# Cours MASS 31

Yoann Dabrowski

8 octobre 2013

# Chapitre 1

# Espaces Vectoriels Normés

## 1 Rappels sur les espaces vectoriels

Dans tout le cours, $\mathbb{K} = \mathbb{R}$, (le corps des nombres réels) ou $\mathbb{C}$ (le corps des nombres complexes). $|\lambda|$ est la valeur absolue ou le module de $\lambda \in \mathbb{K}$.

**Définition 1.** Soit $E \neq \emptyset$ un ensemble muni de deux lois notées $+$ et $\cdot$. On dit que $(E, +, \cdot)$ est un $\mathbb{K}$-e.v. si :

a $+$ est une loi de composition interne $+ : E \times E \to E$ tel que $(E, +)$ est un groupe abélien, c'est-à-dire :

i $+$ est associative $\forall x, y, z \in E, x + (y + z) = (x + y) + z$

ii $+$ est commutative $\forall x, y \in E, x + y = y + x$

iii il existe un élément neutre $0$ : $0 + x = x$

iv tout $x$ admet un opposé $(-x)$ unique tel que $x + (-x) = 0$

b $\cdot : \mathbb{K} \times E \to E$ est une loi externe, telle que $\forall \lambda, \mu \in \mathbb{K} \forall x, y \in E$ :

i $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$

ii $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$

iii $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$

iv $(1 \cdot x) = x$

On note $\lambda \cdot x = \lambda x$

**Exemple fondamental $\mathbf{E = \mathbb{K}^n}$**, Soit $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in \mathbb{K}^n$ on pose $0_E = 0 = (0, ..., 0)$, $x + y = (x_1 + y_1, ..., x_n + y_n), \lambda x = (\lambda x_1, ..., \lambda x_n)$.

*Rappel* 1. On ne multiplie pas deux vecteurs $u, v \in E$ d'un espace vectoriel ($u \cdot v$ n'est pas défini, sauf si on ajoute un produit scalaire). On ne divise pas deux vecteurs.

*Rappel* 2. Si $\lambda \in \mathbb{R}$ on rappelle que $|\lambda| = \sqrt{\lambda^2}$ et si $\lambda = a + ib \in \mathbb{C}$ $(a, b \in \mathbb{R})$ $|\lambda| = \sqrt{a^2 + b^2}$, vérifient :

- $|\lambda\mu| = |\lambda||\mu|$
- $|\lambda| = 0 \iff \lambda = 0$
- $|\lambda + \mu| \leq |\lambda| + |\mu|$

On va généraliser aux espaces vectoriels les notions de valeur absolue et de module.

## 2 Norme sur un espace vectoriel

**Définition 2.** Soit $E$ un $\mathbb{K}$-e.v. Une norme sur $E$ est une application $n : E \to [0, +\infty[$ telle que :

i $\forall x \in E,\ \lambda \in \mathbb{K}\ n(\lambda x) = |\lambda| n(x)$ (homogénéité)

ii $\forall x, y \in E\ n(x+y) \leq n(x) + n(y)$ (inégalité triangulaire ou sous-additivité)

iii $\forall x \in E\ n(x) = 0 \iff x = 0$ (séparation)

Souvent on note $n(x) = ||x||$, sauf dans l'exemple $E = \mathbb{K}, n(x) = |x|$. Un couple $(E, ||.||)$ est appelé espace vectoriel normé (evn).

*Exemple* 1. Si $E = \mathbb{R}^2$, soit

$$N((x,y)) = \max(|x|, \frac{1}{2}|x| + |y|).$$

Vérifions que $N$ est une norme sur $E$. $N$ est bien positive. Vérifions l'homogénéité, soit donc $\lambda \in \mathbb{R}$ :

$$N(\lambda(x,y)) = N((\lambda x, \lambda y)) = \max(|\lambda x|, \frac{1}{2}|\lambda x| + |\lambda y|) = \max(|\lambda||x|, |\lambda|(\frac{1}{2}|x| + |y|)) = |\lambda| N((x,y)).$$

La séparation est facile à vérifier car si $N((x,y)) = 0$ alors $|x| = 0$ et $\frac{1}{2}|x| + |y| = 0$ d'où $|y| = 0$ et donc $(x,y) = (0,0)$.

Il reste l'inégalité triangulaire :

$$\begin{aligned} N((x,y) + (x',y')) &= \max(|x+x'|, \frac{1}{2}|x+x'| + |y+y'|) \\ &\leq \max(|x| + |x'|, \frac{1}{2}(|x| + |x'|) + |y| + |y'|) \\ &\leq N((x,y)) + N((x',y')). \end{aligned}$$

*Exemple* 2. Si $E = \mathbb{R}^n$ on a trois normes classiques, si $X = (x_1, ..., x_n)$ :

$$||X||_1 = \sum_{i=1}^{n} |x_i|$$

$$||X||_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \text{ (norme euclidienne)}$$

$$||X||_\infty = \max_{i=1...n} |x_i|$$

*Exercice* 1. Montrer que ce sont des normes (cf. TD).

*Exemple* 3. Si $E = C^0([a,b], \mathbb{R})$ l'ensemble des fonctions continues sur $[a,b]$, on a trois normes :

$$||f||_1 = \int_a^b |f(t)| dt$$

$$||f||_2 = \sqrt{\int_a^b |f(t)|^2 dt}$$

$$||f||_\infty = \sup_{t \in [a,b]} |f(t)|$$

Cette dernière norme est la norme de la convergence uniforme (une fois que l'on aura vu la notion de convergence pour une norme, la convergence pour $||.||_\infty$ coïncidera avec la convergence uniforme)

*Exemple* 4. Si $G = E \times F$ avec $(E, ||.||_E)$, $(F, ||.||_F)$ des evn. On définit : $||(x,y)||_G = max(||x||_E, ||y||_F)$. C'est une norme sur $G$ (exo) que l'on utilisera dans cette situation ultérieurement (norme produit).

*Remarque* 3. Pour $x, y \in E$, on pose $d(x,y) = ||x - y||$ et on dit que d(x,y) est la distance de $x$ à $y$. Dans le cas de la norme $||.||_2$ on retrouve la distance euclidienne. On utilise souvent $||.||_2$ quand la géométrie intervient. En analyse l'emploi de $||.||_1$ et $||.||_\infty$ est aussi courant.

**Proposition 1.** *(Inégalité triangulaire inverse) Soit* $(E, ||.||)$ *un evn.*

$$\forall x, y \in E \left| \ \|x\| - \|y\| \ \right| \leq ||x - y||$$

.

*Démonstration.* Cas $||x|| \geq ||y||$ : Comme $||x|| = ||x - y + y|| \leq ||x - y|| + ||y||$ par l'inégalité triangulaire, on en déduit $\left| \ \|x\| - \|y\| \ \right| = \|x\| - \|y\| \leq ||x - y||$.

Dans le cas $||y|| \geq ||x||$, on échange $x$ et $y$ par symétrie. □

# 3 Normes équivalentes

**Définition 3.** Soit $E$ un e.v. deux normes $n_1$ et $n_2$ sur E sont dites *équivalentes* si

$$\exists c, C > 0, \forall x \in E, \quad cn_1(x) \leq n_2(x) \leq Cn_1(x).$$

On note alors $n_1 \sim n_2$.

*Remarque* 4. L'équivalence des normes est une relation d'équivalence, c'est à dire qu'elle est réflexive ($n_1 \sim n_1$), symétrique ($n_1 \sim n_2 \Rightarrow n_2 \sim n_1$) et transitive ( $n_1 \sim n_2, n_2 \sim n_3 \Rightarrow n_1 \sim n_3$). Si deux normes sont équivalentes les notions d'analyses (limite, continuité, ...) sont les mêmes pour les deux normes.

*Exemple* 5. Dans $\mathbb{R}^n$, $||.||_1, ||.||_2, ||.||_\infty$ sont équivalentes (cf. TD). On verra plus tard qu'en dimension finie toutes les normes sont équivalentes.

# 4 Boules dans un e.v.n

**Définition 4.** Soient $a \in E$ et $r \in [0, \infty[$.

On appelle *boule ouverte de centre a et de rayon r* de $E$ la partie :

$$B(a,r) = \{x \in E, \mid ||x - a|| < r\}.$$

et *boule fermée de centre a et de rayon r* de $E$ la partie :

$$B_F(a,r) = \{x \in E, \mid ||x - a|| \leq r\}.$$

On appelle *sphère de centre a et de rayon r* de $E$ la partie :

$$S(a,r) = \{x \in E, \mid ||x - a|| = r\}.$$

Dans le cas $r = 0$, $B(x,0) = \emptyset, B_F(x,0) = \{x\}$.

*Exemple* 6. Dessinons la boule fermée de centre 0 et de rayon 1 pour la norme $N$ de l'exemple 1. Comme cette norme ne change pas si $x \to -x$,ou $y \to -y$, il suffit de considérer le cas $x \geq 0, y \geq 0$. Alors $N((x,y)) \leq 1$ équivaut à $x \leq 1$ et $\frac{1}{2}x + y \leq 1$. Donc

$$B_{FN}(0,1) \cap \mathbb{R}_+^2 = \{x \in [0,1], 0 \leq y \leq 1 - x/2\}.$$

Cf cours pour le dessin, c'est l'intersection de 4 demi plans. La boule complète est l'union de 4 telles intersections, au total c'est l'intersection de 6 demi-plans :

$$B_{FN}(0,1) = \{x \in [-1,1], -1 - x/2 \leq y \leq 1 - x/2, -1 + x/2 \leq y \leq 1 + x/2\}.$$

*Exercice* 2. Dessiner les boules de $\mathbb{R}^2$ pour $||.||_1, ||.||_2, ||.||_\infty$

## 4.1 Parties bornées

**Définition 5.** Un ensemble $A \subset E$ est dit *borné* si $\exists M \in [0, \infty[, \forall x \in A, ||x|| \leq M$, c'est à dire s'il est contenu dans une boule.

## 4.2 Convexité

Soit $x, y \in E$, on appelle segment d'extrémité x et y la partie

$$[x,y] = \{\lambda x + (1-\lambda)y, \lambda \in [0,1]\}.$$

On retrouve bien sûr la définition usuelle dans $\mathbb{R}$.

**Définition 6.** Un ensemble $C \subset E, C \neq \emptyset$ est dit *convexe* si $\forall x, y \in C, [x,y] \subset C$.

**Proposition 2.** *Les boules sont des convexes.*

*Démonstration.* Considérons le cas des boules ouvertes. Soient $x, y \in B(a,r)$, $z = \lambda x + (1-\lambda)y$, $\lambda \in [0,1]$.

Par l'inégalité triangulaire et homogénéité, on a :

$$||z - a|| = ||\lambda(x-a) + (1-\lambda)(y-a)|| \leq |\lambda| ||x-a|| + |1-\lambda| ||y-a|| < |\lambda| r + |1-\lambda| r = r.$$

Donc $z \in B(a,r)$. Le cas des boules fermées est similaire. □

# 5 Suites dans un e.v.n.

On rappelle qu'une suite de $E$ est une application $u : \mathbb{N} \to E$ notée $(u_n)_{n \geq 0}$.

## 5.1 Convergence

**Définition 7.** (Convergence) Soit $(u_n)$ une suite d'un evn $(E, ||.||)$. On dit que $u_n$ *converge vers* $l \in E$ (et on note $l = \lim_{n\to\infty} u_n$ ou $u_n \to_{n\to\infty} l$) si la suite numérique $||u_n - l||$ converge vers 0, c'est-à-dire :

$$\forall \epsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \quad ||u_n - l|| \leq \epsilon.$$

*Remarque* 5. Ceci équivaut à $\forall\epsilon > 0, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, \quad u_n \in B(l, \epsilon)$. Comme dans $\mathbb{R}$ on a unicité de la limite (justifiant la notation). En effet si on a deux limites $l_1, l_2$ pour $n$ grand $u_n \in B(l_1, \epsilon) \cap B(l_2, \epsilon)$ donc par inégalité triangulaire $||l_1 - l_2|| \leq ||l_1 - u_n|| + ||u_n - l_2|| \leq 2\epsilon$ Comme $\epsilon > 0$ arbitraire $||l_1 - l_2|| = 0$, soit par l'axiome de séparation $l_1 = l_2$.

**Proposition 3.** *(i) Si $u_n \to u$, alors $||u_n|| \to ||u||$ (réciproque fausse).*
*(ii) Toute suite convergente est bornée (réciproque fausse).*
*(iii) Si $u_n \to u, v_n \to v$ alors pour toute suite $\lambda_n \in \mathbb{K}$, tel que $\lambda_n \to \lambda$ on a $\lambda_n u_n + v_n \to \lambda u + v$.*

*Démonstration.* (i) Par l'inégalité triangulaire inverse $|||u_n||_{|}|u||| \leq ||u_n - l||$
(ii) Par (i) et le cas réel.
(iii) Vu $\lambda_n u_n + v_n - (\lambda u + v) = \lambda_n(u_n - u) + (v_n - v) + (\lambda_n - \lambda)u$, homogénéité et inégalité triangulaire implique :

$$||\lambda_n u_n + v_n - (\lambda u + v)|| \leq |\lambda_n|||u_n - u|| + ||v_n - v|| + |\lambda_n - \lambda|||u|| \to 0.$$

□

## 5.2 Suite extraite, valeur d'adhérence

**Définition 8.** Soit $(u_n)$ une suite de $E$ on appelle *suite extraite* ou *sous-suite* une suite de la forme $v_n = u_{\phi(n)}$, pour $\phi : \mathbb{N} \to \mathbb{N}$ une application strictement croissante

**Définition 9.** On appelle *valeur d'adhérence* d'une suite $(u_n)$ toute limite d'une suite extraite convergente.

**Proposition 4.** *Toute suite extraite d'une suite convergente converge vers la même limite. (Autrement dit, toute suite convergente n'a qu'une seule valeur d'adhérence, sa limite.)*

*Démonstration.* Supposons $u_n \to l$ et si $v_n$ une suite extraite, $||v_n - l||$ est extraite de $||u_n - l||$ ( le résultat est donc une conséquence du cas réel). □

## 5.3 Suite de Cauchy, Complétude (Preuves facultatives)

**Définition 10.** Une suite $(u_n)$ de $E$ est dite de Cauchy si :

$$\forall\epsilon > 0, \exists N \in \mathbb{N}, \forall (p, q) \in \mathbb{N}^2, p \geq N \text{ et } q \geq N \to ||u_p - u_q|| \leq \epsilon.$$

La proposition suivante est similaire au cas réel.

**Proposition 5.** *Toute suite convergente est de Cauchy. Toute suite de Cauchy est bornée. Toute suite de Cauchy possédant une valeur d'adhérence est convergente.*

**Définition 11.** Une partie $A$ de $E$ est dite *complète* si toute suite de Cauchy de $A$ converge dans $A$. Si E est complet on dit que c'est un *espace de Banach.*

On a vu en première année que $\mathbb{K}$ est complet (mais pas $\mathbb{Q}$). On verra que tout evn de dimension finie est complet. Dans le cadre de l'exemple 2, $(C^0([a, b], \mathbb{R}), ||.||_\infty)$ est aussi complet (i.e. un espace de Banach), ce n'est pas le cas de $(C^0([a, b], \mathbb{R}), ||.||_i)$, $i = 1, 2$.

**Proposition 6.** *Si $E, F$ sont des evn complets. Alors $E \times F$ (munie de la norme produit) est complet.*

*Démonstration.* Si $(u_n, v_n)$ est de Cauchy dans $E \times F$, de même, $(u_n)$ est de Cauchy dans $E$, et $(v_n)$ dans $F$, donc par complétude $(u_n)$ converge vers $u$ et $(v_n)$ vers $v$. En conséquence $(u_n, v_n)$ converge vers $(u, v)$ vu $||(u_n, v_n) - (u, v)|| = \max(||u_n - u||, ||v_n - v||) \to 0$. □

# 6 Ouverts dans un evn.

Soit $(E, ||.||)$ un evn.

**Définition 12.** Une partie $O \subset E$ est un *ouvert* (ou une partie ouverte) si

$$\forall x \in O, \exists r > 0, \quad B(x,r) \subset O.$$

## 6.1 Exemples d'ouverts et propriétés

$E, \emptyset$ sont des ouverts de $E$. $[a,b], [a,b[$ ne sont pas ouverts dans $\mathbb{R}$ mais $]a,b[$ l'est.

**Proposition 7.** *Les boules ouvertes sont ouvertes.*

On remarquera que le mot ouvert a deux sens dans "boules ouvertes" et "parties ouvertes" mais qu'ils sont cohérents grâce à la proposition (les boules fermées ne sont pas des ouverts, cf TD).

*Démonstration.* Soit $a \in E, r > 0$ montrons que $B(a,r)$ est un ouvert ($B(a,0)$ est vide donc ouvert). Soit $x \in B(a,r)$, $r - ||x-a|| > 0$, il suffit donc de montrer que :

$$B(x, r - ||x-a||) \subset B(a,r).$$

C'est une conséquence de l'inégalité triangulaire. En effet, si $y \in B(x, r-||x-a||)$, alors $||y-a|| \leq ||y-x|| + ||x-a|| < (r - ||x-a||) + ||x-a|| = r$, donc $y \in B(a,r)$. □

**Proposition 8.**
1. *La partie vide $\emptyset$ et $E$ sont des ouverts.*
2. *la réunion d'une famille d'ouverts est ouverte.*
3. *l'intersection d'une famille finie d'ouverts est ouverte.*

*Remarque* 6. On appelle *topologie* une famille de parties d'un ensemble, qui, comme la famille des ouverts d'un evn, vérifie ces trois propriétés. La famille des ouverts de $E$ est donc appelée topologie (normique) de $E$. $\cap_{n\in\mathbb{N}} B(a, 1/n) = \{a\}$ qui n'est pas ouvert dans $E$ montre que l'hypothèse "finie" est cruciale dans 3.

*Démonstration.*
1. évident.
2. Soit $(O_i)_{i\in I}$ une famille d'ouverts. On peut supposer I non vide (sinon l'union vide étant vide on est ramené à 1). Soit $x \in O = \cup_{i\in I} O_i$, donc il existe $j \in I$, $x \in 0_j$. Comme $O_j$ est ouvert il existe $r > 0$, $B(x,r) \subset O_j \subset O$. Donc $O$ est ouvert.
3. Soit $O_1, ..., O_n$ une famille finie d'ouverts. Soit $x \in O = O_1 \cap \cdots \cap O_n$. Comme $x \in O_i$, et $O_i$ ouvert, il existe $r_i > 0, B(x, r_i) \subset O_i$. Soit $r = \min_{i=1\dots n} r_i > 0$. On déduit de la définition que $B(x,r) \subset B(x,r_i) \subset O_i$ donc $B(x,r) \subset O$, ce qui montre que $O$ est ouvert.

□

*Exemple* 7. Soit $O = \{(x,y), x > 0\}$. Montrons que c'est un ouvert de $\mathbb{R}^2$ pour la norme $||.||_\infty$. En effet

$$O = \cup_{(x,y)\in O} ]0, 2x[\times]y-x, y+x[ = \cup_{(x,y)\in O} B_{||\cdot||_\infty}((x,y), x),$$

est ouvert comme union d'ouverts.

## 6.2 Intérieur

**Définition 13.** Soit $A \subset E$, on dit que x est *intérieur* à A (ou A est un voisinage de x) si $\exists r > 0, B(x,r) \subset A$.

On note $\mathrm{Int}(A)$ ou $\mathring{A}$ l'ensemble des points intérieurs à A.

**Proposition 9.** $\mathrm{Int}(A)$ *est le plus grand ouvert contenu dans A.*

*Démonstration.*

1. $\mathrm{Int}(A)$ **contient tous les ouverts inclus dans A.**
   Soit $U$ un ouvert contenu dans A. Soit $x \in U$, alors comme U est ouvert, $\exists r > 0, B(x,r) \subset U \subset A$, donc $x$ est intérieur à A. Ainsi $U \subset \mathrm{Int}(A)$
2. $\mathrm{Int}(A)$ **est un ouvert.** Soit $x \in \mathrm{Int}(A)$. Soit donc $r > 0$ tel que $B(x,r) \subset A$. Comme $B(x,r)$ est ouvert, tout $y \in B(x,r)$ est intérieur à $B(x,r)$ donc intérieur à $A$. En bilan, $\forall x \in \mathrm{Int}(A), \exists r > 0,\ B(x,r) \subset \mathrm{Int}(A)$, ce qui conclut.

□

**Corollaire 10.** *(exo, cf TD)*

1. *A ouvert si et seulement si* $A = \mathrm{Int}(A)$.
2. $A \subset B \Rightarrow \mathrm{Int}(A) \subset \mathrm{Int}(B)$
3. $\mathrm{Int}(A) \cup \mathrm{Int}(B) \subset \mathrm{Int}(A \cup B)$
4. $\mathrm{Int}(A) \cap \mathrm{Int}(B) = \mathrm{Int}(A \cap B)$

*Exemple* 8. Soit $F = \{(x,y), x \geq 0\}$. Montrons que $Int(F) = O := \{(x,y), x > 0\}$. On a vu à l'exemple 7 que $O$ est ouvert, donc comme $O \subset F$, on a $O \subset Int(F)$. Il reste à voir que $Int(F) \cap \{(x,y), x = 0\} = \emptyset$ (car alors $Int(F) \subset F - \{(x,y), x = 0\} = O$). Mais soit $(-\epsilon, y) \in B_{||.||_\infty}((0,y),\epsilon) \cap F^c$ pour tout $\epsilon > 0$, donc $B_{||.||_\infty}((0,y),\epsilon) \not\subset F$ donc $(0,y)$ n'est pas intérieur à $F$, ce qu'il fallait démontrer.

# 7 Fermés dans un evn.

Soit $(E, ||.||)$ un evn.p

*Rappel* 7. Soit $A \subset E$, on note $A^c = \{x \in E \mid x \notin A\}$ le complémentaire de A. On rappelle que $\emptyset^c = E, E^c = \emptyset, (A^c)^c = A$, $A \cup A^c = E, A \cap A^c = \emptyset$. Les lois de De Morgan impliquent que pour une famille $(A_i)_{i \in I}$

$$\left(\bigcup_{i \in I} A_i\right)^c = \bigcap_{i \in I} A_i^c,$$

$$\left(\bigcap_{i \in I} A_i\right)^c = \bigcup_{i \in I} A_i^c.$$

**Définition 14.** Soit $F \subset E$. On dit que $F$ est un *fermé* de $E$ si $F^c$ est un ouvert de $E$.

Le résultat suivant est obtenu en passant au complémentaire le résultat sur les ouverts.

**Proposition 11.**

1. *La partie vide* $\emptyset$ *et* $E$ *sont des fermés.*
2. *l'intersection d'une famille de fermés est fermée.*
3. *l'union d'une famille finie de fermés est fermée.*

**Proposition 12.** *(Caractérisation séquentielle des fermés) Une partie $F$ d'un evn $E$ est fermée si et seulement si toute suite convergente $(x_n)$ d'éléments de $F$ a sa limite dans $F$.*

*Démonstration.* Supposons $F$ fermé. Soit $(x_n)$ une suite d'éléments de $F$, convergente vers $x$. Soit $y \in F^c$, comme $F^c$ est ouvert il existe $\epsilon > 0$ $B(y, \epsilon) \subset F^c$, d'où $x_n \notin B(y, \epsilon)$ Donc $||x_n - y|| \geq \epsilon$. En passant à la limite on déduit

$$||x - y|| \geq |\ ||x_n - x|| - ||x_n - y||\ | \geq \epsilon - ||x_n - x|| \to_{n \to \infty} \epsilon > 0,$$

Donc $||x - y|| \geq \epsilon$ donc $x \neq y$. Comme $y$ était arbitraire dans $F^c$, $x \in F$.

Réciproquement, supposons que $F$ n'est pas fermé et montrons que la seconde caractérisation est fausse. Soit $x \in F^c$ montrant que $F^c$ n'est pas ouvert, donc pour tout $n \in \mathbb{N}$, $B(x, 1/n) \cap F \neq \emptyset$. Soit $x_n \in B(x, 1/n) \cap F$ $||x_n - x|| \leq 1/n \to_{n \to \infty} 0$, donc $(x_n)$ est une suite d'éléments de $F$ qui converge vers $x \in F^c$. □

*Exemple* 9. Montrons avec la caractérisation séquentielle que $A = \{(x, y), x > 0, y > 0\}$ n'est pas fermé pour la norme $||.||_\infty$. En effet $A \ni (1/n, 1/n) \to (0, 0) \notin A$, ce qui contredirait l'hypothèse que $A$ fermé. Montrons de même que $B = \{(x, y), x \geq 0, y \geq 0\}$ est fermé. En effet, Soit $(x_n, y_n) \in B$ tel que $(x_n, y_n) \to (x, y)$ on a $x_n \to x, y_n \to y$ donc comme $x_n \geq 0$, on déduit $x \geq 0$, et de même $y \geq 0$ donc $(x, y) \in B$ . Ainsi, comme toute limite de suite de $B$ est dans $B$, on déduit que $B$ est fermé.

**Proposition 13.** *(Relations Fermé-Complet)*

1. *Si $C \subset E$ est complet alors il est fermé.*
2. *Si $C \subset E$ est complet et $F \subset C$ est un fermé de $E$, alors $F$ est complet.*

*Démonstration.*

1. Si $C \subset E$ est complet alors si on considère une suite $(x_n)$ convergente vers $x$ dans $E$, elle est de Cauchy, donc converge dans $C$, donc $x \in C$ par unicité de la limite.
2. Si $C \subset E$ est complet et $F \subset C$. Soit $x_n$ une suite de Cauchy de $F$, elle converge dans $C$, donc comme $F$ est fermé, la limite est dans $F$, donc toute suite de Cauchy de $F$ converge dans $F$.

□

## 7.1 Adhérence

**Définition 15.** Soit $A \subset E$. Un point $x \in E$ est dit *adhérent* à A si $\forall \epsilon > 0 B(x, \epsilon) \cap A \neq \emptyset$.

On note $\bar{A}$ (ou $\mathrm{Adh}(A)$) l'ensemble des points adhérents à $A$.

*Exemple* 10. $\bar{E} = E, \bar{\emptyset} = \emptyset, A \subset \bar{A}$. Si $A = \{x_n\}_{n \in \mathbb{N}}$ les valeurs d'adhérence de la suite $(x_n)$ sont dans $\bar{A}$ qui est l'union de l'ensemble des valeurs d'adhérence et de $A$ (exo).

**Proposition 14.**

$$(\mathrm{Adh}(A))^c = \mathrm{Int}(A^c).$$
$$(\mathrm{Int}(B))^c = \mathrm{Adh}(B^c).$$

*Démonstration.* Un point $x \in E$ n'appartient pas à $\mathrm{Adh}(A)$ si et seulement si $\exists \epsilon > 0, B(x, \epsilon) \cap A = \emptyset \iff \exists \epsilon > 0, B(x, \epsilon) \subset A^c$. C'est par définition équivalent à dire que $x$ est un point adhérent à $A^c$.En appliquant le premier résultat à $A = B^c$, on en déduit le second. □

On en déduit toutes les propriétés en passant au complémentaire celles de l'intérieur.

**Corollaire 15.** *1. $\bar{A}$ est le plus petit fermé contenant $A$.*

2. *A fermé si et seulement si $A = \bar{A}$.*
3. $A \subset B \Rightarrow \bar{A} \subset \bar{B}$
4. $\bar{A} \cap \bar{B} \supset \overline{A \cap B}$
5. $\bar{A} \cup \bar{B} = \overline{A \cup B}$

*Démonstration.* 1. $\bar{A}$ est fermé vu que son complémentaire est l'ouvert $\mathrm{Int}(A^c)$. Si $F$ est un fermé contenant $A$, $F^c$ est un ouvert contenu dans $A^c$ donc dans $\mathrm{Int}(A^c)$ le plus grand ouvert contenant $A^c$. En passant au complémentaire, $F \supset \bar{A}$. Les résultats 2.3.4.5 sont analogues, par passage au complémentaire, de résultats sur l'intérieur. □

**Proposition 16.** *$x \in \bar{A}$ si et seulement si il existe une suite $(a_n)$ d'éléments de $A$ vérifiant $a_n \to x$.*

*Démonstration.* Si x est adhérent à $A$ pour tout entier $n$ $B(x, 1/n) \cap A$ est non vide donc contient un élément $a_n$. La suite $(a_n) \in A^{\mathbb{N}}$ converge vers $x$ vu $||a_n - x|| \leq 1/n \to 0$. La réciproque vient de la caractérisation séquentielle des fermés vu $\bar{A}$ fermé. □

*Exemple* 11. Montrons que si $A = \{(x, y), x > 0, y > 0\}$ alos $\overline{A} = B = \{(x, y), x \geq 0, y \geq 0\}$. On a vu à l'exemple 9 que $B$ est fermé, donc comme $A \subset B$, on en déduit $\overline{A} \subset B$

Il reste à montrer que $B - A = \{(x, y), x = 0, y \geq 0$ ou $y = 0, x \geq 0\} \subset \overline{A}$. Or $(0, y) = \lim_{n\to\infty}(1/n, y + 1/n)$ et si $y \geq 0$, $(1/n, y + 1/n) \in A$, donc $(0, y) \in \overline{A}$. De même $(x, 0) = \lim_{n\to\infty}(x + 1/n, 1/n) \in \overline{A}$ si $x \geq 0$.

## 7.2 Densité, Frontière

**Définition 16.** Une partie $A$ est dite *dense* dans $E$ si $\bar{A} = E$.

*Exemple* 12. $\mathbb{Q}$ et $\mathbb{Q}^c$ sont denses dans $\mathbb{R}$.

**Définition 17.** Un point $x \in E$ est dit point frontière d'une partie $A$ si pour tout $r > 0$, $B(x, r)$ est d'intersection non vide avec $A$ et $A^c$. On note $\mathrm{Fr}(A)$ l'ensemble des points frontières de A.

*Remarque* 8. D'après la définition, $\mathrm{Fr}(A) = \mathrm{Fr}(A^c) = \bar{A} \cap \overline{A^c}$ est un fermé.

*Exercice* 3. Montrer que $\mathrm{Int}(A^c), \mathrm{Fr}(A), \mathrm{Int}(A)$ forment une partition de $E$ (i.e. sont disjoints deux à deux et leur union est $E$).

# 8 Limites de fonctions

Soit $(E, ||.||_E)$ et $(F, ||.||_F)$ deux evn.

**Définition 18.** Soit $X \subset E$, $f : X \to F$ une application et $a \in \bar{X}$. On dit que $f$ *tend vers l'élément $l$ de $F$ en $a$* et on note $l = \lim_{x\to a} f(x)$ si :

$$\forall \epsilon > 0, \exists \eta = \eta(\epsilon, a) > 0, \forall x \in X, ||x - a||_E \leq \eta \Rightarrow ||f(x) - l||_F \leq \epsilon.$$

Ceci équivaut à :

$$\forall \epsilon > 0, \exists \eta = \eta(\epsilon, a) > 0, f(X \cap B(a, \eta)) \subset B(l, \epsilon).$$

*Remarque* 9. Comme $a \in \bar{X}$, $X \cap B(a, \eta) \neq \emptyset$. Par exemple, si $E = F = \mathbb{R}$ et $f : X = ]0, 1[ \to \mathbb{R}$ est tel que $f(x) = sin(x)/x$, $0 \in \bar{X}$ et $\lim_{x\to 0} f(x) = 1$.

