h), \ldots, \mathrm{D}f_n(p)(h)\right)|_F \\ = & \max_{i \in \{1, \ldots, n\}} \|f_i(p+h) - f_i(p) - \mathrm{D}f_i(p)(h)\|_{F_i} \\ = & o\left(\|h\|_E\right). \end{aligned}$$

So f is differentiable at p and

$$\mathrm{D}f(p)(h) = (\mathrm{D}f_1(p)(h), \dots, \mathrm{D}f_n(p)(h)) .$$

(4) Suppose that $E = K$. If $U \subseteq K$ is open and $f : U \longrightarrow F$ is differentiable at $p \in U$. We denote by $f'(p)$ the element $\mathrm{D}f(p)(1) \in F$.

$$f(p+h) - f(p) - \mathrm{D}f(p)(h) = o(\|h\|_E) .$$

So

$$f(p+h) - f(p) - hf'(p) = o(\|h\|_E) ,$$
$$\frac{f(p+h) - f(p)}{h} - f'(p) = o(1).$$

That is,

$$\lim_{h \to 0} \frac{f(p+h) - f(p)}{h} = f'(p).$$

Theorem 1.2.6 Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$, $(G, \|\cdot\|_G)$ be normed vector spaces over a complete valued field $(K, |\cdot|)$. Let $U \subseteq E$ and $V \subseteq F$ be open subsets, $f : U \longrightarrow F$ and $g : V \longrightarrow G$ be mappings such that $f(U) \subseteq V$. Let $p \in U$. If f is differentiable at p and g is differentiable at $f(p)$, then $g \circ f : U \longrightarrow G$ is differentiable at p and

$$\mathrm{D}(g \circ f)(p)(h) = \mathrm{D}g(f(p)) (\mathrm{D}f(p)(h)) .$$

Proof

$$f(p+h) - f(p) - \mathrm{D}f(p)(h) = o(\|h\|_E) ,$$

so,

$$f(p+h) - f(p) = \mathcal{O}(\|h\|_E) .$$

$$\begin{aligned} &g(f(p+h)) - g(f(p)) - \mathrm{D}g(f(p)) (f(p+h) - f(p)) \\ =&o(\|f(p+h) - f(p)\|_F) = o(\mathcal{O}\|h\|_E) = o(\|h\|_E) . \end{aligned}$$

$$\begin{aligned} &\mathrm{D}g(f(p)) (f(p+h) - f(p)) - \mathrm{D}g(f(p)) (\mathrm{D}f(p)(h)) \\ =&\mathrm{D}g(f(p)) (f(p+h) - f(p) - \mathrm{D}f(p)(h)) \\ =&\mathcal{O}(o(\|h\|_E)) = o(\|h\|_E) . \end{aligned}$$

So,

$$g(f(p+h)) - g(f(p)) - \mathrm{D}g(f(p)) (\mathrm{D}f(p)(h)) = o(\|h\|_E) .$$

□

Remark 1.2.7 If $(E, \|\cdot\|_E) = (K, |\cdot|)$,

$$(g \circ f)'(p) = \mathrm{D}g(f(p))(f'(p)).$$

If $E = F = K$, $\|\cdot\|_E = \|\cdot\|_F = |\cdot|$.

$$(g \circ f)'(p) = g'(f(p)) \cdot f'(p).$$

Remark 1.2.8 Let $U \subseteq E$ be open. $f : U \longrightarrow F_1 \times \cdots \times F_n$. If f is differentiable at $p \in U$, for any $i \in \{1, \ldots, n\}$, the mapping

$$f_i := \pi_i \circ f : U \longrightarrow F_i$$

is differentiable at p and

$$\mathrm{D}(f_i)(p)(h) = \mathrm{D}\pi_i(f(p))\left(\mathrm{D}f(p)(h)\right) = \pi_i\left(\mathrm{D}f(p)(h)\right).$$

1.3 Multilineal Mappings

Definition 1.3.1 Let K be a commutative unitary ring. Let $E_1, \ldots, E_n; F$ be K-modules. We say that

$$\varphi : E_1 \times \ldots \times E_n \longrightarrow F$$

is n-linear if for any $i \in \{1, \ldots, n\}$ and any $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in E_1 \times \ldots \times E_{i-1} \times E_{i+1} \times \ldots \times E_n$, the mapping

$$E_i \longrightarrow F,\ x_i \mapsto \varphi(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)$$

is a homomorphism of K-modules. (K-linear mapping)
If $n = 1$, 1-linear is also called linear.
If $n = 2$, 2-linear is also called bilinear.

Example 1.3.2
(1) $K \times K \longrightarrow K$ $(a, b) \longmapsto ab$ is bilinear.
(2) $K^n \times K^n \longrightarrow K$ $(x, y) \longmapsto x \cdot y = \sum_{i=1}^n x_i y_i$ is bilinear.
(3) $K \times \ldots \times K \longrightarrow K$ $(x_1, \ldots, x_n) \longmapsto x_1 \cdots x_n$ is n-linear.

Definition 1.3.3 We denote by $\mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F)$ the set of n-linear mappings from $E_1 \times \ldots \times E_n$ to F.

Definition 1.3.4 Let $(K, |\cdot|)$ be a complete valued field.
Let $(E_1, \|\cdot\|_{E_1}), \ldots, (E_n, \|\cdot\|_{E_n}), (F, \|\cdot\|_F)$ be normed vector spaces over K. For any $\varphi \in \mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F)$, we define

$$\|\varphi\| := \sup_{\substack{x_i \in E_i \setminus \{0\} \\ i \in \{1,\ldots,n\}}} \frac{\|\varphi(x_1, \ldots, x_n)\|_F}{\|x_1\|_{E_1} \cdots \|x_n\|_{E_n}}.$$

We denote by $\mathscr{L}(E_1 \times \cdots \times E_n, F)$ the set

$$\{\varphi \in \mathrm{Hom}_K^{(n)}(E_1 \times \cdots \times E_n, F) \mid \|\varphi\| < +\infty\}.$$

$\mathscr{L}^{(n)}(E_1 \times \ldots \times E_n, F)$ is a normed vector space of $\mathrm{Hom}_K^{(n)}(E_1 \times \ldots \times E_n, F)$, and the norm is $\|\cdot\|$.

Theorem 1.3.5 Let $(E_1, \|\cdot\|_{E_1}), \ldots, (E_n, \|\cdot\|_{E_n}), (F, \|\cdot\|_F)$ be normed vector spaces over K. Let $\varphi \in \mathscr{L}^{(n)}(E_1 \times \ldots \times E_n, F)$. For any $p = (p_1, \ldots, p_n) \in E_1 \times \ldots \times E_n$, φ is differentiable at p and

$$\mathrm{D}\varphi(p)(h_1, \ldots, h_n) = \sum_{i=1}^{n} \varphi(p_1, \ldots, p_{i-1}, h_i, p_{i+1}, \ldots, p_n).$$

Proof

$$\begin{aligned}\varphi(p+h) - \varphi(p) = \sum_{i=1}^{n} &\varphi(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, p_i + h_i, p_{i+1}, \ldots, p_n) \\ &-\varphi(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, p_i, p_{i+1}, \ldots, p_n)\end{aligned}$$

$$\begin{aligned}&\varphi(p_1 + h_1, \ldots, p_{i-1} + h_{i-1}, h_i, p_{i+1}, \ldots, p_n) \\ &-\varphi(p_1, \ldots, p_{i-1}, h_i, p_{i+1}, \ldots, p_n) \\ &= \sum_{j=1}^{i-1} \varphi(p_1 + h_1, \ldots, p_{j-1} + h_{j-1}, h_j, p_{j+1}, \ldots, h_i, \ldots, p_n).\end{aligned}$$

$$\begin{aligned}
&\|\varphi(p_1+h_1,\ldots,p_{j-1}+h_{j-1},h_j,p_{j+1},\ldots,h_i,\ldots,p_n)\|_F\\
\leq&\|\varphi\|\cdot\prod_{k=1}^{j-1}\|p_k+h_k\|_{E_k}\cdot\|h_j\|_{E_j}\cdot\prod_{k=j+1}^{i-1}\|p_k\|_{E_k}\cdot\|h_i\|_{E_i}\cdot\prod_{k=i+1}^{n}\|p_k\|_{E_k}\\
=&\mathcal{O}\left(\|h\|^2\right)=o(h),\ h\to 0.
\end{aligned}$$

□

Definition 1.3.6 Let K be a commutative unitary ring. $n\in\mathbb{N}_{\geq 1}$, E and F be K-modules. We say that

$$\varphi\in\mathrm{Hom}_K^{(n)}(E_1\times\ldots\times E_n,F)$$

is **symmetric** if

$$\forall\sigma\in\mathfrak{S}_{\{1,\ldots,n\}},\ \forall(x_1,\ldots,x_n)\in E^n,\ \varphi(x_{\sigma(1)},\ldots,x_{\sigma(n)})=\varphi(x_1,\ldots,x_n).$$

Let $P:E\longrightarrow F$ be a mapping. If there exists a symmetric $\varphi\in\mathrm{Hom}_K^{(n)}(E\times\ldots\times E,F)$ such that

$$\forall x\in E,\ P(x)=\varphi(x,\ldots,x),$$

we say that P is a **homogeneous polynomial mapping of degree** n.
If $F=K$, P is called a **homogeneous polynomial** on E. The symmetric polynomial mapping φ is called the **polarization** of P.

Proposition 1.3.7 Let $(K,|\cdot|)$ be a complete valued field that is non-trivial. Let $(E,\|\cdot\|_E)$ and $(F,\|\cdot\|_F)$ be normed vector spaces over K. Assume that $P:E\longrightarrow F$ is a homogeneous polynomial mapping of degree n. Which admits a bounded polarization φ. Then P is differentiable on E and,

$$\forall(x,h)\in E\times E,\ \mathrm{D}P(x)(h)=n\varphi(x,\ldots,x,h).$$

Proof Let

$$\begin{array}{rccc}\Delta: & E & \longrightarrow & E^n,\\ & x & \longmapsto & (x,\ldots,x).\end{array}$$

Then $P=\varphi\circ\Delta$. Since φ and Δ are differentiable, so it is P.

