Janvier 2017

Cours de Calcul Différentiel

Licence de Mathématiques

Semestre 6

Cédric Dupaix

Université de Franche-Comté

Table des matières

Préambule

Objectifs de ce cours :

- genéraliser la notion de dérivée (fonction d'une variable réelle) et introduire les règles algébriques de calcul rendant aisée l'utilisation de cette (nouvelle) notion ;
- reprendre en les étendant à ce nouveau cadre, les résultats de base vus pour les dérivées : théorème des accroissements finis, formules de Taylor, théorème de la bijection

Cadre

Tous les espaces vectoriels (e.v) considérés dans ce cours seront des $\mathbb{R}$-e.v.
Sauf exception, on se placera en dimension infinie d'espace, les énoncés et les preuves n'étant en général, pas plus compliquées.

Index des principales notations utilisées dans ce cours :

- ⋆ $E = (E, \| \ \|_E)$, $F = (F, \| \ \|_F)$ des $\mathbb{R}$-e.v.n
- ⋆ $B_\rho(x_0) = \{x \in E,\ \|x - x_0\| < \rho\} \subset E$: boule ouverte de centre $x_0 \in E$, de rayon $\rho > 0$
- ⋆ $\mathscr{L}(E, F)$: e.v des applications linéaires de E dans F (avec $\mathscr{L}(E) = \mathscr{L}(E, E)$)
- ⋆ $\mathbb{B}(E, F)$: e.v des applications linéaires et continues de E dans F (avec $\mathbb{B}(E) = \mathbb{B}(E, E)$)
- ⋆ Pour $L \in \mathbb{B}(E, F)$:
  ⋄ $\|L\|_{\mathbb{B}} = \|L\|_{\mathbb{B}(E,F)}$: norme de l'application linéaire L
  ⋄ $\forall h \in E$, Lh ou $L.h$ la quantité $L(h)$
- ⋆ $\mathbb{B}^n(E_1, \dots, E_n; F)$: e.v des applications n-linéaires continues de $E_1 \times \cdots \times E_n$ dans F ($\mathbb{B}^n(E; F) = \mathbb{B}^n(E, \dots, E; F)$)
- ⋆ $\|L\|_{\mathbb{B}^n} = \|L\|_{\mathbb{B}^n(E_1,\dots,E_n;F)}$: norme de l'application n-linéaire L
- ⋆ $\mathrm{Isom(E,F)} \subset \mathbb{B}(E, F)$: ensemble des isomorphismes de E dans F
- ⋆ $\mathscr{M}_{n,p} = \mathscr{M}_{n,p}(\mathbb{R})$ e.v des matrices réelles à n-lignes et p colonnes
- ⋆ Pour $f : E \longrightarrow F$:
  ⋄ $df(a)h$: différentielle de f en $a \in E$ prise en $h \in E$
  ⋄ $df^n(a)(h)^n$: différentielle n-ième de f en $a \in E$ prise au n-uplet $(h)^n := (h, \dots, h) \in E^n$
  ⋄ $f'(a)(h)$: dérivée de f en $a \in E$ dans la direction $h \in E\backslash\{0\}$
  ⋄ $\partial_1 f(a_1, a_2) \in \mathbb{B}(E_1, F)$: différentielle partielle en $(a_1, a_2) \in E_1 \times E_2$ de $f : E_1 \times E_2 \longrightarrow F$ par rapport à sa première variable
- ⋆ Pour $f : \mathbb{R}^n \longrightarrow F$:
  ⋄ $\frac{\partial f}{\partial x_j}(a)$: dérivée partielle de f suivant sa j-ième variable en $a \in \mathbb{R}^n$
- ⋆ Pour $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$:
  ⋄ $J_f(a) \in \mathscr{M}_{p,n}$: matrice jacobienne de f en a
  ⋄ $\nabla f(a) \in \mathbb{R}^n$: vecteur gradient de f en a (cas $p = 1$)
  ⋄ $H_f(a) \in \mathscr{M}_n$: matrice hessienne de f en a (cas $p = 1$)
- ⋆ $Inv(u) = u^{-1}$pour $u \in \mathrm{Isom(E,F)}$.

Chapitre I

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.

1 Définitions

Commençons par rappeler ce qui est déjà connu.

A Cas des fonctions de $\mathbb{R} \longrightarrow \mathbb{R}^n$: notion de dérivée

Définition I.1 (dérivée)

Soit $f : \mathbb{R} \longrightarrow \mathbb{R}^n$ une fonction définie sur un ouvert $U \subset \mathbb{R}$.
On dit que f est dérivable au point $t_0 \in U$ s'il existe $\ell \in \mathbb{R}^n$ telle que

$$\lim_{h \longrightarrow 0} \frac{f(t_0+h) - f(t_0) - \ell h}{h} = 0. \tag{I.1}$$

On note $f'(t_0) = \ell \in \mathbb{R}^n$ la dérivée de f en t_0 et on dit que f est dérivable sur U si elle l'est en tout t_0 de U.
$\diamond$ La convergence dans $\mathbb{R}^n$ étant équivalente à la convergence par composantes, notant $f = (f_i)$, la relation (I.1) montre que f est dérivable en t_0 si et seulement si pour chaque i, f_i l'est et alors $f'(t_0) = (f_i'(t_0))$.

Remarque I.2

1) On a (I.1)$\Longleftrightarrow$ $f(t_0+h) = f(t_0) + \ell h + o(h)$ au voisinage de $h = 0$: f est dérivable en t_0 si et seulement si 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 ne s'étend pas directement au cas $f : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ avec $n > 1$: on ne peut pas diviser par $h \in \mathbb{R}^n$ et surtout quel sens peut-on donner à ℓh ?
C'est ce dernier point qui conduit à la notion de différentielle que nous allons maintenant introduire.

B Cas des fonctions de $E \longrightarrow F$: notion de différentielle

Dans la suite E, F sont deux e.v.n (de dimension finie ou non), $U \subset E$ un ouvert et $f : U \longrightarrow F$ une application.

Définition I.3 (différentielle)

On dit que f est différentiable en $a \in U$ s'il existe une application linéaire *et continue* $L : E \longrightarrow F$ telle que

$$\lim_{h \longrightarrow 0} \frac{f(a+h) - f(a) - Lh}{\|h\|_E} = 0, \tag{I.2}$$

c'est-à-dire

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

On note alors $df(a) = L \in \mathbb{B}(\mathrm{E}, \mathrm{F})$ la différentielle de f en a.
On dit que f est différentiable sur U ouvert de E si elle est différentiable en tout point de U.

Remarque I.4

1) Si $dim(E) < +\infty$, la continuité de l'application linéaire L est automatique.
Si $dim(E) = +\infty$, la continuité de l'application linéaire L est à vérifier.

2) La relation (I.2) peut se réécrire sous la forme suivante :

$$\exists \rho > 0, \exists \varepsilon : B_\rho(0) \subset E \longrightarrow F \text{ vérifiant } \lim_0 \varepsilon = 0 \text{ tq}$$
$$\forall h \in B_\rho(0),\ f(a+h) = f(a) + Lh + \|h\|_E \,\varepsilon(h),$$

Lorsqu'il n'est pas nécessaire de préciser, on écrira simplement que

$$f(a+h) = f(a) + Lh + \|h\|_E \,\varepsilon(h), \text{ au voisinage de } h = 0,$$

ou encore

$$f(a+h) = f(a) + Lh + o(\|h\|_E), \text{ au voisinage de } h = 0$$

(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) L'application $h \mapsto f(a) + df(a)h$ est appelée *application affine tangente à f en a.*

4) Si f est différentiable sur U alors $df : U \longrightarrow \mathbb{B}(\mathrm{E}, \mathrm{F})$.

Exemple 1 *Etude de la différentiabilité de $f : \mathbb{R}_n[X] \longrightarrow \mathbb{R}_{2n}[X]$ tq $f(P) = P^2$ et expression de sa différentielle. On munit $\mathbb{R}_n[X]$ de la norme définie pour tout $P \in \mathbb{R}_n[X]$ par $\|P\| = \underset{x \in [0,1]}{Sup} |P(x)|$ (idem pour $\mathbb{R}_{2n}[X]$).*

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

Proposition I.5 (unicité de la différentielle)

Si f est différentiable en $a \in U$ alors cette différentielle est unique.

Preuve de la Proposition I.5.
Supposons que f admettent deux différentielles L_1 et L_2 en a. Alors par définition, il existe $\rho > 0$ et ε_1, ε_2 tq pour tout $h \in B_\rho(0)$, on ait

$$f(a+h) = f(a) + L_1 h + \|h\|_E \,\varepsilon_1(h) \quad \text{et} \quad f(a+h) = f(a) + L_2 h + \|h\|_E \,\varepsilon_2(h),$$

où $\lim\limits_{0} \varepsilon_i = 0$.
Donc prenant la différence de ces deux égalités, on trouve $L_1h - L_2h = \|h\|_E\,\varepsilon(h)$.
Fixons alors $x \in E$ non nul. Pour tout $t \in \mathbb{R}^*$, vérifiant $|t| < \frac{\rho}{\|x\|_E}$, on peut prendre $h = tx$ et on obtient ainsi $t\,(L_1x - L_2x) = \|tx\|_E\,\varepsilon(tx)$ par linéarité de L_1 et L_2.
Il suffit alors de diviser cette égalité par $t \neq 0$ puis de faire tendre t vers 0 pour trouver $L_1(x) = L_2(x)$ pour tout $x \in E$ (puisque de plus $L_1(0) = L_2(0) = 0$ par linéarité).

■

Proposition I.6 (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 I.6.
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 (I.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 $\frac{1}{\|h\|_E^\#} \leq C_1 \frac{1}{\|h\|_E}$ et $\|f(a+h) - f(a) - Lh\|_F^\# \leq C_2\,\|f(a+h) - f(a) - Lh\|_F$ pour tout $h \in E$. Donc $\dfrac{\|f(a+h) - f(a) - Lh\|_F^\#}{\|h\|_E^\#} \leq C_1\,C_2\,\dfrac{\|f(a+h) - f(a) - Lh\|_F}{\|h\|_E}$ et on conclut donc par (I.2) puisque, toujours par équivalence des normes, $L \in \mathbb{B}(E,F) \Longrightarrow L \in \mathbb{B}(E^\#, F^\#)$.

■

Corollaire I.7

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

Mentionnons, pour terminer ce paragraphe, la propriété suivante.

Proposition I.8 (différentiabilité et continuité)

Si f est différentiable en $a \in U$ alors f est continue en a,
et donc par contraposée :
si f n'est pas continue en a alors$f$ n'est pas différentiable en a.

Preuve de la Proposition I.8.
Si f est différentiable en a alors au voisinage de $h = 0$: $\|f(a+h) - f(a)\|_F = \|df(a)h + \|h\|_E\,\varepsilon(h)\|_F \leq \|df(a)h\|_F + \|h\|_E\,\|\varepsilon(h)\|_F$ et la continuité de $df(a)$ permet de conclure quand $h \longrightarrow 0$.

■

Remarque I.9

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

C Dérivée directionnelle d'une fonction

Définition I.10 (dérivée directionnelle)

Soit E et F deux e.v.n, U un ouvert de E, $f : U \longrightarrow F$ 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

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

On notera $f^{'}(a)(b) \in F$ la dérivée directionnelle de f en a dans la direction b.

Remarque I.11

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. En d'autres termes, on étudie la dérivabilté de la restriction de f à la droite $\{a+tb\}$ en $t=0$ (on étudie donc la dérivabilité d'une fonction d'une seule variable *réelle* t).

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 I.12

Si f est différentiable en a, alors f admet en a des dérivées directionnelles dans toutes les directions. De plus dans ce cas, $\forall b \in E$,

$$df(a)b = f^{'}(a)(b) = \lim_{t \longrightarrow 0} \frac{f(a+tb)-f(a)}{t} \tag{I.3}$$

Egalité utile pour calculer une différentielle lorsque l'on sait que la fonction est différentiable.

Preuve de la Proposition I.12
Il suffit de prendre $h = tb$ avec $b \in E$ fixé et $t \longrightarrow 0$ dans ($I.2$).

■

On a le corollaire suivant :

Corollaire I.13

S'il existe $b \in E$ tel 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.

Exemple 2 *Montrer qu'une norme* $\| \ \|$ *sur un e.v* E *n'est jamais différentiable en* 0_E.

Remarque I.14

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 $f(x,y)=\dfrac{x^2y}{x^4+y^2}$ 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 $(0,0)$ puisque non continue.

2) **Lorsque la fonction est différentiable**, mais qu'on ne connaît pas l'expression de sa différentielle, la Proposition I.12 fournit un moyen pour la déterminer.

3) Lorsque la fonction admet des DD en a alors $L : h \mapsto f'(a)(h)$ fournit, par unicité de la différentielle en a, le seul candidat possible pour $df(a)$. Il suffit alors de montrer que ce L vérifie (cas différentiable) ou ne vérifie pas (cas non différentiable) (I.2).

2 Premiers exemples

Voyons tout de suite quelques exemples simples que vous devrez connaitre (ils constituent les briques élémentaires sur lesquelles nous nous reposerons). Le premier montre que la définition est bien, en un certain sens, la généralisation naturelle de la notion de dérivée pour les fonctions réelles.

A Différentielle de $f:\mathbb{R}\longrightarrow\mathbb{R}^n$

Si $f:\mathbb{R}\longrightarrow\mathbb{R}^n$, alors elle est différentiable en $t_0\in\mathbb{R}$ si et seulement si elle est dérivable en $t_0\in\mathbb{R}$, et alors $df(t_0)$ est définie sur $\mathbb{R}$ par $df(t_0): t\mapsto f'(t_0)t$ pour tout $t\in\mathbb{R}$ ou de manière équivalente $df(t_0)1=f'(t_0)$.

En effet, f est dérivable en $t_0\in\mathbb{R}$, si et seulement si $\lim\limits_{t\longrightarrow 0}\dfrac{f(t_0+t)-f(t_0)-f'(t_0)t}{t}=0$ donc si et seulement si $\lim\limits_{t\longrightarrow 0}\dfrac{f(t_0+t)-f(t_0)-L(t)}{t}=0$ avec $L\in\mathbb{B}(\mathbb{R},\mathbb{R}^n)$ définie par $L: t\mapsto f'(t_0)t$.

B Différentielle d'une constante

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

C Différentielle d'une application linéaire continue

Soient E et F des e.v.n. 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 $L=f$.

D Différentielle d'une application bilinéaire continue

Même si ce cas est un cas particulier de l'exemple suivant, il est très utilisé et nous préférons lui consacré son propre énoncé.

Soient E_1, E_2 et F des e.v.n. Si $f\in\mathbb{B}^2(E_1,E_2;F)$ 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

$$df(x_1,x_2)(h_1,h_2)=f(x_1,h_2)+f(h_1,x_2). \tag{I.4}$$

Preuve : voir le cas n-linéaire continue ci-dessous.
Plus généralement, on a le résultat suivant.

