CALCUL DIFFÉRENTIEL
Cours L3-S6

Université de Frznche-Comté - UFR ST
Année 2010-11
C. Dupaix

4 mai 2011

Table des matières.

Chapitre 1

Introduction - Rappels

1.1 Espaces vectoriels normés (e.v.n.).

Tous les espaces vectoriels considérés dans ce cours sont construits sur le corps $\mathbb{R}$ des réels (typiquement $\mathbb{R}^N$ en dimension fine).
Si une bonne partie des résultats du cours sont énoncés en dimension infinie, l'essentiel des TDs et la totalité de l'examen porteront sur la dimension finie.
Rappelons les définitions suivantes (dans tout ce qui suit E désigne un $\mathbb{R}$ espace vectoriel) :

• **Norme sur** E.
On dit qu'une application $\| \ \|_E$ est une *norme sur* E si elle vérifie les quatre conditions :

1. $\| \ \|_E : E \longrightarrow \mathbb{R}^+$
2. $\|x\|_E = 0 \Longleftrightarrow x = 0$
3. $\forall (x,y) \in E \times E,\ \|x+y\|_E \leqslant \|x\|_E + \|y\|_E$ (inégalité triangulaire)
4. $\forall x \in E,\ \forall t \in \mathbb{R},\ \|tx\|_E = |t|\,\|x\|_E$ (homogénéité).

Rappelons que dans ce cas (seconde inégalité triangulaire) : $|\,\|x\|_E - \|y\|_E\,| \leqslant \|x-y\|_E$.

• **Espace vectoriel normé**.
Un couple $(E, \| \ \|_E)$ est appelé *espace vectoriel normé* et sera noté en abrégé **e.v.n.**
Si aucune ambiguïté n'est à craindre on notera $\| \ \|$ au lieu de $\| \ \|_E$ et E au lieu de $(E, \| \ \|_E)$.
Notez que sur un même espace vectoriel, il y a plusieurs normes possibles.

Exemple 1.1
Soit $B = (e_1, ..., e_N)$ une base de $E = \mathbb{R}^N$. Pour $x \in \mathbb{R}^N$, on note (x_i) les coordonnées de x dans la base B. Alors les applications suivantes définissent des normes (les cas $p = 1, 2$ sont les plus utilisés) sur $\mathbb{R}^N$:

$$\| \ \|_p : (x_1, ..., x_N) \mapsto \left(\sum_{k=1}^{N} |x_k|^p \right)^{1/p},\ \textit{pour tout } p \in \mathbb{R},\ p \geqslant 1 \textit{ fixé.}$$
$$\| \ \|_\infty : (x_1, ..., x_N) \mapsto Max(|x_1|, ..., |x_N|).$$

Pour $p = 2$ (norme euclidienne), la norme est associée à un produit scalaire (euclidien) noté $< \, , \, >_2$.
Ainsi $\forall x = (x_i)_{i=1}^N,\ y = (y_i)_{i=1}^N \in \mathbb{R}^N$ on a :

$$< x, y >_2 = \sum_{k=1}^{N} x_k y_k \ \textit{et} \ \|x\|_2 = \sqrt{< x, x >_2}.$$

Exemple 1.2
Soient $(E_k, \| \|_{E_k})_{1\leqslant k\leqslant n}$ n e.v.n. Alors pour $x = (x_1, ..., x_n) \in E_1 \times ... \times E_n$ les applications

$$\|x\|_p = \left(\sum_{k=1}^{n} \|x_k\|_{E_k}^p\right)^{1/p}, \text{ pour tout } p \in \mathbb{R},\ p \geqslant 1 \text{ fixé, et } \|x\|_\infty = Max(\|x_1\|_{E_1}, ..., \|x_n\|_{E_n})$$

définissent des normes sur $E_1 \times ... \times E_n$ appelées **normes produits***.*

->EXERCICE 1 : montrer que pour $p = 1, 2, \infty$, ces applications sont bien des normes.

• **Boule ouverte de** $E = (E, \| \|_E)$.
Pour $x_0 \in E$ et $r > 0$, on appelle *boule ouverte* de centre x_0 et de rayon r pour la norme $\| \|_E$ l'ensemble défini par $\{x \in E,\ \|x - x_0\|_E < r\}$.
On voit évidement qu'une boule dépend de la norme considérée.

-> EXERCICE 2 : représenter la boule unité de $E = \mathbb{R}^2$ pour $\| \|_2$, $\| \|_1$ et $\| \|_\infty$.

• **Ouvert de** E **de** $E = (E, \| \|_E)$.
$U \subset E$ est un *ouvert de* E si pour tout $x \in U$, il existe une boule ouverte de centre x entièrement contenue dans E.
Une boule ouverte de E est donc un ouvert de E.

Rappelons au passage qu'une réunion quelconque (fine ou non) d'ouverts est un ouvert et qu'une intersection finie d'ouverts est un ouvert.

• **Fermé de de** $E = (E, \| \|_E)$.
$U \subset E$ est un *fermé de* E si son complémentaire $U^C = \{x \in E,\ x \notin U\}$ est un ouvert de E.

• **Borné de** E.
$U \subset E$ est un *borné de* E s'il existe $r > 0$ tels que $U \subset B_r(0)$.

• **Voisinage d'un point de de** $E = (E, \| \|_E)$.
$U \subset E$ est un *voisinage de* $x \in E$, s'il existe une boule ouverte de centre x entièrement contenue dans U (il se peut donc que $x \notin U$).
On dit que U est un *voisinage ouvert* de x si de plus U est un ouvert de E.

• **Normes équivalentes sur** E.
On dit que deux normes N_1 et N_2 sur E sont *équivalentes* s'il existe deux constantes $\alpha > 0$ et $\beta > 0$ telles que pour tout $x \in E$, $\alpha N_1(x) \leqslant N_2(x) \leqslant \beta N_1(x)$.

Exemple 1.3
Si $(E_k, \| \|_{E_k})_{1\leqslant k\leqslant n}$ sont n e.v.n, alors les normes produit de l'Exemple 1.2 sont toutes (deux à deux) équivalentes.

-> EXERCICE 3 : montrer que les normes produits $\| \|_1$, $\| \|_2$ et $\| \|_\infty$ sont 2 à 2 équivalentes.

On a la propriété suivante qui est d'un intérêt pratique évident.

Théorème 1.1 - Equivalence des normes en dimension finie
Si E est un espace vectoriel de **dimension finie** *alors toutes les normes sur E sont deux à deux équivalentes.*

Preuve du Théorème 1.1.
Notons $\| \|_E$ une norme quelconque sur E. On va montrer que $\| \|_E$ et $\| \|_1$ (voir Exemple 1.1) sont équivalentes et donc par transitivité que toutes les normes sont deux à deux équivalentes.

Supposons que $dim(E) = N$ et notons $(e_1, ..., e_N)$ une base de E. Donc tout $x \in E$ s'écrit $x = \sum_{k=1}^{N} x_k e_k$ avec $x_k \in \mathbb{R}$ coordonnées de x dans la base.

$\diamond$ On a alors pour tout $x \in E$

$\|x\|_E = \left\| \sum_{k=1}^{N} x_k e_k \right\|_E \leqslant \sum_{k=1}^{N} |x_k| \, \|e_k\|_E \leqslant \beta \, \|x\|_1$ où $\beta = Max(\|e_1\|_E, ..., \|e_N\|_E) > 0$ qui ne dépend pas de x.

$\diamond$ Introduisons l'ensemble compact $A = \{x \in E, \ \|x\|_1 = 1\}$ (sphère de rayon 1 dans $(E, \| \ \|_1)$) : fermé borné donc compact dans $(E, \| \ \|_1)$ puisque E est de dimension finie).

Or l'application $\| \ \|_E$ est continue de $(E, \| \ \|_1)$ dans $(E, \| \ \|_E)$ (conséquence de l'étape 1) et donc elle admet un minimum sur A (fct continue sur une compact) : il existe $x_0 \in A$ (donc non nul) tel que pour tout $x \in A$, $\|x\|_E \geqslant \alpha = \|x_0\|_E > 0$.

Soit alors $y \in E$ non nul. On a $x = \dfrac{y}{\|y\|_1} \in A$ et donc $\left\| \dfrac{y}{\|y\|_1} \right\|_E \geqslant \alpha$, c'est-à-dire $\|y\|_E \geqslant \alpha \, \|y\|_1$ pour tout $y \in E$.

■

Terminons ce paragraphe par une inégalités fondamentale concernant la norme euclidienne. On sait dans ce cas que la norme dérive du produit scalaire euclidien.

Proposition 1.1 - Inégalité de Cauchy-Schwarz
Pour tout $x, y \in \mathbb{R}^N$, on a $< x, y >_2 \leqslant \|x\|_2 \, \|y\|_2$ avec égalité ssi $\exists \alpha \in \mathbb{R}$ tel que $x = \alpha y$.

1.2 Suites

Dans toute cette partie $E = (E, \| \ \|_E)$ e.v.n.

• **Suite convergente de E.**

Une suite $(x_n)_{n \in \mathbb{N}} \subset E$ *converge* dans E vers $a \in E$, et on note $\lim_{n \longrightarrow \infty} x_n = a$ si $\lim_{n \longrightarrow \infty} \|x_n - a\|_E = 0$.

Cas de la dimension finie ($dim(E) = N$).

B base de E, on note $x_n = \sum_{i=1}^{N} x_{n,i} e_i$. Par Théorème 1.1, choisissant pour $\| \ \|_E$ la norme $\| \ \|_1$, on a $\|x_n - a\|_E = |x_{n,1} - a_1| + \cdots + |x_{n,N} - a_N|$ et on voit donc que la convergence de (x_n) est équivalente à la convergence de ses N coordonnées.

Cas d'un produit.

Si $E = E_1 \times ... \times E_n$ où $(E_k, \| \ \|_{E_k})_{1 \leqslant k \leqslant n}$ sont n e.v.n. La convergence de $(x_n) \subset E$ équivaut à celle de chacune de ses composantes $x_{n,i}$ dans $(E_i, \| \ \|_{E_i})$ pour $i = 1, .., N$ (choisir la norme produit $\| \ \|_1$).

On a alors la caractérisation suivante pour les sous-emsembles fermés d'un e.v.n.

• **Caractérisation des fermés de E.**

$U \subset E$ est *fermé* si et seulement si toute suite d'élements de U qui converge a sa limite dans U.

• **Suite de Cauchy de E.**

$(x_n)_{n \in \mathbb{N}} \subset E$ est de *Cauchy* dans E si $\forall \varepsilon > 0, \ \exists n_0 \geqslant 0, \ n, \ p \geqslant n_0 \Longrightarrow \|x_n - x_p\|_E \leqslant \varepsilon$.

Rappelons que toute suite convergente est de Cauchy.

Cette définition permet une caractérisation des suites convergentes de E dans le cas de la dimension finie et conduit à introduire une définition dans le cas de la dimension infinie..

Proposition 1.2
Soit E un e.v.n.
Si $dim(E) < +\infty$ alors une suite de E converge si et seulement si elle est de Cauchy dans E.
Si $dim(E) = +\infty$ alors si toute suite de Cauchy de E est convergente, on dit que E est complet.
Si $(E, \| \ \|_E)$ est complet, on dit que E est un espace de Banach (en dim finie, tout e.v.n est donc un espace de Banach).

Preuve de la Proposition 1.2.
C'est une conséquence du fait que $\mathbb{R}$ est complet et que la convergence est équivalente à la convergence composante par composante.

■

1.3 Continuité

• **Application continue.**
Soit U ouvert de E et $f : E \longrightarrow F$ définie sur U. Alors f est *continue* en $x_0 \in U$ si $\lim\limits_{\|x-x_0\|_E \longrightarrow 0} \|f(x) - f(x_0)\|_F = 0$ ce que l'on note $\lim\limits_{x \longrightarrow x_0} f(x) = f(x_0)$.
• **Caractérisation séquentielle de la continuité**.
f est continue en $x_0 \in U$ ssi $\forall \, (y_n)_{n \in \mathrm{N}} \subset U$ vérifiant $\lim\limits_{n \longrightarrow \infty} y_n = x_0$, on a $\lim\limits_{n \longrightarrow \infty} f(y_n) = f(x_0)$.
• **Caractérisation topologique de la continuité**.
Si $V \subset F$, l'image réciproque de V par f, notée $f^{-1}(V)$ est l'ensemble $\{x \in E, \ f(x) \in V\}$.

f est continue en $x_0 \in U$ si et seulement si l'image réciproque par f de tout voisinage ouvert (respectivement fermé) de $f(x_0)$ est un voisinage ouvert (resp. fermé) de x_0.

En particulier, on a f continue sur E si et seulement si l'image réciproque par f de tout ouvert (resp. fermé) de F est un ouvert de E (resp. fermé).

Nous en arrivons maintenant aux applications linéaires, puis multi-linéaires, qui constituent un des outils de base de ce cours.

1.4 Applications linéaires - Applications multi-linéaires

Commençons par une notation.
• **Notation**.
Pour E, F e.v.n., on notera $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E dans F. Si $E = F$, on notera $\mathcal{L}(E)$ l'ensemble $\mathcal{L}(E, E)$ c'est-à-dire l'ensemble des endomorphismes de E.

On a alors la caractérisation suivante pour la continuité des ces applications.

Proposition 1.3 - Continuité des applications linéaires
Soit $f : E \longrightarrow F$ une application linéaire. Les assertions suivantes sont équivalentes :

1. *f est continue sur E.*
2. *f est continue en 0.*
3. *Il existe $C > 0$ telle que $\forall x \in E$, $\|f(x)\|_F \leqslant C\|x\|_E$.*

On note $\mathbb{B}(\mathrm{E},\mathrm{F})$ (resp. $\mathbb{B}(\mathrm{E})$) l'e.v des applications linéaires et continues de E dans F (E resp.)

Preuve de la Proposition 1.3.
• 1. $\Longrightarrow$ 2. : évident.
• 2. $\Longrightarrow$ 3.
Si f est continue en 0 alors il existe $r > 0$ tel que pour tout $y \in E$ vérifiant $\|y\|_E < r$, on ait $\|f(y)\|_F \leqslant 1$ (conséquence de la définition de la continuité en 0 et du fait que $f(0) = 0$ puisque f est linéaire).
Soit alors $x \in E$ quelconque mais non nul (si $x = 0$ l'inégalité cherchée est évidente). Si on pose $y = \frac{r}{2\|x\|_E}x$, on a $\|y\|_E = \frac{r}{2} < r$ et donc $\|f(y)\|_F \leqslant 1$ d'après ce qui précède. Mais $f(y) = \frac{r}{2\|x\|_E}f(x)$ et donc $\|f(y)\|_F \leqslant 1 \Longleftrightarrow \|f(x)\|_F \leqslant \frac{2}{r}\|x\|_E$ d'où l'implication.
• 3. $\Longrightarrow$ 1.
Si $x_0 \in E$ alors $\|f(x) - f(x_0)\|_F = \|f(x - x_0)\|_F \leqslant C\|x - x_0\|_E$ et donc f est continue en x_0.
■

On a alors le résultat suivant qui simplifie grandement le calcul différentiel en dimension finie.

Théorème 1.2 - Continuité des applications linéaires
Soit E et F deux e.v.n. de dimensions finies. Alors toute application linéaire de E dans F est continue (donc $\mathbb{B}(\mathrm{E},\mathrm{F}) = \mathcal{L}(E,F)$).

FIN COURS 1
Preuve du Théorème 1.2.
Supposons que $dim(E) = N$ et notons $(e_k)_{1\leqslant k\leqslant N}$ une base de E. Si f est linéaire de E dans F, alors pour tout $x \in E$, $\|f(x)\|_F = \left\|f\left(\sum_{k=1}^{N} x_k e_k\right)\right\|_F = \left\|\sum_{k=1}^{N} x_k f(e_k)\right\|_F \leqslant \|x\|_\infty \sum_{k=1}^{N}\|f(e_k)\|_F$. Ainsi f est continue en 0, donc continue sur E d'après la Proposition 1.3. Noter qu'en fait on obtient que f est continue sur $(E, \|\ \|_\infty)$ et c'est le Théorème 1.1 qui permet de conclure.
■

On peut associer à $\mathbb{B}(\mathrm{E},\mathrm{F})$ une norme.

Proposition 1.4 - Norme sur $\mathbb{B}(\mathrm{E},\mathrm{F})$
Soient E et F deux e.v.n.
1. *L'application de $\mathbb{B}(\mathrm{E},\mathrm{F}) \longrightarrow \mathbb{R}$ définie pour tout $f \in \mathbb{B}(\mathrm{E},\mathrm{F})$ par $\|f\|_{\mathbb{B}(\mathrm{E},\mathrm{F})} = \displaystyle\operatorname*{Sup}_{\substack{x \in E \\ x \neq 0}} \frac{\|f(x)\|_F}{\|x\|_E}$ est une norme sur $\mathbb{B}(\mathrm{E},\mathrm{F})$.*

2. *Pour tout $f \in \mathbb{B}(\mathrm{E},\mathrm{F})$, $\|f\|_{\mathbb{B}(\mathrm{E},\mathrm{F})} = \displaystyle\operatorname*{Sup}_{\substack{\|x\|_E \leqslant 1 \\ x \neq 0}} \frac{\|f(x)\|_F}{\|x\|_E} = \operatorname*{Sup}_{\|x\|_E = 1} \|f(x)\|_F$.*

Preuve de la Proposition 1.4.
◇ Pour le **1.** il suffit de remarquer que cette définition à bien un sens pour f linéaire. En effet, si E est non vide, ce qui semble raisonnable, l'ensemble $A_f = \left\{ \frac{\|f(x)\|_F}{\|x\|_E};\ x \in E,\ x \neq 0 \right\}$ l'est également. De plus d'après la Proposition 1.3 (partie 3.) A_f est majoré et donc A_f qui est un ensemble non vide et majoré admet une borne suérieure. Le reste est évident.
◇ Pour **2.** il suffit de constater que $B_f = \{\|f(x)\|_F\,;\ x \in E,\ \|x\|_E = 1\} \subset A_f$ d'une part et que d'autre part, si $\alpha \in A_f$ alors $\alpha = \frac{\|f(x)\|_F}{\|x\|_E} = \|f(y)\|_F$ $(x \in E,\ x \neq 0)$ avec $y = \frac{x}{\|x\|_E}$ tq $\|y\|_E = 1$ et donc $\alpha \in B_f$. Donc $A_f = B_f$ et il ont donc même borne sup. On fait de même pour le dernier.
■

Remarque 1.1
Pour $f \in \mathbb{B}(\mathrm{E},\mathrm{F})$, la quantité $\|f\|_{\mathbb{B}(\mathrm{E},\mathrm{F})}$ est donc la plus petite constante $C \geqslant 0$ telle que $\|f(x)\|_F \leqslant C\,\|x\|_E$ pour tout $x \in E$ (borne supérieure de l'ensemble A_f). Cette remarque sert lorsque l'on veut calculer explicitement la norme d'une application linéaire continue.

• **Notation**
Dans tout ce cours, lorsque $f \in \mathcal{L}(E,F)$, et s'il n'y a pas d'ambiguïté possible, pour $h \in E$, on notera fh la quantité $f(h)$.

Pour terminer cette partie, on va redonner quelques propriétés et notations relatives aux applications n-linéaires. Commençons par rappeler la définition suivante.

• **Applications n-linéaires continues**.
Soient $(E_k, \|\ \|_{E_k})_{1 \leqslant k \leqslant n}$ et F $(n+1)$ e.v.n. On dit que $f : E_1 \times ... \times E_n \longrightarrow F$ est *n-linéaire* si elle est linéaire par rapport à chacune de ses n variables :

$$\left|\begin{array}{l} \forall i \in \{1, \cdots, n\},\ \forall (x_1, ..., x_n) \in E_1 \times ... \times E_n, \forall \alpha \in \mathbb{R}, \forall h_i \in E_i, \\ f(x_1, ..., x_i + h_i, ..., x_n) = f(x_1, ..., h_i, ..., x_n) + f(x_1, ..., x_i, ..., x_n), \\ f(x_1, ..., \alpha x_i, ..., x_n) = \alpha f(x_1, ..., x_i, ..., x_n). \end{array}\right.$$

• **Notation**.
On note $\mathbb{B}^{\mathrm{n}}(E_1, ..., E_n; F)$ le $\mathbb{R}$-espace vectoriel des applications n-linéaires continues de $E_1 \times ... \times E_n$ (muni de la norme produit) dans F.
Si $E_1 = ... = E_n = E$, on note $\mathbb{B}^{\mathrm{n}}(E; F)$ l'ensemble $\mathcal{L}^n(E, ..., E; F)$.

Remarque 1.2
Notons sans l'énoncer que l'analogue de la Proposition 1.3 reste vraie en remplaçant la propriété 3. par $\|f(x_1,...,x_n)\|_F \leqslant C\,\|x_1\|_{E_1} ... \|x_n\|_{E_n}$.
Le théorème suivant regroupe différentes propriétés qui nous seront utiles lors de ce cours.

Théorème 1.3 - Norme sur $\mathbb{B}^n(E_1,...,E_n;F)$ - Continuité
Soient $(E_k, \|\ \|_{E_k})_{1\leqslant k\leqslant n}$ *et* F $(n+1)$ *e.v.n.*

1. *L'application de* $\mathbb{B}^n(E_1,...,E_n;F) \longrightarrow \mathbb{R}$ *définie pour tout* $f \in \mathbb{B}^n(E_1,...,E_n;F)$ *par* $\|f\|_{\mathbb{B}^n(E_1,...,E_n;F)} = \underset{\substack{x_i \in E_i \\ x_i \neq 0}}{Sup} \dfrac{\|f(x_1,...,x_n)\|_F}{\|x_1\|_{E_1} ... \|x_n\|_{E_n}}$ *est une norme sur* $\mathbb{B}^n(E_1,...,E_n;F)$.

2. *Une application n-linéaire de* $E_1 \times ... \times E_n \longrightarrow F$ *est continue si et seulement si la quantité* $\|f\|_{\mathbb{B}^n(E_1,...,E_n;F)}$ *est finie.*

3. *En dimension finie toute application n-linéaire est continue.*

Preuve du Théorème 1.3.
Pour **1.** et **3.** : même principe que pour le cas linéaire. Seule le **2.** mérite d'être détaillé.
Supposons f continue (en 0) et prenons sur $E_1 \times ... \times E_n$ la norme produit indicée par $p = \infty$. Alors il existe $\eta > 0$ tel que $\|(y_1,...,y_n)\|_\infty \leqslant \eta$ entraine $\|f(y_1,...,y_n)\|_F \leqslant 1$. En particulier, si on choisit y tel que $y_k = \eta\, \frac{x_k}{\|x_k\|_{E_k}}$ on obtient $\|(y_1,...,y_n)\|_\infty = \eta$ ce qui entraine $1 \geqslant \|f(y_1,...,y_n)\|_F = \dfrac{\eta^n}{\|x_1\|_{E_1} ... \|x_n\|_{E_n}} \|f(x_1,...,x_n)\|_F$ ce qui montre que la borne supérieure est finie.
Réciproquement, si la borne supérieure est finie notons S sa valeur et remarquons que pour tout $(y_1,...,y_n) \in E_1 \times ... \times E_n$ on a $\|f(y_1,...,y_n)\|_F \leqslant S\,\|y_1\|_{E_1} ... \|y_n\|_{E_n}$. Ainsi puisque

$$\begin{aligned}
f(x_1,...,x_n) - f(y_1,...,y_n) = \ & f(x_1,x_2,...,x_n) - f(y_1,x_2,...,x_n) \\
& + f(y_1,x_2,...,x_n) - f(y_1,y_2,x_3,...,x_n) \\
& + f(y_1,x_2,...,x_n) - f(y_1,y_2,x_3,...,x_n) + ... \\
& + f(y_1,...,y_i,x_{i+1},x_n) - f(y_1,...,y_i,y_{i+1},x_{i+2},...,x_n) + ... \\
& + f(y_1,y_2,...,y_{n-1},x_n) - f(x_1,,...,x_n)
\end{aligned}$$

$$\begin{aligned}
\text{on a } \|f(x_1,...,x_n) - f(y_1,...,y_n)\|_F \leqslant \ & \|f(x_1,x_2,...,x_n) - f(y_1,x_2,...,x_n)\|_F \\
& + \|f(y_1,x_2,...,x_n) - f(y_1,y_2,x_3,...,x_n)\|_F + ... \\
& + \|f(y_1,y_2,...,y_{n-1},x_n) - f(x_1,,...,x_n)\|_F \\
\leqslant \ & S\ \|x_1 - y_1\|_{E_1} \|x_2\|_{E_2} ... \|x_n\|_{E_n} \\
& + S\ \|y_1\|_{E_1} \|x_2 - y_2\|_{E_2} ... \|x_n\|_{E_n} + ... \\
& + S\ \|y_1\|_{E_1} ... \|y_{n-1}\|_{E_{n-1}} \|x_n - y_n\|_{E_n}.
\end{aligned}$$

Donc que pour tout $R > 0$, et tout (x_k, y_k) vérifiant $\|x_k\|_{E_k} \leqslant R$ et $\|y_k\|_{E_k} \leqslant R$, on a $\|f(x_1,...,x_n) - f(y_1,...,y_n)\|_F \leqslant S\ R^{n-1} \displaystyle\sum_{k=1}^{n} \|x_k - y_k\|_{E_k}$ ce qui entraine la continuité de f.

