Chapter 2: Part 1

10:14 PM

$\mathcal{H}$: real Hilbert space with inner product $\langle \cdot | \cdot \rangle$, norm $\|\cdot\|$, distance d, identity

$\forall_{x,y \in \mathcal{H}}\ \|x\| = \sqrt{\langle x|x \rangle}$ $\quad d(x,y) = \|x-y\|$ $\quad 1$ /* in the textbook it is Id */

2.1. Notation and examples.

$\underbrace{C^{\perp}}_{\text{orthogonal complement of subset } C \text{ of } \mathcal{H}} = \{u \in \mathcal{H} \mid \forall_{x \in C}\ \langle x|u \rangle = 0\}$

$\underbrace{C}_{\subseteq \mathcal{H}}$: orthonormal basis of $\mathcal{H}$ $\overset{\text{def}}{\leftrightarrow}$ $\overline{\operatorname{span}}\, C = \mathcal{H}$ /* $\overline{\operatorname{span}}\, C$: smallest closed linear subspace of χ containing C */

$\mathcal{H}$: seperable Hilbert space $\overset{\text{def}}{\leftrightarrow}$ $\mathcal{H}$ has countable orthonormal basis

[

$(x_i)_{i \in I}$: family of vectors in $\mathcal{H}$ /* it is a net */

$\mathfrak{I}$: class of nonempty finite subsets of I, $\subseteq I$

]

$(x_i)_{i \in I}$:summable $\overset{\text{def}}{\leftrightarrow}$ $\exists_{x \in \mathcal{H}}\ \underbrace{\left(\sum_{i \in J} x_i\right)_{J \in \mathfrak{I}} \text{ converges to } x}_{\left(\forall_{\varepsilon \in \mathbb{R}_{++}}\ \exists_{K \in \mathfrak{I}}\ \forall_{J \in \mathfrak{I} : J \supseteq K}\ \|x - \sum_{i \in J} x_i\| \leq \varepsilon\right)}$

2.2. Basic Identities and inequalities:

Fact 2.10. (Cauchy-Schwarz)

[$x, y \in \mathcal{H}$]

- $|\langle x|y \rangle| \leq \|x\|\,\|y\|$
- $\langle x|y \rangle = \|x\|\,\|y\| \leftrightarrow \exists_{\alpha \in \mathbb{R}_+} (x = \alpha y \ \vee\ y = \alpha x)$

Lemma 2.11.

[$x, y, z \in \mathcal{H}$]

(i) $\|x+y\|^2 = \|x\|^2 + 2\langle x|y \rangle + \|z\|^2$

(ii) Parallelogram identity:

$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$

(iii) Polarization identity:

$4\langle x|y \rangle = \|x+y\|^2 - \|x-y\|^2$

(iv)

$\|x-y\|^2 = 2\|z-x\|^2 + 2\|z-y\|^2 - 4\left\|z - \frac{x+y}{2}\right\|^2$

Lemma 2.12. /* Different ways of expressing orthogonal and obtuse vectors */

[$x, y \in \mathcal{H}$]

(i) $\langle x|y \rangle \leq 0 \leftrightarrow \forall_{\alpha \in \mathbb{R}_+} \|x\| \leq \|x - \alpha y\| \leftrightarrow \forall_{\alpha \in [0,1]} \|x\| \leq \|x - \alpha y\|$

(ii) $x \perp y \leftrightarrow \forall_{\alpha \in \mathbb{R}} \|x\| \leq \|x - \alpha y\| \leftrightarrow \forall_{\alpha \in [-1,1]} \|x\| \leq \|x - \alpha y\|$

Lemma 2.13.

[

$(x_i)_{i\in I}, (u_i)_{i\in I}$: finite families in $\mathcal{H}$

$(\alpha_i)_{i\in I}$: family in $\mathbb{R}$, $\sum_{i\in I} \alpha_i = 1$

]

(i)

$$\left\langle \sum_{i\in I} \alpha_i x_i \,\Big|\, \sum_{j\in I} \alpha_j u_j \right\rangle + \sum_{i\in I}\sum_{j\in I} \frac{\alpha_i\alpha_j}{2} \langle x_i - x_j \mid u_i - u_j \rangle = \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle$$

(ii)

$$\Big\| \sum_{i\in I} \alpha_i x_i \Big\|^2 + \sum_{i\in I}\sum_{j\in I} \frac{\alpha_i\alpha_j}{2} \|x_i - x_j\|^2 = \sum_{i\in I} \alpha_i \|x_i\|^2$$

Proof:

(i)

$$2 \left\langle \sum_{i\in I} \alpha_i x_i \,\Big|\, \sum_{j\in I} \alpha_j u_j \right\rangle$$

/* $\langle \sum_i a_i | b \rangle = \langle a_1 + a_2 + \dots + a_{final} | b \rangle = \langle a_1 | b \rangle + \dots + \langle a_{final} | b \rangle$

$= \sum_i \langle a_i | b \rangle$

say, $b = \sum_j c_j$ then $\langle a_i | \sum_j c_j \rangle = \langle \sum_j c_j | a_i \rangle = \sum_j \langle c_j | a_i \rangle$

$\therefore \langle \sum_i a_i | \sum_j c_j \rangle = \sum_i \sum_j \langle a_i | c_j \rangle$ */

$$= 2 \sum_{i\in I}\sum_{j\in I} \langle \alpha_i x_i | \alpha_j u_j \rangle$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i \alpha_j \underbrace{2\langle x_i | u_j \rangle}_{\langle x_i | u_j \rangle + \underbrace{\langle x_i | u_j \rangle}_{\langle u_j | x_i \rangle}}$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i | u_j \rangle + \underbrace{\sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i | u_j \rangle}$$

/* interchanging index i,j */

$$= \sum_{j\in I}\sum_{i\in I} \alpha_j\alpha_i \langle x_j | u_i \rangle = \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_j | u_i \rangle$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \Big(\langle x_i | u_j \rangle + \langle x_j | u_i \rangle \Big)$$

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \Big(\langle x_i | u_i \rangle + \langle x_j | u_j \rangle - \langle x_i - x_j | u_i - u_j \rangle \Big)$$

/* $\langle x_i - x_j | u_i - u_j \rangle = \langle x_i | u_i \rangle - \langle x_i | u_j \rangle - \langle x_j | u_i \rangle + \langle x_j | u_j \rangle$

$\therefore \langle x_i | u_j \rangle + \langle x_j | u_i \rangle = \langle x_i | u_i \rangle + \langle x_j | u_j \rangle - \langle x_i - x_j | u_i - u_j \rangle$ */

$$= \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i | u_i \rangle$$

$$+ \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_j | u_j \rangle$$

$$- \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i - x_j | u_i - u_j \rangle$$