**Proposition 17.** *(Unicité de la limite) Soient $f : X \to F$, $a \in \bar{X}$. L'application $f$ tend vers au plus un élément de $F$ en $a$.*

*Démonstration.* Supposons $f(x) \to_{x\to a} l_1$ et $f(x) \to_{x\to a} l_2$. Par définition pour tout $\epsilon > 0$ il existe $\eta > 0$ tel que $f(X \cap B(a,\eta)) \subset B(l_1,\epsilon) \cap B(l_2,\epsilon)$. Donc vu $X \cap B(a,\eta)$ non vide $\exists y \in B(l_1,\epsilon) \cap B(l_2,\epsilon)$ donc $||l_1 - l_2|| \leq 2\epsilon$. Comme $\epsilon$ est arbitraire, $||l_1 - l_2|| = 0$ donc $l_1 = l_2$. $\square$

*Exemple* 13. Soit $E = \mathbb{R}^2$, $X = \mathbb{R}^2 - (0,0)$, $a = (0,0) \in \bar{X}$, $F = \mathbb{R}$. Soit $f : X \to \mathbb{R}$ définie par $f(x,y) = \frac{x^5+y^4}{x^2+y^2}$. Montrons que $\lim_{(x,y)\to(0,0)} f(x,y) = 0$.

$$|f(x,y)| \leq \frac{x^2|x^3| + y^4}{x^2+y^2} \leq \frac{(x^2+y^2)|x^3| + (x^2+y^2)y^2}{x^2+y^2} = |x^3| + y^2 \to_{(x,y)\to(0,0)} 0.$$

**Proposition 18.** *(Caractérisation séquentielle des limites) Soient $f : X \to F$, $a \in \bar{X}$. L'application $f$ tend vers $l \in F$ en $a$ si et seulement si pour toute suite $(x_n)$ d'éléments de $X$ : si $x_n$ converge vers $a$, alors $f(x_n)$ converge vers $l$.*

*Démonstration.* Supposons que $f$ tend vers $l$ en $a$. Soit $\epsilon > 0$ il existe $\eta > 0$ tel que $f(X \cap B(a,\eta)) \subset B(l,\epsilon)$. Vu que $x_n \to a$ il existe $N$, tel que $\forall n \geq N, ||x_n - a|| \leq \eta$ donc $\forall n \geq N, ||f(x_n) - l|| \leq \epsilon$. Ceci indique que $f(x_n) \to l$.

Réciproquement, supposons par contraposition, qu'il existe $\epsilon > 0$ tel que pour tout $\eta > 0$ $f(X \cap B(a,\eta)) \cap B(l,\epsilon)^c \neq \emptyset$. Donc, en prenant, $\eta = 1/n$, on obtient $x_n \in X \cap B(a,1/n)$, tel que $||f(x_n) - l|| \geq \epsilon$. pour tout $n$, donc $x_n \to a$ et $f(x_n)$ ne converge pas vers $l$ comme voulu. $\square$

Le résultat suivant est une conséquence immédiate :

**Corollaire 19.** *Soient $f, g : X \to F$, $a \in \bar{X}$, $\lambda : X \to \mathbb{K}$.*

1. *Si $f$ tend vers $l$ en $a$, $l \in \overline{f(A)}$.*
2. *Si $f$ tend vers $l = \lim_{x\to a} f(x)$ en $a$, $g$ tend vers $m = \lim_{x\to a} g(x)$ en $a$ et $\lambda$ tend vers $\mu = \lim_{x\to a} \lambda(x) \in \mathbb{K}$ en $a$, alors $\lambda f + g : x \mapsto \lambda(x)f(x) + g(x)$ tend vers $\mu\, l + m$ en $a$.*
3. *Si de plus $\mu \neq 0$ $f/\lambda$ tend vers $l/\mu = \lim_{x\to a} f(x)/\lambda(x)$.*

**Théorème 20.** *(Composition des limites) Soit $E, F, G$ 3 evn, $X \subset E$, $Y \subset F$.*

*Soit $f : X \to Y$ et $a \in \bar{X}, b \in \bar{Y}$. Soit $g : Y \to G$ de sorte que $g \circ f : X \to G$ est bien définie par $g \circ f(x) = f(g(x))$. On suppose que $b = \lim_{x\to a} f(x), c = \lim_{x\to b} g(x)$ alors :*

$$c = \lim_{x\to a} g \circ f(x).$$

*Démonstration.* Soit $(x_n)$ une suite d'éléments de $X$ convergeant vers $a$. Par hypothèse $(y_n) = (f(x_n))$ (et caractérisation séquentielle des limites) est une suite d'éléments de $Y$ convergeant vers $b$. Donc de même $g(y_n)$ converge vers $c$ par hypothèse. En conclusion, $g \circ f(x_n) = g(y_n)$ converge vers $c$. Par caractérisation séquentielle, on déduit le résultat. $\square$

*Exemple* 14. Soit $E = \mathbb{R}^2, X = \mathbb{R}^2 - \{(0,0)\}$, $f(x,y) = \frac{xy}{x^2+y^2}$. Étudions la limite de de $f(x,y)$ quand $(x,y) \to (0,0)$. Cette limite n'existe pas. Raisonnons par l'absurde et supposons que la limite est $l$. Supposons que $f(x,y) \to \ell$. Soient $\phi_i : \mathbb{R} - \{0\} \to \mathbb{R}^2 - \{(0,0)\}$ définies par $\phi_1(t) = (t,t), \phi_2(t) = (t,2t)$. Dans les deux cas on a $\lim_{t\to 0} \phi_i(t) = (0,0)$. Par le théorème de composition des limites, on devrait avoir $\lim_{t\to 0} f \circ \phi_i(t) = l$. Or $f \circ \phi_1(t) = \frac{t^2}{t^2+t^2} = 1/2 \to_{t\to 0} 1/2$ et $f \circ \phi_2(t) = \frac{2t^2}{t^2+4t^2} = 2/5 \to_{t\to 0} 2/5$. Contradiction.

**Définition 19.** (Autres limites) Soit $f : X \to F$, $a \in \bar{X}$. On dit que $\lim_{x\to a} f(x) = +\infty$ si

$$\forall A \in \mathbb{R}_+, \exists \eta > 0 : \forall x \in X,\ ||x-a|| \leq \eta \Rightarrow ||f(x)|| \geq A.$$

Si $X$ n'est pas borné et $l \in F$, on dit que $\lim_{x\to\infty} f(x) = l$ si :

$$\forall \epsilon > 0, \exists B > 0 : \forall x \in X,\ ||x|| \geq B \Rightarrow ||f(x) - l|| \leq \epsilon.$$

Si $X$ n'est pas borné, on dit que $\lim_{x\to\infty} f(x) = \infty$ si :

$$\forall A \in \mathbb{R}_+, \exists B > 0 : \forall x \in X,\ ||x|| \geq B \Rightarrow ||f(x)|| \geq A.$$

On a les résultats habituels relatifs à ces notions.

# 9 Fonctions continues

## 9.1 Définitions équivalentes

**Définition 20.** Soient $E, F$ deux evn. $A \subset E$ et $f : A \to F$.

1. Soit $a \in A$, f est dit *continue en a* si $\lim_{x\to a} f(x) = f(a)$.
2. $f$ est continue sur $A$ si $f$ est continue en tout point de $A$.

**Proposition 21.** *Soit $f : A \to F$, $A \subset E$. Les assertions suivantes sont équivalentes :*

1. *$f$ est continue sur $A$.*
2. *Pour tout ouvert $O$ de $F$, il existe un ouvert $V$ de $E$ tel que l'image inverse $f^{-1}(O) = A \cap V$.*
3. *Pour tout fermé $C$ de $F$, il existe un fermé $D$ de $E$ tel que l'image inverse $f^{-1}(C) = A \cap D$.*

*Démonstration.* 2. $\Longleftrightarrow$ 3. vient de $(f^{-1}(B))^c = (f^{-1}(B^c))$ et de la relation fermés/ouverts.

1. $\Rightarrow$ 2. Soit $O$ un ouvert de $F$ et $x \in O$, il existe et on choisit $\epsilon(x) > 0$ tel que $B(x, \epsilon(x)) \subset O$. Par continuité de $f$, soit $y \in f^{-1}(O)$, $f(y) = x \in O$, il existe $\eta(y) > 0$ tel que $f(A \cap B(y, \eta(y))) \subset B(x, \epsilon(f(y))) \subset O$. Posons $V = \cup_{y \in f^{-1}(O)} B(y, \eta(y))$, c'est un ouvert de $E$ comme union d'ouverts. Évidemment, si $y \in f^{-1}(O)$, on a $y \in B(y, \eta(y)) \subset V$, d'où $f^{-1}(O) \subset V$ (on sait aussi $f^{-1}(O) \subset A$). De plus si $x \in V \cap A$, il existe $y \in f^{-1}(O)$ tel que $x \in B(y, \eta(y)) \cap A \subset f^{-1}(O)$ par définition de $\epsilon(y)$. Donc $V \cap A = f^{-1}(O)$.

1. $\Rightarrow$ 3. On prend $D = \overline{f^{-1}(C)}$ qui est fermé. On a $D \cap A \supset f^{-1}(C)$, il reste à montrer la réciproque. Soit $x \in D \cap A$, donc il existe $x_n \in f^{-1}(C)$ avec $x_n \to x$ donc $f(x_n) \in C$ et comme $x \in A$ et f continue en x, $f(x_n) \to f(x)$ et comme $C$ fermé, $f(x) \in C$ donc $x \in f^{-1}(C)$ et donc $f^{-1}(C) = A \cap D$.

2. $\Rightarrow$ 1. Soit $a \in A$. Montrons que $\lim_{x\to a} f(x) = f(a)$. Soit $\epsilon > 0$. Par 1. $f^{-1}(B(f(a), \epsilon)) = V \cap A$ pour un ouvert $V$ de $E$. Or $a \in V \cap A$ donc $\exists \eta > 0$ tel que $B(a, \eta) \subset V$. En conséquence

$$f(B(a,\eta) \cap A) \subset f(V \cap A) = f(f^{-1}(B(f(a), \epsilon))) \subset B(f(a), \epsilon),$$

ce qui conclut.

$\square$

*Exemple* 15. 1. $f : E \to \mathbb{R}$ définit par $f(x) = ||x||$ est continue sur $E$ car $|f(x) - f(x_0)| \leq ||x - x_0||$ (inégalité triangulaire inverse).

2. Soit $0 \leq p \leq n = r + s$, $p : R^n \to \mathbb{R}^s$ définie par si $x = (y, z) \in \mathbb{R}^n = \mathbb{R}^r \times \mathbb{R}^s$, $p(x) = z$. On munit $\mathbb{R}^n$ et $\mathbb{R}^s$ des normes $||.||_1$, on voit $||p(x)||_1 \leq ||x||_1$, donc comme $p$ est linéaire, $p$ est continue car $||p(x) - p(y)||_1 = ||p(x - y)||_1 \leq ||x - y||_1$.

*Remarque* 10. Il résulte des théorèmes sur les limites que les opérations algébriques usuelles (somme, produit, composition) préservent la continuité. En particulier si $P$ est une fonction polynomiale $P : \mathbb{R}^n \to \mathbb{R}$ c'est à dire de la forme $P(x) = \sum_{\text{finie}} a_{i_1,\dots,i_n} x_1^{i_1} \dots x_n^{i_n}$ est continue comme somme et produits des projections $(x_1, ..., x_n) \mapsto x_i$.

## 9.2 Homéomorphismes, Continuité uniforme, Applications linéaires continues

On cosidère $(E, ||.||)$ et $(F, ||.||)$ deux evn.

**Définition 21.** Soit $A \subset E$, et $B \subset F$ une application $f : A \to B$ est dite un *homéomorphisme* (ou une application bicontinue) si elle est bijective et si $f : A \to B \subset F$ et $f^{-1} : B \to A \subset E$ sont continues.

**Définition 22.** Soit $A \subset E$, une application $f : A \to F$ est *uniformément continue* si :

$$\forall \epsilon > 0, \exists \eta > 0 : \forall (x, y) \in A^2, ||x - y|| \leq \eta \Rightarrow ||f(x) - f(y)|| \leq \epsilon.$$

Une application $f : A \to F$ est *K-lipschitzienne* avec $K \in [0, +\infty[$ si :

$$\forall (x, y) \in A^2, ||f(x) - f(y)|| \leq K||x - y||.$$

**Théorème 22.** *Un application K-lipschitzienne est uniformément continue.*

*Démonstration.* Pour $\epsilon > 0$ dans la définition il suffit de prendre $\eta = \epsilon/K$. □

*Exemple* 16. $f : \mathbb{R}_+ \to \mathbb{R}$ $f(x) = \sqrt{x}$ est uniformément continue mais pas lipschitzienne (cf TD.). Toute application uniformément continue est continue mais la réciproque est fausse : $g : \mathbb{R} \to \mathbb{R}$ $g(x) = x^2$ n'est pas uniformément continue (cf TD.).

$x \mapsto ||x||$ est 1-lipschitzienne $E \to \mathbb{R}$, $(x, y) \mapsto x + y$ est 2-lipschitzienne $E \times E \to E$.

*Rappel* 11. Une application $u : E \to F$ est dite *linéaire* si :

(i) $\forall x, y \in E, u(x + y) = u(x) + u(y)$

(ii) $\forall x \in E, \lambda \in \mathbb{K}, u(\lambda x) = \lambda u(x)$.

**Proposition 23.** *Si $u : E \to F$ est une application linéaire, les assertions suivantes sont équivalentes :*

1. *$u$ est lipschitzienne.*
2. *$u$ est continue.*
3. *$u$ est continue en 0.*
4. *$u$ est continue en un point.*
5. *Il existe $a \in E$,$\eta > 0$ tel que $u(B(a, \eta)) \subset B(u(a), 1)$.*

*Démonstration.* (Preuve facultative) 1. $\Rightarrow$ 2., 2. $\Rightarrow$ 3.,3. $\Rightarrow$ 4.,4. $\Rightarrow$ 5. sont évidentes. Si on suppose 5., il existe $\eta > 0$ tel que si $||x - a|| \leq \eta$ alors $||u(x) - u(a)|| \leq 1$. Soit $h \in E$, $x = a + h\eta/||h||$ de sorte que $||x - a|| \leq \eta$, on déduit donc $||u(h)||\eta/||h|| = ||u(x - a)|| \leq 1$ c'est-à-dire $||u(h)|| \leq ||h||/\eta$ et donc pour tout $x, y$ en utilisant encore la linéarité $u(x - y) = u(x) - u(y)$, on obtient :

$$||u(x) - u(y)|| \leq \frac{1}{\eta}||x - y||,$$

donc $u$ est $1/\eta$ lipschitzienne. □

**Proposition 24.** *Si $\phi : E \to \mathbb{K}$ est une application linéaire (forme linéaire), $\phi$ est continue si et seulement si son noyau $H = \operatorname{Ker} \phi = \phi^{-1}(\{0\})$ est fermé.*

*Démonstration.* (Preuve facultative) Si $\phi$ est continue, $\phi^{-1}(\{0\})$ est fermé comme image inverse d'un singleton, qui est fermé. Réciproquement, supposons $\phi$ non nulle, soit $e$ tel que $\phi(e) = 1$. Comme le complémentaire de $H$ est ouvert soit $r > 0$ tel que $B(e, r) \subset H^c$.

Montrons par l'absurde que pour tout $x \in B(e, r)$, $\phi(x) \in B(1, 1)$. En effet, sinon soit $x$ avec $|\phi(x) - 1| \geq 1$. Si $t = -\phi(x)/(1 - \phi(x))$, on $\phi(te + (1-t)x) = t1 + (1-t)\phi(x) = t(1 - \phi(x)) + \phi(x) = 0$. Or $||te + (1 - t)x - e|| = |1 - t|||x - e|| = ||x - e||/|\phi(x) - 1| \leq r$ une contradiction car alors $y = te + (1 - t)x \in B(e, r) \cap H$.

On a donc vu $\phi(B(e, r)) \subset B(\phi(e), 1)$ d'où $\phi$ continue par la proposition précédente.

□

# 10 Propriétés particulières des evn de dimension finie.

## 10.1 Complétude (Preuves facultatives)

**Théorème 25.** *Tout evn de dimension finie est complet.*

*Démonstration.* C'est évident en dimension 1. On montre donc le résultat par récurrence sur la dimension. On suppose donc le résultat acquis en dimension strictement inférieure à $n$, soit $(E, ||.||)$ de dimension $n$. Soi $\phi$ une forme linéaire non nulle sur $E$, son noyau $F$ est de dimension $(n - 1)$, donc par hypothèse de récurrence $(F, ||.||)$ (muni de la restriction de la norme de $E$) est complet. Par conséquent $F$ est fermé dans $E$, donc $\phi$ est continue.

Soit $e \in E$ avec $\phi(e) = 1$. L'isomorphisme linéaire $u : (\lambda, f) \to \lambda e + f$ de $\mathbb{K} \times F$ (avec la norme produit donc complet par la proposition 6) sur $E$ est continue (2-lipschitzien). Son isomorphisme réciproque est donné par :

$$\forall x \in E, \quad u^{-1}(x) = (\phi(x), x - \phi(x)e).$$

$u^{-1}$ est donc aussi continue comme $\phi$. $u$ étant lipschitzienne (car linéaire continue), si $(x_n)$, suite de $E$, est de Cauchy $u(x_n)$ l'est aussi donc converge par complétude de $\mathbb{K} \times F$, d'où $x_n = u^{-1}(u(x_n))$ converge aussi par continuité de $u^{-1}$. □

## 10.2 Applications linéaires (Preuves facultatives)

*Rappel* 12. Si $E$ de dimension $n$ et $F$ de dimension $p$. Soit $(e_1, ..., e_n)$ une base de $E$, $(f_1, ..., f_p)$ une base de $F$. Une application linéaire $u$ est décrite par sa matrice $A = (a_{ij})_{i \in [1,p], j \in [1,n]}$ dans ces bases.

Alors, si $x = \sum_{j=1}^{n} x_j e_j$ et $y = u(x) = \sum_{i=1}^{p} y_i f_i$, on rappelle que :

$$y_i = \sum_{j=1}^{n} a_{ij} x_j.$$

On définit aussi la base duale $(e_1^*, ..., e_n^*)$ de l'ev des formes linéaires sur $E$ caractérisés par $e_j^*(e_k) = 1$ si $j = k$ et 0 sinon. En conséquence, pour tout $x \in E$ :

$$u(x) = \sum_{j=1}^{n} x_j u(e_j) = \sum_{j=1}^{n} e_j^*(x) u(e_j).$$

**Théorème 26.** *Toute application linéaire entre evn de dimensions finies est continue (et même lipschitzienne).*

*Démonstration.* En utilisant la représentation du rappel

$$u = \sum_{i=1}^{n} u(e_i) e_i^*,$$

il suffit de montrer que les formes linéaires $e_i^*$ sont continues. Mais $\operatorname{Ker} e_i^*$ est un sous-espace vectoriel de dimension fini donc complet (Théorème 25), donc fermé (proposition 13) dans $E$, d'où la continuité voulue (proposition 24). La lipschitzianité vient de la proposition 23. □

## 10.3 Équivalence des normes et conséquences.

**Théorème 27.** *Toutes les normes d'un espace vectoriel normé de dimension finie sont équivalentes.*

*Démonstration.* Si $||.||_1$, et $||.||_2$ sont deux normes sur $E$. l'application linéaire identité $u = Id_E$ vu de $(E, ||.||_1)$ vers $(E, ||.||_2)$ est continue ainsi que son inverse $u^{-1}$ (théorème 26), donc elles sont $C$ et $1/c$-lipschitzienne respectivement (proposition 23). On en déduit, pour tout $x \in E$ :

$$||x||_2 = ||u(x) - u(0)||_2 \leq C||x||_1,$$
$$||x||_1 = ||u^{-1}(x) - u^{-1}(0)||_1 \leq \frac{1}{c}||x||_2,$$

d'où l'équivalence des normes souhaitée. □

*Remarque* 13. Sur $R^n$ on peut donc parler de continuité, limite etc. sans préciser la norme.

**Proposition 28.** *Soient $E$ un evn, $A \subset E$, $f : A \to I\!R^n$. Si $x \in A$, on note $f(x) = (f_1(x), ..., f_n(x))$ où les $f_i$ sont les fonctions composantes de $f$ : $f_i : A \to I\!R$.*

*Soit $x \in \bar{A}$ et $b = (b_1, ..., b_n) \in I\!R^n$, alors on a l'équivalence :*

$$\lim_{x \to a} f(x) = b \Longleftrightarrow \forall i = 1...n \lim_{x \to a} f_i(x) = b_i.$$

*Démonstration.* On a $f_i = p_i \circ f$, où $p_i$ est i-ème projection $p_i : \mathbb{R}^n \to \mathbb{R}$ définie par $p_i(x_1, ..., x_n) = x_i$. $p_i$ est continue d'après l'exemple 8.2.

Si $\lim_{x \to a} f(x) = b$, on déduit $\lim_{x \to a} f_i(x) = b_i$ d'après le Théorème de composition des limites.

Réciproquement, on munit $\mathbb{R}^n$ de la norme $||.||_\infty$. Si pour tout $i$ $\lim_{x \to a} f_i(x) = b_i$ on a donc pour $\epsilon > 0$, l'existence de $\eta_i > 0$ tel que si $||x - a|| \leq \eta_i$, $||f_i(x) - b_i|| \leq \epsilon$. On pose $\eta = \min_{i=1...n}(\eta_i) > 0$. Donc si $||x - a|| \leq \eta$, pour tout $i$ $||f_i(x) - b_i|| \leq \epsilon$ donc $||f(x) - b||_\infty = \max ||f_i(x) - b_i|| \leq \epsilon$.

□

**Corollaire 29.** *Soient $E$ un evn, $A \subset E$, $f : A \to \mathbb{R}^n$. Si $x \in A$, on note $f(x) = (f_1(x), ..., f_n(x))$ où les $f_i$ sont les fonctions composantes de $f$ : $f_i : A \to \mathbb{R}$. $f$ est continue sur $A$ (resp. en $a \in A$) si et seulement si les $f_i$ sont continues sur $A$ (en resp. $a \in A$).*

La preuve du résultat suivant est semblable et omise.

**Proposition 30.** *Soit $X_n = (x_n^{(1)}, ..., x_n^{(p)})$ une suite de $\mathbb{R}^p$ et soit $L = (\ell_1, ..., \ell_p)$. Alors $X_n$ converge vers $L$ si et seulement si pour tout $i = 1...p$ $x_n^{(i)} \to \ell_i$.*

**Proposition 31.** *Soient $A \subset \mathbb{R}^n$, $p_i : \mathbb{R}^n \to \mathbb{R}$ la i-ème projection définie par $p_i(x_1, ..., x_n) = x_i$. Alors $A$ est bornée dans $\mathbb{R}^n$ si et seulement si pour tout $i$, $p_i(A)$ est bornée dans $\mathbb{R}$.*

*Démonstration.* □

# 11 Compacité dans les evn (Preuves facultatives)

**Définition 23.** Soit $K \subset E$, E evn. $K$ est dite *compacte* si elle possède la propriété suivante (dite de Bolzano-Weierstrass) : De toute suite de $K$, on peut extraire une suite convergente dans $K$.

*Rappel* 14. Dans $\mathbb{R}$ le théorème de Bolzano-Weierstrass indique que toute suite bornée admet une sous-suite convergente et donc que tout fermé borné est compact.

**Proposition 32.** *Un compact est un fermé borné. Un sous-ensemble fermé d'un compact est compact.*

*Démonstration.*

1. Un compact $K$ est fermé, car si une suite $(u_n)$ converge vers $l$ dans $E$, elle admet une sous-suite convergeant vers $k \in K$, dont la limite est nécessairement $l = k$ (proposition 4), donc $l \in K$.
2. On montre par contraposée qu'un ensemble non borné $A$ ne peut pas être compact. Si $A$ non-borné, soit $x_n \in A$ tel que $||x_n|| \geq n$, si une suite extraite $x_{\phi(n)} \to x$ convergeait, elle serait bornée, ce qui n'est pas le cas car $||x_{\phi(n)}|| \geq \phi(n) \to_{n\to\infty} \infty$.
3. Si $F \subset K$ avec $K$ compact, $F$ fermé, une suite de $F$ admet une sous suite convergeant dans $K$ par compacité, donc sa limite est dans $F$ par fermeture, d'où $F$ compacte.

□

*Exemple* 17. Soit $F = \{(x, y) \in \mathbb{R}^2 xy = 1\}$ est fermé mais pas compact. En effet, si $f(x, y) = xy$ est polynomiale donc continue $\mathbb{R}^2 \to \mathbb{R}$ donc $F = f^{-1}(\{1\})$ est fermé comme image réciproque d'un fermé par une application continue. Mais $F$ n'est pas compact car pas borné. $x_n = (1/n, n) \in F$ et $||x_n||_\infty = n \to \infty$.

*Remarque* 15. En général dans un evn un fermé borné n'est pas toujours compact. Dans $C^0([0, 1], \mathbb{R})$, $f_n(x) = x^n$ vérifie $||f_n||_\infty = 1$, mais comme $f_n(x) \to f(x)$ (on dit converge simplement vers f) avec $f(x) = 0$ si $x < 1$, $f(1) = 1$, donc $f$ non continue. Toute suite extraite de $f$ devrait converger vers cette limite qui n'est pas continue, donc elle ne peut pas converger dans $C^0([0, 1], \mathbb{R})$ vers cette limite qui n'est pas dans $C^0([0, 1], \mathbb{R})$. En général, on peut montrer que les boules fermées d'evn sont compactes si et seulement si l'evn est de dimension finie, on montre une implication ci-dessous.

**Théorème 33.** *Si $u : E \to F$ est continue et $K \subset E$ est compacte alors $u(K)$ est compacte.*

*Démonstration.* Soit $y_n$ une suite de $u(K)$ donc $y_n = u(x_n)$ ,avec $(x_n)$ suite de $K$, on extrait donc une suite $x_{\phi(n)}$ convergeant vers $x \in K$. Par continuité, la suite extraite $y_{\phi(n)} = u(x_{\phi(n)}) \to u(x) \in u(K)$. □

**Corollaire 34.** *$E$ evn. Si $K \subset E$ est compacte et $f : K \to I\!R$ est continue, alors la fonction $f$ est bornée et atteint ses bornes : $\exists x_0, x_1 \in K, \forall x \in K f(x_0) \leq f(x) \leq f(x_1)$.*

*Démonstration.* $f(K)$ est compacte donc fermée et bornée. Donc $f$ est bornée, et le $f(K)$ contient son *sup* et son *inf* (par fermeture) c'est-à-dire, il existe $y_0, y_1 \in f(K)$ $y_0 = \inf_{x\in K} f(x)$, $y_1 = \sup_{x\in K} f(x)$. Finalement $y_i = f(x_i)$ avec $x_i \in K$. □

**Théorème 35.** *Dans un evn de dimension finie, les fermés bornés sont compacts.*

*Démonstration.* D'après le théorème 26 un isomorphisme linéaire $u$ de $E$ sur $\mathbb{K}^n$ est continu de $(E, ||.||)$ sur $(K^n, ||.||_\infty)$, et $u^{-1}$ également. $u(K)$ est fermé comme image réciproque d'un fermé par $u^{-1}$ continue, $u(K)$ est borné comme image d'un borné par une application lipschitzienne. Donc $L = u(K)$ est un fermé borné de $(K^n, ||.||_\infty)$. Il suffit de voir que c'est un compact, car alors $K = u^{-1}(L)$ est compact comme image continue d'un compact (theorème 33). Soit $(x_p) = (x_p^{(1)}, ..., x_p^{(n)})$ une suite de $L$, par définition de la norme $(x_p^{(i)})$ sont bornés, elles admettent donc, par le théorème de Bolzano-Weierstrass dans $\mathbb{K}$, une sous-suite simultanément convergente. $x_{\phi(p)}^{(i)} \to x^{(i)}$ Donc si $x = (x^{(1)}, ..., x^{(n)})$, on a $||x_{\phi(p)} - x|| = \max_{i=1...n} |x_{\phi(p)}^{(i)} - x^{(i)}| \to 0$ et comme $L$ est fermé ; $x \in L$ ce qui conclut. □

*Exemple* 18. Soit $K = \{(x, y) \in \mathbb{R}^2 x^2 + y^2/2 = 1\}$ est compact. En effet, si $f(x, y) = x^2 + y^2/2$ est polynomiale donc continue $\mathbb{R}^2 \to \mathbb{R}$ donc $F = f^{-1}(\{1\})$ est fermé comme image réciproque d'un fermé par une application continue. De plus $K \subset B_{||.||_\infty, F}(0, \sqrt{2})$ donc $K$ est borné, donc fermé borné dans $\mathbb{R}^2$ de dimendion finie, donc $K$ est compact.

*Exemple* 19. Soit $g : K \to \mathbb{R}$ définie par $g(x, y) = x^2 + y^2$ $g$ est continue donc atteint ses bornes sur $K$ compact. En effet $g$ est la distance euclidienne à l'origine, il est facile de voir qu'elle atteint son maximum 2 en $(0, \pm\sqrt{2})$ sur $K$ et son minimum 1 en $(\pm 1, 0)$ sur $K$. Le théorème des extremas liés permettra de retrouver ce résultat pour des $g$ et des $K$ plus généraux.

**Théorème 36.** *(de Heine) Toute fonction continue $f$ sur un compact $K \subset E$ est uniformément continue.*

*Démonstration.* Soit $g : (x, y) \to ||f(x) - f(y)||$ de $K^2$ dans $\mathbb{R}$ elle est continue (pour la norme produit sur $E^2$ par composition) donc $g(K^2)$ est compact. Soit $\epsilon > 0$ reste à trouver un $\eta$ de continuité uniforme.

$$A = \{(x, y) \in K^2 \mid ||f(x) - f(y)|| \geq \epsilon\} = g^{-1}([\epsilon, +\infty[)$$

est fermé dans $K^2$ donc compact. Donc l'application continue $(x, y) \mapsto ||x - y||$ atteint sa borne inférieure $m$. On a $m \neq 0$ car sinon on aurait un $(x, x) \in A$, ce qui n'est pas possible vu $\epsilon > 0$.

Finalement si $\eta > 0$ est tel que $\eta < m$, si $||x - y|| \leq \eta$, on a $(x, y) \notin A$, donc $||f(x) - f(y)|| < \epsilon$. □

# Chapitre 2

# Calcul Différentiel, Extrema libres

Dans toute la suite du cours $E = \mathbb{R}^m$ est muni d'une des normes $||.||_1$,$||.||_2$,$||.||_\infty$ (que l'on note simplement $||.||$ si aucune précision n'est nécessaire). On rappelle qu'en dimension finie, toutes les normes sont équivalentes.

## 1 Dérivées des fonctions vectorielles

**Définition 24.** Soit $I$ un intervalle de $\mathbb{R}$ et $X : I \to \mathbb{R}^n$. Soit $t_0 \in I$. On dit que $X$ est dérivable en $t_0$ si $\lim_{h\to 0} \frac{X(t_0+h)-X(t_0)}{h}$ existe. Dans ce cas, elle est notée $X'(t_0)$ ou $\frac{dX}{dt}(t_0)$.

*Remarque* 16.

1. Si $I = [a, b]$ et $t_0 = a$, dans la limite précédente on a $h \geq 0$ (limite à droite). Si $t_0 = b$, on a $h \leq 0$ (limite à gauche).
2. On a $X(t) = (x_1(t), ..., x_n(t))$ avec les fonctions composantes $x_i : I \to \mathbb{R}$. Alors $X$ est dérivable en $t_0$ si et seulement si toutes les fonctions $x_i$ sont dérivables en $t_0$.