■

FIN DU CHAPITRE 0

Chapitre 2

Différentielle d'une application

2.1 De la notion de dérivée à celle de différentielle

Le début de ce cours est consacré à l'introduction des différentes notions de "dérivées". Partant de la notion classique et connue de dérivée, nous allons voir pourquoi il est nécesssaire d'introduire de nouvelles notions.
Commençons par rappeler ce qui est déja connu.

2.1.1 Cas d'une application de $\mathbb{R} \longrightarrow \mathbb{R}^n$: notion de dérivée

Définition 2.1 - Dérivée
Soit $f : \mathbb{R} \longrightarrow \mathbb{R}^n$ une application définie sur un ouvert $U \subset \mathbb{R}$. On dit que f admet une dérivée au point $t_0 \in U$ s'il existe $\ell \in \mathbb{R}^n$ telle que

$$\lim_{h \longrightarrow 0,\ h \neq 0} \frac{f(t_0 + h) - f(t_0) - \ell h}{h} = 0. \tag{2.1}$$

On note alors $f'(t_0) = \ell \in \mathbb{R}^n$ la dérivée de f en t_0 et on dit que f est dérivable en t_0.
Si on se donne une base de $\mathbb{R}^n$, on sait que $y \longrightarrow b$ dans $\mathbb{R}^n$ ssi $y_i \longrightarrow b_i$ dans $\mathbb{R}$, $(i = 1, ..., n)$ et donc notant $f = (f_i)$, la relation (2.1) montre que f est dérivable en t_0 ssi chacun des f_i l'est et alors $f'(t_0) = (f_i'(t_0))$.

Remarque 2.1
1) La relation (2.1) est équivalente à $f(t_0 + h) = f(t_0) + h\ell + o(h)$ au voisinage de $h = 0$ et on retrouve le fait que f est dérivable en t_0 ssi elle y admet un développement limité d'ordre 1. En particulier, si f est dérivable en t_0 alors elle y est continue
2) Cette définition s'étend naturellement au cas où $f : \mathbb{R} \longrightarrow F$ avec F e.v.n
3) En revanche, elle n'a aucun sens si la fonction considérée $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ avec $n > 1$ (comment diviser par un vecteur $h \in \mathbb{R}^n$? ? ?). Néanmoins (puisque $h \in \mathbb{R}^*$), la question de la dérivabilité de f en t_0 équivaut à montrer qu'il existe $\ell \in \mathbb{R}$ tel que $\displaystyle\lim_{h \longrightarrow 0, h \neq 0} \frac{f(t_0 + h) - f(t_0) - \ell h}{|h|} = 0$. Cette question peut, elle, être généralisée ce qui conduit à la notion de différentielle que nous allons maintenant définir.

2.1.2 Cas d'une application de $E \longrightarrow F$: notion de différentielle

Dans ce qui suit E et F désignent deux e.v.n (de dimension finie ou non), U un ouvert de E et f une application définie sur U et à valeurs dans F.

Définition 2.2 - Différentielle
Pour $a \in U$, on dit que f admet une différentielle au point a s'il existe une application linéaire et continue $L : E \longrightarrow F$ *telle que*

$$\lim_{h \longrightarrow 0,\ \|h\|_E \neq 0} \frac{f(a+h) - f(a) - Lh}{\|h\|_E} = 0, \tag{2.2}$$

c'est-à-dire

$$\lim_{h \longrightarrow 0,\ \|h\|_E \neq 0} \frac{\|f(a+h) - f(a) - Lh\|_F}{\|h\|_E} = 0.$$

On note alors $df(a) = L \in \mathbb{B}(\mathrm{E},\mathrm{F})$ cette différentielle et on dit que f est différentiable en a.

On dit que f est différentiable sur U ouvert de E si elle est différentiable en tout point de U.

Remarque 2.2
1) *Si $dim(E) < +\infty$, la continuité de l'application linéaire L est automatique (théorème 1.2 chapI).*
2) *La relation (2.2) est équivalente à dire que la quantité $\|f(a+h) - f(a) - Lh\|_F$ est négligeable devant $\|h\|_E$ au voisinage de $h = 0$ ce que l'on écrit*

$$f(a+h) = f(a) + Lh + o(\|h\|_E),$$

(sous-entendu pour tout $h \in E$ tel que $a + h \in U$ ce qui est licite puisque U est ouvert).
Ainsi, lorsque f est différentiable en a, $f(a) + df(a)h$ apparaît comme une approximation de f au premier ordre en h au voisinage de a.
3) *Si f est différentiable sur U alors $df : U \longrightarrow \mathbb{B}(\mathrm{E},\mathrm{F})$.*

L'objet ainsi défini a bien un sens comme le montrent les deux propositions suivantes.

Proposition 2.1 - Unicité de la différentielle
Si f est différentiable en $a \in U$ alors cette différentielle est unique.

Preuve de la Proposition 2.1.
Supposons que f admettent deux différentielles L_1 et L_2 en a. Alors par définition

$$f(a+h) = f(a) + L_1 h + o(\|h\|_E) \quad et \quad f(a+h) = f(a) + L_2 h + o(\|h\|_E),$$

et donc, prenant la différence de ces deux égalités, on trouve $L_1 h - L_2 h = o(\|h\|_E)$.
Fixons alors $x \in E$. Puisque U est ouvert, on a pour $t \in \mathbb{R}$, $a + tx \in U$ si $|t|$ est assez petit. On peut donc prendre $h = tx$ et on obtient ainsi $t\,(L_1 x - L_2 x) = \|h\|_E \; o(t)$ par linéarité de L_1 et L_2.
Il suffit alors de diviser cette égalité par t puis de faire tendre t vers 0 pour trouver $L_1(x) = L_2(x)$ pour tout $x \in E$.

■

Proposition 2.2 - Invariance de la différentielle pour des normes équivalentes
Soient $\| \ \|_E$ et $\| \ \|_E^{\#}$ deux normes équivalentes sur E, et $\| \ \|_F$ et $\| \ \|_F^{\#}$ deux normes équivalentes sur F. Alors si f est différentiable en $a \in U$ pour $\| \ \|_E$ et $\| \ \|_F$, elle l'est également en $a \in U$ pour $\| \ \|_E^{\#}$ et $\| \ \|_F^{\#}$, et sa différentielle est la même.

Preuve de la Proposition 2.2.
Par définition, si f est différentiable en $a \in U$ pour $\| \ \|_E$ et $\| \ \|_F$, il existe une application linéaire continue de $(E, \| \ \|_E)$ dans $(F, \| \ \|_F)$ telle que (2.2) soit vérifiée.
Mais $\| \ \|_E$ et $\| \ \|_E^{\#}$ d'une part, et $\| \ \|_F$ et $\| \ \|_F^{\#}$ d'autre part étant équivalentes, il existe deux constantes $C_1, C_2 > 0$ telles que $\frac{1}{\|h\|_E^{\#}} \leqslant C_1 \frac{1}{\|h\|_E}$ et $\|f(a+h) - f(a) - Lh\|_F^{\#} \leqslant C_2 \|f(a+h) - f(a) - Lh\|_F$ pour tout $h \in E$. Donc $\dfrac{\|f(a+h) - f(a) - Lh\|_F^{\#}}{\|h\|_E^{\#}} \leqslant C_1 \ C_2 \ \dfrac{\|f(a+h) - f(a) - Lh\|_F}{\|h\|_E}$ et on conclut donc par (2.2).

■

Corollaire 2.2.1
Lorsque E et F sont de dimension finie, il suffit de montrer la différentiabilité pour une norme particulière pour l'avoir pour toute norme.

Mentionons pour terminer ce paragraphe la propriété suivante.

Proposition 2.3 - Différentiabilité et continuité
Si f est différentiable en $a \in U$ alors elle y est continue.

Preuve de la Proposition 2.3.
Si f est différentiable en a alors $\|f(a+h) - f(a)\|_F = \|df(a)h + o(\|h\|_E)\|_F \leqslant \|df(a)h\|_F + \|h\|_E \|o(1)\|_F$ et la continuité de $df(a)$ permet de conclure.

■

Remarque 2.3
La réciproque est évidement fausse...une application continue n'a pas de raison d'être différentiable.

2.1.3 Dérivée directionnelle d'une application

Définition 2.3 - Dérivée directionnelle
Soit E et F deux e.v.n, U un ouvert de E, f une application de U dans F et $a \in U$, $b \in E$, $b \neq 0$. On dit que f admet en a une dérivée directionnelle dans la direction $b \neq 0$ ou que f est dérivable en a dans la direction b si $(t \in \mathbb{R} \mapsto f(a+tb))$ est dérivable en 0 c'est-à-dire si

$$\lim_{t \longrightarrow 0} \frac{f(a+tb) - f(a)}{t} \ \textit{existe.}$$

On notera $f'(a)(b)$ la dérivée directionnelle de f en a dans la direction b.

Remarque 2.4
Quand a *et* b *sont fixés, et que* t *décrit* $\mathbb{R}$*, alors* $(a+tb)$ *décrit la droite (affine) de vecteur directeur* b *et passant par* a*. Ainsi la limite précédente revient à étudier ce qui se passe pour des points situés sur cette droite et tendant vers* a*.*

FIN COURS 2
Les dérivées directionnelles étant obtenues pour une convergence particulière vers le point considéré (sur des droites), la notion obtenue est (strictement) plus faible que la notion de différentielle. On a évidement la propriété suivante.

Proposition 2.4
Si f *est différentiable en* a*, alors* f *admet en* a *des dérivées directionnelles dans toutes les directions.*
De plus dans ce cas, $\forall b \in E$*,* $f'(a)(b) = df(a)b$*.*
En particulier, on a l'égalité suivante (très utile pour calculer une différentielle lorsque l'on sait que la fonction est différentiable) :

$$df(a)b = \lim_{t \longrightarrow 0} \frac{f(a+tb)-f(a)}{t}$$

Preuve de la Proposition 2.4
Partie *1* : évident.
Partie *2* :
Il suffit de prendre $h = tb$ avec $b \in E$ fixé et $t \longrightarrow 0$ dans (2.2).

■

On a le corollaire suivant :

Corollaire 2.4.1
S'il existe $b \in E$ *tel que que* f *n'admette pas de dérivée directionnelles en* $a \in E$ *dans la direction* $b \neq 0$ *alors* f *n'est pas différentiable en* a*.*

Remarque 2.5
1) *La notion de dérivée directionnelle est strictement plus faible et elle ne permet donc en aucun cas d'en déduire la différentiabilité comme le montre l'exemple de la fonction* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *définie par* $f(x,y) = \dfrac{x^3y}{x^4+y^2}$ *si* $(x,y) \neq (0,0)$ *et* $f(0,0) = 0$ *qui admet en* $(0,0)$ *des dérivées directionnelles nulles dans toutes les directions mais qui n'est pas différentiable en* $(0,0)$*.*
2) *S'il n'est pas possible d'étudier la différentiabilité d'une application à l'aide de ses dérivées directionnelles, en revanche,* **quand on sait qu'une application est différentiable***, mais qu'on ne connaît pas l'expression de sa différentielle, les dérivées directionnelles peuvent fournir un moyen simple pour le calcul de cette expression.*

2.2 Premiers exemples

Voyons tout de suite trois exemples simples que vous devrez connaitre. Le premier montre que la définition est bien, en un certain sens, la généralisation naturelle de la notion de dérivée pour les fonctions réelles.

2.2.1 Différentielle de $f : \mathbb{R} \longrightarrow \mathbb{R}$

Si $f : \mathbb{R} \longrightarrow \mathbb{R}$, alors elle est différentiable en $a \in \mathbb{R}$ si et seulement si elle est dérivable en $a \in \mathbb{R}$, et alors $df(a)$ est définie sur $\mathbb{R}$ par $df(a) : t \mapsto f^{'}(a)t$ pour tout $t \in \mathbb{R}$.

En effet, si f est dérivable en $a \in \mathbb{R}$, alors pour tout $t \in \mathbb{R}$, $f(a+t) = f(a) + f^{'}(a)t + o(t)$ et donc $L : \mathbb{R} \longrightarrow \mathbb{R}$ est définie par $L : t \mapsto f^{'}(a)t$.

Réciproquement, s'il existe $L : \mathbb{R} \longrightarrow \mathbb{R}$ linéaire telle que $\displaystyle\lim_{t \longrightarrow 0,\ t \neq 0} \frac{f(a+t) - f(a) - L(t)}{t} = 0$ alors $\displaystyle\lim_{t \longrightarrow 0,\ t \neq 0} \frac{f(a+t) - f(a)}{t} = L(1)$ (ici $L(t) = tL(1)$) et donc f est dérivable en a avec $f^{'}(a) = L(1)$.

2.2.2 Différentielle d'une constante

Si $f : E \longrightarrow F$ est constante alors elle est différentiable sur E et $\forall x \in E$, $df(x) = 0$.

2.2.3 Différentielle d'une application linéaire continue

Si $f \in \mathbb{B}(\mathrm{E}, \mathrm{F})$ alors elle est différentiable sur E et $\forall x \in E$, $df(x) = f$.

En effet, par linéarité de f, on a pour tout $x, h \in E$, $f(a+h) - f(a) - fh = 0$ et donc $L = f$.

2.2.4 Différentielle d'une application bilinéaire continue

Si $f \in \mathbb{B}(E_1, E_2; F)$ alors f est différentiable sur $E_1 \times E_2$. Sa différentielle est alors donnée pour tout $x = (x_1, x_2) \in E_1 \times E_2$ et tout $h = (h_1, h_2) \in E_1 \times E_2$ par

$$df(x_1, x_2)(h_1, h_2) = f(x_1, h_2) + f(h_1, x_2). \tag{2.3}$$

En effet, on a $f(x+h) - f(x) = f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2) = \underbrace{f(x_1, h_2) + f(h_1, x_2)}_{Lh} + \underbrace{f(h_1, h_2)}_{R(h)}$.

Comme $h \mapsto Lh$ est linéaire et continue (par Remarque 1.2 puisque f bili. cont.), f est différentiable en x ssi $R(h) = o(\|h\|_E)$. Et puisque $\|R(h)\|_F \leqslant C\|h_1\|_{E_1} \|h_2\|_{E_2} \leqslant C\|h\|_\infty^2$, on a bien $R(h) = o(\|h\|_E)$. Donc f est différentiable en tout point x, et sa différentielle est donnée par $df(x)h = Lh$.

2.3 Cas des applications de $\mathbb{R}^n \longrightarrow \mathbb{R}^p$

Dans ce paragraphe $E = \mathbb{R}^n$, $F = \mathbb{R}^p$ et on note $B_E = (e_i)_1^n$ une base de E, $(x_i) \in \mathbb{R}^n$ coordonnées de $x \in E$ dans B_E, et $B_F = (\tilde{e}_i)_1^p$ une base de F.

2.3.1 Rappel : identifications d'e.v

On dit que 2 e.v sont isomorphes s'il existe une bijection linéaire continue ψ de l'un dans l'autre. Dans ce cas deux tels espaces sont des copies l'un de l'autre et rien ne les distingue au niveau de la structure d'espace vectoriel.

• Exemple déja utilisé plusieurs fois :

si $dim(E) = n$ et B donnée alors $E \approx R^n$ avec
$$\begin{array}{rcl} \psi : E & \longrightarrow & \mathbb{R}^n \\ x & \mapsto & (x_1, ..., x_n) \end{array}$$
tq $x = \sum_{i=1}^n x_i e_i$: correspondance entre E et $\mathbb{R}^n$
(on note $x \sim (x_i)_i$ la relation $\psi(x) = (x_i)_i$).

• La relation d'identification "$\approx$" est une relation d'équivalence (elle est donc en particulier transitive).

• **L'identification** $\mathcal{L}(\mathbb{R}^n,\mathbb{R}) \approx \mathbb{R}^n$ (formes linéaires).

Si $L \in \mathcal{L}(\mathbb{R}^n,\mathbb{R})$ alors $\forall x \in \mathbb{R}^n,\ L(x) = \sum_{j=1}^{n} x_j L(e_j)$.

Donc L est entièrement déterminée par les n-scalaires $(L(e_1), ..., L(e_n)) \in \mathbb{R}^n$.

Ainsi $\mathcal{L}(\mathbb{R}^n,\mathbb{R}) \approx \mathbb{R}^n$ avec $\begin{array}{l}\psi : \mathcal{L}(\mathbb{R}^n,\mathbb{R}) \longrightarrow \mathbb{R}^n \\ L \mapsto (L(e_1), ..., L(e_n))\end{array}$ et la relation de correspondance $L(x) = \sum_{j=1}^{n} x_j L(e_j),\ \forall x \in \mathbb{R}^n$.

On peut généraliser ce procédé de la manière suivante.

• **L'identification** $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p) \approx M_{p,n}(\mathbb{R})$.

Si $L \in \mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)$ alors

$$L(x) = \sum_{i=1}^{p} L_i(x)\widetilde{e_i} = \sum_{i=1}^{p}\sum_{j=1}^{n} x_j L_i(e_j)\widetilde{e_i} = \begin{pmatrix} L_1(e_1) & \cdots & L_1(e_n) \\ \vdots & & \vdots \\ L_p(e_1) & \cdots & L_p(e_n) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = Ax$$

avec $A \in M_{p,n}(\mathbb{R})$ matrice à p-lignes et n-colonnes à coefficients réels tq $(A)_{ij} = L_i(e_j)$.

Ainsi $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p) \approx M_{p,n}(\mathbb{R})$ avec $\begin{array}{rcl}\psi : \mathcal{L}(\mathbb{R}^n,\mathbb{R}^p) & \longrightarrow & M_{p,n}(\mathbb{R}) \\ L & \mapsto & A\end{array}$ et la relation de correspondance $L(x) = Ax,\ \forall x \in \mathbb{R}^n$ $(L \sim A)$.

Classiquement, on note : $$A = \begin{array}{cc} \begin{array}{ccc} L(e_1) & \cdots & L(e_n) \end{array} & \\ \begin{pmatrix} a_{11} & \cdots & \cdots & a_{1n} \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{pmatrix} & \begin{array}{c} \widetilde{e_1} \\ \\ \vdots \\ \\ \widetilde{e_p} \end{array} \end{array}.$$

2.3.2 Dérivée partielle - Lien avec la différentielle

Cas particulier de DD.

Définition 2.4 - Dérivée partielle

Soit $U \subset E$ un ouvert, $f : U \longrightarrow F$ une application, $a \in U$ et $j \in \{1, ..., n\}$ un entier. On dit que f admet en a une dérivée partielle par rapport à x_j si $t \in \mathbb{R} \mapsto f(a + te_j)$ est dérivable en $t = 0$, c'est-à-dire si $\lim_{t \longrightarrow 0} \frac{f(a + te_j) - f(a)}{t}$ existe (dans F). Dans ce cas on note

$$\frac{\partial f}{\partial x_j}(a) = \lim_{t \longrightarrow 0} \frac{f(a + te_j) - f(a)}{t}.$$

On dira que f admet une dérivée partielle par rapport à x_j sur U si $\frac{\partial f}{\partial x_j}(a)$ en tout $a \in U$.

Remarque 2.6
La dérivée partielle suivant x_j ne voit que ce qui se passe suivant e_j (donc sur les axes) et en conséquence :
1) *d'un point de vue calculatoire, tout se passe comme si toutes les autres variables étaient fixées et considérées comme constantes et les règles de calcul en découlent ;*
2) *une fonction peut admettre des dérivées partielles en un point sans y être continue. En fait l'existence de ces dérivées partielles entrainera "la continuité correspondante" c'est-à-dire celle sur les axes i.e. celle de $t \mapsto f(a + te_j)$;*
3) *en particulier, l'existence des dérivées partielles n'entraîne pas la différentiabilité.*

On a toutefois le lien suivant.

Proposition 2.5 - Lien entre dérivées partielles et différentielle
Si f est différentiable en $a \in U$ alors elle admet en a des dérivées partielles par rapport x_j pour tout $j \in \{1, ..., n\}$ et de plus, on a pour tout $h \in E$

$$df(a)h = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\ h_j \tag{2.4}$$

En particulier, on notera que

$$df(a)e_j = \frac{\partial f}{\partial x_j}(a).$$

Preuve de la Proposition 2.5 : cas particulier de la Proposition 2.4 et de la Définition 2.3.

■

Remarque 2.7
Lorsque la fonction est différentiable, *la formule (2.4) fournit un moyen pour calculer cette différentielle en calculant les dérivées partielles.*
Une autre manière fort utile d'énoncer la Proposition 2.5 pour l'étude de la différentiabilité d'une fonction est la suivante.

Corollaire 2.5.1 - Non différentiabilité
S'il existe $i \in \{1, \cdots, n\}$ tq $\frac{\partial f}{\partial x_i}(a)$ n'existe pas alors f n'est pas différentiable en a.

Terminons ce paragraphe par une propriété très utile lors de l'étude de la différentiabilité de fonctions.

Proposition 2.6
Si f admet en $a \in U$ des dérivées partielles continues alors f est différentiable en a et sa différentielle en a est donnée par la formule (2.4).

Preuve de la Proposition 2.6.
Bien qu'il soit possible de le démontrer à ce niveau du cours, ce résultat est un corollaire d'un des résultats du chapitre suivant et nous y reviendrons donc plus tard. Nous l'admettrons donc provisoirement.

■

2.3.3 Matrice jacobienne - Gradient

Voyons tout de suite le lien existant entre le rappel d'algèbre sur les identifications et notre cours. Puisque si f est différentiable en a, sa différentielle en a, $df(a)$ est une application linéaire de E dans F, on peut donc chercher l'expression de sa matrice relativement aux bases canonique B_E et B_F. La i-ème colonne de cette matrice sera alors donné par $df(a)e_i$ c'est-à-dire la dérivée partielle de f par rapport à x_i au point a !

Formalisons tout cela dans une définition.

Définition 2.5 - Matrice Jacobienne
Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ avec U ouvert, différentiable en $a \in U$.
Alors la matrice de $df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ relativement aux base canoniques B_E et B_F s'appelle **matrice jacobienne** *(ou plus simplement* **jacobienne***) de f en a. On la note $J_f(a)$. Ses coefficients sont les $(J_f(a))_{ij} = \dfrac{\partial f_i}{\partial x_j}(a)$, c'est-à-dire que $df(a)$ s'identifie à la matrice $J_f(a)$, ce que l'on écrit $df(a) \sim \left[\dfrac{\partial f_i}{\partial x_j}(a)\right]$ pour signifier que $\forall h \in \mathbb{R}^n$, $df(a)h = J_f(a)h$*

Dans la pratique, le cas particulier où la différentielle est une forme linéaire (i.e. $F = \mathbb{R}$ soit $p = 1$) porte un nom spécial et possède sa propre notation. Dans ce cas, la matrice jacobienne de f en a possède une seule ligne et $df(a)$ est alors identifiée à un vecteur de $\mathbb{R}^n$.

Définition 2.6 - Gradient
Soit $f : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ avec U ouvert, différentiable en $a \in U$.

On appelle **gradient** *de f en a le vecteur noté $\nabla f(a) \in \mathbb{R}^n$ défini par $\nabla f(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}$.*
La différentielle de f en a s'identifie alors au vecteur $\nabla f(a)$, ce que l'on écrit $df(a) \sim \nabla f(a)$ pour signifier que $\forall h \in \mathbb{R}^n$, $df(a)(h) =< \nabla f(a), h >_2$

2.4 Opérations sur les applications différentiables

Les deux propriétés que nous allons voir maintenant sont les généralisations de propriétés bien connues.

Proposition 2.7 - Stabilité linéaire
Soient $f, g : U \longrightarrow F$ deux applications différentiables en $a \in U$, et soit $\alpha \in \mathbb{R}$. Alors les applications αf et $f + g$ sont différentiables en a avec $d(\alpha f)(a) = \alpha df(a)$ et $d(f + g)(a) = df(a) + dg(a)$.

Preuve de la Proposition 2.7. Evident.

■

Le théorème suivant est d'une très grande importance pour établir par décomposition la différentiabilité de certaines applications ainsi que pour calculer leur différentielle.