/* $\sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i | u_i \rangle = \sum_{j\in I} \alpha_j \Big(\sum_{i\in I} \alpha_i \langle x_i | u_i \rangle \Big) = \Big(\sum_{i\in I} \alpha_i \langle x_i | u_i \rangle \Big) \underbrace{\Big(\sum_{j\in I} \alpha_j \Big)}_{=1} = \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle$

similarly, $\sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_j | u_j \rangle = \sum_{j\in I} \alpha_j \langle x_j | u_j \rangle$

$$= \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle + \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle - \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i - x_j | u_i - u_j \rangle$$

$$= 2 \sum_{i\in I} \alpha_i \langle x_i | u_i \rangle - \sum_{i\in I}\sum_{j\in I} \alpha_i\alpha_j \langle x_i - x_j | u_i - u_j \rangle \quad \dots (2.15)$$

$$\therefore \sum_{i\in I}\alpha_i\langle x_i|u_i\rangle = \Big\langle \sum_{i\in I}\alpha_i x_i \,\Big|\, \sum_{j\in I}\alpha_j u_j\Big\rangle + \sum_{i\in I}\sum_{j\in I}\frac{\alpha_i\alpha_j}{2}\langle x_i-x_j|u_i-u_j\rangle$$

(ii) In (2.15) we put: $(u_i)_{i\in I} := (x_i)_{i\in I}$ then we have

$$\underbrace{2\Big\langle \sum_{i\in I}\alpha_i x_i \,\Big|\, \underbrace{\sum_{j\in I}\alpha_j \underset{x_j}{u_j}}_{\sum_{j\in I}\alpha_j x_j = \sum_{i\in I}\alpha_i x_i}\Big\rangle}_{\langle(\sum_{i\in I}\alpha_i x_i)|(\sum_{i\in I}\alpha_i x_i)\rangle = \|\sum_{i\in I}\alpha_i x_i\|^2} = 2\sum_{i\in I}\alpha_i\underbrace{\langle x_i|\underset{x_i}{u_i}\rangle}_{\langle x_i|x_i\rangle=\|x_i\|^2} - \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\underbrace{\langle x_i-x_j|\underset{x_i}{u_i}-\underset{x_j}{u_j}\rangle}_{\|x_i-x_j\|^2}$$

$$\Leftrightarrow 2\Big\|\sum_{i\in I}\alpha_i x_i\Big\|^2 = 2\sum_{i\in I}\alpha_i\|x_i\|^2 - \sum_{i\in I}\sum_{j\in I}\alpha_i\alpha_j\|x_i-x_j\|^2$$

$$\Leftrightarrow \Big\|\sum_{i\in I}\alpha_i x_i\Big\|^2 + \sum_{i\in I}\sum_{j\in I}\frac{\alpha_i\alpha_j}{2}\|x_i-x_j\|^2 = \sum_{i\in I}\alpha_i\|x_i\|^2$$ ☺

Corollary 2.14.

[$x,y\in\mathcal{H}$, $\alpha\in\mathbb{R}$]

$$\|\alpha x+(1-\alpha)y\|^2 + \alpha(1-\alpha)\|x-y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2$$

2.3. Linear operators and functionals.

$\mathcal{X},\mathcal{Y}$: normed vector space

$\mathcal{B}(\mathcal{X},\mathcal{Y}) = \{T:\mathcal{X}\to\mathcal{Y} \mid T: \text{linear, continuous}\} \rightarrow$ a normed vector space

/* $\mathcal{B}(\mathcal{X},\mathcal{Y})$: set of all linear and continuous operator */

$\mathcal{B}(\mathcal{X}) = \mathcal{B}(\mathcal{X},\mathcal{X})$

norm of linear and continuous operator:

$\forall_{T\in\mathcal{B}(\mathcal{X},\mathcal{Y})}$ $\|T\| = \sup\|T(B(0;1))\| = \sup_{x\in\mathcal{X},\ \|x\|\le 1}\|Tx\|$ /* if T is a matrix then $\|T\|$ is the usual matrix norm */

$\mathcal{B}(\mathcal{X},\mathcal{Y})$: Banach space if $\mathcal{Y}$ is a Banach space

• Fact 2.15.

[$\mathcal{X},\mathcal{Y}$: real normed vector space

$\underset{:\mathcal{X}\to\mathcal{Y}}{T}$: linear]

T: continuous at a point in $\mathcal{X}$ $\Leftrightarrow$ T: Lipschitz continuous

/* for a linear operator continuity and Lipschitz continuity are same */

• Lemma 2.16. (uniform boundedness) /* called Banach-Steinhaus theorem */

[$\mathcal{X}$: real Banach space, /* Banach space is a $\underbrace{\text{complete}}_{\text{every cauchy sequence converges}}$ normed vector space */

$\mathcal{Y}$: real normed vector space

$(T_i)_{i\in I}$: family of operators in $\mathcal{B}(\mathcal{X},\mathcal{Y})$.

T_i : pointwise bounded $\leftrightarrow \forall_{x\in\mathcal{X}} \sup_{i\in I} \|T_i x\| < \infty$

]

$\Rightarrow$

$\sup_{i\in I} \|T_i\| < +\infty$ /* seems intuitive though */

[Fact 2.17: Riesz-Frechet representation theorem]

Fact 2.17.

return to [Gateaux gradient]

• Riesz-Fréchet representation theorem. /* Says that any continuous linear functional on the real Hilbert space can be identified with a vector in that Hilbert space */

$f\in\mathcal{B}(\mathcal{H},\mathbb{R}) \Rightarrow \left(\exists_{1 u\in\mathcal{H}} \ \forall_{x\in\mathcal{H}} \ f(x) = \langle x|u\rangle\right) \wedge \left(\|f\| = \|u\|\right)$

means unique

[Adjoint operator]

return [Frechet differentiable]

* Adjoint operator.

$\mathcal{K}$: real Hilbert space

$T\in\mathcal{B}(\mathcal{H},\mathcal{K})$

$\underbrace{T^*}_{\in\mathcal{B}(\mathcal{K},\mathcal{H})}$: adjoint of $T \overset{\text{def}}{\leftrightarrow} \forall_{x\in\mathcal{H}} \forall_{y\in\mathcal{K}} \langle Tx|y\rangle = \langle x|T^*y\rangle$

/* adjoint operator is a unique operator */

* Properties of adjoint operator.