   Dans ce cas, on a $X'(t_0) = (x_1'(t_0), ..., x_n'(t_0))$. Autrement dit *on peut dériver composante par composante.*
3. Si $X$ est dérivable en $t_0$ alors $X$ est continue en $t_0$ (car chacune des fonctions composantes sont dérivables en $t_0$ donc continues en $t_0$).
4. ***Interprétation géométrique*** Si $X : I \to \mathbb{R}^2$, $X$ est une courbe. $\frac{X(t_0+h)-X(t_0)}{h}$ est un vecteur porté par la droite $(AB)$ reliant $A = X(t_0)$ et $B_h = X(t_0 + h)$. Si $X$ est dérivable en $t_0$ et $X'(t_0) \neq 0$ (point régulier), la droite $(AB_h)$ tend pour $h \to 0$ vers la tangente $T$ à la courbe en $X(t_0)$. On dit que $X'(t_0)$ est *tangent* à la courbe en $X(t_0)$.

*Exemple* 20. *Dérivée d'un vecteur unitaire*

On munit $\mathbb{R}^n$ de la norme $||.||_2$. On suppose que $X : I \to \mathbb{R}$. est dérivable et que pour tout $t \in I$, $||X(t)||_2 = 1$. Donc si $X(t) = (x_1(t), ..., x_n(t))$, on a :

$$||X(t)||_2^2 = \sum_{i=1}^{n} x_i(t)^2 = 1.$$

En dérivant, on obtient :

$$\sum_{i=1}^{n} 2x_i(t)x_i'(t) = 0.$$

On rappelle que si $U = (u_1, ..., u_n), V = (v_1, ..., v_n)$ alors le produit scalaire est défini par :

$$\langle U, V\rangle = \sum_{i=1}^{n} u_i v_i.$$

On vient donc de montrer $\langle X(t_0), X'(t_0)\rangle = 0$ c'est-à-dire que $X(t_0)$ est orthogonal à $X'(t_0)$.

## 2 Dérivées partielles des fonctions de plusieurs variables

**Définition 25.** Soit $U$ un ouvert de $\mathbb{R}^n$, $h \in \mathbb{R}^n$ $f : U \to E = \mathbb{R}^m$ une fonction et $A = (a_1, ..., a_n) \in U$. On dit que $f$ *est dérivable en $A$ suivant le vecteur $h$* si la fonction $t \to f(A + th)$ est dérivable en zéro.

On appelle alors *dérivée de $f$ en $A$ suivant $h$* et on note $D_h f(A) = \lim_{t\to 0} \frac{f(A+th)-f(A)}{t}$ cette dérivée.

On note $(e_1, ..., e_n)$ la base canonique de $R^n$.

**Définition 26.** Soit $U$ un ouvert de $\mathbb{R}^n$, $f : U \to E = \mathbb{R}^m$ une fonction et $A = (a_1, ..., a_n) \in U$. On dit que $f$ possède une *dérivée partielle d'indice $j$ (dans la base canonique) en $A$* (ou une dérivée partielle par rapport à la variable $x_j$) si elle est dérivable en $A$ suivant le vecteur $e_j$.

On appelle *dérivée partielle d'indice $j$ de $f$ en $A$*, et on note $D_j f(A)$ ou $\frac{\partial f}{\partial x_j}(A)$ cette dérivée, c'est à dire :

$$\frac{\partial f}{\partial x_j}(A) = \lim_{t\to 0} \frac{f(a_1, ..., a_i + t, ..., a_n) - f(A)}{t}.$$

*Remarque* 17. (i) $U$ est un ouvert, $A \in U$ donc $\exists r > 0, B_\infty(A, r) \subset U$ (avec la boule pour la norme infinie : $B_\infty(A, r) = ]a_1 - r, a_1 + r[\times ... \times]a_n - r, a_n + r[$. Donc, pour tout $h \in ]-r, r[$, on a $(a_1, ..., a_j + h, ..., a_n) \in U$, donc $f(a_1, ..., a_j + h, ..., a_n)$ est bien définie.

(ii) Soit $f_j : ]a_j - r, a_j + r[ \to E$ la fonction définie par $f_i(t) = (a_1, ..., t, ..., a_n)$ ($t$ dans la $j$-ème composante). Alors par définition $\frac{\partial f}{\partial x_j}(A)$ existe si et seulement si $f_j$ est dérivable en $a_j$ et on a : $f_j'(a_j) = \frac{\partial f}{\partial x_j}(A)$, d'où la formule de la fin de la définition.

(iii) Si $f = (f^{(1)}, ..., f^{(m)})$ avec les fonctions composantes $f^{(k)}$ on a $\frac{\partial f}{\partial x_j}(A) = (\frac{\partial f^{(1)}}{\partial x_j}(A), ..., \frac{\partial f^{(m)}}{\partial x_j}(A))$. Il suffit donc de traiter le cas $E = \mathbb{R}$ pour obtenir le cas général composante par composante.

(iv) Pour $n = 1$ on a une dérivée ordinaire.

(v) Les règles usuelles de dérivation (d'une somme, d'un produit) s'appliquent aux dérivées partielles. Par exemple, pour $f, g : U \to E, \lambda : U \to \mathbb{R}$ dérivables en $A$, on a

$$\frac{\partial}{\partial x_j}(\lambda f + g)(A) = \frac{\partial \lambda}{\partial x_j}(A) f(A) + \lambda(A)\frac{\partial f}{\partial x_j}(A) + \frac{\partial g}{\partial x_j}(A).$$

*Exemple* 21. (i) Soit $f(x, y) = x \sin(xy)$. On a

$$\frac{\partial f}{\partial x}(x, y) = sin(xy) + x(ycos(xy)),$$

$$\frac{\partial f}{\partial y}(x, y) = x^2 cos(xy).$$

(ii) Soit $f(x,y) = \frac{1}{x^2+y^2}, (x,y) \neq (0,0)$. On a

$$\frac{\partial f}{\partial x}(x,y) = \frac{-2x}{(x^2+y^2)^2},$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{-2y}{(x^2+y^2)^2}.$$

(iii) Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $f(x,y) = \frac{xy}{x^2+y^2}$ si $(x,y) \neq (0,0)$ et $f(0,0) = 0$. On étudie l'existence des dérivées partielles premières en $(0,0)$

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = \lim 0 = 0,$$

De même,

$$\frac{\partial f}{\partial y}(0,0) = 0.$$

*Remarque* 18. On a vu au chapitre 1 (exemple 14) que $\frac{xy}{x^2+y^2}$ n'a pas de limite en $(0,0)$, en particulier, $f$ ne peut pas être continue en $(0,0)$. On voit donc qu'**en dimension $n \geq 2$, l'existence de dérivées partielles en un point n'entraîne pas la continuité**.

# 3 Dérivées partielles successives

Soit $f : U \to E$ ($U$ ouvert de $\mathbb{R}^n$). Supposons que $\frac{\partial f}{\partial x_i}(x)$ existe pour tout $x \in U$. On définit ainsi une fonction : $\frac{\partial f}{\partial x_i} : U \to E$,

$$\frac{\partial f}{\partial x_i} : X \mapsto \frac{\partial f}{\partial x_i}(X).$$

Si cette fonction possède une dérivée partielle première en $A \in U$ par rapport à la $j$-ème variable, on pose :

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(A) := \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(A).$$

On dit que $\frac{\partial^2 f}{\partial x_j \partial x_i}$ est une dérivée partielle seconde. Si $i = j$, on note $\frac{\partial^2 f}{\partial x_i^2}(A)$.

De même pour les dérivées suivant les vecteurs $h, k$. Si $D_h f$ existe sur $U$ de sorte que l'on puisse définir une fonction $D_h f : U \to \mathbb{R}$ par $D_h f : a \mapsto (D_h f)(a)$. Alors si $D_h f$ est dérivable suivant le vecteur $k$ en $A$ on note :

$$(D_k D_h f)(A) = D_k(D_h f)(A).$$

Plus généralement (sous réserve que ceci ait un sens, c'est-à-dire que les dérivées successives de $f$ d'ordre inférieur à $p-1$ écrites existent dans $U$, puis que la dérivée partielle de la fonction $\frac{\partial^{p-1} f}{\partial x_{i_2} \ldots \partial x_{i_p}}$ ainsi définie sur $U$ existe en $A$), on peut définir :

$$\frac{\partial^p f}{\partial x_{i_1} \ldots \partial x_{i_p}}(A) := \frac{\partial}{\partial x_{i_1}}(\frac{\partial^{p-1} f}{\partial x_{i_2} \ldots \partial x_{i_p}})(A).$$

*Remarque* 19. Il y a des exemples (cf. TD et exemple suivant) où $\frac{\partial^2 f}{\partial x_j \partial x_i}(A) \neq \frac{\partial^2 f}{\partial x_i \partial x_j}(A)$. Donc en général, la valeur de la dérivée partielle dépend de l'ordre dans lequel on dérive.

*Exemple* 22. Soit $f$ définie sur $\mathbb{R}^2$ par $f(0,0) = 0$ et

$$f(x,y) = \frac{xy(x^2-y^2)}{x^2+y^2}, (x,y) \neq (0,0).$$

En dehors de (0,0, on a

$$\frac{\partial f}{\partial x}(x,y) = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2}, (x,y) \neq (0,0).$$

$$\frac{\partial f}{\partial y}(x,y) = \frac{-x(y^4+4x^2y^2-x^4)}{(x^2+y^2)^2}, (x,y) \neq (0,0).$$

et $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial x}(0,0) = 0$ On trouve ensuite

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x\to 0} \frac{1}{x}\frac{\partial f}{\partial y}(x,0) = 1$$

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y\to 0} \frac{1}{y}\frac{\partial f}{\partial x}(0,y) = -1.$$

## 3.1 Théorème de Schwarz

**Théorème 37.** *(Théorème de Schwarz) Soient $U$ un ouvert de $\mathbb{R}^n$, $f : U \to E$ une fonction et $h, k \in \mathbb{R}^2, A \in U$. On suppose que les dérivées suivant les vecteurs $h$ et $k$ $D_h f, D_k f$ existent dans $U$ et si $D_h(D_k f)$ et $D_k(D_h f)$ existent sur $U$ et sont* ***CONTINUES*** *en $A$. Alors on a :*

$$D_h(D_k f)(A) = D_k(D_h f)(A).$$

*Démonstration.* En traitant coordonnée par coordonnée, on peut supposer $E = \mathbb{R}$. Soit $A \in U$ tel que $B_F(A, 2R) \subset U$. Soit $r > 0$ tel que $||rh|| < R$, $||rk|| < R$ . On note pour $t \in ]0,r]$ :

$$\Delta(t) = \frac{1}{t^2}\left(f(A+th+tk) - f(A+tk) - f(A+th) + f(A)\right).$$

(remarquez que par nos choix $A+th+tk, A+th, A+tk, A \in B_F(A,2R) \subset U$)

Par l'hypothèse de dérivabilité, $\phi(s) = f(A+th+sk) - f(A+sk)$ est dérivable pour $s \in [0,t]$ de dérivée : $\phi'(s) = D_k f(A+th+sk) - D_k f(A+sk)$.

D'après le théorème fondamental du calcul différentiel d'une variable, on a alors :

$$\phi(t) - \phi(0) = \int_0^t ds(D_k f(A+th+sk) - D_k f(A+sk)).$$

De même si on pose $\psi(r) = (D_k f(A+rh+sk))$ de dérivée, $\psi'(r) = D_h(D_k f)(A+rh+sk))$, on déduit :

$$D_k f(A+th+sk) - D_k f(A+sk) = \int_0^t dr D_h(D_k f)(A+rh+sk)).$$

Soit après un changement de variables (linéaire) et en combinant les deux égalités précédentes :

$$\Delta(t) = \int_0^1 ds \int_0^1 dr D_h(D_k f)(A+rth+stk)).$$

Par hypothèse, on connaît la continuité de $D_h(D_k f)$ en $A$ donc il existe $\eta > 0$ tel que si $t \in ]0, \eta]$ :

$$\forall (r,s) \in [0,1]^2 \quad ||D_h(D_k f)(A + rth + stk)) - D_h(D_k f)(A))|| \leq \epsilon.$$

Ainsi, on a :

$$||\Delta(t) - D_h(D_k f)(A))|| \leq \int_0^1 ds \int_0^1 dr ||D_h(D_k f)(A + rth + stk)) - D_h(D_k f)(A))|| \leq \epsilon.$$

En conséquence, pour $t \to 0$ $\Delta(t) \to D_h(D_k f)(A))$. Par symétrie, $\Delta(t) \to_{t \to 0} D_k(D_h f)(A))$, d'où l'égalité par unicité de la limite. □

**Corollaire 38.** *(variante du Théorème de Schwarz) Soient $U$ un ouvert de $\mathbb{R}^n$, $f : U \to E$ une fonction et $A \in U$. On suppose que les dérivées $\frac{\partial^2 f}{\partial x_j \partial x_i}$ et $\frac{\partial^2 f}{\partial x_i \partial x_j}$ existent sur $U$ pour deux indices $i \neq j$ et sont **CONTINUES** en $A$. Alors on a :*

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(A) = \frac{\partial^2 f}{\partial x_i \partial x_j}(A).$$

*Remarque* 20.
1. Dans l'énoncé du résultat précédent, $\frac{\partial f}{\partial x_j}$ et $\frac{\partial f}{\partial x_i}$ existent sur $U$.
2. Il existe des exemples de fonctions ne vérifiant pas les hypothèses du Théorème de Schwarz (cf. TD)
3. La quasi-totalité des fonctions que nous utiliserons vérifie les hypothèses du Théorème de Schwarz.

## 3.2 Fonctions de classe $\mathcal{C}^p$.

**Définition 27.** Soit $f : U \to E$ une fonction définie sur un ouvert $U$ de $\mathbb{R}^n$. On dit que $f$ est de classe $\mathcal{C}^p$ ($p \in \mathbb{N}$) si $f$ ainsi que toutes ses dérivées partielles jusqu'à l'ordre $p$ sont continues. On dit que $f$ est de classe $\mathcal{C}^\infty$ si elle est de classe $\mathcal{C}^p$ pour tout $p \in \mathbb{N}$.

*Remarque* 21. Pour de telles fonctions de classe $\mathcal{C}^p$, les dérivées partielles sont notées pour $i_k \in [0, p]$, $k = i_1 + ... + i_n \leq p$ :

$$\frac{\partial^p f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}.$$

Si $i_k = 0$ on ne dérive pas par rapport à $x_k$. La notation fait bien sens car les dérivées partielles obtenues ne dépendent pas de l'ordre de dérivation grâce au théorème de Schwarz.

De plus soit $h$ un vecteur de $\mathbb{R}^n$ $h = (h_1, ..., h_n)$, on a $D_h f(a) = \sum_{k=1}^n h_k \frac{\partial f}{\partial x_k}(a)$ (cf partie 4.1). Si $f$ est de classe $\mathcal{C}^p$ on a une formule pour la dérivée itérée selon le vecteur $h$ [1] :

$$D_h^k f(a) := (D_h(D_h^{k-1} f))(a) = \sum_{i_1 + ... + i_n = p} \binom{p}{i_1 ... i_n} h_1^{i_1} ... h_n^{i_n} \frac{\partial^p f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(a).$$

1. analogue à la formule du multinôme qui se ramène au binôme de Newton pour $p = 2$ avec les coefficients multinomiaux $\binom{p}{i_1 ... i_k} := \frac{p!}{(i_1)! ... (i_n)!}$ :

$$(x_1 + ... + x_n)^p = \sum_{i_1 + ... + i_n = p} \binom{p}{i_1 ... i_n} x_1^{i_1} ... x_n^{i_n}$$

### 3.3 Un exemple d'Équation aux Dérivées Partielles (EDP)

Soit $f : \mathbb{R} \times \mathbb{R}_+^* \to \mathbb{R}$ définie par

$$f(x,t) = \frac{1}{\sqrt{t}} exp(-\frac{x^2}{4t}).$$

On calcule $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}, \frac{\partial^2 f}{\partial x^2}$ :

$$\frac{\partial f}{\partial x}(x,t) = \frac{-x}{2t\sqrt{t}} exp(-\frac{x^2}{4t}).$$

(on utilise la formule pour la dérivée d'un produit ici)

$$\frac{\partial^2 f}{\partial x^2}(x,t) = \frac{-1}{2t\sqrt{t}} exp(-\frac{x^2}{4t}) + \frac{x^2}{4t^2\sqrt{t}} exp(-\frac{x^2}{4t}).$$

$$\frac{\partial f}{\partial t}(x,t) = \frac{-1}{2t\sqrt{t}} exp(-\frac{x^2}{4t}) + \frac{x^2}{4t^2\sqrt{t}} exp(-\frac{x^2}{4t}).$$

On remarque que $f$ satisfait l'équation aux dérivées partielles suivante (dite "équation de la chaleur")

$$\frac{\partial f}{\partial t}(x,t) = \frac{\partial^2 f}{\partial x^2}(x,t).$$

# 4 Applications Différentiables.

### 4.1 Définition et premières propriétés

*Rappel* 22. On a vu au chapitre 1 (Théorème 26) que toutes les applications linéaires $T : \mathbb{R}^n \to \mathbb{R}^p$ sont continues et qu'il existe une constante $C$ tel que $||T(X)|| \leq C||X||$ pour tout $X \in \mathbb{R}^n$.

**Définition 28.** Soient $f : U \to \mathbb{R}^m$ une fonction définie sur un ouvert $U \subset \mathbb{R}^n$ et $x_0 \in U$. On dit que $f$ est *différentiable en* $x_0$ si il existe une application linéaire $T : \mathbb{R}^n \to \mathbb{R}^m$ telle que :

$$(*) \quad f(x_0 + H) = f(x_0) + T(H) + ||H||\epsilon(H), \ \epsilon(H) \to_{H\to 0} 0$$

pour tout $H \in \mathbb{R}^n$ tel que $x_0 + H \in U$.

Quand $T$ existe, on dit que $T$ est la *différentielle de* $f$ *en* $x_0$ (ou $T$ est tangente à $f$ en $x_0$) et on note : $T = df(x_0)$.

**Proposition 39.** *Si l'application linéaire $T$ est tangente à $f$ en $x_0$, alors, $f$ est dérivable suivant tout vecteur $h$ en $x_0$ et :*

$D_h f(x_0) = T(h)$. *En particulier, la différentielle de $f$ en $x_0$ est unique.*

*Démonstration.* En explicitant $(*)$ pour $H = th$ tel que $B_F(x_0, ||h||) \subset U$ on a $f(x_0 + th) = f(x_0) + tT(h) + t||h||\epsilon(th)$.

Donc $\frac{f(x_0+th)-f(x_0)}{t} = T(h) + ||h||\epsilon(th) \to_{t\to 0} T(h)$ d'où le résultat.

□

*Remarque* 23. (i) $(*)$ équivaut à $\epsilon(H) = \frac{f(x_0+H)-f(x_0)-df(x_0).H}{||H||} \to_{h\to 0} 0$

(ii) Si $f(x) = (f_1(x), ..., f_n(x))$. Comme on peut calculer les limites composante par composante, on déduit que $f$ est différentiable en $x_0$ si et seulement si pour tout $i$ $f_i$ est différentiable en $x_0$ et on a dans ce cas :

$$df(x_0).h = (df_1(x_0).h, ..., df_n(x_0).h).$$

(iii) *Notation de Landau* : Soit $f : U \to \mathbb{R}^p, g : U \to \mathbb{R}$, on note $f = o(g)$ en $x_0 \in U$ si $f(u) = g(u)\epsilon(u)$ avec $\epsilon(u) \to_{u \to x_0} 0$. Par exemple, la relation $(*)$ se réécrit :

$$f(x_0 + H) = f(x_0) + df(x_0).H + o(||H||).$$

**Corollaire 40.** *(Relation différentielle/Dérivées partielles) Soit $f : U \to \mathbb{R}^m$ une fonction définie sur un ouvert $U \subset \mathbb{R}^n$ et différentiable en $x_0 \in U$, alors $f$ admet des différentielles partielles premières en $x_0$ et si $H = (h_1, ..., h_n)$ :*

$$df(x_0).H = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i.$$

*Démonstration.* L'existence des dérivées partielles est conséquence de la proposition précédente. De plus, comme $df(x_0)$ est linéaire, on déduit en utilisant encore le calcul de la proposition précédente : $df(x_0).H = \sum_{i=1}^{n}(df(x_0).e_i)h_i = \sum_{i=1}^{n}(D_{e_i}f(x_0))h_i$. □

**Proposition 41.** *Si $f$ est différentiable en $x_0$, elle est continue en $x_0$.*

*Démonstration.* C'est évident vu $f(x_0+H) = f(x_0)+df(x_0).H+o(||H||)$ quand $H \to 0$ l'application linéaire (en dimension finie donc continue) vérifie $df(x_0).H \to df(x_0).0 = 0$ . □

*Remarque* 24. Les réciproques des deux résultats précédents sont fausses

(i) $f$ peut être continue en $x_0$ sans être différentiable (comme en dimension 1, ex $x \to |x|$ puisqu'elle n'admet pas de dérivées partielles en 0).

(ii) $f$ peut admettre des dérivées partielles sans être différentiable et même sans être continue en $x_0$ (cf. ex 21(iii)). La proposition précédente montre que la notion de différentiabilité plus forte évite ce résultat inhabituel en dimension $n = 1$.

## 4.2 Exemples et opérations de base

**Proposition 42.** *(Cas $n = 1$) Une application $f : U \to \mathbb{R}^m$ avec $U \subset \mathbb{R}$ est différentiable en $a \in U$ si et seulement si elle est dérivable en ce point. Sa différentielle est alors l'application linéaire :*

$$df(a) : h \in \mathbb{R} \mapsto hf'(a) \in F.$$

*En particuler , on a $f'(a) = (df(a).1)$.*

*Démonstration.* Si $f$ est dérivable si et seulement si $\frac{f(a+h)-f(a)-f'(a)h}{h} \to 0$ ssi $\frac{f(a+h)-f(a)-f'(a)h}{|h|} \to 0$ c'est à dire $f(a+h) - f(a) - f'(a)h = o(|h|)$. Par définition, ceci est équivalent à $f$ différentiable en $a$ avec la différentielle indiquée. □

**Proposition 43.** *(Cas des applications linéaires) La différentielle d'une application linéaire $T$ en tout point $t_0$ est $T$ elle même.*

*Démonstration.* (évident) $T(a + h) = T(a) + T(h)$ □

*Remarque* 25. Soit $(e_1, ..., e_n)$ la base canonique de $\mathbb{R}^n$, on note $(e_1^*, ..., e_n^*)$ la base duale de l'espace des formes linéaires sur $\mathbb{R}^n$ caractérisée par $e_j^*(e_i) = 1$ si $i = j$ et 0 si $i \neq j$. On peut aussi voir $e_i^*$ comme la i-ème projection $p_i$ parfois aussi notée $x_i$. Comme elle est linéaire on a alors $dx_j = x_j = e_j^*$. On note donc parfois des façons suivantes la relation du corollaire :

$$df(x_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i} e_i^*.$$

## 4.3 Applications continument différentiables

**Définition 29.** On dit que $f : U \to E$ (avec $U \subset \mathbb{R}^n$ un ouvert) est continument différentiable si elle est différentiable en tout point de $U$ et si pour tout $h$ de $E$ l'application $x \to D_h f(x)$ est continue.

*Rappel* 26. (Théorème des accroissements finis) Soit $f$ une fonction continue sur $[a, b]$ et dérivable sur $]a, b[$ alors il existe $c \in ]a, b[$ tel que $f'(c) = \frac{f(b)-f(a)}{b-a}$.

En particulier on en déduit l'inégalité des accroissements finis :

$$|f(b) - f(a)| \leq (\sup_{x \in ]a,b[} |f'(x)|)|b - a|.$$

**Théorème 44.** *Toute application est continument différentiable si et seulement si elle est de classe $\mathcal{C}^1$ (c'est-à-dire si elle admet des dérivées partielles continues.)*

*Démonstration.* Le sens direct est une conséquence immédiate de la définition vu que les dérivées partielles sont les $D_{e_i} h$.

Réciproquement supposons que $f : U \to E$ admet des dérivées partielles continues. Soit $a \in U$. Si $f$ est différentiable en $a$, nécessairement par les résultats précédents, sa différentielle doit être $T(h_1, ..., h_n) := \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a)$. On note $h = (h_1, ...h_n)$ et $H_k = (h_1, ...h_k, 0, ..., 0)$ ($H_0 = 0$). On va donc montrer que $r(h) := [f(a + h) - f(a) - T(h)]/||h|| \to_{h \to 0} 0$.

On muni $\mathbb{R}^n$ de la norme $||.|| = ||.||_\infty$, on a donc $||H_k|| \leq ||h||$.

On remarque qu'on a la somme télescopique suivante :

$$f(a + h) = f(a) + \sum_{i=1}^n (f(a + H_i) - f(a + H_{i-1})).$$

On a donc :

$$r(h) = \sum_{i=1}^n \frac{f(a + H_i) - f(a + H_{i-1}) - h_i \frac{\partial f}{\partial x_i}(a)}{||h||}.$$

Notons, $s_i(h) = f(a + H_i) - f(a + H_{i-1}) - h_i \frac{\partial f}{\partial x_i}(a)$. Par définition de la notion de dérivée partielle, la fonction $g_i : s \mapsto f(a + H_{i-1} + se_i) - s\frac{\partial f}{\partial x_i}(a)$ est dérivable en $t \in [0, h_k]$ de dérivée : $\frac{\partial f}{\partial x_i}(a + H_{i-1} + te_i) - \frac{\partial f}{\partial x_i}(a)$.

Par l'inégalité des accroissements finis (une variable), on déduit :

$$||s_i(h)|| = ||g_i(h_k) - g_i(0)|| \leq |h_k| \sup_{t \in [0,h_i]} \left\| \frac{\partial f}{\partial x_i}(a + H_{i-1} + te_i) - \frac{\partial f}{\partial x_i}(a) \right\|.$$

Comme $||H_{i-1} + te_i|| \leq ||h||$ pour $t \in [0, h_k]$ et vu $|h_k| \leq ||h||$, on déduit :

$$||r(h)|| \leq \frac{1}{||h||}\sum_{k=1}^{n}||s_k(h)|| \leq \sup_{||\ell||\leq||h||}\left\|\frac{\partial f}{\partial x_i}(a+\ell)-\frac{\partial f}{\partial x_i}(a)\right\|.$$

Or par l'hypothèse de continuité des $\frac{\partial f}{\partial x_i}$, le membre de droite ci-dessus tend vers 0 pour $||h|| \to 0$. □

# 5 Dérivation des fonctions composées

**Théorème 45.** *Soit $f : U \to E$ une application différentiable en $x_0 \in U$, un ouvert de $I\!R^n$. Soient $V \subset I\!R^p$ un ouvert et $g : V \to I\!R^n$ une fonction différentiable en $y_0 \in V$ telle que $g(V) \subset U$ et $g(y_0) = x_0$.*

*On peut définir $f \circ g : V \to E$.*

*Alors la fonction $f \circ g$ est différentiable en $y_0 \in U$ et on a :*

$$d(f \circ g)(y_0) = df(g(y_0)) \circ dg(y_0).$$

*Démonstration.* Les hypothèses de différentiabilités donnent pour tous $h \in U - \{x_0\}$ et $\ell \in V - \{y_0\}$ :

$$f(x_0+h) = f(x_0) + df(x_0).h + ||h||\epsilon_1(h), \quad g(y_0+\ell) = g(y_0) + dg(y_0).\ell + ||\ell||\epsilon_2(\ell),$$

avec $\epsilon_i$ tendant vers 0 en 0. Considérons $h(\ell) = g(y_0+\ell) - g(y_0) = g(y_0+\ell) - x_0$, de sorte que par hypothèse pour $\ell \in V - \{y_0\}$ (vu $g(y_0+\ell) \in U$) on a $h(\ell) \in U - \{x_0\}$. De plus $(f \circ g)(y_0+\ell) = f(x_0 + h(\ell))$ de sorte qu'en appliquant les formules précédentes :

$$\begin{aligned}(f \circ g)(y_0+\ell) &= f(x_0) + df(x_0).(h(\ell)) + ||h(l)||\epsilon_1(h(l)) \\ &= f(x_0) + df(x_0).(dg(y_0)(\ell)) + ||\ell||df(x_0).(\epsilon_2(\ell)) + ||h(l)||\epsilon_1(h(l)) \\ &= f(x_0) + df(x_0).(dg(y_0).(\ell)) + ||\ell||r(\ell).\end{aligned}$$

avec $r(\ell) = df(x_0).(\epsilon_2(\ell)) + \frac{||h(\ell)||}{||\ell||}\epsilon_1(h(\ell)) = df(x_0).(\epsilon_2(\ell)) + ||dg(y_0).(\ell/||\ell||) + \epsilon_2(\ell)||\epsilon_1(h(\ell))$.
Or on a la majoration

$$||r(\ell)|| \leq C||\epsilon_2(\ell)|| + D||\epsilon_1(h(\ell))|| + ||\epsilon_2(\ell)||\ ||\epsilon_1(h(\ell))|| \to_{\ell\to 0} 0$$

vu $df(x_0)$ (resp. $dg(y_0)$= donc $C$ (resp. $D$)-lipschitzienne. Donc par définition $(f \circ g)$ est différentiable en $y_0$ avec la différentielle annoncée. □

**Théorème 46.** *Soit $f : U \to E$ une application différentiable en $x_0 \in U$, un ouvert de $I\!R^n$. Soient $I$ un intervalle de $I\!R$ et $\Phi : I \to I\!R^n$ une fonction dérivable en $t_0 \in I$ tel que $\Phi(I) \subset U$ et $\Phi(t_0) = x_0$.*

*On pose $\Phi(t) = (\varphi_1(t), ..., \varphi_n(t))$. On peut définir $f \circ \Phi : I \to E$ par $f \circ \Phi(t) = f(\varphi_1(t), ..., \varphi_n(t))$.*

*Alors la fonction $f \circ \Phi$ est dérivable en $t_0$ et on a :*

$$(f \circ \Phi)'(t_0) = df(\Phi(t_0))(\Phi'(t_0)) = \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\Phi(t_0))\varphi_i'(t_0).$$

*Remarque* 27. Géométriquement, $\Phi$ et $f \circ \Phi$ définissent des courbes (ou "arcs paramétrés"). Le théorème dit donc que l'application linéaire $df(\Phi(t_0))$ transforme le vecteur $\Phi'(t_0)$ tangent à $\Phi$ en $t_0$ pour l'envoyer vers le vecteur $(f \circ \Phi)'(t_0)$ tangent à $f \circ \Phi$ en $t_0$.

*Démonstration.* On prend $g = \Phi$ dans le théorème précédent et on se souvient que pour une application partant de $\mathbb{R}$, $d(f \circ \Phi)(t_0).1 = (f \circ \Phi)'(t_0)$ et

$$df(\Phi(t_0)) \circ d\Phi(t_0).1 = df(\phi(t_0)).(\Phi'(t_0)1) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\Phi(t_0))\varphi_i'(t_0),$$

par la proposition 42. □

**Théorème 47.** *Soit $f : U \to E$ une application différentiable en $X = (x_1, ..., x_n) \in U$ (par exemple $f$ $\mathcal{C}^1$) pour $U$ un ouvert de $\mathbb{R}^n$. Soient $V \subset \mathbb{R}^p$ un ouvert et $g : V \to \mathbb{R}^n$ une fonction possédant des dérivées partielles en $Y = (y_1, ..., y_p) \in V$ tel que $g(V) \subset U$ et $g(Y) = V$.*

*On peut définir $f \circ g : V \to E$.*

*Alors la fonction $f \circ g$ possède des dérivées partielles en $Y \in U$ et on a :*

$$\frac{\partial f \circ g}{\partial y_j}(y_1, ..., y_p) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g(y_1, ..., y_p)) \times \frac{\partial g_i}{\partial y_j}(y_1, ..., y_p).$$

*Remarque* 28.

