CALCUL DIFFÉRENTIEL

L. H. ELIASSON

UNIVERSITÉ DE PARIS, 2020

1. Espaces vectoriels normés

Soit X un espace vectoriel (sur $\mathbb{R}$).

CM1

1.1. **Normes.**

Définition. Une fonction $N : X \to \mathbb{R}$ est une *norme sur* X ssi
(i) $N(x) = 0$ ssi $x = 0$ – *séparation*
(ii) $N(\lambda x) = |\lambda|\, N(x)\ \forall x \in X,\ \lambda \in \mathbb{R}$ – *homogénité*
(iii) $N(x+y) \le N(x) + N(y)\ \forall x, y \in X$ – *inégalité*-Δ.

Un *espace vectoriel normé* est un couple (X, N) où X est un espace vectoriel et N est une norme sur X.

Exemple. Soit $x = (x_1, \dots, x_m) \in \mathbb{R}^m$. Alors

$$\|x\|_1 = \sum_{1 \le k \le m} |x_k|$$

est une norme sur $\mathbb{R}^m$ – *la norme* l^1. Aussi

$$\|x\|_\infty = \max_{1 \le k \le m} |x_k|$$

est une norme sur $\mathbb{R}^m$ – *la norme* l^∞.

Notez, en particulier, que si la dimension $m = 1$, alors $\|x\|_1 = \|x\|_\infty$ n'est autre chose que la valeure absolue $|x|$.

Proposition 1.1. *Soit N une norme sur X. Alors*
(iv) $N(-x) = N(x)\ \forall x \in X$
(v) $N(x) \ge 0\ \forall x \in X$
(iii) $N(x-y) \ge |N(x) - N(y)|\ \forall x, y \in X$ – *inégalité-Δ renversée.*

Démonstration. (iv) $N(-x) = N((-1)x) = |-1|\, N(x) = N(x)$. (v) $0 = N(0) = N(x-x) \le N(x) + N(-x) = 2N(x)$ ce qu'implique que $N(x) \ge 0$. (vi) $N(x) = N(x-y+y) \le N(x-y) + N(y)$ ce que implique que $N(x-y) \ge N(x) - N(y)$. Interchangeant x et y donne aussi que $N(x-y) = N(y-x) \ge N(y) - N(x)$.

Date: 2 septembre 2021.

□

Proposition 1.2. *Soient X et Y deux e.v.'s, et soit $L : Y \to X$ une application linéaire et injective. Si N est une norme sur X, alors $N \circ L$ est une norme sur Y.*

Démonstration. (i) $N \circ L(x) = 0 \iff N(L(x))) = 0 \iff L(x) = 0 \iff x = 0$ où la dernière équivalence suivi parce que L est injective.
(ii) $N \circ L(\lambda x) = N(L(\lambda x)) = N(\lambda L(x)) = |\lambda| N(L(x)) = |\lambda| N \circ (x)$.
(iii) $N \circ L(x + y) = N(L(x + y)) = N(L(x) + L(y)) \leq N(L(x)) + N(L(y)) \leq N \circ L(x) + N \circ L(y)$. □

Exemple. L'application

$$L : \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \mapsto (x_1, x_2, x_3, x_4)$$

est une bijection linéaire entre l'espace vectoriel $\mathcal{M}_2$ des matrices 2×2 à coefficients réels et $\mathbb{R}^4$. Les applications

$$\mathcal{M}_2 \ni A \mapsto \|L(A)\|_p, \quad p = 1, \infty$$

sont donc deux normes sur $\mathcal{M}_2$.

1.1.1. *Normes euclidiennes.* Soit f un produit scalaire sur X, i.e. f est une application $X \times X \to \mathbb{R}$ t.q.

(i) f est bi-linéaire
(ii) f est symétrique
(iii) f est définie positive.

Proposition 1.3. *(Inégalité de Cauchy-Schwarz.) Soit f un produit scalaire sur X. Alors*

$$f(x, y)^2 \leq f(x, x) f(y, y) \quad \forall x, y \in X,$$

avec égalité ssi x et y sont co-linéaires.

Démonstration. Soient $x, y \in X$.