[$\mathcal{K}$: real Hilbert space

$T\in\mathcal{B}(\mathcal{H},\mathcal{K})$

$\ker T = \{x\in\mathcal{H} : Tx = 0\}$

] $\Rightarrow$

(i) $T^{**} = T$ /* Double adjoint gives the same operator */

(ii) $\|T^*\| = \|T\| = \sqrt{\|T^*T\|}$ /* linear continuous operator and its adjoint has same norm */

(iii) $(\ker T)^\perp = \overline{\operatorname{ran}} T^*$

(iv) $(\operatorname{ran} T)^\perp = \ker T^*$

(v) $\ker T^*T = \ker T$

$\overline{\operatorname{ran}}\, TT^* = \overline{\operatorname{ran}}\, T$ } /* Finding range closure and kernel of a linear continuous operator using its adjoint */

Fact 2.19.

[$\mathcal{K}$: real Hilbert space,

$T\in\mathcal{B}(\mathcal{H},\mathcal{K})$]

$\operatorname{ran} T$: closed $\Leftrightarrow \operatorname{ran} T^*$: closed $\Leftrightarrow \operatorname{ran} TT^*$: closed

$\Leftrightarrow \operatorname{ran} TT^*$: closed $\Leftrightarrow \exists_{\alpha\in\mathbb{R}_{++}} \forall_{x\in(\ker T)^\perp} \|Tx\| \geq \alpha\|x\|$

• Hyperplane, halfspace etc.

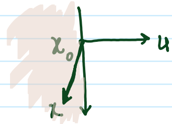
$[u \in \mathcal{H} \setminus \{0\}, \eta \in \mathbb{R}]$

$(\text{closed hyperplane in } \mathcal{H}) = \{x \in \mathcal{H} : \langle x | u \rangle = \eta\}$

$(\text{closed halfspace in } \mathcal{H} \text{ with outer normal } u) = \{x \in \mathcal{H} : \langle x | u \rangle \leq \eta\}$ // in $\mathbb{R}^n$: $u^T x \leq \eta = u^T x_0$

$\to u^T(x - x_0) \leq 0$

$(\text{open halfspace in } \mathcal{H} \text{ with outer normal } u) = \{x \in \mathcal{H} : \langle x | u \rangle < \eta\}$



$d_C : (\text{distance function to } C = \{x \in \mathcal{H} : \langle x | u \rangle = \eta\}), \mathcal{H} \to \mathbb{R}_+,$

$$x \mapsto \underbrace{\frac{|\langle x | u \rangle - \eta|}{\|u\|}}$$

this is the minimum distance of point x from the hyperplane $C = \{x \in \mathcal{H} : \langle x | u \rangle = \eta\}$

Chapter 2: Part 2

10:27 PM

2.4. Strong and weak topology:

$(\mathcal{H}, d)$: strong topology
(or norm topology)

$(x_a)_{a\in A}$ in $\mathcal{H}$ converges strongly to a point x $\overset{def}{\Leftrightarrow}$ $\|x_a - x\| \to 0$
$x_a \to x$

Fact 2.21.
[U: closed linear subspace of $\mathcal{H}$
V: " " " " $\mathcal{H}$, has finite dimension or codimension
]
$U+V$: closed linear subspace.

Weak topology: The family of all finite intersections of open halfspaces of $\mathcal{H}$ forms the base of the weak topology of $\mathcal{H}$. the resulting topolgical space ($\mathcal{H}$, weak topology of $\mathcal{H}$) is called $\mathcal{H}_{weak}$ / we consider only sets in the weak topology in our consideration while dealing with $\mathcal{H}_{weak}$ */

topology: family of subsets of X that contains $X, \emptyset$, arbitrary union and finite intersections of elements of X

/* base. $\underbrace{B}_{\text{subfamily of } T}$: base of T $\overset{def}{\Leftrightarrow}$ $\forall_{x\in X}$ $\forall_{V\in \mathcal{V}(x)}$ $\exists_{S\in B}$ $(x\in S \wedge S\subseteq V)$

importance of base:
B: base of $T \Rightarrow \forall_{S\in T}$ S = union of elements of B */

Lemma 2.23. $\mathcal{H}^{weak}$: Hausdorff space. [H^weak is Hausdorff]

/* $\overset{\text{topology of } X}{T}$: Hausdorff space $\Leftrightarrow \forall_{x,y\in X: x\neq y} \exists_{V\in\mathcal{V}(x), W\in\mathcal{V}(y): V,W \text{ elements of base } T}$ $V\cap W = \emptyset$ */

/* What does this mean?
Recall the definition of Hausdorff space: for any two distinct points $x_1, x_2 \in X$ there exists $V_1 \in \mathcal{V}(x_1), V_2\in\mathcal{V}(x_2)$ such that $V_1 \cap V_2 \neq \emptyset$
now if we confine the neighborhoods are associated with open halfspaces of $\mathcal{H}$ (more specifically union of finite intersections of $\mathcal{H}$, i.e., the weak topology of $\mathcal{H}$), and show that the Hausdorff condition holds, then we have $\mathcal{H}^{weak} \equiv (\mathcal{H}$, weak topology of $\mathcal{H})$ is a Hausdorff space */

Proof.
x, y: distinct points in $\mathcal{H}$, our goal is to show that the elements of $\mathcal{H}_{weak}$ = union of members of its base
construct disjoint neighborhouds for different point.
$u := x - y$
$w = \frac{x+y}{2}$

Consider the set: $\{z\in\mathcal{H} \mid \langle z-w | u\rangle > 0\}$: this is a neighborhood of x, also a member of $\mathcal{H}_{weak}$'s base and hence $\mathcal{H}_{weak}$ because:

$$\langle \underbrace{z}_{:=x} - w | u\rangle = \langle x - \frac{x+y}{2} | x-y\rangle = \langle \frac{x-y}{2} | x-y\rangle = \frac{1}{2}\langle x-y | x-y\rangle = \frac{1}{2}\|x-y\|^2 > 0 \quad (\because x\neq y)$$

$$\therefore x \in \{z\in\mathcal{H} \mid \langle z-w | u\rangle > 0\}$$

Similarly the set: $\{z\in\mathcal{H} \mid \langle z-w | u\rangle < 0\}$ is a neighborhood of y

$$\text{as } \langle \underbrace{z}_{:=y} - w | u\rangle = \langle y - \frac{x+y}{2} | x-y\rangle = \langle \tfrac{1}{2}(y-x) | x-y\rangle = \langle (-\tfrac{1}{2})(x-y) | (x-y)\rangle$$

$$= (-\tfrac{1}{2})\langle x-y | x-y\rangle = -\tfrac{1}{2}\|x-y\|^2 < 0$$

Clearly $\{z\in\mathcal{H} : \langle z-w|u\rangle > 0\}$ and $\{z\in\mathcal{H} : \langle z-w|u\rangle < 0\}$ are disjoint, so we have two disjoint neighborhoods of any two distinct points (definition of a Hausdorff space)

$\therefore \mathcal{H}^{weak}$: Hausdorff space. ∎

*Weakly open. (A subset of $\mathcal{H}$): weakly open $\overset{def}{\Leftrightarrow}$ it is union of finite intersections of open half-spaces

*Weakly convergent net.