1. Pour simplifier les notations si on écrit $X = g(y_1, ..., y_p)$, $Y = (y_1, ..., y_p)$, alors on a :

$$\frac{\partial f \circ g}{\partial y_j}(Y) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X)\frac{\partial g_i}{\partial y_j}(Y).$$

2. Dans la formule précédente, on a

$$\frac{\partial f}{\partial x_i}(X) = \frac{\partial f}{\partial x_i}(g_1(Y), ..., g_n(Y)).$$

   Ici $\frac{\partial f}{\partial x_i}$ désigne la dérivée partielle de $f$ par rapport à la $i$-ème variable (parfois écrite $D_i f$) Donc $\frac{\partial f}{\partial x_i}$ est un symbole et le $x_i$ qui y figure N'EST PAS UNE VARIABLE. Il est donc INTERDIT d'écrire :

$$\frac{\partial f}{\partial g_i(Y)} \quad \text{PAS DE SENS, ATTENTION!!!.}$$

*Démonstration.* Pour calculer les dérivées partielles on dérive $h_j(t) = (f \circ g)(y_1, ...y_j + t, ..., y_p) = f \circ \Phi(t)$ avec $\Phi(t) = g(y_1, ...y_j + t, ..., y_p)$, il suffit donc d'appliquer le résultat précédent à ce $\Phi$. □

**Corollaire 48.** *Soit $f : U \to E$ une application $\mathcal{C}^p$ ($p \in \mathbb{N}$,ou $p = \infty$) sur un ouvert $U$ de $\mathbb{R}^n$. Soient $V \subset \mathbb{R}^p$ un ouvert et $g : V \to \mathbb{R}^n$ une fonction $\mathcal{C}^p$ avec $g(V) \subset U$.*

*Alors $f \circ g : V \to E$ est de classe $\mathcal{C}^p$.*

*Démonstration.* On montre donc le résultat par récurrence et il reste à voir pourquoi si le résultat est vrai au rang $p \geq 0$ pourquoi il est vrai au rang $p + 1$.

Le cas $p = 0$ est déjà connu et le cas $p = \infty$ vient de l'ensemble des autres cas. Or par le théorème précédent, les dérivées partielles existent et sont :

$$\frac{\partial f \circ g}{\partial y_j}(y_1, ..., y_p) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g(y_1, ..., y_p)) \times \frac{\partial g_i}{\partial y_j}(y_1, ..., y_p).$$

Comme ce sont des sommes, produits et composées de fonctions $\mathcal{C}^p$ on déduit que $f \circ g$ est de classe $\mathcal{C}^{p+1}$ puisque ces dérivées partielles sont de classe $\mathcal{C}^p$. □

## 5.1 Exemples

*Exemple* 23. (Pour le théorème 46)

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^2y + y^3$, $\phi : \mathbb{R} \to \mathbb{R}^2$ définie par $\phi(t) = (\cos(t), t^4)$. On note $\phi = (\phi_1, \phi_2)$. $f$ est $\mathcal{C}^1$ par les opérations usuelles et $\phi$ dérivable en tout $t \in \mathbb{R}$. Donc $f \circ \phi$ est dérivable en $t$ et :

$$(f \circ \phi)'(t) = \frac{\partial f}{\partial x}(\phi(t))\phi_1'(t) + \frac{\partial f}{\partial y}(\phi(t))\phi_2'(t).$$

On peut calculer séparément :

$$\frac{\partial f}{\partial x} = 2xy, \quad \frac{\partial f}{\partial y} = x^2 + 3y^2,$$

$$\phi_1'(t) = -\sin(t), \quad \phi_2'(t) = 4t^3.$$

On trouve

$$(f \circ \phi)'(t) = -2\cos(t)t^4\sin(t) + (\cos(t)^2 + 3t^8)4t^3.$$

On peut retrouver cette valeur en dérivant directement :$(f \circ \phi)(t) = \cos^2(t)t^4 + t^{12}$.

*Exemple* 24. (Pour le théorème 47) Soient $f : \mathbb{R}^2 \to \mathbb{R}$, $g : \mathbb{R}^2 \to \mathbb{R}^2$ définies par $f(x,y) = \sin(x+y)\cos(x-y)$, $g(u,v) = (\frac{uv}{u^2+v^2}, u^2+v^2)$ si $(u,v) \neq (0,0)$ et $g(0,0) = (0,0)$, donc on note $g_1(u,v) = \frac{uv}{u^2+v^2}$si $(u,v) \neq (0,0)$ et $g_1(0,0) = 0$, $g_2(u,v) = u^2+v^2$.

On pose $H(u,v) = f \circ g(u,v) = f(g_1(u,v), g_2(u,v))$. Le but est de calculer $\frac{\partial H}{\partial u}, \frac{\partial H}{\partial v}$.

On calcule les dérivées partielles de $f$ et $g$ :

$$\frac{\partial f}{\partial x}(x,y) = -\sin(x+y)\sin(x-y) + \cos(x+y)\cos(x-y) = \cos(2x)$$
$$\frac{\partial f}{\partial y}(x,y) = \sin(x+y)\sin(x-y) + \cos(x+y)\cos(x-y) = \cos(2y).$$

Par composition, ces dérivées partielles sont bien continues donc $f$ est bien de classe $\mathcal{C}^1$.

$$\frac{\partial g_2}{\partial u}(u,v) = 2u, \quad \frac{\partial g_2}{\partial v}(u,v) = 2v$$

Si $(u,v) \neq (0,0)$

$$\frac{\partial g_1}{\partial u}(u,v) = \frac{v}{u^2+v^2} - \frac{2u^2v}{(u^2+v^2)^2} = \frac{v(v^2-u^2)}{(u^2+v^2)^2}$$
$$\frac{\partial g_1}{\partial v}(u,v) = \frac{u}{u^2+v^2} - \frac{2v^2u}{(u^2+v^2)^2} = \frac{u(u^2-v^2)}{(u^2+v^2)^2}$$

et on a vu à l'exemple 21 que $g_1$ admet aussi des dérivées partielles en $(0,0)$ (sans être continue en $(0,0)$) :

$$\frac{\partial g_1}{\partial u}(0,0) = 0, \quad \frac{\partial g_1}{\partial v}(0,0) = 0.$$

On peut donc appliquer le théorème partout sur $\mathbb{R}^2$.

Pour $(u,v) \neq (0,0)$, on a donc :

$$\frac{\partial H}{\partial u}(u,v) = \cos(2\frac{uv}{u^2+v^2})\frac{v(v^2-u^2)}{(u^2+v^2)^2} + \cos(2(u^2+v^2))2u,$$

$$\frac{\partial H}{\partial u}(u,v) = \cos(2\frac{uv}{u^2+v^2})\frac{u(u^2-v^2)}{(u^2+v^2)^2} + \cos(2(u^2+v^2))2v,$$

$$\frac{\partial H}{\partial u}(0,0) = 0,$$

$$\frac{\partial H}{\partial u}(0,0) = 0.$$

*Exemple* 25. (Pour le théorème 47 et son corollaire) Soient $f : \mathbb{R}^2 \to \mathbb{R}$ et $g : \mathbb{R}^2 \to \mathbb{R}^2$ définies par $f(x,y) = x^2y + y^3$ et $g(u,v) = (\sin(u), uv) = (g_1(u,v), g_2(u,v))$.

On a donc $f \circ g : \mathbb{R}^2 \to \mathbb{R}$ et $f \circ g(u,v) = \sin^2(u)uv + u^3v^3$.

$f$ et $g$ sont $\mathcal{C}^\infty$ vu que leurs composantes sont soit polynomiales, soit à une variable et $\mathcal{C}^\infty$. Donc $f \circ g$ est $\mathcal{C}^\infty$ dans $\mathbb{R}^2$ et on a :

$$\begin{aligned}\frac{\partial f \circ g}{\partial u}(u,v) &= \frac{\partial f}{\partial x}(g(u,v)).\frac{\partial g_1}{\partial u}(u,v) + \frac{\partial f}{\partial y}(g(u,v)).\frac{\partial g_2}{\partial u}(u,v) \\ &= 2uv\sin(u).\cos(u) + (\sin^2 + 3u^2v^2).v,\end{aligned}$$

$$\begin{aligned}\frac{\partial f \circ g}{\partial v}(u,v) &= \frac{\partial f}{\partial x}(g(u,v)).\frac{\partial g_1}{\partial v}(u,v) + \frac{\partial f}{\partial y}(g(u,v)).\frac{\partial g_2}{\partial v}(u,v) \\ &= 2uv\sin(u).0 + (\sin^2 + 3u^2v^2).u = (\sin^2 + 3u^2v^2).u.\end{aligned}$$

# 6 Théorème des Accroissements finis

Pour un Rappel en dimension 1 cf. Rappel 26.

**Théorème 49.** *Soit $U$ un ouvert* convexe *de $\mathbb{R}^n$, soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Alors*

$$\forall a = (a_1, ..., a_n), b = (b_1, ..., b_n) \in U, \exists \theta \in ]0,1[ \quad f(b) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + \theta(b-a)).(b_i - a_i),$$

*Remarque* 29. (i) En posant $h = b - a = (h_1, ..., h_n)$ la relation s'écrit :

$$f(a+h) - f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + \theta h).(h_i).$$

(ii) U étant convexe $a + \theta(b-a) \in [a,b] \subset U$.

(iii) La relation s'écrit aussi $f(a+h) - f(a) = df(a+\theta h).h$

(iv) Ici $f$ fonction numérique, $f : U \to \mathbb{R}^p$ $p > 1$ ne marche pas. (ex prendre $f(t) = (\cos(t), \sin(t))$)

*Démonstration.* Soient $a, b \in U$. On définit $\phi : [0,1] \to \mathbb{R}^n$ par $\phi(t) = a + t(b-a) = tb + (1-t)a$. $\phi$ est dérivable et $\phi'(t) = (b-a)$. $\phi(t) \in U$ car $U$ et convexe. Donc $\phi : [0,1] \to U$.

On peut poser $F(t) = f(\phi(t)) = f(a + t(b-a))$. D'après le théorème des fonctions composées (Th 46) , on sait que $F$ est dérivable et

$$F'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + t(b-a)).(b_i - a_i).$$

Enfin, $F$ vérifie les hypothèses du TAF habituel (F continue sur $[0,1]$ dérivable sur $]0,1[$) donc il existe $\theta \in ]0,1[$ $F(1)-F(0)=F'(\theta)(1-0)=F'(\theta)$. C'est exactement le résultat. $\square$

**Proposition 50.** *(Inégalité des accroissements finis) Soient $U$ un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$.*

*Supposons que $M = \max_{i=1,\dots,n} \sup_{x\in U} |\frac{\partial f}{\partial x_i}(x)| < \infty$, alors :*

$$\forall a,b \in U |f(b)-f(a)| \leq M||b-a||_1.$$

*En particulier, $f$ est lipschitzienne.*

*Démonstration.* En appliquant TAF, on obtient

$$|f(b)-f(a)| \leq \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i}(a+\theta(b-a)) \right| |b_i - a_i| \leq M \sum_{i=1}^{n} |b_i - a_i| = M||b-a||_1.$$

$\square$

**Proposition 51.** *(Inégalité des accroissements finis) Soient $U$ un ouvert convexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^m$ de classe $\mathcal{C}^1$. si pour tout $i$, pour tout $x \in U$, $\frac{\partial f}{\partial x_i}(x) = 0$ alors $f$ est constante dans $U$.*

*Démonstration.* La proposition se réduit au cas $m = 1$ en prenant les coordonnées. On a $M = 0$ dans le théorème précédent donc pour tout $a, b \in U$, $|f(b)-f(a)| \leq 0$ $\square$

*Remarque* 30. Ce résultat est vrai dans des ouverts plus généraux que les ouverts convexes. Mais elle est fausse sur un ouvert quelconque par ex. sur la réunion de 2 ouverts disjoints.

# 7 Formules de Taylor

## 7.1 Rappel : Formule de Taylor avec reste intégral en dimension 1

**Proposition 52.** *Soit $f : I \to E$ une fonction $\mathcal{C}^{n+1}$ sur un intervalle $I$. Soit $a, b \in I$ alors :*

$$f(b) = \sum_{k=0}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)dt$$

*Démonstration.* En considérant coordonnée par coordonnée, on peut considérer $E = \mathbb{R}$.

Pour $n = 0$, c'est le théorème fondamental du calcul intégro-differentiel. On montre le résultat par récurrence sur $n$. Pour $n \geq 1$, on suppose le résultat au rang $n-1$ et on intègre par partie :

$$\begin{aligned}\int_a^b \frac{(b-t)^{n-1}}{(n-1)!} f^{(n)}(t)dt &= \left[ -\frac{(b-t)^n}{n!} f^{(n)}(t) \right]_a^b + \int_a^b \frac{(b-t)^n}{(n)!} f^{(n+1)}(t)dt \\ &= \frac{(b-a)^n}{n!} f^{(n)}(t) + \int_a^b \frac{(b-t)^n}{(n)!} f^{(n+1)}(t)dt.\end{aligned}$$

En remplaçant dans l'hypothèse de récurrence on obtient le résultat. $\square$

## 7.2 Formule de Taylor avec reste intégral : cas général

**Théorème 53.** *Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to E$ de classe $\mathcal{C}^{k+1}$ ($k \geq 0$). Si $[a, a+h] \subset U$, on a :*

$$f(a+h) = \sum_{p=0}^{k} \frac{1}{p!} D_h^p f(a) + \int_0^1 D_h^{k+1} f(a+sh) \frac{(1-s)^k}{k!} ds.$$

*Démonstration.* Notons $\phi(s) = f(a+sh)$. On a $\phi'(s) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a+sh)h_i = D_h f(a+sh)$ par dérivation des fonctions composées. En itérant ce résultat, on obtient $\phi^{(p)}(s) = (D_h^p f)(a+sh)$. On obtient le résultat en appliquant la formule de Taylor avec reste intégral de dimension 1 à $\phi$. □

## 7.3 Formule de Taylor Young

**Théorème 54.** *Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to E$ de classe $\mathcal{C}^k$ ($k \geq 1$). Si $[a, a+h] \subset U$, on a :*

$$f(a+h) = \sum_{p=0}^{k} \frac{1}{p!} D_h^p f(a) + o(||h||^k).$$

*Démonstration.* Soit $r > 0$ tel que $B(a,r) \subset U$. La formule de Taylor avec reste intégral à l'ordre $(k-1)$ montre que :

$$f(a+h) = \sum_{p=0}^{k} \frac{1}{p!} D_h^p f(a) + \int_0^1 (D_h^k f(a+sh) - D_h^k f(a)) \frac{(1-s)^{k-1}}{(k-1)!} ds.$$

(On a remarqué que $\int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} = \frac{1}{(k)!}$) Comme on a la formule,

$$D_h^p f(a) := (D_h(D_h^{p-1} f))(a) = \sum_{i_1+\ldots+i_n=p} \begin{pmatrix} p \\ i_1 \ldots i_n \end{pmatrix} h_1^{i_1} \ldots h_n^{i_n} \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a),$$

on déduit :

$$r_s(h) := ||(D_h^p f(a+sh) - D_h^p f(a))|| \leq \sum_{i_1+\ldots+i_n=p} \begin{pmatrix} p \\ i_1 \ldots i_n \end{pmatrix} |h_1|^{i_1} \ldots |h_n|^{i_n} || \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a+sh)) - \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a))||.$$

D'où on déduit :

$$r_s(h) \leq \left( \sum_{i_1+\ldots+i_n=p} \begin{pmatrix} p \\ i_1 \ldots i_n \end{pmatrix} |h_1|^{i_1} \ldots |h_n|^{i_n} \right) \max_{i_1+\ldots+i_n=p} \sup_{s \in [0,1]} || \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a+sh) - \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a)||$$

Soit encore en appliquant le multinôme de Newton :

$$||(D_h^k f(a+sh) - D_h^k f(a))|| \leq ||h||_1^k \max_{i_1+\ldots+i_n=k} \sup_{s \in [0,1]} || \frac{\partial^k f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a+sh)) - \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a))|| = o(||h||_1^k).$$

La dernière égalité vient de la continuité des dérivées jusqu'à l'ordre $p \leq k$ d'où

$$\max_{i_1+\ldots+i_n=p} \sup_{s \in [0,1]} || \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a+sh)) - \frac{\partial^p f}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}}(a))|| \to_{h \to 0} 0.$$

□

**Corollaire 55.** *Soit $U \subset I\!R^n$ un ouvert et $f : U \to E$ de classe $\mathcal{C}^2$. Si $[a, a+h] \subset U$, on a :*

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_j + \frac{1}{2!} \sum_{j,k=1}^{n} \frac{\partial^2 f}{\partial x_k \partial x_j}(a) h_j h_k + o(||h||^2).$$

## 7.4 Théorème de Taylor-Lagrange

**Lemme 56.** *(Formule de la moyenne) Soit $\phi : [a,b] \to I\!R$ et $\psi : [a,b] \to I\!R$ deux fonctions continues. On suppose que $\forall t \in [a,b]$ $\phi(t) \geq 0$. Alors on a :*

$$\exists c \in [a,b] \int_a^b \phi(t)\psi(t)dt = \psi(c) \int_a^b \phi(t)dt.$$

*Démonstration.* Si $\int_a^b \phi(t)dt = 0$ on a $\phi(t) = 0$ pour tout $t$ donc le résultat est évident. On suppose $\int_a^b \phi(t)dt \neq 0$ Comme $\phi$ est continue sur $[a,b]$ elle est bornée et atteint ses bornes. On a donc $x_1, x_2 \in [a,b]$ $m = \psi(x_1) \leq \psi(x) \leq M = \psi(x_2)$ pour tout $x \in [a,b]$.

Vu $\phi(t) \geq 0$, on a :

$$m \int_a^b \phi(t)dt \leq \int_a^b \phi(t)\psi(t)dt \leq M \int_a^b \phi(t)dt.$$

D'où on a $\psi(x_1) \leq \frac{\int_a^b \phi(t)\psi(t)dt}{\int_a^b \phi(t)dt} \leq \psi(x_2)$.

Comme $\psi$ est continue, par le TVI, il existe $c \in [x_1, x_2]$ tel que $\psi(c) = \frac{\int_a^b \phi(t)\psi(t)dt}{\int_a^b \phi(t)dt}$. □

**Théorème 57.** *Soient $U \subset I\!R^n$ un ouvert et $f : U \to I\!R$ de classe $\mathcal{C}^p$ $(p \geq 1)$. Si $[a, a+h] \subset U$, il existe $\theta \in [0,1]$ :*

$$f(a+h) = \sum_{p=0}^{k-1} \frac{1}{p!} D_h^p f(a) + \frac{1}{k!} D_h^k f(a + \theta h).$$

*Remarque* 31. En appliquant le TAF à une fonction appropriée, on peut en fait montrer que l'on peut choisir $\theta \in ]0,1[$.

*Démonstration.* On applique le lemme précédent avec $a = 0, b = 1$ aux fonctions $\psi(s) = D_h^k(a + s(b-a))$ et $\phi(s) = \frac{(1-s)^{k-1}}{(k-1)!}$. L'intégrale correspondante apparaît dans la formule de Taylor avec reste intégral et $\int_0^1 \phi(s)ds = \frac{1}{k!}$. D'où le résultat. □

# 8 Fonctions homogènes, Théorème d'Euler

**Définition 30.** Soit $C \subset \mathbb{R}^n$, on dit que $C$ est **symétrique** (autour de l'origine) si $x \in C \Rightarrow -x \in C$.

On dit que que $C$ est un **cône** si $\forall x \in C, \forall t > 0, tx \in C$.

.

*Exercice* 4. Les boules ouvertes et fermées sont symétriques. $(\mathbb{R}_+^*)^n$, $\mathbb{R}_+^n$ sont des cônes. Un sous espace vectoriel ou $(\mathbb{R}^*)^n$ sont des cônes symétriques. $\{(x,y) \mid y \geq 0, x \geq y\}$ est un cône. $\{(x,y) \mid |x| \geq |y|\}$ est un cône symétrique.

**Définition 31.** Soit $f : C \to \mathbb{R}$ une fonction définie sur un cône symétrique $C$ est dite **homogène** de degré $k \in \mathbb{Z}$ si pour tout $t \in \mathbb{R}^*$, pour tout $x \in C$ $f(tx) = t^k f(x)$.

Soit $f : C \to \mathbb{R}$ une fonction définie sur un cône $C$ est dite **positivement homogène** de degré $\alpha \in \mathbb{R}$ si pour tout $t \in ]0, \infty[$, pour tout $x \in C$ $f(tx) = t^\alpha f(x)$.

*Exemple* 26.
1. Une norme est positivement homogène de degré 1.
2. Une fonction polynomiale est homogène de degré $k$ si et seulement si tous ses monômes sont de degré $k$.
3. (Fonctions de Cobb-Douglas) Soient $(\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$ et $f : (\mathbb{R}_+^*)^n \to \mathbb{R}$ définie par : $f(x_1, ..., x_n) = x_1^{\alpha_1} ... x_n^{\alpha_n}$. $f$ est positivement homogène de degré $\alpha_1 + ... + \alpha_n$.

**Proposition 58.** *Soit $C \subset \mathbb{R}^n$, $D \subset \mathbb{R}^p$ des cônes (resp. des cônes symétriques), $f, g : C \to \mathbb{R}$, $h = (h_1, ...., h_n) : D \to C$.*

1. *Si $f, g$ sont positivement homogènes (resp. homogènes) de degré $d$ alors $f + g$ est positivement homogène (resp. homogène) de degré $d$.*
2. *Si $f, g$ sont positivement homogènes (resp. homogènes) de degré $d_1$ et $d_2$ alors $fg$ est positivement homogène (resp. homogène) de degré $d_1 + d_2$.*
3. *Si $f$ est positivement homogène (resp. homogène) de degré $d_1$ et $h_1, ..., h_n$ sont positivement homogènes (resp. homogènes) de degré $d_2$ alors $f \circ h$ est positivement homogène (resp. homogène) de degré $d_1 d_2$.*
4. *Si $C$ est ouvert et si $f$ est positivement homogène (resp. homogène) de degré $d$ et admet une dérivée selon le vecteur $h \in \mathbb{R}^n$ alors $D_h f$ est positivement homogène (resp. homogène) de degré $d - 1$.*

*Démonstration.* Les preuves sont faciles. Si $f(tx) = t^d f(x)$, $g(tx) = t^d g(x)$ alors $f(tx) + g(tx) = t^d(f(x) + g(x))$.

Si $f(tx) = t^{d_1} f(x)$, $g(tx) = t^{d_2} g(x)$ alors $f(tx)g(tx) = t^{d_1} f(x) t^{d_2} g(x) = t^{d_1+d_2} f(x)g(x)$.

Si $f(tx) = t^{d_1} f(x)$, $h_i(tx) = t^{d_2} h_i(x)$ alors

$$f(h_1(tx), ..., h_n(tx)) = f(t^{d_2} h_1(x), ..., t^{d_2} h_n(x)) = (t^{d_2})^{d_1} f(h_1(x), ..., h_n(x)) = (t^{d_1 d_2}) f(h_1(x), ..., h_n(x)).$$

Enfin si $f$ admet une dérivée selon le vecteur $h$ on a :

$$D_h f(tx) = \lim_{u \to 0} \frac{f(tx + uh) - f(tx)}{u} = \lim_{u \to 0} \frac{f(tx + uth) - f(tx)}{ut} = \lim_{u \to 0} t^d \frac{f(x + uh) - f(x)}{ut} = t^{d-1} D_h f(x).$$

□

**Théorème 59.** *(Théorème d'Euler) Soit $C$ un cône ouvert et $f : C \to \mathbb{R}$.*

1. *Soit $f : C \to \mathbb{R}$ une fonction positivement homogène de degré $\alpha$ et si $f$ est différentiable en $X = (x_1, ..., x_n) \in C$, alors $f$ vérifie l'identité d'Euler :*

$$x_1 \frac{\partial f}{\partial x_1}(X) + \cdots + x_n \frac{\partial f}{\partial x_n}(X) = \alpha f(X).$$

2. *Réciproquement, si $f$ est différentiable en tout $X \in C$ et vérifie l'identité d'Euler, alors $f$ est positivement homogène de degré $\alpha$ sur $C$.*

*Démonstration.* Soit $f$ positivement homogène de degré $\alpha$, on considère $\phi : ]0,\infty[ \to \mathbb{C}$ définie par $\phi(t) = (tx_1, ..., tx_n)$. en appliquant le théorème de dérivation des fonctions composées à $f \circ \phi$ on obtient

$$(f \circ \phi)'(t) = x_1 \frac{\partial f}{\partial x_1}(tX) + \cdots + x_n \frac{\partial f}{\partial x_n}(tX).$$

Par ailleurs, comme $f$ est positivement homogène, on a $f \circ \phi(t) = t^\alpha f(X)$, d'où $(f \circ \phi)'(t) = \alpha t^{\alpha-1} f(X)$ ce qui conclut en évaluant en $t = 1$.

Réciproquement $g(t) = t^{-\alpha} f \circ \phi(t)$ est dérivable sur $]0,\infty[$ de dérivée :

$$g'(t) = -\alpha t^{-\alpha-1}(f \circ \phi)'(t) + t^{-\alpha-1}\left(tx_1 \frac{\partial f}{\partial x_1}(tX) + \cdots + tx_n \frac{\partial f}{\partial x_n}(tX)\right) = 0.$$

où la dernière égalité vient de l'identité d'Euler. Donc $g$ est constante, c'est à dire $g(t) = g(1)$ ce qui équivaut à $f$ homogène de degré $\alpha$. $\square$

# 9 Extrema libres

Le but de cette partie est d'étudier l'optimisation sans contrainte d'une fonction numérique.

**Définition 32.** Soit $f : A \to \mathbb{R}$ une fonction définie sur $A \subset \mathbb{R}^n$ et soit $a \in A$.

- $f$ possède un **maximum local (ou relatif) en** $a$ (resp. maximum local strict) si

$$\exists \rho > 0 : \forall x \in B(a,\rho) \cap A,\ f(x) \leq f(a),\ (\text{resp. } f(x) < f(a)).$$

- $f$ possède un **minimum local (ou relatif) en** $a$ (resp. minimum local strict) si

$$\exists \rho > 0 : \forall x \in B(a,\rho) \cap A,\ f(x) \geq f(a),\ (\text{resp. } f(x) > f(a)).$$

- $f$ possède un **extremum local (ou relatif) en** $a$ si elle possède soit un minimum local en $a$ soit en maximum local en $a$.
- $f$ possède un **maximum global en** $a$ si

$$\forall x \in A f(x) \leq f(a).$$

- $f$ possède un **minimum global en** $a$ si

$$\forall x \in A f(x) \geq f(a).$$

- $f$ possède un **extremum global en** $a$ si elle possède soit un minimum global en $a$ soit en maximum global en $a$.

## 9.1 Conditions nécessaires du premier ordre pour un extremum local

**Théorème 60.** *Soit $U$ un ouvert de $\mathbb{R}^n$. Soit $f : U \to \mathbb{R}$ une fonction possédant des dérivées partielles premières sur $U$. Alors si $a$ est un extremum local de $f$, on a :*

$$\forall i = 1, ..., n\ \frac{\partial f}{\partial x_i}(a) = 0.$$

*Remarque* 32. Si $f$ est différentiable en $a$, cela s'écrit $df(a) = 0$. Par ailleurs la réciproque est fausse.

*Démonstration.* Soit $a = (a_1, ..., a_n) \in U$, $U$ étant ouvert, il existe $\rho > 0$ tel que $B = B_{||.||_\infty}(a, \rho) = ]a_1 - \rho, a_1 + \rho[\times ... \times]a_n - \rho, a_n + \rho[\subset U$, et tel que $a$ soit un extremum de $f$ sur $B$. Donc $f_i : ]-\rho, \rho[\to \mathbb{R}$ est bien définie par $f_i(t) = f(a_1, ..., a_i + t, ..., a_n)$ et est dérivable sur $] - \rho, \rho[$

$t \mapsto f_i(t)$ possède encore un extremum en $t = 0$. D'où $f_i'(0) = 0$.

(Rappel première année : quitte à prendre $-f$, on suppose $f$ a un minimum, vu pour $x \geq 0$, $\frac{f_i(x)-f_i(0)}{x} \geq 0$ la dérivée à droite est positive, et vu pour $x \leq 0$, $\frac{f_i(x)-f_i(0)}{x} \leq 0$, la dérivée à gauche est négative, donc la dérivée, que l'on suppose exister, doit être positive et négative, donc nulle).

Or par définition $\frac{\partial f}{\partial x_i}(a) = f_i'(0) = 0$. □

**Définition 33.** Soit $U$ un ouvert de $\mathbb{R}^n$. Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et soit $a \in U$. On dit que $a$ est un **point critique** de $f$ si

$$\forall i = 1, ..., n \ \frac{\partial f}{\partial x_i}(a) = 0.$$

**Définition 34.** Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ sur l'ouvert $U \subset \mathbb{R}^n$. On pose, pour $x \in U$ :

$$(\nabla f)(x) = \left(\frac{\partial f}{\partial x_i}(x)\right)_i \in \mathbb{R}^n.$$

On dit que $(\nabla f)(x)$ est le **gradient de $f$ en $x$**.

*Remarque* 33. Autrement dit, le théorème précédent indique qu'un extremum local est nécessairement un point critique. Pour trouver les extrema locaux, on commence donc par chercher les points critiques de $f$. Attention, $f$ est définie sur un ouvert, si $f$ est définie sur un ensemble non ouvert $F$, $f$ peut avoir des extrema sur la frontière de $F$ qui ne sont pas des points critiques.

## 9.2 Conditions suffisantes du second ordre pour un extremum local

**Définition 35.** Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^2$ sur l'ouvert $U \subset \mathbb{R}^n$. On pose, pour $x \in U$ :

$$H(f)(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{ij} \in M_n(\mathbb{R}).$$

On dit que $H(f)(x)$ est la **matrice hessienne de $f$ en $x$**.

*Remarque* 34. Par le théorème de Schwarz, la matrice hessienne est symétrique (Une matrice $A = (a_{ij})$ est dite symétrique si $a_{ij} = a_{ji}$.)

*Rappel* 35. (Forme bilinéaire symétrique/positivité).

(i) Soit $M = (M_{ij})$ une matrice symétrique ($M = M^T$ M égale sa transposée $(M^T)_{ij} = M_{ji}$). On définit une application $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ en posant $B(H, K) = H^T M K$, où on voit $K, H \in \mathbb{R}^n \simeq M_{n,1}(\mathbb{R})$ comme matrice colonne (donc $H^T$ comme matrice ligne).

(ii) Concrètement, on a

$$B(H, K) = H^T [\sum_{j=1}^n M_{ij} K_j]_i = \sum_{i,j=1}^n H_i M_{ij} K_j.$$

$H \mapsto B(H, K)$ et $K \mapsto B(H, K)$ sont linéaires, on dit que $B$ est **bilinéaire**.

Par ailleurs, comme $M$ est supposée symétrique, on a

$$B(H,K) = \sum_{i,j=1}^{n} H_i M_{ij} K_j = \sum_{i,j=1}^{n} H_i M_{ji} K_j = B(K,H).$$

On dit que $B$ est **symétrique**

(iii) On pose $Q(H) = B(H,H) = \sum_{i,j=1}^{n} M_{ij} H_i H_j$. $Q : \mathbb{R}^n \to \mathbb{R}$ est une application polynomiale homogène de degré 2, on dit que $Q$ est une **forme quadratique**.