Supposons d'abord que $f(y, y) = 1$. Alors

$$\begin{aligned} 0 \leq f(x + \lambda y, x + \lambda y) &= f(x, x) + 2\lambda f(x, y) + \lambda^2 f(y, y) \\ &= (\lambda + f(x, y))^2 + f(x, x) - f(x, y)^2 \quad \forall \lambda \in \mathbb{R}. \end{aligned}$$

Si $\lambda = -f(x, y)$, alors

$$f(x, y)^2 \leq f(x, x) = f(x, x) f(y, y),$$

avec égalité ssi $f(x + \lambda y, x + \lambda y) = 0$, i.e. $x = -\lambda y$.

Supposons maintenant que $f(y, y) > 0$ et soit $\tilde{y} = \frac{1}{\sqrt{f(y,y)}} y$. Alors $f(\tilde{y}, \tilde{y}) = 1$ et, donc,

$$f(x, \tilde{y})^2 \leq f(x, x) = f(x, x) f(\tilde{y}, \tilde{y}),$$

avec égalité ssi x et $\tilde{y}$ sont co-linéaires.

Finalement, $f(y, y) = 0$, alors $y = 0$ et $f(x, y) = 0$.

□

Corollaire 1.4. *Si f est un produit scalaire sur X, alors*

$$N(x) = \sqrt{f(x,x)} \quad \forall x \in X$$

est une norme sur X – la norme sur X induite par f.

Démonstration. Les conditions (i) et (ii) sont immédiates. (iii) Par l'inégalité de Cauchy-Schwarz on a $N(x+y)^2 = f(x+y, x+y) = f(x,x) + f(y,y) + 2f(x,y) \leq N(x)^2 + N(y)^2 + 2N(x)N(y) = (N(x) + N(y))^2$. Comme $N \geq 0$ ceci donne $N(x+y) \leq N(x) + N(y)$.

□

Exemple. Soit $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$. Alors

$$\|x\|_2 = \sqrt{\sum_{1 \leq k \leq m} |x_k|^2}, \quad \forall x \in \mathbb{R}^m$$

est la norme sur $\mathbb{R}^n$ induite par le produit scalaire canonique

$$f(x,y) = \sum_{1 \leq k \leq n} x_k y_k \quad \forall x, y \in \mathbb{R}^m$$

– on l'appelle *la norme l^2*.

Définition. Une norme sur X est une *norme euclidienne* ssi elle est induite par un produit scalaire sur X.

Proposition 1.5. *Soit N une norme euclidienne sur X. Alors*

$$N(x+y)^2 + N(x-y)^2 = 2N(x)^2 + 2N(y)^2 \quad \forall x, y \in \mathbb{R}^n$$

– l'identité du parallélogramme.

Démonstration. Soit $N(x) = \sqrt{f(x,x)}$ où f est un produit scalaire sur X. Alors

$$\begin{aligned} N(x+y)^2 + N(x-y)^2 &= f(x+y, x+y) + f(x-y, x-y) \\ &= f(x,x) + 2f(x,y) + f(y,y) + f(x,x) + 2f(x,-y) + f(-y,-y) \\ &= f(x,x) + 2f(x,y) + f(y,y) + f(x,x) - 2f(x,y) + f(y,y) \\ &= 2f(x,x) + 2f(y,y) = 2N(x)^2 + 2N(y)^2. \end{aligned}$$

□

Remarque. La réciproque de cette proposition est aussi vraie : si une norme vérifie l'identité du parallélogramme, alors elle est une norme euclidienne.

Les normes l^1 et l^∞ ne sont pas des normes euclidiennes sur $\mathbb{R}^m$. En effet, si $x = (1, 0, \ldots, 0)$ et $y = (0, \ldots, 0, 1)$, alors

$$\|x+y\|_1^2 + \|x-y\|_1^2 = 8 \neq 4 = 2\|x\|_1^2 + 2\|y\|_1^2$$

et

$$\|x+y\|_\infty^2 + \|x-y\|_\infty^2 = 2 \neq 4 = 2\|x\|_\infty^2 + 2\|y\|_\infty^2 .$$

1.1.2. *Equivalence.*

Définition. Deux normes N_1 et N_2 sur X sont dite *équivalentes* ssi il existe une constante $C \geq 1$ t.q.

$$\frac{1}{C} N_2(x) \leq N_1(x) \leq C N_2(x) \quad \forall x \in X.$$

Notez bien que "d'être équivalentes" est une relation d'équivalence. En particulier, si N_1 et N_2 sont équivalentes et si N_2 et N_3 sont équivalentes, alors N_1 et N_3 sont équivalentes ;

Proposition 1.6. *On a*

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{m}\,\|x\|_2 \leq m\,\|x\|_\infty \quad \forall x \in \mathbb{R}^m.$$

En particulier, les normes l^1, l^2 et l^∞ sur $\mathbb{R}^m$ sont toutes équivalentes entre elles.