$\underbrace{(x_a)_{a\in A}}_{\text{net in } \mathcal{H}}$: converges weakly to x $\overset{def}{\Leftrightarrow}$ $\forall_{u\in\mathcal{H}}$ $\langle x_a | u\rangle \to \langle x_a | x\rangle$ /* x is called the weak limit of $(x_a)_{a\in A}$ */

/* $x_a \rightharpoonup x$ */

*Weakly closed set. Weakly compact set

$\underbrace{C}_{\text{subset of } \mathcal{H}}$: weakly closed $\overset{def}{\Leftrightarrow}$ weak limit of every weakly convergent net in C is also in C.

$\underbrace{C}_{\text{subset of } \mathcal{H}}$: weakly compact $\overset{def}{\Leftrightarrow}$ every net in C has a weak cluster point in C

/* Cluster point: $x \in X$, cluster point of $(x_a)_{a\in A}$ $\overset{def}{\Leftrightarrow}$ $(x_a)_{a\in A}$ lies frequently in every neighborhood of x
$\Leftrightarrow \forall_{V\in\mathcal{V}(x)} \forall_{b\in A} \exists_{a\in A: a\succeq b}$ $x_a \in V$ */

[D: nonempty set of $\mathcal{H}$.

$\mathcal{H}$: real Hilbert space,

$T: D \to \mathcal{H}$

$f: \mathcal{H} \to [-\infty, +\infty]$

]

T: weakly continuous $\overset{\text{def}}{\Longleftrightarrow} \forall_{\underbrace{(x_a)_{a\in A}}_{\text{net in } D} : x_a \rightharpoonup x \wedge x \in D} \quad Tx_a \rightharpoonup Tx$ $\quad$ /* $\underline{\lim}\, \xi_a = \left[\left(a \in A \,//\, (\xi_b : b \succcurlyeq a) \right) // \inf_{b\in A} (\cdot) \right) // \sup_{a\in A} (\cdot) \right]$ */

f: weakly lower semicontinuous at $x \in \mathcal{H} \overset{\text{def}}{\Longleftrightarrow} \forall_{\underbrace{(x_a)_{a\in A}}_{\text{net in } \mathcal{H}} : x_a \rightharpoonup x} \quad \underline{\lim}\, f(x_a) \geq f(x)$

Lemma 2.35. [Norm in H: weakly lower semicontinuous]

Norm of $\mathcal{H}$: weakly lower semicontinuous, i.e.,

$$\forall_{x\in\mathcal{H}} \ \forall_{\underbrace{(x_a)_{a\in A}}_{\text{net in } \mathcal{H}} : x_a \rightharpoonup x} \quad \underline{\lim}\, \|x_a\| \geq \|x\|$$

Proof:

take $x \in \mathcal{H}$, $(x_a)_{a\in A}: x_a \rightharpoonup x \overset{\text{def}}{\Longleftrightarrow} \forall_{u\in\mathcal{H}} \quad \langle x_a | u\rangle \to \langle x | u \rangle \Leftrightarrow \lim \langle x_a | u \rangle = \langle x | u\rangle \Rightarrow \lim |\langle x_a | u\rangle| = |\langle x | u \rangle|$

$\overset{u:=x}{\Rightarrow} \lim |\langle x_a | x \rangle| = \|x\|^2$

now $\|x\|^2 = \lim |\langle x_a | x\rangle| = \underline{\lim}\, \underbrace{|\langle x_a | x \rangle|}$ /*recall that unlike $\lim \|x_a\|$ which has to be in finite norm, i.e. in $\mathbb{R}$, $\underline{\lim}\, \|x_a\| \in [-\infty, +\infty]$ and always exists in that sense */

$\leq \|x_a\| \, \|x\|$ // using Cauchy-Schwarz

$\leq \underline{\lim}\, \|x_a\| \, \|x\| = \|x\| \, \underline{\lim}\, \|x_a\|$

when $x = 0$, $\|x\| \leq \underline{\lim}\, \|x_a\|$ trivially holds

when $x \neq \emptyset$, $\|x\|^2 \leq \|x\| \, \underline{\lim}\, \|x_a\|$

$\Leftrightarrow \|x\| \leq \underline{\lim}\, \|x_a\|$

$\therefore$ Norm of $\mathcal{H}$: weakly lower semicontinuous. ▮

Lemma 2.36. /* An important result */

$\left[(x_a)_{a\in A} \right.$: net in $\mathcal{H}$, bounded, $x_a \rightharpoonup \underbrace{x}_{\in\mathcal{H}}$

$(u_a)_{a\in A}$: net in $\mathcal{H}$, $u_a \to \underbrace{u}_{\in\mathcal{H}}$]

$\Rightarrow$

$\langle x_a | u_a \rangle \to \langle x | u \rangle$

Proof:

$(x_a)_{a\in A}$: net in $\mathcal{H}$, bounded, $x_a \rightharpoonup \underbrace{x}_{\in\mathcal{H}}$

$\Rightarrow \sup_{a\in A} \|x_a\| < +\infty \ldots$ (i)

$u_a \to u \overset{\text{def}}{\Longleftrightarrow} \|u_a - u\| \to 0 \ldots$ (ii)

$x_a \rightharpoonup x \overset{\text{def}}{\Longleftrightarrow} \forall_{y\in\mathcal{H}} \quad \langle x_a | y \rangle \to \langle x | y \rangle \Rightarrow \langle x_a - x | y \rangle \to 0$

set $y := u \Rightarrow \langle x_a - x | u \rangle \to 0 \ldots$ (iii)

$\forall_{a\in A} \quad 0 \leq |\langle x_a | u_a \rangle - \langle x | u \rangle|$

$= |\langle x_a | u_a \rangle - \langle x_a | u \rangle + \langle x_a | u \rangle - \langle x | u \rangle|$ /* this is the famous "one change at a time" trick in real analysis */

$= |\langle x_a | u_a - u \rangle + (-\langle x_a - x | u \rangle)|$

$\leq |\langle x_a | u_a - u \rangle| + |\langle x_a - x | u \rangle|$

$\leq \|x_a\| \, \|u_a - u\| + |\langle x_a - x | u \rangle|$

$$\leq \underbrace{\underbrace{\underbrace{\left(\sup_{a\in A} \|x_a\|\right)}_{<+\infty} \underbrace{\|u_a - u\|}_{\to 0 \ /\text{* by (ii) *}/}}_{\to 0} + \underbrace{|\langle x_a - x | u \rangle|}_{\to 0}}_{\to 0}$$

/* now use

Result 1.2.5. (Another important result, to establish convergence of a net indirectly)

$\left[(x_a)_{a\in A}, (y_a)_{a\in A} \right] \ (x \leq x_a \leq y_a, \ y_a \to x) \Rightarrow x_a \to x$

*/

$\Rightarrow \langle x_a | u_a \rangle - \langle x | u \rangle \to 0$

▮

2.5. Weak convergence of sequences.

Lemma 2.37.