Théorème 2.1 - Différentielle d'une composée
Soient E, F et G trois e.v.n., U un ouvert de E, V un ouvert de F et $f : U \longrightarrow F$, $g : V \longrightarrow G$ deux applications vérifiant $f(U) \subset V$. On suppose que f est différentiable en $a \in U$ et que g est différentiable en $f(a) \in V$.
Alors l'application $g \circ f : U \longrightarrow G$ est différentiable en $a \in U$ et

$$d(g \circ f)(a) = dg(f(a)) \circ df(a) \tag{2.5}$$

Preuve du Théorème 2.1.
Posant $b = f(a)$, la différentiabilité de f en a et de g en b se traduit par l'existence de deux fonctions ε et φ définies dans un voisinage de 0_E et de 0_F respectivement telles que $\lim\limits_{h \longrightarrow 0, \|h\|_E \neq 0} \varepsilon(h) = 0$, $\lim\limits_{u \longrightarrow 0, \|u\|_F \neq 0} \varphi(u) = 0$ et

$$f(a+h) = f(a) + df(a)h + \|h\|_E \ \varepsilon(h) \tag{2.6}$$
$$g(b+u) = g(b) + dg(b)u + \|u\|_F \ \varphi(u). \tag{2.7}$$

On a alors $(g \circ f)(a+h) = g(f(a+h)) = g\Big(f(a) + df(a)h + \|h\|_E \ \varepsilon(h)\Big)$ par (2.6)
et donc en utilisant (2.7) avec $u = df(a)h + \|h\|_E \ \varepsilon(h)$ (qui tend bien vers 0 lorsque $h \longrightarrow 0$ par continuité de $df(a)$), on obtient

$$g(f(a+h)) = g(b) + dg(b)\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big) + \Big\|\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big)\Big\|_F \ \varphi\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big).$$

Considérons alors séparément les deux derniers termes de cette égalité apparaissant dans le produit.
Pour le premier, utilisant la linéarité de $dg(b)$, on peut écrire que

$$\begin{aligned} dg(b)\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big) &= dg(b)\Big(df(a)h\Big) + \|h\|_E \ dg(b)\Big(\varepsilon(h)\Big) \\ &= dg(b)\Big(df(a)h\Big) + o(\|h\|_E) \end{aligned} \tag{2.8}$$

puisque $\Big\| \|h\|_E \ dg(b)\Big(\varepsilon(h)\Big)\Big\|_F = \|h\|_E \ \Big\|dg(b)\Big(\varepsilon(h)\Big)\Big\|_F \leqslant C_1 \ \|h\|_E \ \|\varepsilon(h)\|_F$ par continuité de $dg(b)$.
Pour le second, on remarque que d'une part

$$\Big\|\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big)\Big\|_F \leqslant (C_1 + \|\varepsilon(h)\|_F) \ \|h\|_E \leqslant C_2 \|h\|_E \, ,$$

et que d'autre part $df(a)h + \|h\|_E \ \varepsilon(h) \longrightarrow 0$ qiand $h \longrightarrow 0$ et donc que $\varphi\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big)$ lorsque $h \longrightarrow 0$.
Ainsi, on a $\Big\|\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big)\Big\|_F \ \varphi\Big(df(a)h + \|h\|_E \ \varepsilon(h)\Big) = o(\|h\|_E)$ quand $\|h\|_E \longrightarrow 0$ ce qui compte tenu de (2.8) permet d'obtenir

$$g(f(a+h)) = g(b) + dg(b)\Big(df(a)h\Big) + o(\|h\|_E).$$

Ceci montre que l'application $g \circ f$ est différentiable en a et que sa différentielle est donnée par (2.5).
■

FIN COURS 3

Remarque 2.8 Cas de la dimension finie
Dans le cas particulier de fonctions différentiables $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ *et* $f : \mathbb{R}^p \longrightarrow \mathbb{R}^n$*, la formule (2.5) pour* $\varphi = g \circ f : \mathbb{R}^p \longrightarrow \mathbb{R}$*, à savoir* $d\varphi(a) = dg(f(a))df(a)$ *donne pour les jacobiennes*

$$\underbrace{J_\varphi(a)}_{\in M_{1,n}(\mathbb{R})} = \underbrace{J_g(f(a))}_{\in M_{1,n}(\mathbb{R})} . \underbrace{J_f(a)}_{\in M_{n,p}(\mathbb{R})}$$

et donc pour $j = 1, ..., p$

$$\begin{aligned}\Big(J_\varphi(a)\Big)_{1j} = \Big(J_g(f(a)).J_f(a)\Big)_{1j} &\iff \frac{\partial \varphi}{\partial x_j}(a) = \sum_{k=1}^{n} \Big(J_g(f(a))\Big)_{1k}\Big(J_f(a)\Big)_{kj} \\ &\iff \frac{\partial \varphi}{\partial x_j}(a) = \sum_{k=1}^{n} \frac{\partial g}{\partial x_k}(f(a))\frac{\partial f_k}{\partial x_j}(a)\end{aligned}$$

pour les dérivées partielles

2.5 Applications de classe C^1

Définition 2.7 - Classe C^1
L'application $f : U \longrightarrow F$ *est dite continûment différentiable sur* U *ou de classe* C^1 *sur* U *(on note alors* $f \in C^1(U)$*) si :*

(i) f *est différentiable sur* U *;*

(ii) *sa différentielle* $df : U \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$ *est continue.*

La propriété suivante est évidente compte tenu de ce que nous avons vu mais elle est d'un intérêt pratique certain.

Proposition 2.8 - Stabilité linéaire de la classe C^1
Soient f, $g : U \longrightarrow F$ *deux applications de classe* C^1 *sur* U*, et soit* $\alpha \in \mathbb{R}$*. Alors les applications* αf *et* $f + g$ *sont de classe* C^1 *sur* U*.*

Preuve de la Proposition 2.8
D'après la Proposition 2.7 la différentiation est linéaire et comme la continuité l'est également, la preuve est directe.

■

On déduit immédiatement du Théorème 2.1 la propriété suivante

Corollaire 2.1.1 - Composée de fonctions C^1
Soient E*,* F *et* G *trois e.v.n.,* U *un ouvert de* E*,* V *un ouvert de* F *et* $f : U \longrightarrow F$*,* $g : V \longrightarrow G$ *deux applications vérifiant* $f(U) \subset V$*. On suppose que* f *est* C^1 *sur* U *et que* g *est* C^1 *sur* $f(U) \subset V$*. Alors l'application* $g \circ f : U \longrightarrow G$ *est* C^1 *sur* U

2.6 Quelques exemples importants

Les exemples suivants sont plus que des simples illustrations de l'utilisation des notions et propriétés vues précédement. Ils sont en effet très souvent utilisés et ils constituent en quelque sorte, des briques de base pour le calcul de différentielles plus compliquées.

2.6.1 Applications à valeurs dans un espace produit

On se donne un entier $n > 1$. Soit E et $F_1, ..., F_n$ des e.v.n. et U un ouvert de E. On considère alors $f = (f_1, ..., f_n)$ une application définie sur U et à valeurs dans $F = F_1 \times ... \times F_n$. On a alors le résultat suivant qui est, comme le montre la preuve, une application du Théorème 2.1 compostion.

Théorème 2.2 - Différentielle d'applications à valeurs dans un produit
L'application f précédente est différentiable en $a \in U$ si et seulement si chaque f_k l'est ($k = 1, ..., n$). Si tel est le cas alors on a la formule

$$df(a) = (df_1(a), ..., df_n(a)) \tag{2.9}$$

Preuve du Théorème 2.2.
Pour chaque k entre 1 et n, on définit les applications $p_k : F \longrightarrow F_k$ et $i_k : F_k \longrightarrow F$ en posant pour tout $(x_1, ..., x_n) \in F$, $p_k(x_1, ..., x_n) = x_k$ (projection du produit sur son k-ième facteur) et $i_k(x_k) = (0, ..., 0, x_k, 0, ..., 0)$ (k-ième application coordonnée : 0 partout sauf en k-ième position).
Remarquons tout de suite que toutes ces applications sont linéaires continues et donc différentiables (sur F pour les p_k et sur F_k pour les i_k).
Supposons donc que f soit différentiable en a. Alors puisque p_k est différentiable sur F, que $f(a) \in F$ et que $f_k = p_k \circ f$, on déduit du Théorème 2.1 de différentiabilité de la composée que f_k est différentiable en a.
Réciproquement, si chaque f_k est différentiable en a, alors puisque $f = \sum_{k=1}^{n} i_k \circ f_k$ avec $f_k(a) \in F_k$ et i_k différentiable sur F_k, f est différentiable en a comme somme et composée d'applications qui le sont (Théorème 2.1 et Proposition 2.7) et $df(a) = \sum_{k=1}^{n} di_k(f_k(a)) \circ df_k(a) = \sum_{k=1}^{n} i_k \circ df_k(a)$ (c'est-à-dire (2.9)).

■

2.6.2 Différentielle de l'application inverse $Inv : u \mapsto u^{-1}$

⚠ Dans tout ce paragraphe, E et F sont deux espaces de Banach (e.v.n complets) vérifiant de plus, si la dimension est fine, $dim(E) = dim(F)$.
On rappelle la définition suivante :

Définition 2.8 - Isomorphisme
On dit que $f : E \longrightarrow F$ est un isomorphisme de E dans F si

1. $f \in \mathbb{B}(\mathrm{E},\mathrm{F})$.
2. *il existe* $g \in \mathbb{B}(\mathrm{F},\mathrm{E})$ *telle que* $g \circ f = id_E$ *et* $f \circ g = id_F$.

On note $\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$ *l'ensemble des isomorphismes de E dans F et si $E = F$ on note* $\mathrm{G}\mathbb{B}(\mathrm{E})$ *cet ensemble.*

Remarque 2.9
1) *Un isomorphisme est donc une application linéaire bijective de $E \longrightarrow F$ et bicontinue.*
2) *En dimension finie, la bijection réciproque est automatiquement linéaire et continue (facile).*
3) *En dimension infinie, c'est encore vrai mais la continuité de la bijection réciproque est non triviale à montrer (théorème de Banach).*
On a alors le résultat suivant qui est une généralisation d'un résultat bien connu sur les séries géométriques de raison absolument plus petite que 1 strictement.

Lemme 2.1
On note $id \in \mathrm{G}\mathbb{B}(\mathrm{E})$ l'endomorphisme identité et $f^k = \underbrace{f \circ \cdots \circ f}_{k \text{ termes}}$ avec la convention $f^0 = id$.

Si $f \in \mathbb{B}(\mathrm{E})$ avec $\|f\|_{\mathbb{B}(\mathrm{E})} < 1$ alors $(id - f) \in \mathrm{G}\mathbb{B}(\mathrm{E})$ et $(id - f)^{-1} = \sum_{k \geqslant 0} f^k$.

Preuve du Lemme 2.1.
Soit $f \in \mathbb{B}(\mathrm{E})$ vérifiant $\|f\|_{\mathbb{B}(\mathrm{E})} < 1$.
On a puisque f est continue $\sum_{k \geqslant 0} \left\|f^k\right\|_{\mathbb{B}(\mathrm{E})} \leqslant \sum_{k \geqslant 0} \|f\|^k_{\mathbb{B}(\mathrm{E})} = \dfrac{1}{1 - \|f\|_{\mathbb{B}(\mathrm{E})}}$ puisque $\|f\|_{\mathbb{B}(\mathrm{E})} < 1$ (série géométrique de raison $\|f\|_{\mathbb{B}(\mathrm{E})}$). Donc la série de terme générale f^k est normalement convergente, donc convergente.
Notant $S = \sum_{k \geqslant 0} f^k$, on voit que $Sf = fS$ et donc que $S(id - f) = (id - f)S = id$ ce qui montre que $(id - f)$ est inversible, d'inverse S.

■

On s'intéresse alors aux propriétés de l'application Inv définie par

$$\begin{array}{rccl} Inv : & \mathbb{B}(\mathrm{E},\mathrm{F}) & \longrightarrow & \mathbb{B}(\mathrm{F},\mathrm{E}) \\ & u & \mapsto & Inv(u) = u^{-1} \end{array} \tag{2.10}$$

On a la proposition suivante
.

Proposition 2.9 - Topologie de $\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$ - Continuité de l'inverse

(a) *L'ensemble* $\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$ *est un ouvert de* $\mathbb{B}(\mathrm{E},\mathrm{F})$.

(b) *L'application Inv est définie sur $D_{Inv} = \mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$ et $Im(Inv) = Inv(\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})) = \mathrm{G}\mathbb{B}(\mathrm{F},\mathrm{E})$*

(c) *L'application Inv est continue sur $D_{Inv} = \mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$.*

Preuve de la Proposition 2.9.
On note pour alléger, $\|\ \| = \|\ \|_{\mathbb{B}(\mathrm{E})}$.

(a) On peut supposer que $\mathrm{G\mathbb{B}(E,F)} \neq \emptyset$ (sinon la proposition est évidente).
Soit $u_0 \in \mathrm{G\mathbb{B}(E,F)}$.
On veut donc montrer que tout $u \in \mathrm{G\mathbb{B}(E,F)}$ suffisament proche de u_0 est encore un isomorphisme de E dans F. Mais puisque $u_0 \in \mathrm{G\mathbb{B}(E,F)}$ alors u est un isomorphisme de E dans F si et seulement si $(u_0)^{-1}u$ est un isomorphisme de E dans E.
D'autre part, on sait par le Lemme 2.1, que $(u_0)^{-1}u \in \mathrm{G\mathbb{B}(E)}$ dès que $\|id - (u_0)^{-1}u\| < 1$. Comme par continuité de $(u_0)^{-1}$ on a

$$\|id - (u_0)^{-1}u\| = \|(u_0)^{-1}(u_0 - u)\| \leqslant \|(u_0)^{-1}\|_{\mathbf{B(F,E)}} \ \|u_0 - u\|_{\mathbf{B(E,F)}} \tag{2.11}$$

il suffit de prendre u tel que $\|u - u_0\|_{\mathbf{B(E,F)}} < \frac{1}{\|(u_0)^{-1}\|_{\mathbf{B(F,E)}}}$. Ainsi pour tout $u_0 \in \mathrm{G\mathbb{B}(E,F)}$, la boule ouverte de centre u_0 et de rayon $\frac{1}{\|(u_0)^{-1}\|_{\mathbf{B(F,E)}}}$ est contenue dans $\mathrm{G\mathbb{B}(E,F)}$; donc $\mathrm{G\mathbb{B}(E,F)}$ est un ouvert.

(b) Evident.
(c) Montrons que pour tout $u_0 \in \mathrm{G\mathbb{B}(E,F)}$, l'application Inv est continue en u_0, c'est-à-dire que $\|u^{-1} - u_0^{-1}\|_{\mathbf{B(F,E)}}$ tend vers 0 lorsque $\|u - u_0\|_{\mathbf{B(E,F)}}$ tend vers 0.
Posant $f = id - (u_0)^{-1}u$, on a par (2.11), $\|f\| \leqslant \|u_0^{-1}\|_{\mathbf{B(F,E)}} \ \|u_0 - u\|_{\mathbf{B(E,F)}}$ et donc $\|f\|$ tend vers 0 quand $\|u - u_0\|_{\mathbf{B(E,F)}}$ tend vers 0 et $\|f\| < 1$ en prenant des u comme en **(a)**.
Ceci étant dit, on a alors

$$\begin{aligned}
u^{-1} - u_0^{-1} &= \{\left(u_0^{-1}u\right)^{-1} - id\}u_0^{-1} \\
&= \{\left(id - id + u_0^{-1}u\right)^{-1} - id\}u_0^{-1} \\
&= \{(id - f)^{-1} - id\}u_0^{-1} \\
&= \{\sum_{k\geqslant 0} f^k - id\}u_0^{-1} \\
&= \{\sum_{k\geqslant 1} f^k\}u_0^{-1}
\end{aligned}$$

et on en déduit alors $\left\|u^{-1} - u_0^{-1}\right\|_{\mathbf{B(F,E)}} \leqslant \sum_{k\geqslant 1} \|f\|^k \ \left\|u_0^{-1}\right\|_{\mathbf{B(F,E)}} = \left\|u_0^{-1}\right\|_{\mathbf{B(F,E)}} \ \dfrac{\|f\|}{1 - \|f\|} \longrightarrow 0$ quand $\|u_0 - u\|_{\mathbf{B(E,F)}} \longrightarrow 0$. Ce qui termine la preuve de la Proposition 2.9.

■

On vient donc de voir que l'application Inv définie précédement est continue de $\mathrm{G\mathbb{B}(E,F)}$, ouvert de l'e.v.n. $\mathrm{\mathbb{B}(E,F)}$, dans $\mathrm{G\mathbb{B}(F,E)}$, ouvert de l'e.v.n. $\mathrm{\mathbb{B}(F,E)}$. On peut donc considérer Inv comme fonction de $\mathrm{\mathbb{B}(E,F)}$ à valeurs dans $\mathrm{\mathbb{B}(F,E)}$ (dans ce cas $\mathrm{G\mathbb{B}(E,F)}$ est l'ensemble de définition de la fonction et $\mathrm{G\mathbb{B}(F,E)}$ son image). Le fait que $\mathrm{G\mathbb{B}(E,F)}$ soit un ouvert de $\mathrm{\mathbb{B}(E,F)}$, nous permet d'étudier la différentiabilité de l'application Inv. Si elle existe, cette différentielle sera donc un élément de $\mathrm{\mathbb{B}(\mathbb{B}(E,F),\mathbb{B}(F,E))}$. Plus précisément, on a le résultat suivant.

Théorème 2.3 - Régularité C^1 de l'inverse
L'application Inv est de classe C^1 sur $\mathrm{G\mathbb{B}(E,F)}$, et sa différentielle en $u \in \mathrm{G\mathbb{B}(E,F)}$ est donnée pour tout $h \in \mathrm{\mathbb{B}(E,F)}$ par

$$dInv(u)h = -u^{-1} \circ h \circ u^{-1} \tag{2.12}$$

Preuve du Théorème 2.3.
Pour $u \in \mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$, on revient à la définition de la différentielle de Inv en u.
Pour cela, on se donne $h \in \mathbb{B}(\mathrm{E},\mathrm{F})$ et on va considérer la différence $Inv(u+h) - Inv(u)$. Commençons par remarquer que, d'après le Lemme 2.1, pour tout $h \in \mathbb{B}(\mathrm{E},\mathrm{F})$ vérifiant $\|hu^{-1}\| < 1$ (il suffit que $\|h\| < 1/\|u^{-1}\|$), on a $(id + hu^{-1}) \in GB(F)$ et

$$Inv(u+h) = (u+h)^{-1} = \Big((id + hu^{-1})u\Big)^{-1} = u^{-1}\left(id + hu^{-1}\right)^{-1} = u^{-1}\left(id - hu^{-1} + \sum_{k\geqslant 2}(-hu^{-1})^k\right),$$

que l'on récrit sous la forme

$$Inv(u+h) - Inv(u) + u^{-1}hu^{-1} = u^{-1}\sum_{k\geqslant 2}(-hu^{-1})^k.$$

Donc, par définition de la différentielle pour avoir la relation (2.12), il suffit de montrer que $u^{-1}\sum_{k\geqslant 2}(-hu^{-1})^k = o(\|h\|)$. Mais, si h vérifie $\|h\| \leqslant 1$, on peut écrire que

$$\left\|u^{-1}\sum_{k\geqslant 2}(-hu^{-1})^k\right\| \leqslant \left\|u^{-1}\right\|\sum_{k\geqslant 2}\left\|hu^{-1}\right\|^k \leqslant \|h\|^2\sum_{k\geqslant 2}\left\|u^{-1}\right\|^{k+1} = o(\|h\|),$$

ce qui termine la preuve de la différentiabilité de Inv et de (2.12).
Il reste alors à montrer que $u \mapsto dInv(u)$ est continue sur $\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$.
Introduisons les applications défines par

$$\begin{array}{rcl} \varphi : \mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F}) & \longrightarrow & \mathbb{B}(\mathrm{F},\mathrm{E}) \times \mathbb{B}(\mathrm{F},\mathrm{E}) \\ u & \mapsto & (u^{-1}, u^{-1}) \end{array} \quad \text{et} \quad \begin{array}{rcl} \psi : \mathbb{B}(\mathrm{F},\mathrm{E}) \times \mathbb{B}(\mathrm{F},\mathrm{E}) & \longrightarrow & \mathbb{B}\Big(\mathbb{B}(\mathrm{E},\mathrm{F}), \mathbb{B}(\mathrm{F},\mathrm{E})\Big) \\ (v,w) & \mapsto & \psi(v,w) : h \mapsto -v \circ h \circ w \end{array}$$

Noter $\psi(v,w) \in \mathbb{B} = \mathbb{B}\Big(\mathbb{B}(\mathrm{E},\mathrm{F}), \mathbb{B}(\mathrm{F},\mathrm{E})\Big)$ puisqu'elle est linéaire et vérifie

$$\|\psi(v,w)\|_{\mathbb{B}} = \underset{\|h\|_{\mathbb{B}(\mathrm{E},\mathrm{F})} = 1}{Sup} \|\psi(v,w)h\|_{\mathbb{B}(\mathrm{F},\mathrm{E})} \leqslant \|v\|_{\mathbb{B}(\mathrm{E},\mathrm{F})}\ \|w\|_{\mathbb{B}(\mathrm{E},\mathrm{F})}$$

ce qui prouve aussi que ψ qui est bilinéaire, est continue.
Ainsi, φ est continue sur $\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$ car chaque application coordonnée l'est et vérifie $\varphi(\mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})) \subset \mathbb{B}(\mathrm{F},\mathrm{E}) \times \mathbb{B}(\mathrm{F},\mathrm{E})$, ensemble sur lequel ψ est continue.
Puisqu'on a $dInv = \psi \circ \varphi$, on en déduit par composition d'applications continues que $dInv$ l'est aussi.

■

FIN COURS 4

2.7 Variété tangente

E, F e.v.n, $U \subset E$ ouvert $f : U \longrightarrow F$.
On considère les ensembles de la forme $V = V(f) = \{(x,y) \in U \times F\ ;\ y = f(x)\}$.

Remarque
Lorsque $f : \mathbb{R} \longrightarrow \mathbb{R}$, l'ensemble V n'est rien d'autre que la courbe représentative de f.
Lorsque $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, l'ensemble V est une surface de $\mathbb{R}^3$.

Définition 2.9 Variété différentiable
Si f est différentiable sur U l'ensemble $V = V(f) = \{(x, y) \in U \times F \ ; \ y = f(x)\}$ est appelé variété différentiable (associée à f sur U).

La notion de tangente à une courbe (cas $E = F = \mathbb{R}$) se généralise de la manière suivante.

Définition 2.10 Vecteur tangent
Soit $M_0 = (x_0, f(x_0)) \in V$.
On dit que le vecteur $H \in E \times F$ est tangent à V en M_0 si

$$\exists (A_n)_n \subset V, \ \exists (\lambda_n) \subset \mathbb{R}^+ \text{ telles que } \left| \begin{array}{l} \lim\limits_{n \longrightarrow +\infty} A_n = M_0 \\ \lim\limits_{n \longrightarrow +\infty} \lambda_n (A_n - M_0) = H \end{array} \right.$$

On note $T_{M_0}(V)$ l'ensemble des vecteurs tangents à V en M_0.

Exercice
Déterminer $T_0(V)$ pour $f : \mathbb{R} \longrightarrow \mathbb{R}$ dans les cas suivants :
1. $f(x) = \sin(x)$ **2.** $f(x) = |x|$ **3.** $f(x) = E(x)$.

Remarque
1. $T_{M_0}(V)$ est un cône de sommet 0 (si $H \in T_{M_0}(V)$ alors $\alpha H \in T_{M_0}(V)$ pour tout $\alpha \geqslant 0$).
2. Si $H \in T_{M_0}(V)$ avec $H \neq 0$ alors $\lim\limits_{n \longrightarrow +\infty} \lambda_n = +\infty$.

On a alors le résultat suivant.

Théorème 2.4
Soit $M_0 \in V$.
Si f est différentiable en $x_0 \in U$ alors $T_{M_0}(V) = \{(h, k) \in E \times F \ ; \ k = df(x_0)h\}$.
On dit dans ce cas que $T_{M_0}(V)$ est une variété linéaire tangente à V en M_0.

Preuve du Théorème 2.4.
• Montrons que $T_{M_0}(V) \subset \{(h, k) \in E \times F \ ; \ k = df(x_0)h\}$
Soit $H = (h, k) \in T_{M_0}(V)$.
Donc il existe $(a_n) \subset U$, $(\lambda_n) \subset \mathbb{R}^+$ telles que Puisque f est différentiable en x_0, on a

$$\begin{array}{rl} & f(a_n) - f(x_0) = df(x_0)(a_n - x_0) + \|a_n - x_0\| \, \varepsilon(a_n - x_0) \text{ avec } \lim\limits_{0} \varepsilon = 0 \\ \Longrightarrow & \lambda_n (f(a_n) - f(x_0)) = df(x_0)(\lambda_n(a_n - x_0)) + \|\lambda_n(a_n - x_0)\| \, \varepsilon(a_n - x_0) \\ \Longrightarrow & k = df(x_0)h \text{ car } df(x_0) \text{ continue} \end{array}$$

• Montrons que $\{(h, k) \in E \times F \ ; \ k = df(x_0)h\} \subset T_{M_0}(V)$
Soit $(h, k) \in E \times F$ tels que $k = df(x_0)h$.
Soit $(t_n) \subset \mathbb{R}^+$ telle que $\lim\limits_{n \longrightarrow +\infty} t_n = 0^+$. Alors posant pour tout n, $a_n = x_0 + t_n h$ et $\lambda_n = \frac{1}{t_n}$, on a

$$(a_n, f(a_n))_n \subset V, (\lambda_n) \subset \mathbb{R}^+ \text{ et } \left| \begin{array}{l} \lim\limits_{n \longrightarrow +\infty} a_n = x_0 \\ \lim\limits_{n \longrightarrow +\infty} \lambda_n (a_n - x_0) = h \\ \lim\limits_{n \longrightarrow +\infty} \lambda_n (f(a_n) - f(x_0)) = \lim\limits_{n \longrightarrow +\infty} \dfrac{(f(x_0 + t_n h) - f(x_0))}{t_n} = df(x_0)h = k \end{array} \right.$$

puisque f différentiable en x_0. ■

Définition 2.11 Variété affine tangente
Soit $M_0 = (x_0, f(x_0)) \in V$ *avec* f *différentiable en* x_0.
L'ensemble $T^a_{M_0}(V) = \{(h,k) \in E \times F \; ; \; k = f(x_0 + df(x_0)(h - x_0)\}$ *est appelé variété affine tangente à* V *en* M_0.
Elle vérifie $T^a_{M_0}(V) = M_0 + T_{M_0}(V)$.