E Différentielle d'une application n-linéaire continue

Soient $(E_i,\|\ \|_{E_i})_{i=1}^n$ et F des e.v.n. On pose $E=E_1\times\cdots\times E_n$. Si $f\in\mathbb{B}^n(E_1,\ldots,E_n;F)$ alors f est différentiable sur E. Sa différentielle est alors donnée pour tout $X=(x_i)_1^n\in E$, et tout $H=(h_i)_1^n\in E$ par

$$df(X)H=\sum_{i=1}^n f(x_1,..,x_{i-1},h_i,x_{i+1},..,x_n) \tag{I.5}$$

Preuve :
On raisonne par récurrence :
• On sait déjà que la propriété est vraie au rang 1 (linéaire continue) et au rang 2 (bilinéaire continue).
• Avant de prouver l'hérédité et afin de simplifier un peu les notations, on pose $E' = E_1 \times \cdots \times E_{n-1}$ et ainsi $E = E_1 \times \cdots \times E_{n-1} \times E_n = E' \times E_n$. On peut alors écrire tout $Y = (y_1, ..., y_n) \in E$ sous la forme $Y = (Y', y_n)$ avec $Y' = (y_1, ..., y_{n-1}) \in E'$ et $y_n \in E_n$. Pour fixer les idées, on munit les espaces produits de la norme produit $\| \ \|_1$ que l'on notera $\| \ \|$. Compte tenu des notations introduites, on aura donc pour tout $Y \in E$, $\|Y\|_E = \|Y'\|_{E'} + \|y_n\|_{E_n}$

Supposons alors qu'il existe un entier $n \geq 2$ tel que toute application $(n-1)$-linéaire continue soit différentiable sur E' et vérifie pour tout $X' \in E'$, et tout $H' \in E'$ l'égalité

$$(a) \qquad df(X')H' = \sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1})$$

Soit alors f une application n-linéaire continue.
Montrons que f est différentiable en $X \in E$ fixé.
Pour $H = (H', h_n) \in E$ au voisinage de $0_E = (0_{E'}, 0_{E_n})$, on a

$$(b) \qquad \begin{aligned} f(X+H) - f(X) &= f(X' + H', x_n + h_n) - f(X', x_n) \\ &= f(X' + H', x_n + h_n) - f(X', x_n + h_n) + f(X', h_n). \end{aligned}$$

Or, l'application $g_n : y = (y_1, ..., y_{n-1}) \in E' \mapsto f(y_1, ..., y_{n-1}, x_n + h_n)$ est $(n-1)$-linéaire continue et donc, par hypothèse de récurrence, différentiable avec d'après (a), pour tout $H' \in E'$ au voisinage de $0_{E'}$:

$$g_n(X' + H') - g_n(X') = \sum_{i=1}^{n-1} g_n(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}) + o(\|H'\|_{E'})$$

c'est-à-dire

$$f(X' + H', x_n + h_n) - f(X', x_n + h_n) = \sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, x_n + h_n) + o(\|H'\|_{E'})$$

On déduit donc de (b) que

$$\begin{aligned} f(X+H) - f(X) &= \sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, x_n + h_n) + o(\|H'\|_{E'}) + f(X', h_n) \\ &= \sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, x_n) + f(X', h_n) \\ &\quad + \sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, h_n) + o(\|H'\|_{E'}) \\ &= \sum_{i=1}^{n} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, x_n) + \sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, h_n) + o(\|H'\|_{E'}) \end{aligned}$$

Pour conclure, il suffit donc de montrer que $\displaystyle\sum_{i=1}^{n-1} f(x_1, .., x_{i-1}, h_i, x_{i+1}, .., x_{n-1}, h_n) + o(\|H'\|_{E'}) = o(\|H\|_E)$.

Or, puisque $\|H'\|_{E'} \leq \|H\|_E$, on a $o(\|H'\|_{E'}) = o(\|H\|_E)$ et d'autre part, puisque l'application n-linéaire f est continue, on a aussi

$$\begin{aligned} \|f(x_1, ..., x_{i-1}, h_i, x_{i+1}, ..., x_{n-1}, h_n)\|_F &\leq C\|x_1\|_{E_1} ... \|x_{i-1}\|_{E_{i-1}} \|h_i\|_{E_i} \|x_{i+1}\|_{E_{i+1}} ... \|x_{n-1}\|_{E_{n-1}} \|h_n\|_{E_n} \\ &\leq C\|X\|^{n-2}\|H\|^2 \end{aligned}$$

ce qui prouve l'hérédité (la continuité de l'application linéaire est évidente).

3 Cas des fonctions de $\mathbb{R}^n \longrightarrow \mathbb{R}^p$

Dans tout ce paragraphe $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ et on note $B_E = (e_i)_1^n$ la base canonique de E.

A Dérivée partielle - Lien avec la différentielle

Les dérivées partielles constituent un cas particulier de dérivées directionnelles comme le montre la définition suivante.

Définition I.15 (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 $\lim\limits_{t \longrightarrow 0} \dfrac{f(a + te_j) - f(a)}{t}$ 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 $\frac{\partial f}{\partial x_j}(a)$ existe pour tout $a \in U$.

Remarque I.16

La dérivée partielle suivant x_j ne voit que ce qui se passe sur la droite $a + te_j$ (donc paralèllement aux j-ième axe de coordonées) 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 : les règles algébriques 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 de la fonction restreinte à la droite $(a + te_j)$ i.e. celle de $t \mapsto f(a + te_j)$;

3) en particulier, l'existence des dérivées partielles n'entraîne ni la différentiabilité, ni l'existence de DD.

On a toutefois le lien suivant.

Proposition I.17 (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 $h \in E$

$$df(a)h = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\ h_j \tag{I.6}$$

En particulier, on notera que

$$df(a)e_j = \frac{\partial f}{\partial x_j}(a). \tag{I.7}$$

Preuve de la Proposition I.17 : cas particulier de Dérivées directionnelles (Définition I.10 et Proposition I.12).

■

Remarque I.18

1) **Lorsque la fonction est différentiable**, la formule (I.6) fournit un moyen pour calculer cette différentielle en calculant les dérivées partielles.

2) Lorsque la fonction admet des dérivées partielles en a alors $L : h \mapsto \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, h_j$ fournit, par unicité de la différentielle en a, le seul candidat possible pour $df(a)$. Il suffit alors de montrer que ce L vérifie (cas différentiable) ou ne vérifie pas (cas non différentiable) (I.2).

Une autre manière fort utile d'énoncer la Proposition I.17 pour l'étude de la différentiabilité d'une fonction est la suivante.

Corollaire I.19 (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 I.20

Si f admet dans un voisinage de $a \in U$ des dérivées partielles et que celles-ci sont continues en a alors f est différentiable en a et sa différentielle en a est donnée par la formule (I.6).

Preuve de la Proposition I.20.
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. ■

B Matrice jacobienne - Gradient

Rappelons qu'on est toujours dans le cas où $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$.
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 (par (I.7)) la dérivée partielle de f par rapport à x_i au point a !

Formalisons tout cela dans une définition.

Définition I.21 (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 $df(a) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$ relativement aux bases canoniques B_E et B_F s'appelle **matrice jacobienne** (ou plus simplement **jacobienne**) de f en a. On la note $J_f(a) \in \mathscr{M}_{p,n}$. Ses coefficients sont les $(J_f(a))_{ij} = \frac{\partial f_i}{\partial x_j}(a)$, c'est-à-dire que $df(a)$ s'identifie à la matrice $J_f(a)$, ce que l'on écrit $df(a) \sim J_f(a)$ 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 I.22 (vecteur 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 $\forall h \in \mathbb{R}^n$, $df(a)(h) =< \nabla f(a), h >_2$

4 Opérations sur les fonctions différentiables

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

Proposition I.23 (stabilité par combinaison linéaire)

Soient E, F deux e.v.n., $U \subset E$ ouvert, f, $g : U \longrightarrow F$ deux applications différentiables en $a \in U$, et $\alpha \in \mathbb{R}$. 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 I.23. Evident. ∎

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

Théorème I.24 (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

$$d(g \circ f)(a) = dg(f(a)) \circ df(a) \tag{I.8}$$

Preuve du Théorème I.24.
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 telles que

$$f(a + h) = f(a) + df(a)h + \|h\|_E \ \varepsilon(h) \text{ au voisinage de } h = 0_E \tag{a}$$

$$g(b + u) = g(b) + dg(b)u + \|u\|_F \ \varphi(u) \text{ au voisinage de } u = 0_F. \tag{b}$$

avec $\lim\limits_{h \longrightarrow 0_E} \varepsilon(h) = 0_F$ et $\lim\limits_{u \longrightarrow 0_F} \varphi(u) = 0_G$.
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 (a)
Or $df(a) \in \mathbb{B}(E, F)$ et donc $u = df(a)h + \|h\|_E \ \varepsilon(h)$ vérifie

$$\|u\|_F = \|df(a)h + \|h\|_E \ \varepsilon(h)\|_F \leq (\|df(a)\|_{\mathbb{B}} + \|\varepsilon(h)\|_F) \ \|h\|_E \leq C_1 \|h\|_E \,. \tag{c}$$

Ainsi, $\lim\limits_{h \longrightarrow 0_E} u(h) = 0_F$ et on peut donc utiliser (b) avec ce u, ce qui implique

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

Utilisant alors la continuité de $dg(b)$ et (c), on a

$$\begin{aligned}\left\|dg(b)\Big(\|h\|_E\ \varepsilon(h)\Big)+\|u\|_F\ \varphi(u)\right\|_G &\leq \|h\|_E\ \|dg(b)(\varepsilon(h)\|_G+\|u\|_F\ \|\varphi(u)\|_G\\ &\leq \|h\|_E\ \|dg(b)\|_{\mathbb{B}}\ \|\varepsilon(h)\|_F+C_1\,\|h\|_E\ \|\varphi_1(h)\|_G\end{aligned}$$

et donc $dg(b)\Big(\|h\|_E\ \varepsilon(h)\Big)+\|u\|_F\ \varphi(u)=o(\|h\|_E)$ ce qui permet d'obtenir

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

Puisque $dg(b)\circ df(a)\in\mathbb{B}(E,G)$, ceci montre que l'application $g\circ f$ est différentiable en a et que sa différentielle est donnée par ($I.8$). ■

Remarque I.25 (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 (I.8) 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

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

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

5 Fonctions de classe $\mathscr{C}^1$

Définition I.26 (classe $\mathscr{C}^1$)

L'application $f:U\longrightarrow F$ est dite continûment différentiable sur U ou de classe $\mathscr{C}^1$ sur U (on note alors $f\in\mathscr{C}^1(U)$) 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 I.27 (stabilité linéaire de la classe $\mathscr{C}^1$)

Soient f, $g:U\longrightarrow F$ deux applications de classe $\mathscr{C}^1$ sur U, et soit $\alpha\in\mathbb{R}$. Alors les applications αf et $f+g$ sont de classe C^1 sur U (en d'autres termes $\mathscr{C}^1(U)$ est un e.v).

Preuve de la Proposition I.27
D'après la Proposition I.23 la différentiation est stable par combinaison linéaire et comme la continuité l'est également, la preuve est directe.

■

On déduit immédiatement du Théorème I.24 la propriété suivante

Corollaire I.28 (composée de fonctions $\mathscr{C}^1$)

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 $\mathscr{C}^1$ sur U et que g est $\mathscr{C}^1$ sur $f(U) \subset V$. Alors l'application $g \circ f : U \longrightarrow G$ est $\mathscr{C}^1$ sur U

6 Deux 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.

A Fonctions 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 fonction 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 I.24 de composition.

Théorème I.29 (différentielle d'applications à valeurs dans un produit)

La fonction 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

$$df(a) = (df_1(a), ..., df_n(a)) \tag{I.9}$$

Preuve du Théorème I.29.
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, muni par exemple de la norme produit $\| \|_1$, pour les p_k et sur F_k pour les i_k).
$\Longrightarrow$ / 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 I.24 de différentiabilité de la composée que f_k est différentiable en a.
$\Longleftarrow$ / 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 I.24 et Proposition I.23) 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 (I.9)).

■

B 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 finie, $dim(E) = dim(F)$.
On rappelle la définition suivante :

Définition I.30 (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 $\mathrm{Isom}(\mathrm{E},\mathrm{F})$ l'ensemble des isomorphismes de E dans F et si $E = F$ on note $\mathrm{Isom}(\mathrm{E})$ cet ensemble.

Remarque I.31

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 I.32

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

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

Preuve du Lemme I.32.
Soit $f \in \mathbb{B}(\mathrm{E})$ vérifiant $\|f\|_{\mathbf{B}(\mathrm{E})} < 1$.
On a puisque f est continue $\sum_{k\geq 0} \|f^k\|_{\mathbf{B}(\mathrm{E})} \leq \sum_{k\geq 0} \|f\|^k_{\mathbf{B}(\mathrm{E})} = \dfrac{1}{1 - \|f\|_{\mathbf{B}(\mathrm{E})}}$ puisque $\|f\|_{\mathbf{B}(\mathrm{E})} < 1$ (série géométrique de raison $\|f\|_{\mathbf{B}(\mathrm{E})}$). Donc la série de terme générale f^k est normalement convergente (ie. $\forall k$, $\|f^k\| \leq u_k$ avec $\sum u_k$ convergente), donc convergente.
Notant $S = \sum_{k\geq 0} f^k$, on voit que $Sf = fS$ et donc que $S \circ (id - f) = (id - f) \circ S = id$ ce qui montre que $(id - f)$ est inversible, d'inverse S.

■

On s'intéresse alors aux propriétés de la fonction Inv définie par

$$Inv : \begin{array}{ccc} \mathbb{B}(\mathrm{E},\mathrm{F}) & \longrightarrow & \mathbb{B}(\mathrm{F},\mathrm{E}) \\ u & \mapsto & Inv(u) = u^{-1} \end{array} \tag{I.10}$$

On a la proposition suivante
.

Proposition I.33 (topologie de $\mathrm{Isom}(\mathrm{E},\mathrm{F})$ - Domaine de l'inverse)

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

(b) La fonction Inv est définie sur $D_{Inv} = \mathrm{Isom}(\mathrm{E},\mathrm{F})$ et $Im(Inv) = Inv(\mathrm{Isom}(\mathrm{E},\mathrm{F})) = \mathrm{Isom}(\mathrm{F},\mathrm{E})$.

Preuve de la Proposition I.33.
(a) On peut supposer que $\mathrm{Isom}(\mathrm{E},\mathrm{F}) \neq \emptyset$ (sinon la proposition est évidente).

Soit $u_0 \in \mathrm{Isom}(\mathrm{E},\mathrm{F})$.
On veut donc montrer que tout $u \in \mathrm{Isom}(\mathrm{E},\mathrm{F})$ suffisamment proche de u_0 est encore un isomorphisme de E dans F. Mais puisque $u_0 \in \mathrm{Isom}(\mathrm{E},\mathrm{F})$ alors u est un isomorphisme de E dans F si et seulement si $(u_0)^{-1} \circ u$ est un isomorphisme de E dans E.
D'autre part, on sait par le Lemme I.32, que $(u_0)^{-1} \circ u \in \mathrm{Isom}(\mathrm{E})$ dès que $\|id - (u_0)^{-1} \circ u\| < 1$. Comme par continuité de $(u_0)^{-1}$ on a

$$\|id - (u_0)^{-1}u\|_{\mathbb{B}(\mathrm{E})} = \|(u_0)^{-1}(u_0 - u)\|_{\mathbb{B}(\mathrm{E})} \leq \|(u_0)^{-1}\|_{\mathbb{B}(\mathrm{F},\mathrm{E})} \, \|u_0 - u\|_{\mathbb{B}(\mathrm{E},\mathrm{F})}$$

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

(b) Evident. $\blacksquare$

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

Théorème I.34 (régularité $\mathscr{C}^1$ de l'inverse)

L'application Inv est de classe $\mathscr{C}^1$ sur $\mathrm{Isom}(\mathrm{E},\mathrm{F})$, et sa différentielle en $u \in \mathrm{Isom}(\mathrm{E},\mathrm{F})$ est donnée pour tout $h \in \mathbb{B}(\mathrm{E},\mathrm{F})$ par

$$dInv(u)h = -u^{-1} \circ h \circ u^{-1} \tag{I.11}$$

Preuve du Théorème I.34.
Pour alléger un peu les notations, on note la composée comme un produit : $vw = v \circ w$.
Pour $u \in \mathrm{Isom}(\mathrm{E},\mathrm{F})$, 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 considèrer la différence $Inv(u+h) - Inv(u)$. Commençons par remarquer que, d'après le Lemme I.32, pour tout $h \in \mathbb{B}(\mathrm{E},\mathrm{F})$ vérifiant $\|hu^{-1}\|_{\mathbb{B}(F)} < 1$ (il suffit que $\|h\| < 1/\|u^{-1}\|$), on a $(id + hu^{-1}) \in \mathrm{Isom}(\mathrm{F})$ et

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

que l'on récrit sous la forme

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

Donc, par définition de la différentielle pour avoir la relation ($I.11$), il suffit de montrer que $u^{-1} \sum_{k\geq 2}(-hu^{-1})^k = o(\|h\|_{\mathbb{B}(\mathrm{E},\mathrm{F})})$. Mais, si h vérifie $\|hu^{-1}\|_{\mathbb{B}(F)} < 1$, on peut écrire que

$$\left\| u^{-1} \sum_{k\geq 2}(-hu^{-1})^k \right\|_{\mathbb{B}(\mathrm{F},\mathrm{E})} \leq \|u^{-1}\|_{\mathbb{B}(\mathrm{F},\mathrm{E})} \sum_{k\geq 2} \|hu^{-1}\|^k_{\mathbb{B}(F)} = \frac{\|u^{-1}\|_{\mathbb{B}(\mathrm{F},\mathrm{E})} \, \|hu^{-1}\|^2_{\mathbb{B}(F)}}{1 - \|hu^{-1}\|_{\mathbb{B}(F)}} \leq C\|h\|^2_{\mathbb{B}(\mathrm{E},\mathrm{F})} = o(\|h\|_{\mathbb{B}(\mathrm{E},\mathrm{F})}),$$

ce qui termine la preuve de la différentiabilité de Inv et de ($I.11$).
Il reste alors à montrer que $u \mapsto dInv(u)$ est continue sur $\mathrm{Isom}(\mathrm{E},\mathrm{F})$.
Introduisons les applications définies par

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

Noter pour tout $(v,w) \in \mathbb{B}(\mathrm{F},\mathrm{E}) \times \mathbb{B}(\mathrm{F},\mathrm{E})$, l'application $\psi(v,w) \in \mathbb{B} = \mathbb{B}\Big(\mathbb{B}(\mathrm{E},\mathrm{F}), \mathbb{B}(\mathrm{F},\mathrm{E})\Big)$ vérifie

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

ce qui prouve aussi que ψ qui est bilinéaire, est continue.
Ainsi, φ est continue sur $\mathrm{Isom}(\mathrm{E},\mathrm{F})$ car chaque application coordonnée l'est et vérifie $\varphi(\mathrm{Isom}(\mathrm{E},\mathrm{F})) \subset \mathbb{B}(\mathrm{F},\mathrm{E}) \times \mathbb{B}(\mathrm{F},\mathrm{E})$, 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. ■

Une conséquence directe de la Proposition I.33 et du Théorème I.34 est le résultat suivant

Corollaire I.35

Pour $n \in \mathbb{N}^*$, on note $\mathscr{M}_n$ l'e.v des matrices carrées de taille n à coefficients réels et $GL_n \subset \mathscr{M}_n$ le sous-ensemble des matrices inversibles. L'application $INV : A \in GL_n \mapsto A^{-1} \in GL_n$

(a) L'ensemble GL_n est un ouvert de $\mathscr{M}_n$.