(iv) On admettra que toute forme bilinéaire symétrique vient d'une matrice symétrique $M$ comme ci-dessus. (Idée de preuve : identifier le deuxième $\mathbb{R}^n$ au dual de $E \simeq \mathbb{R}^n$ alors $B(H,.) \in E^* = \mathbb{R}^n$ on définit ainsi une application $H \to B(H,.)$ linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$, sa matrice est alors $M$ et on vérifie que la condition de symétrie de $B$ implique que $M$ est symétrique)

**Définition 36.** Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^2$ sur l'ouvert $U \subset \mathbb{R}^n$. On pose, pour $x \in U$, la **forme hessienne (ou différentielle seconde) de f en x)** est la forme bilinéaire symétrique suivante notée $D^2 f(x)$ :

$$(D^2 f(x))(H,K) = D_H D_K f(x) = \sum_{i,j=1}^{n} H_i \frac{\partial^2 f}{\partial x_i \partial x_j}(x) K_j$$

Elle vient de la matrice hessienne de $f$ comme précédemment.

**Définition 37.** Soit $A = (a_{ij}) \in M_n(\mathbb{R})$ une matrice symétrique. On rappelle que $B(H,K) = \sum_{i,j=1}^{n} H_i a_{ij} K_j$ est la forme bilinéaire associée à $A$. On dit que

1. $B$ est **positive** si $\forall H \in \mathbb{R}^n, B(H,H) \geq 0$
2. $B$ est **définie positive** si $\forall H \in \mathbb{R}^n - \{0\}, B(H,H) > 0$
3. $B$ est **négative** si $\forall H \in \mathbb{R}^n, B(H,H) \leq 0$
4. $B$ est **définie négative** si $\forall H \in \mathbb{R}^n - \{0\}, B(H,H) < 0$
5. $B$ est **non-dégénérée** si $det(A) \neq 0$ (c'est-à-dire $A$ inversible).

*Remarque* 36. Si $A = B^T B$, on a $B(H,H) = \sum_{ij} H_i \sum_k B_{ki} B_{kj} H_j = \sum_k \left(\sum_i B_{ki} H_i\right)^2 \geq 0$. Vous verrez en MASS 41 que toute matrice positive est de la forme $A = B^2$ avec $B = B^T$ dite racine carrée de $A$. Dire qu'une matrice est positive reviendra à dire que ses valeurs propres (les racines du polynôme dit caractéristique $\lambda \mapsto det(M - \lambda I_n)$) sont positives. On verra le cas $n = 2$ en détail plus bas.

**Théorème 61.** *Soient $f : U \to \mathbb{R}$ de classe $\mathcal{C}^2$ définie sur un ouvert $U$ et $a$ un point critique de $f$.*

1. *Si $D^2 f(a)$ définie positive, alors $a$ est un minimum local strict de $f$.*
2. *Si $D^2 f(a)$ définie négative, alors $a$ est un maximum local strict de $f$.*
3. *Si $a$ est un minimum local de $f$, alors $D^2 f(a)$ est positive.*
4. *Si $a$ est un maximum local de $f$, alors $D^2 f(a)$ est négative.*
5. *Si $D^2 f(a)$ n'est ni négative, ni positive alors $a$ n'est pas un extremum de $f$.*
6. *sinon (pas dans les cas (1),(2),(5), $D^2 f(a)$ n'est ni définie positive ni définie négative, ni à la fois non négative et non positive), le théorème ne permet pas de conclure.*

*Remarque* 37. Dans le dernier cas, on fait une étude directe.

*Démonstration.* D'après la formule de Taylor-Young, on a :

$$f(a+H)-f(a)=df(a)(H)+\frac{1}{2!}D^2f(a)(H,H)+||H||^2\epsilon(H),$$

avec $\epsilon(H)\to_{H\to 0}0$. Donc on a :

$$f(a+H)-f(a)=0+\frac{||H||^2}{2}\left(D^2f(a)(\frac{H}{||H||},\frac{H}{||H||})+2\epsilon(H)\right).$$

Le cas (2) se déduit du cas (1) en considérant $-f$ et de même le (4) du (3). Pour le (1), on suppose donc $D^2f(a)$ définie positive. $K\to D^2f(a)(K,K)$ est continue, donc elle atteint son minimum en $K_0$ sur le compact $S(0,1)$ (la sphère de centre 0 et de rayon 1 qui est un fermé du compact $B(0,1)$ donc compact.)

On déduit donc $\forall K\in S(0,1)$, $D^2f(a)(K,K)\geq D^2f(a)(K_0,K_0)=a_0>0$ d'après l'hypothèse (pour éviter la compacité on peut aussi montrer que $N(H)=\sqrt{D^2f(a)(H,H)}$ est une norme et utiliser l'équivalence des normes). Soit $\rho>0$ tel que pour tout $H\in B(0,\rho)$, $2|\epsilon(H)|\leq a_0/2$, on déduit :

$$f(a+H)-f(a)\geq\frac{a_0||H||^2}{4}.$$

d'où pour tout $H\neq 0$, $f(a+H)>f(a)$.

Considérons maintenant le cas (3). Si $D^2f(a)$ n'est pas positive, il existe $H\neq 0$ tel que

$D^2f(a)(H/||H||,H/||H||)=-a_0<0$, Ceci est encore vrai pour $H$ multiplié par $\lambda>0$. Pour $\lambda$ petit ($\forall\lambda\in[0,t_0]$, tel que $\epsilon(\lambda H)\leq a_0/4$), on trouve

$$f(a+\lambda H)-f(a)=\frac{\lambda^2||H||^2}{2}\left(D^2f(a)(\frac{H}{||H||},\frac{H}{||H||})+2\epsilon(\lambda H)\right)\leq\frac{a_0\lambda^2||H||^2}{4}<0$$

Donc $a$ n'est pas un minimum local (vu $a+\lambda H,\lambda\in]0,t_0]$ arbitrairement proche de $a$).

Le cas (5) est le résumé des cas (3),(4). Dans le cas (6), par exemple $f(x,y)=x^4+y^4$, ou $g(x,y)=x^3+y^3$ de hessiennes nulles en 0, on ne peut rien conclure a priori, $f$ a un minimum global strict alors que $g$ (cf restriction $g(x,0)$) n'a ni maximum ni minimum). □

**Corollaire 62.** *Soit $f:U\to I\!R$ de classe $\mathcal{C}^2$ définie sur un ouvert $U\subset I\!R^2$ et soit $a$ un point critique de $f$.*

*Soit $H(f)(a)=\begin{pmatrix} r & s\\ s & t\end{pmatrix}$ (c'est à dire on prend les notations de Monge $r=\frac{\partial^2 f}{\partial x^2}(a),s=\frac{\partial^2 f}{\partial x\partial y}(a),t=\frac{\partial^2 f}{\partial y^2}(a)$)*

1. *Si $rt-s^2>0$ et $r>0$, alors $a$ est un minimum local strict de $f$.*
2. *Si $rt-s^2>0$ et $r<0$, alors $a$ est un maximum local strict de $f$..*
3. *Si $rt-s^2<0$ alors $a$ n'est pas un extremum de $f$ (et c'est même un point selle cf sous-section 9.4).*
4. *Si $rt-s^2=0$ cas douteux.*

*Démonstration.* $D^2f(a)((h_1,h_2),(h_1,h_2))=rh_1^2+2sh_1h_2+th_2^2=(h_1^2)P(h_2/h_1)$ si $h_1\neq 0$, avec $P(\lambda)=r+2s\lambda+t\lambda^2$ le polynôme de second degré de discriminant $\Delta=4(s^2-4rt)$. Si $\Delta<0$ pas de racine et selon le signe de $r$, P est soit toujours positif (cas $D^2f(a)$ définie positive) soit toujours négative ($D^2f(a)$ définie négative).

Si $\Delta<0$, il a deux racines donc $P$ et $D^2f(a)$ ont des valeurs positives et négatives (donc $D^2f(a)$ n'est ni positive ni négative). Enfin, $\Delta=0$ on est dans le cas (6) du théorème. □

## 9.3 Convexité, Critère d'extremum global (Preuves facultatives)

**Définition 38.** Soit $C \subset \mathbb{R}^n$ un ensemble convexe, $f : C \to \mathbb{R}$ est une **fonction convexe** si l'ensemble $\mathrm{Epi}(f) = \{(x, y) \in C \times \mathbb{R} \ : \ y \geq f(x)\}$ (l'épigraphe de $f$) est convexe. $f : C \to \mathbb{R}$ est une **fonction concave** si $-f$ est convexe.

*Remarque* 38. Si $(a, y_1), (b, y_2) \in Epi(f)$, $[(a, y_1), (b, y_2)] \subset Epi(f) \iff \forall\lambda \in [0, 1]\ f(\lambda a+(1-\lambda)b) \leq \lambda y_1 + (1-\lambda)y_2$. on déduit donc si $f$ convexe en prenant $y_1 = f(a), y_2 = f(b)$ :

$$\forall a, b \in C \forall\lambda \in [0, 1], f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b).$$

Réciproquement cela caractérise les fonctions convexes car pour $y_1$ et $y_2$ comme ci-dessus $\lambda f(a) + (1-\lambda)f(b) \leq \lambda y_1 + (1-\lambda)y_2$.

En conséquence, une application linéaire est convexe et une somme d'applications convexes est convexe.

Une application est dite strictement convexe si de plus :

$$\forall a \neq b \in C \forall\lambda \in ]0, 1[, f(\lambda a + (1-\lambda)b) < \lambda f(a) + (1-\lambda)f(b).$$

On remarque de même que la somme d'une fonction convexe et d'une fonction strictement convexe est strictement convexe. Par exemple $X \mapsto ||X||_2^2$ est strictement comvexe.

**Proposition 63.** *Si $n = 1$ et $f$ est de classe $\mathcal{C}^1$ sur un ouvert convexe $U$ alors $f$ est convexe si et seulement si $f'$ est croissante.*

*Démonstration.* Supposons $f$ convexe. Soit $a < a + h < b < b + h$, montrons que $\frac{f(a+h)-f(a)}{h} \leq \frac{f(b+h)-f(b)}{h}$ (ce qui montre en prenant $h \to 0$ la croissance demandée). En découpant les intervalles, il suffit d'avoir pour $a < b < c$, $\frac{f(b)-f(a)}{b-a} \leq \frac{f(c)-f(b)}{c-b} (*)$

C'est une conséquence de $b = \frac{c-b}{c-a}a + c\frac{b-a}{c-a}$ d'où par l'inégalité de la remarque précédente caractérisant la convexité :

$$f(b) \leq \frac{c-b}{c-a}f(a) + f(c)\frac{b-a}{c-a} = f(a) + (f(c) - f(a))\frac{b-a}{c-a}$$

d'où $(*)$.

Réciproquement si $f'$ croissante, montrons que $f$ convexe, on veut voir $f(\lambda a + (1-\lambda)b) \leq \lambda f(a)+(1-\lambda)f(b)$ pour $a < b, \lambda \in ]0, 1[$. Par l'égalité des accroissements fins, la pente $\frac{f(\lambda a+(1-\lambda)b)-f(a)}{(1-\lambda)(b-a)}$ est atteinte par $f'$ en un point de $]a, \lambda a + (1-\lambda)b[$, et de même $\frac{f(b)-f(\lambda a+(1-\lambda)b)}{(\lambda(b-a)}$ est atteinte par $f'$ en un point de $]\lambda a + (1-\lambda)b, b[$ donc par croissance de la dérivée :

$$\frac{f(\lambda a + (1-\lambda)b) - f(a)}{(1-\lambda)(b-a)} \leq \frac{f(b) - f(\lambda a + (1-\lambda)b)}{\lambda(b-a)}$$

$$\iff f(\lambda a + (1-\lambda)b)(\frac{1}{(1-\lambda)(b-a)} + \frac{1}{\lambda(b-a)}) \leq \frac{f(a)}{(1-\lambda)(b-a)} + \frac{f(b)}{\lambda(b-a)}$$

$$\iff f(\lambda a + (1-\lambda)b)(\frac{1}{\lambda(1-\lambda)}) \leq \frac{f(a)}{(1-\lambda)} + \frac{f(b)}{\lambda}.$$

Ceci conclut. □

**Proposition 64.** *Si $f$ est de classe $\mathcal{C}^1$ sur un ouvert convexe $U$ et $f$ est convexe, alors tout point critique $a$ de $f$ est un minimum global de $f$.*

*Remarque* 39. Un minimum global qui est local strict est aussi global strict.

*Démonstration.* Soit $f$ convexe. En considérant, $t \mapsto f(a+t(b-a))$ qui est aussi convexe, il suffit du cas $n=1$. Comme la dérivée est croissante et s'annule en $a$, pour $b < a$ on a $f(a) = f(b) + \int_b^a f'(t)dt \leq f(b)$ vu $f'(t) \leq 0$ sur $[a,b]$ et de même pour $b > a$ d'où $a$ un un minimum global. $\square$

**Proposition 65.** *Si $f$ est de classe $\mathcal{C}^2$ sur un ouvert convexe $U$. $f$ est convexe si et seulement si $D^2 f(a)$ est positive pour tout $a$ de $C$.*

*Démonstration.* $g : x \mapsto f(x) - f(a) - df(a).(x-a)$ est convexe comme somme d'une application convexe et d'une application linéaire. Or $a$ est critique pour $g$ donc par la proposition, $a$ est un minimum global donc local, donc $D^2 f(a)$ est positive par le théorème 61.

Réciproquement, si $D^2 f(a)$ positive pour tout $a$. Soit $a, b \in C$ pour vérifier l'inégalité il suffit de vérifier que $\phi : t \mapsto f(a+t(b-a))$ est convexe. Or sa dérivée première par le théorème des fonctions composées est :$\phi'(t) = Df(a+t(b-a)).(b-a) = D_{b-a}f(a+t(b-a))$ et sa dérivée seconde est de même $\phi''(t) = D_{b-a}D_{b-a}f(a+t(b-a)) = D^2 f(a+t(b-a))(b-a, b-a) \geq 0$ par hypothèse. Donc $\phi''(t) \geq 0$ d'où $\phi'(t)$ croissante donc $\phi$ convexe. $\square$

## 9.4 Points selles (partie facultative)

Les points critiques $a$ qui ne sont pas des extrema peuvent être de différents types. L'absence d'extrema peut être visible sur une droite passant par $a$ s'il y a un point d'inflexion (comme pour $x \mapsto x^3$ dans $\mathbb{R}$) et il peut y avoir des points critiques qui sont des maxima dans certaines directions et des minima dans d'autres. Ces points ont un certain intérêt et seront nommés points selles.

**Définition 39.** Soit $U \subset \mathbb{R}^n$ et $f : U \to \mathbb{R}$ et $a \in U$.

1. Soient deux sous-espaces vectoriels $F$ et $G$ supplémentaires $\mathbb{R}^n = F \oplus G$ (c'est à dire $F \cap G = \{0\}$ et $\mathbb{R}^n = F + G$) On dit que $a$ est un *point selle (resp. point selle local)* de $f$ selon la décomposition $\mathbb{R}^n = F \oplus G$ si $a$ est un minimum (resp. minimum local) pour la restriction $f_{|a+F}$ de $f$ au sous espace affine $a+F$, et si $a$ est un maximum (resp. maximum local) pour la restriction $f_{|a+G}$ de $f$ au sous espace affine $a+G$. On parle de point selle si il existe une telle décomposition.
2. Si $f$ de classe $\mathcal{C}^1$. Soit $a$ un point critique de $f$, un sous espace vectoriel $H \subset \mathbb{R}^n$ est un plan d'inflexion si pour toute droite $\Delta$ passant par $a$ inclus dans $a+H$, $f_{|\Delta}$ n'a pas d'extrema local en $a$.

*Remarque* 40. La décomposition $F \oplus G$ d'un point selle n'est pas forcément unique et on ne demande rien en dehors $(a+G) \cup (F+a)$, en particulier, il peut y avoir des plans d'inflexion en un point selle (ex $f(x,y) = x^2 - y^2 + (x-y)^3$, $(0,0)$ est un point selle local dans la décomposition $(\mathbb{R}, 0) \oplus (0, \mathbb{R})$ car $x^2 + x^3$ a un minimum local en 0 et $-y^2 - y^3$ un maximum local, de même $(0,0)$ est un point selle dans la décomposition $\mathbb{R}(1, 1/2) \oplus \mathbb{R}(1/2, 1)$ mais $\mathbb{R}(1, -1)$ est une droite d'inflexion)

**Proposition 66.** *Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$*

1. *Si $a$ est un point selle de $f$, c'est un point critique de $f$.*
2. *Si $f$ est $\mathcal{C}^2$ et $a$ est un point critique de $f$. Si $D^2 f(a)$ est non-dégénérée, ni positive ni négative, alors $a$ est un point selle local de $f$.*

3. *Si $a$ est un point critique de $f$ $H$ est un plan d'inflexion en $a$ de dimension $dim(H) > n/2$ alors $a$ n'est pas un point selle local. De plus si $f$ est $\mathcal{C}^2$ pour tout $h \in H$, $D^2f(a)(H,H) = 0$.*

*Démonstration.* Pour (1) on remarque qu'il suffit de montrer $df(a) = 0$ ce qui ne dépend pas de la base de $\mathbb{R}^n$ on peut donc supposer $a$ point selle pour la décomposition $F = \mathbb{R}^k \times \{0\}$, $G = \{0\} \times \mathbb{R}^{n-k}$. Comme f restreint à $a+F$ à un minimum local, les $k$ premières dérivées partielles s'annulent, les n-k dernières s'annulent à cause du maximum sur $a+G$, d'où $df(a) = 0$.

La preuve de (2) nécessite MASS41. Pour (2), comme $D^2f(a)$ est non dégénérée, les valeurs propres de $H(f)(a)$ (les racines du polynôme $X \mapsto det(H(f)(a) - Xid))$ sont non nulles. Comme elle est ni positive ni négative, il y a à la fois des valeurs propres $\lambda$ positives et négatives. Soit $F$ l'espace vectoriel engendré par les vecteurs propres $u$ (les $u \in \mathbb{R}^n$ tels que $H(f)(a)u = \lambda u$ qui existent car si $det(H(f)(a) - \lambda id) = 0$, $H(f)(a) - \lambda id$ n'est pas injective donc à un noyau) des valeurs propres $\lambda$ strictement positives, et de même $G$ avec les négatives. $D^2f(a)$ restreint à $F$ est positive donc $f|_{a+F}$ admet un minimum local et de même pour $G$.

Pour (3), si $dim(H) > n/2$ et supposons par l'absurde $a$ point selle, on a $dim(F) + dim(G) = n$, on a soit $dim(F) \geq n/2$, soit $dim(G) \geq n/2$, disons qu'on se trouve dans le premier cas, alors $n \geq dim(H+F) = dim(F) + dim(H) - dim(F \cap H)$ implique $dim(F \cap H) \geq dim(F) + dim(H) - n > n/2 + n/2 - n = 0$ donc $F \cap H \neq \{0\}$ une contradiction car la restriction de $f$ à toute droite dans $a + F \cap H$ devrait avoir un minimum local en $a$ et un point d'inflexion à la fois. Si $D^2f(a)(H,H) \neq 0$, on a vu que cela suffit à ce que $f$ ait un extremum local sur la droite $a + \mathbb{R}H$, vu si $\phi(\lambda) = f(a + \lambda H)$, $\phi''(0) = D^2f(a)(H,H)$. □

**Théorème 67.** *Soient $A \subset \mathbb{R}^{n-k}, B \subset \mathbb{R}^k$ des compacts convexes et $K : C = A \times B \to \mathbb{R}$ continue. Si pour tout $(a,b) \in C, a \in \mathbb{R}^{n-k}, b \in \mathbb{R}^k$, $x \mapsto K(x,b)$ est convexe et $y \mapsto K(a,y)$ est concave, alors il existe un point de $C$ qui soit un point selle $(x_0, y_0)$ selon la décomposition $\mathbb{R}^{n-k} \times \{0\} \oplus \{0\} \times \mathbb{R}^k$ autrement dit :*

$$\forall x \in A, y \in B \quad K(x_0, y) \leq K(x_0, y_0) \leq K(x, y_0). \tag{2.1}$$

*De plus, (2.1) est équivalente à l'égalité :*

$$\mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B}K(x,y) = \mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K(x,y). \tag{2.2}$$

*Remarque* 41. On a des $Min$ et $Max$ au lieu d'inf et sup car des fonctions continues sur des compacts atteignent leur bornes (cf. la preuve pour la continuité de $x \mapsto \mathrm{Max}_{y\in B}K(x,y)$ et de façon similaire de $y \mapsto \mathrm{Max}_{x\in A}K(x,y)$.

Dans le cas où $f$ est bilinéaire, ce résultat s'appelle le théorème du min-max de von Neumann. Il a une signification en théorie des jeux. Si $f$ donne la valeur que gagne un joueur $A$ en position $x \in U$ si $f(x) \geq 0$ et $-f(x)$ la valeur que gagne le joueur $B$ (et perd le joueur A) si $f(x) \leq 0$. Si $A$ ne peut influencer que la direction $\{0\} \times \mathbb{R}^k$ et $B$ seulement la direction $\mathbb{R}^{n-k} \times \{0\}$. Alors un point selle est un "équilibre de Nash" c'est-à-dire un point où ni $A$ ni $B$ n'ont intérêt à changer leur stratégie, car si $A$ change sa stratégie celle de $B$ étant constante, étant donné que le point selle est un maximum, $A$ va perdre en gain, et de même si $B$ change sa position avec celle de $A$ constante, le caractère de minimum dans la direction du changement de $B$ montre que $B$ ne peut que perdre plus.

*Démonstration.*

- $\mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K(x,y) \leq \mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B}K(x,y)$ est toujours vrai. Comme pour tout $x \in A, y \in B$, $\mathrm{Min}_{x\in A}K(x,y) \leq K(x,y) \leq \mathrm{Max}_{y\in A}K(x,y)$, on déduit en prenant le max : $\mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K(x,y) \leq \mathrm{Max}_{y\in B}K(x,y)$ soit en prenant un $Min$ en $x$ :

$$\mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K(x,y) \leq \mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B}K(x,y).$$

- (2.1) $\Rightarrow$ (2.2)
  De plus, en considérant $(x_0, y_0)$ de (2.1), on a :

$$K(x_0, y_0) \leq \mathrm{Min}_{x\in A}K(x, y_0) \leq \mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K(x, y),$$
$$K(x_0, y_0) \geq \mathrm{Max}_{y\in B}K(x_0, y) \geq \mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B}K(x, y),$$

  d'où l'égalité complète en rassemblant les 3 dernières inégalités.
- $g : x \mapsto \mathrm{Max}_{y\in B}K(x, y)$ est continue.
  Soit $x, x_n \in A$, $x_n \to x$, soit $y_n$ (resp $t$) atteignant le *max* pour $x_n$ (resp $x$) c'est à dire : $\mathrm{Max}_{y\in B}K(x_n, y) = K(x_n, y_n)$. Supposons que $g(x_n) = K(x_n, y_n)$ ne converge pas vers $g(x)$. Par compacité, on peut extraire une suite telle que $y_{\phi(n)} \to Y$. Par continuité de $K$ :

$$g(x_{\phi(n)}) = K(x_{\phi(n)}, y_{\phi(n)}) \to K(x, Y) < K(x, t) = \mathrm{Max}_{y\in B}K(x, y) = g(x).$$

  Or $K(x_{\phi(n)}, t) \leq K(x_{\phi(n)}, y_{\phi(n)})$ donc en passant à la limite par continuité de $K$, $K(x, t) \leq K(x, Y) < K(x, t)$, une contradiction.
- (2.1) $\Leftarrow$ (2.2) On prend $x_0 \in A$ réalisant le minimum c'est à dire tel que :

$$\alpha = \mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B}K(x, y) = \mathrm{Max}_{y\in B}K(x_0, y)$$

  Il existe par la continuité du point précédent et par compacité. De même, il existe $y_0 \in B$ réalisant le maximum :

$$\mathrm{Min}_{x\in A}K(x, y_0) = \mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K(x, y) = \alpha.$$

  Donc pour tout $x \in A, y \in B$, en utilisant (2.2) pour l'égalité du milieu, on obtient :

$$K(x_0, y) \leq \mathrm{Max}_{Y\in B}K(x_0, Y) = \alpha = \mathrm{Min}_{X\in A}K(X, y_0) \leq K(x, y_0).$$

  En prenant $x = x_0$, $y = y_0$, on voit $\alpha = K(x_0, y_0)$, ce qui dit donc que $(x_0, y_0)$ est un point selle.
- Montrons (2.2). Considérons, pour $\epsilon > 0$,

$$K_\epsilon(x, y) = K(x, y) + \epsilon||x||_2^2.$$

  Comme $x \mapsto \epsilon||x||_2^2$ est strictement convexe, il en est de même de $K_\epsilon(., y)$ pour tout $y \in B$ (convexe plus strictement convexe donne strictement convexe).
  Montrons que pour tout $y$, la fonction $K_\epsilon(., y)$ a un unique minimum. En effet, si $x_1 \neq x_2$ sont deux minima, par stricte convexité : $K_\epsilon((x+y)/2, y) < K_\epsilon(x_1, y)/2 + K_\epsilon(x_2, y)/2 = K(x_i, y)$ en contradiction avec le caractère de minimum. Donc on a un unique $E(y)$ atteignant le minimum de $K_\epsilon(., y)$ Par le deuxième point (appliqué à $-K_\epsilon(y, x)$) $f_\epsilon(y) = K_\epsilon(E(y), y)$ est continue, donc atteint son maximum en $y^*$. En conséquence, par la définition de $f_\epsilon$ et le choix de $y^*$

$$f_\epsilon(y^*) = \mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A}K_\epsilon(x, y) = K_\epsilon(E(y^*), y^*) = \mathrm{Min}_{x\in A}K_\epsilon(x, y^*).$$

  Soit $x \in A, y \in B, t \in ]0, 1[$, on a par concavité :

$$K_\epsilon(x, (1-t)y^* + ty) \geq (1-t)K_\epsilon(x, y^*) + tK_\epsilon(x, y) \geq (1-t)f_\epsilon(y^*) + tK_\epsilon(x, y).$$

En prenant $x = E((1-t)y^* + ty)$, on obtient $f_\epsilon((1-t)y^* + ty) \geq (1-t)f_\epsilon(y^*) + tK_\epsilon(E((1-t)y^* + ty), y)$.
Vu que $y^*$ maximise $f_\epsilon$, en soustrayant et divisant par $t$, on a :

$$f_\epsilon(y^*) \geq K_\epsilon(E((1-t)y^* + ty), y) \quad (*).$$

On veut prendre $t \to 0$, voyons que $y \mapsto E(y)$ est continue. Supposons $y_n \to y$, et supposons $E(y_n) \not\to E(y)$ par compacité, on a une suite extraite $y_{\phi(n)}$ telle que $E(y_{\phi(n)}) \to Z \neq E(y)$. Par continuité $K_\epsilon(E(y_{\phi(n)})), y_{\phi(n)}) \to K_\epsilon(Z, y) > K_\epsilon(E(y), y)$,
l'inégalité stricte venant de l'unicité du minimum d'une fonction strictement convexe.
Or par définition $K_\epsilon(E(y)), y_{\phi(n)}) \geq K_\epsilon(E(y_{\phi(n)})), y_{\phi(n)})$ donc en passant à la limite $K_\epsilon(E(y), y) \geq K_\epsilon(Z, y) > K_\epsilon(E(y), y)$, une contradiction.
On a donc montré la continuité de $y \mapsto E(y)$.
Donc en passant à la limite dans l'inégalité $(*)$, on obtient : $f_\epsilon(y^*) \geq K_\epsilon(E(y^*), y)$ et ce pour tout $y \in B$ Par ailleurs par définition de $f_\epsilon$, $f_\epsilon(y^*) \leq K_\epsilon(x, y^*)$. Autrement dit $(E(y^*), y^*)$ est un point selle de $K_\epsilon$. Par l'implication (2.1) $\Rightarrow$ (2.2), on déduit, vu $K(x,y) \leq K_\epsilon(x,y) \leq K(x,y) + \epsilon D$ (avec $D = Max_{x \in A} ||x||_2^2 < \infty$ par compacité) :

$$\begin{aligned}\mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B} K(x,y) &\leq \mathrm{Min}_{x\in A}\mathrm{Max}_{y\in B} K_\epsilon(x,y) \\ &= \mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A} K_\epsilon(x,y) \leq \epsilon C + \mathrm{Max}_{y\in B}\mathrm{Min}_{x\in A} K(x,y).\end{aligned}$$

En prenant $\epsilon \to 0$, on obtient l'inégalité qui manque pour avoir (2.2) pour $K$.

$\square$

# Chapitre 3

# Notions de géométrie différentielle, Extrema liés

## 1 Formes différentielles

Cette partie inscrit la notion de différentielle d'une fonction dans le cadre plus générale des formes différentielles.

### 1.1 Rappels d'algèbre linéaire et notation différentielle

Ce paragraphe complète la remarque 25.

Soit l'espace vectoriel $E = \mathbb{R}^n$. On rappelle que l'on note $e_i = (0, ..., 0, 1, 0, ..., 0) \in E$ où le 1 est en i-ème position.

1. (base canonique) Si $X = (x_1, ..., x_n) \in E$n on a $X = x_1 e_1 + ... + x_n e_n$ (écriture unique). $(e_1, ..., e_n)$ est la base canonique de $E$.
2. (forme linéaire base duale) On note $(\mathbb{R}^n)^* = \mathcal{L}(R^n, R)$ l'espace vectoriel des formes linéaires sur $\mathbb{R}^n$ (c'est à dire des application linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$.
   Si $\alpha \in (\mathbb{R}^n)^* = E^*$, on note $(a_1, ..., a_n)$ la matrice (matrice ligne) de $\alpha$ par rapport aux bases canoniques de $\mathbb{R}^n$ et de $\mathbb{R}$. On a $\alpha(x_1, ..., x_n) = a_1 x_1 + ... + a_n x_n$.
   En algèbre linéaire on note $e_i^*$ la forme linéaire de matrice ligne $(0, ..., 0, 1, 0, ..., 0) \in E^*$ où le 1 est en i-ème position. On a alors $\alpha = a_1 e_1^* + ... + a_n e_n^*$. $(e_1^*, ..., e_n^*)$ est une base de $E^*$, on dit que c'est la **base duale de la base canonique** de $\mathbb{R}^n$. On remarque que $e_i^*(x_1, ..., x_n) = x_i$ n'est rien d'autre que la projection sur la $i$-ème coordonnée (ailleurs notée $p_i$ ou $x_i$)
3. (base duale en notation différentielle) Étant donné que la fonction $x_i = p_i = e_i^*$ est linéaire, on la note aussi de façon cohérente $dx_i : \mathbb{R}^n \to \mathbb{R}$. On a donc $dx_i(h_1, ..., h_n) = h_i$. Par suite, on a :
   $$\alpha(h_1, ..., h_n) = a_1 h_1 + ... + a_n h_n = a_1 dx_1(h_1, ..., h_n) + ... + a_n dx_n(h_1, ..., h_n) = (a_1 dx_1 + ... + a_n dx_n)(h_1, ..., h_n)$$
   Ceci montre la relation mentionnée ci-dessus :$\alpha = a_1 dx_1 + ... + a_n dx_n$. On remarque que $dx_i(e_i) = 1$ et $dx_i(e_j) = 0$ si $j \neq i$.

### 1.2 Définition des formes différentielles, différentielles totales, formes exactes et fermées

**Définition 40.** Soit $U$ un ouvert de $(\mathbb{R}^n)$. Une application $\omega : U \to (\mathbb{R}^n)^*$ est appelée forme différentielle sur $U$. Elle est donnée par $n$ applications $\omega_1, ..., \omega_n : U \to \mathbb{R}$ tel que $\omega(x) = \omega_1(x)dx_1 + ... + \omega_n(x)dx_n$.