$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$

$\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: posses a weakly convergent subsequence.

Proof:

$\mathcal{H}^{weak}$: Hausdorff space

set $\rho = \sup_{n\in\mathbb{N}} \|x_n\| \in (-\infty, \infty)$

$C = B(0;\rho)$: closed ball with 0 center and ρ radius. i.e., $B(0;\rho) = \{x \in \mathcal{H} \mid \|x\| < +\infty\}$

Fact 2.27. (Banach-Alaoglu): The closed unit ball $B(0;1)$ of $\mathcal{H}$ is weakly compact

Fact 2.30. (Eberlein-Šmulian): $[C \subseteq \mathcal{H}]$ (C: weakly compact $\Leftrightarrow$ C: weakly sequentially compact)

every net has a weak cluster point, i.e., every net has a subnet that weakly converges to a point in $\mathcal{H}$.

every sequence in C has a weak sequential cluster point, i.e., every sequence has a subsequence that weakly converges to a point in C.

$C = B(0;\rho)$: weakly sequentially compact

now $(x_n)_{n\in\mathbb{N}}$ lies in C by construction

/* Definition 1.32.

(C: subset of Hausdorff space X): $\overset{weakly}{\text{sequentially}}$ compact

$\overset{def}{\longleftrightarrow}$

every sequence in C has a subsequence which converges $\overset{weakly}{\text{to}}$ a point in C

*/

as C is in $\underbrace{\mathcal{H}}_{\text{Hausdorff space}}$ and is weakly sequentially compact $\Rightarrow (x_n)_{n\in\mathbb{N}}$ has a weakly convergent subsequence.

■

Lemma 2.38.

$[(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}]$

$(x_n)_{n\in\mathbb{N}}$: converges weakly $\Leftrightarrow$
- $(x_n)_{n\in\mathbb{N}}$: bounded
- $(x_n)_{n\in\mathbb{N}}$: posseses atmost one weak sequential cluster point.

Proof: $(\Rightarrow)$

take $\exists_{x\in\mathcal{H}}\ x_n \rightharpoonup x$ /* now,

*Lemma 2.23.

$\mathcal{H}^{weak}$: Hausdorff space.

• Fact 1.9. * *

[$(x_a)_{a\in A}$: net in Hausdorff space X, $x_a \overset{\text{def}}{\to} \bar{x}$

$(x_{k(b)})_{b\in B}$: subnet of $(x_a)_{a\in A}$] $\Rightarrow x_{k(b)} \to \bar{x}$ /* If a net in a Hausdorff space converges, then so does any of its subnet and to the same point. Hilbert space is Hausdorff. */ */

$\therefore$ x: unique sequential cluster point of $(x_n)_{n\in\mathbb{N}}$

now $x_n \rightharpoonup x \overset{def}{\Leftrightarrow}$ ...(1) $\forall_{u\in\mathcal{H}}\ \langle x_n | u\rangle \to \langle x | u\rangle$

$\Rightarrow (\langle x_n|u\rangle)_{n\in\mathbb{N}}$: bounded $\forall_{u\in\mathcal{H}}$

$\therefore \forall_{u\in\mathcal{H}} \sup_{n\in\mathbb{N}} \|\langle x_n|u\rangle\| < +\infty$

/*

Lemma 2.16. (Banach-Steinhaus uniform boundedness theorem)

[$\mathcal{X}$: real Banach space; $\mathcal{Y}$: real normed vector space; $(T_i)_{i\in I}$: family of operators in $\mathcal{B}(\mathcal{X},\mathcal{Y})$. pointwise bounded $\overset{def}{\Leftrightarrow} \forall_{\tilde{x}\in\mathcal{X}} \sup_{i\in I} \|T_i\tilde{x}\| < +\infty$] $\Rightarrow \sup_{i\in I} \|T_i\| < +\infty$

set. $I = \mathbb{N}$, $T_i = \langle x_i | \cdot\rangle : \mathcal{H} \to \mathbb{R}$ $\in \mathcal{B}(\mathcal{H},\mathbb{R})$; $\underbrace{\mathcal{X} := \mathcal{H}, \mathcal{Y} := \mathbb{R}}_{\text{real normed space}}$; $\therefore \forall_{\tilde{x}\in\mathcal{H}} \sup_{i\in I} \|\langle x_i|\tilde{x}\rangle\| = \|T_i u\| < +\infty$

$\sup_{n\in\mathbb{N}} \|T_n\| = \|\langle x_n|\cdot\rangle\| = \|x_n\| < +\infty$

$\leftrightarrow (x_n)_{n\in\mathbb{N}}$: bounded ...(2)

From (1), (2) we have proven $\Rightarrow$:)

$(\Leftarrow)$ given: $(x_n)_{n\in\mathbb{N}}$: bounded. posseses atmost one sequential cluster point. Say x

*Lemma 2.37. *

[$(x_n)_{n\in\mathbb{N}}$: bounded sequence in $\mathcal{H}$] $\Rightarrow \exists_{(x_{k_n})_{n\in\mathbb{N}}:\ \text{subsequence of } (x_n)_{n\in\mathbb{N}}}\ (x_{k_n})_{n\in\mathbb{N}}$: weakly convergent

$\Leftrightarrow (x_n)_{n\in\mathbb{N}}$: posseses a weak sequential cluster point.

$\therefore (x_n)_{n\in\mathbb{N}}$: posses exactly one sequential cluster point (3)

now $(x_n)_{n\in\mathbb{N}}$: bounded $\Rightarrow \sup_{n\in\mathbb{N}} \|x_n\| = \rho < +\infty$, now construct closed ball $B(0;\rho)$ and $(x_n)_{n\in\mathbb{N}}$ lies in $B(0;\rho) = C$

$\therefore (x_n)_{n\in\mathbb{N}} \subseteq B(0;\rho) = C$

Fact 2.27. (Banach-Alaoglu) * *

The closed unit ball $B(0;1)$ of $\mathcal{H}$ is weakly compact.

now $B(0;\rho)$: weakly compact

$\Updownarrow$

$C = B(0;\rho)$: weakly sequentially compact (4)

Fact 2.30 (Eberlein-Šmulian)

[C: subset of $\mathcal{H}$]

C: weakly compact $\Leftrightarrow$ C: weakly sequentially compact

now

* *

Lemma 1.34. [C: $\overset{weakly}{\text{sequentially}}$ compact, // from (4) $\subseteq \underbrace{\cancel{X}}_{\mathcal{H}^{weak}:\ \text{Hausdorff space}}$ (Hausdorff space); $(x_n)_{n\in\mathbb{N}} \subseteq C$, admits a unique $\overset{weak}{\text{sequential}}$ cluster // from (4) point x] $x_n \rightharpoonup x$

$\therefore x_n \rightharpoonup x$

■

Lemma 2.39.

