DIFFERENTIABILITY

1. Introduction

In this notes we collect basic results on derivatives of functions defined on open subsets of real or complex Banach spaces.
Symbol $\mathbb{K}$ denotes the base field which is either $\mathbb{R}$ or $\mathbb{C}$.

2. Preliminaries on bounded multilinear forms on normed spaces over $\mathbb{K}$

In this section we fix a positive integer n and consider normed spaces $\mathfrak{D}_1, ..., \mathfrak{D}_n, \mathfrak{X}$ over $\mathbb{K}$.

Definition 2.1. Let $L : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$ be a $\mathbb{K}$-multilinear form. Suppose that there exists $C > 0$ such that

$$||L(x_1, ..., x_n)|| \leq C \cdot ||x_1|| \cdot ... \cdot ||x_n||$$

for every $x_i \in \mathfrak{D}_i$ for $i \in \{1, ..., n\}$. Then L is *bounded*.

The following result characterizes bounded multilinear forms.

Theorem 2.2. *Let $L : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$ be a $\mathbb{K}$-multilinear form. Then the following assertions are equivalent.*

- **(i)** *L is continuous.*
- **(ii)** *L is continuous at zero n-tuple in $\mathfrak{D}_1 \times ... \times \mathfrak{D}_n$.*
- **(iii)** *L is bounded.*

Proof. The implication **(i)** $\Rightarrow$ **(ii)** is obvious.
Suppose that L is continuous at zero n-tuple in $\mathfrak{D}_1 \times ... \times \mathfrak{D}_n$. Assume that there exists a sequence $\big\{(x_{m,1}, ..., x_{m,n})\big\}_{m \in \mathbb{N}_+}$ such that $x_{m,i} \in \mathfrak{D}_i$, $||x_{m,i}|| = 1$ for each i and

$$||L(x_{m,1}, ..., x_{m,n})|| \geq m$$

for each $m \in \mathbb{N}_+$. Define $y_{m,i} = \frac{1}{\sqrt[n]{m}} \cdot x_{m,i}$ for every $i \in \{1, ..., n\}$ and every $m \in \mathbb{N}_+$. Then $\{y_{m,i}\}_{m \in \mathbb{N}_+}$ tends to zero for every $i \in \{1, ..., n\}$ and

$$||L(y_{m,1}, ..., y_{m,n})|| \geq 1$$

for every $m \in \mathbb{N}_+$. This is a contradiction with the assumption that L is continuous at zero n-tuple in $\mathfrak{D}_1 \times ... \times \mathfrak{D}_n$. Therefore, there exists $C > 0$ such that

$$||L(x_1, ..., x_n)|| \leq C$$

for every $x_i \in \mathfrak{D}_i$ with $||x_i|| = 1$ for $i \in \{1, ..., n\}$. Thus the implication **(ii)** $\Rightarrow$ **(iii)** holds.
Assume that L is bounded. Pick $x_i \in \mathfrak{D}_i$ and $h_i \in \mathfrak{D}_i$ for $i \in \{1, ..., n\}$. Define

$$z_0 = (x_1, ..., x_n),\ z_i = (x_1 + h_1, ..., x_i + h_i, x_{i+1}, ..., x_n)$$

for $i \in \{1, ..., n\}$. Then

$$\big|\big|L\left(x_1 + h_1, ..., x_n + h_n\right) - L\left(x_1, ..., x_n\right)\big|\big| = \left|\left|\sum_{i=1}^{n}\left(L(z_i) - L(z_{i-1})\right)\right|\right| \leq$$

$$\leq \sum_{i=1}^{n}\big|\big|L(z_i) - L(z_{i-1})\big|\big| \leq \sum_{i=1}^{n} C \cdot ||x_1|| \cdot ... \cdot ||x_{i-1}|| \cdot ||h_i|| \cdot ||x_{i+1}|| \cdot ... \cdot ||x_n||$$

Thus if $(h_1, ..., h_n) \to 0$, then $L\left(x_1 + h_1, ..., x_n + h_n\right) - L\left(x_1, ..., x_n\right) \to 0$. This shows that L is continuous. Hence the proof of the implication **(iii)** $\Rightarrow$ **(i)** is completed. $\square$

Definition 2.3. Let $L : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$ be a bounded $\mathbb{K}$-multilinear form. Then we define

$$\|L\| = \sup\left\{\|L\left(x_1, ...x_n\right)\| \,\middle|\, \forall_{i\in\{1,...,n\}}\, x_i \in \mathfrak{D}_i \text{ and } \|x_i\| = 1\right\}$$

and call it *the operator norm of* L.

Fact 2.4. *Let $L : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$ be a bounded $\mathbb{K}$-multilinear form. Then*

$$\|L(x_1, ..., x_n)\| \leq \|L\| \cdot \|x_1\| \cdot ... \cdot \|x_n\|$$

for every $(x_1, ..., x_n) \in \mathfrak{D}_1 \times ... \times \mathfrak{D}_n$.

Proof. Left for the reader as an exercise. □

Theorem 2.5. *Let $L\left(\mathfrak{D}_1, ..., \mathfrak{D}_n; \mathfrak{X}\right)$ be a $\mathbb{K}$-vector space of bounded multilinear forms $\mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$ with respect to operations defined pointwise. Suppose that $\mathfrak{X}$ is a Banach space over $\mathbb{K}$. Then*

$$L\left(\mathfrak{D}_1, ..., \mathfrak{D}_n; \mathfrak{X}\right) \ni L \mapsto \|L\| \in [0, +\infty)$$

is a norm which makes $L\left(\mathfrak{D}_1, ..., \mathfrak{D}_n; \mathfrak{X}\right)$ into a Banach space over $\mathbb{K}$.