(b) La fonction $INV : A \in GL_n \mapsto A^{-1} \in GL_n$ est de classe $\mathscr{C}^1$ sur GL_n, et sa différentielle en $A \in GL_n$ est donnée pour tout $H \in \mathscr{M}_n$ par

$$dINV(A)H = -A^{-1} \times H \times A^{-1} \tag{I.12}$$

Preuve du corollaire I.35.
(a) La preuve de la Proposition I.33 s'adapte directement. Sinon, on peut aussi utiliser le fait que le déterminant sur $\mathscr{M}_n$ est une application continue (considérée comme fonction des vecteurs colonnes de $A \in \mathscr{M}_n$, elle est n-linéaire et donc continue puisqu'on est en dimension finie) et donc $GL_n = det^{-1}(\mathbb{R}^*)$ est ouvert en tant qu'image réciproque d'un ouvert par une application continue.
(b) On comence par justifier la régularité $\mathscr{C}^1$ et on établit dans un second temps la formule (I.12).
◇ On a le diagramme suivant :

$$\begin{tikzcd} A \in GL_n \arrow[r, "INV"'] \arrow[d, "\varphi"'] & A^{-1} \in GL_n \\ u \in Isom(\mathbb{R}^n) \arrow[r, "Inv"'] & u^{-1} \in Isom(\mathbb{R}^n) \arrow[u, "\psi"'] \end{tikzcd}$$

où $\forall A \in \mathscr{M}_n$, $\varphi(A)$ est l'endomorphisme canoniquement associé à A tandis que $\forall w \in \mathscr{M}_n$, $\psi(w)$ est la matrice de w relativement à la base canonique de $\mathbb{R}^n$.
Ces deux applications vérifient $\varphi \in \mathbb{B}(\mathscr{M}_n, \mathbb{B}(\mathbb{R}^n))$, $\psi \in \mathbb{B}(\mathbb{B}(\mathbb{R}^n), \mathscr{M}_n)$ et sont donc $\mathscr{C}^1$ (on sait qu'elle sont différentiables car linéaires et continues et la continuité de la différentielle d'une telle application est immédiate).
On en déduit que $INV = \psi \circ Inv \circ \varphi$ est $\mathscr{C}^1$, par composition d'applications $\mathscr{C}^1$, puisque $\varphi(GL_n) \subset Isom(\mathbb{R}^n)$.
◇ La formule (I.12) peut s'obtenir en différentiant la formule précédente mais aussi plus simplement en différentiant la formule $A \times INV(A) = I$ valable pour tout $A \in GL_n$. On obtient $\forall A \in GL_n$, $\forall H \in \mathscr{M}_n$, $A \times dINV(A)H + H \times INV(A) = 0$ et donc (I.12).

FIN DU CHAPITRE I

Chapitre II

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.

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

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

Théorème II.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

$$f^{'}(c) = \frac{f(b) - f(a)}{b - a}, \tag{II.1}$$

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 II.1 la généralisation suivante moyennant l'introduction de la notation : $[a, b] = \{(1 - t)a + tb,\ t \in [0, 1]\}$, $]a, b[= \{(1 - t)a + tb,\ t \in]0, 1[\}$.

Corollaire II.2 (Cas des fonctions 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, (II.1) n'a pas de sens puisqu'on ne peut pas diviser par (le vecteur) $b - a$.

Preuve du Corollaire II.2
Appliquer le Théorème II.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 fonctions $f : E \longrightarrow F$, $dim(F) > 1$. La réponse est **non** comme le montre l'exemple suivant.

Exemple II.3

Voyons pourquoi ce résultat ne se généralise pas aux fonctions à valeurs dans F, $dim(F) > 1$.

Supposons d'abord qu'un funambule (qui se déplace donc sur un fil...) parte de $f(a)$ à l'instant a et notons $f(t)$ sa position à l'instant $t \geq a$ (donc $f : \mathbb{R} \longrightarrow \mathbb{R}$). On a alors l'évidence suivante : si à un instant $b > a$ le funambule est revenu en $f(a)$, alors il s'est arrêté pour 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 s'arrêter pour faire demi-tour et 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 II.1 aux fonctions à valeurs dans un e.v.n admet une autre formulation connue sous le nom d'inégalité des accroissements finis.

Théorème II.4 (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 $t \in]a, b[$, $\|f^{'}(t)\| \leq g^{'}(t)$ alors

$$\|f(b) - f(a)\| \leq g(b) - g(a). \tag{II.2}$$

Idée de la preuve du Théorème II.4
Idéalement on aimerait obtenir l'inégalité (équivalente) : $\forall t \in [a, b]$, $\|f(t) - f(a)\| - (g(t) - g(a)) \leq 0$.
Pour cela deux étapes :
(1) d'abord montrer que cette inégalité est vérifiée sur un voisinage $[a, t_0]$ de a ;
(2) puis "propager l'inégalité" sur $[t_0, t_0 + h]$ ($h > 0$ petit) puis "repropager" à partir de de $t_0 + h$.
Dans la vraie vie, c'est un peu plus technique :
(1) utilise un argument de continuité : l'inégalité est vraie pour $t = a$ mais on ne peut rien en dire si $t > a$.
Pour cette raison, on va plutôt montrer que :
$\forall \varepsilon > 0$, $\forall t \in [a, b]$, $\|f(t) - f(a)\| - (g(t) - g(a)) - \varepsilon \leq 0$.
Cette inégalité, vérifiée en $t = a$, l'est aussi sur un voisinage $[a, t_0]$ de a par continuité.
(2) va évidement utiliser l'hypothèse sur les dérivées en écrivant qu'au voisinage de $h = 0^+$
$\frac{f(t_0+h)-f(t_0)}{h} = f'(t_0) + o(1)$ d'où $\|\frac{f(t_0+h)-f(t_0)}{h}\| \leq \|f'(t_0)\| + \|o(1)\| \leq g'(t_0) + \|o(1)\|$
et donc $\|f(t_0 + h) - f(t_0)\| \leq g(t_0 + h) - g(t_0) + \|o(h)\|$.
C'est la prise en compte de ces 2 étapes qui conduit à introduire la fonction auxiliaire ψ ci-dessous .

Preuve du Théorème II.4
On va montrer que pour tout réel $\varepsilon > 0$ donné, on a $\forall t \in [a, b]$ l'inégalité

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

Il suffira alors de prendre $t = b$ puis $\varepsilon \downarrow 0$ pour avoir le Théorème II.4 grace à la continuité de f et g.

Soit donc $\varepsilon > 0$ fixé.
Introduisons la fonction continue (par composition) $\psi : [a, b] \longrightarrow \mathbb{R}$ définie sur $[a, b]$ par

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

Il s'agit donc de montrer que $\forall t \in [a, b]$, on a $\psi(t) \leq 0$.

⋄ Commençons par remarquer que l'inégalité (a) est vérifié pour $0 < t - a$ suffisament petit.
En effet, puisque $\psi(a) = -\varepsilon < 0$, par continuité de ψ en a, $\exists \alpha > 0$, $\forall t \in [a, a+\alpha]$, $\psi(t) \leq 0$.

⋄ Montrons alors que l'inégalité (a) est vraie sur $[a, b]$ en entier.
Considérons pour cela l'ensemble des $t \in [a, b]$ pour lesquels elle n'est pas vérifiée, soit

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

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

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

- *(i)* $c > a$ puisque (a) est vérifiée pour $t \in [a, a+\alpha]$.
- *(ii)* $c \notin Z$ car Z est ouvert et donc si c appartenait à Z, il existerait un intervalle ouvert centré en c et contenu dans Z, donc des $t \in Z$ pour lesquels $a < t < c$ et $t \in Z$ ce qui contredirait le fait que c est borne inférieure de Z.
- *(iii)* $c < b$ car sinon on aurait $Z = \{b\}$ qui serait fermé.

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

$$\|f'(c)\| \leq g'(c). \tag{b}$$

Or, par dérivabilité de f et g en c, il existe $\beta > 0$ tel que pour tout t vérifiant $0 < t - c \leq \beta$

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

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

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

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

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

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

Utilisant (c) et (d), on obtient alors pour $c \leq t \leq \beta + c$

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

Ainsi les t qui vérifient $c \leq t \leq \beta + c$ satisfont (a) et ne sont donc pas dans Z. Comme on sait déja que $[a, c[\cap Z = \emptyset$, on en déduit que $[a, c+\beta] \cap Z = \emptyset$ avec $\beta > 0$ et donc la borne inférieure de Z est forcément supérieure à $c + \beta$ ie. $c \geq c + \beta$ ce qui est absurde et donc $Z = \emptyset$.

■

On peut alors déduire du Théorème II.4 le corollaire suivant qui est au moins aussi important puisque c'est très souvent cette forme 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 majorée, 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 II.5 (Cas des fonctions 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

$$\|f(x) - f(y)\|_F \leq \|x - y\|_E \underset{z \in]x,y[}{Sup} \|df(z)\|_{\mathbf{B(E,F)}} \tag{II.3}$$

En particulier, si $E = \mathbb{R}$, l'inégalité s'écrit

$$\|f(x) - f(y)\|_F \leq |x - y| \underset{t \in]x,y[}{Sup} \|f^{'}(t)\|_F .$$

puisque $\forall h \in \mathbb{R}$, $df(t)h = f'(t)h$ et donc $\|df(t)\|_{\mathbf{B}(\mathbb{R},F)} = \|f'(t)\|_F$.

Preuve du Corollaire II.5
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[1].
- 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 continue sur $[0, 1]$, 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 $\|df(z)\|_{\mathbf{B(E,F)}} \leq C$ pour tout $z \in [x, y]$, on a

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

On peut donc appliquer le Théorème II.4 aux fonctions ψ et $g : t \mapsto C \|x - y\|_E t$ sur $[0, 1]$ puisque $\forall t \in]0, 1[$, $\left\|\psi^{'}(t)\right\|_F \leq g^{'}(t)$. On en déduit alors que $\|\psi(1) - \psi(0)\|_F \leq g(1) - g(0)$ c'est-à-dire (II.3). ■

Remarque II.6

1. Noter que si U est un ouvert convexe alors par définition $\forall x, y \in U$, $[x, y] \subset U$ et l'hypothèse du Corollaire II.5 est donc systématiquement vérifiée.
2. Le cas $f : \mathbb{R} \longrightarrow F$ (Théorème II.1 et son Corollaire) contient l'hypothèse de convexité de l'ouvert puisqu'on se place sur des *intervalles ouverts* qui sont les seuls ouverts convexes de $\mathbb{R}$.

Nous allons terminer ce chapitre en voyant quelques conséquences des résultats sur les accroissements finis.

2 Applications

Dans toute cette partie, et sauf précision supplémentaire, on supposera que E et F sont deux e.v.n, que U est un ouvert non vide de E et que $f : E \longrightarrow F$ est une fonction définie sur U.

A Caractérisation des fonctions différentiables lipschitziennes

Voyons tout de suite une première conséquence du Corollaire II.5. Pour cela nous avons besoin de la définition suivante que vous connaissez très probablement déja.

[1]Bien que f soit différentiable sur U qui contient $[x, y]$, cette différentielle peut ne pas être bornée sur le segment : considérer par exemple $f : \mathbb{R} \longrightarrow \mathbb{R}$ avec $f(x) = x^2 \sin \frac{1}{x^2}$, $f(0) = 0$

Définition II.7 (fonction lipschitzienne)

On dit que f est **Lipschitzienne** de rapport $K > 0$ sur U, s'il existe $K > 0$ tel que

$$\forall x, y \in U, \quad \|f(x) - f(y)\|_F \leq K \|x - y\|_E . \tag{II.4}$$

Dans le cas où $K < 1$, on dit que f est **contractante** .

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

Proposition II.8 (fonction différentiable Lipschitzienne sur un convexe)

On suppose que l'ouvert U est **convexe** et que f est différentiable sur U.
Alors les deux propriétés suivantes sont équivalentes :

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

Preuve de la Proposition II.8
$\Longleftarrow$ / Conséquence du Corollaire II.5 (en utilisant la convexité de U) et de la Définition II.7.
$\Longrightarrow$ / Supposons f Lipschitzienne.
Soit $h \in E$ et $x_0 \in U$ donnés. Puisque U est ouvert, pour $t \in \mathbb{R}^*$ au voisinage de 0, $x_0 + th \in U$. Donc on peut prendre $x = x_0 + th$ et $y = x_0$ dans (II.4), ce qui donne

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

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

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

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

■

B Applications de différentielle nulle sur un ouvert connexe

Une autre conséquence immédiate est la suivante.

Définition II.9 (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 utilisera la définition suivante (connexité par arcs affines):
Un ouvert U d'un e.v.n E est **connexe par arcs affines** si pour tout $a, b \in U$, il existe une famille finies de 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 II.10

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. Pour cette raison et puisque nous ne travaillons que dans ce cadre, dans ce qui suit, on parlera de connexité sans préciser qu'elle est par arcs.

On peut alors énoncer la propriété suivante.

Proposition II.11 (différentielle nulle sur un connexe)

On suppose que l'ouvert U est **connexe** et que f est différentiable sur U.
Si la différentielle de f est nulle sur U alors f est constante sur U.

Idée de la preuve de la Proposition II.11
Elle est relativement simple :
montrer que $\forall a, b \in U$, $f(a) = f(b)$. Pour cela on va utiliser la connexité par arcs et le Corollaire II.5 sur chaque sous-segment permettant de joindre a à b.

Preuve de la Proposition II.11
On utilise les notations de la Définition II.9.
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 II.5 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 II.12

Noter que la Proposition II.11 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 II.6 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 $U =]0, 1[\cup]2, 3[$).

C Caractérisation des fonctions $\mathscr{C}^1(U)$ lorsque $U \subset \mathbb{R}^n$

Si on ne peut espérer caractériser la différentiabilité en terme de dérivées partielles, on a néanmoins une caractérisation pour la classe $\mathscr{C}^1$. 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 II.13 (caractérisation des fonctions $\mathscr{C}^1$)

On suppose que $E = \mathbb{R}^n$. Les deux propriétés suivantes sont alors équivalentes :

1. f est de classe $\mathscr{C}^1$ sur U (voir Définition I.26)
2. f admet sur U des dérivées partielles continues.