$\Big[$ $(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$

C : nonempty, $\subseteq \mathcal{H}$

$\forall_{x\in C}$ $(\|x_n - x\|)_{n\in\mathbb{N}}$: converges

$\forall_{y:\text{ sequential cluster point of } (x_n)_{n\in\mathbb{N}}}$ $y \in C$

$\Big]$

$\Rightarrow$

$(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in C.

Proof: /* proof strategy: want to show $P \to (Q \wedge R)$: show $(P\to Q)$ and $(P \to R)$ seperately */

given that $\forall_{x\in C}$ $\|x_n - x\|_{n\in\mathbb{N}}$: converges $\Rightarrow$ $\boxed{(x_n)_{n\in\mathbb{N}} : \text{bounded}}$ $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: will posses atleast one weak sequential cluster point // using Lemma 2.37

so we have to show that given $\Rightarrow$ $(x_n)_{n\in\mathbb{N}}$: almost one weak sequential cluster point in C

/* By Lemma 2.38 */

$\Updownarrow$

all the subsequences converges weakly almost one point in C

$\Updownarrow$

$(x_n)_{n\in\mathbb{N}}$ cannot have two distinct weak sequential cluster points in C

x, y : weak sequential cluster points of $(x_n)_{n\in\mathbb{N}}$ in C

$\Updownarrow$

$$x_{k_n} \rightharpoonup \underbrace{x}_{\in C}, \quad x_{l_n} \rightharpoonup \underbrace{x}_{\in C}$$

$x, y \in C \Rightarrow$ /* by given */

$(\|x_n - x\|)_{n\in\mathbb{N}}$: converges

$(\|x_n - y\|)_{n\in\mathbb{N}}$: converges

/* possibly to different numbers */

again,

$$\forall_{n\in\mathbb{N}} \quad \underbrace{\|x_n - y\|^2}_{\|x_n\|^2 + \|y\|^2 - 2\langle x_n | y\rangle} - \underbrace{\|x_n - x\|^2}_{\|x_n\|^2 + \|x\|^2 - 2\langle x_n | x\rangle} + \|x\|^2 - \|y\|^2$$

$$= \cancel{\|x_n\|^2} + \cancel{\|y\|^2} - 2\langle x_n | y\rangle - \cancel{\|x_n\|^2} - \cancel{\|x\|^2} + 2\langle x_n | x\rangle + \cancel{\|x\|^2} - \cancel{\|y\|^2}$$

$$= 2\langle x_n | x - y\rangle$$

$$\therefore \underbrace{\lim_{n\to\infty} \|x_n - y\|^2}_{\text{by given, converges to some number say } v_1} - \underbrace{\lim_{n\to\infty} \|x_n - x\|^2}_{\text{by given, converges to some number in } C\text{, say } v_2} + \underbrace{\lim_{n\to\infty} (\|x\|^2 - \|y\|^2)}_{\|x\|^2 - \|y\|^2 \text{ /* fixed term */}} = \lim_{n\to\infty} 2\langle x_n | x - y\rangle$$

$$\leftrightarrow \underbrace{\tfrac{1}{2}(v_1 - v_2 + \|x\|^2 - \|y\|^2)}_{\text{say } l} = \lim_{n\to\infty} \langle x_n | x - y\rangle$$

$$\leftrightarrow \lim_{n\to\infty} \langle x_n | x - y\rangle = l$$

$$\leftrightarrow \langle x_n | x - y\rangle \to l$$

* recall: $(y_b)_{b\in B}$: subnet of net $(x_a)_{a\in A}$ $\overset{\text{def}}{\leftrightarrow}$ $\exists_{k: B\to A}$ $\begin{cases} \bullet \forall_{b\in B}\ y_b = x_{k(b)} \\ \bullet \forall_{a\in A} \exists_{a\in B} \forall_{b\in B: b\succcurlyeq d}\ k(b) \succcurlyeq a \end{cases}$

as the sequence $(x_n)_{n\in\mathbb{N}}$ satisfies $\langle x_n | x - y\rangle \to l$, any two subsequences ($(x_{k_n})_{n\in\mathbb{N}}$, $(x_{l_n})_{n\in\mathbb{N}}$) of it will also satisfy : $\langle x_{k_n} | x - y\rangle \to l$, $\langle x_{l_n} | x - y\rangle \to l$. /* Fact 1.9. $(x_a)_{a\in A}$: net in a Hausdorff space, converges to $x \in X$ $\Rightarrow$ $\underbrace{x_{k(b)}}_{(x_{k(b)})_{b\in B} : \text{ subnet of } (x_a)_{a\in A}} \to x$; we apply this for sequence $(\langle x_n | x - y\rangle)_{n\in\mathbb{N}}$ */

using

but $\begin{cases} x_{k_n} \rightharpoonup x \leftrightarrow \forall_{u\in\mathcal{H}} \langle x_{k_n} | u\rangle \to \langle x | u\rangle \ : \text{set } u := x - y \ \therefore \langle x_{k_n} | x - y\rangle \to \langle x | x - y\rangle \\ u \to u \\ x_{l_n} \rightharpoonup y \leftrightarrow \forall_{u\in\mathcal{H}} \langle x_{l_n} | u\rangle \to \langle x | u\rangle \ : \text{set } u := x - y \ \therefore \langle x_{l_n} | x - y\rangle \to \langle y | x - y\rangle \end{cases}$

/* if we have one weakly convergent bounded net, and one strongly convergent net with same index set, then the "inner product" net strongly converges */

** Lemma 2.36.

$\Big[$ $(x_a)_{a\in A}$: net in $\mathcal{H}$, bounded, $x_a \rightharpoonup \underbrace{x}_{\in\mathcal{H}}$

$(u_a)_{a\in A}$: net in $\mathcal{H}$, $u_a \to \underbrace{u}_{\in\mathcal{H}}$ $\Big] \Rightarrow \langle x_a | u_a\rangle \to \langle x | u\rangle$

$\langle x | x - y\rangle = l$

$\langle y | x - y\rangle = l$

subtracting both: $l - l = \langle x | x - y\rangle - \langle y | x - y\rangle = \langle x - y | x - y\rangle = \|x - y\|^2 \leftrightarrow \|x - y\|^2 = 0 \leftrightarrow x = y.$ ∎

$\boxed{(x_n)_{n\in\mathbb{N}} : \text{has almost 1 sequential cluster point}}$

so we have shown : $(x_n)_{n\in\mathbb{N}}$: bounded, and has almust 1 sequential cluster point $\Leftrightarrow$ $(x_n)_{n\in\mathbb{N}}$: converges weakly

/* By Lemma 2.38 */ ∎

Proposition 2.40.