*Remarque* 42. Pour tout x $\omega(x) : \mathbb{R}^n \to \mathbb{R}$ est linéaire. Si $H = (h_1, ..., h_n), \omega(x).H = \omega_1(x)h_1 + ... + \omega_n(x)h_n$.

*Exemple* 27. Si $f$ est différentiable sur $U$, alors $x \mapsto df(x)$ est une forme différentielle. On verra en TD que toute forme différentielle n'est pas de cette forme. On note que $df(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)dx_i$.

**Définition 41.** Soit $f : U \to \mathbb{R}$ une application de classe $\mathcal{C}^1$ sur l'ouvert $U \subset \mathbb{R}^n$. On appelle *différentielle totale* de $f$ la forme différentielle $df$. Une forme différentielle $\omega$ sur $U$ est dite *exacte* si $\omega = df$ pour une application $\mathcal{C}^1$. On dit alors que $f$ est une *primitive* de $\omega$.

**Définition 42.** Une forme différentielle $\omega$ sur $U$ de classe $\mathcal{C}^1$ est dite *fermée* si :

$$\forall i, j \in [1, n], \forall x \in U \quad \frac{\partial \omega_j}{\partial x_i}(x) = \frac{\partial \omega_i}{\partial x_j}(x).$$

*Remarque* 43. Le Théorème de Schwarz implique que toute forme différentielle exacte de classe $\mathcal{C}^1$ est fermée. On verra en TD que la réciproque est fausse en général. Vous verrez en MASS 41 une condition sur $U$ pour avoir la réciproque (Théorème de Poincaré), par exemple $U$ convexe suffit.

# 2 Difféomorphismes

**Définition 43.** Soient $U \subset \mathbb{R}^n, V \subset \mathbb{R}^p$. Une application $f : U \to V$ une fonction différentiable. $f$ est un *difféomorphisme* si $f$ est bijective et que $f^{-1}$ est différentiable.

On dit que $f$ est un $\mathcal{C}^k$*-difféomorphisme* ($k \in \mathbb{N}^* \cup \infty$) si de plus $f$ et $f^{-1}$ sont de classe $\mathcal{C}^k$.

**Proposition 68.** *Soit $f : U \to V$ un difféomorphisme, alors $\forall x \in U$, $df(x) : \mathbb{R}^n \to \mathbb{R}^p$ est un isomorphisme linéaire (en particulier nécessairement $n = p$) et on a :*

$$(df(x))^{-1} = df^{-1}(f(x)).$$

*Remarque* 44.
1. Le résultat précédent montre que la dimension est invariante par difféomorphisme. De même des ouverts de $\mathbb{R}^n$ et $\mathbb{R}^p$ ne peuvent être homéomorphes que si $n = p$ mais c'est beaucoup plus dur (Théorème d'invariance du domaine de Brouwer). Par contre, il existe des applications continues surjectives de $[0, 1]$ dans $[0, 1]^2$.
2. Le théorème d'inversion locale va donner des conditions pour la réciproque de la proposition précédente

*Démonstration.* Comme $f^{-1} \circ f(y) = y$, en différenciant $f^{-1} \circ f$ par le théorème des fonctions composées en $x$, on obtient : $df^{-1}(f(x)) \circ df(x) = id$.

De même en différenciant $f \circ f^{-1}(y) = y$ en $z = f(x)$ on obtient : $df(f^{-1}(z)) \circ df^{-1}(z) = Id$. Donc $df(x)$ et $df^{-1}(f(x))$ sont inverses l'un de l'autre, ce qui conclut. $\square$

**Définition 44.** Soit $f : U \to \mathbb{R}^p$ une application différentiable sur un ouvert $U \subset R^n$. $f(x) = (f_1(x), ..., f_p(x))$. La matrice de l'application linéaire $df(x)$ dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ est appelée, *matrice jacobienne* de $f$ et notée $J(f)(x)$ :

$$(J(f)(x))_{ij} = (\frac{\partial f_i}{\partial x_j}(x)).$$

*Remarque* 45. Le théorème de dérivation des fonctions composées donne donc :

$$J(g \circ f)(x_0) = J(g)(f(x_0))J(f)(x_0),$$

et le résultat pour les inverses de la proposition précédente s'écrit :

$$J(f^{-1})(y_0) = [J(f)(f^{-1}(y_0))]^{-1}.$$

## 2.1 Théorème d'inversion locale

**Théorème 69.** ***(d'inversion locale)*** *Soit $f : U \to \mathbb{R}^n$ une fonction de classe $\mathcal{C}^k$ $(k \in \mathbb{N}^* \cup \infty)$ définie sur un ouvert $U$ de $\mathbb{R}^n$. Soit $x_0 \in U$. On suppose que $df(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ est un isomorphisme linéaire. Alors il existe une boule ouverte non vide $V = B(x_0, r)$ et un ouvert $W \subset \mathbb{R}^n$ tel que $f_{|V} : V \to W$ soit un $\mathcal{C}^k$-difféomorphisme.*

*Remarque* 46. $df(x_0)$ est un isomorphisme si et seulement si $det(Jf(x_0)) \neq 0$.

*Démonstration.* On munit $\mathbb{R}^n$ de la norme $||.||_1$.

- On considère $F = df(x_0)^{-1} \circ f$. On a donc par le théorème de différentiation des fonctions composées (et la différentielle des applications linéaires) $dF(x_0) = d(df(x_0)^{-1})(f(x_0)) \circ df(x_0) = df(x_0)^{-1} \circ df(x_0) = Id$. Il suffit de montrer le théorème pour $F$.
- Choisissons $r$ tel que $B(x_0, 2r) \subset U$ et vu la continuité de la différentielle telle que $|||dF(x) - Id||| := \sum_j \max_i |\frac{\partial F_j}{\partial x_i}(x) - 1_{\{i=j\}}| \leq 1/2n$ pour tout $x \in B(x_0, 2r)$.
  Cela montre en particulier que $dF(x)$ est inversible pour tout $x \in B(x_0, 2r)$. En effet considérons $T_n = \sum_{k=0}^n (Id - dF(x))^k$ Vu $|||AB||| \leq |||A|||\ |||B|||$ (cf TD 1 ex 3), on a $|||T_n - T_m||| \leq \sum_{k=m+1}^n |||Id - dF(x)|||^k \leq \sum_{k=m+1}^\infty 1/(2^{m+1}) = 1/2^m \to 0$ pour $n \geq m$. Cela montre en particulier que $T_n$ est de Cauchy donc converge disons vers $T$ (complétude en dimension finie). De plus $T_n dF(x) = \sum_{k=0}^n (Id - dF(x))^k (Id - (Id - dF(x))) = \sum_{k=0}^n (Id - dF(x))^k - \sum_{k=1}^{n+1} (Id - dF(x))^k = Id - (Id - dF(x))^{k+1}$ d'où $|||T_n dF(x) - Id||| \leq 1/2^{n+1} \to 0$ d'où en passant à la limite $T dF(x) = Id$. De me on voit $dF(x)T = Id$ d'où $dF(x)$ est inversible pour tout $x \in B(x_0, 2r)$ comme annoncé.
- D'après l'inégalité des accroissements finis

$$\begin{aligned}
\forall (u,v) \in B(x_0, 2r),\ ||F(v) - F(u) - (v-u)||_1 &= \sum_j |F_j(v) - F_j(u) - (v_j - u_j)| \\
&\leq n \max_j \max_i \sup_{x \in B(x_0,2r)} |\frac{\partial F_j}{\partial x_i}(x) - 1_{\{i=j\}}|||v-u||_1 \\
&\leq \sup_{x \in B(x_0,2r)} n \max_j \max_i |\frac{\partial F_j}{\partial x_i}(x) - 1_{\{i=j\}}|\ ||v-u||_1 \\
&\leq \sup_{x \in B(x_0,2r)} n \sum_j \max_i |\frac{\partial F_j}{\partial x_i}(x) - 1_{\{i=j\}}|\ ||v-u||_1 \\
&\leq \frac{1}{2}||v-u||_1 \qquad (*).
\end{aligned}$$

En particulier $F$ est injective sur $B(x_0, 2r)$ car $F(x) = F(y)$ implique $||v-u||_1 \leq \frac{1}{2}||v-u||_1$ ce qui n'est possible par séparation que si $v = u$.

- Soit $y_0 = F(x_0)$ Montrons que $B_F(y_0, r/2) \subset F(B_F(x_0, r))$. Soit $y \in B_F(y_0, r/2)$.
On construit par la méthode des itérations de Picard, un point fixe de l'application $G : z \mapsto z + y - F(z)$, c'est à dire un $X$ tel que $G(X) = X \Leftrightarrow F(X) = y$, ce qui montrera le résultat.
On définit par récurrence la suite (débutant à $x_0$) $x_{k+1} = x_k + y - F(x_k)$.
Montrons par récurrence que pour tout $k$, $||x_{k+1} - x_0|| \leq r$ et $||x_{k+1} - x_k||_1 \leq (||x_1 - x_0||_1)/2^k$.
On remarque que $x_{k+1} - x_k = y - F(x_k) = (x_k - x_{k-1} + F(x_{k-1})) - F(x_k)$ En utilisant (*) (applicable car $x_k, x_{k-1} \in B(x_0, r)$ par l'hypothèse de récurrence), on déduit

$$||x_{k+1} - x_k||_1 = ||F(x_k) - F(x_{k-1}) - (x_k - x_{k-1})||_1 \leq ||x_k - x_{k-1}||_1/2 \leq (||x_1 - x_0||_1)/2^k.$$

En sommant ces inégalités, on trouve :

$$||x_{k+1} - x_0||_1 \leq \sum_{j=0}^{k} ||x_{j+1} - x_j||_1 \leq \sum_{j=0}^{k} \frac{1}{2^j}(||x_1 - x_0||_1) \leq 2r/2 = r.$$

Ceci conclut la récurrence.
- Montrons que $x_k$ est de Cauchy, en effet pour $p \geq q$ :

$$||x_p - x_q||_1 \leq \sum_{j=q}^{p-1} ||x_{j+1} - x_j||_1 \leq \sum_{j=q}^{\infty} \frac{1}{2^j}(||x_1 - x_0||_1) \leq r/2^q \to_{q\to\infty} 0.$$

(On a utilisé la somme de la série géométrique) La suite $x_k$ est donc convergente vers $X$ et par continuité de $F$, on déduit $X = X + y - F(X)$, ce que l'on voulait.
On a donc montré que $F$ est une bijection de $V = F^{-1}(B(y_0, r/2)) \subset B_F(x_0, r)$ sur $W = B(y_0, r/2)$. Il reste à voir que $f^{-1}$ est $\mathcal{C}^k$. Commençons par le cas $k = 1$. On note $G$ l'inverse de $F$.
- Soit $y \in W$, $x = G(y)$ et $l \in R^n$. On note $h(l) = G(y+l) - G(y)$. En utilisant la différentiabilité de $f$ :

$$y + l = F(G(y + l)) = F(x + h(l)) = F(x) + dF(x).h(l) + o(||h(l)||),$$

d'où $l = F(x + h(l)) - F(x) = dF(x).h(l) + ||h(l)||\epsilon(h(l))$. En utilisant (*), on obtient

$$||h(l) - l|| = ||F(x + h(l)) - F(x) + (x + h(l) - x)|| \leq ||h(l)||/2.$$

En utilisant l'inégalité triangulaire inverse, $|||h(l)|| - ||l||| \leq ||h(l) - l|| \leq ||h(l)||/2$, on déduit $||h(l)||/2 \leq ||l|| \leq 3||h(l)||/2$.
Enfin en utilisant l'inversibilité de $df(x)$ pour $x \in B(x_0, 2r)$, on obtient :

$$G(y + l) - G(y) = h(l) = dF(x))^{-1}(l) + dF(x))^{-1}(||h(l)||\epsilon(h(l)))$$

et

$$||dF(x))^{-1}(||h(l)||\epsilon(h(l)))||_1 \leq 2||l||\ |||dF(x))^{-1}(l)|||\ ||\epsilon(h(l))||_1 = o(||l||).$$

En conséquence $g$ est différentiable de différentielle $dG(y) = dF(G(y))^{-1}$. Si $F$ est $\mathcal{C}^k$, on montre que $G$ est $\mathcal{C}^k$ par récurrence sur $k$ en utilisant la formule de compositions (si G et dF $\mathcal{C}^{k-1}$ par l'hypothèse de récurrence et l'hypothèse, il en va de même de dG, donc $G$ est $\mathcal{C}^k$.)

$\square$

**Définition 45.** Soit $f : U \to \mathbb{R}^p$ une fonction différentiable sur $U$. On dit que $f$ est un difféomorphisme local (resp. $\mathcal{C}^k$-difféomorphisme local) si pour tout $x \in U$ il existe un ouvert $V \subset U$ $x \in V$ et $W$ ouvert de $\mathbb{R}^p$ tel que $f_{|V} : V \to W$ soit un difféomorphisme (resp. $\mathcal{C}^k$-difféomorphisme).

**Définition 46.** Soit $f : U \to \mathbb{R}^p$ une fonction de classe $\mathcal{C}^1$ définie sur un ouvert $U$. On dit que $f$ est *étale* si $\forall x \in U$ $df(x)$ est un isomorphisme.

**Corollaire 70.** *Soit $f : U \to \mathbb{R}^n$ une fonction de classe $\mathcal{C}^k$. Alors $f$ est un $\mathcal{C}^k$-difféomorphisme local si et seulement sif est étale (c'est à dire pour tout $x$ dans $U$, $df(x) : \mathbb{R}^n \to \mathbb{R}^n$ est un isomorphisme linéaire).*

*Exemple* 28. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(\rho, \theta) = (\rho \cos(\theta), \rho \sin(\theta))$. $f$ est $\mathcal{C}^\infty$ et

$$Jf(\rho, \theta) = \begin{pmatrix} \cos(\theta) & -\rho \sin(\theta) \\ \sin(\theta) & \rho \cos(\theta) \end{pmatrix}.$$

$det(Jf(\rho, \theta)) = \rho(\sin^2(\theta) + \cos^2(\theta)) = \rho.$

Donc $f$ est un difféomorphisme local sur $\mathbb{R}^* \times \mathbb{R}$.

Ce n'est pas un difféomorphisme global car $f$ n'est pas injectif $f(\rho, \theta) = f(\rho, \theta + 2k\pi) = f(-\rho, \theta + \pi)$.

*Exemple* 29.

- Soit $E = \{P(X) = X^n + \sum_{i=0}^{n-1} y_i X^i\} \simeq \mathbb{R}^n$ l'ensemble des polynômes unitaires. Pour $Y = (y_0, ..., y_{n-1}) \in \mathbb{R}^n$, on notera $P_Y = X^n + \sum_{i=0}^{n-1} y_i X^i$.
- Soit $\varphi : \mathbb{R}^n \to E$ l'application définie par

$$\varphi(x_1, ..., x_n) = \prod_{i=1}^{n} (X - x_i).$$

C'est une application de classe $\mathcal{C}^\infty$ de dérivées partielles :
$\frac{\partial \varphi}{\partial x_i} = -\prod^n_{\substack{j=1 \\ j \neq i}} (X - x_j).$

Le déterminant de $d\varphi(x)$ est nul si et seulement si les vecteurs $\frac{\partial \varphi}{\partial x_i}$ ne sont pas linéairement indépendants. Si deux des coordonnées $x_i$ sont égales, deux des dérivées partielles sont égales et donc le déterminant est zéro. Sinon, si tous les $x_i$ sont deux à deux distincts, rappelons pourquoi les polynômes $\frac{\partial \varphi}{\partial x_i}$, qui sont des polynômes d'interpolations de Lagrange, sont linéairement indépendants. En effet, si $\sum \lambda_i \frac{\partial \varphi}{\partial x_i} = 0 (*)$, en évaluant le polynôme en $x_i$, on trouve $\frac{\partial \varphi}{\partial x_j}(x_i) = 0$ si $j \neq i$ et

$$\frac{\partial \varphi}{\partial x_i}(x_i) = -\prod^n_{\substack{j=1 \\ j \neq i}} (x_i - x_j) \neq 0,$$

d'où la relation $(*)$ implique $\lambda_i \prod^n_{\substack{j=1 \\ j \neq i}} (x_i - x_j) = 0$, c'est à dire $\lambda_i = 0$ pour tout $i$, donc les $\frac{\partial \varphi}{\partial x_j}$ sont linéairement indépendants. En bilan, $\varphi$ est étale en $(x_1, ..., x_n)$ si et seulement si tous les $x_i$ sont disjoints.

- Si $a = (a_i) \in \mathbb{R}^n$ est tel que les $a_i$ sont deux à deux distincts et si $c = (c_i) \in \mathbb{R}^n$ tel que

$$\varphi(a_1, ..., a_n) = P_c,$$

le théorème d'inversion locale montre que $\varphi$ est un $C^\infty$-difféomorphisme sur un voisinage $V$ de $a$, ce qui fournit un voisinage $U$ de $c$ tel que racine $a_i$ de $P_x$ sont des fonctions $\mathcal{C}^\infty$ des coefficients du polynômes $x \in U$.. On dit que les racines d'un polynôme sont des fonctions $\mathbb{C}^\infty$ des coefficients au voisinage des racines simples. On se souviendra qu'il n'y a pas de formule explicite pour les racines en degré supérieur ou égal à 5.

## 2.2 Théorème d'inversion globale

**Définition 47.** Soit $f : U \to \mathbb{R}^p$ une fonction définie sur un ouvert $U$. On dit que $f$ est *ouverte* si pour tout ouvert $V \subset U$, $f(V)$ est un ouvert de $\mathbb{R}^p$.

**Proposition 71.** *Une application étale est ouverte.*

*Démonstration.* Soit $f : U \to \mathbb{R}^n$ étale. Par le théorème d'inversion locale, pour tout $x \in U$, il existe $r_x > 0$ tel que $f$ restreint à $B(x, r_x)$ et un $\mathcal{C}^1$-difféomorphisme, en particulier un homéomorphisme, donc en particulier est ouverte.

Comme $V = \bigcup_{x \in V} B(x, min(r_x, d(x, V^c)/2))$, vu que $V$ est ouvert et donc $V^c$ fermé, d'où $B(x, d(x, V^c)/2) \subset V$, on a :

$$f(V) = \bigcup_{x \in V} f(B(x, min(r_x, d(x, V^c)/2))].$$

Donc $f(V)$ est ouvert comme union d'ouverts. □

**Théorème 72.** ***(d'inversion globale)*** *Soit $f : U \to \mathbb{R}^n$ une application de classe $\mathcal{C}^k$ étale et injective, alors $f : U \to f(U)$ est un $\mathcal{C}^k$-difféomorphisme.*

*Démonstration.* Par la proposition précédente $f(U)$ est ouvert, donc $f$ est une bijection entre $U$ et $f(U)$, comme localement l'inverse de $f$ est $\mathcal{C}^k$, vu $f$ $\mathcal{C}^k$-difféomorphisme local par le théorème d'inversion local, on déduit que $f^{-1}$ est $\mathcal{C}^k$. □

*Remarque* 47. Comme l'inverse n'a pas de formule explicite en général, on peut calculer sa différentielle en utilisant la formule $df^{-1}(y) = (df(f^{-1}(y)))^{-1}$.

*Exemple* 30. Soit $A = (a, 0), B = (b, 0) \in \mathbb{R} \times \{0\} \subset \mathbb{R}^2, A \neq B$. Soit $\phi : \mathbb{H} = \mathbb{R} \times \mathbb{R}_+^* \to \mathbb{R}^2$ l'application définie par

$$\phi(x, y) = (||(x - a, y)||_2, ||(x - b, y)||_2).$$

$\phi$ est de classe $\mathcal{C}^\infty$ (vu $y \neq 0$ sur son domaine, et par composition vu $x \mapsto \sqrt{x}$ $\mathcal{C}^\infty$ sur $\mathbb{R}_+^*$). Sa différentielle est

$$d\phi(X).H = \left( \frac{\langle (x - a, y), H \rangle}{||(x - a, y)||_2}, \frac{\langle (x - b, y), H \rangle}{||(x - b, y)||_2} \right).$$

Comme les vecteurs $(x - a, y)$ et $(x - b, y)$ sont linéairement indépendants (vu $a \neq b, y \neq 0$), $d\phi(X)$ est bijective.

La caractérisation des triangles par les longueurs des côtés montre que $\phi$ est injective. Son image est $\phi(\mathbb{H}) = \{(u, v) \mid |u - v| < |a - b| < u + v\}$. Donc $\phi$ est un $\mathcal{C}^\infty$-difféomorphisme de $\mathbb{H}$ sur $\phi(\mathbb{H})$. Les images des ellipses et hyperboles de foyers $A$ et $B$ sont les droites $u \pm v = cst$.

# 3 Théorème des fonctions implicites

## 3.1 Nature du problème

Soit $f : U \to \mathbb{R}$ une fonction définie sur un ouvert $U$ de $\mathbb{R}^2$ et de classe $\mathcal{C}^1$. Si $a \in \mathbb{R}$, on considère la courbe de niveau pour $f$ de valeur $a$ : $C_a = \{(x,y) \in U \mid f(x,y) = a\}$. En général, si $C_a$ est non vide, $C_a$ est une courbe contenue dans $U$.

*Exemple* 31. Si $f(x,y) = x^2 + y^2$, $C_{R^2}$ est un cercle de centre 0 et de rayon $R$.

Supposons que la courbe admette une tangente non verticale en un point $X_0 = (x_0, y_0)$, Au voisinage de $X_0$, $C_a$ est le graphe d'une fonction ($\mathcal{C}^1$) $\varphi$. Autrement dit :

$\exists \alpha > 0, \exists \varphi : ]x_0 - \alpha, x_0 + \alpha[ \to \mathbb{R}$ tels que $\forall x \in ]x_0 - \alpha, x_0 + \alpha[$ on a $(x, \phi(x)) \in C_a$. On a donc $\forall x \in ]x_0 - \alpha, x_0 + \alpha[, \quad f(x, \phi(x)) = a$ et de plus il existe $\beta > 0$ tel que $\phi(x)$ est l'unique solution de $f(x,y) = a$ avec $y \in ]y_0 - \beta, y_0 + \beta[$.

Une telle fonction $\varphi$ s'appelle une "fonction implicite" (ou plus précisément une fonction définie implicitement par l'équation $f(x, \phi(x)) = a$).

*Remarque* 48. Si $f$ admet une tangente verticale en $X_1 = (x_1, y_1)$, on voit qu'une telle fonction implicite $\varphi(x)$ n'existe pas. Par contre au voisinage de $y_1$ il existe une fonction $\psi(y)$ telle que $f(\psi(y), y) = a$ $\psi$ est aussi une fonction implicite.

Obtenir $\phi$ permet de passer d'une courbe définie par une équation à une courbe définie par un paramétrage (ici une application injective de $\mathbb{R}$ vers $\mathbb{R}^2$). En considérant l'inverse de $\varphi$ (définie sur la courbe $C_a \cap ]x_0 - \alpha, x_0 + \alpha[ \times ]x_0 - \beta, y_0 + \beta[ = \varphi(]x_0 - \alpha, x_0 + \alpha[)$), on obtient une image de la courbe dans $\mathbb{R}$, ce que l'on appelle une "carte". Si on construit différentes cartes de cette façon, $\phi_1^{-1}$, $\phi_2^{-1}$, alors le changement de carte $\phi_1^{-1} \circ \phi_2 : \mathbb{R} \to \mathbb{R}$ peut être montré être un $\mathcal{C}^1$ difféomorphisme (en utilisant le théorème d'inversion locale).

Le passage et l'exploitation de ces différentes représentations, par paramétrage, par carte, par équation est la base de la géométrie différentielle.

Pour notre problème de définition de fonctions implicites, le théorème des fonctions implicites va rendre compte de cette situation dans un cadre plus général.

## 3.2 Énoncé du théorème (preuve facultative)

**Théorème 73.** *Soient $U \subset \mathbb{R}^n \times \mathbb{R}^q$ un ouvert et $f : U \to \mathbb{R}^q$ une fonction de classe $\mathcal{C}^k$ ($k \in \mathbb{N}^* \cup \{\infty\}$). Soit $(a,b) \in U$ tel que $f(a,b) = 0$.*

*Si $df(a,b)_{|\{0\} \times \mathbb{R}^q} : \{0\} \times \mathbb{R}^q \to \mathbb{R}^q$ est un* ***isomorphisme*** *linéaire, alors il existe $r, s > 0$ tels que $B(a,r) \times B(b,s) \subset U$ et une application de classe $\mathcal{C}^k$ $\varphi : B(a,r) \to \mathbb{R}^q$ telle que :*

*$\forall x \in B(a,r)$, le vecteur $\varphi(x)$* ***est l'unique*** *élément de $B(b,s)$* ***solution*** *de :*

$$f(x, \varphi(x)) = 0.$$

*En particulier on a $\varphi(a) = b$ et de plus $\forall (x,y) \in B(a,r) \times B(b,s)$, $df(x,y)_{|\{O\} \times \mathbb{R}^q}$ est un isomorphisme.*

*Démonstration.* C'est une conséquence du théorème d'inversion locale (en fait les deux résultats sont équivalents). On munit tous les espaces de la norme $||.||_\infty$. Soit $g = (g_1, g_2) : R^{n+q} \to R^{n+q}$ définit par $g_1(x,y) = x, g_2(x,y) = f(x,y)$. $g$ est $\mathcal{C}^1$ et sa différentielle est :

$dg(a,b).(h,l) = (h, df(a,b)(h,l))$.

Voyons que $dg(a,b)$ est inversible. Comme c'est une application entre deux espaces de même dimension, il suffit (théorème du rang) de montrer qu'elle est injective. Soit $(h,l) \in Kerdg(a,b)$, on a $h=0$ puis $df(a,b)(0,l)=0$, d'où par inversibilité de $df(a,b)_{|\{0\}\times\mathbb{R}^q}$ on déduit $l=0$, ce qui conclut.

Le théorème d'inversion locale fournit donc un $\mathcal{C}^k$-difféomorphisme de la boule $V=B((a,b),s)$ sur un ouvert $W \subset \mathbb{R}^{n+q}$ voisinage de $(a,f(a,b))$. Notons $h=(h_1,h_2)=g_{|V}^{-1}$ le $\mathcal{C}^k$-difféomorphisme réciproque.

On a donc

$$g(h_1(u,v),h_2(u,v)) = (h_1(u,v), f(h_1(u,v),h_2(u,v)) = (u,v),$$

d'où $h_1(u,v)=u, \quad f(u,h_2(u,v))=v$.

Soit $r$ tel que $B((a,0),r) \subset W$, on définit l'application $\varphi(u)=h_2(u,0)$ de $B(a,r)$ vers $\mathbb{R}^q$.

Par composition, $\varphi$ est de classe $\mathcal{C}^k$. Si $u \in B(a,r)$, on a $(u,0) \in B((a,0),r)$ d'où $h(u,0) \in B((a,b),s)$, donc $\varphi(u) \in B(b,s)$ et $f(u,\varphi(u))=0$ comme souhaité.

Si $(u,v) \in B(a,r) \times B(b,s)$ (remarquer que nécessairement $r \leq s$) on a donc $(u,v) \in V$ , d'où $dg(u,v)$ est un isomorphisme, donc $df(u,v)_{|\{0\}\times\mathbb{R}^q}$ est aussi un isomorphisme (pour une raison de dimension il suffit de vérifier son injectivité qui découle de celle de $(0, df(u,v)_{|\{0\}\times\mathbb{R}^q}(l)) = dg(u,v)(0,l)=0$ implique $l=0$)

Enfin, si $v \in B(b,s)$ est solution de $f(u,v)=0$ pur $u \in B(a,r)$ on $(u,v) \in V$, donc $g(u,v)=(u,0)$ implique $h(u,0)=(u,v)$ d'où $\varphi(u)=v$, d'où l'unicité de la solution. □

*Remarque* 49.
1. L-unicité de la solution de l'équation selon l'énoncé du théorème garantit l'unicité de la fonction implicite $\varphi$ (une fois fixé $B(a,r)$).
2. La restriction $df(a,b)_{|\{0\}\times\mathbb{R}^q}$ est aussi appelée différentielle partielle de $f$ par rapport à $Y$ (si on note (X,Y) les variables dans $U$) et notée $d_Y f(a,b) : \mathbb{R}^q \to \mathbb{R}^q$. Soit $Z_0=(X_0,Y_0) \in U$, et une boule pour la norme produit $B(X_0,r) \times B(X_0,s) \subset U$. On peut considérer la restrictions $f_{Y_0}(X)=f(X,Y_0), X \in B(X_0,r)$, alors : $d_X f(X_0,Y_0) := df_{Y_0}(X) = df(X_0,Y_0)_{|\{0\}\times\mathbb{R}^q} \circ J_{q,q+n}$ avec $J_{q,q+N} : \mathbb{R}^q \to \{0\}\times\mathbb{R}^q$ l'isomorphisme canonique.
   On a alors
   $$df(X_0,Y_0).(H,K) = df(X_0,Y_0).(0,K) + df(X_0,Y_0).(H,0) = d_X f(X_0,Y_0).H + d_Y f(X_0,Y_0).K$$
3. Si les variables $(X,Y) \in U$ sont détaillées en $X=(x_1,...,x_n)$, $Y=(y_1,...,y_q)$, la condition selon laquelle isomorphisme $df(a,b)_{|\{0\}\times\mathbb{R}^q}$ est isomorphisme linéaire est équivalent à l'inversibilité de la matrice :
   $$\left(\frac{\partial f_i}{\partial y_j}\right)_{\substack{1\leq i\leq q \\ 1\leq j\leq q}}.$$
   C'est à dire à $det\left(\frac{\partial f_i}{\partial y_j}\right)_{\substack{1\leq i\leq q \\ 1\leq j\leq q}} \neq 0$.
4. La preuve est une application du Théorème d'inversion locale. En fait, les deux théorèmes sont équivalents.

**Corollaire 74.** *(Cas q=1) Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^k$ ($k \in \mathbb{N}^* \cup \{\infty\}$). Soit $A=(a_1,...,a_n) \in U$ tel que $f(A)=0$ et $\frac{\partial f}{\partial x_n}(A) \neq 0$.*

*Alors il existe $r,s>0$ tel que $B((a_1,...,a_{n-1}),r)\times ]a_n-s,a_n+s[ \subset U$ et une application de classe $\mathcal{C}^k$ $\varphi : B((a_1,...,a_{n-1}),r) \to ]a_n-s,a_n+s[$ tel que :*

*$\forall x \in B((a_1,...,a_{n-1}),r)$, le $\varphi(x_1,...,x_{n-1})$ **est l'unique** élément $x_n$ de $]a_n-s,a_n+s[$ **solution** de :*

$$f(x_1,...,x_{n-1},x_n)=0.$$

*En particulier : $f(x_1,...,x_{n-1},\varphi(x_1,...,x_{n-1}))=0$.*

*De plus on peut choisir $r,s$ pour avoir $\forall(x_1,...,x_n)\in B((a_1,...,a_{n-1}),r)\to]a_n-s,a_n+s[$ $\frac{\partial f}{\partial x_n}(x_1,...,x_n)\neq 0$.*

## 3.3 Différentielles et dérivées partielles d'une fonction implicite.

**Corollaire 75.** *Sous les conditions du théorème 73, la différentielle de $\varphi$ est donnée pour tout $x\in B(a,r)$ par :*

$$d\varphi(x)=-df(x,\varphi(x))^{-1}_{|\{0\}\times\mathbb{R}^q}\circ df(x,\varphi(x))_{|\mathbb{R}^n\times\{0\}}=-(d_Yf(x,\varphi(x)))^{-1}\circ d_Xf(x,\varphi(x))).$$

*Démonstration.* En différenciant la relation $f(x,\varphi(x))=0$, on obtient pour $h\in\mathbb{R}^n$ :

$$0=df(x,\varphi(x))(h,d\varphi(x)(h))=df(x,\varphi(x))(h,0)+df(x,\varphi(x))(0,d\varphi(x)(h)).$$

On trouve donc :

$$df(x,\varphi(x))_{|\{0\}\times\mathbb{R}^q}(d\varphi(x)(h))=-df(x,\varphi(x))_{|\mathbb{R}^n\times\{0\}}(h),$$

d'où le résultat en appliquant l'inversibilité présente dans l'hypothèse. □

**Corollaire 76.** *Sous les conditions du corollaire 74, pour tout $i=1,...,n-1$, tout $(x_1,...,x_n)\in B((a_1,...,a_{n-1}),r)\to]a_n-s,a_n+s[$ :*

$$\frac{\partial\varphi}{\partial x_i}(x_1,...,x_{n-1})=-\frac{1}{\frac{\partial f}{\partial x_n}(x_1,...,x_n)}\times\frac{\partial f}{\partial x_i}(x_1,...,x_n).$$

## 3.4 Application : Tangente à une courbe.

Soit $f:U\to\mathbb{R}$, de classe $\mathcal{C}^k$ $k\geq 1$ sur l'ouvert $U\subset\mathbb{R}^2$.