Proof. We left as an exercise the proof that operator norm is well defined vector space norm on $L\left(\mathfrak{D}_1, ..., \mathfrak{D}_n; \mathfrak{X}\right)$. Consider a Cauchy's sequence $\{L_m\}_{m\in\mathbb{N}}$ with respect to operator norm. Then $\{\|L_m\|\}_{m\in\mathbb{N}}$ is Cauchy's sequence and hence is convergent in $\mathbb{R}$. Fix $(x_1, ...x_n) \in \mathfrak{D}_1 \times ... \times \mathfrak{D}_n$. Then by Fact 2.4

$$\|\left(L_m - L_k\right)\left(x_1, ..., x_n\right)\| \leq \|L_m - L_k\| \cdot \|x_1\| \cdot ... \cdot \|x_n\|$$

for every $m, k \in \mathbb{N}$. This implies that $\left\{L_m\left(x_1, ..., x_n\right)\right\}_{m\in\mathbb{N}}$ is a Cauchy's sequence in $\mathfrak{X}$. Since $\mathfrak{X}$ is a Banach space over $\mathbb{K}$, we derive that this sequence is convergent. We define

$$L\left(x_1, ..., x_n\right) = \lim_{m\to+\infty} L_m\left(x_1, ..., x_n\right)$$

Note that we have

$$\|L(x_1, ..., x_n)\| = \lim_{m\to+\infty} \|L_m\left(x_1, ..., x_n\right)\| \leq \left(\lim_{m\to+\infty} \|L_m\|\right) \cdot \|x_1\| \cdot ... \cdot \|x_n\|$$

Therefore, $L : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$ is a bounded $\mathbb{K}$-multilinear form. We claim that L is the limit of $\{L_m\}_{m\in\mathbb{N}}$ with respect to operator norm. For the proof fix $(x_1, ...x_n) \in \mathfrak{D}_1 \times ... \times \mathfrak{D}_n$ such that $\|x_1\| = ... = \|x_n\| = 1$. Then

$$\|\left(L - L_m\right)\left(x_1, ...x_n\right)\| \leq \|L(x_1, ..., x_n) - L_k(x_1, ..., x_m)\| + \|L_k - L_m\|$$

Thus we have

$$\|\left(L - L_m\right)\left(x_1, ...x_n\right)\| \leq \limsup_{k\to+\infty} \|L_k - L_m\|$$

The left hand side does not depend on $x_1, ..., x_n$ and we deduce that

$$\|L - L_m\| \leq \limsup_{k\to+\infty} \|L_k - L_m\|$$

Invoking once again the assumption that $\{\|L_m\|\}_{m\in\mathbb{N}}$ is Cauchy's sequence we infer

$$\lim_{m\to+\infty} \|L - L_m\| \leq \lim_{m\to+\infty} \limsup_{k\to+\infty} \|L_k - L_m\| = 0$$

This completes the proof. □

Proposition 2.6. *The canonical map $L\left(\mathfrak{D}_1, ..., \mathfrak{D}_{n-1}; L\left(\mathfrak{D}_n, \mathfrak{X}\right)\right) \to L\left(\mathfrak{D}_1, ..., \mathfrak{D}_n; \mathfrak{X}\right)$ which sends L in $L\left(\mathfrak{D}_1, ..., \mathfrak{D}_{n-1}; L\left(\mathfrak{D}_n, \mathfrak{X}\right)\right)$ to a $\mathbb{K}$-multilinear form given by formula*

$$\mathfrak{D}_1 \times ... \times \mathfrak{D}_n \ni \left(x_1, ..., x_n\right) \mapsto L(x_1, ..., x_{n-1})(x_n) \in \mathfrak{X}$$

is an isometry of normed spaces.

Proof. Left for the reader as an exercise. □

3. Notion of Fréchet derivatives

In this section we introduce derivatives and prove their basic properties. We fix Banach spaces $\mathfrak{D}, \mathfrak{X}$ over $\mathbb{K}$. Let U be an open subset of $\mathfrak{D}$ and let V be an open subset of $\mathfrak{X}$.

Fact 3.1. *Let x be a point in U and let $f : U \to V$ be a function. Suppose that there are continuous $\mathbb{K}$-linear maps $L_i : \mathfrak{D} \to \mathfrak{X}$ for $i = 1, 2$. If both functions*

$$\left\{h \in \mathfrak{D} \,\middle|\, h \neq 0 \text{ and } x + h \in U\right\} \ni h \mapsto \frac{f(x+h) - f(x) - L_i(h)}{||h||} \in \mathfrak{X}$$

tend to zero as $h \to 0$, then $L_1 = L_2$.

Proof. By assumption the function

$$\left\{h \in \mathfrak{D} \,\middle|\, h \neq 0 \text{ and } x + h \in U\right\} \ni h \mapsto (L_1 - L_2)\left(\frac{h}{||h||}\right) \in \mathfrak{X}$$

tends to zero as $h \to 0$. This implies that $L_1 - L_2$ sends each vector of the unit sphere in $\mathfrak{D}$ to zero. Thus $L_1 - L_2 = 0$ and hence $L_1 = L_2$. □

Definition 3.2. Let x be a point in U. A function $f : U \to V$ is *differentiable at point x* if there exists a continuous $\mathbb{K}$-linear map $L : \mathfrak{D} \to \mathfrak{X}$ such that the function

$$\left\{h \in \mathfrak{D} \,\middle|\, h \neq 0 \text{ and } x + h \in U\right\} \ni h \mapsto \frac{f(x+h) - f(x) - L(h)}{||h||} \in \mathfrak{X}$$

tends to zero as $h \to 0$. Moreover, the unique continuous $\mathbb{K}$-linear map L is *the derivative of f at x*.

Remark 3.3. Notion of differentiability defined above is named by some authors *Fréchet differentiability* after french mathematician Maurice Fréchet.