Exemples (f fonction différentiable)
1. Dans $\mathbb{R}^2$, V est donc une courbe et $T^a_{M_0}(V)$ la droite d'équation $(y = f(x_0 + df(x_0)(x - x_0))$ (tangente).
2. Dans $\mathbb{R}^3$, V est une surface et $T^a_{M_0}(V)$ le plan d'équation $(z = f(x_0, y_0) + \dfrac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + $ $+ \dfrac{\partial f}{\partial y}(x_0, y_0)(y - y_0))$ (plan tangent).

$\boxed{\text{FIN DU CHAPITRE 2}}$

Chapitre 3

Théorèmes des accroissements finis - Applications.

Les résultats de ce chapitre sont essentiels tant leur utilisation est fréquente. Ils constituent non seulement des outils de base pour ce cours mais également pour l'analyse en général.

3.1 Théorèmes des accroissements finis (T.A.F)

Commençons par rappeler ce que vous devez déja connaître (cf cours de première année).

Théorème 3.1 - Cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$
Soit $a, b \in \mathbb{R}$, $a < b$ et $f : \mathbb{R} \longrightarrow \mathbb{R}$ continue sur $[a, b]$ et dérivable sur $]a, b[$. Alors $\exists c \in]a, b[$ tel que

$$f^{'}(c) = \frac{f(b) - f(a)}{b - a}, \tag{3.1}$$

ou de manière équivalente, il existe $t_0 \in]0, 1[$ tel que

$$f^{'}(t_0 a + (1 - t_0)b)) = \frac{f(b) - f(a)}{b - a}.$$

Lorsque $f(a) = f(b)$, on obtient $f'(c) = 0$: c'est le théorème de Rolle.

On déduit facilement du Théorème 3.1 la généralisation suivante moyennant l'introduction de la notation : $[a, b] = \{ta + (1 - t)b,\ t \in [0, 1]\}$, $]a, b[= \{ta + (1 - t)b,\ t \in]0, 1[\}$.

Corollaire 3.1.1 - Cas des applications de E dans $\mathbb{R}$
Soit E un e.v.n, U un ouvert de E et $a, b \in E$ tels que $[a, b] \subset U$. Si $f : U \longrightarrow \mathbb{R}$ est continue sur $[a, b]$ et différentiable sur $]a, b[$ alors il existe $t_0 \in]0, 1[$ tel que

$$f(b) - f(a) = df((1 - t_0)a + t_0 b))(b - a).$$

Noter que dans ce cas, (3.1) n'a pas de sens puisqu'on ne peut pas diviser par (le vecteur) $b - a$.

Preuve du Corollaire 3.1.1
Appliquer le Théorème 3.1 à $\psi : t \in [0, 1] \mapsto f((1 - t)a + tb)$ (décomposer pour calculer sa diff.).

■

On peut alors se demander si ce résultat se généralise sous sa forme actuelle aux applications $f : E \longrightarrow F$, $dim(F) > 1$. La réponse est **non** comme le montre l'exemple suivant.

Exemple 3.1
Voyons pourquoi ce résultat ne se généralise pas directement aux applications à valeurs dans F, $dim(F) > 1$.
Supposons d'abord qu'un funambule (qui se déplace donc sur un fil...) parte d'un point A à l'instant a et notons $f(t)$ sa position à l'instant $t \geqslant a$ (donc $f : \mathbb{R} \longrightarrow \mathbb{R}$). On a alors l'évidence suivante : si à un instant $b > a$ le funambule est revenu en A, alors il a fait demi-tour, et donc il a annulé sa vitesse à un moment donné. D'un point de vue mathématique cette affirmation se traduit par : si $f(a) = f(b)$ alors il existe $c \in]a, b[$ tel que $f'(c) = \frac{f(b)-f(a)}{b-a} = 0$ (f' est la vitesse du funambule).
Si maintenant le funambule descend de son fil, partant d'un point, il n'est plus obligé de faire demi-tour pour y revenir, il suffit en effet que son trajet soit une boucle : ici sa position $f : \mathbb{R} \longrightarrow \mathbb{R}^2$!.
Ecrivons cela de manière plus rigoureuse :
Soit la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}^2$ définie par $f(t) = (\cos(t), \sin(t)$. Elle est continue sur $[0, 2\pi]$, dérivable sur $]0, 2\pi[$ mais $f^{'}(c) = f(2\pi) - f(0) \Longleftrightarrow \begin{cases} -\cos(c) = 0 \\ \sin(c) = 0 \end{cases}$ qui n'admet pas de solution et donc $\forall c \in]0, 2\pi[,\ f^{'}(c) \neq f(2\pi) - f(0)$.
La généralisation du Théorème 3.1 aux fonctions à valeurs dans un e.v.n admet une autre formulation connue sous le nom d'inégalité des accroissements finis.

Théorème 3.2 - Inégalité des accroissements finis
Soit $a, b \in \mathbb{R}$ tels que $a < b$, F un e.v.n et $f : \mathbb{R} \longrightarrow F$, $g : \mathbb{R} \longrightarrow \mathbb{R}$ deux fonctions continues sur $[a, b]$ et dérivables sur $]a, b[$.

Si pour tout $t \in]a, b[$, $\|f^{'}(t)\| \leqslant g^{'}(t)$ alors

$$\|f(b) - f(a)\| \leqslant g(b) - g(a). \tag{3.2}$$

Preuve du Théorème 3.2
On va montrer que pour tout réel $\varepsilon > 0$ donné, on a $\forall t \in [a, b]$ l'inégalité

$$\|f(t) - f(a)\| \leqslant g(t) - g(a) + (t - a)\varepsilon + \varepsilon. \tag{3.3}$$

Il suffira alors de prendre $t = b$ puis $\varepsilon \downarrow 0$ pour avoir le Théorème 3.2 grace à la continuité de f et g.
Soit donc $\varepsilon > 0$ fixé.
⋄ Commençons par remarquer que l'inégalité (3.3) est vérifié pour $0 < t - a$ suffisament petit.
En effet,
- f est continue en a donc : $\exists \alpha > 0$, $\forall t \in \mathbb{R}$, $0 \leqslant t - a \leqslant \alpha \Longrightarrow \|f(t) - f(a)\| \leqslant \frac{\varepsilon}{2}$.
- $g^{'} \geqslant 0$ sur $]a, b[$ donc $g(t) - g(\tau) \geqslant 0$ pour t et τ vérifiant $a < \tau < t < b$, soit par continuité en a, $g(t) - g(a) \geqslant 0$, $\forall t \in [a, b[$.

Ainsi, on obtient qu'il existe $\alpha > 0$ tel que pour tout $t \in [a, a + \alpha[$

$$\|f(t) - f(a)\| \leqslant \frac{\varepsilon}{2} < \varepsilon \leqslant g(t) - g(a) + (t - a)\varepsilon + \varepsilon$$

On en conclut que l'inégalité est vérifié pour t assez proche de a.

$\diamond$ Montrons alors que l'inégalité est vraie sur $[a, b]$ en entier.
Introduisons pour cela la fonction continue (par composition) $\psi : [a, b] \longrightarrow \mathbb{R}$ définie sur $[a, b]$ par

$$\psi(t) = \|f(t) - f(a)\| - (g(t) - g(a) + (t-a)\varepsilon + \varepsilon)\,,$$

et définissons des éléments t de $[a, b]$ pour lesquels l'inégalité (3.3) n'est pas vérifiée, soit

$$Z_\varepsilon = \{t \in [a, b],\ \psi(t) > 0\}\,.$$

Montrons que cet ensemble est vide.
On a $Z_\varepsilon = \psi^{-1}(]0, +\infty[)$ ouvert (image réciproque d'un ouvert par une application continue).
Supposons que $Z_\varepsilon \neq \emptyset$.

Alors $Z_\varepsilon \subset [a, b]$ et est non vide donc il possède une borne inférieure c dont on peut dire trois choses :

(i) $c > a$ puisque (3.3) est vérifiée pour $t \in [a, a+\alpha[$.

(ii) $c \notin Z_\varepsilon$ car Z_ε est ouvert. En effet si c appartenait à Z_ε, il existerait une boule de centre c entièrement contenue dans Z_ε donc des $t \in Z_\varepsilon$ pour lesquels $a < t < c$ et $t \in Z_\varepsilon$ ce qui contredirait le fait que c est borne inférieure de Z_ε.

(iii) $c < b$ car sinon on aurait $Z_\varepsilon = \{b\}$ qui serait fermé.

Alors de *(i)* et *(iii)*, on déduit que $a < c < b$ et donc par hypothèse que

$$\|f'(c)\| \leqslant g'(c). \tag{3.4}$$

Or, il existe $\beta > 0$ tel que pour tout t vérifiant $0 < t - c \leqslant \beta$

$$\left\|\frac{f(t)-f(c)}{t-c}\right\| - \left\|f'(c)\right\| \leqslant \left\|\frac{f(t)-f(c)}{t-c} - f'(c)\right\| \leqslant \frac{\varepsilon}{2}$$

$$\text{et} \quad g'(c) - \frac{g(t)-g(c)}{t-c} \leqslant \left|\frac{g(t)-g(c)}{t-c} - g'(c)\right| \leqslant \frac{\varepsilon}{2},$$

soit $\left\|\dfrac{f(t)-f(c)}{t-c}\right\| - \dfrac{\varepsilon}{2} \leqslant \left\|f'(c)\right\|$ et $g'(c) \leqslant \dfrac{g(t)-g(c)}{t-c} + \dfrac{\varepsilon}{2}$.
Compte tenu de (3.4), on en déduit que pour $c \leqslant t \leqslant \beta + c$ on a

$$\|f(t) - f(c)\| \leqslant g(t) - g(c) + (t-c)\varepsilon. \tag{3.5}$$

D'autre part, on a vu en *(ii)* que $c \notin Z_\varepsilon$, et puisque $a < c < b$, on a donc

$$\|f(c) - f(a)\| \leqslant g(c) - g(a) + (c-a)\varepsilon + \varepsilon. \tag{3.6}$$

Utilisant (3.5) et (3.6), on obtient alors pour $c \leqslant t \leqslant \beta + c$

$$\begin{aligned}\|f(t) - f(a)\| &\leqslant \|f(t) - f(c)\| + \|f(c) - f(a)\| \\ &\leqslant g(t) - g(c) + (t-c)\varepsilon + g(c) - g(a) + (c-a)\varepsilon + \varepsilon \\ &\leqslant g(t) - g(a) + (t-a)\varepsilon + \varepsilon.\end{aligned}$$

Ainsi les t qui vérifient $c \leqslant t \leqslant \beta + c$ satisfont (3.3) et ne sont donc pas dans Z_ε. Comme on sait déja que $[a, c[\cap Z_\varepsilon = \emptyset$, on en déduit que $[a, c+\beta] \cap Z_\varepsilon = \emptyset$ avec $\beta > 0$. Ceci entraine nécessairement que la borne inférieure de Z_ε est supérieure ou égale à $c + \beta$ ce qui contredit le fait que c est le plus grand des minorants. On obtient donc ainsi une contradiction avec l'hypothèse de départ et donc $Z_\varepsilon = \emptyset$. ■

FIN COURS 5

On peut alors déduire du Théorème 3.2 plusieurs corollaires qui sont au moins aussi importants que ce théorème puisque ce sont eux que nous utiliserons dans la pratique.

Convention. Soit $A \subset \mathbb{R}$ non vide
Si A est non majorée, on pose $\overline{Sup}(A) = +\infty$.
Si A est mojorée, on pose $\overline{Sup}(A) = Sup(A) \in \mathbb{R}$.
On conviendra de noter $Sup(A)$ la quantité $\overline{Sup}(A)$ que l'on continuera à appelé borne supérieure de A.
Noter qu'avec cette convention classique, toute partie non vide de $\mathbb{R}$ admet une borne supérieure.

Corollaire 3.2.1 - Cas des applications de $\mathbb{R}$ dans F
Soit $a, b \in \mathbb{R}$, $a < b$, F e.v.n et $f : \mathbb{R} \longrightarrow F$, continue sur $[a, b]$ et dérivable sur $]a, b[$. Alors on a

$$\|f(b) - f(a)\| \leqslant (b - a) \underset{t \in]a,b[}{Sup} \|f'(t)\|.$$

Preuve du Corollaire 3.2.1

• Si $\underset{]a,b[}{Sup} \|f'\| = +\infty$ c'est évident (inégalité dégénérée).

• Sinon $\underset{]a,b[}{Sup} \|f'\| = C \in \mathbb{R}^+$ et on peut appliquer le Théorème 3.2 avec $g(t) = Ct$. ■

On peut alors énoncer l'inégalité des accroissements finis sous sa forme la plus générale.

Corollaire 3.2.2 - Cas des applications de E dans F
Soit E et F deux e.v.n, U un ouvert de E et $f : U \longrightarrow F$ différentiable.
Alors pour tout $x, y \in U$ tels que $[x, y] \subset U$ on a

$$\|f(x) - f(y)\|_F \leqslant \|x - y\|_E \underset{z \in]x,y[}{Sup} \|df(z)\|_{\mathbf{B(E,F)}} \tag{3.7}$$

Preuve du Corollaire 3.2.2

On fixe $x, y \in U$ tels que $[x, y] \subset U$.

• Si $\underset{z \in]x,y[}{Sup} \|df(z)\|_{\mathbf{B(E,F)}} = +\infty$: évident.

• Sinon $\underset{z \in]x,y[}{Sup} \|df(z)\|_{\mathbf{B(E,F)}} = C \in \mathbb{R}^+$. On définit alors ψ en posant pour tout $t \in [0, 1]$, $\psi(t) = f((1 - t)x + ty)$ qui est dérivable sur $]0, 1[$ (par composition). De plus $\psi'(t) = df((1 - t)x + ty)(y - x)$ pour tout $t \in]0, 1[$ et donc puisque $\|df(z)\|_{\mathbf{B(E,F)}} \leqslant C$ pour tout $z \in [x, y]$, on a

$$\left\|\psi'(t)\right\|_F = \|df((1 - t)x + ty)(y - x)\|_F \leqslant \|x - y\|_E \|df((1 - t)x + ty)\|_{\mathbf{B(E,F)}}. \leqslant C \|x - y\|_E$$

On peut donc appliquer le Théorème 3.2 aux fonctions ψ et $g : t \mapsto C \|x - y\|_E t$ sur $[0, 1]$ puisque

$\forall t \in]0,1[,\ \left\|\psi^{'}(t)\right\|_F \leqslant g^{'}(t)$. On en déduit alors que $\|\psi(1)-\psi(0)\|_F \leqslant g(1)-g(0)$ c'est-à-dire (3.7). ■

Remarque 3.1
Noter que si U est un ouvert convexe pour tout $x, y \in U$, on a $[x,y] \subset U$ et l'hypothèse du Corollaire 3.2.2 est systématiquement vérifiée.
Nous allons terminer ce chapitre en voyant quelques conséquences des résultats sur les accroissements finis.

3.2 Applications

3.2.1 Caractérisation des applications différentiables Lipschitziennes

Voyons tout de suite une première conséquence du Corollaire 3.2.2. Pour cela nous avons besoin de la définition suivante que vous connaissez très probablement déja.

Définition 3.1 - Application Lipschitzienne
Soit E et F deux e.v.n, U un ouvert de E et $f : E \longrightarrow F$ définie sur U. On dit que f est **Lipschitzienne** *de rapport $K>0$ sur U, s'il existe $K>0$ tel que*

$$\forall x, y \in U, \quad \|f(x)-f(y)\|_F \leqslant K\,\|x-y\|_E\,. \tag{3.8}$$

Si la constante vérifie de plus $K<1$, on dit que f est **contractante**.

On a alors la caractérisation suivante pour les fonctions différentiables et Lipschitziennes sur un ouvert convexe.

Proposition 3.1 - Application différentiable Lipschitzienne sur un convexe
Soit E et F deux e.v.n, U un ouvert **convexe** *de E et $f : U \longrightarrow F$ différentiable.*
Alors les deux propriétés suivantes sont équivalentes :

1. *f est Lipschitzienne sur U*
2. *df est bornée sur U (i.e. $\exists C>0,\ \underset{z\in U}{Sup}\,\|df(z)\|_{\mathbf{B(E,F)}} \leqslant C$).*

Preuve de la Proposition 3.1
$\Longleftarrow$ / Conséquence du Corollaire 3.2.2 (en utilisant la convexité de U) et de la Définition 3.1.
$\Longrightarrow$ / Supposons f Lipschitzienne.
Soit $h \in E$ et $x_0 \in U$ donnés. Puisque U est ouvert, pour $t \in \mathbb{R}$ assez petit, $x_0 + th \in U$. Donc on peut prendre $x = x_0 + th$ et $y = x_0$ dans (3.8), ce qui donne

$$\frac{\|f(x_0+th)-f(x_0)\|_F}{|t|} \leqslant K\,\|h\|_E\,.$$

Mais f est différentiable sur U, donc faisant tendre t vers 0, on obtient $\|df(x_0)h\|_F \leqslant K\,\|h\|_E$ pour tout $h \in E$ et tout $x_0 \in U$. On trouve ainsi pour tout $x_0 \in U$

$$\|df(x_0)\|_{\mathbf{B(E,F)}} = \underset{h\in E}{Sup}\,\frac{\|df(x_0)(h)\|_F}{\|h\|_E} \leqslant K,$$

et donc *2.* en prenant $\underset{x_0 \in U}{Sup}$.

■

3.2.2 Applications de différentielle nulle sur un ouvert connexe

Une autre conséquence immédiate est la suivante.

Définition 3.2 - Ouvert connexe
Un ouvert U d'un e.v.n E est **connexe** *s'il n'est pas union de deux ouverts non vides disjoints.*
Pour ce qui nous intéresse, on peut également adopter la définition équivalente (connexité par arcs) :
Un ouvert U d'un e.v.n E est **connexe** *si pour tout $a, b \in U$, il existe des points $a_1, ..., a_N \in U$ tels que $a_1 = a$, $a_N = b$ et $[a_i, a_{i+1}] \subset U$ pour tout $i = 1, ..., N-1$.*
Noter que l'entier N dépend à priori des points a et b considérés.

Remarque 3.2
1) *Sur $\mathbb{R}$, les seuls ouverts connexes sont les intervalles ouverts.*
2) *En particulier, tout ouvert convexe est connexe.*
3) *De manière générale, la connexité par arcs entraîne la connexité. La réciproque, qui n'est pas toujours vraie, l'est toutefois pour $U \subset E$ ouvert avec E e.v.n.*
On peut alors énoncer la propriété suivante.

Proposition 3.2 - Différentielle nulle sur un connexe
Soit E et F deux e.v.n, U un ouvert **connexe** *de E et $f : U \longrightarrow F$ différentiable.*
Si la différentielle de f est nulle sur U alors f est constante sur U.

Preuve du Proposition 3.2
On utilise les notations de la Définition 3.2.
Soient $a, b \in U$. Fixons i entre 1 et $N-1$. On a $I_i = [a_i, a_{i+1}] \subset U$ et on peut donc appliquer le Corollaire 3.2.2 à f sur I_i et puisque $df = 0$, on trouve $f(a_i) = f(a_{i+1})$. Ainsi on en déduit $f(a) = f(b)$ et donc que f est constante sur U puisque ceci est valable pour $a, b \in U$ arbitraires.

■

Remarque 3.3
Noter que la Proposition 3.2 ne fait qu'étendre la propriété que vous connaissez déja lorsque $f : \mathbb{R} \longrightarrow \mathbb{R}$: si f est continue sur $[a, b]$ et dérivable, de dérivée nulle sur $]a, b[$ alors elle est constante sur $[a, b]$. On voit en particulier que, là encore (comme pour la Remarque 3.1 partie 2.), l'hypothèse géométrique "U connexe" était contenue mais cachée en dimension 1. En effet si f est dérivable de dérivée nulle sur deux intervalles ouverts disjoints (donc sur un ouvert non connexe), on peut seulement en déduire qu'elle est constante par morceaux. Il n'y a aucune raison pour que les constantes soient les mêmes (exemple : $f = 0$ sur $[0, 1]$ et $f = 1$ sur $[2, 3]$ avec $U =]0, 1[\cup]2, 3[$).

3.2.3 Caractérisation des fonctions C^1

Si on ne peut espérer caractériser la différentiabilité en terme de dérivées partielles, on a néanmoins une caractérisation pour la classe C^1. Celle-ci s'obtient grace à l'inégalité des accroissements finis, et elle est d'une importance pratique considérable. Elle s'énonce de la manère suivante.

Proposition 3.3 - Caractérisation des fonctions C^1
Soit F e.v.n, $U \subset \mathbb{R}^n$ un ouvert et $f : \mathbb{R}^n \longrightarrow F$ une application définie sur U. Les deux propriétés suivantes sont équivalentes :

1. *f est de classe C^1 sur U (voir Définition 2.7)*
2. *f admet sur U des dérivées partielles continues.*

Preuve du Proposition 3.3
• *1.$\Longrightarrow$2.* est directe puisque $\frac{\partial f}{\partial x_i}(x) = df(x)e_i$.
• *2.$\Longrightarrow$1.* Soit $a \in U$ tel que f admette en a des dérivées partielles continues.
◇ Commençons par mq si f est différentiable en a alors, si ses DPs sont continues, elle est C^1.
Si f est différentiable en a alors on a $df(a)h = \displaystyle\sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\ h_j$ pour tout $h \in \mathbb{R}^n$.
Par conséquent, pour tout $x, a \in U$ et tout $h \in \mathbb{R}^n$

$$\begin{aligned} \|df(x)h - df(a)h\|_F &= \left\| \sum_{j=1}^{n} \left(\frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a) \right) h_j \right\|_F \\ &\leqslant \|h\|_\infty \sum_{j=1}^{n} \left\| \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a) \right\|_F \end{aligned}$$

et donc $\|df(x) - df(a)\|_{\mathbf{B}(\mathbb{R}^n,F)} = \underset{h\in\mathbb{R}^n}{Sup} \dfrac{\|(df(x) - df(a))(h)\|_F}{\|h\|_\infty} \leqslant \displaystyle\sum_{j=1}^{n} \left\| \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a) \right\|_F$ ce qui permet de conclure que $\lim\limits_{x\longrightarrow a} \|df(x) - df(a)\|_{\mathbf{B}(\mathbb{R}^n,F)} = 0$. Ainsi, si les dérivées partielles sont continues en a et que f y est différentiable alors df est continue en a.
◇ Toute la question est alors de montrer que f est différentiable en a et tout se ramène donc finalement à montrer qu'au voisinage de $x = a$ (seul candidat possible pour la différentielle en a)

$$\left\| f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j) \right\|_F = o(\|x - a\|_1) \tag{3.9}$$

On munit $\mathbb{R}^n$ de la norme $\|\ \|_1$.
Puisque $a \in U$ ouvert, il existe $\eta_0 > 0$ tel que $B_{\eta_0}(a) \subset U$ et donc pour tout $x \in B_{\eta_0}(a)$ on peut donc écrire

$$\begin{aligned} f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j) &= f(x_1, \cdots, x_n) - f(a_1, x_2, \cdots, x_n) - \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) \\ &\quad + f(a_1, x_2, \cdots, x_n) - f(a_1, a_2, \cdots, x_n) - \frac{\partial f}{\partial x_2}(a)(x_2 - a_2) \\ &\quad + \\ &\quad \vdots \\ &\quad + f(a_1, \cdots, a_{n-1}, x_n) - f(a_1, \cdots, a_n) - \frac{\partial f}{\partial x_n}(a)(x_n - a_n) \end{aligned} \tag{3.10}$$

Puisque pour tout $x \in B_{\eta_0}(a)$ on a $[a,x] \subset B_{\eta_0}(a) \subset U$, on peut introduire la fonction g définie pour tout $t \in [0,1]$ par $g(t) = f(tx_1 + (1-t)a_1, x_2, \cdots, x_n) - t\,\dfrac{\partial f}{\partial x_1}(a)(x_1 - a_1)$.
Par composition, cette fonction est continue sur $[0,1]$ et dérivable sur $]0,1[$ de dérivée

$$g^{'}(t) = \left(\frac{\partial f}{\partial x_1}(tx_1 + (1-t)a_1, x_2, \cdots, x_n) - \frac{\partial f}{\partial x_1}(a)\right)(x_1 - a_1). \tag{3.11}$$

Or par continuité de $\frac{\partial f}{\partial x_1}$ en a (c'est l'hypothèse), on a

$$\forall \varepsilon > 0,\ \exists \eta_1 > 0,\ \|x-a\|_1 \leqslant \eta_1 \Longrightarrow \left\|\frac{\partial f}{\partial x_1}(x) - \frac{\partial f}{\partial x_1}(a)\right\|_F \leqslant \varepsilon.$$

En particulier puisque pour $t \in]0,1[$, on a :
$\|(tx_1 + (1-t)a_1, x_2, \cdots, x_n) - (a_1, a_2, \cdots, a_n)\|_1 = \|(t(x_1 - a_1), x_2 - a_2, \cdots, x_n - a_n)\|_1 \leqslant \eta_1$
on en déduit pour (3.11) que pour tout $t \in]0,1[$,

$$\left\|g^{'}(t)\right\|_F \leqslant \varepsilon |x_1 - a_1|.$$

On peut donc appliquer le Corollaire 3.2.1 à g sur $[0,1]$, et on obtient ainsi que $\|g(1) - g(0)\|_F \leqslant \varepsilon |x_1 - a_1|$, soit en revenant à f

$$\left\|f(a_1, x_2, \cdots, x_n) - f(x_1, x_2, \cdots, x_n) - \frac{\partial f}{\partial x_1}(a)(x_1 - a_1)\right\|_F \leqslant \varepsilon |x_1 - a_1|.$$

On fait de même pour les autres termes de (3.10) et on obtient ainsi que pour tout $\varepsilon > 0$, il existe $\eta = Min(\eta_0, \eta_1, \cdots, \eta_n)$ (où η_i correspond au η_1 pour i-ème terme) tel que

$$\|x-a\|_1 \leqslant \eta \Longrightarrow \left\|f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j)\right\|_F \leqslant \varepsilon \|x-a\|_1$$

Ceci montre (3.9) et conclut la preuve de la Proposition 3.3.