Moreover,

$$\begin{aligned}
\mathrm{D}P(x)(h) &= \mathrm{D}\varphi(\Delta(x))\,(\mathrm{D}\Delta(x)(h)) \\
&= \mathrm{D}\varphi(x,\ldots,x)(h,\ldots,h) \\
&= \sum_{i=1}^{n} \varphi(x,\ldots,x,h,x,\ldots,x) \\
&= n\varphi(x,\ldots,x,h).
\end{aligned}$$

□

Remark 1.3.8 Assume that $E = K$. Let $P : K \longrightarrow F$ be a homogeneous polynomial mapping of degree n of form $P(x) = x^n s$, where $s \in F$. Its polarization is of the form

$$\varphi(a_1,\ldots,a_n) = a_1 \cdots a_n s.$$

$$P'(x) = \mathrm{D}P(x)(1) = n\varphi(x,\ldots,x,1) = nx^{n-1}s.$$

Proposition 1.3.9 Let n be a positive integer $n \geq 2$. Let $(E_1,\|\cdot\|_1),\ldots,(E_n,\|\cdot\|_n),(F,\|\cdot\|_F)$ be normed vector spaces. For any $i \in \{1,\ldots,n_1\}$, the mapping

$$\mathscr{L}^{(n)}(E_1,\ldots,E_n;F) \xrightarrow{f} \mathscr{L}^{(i)}(E_1,\ldots,E_i,\mathscr{L}^{(n-i)}(E_{i+1},\ldots,E_n;F))$$

$$\varphi \longmapsto \left(\begin{smallmatrix} E_1\times\ldots e_n \longrightarrow \mathscr{L}^{(i)}(E_{i+1},\ldots,E_n;F) \\ (x_1,\ldots,x_i)\longmapsto \left(\begin{smallmatrix}(x_{i+1},\ldots,x_n)\longmapsto\varphi(x_1,\ldots,x_n) \\ E_{i+1}\times\ldots\times E_n\in F\end{smallmatrix}\right)\end{smallmatrix}\right)$$

is an isomorphism of vector spaces over K, and in the same time an isometry, ($\|f(\varphi)\| = \|\varphi\|$).

Remark 1.3.10
$\forall (x_1,\ldots,x_n) \in E_1 \times \ldots \times E_n$, $f(\varphi)(x_1,\ldots,x_i)(x_{i+1},\ldots,x_n) = \varphi(x_1,\ldots,x_n)$

Proof $\forall (x_1,\ldots,x_n) \in E_1 \times \ldots \times E_n$,

$$\begin{aligned}
\varphi(x_1,\ldots,x_n) :&E_{i+1} \times \ldots \times E_n \longrightarrow F \text{ is bounded} \\
&(x_{i+1},\ldots,x_n) \longmapsto \varphi(x_1,\ldots,x_n)
\end{aligned}$$

Since

$$\|\varphi(x_1,\ldots,x_n)\|_F \leq (\|\varphi\|\cdot\|x_1\|\ldots\|x_i\|)\,\|x_{i+1}\|\ldots\|x_n\|.$$

$$\begin{aligned}\|f(\varphi)\| &= \sup_{x_j \in E_j\setminus\{0\}, j=1,\dots,i} \frac{\|\varphi(x_1,\dots,x_i,\cdot)\|}{\|x_1\|_{E_1}\cdots\|x_n\|_{E_i}} \\ &= \sup_{x_j \in E_j\setminus\{0\}, j=1,\dots,i} \sup_{x_k \in E_k\setminus\{0\}, k=i+1,\dots,n} \frac{\|\varphi(x_1,\dots,x_n)\|_F}{\|x_1\|_{E_1}\cdots\|x_n\|_{E_n}} \\ &= \|\varphi\|.\end{aligned}$$

Hence f is injective. ($\ker(f) = \{0\}$)
For any $\psi \in \mathscr{L}^{(i)}(E_1,\dots,E_i,\mathscr{L}^{(n-i)}(E_{i+1},\dots,E_n))$,

$$\begin{aligned}\varphi:\ E_1\times\ldots\times E_n &\longrightarrow F \\ (x_1,\dots,x_n) &\longmapsto \psi(x_1,\dots,x_i)(x_{i+1},\dots,x_n)\end{aligned}$$

belongs to $\mathscr{L}^{(n)}(E_1,\dots,E_n;F)$ and $f(\varphi)=\psi$. So f is surjective. □

Corollary 1.3.11 If $E_1,\dots,E_n$ are all finite dimensional, then

$$\mathscr{L}^{(n)}(E_1,\dots,E_n;F) = \mathrm{Hom}_K^{(n)}(E_1\times\ldots\times E_n,F).$$

Proof If $n=1$, $\mathscr{L}(E_1,F)=\mathrm{Hom}_K(E_1,F)$.

$$\begin{aligned}\mathscr{L}^{(n)}(E_1,\dots,E_n;F) &\cong \mathscr{L}(E_1,\mathscr{L}^{(n-1)}(E_2,\dots,E_n;F)) \\ &= \mathrm{Hom}_K(E_1,\mathrm{Hom}_K^{(n-1)}(E_2\times\ldots\times E_n,F)) \\ &\cong \mathrm{Hom}_K^{(n)}(E_1\times\ldots\times E_n,F).\end{aligned}$$

□

Let $(K,|\cdot|)$ be a complete nontrivial valued field. Let $(E,\|\cdot\|_E)$ and $(F,\|\cdot\|_F)$ be normed vector spaces over K.

Definition 1.3.12 Let $U\subseteq E$ be an open subset of E, $f: U\longrightarrow F$ be a mapping.
If f is continuous on U, we say that f is **of class** $\mathcal{C}^0$ and we denote by

$$\mathrm{D}^0 f$$

the mapping $f: U\longrightarrow F$. Denote by

$$\mathcal{C}^0(U,F)$$

the set of mappings from U to F.

$$U \xrightarrow{(f,g)} K \times K \xrightarrow{\times} K$$

$$p \longmapsto (f(p), g(p)) \longmapsto f(p) \times g(p)$$

Let $p \in U$. If f is differentiable on an open neighborhood V of p such that $V \subseteq U$. Then

$$\begin{aligned} \mathrm{D}f : V &\longrightarrow \mathscr{L}(E, F) \\ x &\longmapsto \mathrm{D}f(x) \end{aligned}$$

is a mapping. If $\mathrm{D}f$ is $(n-1)$-times differentiable at p, we say that f is **of class** $\mathcal{C}^n$ at p. If f is of class $\mathcal{C}^n$ at every point of U, we say that f is **n-times differentiable** at p. We denote by

$$\mathrm{D}^n f(p) \in \mathscr{L}^{(n)}(E, \dots, E, F)$$

the n-linear mapping that sends $(h_1, \dots, h_n) \in E^n$ to

$$\mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \dots, h_{n-1})(h_n) \in F.$$

Remark 1.3.13

$$\mathrm{D}^n f(p)(h_1, \dots, h_n) = \mathrm{D}^i(\mathrm{D}^{n-i} f)(p)(h_1, \dots, h_i)(h_{i+1}, \dots, h_n).$$

1.4 Convexity

Definition 1.4.1 Let E be a vector space over a field K. $S \subseteq E$ be a non-empty subset.
We call affine combination of elements of S any element of E of the form

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n,$$

where $n \in \mathbb{N}_{\geq 1}$, $s_1, \ldots, s_n \in S$, $a_1, \ldots, a_n \in K$ such that

$$a_1 + a_2 + \cdots + a_n = 1.$$

We denote by $\mathrm{Aff}(S)$ the set of all affine combinations of elements of S. One has $S \subseteq \mathrm{Aff}(S)$. $\mathrm{Aff}(S)$ is called the affine hull of S.
If $S = \mathrm{Aff}(S)$, we say that S is an affine subspace of E.

Proposition 1.4.2
(1) If F is a vector subspace of E, $\forall p \in E$,

$$p + F = \{p + x \mid x \in F\}$$

is an affine subspace of E.
(2) If $A \subseteq E$ is an affine subspace of E. For any $p \in A$,

$$A - p := \{x - p \mid x \in A\}$$

is a vector subspace of E, which is not dependent on the choice of p. We call it the vector space **associated** with A.

Proof
(1) Let $(x_1, \ldots, x_n) \in F^n$, $(a_1, \ldots, a_n) \in K^n$, such that $\sum_{i=1}^n a_i = 1$. Then

$$\begin{aligned}\sum_{i=1}^n a_i(p + x_i) &= p \cdot \sum_{i=1}^n a_i + \sum_{i=1}^n a_i x_i \\ &= p + \sum_{i=1}^n a_i x_i \in p + F.\end{aligned}$$

(2) Let $(x_1, \ldots, x_n) \in A^n$, $(b_1, \ldots, b_n) \in K^n$.

$$\begin{aligned}\sum_{i=1}^{n} b_i(x_i - p) &= \sum_{i=1}^{n} b_i x_i - \left(\sum_{i=1}^{n} b_i\right) p \\ &= \left(\sum_{i=1}^{n} b_i x_i + \left(1 - \sum_{i=1}^{n} b_i\right) p\right) - p \\ &\in A - p.\end{aligned}$$

Let $q \in A$, $\forall x \in A$, $x - p = (x - q) + (q - p) \in A - q$. So $A - p \subseteq A - q$. By symmetry, $A - q \subseteq A - p$. Hence $A - p = A - q$. □

Example 1.4.3 Let A be an m by p matrix with coefficients in $\mathbb{R}$. Let $(b_1, \ldots, b_n) \in E^m$. Consider the linear equation

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

The solution set is

$$S \coloneqq \{(x_1, \ldots, x_p) \in E^p \mid A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}\}.$$

Claim: S is an affine subspace of E^p.

Proof Let $\underline{x}^{(1)}, \ldots, \underline{x}^{(n)}$ be elements of S, where $\underline{x}^{(i)} = (x_1^{(i)}, \ldots, x_p^{(i)})$. Let $(a_1, \ldots, a_n) \in \mathbb{R}^n$, $\underline{x} = a_1 x^{(1)} + \cdots + a_n x^{(n)}$.

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} = A \left(a_1 \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_p^{(1)} \end{pmatrix} + \cdots + a_n \begin{pmatrix} x_1^{(n)} \\ \vdots \\ x_p^{(n)} \end{pmatrix} \right).$$

$$a_1 A \begin{pmatrix} x_1^{(1)} \\ \vdots \\ x_p^{(1)} \end{pmatrix} + \ldots a_n A \begin{pmatrix} x_1^{(n)} \\ \vdots \\ x_p^{(n)} \end{pmatrix} = (a_1 + \cdots + a_n) \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

$$x_j = a_1 x_j^{(1)} + \cdots + a_n x_j^{(n)}.$$

□

Proposition 1.4.4 Let $S \subseteq E$. Then $\mathrm{Aff}(S)$ is the smallest affine subspace of E containing S.