Idée de la preuve de la Proposition II.13
Pour *2.*$\Longrightarrow$*1.*, les dérivées partielles fournissant le seul candidat possible pour la différentielle, et tout se ramène alors à montrer que ce dernier satisfait la définition de la différentielle.
On fait donc apparaître les "variations de la fonction par rapport à chacune de ses variables" et on applique dans chaque direction ainsi obtenue le Corollaire II.5.

Preuve du Proposition II.13

• *1.*$\Longrightarrow$*2.* est directe si df est continue sur U puisque, $\forall x\in U$, $\frac{\partial f}{\partial x_i}(x)=df(x)e_i$.
• *2.*$\Longrightarrow$*1.* Soit $a\in U$ tel que f admette en a des dérivées partielles continues.
◇ Commençons par montrer que 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 $h\in\mathbb{R}^n$

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

et donc $\|df(x)-df(a)\|_{\mathbf{B}(\mathbb{R}^n,F)}\le\displaystyle\sum_{j=1}^{n}\left\|\frac{\partial f}{\partial x_j}(x)-\frac{\partial f}{\partial x_j}(a)\right\|_F$ 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)

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

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

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

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

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

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

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

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

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

On peut donc appliquer le Corollaire II.5 à g sur $[0,1]$, et on obtient ainsi que $\|g(1)-g(0)\|_F\le\varepsilon|x_1-a_1|$, soit en revenant à f

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

On fait de même pour les autres termes de (b) 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

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

Ceci montre (a) et conclut la preuve de la Proposition II.13.

■

Remarque II.14

La Proposition I.20 est un corollaire de la Proposition II.13 que nous venons de montrer.

D Suite de fonctions différentiables

Commençons ce paragraphe par deux petits rappels.
Rappels
Soit $(f_n)_n$ une suite de fonctions définies sur $W \subset E$ e.v.n et à valeurs dans F e.v.n.

(1) On dit que $f_n \longrightarrow f$ uniformément sur W si $\lim\limits_{n \longrightarrow \infty} \underset{x \in W}{Sup} \|f_n(x) - f(x)\|_F = 0$.

(2) Si $(f_n)_n \subset \mathscr{C}^0(W)$ et que $f_n \longrightarrow f$ uniformément sur W alors $f \in \mathscr{C}^0(W)$.

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

Proposition II.15 (suite de fonctions différentiables)

Soit $(f_n)_n$ 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. $\forall x \in U,\ \lim\limits_{n \longrightarrow \infty} f_n(x) = f(x)$;

(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, $\lim\limits_{n \longrightarrow \infty} \underset{x \in B}{Sup} \|df_n(x) - g(x)\|_{\mathbf{B}(E,F)} = 0$;

alors la limite f est différentiable sur U et $df = g$.
En d'autre termes : $\lim\limits_{n \longrightarrow \infty} df_n = d \lim\limits_{n \longrightarrow \infty} f_n$

Idée de la preuve de la Proposition II.15
On commence par déduire de *(H2)* que pour tout n assez grand, au voisinage de a, on a $f_n(x) - f_n(a) - g(a)(x-a) = o(\|x-a\|)$ et on conclut grace à *(H1)*.

Preuve de la Proposition II.15
Soit $a \in U$.
Fixons $\varepsilon > 0$.
Puisque U est ouvert, il existe $\rho > 0$ tel que $B_\rho(a) \subset U$ et $B_\rho(a)$ étant bornée, on a par *(H2)*

$$(a) \qquad \exists N \in \mathbb{N}, \forall k \geq N, \forall x \in B_\rho(a), \|df_k(x) - g(x)\|_{\mathbf{B}(\mathrm{E},\mathrm{F})} \leq \varepsilon.$$

et donc en paticulier $\forall n, p \geq N$, $\forall x \in B_\rho(a)$,

$$\|df_n(x) - df_p(x)\|_{\mathbf{B}(\mathrm{E},\mathrm{F})} \leq \|df_n(x) - g(x)\|_{\mathbf{B}(\mathrm{E},\mathrm{F})} + \|df_p(x) - g(x)\|_{\mathbf{B}(\mathrm{E},\mathrm{F})} \leq 2\varepsilon,$$

ce qui implique

$$(b) \qquad \underset{x \in B_\rho(a)}{Sup} \|df_n(x) - df_p(x)\|_{\mathbf{B}(\mathrm{E},\mathrm{F})} \leq 2\varepsilon.$$

Mais alors, puisque $B_\rho(a)$ est un ouvert convexe, pour tout $x \in B_\rho(a)$ on a $[a,x] \subset B_\rho(a)$ et on peut appliquer le Corollaire II.5 à $f_n - f_p$ sur $[a,x]$. Compte tenu de (b), on obtient ainsi

$$\|f_n(x) - f_n(a) - f_p(x) + f_p(a)\|_F \leq 2\varepsilon \|x-a\|_E .$$

Fixons alors $p \geq N$.
Par différentiabilité de f_p en a, il existe $r < \rho$ tel que

$$\forall x \in B_r(a), \|f_p(x) - f_p(a) - df_p(a)(x-a)\|_F \leq \varepsilon \|x-a\|_E$$

et de plus par (a), on a $\|df_p(x)(x-a) - g(x)(x-a)\|_F \leq \varepsilon \|x-a\|_E$.
De ces trois dernières inégalités, on obtient que pour tout $n \geq N$ et $x \in B_r(a)$,

$$\begin{aligned} \|f_n(x) - f_n(a) - g(a)(x-a)\|_F &\leq \|f_n(x) - f_n(a) - f_p(x) + f_p(a)\|_F \\ &\quad + \|f_p(x) - f_p(a) - df_p(a)(x-a)\|_F \\ &\quad + \|df_p(a)(x-a) - g(a)(x-a)\|_F \\ &\leq 4\varepsilon \|x-a\|_E \end{aligned}$$

On fait alors tendre n vers $+\infty$, ce qui par *(H1)*, implique finalement que

$$\forall \varepsilon > 0,\ \exists r > 0,\ \forall x \in B_r(a),\ \|f(x) - f(a) - g(a)(x-a)\|_F \leq 4\varepsilon \|x-a\|_E .$$

Ainsi, f est différentiable en a et $df(a) = g(a)$ ce qui permet de conclure. ∎

Remarque II.16

Si dans le théorème précédent $(f_n)_n \subset \mathscr{C}^1(U)$ alors $f \in \mathscr{C}^1(U)$. En effet, si $a \in U$ il existe $r > 0$ tel que g donc df soit limite uniforme de $(df_n)_n \subset \mathscr{C}^0(B_r(a))$ sur $B_r(a)$ et on utilise le rappel (2).

FIN DU CHAPITRE II

Chapitre III

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

De nouveau dans tout ce chapitre, et sauf précision supplémentaire, on supposera que E et F sont deux e.v.n, que U est un ouvert non vide de E et que $f : E \longrightarrow F$ est une fonction définie sur U.

1 Différentielle seconde

A Définition - Théorème de Schwarz

Définition III.1 (différentielle seconde)

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

Remarque III.2

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

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

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 III.3 (Théorème de Schwarz)

Si f est 2 fois différentiable en $x_0 \in U$ alors $d^2 f(x_0) : E \times E \longrightarrow F$ est symétrique, i.e.

$$\forall h, k \in E,\ d^2 f(x_0)(h, k) = d^2 f(x_0)(k, h). \qquad \text{(III.1)}$$

Idée de la preuve du Théorème III.3
f étant deux fois différentiable en a, elle est différentiable au voisinage a donc (h et k proches de 0) :
en $x_0 + h$: $f(x_0 + h + k) = f(x_0 + h) + df(x_0 + h)k + (reste)$
et en x_0 : $f(x_0 + k) = f(x_0) + df(x_0)k + (reste)$
Par différence, on a alors :
$f(x_0 + h + k) - f(x_0 + h) - f(x_0 + k) + f(x_0) = (df(x_0 + h) - df(x_0))k + (reste)$
et puisque df est différentiable en x_0 :
$f(x_0 + h + k) - f(x_0 + h) - f(x_0 + k) + f(x_0) = d^2 f(x_0)(h, k) + (reste)$.
La partie gauche de cette égalité étant symétrique, celle de droite l'est donc aussi.

Preuve du Théorème III.3
On obtient le résultat comme conséquence de la propriété suivante : au voisinage de $h, k = 0$

$$(a) \qquad \begin{array}{l} f(x_0 + h + k) - f(x_0 + h) - f(x_0 + k) + f(x_0) - d^2 f(x_0)(h, k) \\ = (\|h\|_E + \|k\|_E)^2 \ \eta(\|h\|_E , \|k\|_E), \end{array}$$

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 (a) que

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

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

• Il reste donc à prouver (a).
Puisque f est 2 fois différentiable en x_0, elle est différentiable dans un voisinage de x_0 et df est différentiable en x_0. Ainsi, on a
⋄ d'une part qu'il existe $\alpha > 0$ tel que pour $h \in B_{\alpha/2}(0)$ fixé, la fonction b_h définie sur $B_{\alpha/2}(0)$ par

$$b_h(k) = f(x_0 + h + k) - f(x_0 + h) - f(x_0 + k) + f(x_0) - d^2 f(x_0)(h, k)$$

soit différentiable avec
$$\begin{aligned} db_h(k) &= df(x_0 + h + k) - df(x_0 + k) - d^2 f(x_0)(h, .) \\ &= df(x_0 + h + k) - df(x_0) - d^2 f(x_0)(h + k, .) \\ &\quad -df(x_0 + k) + df(x_0) + d^2 f(x_0)(k, .) \end{aligned}$$

⋄ et que d'autre part fixant $\varepsilon > 0$, il existe $\beta \in]0, \alpha[$ tel qu'on ait pour tout $\ell \in B_\beta(0)$

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

Ainsi choisissant h et k tels que $\|h\|_E$, $\|k\|_E < \beta/2$, on en déduit

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

Appliquons alors le Théorème des accroissements finis (Corollaire II.5) à b_h sur $[0, k] \subset B_{\beta/2}(0)$. On trouve que pour tout $\|h\|_E , \|k\|_E < \beta/2$,

$$\begin{aligned} \|b_h(k) - b_h(0)\|_F &= \|b_h(k)\|_F \leq 2\varepsilon \|k\|_E \ \underset{z\in[0,k]}{Sup} (\|h\|_E + \|z\|_E) \leq 2\varepsilon \|k\|_E (\|h\|_E + \|k\|_E) \\ &\leq \ \leq 2\varepsilon(\|h\|_E + \|k\|_E)^2 \end{aligned}$$

ce qui conclut la preuve de (a). ■

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 III.4 (calcul d'une différentielle seconde)

On suppose f différentiable sur U.
Pour $h \in E$ fixé, on définit $g_h : x \in U \mapsto g_h(x) = df(x)h \in F$.
Alors, si f est deux fois différentiable en $x_0 \in U$, g_h est différentiable en x_0 et

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

Preuve du Théorème III.4.
Décomposer l'application à h fixé.

■

Remarque III.5

Tel que formulé, pour pouvoir appliquer le Théorème III.4, il faut déja savoir que f est deux fois différentiable (à justifier par les résultats généraux).
Noter que la réciproque est tout de même vraie.

Voyons tout de suite sur deux exemples comment utiliser la relation (III.2).

Exemple III.6

1. *Cas 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 III.4, 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 $f^{''}(a)$: $\forall h, k \in \mathbb{R}$, $d^2f(a)(h, k) = f^{''}(a)hk$.
2. *Cas d'une application $f \in \mathbb{B}^2(E_1, E_2; F)$.*
   Pour E_1, E_2 et F e.v.n, on a vu dans l'exemple D que si $f \in \mathbb{B}^2(E_1, E_2; F)$ alors elle est différentiable sur $E_1 \times E_2$ avec pour tout $x, h \in E_1 \times E_2$, $df(x)h = f(x_1, h_2) + f(h_1, x_2)$.
   Cette expression permet de justifier la différentiabilité de df : elle montre que c'est une application linéaire continue de $E_1 \times E_2$ dans $\mathbb{B}(E_1 \times E_2, F)$ (ie. $df \in \mathbb{B}(E_1 \times E_2, \mathbb{B}(E_1 \times E_2, F))$.
   Ainsi f est donc deux fois différentiable sur $E_1 \times E_2$ et on peut donc utiliser la formule (III.2) qui donne (g_h est ici linéaire) :
   $$\forall k, h \in E_1 \times E_2, \; d^2f(x)(k, h) = dg_h(x)k = g_h(k) = f(k_1, h_2) + f(h_1, k_2).$$

B Dérivées partielles secondes

On suppose dans toute cette partie que $E = \mathbb{R}^n$.

Définition III.7 (dérivée partielle seconde)

Pour i et j entiers donnés vérifiant $1 \leq i, j \leq n$, 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 III.8 (Lien entre dérivées partielles et différentielle - Formule de Schwarz)

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

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = d^2 f(x_0)(e_j, e_i). \tag{III.3}$$

De plus, dans ce cas (Théorème de Schwarz)

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_j x_i}(x_0). \tag{III.4}$$

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 III.8.
Si f est deux fois différentiable en $x_0 \in U$, on a au voisinage de $h = 0$

$$df(x_0 + h) = df(x_0) + d^2 f(x_0)(h, .) + o(\|h\|_E).$$

En particulier pour $h = te_i$ ($t \in \mathbb{R}^*$ proche de 0), on obtient

$$df(x_0 + te_i)e_j = df(x_0)e_j + d^2 f(x_0)(te_i, e_j) + o(t)$$

et donc, puisque $df(x)e_j = \frac{\partial f}{\partial x_j}(x)$

$$\lim_{t \longrightarrow 0} \frac{\frac{\partial f}{\partial x_j}(x_0 + te_i) - \frac{\partial f}{\partial x_j}(x_0)}{t} = d^2 f(x_0)(e_i, e_j)$$

c'est-à-dire l'égalité (III.3).
L'égalité (III.4) découle quant à elle directement de la formule (III.1) du Théorème III.3.

■

Remarque III.9

En outre, on voit que la formule (III.3) 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

$$d^2 f(x_0)(h,k) = \sum_{i,j=1}^{n} h_i k_j \frac{\partial^2 f}{\partial x_j x_i}(x_0). \tag{III.5}$$

C Propriétés - Application de classe $\mathscr{C}^2$

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 III.10

1. Toute combinaison linéaire d'applications 2 fois différentiables en $a \in U$ est encore 2 fois différentiable en a.

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 I.26 en posant :

Définition III.11 (classe $\mathscr{C}^2$)

Une application f est dite de classe $\mathscr{C}^2$ sur U (noté $f \in \mathscr{C}^2(U)$) 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 III.10 on a évidement que toute combinaison linéaire de fonctions $\mathscr{C}^2$ est encore $\mathscr{C}^2$; et que toute composée de fonctions $\mathscr{C}^2$ est $\mathscr{C}^2$.

On a alors la caractérisation suivante.

Proposition III.12 (caractérisation des fonctions $\mathscr{C}^2$)

On suppose que $E = \mathbb{R}^n$. Alors les deux assertions suivantes sont équivalentes :

1. f est de classe $\mathscr{C}^2$ sur U ;
2. les dérivées partielles secondes de f existent et sont continues sur U.

Preuve du Proposition III.12.
L'implication 1. $\Longrightarrow$ 2. est une conséquence directe de la formule (III.3) du Corollaire III.8.
• Montrons alors que 2. $\Longrightarrow$ 1.
Pour tout $i, j = 1, \ldots, n$, $\frac{\partial^2 f}{\partial x_j \partial x_i} \in \mathscr{C}^0(U)$: en d'autres termes, pour chaque i la fonction $\frac{\partial f}{\partial x_i}$ admet des dérivées

partielles premières qui sont continues et elle est donc de classe $\mathscr{C}^1$ sur U d'après la Proposition II.13.
Or, pour tout $x \in U$ et tout $h \in E$, on a

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

et donc $df = \varphi \circ \psi$ est de classe $\mathscr{C}^1$ sur U par composition en posant

$$\begin{array}{rccc} \psi : & U & \longrightarrow & F^n \\ & x & \mapsto & \left(\frac{\partial f}{\partial x_i}(x)\right)_{i=1}^{n} \end{array}$$

$\mathscr{C}^1$ sur U car chaque composante l'est, et

$$\begin{array}{rccl} \varphi : & F^n & \longrightarrow & \mathbb{B}(\mathbb{R}^n, F) \\ & u = (u_k) & \mapsto & \varphi(u) \text{ où } \varphi(u) : h \in \mathbb{R}^n \mapsto \sum_{i=1}^{n} u_i h_i \in F \end{array}$$

$\mathscr{C}^1$ sur F^n car linéaire continue ($\varphi \in \mathbb{B}(F^n, \mathbb{B}(\mathbb{R}^n, F))$).