On considère la courbe $C=\{(x,y)\in U\mid f(x,y)=0\}$.

Soit $(a,b)\in\mathbb{C}$ tel que $\frac{\partial f}{\partial y}(a,b)\neq 0$. D'après le théorème des fonctions implicites, il existe $r,s>0$ et une fonction $\varphi:O=]a-r,a+r[\to]b-s,b+s[$ de classe $\mathcal{C}^k$ tel que $f(x,\varphi(x))=0$, $\varphi(a)=b$, $(x,\varphi(x))$ est l'unique solution de $f(x,y)=0$ dans $O$.

On voit que $O\cap C$ est le graphe de $\varphi$ $\{(x,\varphi(x)),x\in]a-r,a+r[$. L'application $\theta:]a-r,a+r[\to O$ définie par $\theta(x)=(x,\varphi(x))$ a pour image $C\cap O$, on dit que c'est une paramétrisation locale de la courbe $C$.

$\theta'(x)=(1,\varphi'(x))$ est par définition un vecteur tangent à $C\cap O$.

Or on a vu qu'en différenciant l'équation $f(x,\varphi(x))=0$, on obtient :

$\frac{\partial f}{\partial x}(x,\phi(x))+\frac{\partial f}{\partial y}(x,\phi(x))\phi'(x)=0$ c'est à dire $\langle\nabla f(x,\phi(x)),(1,\phi'(x))\rangle=0$. Autrement dit la direction de la tangente à $C\cap O$ en un point $(x,\varphi(x))$ est donnée par la droite orthogonale au gradient $\nabla f(x,\phi(x))$. Sa direction est aussi le noyau de la forme linéaire $df(x,\varphi(x))$.

Ex : $f(x,y)=x^2+y^2-R^2$, $\nabla f(x,y)=2(x,y)$

# 4 Extrema liés (optimisation avec contraintes d'égalités)

## 4.1 Introduction

Soit $f : U \to \mathbb{R}$ une fonction définie sur un ouvert $U$ de $\mathbb{R}$. et $h_i : U \to \mathbb{R}$ $q$ autres fonctions.

On dit que les fonctions $h_i$ sont les contraintes. On va vouloir trouver les extrema de $f$ sur l'ensemble des contraintes $C = \{X \in U \mid h_i(x) = 0, \forall i = 1, ..., q\}$.

**Problème** : Trouver les extrema (locaux) $z^*$ de la fonction $f$ soumise aux contraintes $h_i(z^*) = 0$ $\forall i = 1, ..., q$. C'est à dire les extrema (locaux) de la fonction $f_{|C} : C \to \mathbb{R}$ tel que $f_{|C}(x) = f(x), \forall x \in C$.

En général, on suppose les fonctions $h_i$ continues donc $C$ est fermé, et l'hypothèse du théorème des extrema libres ne s'applique pas (on n'est pas sur un ouvert).

## 4.2 Théorème des multiplicateurs de Lagrange

**Théorème 77.** *Soit $U$ un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et $h_i : U \to \mathbb{R}$ des fonctions de classes $\mathcal{C}^1$.*

*Soit $z^* \in C = \{X \in U \mid h_i(x) = 0, \forall i = 1, ..., q\}$. Supposons que :*

1. *les vecteurs $\nabla h_i(z^*)$ sont linéairement indépendants.*
2. *$z^*$ est un extremum local de $f_{|C}$.*

*Alors il existe $\lambda_1, ...\lambda_q \in \mathbb{R}$ tel que*

$$\nabla f(z^*) + \sum_{i=1}^{q} \lambda_i \nabla h_i(z^*) = 0.$$

On dit que les nombres $\lambda_1, ..., \lambda_q$ sont des multiplicateurs de Lagrange.

*Remarque* 50.
1. Comme les $\nabla h_i(z^*)$ sont linéairement indépendants, les $\lambda_i$ sont uniques. Cette condition implique $q \leq n$.
2. On peut aussi énoncer le théorème avec les différentielles, si les $dh_i(z^*)$ sont des formes différentielles linéairement indépendantes et $z^*$ est un extremum local de $f_{|C}$. Alors il existe $\lambda_1, ...\lambda_q \in \mathbb{R}$ tel que
$$df(z^*) = \sum_{i=1}^{q} \lambda_i dh_i(z^*).$$
3. Le théorème a une interprétation géométrique, il dit que $\nabla f(z^*)$ est orthogonal au sous-espace vectoriel tangent à la "surface" (on dit parfois nappe) $C$ en $z^*$. Ceci est détaillé dans la remarque après la preuve.

*Démonstration.* L'indépendance linéaire des $\nabla h_i(z^*)$ signifie que la matrice

$$\left(\frac{\partial h_i}{\partial x_j}(z^*)\right)_{\substack{i=1...q \\ j=1...n}}$$

est de rang $q$. Comme le rang d'une matrice est le même que sa transposée, cela signifie que les vecteurs $(\frac{\partial h_i}{\partial x_1}(z^*))_{i=1,...,q}, ..., (\frac{\partial h_i}{\partial x_n}(z^*))_{i=1,...,q}$ forment une famille de rang $q$ (i.e. l'espace vectoriel qu'ils engendrent est de dimension $q$) donc on peut en extraire une base ayant $q$ vecteurs. Quitte à permuter les coordonnées ,on peut donc supposer que $(\frac{\partial h_i}{\partial x_{n-q+1}}(z^*))_{i=1,...,q}$,..., $(\frac{\partial h_i}{\partial x_n}(z^*))_{i=1,...,q}$ sont linéairement indépendants.

On considère donc la fonction $h = (h_1, ..., h_q) : \mathbb{R}^{n-q} \times \mathbb{R}^q \to \mathbb{R}^q$. Disons que $z^* = (z_0, z_1) \in \mathbb{R}^{n-q} \times \mathbb{R}^q$ selon cette décomposition. Selon la remarque 49. (3), ceci nous dit que $dh(z^*)_{|\{0\}\times\mathbb{R}^q}$ est un isomorphisme linéaire (car la matrice correspondante

$$\left(\frac{\partial h_i}{\partial x_j}(z^*)\right)_{\substack{i=1...q \\ j=(n-q+1)...n}}$$

est de déterminant non nul puisque ses vecteurs lignes ont rang $q$). On peut donc appliquer le théorème des fonctions implicites à $h$. Il existe donc un produit de boules $B(z_0, r) \times B(z_1, s) \subset U$ et une fonction $\varphi : B(z_0, r) \to B(z_1, s)$ de classe $\mathcal{C}^1$ telle que

$$h(x, \varphi(x)) = 0, \quad \varphi(z_0) = z_1.$$

De plus le corollaire 75 du théorème des fonctions implicites nous dit que :

$$d\varphi(x) = -d_Y h(x, \varphi(x))^{-1} \circ d_X h(x, \varphi(x)).$$

On a donc $\forall x \in B(z_0, r)(x, \varphi(x)) \in C$ d'où la fonction $g(x) = f(x, \varphi(x)) = f_{|C}(x, \varphi(x))$ atteint donc un extremum local en $z_0$. Par la condition nécessaire, on déduit donc $dg(z_0) = 0$. Or par le théorème de dérivation des fonctions composées,

$$dg(x) = df(x, \varphi(x)) \circ (Id_{n-q} \oplus d\varphi(x)) = d_X f(x, \varphi(x)) - d_Y f(x, \varphi(x)) \circ d_Y h(x, \varphi(x))^{-1} \circ d_X h(x, \varphi(x)).$$

En $z^*$ on obtient donc :

$$d_X f(z^*) = d_Y f(z^*) \circ d_Y h(z^*)^{-1} \circ d_X h(z^*).$$

Par ailleurs, on a évidemment, par la propriété de l'inverse :

$$d_Y f(z^*) = d_Y f(z^*) \circ d_Y h(z^*)^{-1} \circ d_Y h(z^*).$$

Donc si

$$\Lambda = d_Y f(z^*) \circ d_Y h(z^*)^{-1} = (-\lambda_1, ..., -\lambda_q),$$

matrice ligne on a obtenu :

$$df(z^*) = \Lambda \circ dh(z^*).$$

Ceci s'interprète comme la relation cherchée :

$$df(z^*) + \sum_{i=1}^{q} \lambda_i dh_i(z^*) = 0.$$

$\square$

*Remarque* 51. 1. Soit le sous espace vectoriel $T_z(C) = \{H \in \mathbb{R}^n \mid dh_i(z).H = 0 \forall i = 1...q\}$. Voyons que, sous les hypothèses du théorème (pour $z$ au lieu de $z^*$) $z + T_z(C)$ est l'espace affine tangent à la surface c'est à dire que l'espace vectoriel $T_z(C)$ est engendré par les vecteurs tangents aux courbes $\gamma : \mathbb{R} \to C$ $\mathcal{C}^1$ passant par $z$ disons en 0, $\gamma(0) = z$.

Soit $\gamma$ une telle courbe on a $h_i(\gamma(t)) = 0$ pour tout $t, i$ donc en dérivant $dh_i(\gamma(t)).\gamma'(t) = 0$ c'est-à-dire en $t = 0$ $dh_i(z).\gamma'(0) = 0$ autrement dit $\gamma'(0) \in T_z(C)$, donc la droite tangente à $\gamma$ en $t = 0$ $z + \mathbb{R}\gamma'(0) \in z + T_z(C)$.

Réciproquement soit $H = (h_1, ..., h_n) \in T_z(C)$. Voyons qu'il existe une courbe $\gamma :] - \epsilon, \epsilon[\to \mathbb{C}$ avec $\gamma(0) = z, \gamma'(0) = H$. Par la preuve du théorème, on peut supposer que le théorème des fonctions implicites nous donne $\varphi$ tel que $(x, \varphi(x)) \in C$ pour tout $x \in B(z_0, r)$. Prenons $\gamma(t) = (z_0 + Gt, \varphi(z_0 + Gt))$ avec $G = (h_1, ..., h_{n-q}) \in \mathbb{R}^{n-q}$ (on prend $\epsilon = r/||G||$.)

On a donc par dérivation des fonctions composées :

$$\gamma'(0) = (G, d\varphi(z_0).G)$$

Or $0 = dh(z^*).H = d_X h(z^*).G + d_Y h(z^*).K$ avec $K = (h_{n-q+1}, ..., h_n)$ d'où $K = -d_Y h(z^*)^{-1} \circ d_X h(z^*).G = d\varphi(z_0).G$. Donc $\gamma'(0) = H$ comme voulu.

2. Sous les hypothèses du théorème, on veut voir que $Vect(\nabla h_i(z^*), i = 1, ..., q)$ est l'ensemble $N_{z^*}(C)$ des vecteurs orthogonaux à $T_{z^*}C$.

$$N_{z^*}(C) = \{G \in \mathbb{R}^n \mid \langle G, H\rangle, \forall H \in T_{z^*}C.\}$$

En effet, $\nabla h_i(z^*) \in N_{z^*}(C)$ car si $H \in T_{z^*}C$ on a $\langle \nabla h_i(z^*), H\rangle = dh_i(z).H = 0$.

Enfin Comme les $\nabla h_i(z^*)$ d'après l'hypothèse du théorème, on a $dim N_{z^*}(C) \geq q$ et vu que $N_{z^*}(C) \cap T_{z^*}C = \{0\}$ d'où $dim(N_{z^*}(C)) + dim(T_{z^*}C) \leq n$ et que $dim(T_{z^*}C) \geq n - q$ (lieu d'annulation de $q$ forme linéaire, on a donc égalité dans toutes les inégalités ci-dessus, donc $dim N_{z^*}(C) = q$ et les $\nabla h_i(z^*)$ engendrent donc $N_{z^*}(C)$.

## 4.3 Exemple et Méthode

*Exemple* 32. On pose $U = \mathbb{R}^2 - \{0, 0\}$. On veut trouver les extrema de $f(x, y) = \sqrt{x^2 + y^2}$ (la distance euclidienne du point $M = (x, y)$ au point $(0, 0)$) sous la contrainte d'annulation de $h(x, y) = x^2 + y^2 - 4x + 3$.

La courbe $C = \{(x, y) \mid x^2 + y^2 - 4x + 3 = 0\} = \{(x, y) \mid (x - 2)^2 + y^2 = 1\}$ est le cercle de centre $A = (2, 0)$ et de rayon 1. On a $\nabla h(x, y) = 2(x - 2, y)$ Si $M = (x, y) \in C$ $\nabla h(x, y) = 2\overrightarrow{AM} \neq (0, 0)$ et est bien normal à $C$.

Comme $\nabla h(x, y) \neq 0$ c'est une famille libre et on peut appliquer le théorème.

Pour chercher les extrema de $f$ il revient au même de chercher ceux de $g(x, y) = x^2 + y^2$.

On a $\nabla g(x, y) = (2x, 2y)$. En un extremum local $(x_0, y_0)$, le théorème nous dit qu'il y a multiplicateur de Lagrange $\lambda$ tel que : $\nabla g(x_0, y_0) = \lambda \nabla h(x_0, y_0)$, c'est à dire :

$2x_0 = \lambda 2(x_0 - 2), 2y_0 = 2\lambda y_0$ et de plus on a la contrainte $(x_0 - 2)^2 + y_0^2 = 1$. On obtient un système de 3 équations à 3 inconnus, $x_0, y_0$ et le multiplicateur de Lagrange $\lambda$.

On obtient ici $\lambda = 1$ ou $y = 0$. La première condition est impossible car elle implique $2 - x_0 = 0 - x_0$. Donc $y = 0$ d'où $x_0 = 1$ (alors $\lambda = -1$) ou $x_0 = 3$ (alors $\lambda = 3$)

Comme la courbe $C$ est compacte, $f$ atteint un maximum et un minimum, $f(1, 0) = 1$ est le minimum et $f(3, 0) = 3$ est le maximum.

**Méthode générale** Les contraintes et la condition des multiplicateurs de Lagrange forment un système de $n + q$ équations à $n + q$ inconnus (les $n$ coordonnées d'un extrema, et les $q$ multiplicateurs de Lagrange)

$$\begin{cases} h_i(x_1, ..., x_n) = 0, i = 1, ..., q \\ \nabla f(x_1, ..., x_n) + \sum_{i=1}^{q} \lambda_i \nabla h_i(x_1, ..., x_n) = 0 \end{cases}$$

Parmi les candidats à être des extrema, soit on regarde le comportement de $f_{|C}$ pour voir si il y a un extremum local (cf. conditions du second ordre), soit on cherche les valeurs $f(z^*)$ et on trouve les extrema globaux parmi les $z^*$ trouvés (si $f$ est bornée sur $C$, par exemple si $C$ est compact, c'est toujours vrai. Sinon, on fait attention aux limites à l'infini).

*Exemple* 33. (Cas quadratique, introduction à la notion de vecteurs propres et de valeurs propres) Soit $A = A^T \in M_n(\mathbb{R})$ une matrice symétrique. On cherche à trouver les extrema de $f(X) = \langle AX, X\rangle$ sous la contrainte $g(X) = ||X||_2^2 - 1 = 0$. Remarquez que si le minimum est positif (resp. strictement positif) on aura alors par définition que $A$ est positive (resp. défini positive) (et de même avec la négativité du maximum et négative). La positivité d'une matrice s'exprime donc grâce à un problème d'extrémisation sous contrainte.

On a vu en TD que $df(X).H = 2\langle AX, H\rangle$ et $dg(X).H = \langle X, H\rangle$ (on a bien $dg(X) \neq 0$ sous la contrainte $||X||_2 = 1$). On déduit donc du théorème qu'en un extremum $X$, il existe un multiplicateur de Lagrange $\lambda$ tel que :

$$AX = -\lambda X.$$

$-\lambda$ est donc ce qu'on appelle une *valeur propre* de $A$ (un nombre $\mu$ tel qu'il existe $X \in \mathbb{R}^n$ avec $AX = \mu X$.) et $X$ est ce qu'on appelle un *vecteur propre* associé.

Remarquez que $AX + \lambda X = 0$ implique que $A + \lambda I_n$ n'est pas injective ce qui équivaut (en dimension finie) à $A + \lambda I_n$ n'est pas inversible c'est à dire $\det(A + \lambda I_n) = 0$. Les multiplicateurs de Lagrange (valeurs propres) sont donc les racines du polynôme : $T \mapsto \det(A + TI_n)$ (resp. $T \mapsto \det(A - TI_n)$). C'est la façon générale de trouver les valeurs propres.

En un extremum $X$ avec multiplicateur de Lagrange $\lambda$, on a $f(X) = -\lambda||X||_2^2 = -\lambda$. Un minimum est donc atteint en un vecteur propre pour la plus petite valeur propre $(-\lambda)$. On a donc obtenu le résultat suivant.

**Corollaire 78.** *Une matrice symétrique $A = A^T \in M_n(\mathbb{R})$ est positive (resp . définie positive) si et seulement si toutes les racines du polynôme $X \mapsto \det(A - XI_n)$ sont positives (resp. strictement positive).*

*Une matrice symétrique $A = A^T \in M_n(\mathbb{R})$ est négative (resp . définie négative) si et seulement si toutes les racines du polynôme $X \mapsto \det(A - XI_n)$ sont négatives (resp. strictement négative).*

## 4.4 Lagrangien et conditions du second ordre (Preuves facultatives)

Introduisons le lagrangien du problème de minimisation de $f : U \to \mathbb{R}$ sous les contraintes $h_i : U \to \mathbb{R}$ $i = 1, ..., q$ ($U$ ouvert de $\mathbb{R}^n$) :

$$L(x, \lambda_1, ..., \lambda_q) := f(x) + \sum_{i=1}^{q} \lambda_i h_i(x).$$

On remarque que $\nabla_x L = \nabla f + \sum_{i=1}^q \lambda_i \nabla h_i(x)$ et $\frac{\partial L}{\partial \lambda_j} = h_j(x)$.

En un extremum $z^*$ le théorème des multiplicateurs de Lagrange énonce donc qu'il existe des multiplicateurs de Lagrange $\lambda^* = (\lambda_1, ..., \lambda_q)$ tel que $\nabla L(z^*, \lambda^*) = 0$ c'est à dire $(z^*, \lambda^*)$ est un point critique de $L$. On va obtenir des conditions du second ordre en utilisant les dérivées secondes du Lagrangien.

**Définition 48.** Soit $U \subset \mathbb{R}^n$ et $L : U \times \mathbb{R}^q \to \mathbb{R}$ un Lagrangien de classe $\mathcal{C}^2$. La **Hessienne partielle** du lagrangien $L$ est la matrice symétrique

$$H_1(L) = \left(\frac{\partial^2 L}{\partial x_i \partial x_j}\right)_{\substack{i=1...n \\ j=1...n}}.$$

**Théorème 79.** *(Condition nécessaire du second ordre pour minima sous contraintes d'égalité) Soit $U$ un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et $h_i : U \to \mathbb{R}$ des fonctions de classes $\mathcal{C}^2$. Soit le lagrangien du problème :*

$$L(x, \lambda_1, ..., \lambda_q) := f(x) + \sum_{i=1}^{q} \lambda_i h_i(x).$$

*Soit $H_1(L)$ la hessienne partielle du lagrangien et $z^* \in C = \{X \in U \mid h_i(x) = 0, \forall i = 1, ..., q\}$. Supposons que :*

1. *les vecteurs $\nabla h_i(z^*)$ sont linéairement indépendants.*
2. *$z^*$ est un minimum local de $f_{|C}$.*

*Alors il existe $\lambda^* = (\lambda_1, ...\lambda_q) \in \mathbb{R}^q$ tel que $\nabla L(z^*, \lambda^*) = 0$, et pour tout $K \in T_{z^*}(C) = \{K \in \mathbb{R}^n \mid dh_i(z^*).K = 0 \forall i = 1...q\}$, on ait :*

$$\langle H_1(L)(z^*, \lambda^*)K, K\rangle \geq 0.$$

La conclusion du théorème dit donc que la hessienne partielle $H_1(L)(x^*, \lambda^*)$ restreinte au plan tangent $T_{z^*}(C)$ est positive, on va obtenir une condition suffisante si on la suppose en plus définie positive.

*Démonstration.* L'existence (et l'unicité) de $\lambda^*$ satisfaisant la condition de premier ordre est donnée par le théorème des multiplicateurs de Lagrange.

Soit $\gamma : ]-\epsilon, \epsilon[ \to C$ la courbe construite à la remarque 51.(1) telle que $\gamma(0) = z^*$, $\gamma'(0) = K \in T_{z^*}(C)$. 0 est minimum local de $f \circ \gamma$ vu que $\gamma$ est à valeur dans $C$ et que $z^* = \gamma(0)$ est un minimum local de $f_{|C}$. Il en est de même de $L_{\lambda^*} \circ \gamma = f \circ \gamma$. avec $L_{\lambda^*}(x) = L(x, \lambda^*)$

De plus, comme $h$ est $\mathcal{C}^2$ le théorème d'inversion locale appliquée pour construire $\gamma$ implique que $\gamma$ est de classe $\mathcal{C}^2$

Calculons les deux premières dérivées :

$$(L_{\lambda^*} \circ \gamma)'(t) = \langle \nabla_x L(\gamma(t), \lambda^*), \gamma'(t)\rangle.$$

$$(L_{\lambda^*} \circ \gamma)''(t) = \langle \nabla_x L(\gamma(t), \lambda^*), \gamma''(t)\rangle + \langle H_1 L(\gamma(t), \lambda^*)\gamma'(t), \gamma'(t)\rangle.$$

En prenant en compte que $\nabla_x L(\gamma(0), \lambda^*) = 0$ par la condition des multiplicateurs de Lagrange, on obtient en $t = 0$ :$(L_{\lambda^*} \circ \gamma)'(0) = 0$,

$$(L_{\lambda^*} \circ \gamma)''(0) = \langle H_1 L(z^*, \lambda^*)K, K\rangle.$$

En appliquant Taylor Young à $L_{\lambda^*} \circ \gamma$ on obtient :

$$L_{\lambda^*} \circ \gamma(t) = f(z^*) + \frac{t^2}{2}\langle H_1 L(z^*, \lambda^*)K, K\rangle + o(t^2) \geq f(z^*).$$

En divisant par $t^2$ et en faisant tendre $t \to 0$, on obtient donc la conclusion :

$$\langle H_1 L(z^*, \lambda^*)K, K\rangle \geq 0.$$

□

**Théorème 80.** *(Conditions suffisantes du second ordre pour minima sous contrainte d'égalité) Soit $U$ un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et $h_i : U \to \mathbb{R}$ des fonctions de classes $\mathcal{C}^2$. Soit le lagrangien du problème :*

$$L(x, \lambda_1, ..., \lambda_q) := f(x) + \sum_{i=1}^{q} \lambda_i h_i(x).$$

*Soit $H_1(L)$ la hessienne partielle du lagrangien et $z^* \in C = \{X \in U \mid h_i(x) = 0, \forall i = 1, ..., q\}$. Supposons qu'il existe $\lambda^* = (\lambda_1, ..., \lambda_q) \in \mathbb{R}^q$ tel que $\nabla L(z^*, \lambda^*) = 0$, et tel que pour tout $K \in T_{z^*}(C) = \{K \in \mathbb{R}^n \mid dh_i(z^*).K = 0 \forall i = 1...q\}$, avec $K \neq 0$ on ait :*

$$\langle H_1(L)(z^*, \lambda^*)K, K\rangle > 0,$$

*alors $z^*$ est un minimum local de $f_{|C}$.*

*Démonstration.* Supposons par l'absurde que $z^*$ n'est pas un minimum local, il existe donc une suite $z_n \in C$, $z_n \to z^*$ tel que $f(z_n) < f(z^*)$. Soit $t_n = ||z_n - z|| > 0$. Comme le Lagrangien $L$ est $\mathcal{C}^2$, pour $n$ assez grand, Taylor Young donne :

$$f(z_n) = L(z_n, \lambda^*) = L(z^*, \lambda^*) + \langle \nabla_x L(z^*, \lambda^*), \frac{z_n - z^*}{t_n}\rangle t_n + \frac{t_n^2}{2}\langle H_x L(z^*, \lambda^*)\frac{z_n - z^*}{t_n}, \frac{z_n - z^*}{t_n}\rangle + o(t_n^2)$$

Comme par hypothèse, $\nabla_x L(z^*, \lambda^*) = 0$ le développement commence à l'ordre 2. Par ailleurs, si $u_n = \frac{z_n - z^*}{t_n}$ on a $||u_n|| = 1$, donc par compacité de la sphère unité, on extrait une sous-suite telle que $u_{\varphi(n)} \to K$. Par ailleurs, $h_i(z_n) = h(z^*) = 0$, donc par la définition de la différentielle $h_i(z_n) = h_i(z^*) + dh_i(z^*).u_n t_n + o(t_n)$, on a donc en divisant par $t_n$ $dh_i(z^*).u_n + o(1) = 0$ soit en faisant tendre $n \to \infty$ après extraction $dh_i(z^*).K = 0$. Autrement dit, comme cela est vrai pour tout $i = 1...q$, $K \in T_{z^*}(C)$.

Vu que $L(z^*, \lambda^*) = f(z^*)$, on déduit :

$$\frac{t_n^2}{2}\langle H_x L(z^*, \lambda^*)\frac{z_n - z^*}{t_n}, \frac{z_n - z^*}{t_n}\rangle + o(t_n^2) = f(z_n^*) - f(z^*) < 0.$$

Donc en divisant par $t_n^2$ extrayant et prenant la limite : $\frac{1}{2}\langle H_1 L(z^*, \lambda^*)K, K\rangle \leq 0$.

Ceci contredit l'hypothèse.

□

**Corollaire 81.** *(Cas de la dimension 2) Soit $U$ un ouvert de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ et $h : U \to \mathbb{R}$ des fonctions de classes $\mathcal{C}^2$. Soit le lagrangien du problème : $L(x, \lambda) := f(x) + \lambda h(x)$. Soit $H(L)$ la hessienne (totale) du lagrangien et $z^* \in C = \{X \in U \mid h(x) = 0\}$, tel que $\nabla h(z^*) \neq 0$.*

*Soit $(z^*, \lambda^*)$ un point critique du Lagrangien :*

1. *Si $\det(H(L)(z^*, \lambda^*)) < 0$ alors $z^*$ est un minimum local de $f|_C$.*
2. *Si $\det(H(L)(z^*, \lambda^*)) > 0$ alors $z^*$ est un maximum local de $f|_C$.*

Le résultat ci-dessus est en fait valable dés qu'on a $(n-1)$ contraintes en dimension $n$ sous l'hypothèse du théorème des multiplicateurs de Lagrange en remplaçant le déterminant par $(-1)^n \det H(L)$.

*Démonstration.* On rappelle que si $u = (a, b) \in R^2 - \{(0,0)\}$, $u^\perp = (-b, a)$ est orthogonal à $u$. D'après le critère général obtenir un minimum revient à montrer la positivité de :

$$\begin{aligned}
\langle H_1L(z^*,\lambda^*)(\nabla h(z^*))^\perp, (\nabla h(z^*))^\perp\rangle &= \frac{\partial^2 L}{\partial x^2}(z^*,\lambda^*)\left(\frac{\partial h}{\partial y}(z^*,\lambda^*)\right)^2 + \frac{\partial^2 L}{\partial y^2}(z^*,\lambda^*)\left(\frac{\partial h}{\partial x}(z^*,\lambda^*)\right)^2 \\
&\quad - 2\frac{\partial^2 L}{\partial x \partial y}(z^*,\lambda^*)\left(\frac{\partial h}{\partial x}(z^*,\lambda^*)\right)\left(\frac{\partial h}{\partial y}(z^*,\lambda^*)\right) \\
&= -\det\begin{pmatrix} \frac{\partial^2 L}{\partial x^2}(z^*,\lambda^*) & \frac{\partial^2 L}{\partial x\partial y}(z^*,\lambda^*) & \frac{\partial h}{\partial x}(z^*) \\ \frac{\partial^2 L}{\partial x\partial y}(z^*,\lambda^*) & \frac{\partial^2 L}{\partial y^2}(z^*,\lambda^*) & \frac{\partial h}{\partial y}(z^*) \\ \frac{\partial h}{\partial x}(z^*) & \frac{\partial h}{\partial y}(z^*) & 0 \end{pmatrix} \\
&= -\det\left(H(L)(z^*,\lambda^*)\right).
\end{aligned}$$

□

*Exemple* 34. En considérant l'exemple 22 de la section 4.3. Le lagrangien est $L(x, y, \lambda) = x^2 + y^2 - \lambda((x-2)^2 + y^2 - 1)$. La matrice hessienne du Lagrangien est :

$$HL(x,y,\lambda) = \begin{pmatrix} 2-2\lambda & 0 & 2(x-2) \\ 0 & 2-2\lambda & 2y \\ 2(x-2) & 2y & 0 \end{pmatrix}$$

de déterminant $\det(HL(x, y, \lambda)) = -8(1-\lambda)[(x-2)^2 + y^2]$ Aux points critiques du Lagrangien trouvés à l'exemple 22, c'est à dire $(1, 0, -1)$ et $(3, 0, 3)$, on a $\det(HL(1, 0, -1)) = -16 < 0$ donc on retrouve que $(1, 0)$ est un minimum local de $f_{|C}$ et on a $\det(HL(3, 0, 3)) = 16 > 0$ donc on retrouve que $(3, 0)$ est un maximum local de $f_{|C}$. Directement avec le théorème 79, on peut aussi dire que $H_1L = 2 - 2\lambda$ donc il en est de même de la restriction ay plan tangent à la contrainte et $2 - 2\lambda > 0$ ssi on a un minimum.

On rappelle que si $u = (a, b, c) \in R^3$ avec $c \neq 0$, $u \wedge e_1 = (0, c, -b)$ et $u \wedge e_2 = (-c, 0, a)$ sont orthogonaux à $u$, non nuls, et forment une base du plan orthogonal à $u$. Bien sûr si $c = 0$, $(0, 0, 1)$ fournit un vecteur orthogonal à $u$ très pratique.