Proof
Let $A \subseteq E$ be an affine subspace containing S. $\forall n \in \mathbb{N}_{\geq 1}$, $\forall (x_1, \ldots, x_n) \in S^n \subseteq A^n$, $(a_1, \ldots, a_n) \in \mathbb{R}$, $a_1 + \cdots + a_n = 1$, one has

$$\sum_{i=1}^{n} a_i x_i \in A.$$

So $\mathrm{Aff}(S) \subseteq A$.

To show that $\mathrm{Aff}(S)$ is an affine subspace containing S, it is sufficient to check that $\mathrm{Aff}(S)$ is an affine subspace.

If $S = \varnothing$, then $\mathrm{Aff}(S) = \varnothing$. It is an affine subspace.
Suppose that $S \neq \varnothing$, $p \in S$. We prove that $\mathrm{Aff}(S) - p$ is equal to $\mathrm{Span}_{\mathbb{R}}(S - p)$. Let $y = a_1 x_1 + \cdots + a_n x_n \in \mathrm{Aff}(S)$.

$$y - p = a_1(x_1 - p) + \cdots + a_n(x_n - p) \in \mathrm{Span}_{\mathbb{R}}(S - p).$$

Let $(x_1, \ldots, x_n) \in S^n$, $(b_1, \ldots, b_n) \in \mathbb{R}^n$.

$$\sum_{i=1}^{n} b_i(x_i - p) = \left(\sum_{i=1}^{n} b_i x_i + \left(1 - \sum_{i=1}^{n} b_i \right) p \right) - p \in \mathrm{Aff}(S) - p.$$

□

Definition 1.4.5 Let $S \subseteq E$. We call **convex combination** of elements of S any element of E of the form

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n,$$

where $n \in \mathbb{N}_{\geq 1}$, $s_1, \ldots, s_n \in S$, $a_1, \ldots, a_n \in \mathbb{R}_{\geq 0}$ such that

$$a_1 + a_2 + \cdots + a_n = 1.$$

We denote by $\mathrm{Conv}(S)$ the set of all convex combinations of elements of S. $\mathrm{Conv}(S)$ is called the **convex hull** of S. One has $S \subseteq \mathrm{Conv}(S) \subseteq \mathrm{Aff}(S)$.

Proposition 1.4.6 Let E be a vector space over $\mathbb{R}$ and $C \subseteq E$. Then C is convex if and only if

$$\forall (x,y) \in C^2,\ \forall \lambda \in [0,1],\ \lambda x + (1-\lambda)y \in C.$$

Proof It is sufficient to check "$\Leftarrow$". We prove by induction on n that

$$\forall n \in \mathbb{N}_{\geq 1},\ \forall (x_1,\ldots,x_n) \in C^n,\ \forall (a_1,\ldots,a_n) \in \mathbb{R}^n_{\geq 0},\ \sum_{i=1}^n a_i = 1,\ \sum_{i=1}^n a_i x_i \in C.$$

The case where $n = 1$ is trivial. The case where $n = 2$ comes from the hypothesis. Suppose $n \geq 3$ in assuming that the statement holds for any integer less than n. If $a_n = 1$, then $a_1 = \cdots = a_{n-1} = 0$, so $\sum_{i=1}^n a_i x_i = x_n \in C$. If $a_n < 1$, we have $a_1 + \cdots + a_{n-1} = 1 - a_n > 0$. By the induction hypothesis,

$$x := \sum_{i=0}^{n-1} \frac{a_i}{1-a_n} x_i \in C.$$

Taking $y = x_n$, $t = 1 - a_n$,

$$C \ni tx + (1-t)y = \sum_{i=1}^n a_i x_i.$$

$\square$

1.5 Mean Value Theorems

Theorem 1.5.1 (Mean Value Inequality) Let $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$. Let $(a,b) \in \mathbb{R}^2$ such that $a < b$. Let $f : [a,b] \longrightarrow F$ be a continuous mapping that is differentiable on $]a,b[$. Then

$$\|f(b) - f(a)\|_F \leq (b-a) \cdot \sup_{t \in]a,b[} \|f'(t)\|_F.$$

Proof We may suppose that $\sup_{t\in]a,b[} \|f'(t)\|_F < +\infty$. Take

$$M > \sup_{t \in]a,b[} \|f'(t)\|_F.$$

Let $m = \frac{a+b}{2}$. Let

$$J = \{x \in [m,b] \mid \forall t \in [m,x], \|f(t) - f(m)\|_F \leq M(t-m)\}.$$

It is an interval containing m. So it is of the form

$$[m, c[\text{ or } [m, c]$$

$$\forall t \in [m, c[, \ \|f(t) - f(m)\|_F \leq M(t - m).$$

Taking the limit $t < c$, $t \to c$, we get $c \in J$. So $J = [m, c]$. We then check $c = b$.

If $c \neq b$, then $c \in]a, b[$, so f is differentiable at c. That is

$$\|f(c + h) - f(c)\|_F = \|f'(c)h + o(\|h\|)\|_F \leq \|f'(c)\|_F h + o(\|h\|), \ h \to 0.$$

Since $M > \|f'(c)\|_F$, $\exists h_0 > 0$ such that

$$\forall h \in]0, h_0], \ \|f(c + h) - f(c)\|_F \leq Mh.$$

$$\begin{aligned}\|f(c + h) - f(m)\| &\leq \|f(c + h) - f(c)\| + \|f(c) - f(m)\| \\ &\leq Mh + M(c - m) = M(c + h - m).\end{aligned}$$

So $[m, c + h_0] \subseteq J$, contradiction. Thus $b = c$. $\|f(b) - f(m)\|_F \leq M(b - m)$.

By the same reason, $\|f(m) - f(a)\|_F \leq M(m - a)$. So

$$\|f(b) - f(a)\|_F \leq \|f(b) - f(m)\|_F + \|f(m) - f(a)\|_F \leq M(b - a).$$

Taking the limit when $M \to \sup\limits_{t \in]a,b[} \|f'(t)\|_F$, we get the announced result. □

Corollary 1.5.2 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$. $U \subseteq E$ be an open subset, and $(x, y) \in U^2$ such that

$$[x, y] = \{tx + (1 - t)y \mid t \in [0, 1]\} \subseteq U.$$

Let $f : U \longrightarrow F$ be a differentiable mapping. Then

$$\|f(x) - f(y)\|_F \leq \left(\sup_{z \in]x,y[} \|\mathrm{D}f(z)\| \right) \cdot \|x - y\|_E.$$

Proof Let

$$\begin{aligned} g : \ [0, 1] &\longrightarrow U \\ t &\longmapsto tx + (1 - t)y. \end{aligned}$$

$$g(0) = x, \ g(1) = y, \ g'(t) = x - y.$$

Then,

$$(f \circ g)'(t) = \mathrm{D}f(g(t))(x - y),$$
$$\mathrm{D}(f \circ g)(t)(1) = \mathrm{D}f(g(t))(\mathrm{D}g(t)(1)).$$

By the theorem,

$$\begin{aligned}
\|f(x) - f(y)\|_F &= \|f(g(1)) - f(g(0))\|_F \\
&\le \sup_{t \in]0,1[} \|\mathrm{D}f(g(t))(x - y)\|_F \\
&\le \sup_{t \in]0,1[} |\mathrm{D}f(g(t))| \cdot \|x - y\|_E \\
&= \sup_{z \in]x,y[} \|\mathrm{D}f(z)\| \cdot \|x - y\|_E .
\end{aligned}$$

$\square$

Definition 1.5.3 Let $(X, \mathscr{T})$ be a topological space, $p \in X$. Let U be a neighborhood of p and $f : U \longrightarrow \mathbb{R}$ be a mapping. If there exists a neighborhood V of p such that $p \in V \subseteq U$ and

$$\forall x \in V, \ f(p) \ge f(x),$$

we say that p is a **local maximum point** of f on U.

If p is a local maximum point or a local minimum point, we say that p is a **local extremum** of f on U.

If $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ are normed vector spaces. $U \subseteq E$ open, $f : U \to F$ is differentiable. If $p \in U$ is such that

$$\mathrm{D}f(p) = 0 \in \mathscr{L}(E, F),$$

we say that p is a **critical point** of f.

Theorem 1.5.4 Let $(E, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$. $U \subseteq E$ be an open subset, $f : U \longrightarrow \mathbb{R}$ be a differentiable mapping. If $p \in U$ is a local extremum point of f, then it is a critical point $(\mathrm{D}f(p) = 0)$.

Proof There exists $r > 0$ such that $p + B(0, r) \subseteq U$ and

$$(h \in B(0, r)) \longmapsto f(p + h) - f(p) \in \mathbb{R}$$

does not change the sign.

$\forall h \in B(0,r),\ \forall \in [0,1],$

$$(f(p+th)-f(p))\,(f(p-th)-f(p)) \geq 0.$$

Taking the limit when $t \to 0$, $-\mathrm{D}f(p)(h)^2 \geq 0$. So $\mathrm{D}f(p)(h) = 0$. □

Theorem 1.5.5 (Rolle) Let $(a,b) \in \mathbb{R}^2$, $a < b$. Let $f : [a,b] \longrightarrow \mathbb{R}$ be a continuous mapping that is differentiable on $]a,b[$. If $f(a) = f(b)$, then

$$\exists t \in]a,b[,\ f'(t) = 0.$$

Proof If there exists t which is in $]a,b[$ and is an extremum point of f, then $f'(t) = 0$. Since $[a,b]$ is compact and f is continuous, so f attains its maximum and minimum.