■

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 III.13 (différentielle d'ordre k)

Soit $k \geq 2$ un entier.
On dit que 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'identifier à une application k-linéaire. 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.

Remarque III.14 (dérivée partielle d'ordre k)

On étend de manière similaire la notion de dérivée partielle à l'ordre k

On peut alors étendre à l'ordre k la Définition I.26 en posant :

Définition III.15 (classes $\mathscr{C}^k$ et $\mathscr{C}^\infty$)

Soit $k \geq 2$ un entier. On dit que f est de classe $\mathscr{C}^k$ sur U ($f \in \mathscr{C}^k(U)$) 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 $\mathscr{C}^k$ pour tout k sur U, on dit que f est de classe $\mathscr{C}^\infty$ sur U et on note $f \in \mathscr{C}^\infty(U)$.

On a alors les propriétés de base suivantes :

Proposition III.16

Soit $k \geq 2$ un entier.
Si f est k fois différentiable sur U (respectivement $\mathscr{C}^k$ sur U) alors

1. si $\forall x \in U$, $\forall H = (h_i)_{i=1}^{k-1} \in E^{k-1}$, $g_H(x) = d^{k-1}f(x)(h_1, \dots, h_{k-1})$ alors g_H est différentiable sur U et $\forall x \in U$, $\forall h_k \in E$, $dg_H(x)h_k = d^k f(x)(h_1, \dots, h_k)$;
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 $\mathscr{C}^k$ sur U) alors pour tout $\alpha \in \mathbb{R}$, $(f + \alpha g)$ est k fois différentiable sur U (respectivement $\mathscr{C}^k$ 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 $\mathscr{C}^k$ sur V) alors $g \circ f : U \longrightarrow F$ est k fois différentiable sur U (respectivement $\mathscr{C}^k$ 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 III.17

On suppose que $E = \mathbb{R}^n$. Alors pour $k \geq 2$ entier

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 $1 \leq i_1, .., i_k \leq n$ sont des entiers ;
2. f est de classe $\mathscr{C}^k$ sur U si et seulement si ses dérivées partielles d'ordre k sont toutes continues sur U

Terminons ce paragraphe par deux propriétés.

Proposition III.18

Soit $(E_i)_{i=1}^n$ et F des e.v.n et $E = E_1 \times \cdots \times E_n$.
Si $f \in \mathbb{B}^n(E; F)$ alors $f \in \mathscr{C}^\infty(E)$

Preuve de la Propriété III.18.
Technique mais facile...sans intérêt.

Théorème III.19 (régularité $\mathscr{C}^\infty$ de l'inverse)

Soient E et F de Banach ($dim(E) = dim(F)$ si la dimension est finie). L'application Inv, définie en (I.10) est de classe $\mathscr{C}^\infty$ sur $\mathrm{Isom(E, F)}$.

Preuve du Théorème III.19.
Utiliser la décomposition de $dInv(u)$ de la preuve du théorème I.34

■

3 Formules de Taylor

Elles sont au nombre de trois. Elles se distinguent par la forme du "reste", les parties régulières étant identiques. Nous aurons besoin dans ce paragraphe du résultat auxiliaire suivant.

Lemme III.20

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

$$\text{(III.6)} \qquad \forall t \in I, \quad \left(\sum_{i=0}^{k} \frac{(1-t)^i}{i!} \; g^{(i)}(t) \right)' = \frac{(1-t)^k}{k!} \; g^{(k+1)}(t).$$

Preuve du Lemme III.20.
Par récurrence.
• Initialisation (k=0) :
pour toute fonction dérivable, (III.6) s'écrit $\forall t \in I$, $g'(t) = g'(t)$ (en utilisant $g^{(0)} = g$).
• Supposons que pour toute fonction $g : \mathbb{R} \longrightarrow F$ 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 $g : \mathbb{R} \longrightarrow F$ 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 $k \geq 0$. ■

Remarque III.21 (dans le cas où $F = \mathbb{R}^p$)

Si $[0,1] \subset I$ et si $g \in \mathscr{C}^{k+1}(I)$ avec $F = \mathbb{R}^p$, intégrant (III.6) entre 0 et 1, on trouve la formule

$$\text{(III.7)} \qquad g(1) - g(0) - g^{'}(0) - \cdots - \frac{1}{k!} g^{(k)}(0) = \int_0^1 \frac{(1-t)^k}{k!} \; g^{(k+1)}(t) dt.$$

Théorème III.22 (formule de Taylor avec reste intégral : $f : E \longrightarrow \mathbb{R}^p$)

Soit $k \in \mathbb{N}$. On suppose que $F = \mathbb{R}^p$ et que $f \in \mathscr{C}^{k+1}(U)$.
Alors pour tout $x, h \in E$ 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 $\mathscr{C}^1$ (i.e. $k = 0$), on trouve

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

Preuve du Théorème III.22.
Pour $x, h \in E$, on pose pur tout $t \in [0,1]$, $\psi(t) = x + th$: fonction $\mathscr{C}^\infty$ 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 $\mathscr{C}^{k+1}([0,1])$ puisque f est $\mathscr{C}^{k+1}$ sur U.
Il suffit alors de remarquer que $g^{(q)}(t) = d^q f(x+th)(h)^q$ pour tout $q \le k+1$ pour conclure par (III.7).
■

Théorème III.23 (formule de Taylor-Lagrange : $f : E \longrightarrow F$)

On suppose que f est $(k+1)$ fois différentiable sur U $(k \in \mathbb{N})$.
S'il existe $C > 0$ tel que pour tout $x \in U$, $\left\| d^{k+1} f(x) \right\|_{\mathbf{B}^{k+1}(E;F)} \le C$ alors pour tout $[x, x+h] \subset U$ on a l'inégalité

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

Preuve du Théorème III.23.
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 III.20 que pour tout $t \in]0,1[$

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

en posant pour tout $t \in [0,1]$, $w(t) = -C \left\| h \right\|^{k+1} \frac{(1-t)^{k+1}}{(k+1)!}$.
Les fonctions r et w sont continues sur sur $[0,1]$, dérivables sur $]0,1[$ et vérifient pour tout $t \in]0,1[$, $\|r(t)\| \le w'(t)$.
On peut donc appliquer l'inégalité des accroissements finis (Théorème II.4) :

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

■

Théorème III.24 (formule de Taylor-Young : $f : E \longrightarrow F$)

On suppose que f est k fois différentiable en un point $x_0 \in U$ ($k \geq 1$).
Alors, au voisinage de $h = 0$, on a l'égalité

$$f(x_0 + h) = f(x_0) + df(x_0)h + \cdots + \frac{d^k f(x_0)}{k!}(h)^k + o(\|h\|_E^k). \tag{III.8}$$

Preuve du Théorème III.24.
Dans ce qui suit, on prendra $h \in E$ dans un voisinage ouvert de 0 suffisament petit 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 (III.8) soit vérifiée au rang $k - 1$, $k \geq 2$ 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 $1 \leq p \leq k - 1$, 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 $d\xi(h) = df(x_0 + h) - df(x_0) - d^2 f(x_0)(h) - \cdots - \dfrac{d^{k-1} df(x_0)}{(k-1)!}(h, \cdots, h)$.
Si on applique alors l'hypothèse de récurrence à df, on obtient $d\xi(h) = o(\|h\|_E^{k-1})$ soit

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

On peut donc appliquer à ξ le théorème des accroissements finis sur $B(0, \eta)$ (Corollaire II.5), ce qui donne $\|\xi(h) - \xi(0)\|_F \leq \varepsilon\ \|h\|_E^{k-1}\ \|h\|_E$. Soit encore, puisque $\xi(0) = 0$, que $\|\xi(h)\|_F \leq \varepsilon\ \|h\|_E^k$ pour h dans un voisinage de 0. La propriété est donc vraie au rang k.

■

Remarque III.25

La formule de Taylor-Young est **locale** (valable au voisinage de $x_0 \in U$) contrairement aux deux autres formules de Taylor qui sont **globales** (il suffit d'être dans U : inutile de se placer au voisinage de $x_0 \in U$).

4 Extrema libres

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

Définition III.26 (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 $\forall x \in V$, $f(x) \geq f(x_0)$ (respectivement $f(x) \leq f(x_0)$).
Si x_0 est un *minimum* ou un *maximum*, on dit que c'est un *extremum.*

Remarque III.27

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 III.28 (point critique)

On suppose f différentiable sur U.
Si $x_0 \in U$ vérifie $df(x_0) = 0$ alors on dit que c'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 III.29 (condition nécessaire d'ordre 1)

On suppose f différentiable sur U.
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 III.29.
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 $f(x_0 + h) \leq f(x_0)$.
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 $t > 0$

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

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

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

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

Remarque III.30

1. Le Théorème III.29 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 $f(x) = x^3$.
2. Néanmoins, grace à la forme contraposée du Théorème III.29, on sait où chercher les extrema : parmi les points critiques.
3. L'hypothèse U ouvert est esentielle : la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ telle que $f(x) = x$ et considérée sur $U = [0, 1]$ admet un maximum et un minimum bien que sa dérivée ne s'annulle jamais.

Nous aurons ensuite besoin de la définition suivante.

Définition III.31 (forme bilinéaire symétrique définie positive (ou négative))

Soit $\varphi : E \times E \longrightarrow \mathbb{R}$ une forme bilinéaire continue et symétrique.
On dit que φ est

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

On dit φ est définie positive si elle vérifie les deux propriétés.

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

• Pour nous, la forme bilinéaire symétrique φ sera, quand elle existe $d^2f(x_0)$, avec $x_0 \in U$ ouvert de E et $f : U \longrightarrow \mathbb{R}$ fonction 2 fois différentiable.

Remarque III.32 (cas $E = \mathbb{R}^n$)

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 dans une base orthonormée $(u_1, \cdots, u_n)$ 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) ssi $\forall 1 \leq i \leq n$, $\lambda_i \geq 0$ (resp. $\lambda_i \leq 0$).
2. φ est définie positive (resp. négative) ssi $\forall 1 \leq i \leq n$, $\lambda_i > 0$ (resp. $\lambda_i < 0$).
3. φ est définie positive (resp. négative) ssi il existe $\alpha > 0$ tel que $\forall h \in E$, $\varphi(h, h) \geq \alpha \|h\|_E^2$ (resp. $\varphi(h, h) \leq -\alpha \|h\|_E^2$)

Cette dernière caractérisation reste valable lorsque E n'est plus de dimension finie.

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

Définition III.33

On suppose que $E = \mathbb{R}^n$ et que f est 2 fois différentiable en $x_0 \in U$.
La matrice $(n \times n)$ de la forme bilinéaire symétrique $d^2 f(x_0)$ dans la base canonique de $\mathbb{R}^n$, notée $H_f(x_0)$, est appelée **matrice hessienne de f en x_0**.
Par le Corollaire III.8, ses coefficients sont données par $\Big(H_f(x_0)\Big)_{ij} = d^2 f(x_0)(e_i, e_j) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$ pour tout $i, j \in \{1, \dots, n\}$ et elle est **symétrique** par le Théorème III.3.

On peut énoncer deux nouveaux résultats qui termineront ce chapitre.

Théorème III.34 (condition nécessaire d'ordre 2)

On suppose que f est 2 fois différentiable en $x_0 \in U$.
Si x_0 est un minimum (respectivement maximum) de f alors $d^2 f(x_0)$ est positive (resp. négative).

Preuve du Théorème III.34.
Supposons que $x_0 \in U$ soit un minimum de f. Alors d'après le Théorème III.29 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 III.24), soit

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

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 $f(x_0 + h) - f(x_0) \geq 0$.
Ainsi, $\|h\|_E < \alpha$ entraine $d^2 f(x_0)(h, h) + o(\|h\|_E^2) \geq 0$.
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

$$\begin{aligned} & d^2 f(x_0)(tx, tx) + o(\|tx\|_E^2) \geq 0 \\ \Longrightarrow \quad & t^2 \left(d^2 f(x_0)(x, x) + o(1) \right) \geq 0 \\ \Longrightarrow \quad & d^2 f(x_0)(x, x) + o(1) \geq 0, \end{aligned}$$

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

■

Théorème III.35 (condition suffisante d'ordre 2)

On suppose que f est 2 fois différentiable en $x_0 \in U$.
Si $x_0 \in U$ vérifie $df(x_0) = 0$ avec $d^2 f(x_0)$ définie positive (respectivement définie négative) alors x_0 est un minimum (resp. maximum) local de f.

Preuve du Théorème III.35.
C'est une conséquence plus ou moins directe de la formule de Taylor-Young à l'ordre 2. En effet, d'après le Théorème III.24, $\exists \beta_0 > 0$ tel que pour tout $h \in E$ vérifiant $\|h\|_E < \beta_0$ on ait $(x_0 + h) \in U$, et

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

Mais si $d^2 f(x_0)$ est définie positive, d'après la Remarque III.32, il existe $\alpha > 0$ tel que $\forall h \in E$, $d^2 f(x_0)(h, h) \geq \alpha \|h\|_E^2$ et on en déduit donc que pour tout $\|h\|_E < \beta_0$

$$(a) \qquad f(x_0 + h) - f(x_0) \geq \frac{\alpha}{2} \|h\|_E^2 + o(\|h\|_E^2).$$

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 $\dfrac{o(\|h\|_E^2)}{\|h\|_E^2} \geq -\frac{\alpha}{4}$. Ainsi on déduit de ($a$) qu'il existe $\beta = Min(\beta_0, \beta_1) > 0$ tel que $\|h\|_E < \beta$ entraine

$$(b) \qquad f(x_0 + h) - f(x_0) \geq \left(\frac{\alpha}{2} - \frac{\alpha}{4} \right) \|h\|_E^2 \geq 0,$$

ce qui termine la preuve.

■

Remarque III.36

1) Comme le montre l'inégalité (b) dans la preuve précédente, sous les hypothèses du Théorème III.35, 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 d'étude pour la recherche des extrema locaux d'une fonction (2 fois différentiable):

1. On cherche les points critiques de f (les extrema en sont nécessairement par le Théorème III.29)
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 III.35.
   Si elle n'est ni positive, ni négative on conclut par le Théorème III.34.
   Sinon il faut mener une étude spécifique pour ce point.

FIN DU CHAPITRE III

Chapitre IV

Théorème d'inversion locale - Théorème des fonctions implicites

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, $U \subset E$ et $V \subset F$ sont des ouverts, et $f : E \longrightarrow F$ est définie sur U.
Nous avons d'abord besoin d'une définition.

Définition IV.1 (homéomorphisme - difféomorphisme)

On dit que

1. f est un homéomorphisme de U dans V si *(a)* f est continue sur U et bijective de U dans V ; *(b)* la bijection réciproque de $f_{|U}$, notée f^{-1}, est continue sur V ;
2. f est un difféomorphisme de U dans V si *(a)* f est $\mathscr{C}^1$ sur U et bijective de U dans V ; *(b)* f^{-1} est de classe $\mathscr{C}^1$ sur V ;
3. f est un $\mathscr{C}^{k+1}$ difféomorphisme de U dans V si *(a)* f est $\mathscr{C}^{k+1}$ sur U et bijective de U dans V ; *(b)* f^{-1} est de classe $\mathscr{C}^{k+1}$ sur V.

Dans tous les cas, on a évidement $V = f(U)$ (et $f^{-1}(V) = U$).

Remarque IV.2

Un homéomorphisme $\mathscr{C}^1$ n'est pas forcément un difféomorphisme : une fonction peut être bijective et $\mathscr{C}^1$ sans 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 $\mathscr{C}^{k+1}$ alors la bijection réciproque de f est automatiquement $\mathscr{C}^{k+1}$ comme le montre la seconde partie du résultat suivant qui nous sera utile par la suite.