∎

Remarque 3.4
La Proposition 2.6 est un corollaire de la Proposition 3.3 que nous venons de montrer.
FIN COURS 6

3.2.4 Suite d'applications différentiables

Commençons ce paragraphe par deux petits rappels.
Rappels
Soit $(f_n)_n$ suite d'applications définies sur $U \subset E$ e.v.n et à dans F e.v.n.
(1) On dit que $f_n \longrightarrow f$ uniformément sur U si $\lim\limits_{n \longrightarrow \infty} \underset{x \in U}{Sup} \|f_n(x) - f(x)\|_F = 0$.
(2) Si $(f_n)_n \subset C^0(U)$ et que $f_n \longrightarrow f$ uniformément sur U alors $f \in C^0(U)$.

Une autre conséquence de l'inégalité des accroissements finis est la suivante.

Théorème 3.3 - Suite d'applications différentiables
Soit E et F deux Banach, U un ouvert **convexe** *de E et $(f_n)_n$ une suite d'applications de $U \longrightarrow F$ toutes différentiables.*
On suppose que

(H1) il existe une application $f : U \longrightarrow F$ telle que $(f_n)_n$ converge simplement vers f sur U i.e. $\forall x \in U$, $\lim_{n \longrightarrow \infty} f_n(x) = f(x)$;

(H2) il existe une application $g : U \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$ telle que $(df_n)_n$ converge uniformément vers g sur tout borné de U i.e. pour tout $B \subset U$ borné,

alors on a

(i) (f_n) converge uniformément vers f sur tout borné de U ;

(ii) la limite f est différentiable sur U et $df = g$.

Preuve du Théorème 3.3
Remarquons que si $B \subset U$ est borné, alors on peut le supposer en plus ouvert et convexe sans restreindre la généralité puisque U est lui même ouvert et convexe. En effet, si tel n'est pas le cas pour $a \in B$ arbitraire, il existe $\rho > 0$ tel que $B \subset B_\rho(a)$, et il suffit alors de remplacer B par $U \cap B_\rho(a)$ qui est ouvert et convexe avec $B \subset U \cap B_\rho(a) \subset U$.
• *Preuve de (i).*
Soit $B \subset U$ borné (i.e. $\exists R > 0$, $B \subset B_R(0)$) et soit $a \in B$ fixé. On a pour tout $x \in B$,

$$\|f_n(x) - f_p(x)\|_F \leqslant \|f_n(a) - f_p(a)\|_F + \|f_n(x) - f_p(x) - f_n(a) + f_p(a)\|_F . \tag{3.12}$$

Puisque par *(H2)*, $(df_n)_n$ converge uniformément sur B, elle est uniformément de Cauchy sur B et donc pour tout $\varepsilon > 0$, il existe $N \geqslant 1$ tel que pour tout $n, p \geqslant N$,

$$\underset{y \in B}{Sup} \|df_n(y) - df_p(y)\|_{\mathbb{B}(\mathrm{E},\mathrm{F})} \leqslant \varepsilon. \tag{3.13}$$

Donc fixant n et p tels que $n, p \geqslant N$, on peut appliquer le Corollaire 3.2.2 à la fonction $x \mapsto f_n(x) - f_p(x) - f_n(a) + f_p(a)$ sur $[a, x] \subset B$ (qui est ouvert convexe). La différentielle de cette fonction vérifie (3.13) et on en déduit donc que

$$\|f_n(x) - f_p(x) - f_n(a) + f_p(a)\|_F \leqslant \|x - a\|_E \ \varepsilon \leqslant 2R\, \varepsilon. \tag{3.14}$$

D'autre part, par *(H1)*, $(f_n(a))_n$ converge et elle est donc de Cauchy. Ainsi quitte à choisir un N plus grand, on peut supposer que $n, p \geqslant N$ entraine

$$\|f_n(a) - f_p(a)\|_F \leqslant \varepsilon. \tag{3.15}$$

Revenant alors à (3.12), on déduit ainsi de (3.14), (3.15) que pour tout $\varepsilon > 0$, il existe $N \geqslant 1$ tel que pour tout $n, p \geqslant N$ $\|f_n(x) - f_p(x)\|_F \leqslant (1 + 2R)\varepsilon$ et faisant tendre p vers $+\infty$, on trouve $\underset{x \in B}{Sup} \|f_n(x) - f(x)\|_F \leqslant (1 + 2R)\varepsilon$ c'est-à-dire l'affirmation *(i)* du Théorème.
• *Preuve de (ii).*

On conserve les notations du début de la preuve. On écrit, avec n que l'on va fixer, et pour tout $h \in E$ tel que $a + h \in B_\eta(a)$ avec $\eta > 0$ à choisir

$$\begin{aligned}\|f(a+h) - f(a) - g(a)h\|_F \leqslant\ & \|f(a+h) - f_n(a+h) - f(a) + f_n(a)\|_F + \|f_n(a+h) - f_n(a) - df_n(a)h\|_F \\ & + \|df_n(a)h - g(a)h\|_F\end{aligned}$$

Considérons alors successivement les trois termes du membre de droite de cette inégalité.
Pour cela, on se donne $\varepsilon > 0$ arbitraire fixé, et on raisonne à n fixé ($n = N$ ds ce qui précède).
Pour le premier des trois termes, si on prend $x = a + h$ dans (3.14) et qu'on fait tendre p vers $+\infty$, on déduit de (H_1) que

$$\|f(a+h) - f_n(a+h) - f(a) + f_n(a)\|_F \leqslant \|h\|_E\ \varepsilon.$$

Concernant le second terme, puisque f_n est différentiable, il existe $\eta > 0$ tel que pour $\|h\|_E < \eta$, on peut écrire que

$$\|f_n(a+h) - f_n(a) - df_n(a)h\|_F \leqslant \|h\|_E\ \varepsilon.$$

Enfin quitte à prendre n toujours fixé mais suffisament grand, on à

$$\|df_n(a)h - g(a)h\|_F \leqslant \|df_n(a) - g(a)\|_{\mathbf{B(E,F)}}\ \|h\|_E \leqslant \underset{a\in B}{Sup}\, \|df_n(a) - g(a)\|_{\mathbf{B(E,F)}}\ \|h\|_E \leqslant \|h\|_E\ \varepsilon,$$

d'après (H_2).
Ainsi, en recollant les morceaux, on trouve que pour tout $\|h\|_E < \eta$,

$$\|f(a+h) - f(a) - g(a)h\|_F \leqslant 3\,\|h\|_E\ \varepsilon$$

ce qui permet de conclure.

■

Remarque 3.5
Si dans le théorème précédent on suppose les $(f_n)_n$ de clssse C^1 sur U alors f l'est aussi (voir rappel (2)).

FIN DU CHAPITRE 3

Chapitre 4

Différentielles d'ordre supérieur - Formules de Taylor - Extrema.

4.1 Différentielle seconde

4.1.1 Définition - Théorème de Schwarz

Définition 4.1
Soit E et F e.v.n, $U \subset E$ ouvert et $f : U \longrightarrow F$. On dit que f est 2 fois différentiable en $x_0 \in U$ si :

(i) *f est différentiable sur un voisinage ouvert $V \subset U$ de x_0 ;*

(ii) *L'application $df : V \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$ est différentiable en x_0.*

On note $d^2f(x_0)$ la différentielle seconde de f en x_0 (c'est-à-dire $d^2f(x_0) = (d(df))(x_0)$).
On dit que f est 2 fois différentiable sur U si elle l'est en chaque point de U.

Remarque 4.1
1) *Si f est 2 fois différentiable en x_0, puisque $df : V \subset E \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$, on a :*

- *$d^2f(x_0)$ linéaire de $E \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$ ie. pour tout $h \in E$, $d^2f(x_0)h$ est linéaire de E dans F ; ainsi $(h,k) \in E^2 \mapsto (d^2f(x_0)h)k$ est une application* BILINÉAIRE
- *compte tenu de cela, on convient de noter $d^2f(x_0)(h,k)$ la quantité $(d^2f(x_0)h)k$ ce qui signifie qu'on identifie l'application $d^2f(x_0) \in \mathbb{B}\left(E; \mathbb{B}(\mathrm{E}, \mathrm{F})\right)$ à une application bilinéaire, soit $d^2f(x_0) \in \mathbb{B}^2(E, F)$.*

Compte tenu de cette identification, on conviendra de noter $d^2f(x_0)(h,k)$ la quantité $(d^2f(x_0)h)k$ et $d^2f(x_0)(h, .)$ la quantité $d^2f(x_0)h$.
2) *En fait, on peut montrer que $\mathbb{B}\left(E; \mathbb{B}(\mathrm{E}, \mathrm{F})\right) \approx \mathbb{B}^2(E, F)$ et même, plus généralement que pour tout $p \in \mathbb{N}^*$, $\mathbb{B}\left(E; \mathbb{B}^p(E, F)\right) \approx \mathbb{B}^{p+1}(E, F)$.*

Théorème 4.1 Théorème de Schwarz
Soit E et F deux e.v.n, U un ouvert de E et $f : U \longrightarrow F$ 2 fois différentiable en $x_0 \in U$. Alors $d^2f(x_0) : E \times E \longrightarrow F$ est symétrique, i.e.

$$\forall h, k \in E, \; d^2f(x_0)(h,k) = d^2f(x_0)(k,h). \tag{4.1}$$

Preuve du Théorème 4.1
On obtient le résultat comme conséquence de la propriété suivantes : au voisinage de $h, k = 0$

$$\begin{aligned} &f(x_0+h+k) - f(x_0+h) - f(x_0+k) + f(x_0) - d^2f(x_0)(h,k) \\ &= (\|h\|_E + \|k\|_E)^2 \ \eta(\|h\|_E, \|k\|_E), \end{aligned} \tag{4.2}$$

où $\eta(\|h\|_E, \|k\|_E) \longrightarrow 0$ lorsque $(h,k) \longrightarrow 0$.
• En effet supposons provisoirement ceci démontré.

Fixons $h, k \in E$. Pour tout $t > 0$ au voisinage de 0, on déduit de l'égalité (4.2) que

$$\left\| d^2f(x_0)(th, tk) - d^2f(x_0)(tk, th) \right\|_F = (\|th\|_E + \|tk\|_E)^2 \eta(\|th\|_E, \|tk\|_E).$$

et ainsi on obtient $t^2 \left\| d^2f(x_0)(h,k) - d^2f(x_0)(k,h) \right\|_F \leqslant t^2(\|h\|_E + \|k\|_E)^2 \eta(t(\|h\|_E, \|k\|_E))$.
Simplifiant par t^2, on en déduit alors le résultat en faisant tendre t vers 0.

FIN COURS 7

• Il reste donc à prouver (4.2).
On pose $a(h,k) = f(x_0+h+k) - f(x_0+h) - f(x_0+k) + f(x_0) - d^2f(x_0)(h,k)$ et on le réécrit sous la forme $a(h,k) = a_1(h,k) + a_2(h,k)$ avec

$$a_1(h,k) = f(x_0+h+k) - f(x_0+h) - f(x_0+k) + f(x_0) - df(x_0+h)k + df(x_0)k,$$

et

$$a_2(h,k) = df(x_0+h)k - df(x_0)k - d^2f(x_0)(h,k).$$

On a alors par définition de la différentielle (propriété de continuité)

$$\|a_2(h,k)\|_F \leqslant \|k\|_E \left\| df(x_0+h) - df(x_0) - d^2f(x_0)(h,.) \right\|_{\mathbf{B(E,F)}}$$

mais df est différentiable en x_0 et donc

$$\left\| df(x_0+h) - df(x_0) - d^2f(x_0)(h,.) \right\|_{\mathbf{B(E,F)}} \leqslant \|h\|_E \, \eta_1(\|h\|_E)$$

avec $\eta_1(\|h\|_E) \longrightarrow 0$ lorsque $h \longrightarrow 0$. On obtient ainsi pour a_2

$$\|a_2(h,k)\|_F \leqslant \|k\|_E \, \|h\|_E \, \eta_1(\|h\|_E) \leqslant (\|h\|_E + \|k\|_E)^2 \ \eta_1(\|h\|_E). \tag{4.3}$$

Considérons maintenant a_1. On peut le récrire sous la forme $a_1(h,k) = b_h(k) - b_h(0)$ où on a posé

$$b_h(k) = f(x_0+h+k) - f(x_0+k) - df(x_0+h)k + df(x_0)k.$$

Mais puisque f est différentiable dans un voisinage de x_0, il existe $\alpha > 0$ tel que pour tout $k \in E$ vérifiant $\|k\|_E < \alpha$, b_h est différentiable sur $B_\alpha(0)$ et

$$db_h(k) = df(x_0+h+k) - df(x_0+k) - df(x_0+h) + df(x_0).$$

Soit alors $\varepsilon > 0$.
Puisque f est 2 fois différentiable en x_0, il existe $\beta > 0$ tq pour tout $\ell \in E$ vérifiant $\|\ell\|_E < \beta$ on ait

$$\left\| df(x_0+\ell) - df(x_0) - d^2f(x_0)(\ell,.) \right\|_{\mathbf{B(E,F)}} \leqslant \varepsilon \ \|l\|_E \,.$$

Donc si on choisit h et k tels que $\|h\|_E\,,\ \|k\|_E < \gamma = Min(\alpha, \beta/2)$, on peut écrire que

$$
\begin{aligned}
\|db_h(k)\|_{\mathbf{B(E,F)}} &\leqslant \left\|df(x_0+h+k) - df(x_0) - d^2f(x_0)(h+k,.)\right\|_{\mathbf{B(E,F)}} \\
&\quad + \left\|df(x_0+h) - df(x_0) - d^2f(x_0)(h,.)\right\|_{\mathbf{B(E,F)}} \\
&\quad + \left\|df(x_0+k) - df(x_0) - d^2f(x_0)(k,.)\right\|_{\mathbf{B(E,F)}} \\
&\leqslant \varepsilon\|h+k\|_E + \varepsilon\|h\|_E + \varepsilon\|k\|_E \leqslant 2\varepsilon(\|h\|_E + \|k\|_E).
\end{aligned}
$$

Donc par Corollaire 3.2.2 appliqué à b_h sur $B_\gamma(0)$, on trouve que pour tout $\|h\|_E\,, \|k\|_E < \gamma$,

$$
\|a_1(h,k)\|_F \leqslant 2\varepsilon\|k\|_E(\|h\|_E + \|k\|_E) \leqslant 2\varepsilon(\|h\|_E + \|k\|_E)^2
$$

ce qui compte tenu de (4.3) conclut la preuve de (4.2). ■

Nous venons de voir comment étendre la notion de différentielle à l'ordre deux. Nous allons maintenant voir comment faire pour les dérivées partielles, mais avant cela, voyons comment calculer de manière pratique une différentielle seconde.

Théorème 4.2 - Calcul d'une différentielle seconde
Soit E et F deux e.v.n, U un ouvert de E et $f : U \longrightarrow F$ une application différentiable sur U.

1. *Si f est deux fois différentiable en $x_0 \in U$ alors*

   $$\forall h \in E, \quad d^2f(x_0)(h,h) = g''(0), \tag{4.4}$$

   où $g : t \mapsto f(x_0+th)$ est définie sur l'ensemble des $t \in \mathbb{R}$ pour lesquels $(x_0+th) \in U$.

2. *Pour $h \in E$ fixé, on pose $g_h(x) = df(x)h$ pour tout $x \in U$. Si f est deux fois différentiable sur U alors g_h est différentiable sur U et*

   $$\forall h,k \in E, \quad dg_h(x)k = d^2f(x)(k,h). \tag{4.5}$$

Preuve du Théorème 4.2.
Décomposer les applications : exercice à faire en TD.

■

Voyons tout de suite sur un exemple comment utiliser la relation (4.5).

Exemple 4.1 - Différentielle seconde de $f : \mathbb{R} \longrightarrow \mathbb{R}$
Reprenons l'exemple d'une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$. Supposons f deux fois dérivable sur un voisinage ouvert de $a \in \mathbb{R}$.
On a vu que $df(x)h = f'(x)h$, $\forall h \in \mathbb{R}$. Posons donc pour x suffisament proche de a, $g_h(x) = f'(x)h$ (ce qui a bien un sens puisque df et donc f' est définie dans un voisinage de a).
Alors, f étant deux fois dérivable en a, f' est dérivable donc différentiable en a puisque c'est une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Ainsi f est 2 fois différentiable en a et donc par le Théorème 4.2, g_h est différentiable en a.
On a alors $dg_h(a)k = \left(df'(x)h\right) k$ pour tout $k \in \mathbb{R}$. Or f' étant de $\mathbb{R}$ dans $\mathbb{R}$, sa différentielle en a s'identifie avec sa dérivée en ce point. Ainsi, on a $dg_h(a)k = \left(\left(f'\right)'(x)h\right) k = (f''(x)h)k = f''(x)hk$.
La différentielle seconde de f en a s'identifie donc au nombre $f''(a)$: $\forall h, k \in \mathbb{R}$, $d^2f(a)(h,k) = f''(a)hk$.

4.1.2 Dérivées partielles secondes

Définition 4.2 - Dérivée partielle seconde
Soit $U \subset \mathbb{R}^n$ un ouvert, F e.v.n et $f : U \longrightarrow F$ une application.
Pour i et j entiers donnés vérifiant $1 \leqslant i, j \leqslant n$, on dit que f admet en $x_0 \in U$ une dérivée partielle seconde par rapport à (x_i, x_j) si

1. $\dfrac{\partial f}{\partial x_j}$ *existe sur $V \subset U$ voisinage ouvert de x_0 ;*
2. $\dfrac{\partial f}{\partial x_j}$ *admet en x_0 une dérivée partielle par rapport à x_i : $\dfrac{\partial}{\partial x_i}\left(\dfrac{\partial f}{\partial x_j}\right)(x_0)$.*

Si tel est le cas on note $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$ la quantité $\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(x_0)$ et $\frac{\partial^2 f}{\partial x_i^2}(x_0)$ la quantité $\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right)(x_0)$.
Là encore, on dira que f admet sur U une dérivée partielle seconde par rapport à (x_i, x_j) si cette dernière existe en tout point de U.

On peut alors montrer une première propriété qui est une conséquence directe des résultats vus précédement.

Corollaire 4.2.1 - Lien entre dérivées partielles et différentielle - Formule de Schwarz
Soit $U \subset \mathbb{R}^n$ un ouvert, F e.v.n et $f : U \longrightarrow F$ une application.
Si f est deux fois différentiable en $x_0 \in U$, alors elle admet en x_0 des dérivées partielles secondes par rapport à (x_i, x_j) pour tout $i, j = 1, .., n$. On a alors

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = d^2 f(x_0)(e_i, e_j). \tag{4.6}$$

De plus, dans ce cas

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j x_i}(x_0). \tag{4.7}$$

De manière équivalente, s'il existe $i, j \in \{1, .., n\}$ tels que $\frac{\partial^2 f}{\partial x_i x_j}(x_0) \neq \frac{\partial^2 f}{\partial x_j x_i}(x_0)$ alors f n'admet pas de différentielle seconde en x_0.

Preuve du Corollaire 4.2.1.
L'existence des dérivées partielles et la formule (4.6) résultent de la définition de la différentielle seconde et de la Propsition 2.5.
L'égalité (4.7) découle quant à elle du Théorème 4.1.

■

Remarque 4.2
En outre, on voit que la formule (4.6) fournit un moyen d'expliciter la différentielle seconde de f en un point losqu'elle existe, en fonction des dérivées partielles secondes de f. Dans ce cas pour tout $h, k \in E$, on a la relation

$$d^2 f(x_0)(h,k) = \sum_{i,j=1}^{n} h_i k_j \frac{\partial^2 f}{\partial x_j x_i}(x_0). \tag{4.8}$$

4.1.3 Propriétés - Application de classe C^2

Puisque la différentielle seconde n'est rien d'autre que la différentielle première d'une différentielle, elle hérite des propriétés de cette dernière que nous avons vues précédement. En particulier, on a :

Proposition 4.1
1. Soient E et F deux e.v.n, U un ouvert de E et $f, g : U \longrightarrow F$ des applications 2 fois différentiables en $a \in U$. Si $\alpha \in \mathbb{R}$, alors les applications αf et $f + g$ sont 2 fois différentiables.
2. Soient E, F et G trois e.v.n., U un ouvert de E, V un ouvert de F et $f : U \longrightarrow F$, $g : V \longrightarrow G$ deux applications vérifiant $f(U) \subset V$. On suppose que f est deux fois différentiable en $a \in U$ et que g est deux fois différentiable en $f(a) \in V$.
Alors l'application $g \circ f : U \longrightarrow G$ est deux fois différentiable en $a \in U$.

Etendons alors la Définition 2.7 en posant :

Définition 4.3 - Classe C^2
Soient E et F deux e.v.n et U un ouvert de E. Une application $f : U \longrightarrow F$ est dite de classe C^2 sur U ou simplement C^2 sur U (on note alors $f \in C^2(U)$) si :

(i) *f est 2 fois différentiable sur U ;*

(ii) *sa différentielle seconde $d^2 f : U \longrightarrow \mathbb{B}^2(E, F)$ est continue.*

Compte tenu de la Proposition 4.1 on a évidement que toute combinaison linéaire de fonctions C^2 est encore C^2 ; et que toute composée de fonctions C^2 est C^2.

On a alors la caractérisation suivante.

Proposition 4.2 - Caractérisation des fonctions C^2
On suppose que $E = \mathbb{R}^n$ et F e.v.n. Soit U un ouvert de E et f une application de U dans F. Alors les deux assertions suivantes sont équivalentes :

1. *f est de classe C^2 sur U ;*
2. *les dérivées partielles secondes de f existent et sont continues sur U.*

Preuve du Proposition 4.2.
L'implication 1. $\Longrightarrow$ 2. est une conséquence directe de la formule (4.6) du Corollaire 4.2.1.
Montrons alors que 2. $\Longrightarrow$ 1.
Supposant 2., on a en particulier que les dérivées partielles premières de f existent et sont différentiables donc continues. Ainsi, on déduit de la Proposition 3.3 que f est C^1 sur U. De plus, d'après la Proposition 2.5 on a pour tout $x \in U$ et tout $h \in E$

$$df(x)h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)\ h_i.$$

Ainsi, df est une somme de fonctions admettant toutes des dérivées partielles premières et donc pour tout $j = 1, .., n$, les $\frac{\partial}{\partial x_j}(df)$ existent et vérifient pour tout $x \in U$ et tout $h \in E$ (utiliser la définition de le DP première)

$$\frac{\partial}{\partial x_j}(df)(x)h = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(x)\ h_i.$$

Mais par hypothèse les dérivées partielles secondes de f existent et sont continues sur U et donc les $\frac{\partial}{\partial x_j}(df) : U \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$ existent et sont continues. Donc on peut conclure en appliquant à nouveau, cette fois à df, la Proposition 3.3. ■
FIN COURS 8

4.2 Différentielles d'ordre supérieur à deux

Les notions et résultats précédents s'étendent aux ordres supérieurs par récurrence. On donne ci-dessous les principales définitions et propriétés sans démonstrations. Celles-ci consistent à appliquer les résultats et les méthodes vues précédement et une récurrence.