If the extremum points of f are in $\{a,b\}$. Since $f(a) = f(b)$, f is compact, so $f'(t) = 0$ on $]a,b[$. □

Theorem 1.5.6 (Gronwall inequality) Let $(F, \|\cdot\|)$ be a normed vector space over $\mathbb{R}$, $(a,b) \in \mathbb{R}^2$, $a < b$. Let $f : [a,b] \longrightarrow F$ and $g : [a,b] \longrightarrow \mathbb{R}$ be differentiable mappings on $]a,b[$. If $\forall t \in]a,b[$, $\|f'(t)\| \leq g'(t)$, then

$$\|f(b)-f(a)\|_F \leq g(b)-g(a).$$

Proof Let $m \in]a,b[$. Let $\varepsilon > 0$,

$$J := \{t \in [m,b] \mid \forall s \in [m,t],\ \|f(s)-f(m)\|_F \leq g(s)-g(m)+\varepsilon(s-m)\}.$$

Since f and g are continuous, J is a closed interval of the form $[m,c]$.
If $c < b$,

$$\begin{aligned} f(c+h) &= f(c)+hf'(c)+o(h), \\ g(c+h) &= g(c)+hg'(c)+o(h), \quad h > 0, h \to 0. \end{aligned}$$

$\exists \delta > 0$, such that $[c,c+\delta] \subseteq [c,b]$ and $\forall h \in [0,\delta]$,

$$\|f(c+h)-f(c)\| \leq h\|f'(c)\| + \frac{\varepsilon}{2}h.$$

$$g(c+h)-g(c) \geq hg'(c) - \frac{\varepsilon}{2}h.$$

So,

$$\|f(c+h)-f(c)\| \leq g(c+h)-g(c)+\varepsilon h.$$

By the triangle inequality,

$$\|f(c+h) - f(m)\| \leq g(c+h) - g(m) + \varepsilon(c+h-m).$$

So $J \supseteq [m, c+\delta]$, contradiction.
Therefore $c = b$.

$$\|f(b) - f(m)\| \leq g(b) - g(m) + \varepsilon(b-m).$$

A similar argument shows that

$$\|f(m) - f(a)\| \leq g(m) - g(a) + \varepsilon(m-a).$$

Hence,

$$\|f(b) - f(a)\| \leq g(b) - g(a) + \varepsilon(b-a).$$

$$\|f(c+h) - f(c) + hf'(c)\| \leq \varphi(h)h, \ \lim_{h\to 0} \varphi(h) = 0.$$

$$\exists \delta > 0, \ \forall h > 0, \ 0 \leq h < \delta \Rightarrow |\varphi(h)| \leq \frac{\varepsilon}{2}.$$

□

Theorem 1.5.7 (Mean value theorem of Lagrange) Let $(a, b) \in \mathbb{R}^2$, $a < b$. Let $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous mapping that is differentiable on $]a, b[$. Then

$$\exists \xi \in\ ]a, b[, \ f(b) - f(a) = f'(\xi)(b-a).$$

Proof Let $g : [a, b] \longrightarrow \mathbb{R}$.

$$g(t) := f(b) - f(t) + C(b-t), \text{ where } C = -\frac{f(b) - f(a)}{b-a}.$$

Then $g(a) = g(b) = 0$, $g'(t) = -f'(t) - C$.

$$\exists \xi \in\ ]a, b[, \ g'(\xi) = 0, \ f'(\xi) = -C = \frac{f(b) - f(a)}{b-a}.$$

□

Theorem 1.5.8 (Darboux) Let I be an open interval in $\mathbb{R}$ and $f : I \longrightarrow \mathbb{R}$ be a differentiable mapping. Then $f'(I)$ is an interval.

Proof Let a, b be two elements in I such that $a < b$. Let

$$\begin{aligned} g: \ [a,b] &\longrightarrow \mathbb{R} \\ t &\longmapsto \begin{cases} \frac{f(t)-f(a)}{t-a}, \ t \neq a \\ f'(a), \ t = a \end{cases} \end{aligned}$$

g is continuous, and $g([a,b])$ is an interval. By the mean value theorem of Lagrange, $g([a,b]) \subseteq f'(I)$.
Let

$$\begin{aligned} h: \ [a,b] &\longrightarrow \mathbb{R} \\ t &\longmapsto \begin{cases} \frac{f(t)-f(b)}{t-b}, \ t \neq b \\ f'(b), \ t = b \end{cases} \end{aligned}$$

$h([a,b])$ is an interval contained in $f'(I)$.
$h([a,b]) \cup g([a,b])$ is an interval since

$$\frac{f(b)-f(a)}{b-a} \in h([a,b]) \cap g([a,b]),$$

$$\{f'(a), f'(b)\} \subseteq h([a,b]) \cup g([a,b]).$$

So the interval liking $f'(a)$, $f'(b)$ is contained in $f'(I)$. Hence, $f'(I)$ is an interval.
□

1.6 Higher Differential

We fix a complete non-trivially valued field $(K, |\cdot|)$. Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over K.

Definition 1.6.1 Let $U \subseteq E$ be an open subset, $f : U \longrightarrow F$ be a mapping, $p \in U$.
(1) If f is continuous at p, we say that f is 0-time differentiable at p, and we let

$$\mathrm{D}^0 f(p) := f(p).$$

(2) If f is differentiable at p, we say that f is 1-time differentiable at p, and we let

$$\mathrm{D}^1 f(p) := \mathrm{D}f(p).$$

(3) Let $n \geq 2$. If exists open neighborhood V of p such that $V \subseteq U$ and f is differentiable on V and $\mathrm{D}f$ is $n-1$-time differentiable on V, we say that f is

n-time differentiable at p, and we let

$$\mathrm{D}^n f(p) \in \mathscr{L}(E, \ldots, E, F)$$

be the multilineal mapping sending $(h_1, \ldots, h_n) \in E^n$ to

$$\mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \ldots, h_{n-1})(h_n).$$

If $E = K$, $\mathrm{D}^n f(p)(1, \ldots, 1)$ is denoted as $f^{(n)}(p) \in F$. $f^{(0)}(p)$ is often denoted as $f(p)$.

Remark 1.6.2 $\forall i \in \{1, \ldots, n\}$,

$$\mathrm{D}^n f(p)(h_1, \ldots, h_n) = \mathrm{D}^i(\mathrm{D}^{n-i} f)(p)(h_1, \ldots, h_i)(h_{i+1}, \ldots, h_n).$$

If $E = K$,

$$f^{(n)}(p)(h_1, \ldots, h_n) = \mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \ldots, h_{n-1})(h_n).$$

Definition 1.6.3 Let X be a set, we denote by $\mathfrak{S}_X$ the element of all bijection from X to X. $(\mathfrak{S}_X, \circ)$ forms a group. The identity mapping Id_X is the neutral element of $(\mathfrak{S}_X, \circ)$. $(\mathfrak{S}_X, \circ)$ is called the symmetric group of X. The elements of $(\mathfrak{S}_X, \circ)$ are called permutations of X.

Let $n \in \mathbb{N}_{\geq 2}$, $x_1, \ldots, x_n$ be distinct elements of X. We denote by $(x_1\, x_2 \cdots x_n)$ the element of $\mathfrak{S}_X$ that sends x_i to x_{i+1}, $(i \in \{1, \ldots, n-1\})$, x_n to x_1, $y \in X \backslash \{x_1, \ldots, x_n\}$ to y itself. This element is called an n-cycle. A 2-cycle is also called a transposition.

Remark 1.6.4 $\mathfrak{S}_X$ acts on X.

$$\begin{aligned} \mathfrak{S}_X \times X &\longrightarrow X \\ (\sigma, x) &\longmapsto \sigma(x). \end{aligned}$$

If $\sigma \in \mathfrak{S}_X$, $x \in X$, we denote by $\mathrm{orb}_\sigma(x)$ the set $\{\sigma^n(x) \mid n \in \mathbb{Z}\}$.

$$\langle \sigma \rangle := \{\sigma^n \mid n \in \mathbb{Z}\} \subseteq \mathfrak{S}_X$$

is a group. $\mathrm{orb}_\sigma(x)$ is the orbit of x under the action of $\langle \sigma \rangle$.

Proposition 1.6.5 If $\mathrm{orb}_\sigma(x)$ is finite of d elements, then $\sigma^d(x) = x$, and $\mathrm{orb}_\sigma(x) = \{x, \sigma(x), \ldots, \sigma^{d-1}(x)\}$. Moreover, the restriction of σ to $\mathrm{orb}_\sigma(x)$ identifies to the restriction of the cycle $(x, \sigma(x), \ldots, \sigma^{d-1}(x))$.

Proof Since $\mathrm{orb}_\sigma(x)$ is finite,

$$\{(n,m) \in \mathbb{Z}^2 \mid n < m,\ \sigma^n(x) = \sigma^m(x)\}$$

Let

$$l = \min\{m - n \mid (n,m) \in \mathbb{Z}^2\ n < m,\ \sigma^n(x) = \sigma^m(x)\}.$$

Then $x, \sigma(x), \ldots, \sigma^{l-1}(x)$ are distinct, and $\sigma^l(x) = x$. $\forall n \in \mathbb{Z}$, then n can be written as $n = lp + r$, where $p \in \mathbb{Z}$, $r \in \{0, \ldots, l-1\}$.

$$\sigma^n(x) = \sigma^r(\sigma^{lp}(x)) = \sigma^r((\sigma^l \circ \ldots \sigma^l)(x)) = \sigma^r(x).$$

So, $\mathrm{orb}_\sigma(x) = \{x, \sigma(x), \ldots, \sigma^{l-1}(x)\}$, ($l = d$). □

Remark 1.6.6 If X is finite, then X can be written as a distinct union of orbits (under the action of $\langle\sigma\rangle$). Let $d_i = \#(\mathrm{orb}_\sigma(x_i))$, $i = 1, \ldots, n$, then

$$\sigma|_{\mathrm{orb}_\sigma(x^{(i)})} = \left(x^{(i)}, \sigma(x^{(i)}), \ldots, \sigma^{d_i-1}(x^{(i)})\right)|_{\mathrm{orb}_\sigma(x^{(i)})}.$$

So $\sigma = \tau_1 \circ \cdots \circ \tau_n$, where $\tau_i := (x^{(i)}, \sigma(x^{(i)}), \ldots, \sigma^{d_i-1}(x^{(i)}))$.

Corollary 1.6.7 Suppose that X is finite. Any $\sigma \in \mathfrak{S}_X$ can be written as a composition of transpositions.

Proof

$$(x_1\ \ldots\ x_n) = (x_1\ x_2) \circ (x_2 \ldots\ x_n),$$

So,

$$(x_1\ \ldots\ x_n) = (x_1\ x_2) \circ (x_2\ x_3) \circ \cdots \circ (x_{n-1}\ x_n).$$

□

Definition 1.6.8 Denote by $\mathfrak{S}_n$ the symmetric group $\mathfrak{S}_{\{1,\ldots,n\}}$. A composition of the form $(i\ i+1)$, $i \in \{1, \ldots, n-1\}$ is called an adjacent transposition.