Démonstration. Les inégalités $\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1$ sont immédiates. L'inégalité $\|x\|_1 \leq m\,\|x\|_\infty$ est aussi immédiate, et elle donne $\|x\|_2 \leq \sqrt{m}\,\|x\|_\infty$.

On démontre (par induction sur m) que

$$m\,\|x\|_2^2 = \sum_{1\leq i<j\leq m} (x_i^2 + x_j^2) + \sum_{1\leq j\leq m} x_j^2.$$

Comme $a^2 + b^2 \geq 2ab$ on a

$$\sum_{i<j}(x_i^2 + x_j^2) + \sum_j x_j^2 \geq 2\sum_{i<j} x_i x_j + \sum_j x_j^2 = (x_1 + \cdots + x_m)^2 = \|x\|_1^2 .$$

D'où $\|x\|_1 \leq \sqrt{m}\,\|x\|_2$.

□

Sur certain e.v. il existe des normes non-équivalentes.

Exemple. Soit $X = \mathcal{C}^0([0,1])$ l'espace de toutes les fonctions $f : [0,1] \to \mathbb{R}$ qui sont continues. Soient

$$\|f\|_{L^1} = \int_0^1 |f(t)|\, dt$$

et

$$\|f\|_{\mathcal{C}^0} = \sup_{0\le t\le 1} |f(t)| \,.$$

Alors $\| \ \|_{L^1}$ et $\|f\|_{\mathcal{C}^0}$ sont deux normes sur X qui, en plus, vérifie

$$\|f\|_{L^1} \le \|f\|_{\mathcal{C}^0} \quad \forall f \in X.$$

Mais elles ne sont pas équivalentes. En effet, si $f_n(t) = t^n$ alors

$$\|f_n\|_{L^1} = \frac{1}{n+1} \to 0, \quad n \to \infty$$

tandis que $\|f_n\|_{\mathcal{C}^0} = 1 \ \forall n$.

Dans cet exemple l'espace vectoriel est de dimension ∞ et ceci est essentiel parce que

Théorème. *Toutes les normes sur un espaces vectoriel X de dimension finie sont équivalentes.*

On acceptera ce résultat qui sera démontré dans le cours de topologie.

1.1.3. *Les normes l^p, $1 \le p < \infty$.* Soit $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$. Alors

$$\|x\|_p = (\sum_{1\le k\le m} |x_k|^p)^{\frac{1}{p}}, \quad \forall x \in \mathbb{R}^m$$

est une norme sur $\mathbb{R}^m$ – on l'appelle *la norme l^p*. La séparation et l'homogénité sont faciles, mais l'inégalité-Δ est plus difficile de montrer.

1.2. **Les ouverts.** Soit N une norme sur X.

1.2.1. *Les boules.*

Définition. Soit $a \in X$ et de rayon $r \ge 0$. L'ensemble

$$B_<(a,r) = \{x \in X : N(x-a) < r\}$$

s'appelle *la boule stricte* (p.r.à N) de centre a et de rayon r. L'ensemble

$$B_\le(a,r) = \{x \in X : N(x-a) \le r\}$$

s'appelle *la boule élargie* (p.r.à N) de centre a et de rayon r.

Exemple. Si $X = \mathbb{R}$, alors chaque boule est un *intervalle*. Par exemple, si N est la norme “valeure absolue”, alors $B_<(a,r)$ est l'intervalle ouvert $]a-r, a+r[$ et $B_\le(a,r)$ est l'intervalle fermé $[a-r, a+r]$.

Nota Bene. Les boules dans un e.v. X sont toujours définies p.r. à une norme N qui n'est pas explicitée, mais sous-entendues, dans les notations $B_<(a,r)$ et $B_\le(a,r)$. Une notation sans ambiguité serait $B^N_<(a,r)$ et $B^N_\le(a,r)$.