$\Big[$ $(e_i)_{i\in I}$: totally ordered family in $\mathcal{H}$, $\overline{\text{span}}\{e_i\}_{i\in I} = \mathcal{H}$ /* $\overline{\text{span}}\, C$: smallest linear subspace of $\mathcal{X}$ containing C */

/* By Lemma 2.38 */ ∎

Proposition 2.40.

[$(e_i)_{i\in I}$: totally ordered family in $\mathcal{H}$, $\overline{\text{span}}\{e_i\}_{i\in I}=\mathcal{H}$ /* $\overline{\text{span}}\,\underbrace{C}_{C\subseteq X}$: smallest linear subspace of X containing C */

$(x_n)_{n\in\mathbb{N}}$: a sequence in $\mathcal{H}$] $\Rightarrow$

(i) $x_n \rightharpoonup x$

(ii) $(x_n)_{n\in\mathbb{N}}$: bounded $\wedge$

$\forall_{i\in I}$ $\langle x_n | e_i\rangle \to \langle x | e_i\rangle$ as $n\to\infty$

• Lemma 2.41.

[$(x_n)_{n\in\mathbb{N}}, (u_n)_{n\in\mathbb{N}}$: sequences in $\mathcal{H}$

$x, u\in\mathcal{H}$] $\Rightarrow$

(i) $\left(x_n \rightharpoonup x \wedge \overline{\lim}\, \|x_n\| \leq \|x\|\right) \Leftrightarrow x_n \to x$

(ii) $\mathcal{H}$: finite dimensional $\Rightarrow$ $\left(x_n \rightharpoonup x \Leftrightarrow x_n \to x\right)$

(iii) $x_n \rightharpoonup x$, $u_n \to u$ $\Rightarrow$ $\langle x_n | u_n\rangle \to \langle x|u\rangle$

Corollary 2.42. /* Charecterization of strong convergence */

[$(x_n)_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$

$x\in\mathcal{H}$]

$\Rightarrow$

$(x_n \to x) \Leftrightarrow \left(x_n \rightharpoonup x \wedge \|x_n\| \to \|x\|\right)$

Chapter 2: Part 3

5:45 PM

2.6. Differentiability.

$\mathcal{K}$: real Banach space.

A Banach space is a vector space X over the field $\mathbf{R}$ of real numbers, or over the field $\mathbf{C}$ of complex numbers, which is equipped with a norm and which is complete with respect to that norm, that is to say, for every Cauchy sequence $\{x_n\}$ in X, there exists an element x in X such that

$$\lim_{n\to\infty} x_n = x,$$

or equivalently:

$$\lim_{n\to\infty} \|x_n - x\|_X = 0.$$

$(x_n)_{n\in\mathbb{N}}$: Cauchy sequence $\overset{\text{def}}{\leftrightarrow}$ $d(x_m, x_n) \to 0$ as $m, n \to \infty$

The vector space structure allows one to relate the behavior of Cauchy sequences to that of converging series of vectors. A normed space X is a Banach space if and only if each absolutely convergent series in X converges,[2]

$$\sum_{n=1}^{\infty} \|v_n\|_X < \infty \quad \text{implies that} \quad \sum_{n=1}^{\infty} v_n \quad \text{converges in} \quad X.$$

Completeness of a normed space is preserved if the given norm is replaced by an equivalent one.

All norms on a finite-dimensional vector space are equivalent. Every finite-dimensional normed space over $\mathbf{R}$ or $\mathbf{C}$ is a Banach space.[3]

https://en.wikipedia.org/wiki/Banach_space#Definition

Definition 2.43. (Gâteaux differentiability)

[C: subset of $\mathcal{H}$

$T: C \to \mathcal{K}$

$x \in C$: $\forall_{y\in\mathcal{H}}\ \exists_{\alpha\in\mathbb{R}_{++}}$ $[x, x+\alpha y] \subseteq C$ /* this condition just keeps the points involved in the definition of derivative in C */

]

unique when exists

T: Gâteaux differentiable at x $\overset{\text{def}}{\leftrightarrow}$ $\exists_{\underbrace{DT(x)\in\mathcal{B}(\mathcal{H},\mathcal{K})}_{\text{i.e. } DT(x) \text{ linear and continuous}}}$ $\forall_{y\in\mathcal{H}}$ $DT(x)\,y = \lim_{\alpha\downarrow 0} \dfrac{T(x+\alpha y) - T(x)}{\alpha}$

• Defining Higher order Gâteaux derivative :

e.g. for second order:

$D^2T(x) \in \mathcal{B}(\mathcal{H}, \mathcal{B}(\mathcal{H},\mathcal{K}))$: $\forall_{y\in\mathcal{H}}$ $D^2T(x)\,y = \lim_{\alpha\downarrow 0} \dfrac{DT(x+y) - DT(x)}{\alpha}$

[Gateaux gradient]

• Notion of Gâteaux gradient:

[C: subset of $\mathcal{H}$,

$f: C \to \mathbb{R}$, Gâteaux differentiable at $x \in C$

]

/* recall Fact 2.17. [Fact 2.17: Riesz-Frechet] */

(using Fact 2.17)

$\Rightarrow$

$\exists_{\nabla f(x):\ \text{unique}, \in\mathcal{H}}$ $\forall_{y\in\mathcal{H}}$ $Df(x)\,y = \langle y \mid \nabla f(x)\rangle$

$\exists_{\nabla f(x):\ \text{unique},\ \in \mathcal{H}}\ \forall_{y\in\mathcal{H}}\ Df(x)y = \langle y \mid \nabla f(x)\rangle$

$\nabla f(x)$: called Gâteaux gradient

• Gâteaux Hessian.

f: twice Gâteaux differentiable at $x \in C$

$\Rightarrow \exists_{\nabla^2 f(x) \in \mathcal{B}(\mathcal{H}):\ \text{unique}}\ \forall_{y\in\mathcal{H}}\ \forall_{z\in\mathcal{H}}\ (D^2 f(x) y) z = \langle z \mid \underbrace{\nabla^2 f(x)}_{\text{Gâteaux Hessian of } f \text{ at } x} y\rangle$

[Frechet differentiable]

* Fréchet differentiable:

[$x \in \mathcal{H}$,
$C \in \mathcal{V}(x)$
$T: C \to \mathcal{K}$]

T: Fréchet differentiable at $x \overset{\text{def}}{\leftrightarrow}$

$$\exists_{\underbrace{DT(x) \in \mathcal{B}(\mathcal{H},\mathcal{K})}_{\text{linear and continuous}}}\ \lim_{0 \neq \|y\| \to 0} \frac{\|T(x+y) - Tx - DT(x)y\|}{\|y\|} = 0 \qquad \{2.36\}$$

Example 2.46.