**Corollaire 82.** *(Cas de la dimension 3) Soit $U$ un ouvert de $I\!R^2$, $f : U \to I\!R$ et $h : U \to I\!R$ des fonctions de classes $\mathcal{C}^2$. Soit le lagrangien du problème : $L(x, \lambda) := f(x) + \lambda h(x)$. Soit $H_x(L)$ la hessienne partielle du lagrangien et $H(L)$ sa hessienne (totale). Soit $X^* \in C = \{X \in U \mid h(x) = 0\}$, tel que $\nabla h(X^*) \neq 0$ et soit $e \neq 0$ orthogonal à $\nabla h(X^*)$.*

*On suppose que $(X^*, \lambda^*)$ un point critique du Lagrangien :*

1. *Si* $\det(H(L)(X^*, \lambda^*)) < 0$ *et* $\langle H_1(L)e, e\rangle > 0$*,* $X^*$ *est un minimum local de* $f|_C$*.*
2. *Si* $\det(H(L)(X^*, \lambda^*)) < 0$ *et* $\langle H_1(L)e, e\rangle < 0$ $X^*$ *est un maximum local de* $f|_C$*.*
3. *Si* $\det(H(L)(X^*, \lambda^*)) > 0$*,* $X^*$ *n'est pas un extremum local de* $f|_C$*.*
4. *Si* $\det(H(L)(X^*, \lambda^*)) = 0$*, on ne peut pas conclure.*

*Remarque* 52. Sous les hypothèses du théorème on a une formule pour le déterminant qu'il n'est pas besoin d'apprendre :

$$\begin{aligned}\det\left(H(L)(z^*,\lambda^*)\right) = {} & \left(\frac{\partial h}{\partial x}(X^*,\lambda^*)\right)^2 \left[\left(\frac{\partial^2 L}{\partial y\partial z}(X^*,\lambda^*)\right)^2 - \frac{\partial^2 L}{\partial y^2}(X^*,\lambda^*)\frac{\partial^2 L}{\partial z^2}(X^*,\lambda^*)\right] \\ & + \left(\frac{\partial h}{\partial y}(X^*,\lambda^*)\right)^2 \left[\left(\frac{\partial^2 L}{\partial x\partial z}(X^*,\lambda^*)\right)^2 - \frac{\partial^2 L}{\partial x^2}(X^*,\lambda^*)\frac{\partial^2 L}{\partial z^2}(X^*,\lambda^*)\right] \\ & + \left(\frac{\partial h}{\partial z}(X^*,\lambda^*)\right)^2 \left[\left(\frac{\partial^2 L}{\partial x\partial y}(X^*,\lambda^*)\right)^2 - \frac{\partial^2 L}{\partial x^2}(X^*,\lambda^*)\frac{\partial^2 L}{\partial y^2}(X^*,\lambda^*)\right] \\ & +2\left(\frac{\partial h}{\partial x}(X^*,\lambda^*)\frac{\partial h}{\partial y}(X^*,\lambda^*)\right)\left[\frac{\partial^2 L}{\partial x\partial y}(X^*,\lambda^*)\frac{\partial^2 L}{\partial z^2}(X^*,\lambda^*) - \frac{\partial^2 L}{\partial x\partial z}(X^*,\lambda^*)\frac{\partial^2 L}{\partial y\partial z}(X^*,\lambda^*)\right] \\ & +2\left(\frac{\partial h}{\partial x}(X^*,\lambda^*)\frac{\partial h}{\partial z}(X^*,\lambda^*)\right)\left[\frac{\partial^2 L}{\partial x\partial z}(X^*,\lambda^*)\frac{\partial^2 L}{\partial y^2}(X^*,\lambda^*) - \frac{\partial^2 L}{\partial x\partial y}(X^*,\lambda^*)\frac{\partial^2 L}{\partial z\partial y}(X^*,\lambda^*)\right] \\ & +2\left(\frac{\partial h}{\partial y}(X^*,\lambda^*)\frac{\partial h}{\partial z}(X^*,\lambda^*)\right)\left[\frac{\partial^2 L}{\partial y\partial z}(X^*,\lambda^*)\frac{\partial^2 L}{\partial x^2}(X^*,\lambda^*) - \frac{\partial^2 L}{\partial x\partial y}(X^*,\lambda^*)\frac{\partial^2 L}{\partial x\partial z}(X^*,\lambda^*)\right].\end{aligned}$$

*Démonstration.* Comme $\nabla h(X^*) \neq 0$, on peut supposer, quitte à permuter les coordonnées, que $\frac{\partial h}{\partial z}(X^*) \neq 0$. $\nabla h(X^*) \wedge e_1, \nabla h(X^*) \wedge e_2$ est donc une base du plan orthogonal de $\nabla h(X^*)$. Le critère du théorème nous dit de considérer la matrice :

$$\begin{aligned}&\begin{pmatrix} r & s \\ s & t \end{pmatrix} \\ &= \begin{pmatrix} \langle H_1L(X^*,\lambda^*)(\nabla h(X^*)\wedge e_1), (\nabla h(X^*)\wedge e_1)\rangle & \langle H_1L(X^*,\lambda^*)(\nabla h(X^*)\wedge e_1), (\nabla h(X^*)\wedge e_2)\rangle \\ \langle H_1L(X^*,\lambda^*)(\nabla h(X^*)\wedge e_2), (\nabla h(X^*)\wedge e_1)\rangle & \langle H_1L(X^*,\lambda^*)(\nabla h(X^*)\wedge e_2), (\nabla h(X^*)\wedge e_2)\rangle \end{pmatrix}\end{aligned}$$

Plus précisément, nous pouvons calculer :

$$\begin{aligned}\langle H_1L(X^*,\lambda^*)(\nabla h(X^*)\wedge e_1), (\nabla h(X^*)\wedge e_1)\rangle &= \frac{\partial^2 L}{\partial y^2}(X^*,\lambda^*)\left(\frac{\partial h}{\partial z}(X^*,\lambda^*)\right)^2 + \frac{\partial^2 L}{\partial z^2}(X^*,\lambda^*)\left(\frac{\partial h}{\partial y}(X^*,\lambda^*)\right)^2 \\ &\quad - 2\frac{\partial^2 L}{\partial y\partial z}(X^*,\lambda^*)\left(\frac{\partial h}{\partial z}(X^*,\lambda^*)\right)\left(\frac{\partial h}{\partial y}(X^*,\lambda^*)\right) \\ &= -\det\begin{pmatrix} \frac{\partial^2 L}{\partial x^2}(X^*,\lambda^*) & \frac{\partial^2 L}{\partial x\partial y}(X^*,\lambda^*) & \frac{\partial h}{\partial x}(z^*) \\ \frac{\partial^2 L}{\partial x\partial y}(X^*,\lambda^*) & \frac{\partial^2 L}{\partial y^2}(X^*,\lambda^*) & \frac{\partial h}{\partial y}(z^*) \\ \frac{\partial h}{\partial x}(z^*) & \frac{\partial h}{\partial y}(z^*) & 0 \end{pmatrix} \\ &= -\det\left(H(L)(X^*,\lambda^*)\right).\end{aligned}$$

Pour simplifier on ne mentionne plus l'évaluation en $(X^*,\lambda^*)$ et on calcule

$$\begin{aligned}&rt - s^2 = \\ &\left[\frac{\partial^2 L}{\partial y^2}\left(\frac{\partial h}{\partial z}\right)^2 + \frac{\partial^2 L}{\partial z^2}\left(\frac{\partial h}{\partial y}\right)^2 - 2\frac{\partial^2 L}{\partial y\partial z}\left(\frac{\partial h}{\partial z}\right)\left(\frac{\partial h}{\partial y}\right)\right]\left[\frac{\partial^2 L}{\partial x^2}\left(\frac{\partial h}{\partial z}\right)^2 + \frac{\partial^2 L}{\partial z^2}\left(\frac{\partial h}{\partial x}\right)^2 - 2\frac{\partial^2 L}{\partial x\partial z}\left(\frac{\partial h}{\partial z}\right)\left(\frac{\partial h}{\partial x}\right)\right] \\ &- \left[\frac{\partial^2 L}{\partial x\partial z}\left(\frac{\partial h}{\partial z}\right)\left(\frac{\partial h}{\partial y}\right) + \frac{\partial^2 L}{\partial y\partial z}\left(\frac{\partial h}{\partial z}\right)\left(\frac{\partial h}{\partial x}\right) - \frac{\partial^2 L}{\partial x\partial y}\left(\frac{\partial h}{\partial z}\right)^2 - \frac{\partial^2 L}{\partial z^2}\left(\frac{\partial h}{\partial x}\right)\left(\frac{\partial h}{\partial y}\right)\right]^2 \\ &= -\left(\frac{\partial h}{\partial z}\right)^2 \det\left(H(L)(X^*,\lambda^*)\right).\end{aligned}$$

Le dernier résultat est le fruit d'un calcul fastidieux (en utilisant la formule de la remarque précédente).

On voit donc que $rt - s^2$ est du signe opposé à $\det\left(H(L)(X^*, \lambda^*)\right)$. Si ce dernier est positif, on a $rt - s^2 < 0$ et la matrice n'est ni positive ni négative, par le critère usuel et on n'a pas d'extremum, s'il est nul on ne peut pas conclure.

Si $\det\left(H(L)(X^*, \lambda^*)\right) < 0$ c'est à dire $rt - s^2 > 0$, on sait que la matrice est soit définie positive, soit définie négative donc tous les $\langle H_1(L)e, e\rangle$ sont du même signe pour $e$ orthogonal à $\nabla h(X^*)$, donc du signe de $r$, le critère usuel conclut. $\square$

# Chapitre 4

# Intégrales dépendant d'un paramètre

Dans cette partie $E$ est un espace vectoriel normé (souvent de dimension finie). Comme à l'exemple 3, on munit $E \times \mathbb{R}$ de la norme produit $||(x,t)|| = \max(||x||, |t|)$.

## 1 Intégrales définies dépendant d'un paramètre

**Définition 49.** Soit $A \subset E$ une partie d'un evn $E$. Soit $f : A \times [a,b] \to \mathbb{R}$ un fonction. On suppose que pour tout $x \in A$ la fonction $t \mapsto f(x,t)$ est continue par morceau (ou plus généralement Riemann intégrable). Elle est donc intégrable et on peut poser :

$$F(x) = \int_a^b f(x,t)dt.$$

On obtient ainsi une fonction $f : A \to \mathbb{R}$.

*Exemple* 35. Soit $F : \mathbb{R} \to \mathbb{R}$. $F(x) = \int_0^1 \frac{dt}{1+x^2t^2}$

Dans ce cas on peut calculer en changeant de variable pour $x \neq 0$ ($F(0) = 1$ est évident) :

$$F(x) = \frac{1}{x}\int_0^1 \frac{d(xt)}{1+x^2t^2} = \frac{Arctan(x)}{x}.$$

L'étude des intégrales dépendant d'un paramètre est surtout utile quand on ne peut pas calculer l'intégrale. En fait, On peut voir *Arctan* comme une fonction définie en utilisant une intégrale dépendant d'un paramètre (si on ne connaît pas la fonction tangente pour la définir comme son inverse).

**Théorème 83.** *(Théorème de continuité, cas continuité uniforme) Soit $A \subset E$. Soit $f : A \times [a,b] \to \mathbb{R}$ une application uniformément continue alors $x \mapsto F(x) = \int_a^b f(x,t)dt$ est (uniformément) continue sur $A$.*

*Remarque* 53. On remarquera que si $A$ est compact et $f$ continue, le théorème s'applique (car par le Théorème de Heine toute application continue sur un compact est uniformément continue).

*Démonstration.* Soit $\epsilon > 0$ fixé. Comme $f$ est uniformément continue, il existe $\eta > 0$ tel que si $||(x,t)-(x',t')|| < \eta$ (c'est à dire $||x-x'|| < \eta$, et $|t-t' < \eta$) alors $|f(x,t)-f(x',t')| < \epsilon/(b-a)$.
On a $F(x) - F(x_0) = \int_a^b [f(x,t) - f(x_0,t)]dt$

Donc si $|x - x_0| < \eta$, comme $|f(x,t) - f(x_0,t)| < \epsilon/(b-a)$, on obtient

$$|F(x) - F(x_0)| \le \int_a^b |f(x,t) - f(x_0,t)|dt \le \int_a^b \epsilon/(b-a)dt = \epsilon.$$

On a donc montré que $f$ est uniformément continue. $\square$

**Corollaire 84.** *(Théorème de continuité) Soit $U \subset I\!R^n$ un ouvert. Soit $f : U \times [a,b] \to I\!R$ une application continue alors $x \mapsto F(x) = \int_a^b f(x,t)dt$ est continue sur $U$.*

*Démonstration.* Soit $x \in U$. Pour montrer la continuité en $x$, il suffit de se restreindre à une boule $B_F(x,r) \subset U$ $r > 0$ qui existe car $U$ ouvert. Or la remarque précédente montre que $F$ est uniformément continue sur $B_F(x,r)$ qui est compact (car fermé borné de $\mathbb{R}^n$). $\square$

**Théorème 85.** *(Théorème de dérivabilité) Soit $U \subset I\!R^n$ un ouvert. Soit $f : U \times [a,b] \to I\!R$. On suppose*

1. *$t \mapsto f(x,t)$ est continue*
2. *$\frac{\partial f}{\partial x_i}$ existe sur $U \times [a,b]$ et est continue.*

*Alors $x \mapsto F(x) = \int_a^b f(x,t)dt$ admet une i-ème dérivée partielle et $\frac{\partial F}{\partial x_i}$ est continue sur $U$ et pour tout $x \in U$ :*

$$\frac{\partial F}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(x,t)dt.$$

*Démonstration.* Comme les dérivées partielles ne dépendent que de fonctions à 1 variable, on peut supposer $n = 1$.

Par le théorème de continuité si on pose $G(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)dt$, on sait que $G$ est continue sur $U$. Il nous reste à voir que pour $x_0 \in U$ :

$$\lim_{x \to x_0} \frac{F(x) - F(x_0)}{x - x_0} - G(x_0) = 0.$$

Soit $\epsilon > 0$. On fixe $r > 0$ tel que $B_F(x_0,r) \subset U$ qui est compact, de sorte que par le Th de Heine, $\frac{\partial f}{\partial x}$ est uniformément continue sur $B_F(x_0,r) \times [a,b]$.

On déduit donc qu'il existe $\eta > 0$ tel que si $||(x,t) - (x_0,t')|| < \eta$ alors $|\frac{\partial f}{\partial x}(x,t) - \frac{\partial f}{\partial x}(x_0,t')| < \epsilon/(b-a)$.

Par le TAF appliquée pour chaque $t$ à la fonction $\mathcal{C}^1$ $x \mapsto f(x,t)$, il existe $\xi(t)$ entre $x_0$ et $x$ tel que :

$$f(x,t) - f(x_0,t) = (x - x_0)\frac{\partial f}{\partial x}(\xi(t),t).$$

Si $|x - x_0| < \eta$ de sorte que $|\xi(t) - x_0| < \eta$ on en déduit :

$$|f(x,t) - f(x_0,t) - (x-x_0)\frac{\partial f}{\partial x}(x_0,t)| \le |x - x_0|\epsilon/(b-a).$$

Donc

$$|F(x) - F(x_0) - (x-x_0)G(x_0)| = \int_a^b dt|f(x,t) - f(x_0,t) - (x-x_0)\frac{\partial f}{\partial x}(x_0,t)| \le \epsilon|x - x_0|.$$

Ceci conclut. $\square$

# 2 Complément : Intégrales généralisées à valeur Evn complet et théorèmes de convergence (preuves facultatives ou admises)

## 2.1 Intégrale à valeur Evn complet

Soit $F$ un evn complet, $G$ evn. Soit $A \subset G$. On définit $E = \mathcal{C}M(A, F)$ l'ensemble des fonctions continues par morceaux sur $A$ à valeur $F$.

C'est un Evn pour la norme de la convergence uniforme, si $f \in E$ :

$$||f||_E = \sup_{t \in A} ||f(t)||_F.$$

On va utiliser le théorème suivant de prolongement des applications linéaires continues pour définir l'intégrale à valeur dans $F$.

**Proposition 86.** *Toute application linéaire $u$ d'un sous-espace vectoriel dense $D$ d'un evn $E$ vers un evn complet $F$ se prolonge en une unique application linéaire continue $v : E \to F$, ayant la même constante de lipschitzianité que $u$.*

*Démonstration.* Soit $x \in E$, et par densité $x_n \in D$, $x_n \to x$. Comme $u$ est continue donc $K$-lipschitzienne (par proposition 23), $u(x_n)$ est de Cauchy, donc converge vers $z \in F$ par complétude.

Soit $y_n \to x$ une autre telle suite, alors $||u(y_n)-z|| \leq ||u(x_n)-u(y_n)||+||u(x_n)-z|| \leq K||x_n-y_n||+ ||u(x_n)-z|| \to 0$. Donc la limite $z$ ne dépend pas de la suite choisie. On pose $v(x) = z$. En particulier, $v$ étend $u$ (en considérant la suite constante). Si $x_n \to x, y_n \to y$ en passant à la limite dans la relation $u(\alpha x_n + \beta y_n) = \alpha u(x_n) + \beta u(y_n)$, on déduit que $v$ est linéaire et avec $||u(x_n - y_n)|| \leq K||x_n - y_n||$, on déduit que $v$ est $K$-lipschitzienne. Par densité de $D$ on déduit l'unicité de $v$. □

Soit $\phi$ une fonction en escalier de $\phi : [a, b] \to F$ evn complet, c'est à dire qu'il existe une suite $a_0 = a < a_< ... < a_n = b$ telle que $\phi(x) = c_{i-1}$ pour $x \in [a_{i-1}, a_i)$. On définit $I(\phi) = \int_{[a,b]} \phi(t)dt = \sum_{i=1}^n (a_i - a_{i-1})\phi(a_{i-1})$. $I$ est une application linéaire continue, car par l'inégalité triangulaire

$$I(\phi) \leq \sum_{i=1}^{n} |a_i - a_{i-1}|||\phi(a_{i-1})||_F \leq ||\phi||_E|b - a|.$$

Comme les fonctions en escalier sont denses dans les fonctions continues par morceaux, la proposition précédente permet d'étendre l'intégrale comme quant $F = \mathbb{R}$ et on a :

**Définition 50.** L'intégrale des fonctions continues par morceaux est l'unique prolongement linéaire continu de l'intégrale des fonctions en escaliers, noté $\int_a^b dt f(t)dt$.

**Proposition 87.** *(Inégalité triangulaire)* $|| \int_a^b dt f(t)||_F \leq \int_a^b dt||f(t)||_F$.

*Démonstration.*

$$I(\phi) \leq \sum_{i=1}^{n} |a_i - a_{i-1}|||\phi(a_{i-1})||_F = \int_a^b ||\phi(t)||_F dt$$

pour $\phi$ en escalier et on prolonge par continuité. □

On a toutes les propriétés usuelles, Chasles, linéarité, en particulier si $F = \mathbb{R}^n$ et $f = (f_1, ..., f_n)$ $\int_a^b f(t)dt = (\int_a^b f_1(t)dt, ..., \int_a^b f_n(t)dt)$.

## 2.2 Intégrale généralisée des fonctions intégrables (Preuves admises)

Soit $I$ un intervalle de $\mathbb{R}$, $F$ un evn complet.

**Définition 51.** Une fonction $f : I \to F$ continue par morceaux est dite *intégrable* si il existe $M$ tel que pour tout segment $J = [a, b] \subset I$ on ait :

$$\int_a^b ||f(t)||_F dt \leq M.$$

On note alors une norme sur cet espace de fonctions intégrables : $||f||_1 = \int_I ||f(t)||dt = \sup_{J\subset I} \int_J ||f(t)||_F dt$.

**Proposition 88.** *(Intégrale des fonctions intégrables) Soit $f : I \to F$ intégrable. Pour toute suite croissante $J_n$ de segment de $I$ tel que $\cup J_n = I$, la suite $\int_{J_n} f(t)dt$ converge dans $F$. La limite ne dépend pas de la suite $J_n$, on l'appelle l'intégrale de $f$ sur $I$ et elle est notée (si $a = \inf I \in \mathbb{R}\cup\{-\infty\}$, $b = \sup I \in \mathbb{R} \cup \{+\infty\}$) :*

$$\int_I f = \int_a^b f(t)dt.$$

**Définition 52.** (Intégrale généralisée ou semi-convergente) Soit $I = [a, b[, b \in \mathbb{R}\cup\{\infty\}$ un intervalle, et $f : I \to F$ (généralement $F = \mathbb{R}$) $f$ continue par morceau. Si $\lim_{L\to b^-} \int_a^L f(t)dt$ existe, on dit que $\int_a^b f(t)dt$ converge et on pose :

$$\int_a^b f(t)dt = \lim_{L\to b} \int_a^L f(t)dt.$$

On définit de même avec les intervalles $]-\infty, b], ]a, b], ]-\infty, +\infty[$. On remarquera que pour les fonctions intégrables, l'intégrale généralisée coïncide avec la définition précédente. En pratique, si $F = \mathbb{R}$, on se ramène souvent à des fonctions intégrables après intégration par partie.

## 2.3 Théorèmes de convergences (Preuves admises)

Soit $I$ un intervalle de $\mathbb{R}$, $F$ un evn complet.

**Théorème 89.** *(Théorème de convergence monotone) Soit $f_n : I \to \mathbb{R}$ une suite croissante de fonctions intégrables qui converge simplement vers une fonction continue par morceaux $f$ sur $I$ (c'est à dire $\forall t \in I f_n(t) \to f(t)$.)*

*Alors la fonction $f$ est intégrable sur $I$ si et seulement si la suite $\int_I f_n$ est majorée et alors :*

$$\int_I f = \lim_{n\to\infty} \int_I f_n = \sup \int_I f_n.$$

**Théorème 90.** *(Théorème de convergence dominée) Soit $f_n : I \to F$ une suite de fonctions intégrables qui converge simplement vers une fonction continue par morceau $f$ (c'est à dire $\forall t \in I f_n(t) \to f(t)$.). On suppose qu'il existe $g : I \to \mathbb{R}_+$ intégrable telle que :*

$$\forall t \in I ||f_n(t)||_F \leq g(t) \quad \textit{(Condition de domination)},$$

*alors $f$ est intégrable et :*

$$\int_I f = \lim_{n\to\infty} \int_I f_n.$$

# 3 Intégrales généralisées dépendant d'un paramètre (Preuves facultatives)

Soient $E$ un evn et $G$ un evn complet. Soit $I$ un intervalle de $\mathbb{R}$. Soit finalement $A \subset E$ une partie de $E$.

**Définition 53.** Soit $f : A \times I \to G$. On suppose que pour tout $x \in X$, $t \mapsto f(x,t)$ est intégrable. Dans ce cas, on peut poser :

$F(x) = \int_I f(x,t)dt$. On définit ainsi une intégrale dépendant d'un paramètre la fonction $F : A \to G$.

**Théorème 91.** *(Théorème de continuité avec hypothèse de domination)*

*Soit $f : A \times I \to G$.*

*On suppose :*

1. *Pour tout $x \in A$, $t \mapsto f(x,t)$, est continue par morceau sur $I$.*
2. *Pour tout $t \in I$,$x \mapsto f(x,t)$ est continue en $x_0 \in A$.*
3. *(Hypothèse de domination) Il existe une fonction intégrable $g : I \to \mathbb{R}_+$ telle que*

$$\forall t \in I, \forall x \in A, \quad ||f(x,t)||_G \leq g(t).$$

*Alors la fonction $x \mapsto F(x) = \int_I f(x,t)dt$ est continue en $x_0$.*

On remarquera que dans l'hypothèse de domination, la fonction $g$ ne dépend pas de $x$.

*Démonstration.* L'hypothèse de domination garantit que $t \mapsto f(x,t)$ est intégrable. Soit $x_n \in A$ tel que $x_n \to x_0$. Par continuité de $x \mapsto f(x,t)$, pour chaque $t$, $f(x_n,t) \to f(x_0,t)$. On peut donc appliquer le théorème de convergence dominée pour conclure

$$\lim_{n\to\infty} \int_I f(x_n,t)dt = \int_I f(x_0,t)dt.$$

$\square$

*Exemple* 36. Soit $f : \mathbb{R} \to \mathbb{C}$ intégrable sur $\mathbb{R}$. Sa *transformée de Fourier* est définie par :

$$\hat{f}(x) = \int_{\mathbb{R}} f(t)e^{itx}dt.$$

Elle est continue sur $\mathbb{R}$ en utilisant une domination par $|f|$.

**Théorème 92.** *(Théorème de dérivabilité avec hypothèse de domination) Soit $f : U \times I \to \mathbb{R}^m$ avec $U \subset \mathbb{R}^n$ un ouvert.*

*On suppose :*

1. *Pour tout $x \in U$, $t \mapsto f(x,t)$, est intégrable sur $I$.*
2. *Pour tout $t \in I$, la fonction $x \mapsto f(x,t)$ admet une i-ème dérivée partielle sur $U$ et $t \mapsto \frac{\partial f}{\partial x_i}(x,t)$ est continue par morceaux.*
3. *(Hypothèse de domination) Il existe une fonction intégrable $g : I \to \mathbb{R}_+$ telle que*

$$\forall t \in I, \forall x \in U, \quad \left\| \frac{\partial f}{\partial x_i}(x,t) \right\| \leq g(t).$$

*Alors la fonction $x \mapsto F(x) = \int_I f(x,t)dt$ admet une i-ème dérivée partielle sur $U$ et :*

$$\frac{\partial F}{\partial x_i}(x) = \int_I \frac{\partial f}{\partial x_i}(x,t)dt.$$

*Démonstration.* On peut supposer $n = m = 1$. On fixe $x_0$ et montre la dérivabilité en $x_0$. On pose

$$h(x,t) = \frac{f(x,t) - f(x_0,t)}{x - x_0}, \quad \text{si } x \neq x_0 \quad \text{et} \quad h(x_0,t) = \frac{\partial f}{\partial x}(x,t).$$

Pour $x \neq x_0$,

$$\frac{F(x) - F(x_0)}{x - x_0} = \int_I h(x,t)dt.$$

Il suffit donc de prouver que $x \mapsto \int_I h(x,t)dt$ est continue en $x_0$. Par hypothèse, $t \mapsto h(x,t)$ est continue par morceau et $x \mapsto h(x,t)$ est continue. Enfin l'inégalité des accroissements finies donne, pour $x \neq x_0$ :

$$||h(x,t)|| \leq \sup_{u \in [x_0,x]} \left\| \frac{\partial f}{\partial x_i}(u,t) \right\| \leq g(t).$$

La même inégalité étant évidente en $x_0$, on a la condition de domination et le théorème de continuité conclut. □

**Corollaire 93.** *(Théorème de dérivation successive) Soit $f : U \times I \to I\!R^m$ avec $U \subset I\!R^n$ un ouvert une fonction $\mathcal{C}^k$ $(k \in I\!N \times \{\infty\})$.*

*On suppose qu'il existe $\phi_0, \phi_1, ..., \phi_k$ intégrables sur $I$ telles que*

$$\forall (i_1, ..., i_n), i_1 + ... + i_n = p \leq k, \forall x \in U \forall t \in I \left\| \frac{\partial^p f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x,t) \right\| \leq \phi_p(t).$$

*Alors la fonction $x \mapsto F(x) = \int_I f(x,t)dt$ est de classe $\mathcal{C}^k$ sur $U$ et pour $p = i_1 + ... + i_n \leq k$ :*

$$\frac{\partial^p F}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x) = \int_I \frac{\partial^p f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x,t)dt.$$

# 4 Un exemple : la fonction $\Gamma$

**Proposition 94.** *Pour $x \in I\!R_+^*$, la fonction $t \mapsto t^{x-1}e^{-t}$ est intégrable sur $I\!R_+^*$. La fonction $\Gamma$ est définie par :*

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt$$

*est de classe $\mathcal{C}^\infty$ sur $I\!R_+^*$ et l'on a :*

$$\Gamma^{(k)}(x) = \int_0^{+\infty} (\ln(t))^k t^{x-1}e^{-t}dt.$$

*De plus, pour $x > 0$ $\Gamma(x+1) = x\Gamma(x)$, pour $n \in I\!N$, $\Gamma(n+1) = n!$ et $\Gamma(1/2) = \sqrt{\pi}$.*

*Remarque* 54. $\Gamma$ est donc un prolongement de la factorielle (en fait elle est analytique sur $z$ $\mathrm{Re}(z) > 0$ ce qui montre que ceci la caractérise). En probabilité, la fonction $\Gamma$ intervient comme constante de normalisation dans la définition de la loi $\Gamma$, de densité $\frac{1}{\Gamma(n)} 1_{\{[0,\infty[\}}(x)t^{x-1}e^{-t}$ intervient comme la loi de la somme de $n$ variables exponentielles indépendantes (leur densité est $1_{\{[0,\infty[\}}(x)e^{-t}$).

*Démonstration.* Comme $(x,t) \mapsto e^{(x-1)\ln(t)}e^{-t}$ est $\mathcal{C}^\infty$ sur $(\mathbb{R}_+^*)^2$ on va appliquer le théorème de dérivation successive.

On doit donc dominer les dérivées pour $x \in ]\alpha,\beta[$ (ce qui suffit par localité de la différentiation) :

$$\frac{\partial^k}{\partial x^k} e^{(x-1)\ln(t)}e^{-t} = (\ln(t))^k t^{x-1}e^{-t}.$$

Or pour $t \leq 1$, $|(\ln(t))^k|t^{x-1}e^{-t} \leq (|\ln(t)|^k t^{x/2})t^{x/2-1} \leq C_1 t^{x/2-1}$ (en effet $(\ln(t))^k t^{x/2})$ est continue sur $[0,1]$ vu $x>0$ et la décroissance en $x$, en fait on peut calculer le maximum $(2k/x)^k e^{-k} \leq C_1 = (2k/\alpha)^k e^{-k}$).

De même, pour $t \geq 1$, $|(\ln(t))^k|t^{x-1}e^{-t} \leq (t^{x+k-1}e^{-t/2})e^{-t/2} \leq (t^{\beta+k-1}e^{-t/2})e^{-t/2} \leq C_2 e^{-t/2}$. ( $C_2$ existe car $(t^{\beta+k-1}e^{-t/2})$ est continue et tend vers 0 en $+\infty$, on peut prendre en calculant le max $C_2 = (\beta+2k)^{\beta/2+k}e^{-\beta/2-k}$)

En bilan Si on pose $g_k(t) = 1_{[0,1]}(t)C_1 t^{\alpha/2-1} + 1_{[1,\infty[}(t)C_2 e^{-t}/2$ qui est intégrable car $\alpha/2-1 > -1$, et on a la domination souhaitée : :

$$\left|\frac{\partial^k}{\partial x^k} e^{(x-1)\ln(t)}e^{-t}\right| \leq g_k(t).$$

On déduit donc que $\Gamma$ est $\mathcal{C}^\infty$ d'abord sur $]\alpha,\beta[$ puis comme $\alpha,\beta>0$ arbitraire sur $\mathbb{R}_+^*$.

Pour voir les identités, on intègre par partie : $\int_\alpha^\beta t^x e^{-t}dt = [-e^{-t}t^x]_\alpha^\beta + x\int_\alpha^\beta t^{x-1}e^{-t}dt$, donc pour $\alpha \to 0, \beta \to +\infty$ on obtient $\Gamma(x+1) = x\Gamma(x)$. La relation de la factorielle vient de $\Gamma(1)=1$.

En faisant le changement de variable $u^2 = t$, $dt = 2udu$ (puis en utilisant la parité et les coordonnées polaires, $\Gamma(\frac{1}{2}) = \int_0^{+\infty} t^{-1/2}e^{-t}dt = 2\int_0^{+\infty} e^{-u^2}du = \left(\int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy e^{-x^2-y^2}\right)^{1/2} = \left(\int_0^{2\pi} d\theta \int_0^{+\infty} dr e^{-r^2}2r/2\right)^{1/2} = \sqrt{\pi}$.

$\square$