Proposition IV.3

1. Si f est un difféomorphisme de U dans V (donc nécessairement $V = f(U)$).
   Alors pour tout $x \in U$, $df(x) \in \mathrm{Isom(E,F)}$ et pour tout $y \in V$,

   $$df^{-1}(y) = \big(df(f^{-1}(y))\big)^{-1}. \tag{IV.1}$$

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

Preuve de la Proposition IV.3.
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 (IV.1).
Concernant la partie 2 de la Proposition, il suffit de remarquer que si f est de plus $\mathscr{C}^{k+1}$, la formule (IV.1) s'écrit pour tout $y \in V$

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

Donc puisque df et Inv sont $\mathscr{C}^k$ (voir Théorème III.19) et f^{-1} est $\mathscr{C}^1$ la composée i.e. df^{-1} est $\mathscr{C}^1$ i.e. f^{-1} est $\mathscr{C}^2$ et de proche en proche on obtient finalement que df^{-1} est $\mathscr{C}^k$, soit f^{-1} est $\mathscr{C}^{k+1}$. ■

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

Théorème IV.4 (caractérisation des difféomorphismes)

Soit $f : U \longrightarrow V$ un homéomorphisme $\mathscr{C}^1$.
Les propriétes suivantes sont alors équivalentes :

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

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

FIN COURS 9

Preuve du Théorème IV.4.
• L'implication (*1* $\Longrightarrow$ *2*) est une conséquence directe de la Proposition IV.3.

• 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

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

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 (a) 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

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

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

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

et donc que

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

Ceci montre que $\|x-a\|_E \left\|(df(a))^{-1}(\varphi(x))\right\|_E = o(\|y-b\|_F)$ 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

$$\begin{aligned}
&f^{-1} : V \subset F \longrightarrow U \subset E \ \text{ continue car } f \text{ homéomorphisme ;}\\
&df : U \subset E \longrightarrow \mathrm{Isom(E,F)} \ \text{ continue car } f \in \mathscr{C}^1(U)\\
&Inv : \mathrm{Isom(E,F)} \longrightarrow \mathrm{Isom(F,E)} \ \text{ continue par la Proposition I.33,}
\end{aligned}$$

et elle est donc continue sur V.

■

Remarque IV.5

1) En dimension finie, on a $df(x) \in \mathrm{Isom}(\mathbb{R}^n)$ 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 :

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

2 Théorème d'inversion locale - Théorème d'inversion globale

A Motivations

Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ 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 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 $\mathscr{C}^1$, la fonction $f^{'}$ est continue et comme $f^{'}(x_0) > 0$, il existe un intervalle ouvert I_{x_0} contenant x_0 tel que $\forall x \in I_{x_0}$, $f^{'}(x) > 0$. Ainsi posant $W_{y_0} = f(I_{x_0})$, on en déduit que $f : I_{x_0} \longrightarrow W_{y_0}$ est continue, strictement croissante et donc bijective de I_{x_0} dans W_{y_0}. On note $f^{-1} : W_{y_0} \longrightarrow I_{x_0}$ 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 $x = f^{-1}(y) \in I_{x_0}$ tel que $f(x) = y$.

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

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

B Le Théorème d'inversion locale

La preuve du théorème d'inversion locale utilise le résultat suivant.

Théorème IV.6 (théorème du point fixe de Banach)

Soit E un espace de Banach et $f : E \longrightarrow E$ une application *contractante* sur E ie.

$$\exists K \in]0,1[,\ \forall x, x' \in E,\ \|f(x) - f(x')\| \leq K\|x - x'\|.$$

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

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

Preuve du Théorème IV.6.
Pour $x_0 \in E$ donné, on définit par récurrence la suite (x_n) en posant $\forall n \in \mathbb{N},\ x_{n+1} = f(x_n)$.
L'application f étant contractante sur E, elle l'est sur $[x_n, x_{n+1}]$ et donc

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

et de proche en proche, on obtient ainsi

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

On en déduit donc que pour tous entiers n et p

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

Puisque $K \in]0,1[$, ceci 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 et x_2 points fixes de f.
Alors, par contraction, $\|x_1 - x_2\| = \|f(x_1) - f(x_2)\| \leq K\|x_1 - x_2\|$ 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 nous l'utiliserons.

Corollaire IV.7

Soit E Banach, $U \subset E$ un ouvert, $C \subset U$ un fermé non vide 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 IV.7.
Utilisant les notations de la preuve du Théorème IV.6, il suffit de noter pour $x_0 \in C$, on a $(x_n) \subset C$ (car $f(C) \subset C$). La limite $a \in E$ que l'on obtient alors appartient bien à C puisque ce dernier est fermé.
■

On peut maintenant énoncer le résultat principal de ce chapitre.

Théorème IV.8 (Théorème d'inversion locale)

Soient E et F deux espaces de Banach ($dim(E) = dim(F)$ si la dimension est finie), Ω un ouvert de E et $f : \Omega \longrightarrow F$ une application de classe $\mathscr{C}^{k+1}$ sur Ω avec $k \geq 0$ entier.
S'il existe $x_0 \in \Omega$ tel que $df(x_0) \in \mathrm{Isom(E, F)}$ alors :

- ⋄ il existe $U \subset \Omega$ voisinage ouvert de x_0
- ⋄ il existe $V \subset F$ voisinage ouvert de $y_0 = f(x_0)$

tels que f soit un $\mathscr{C}^{k+1}$ difféomorphisme de U dans V.

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

Preuve du Théorème IV.8.
La preuve se compose de 3 étapes.

Etape 1 : *simplifications.*
Il suffit de montrer le résultat pour $k = 0$, tout difféomorphisme de classe $\mathscr{C}^{k+1}$ étant un $\mathscr{C}^{k+1}$ difféomorphisme (Proposition IV.3 partie 2).

On observe ensuite qu'on peut supposer que

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

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 (a).

Etape 2 : *f est localement un homéomorphisme.*
On pose pour tout $x \in \Omega$, $g(x) = x - f(x)$ qui vérifie donc $g \in \mathscr{C}^1(\Omega)$, $g(0) = 0$ et $dg(0) = 0$.
La continuité de dg en 0 montre que

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

puis le T.A.F à g sur $B_{2\rho}(0)$ (ouvert convexe) que $\forall x, x' \in B_{2\rho}(0)$, $\|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|$ et donc par continuité de g que

$$(b) \qquad \forall x, x' \in \overline{B_{2\rho}(0)}, \ \|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|$$

Montrons alors que :

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

On fixe $y \in B_\rho(0)$.
On va alors obtenir x comme point fixe de l'application h définie pour tout $x \in \Omega$ par $h(x) = y + g(x)$ (on a alors $h(x) = x$ ssi $y = f(x)$).

L'application h qui est de classe $\mathscr{C}^1$ et vérifie :
- pour tout $x, x' \in \overline{B_{2\rho}(0)}$, $\|h(x) - h(x')\| = \|g(x) - g(x')\| \leq \frac{1}{2}\|x - x'\|$ par (b)
- pour tout $x \in \overline{B_{2\rho}(0)}$, $\|h(x)\| = \|g(x) + y\| \leq \|y\| + \frac{1}{2}\|x\| < 2\rho$ par (b) (avec $x' = 0$ et car $g(0) = 0$)

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 (c), ce que l'on peut réécrire sous la forme

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

Ainsi posant $U = B_{2\rho}(0) \cap f^{-1}(B_\rho(0))$ et $V = B_\rho(0)$, on obtient que f est bijective de U sur $f(U) = V$ (puisque f étant surjective, on a $f(f^{-1}(B_\rho(0)) = B_\rho(0)$ et donc d'une part, $f(U) = f(B_{2\rho}(0)) \cap B_\rho(0) \subset B_\rho(0) = V$ et d'autre part par (c) $B_\rho(0) \subset f(B_{2\rho}(0))$ et donc $V = B_\rho(0) \subset f(U) = f(B_{2\rho}(0)) \cap B_\rho(0)$).

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 (b) que $\|f^{-1}(y) - f^{-1}(y')\| \leq 2\|y - y'\|$.
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 $\mathscr{C}^1$ de l'inverse local.*
On veut $f^{-1} \in \mathscr{C}^1(V)$, ce qui compte tenu du Théorème IV.4, revient à prouver que $\forall x \in U$, $df(x) \in \mathrm{Isom(E)}$.

Or $df(0) \in \mathrm{Isom(E)}$ qui est ouvert dans $\mathbb{B}(E)$ et donc $\exists r_1 > 0$ tel que $B_{r_1}(df(0)) \subset \mathrm{Isom(E)}$.
Mais par continuité de df en 0 (f est $\mathscr{C}^1$) on a $\exists r_2 > 0$, $\forall x \in B_{2r_2}(0)$, $\|df(x) - df(0)\| \leq r_1$.
Ainsi, $\exists r_2 \in]0, \rho]$, $\forall x \in B_{2r_2}(0)$, $df(x) \in \mathrm{Isom(E)}$.
On obtient donc le résultat en prenant ρ dans l'étape 2 éventuellement plus petit (le remplacer par r_2 et adapter les définitions de U et V) et en tuilisant le Théorème IV.4.

■

FIN COURS 10

C Le Théorème d'inversion globale

On peut se demander à quelle condition le T.I.L devient global (ie. f difféomorphisme sur Ω en entier)
Si on prend $f : \mathbb{R} \longrightarrow \mathbb{R}$ vérifiant les hypothèses du T.I.L en tout point, on voit que l'homéomorphisme local deviendra global dès que f est injective. Si de plus, on veut la régularité $\mathscr{C}^1$ de la bijection réciproque il suffit de plus que f' ne s'annule pas (donc soit inersible).
C'est exactement ce que donne dans un cadre plus général le résultat suivant.

Théorème IV.9 (théorème d'inversion globale)

Soient E et F deux espaces de Banach ($dim(E) = dim(F)$ si la dimension est finie), $U \subset E$ ouvert et $f : U \longrightarrow F$ une application $\mathscr{C}^1$.
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 $x \in U$, $df(x) \in \mathrm{Isom(E,F)}$ et .

Preuve du Théorème IV.9.
• *1)* $\Longrightarrow$ *2)* est évidente par Théorème IV.4.
• Montrons *2)* $\Longrightarrow$ *1)*.
On suppose donc que f est $\mathscr{C}^1$ et injective sur U et que pour tout $x \in U$, $df(x) \in \mathrm{Isom(E,F)}$.
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 $U_0 \subset U$ ouvert de E).
Soit $U_0 \subset U$ 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 $x_0 \in U_0$, $f \in \mathscr{C}^1(U_0)$ et $df(x_0) \in \mathrm{Isom(E,F)}$.
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 ouvert implique $V = f(U)$ 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 : $f \in \mathscr{C}^1$ est bijective de U sur $f(U)$ d'inverse continue sur $f(U)$ ie. f homéomorphisme $\mathscr{C}^1$ de U sur V.
Puisque pour tout $x \in U$, $df(x) \in \mathrm{Isom(E,F)}$, on déduit le résultat (f difféo) du théorème IV.4.

■

3 Théorème des fonctions implicites

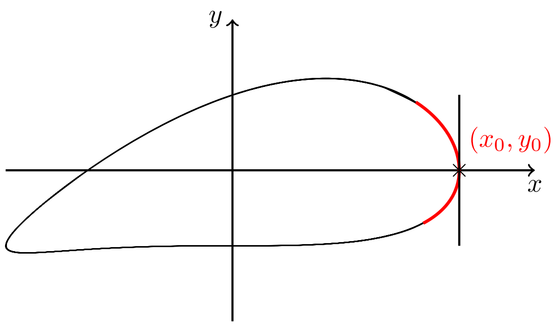
Commençons par motiver ce résultat avant d'en donner l'énoncé.

A Motivation

Là encore, pour simplifier, considérons une fonction $f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ de classe $\mathscr{C}^1$ et intéressons nous à la courbe de niveau $\alpha \in \mathbb{R}$ de f c'est-à-dire à l'ensemble $\{f = \alpha\} = \{(x, y) \in \mathbb{R}^2,\ f(x, y) = \alpha\}$. Supposons que cet ensemble soit non vide, non réduit à un point et que l'on en connaisse un point (x_0, y_0).
Que peut-on alors dire de la courbe $\{f = \alpha\}$ au voisinage de (x_0, y_0) ? Plus précisément, cette courbe peut-elle être décrite localement comme le graphe d'une fonction φ de la seule variable x i.e

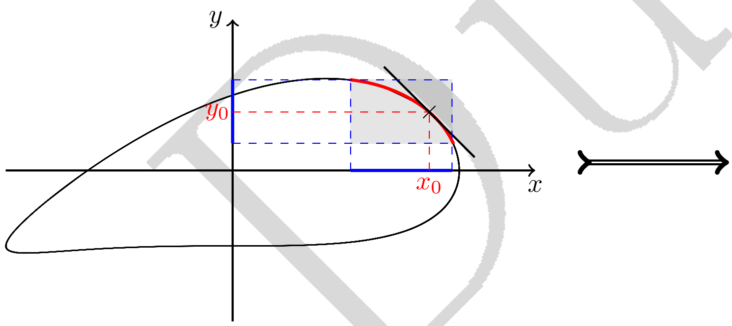
existe-t-il une fonction φ telle que $\{f = \alpha\} = \{(x, \varphi(x)),\ x$ dans un voisinage de $x_0\}$?

Commençons par lever une obstruction. Si on est dans le cas suivant, la réponse à la question posée sera visiblement non :

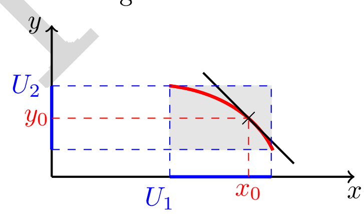


Dans cette configuration, la courbe de niveau $\{f = \alpha\}$ ne pourra jamais être décrite par une relation de la forme "$y = \varphi(x)$" au voisinage de (x_0, y_0) (partie rouge de la courbe) car alors φ serait multivoque (plusieurs images pour un même x). En outre, on peut réduire autant qu'on veut la partie rouge (ie le voisinage), le problème perdure.
En fait, la fonction f étant régulière, cette configuration se produit car la tangente à la courbe de niveau est verticale au point (x_0, y_0), c'est-à-dire $\frac{\partial f}{\partial y}(x_0, y_0) = 0$. La condition $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ apparait donc comme nécessaire.

Considérons donc un autre point $(x_0, y_0) \in \{f = \alpha\}$ pour lequel $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ ou en d'autres termes la tangente à la courbe de niveau n'est pas verticale au point (x_0, y_0).



Regardons uniquement ce qui se passe dans la fenêtre grise :



On voit dans ce cas que la partie rouge de la courbe de niveau, peut effectivement s'obtenir comme graphe d'une certaine application $\varphi : U_1 \longrightarrow U_2$.
De plus, la régularité de φ coïncide avec celle de la courbe de niveau, donc avec celle de f.

C'est ce type de résultat qui est donné, dans un cadre plus général, par le théorème des fonctions implicites.
Dans l'exemple précédent, la dérivée partielle de f par rapport à y jouait un rôle central. Si on veut pouvoir énoncer ce type de résultat dans le cas général où $f : E_1 \times E_2 \longrightarrow F$ avec E_1, E_2 et F e.v.n, il nous donc faut généraliser la notion de dérivée partielle. C'est ce qui conduit à la notion de différentielle partielle que nous allons maintenant voir.

B 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 $f_{x_1} : \begin{array}{ccc} U_{x_1} & \longrightarrow & F \\ y & \longmapsto & f(x_1, y) \end{array}$ 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

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

Résumons tout cela dans une définition/proposition :

Définition IV.10 (différentielle partielle)

On suppose que $f : U \subset E_1 \times E_2 \longrightarrow F$ est différentiable.
L'application $(x_1, x_2) \in U \mapsto df_{x_1}(x_2) \in \mathbb{B}(E_2, F)$ est appelée différentielle partielle de f relativement à sa seconde variable. On la note $\partial_2 f$.
On définit de même l'application $\partial_1 f : (x_1, x_2) \in U \mapsto df_{x_2}(x_1) \in \mathbb{B}(E_1, F)$ qui est appelée différentielle partielle de f relativement à sa première variable.
Pour tout $(x_1, x_2) \in U$ et tout $h = (h_1, h_2) \in E_1 \times E_2$, on a

$$df(x_1, x_2)h = df(x_1, x_2)(h_1, 0) + df(x_1, x_2)(0, h_2) = \partial_1 f(x_1, x_2)h_1 + \partial_2 f(x_1, x_2)h_2 \tag{IV.2}$$

Remarque IV.11

La notion de différentielle partielle telle que nous l'avons définie n'a donc de sens que pour les f différentiables ce qui sera toujours le cadre dans lequel nous nous placerons.
Ce n'est évidement pas nécessaire : pour que f admette en $(x_1, x_2) \in U$ une différentielle partielle par rapport à sa seconde variable il suffit qu'il existe $L_2 \in \mathbb{B}(E_2, F)$ telle que pour tout $h_2 \in E_2$ au vosiinage de $h_2 = 0$, on ait

$$f(x_1, x_2 + h_2) = f(x_1, x_2) + L_2(h_2) + o(\|h_2\|_{E_2})$$

Dans ce cas, f peut évidement admettre des différentielles partielles sans être différentiable et la formule (IV.2) n'est alors plus valide.