Exemple. Si $X = \mathbb{R}^2$, alors la forme d'une boule dépend de la norme : si N est la norme l^2, alors chaque boule est un *disque* ; si N est la norme l^∞, alors chaque boule est un *cube* ; si N est la norme l^1, alors chaque boule est un *losange*.

C'est immédiat que pour $a, b \in X$ et tout $r \geq 0$

$$B_<(a+b, r) = b + B_<(a, r),$$

et la même chose pour les boules élargies.

Proposition 1.7. *Soit $a \in X$.*
(i) Pour tout $t > r \geq s > 0$

$$B_<(a, s) \subseteqq B_<(a, r) \subseteqq B_\leq(a, r) \subseteqq B_<(a, t).$$

(ii) Pour tout $r \geq 0$

$$\bigcup_{s<r} B_\leq(a, s) = B_<(a, r) \subseteqq B_\leq(a, r) = \bigcap_{t>r} B_<(a, t).$$

Démonstration. (i) On a

$$N(a-x) < s \implies N(a-x) < r \implies N(a-x) \leq r \implies N(a-x) < t.$$

(ii) On a, par (i), que

$$\bigcup_{s<r} B_\leq(a, s) \subseteqq B_<(a, r).$$

D'autre part, si $x \in B_<(a, r)$, alors $r_x = N(a - x) < r$ et donc

$$x \in B_\leq(a, r_x) \subseteqq \bigcup_{s<r} B_\leq(a, s).$$

On a, par (i), que

$$B_\leq(a, r) \subseteqq \bigcap_{t>r} B_<(a, t).$$

D'autre part, si $x \in \bigcap_{t>r} B_<(a, t)$, alors $N(a - x) < t$ pour tout $t > r$. Ceci implique que $N(a - x) \leq r$, i.e.

$$x \in B_\leq(a, r).$$

$\square$

1.2.2. *Les ouverts.*

Définition. Une partie A de X est dite *ouverte* ou *un ouvert* (p.r.à N) ssi pour tout $a \in A$ il existe $r > 0$ t.q.

$$B_<(a,r) \subset A.$$

Notez que les parties $A = \emptyset$ et $A = X$ sont ouvertes (p.r.à n'importe quelle norme sur X).

Proposition 1.8. *$B_<(a,r)$ est ouverte pour tout $a \in X$, $r > 0$.*

Les boules strictes sont d'habitude appelées les *boules ouvertes*.

Démonstration. Soit $b \in B_<(a,r)$ et soit $s = r - N(a-b)$. Alors $s > 0$ et

$$B_<(b,s) \subset B_<(a,r).$$

En effet, si $x \in B_<(b,s)$, alors

$$N(a-x) = N(a-b+b-x) \leq N(a-b)+N(b-x) < N(a-b)+s = r,$$

i.e. $x \in B_<(a,r)$. □

Exemple. Soit $X = \mathbb{R}$ et soit N est la norme "valeure absolue". Alors (p.r. à N) $]a,b[$ est ouverte parce que $= B_<(\frac{a+b}{2}, \frac{b-a}{2})$. Par contre, $[a,b[$ n'est pas ouverte parce que une boule $B_<(a,r)$ n'est jamais (pour $r > 0$) $\subseteqq [a,b[$.

Les parties $[a,b]$ et $]a,b]$ ne sont pas ouvertes non plus.

Proposition 1.9.
(i) Une réunion (même non dénombrable) d'ouverts est ouverte.
(ii) Une intersection finie d'ouverts est ouverte.

Démonstration. Soient A_ι un ouvert dans X pour chaque $\iota \in I$.