[$L \in \mathcal{B}(\mathcal{H})$
$u \in \mathcal{H}$
$x \in \mathcal{H}$
$f: \mathcal{H} \to \mathbb{R},\ f(y) = \langle Ly \mid y\rangle - \langle y \mid u\rangle$]

$\Rightarrow$ f: twice Fréchet differentiable on $\mathcal{H}$, $\nabla f(x) = (L + L^*)x - u$, $\nabla^2 f(x) = L + L^*$

/* L^*: adjoint operator */ [Adjoint operator]

Proof:

$\forall_{y \in \mathcal{H}}$

$L \in \mathcal{B}(\mathcal{H}) = \mathcal{B}(\mathcal{H},\mathcal{H})$

$f(x+y) - f(x) = \langle L(x+y) \mid (x+y)\rangle - \langle x+y \mid u\rangle - \langle Lx \mid x\rangle + \langle x \mid u\rangle$

$= \langle Lx + Ly \mid x+y\rangle - \langle x+y \mid u\rangle - \langle Lx \mid x\rangle + \langle x \mid u\rangle$

$= \langle Lx \mid x+y\rangle + \langle Ly \mid x+y\rangle - \langle x \mid u\rangle - \langle y \mid u\rangle - \langle Lx \mid x\rangle + \langle x \mid u\rangle$

$= \cancel{\langle Lx \mid x\rangle} + \langle Lx \mid y\rangle + \langle Ly \mid x\rangle + \langle Ly \mid y\rangle - \cancel{\langle x \mid u\rangle} - \langle y \mid u\rangle - \cancel{\langle Lx \mid x\rangle} + \cancel{\langle x \mid u\rangle}$

$= \underbrace{\langle Lx \mid y\rangle}_{\langle y \mid Lx\rangle} + \underbrace{\langle Ly \mid x\rangle}_{= \langle y \mid L^* x\rangle} + \langle Ly \mid y\rangle - \langle y \mid u\rangle$

$= \langle y \mid \underbrace{Lx + L^* x}_{(L+L^*)(x)}\rangle + \langle Ly \mid y\rangle - \langle y \mid u\rangle$

/* note that $(A+B)x = Ax + Bx$ is valid but $A(x+y)$ is not necessarilly $Ax + Ay$! */

$$= \langle y \mid (L+L^*)(x) \rangle + \langle Ly \mid y \rangle - \langle y \mid u \rangle = \langle y \mid (L+L^*)x - u \rangle + \langle Ly \mid y \rangle$$

$$\therefore \quad f(x+y) - f(x) - \langle y \mid (L+L^*)x - u \rangle = \langle Ly \mid y \rangle$$

$$\to \left| f(x+y) - f(x) - \underbrace{\langle y \mid (L+L^*)x - u \rangle}_{DT(x)\,y} \right| = \left| \langle Ly \mid y \rangle \right| \leq \underbrace{\|Ly\|}\, \|y\| \quad \text{/* Cauchy-Schwartz */}$$

$$\leq \|L\|\,\|y\|\,\|y\| \quad \text{/* By definition, } \|L\| = \sup_{y \neq 0} \frac{\|Ly\|}{\|y\|} \to \|Ly\| \leq \|L\|\,\|y\| \text{ */}$$

$$\leq \|L\|\,\|y\|^2$$

$$\to \frac{\| f(x+y) - f(x) - \langle (L+L^*)x - u \mid y \rangle \|}{\|y\|} \leq \|L\|\,\|y\|$$

$$\to \lim_{0 \neq \|y\| \to 0} \frac{\| f(x+y) - f(x) - \langle (L+L^*)x - u \mid y \rangle \|}{\|y\|} \leq \lim_{0 \neq \|y\| \to 0} \|L\|\,\|y\| = 0$$

$\therefore$ By definition, {2-36} $DT(x)y = \langle (L+L^*)x - u \mid y \rangle$ ■

Lemma 2-49.

[$x \in \mathcal{H}$,
$C \in \mathcal{V}(x)$
$T: C \to \mathcal{K}$, Frechet differentiable at x]

(i) T: Gâteaux differentiable and the two derivatives coincide

(ii) T: continuous at x

Fact 2-50.

[$T: \mathcal{H} \to \mathcal{K}$,
$x \in \mathcal{H}$
$\exists$ Gateaux derivative T in neighborhood of x
DT: continuous at x]

$\Rightarrow$

T: Fréchet differentiable at x

Fact 2-51. /* Chain rule */

[
$x \in \mathcal{H}$,
U: neighborhood of x
$\mathcal{G}$: real Banach space
$T: U \to \mathcal{G}$, Fréchet differentiable
V: neighborhood of Tx
$R: V \to \mathcal{K}$, Gâteaux differentiable at Tx
]

- $R \circ T$: Gâteaux differentiable at x /* So the differentiability type of the composition is same as the outer operator */
- $D(R \circ T)(x) = (DR(Tx)) \circ (DT(x))$

$$= \left(DR(y)\big|_{y=Tx} \circ DT(x) \right)$$

- R: Fréchet differentiable at $x \Rightarrow R \circ T$: Fréchet differentiable at x

$= \left(DR(y)\big|_{y=Tx} \circ DT(x) \right)$

• R : Fréchet differentiable at $x \Rightarrow R \circ T$: Fréchet differentiable at x

Fact 2.53.

[$x \in \mathcal{H}$

U : neighborhood of x

K : real Banach space

$T : U \to K$, twice Fréchet differentiable at x]

$\Rightarrow$

$$\forall_{y,z \in \mathcal{H}\times\mathcal{H}} \quad \underbrace{\left(D^2T(x)y\right)z}_{\langle z \,|\, \nabla^2 T(x) y\rangle} = \underbrace{\left(D^2T(x)z\right)y}_{\langle y \,|\, \nabla^2 T(x) z\rangle}$$

/*

This is a generalization of the symmetric nature of Hessian

$\because \quad z^T \nabla^2 T(x)\, y = z^T \nabla^2 T(x)^T y = (\nabla^2 T(x) z)^T y = y^T \nabla^2 T(x) z$ */