Corollary 1.6.9 Any $\sigma \in \mathfrak{S}_n$ can be written as a composition of adjacent transpositions.

Proof Let $(j,k) \in \{1,\dots,n\}^2$, $j<k$,

$$(j-1\ j) \circ (j\ k) \circ (j-1\ j) = (j-1\ k).$$

$$(j\ k) = (j\ j+1) \circ (j+1\ j+2) \circ \dots \circ (k-1\ k) \circ \dots (j\ j+1).$$

$\square$

Theorem 1.6.10 (Schwarz) Let $U \subseteq E$ be an open subset, $f : U \longrightarrow F$ be a mapping. $n \in \mathbb{N}_{\geq 1}$, $p \in U$. Assume that f is n-times differentiable at p. Then $\forall \sigma \in \mathfrak{S}_n$, $\forall (h_1, \dots h_n) \in E^n$,

$$\mathrm{D}^n f(p)(h_1, \dots, h_n) = \mathrm{D}^n f(p)(h_{\sigma(1)}, \dots, h_{\sigma(n)}).$$

Proof (By induction) The case where $n=1$ is trivial. Case $n=2$: Exists V open, $p \in V \subseteq U$. f is differentiable on V and $\mathrm{D}f$ is differentiable at p.

$$\mathrm{D}f(p+h)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(h,\cdot) = o\left(\|h\|_E\right).$$

Let $\varepsilon > 0$, $\exists \delta > 0$, $\forall h \in E$, $\|h\|_E \leq 2\delta \Rightarrow p+h \in V$ and

$$\|\mathrm{D}f(p+h)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(h,\cdot)\| \leq \varepsilon \|h\|_E.$$

Let $h \in E$ such that $\|h\|_E \leq \delta$. Define $g_h : B(0,\delta) \longrightarrow F$ as

$$g_h(k) = f(p+h+k) - f(p+h) - f(p+k) + f(p) - \mathrm{D}^2 f(p)(h,k).$$

Then,

$$\begin{aligned}
\mathrm{D}g_h(k)(\cdot) =& \mathrm{D}f(p+h+k)(\cdot) - \mathrm{D}f(p+k)(\cdot) - \mathrm{D}^2 f(p)(h,\cdot)\\
=& \mathrm{D}f(p+h+k)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(h+k,\cdot)\\
& -(\mathrm{D}f(p+k)(\cdot) - \mathrm{D}f(p)(\cdot) - \mathrm{D}^2 f(p)(k,\cdot))
\end{aligned}$$

$$\|\mathrm{D}g_h(k)(\cdot)\| \leq \varepsilon \|h+k\|_E + \varepsilon \|k\|_E \leq 3\varepsilon \max\{\|k\|_E, \|h\|_E\}.$$

$g_h(0) = 0$. Therefore, $\|g_h(k)\| \leq 3\varepsilon \max \|k\|_E, \|h\|_E{}^2$ (mean value inequality).

$$\|g_h(k) - g_h(0)\| \leq \left(\sup_{t \in]0,1[} \|\mathrm{D}g_h(tk)\|\right) \cdot \|k\|.$$

Therefore,

$$f(p+h+k)-f(p+h)-f(p+k)+f(p)-\mathrm{D}^2f(p)(h,k) = o\left(\max\{\|k\|_E, \|h\|_E\}^2\right).$$

By symmetry,

$$f(p+h+k)-f(p+h)-f(p+k)+f(p)-\mathrm{D}^2f(p)(k,h) = o\left(\max\{\|k\|_E, \|h\|_E\}^2\right).$$

$$\mathrm{D}^2f(p)(h,k) - \mathrm{D}^2f(p)(k,h) = o\left(\max\{\|h\|_E, \|k\|_h\}^2\right).$$

$$\mathrm{D}^2f(p)(th,tk) - \mathrm{D}^2f(p)(tk,th) = o(|t|^2),\ t \to 0.$$

$$\mathrm{D}^2f(p)(h,k) - \mathrm{D}^2f(p)(k,h) = o(1),\ t \to 0.$$

Suppose $n \geq 3$.

$$\mathrm{D}^n f(p)(h_1, \ldots, h_n) = \mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \ldots, h_{n-1})(h_n).$$

If $\sigma = (j\ j+1)$, $j \leq 2$,

$$\mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_1, \ldots, h_{n-1})(h_n) = \mathrm{D}^{n-1}(\mathrm{D}f)(p)(h_{\sigma(1)}, \ldots, h_{\sigma(n-1)})(h_n)$$

by the induction hypothesis, if $\sigma = (n-1\ n)$,

$$\mathrm{D}^n f(p)(h_1, \ldots, h_n) = \mathrm{D}^2\left((\mathrm{D}^{n-2}f)(h_1, \ldots, h_{n-2})(h_{n-1}, h_n)\right)$$

$$\begin{aligned}\mathrm{D}^n f(p)(h_{\sigma(1)}, \ldots, h_{\sigma(n)}) &= \mathrm{D}^n f(p)(h_1, \ldots, h_{n-1}) \\ &= \mathrm{D}^2\left((\mathrm{D}^{n-2}f)(h_1, \ldots, h_{n-2})(h_n, h_{n-1})\right) \\ &= \mathrm{D}^n f(p)(h_1, \ldots, h_n).\end{aligned}$$

$\square$

1.7 Taylor's Formula

Theorem 1.7.1 (Toylor-Young) Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{R}$, $U \subseteq E$ open, $n \in \mathbb{N}$, $f : U \longrightarrow F$ be a mapping, $p \in U$. Suppose that f is n-times differentiable at p. Then

$$f(x) = f(p) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(x-p, \dots, x-p) + o(\|x-p\|^n), \; x \to p.$$

Proof (By induction on n)
$n = 0$, $f(x) = f(p) + o(1)$ follows by continuity of f; $n = 1$ follows by the differentiability of f.

From $n-1$ to n. Let $g : U \longrightarrow F$

$$g(x) = f(x) - f(p) - \sum_{k=1}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(x-p, \dots, x-p).$$

g is differentiable on an open neighborhood of p,

$$\mathrm{D}g(x)(h) = \mathrm{D}f(x)(h) - \sum_{k=1}^{n} \frac{1}{k!} k \mathrm{D}^k f(p)(x-p, \dots, x-p, h)$$

$$\mathrm{D}g(x) = \mathrm{D}f(x) - \sum_{l=0}^{n-1} \frac{1}{l!} \mathrm{D}^l (\mathrm{D}f)(x-p, \dots, x-p) \overset{\text{hyp.}}{=\!=} o(\|x-p\|^{n-1}), x \to p.$$

So $g(x) = o(\|x-p\|^n)$.

$$\forall \varepsilon > 0, \; \exists \delta > 0, \forall x \in B(p, \delta), \; \|\mathrm{D}g(x)\| \leq \varepsilon \|x-p\|^{n-1}.$$

$g(p) = 0$, so

$$\|g(x) - g(p)\| \leq \varepsilon \|x-p\|^{n-1} \cdot \|x-p\| = \varepsilon \|x-p\|^n.$$

□

Theorem 1.7.2 (Taylor-Lagrange) Let $(a, b) \in \mathbb{R}^2$, $a < b$. $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous mapping. Suppose that f is $(n+1)$-times differentiable on $]a, b[$ and $\forall k \in \{3, \dots, n\}$, $f^{(k)} :]a, b[\longrightarrow \mathbb{R}$ tends to a continuous mapping $[a, b] \longrightarrow \mathbb{R}$.

Then

$$\exists \xi \in]a,b[, f(b) - \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) = \frac{f^{(n+1)}(\xi)(b-a)^{n+1}}{(n+1)!}.$$

Proof Let $g : [a,b] \longrightarrow \mathbb{R}$.

$$g(t) := \sum_{k=0}^{n} \frac{(b-t)^k}{k!} f^{(k)}(t) - C\frac{(b-t)^{n+1}}{(n+1)!}.$$

Then $g(b) = f(b)$, $g(a) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) - C\frac{(b-a)^{n+1}}{(n+1)!}$.

$$\begin{aligned} g'(t) &= \sum_{k=0}^{n} \frac{(b-t)^k}{k!} f^{(k+1)}(t) - \sum_{k=1}^{n} \frac{(b-t)^{k-1}}{(k-1)!} f^{(k)}(t) + C\frac{(b-t)^n}{n!} \\ &= \frac{(b-t)^n}{n!} f^{(n+1)}(t) + C\frac{(b-t)^n}{n!}. \end{aligned}$$

Take C such that $g(a) = g(b)$. Then by Rolle's theorem, $\exists \xi \in]a,b[, g'(\xi) = 0$, $C = -f^{(n+1)}(\xi)$. Then,

$$g(a) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{f^{(n+1)}(\xi)}{(n+1)!} + \frac{f^{(n+1)}(\xi)}{(n+1)!}(b-a)^{n+1} = f(b) = g(b).$$

□

Theorem 1.7.3 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be two normed vector spaces over $\mathbb{R}$, $U \subseteq E$ be an open subset, and $f : U \longrightarrow F$ be a mapping that is $(n+1)$-times differentiable, where $n \in \mathbb{N}$. Let $p \in U$, $h \in E$ such that $\forall t \in [0,1], p + th \in U$. Let

$$M = \sup_{t \in [0,1]} \|\mathrm{D}^{n+1} f(p+th)\|.$$

Then,

$$\|f(p+h) - \sum_{k=0}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(h, \ldots, h)\|_F \leq \frac{M}{(n+1)!} \|h\|_E^{n+1}.$$

Proof We define $\phi : [0,1] \longrightarrow F$

$$\phi(t) = f(p+th) + \sum_{k=1}^{n} \frac{(1-t)^k}{k!} \mathrm{D}^k f(p+th)(h,\ldots,h).$$

$$\phi(0) = \sum_{k=0}^{n} \frac{1}{k!} \mathrm{D}^k f(p)(h,\ldots,h),\ \phi(1) = f(p+h).$$

$$\begin{aligned}\phi'(t) =& \mathrm{D} f(p+h)(h) + \sum_{k=1}^{n} \frac{(1-t)^k}{k!} \mathrm{D}^{k+1} f(p+th)(h,\ldots,h) \\ &- \sum_{k=1}^{n} \frac{(1-t)^{k-1}}{(k-1)!} \mathrm{D}^k f(p+th)(h,\ldots,h) \\ =& \sum_{k=0}^{n} \frac{(1-t)^k}{k!} \mathrm{D}^{k+1} f(p+th)(h,\ldots,h) \\ &- \sum_{l=0}^{n-1} \frac{(1-t)^l}{(l)!} \mathrm{D}^{l+1} f(p+th)(h,\ldots,h) \\ =& \frac{(1-t)^n}{n!} \mathrm{D}^{n+1} f(p+th)(h,\ldots,h).\end{aligned}$$

So,

$$\|\phi'(t)\| \leq M \|h\|_E^{n+1} \frac{(1-t)^n}{n!},\ t \in [0,1].$$

By Gronwall's inequality,

$$\|\phi(1) - \phi(0)\|_F \leq M \cdot \|h\|^{n+1} \frac{1}{(n+1)!}.$$

□

1.8 Banach Space

Proposition 1.8.1 Let (X,d) be a metric space and $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. If

$$\sum_{n\in NN} \mathrm{d}(x_n, x_{n+1}) < +\infty,$$

then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.