Définition 4.4 Différentielle d'ordre k
Soient E et F des e.v.n et U un ouvert de E. On dit qu'une application $f : U \longrightarrow F$ admet en $x_0 \in U$ une différentielle k-ième que l'on notera $d^k f(x_0)$ si

1. *f est $(k-1)$ fois différentiable dans un voisinage ouvert $V \subset U$ de x_0 ;*
2. *la différentielle $(k-1)$-ième de f, $d^{k-1}f : V \longrightarrow \mathbb{B}^{k-1}(E, F)$ est différentiable en x_0.*

Le même type d'arguments que ceux utilisés pour la différentielle seconde montre que $d^k f(x_0)$ peut s'identidier à une application k-linéaire sur E. On note , $d^k f(x_0) \in \mathbb{B}^k(E, F)$.
On dira que f est k fois différentiable sur U si elle est k fois différentiable en tout point de U.

On peut alors étendre à l'ordre k la Définition 2.7 en posant :

Définition 4.5 Classes C^k et C^∞
Soit $k \geqslant 2$ un entier. On dit que f est de classe C^k sur U ($f \in C^k(U)$) si elle est k fois différentiable sur U et si sa différentielle k-ième $d^k f : E \longrightarrow \mathbb{B}^k(E, F)$ est continue.
Si f est C^k pour tout k sur U, on dit que f est de classe C^∞ sur U et on note $f \in C^\infty(U)$.

On a alors les propriétés de base suivantes :

Proposition 4.3
E et F sont des e.v.n, U un ouvert de E, k un entier et $f : U \longrightarrow F$ une application k fois différentiable sur U (respectivement C^k sur U).

1. *On a $d^k f = d(d^{k-1} f) = d^{k-1}(df)$ sur U.*
2. *Pour $x_0 \in U$, $d^k f(x_0) \in \mathbb{B}^k(E, F)$ est symétrique i.e. pour toute permutation σ à k éléments, $d^k f(x_0)(h_1, .., h_k) = d^k f(x_0)(h_{\sigma(1)}, .., h_{\sigma(k)})$ pour tout $h_1, .., h_k \in E$.*
3. *Si $g : U \longrightarrow F$ est k fois différentiable sur U (respectivement C^k sur U) alors pour tout $\alpha \in \mathbb{R}$, $(f + \alpha g)$ est k fois différentiable sur U (respectivement C^k sur U).*
4. *Si G est un e.v.n, $V \subset F$ un ouvert vérifiant $f(U) \subset V$ et $g : V \longrightarrow G$ une application k fois différentiable sur V (respectivement C^k sur V) alors $g \circ f : U \longrightarrow F$ est k fois différentiable sur U (respectivement C^k sur U).*

Dans le cas $E = \mathbb{R}^n$, on étend également par récurrence la défintion des dérivées partielles. On a alors le résultat suivant.

Proposition 4.4
On suppose donc que $E = \mathbb{R}^n$, F e.v.n, $U \subset E$ ouvert et $f : U \longrightarrow F$ une application. Alors

1. *si f est k fois différentiable en $x_0 \in U$, elle admet en x_0 des dérivées partielles d'ordre k et on a la relation*
$$d^k f(x_0)(e_{i_1}, .., e_{i_k}) = \frac{\partial^k f}{\partial x_{i_1} .. \partial x_{i_k}}(x_0);$$
où $1 \leqslant i_1, .., i_k \leqslant n$ sont des entiers.
2. *f est de classe C^k sur U si et seulement si ses dérivées partielles d'ordre k sont toutes continues sur U*

Terminons ce paragraphe par une propriété.

Théorème 4.3 - Régularité C^∞ de l'inverse
E, F Banach ($dim(E) = dim(F)$ si dim finie). L'application Inv, définie en (2.10) est de classe C^∞ sur $G\mathbb{B}(\mathrm{E}, \mathrm{F})$.

Preuve du Théorème 4.3.

Utiliser la décomposition de $dInv(u)$ de la preuve du théorème 2.3

■

4.3 Formules de Taylor

Dans toute cette Section on suppose que $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$.
Pour une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}^p$, on notera classiquement $f^{(k)}$ sa dérivée k-ième quand elle existe.
Nous aurons besoin dans ce paragraphe du résultat auxiliaire suivant.

Lemme 4.1
Soit I un intervalle ouvert de $\mathbb{R}$ et $g : \mathbb{R} \longrightarrow \mathbb{R}^p$ une fonction $(k+1)$ fois dérivable sur I. Alors

$$\forall t \in I, \quad \left(\sum_{i=0}^{k} \frac{(1-t)^i}{i!}\ g^{(i)}(t)\right)' = \frac{(1-t)^k}{k!}\ g^{(k+1)}(t). \tag{4.9}$$

Preuve du Lemme 4.1.
Par récurrence.
• Initialisation (k=0) :
pour toute fonction dérivable, (4.9) s'écrit $\forall t \in I$, $g'(t) = g'(t)$ (en utilisant $g^{(0)} = g$).
• Supposons que pour toute fonction $g : \mathbb{R} \longrightarrow \mathbb{R}^p$ dérivable k fois sur I on ait

$$\left(\sum_{i=0}^{k-1} \frac{(1-t)^i}{i!}\ g^{(i)}(t)\right)' = \frac{(1-t)^{k-1}}{(k-1)!}\ g^{(k)}(t).$$

Soit alors $g : \mathbb{R} \longrightarrow \mathbb{R}^p$ fonction $(k+1)$fois dérivable sur I.
Pour $i = 1, .., k$, la fonction $t \longrightarrow \dfrac{(1-t)^i}{i!}\ g^{(i)}(t)$ est dérivable sur I et on peut donc écrire

$$\begin{aligned}
\left(\sum_{i=0}^{k} \frac{(1-t)^i}{i!}\ g^{(i)}(t)\right)' &= \left(\sum_{i=0}^{k-1} \frac{(1-t)^i}{i!}\ g^{(i)}(t)\right)' + \left(\frac{(1-t)^k}{k!}\ g^{(k)}(t)\right)' \\
&= \frac{(1-t)^{k-1}}{(k-1)!}\ g^{(k)}(t) + \frac{(1-t)^k}{k!}\ g^{(k+1)}(t) - k\frac{(1-t)^{k-1}}{k!}\ g^{(k)}(t) \\
&= \frac{(1-t)^k}{k!}\ g^{(k+1)}(t).
\end{aligned}$$

On a donc l'hérédité et par récurrence, la propriété est donc vraie pour tout $k \geqslant 0$. ■

Remarque 4.3
Si $[0,1] \subset I$ et si $g \in C^{k+1}(I)$, intégrant (4.9) entre 0 et 1, on trouve la formule

$$g(1) - g(0) - g^{'}(0) - \cdots - \frac{1}{k!} g^{(k)}(0) = \int_0^1 \frac{(1-t)^k}{k!}\ g^{(k+1)}(t)dt. \tag{4.10}$$

Théorème 4.4 - Taylor avec reste intégral
Soit $U \subset \mathbb{R}^n$ ouvert, $k \geqslant 0$ un entier et $f : U \longrightarrow \mathbb{R}^p$ une application de classe C^{k+1} sur U.
Alors pour tout $x, h \in \mathbb{R}^n$ tels que $[x, x+h] \subset U$ on a, notant $(h)^k = \underbrace{(h, \cdots, h)}_{k \text{ termes}}$

$$\begin{aligned} f(x+h) = \;& f(x) + df(x)h + \cdots + \frac{d^k f(x)}{k!}(h)^k \\ & + \frac{1}{k!}\int_0^1 (1-t)^k d^{k+1} f(x+th)(h)^{k+1} dt. \end{aligned}$$

Noter qu'en particulier, dans le cas où f est C^1 (i.e. $k = 0$), on trouve

$$f(x+h) - f(x) = \int_0^1 df(x+th)h dt.$$

Preuve du Théorème 4.4.
Pour $x, h \in \mathbb{R}^n$, on pose pur tout $t \in [0,1]$, $\psi(t) = x + th$: fonction C^∞ sur $[0,1]$. Puisque pour tout $t \in [0,1]$, $\psi(t) \in [x, x+h] \subset U$ la fonction $g : t \mapsto f(x+th)$ est $C^{k+1}([0,1])$ puisque f est C^{k+1} sur U. Il suffit alors de remarquer que $g^{(q)}(t) = d^q f(x+th)(h)^q$ pour tout $q \leqslant k+1$ pour conclure par (4.10).

■

Théorème 4.5 - Taylor-Lagrange
Mêmes notations qu'avant avec $f : U \longrightarrow F$ une application $(k+1)$ fois différentiable.
S'il existe $C > 0$ tel que pour tout $x \in U$, $\left\| d^{k+1} f(x) \right\|_{\mathbf{B}^{k+1}(E;F)} \leqslant C$ alors pour tout $[x, x+h] \subset U$ on a l'inégalité

$$\left\| f(x+h) - f(x) - df(x)h - \cdots - \frac{d^k f(x)}{k!}(h)^k \right\|_F \leqslant \frac{C \|h\|_E^{k+1}}{(k+1)!}.$$

Preuve du Théorème 4.5.
On introduit toujours la même fonction $g : t \mapsto f(x+th)$.
Posant $r(t) = \sum_{i=0}^{k} \frac{(1-t)^i}{i!} \; g^{(i)}(t)$, on déduit du Lemme 4.1 que pour tout $t \in]0,1[$

$$\|r'(t)\| = \left\| \frac{(1-t)^k}{k!} \; g^{(k+1)}(t) \right\| = \left\| \frac{(1-t)^k}{k!} \; d^{k+1} f(x+th)(h)^{k+1} \right\| \leqslant C \|h\|^{k+1} \frac{(1-t)^k}{k!} = w'(t)$$

en posant pour tout $t \in [0,1]$, $w(t) = -C \|h\|^{k+1} \frac{(1-t)^{k+1}}{(k+1)!}$.
Les fonctions r et w sont continues sur sur $[0,1]$, dérivables sur $]0,1[$ et vérifient pour tout $t \in]0,1[$, $\|r(t)\| \leqslant w'(t)$. On peut donc appliquer l'inégalité des accroissements finis (Théorème 3.2) :

$$\begin{aligned} \|r(1) - r(0)\| &= \left\| f(x+h) - f(x) - df(x)h - \cdots - \frac{d^k f(x)}{k!}(h)^k \right\| \\ &\leqslant w(1) - w(0) = \frac{C \|h\|^{k+1}}{(k+1)!}. \end{aligned}$$

■

Théorème 4.6 - Taylor-Young
Soit U un ouvert de $\mathbb{R}^n$, $x_0 \in U$, $k \geqslant 0$ un entier et $f : U \longrightarrow \mathbb{R}^p$ une application k fois différentiable en x_0. Alors on a l'égalité

$$f(x_0+h) = f(x_0) + df(x_0)(h) + \cdots + \frac{d^k f(x_0)}{k!}(h,\cdots,h) + o(\|h\|_E^k). \tag{4.11}$$

Preuve du Théorème 4.6.
Dans ce qui suit on note $E=\mathbb{R}^n$, $F=\mathbb{R}^p$ et on prendra h dans un voisinage ouvert suffisament petit de 0 pour que les différentes quantités aient un sens.

On procède par récurrence.
• Pour $k=1$, ce n'est rien d'autre que la différentiabilité de la fonction.
• Supposons que (4.11) soit vérifiée au rang $k-1$, $k \geqslant 2$ et posons

$$\xi(h) = f(x_0+h) - f(x_0) - \cdots - \frac{d^k f(x_0)}{k!}(h)^k.$$

Il s'agit alors de montrer que $\xi(h) = o(\|h\|_E^k)$.
Pour p entier tel que $1 \leqslant p \leqslant k-1$, on définit $L_p : \begin{array}{ccc} E & \longrightarrow & F \\ h & \mapsto & d^p f(x_0)(h)^p \end{array}$.
On peut alors décompose L_p en $L_p = \Psi \circ \varphi$ où

$$\varphi : \begin{array}{ccc} E & \longrightarrow & E^p \\ h & \mapsto & (h)^p \end{array} \quad et \quad \Psi : \begin{array}{ccc} E^p & \longrightarrow & F \\ (h_1,\cdots,h_p) & \mapsto & d^p f(x_0)(h_1,\cdots,h_p). \end{array}$$

Ainsi on voit que φ est linéaire donc différentiable et Ψ est p-linéaire donc différentiable. Donc L_p est différentiable et pour tout $u \in E$, $dL_p(h)(u) = d\Psi(\varphi(h)) \circ \varphi(u)$. On a donc

$$dL_p(h)(u) = d\psi(h,...,h)(u,...,u)) = \sum_{i=1}^{p} d^p f(x_0) \underbrace{(h,...,h,u,h,...,h)}_{p \text{ termes avec } u \text{ en position } i}$$

Mais $d^p f(x_0)$ est symétrique et donc $dL_p(h)(u) = p\, d^p f(x_0)(h,\cdots,h,u)$. On en déduit donc que ξ est différentiable au voisinage de $h=0$ avec pour tout $u \in E$

$$\begin{aligned} d\xi(h)u &= df(x_0+h)u - df(x_0)u - d^2 f(x_0)(h,u) - \cdots - \frac{d^k f(x_0)}{(k-1)!}(h,\cdots,h,u) \\ &= df(x_0+h)u - df(x_0)u - d(df)(x_0)(h)(u) - \cdots - \frac{d^{k-1}(df)(x_0)}{(k-1)!}(h,\cdots,h)(u), \end{aligned}$$

c'est-à-dire $d\xi(h) = df(x_0+h) - df(x_0) - d^2 f(x_0)(h) - \cdots - \dfrac{d^{k-1} f(x_0)}{(k-1)!}(h,\cdots,h)$.
Si on d'applique alors l'hypothèse de récurrence à df, on obtient $d\xi(h) = o(\|h\|_E^{k-1})$ soit

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \|h\|_E < \eta \Longrightarrow \|d\xi(h)\|_{\mathbf{B(E,F)}} \leqslant \varepsilon\ \|h\|_E^{k-1}.$$

On peut donc appliquer à ξ le théorème des accroissements finis sur $B(0,\eta)$ (Corollaire 3.2.2), ce qui donne $\|\xi(h)-\xi(0)\|_F \leqslant \varepsilon\ \|h\|_E^{k-1}\ \|h\|_E$. Soit encore, puisque $\xi(0)=0$, que $\|\xi(h)\|_F \leqslant \varepsilon\ \|h\|_E^k$ pour h dans un voisinage de 0. La propriété est donc vraie au rang k.

■

Remarque 4.4
Le Théorème 4.6 est **local**. *On se place en effet au voisinage de $x_0 \in U$. C'est-à dire qu'on suppose qu'il existe $\alpha > 0$ tel que pour tout $h \in E$, vérifiant $\|h\|_E < \alpha$ on a $x_0+h \in U$ (i.e. $x_0+B(0,\alpha) \subset U$). En conséquence, les $f(x_0+h)$ sont bien définis.*
FIN COURS 9

4.4 Extrema libres

Pour la fin de cette section $E=\mathbb{R}^n$, $F=\mathbb{R}$, $U \subset \mathbb{R}^n$ est un ouvert et $f : U \longrightarrow \mathbb{R}$.
Commençons par deux définitions.

Définition 4.6 - Extremum local
On dira que f admet en $x_0 \in U$ un minimum local (respectivement maximum local) s'il existe un voisinage ouvert V de x_0 tel que $V \subset U$ et $\forall x \in V$, $f(x) \geqslant f(x_0)$ (respectivement $f(x) \leqslant f(x_0)$).
Si x_0 est un minimum ou un maximum, on dit que c'est un extremum.

Remarque 4.5
On voit que la notion introduite est purement locale : le voisinage V peut être "petit". On ne s'intéresse pas à ce qui se passe en dehors de V. C'est ce qui justifie la terminologie "local" (par opposition à global).

Définition 4.7 - Point critique
Soit $f : U \longrightarrow \mathbb{R}$ différentiable. Si $x_0 \in U$ est tel que $df(x_0) = 0$ alors on dit que x_0 est un point critique de f.

On a alors la propriété suivante qui indique, pour une fonction différentiable, où chercher les extrema de f. Ce résultat est déja bien connue dans le cas d'une fonction de $\mathbb{R}$ dans $\mathbb{R}$.

Théorème 4.7 - Condition nécessaire d'ordre 1
Soit $f : U \longrightarrow \mathbb{R}$ différentiable. Si $x_0 \in U$ est un extremum de f alors c'est un point critique de f i.e. $df(x_0) = 0$.

Preuve du Théorème 4.7.
Supposons par exemple que f admette en x_0 un maximum. Alors $\exists \alpha > 0$ tel que pour tout $h \in E$ vérifiant $\|h\|_E < \alpha$, on ait $f(x_0+h) \leqslant f(x_0)$.
Fixons alors $k \in E^*$ et considérons des $t \in \mathbb{R}^*$ tels que $|t| < \frac{\alpha}{\|k\|_E}$ de telle sorte qu'on ait $\|tk\|_E < \alpha$.
Par différentiabilité de f en x_0, on a pour $t > 0$

$$0 \leqslant f(x_0+tk) - f(x_0) = df(x_0)(tk) + \|k\|_E\, o(t), \tag{4.12}$$

et donc, divisant par t on obtient $df(x_0)(k) + \|k\|_E\, o(1) \geqslant 0$, soit encore pour $t \longrightarrow 0$, $df(x_0)(k) \geqslant 0$.
Mais alors remplçant t par $-t$ dans (4.12), on a pour $t > 0$

$$0 \leqslant f(x_0-tk) - f(x_0) = df(x_0)(-tk) + \|k\|_E\, o(t),$$

et on obtient ainsi $df(x_0)(k) \leqslant 0$, d'où nécessairement $df(x_0)(k) = 0$ avec $k \in E^*$ arbitraire donc $df(x_0) = 0$.

■

Remarque 4.6
Le Théorème 4.7 ne donne évidement qu'une condition nécessaire. Elle n'est en effet pas suffisante : si $df(x_0)$=0 alors x_0 n'est pas forcément un extremum. Considérer par exemple la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ telle que $f(x) = x^3$.

Nous aurons ensuite besoin de la définition suivante.

Définition 4.8 - Forme bilinéaire symétrique définie positive (ou négative)
Soit $\varphi : E \times E \longrightarrow \mathbb{R}$ une forme bilinéaire symétrique.
On dit que φ est

1. *positive si $\forall h \in E$, $\varphi(h, h) \geqslant 0$;*
2. *définie si $\varphi(h, h) = 0$ entraine $h = 0$.*

On dit dans ce cas que φ est définie positive.

Si au lieu de 1, on a $\forall h \in E$, $\varphi(h, h) \leqslant 0$, on dit que φ est négative. Elle sera donc définie négative si elle vérifie de plus 2.

Remarque 4.7
La matrice à coefficients réels A de φ dans la base canonique de E ($(A)_{ij} = \varphi(e_i, e_j)$) est symétrique et donc diagonalisable. Ainsi il existe $(u_1, \cdots, u_n)$ base orthonormée de E (constituée de vecteurs propres de A) telle que[a] $\varphi(h, h) = \sum_{i=1}^{n} \lambda_i \,(< h, u_i >_2)^2$ où $\lambda_1, \cdots, \lambda_n$ sont les valeurs propres de A (avec répétition). On voit ainsi qu'on a les caractérisations suivantes :

1. *φ est positive (respectivement négative) si $\forall 1 \leqslant i \leqslant n$, $\lambda_i \geqslant 0$ (resp. $\lambda_i \leqslant 0$).*
2. *φ est définie positive (resp. négative) si $\forall 1 \leqslant i \leqslant n$, $\lambda_i > 0$ (resp. $\lambda_i < 0$).*
3. *φ est définie positive (resp. négative) s'il existe $\alpha > 0$ tel que $\forall h \in E$, $\varphi(h, h) \geqslant \alpha \|h\|_E^2$ (resp. $\varphi(h, h) \leqslant -\alpha \|h\|_E^2$)*

[a] Décomposant h dans la base des vecteurs propres, on a $h = \sum_{i=1}^{n} h_i u_i$ donc $h_i =< h, u_i >_2$ et donc

$$\varphi(h, h) = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j < Au_i, u_j >_2 = \sum_{i=1}^{n} \lambda_i \, h_i^2 = \sum_{i=1}^{n} \lambda_i \,(< h, u_i >_2)^2$$

Remarque 4.8
Pour nous, la forme bilinéaire symétrique φ de la Remarque 4.7 sera, quand elle existe $d^2f(x_0)$, avec $x_0 \in U$ ouvert de $\mathbb{R}^n$ et $f : U \longrightarrow \mathbb{R}$ fonction 2 fois différentiable. Compte tenu de la formule (4.6) vue dans le Corollaire 4.2.1, la matrice A de $d^2f(x_0)$ dans la base canonique de E sera la matrice **symétrique** *$(n \times n)$ de coefficients $d^2f(x_0)(e_i, e_j) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$. Cette matrice porte le nom de* **matrice hessienne de** f **en** x_0*, on la notera $H_f(x_0)$.*
On peut énoncer deux nouveaux résultats qui terminerons ce chapitre.

Théorème 4.8 - Condition nécessaire d'ordre 2
Soit $f : U \longrightarrow \mathbb{R}$ une application 2 fois différentiable. Si $x_0 \in U$ est un minimum (respectivement maximum) de f alors $d^2f(x_0)$ est positive (resp. négative).

Preuve du Théorème 4.8.
Supposons que $x_0 \in U$ soit un minimum de f. Alors d'après le Théorème 4.7 on a $df(x_0) = 0$ et puisque de plus f est 2 fois différentiable, on peut lui appliquer la formule de Taylor-Young à l'ordre deux en x_0 (Théorème 4.6), soit

$$f(x_0 + h) - f(x_0) = \frac{1}{2}d^2f(x_0)(h, h) + o(\|h\|_E^2)$$

pour tout $h \in E$ suffisament proche de 0 avec par définition $\lim\limits_{h \longrightarrow 0} \dfrac{o(\|h\|_E^2)}{\|h\|_E^2} = 0$.
Mais x_0 étant un minimum de f, on a $\exists \alpha > 0$ tel que $B(x_0, \alpha) \subset U$ et $\|h\|_E < \alpha$ entraine $f(x_0 + h) - f(x_0) \geqslant 0$.
Ainsi, $\|h\|_E < \alpha$ entraine $d^2f(x_0)(h, h) + o(\|h\|_E^2) \geqslant 0$.
Fixons alors $x \in E$, $x \neq 0$.
Pour tout $t \in \mathbb{R}$ vérifiant $0 < |t| < \dfrac{\alpha}{\|x\|_E}$, on a $\|tx\|_E < \alpha$ et donc

$$\begin{array}{ll} & d^2f(x_0)(tx, tx) + o(\|tx\|_E^2) \geqslant 0 \\ \Longrightarrow & t^2\left(d^2f(x_0)(x, x) + \dfrac{1}{t^2}o(\|tx\|_E^2)\right) \geqslant 0 \\ \Longrightarrow & d^2f(x_0)(x, x) + \dfrac{1}{t^2}o(\|tx\|_E^2) \geqslant 0, \end{array}$$

d'où $d^2f(x_0)(x, x) \geqslant 0$ en faisant tendre t vers 0.

■

Théorème 4.9 - Condition suffisante d'ordre 2
Soit $f : U \longrightarrow \mathbb{R}$ une application 2 fois différentiable. Si $x_0 \in U$ est tel que $df(x_0) = 0$ et $d^2f(x_0)$ est définie positive (respectivement définie négative) alors x_0 est un minimum (resp. maximum) local de f.

Preuve du Théorème 4.9.
C'est une conséquence plus ou moins directe de la formule de Taylor-Young à l'ordre 2. En effet, d'après le Théorème 4.6, $\exists \beta_0 > 0$ tel que pour tout $h \in E$ vérifiant $\|h\|_E < \beta_0$ on ait $(x_0 + h) \in U$, et

$$f(x_0 + h) - f(x_0) = \frac{1}{2}d^2f(x_0)(h, h) + o(\|h\|_E^2).$$

Mais si $d^2f(x_0)$ est définie positive, d'après la Remarque 4.7, il existe $\alpha > 0$ tel que $\forall h \in E$, $d^2f(x_0)(h,h) \geqslant \alpha \left\|h\right\|_E^2$ et on en déduit donc que pour tout $\left\|h\right\|_E < \beta_0$

$$f(x_0+h) - f(x_0) \geqslant \frac{\alpha}{2} \left\|h\right\|_E^2 + o(\left\|h\right\|_E^2). \tag{4.13}$$

Mais $\lim\limits_{h \longrightarrow 0} \dfrac{o(\|h\|_E^2)}{\|h\|_E^2} = 0$ et donc il existe $\beta_1 > 0$ tel que $\|h\|_E < \beta_1$ entraine $\left|\dfrac{o(\|h\|_E^2)}{\|h\|_E^2}\right| < \frac{\alpha}{4}$. Ainsi on déduit de (4.13) qu'il existe $\beta = Min(\beta_0, \beta_1) > 0$ tel que $\|h\|_E < \beta$ entraine

$$f(x_0+h) - f(x_0) \geqslant \frac{\alpha}{2} - \frac{\alpha}{4} \geqslant 0, \tag{4.14}$$

ce qui termine la preuve.