Remark 3.4. Let x be a point in U and let $f : U \to V$ be a function differentiable at point x. Then the derivative of f at x is usually denoted by $f'(x)$.

Fact 3.5. *Let x be a point in U and let $f : U \to V$ be a function differentiable at x. Then f is continuous at x.*

Proof. Consider the function $\phi_f(h)$ defined on the set

$$\left\{h \in \mathfrak{D} \,\middle|\, h \neq 0 \text{ and } x + h \in U\right\}$$

by formula $f(x+h) - f(x) = f'(x)(h) + \phi_f(h) \cdot ||h||$. By definition ϕ_f is continuous at zero. In order to complete the argument it suffices to note that the set $\left\{h \in \mathfrak{D} \,\middle|\, x + h \in U\right\}$ contains a neighborhood of zero in $\mathfrak{D}$. □

Definition 3.6. A function $f : U \to V$ is *differentiable* if it is differentiable at each point of U.

4. Chain rule

Chain rule is a basic tools for calculating Fréchet derivatives of a more complex functions.

Theorem 4.1. *Let $U \subseteq \mathfrak{D}$, $V \subseteq \mathfrak{X}$, $W \subseteq \mathfrak{Z}$ be open subsets of Banach spaces over $\mathbb{K}$ and let $f : U \to V$, $g : V \to W$ be functions. Suppose that f is differentiable at some point x in U and g is differentiable at $f(x)$. Then $g \cdot f$ is differentiable at x and the chain rule*

$$(g \cdot f)'(x) = g'(f(x)) \cdot f'(x)$$

holds.

Proof. Let L be derivative of f at x and let K be a derivative of g at $f(x)$. For h in $\mathfrak{D}$ such that $x+h\in U$ define $\phi_f(h)$ by formula

$$f(x+h)-f(x)-L(h)=\phi_f(h)\cdot\|h\|$$

Similarly for s in $\mathfrak{X}$ such that $f(x)+s\in V$ define $\phi_g(s)$ by formula

$$g(f(x)+s)-g(f(x))-K(s)=\phi_g(s)\cdot\|s\|$$

Now pick nonzero h in $\mathfrak{D}$ such that $x+h\in U$ and $f(x+h)\in V$. Then

$$\left\|g\left(f(x+h)\right)-g\left(f(x)\right)-K\left(L(h)\right)\right\|=\left\|\phi_g\big(f(x+h)-f(x)\big)\cdot\|f(x+h)-f(x)\|+K\big(\phi_f(h)\cdot\|h\|\big)\right\|\leq$$

$$\leq\left\|\phi_g\big(f(x+h)-f(x)\big)\right\|\cdot\|f(x+h)-f(x)\|+\left\|K\big(\phi_f(h)\big)\right\|\cdot\|h\|\leq$$

$$\leq\left\|\phi_g\big(f(x+h)-f(x)\big)\right\|\cdot\|f(x+h)-f(x)-L(h)\|+\left\|\phi_g\big(f(x+h)-f(x)\big)\right\|\cdot\|L(h)\|+\left\|K\big(\phi_f(h)\big)\right\|\cdot\|h\|\leq$$

$$\leq\Bigg(\left\|\phi_g\big(f(x+h)-f(x)\big)\right\|\cdot\|\phi_f(h)\|+\left\|\phi_g\big(f(x+h)-f(x)\big)\right\|\cdot\|L\|+\|K\|\cdot\|\phi_f(h)\|\Bigg)\cdot\|h\|$$

According to Fact 3.5 we have $f(x+h)-f(x)\to 0$ as $h\to 0$. Hence by differentiability of f at x and g at $f(x)$ we derive that

$$\phi_g\big(f(x+h)-f(x)\big)\to 0,\ \phi_f(h)\to 0$$

as $h\to 0$. Since

$$\left\{h\in\mathfrak{D}\,\middle|\,x+h\in U\text{ and }f(x+h)\in V\right\}$$

contains open neighborhood of zero in $\mathfrak{D}$, we derive that

$$\left\{h\in\mathfrak{D}\,\middle|\,h\neq 0\text{ and }x+h\in U\right\}\ni h\mapsto\frac{g\left(f(x+h)\right)-g\left(f(x)\right)-K\left(L(h)\right)}{\|h\|}\in\mathfrak{Z}$$

tends to zero as $h\to 0$. This completes the proof. □

5. Mean value inequality

The main topic of this section is extremely useful inequality, which connects derivatives with local change of a function.

Definition 5.1. Let $\mathfrak{D}$ be an affine space over $\mathbb{K}$ and let x_1,x_2 are points in $\mathfrak{D}$. The subsets

$$[x_1,x_2]=\left\{t\cdot x_1+(1-t)\cdot x_2\in\mathfrak{D}\,\middle|\,t\in[0,1]\right\},\ (x_1,x_2)=\left\{t\cdot x_1+(1-t)\cdot x_2\in\mathfrak{D}\,\middle|\,t\in(0,1)\right\}$$

of $\mathfrak{D}$ are called *the closed* and *the open interval with endpoints* x_1,x_2, respectively.

Theorem 5.2. *Let $U\subseteq\mathfrak{D},V\subseteq\mathfrak{X}$ be open subsets of Banach spaces over $\mathbb{K}$ and let $f:U\to V$ be a continuous function. Suppose that x_1,x_2 are points of $\mathfrak{D}$ such that $[x_1,x_2]\subseteq U$ and f is differentiable at every point in (x_1,x_2). Then the mean value inequality*

$$\|f(x_1)-f(x_2)\|\leq\|x_1-x_2\|\cdot\sup_{z\in(x_1,x_2)}\left\|f'(z)\right\|$$

holds.