Proof Let $N \in \mathbb{N}$. If $(n, m) \in \mathbb{N}^2$, $n > m$, by the triangle inequality,

$$\mathrm{d}(x_m, x_n) \leq \sum_{k=m}^{n-1} \mathrm{d}(x_k, x_{k+1}) \leq \sum_{k \geq N} \mathrm{d}(x_k, x_{k+1}).$$

So,

$$0 \leq \sup_{(n,m) \in \mathbb{N}_{\geq N}} \mathrm{d}(x_n, x_m) \leq \sum_{k \geq N} \mathrm{d}(x_k, x_{k+1}).$$

Taking the limit when $N \to +\infty$, we get

$$\lim_{(n,m) \in \mathbb{N}^2_{\geq N}} \mathrm{d}(x_n, x_m) = 0.$$

Let $(a_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. If $\left(\sum_{k=0}^{n} a_k \right)_{n \in \mathbb{N}}$ converges to some l in $\mathbb{R}$. Then, $l - \sum_{k=0}^{N-1} a_k$ converges to 0. If $a_k \leq 0$ for any $k \in \mathbb{N}$, $l - \sum_{k=0}^{N-1} a_k = \sum_{k=N}^{+\infty} a_k$.

$$l - \sum_{k=0}^{N-1} a_k = \lim_{n \to +\infty} \left(\sum_{k=0}^{n} a_k - \sum_{k=0}^{N-1} a_k \right) = \lim_{n \to +\infty} \sum_{k=N}^{n} a_k.$$

□

Definition 1.8.2 Let $(K, |\cdot|)$ be a complete valued field and $(E, \|\cdot\|)$ be a normed vector space over K. If E equipped with the metric

$$\begin{array}{ccc} E \times E & \longrightarrow & \mathbb{R}_{\geq 0} \\ (x, y) & \longmapsto & \|x - y\|_E. \end{array}$$

is complete, we say that $(E, \|\cdot\|)$ is a **Banach space**.
Let $(E, \|\cdot\|)$ be a Banach space. If $(x_n)_{n \in \mathbb{N}}$ is a sequence in E such that $\sum_{n \in \mathbb{N}} \|x_n\| < +\infty$, we say that $\sum_{n \in \mathbb{N}} x_n$ **converges absolutely**.

Remark 1.8.3 Suppose that $\sum_{n \in \mathbb{N}} x_n$ converges absolutely. Then $\left(\sum_{k=0}^{n} x_k \right)_{n \in \mathbb{N}}$ is a Cauchy sequence, since

$$\|x_n\| = \| \sum_{k=0}^{n} x_k - \sum_{k=0}^{n} x_{k-1} \|.$$

So, $\sum_{n \in \mathbb{N}} x_n$ converges.

Theorem 1.8.4 (Root test of Cauchy) Let $(E, \| \cdot \|)$ be a Banach space and $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. Let

$$r = \limsup_{n\to\infty} \|x_n\|^{\frac{1}{n}} \in [0, +\infty]$$

If $r < 1$, then $\sum_{n\in\mathbb{N}} x_n$ converges absolutely.
If $r > 1$, then $\sum_{n\in\mathbb{N}} x_n$ diverges.

Lemma 1.8.5 If a series $\sum_{n\in\mathbb{N}} x_n$ converges, then $\lim_{n\to+\infty} \|x_n\| = 0$.

Proof (of lemma)

$$\|x_n\| = \| \sum_{k=0}^{n} x_k - \sum_{k=0}^{n-1} x_k \|.$$

Since $\sum_k^n x_k$ converges to some $l \in E$.

$$\lim_{n\to+\infty} \|x_n\| = \lim_{n\to+\infty} \| \sum_{k=0}^{n} x_k - \sum_{k=0}^{n-1} x_k \| = \|l - l\| = 0.$$

□

Proof (of theorem) If $r > 1$, $\exists \beta > 1$ such that $r > \beta$. Since $r = \limsup_{n\to+\infty} \|x_n\|^{\frac{1}{n}}$, $\exists I \subseteq \mathbb{N}$ infinite such that $\lim_{n\in I, n\to+\infty} \|x_n\|^{\frac{1}{n}} = r$ (Bolzano-Weierstrass).

$$\exists N \in \mathbb{N}, \forall n \in I_{\geq N}, \|x_n\|^{\frac{1}{n}} \geq \beta.$$

So, $\|x_n\| \geq \beta^n \geq 1$. So $\sum_{n\in\mathbb{N}} x_n$ diverges.
If $r < 1$, $\exists \alpha \in]0, 1[, r < \alpha$. Since $r = \limsup_{n\to+\infty} \|x_n\|^{\frac{1}{n}}$,

$$\exists N \in \mathbb{N}, \ \forall n \geq N, \ \|x_n\|^{\frac{1}{n}} \leq \alpha, \|x_n\| \leq \alpha^n.$$

So,

$$\sum_{n\geq N} \|x_n\| \leq \sum n \geq N\alpha^n = \frac{\alpha^N}{1-\alpha} < +\infty.$$

Therefore, $\sum_{n\in\mathbb{N}} x_n$ converges absolutely. □

Theorem 1.8.6 (Ratio test of D'Alembert) LEt $(E, \| \cdot \|)$ be a Banach space and $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$.

(1) If

$$\limsup_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} < 1,$$

then $\sum_{n\in\mathbb{N}} x_n$ converges absolutely.
(2) If

$$\limsup_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} > 1,$$

then $\sum_{n\in\mathbb{N}} x_n$ diverges.

Proof
(1) Let $0 < \alpha < 1$ such that

$$\limsup_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} < \alpha.$$

$$\exists N \in \mathbb{N}, \forall x \in \mathbb{N}_N, \|x_{n+1}\|\alpha\|x_n\| \leq \alpha^{n+1-N}\|x_N\|.$$

Thus,

$$\sum_{n\geq\mathbb{N}} \|x_n\| \leq \sum_{n\geq\mathbb{N}} \|x_N\|\alpha^{n-N} = \|x_N\|\frac{1}{1-\alpha} < +\infty.$$

(2) Let $\beta > 1$ such that

$$\liminf_{n\to+\infty} \frac{\|x_{n+1}\|}{\|x_n\|} > \beta.$$

$$\exists N \in \mathbb{N}, x_N \neq 0, \text{ and } \forall n \in \mathbb{N}_{\geq N}, \|x_{n+1}\| \geq \beta\|x_n\|$$

$$\forall n \geq N, \|x_n\| \geq \beta^{n-N}\|x_N\| \to +\infty(n \to +\infty)$$

So $\sum_{n\in\mathbb{N}} x_n$ diverges. □

Let $z \in \mathbb{C}$. The series $\sum_{n\in\mathbb{N}} \frac{z^n}{n!}$ converges absolutely since

$$\left|\frac{z^{n+1}/(n+1)!}{z^n/n!}\right| = \frac{|z|}{n+1} \to 0(n \to +\infty).$$

We denote by e^z this limit.

1.9 Local inversion

Definition 1.9.1 Let X be a topological space and $Y \subseteq X$. If $\overline{Y} = X$, we say that Y is dense.

Theorem 1.9.2 (Baire) Let (X, d) be a complete metric space. Let $(\Omega_n)_{n\in\mathbb{N}}$ be a sequence of dense open subset of X. Let $\Omega = \bigcap_{n\in\mathbb{N}} \Omega_n$, then Ω is dense in X.

Proof Suppose that Ω is not dense. Let $x_0 \in X\backslash\overline{\Omega}$, exists $\varepsilon > 0$ such that $B(x_0, \varepsilon) \subseteq X\backslash\overline{\Omega}$.
Let $r_0 = \varepsilon$. We construct in a recursive way sequence $(x_n)_{n\in\mathbb{N}} \in X^{\mathbb{N}}$ and $(r_n)_{n\in\mathbb{N}} \in \mathbb{R}^n_{\geq 0}$ as follows.

Suppose that (x_n, r_n) is chosen. $B(x_n, r_n) \cap \Omega_n \neq \varnothing$. We pick $x_{n+1} \in X$ and $r_{n+1} \leq \frac{x_n}{2}$ such that $B(x_{n+1}, r_{n+1}) \subseteq B(x_n, r_n) \cap \Omega_n$, $\mathrm{d}(x_{n+1}, x_n) < r_n$. $\sum_{n\in\mathbb{N}} r_n < +\infty$ (ratio test).

Then the sequence converges to some l. For any $n \in \mathbb{N}$, $x_n \in B(x_0, \varepsilon)$. So $l \in \overline{B}(x_0, \varepsilon)$.

Moreover, $\forall n \in \mathbb{N}$, $l \in \overline{B}(x_{n+1}, r_{n+1}) \subseteq B_{x_n, r_n} \cap \Omega_n$. Thus $l \in \cap_{n\in\mathbb{N}}\Omega_n = \Omega$. Contradiction. $\square$

Corollary 1.9.3 Let (X, d) be a non-empty complete metric space and $(Y_n)_{n\in\mathbb{N}}$ be a family of closed subsets of X such that $X = \cup_{n\in\mathbb{N}} Y_n$. Then exists $n \in \mathbb{N}$ such that $Y^\circ \neq \varnothing$.