■

Remarque 4.9
1) *Comme le montre l'inégalité (4.14) dans la preuve précédente, sous les hypothèses du Théorème 4.9, le minimum (repectivement maximum) local est* **stricte** *i.e.* $f(x_0+h) - f(x_0) > 0$ *au voisinage de* x_0 *pour* $h \neq 0$.
2) *Les différentes propriétés que nous venons de voir fournissent un plan pour l'étude des extrema d'une fonction :*

1. *On cherche les points critiques de f (les extrema en sont nécessairement par le Théorème 4.7) ;*
2. *On étudie la nature de chaque point critique. Ainsi pour chaque point critique :*
   (a) *On calcule la différentielle seconde ou, ce qui revient au même, la matrice hessienne de f en ce point ;*
   (b) *Si celle-ci est définie positive ou définie négative : on conclut par le Théorème 4.9.*
   Si elle n'est ni positive, ni négative on conclut par le Théorème 4.8.
   Sinon il faut mener une étude spécifique pour ce point.

$\boxed{\text{FIN DU CHAPITRE 4}}$

Chapitre 5

Théorème d'inversion locale - Théorème des fonctions implicites

5.1 Difféomorphismes locaux

Dans tout ce paragraphe, E et F sont des Banach vérifiant de plus $dim(E) = dim(F)$ si la dimension est finie.

Nous avons d'abord besoin d'une définition.

Définition 5.1 Homéomorphisme - Difféomorphisme
Soient U un ouvert de E, V un ouvert de F et $f : U \longrightarrow V$ une application. On a alors les définitions suivantes :

1. *f est un homéomorphisme de U dans V si* (a) *f est continue sur U et bijective de U dans V ;*
   (b) *f^{-1} est continue sur V.*
2. *f est un difféomorphisme de U dans V si* (a) *f est C^1 sur U et bijective de U dans V ;*
   (b) *f^{-1} est de classe C^1 sur V.*
3. *f est un C^{k+1} difféomorphisme de U dans V si* (a) *f est C^{k+1} sur U et bijective de U dans V ;*
   (b) *f^{-1} est de classe C^{k+1} sur V.*

Remarque 5.1
Noter qu'un homéomorphisme C^1 n'est pas forcément un difféomorphisme car une fonction peut être bijective et C^1 sans pour autant que son inverse soit différentiable (considérer par exemple $f(x) = x^3$, $x \in \mathbb{R}$).

En revanche, si f est un difféomorphisme et f est C^{k+1} alors la bijection réciproque de f est automatiquement C^{k+1} comme le montre la seconde partie du résultat suivant qui nous sera utile par la suite.
FIN COURS 10

Proposition 5.1
Soient U un ouvert de E , V un ouvert de F et f une application de E dans F.

1. *Si f est un difféomorphisme de U dans V (donc nécessairement $V = f(U)$). Alors pour tout $x \in U$, $df(x)$ est inversible et pour tout $y \in V$,*

$$df^{-1}(y) = \Big(df(f^{-1}(y))\Big)^{-1}. \tag{5.1}$$

2. *Si f est un difféomorphisme C^{k+1} de U dans V alors c'est un C^{k+1} difféomorphisme.*

Preuve de la Proposition 5.1.
Puisque f et f^{-1} sont différentiables sur U et $V = f(U)$ respectivement, la composée $f^{-1} \circ f$ l'est également sur U. Comme pour tout $x \in U$, $f^{-1} \circ f(x) = x$, on en déduit que pour tout $h \in E$ $d(f^{-1} \circ f)(x)(h) = h$, i.e. $df^{-1}(f(x)) \circ df(x)(h) = h$ soit encore $df^{-1}(f(x)) \circ df(x) = I_E$. Comme de même on obtient que $d(f \circ f^{-1})(f(x)) = df(x) \circ df^{-1}(f(x)) = I_F$, ceci montre d'une part que $df(x)$ est inversible, que d'autre part son inverse est $df^{-1}(f(x))$ et que de plus on a la formule (5.1).
Concernant la partie 2 du Lemme, il suffit de remarquer que si f est de plus C^{k+1}, la formule (5.1) s'écrit pour tout $y \in V$

$$df^{-1}(y) = \Big(df(f^{-1}(y))\Big)^{-1} = Inv \circ df \circ f^{-1}(y).$$

Donc puisque df et Inv sont C^k (voir Théorème 4.3) et f^{-1} est C^1 la composée i.e. df^{-1} est C^1 i.e. f^{-1} est C^2 et de proche en proche on obtient finalement que df^{-1} est C^k, soit f^{-1} est C^{k+1}. ■

Voyons alors une caractérisation des difféomorphismes qui s'avère parfois utile pour vérifier de manière concrête qu'une application est un difféomorphisme.

Théorème 5.1 Caractérisation des difféomorphismes
Soient U un ouvert de E, V un ouvert de F et $f : U \longrightarrow V$ un homéomorphisme C^1. Les deux propriétes suivantes sont alors équivalentes :

1. *f est un difféomorphisme de U dans V ;*
2. *pour tout $x \in U$, $df(x) \in \mathrm{G}\mathbb{B}(\mathrm{E},\mathrm{F})$.*

De plus, si l'équivalence est vérifiée, on a pour tout $x \in U$, $df^{-1}(f(x)) = (df(x))^{-1}$.

Preuve du Théorème 5.1.
• L'implication (*1* $\Longrightarrow$ *2*) est une conséquence directe de la Proposition 5.1.

• Montrons que (2 $\Longrightarrow$ 1.)
Pour $a, x \in U$, on note $b = f(a)$, $y = f(x)$ (qui sont donc dans V) et on note $g = f^{-1}$.
Puisque f est différentiable en a, on peut écrire pour x au voisinage de a que

$$f(x) - f(a) = df(a)(x - a) + o(\|x - a\|_E),$$

ce qui compte tenu des notations introduite s'écrit

$$y - b = df(a)(x - a) + o(\|x - a\|_E).$$

Prenant alors $(df(a))^{-1}$ de cette dernière égalité, on trouve par linéarité

$$(df(a))^{-1}(y-b) - (df(a))^{-1}\left(o(\|x-a\|_E).\right) = x - a = f^{-1}(y) - f^{-1}(b), \tag{5.2}$$

soit

$$g(y) - g(b) = (df(a))^{-1}(y-b) - (df(a))^{-1}\left(o(\|x-a\|_E)\right).$$

Montrons alors que $(df(a))^{-1}\left(o(\|x-a\|_E)\right) = o(\|y-b\|_F)$, ce qui prouvera bien que g est différentiable en b avec $dg(b) = (df(a))^{-1}$.
D'après (5.2) il existe un voisinage ouvert de a et une fonction φ définie dans vérifiant $\lim\limits_{x\longrightarrow a}\varphi(x) = 0$ et tel que

$$\begin{aligned}
\left\|(df(a))^{-1}(y-b)\right\|_E &= \left\|x - a + (df(a))^{-1}\left(\|x-a\|_E\,\varphi(x)\right)\right\|_E \\
&\geqslant \|x-a\|_E - \|x-a\|_E \ \left\|(df(a))^{-1}\left(\varphi(x)\right)\right\|_E \\
&\geqslant \|x-a\|_E\left(1 - \left\|(df(a))^{-1}\left(\varphi(x)\right)\right\|_E\right).
\end{aligned}$$

Or $\lim\limits_{x\longrightarrow a}\varphi(x) = 0$ et $(df(a))^{-1}$ est continue en 0 donc pour $\|x-a\|_E$ suffisament petite, on aura $\left\|(df(a))^{-1}\left(\varphi(x)\right)\right\|_E \leqslant \frac{1}{2}$. On en déduit donc que

$$\|x-a\|_E \leqslant 2\ \left\|(df(a))^{-1}(y-b)\right\|_E \leqslant 2\ \|y-b\|_F\ \left\|(df(a))^{-1}\right\|_{\mathbf{B}(\mathrm{F},\mathrm{E})}$$

et donc que

$$\|x-a\|_E\left\|(df(a))^{-1}\left(\varphi(x)\right)\right\|_E \leqslant 2\ \|y-b\|_F\ \underbrace{\left\|(df(a))^{-1}\right\|_{\mathbf{B}(\mathrm{E},\mathrm{F})}}_{constante}\ \underbrace{\left\|(df(a))^{-1}\left(\varphi(x)\right)\right\|_E}_{\longrightarrow 0 \text{ pour } x\longrightarrow a}.$$

Ceci montre que $\|x-a\|_E\left\|(df(a))^{-1}\left(\varphi(x)\right)\right\|_E = o(\|y-b\|_F)$ et prouve ainsi que f^{-1} est différentiable en $b = f(a)$ avec $df^{-1}(f(a)) = (df(a))^{-1}$.

Montrons alors que df^{-1} est continue.
La formule $df^{-1}(b) = (df(f^{-1}(b)))^{-1}$ valable pour tout $b \in V$, montre que df^{-1} peut se décomposer suivant $df^{-1} = Inv \circ df \circ f^{-1}$ avec

$$\begin{aligned}
&f^{-1} : V \subset F \longrightarrow U \subset E \ \ \text{continue car } f \text{ homéomorphisme}\,; \\
&df : U \subset E \longrightarrow GL(E,F) \ \ \text{continue car } f\ C^1 \\
&Inv : GL(E,F) \longrightarrow GL(F,E) \ \ \text{continue par la Proposition 2.9},
\end{aligned}$$

et elle est donc continue sur V.

■

Remarque 5.2
1. *En dimension finie, on a* $df(x) \in G\mathbb{L}(\mathbb{R}^n)$ *ssi* $J_f(x)$ *inversible ssi* $det(J_f(x)) \neq 0$.
De plus la formule $df^{-1}(f(x)) = (df(x))^{-1}$ *s'écrit en terme de matrices jacobiennes* $J_{f^{-1}}(f(x)) = (J_f(x))^{-1}$.
2. *Si* $f : \mathbb{R} \longrightarrow \mathbb{R}$*, la condition 2 s'écrit* $f^{'}(x) \neq 0$ *pour tout* $x \in U$.
La relation $df^{-1}(f(x)) = (df(x))^{-1}$ *redonne la formule bien connue :*

$$(f^{-1})^{'}(y) = (f^{'})^{-1}(x) = \frac{1}{f^{'}(x)} = \frac{1}{f^{'}(f^{-1}(y))}.$$

5.2 Le théorème d'inversion locale

5.2.1 Motivations

Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction de classe C^1 sur $\mathbb{R}$ et soit $x_0 \in \mathbb{R}$ une valeur donnée telle que $f^{'}(x_0) \neq 0$, disons par exemple pour fixer les idées $f^{'}(x_0) > 0$. On pose $y_0 = f(x_0)$.
On s'intéresse alors aux solutions $(a, b) \in \mathbb{R}^2$ de l'équation $y = f(x)$ i.e. aux couples $(a, b) \in \mathbb{R}^2$ vérifiant $b = f(a)$.
Puisque (x_0, y_0) est une solution de $y = f(x)$, la première question que l'on se pose est la suivante :

lorsque y *est "suffisament proche de* y_0*", peut-on trouver des* $x \in \mathbb{R}$ *tels que* $f(x) = y$ *?*

Un élément de réponse est donné par la constatation suivante :
puisque f est C^1, la fonction $f^{'}$ est continue et comme $f^{'}(x_0) > 0$, il existe un voisinage ouvert V_{x_0} de x_0 tel que $\forall x \in V_{x_0}$, $f^{'}(x) > 0$. Ainsi posant $W_{y_0} = f(V_{x_0})$, on en déduit que $f : V_{x_0} \longrightarrow W_{y_0}$ est continue, strictement croissante et donc bijective de V_{x_0} dans W_{y_0}. On note $f^{-1} : W_{y_0} \longrightarrow V_{x_0}$ la bijection réciproque. Comme conséquence, on a donc :

il existe un voisinage ouvert W_{y_0} *tel que si* $y \in W_{y_0}$ *alors il existe un unique* $x = f^{-1}(y) \in V_{x_0}$ *tel que* $f(x) = y$.

On peut de plus remarquer que grace aux hypothèses faites sur f, l'application $f^{-1} : W_{y_0} \longrightarrow V_{x_0}$ est de classe C^1. Donc $y \mapsto f^{-1}(y) = x$ est C^1 ce qui signifie que "x dépend de manière régulière (C^1) de la donnée y".

C'est ce type de résultats qu'on souhaite généraliser.

5.2.2 Théorème d'inversion locale

La preuve du théorème d'inversion locale nécessite de connaître le théorème suivant qui est un résultat en lui-même.

Théorème 5.2 Théorème de point fixe de Banach
E Banach et $f : E \longrightarrow E$ *une application* contractante *ie.*

$$\exists K > 0,\ \forall x, x' \in E,\ \|f(x) - f(x')\| \leqslant K\|x - x'\|.$$

Alors f *admet un unique point fixe :* $\exists x^* \in E$, $f(x^*) = x^*$.

De plus, posant pour tout $n \in \mathbb{N}^*$, $f^n = f \circ \ldots \circ f$, *on a* $\forall x_0 \in E$, $\lim\limits_{n \longrightarrow +\infty} f^n(x_0) = x^*$.

Preuve du Théorème 5.2.
Pour $x_0 \in E$ donné, on définit par récurrence la suite (x_n) en posant $\forall n \in \mathbb{N}$, $x_{n+1} = f(x_n)$.
Le T.A.F appliqué à f sur $[x_n, x_{n+1}]$ donne, f étant contractante,

$$\|x_{n+1} - x_n\| \leqslant K\|x_n - x_{n-1}\|$$

et de proche en proche, on obtient ainsi

$$\|x_{n+1} - x_n\| \leqslant K^n\|x_1 - x_0\|$$

On en déduit donc que pour tous entiers n et p

$$\|x_{n+p} - x_n\| \leqslant \sum_{q=n}^{n+p-1} \|x_{q+1} - x_q\| \leqslant \|x_1 - x_0\|\frac{1 - K^p}{1 - K}\ K^n \leqslant \frac{\|x_1 - x_0\|}{1 - K}\ K^n$$

Puisque $K \in [0, 1[$, eci montre que la suite (x_n) est de Cauchy dans E qui est supposé complet, donc elle converge vers un certain $a \in E$. La continuité de f et la relation de récurrence montre alors que $a = f(a)$
Concernant l'unicité, on suppose qu'il existe $x_1 \neq x_2$ points fixes de f.
On peut alors appliquer de nouveau le T.A.F à f sur $[x_1, x_2]$ d'où $\|x_1 - x_2\| = \|f(x_1) - f(x_2)\| \leqslant K\|x_1 - x_2\|$ et donc $x_1 = x_2$ puisque $K \in [0, 1[$.

■

Comme conséquence de ce théorème, on a le résultat suivant qui est la forme sous laquelle l'utiliserons.

Corollaire 5.2.1
E Banach, $U \subset E$ ouvert et $C \subset U$ fermé. et $f : U \longrightarrow E$ une application.
Si $f(C) \subset C$ et que f est contractante *sur C alors elle admet un unique point fixe dans C : $\exists x^* \in C$, $f(x^*) = x^*$.*

Preuve du Corollaire 5.2.1.
Utilisant les notations de la preuve du Théorème 5.2, il suffit de noter pour $x_0 \in C$, on a $(x_n) \subset C$ (car $f(C) \subset C$) et la limte obtenue $a \in C$ puisque ce dernier est fermé.

■

FIN COURS 11
On peut maintenant énoncer le résultat principal de ce chapitre.

Théorème 5.3 Théorème d'inversion locale
E, F Banach ($dim(E) = dim(F)$ si dim finie), Ω un ouvert de E et $f : \Omega \longrightarrow F$ une application de classe C^{k+1} sur Ω avec $k \geqslant 0$ entier.
On suppose qu'il existe $x_0 \in \Omega$ tel que $df(x_0) \in \mathrm{G\mathbb{B}(E,F)}$.
Alors il existe $U \subset \Omega$ voisinage ouvert de x_0 et il existe $V \subset F$ voisinage ouvert de $y_0 = f(x_0)$ tels que f soit un C^{k+1} difféomorphisme de U dans V.

En particulier, on a $V = f(U)$ qui est donc ouvert, et de plus pour tout $x \in U$ et $y = f(x) \in V$, on a $df^{-1}(y) = (df(x))^{-1}$.

Preuve du Théorème 5.3.
Il suffit de montrer le résultat pour $k = 0$. Le cas $k \geqslant 1$ résultera alors de la Proposition 5.1 (partie 2).

La preuve se compose de 3 étapes.

Etape 1 : simplification

On commence par simplifier un peu le problème en observant qu'on peut supposer que

$$x_0 = 0, \quad f(0) = 0, \quad E = F \ \text{ et } df(0) = Id_E. \tag{5.3}$$

En effet, si tel n'est pas le cas, il suffit de remplacer f par $x \in E \mapsto (df(x_0))^{-1}(f(x_0+x) - f(x_0)) \in E$ qui vérifie les hypothèses (5.3).

Etape 2 : f est localement un homéomorphisme

On pose pour tout $x \in \Omega$, $g(x) = x - f(x)$ qui vérifie donc $g \in C^1(\Omega)$, $g(0) = 0$ et $dg(0) = 0$.

La continuité de dg en 0 montre que

$$\exists \rho > 0, \ \forall x \in \overline{B_{2\rho}(0)}, \ \|dg(x)\| \leqslant \frac{1}{2}$$

puis le T.A.F à g sur $\overline{B_{2\rho}(0)}$ (convexe) que

$$\forall x, x' \in \overline{B_{2\rho}(0)}, \ \|g(x) - g(x')\| \leqslant \frac{1}{2}\|x - x'\| \tag{5.4}$$

Montrons alors que :

$$\forall y \in B_\rho(0), \ \exists! x \in B_{2\rho}(0), \ y = f(x) \tag{5.5}$$

Pour cela, on utilise le théorème du point fixe vu précédement.

Fixons $y \in B_\rho(0)$ et posons $\forall x \in \Omega$, $h(x) = y + g(x)$ (on a alors $h(x) = x$ ssi $y = f(x)$).

L'application h qui est de classe C^1 vérifie par (5.4) :

- pour tout $x, x' \in \overline{B_{2\rho}(0)}$, $\|h(x) - h(x')\| = \|g(x) - g(x')\| \leqslant \frac{1}{2}\|x - x'\|$
- pour tout $x \in \overline{B_{2\rho}(0)}$, $\|h(x)\| = \|g(x) + y\| \leqslant \|y\| + \frac{1}{2}\|x\| < 2\rho$

En conséquence, h admet un unique point fixe $x \in \overline{B_{2\rho}(0)}$ mais puisque $h(x) = x$ l'inégalité $h(x) < 2\rho$ montre que $x \in B_{2\rho}(0)$ et on a donc (5.5), ce que l'on peut réécrire sous la forme

$$\forall y \in B_\rho(0), \ \exists! x \in B_{2\rho}(0) \cap f^{-1}(B_\rho(0)), \ y = f(x) \tag{5.6}$$

Ainsi posant $U = B_{2\rho}(0) \cap V$ et $V = f^{-1}(B_\rho(0))$, on obtient que f est bijective de U sur V.

Notons $f^{-1} : V \longrightarrow U$ son inverse locale.

Puisque pour tout $y, y' \in V$, avec $y = f(x)$ et $y' = f(x')$ elle vérifie $f^{-1}(y) - f^{-1}(y') = g(x) - g(x') + y - y'$, on déduit de (5.4) que $\|f^{-1}(y) - f^{-1}(y')\| \leqslant 2\|y - y'\|$.

Ainsi f^{-1} est continue sur V et donc f, qui est elle même continue, définit un homéomorphisme de U sur V.

Etape 3 : régularité C^1 de l'inverse local

On montrer que $f^{-1} \in C^1(V)$, ce qui compte tenu du Théorème 5.1, revient à prouver que $\forall x \in U$, $df(x) \in G\mathbb{B}(E)$.

Or $df(0) \in G\mathbb{B}(E)$ qui est ouvert dans $\mathbb{B}(E)$ et donc $\exists r_1 > 0$ tel que $B_{r_1}(df(0)) \subset G\mathbb{B}(E)$.

Mais par continuité de df en 0 (f est C^1) on a $\exists r_2 > 0$, $\forall x \in B_{2r_2}(0)$, $\|df(x) - df(0)\| \leqslant r_1$.

Ainsi, $\exists r_2 > 0$, $\forall x \in B_{2r_2}(0)$, $df(x) \in G\mathbb{B}(E)$.

On obtient donc le résultat en prenant ρ dans l'étape 2 éventuellement plus petit (le remplacer par $Min(\rho, r_2)$ et adapter les définitions de U et V).

■

On peut se demander à quelle condition le théorème devient global.
Si on prend $f : \mathbb{R} \longrightarrow R$ vérifiant les hypothèses du T.I.L, on voit que l'homéomorphisme local deviendra global dès que f ne change pas de monotonie, ce qui est le cas dès qu'elle est injective. Si de plus, on veut la régularité C^1 de la bijection réciproque il faut de plus que f' ne s'annule pas.
C'est exactement ce que donne dans un cadre plus général le résultat suivant.

Théorème 5.4 Théorème d'inversion globale
E, F Banach ($dim(E) = dim(F)$ si dim finie), $U \subset E$ ouvert et $f : U \longrightarrow F$ une application C^1.
Les deux propriétes suivantes sont alors équivalentes :
1) *f est un difféomorphisme de U dans l'ouvert $V = f(U)$ de F.*
2) *f est injective sur U et pour tout $x \in U$, $df(x) \in G\mathbb{B}(E,F)$.*

Preuve du Théorème 5.4.
• *1) $\Longrightarrow$ 2)* est évidente.
• Montrons *2) $\Longrightarrow$ 1)*.
On suppose donc que f est injective sur U et pour tout $x \in U$, $df(x) \in G\mathbb{B}(E,F)$.
Montrons dans un premier temps que $f(U)$ est ouvert.
Montrons qu'en fait l'application f est ouverte (ie. $f(U_0)$ ouvert de F pour tout U_0 ouvert de E).
Soit U_0 ouvert de E.
Alors $f(U_0)$ est ouvert si $\forall y_0 \in f(U_0)$, $\exists V_1 \subset f(U_0)$ voisinage ouvert de y_0.
Fixons $y_0 \in f(U_0)$, donc $\exists x_0 \in U_0$, $y_0 = f(x_0)$ et puisque $U_0 \subset U$, on a $x_0 \in U_0$, $f \in C^1(U_0)$ et $df(x_0) \in G\mathbb{B}(E,F)$.
Le T.I.L montre alors qu'il existe $U_1 \subset U_0$ voisinage ouvert de x_0, $V_1 = f(U_1) \subset f(U_0)$ voisinage ouvert de $f(x_0)$ tels que f soit un difféomorphisme de U_1 sur V_1.
Donc f est ouverte et en particulier, U implique $f(U)$ ouvert.

Si f est injective sur U elle est évidement bijective de U sur $f(U)$. Notons $g : f(U) \longrightarrow U$ sa bijection réciproque.
Puisque pour tout $U_0 \subset U$ ouvert, on a $g^{-1}(U_0) = \{y \in V;\ g(y) \in U_0\} = f(U_0)$ qui est ouvert, g est donc continue sur V.
Résumons : $f \in C^1$ est bijective de U sur $f(U)$ d'inverse continue sur $f(U)$ ie. f homéomorphisme C^1 de U sur V.
Puisque pour tout $x \in U$, $df(x) \in G\mathbb{B}(E,F)$, on déduit le résultat (f difféo) du théorème 5.1.

■

5.3 Théorème des fonctions implicites

Avant d'énoncer le résultat principal de cette Section nous avons besoin d'une notion, certes naturelle, mais non encore vue jusqu'à maintenant.

5.3.1 Différentielle partielle

Introduisons donc E_1, E_2 et F trois e.v.n, U un ouvert de $E_1 \times E_2$ et $f : U \longrightarrow F$ une application différentiable.
Pour $(x_1, x_2) \in U$, l'ensemble $U_{x_1} = \{y \in E_2,\ (x_1, y) \in U\}$ est un ouvert de E_2 contenant x_2.
La fonction définie par $f_{x_1} : \begin{array}{ccc} U_{x_1} & \longrightarrow & F \\ y & \mapsto & f(x_1, y) \end{array}$ est alors différentiable sur U_{x_1}.
En effet, elle peut se décomposer en $f_{x_1} = f \circ \psi$ où $\psi : y \in U_{x_1} \mapsto (x_1, y) \in E_1 \times E_2$ est différentiable car chaque composante l'est ($y \mapsto x_1$ constante et $y \mapsto y$ linéaire), et f qui est différentiable sur $\psi(U_{x_1}) \subset U$ (par définition de U_{x_1}). Ainsi, pour tout $x_2 \in U_{x_1}$ on a

(5.7)
- $df_{x_1}(x_2) \in \mathbb{B}(E_2, F)$
- $\forall h_2 \in E_2,\ df_{x_1}(x_2)(h_2) = df(x_1, x_2) \circ d\psi(x_2)(h_2) = df(x_1, x_2)(0, h_2)$
- l'application $(x_1, x_2) \mapsto df_{x_1}(x_2)$ est définie sur U en entier.