Proof. For every $t\in[0,1]$ we define $x(t)=t\cdot x_1+(1-t)\cdot x_2$. Consider the continuous function $g:[0,1]\to V$ given by formula $g(t)=f\left(x(t)\right)$. Theorem 4.1 together with assumptions imply that g is differentiable over $\mathbb{R}$ at each point of $(0,1)$ as the composition of functions

$$[0,1]\ni t\mapsto x(t)\in U,\ f:U\to V$$

Moreover, its derivative is the map

$$\mathbb{R}\ni h\mapsto h\cdot f'\left(x(t)\right)\left(x_1-x_2\right)\in\mathfrak{X}$$

for every $t \in (0,1)$. Thus the mean value inequality is implied by the inequality

$$\|g(1) - g(0)\| \leq \sup_{t\in(0,1)} \left\|g'(t)\right\|$$

Fix $\epsilon > 0$ and consider the set

$$S = \left\{ s \in [0,1] \,\middle|\, \|g(s) - g(0)\| \leq s \cdot \sup_{t\in(0,s)} \|g'(t)\| + s\cdot\epsilon + \epsilon \right\}$$

We shall prove that S satisfies the following assertions.

(1) There exists $h > 0$ such that $[0,h) \subseteq S$.

(2) For every s in $S \cap (0,1)$ there exists $h > 0$ such that $s + h$ is contained in S.

(3) For every increasing sequence $\{s_n\}_{n\in\mathbb{N}}$ of elements of S its limit is contained in S.

The assertion **(1)** holds by continuity of g at zero.
Let us prove **(2)**. Write

$$g(s+h) - g(s) - g'(s)\cdot h = \phi_g(h)\cdot|h|$$

for $h > 0$ such that $s + h \leq 1$. Then $\phi_g(h)$ tends to zero as $h \to 0$ according to differentiability of g at $(0,1)$. Thus

$$\|g(s+h) - g(0)\| \leq \|g(s+h) - g(s)\| + \|g(s) - g(0)\| \leq$$
$$\leq \|\phi_g(h)\cdot h + g'(s)\cdot h\| + s\cdot \sup_{t\in(0,s)} \|g'(t)\| + s\cdot\epsilon + \epsilon \leq$$
$$\leq h\cdot\left(\|\phi_g(h)\| + \|g'(s)\|\right) + s\cdot \sup_{t\in(0,s)} \|g'(t)\| + s\cdot\epsilon + \epsilon$$

Since $\phi_g(h)$ tends to zero as $h \to 0$, we may pick $h > 0$ such that $s + h \leq 1$ and $\|\phi_g(h)\| \leq \epsilon$. Then

$$\|g(s+h) - g(0)\| \leq h\cdot\left(\epsilon + \|g'(s)\|\right) + s\cdot \sup_{t\in(0,s)} \|g'(t)\| + s\cdot\epsilon + \epsilon \leq$$
$$\leq (s+h)\cdot \sup_{t\in(0,s]} \|g'(t)\| + \epsilon\cdot(s+h) + \epsilon \leq (s+h)\cdot \sup_{t\in(0,s+h)} \|g'(t)\| + \epsilon\cdot(s+h) + \epsilon$$

and hence clearly $s + h$ is in S.
For the proof of **(3)** fix $\{s_n\}_{n\in\mathbb{N}}$ an increasing sequence of elements of S. Let s be its limit. For every $n \in \mathbb{N}$ we have

$$\|g(s_n) - g(0)\| \leq s_n\cdot \sup_{t\in(0,s_n)} \|g'(t)\| + s_n\cdot\epsilon + \epsilon \leq s\cdot \sup_{t\in(0,s)} \|g'(t)\| + s\cdot\epsilon + \epsilon$$

For $n \to +\infty$ we obtain

$$\|g(s) - g(0)\| \leq s\cdot \sup_{t\in(0,s)} \|g'(t)\| + s\cdot\epsilon + \epsilon$$

by continuity of g on $[0,1]$. Thus the proof of **(3)** is complete.
Using these three assertions we complete the proof. Note first that by **(1)** the set S contains some elements of $(0,1)$ and by **(3)** it contains its least upper bound. According to **(2)** the least upper bound of S cannot be contained in $(0,1)$. Thus the least upper bound of S is 1. This proves that

$$\|g(1) - g(0)\| \leq \sup_{t\in(0,1)} \|g'(t)\| + 2\cdot\epsilon$$

for every $\epsilon > 0$ and thus

$$\|g(1) - g(0)\| \leq \sup_{t\in(0,1)} \|g'(t)\|$$

The proof is complete. □

Corollary 5.3. *Let $U \subseteq \mathfrak{D}, V \subseteq \mathfrak{X}$ be open subsets of Banach spaces over $\mathbb{K}$ and let $f : U \to V$ be a differentiable function. If U is connected and derivative of f at each point of U is the zero map, then f is constant.*

Proof. Theorem 5.2 shows that for every open convex set $W \subseteq U$ the restriction $f_{|W}$ is constant. Let y be some element of $f(U)$. It follows that the set $f^{-1}(y)$ is open. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of elements of $f^{-1}(y)$ which is convergent to some point x in U. Pick an open and convex neighborhood W of x. Then for sufficiently large $n \in \mathbb{N}$ we have $x_n \in W$ and thus $x \in f^{-1}(y)$. Therefore, $f^{-1}(y)$ is closed. Hence $f^{-1}(y)$ is a clopen nonempty subset of a connected set U. This shows that $U = f^{-1}(y)$. □

6. Convergence of sequences of differentiable functions

In this section we prove important result concerning convergence of differentiable functions.