C Enoncé du théorème des fonctions implicites

On a donc le résultat suivant :

Théorème IV.12 (théorème des fonctions implicites)

Soit E_1, E_2 et F des espaces de Banach vérifiant $dim(E_2) = dim(F)$ si la dimension est finie, U un ouvert de $E_1 \times E_2$, et $f : U \longrightarrow F$ une application de classe $\mathscr{C}^{k+1}$ sur U pour un $k \geq 0$ 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. $\partial_2 f(x_0, y_0) \in \mathrm{Isom}(E_2, F)$
(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 contenant x_0 ;
il existe $U_2 \subset E_2$ ouvert contenant y_0 ;
il existe $\varphi : U_1 \longrightarrow U_2$ application de classe $\mathscr{C}^{k+1}$ sur U_1, appelée *fonction implicite* ;

tels que

$$\forall (x, y) \in U_1 \times U_2, \ \Big(f(x, y) = 0 \Longleftrightarrow y = \varphi(x)\Big) \tag{IV.3}$$

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

Différentes remarques s'imposent avant de passer à la preuve de ce résultat.

Remarque IV.13

1. L'hypothèse $f(x_0, y_0) = 0$ n'est aucunement restrictive puisque si tel n'est pas le cas, il suffit de translater f.
2. Comme attendu, la fonction implicite φ a la même régularité que f, donc de classe $\mathscr{C}^{k+1}$.
3. Le théorème est évidement symétrique au sens où l'on peut obtenir une fonction implicite de la variable y ie $x = \varphi(y)$. Il faut naturellement adapter les hypothèses du théorème : $dim(E_1) = dim(F)$ en dimension finie et *(ii)* doit être remplacée par $\partial_1 f(x_0, y_0) \in \text{Isom}(E_1, F)$.

Revenons au théorème.

Preuve du Théorème IV.12.

On va appliquer le Théorème d'inversion locale (Th. IV.8) 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 $\mathscr{C}^{k+1}$ 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 de la définition IV.10 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

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

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 IV.8) 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 $\mathscr{C}^{k+1}$ 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 $(x,z) \in W$[1] .

Définissons alors

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

qui est donc un voisinage ouvert de x_0 (puisque $x_0 \in U_1'$ et que $(x_0, 0) \in W$ qui est tout vert) puis

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

Alors on a $\varphi \in \mathscr{C}^{k+1}(U_1)$ 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}

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

---

[1] Puisque $g^{-1} = (g_1^{-1}, g_2^{-1})$, on $g^{-1}(x,z) = (a,b) \iff (x,z) = g(a,b) \iff \big(x = a \text{ et } z = f(a,b)\big)$ et donc en identifiant $g_1^{-1}(x,z) = a = x$

ce qui conclut la preuve du Théorème IV.12.

■

Terminons ce paragraphe par un corollaire direct, mais fort utile, du théorème des fonctions implicites.

Corollaire IV.14

Les hypothèses sont celles du Théorème IV.12.
On a a lors les trois résultats suivants :

1. La relation (IV.3) entraine que :

$$\forall x \in U_1, \quad f(x, \varphi(x)) = 0. \tag{IV.4}$$

2. On peut supposer que les ouverts U_1 et U_2 obtenus dans le Théorème IV.12 sont tels que

$$\forall (x, y) \in U_1 \times U_2, \quad \partial_2 f(x, y) \in \mathrm{Isom}(E_2, F). \tag{IV.5}$$

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

$$d\varphi(x)h = -\left(\partial_2 f(x, \varphi(x))\right)^{-1} \circ \partial_1 f(x, \varphi(x))(h). \tag{IV.6}$$

Preuve du Corollaire IV.14.
• Partie *1.* La relation (IV.4) est contenue dans (IV.3).

• Partie *2.* C'est une conséquence du fait que $\partial_2 f(x_0, y_0) \in \mathrm{Isom}(E_2, F)$ ouvert de $\mathbb{B}(E_2, F)$ et que f étant $\mathscr{C}^1(U)$, 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 étape 3).

• Partie *3.* Il suffit de différentier la relation (IV.4) 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

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

et on conclut alors grace à (IV.5).

■

FIN COURS 11

Remarque IV.15 (cas de la dimension finie)

1. Supposons que $E_1 = \mathbb{R}^n$ et $E_2 = F = \mathbb{R}^p$.
   Alors, on a les équivalences suivantes :

$$\begin{aligned} & \partial_2 f(x_0, y_0) \in \mathrm{Isom}(\mathbb{R}^p) && (1) \\ \iff & J_{f_{x_0}}(y_0) \in \mathscr{M}_p \text{ inversible} && \\ \iff & J_f(x_0, y_0) \in \mathscr{M}_{p,n+p} \text{ admet une sous matrice de rang } p && (2) \\ \iff & df(x_0, y_0) \text{ surjective} && (3) \\ \iff & (\nabla f_1(x_0, y_0), \dots, \nabla f_p(x_0, y_0)) \text{ est une famille libre (de } (\mathbb{R}^{n+p})^p) \text{ par définition} && (4) \\ & \text{du rang (nombre de vecteurs ligne ou colonne linéaire indépendants)} && \end{aligned}$$

   On peut donc reformuler le Théorème IV.12 en remplaçant (1) (hypothèse *(ii)* du théorème) soit par (2), soit par (3), soit par (4).

2. Il se peut qu'on ait à appliquer le Théorème IV.12 avec $\Omega \subset \mathbb{R}^N$ et $f : \Omega \longrightarrow \mathbb{R}^p$, avec $N > p$ et, par exemple, $df(X_0)$ surjective. Dans ce cas, on écrit $\mathbb{R}^N = \mathbb{R}^{N-p} \times \mathbb{R}^p = E_1 \times E_2$ avec $X_0 = (x_0, y_0) \in \mathbb{R}^{N-p} \times \mathbb{R}^p$.

D Espace tangent - Application aux courbes de niveau de fonctions $\mathscr{C}^1$

Cette partie nous servira dans le paragraphe suivant (extrema liés) où nous aurons besoin des notions suivantes. A partir de maintenant et même si ce n'est pas indispensable, jusqu'à la fin de ce cours on se placera en dimension finie et on notera :

$$\Omega \subset \mathbb{R}^n \text{ un ouvert et } g : \mathbb{R}^n \longrightarrow \mathbb{R}^p \text{ vérifiant } g \in \mathscr{C}^1(\Omega).$$

On notera V le graphe de g à savoir l'ensemble

$$V_g = \{(x, y) \in \Omega \times F \ ; \ y = g(x)\}.$$

La notion de tangente à une courbe (cas $n = p = 1$) se généralise de la manière suivante.

Définition IV.16 (espace tangent - espace affine tangent)

Soit $M_0(x_0, g(x_0)) \in V_g$.
Le graphe de $dg(x_0)$, noté $T_{M_0}(V_g) = \{(h, k) \in \mathbb{R}^n \times \mathbb{R}^p \ ; \ k = dg(x_0)h\}$ est appelé *espace tangent* à V_g en M_0.
On lui associe de manière naturelle *l'espace affine tangent* :
$T^a_{M_0}(V_g) = M_0 + T_{M_0}(V_g) = \{(h, k) \in \mathbb{R}^n \times \mathbb{R}^p \ ; \ k = g(x_0) + dg(x_0)(h - x_0)\}$

On retrouve les résultats classiques suivants :

Exemple IV.17

1. Pour $g : \mathbb{R} \longrightarrow \mathbb{R}$:
   - V_g est donc une courbe ;
   - $T^a_{M_0}(V_g)$ la droite tangente d'équation $(y = g(x_0) + g'(x_0)(x - x_0))$.
2. Pour $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$:
   - V_g est une surface ;
   - $T^a_{M_0}(V_g)$ est le plan tangent d'équation $(z = g(x_0, y_0) + \dfrac{\partial g}{\partial x}(x_0, y_0)(x - x_0) + + \dfrac{\partial g}{\partial y}(x_0, y_0)(y - y_0))$.

Venons en maintenant, sous forme de remarque, à ce qui motive ces définitions.
On considère toujours une fonction $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ telle que que $g \in \mathscr{C}^1(\Omega)$ avec $\Omega \subset \mathbb{R}^n$ ouvert et on suppose que de plus $n > p$.
On introduit alors la courbe de niveau 0 de g que l'on supposera non vide et non réduite à un point

$$U = \{x \in \Omega, \ g(x) = 0\}.$$

C'est une partie fermée de $\mathbb{R}^n$ puisque $U = g^{-1}(\{0\})$ avec g continue. Cette courbe de niveau peut alors être décrite localement implicitement de la façon suivante :

Remarque IV.18

S'il existe $a \in \Omega$ tel que $dg(a)$ soit surjective, écrivant $\mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$, on déduit du théorème des fonctions implicites que

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

Dans ce cas, on a donc

$$U \cap (U_1 \times U_2) = \{(x_1, x_2) \in U_1 \times U_2, \ x_2 = \varphi(x_1)\}. \tag{IV.7}$$

En termes imagés cela revient à regarder U à travers la fenêtre $U_1 \times U_2$.

On a alors la définition suivante.

Définition IV.19 (paramétrisation locale)

Soit la fonction $\Psi : \mathbb{R}^{n-p} \longrightarrow \mathbb{R}^n$, $\Psi \in \mathscr{C}^1(U_1)$ 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 IV.20

On utilise les hypothèses et notations précédentes sur g ($g : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p \in \mathscr{C}^1$, $dg(a)$ surjective).
On pose $V_\varphi = U \cap (U_1 \times U_2)$ et on note $T_a(V_\varphi)$ l'espace tangent à V_φ en a.
On a alors les égalités suivantes

$$T_a(V_\varphi) = \{(h_1, h_2) \in \mathbb{R}^{n-p} \times \mathbb{R}^p,\ h_2 = d\varphi(a_1)h_1\} \tag{IV.8}$$

$$= Ker(dg(a)) = \bigcap_{i=1}^{p} Ker(dg_i(a)) \tag{IV.9}$$

$$= Im(d\Psi(a_1)) \tag{IV.10}$$

Noter qu'en particulier (IV.9) montre que $dim(T_a(V_\varphi)) = n - p$ (puisque $dg(a)$ surjective).

Preuve de la Proposition IV.20
Les notations sont celles utilisées au début de ce paragraphe.
• L'égalité (IV.8) est donnée par la définition de $T_a(V_\varphi)$ et l'égalité (IV.7).
• Egalité (IV.9). On a $(h, k) \in Ker(dg(a))$ ssi $\partial_1 g(a)h + \partial_2 g(a)k = 0$ ssi $k = -\Big(\partial_2 g(a)\Big)^{-1} \partial_1 g(a)h = d\varphi(a_1)h$ donc ssi $(h, k) \in T_a(V_\varphi)$.
La seconde égalité vient de ce que $dg(a)h = 0$ ssi $dg_i(a)h = 0$ pour tout $i = 1, .., p$.
• Egalité (IV.10). Par surjectivité, on a $dim(Ker(dg(a))) = n - p$.
Or $d\psi(a_1) \in \mathbb{B}(\mathbb{R}^{n-p}, \mathbb{R}^n)$ est injective puisque $d\psi(a_1)h = 0 \Longrightarrow h = 0$ et donc $dim(Im(d\psi(a_1))) = n - p$.
Mais puisque pour tout $x_1 \in U_1$, $g(\psi(x_1)) = 0$ on a en différentiant, $dg(\psi(a_1))d\psi(a_1) = 0$ 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 IV.21

Si $p = 1$, la proposition IV.20 montre que $\nabla g(a)$ est orthogonal à la courbe de niveau $\{g = g(a)\}$ en a. En effet, si $H \in T_a(V_\varphi)$ alors $H \in Ker(dg(a))$ donc $dg(a)H = 0$ et donc $< \nabla g(a), H >= 0$.

On peut montrer que de plus, $\nabla g(a)$ est orienté dans la direction de plus grande pente de g au point a : pour le montrer on calcule la dérivée de g en a suivant une direction $d = (u, v)$ avec $\|(u, v)\| = 1$. Celle-ci donne la pente de g en a dans la direction d. Or elle est donnée par $g'(a)(d) =< \nabla g(a).d >_2= \|\nabla g(a)\|_2 \|d\|_2 \cos(\alpha)$ où $\alpha \in [0, \pi[$ mesure l'angle fait par les deux vecteurs. Elle est donc maximale lorsque $\cos(\alpha) = 1$ à savoir $\alpha = 0$ c'est-à-dire pour d et $\nabla g(a)$ orientés dans la même direction.

Venons en à la dernière partie de ce cours qui a motivé le paragraphe précédent comme nous allons le voir.

4 Extrema locaux liés : méthode des multiplicateurs de Lagrange

On se donne un ouvert $\Omega \subset \mathbb{R}^n \times \mathbb{R}^p$ ainsi que deux fonctions $f : \Omega \longrightarrow \mathbb{R}$, $g : \Omega \longrightarrow \mathbb{R}^p$ vérifiant $f, g \in \mathscr{C}^1(\Omega)$.
Posons enfin $U = \{x \in \Omega,\ g(x) = 0\}$ et considérons le problème d'optimisation suivant :

Trouver $a \in U$ tel que $f(a)$ soit un extremum local de f sur U

On dit que l'on cherche à *optimiser* f *sous la contrainte* (**s.c**) U (sous-entendu *localement*) .
On a alors la définition suivante

Définition IV.22 (Extrema liés)

On dit que U est la *contrainte*.
On dit que $a \in U$ est un extremum de f sous la contrainte (**s.c**) U si il existe un ouvert $\mathcal{O} \subset \Omega$ tel que $\forall x \in U \cap \mathcal{O}$ $f(x) \leq f(a)$ (ou $f(x) \geq f(a)$).
On dit dans ce cas que a est un *extremum lié*, la liaison étant donnée par la contrainte U.

Remarque IV.23

1. Attention, d'un point de vue topologique, ce problème est vraiment différent de celui de la recherche d'extrema libres car ici U est un fermé (g continue).
   Pour s'en persuader, considérer $f(x) = x \in \mathbb{R}$ s.c $x \in U = [0, 1]$: l'étude des points critiques ne donne rien et pourtant f admet un maximum et un minimum sur U.
2. Puisque $a \in U$ équivaut à $g(a) = 0$, il y a un lien entre les différentes coordonnées de a ce qui justifie la terminologie d'extrema liés.
3. Dans notre cas la contrainte est de la forme $\{g = 0\}$: on parle donc de *contrainte d'égalité*.
   On peut aussi être amené à optimiser sous *contrainte d'inégalité*, l'ensemble U étant dans ce cas de la forme $U = \{g \leq 0\}$.

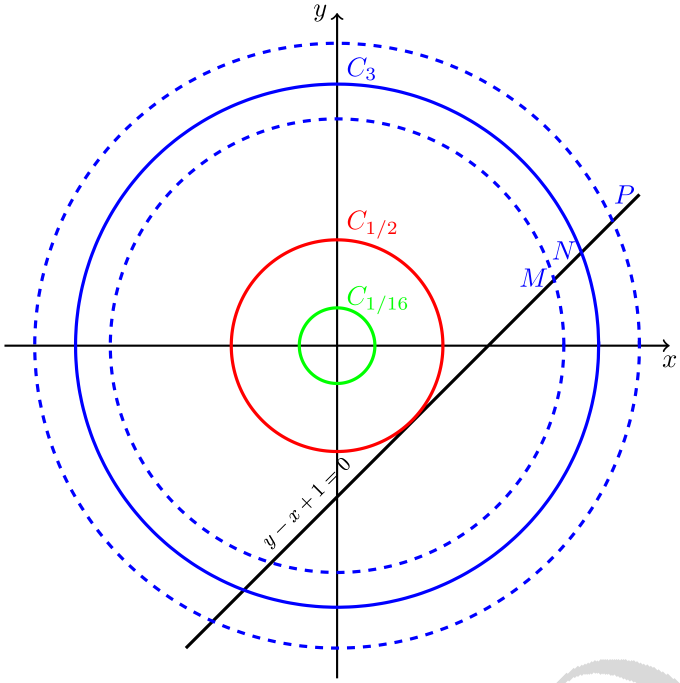
Avant de traiter le cas général, essayons de comprendre ce qui peut se passer à l'aide d'un exemple.