Proof Let $\Omega_n = X\backslash Y_n$. Suppose that $\forall n \in \mathbb{N}$, $Y_n^\circ = \varnothing$. Then $\overline{\Omega}_n = X\backslash Y_n^\circ = X$. Thus $\Omega := \cap_{n\in\mathbb{N}} \Omega_n$ is dense in X. Namely, $X = \overline{\Omega}$. So

$$\varnothing = X\backslash\overline{\Omega} = (X\backslash\Omega)^\circ = \left(X\backslash \bigcap_{n\in\mathbb{N}} \Omega_n\right)^\circ = \left(\bigcup_{n\in\mathbb{N}} Y_n\right)^\circ = X^\circ = X.$$

Contradiction. $\square$

Theorem 1.9.4 (Banach) Let $(K, |\cdot|)$ be a complete non-trivially valued filed, and E be a vector space over K. Let $\|\cdot\|_1$, $\|\cdot\|_2$ be two norms on E such that $(E, \|\cdot\|_1)$ and $(E, \|\cdot\|_2)$ are both Banach spaces.

If $\exists C > 0$ such that $\|\cdot\|_2 \leq C\|\cdot\|_1$. Then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. ($\exists C' > 0$, $\|\cdot\|_1 \leq C'\|\cdot\|_2$)

Proof For $x \in E$ and $r > 0$. Let

$$B_i(x,r) := \{y \in E \mid \|y-x\|_i < r\},\ i = 1,2$$

$\forall y \subseteq E$, let $\overline{Y}^{\|\cdot\|_2}$ be the closure of Y in $(E, \|\cdot\|_2)$.

$$E = \bigcup_{n\geq 1} B_1(0,n) = \bigcup_{n\geq 1} \overline{B_1(0,n)}^{\|\cdot\|_2}.$$

Hence, $\exists n_0 \geq 1$, $p \in E$, $r_0 > 0$ such that

$$B_2(p, r_0) \subseteq \overline{B_1(0,n_0)}^{\|\cdot\|_2}$$

or equivalently,

$$B_2(0,r_0) \subseteq \overline{B_1(-p,n_0)}^{\|\cdot\|_2} \subseteq \overline{B_1(0,n_0+\|p\|_1)}^{\|\cdot\|_2}$$

since $\forall x \in B_1(-p, n_0)$

$$\|x\|_1 = \|x-p+p\|_1 \leq \|x-p\| + \|p\|_1 < n_0 + \|p\|_1.$$

Let $r_1 = n_0 + \|p\|_1$,

$$B_2(0,r_0) \subseteq \overline{B_1(0,r_1)}^{\|\cdot\|_2} \subseteq B_1(0,r_1) + B_2(0, r_0|a|)$$

where $a \in K$, $0 < |a| < \frac{1}{2}$.
In fact, $\forall x \in \overline{B_1(0,r_0)}^{\|\cdot\|_2}$, exists sequence $(x_n)_{n\in\mathbb{N}} \in B_1(0,r_1)^{\mathbb{N}}$, such that $x_n \to x$ in $(E, \|\cdot\|_2)$, $\exists n \in \mathbb{N}$, $\|x_n - x\|_2 < r_0|a|$

$$B_2(0, r_0|a|^n) \subseteq B_1(0, r_1|a|^n) + B_2(0, r_0|a|^{n+1})$$

Let $y \in B_2(0,r_0)$, we choose $(x_0, y_0) \in B_1(0,r_1) \times B_2(0, r_0|a|)$ such that $y = x_0 + y_0$. When (x_n, y_n) si chosen, let $(x_{n+1}, y_[n+1]) \in B_1(0, r_0|a|^{n+1}) \times B_2(0, r_0|a|^{n+2})$, $y_n = x_{n+1} + y_{n+1}$, $y = y_n + \sum_{k=0}^n x_k$. So $\sum_{n\in\mathbb{N}} x_n$ converges to y.
Moreover, $\sum_{n\in\mathbb{N}} \|x_n\| < +\infty$, so it converges in $(E, \|\cdot\|_1)$ to some x. Therefore, $x = y$ since $\|\cdot\|_2 \leq C\|\cdot\|_1$. So $\|y\|_1 \|x\|_1 \leq \sum_{n\in\mathbb{N}} \|x_n\|_1 \leq \frac{r_1}{1-|a|}$.

Therefore $B_2(0,r_0) \subseteq B_1(0, \frac{r}{1-|a|})$. So $\|\cdot\|_1$ is bounded by a constant times $\|\cdot\|_2$. $\square$

Proposition 1.9.5 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be Banach spaces over a complete non-trivially valued filed $(K, |\cdot|)$, and $f : E \longrightarrow F$ be a bounded mapping.
(1) If f is invertible, then f^{-1} is bounded.
(2) If f is surjective, for any $U \subseteq E$ open, $f(U)$ is open in F.

Proof
(1) We define a mapping

$$\begin{array}{rccc} \|\cdot\|'_E : & E & \longrightarrow & \mathbb{R}_{\geq 0} \\ & x & \longmapsto & \|f(x)\|_F. \end{array}$$

This is a norm on E. In fact, if $\|x\|'_E = \|f(x)\|_F = 0$, then $f(x) = 0_F$. So $x = 0_E$. Moreover,

$$\forall x \in E,\ \|x\|'_E = \|f(x)\|_F \leq \|f\|\|x\|_E.$$

So there exists $C > 0$ such that $\|\cdot\|_E \leq C\|\cdot\|'_E$. That is,

$$\forall y \in F,\ \|y\|_F = \|f(f^{-1}(y))\|_F = \|f^{-1}(y)\|'_E \geq C^{-1}\|f^{-1}(y)\|_E.$$

So, $\|f^{-1}\| \leq C$.
(2) Let

$$E_0 = \ker(f) = \{x \in E \mid f(x) = 0_F\}.$$

This is a closed vector subspace of E. $\|\cdot\|_E$ induces by passing to quotient a norm $\|\cdot\|_Q$ on $Q := E/E_0$. Let

$$\begin{array}{rccc} g : & Q & \longrightarrow & F \\ & [x] & \longmapsto & f(x). \end{array}$$

This is a K-linear bijection.

If $\alpha \in Q$,

$$\forall x \in \alpha,\ \|g(\alpha)\|_F = \|f(x)\|_F \leq \|f\| \cdot \|x\|_E.$$

Since $\|\alpha\|_Q := \inf_{x\in\alpha} \|x\|_E$, $\|g(\alpha)\|_F \leq \|f\| \cdot \|\alpha\|_Q$. So $\|g\| \leq \|f\|$. By (1), g^{-1} is bounded (hence is continuous).

If $V \subseteq Q$ is open, then $g(V) \subseteq F$ is open. Let $U \subseteq E$ be an open subset. Let

$$\begin{array}{rccc} \pi : & E & \longrightarrow & Q \\ & x & \longmapsto & [x]. \end{array}$$

Let $x \in U$, $r > 0$ such that $B(x, r) \subseteq U$. For any $\alpha \in Q$, if

$$\|\alpha - [x]\|_Q = \inf_{y\in\alpha} \|y - x\|_E < r,$$

then, exists $y \in \alpha$ such that $\|y - x\|_E < r$.
Therefore,

$$B([x], r) \subseteq \pi(B(x, r)) \subseteq \pi(U).$$

This means that $\pi(U)$ is open. So $f(U) = g(\pi(U))$ is open. □

Definition 1.9.6 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector space over a complete non-trivially valued field $(K, |\cdot|)$, $U \subseteq E$ open, $f : U \longrightarrow F$. If $\forall p \in U$, f is n-times differentiable at p, and $\mathrm{D}^n f : U \longrightarrow \mathscr{L}^{(n)}(E, \dots, E, F)$ is continuous, we say that f is of class $\mathcal{C}^n$.

If $\forall n \in \mathbb{N}$, f is n-times differentiable on U, we say that f is smooth, or of class $\mathcal{C}^\infty$. ($\forall n \in \mathbb{N}$, f is of class $\mathcal{C}^n$.)

Proposition 1.9.7 Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ and $(G, \|\cdot\|_G)$ be normed vector space over a complete non-trivially valued field $(K, |\cdot|)$. $U \subseteq E$, $V \subseteq F$ be open subsets, $f : U \longrightarrow V$, $g : V \longrightarrow G$ be mappings. $n \in \mathbb{N}$.
(1) Let $p \in U$. If f is n-times differentiable at p and g is n-times differentiable at $f(p)$, then $g \circ f$ is n-times differentiable at p.
(2) If f is of class $\mathcal{C}^n$ on U and g is of class $\mathcal{C}^n$ on V, then $g \circ f$ is of class $\mathcal{C}^n$ on U.

Proof (induction on n)
$n = 0$, continuity composition.
$n = 1$, differentiability of composition.
$n \geq 2$,

$$\mathrm{D}(g \circ f)(p)(\cdot) = \mathrm{D}g(f(p))(\mathrm{D}f(p)(\cdot))$$

Let

$$\begin{array}{rccl} \Phi : & \mathscr{L}(F, G) \times \mathscr{L}(E, F) & \longrightarrow & \mathscr{L}(E, G) \\ & (\alpha, \beta) & \longmapsto & \alpha \circ \beta. \end{array}$$

This is a bounded bilinear mapping. $\|\alpha \circ \beta\| \leq \|\alpha\| \cdot \|\beta\|$.

$$(\|\alpha \circ \beta(h)\|_G = \|\alpha(\beta(h))\|_G \leq \|\alpha\| \cdot \|\beta(h)\|_F \leq \|\alpha\|\|\beta\|\|h\|_E)$$

Φ is of class $\mathcal{C}^\infty$.

$$\mathrm{D}(g \circ f) = \Phi(\mathrm{D}g \circ f, \mathrm{D}f).$$

(2) Since $\mathrm{D}g$ and $\mathrm{D}f$ is of class $\mathcal{C}^{n-1}$, we obtain that $\mathrm{D}(g \circ f)$ is of class $\mathcal{C}^{n-1}$, so $g \circ f$ is of class $\mathcal{C}^n$.
(1) If g is n-times differentiable at $f(p)$, $\mathrm{D}g$ is $(n-1)$-times differentiable at $f(p)$.

So $\mathrm{D}g \circ f$ is $(n-1)$-times differentiable at p. $\mathrm{D}f$ is $(n-1)$-times differentiable at p. So $\mathrm{D}(g \circ f)$ is $(n-1)$-times differentiable at p. $\square$

Theorem 1.9.8 Let $(E, \|\cdot\|)$ be a Banach space over a complete non-trivially valued filed $(K, |\cdot|)$. Let

$$\mathrm{GL}(E) := \{\varphi \in \mathscr{L}(E,E) \mid \varphi \text{ is invertible}\}.$$

This set forms a group under $\circ$.
(1) $\forall \varphi \in \mathscr{L}(E,E)$, if $\|\varphi\| < 1$, then $\mathrm{Id}_E + \varphi \in \mathrm{GL}(E)$.
(2) $\mathrm{GL}(E) \subseteq \mathscr{L}(E,E)$ is open.
(3)

$$\begin{array}{rccc} \iota: & \mathrm{GL}(E) & \longrightarrow & \mathrm{GL}(E) \\ & \varphi & \longmapsto & \varphi^{-1} \end{array}$$

is of class $\mathcal{C}^\infty$.