Definition 6.1. Let X be a topological space and let Y be a metric space. Let $\{f_n : X \to Y\}_{n\in\mathbb{N}}$ be a sequence of functions and $f : X \to Y$ be a function. Suppose that for every point x in X there exists an open neighborhood W of x in X such that the sequence $\{f_{n|W}\}_{n\in\mathbb{N}}$ converges uniformly to $f_{|W}$. Then the sequence $\{f_n\}_{n\in\mathbb{N}}$ is *locally uniformly convergent to* f.

Theorem 6.2. *Let $U \subseteq \mathfrak{D}$ be open subset of a Banach space $\mathfrak{D}$ over $\mathbb{K}$, let $\mathfrak{X}$ be a Banach space over $\mathbb{K}$ and let $\{f_n : U \to \mathfrak{X}\}_{n\in\mathbb{N}}$ be a sequence of functions. Assume that the following assertions hold.*

(1) *U is connected.*

(2) *There exists u in U such that the sequence $\{f_n(u)\}_{n\in\mathbb{N}}$ is convergent to some element of $\mathfrak{X}$.*

(3) *f_n is differentiable for every $n \in \mathbb{N}$.*

(4) *The sequence of maps*
$$\left\{U \ni x \mapsto f_n'(x) \in L(\mathfrak{D}, \mathfrak{X})\right\}_{n\in\mathbb{N}}$$
is locally uniformly convergent to a continuous map $g : U \to L(\mathfrak{D}, \mathfrak{X})$.

Then the sequence $\{f_n\}_{n\in\mathbb{N}}$ converges locally uniformly to a differentiable function $f : U \to \mathfrak{X}$ and $f'(x) = g(x)$ for every $x \in U$.

Proof. Suppose that z is a point of U such that $\{f_n(z)\}_{n\in\mathbb{N}}$ is convergent. Let W be a bounded, open and convex neighborhood of z in $\mathfrak{D}$ and assume that
$$\left\{W \ni x \mapsto f_n'(x) \in L(\mathfrak{D}, \mathfrak{X})\right\}_{n\in\mathbb{N}}$$
converges uniformly. Let x be a point of W. Then
$$\begin{aligned}\|f_n(x) - f_m(x)\| &\le \|(f_n(x) - f_m(x)) - (f_n(z) - f_m(z))\| + \|f_n(z) - f_m(z)\| \le \\ &\le \|x - z\| \cdot \sup_{y\in(x,z)} \left\|f_n'(y) - f_m'(y)\right\| + \|f_n(z) - f_m(z)\|\end{aligned}$$
Hence
$$\sup_{x\in W} \|f_n(x) - f_m(x)\| \le \operatorname{diam}(W) \cdot \sup_{x\in W} \left\|f_n'(x) - f_m'(x)\right\| + \|f_n(z) - f_m(z)\|$$
Since $\{f_n(z)\}_{n\in\mathbb{N}}$ is convergent, $\mathfrak{X}$ is complete and
$$\left\{W \ni x \mapsto f_n'(x) \in L(\mathfrak{D}, \mathfrak{X})\right\}_{n\in\mathbb{N}}$$
converges uniformly, we derive that $\{f_{n|W}\}_{n\in\mathbb{N}}$ converges uniformly. This proves that the sequence $\{f_{n|W}\}_{n\in\mathbb{N}}$ is uniformly convergent for every bounded, open and convex subset W of U such that the sequence
$$\left\{W \ni x \mapsto f_n'(x) \in L(\mathfrak{D}, \mathfrak{X})\right\}_{n\in\mathbb{N}}$$
converges uniformly and there exists $z \in W$ such that $\{f_n(z)\}_{n\in\mathbb{N}}$ is convergent.
We define $\mathcal{W}$ as the largest open subset of U such that $\{f_{n|\mathcal{W}}\}_{n\in\mathbb{N}}$ converges locally uniformly. Note that $u \in \mathcal{W}$ according to the first part of the proof. Suppose that $\{z_n\}_{n\in\mathbb{N}}$ is a sequence of

elements of $\mathcal{W}$ convergent to some point z in U. Pick a bounded, open and convex neighborhood W of z such that

$$\left\{W \ni x \mapsto f_n'(x) \in L(\mathfrak{D}, \mathfrak{X})\right\}_{n\in\mathbb{N}}$$

converges uniformly. Then for sufficiently large $n \in \mathbb{N}$ we have $z_n \in W$. Thus $\{f_{n|W}\}_{n\in\mathbb{N}}$ is uniformly convergent by the first part of the proof and hence z is in $\mathcal{W}$. This implies that $\mathcal{W}$ is a closed subset of U. Hence it is a clopen and nonempty subset of U. Since U is connected, we have $U = \mathcal{W}$ and $\{f_n\}_{n\in\mathbb{N}}$ is locally uniformly convergent to some function $f : U \to \mathfrak{X}$.
Fix x in U and let W be an open neighborhood of zero in $\mathfrak{D}$ such that $[x, x+h] \subseteq U$ for every $h \in W$. We apply Theorem 5.2 to a function $k_n : W \to \mathfrak{X}$ given by formula

$$k_n(h) = f_n(x+h) - f_n'(x)(h)$$

with derivative $k_n'(h) = f_n'(x+h) - f_n'(x)$ for all $h \in W$. We deduce that

$$\left\|f_n(x+h) - f_n(x) - f_n'(x)(h)\right\| = \|k_n(h) - k_n(0)\| \le \|h\| \cdot \sup_{z\in(0,h)} \|k_n'(z)\| =$$

$$= \|h\| \cdot \sup_{z\in(0,h)} \|f_n'(x+z) - f_n'(x)\| = \|h\| \cdot \sup_{z\in(x,x+h)} \|f_n'(z) - f_n'(x)\|$$

For $n \to +\infty$ we obtain that

$$\left\|f(x+h) - f(x) - g(x)(h)\right\| \le \|h\| \cdot \sup_{z\in(x,x+h)} \|g(z) - g(x)\|$$

Since g is continuous, we derive that

$$\lim_{h\to 0} \sup_{z\in(x,x+h)} \|g(z) - g(x)\| = 0$$

and thus $f'(x) = g(x)$. □

7. Partial derivatives

We fix Banach spaces $\mathfrak{D}_1, ..., \mathfrak{D}_n, \mathfrak{X}$ over $\mathbb{K}$ for some positive integer n. For each $i \in \{1, ..., n\}$ let $U_i \subseteq \mathfrak{D}_i$ be an open subset and let V be an open subset of $\mathfrak{X}$.