Exemple IV.24

Soit à optimiser, la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *définie par* $f(x_1, x_2) = x_1^2 + x_2^2$ *sous la contrainte* $U = \{(x_1, x_2) \in \mathbb{R}^2,\ x_2 - x_1 + 1 = 0\}$ *(donc* $g(x_1, x_2) = x_2 - x_1 + 1$ *ici).*
Dans ce cas, $U = \{(x_1, x_1 - 1) \in \mathbb{R}^2,\ x_1 \in \mathbb{R}\}$ et donc pour tout $(x_1, x_2) \in U$, on a $f(x_1, x_2) = f(x_1, x_1 - 1)$. Pour ce cas particulier, le problème se ramène donc à un problème d'extrema libres (ce qui n'est pas toujours le cas) pour $F(x_1) = f(x_1, x_1 - 1)$ sur l'ouvert $\mathbb{R}$.

Représentons f dans $\mathbb{R}^3$ sous la forme $z = f(x, y)$ alors pour $z = \alpha$ fixé, on obtient (dans le plan $z = \alpha$) la courbe de niveau α de f que l'on notera C_α. On représente dans ce même plan la contrainte qui sera une droite.
On procède ainsi pour différentes valeurs de α et on représente toutes ces courbes dans un même plan.

On obtient ainsi une représentation de la forme suivante :



Commençons par noter que, pour C_α donné, s'il y a des extrema ce sont nécessairement des éléments de $U \cap C_\alpha$. On peut donc faire les constatations suivantes en fonction des valeurs de α :

- Si $\alpha < \frac{1}{2}$ (courbe verte) alors l'ensemble $U \cap C_\alpha$ est vide et il ne peut donc pas y avoir d'extremum.
- Si $\alpha > \frac{1}{2}$ (courbe bleue), il y a deux points d'intersections mais aucun ne peut-être un extremum. En effet, par exemple pour N : lorsqu'on augmente α pour obtenir P, la valeur de f augmente et donc N ne peut pas être un maximum et de même si on diminue α pour M, ce ne peut pas être un minimum.

La seule possibilité qui reste est la courbe rouge (ce qui est confirmé par le calcul des extrema de F).

Remarque IV.25 (conclusion de cet exemple)

On peut tirer deux enseignements de cet exemple :

1. On peut visiblement, au moins dans certains cas, ramener le problème d'extrema liés à un problème d'extrema libres en substituant la contrainte dans la fonction à optimiser. L'intérêt est alors qu'on dispose de tous les outils mis en place pour la recherche de tels extrema.
2. Dans cet exemple, seule la courbe rouge peut convenir et le raisonnement ayant conduit à cette conclusion semble montrer qu'un point a ne pourra être un extremum de f s.c U que si "*la contrainte est tangente à la courbe de niveau* $\{f = f(a)\}$ *en ce point*". Cette remarque, pour l'instant formelle, s'avèrera être une propriété générale comme nous le verrons par la suite.

Nous allons maintenant voir la méthode dite des multiplicateurs de Lagrange qui formalise la remarque précédente. Elle possède en outre le gros avantage d'avoir une formulation générale implémentable numériquement. Elle repose sur le résultat suivant

Théorème IV.26

Soit $\Omega \subset \mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$ un ouvert, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ et $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ deux fonctions vérifiant $f, g \in \mathscr{C}^1(\Omega)$ et $U = g^{-1}(\{0\})$.
Si $a \in U$ est un extremum de f s.c U et que $\partial_2 g(a) \in \mathrm{Isom}(\mathbb{R}^p)$ (voir remarque IV.15 pour différentes formulations de cette hypothèse) alors il existe $\lambda_a \in \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ tel que

$$df(a) + \lambda_a \circ dg(a) = 0 \tag{IV.11}$$

ou, de manière équivalente, il existe p scalaires λ_i $(i = 1, .., p)$ tels que

$$\nabla f(a) + \sum_{i=1}^{p} \lambda_i \nabla g_i(a) = 0.$$

On dit que l'application $\lambda_a \in \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ ou le vecteur $(\lambda_i)_{i=1}^p \in \mathbb{R}^p$ est le *multiplicateur de Lagrange associé f s.c U.*

Preuve du Théorème IV.26.
Si $a \in U$ est un extremum de f s.c U alors

$$a \in U \text{ et } \exists \mathcal{O} \subset \Omega \text{ voisinage ouvert de } a \text{ tel que } \forall x \in \mathcal{O} \cap U,\ f(a) \leq f(x) \text{ (ou } f(a) \geq f(x))$$

Si de plus $\partial_2 g(a) \in \mathrm{Isom}(\mathbb{R}^p)$ alors (voir Remarque IV.18) :

$$\exists U_1, U_2 \text{ voisinages ouverts de } a_1 \text{ et } a_2 \text{ respectivement, et } \varphi \in \mathscr{C}^1(U_1) \text{ tels que}$$
$$V = U \cap (U_1 \times U_2) = \{(x_1, x_2) \in U_1 \times U_2,\ x_2 = \varphi(x_1)\}.$$

On note (U_1, ψ) la paramétrisation locale associée.

Alors, quitte à prendre U_1 et U_2 plus petits, on peut supposer que $U_1 \times U_2 \subset \mathcal{O}$.

Ainsi, on aura $\forall (x_1, x_2) \in V,\ f(x_1, x_2) = f(x_1, \varphi(x_1)) = f(\psi(x_1)) := F(x_1)$ et donc en particulier, $f(a_1, a_2) \leq f(x_1, x_2) \iff F(a_1) \leq F(x_1)$.

Ainsi, la recherche des extrema de f restreinte à V s.c U se ramène-t-elle à celle des extrema de $F = f \circ \psi$ sur l'ouvert U_1 (en ce sens la méthode du Lagrangien est une méthode de substitution).

Le gros avantage est que, ce faisant, *on s'est ramené à un problème d'extrema libres sur l'ouvert U_1 pour F.*

Ainsi 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 III.29. On a donc

$$dF(a_1) = 0 = df(\psi(a_1)) \circ d\psi(a_1).$$

(on différentie donc $f_{|V}$ en $a \in V$ dans une direction $d\psi(a) \in T_a(V)$ - puisque $Im(d\psi(a)) = T_a(V)$).
Or, puisque $F = f \circ \psi$ et que $\forall x_1 \in U_1,\ \psi(x_1) = (x_1, \varphi(x_1))$ on a puisque $a_1 \in U_1$

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

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

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

et écrivant de plus que

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

on trouve la relation $df(a) = -\lambda_a dg(a)$ en posant

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

■

Remarque IV.27

1. La fonction F introduite dans la preuve précédente montre que la méthode du Lagrangien peut-être vue comme une méthode de substitution (locale) permettant de se ramener du problème d'extrema liés à un problème d'extrema libres.
2. Le multiplicateur de Lagrange est donné "explicitement" (formule (a)) dès lors que a est connu.
3. Puisque par le théorème IV.26 on a $df(a) = -\lambda_a dg(a)$, il en résulte que $\forall h \in Ker(dg(a))$, $df(a)h = -\lambda_a dg(a)h = 0$ et donc $df(a) \in \Big(Ker(dg(a))\Big)^{\perp}$. Puisque par la proposition IV.20 $T_a(V) = Ker(dg(a))$, on obtient

$$df(a) \in \Big(T_a(V)\Big)^{\perp}$$

   La condition nécessaire d'extrema liés du théorème, traduit donc le fait que la forme linéaire $df(a)$ restreinte à l'espace tangent $T_a(V)$ est nulle.
4. En particulier, lorsque $p = 1$, la définition de l'espace tangent et la relation (IV.11) montrent que $T_a(V_f) = T_a(V_g)$. On retrouve donc bien, comme intuité dans la remarque IV.25, qu'en un extremum, la contrainte est tangente à la courbe de niveau correspondante de f.

Afin de pouvoir exploiter ce résultat de manière concrète, nous aurons besoin de la définition/propriété suivante. Théorème IV.26.

Corollaire IV.28

Soit $\Omega \subset \mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$ un ouvert, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ et $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ deux fonctions vérifiant $f, g \in \mathscr{C}^1(\Omega)$ et $U = g^{-1}(\{0\})$.
L'application L définie pour tout $(x, \lambda) \in \Omega \times \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ par

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

est appelée *Lagrangien de f s.c U*.
Si $a \in U$ est un extremum de f s.c U pour lequel $\partial_2 g(a) \in \text{Isom}(\mathbb{R}^p)$ 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 IV.28.
Il suffit que remarquer qu'un point critique $(x, \lambda) \in \Omega \times \mathbb{B}(\mathbb{R}^p, \mathbb{R})$ du Lagrangien L est, par définition, solution de $dL(x, \lambda) = 0$. Il 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 IV.29

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

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

On doit donc résoudre un *système non linéaire de $(n+p)$ équations à $(n+p)$ inconnues* (les coordonnées des vecteurs $x \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}^p$).
C'est le prix à payer pour se ramener du problème d'extrema liés à un problème d'extrema libres que l'on sait traiter et pour lequel le système à résoudre est non linéaire de taille $n \times n$.

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 IV.30

Soit $\Omega \subset \mathbb{R}^n = \mathbb{R}^{n-p} \times \mathbb{R}^p$ un ouvert, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ et $g : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ des fonctions deux fois différentiables sur Ω et $U = g^{-1}(\{0\})$.
Soit (a, λ) un point critique du Lagrangien L pour lequel $\partial_2 g(a) \in \text{Isom}(\mathbb{R}^p)$.
Si l'application bilinéaire symétrique $d^2 f(a) + \lambda d^2 g(a)$ restreinte à $T_a(V) = Ker(dg(a))$ est :

1. définie positive alors a est un minimum de f s.c U ;
2. définie négative alors a est un maximum de f s.c U ;
3. ni positive, ni négative alors a n'est pas un extremum de f s.c U ;
4. dans les autres cas (positive ou négative mais non défine) on ne peut rien dire directement.

Preuve du Théorème IV.30.
On a vu dans la preuve du théorème IV.26 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 (a, λ) est un point critique du Lagrangien L alors $a_1 \in U_1$ est un point critique de F puisque :

$$\begin{array}{rcl} dF(a_1) & = & df(a) \circ d\psi(a_1) = -\lambda \circ dg(a) \circ d\psi(a_1) = -\lambda \circ (\partial_1 g(a) + \partial_2 g(a) \circ d\varphi(a_1)) \\ & = & -\lambda \circ (\partial_1 g(a) + \partial_2 g(a) \circ \{(-(\partial_2 g(a))^{-1}\partial_1 g(a))\}) = 0 \end{array}.$$

Ainsi, on peut appliquer les théorèmes III.34 et III.35 à F et il faut donc pour cela étudier la forme bilinéaire symétrique $d^2F(a_1)$.

$\diamond$ Calculons donc son expression.
Or (a, λ) étant un point critique de L, il vérifie $df(a) + \lambda dg(a) = 0$, et donc on a

$$d^2F(a_1) = d^2f(a)\Big(d\psi(a_1), d\psi(a_1)\Big) + df(a)d^2\psi(a_1) = d^2f(a)\Big(d\psi(a_1), d\psi(a_1)\Big) - \lambda dg(a)d^2\psi(a_1).\lambda_a = -\partial_2 f(\psi(a_1)) \circ \Big(\partial_2 g(a)\Big)^{-1}.$$

Or $\forall x_1 \in U_1$, $g(\psi(x_1)) = 0$, d'où $dg(\psi(x_1))d\psi(x_1) = 0$, et donc $d^2g(a)\Big(d\psi(a_1), d\psi(a_1)\Big) + dg(a)d^2\psi(a_1) = 0$, c'est-à-dire $dg(a)d^2\psi(a_1) = -d^2g(a)\Big(d\psi(a_1), d\psi(a_1)\Big)$ et ainsi $d^2F(a_1) = \Big(d^2f(a) + \lambda d^2g(a)\Big)\Big(d\psi(a_1), d\psi(a_1)\Big)$..
L'application bilinéaire symétrique $Q = d^2F(a_1)$ est donc donnée pour tout $h \in \mathbb{R}^n$ par

$$Q(h) = \Big(d^2f(a) + \lambda d^2g(a)\Big)(d\psi(a_1)h, d\psi(a_1)h) = \Big(d^2f(a) + \lambda d^2g(a)\Big) \circ (d\psi(a_1), d\psi(a_1))h$$

c'est donc la restriction de $d^2f(a) + \lambda d^2g(a)$ à $Im(d\psi(a_1))$ et donc à $T_a(V) = Ker(dg(a))$ (proposition IV.20).
La conclusion résulte alors de l'application des théorèmes III.34 et III.35 à $Q = d^2F(a_1)$.

■

Terminons cette partie en montrant sur un exemple comment appliquer cette méthode.

Exemple IV.31

On veut optimiser $f(x, y) = xy$ sous la contrainte $g(x, y) = x^2 + 4y^2 - 8 = 0$.
• On commence donc par chercher les points critiques du Lagrangien $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ (le multiplicateur $\lambda \in \mathbb{R}$ ici puisque g est à valeur dans $\mathbb{R}$). Ce sont les solutions de $dL(x, y, \lambda) = 0$ donc de $\nabla L(x, y, \lambda) = 0$ donc du système

$$(S) \left\{ \begin{array}{l} \frac{\partial L}{\partial x}(x, y, \lambda) = y + 2\lambda x = 0 \\ \frac{\partial L}{\partial y}(x, y, \lambda) = x + 8\lambda y = 0 \\ \frac{\partial L}{\partial \lambda}(x, y, \lambda) = x^2 + 4y^2 - 8 = 0 \end{array} \right.$$

Noter que par définition de L, la dérivée partielle $\frac{\partial L}{\partial \lambda}$ est toujours égale à g. On a alors

$$(S) \iff \left\{ \begin{array}{l} x + 8\lambda y = 0 \\ (1 - 16\lambda^2)y = 0 \\ x^2 + 4y^2 = 8 \end{array} \right. \iff \left\{ \begin{array}{l} \lambda = \pm\frac{1}{4} \\ x = -8\lambda y \\ y^2 = 1 \end{array} \right.$$

et l'ensemble des points critiques est donc $\Big\{(2, 1, -\frac{1}{4}), (2, -1, \frac{1}{4}), (-2, -1, -\frac{1}{4}), (-2, 1, \frac{1}{4})\Big\}$.
• Etudions en détail le point $(2, 1, -\frac{1}{4})$ (noter que $\frac{\partial g}{\partial y}(2, 1) = 8 \neq 0$ et donc $\partial_2 g(2, 1) \in \mathrm{Isom}(\mathbb{R})$).
On étudie pour cela la forme bilinéaire symétrique $Q = d^2f(2, 1) - \frac{1}{4}d^2g(2, 1)$ sur $Ker(dg(2, 1))$.
Utilisant les dérivées partielles, on trouve que la matrice hessienne de $(f - \frac{1}{4}g)$ en $(2, 1)$ est donnée par $A = \begin{pmatrix} -\frac{1}{2} & 1 \\ 1 & -\frac{1}{2} \end{pmatrix}$ avec pour tout $H = (h, k) \in \mathbb{R}^2$, $(d^2f(2, 1) - \frac{1}{4}d^2g(2, 1))(H, H) =< AH, H >$ où $<\ ,\ >$ désigne le produit sclaire sur $\mathbb{R}^2$.
Déterminons alors $Ker(dg(2, 1))$:
On a $H = (h, k) \in Ker(dg(2, 1)) \iff 4h + 8k = 0 \iff h = -2k$ ie $Ker(dg(2, 1)) = vect(-2, 1)$.
Ainsi, la restriction de Q à $Ker(dg(2, 1))$ vérifie $Q(H) =< A\begin{pmatrix} -2k \\ k \end{pmatrix}, \begin{pmatrix} -2k \\ k \end{pmatrix}) >= -8k^2$. Elle est donc définie négative et le point $(2, 1)$ est donc un maximum de f s.c $\{g = 0\}$.

FIN DU CHAPITRE IV

Index