Proof
(1) The series $\sum_{n\in\mathbb{N}}(-1)^n\varphi^n$ converges absolutely since $\|\varphi^n\| \leq \|\varphi\|^n$.
Let η be the limit of $\sum_{n\in\mathbb{N}}(-1)^n\varphi^n$.

$$(\mathrm{Id} + \varphi) \circ \sum_{k=0}^{n}(-1)^k\varphi^k = \mathrm{Id} + (-1)^n\varphi^{n+1}.$$

Taking the limit when $n \to +\infty$, we get $(\mathrm{Id}_E + \varphi) \circ \eta = \mathrm{Id}_\mathrm{E}$. For the same reason, $\eta \circ (\mathrm{Id}_E + \varphi) = \mathrm{Id}_E$.
(2) If $f \in \mathrm{GL}(E)$, $\forall \varphi \in \mathscr{L}(E,E)$ such that

$$\|\varphi\| < \frac{1}{\|f^{-1}\|},\ f + \varphi = f \circ \left(\mathrm{Id}_E + f^{-1} \circ \varphi\right),\ \|f^{-1} \circ \varphi\| \leq \|f^{-1}\| \cdot \|\varphi\| < 1.$$

So $\mathrm{Id}_E + f^{-1} \circ \varphi \in \mathrm{GL}(E)$. Hence $f + \varphi \in \mathrm{GL}(E)$.
(3) Let $f \in \mathrm{GL}(E)$, $\varphi \in \mathscr{L}(E,E)$. $\|\varphi\| \leq \frac{1}{\|f^{-1}\|}$.

$$\begin{aligned} \iota(f+\varphi) - \iota(f) &= (f+\varphi)^{-1} - f^{-1} \\ &= \left(f \circ \left(\mathrm{Id}_E + f^{-1} \circ \varphi\right)\right)^{-1} - f^{-1} \\ &= \left(\mathrm{Id}_E + f^{-1} \circ \varphi\right)^{-1} \circ f^{-1} - f^{-1} \\ &= \sum_{n\in\mathbb{N}}(-1)^n \left(f^{-1} \circ \varphi\right)^n \circ f^{-1} - f^{-1} \\ &= -f^{-1} \circ \varphi \circ f^{-1} + o(\|\varphi\|) \end{aligned}$$

since

$$\begin{aligned}
&\sum_{n\geq 2}\|(-1)^n\left(f^{-1}\circ\varphi\right)^n\circ f^{-1}\| \\
\leq&\sum_{n\geq 2}\|f^{-1}\|\cdot\left(\|f^{-1}\|\cdot\|\varphi\|\right)^n \\
=&\|\varphi\|^2\left(\|f\|^3\cdot\sum_{n\geq 2}\left(\|f^{-1}\|\cdot\|\varphi\|\right)^{n-2}\right) \\
=&o(\|\varphi\|).
\end{aligned}$$

Let

$$\begin{array}{rccc}\Phi: & \mathscr{L}(E,E)^3 & \longrightarrow & \mathscr{L}(E,E) \\ & (\alpha,\beta,\gamma) & \longmapsto & \alpha\circ\beta\circ\gamma.\end{array}$$

bounded 3-linear mapping.

$$\mathrm{D}\iota(f)(\cdot)=-\Phi\left(\iota(f),\cdot,\iota(f)\right).$$

By induction, we obtain that ι is of class $\mathcal{C}^n$ for any $n\in\mathbb{N}$. □

Definition 1.9.9 Let (X,d) be a metric space, $f:X\longrightarrow X$ be a mapping. If exists $\alpha\in\,]0,1[$, such that f is α-Lipschitzian, we say that f is a **contraction**.

Definition 1.9.10 Let $f:X\longrightarrow X$ be a mapping. If $x\in X$ id such that $f(x)=x$, we say that x is a **fixed point** of f.

Theorem 1.9.11 (Banach fixed point theorem) Let (X,d) be a non-empty complete metric space and $f:X\longrightarrow X$ be a contraction. Then f admits a unique fixed point.

Proof
" Uniqueness": Let $\alpha\in\,]0,1[$, such that f is α-Lipschitzian. If a and b are fixed point of f, then $\mathrm{d}(a,b)=\mathrm{d}(f(a),f(b))\leq\alpha\mathrm{d}(a,b)$. So $\mathrm{d}(a,b)=0$, $a=b$.
" Existence": Let $x_0\in X$. For any $n\in\mathbb{N}$, let $x_n=f^n(x_0)$. Then

$$\mathrm{d}(x_n,x_{n+1})=\mathrm{d}(f(x_{n-1}),f(x_n))\leq\alpha\mathrm{d}(x_{n-1},x_n)\leq\cdots\leq\alpha^n\mathrm{d}(x_0,x_1).$$

So

$$\sum_{n\in\mathbb{N}}\mathrm{d}(x_n,x_{n+1})\leq\sum_{n\in\mathbb{N}}\alpha^n\mathrm{d}(x_0,x_1)=\frac{1}{1-\alpha}\mathrm{d}(x_0,x_1)<+\infty.$$

Hence $(x_n)_{n\in\mathbb{N}}$ converges to some $a \in X$.

$$\mathrm{d}(a, f(a)) = \lim_{n\to+\infty} \mathrm{d}(x_n, f(x_n)) = \lim_{n\to+\infty} \mathrm{d}(x_n, x_{n+1}) = 0.$$

So $a = f(a)$. $\square$

Definition 1.9.12 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over a complete value filed $(K, |\cdot|)$, $U \subseteq E$, $V \subseteq F$ be open subsets, $f : U \longrightarrow V$ be a bijection, $n \in \mathbb{N} \cup \{\infty\}$. If f and f^{-1} are both of class $\mathcal{C}^n$, we say that f is a $\mathcal{C}^n$-diffeomorphism.

Theorem 1.9.13 Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be Banach spaces over $\mathbb{R}$, $U \subseteq E$ open and $f : U \longrightarrow F$ be a mapping of class $\mathcal{C}^n$ $(n \in \mathbb{N} \cup \{\infty\})$. Let $p \in U$. Suppose that $\mathrm{D}f(p) \in \mathscr{L}(E, F)$ is invertible. Then there exists a open neighborhood V of p contained in U, such that $f|_V : V \longrightarrow f(V)$ is a $\mathcal{C}^n$-homeomorphism. Moreover

$$\mathrm{D}f^{-1}(y) = \mathrm{D}f(f(y))^{-1}.$$

Proof By replacing f by

$$\tilde{f} : x \longmapsto \mathrm{D}f(p)^{-1} \left(f(p+x) - f(p)\right).$$

We may assume that $E = F$, $p = f(p) = 0$, $\mathrm{D}f(p) = \mathrm{Id}_E$.

$$\mathrm{D}\tilde{f}(0)(h) = \mathrm{D}f(p)^{-1}(\mathrm{D}f(p)(h)) = h, \mathrm{D}\tilde{f}(0) = \mathrm{Id}_E.$$

Let $\mu : U \longrightarrow E$, $\mu(x) = f(x) - x$, $\mathrm{D}\mu(0) = 0$. Since $\mathrm{D}f$ is continuous, so is $\mathrm{D}\mu$.

$$\exists r > 0,\ \forall x \in \overline{B}(0_E, r),\ \|\mathrm{D}\mu(x)\| \leq \frac{1}{2}.$$

So μ is $\frac{1}{2}$-Lipschitzian on $\overline{B}(0_E, r)$ (mean value inequality).

$$\forall (x, y) \in \overline{B}(0_E, r)^2,\ \|f(x) - f(y)\| \geq \|x - y\| \|\mu(x) - \mu(y)\| \geq \frac{1}{2}\|x - y\|.$$

So f is injective on $\overline{B}(0_E, r)$. Let $a \in \overline{B}(0_E, \frac{r}{2})$.

$$\forall x \in \overline{B}(0_E, r),\ \|a - \mu(x)\| \leq \|a\| + \|\mu(x)\| \leq \frac{1}{2}r + \frac{1}{2}r = r.$$

Let

$$\begin{array}{rccc} \nu: & \overline{B}(0,r) & \longrightarrow & \overline{B}(0,r) \\ & x & \longmapsto & a-\mu(x) \end{array}$$

ν is a contraction. By Banach's fixed point theorem,

$$\exists! g(a) \in \overline{B}(0,r),\ \nu(g(a)) = a-\mu(g(a)) = a - f(g(a)).$$

That is $f(g(a)) = a$. Let $W = B(0, \frac{r}{2})$, $V = f^{-1}(W) \cap B(0,r)$, $f|_V : V \longrightarrow W$ is a bijection.

$\forall z \in B(0,r), \mathrm{D}f(z) = \mathrm{Id}_E + \mathrm{D}\mu(z) \in \mathrm{GL}(E)$.

$\forall (x,x_0) \in V \times V, y = f(x), y_0 = f(x_0), y-y_0 = \mathrm{D}f(x_0)(x-x_0) + o\left(\|x-x_0\|\right).$

$$\|x-x_0\| = \|y-y_0-(\mu(x)-\mu(x_0))\| \leq \|y-y_0\| + \frac{1}{2}\|x-x_0\|,$$

$$\frac{1}{2}\|f^{-1}(y)-f^{-1}(y_0)\| = \frac{1}{2}\|x-x_0\| \leq \|y-y_0\|.$$

So,

$$\mathrm{D}f(x_0)(x-x_0) = y-y_0+o(\|y-y_0\|),$$

$$\begin{aligned} f^{-1}(y)-f^{-1}(y_0) =& x-x_0 = \mathrm{D}f(x_0)^{-1}(y-y_0) + o(\|y-y_0\|) \\ =& \mathrm{D}f(f^{-1}(y_0))^{-1}(y-y_0) + o(\|y-y_0\|) \end{aligned}$$

Thus,

$$\mathrm{D}f^{-1} = \iota \circ \mathrm{D}f \circ f^{-1}.$$

□