Definition 7.1. Consider a function $f : U_1 \times ... \times U_n \to V$ and a point $x = (x^1, ..., x^n) \in U_1 \times ... \times U_n$. Fix i in $\{1, ..., n\}$. Suppose that the restriction

$$f_{|\{(x^1,...,x^{i-1})\}\times U_i \times \{(x^{i+1},...,x^n)\}} : U_i \to V$$

is differentiable at x^i. Then its derivative is *the partial derivative of f at x along i-th axis.*

Remark 7.2. In the situation of the definition above we usually denote the partial derivative of f at x along i-th axis by the symbol

$$\frac{\partial f}{\partial x_i}(x)$$

Note that $\frac{\partial f}{\partial x_i}(x) : \mathfrak{D}_i \to \mathfrak{X}$ is a continuous $\mathbb{K}$-linear map.

Proposition 7.3. *Let $f : U_1 \times ... \times U_n \to V$ be a function differentiable at some point $x \in U_1 \times ... \times U_n$. Fix $i \in \{1, ..., n\}$. Then*

$$\frac{\partial f}{\partial x_i}(x) : \mathfrak{D}_i \to \mathfrak{X}$$

exists and is the composition of the canonical inclusion $\mathfrak{D}_i \hookrightarrow \mathfrak{D}_1 \times ... \times \mathfrak{D}_n$ with $f'(x) : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{X}$.

Proof. Suppose that ϕ_f is a function given by formula

$$\left\{s \in \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \,\middle|\, s \neq 0 \text{ and } x + s \in U\right\} \ni h \mapsto \frac{f(x+s) - f(x) - f'(x)(s)}{\|s\|} \in \mathfrak{X}$$

Denote the restriction $f_{|\{(x^1,...,x^{i-1})\}\times U_i \times \{(x^{i+1},...,x^n)\}}$ by f_i and denote the inclusion $\mathfrak{D}_i \hookrightarrow \mathfrak{D}_1 \times ... \times \mathfrak{D}_n$ by j_i. Recall that for every $h \in \mathfrak{D}_i$ we have $\|h\| = \|j_i(h)\|$. Pick $h \in \mathfrak{D}_i$ such that $h \neq 0$ and $x_i + h \in U_i$. We have

$$\frac{f_i(x_i+h) - f_i(x_i) - (f'(x)\cdot j_i)(h)}{\|h\|} = \frac{f(x+j_i(h)) - f(x) - f'(x)(j_i(h))}{\|j_i(h)\|} = \phi_f(j_i(h))$$

Since by definition of $f'(x)$ the function $\phi_f(s)$ tends to zero as $s \to 0$, we derive that $\phi_f(j_i(h))$ tends to zero as $h \to 0$. Thus the partial derivative of f at x along i-th axis exists and is given by formula $f'(x)\cdot j_i$. □

It is reasonable to ask for the converse of Proposition 7.3. The next theorem gives useful answer to this question under some additional assumptions.

Theorem 7.4. *Let $f : U_1 \times ... \times U_n \to V$ be a function and let x be a point in $U_1 \times .. \times U_n$. Suppose that the following two assertions hold.*

(1) *$\frac{\partial f}{\partial x_i}(u)$ exist for each $i \in \{1,...,n\}$ and every point $u \in U_1 \times ... \times U_n$.*

(2) *For each $i \in \{1,...,n\}$ the map*

$$U_1 \times ... \times U_n \ni u \mapsto \frac{\partial f}{\partial x_i}(u) \in L(\mathfrak{D}_i, \mathfrak{X})$$

is continuous at x.

Then f is differentiable at x and

$$f'(x) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)\cdot \mathrm{pr}_i$$

where $\mathrm{pr}_i : \mathfrak{D}_1 \times ... \times \mathfrak{D}_n \to \mathfrak{D}_i$ is the projection onto i-th axis.

Proof. Pick $h \in \mathfrak{D}_1 \times ... \times \mathfrak{D}_n$ such that $x + h \in U_1 \times ... \times U_n$. Write $x = (x^1,...,x^n)$ and $h = (h^1,...,h^n)$. Let $z_0 = x$ and

$$z_i = (x^n,...,x^{i+1}, x^i + h^i, ..., x^1 + h^1)$$

for each $i \in \{1,...,n\}$. Then

$$f(x+h) - f(x) - \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x)(h^i) =$$

$$= \sum_{i=1}^{n}\left(f(z_i) - f(z_{i-1}) - \frac{\partial f}{\partial x_i}(z_{i-1})(h^i)\right) + \sum_{i=1}^{n}\left(\frac{\partial f}{\partial x_i}(z_{i-1})(h^i) - \frac{\partial f}{\partial x_i}(x)(h^i)\right)$$

For each $i \in \{1,...,n\}$ define $\phi_i(h^i)$ by formula

$$f(z_i) - f(z_{i-1}) - \frac{\partial f}{\partial x_i}(z_{i-1})(h^i) = \phi_i(h^i)\cdot \|h^i\|$$

Then we have

$$\left\|f(x+h) - f(x) - \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x)(h^i)\right\| \leq$$

$$\leq \|h\| \cdot \sum_{i=1}^{n}\left(\|\phi_i(h^i)\| + \left\|\frac{\partial f}{\partial x_i}(z_{i-1}) - \frac{\partial f}{\partial x_i}(x)\right\|\right)$$