Définition 5.2 - Différentielle partielle
L'application $(x_1, x_2) \mapsto df_{x_1}(x_2) \in \mathbb{B}(E_2, F)$ est appelée différentielle partielle de f relativement à sa seconde variable. On la note $\partial_2 f$.
On définit de même l'application $(x_1, x_2) \mapsto df_{x_2}(x_1) \in \mathbb{B}(E_1, F)$ qui est appelée différentielle partielle de f relativement à sa première variable. On la note $\partial_1 f$.

On a alors le résultat suivant.

Lemme 5.1
On suppose que $f : U \subset E_1 \times E_2 \longrightarrow F$ est différentiable.
Alors pour tout $(x_1, x_2) \in U$ et tout $h = (h_1, h_2) \in E_1 \times E_2$, on a

$$df(x_1, x_2)h = \partial_1 f(x_1, x_2)h_1 + \partial_2 f(x_1, x_2)h_2$$

Preuve du Lemme 5.1.
Pour tout $(x_1, x_2) \in U$ et tout $(h_1, h_2) \in E_1 \times E_2$, on a par linéarité de $df(x_1, x_2)$ et par (5.7) :
$df(x_1, x_2)(h_1, h_2) = df(x_1, x_2)(h_1, 0) + df(x_1, x_2)(0, h_2) = \partial_1 f(x_1, x_2)h_1 + \partial_2 f(x_1, x_2)h_2$.
■

5.3.2 Enoncé du théorème des fonctions implicites

On a maintenant tous les éléments pour énoncer le théorème des fonctions implicites.

Théorème 5.5 Théorème des fonctions implicites
Soit E_1, E_2 et F des Banach vérifiant $dim(E_2) = dim(F)$ si la dimension est fine, U un ouvert de $E_1 \times E_2$, et $f : U \longrightarrow F$ une application de classe C^{k+1} sur U pour un $k \geqslant 0$ entier.
On suppose qu'il existe $(x_0, y_0) \in U$ tel que

(i) $f(x_0, y_0) = 0$;
(ii) $\partial_2 f(x_0, y_0) : E_2 \longrightarrow F$ *est inversible i.e.* $\partial_2 f(x_0, y_0) \in G\mathbb{B}(E_2, F)$
(soit encore $h_2 \in E_2 \mapsto df(x_0, y_0)(0, h_2) \in F$ inversible).

Alors | *il existe $U_1 \subset E_1$ ouvert contenannt x_0 ;*
| *il existe $U_2 \subset E_2$ ouvert contenannt y_0 ;*
| *il existe $\varphi : U_1 \longrightarrow U_2$ application de classe C^{k+1} sur U_1 (φ : fct implicite) ;*

tels que

$$\begin{pmatrix} (x, y) \in U_1 \times U_2 \\ f(x, y) = 0 \end{pmatrix} \Longleftrightarrow \begin{pmatrix} x \in U_1 \\ y = \varphi(x) \end{pmatrix} \tag{5.8}$$

En particulier, puisque $f(x_0, y_0) = 0$, l'implication ($\Longrightarrow$) montre que $y_0 = \varphi(x_0)$.

Preuve du Théorème 5.5.
On va appliquer le Théorème 5.1 à la fonction g définie par $g : \begin{array}{ccc} U & \longrightarrow & E_1 \times F \\ (x, y) & \mapsto & (x, f(x, y)) \end{array}$.
Elle est de classe C^{k+1} sur U car chacune de ses composantes l'est.
De plus sa différentielle est donnée pour tout $(x, y) \in U$ et tout $(h, k) \in E_1 \times E_2$ par $dg(x, y)(h, k) = (h, df(x, y)(h, k)) \in E_1 \times F$ et on déduit du Lemme 5.1 que

$$dg(x_0, y_0)(h, k) = \begin{pmatrix} id_{E_1} & 0_{E_2} \\ \partial_1 f(x_0, y_0) & \partial_2 f(x_0, y_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

Ainsi, pour tout $(a_1, a_F) \in E_1 \times F$, on a

$$\begin{aligned} dg(x_0, y_0)(h, k) = (a_1, a_F) &\Longleftrightarrow \begin{pmatrix} id_{E_1} & 0_{E_2} \\ \partial_1 f(x_0, y_0) & \partial_2 f(x_0, y_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} a_1 \\ a_F \end{pmatrix} \\ &\Longleftrightarrow \left\{ \begin{array}{l} h = a_1 \\ \partial_1 f(x_0, y_0)(h) + \partial_2 f(x_0, y_0)(k) = a_F \end{array} \right. \\ &\Longleftrightarrow \left\{ \begin{array}{l} h = a_1 \\ k = -\left(\partial_2 f(x_0, y_0)\right)^{-1} \left(\partial_1 f(x_0, y_0)(a_1)\right). \end{array} \right. \end{aligned}$$

Ceci montre que $dg(x_0, y_0)$ est un isomorphisme de $E_1 \times E_2$ sur $E_1 \times F$.
On peut donc appliquer le Théorème d'inversion locale (Théorème 5.1) à g avec (x_0, y_0), et donc puisque $f(x_0, y_0) = 0$,

| il existe $U_1' \times U_2 \subset U \subset E_1 \times E_2$ voisinage ouvert de (x_0, y_0) ;
| il existe $W \subset E_1 \times F$ voisinage ouvert de $g(x_0, y_0) = (x_0, 0)$;
| tels que $g : U_1' \times U_2 \longrightarrow W = g(U_1' \times U_2)$ soit un C^{k+1} difféomorphisme.

Notons $g^{-1} : W \longrightarrow U_1' \times U_2$ l'inverse local de g.

Alors compte tenu de l'expression de g, son inverse (local) est de la forme $g^{-1}(x,z) = (g_1^{-1}(x,z), g_2^{-1}(x,z)) = (x, g_2^{-1}(x,z))$ pour tout $(x,z) \in W$.
Définissons alors

$$U_1 = \{x \in U_1', \ (x,0) \in W\}$$

qui est donc un voisinage ouvert de a_1 (puisque $a_1 \in U_1'$ et que $(a_1, 0) \in W$ qui est tout vert) puis

$$\begin{array}{rcl} \varphi : U_1 & \longrightarrow & U_2 \\ x & \mapsto & g_2^{-1}(x,0) \end{array}$$

Alors on a $\varphi \in C^{k+1}(U_1)$ et pour tout $(x,y) \in U_1 \times E_2$, $f(x,y) = 0$ ssi $y = \varphi(x)$.

En effet, pour $(x,y) \in U_1 \times U_2$, on a compte tenue de la forme de g^{-1}

$$f(x,y) = 0 \Longleftrightarrow g(x,y) = (x,0) \Longleftrightarrow (x,y) = g^{-1}(x,0) \Longleftrightarrow y = g_2^{-1}(x,0) = \varphi(x)$$

ce qui conclut la preuve du Théorème 5.5.

■

Terminons ce cours par un corollaire direct du théorème des fonctions implicites.

Corollaire 5.5.1
Les hypothèses sont celles du Théorème 5.5. On a a lors les trois résultats suivants :

1. *La relation (5.8) entraine que :*

$$\forall x \in U_1, \quad f(x, \varphi(x)) = 0. \tag{5.9}$$

2. *On peut supposer que les ouverts U_1 et U_2 obtenus dans le Théorème 5.5 sont tels que*

$$\forall (x,y) \in U_1 \times U_2, \quad \partial_2 f(x,y) \in G\mathbb{B}(E_2, F). \tag{5.10}$$

3. *La différentielle de la fonction implicite est donnée pour tout $x \in U_1$ et tout $h \in E_1$ par*

$$d\varphi(x)h = -\left(\partial_2 f(x, \varphi(x))\right)^{-1} \circ \partial_1 f(x, \varphi(x))(h). \tag{5.11}$$

Preuve du Corollaire 5.5.1.
• Partie *1.* La relation (5.9) est contenue dans (5.8).

• Partie *2.* C'est une conséquence du fait que $\partial_2 f(x_0, y_0) \in G\mathbb{B}(E_2, F)$ ouvert de $\mathbb{B}(E_2, F)$ et que f étant $C^1(U)$, on a $\partial_2 f : U \longrightarrow \mathbb{B}(E_2, F)$ continue en $(x_0, y_0) \in U_1 \times U_2$ ouvert (mêmes arguments que dans le preuve du T.I.L étape2).

• Partie *3.* Il suffit de différencier la relation (5.9) en constatant que $g : x \in U_1 \subset E_1 \mapsto g(x) = f(x, \varphi(x)) = f \circ \psi(x)$ avec $\psi : x \mapsto (x, \varphi(x)) \in U_1 \times U_2 \subset U$. On trouve ainsi pour tout $h \in E_1$, $dg(x)h = df(\psi(x)) \circ d\psi(x)h = 0$, soit

$$\begin{aligned} df(x,\varphi(x)) \circ (h, d\varphi(x)(h)) = 0 &\Longleftrightarrow df(x,\varphi(x))(h,0) + df(x,\varphi(x))(0, d\varphi(x)(h)) = 0 \\ &\Longleftrightarrow \partial_2 f(x,\varphi(x))(d\varphi(x)(h)) = -\partial_1 f(x, \varphi(x)(h) \end{aligned}$$

et on conclut alors grace à (5.10).

■

REMARQUES
1) On suppose que $E_1 = \mathbb{R}^n$ et $E_2 = F = \mathbb{R}^p$.

$$
\begin{array}{llr}
(1) : \partial_2 f(x_0, y_0) \in G\mathbb{L}(\mathbb{R}^p) & \Longleftrightarrow J_{f_{x_0}}(y_0) \in M_p \text{ inversible} & \\
& \Longleftrightarrow J_f(x_0, y_0) \in M_{n+p,p} \text{ admet une sous matrice de rang } p & (2) \\
& \Longleftrightarrow df(x_0, y_0) \text{ surjective} & (3) \\
& \Longleftrightarrow (\nabla f_i(x_0, y_0))_{i=1..p} \text{ est une famille libre} & (4) \\
& \quad \text{ie. } \forall h \in \mathbb{R}^n, \displaystyle\sum_{i=1}^{p} \lambda_i < h, \nabla f_i(x_0, y_0) >= 0 \Longrightarrow (\lambda_i)_i = 0 &
\end{array}
$$

On peut donc reformuler le Théorème 5.5 en remplaçant l'hypothèse (1) par (2), (3) ou (4).
2) Il se peut qu'on ait à appliquer le Théorème 5.5 avec $\Omega \subset \mathbb{R}^N$ et $f : \Omega \longrightarrow \mathbb{R}$, $g : \Omega \longrightarrow \mathbb{R}^p$ avec $N > p$ et par exemple, $df(X_0)$ surjective. Dans ce cas, on écrit $\mathbb{R}^N = \mathbb{R}^{N-p} \times \mathbb{R}^p = E_1 \times E_2 \times \mathbb{R}^p$ avec $X_0 = (x_0, y_0) \in \mathbb{R}^{N-p} \times \mathbb{R}^p$.

5.3.3 Application du Théorème 5.5 aux variétés différentiables

Pour $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ $(n > p)$, $\Omega \subset \mathbb{R}^n$ ouvert, $g \in C^1(\Omega)$, on considère l'ensemble $U = \{x \in \Omega,\ g(x) = 0\} = g^{-1}(\{0\})$: fermé (courbe de niveau 0 de g).
Si $\exists a \in \Omega$ tq $dg(a)$ est surjective, écrivant $\mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$, on déduit du théorème des fonctions implicites que

$$
\left|
\begin{array}{l}
\exists U_1 \subset E_1 = \mathbb{R}^{n-p} \text{ ouvert contenant } a_1\,; \\
\exists U_2 \subset E_2 = \mathbb{R}^p \text{ ouvert contenant } a_2\,; \\
\exists \varphi : U_1 \longrightarrow U_2 \text{ de classe } C^1 \text{ sur } U_1\,;
\end{array}
\right.
\quad \text{tels que } \forall x \in U_1 \times U_2,\ (g(x) = 0 \Longleftrightarrow x_2 = \varphi(x_1)).
$$

Dans ce cas, on donc $U \cap (U_1 \times U_2) = \{(x_1, x_2) \in U_1 \times U_2,\ x_2 = \varphi(x_1)\}$ ie. U est localement une variété différentiable (C^1) au voisinage de a.

Définition 5.3
Soit la fonction $\Psi : \mathbb{R}^{n-p} \longrightarrow \mathbb{R}^n$, $\Psi \in C^1(U_1)$ définie par $\Psi(x_1) = (x_1, \varphi(x_1))$.
On dit que (Ψ, U_1) est une paramétrisation locale de U au vosinage de a.

On a alors le résultat suivant

Proposition 5.2
On note $V = U \cap (U_1 \times U_2)$ et $T_a(V)$ la variété linéaire tangente à V en a (voir Définition 2.10 et Théorème 2.4). Sous les hypothèses précédentes, on a

$$
\begin{array}{llll}
(5.12) & T_a(V) & = & \{(h_1, h_2) \in \mathbb{R}^{n-p} \times \mathbb{R}^p,\ h_2 = d\varphi(a_1)h_1\} \\
(5.13) & & = & Ker(dg(a)) = \displaystyle\bigcap_{i=1}^{p} Ker(dg_i(a)) \\
(5.14) & & = & Im(d\Psi(a))
\end{array}
$$

Noter qu'en particulier (5.13) montre que $dim(T_a(V)) = n - p$ (puisque $dg(a)$ surjective).

Preuve de la Proposition 5.2
Les notations sont celles utilisées au début de ce paragraphe.
• L'égalité (5.12) est donnée par le Théorème 2.4.
• Egalité (5.13). On a $(h,k) \in Ker(dg(a))$ ssi $\partial_1 g(a)h + \partial_2 g(a)k = 0$ ssi $k = -\Big(\partial_2 g(a)\Big)^{-1}\partial_1 g(a)h = d\varphi(a)h$ donc ssi $(h,k) \in T_a(V)$.
La seconde égalité vient de ce que $dg(a)h = 0$ ssi $dg_i(a)h = 0$ pour tout $i = 1,..,p$.
• Egalité (5.14). Par surjectivité, on a $dim(Ker(dg(a))) = n - p$.
Or $d\psi(a)$ est injective puisque $d\psi(a)h = 0 \Longrightarrow h = 0$ et donc $dim(Im(d\psi(a))) = n - p$.
Mais puisque pour tout $x_1 \in U_1$, $g(\psi(x_1)) = 0$ on a en différenciant, $dg(\psi(a_1))d\psi(a_1) = 0$ ie. $Im(d\psi(a_1)) \subset Ker(dg(a))$.
Ainsi, $Im(d\psi(a_1))$ est un s.e.v de $Ker(dg(a))$ et les deux ayant même dimension, ils sont égaux.
■

Remarque 5.3
Lorsque $p = 1$, la proposition 5.2 montre que $\nabla g(a)$ est orthogonal à la courbe de niveau $\{g = g(a)\}$ en a.

5.4 Extrema locaux liés

Le problème est le suivant :
$\Omega \subset \mathbb{R}^n \times \mathbb{R}^p$ ouvert, $f : \Omega \longrightarrow \mathbb{R}$, $g : \Omega \longrightarrow \mathbb{R}^p$ avec $f, g \in C^1(\Omega)$, $U = \{x \in \Omega,\ g(x) = 0\}$:

$$\text{Trouver } a \in U \text{ tel que } f(a) = \min_{x\in U} f(x) \text{ (ou max)}, \qquad \text{(P)}$$

on dit que l'on cherche à optimiser f sous la contrainte (**s.c**) U.
Il s'agit d'un problème d'extrema liés (la liaison étant donnée par U).
Remarque 5.4
Attention, ce pb est vraiment différent de celui de la recherche d'extrema libres puisqu'ici U est fermé (g continue). Pour s'en persuader, considérer $f(x) = x \in \mathbb{R}$ s.c $x \in [0,1]$: l'étude des points critiques ne donne rien.

5.4.1 Première approche : méthode de substitution

Si U peut être décrit par une relation de la forme

$$(x \in \Omega, g(x) = 0) \Longleftrightarrow (x_1 \in W,\ x_2 = h(x_1))$$

avec h, W **explicites** et ayant de suffisament bonnes propriétés (typiquement h régulière, W ouvert), alors par substitution, on aura $\min_{x\in U} f(x) = \min_{x_1\in W} f(x_1, h(x_1)))$, ce qui dans les bon cas, se ramène à ce qu'on sait faire (extrema libres).
Exercice
Déterminer sur $\mathbb{R}^2$, les extrema locaux de $f(x,y) = x^2 + y^2$ s.c $y - x + 1 = 0$.

Lorsque cette méthode ne peut pas être utiliser (trop compliquée ou trop implicite ou impossible), on peut chercher à appliquer la méthode dire des multiplicateurs de Lagrange (qui peut être vue comme un raffinement de la méthode de substitution) que nous allons maintenant décrire.

5.4.2 Seconde approche : méthode de Lagrangien

Elle repose sur le résulat suivant

Théorème 5.6
Soit $\Omega \subset \mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$ *ouvert,* $f : \mathbb{R}^n \longrightarrow R$, $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $f, g \in C^1(\Omega)$ *et* $U = g^{-1}(\{0\})$.
Si $a \in U$ *est un extremum de* f *s.c* U *et que* $\partial_2 g(a) \in G\mathbb{B}(\mathbb{R}^p)$ *alors il existe* $\lambda_a \in \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ *tel que*

$$df(a) + \lambda_a dg(a) = 0 \tag{5.15}$$

L'application λ_a *est appelée multiplicateur de Lagrange associé* f *s.c* U.
Puisque $\mathbb{B}(\mathbb{R}^p, \mathbb{R}) \sim \mathbb{R}^p$*, la conclusion équivaut donc à : il existe* p *scalaires* λ_i $(i = 1, .., p)$ *tels que*

$$df(a) + \sum_{i=1}^{p} \lambda_i dg_i(a) = 0.$$

Preuve du Théorème 5.6.
Grace aux hypothèses faites sur g, on sait que U est localement une variété C^1, notée V, au voisinage de a.
Alors, chercher les extrema de f restreinte à V s.c U revient à chercher les extrema de $F : x_1 \mapsto f(x_1, \varphi(x_1))$ sur l'ouvert U_1.
Il s'agit donc d'un problème d'extrema libre pour F.
Or si $a \in V$ est un extremum de f s.c U alors $a_1 \in U_1$ est un extremum de F sur U_1 et c'est donc en particulier un point critique de F par le Théorème 4.7 : $dF(a_1) = 0$.
Or, puisque $F = f \circ \psi$ et que $\forall x_1 \in U_1$, $\psi(x_1) = (x_1, \varphi(x_1))$ on a

$$dF(a_1) = \partial_1 f(\psi(a_1)) + \partial_2 f(\psi(a_1)) d\varphi(a_1)$$

et donc si $dF(a_1) = 0$, utilisant la relation donnant la différentielle de la fonction implicite φ associée à g : $d\varphi(a_1) = -\Big(\partial_2 g(a)\Big)^{-1} \partial_1 g(a)$, on a

$$\partial_1 f(\psi(a_1)) = -\partial_2 f(\psi(a_1)) d\psi(a_1) = \partial_2 f(\psi(a_1)) \Big(\partial_2 g(a)\Big)^{-1} \partial_1 g(a)$$

et écrivant de plus que

$$\partial_2 f(\psi(a_1)) = \partial_2 f(\psi(a_1)) = \partial_2 f(\psi(a_1)) \Big(\partial_2 g(a)\Big)^{-1} \partial_2 g(a)$$

on trouve la relation $df(a) = -\lambda_a dg(a)$ en posant $\lambda_a = -\partial_2 f(\psi(a_1)) \Big(\partial_2 g(a)\Big)^{-1}$.

■

Remarque 5.5
1) La conclusion du théorème 5.6 dit qu'en un point extremal, $df(a) \in \Big(Ker(dg(a))\Big)^{\perp}$ *ou en d'autres termes que* $df(a) = 0$ *sur* $Ker(dg(a))$ *ie. sur* $T_a(V)$.
2) Lorsque $p = 1$*, la conclusion s'crit :* $\nabla f(a)$ *et* $\nabla g(a)$ *colinéaires ce qui conforte l'intuition (faire un dessin).*

Corollaire 5.6.1
On se place sous les hypothèses du Théorème 5.6 et pour $(x, \lambda) \in \Omega \times \mathbb{B}(\mathbb{R}^p, \mathbb{R})$*, on définit le Lagrangien (de* f *s.c* U*) par*

$$L(x, \lambda) = f(x) + \lambda g(x).$$

Si $a \in U$ *est un extremum de* f *s.c* U *pour le quel* $\partial_2 g(a) \in G\mathbb{B}(\mathbb{R}^p)$ *alors* $\exists \lambda_a \in \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ *tel que* (a, λ_a) *est un point critique de* L

Preuve du Corollaire 5.6.1.
Il suffit que remarquer qu'un point critique du Lagrangien L qui est une solution $(x, \lambda) \in \Omega \times \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ de $dL(x, \lambda) = 0$ vérifie donc $\partial_1 L(x, \lambda) = 0$ et $\partial_2 L(x, \lambda) = 0$, c'est-à-dire respectivement $df(x) + \lambda dg(x) = 0$ et $g(x) = 0$.

■

Remarque 5.6
1) Les extrema de f *s.c* U *pour lesquels* $dg(a)$ *est sujective sont donc à chercher parmi les points critiques du Lagrangien* L*.*
2) En terme de dérivées partielles, un point critique $(x, \lambda) \in \Omega \times \mathbb{R}^p$ *de* L *est solution de* $dL(x, \lambda) = 0$*, soit*

$$\begin{cases} \nabla f(x) + \displaystyle\sum_{i=1}^{p} \lambda_i \nabla g_i(x) = 0 & : \quad n - equations \\ g(x) = 0 & : \quad p - equations \end{cases}$$

On doit donc résoudre un système (non linéaire) de (n+p) inconnues à (n+p) équations. C'est en quelque sorte le prix à payer pour se ramener à ce qu'on sait faire (extrema libres).

Comme dans le cas des extrema libres, il faut pouvoir, une fois les candidats potentiels déterminés, savoir ce qu'on peut en dire. On a le résultat suivant.

Théorème 5.7
On se place toujours sous les hypothèses du théorème 5.6 et on suppose de plus que f *et* g *sont deux fois différentiables sur* Ω*.*
On suppose qu'on connait un point critique (a, λ) *du Lagrangien* L *pour lequel* $\partial_2 g(a) \in G\mathbb{B}(\mathbb{R}^p)$*.*
Si l'application bilinéaire symétrique $d^2 f(a) + \lambda d^2 g(a)$ *restreinte à* $T_a(V) = Ker(dg(a))$ *est :*
- définie positive alors a *est un minimum de* f *s.c* U
- définie négative alors a *est un maximum de* f *s.c* U
- ni positive, ni négative alors a *n'est pas un extremum de* f *s.c* U
- dans les autres cas (positive ou négative mais non défine) on ne peut rien dire directement.

Preuve du Théorème 5.7.
On a vu dans la preuve du théorème 5.6 que sous les hypothèses faites sur g, l'optimisation sur V de f s.c U équivaut à la recherche des extrema de la fonction $F = f \circ \psi$ sur l'ouvert U_1.
Or, si on connait un point critique (a, λ) du Lagrangien L alors $a_1 \in U_1$ est un point critique de F puisque :

$$\begin{aligned} dF(a_1) &= df(a)d\psi(a_1) \\ &= -\lambda dg(a)d\psi(a_1) \\ &= -\lambda(\partial_1 g(a) + \partial_2 g(a)d\varphi(a_1)) \\ &= 0 \end{aligned}$$

puisque $d\varphi(a_1) = -\Big(\partial_2 g(a)\Big)^{-1}\partial_1 g(a)$.
Ainsi, on peut appliquer les théorèmes 4.9 et 4.8 à F et il faut donc pour cela étudier la forme bilinéaire symétrique $d^2F(a_1)$.
Or, puisque $df(a) + \lambda dg(a) = 0$, on a

$$d^2F(a_1) = d^2f(a)(d\psi(a_1))^2 + df(a)d^2\psi(a_1) = d^2f(a)(d\psi(a_1))^2 - \lambda dg(a)d^2\psi(a_1) \tag{5.16}$$

mais puisque pour tout $x_1 \in U_1$, $g(\psi(x_1)) = 0$, on trouve par ailleurs que $dg(\psi(x_1))d\psi(x_1) = 0$, et donc $d^2g(a)(d\psi(a_1))^2 + dg(a)d^2\psi(a_1) = 0$, soit $dg(a)d^2\psi(a_1) = -d^2g(a)(d\psi(a_1))^2$ et il suffit alors de remplacer dans (5.16) pour trouver que

$$d^2F(a_1) = \Big(d^2f(a) + \lambda d^2g(a)\Big)(d\psi(a_1))^2.$$

Il suffit alors pour conclure, d'appliquer les théorèmes 4.9 et 4.8 à $d^2F(a_1)$. ■

$\boxed{\text{FIN DU CHAPITRE 5}}$

FIN DU COURS
DE
CALCUL DIFFERENTIEL