By definition of partial derivative along i-th axis at z_{i-1} we derive that $\phi_i(h_i) \to 0$ as h_i tends to zero. If $h \to 0$, then by continuity of partial derivatives we have

$$\frac{\partial f}{\partial x_i}(z_{i-1}) - \frac{\partial f}{\partial x_i}(x) \to 0$$

for every i. These results imply that

$$\sum_{i=1}^{n}\left(\left\|\phi_i(h^i)\right\| + \left\|\frac{\partial f}{\partial x_i}(z_{i-1}) - \frac{\partial f}{\partial x_i}(x)\right\|\right) \to 0$$

for $h \to 0$. This completes the proof. □

8. Higher order derivatives

We introduce higher order Fréchet derivatives. We fix Banach spaces $\mathfrak{D}, \mathfrak{X}$ over $\mathbb{K}$. Let U be an open subset of $\mathfrak{D}$ and let V be an open subset of $\mathfrak{X}$.

Definition 8.1. Let $f : U \to V$ be a function. For each natural number m we define *m-th derivative $f^{(m)}$ of f* by recursive formula

$$f^{(0)} = f,\ f^{(m)} = \left(f^{(m-1)}\right)' \text{ for } m > 1$$

If $f^{(m)}$ exists for some natural number m, then f is *m-times differentiable on U*.

Note that the definition above gives m-th derivative as a function defined on the whole domain.

Remark 8.2. Let $f : U \to V$ be a m-times differentiable function on U. Then $f^{(m)}$ can be identified with a function

$$f^{(m)} : U \to L(\underbrace{\mathfrak{D}, \ldots, \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$$

Indeed, the original codomain of $f^{(m)}$ is $\underbrace{L\left(\mathfrak{D}, L\left(\mathfrak{D}, \ldots, L\left(\mathfrak{D}, \mathfrak{X}\right)\ldots\right)\right)}_{m \text{ times } \mathfrak{D} \text{ symbol}}$ and according to Proposition 2.6 we have canonical isometry

$$\underbrace{L\left(\mathfrak{D}, L\left(\mathfrak{D}, \ldots, L\left(\mathfrak{D}, \mathfrak{X}\right)\ldots\right)\right)}_{m \text{ times } \mathfrak{D} \text{ symbol}} = L(\underbrace{\mathfrak{D}, \ldots, \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$$

Thus we can regard $f^{(m)}$ as a function on U taking values in $L(\underbrace{\mathfrak{D}, \ldots, \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$.

Now we introduce the notion of the higher derivative defined locally for a point in the domain.

Definition 8.3. Let $f : U \to V$ be a function on U and let x be a point of U. Let m be a positive integer. Suppose that f is $(m-1)$-times differentiable on some open neighborhood of x in U. Then *m-th derivative of f at x* is the derivative of $f^{(m-1)}$ at x. If it exists, then f is *m-times differentiable at x*.

Remark 8.4. Let x be a point in U and let $f : U \to V$ be a function. Let m be a positive integer. Assume that f is m-times differentiable at x. Then the m-th derivative of f at x is usually denoted by $f^{(m)}(x)$. Similarly to Remark 8.2 we identify $f^{(m)}(x)$ with an element in $L(\underbrace{\mathfrak{D}, \ldots, \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$.

Theorem 8.5. *Let $f : U \to V$ be a function on U and let x be a point of U. Suppose that f is m-times differentiable at x for some integer m greater or equal 2. Then*

$$f^{(m)}(x) \in L(\underbrace{\mathfrak{D}, \ldots, \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$$

is a symmetric $\mathbb{K}$-multilinear form.

Proof for the second derivative. Pick $h, s \in \mathfrak{D}$. Assume that r is a positive real number greater than norms of h and s. Let I be an open interval in $\mathbb{R}$ containing zero and such that

$$I \subseteq \left\{ t \in \mathbb{R} \,\middle|\, \forall_{\xi,\eta \in \mathfrak{D}} \left(\|\xi\| < r \text{ and } \|\eta\| < r \right) \Rightarrow x + t \cdot \xi + t \cdot \eta \in U \right\}$$

Consider the expression

$$F(t) = f(x + t \cdot h + t \cdot s) - f(x + t \cdot s) - f(x + t \cdot h) + f(x) - t^2 \cdot f''(x)(s, h)$$

defined for $t \in I$. For fixed $t \in I$ define a function

$$g(\xi) = f(x + t \cdot \xi + t \cdot s) - f(x + t \cdot \xi) - t^2 \cdot f''(x)(s, \xi)$$

Then g is defined for each $\xi \in \mathfrak{D}$ with $\|\xi\| < r$. It is differentiable function and we have formula

$$g'(\xi) = t \cdot f'(x + t \cdot \xi + t \cdot s) - t \cdot f'(x + t \cdot \xi) - t^2 \cdot f''(x)(s)$$

which follows from Theorem 4.1. Thus by Theorem 5.2

$$\|F(t)\| = \|g(h) - g(0)\| \leq \|h\| \cdot \sup_{\xi \in (0,h)} \|g'(\xi)\| =$$

$$= t \cdot \|h\| \cdot \sup_{\xi \in (0,h)} \|f'(x + t \cdot \xi + t \cdot s) - f'(x + t \cdot \xi) - t \cdot f''(x)(s)\|$$

We write

$$f'(x + t \cdot \xi + t \cdot s) = f'(x) + f''(x)(t \cdot \xi + t \cdot s) + \phi(t \cdot \xi + t \cdot s) \cdot t \cdot \|\xi + s\|$$

and

$$f'(x + t \cdot \xi) = f'(x) + f''(x)(t \cdot \xi) + \phi(t \cdot \xi) \cdot t \cdot \|\xi\|$$

Therefore, we have

$$\|F(t)\| \leq t \cdot \|h\| \cdot \sup_{\xi \in (0,h)} \|f'(x + t \cdot \xi + t \cdot s) - f'(x + t \cdot \xi) - t \cdot f''(x)(s)\| =$$

$$= t \cdot \|h\| \cdot \sup_{\xi \in (0,h)} \left\| \phi(t \cdot \xi + t \cdot s) \cdot t \cdot \|\xi + s\| - \phi(t \cdot \xi) \cdot t \cdot \|\xi\| \right\| =$$

$$= t^2 \cdot \|h\| \cdot \sup_{\xi \in (0,h)} \left\| \phi(t \cdot \xi + t \cdot s) \cdot \|\xi + s\| - \phi(t \cdot \xi) \cdot \|\xi\| \right\|$$

Since f' is differentiable at x, we derive that

$$\lim_{t \to 0} \phi(t \cdot \xi + t \cdot s) = \lim_{t \to 0} \phi(t \cdot \xi) = 0$$

and hence

$$\lim_{t \to 0} \frac{F(t)}{t^2} = 0$$

This implies that

$$\lim_{t \to 0} \frac{f(x + t \cdot h + t \cdot s) - f(x + t \cdot s) - f(x + t \cdot h) + f(x)}{t^2} = f''(x)(s, h)$$

Since the left hand side is symmetric with respect to s and h, we deduce that it also converges to $f''(x)(h, s)$ as $t \to 0$. Thus $f''(x)(s, h) = f''(x)(h, s)$. According to the fact that h and s are arbitrary we infer that $f''(x)$ is a symmetric $\mathbb{K}$-bilinear form. □

Proof of the general case. We proved the theorem for $m = 2$. Suppose that it holds for some $m \geq 2$. We prove it for $m + 1$. For this assume that f is $(m+1)$-times differentiable at x. By shrinking domain of f we may assume that f is m-times differentiable function on U. Pick elements $h_1, h_2, h_3, ..., h_{m+1} \in \mathfrak{D}$ and fix a permutation σ of the set $\{2, 3, ..., m + 1\}$. Consider the composition of $f^{(m)} : U \to L(\underbrace{\mathfrak{D}, ..., \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$ with the map $\mathrm{ev}_{h_2, h_3, ..., h_{m+1}} : L(\underbrace{\mathfrak{D}, ..., \mathfrak{D}}_{m \text{ times}}; \mathfrak{X}) \to \mathfrak{X}$ given by formula

$L \mapsto L(h_2, h_3, ..., h_{m+1})$. According to Theorem 4.1 we derive that the derivative of this composition at x is a $\mathbb{K}$-linear map

$$f^{(m+1)}(x)(-, h_2, h_3, ..., h_{m+1}) : \mathfrak{D} \to \mathfrak{X}$$

Similarly the derivative at x of the composition of $f^{(m)} : U \to L(\underbrace{\mathfrak{D}, ..., \mathfrak{D}}_{m \text{ times}}; \mathfrak{X})$ with the map $\mathrm{ev}_{h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)}} : L(\underbrace{\mathfrak{D}, ..., \mathfrak{D}}_{m \text{ times}}; \mathfrak{X}) \to \mathfrak{X}$ given by formula $L \mapsto L(h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)})$ is a $\mathbb{K}$-linear map

$$f^{(m+1)}(x)(-, h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)}) : \mathfrak{D} \to \mathfrak{X}$$

Since we have $\mathrm{ev}_{h_2, h_3, ..., h_{m+1}} \cdot f^{(m)} = \mathrm{ev}_{h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)}} \cdot f^{(m)}$ (by inductive assumption), we deduce that $f^{(m+1)}(x)(-, h_2, h_3, ..., h_{m+1}) = f^{(m+1)}(x)(-, h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)})$. In particular, we derive that

$$f^{(m+1)}(x)(h_1, h_2, h_3, ..., h_{m+1}) = f^{(m+1)}(x)(h_1, h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)})$$

for every permutation σ of the set $\{2, ..., m, m+1\}$. Next observe that

$$f^{(m+1)}(x) = \left(f^{(m-1)}\right)''(x)$$

and hence

$$f^{(m+1)}(x)(h_1, h_2, h_3, ..., h_{m+1}) = \left(f^{(m-1)}\right)''(x)(h_1, h_2)(h_3, ..., h_{m+1}) =$$
$$= \left(f^{(m-1)}\right)''(x)(h_2, h_1)(h_3, ..., h_{m+1}) = f^{(m+1)}(x)(h_2, h_1, h_3, ..., h_{m+1})$$

by the symmetry of the second derivative. Let us summarize these results in slightly different form. For every elements $h_1, h_2, h_3, ..., h_{m+1} \in \mathfrak{D}$ we have

$$f^{(m+1)}(x)(h_1, h_2, h_3, ..., h_{m+1}) = f^{(m+1)}(x)(h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)})$$

for each permutation σ of $\{1, ..., m+1\}$ such that $\sigma(1) = 1$ and

$$f^{(m+1)}(x)(h_1, h_2, h_3, ..., h_{m+1}) = f^{(m+1)}(x)(h_2, h_1, h_3, ..., h_{m+1})$$

This implies that

$$f^{(m+1)}(x)(h_1, h_2, h_3, ..., h_{m+1}) = f^{(m+1)}(x)(h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}, ..., h_{\sigma(m+1)})$$

for every permutation σ of $\{1, ..., m+1\}$ and every elements $h_1, h_2, h_3, ..., h_{m+1} \in \mathfrak{D}$. This completes the proof that $f^{(m+1)}(x)$ is a symmetric $\mathbb{K}$-multilinear form. $\square$