(i) Si $x \in \bigcup_{\iota \in I} A_\iota$, alors $x \in A_{\iota_x}$ pour (au moins) un $\iota_x \in I$. Comme A_{ι_x} est un ouvert, il existe $r_x > 0$ t.q.

$$B_<(x, r_x) \subseteqq A_{\iota_x} \subseteqq \bigcup_{\iota \in I} A_\iota.$$

(ii) Si $x \in \bigcap_{\iota \in I} A_\iota$, alors $x \in A_\iota$ pour tout $\iota \in I$. Comme les A_ι sont des ouverts, il existe $r_\iota > 0$ t.q.

$$B_<(x, r_\iota) \subseteqq A_\iota \quad \forall \iota \in I.$$

Si

$$r = \inf_{\iota \in I} r_\iota,$$

alors
$$B_<(x,r) \subseteqq \bigcap_{\iota\in I} A_\iota.$$
Il se peut que $r=0$, mais ceci n'arrive pas si I est un ensemble fini. $\square$

Exemple. Soit $X=\mathbb{R}$ et soit N est la norme "valeure absolue". Alors (p.r.à N) $]a,\infty[$ est ouverte parce que
$$= \bigcup_{b>a}]a,b[.$$
La partie $]-\infty,b[$ est aussi ouverte.

Notez que
$$\bigcap_{n\geq 1}]0,1+\frac{1}{n}[=]0,1].$$
Voici donc une intersection (non-finie) d'ouverts qui n'est pas ouverte.

CM2

1.2.3. *Equivalence.* Une partie de X peut être ouvert p.r.à une norme, sans être ouverte p.r.à une autre norme.

Exemple. Soit $X=\mathcal{C}^0([0,1])$. Soient $N_1=\|\ \|_{\mathcal{C}^0}$ et $N_2=\|\ \|_{L^1}$. Alors $A=B_<^{N_1}(0,1)$ est ouverte p.r.à N_1 par Proposition (1.8). Mais A n'est par ouverte p.r.à N_2.

Mais ceci ne peut pas arriver si les deux normes sont équivalentes.

Lemme. *Soient N_1 et N_2 deux normes sur X et supposons qu'il existe $C>0$ t.q.*
$$N_2(x)\leq CN_1(x) \quad \forall x\in X.$$
Alors
$$B_<^{N_1}(a,\frac{r}{C}) \subseteqq B_<^{N_2}(a,r) \quad \forall a,r.$$

Démonstration.
$$x\in B_<^{N_1}(a,\frac{r}{C}) \iff N_1(a-x)<\frac{r}{C} \implies N_2(a-x)<r \iff x\in B_<^{N_2}(a,r).$$
$\square$

Proposition 1.10. *Soient N_1 et N_2 deux normes équivalentes sur X. Alors une partie A de X est ouverte p.r.à N_1 ssi elle est ouverte p.r.à N_2.*

Démonstration. Soit A ouverte p.r.à N_2, et soit $a\in A$. Alors il existe $r>0$ t.q.
$$B_<^{N_2}(a,r)\subseteqq A.$$
Comme N_1 et N_2 sont équivalentes, il existe $C>0$ t.q.
$$N_2(x)\leq CN_1(x) \quad \forall x\in X.$$

Par le Lemme

$$B_<^{N_1}(a, \frac{r}{C}) \subseteqq B_<^{N_2}(a, r) \subseteqq A.$$

A est donc ouverte p.r.à N_1. □

Dans un e.v. de dimension finie on peut donc parler des parties ouvertes *sans référence à une norme quelconque* parce que toutes les normes sont équivalentes. Par contre, dans un e.v. de dimension ∞ il est essentiel de préciser de quelle norme on parle.

1.2.4. *L'intérieure.*

Définition. Soit A une partie de X. *L'intérieure de* A (p.r.à N) est l'ensemble

$$\mathrm{int}(A) = \bigcup_{\substack{B \subseteq A \\ B \text{ ouvert}}} B,$$

i.e. c'est la réunion de tous les ouverts (p.r.à N) contenus dans A.

Notez bien que l'intérieure dépend du choix de norme N qui n'est pas explicité, mais sous-entendu, dans la notation $\mathrm{int}(A)$.

Par Proposition (1.9) $\mathrm{int}(A)$ est lui-même un ouvert, c'est donc *le plus grand ouvert contenu dans* A.

Exemple. Soit $X = \mathbb{R}$ et soit N est la norme "valeure absolue". Alors (p.r. à N) les ensembles $]0,1[$, $[0,1[$, $[0,1]$ et $]0,1]$ ont tous la même intérieure, à savoir $]0,1[$.

Proposition 1.11.

$$int(B_\leq(a, r)) = B_<(a, r).$$

En particulier, $B_\leq(a, r)$ *n'est pas un ouvert.*

Démonstration. Comme $B_<(a, r)$ est un ouvert qui est contenu dans $B_\leq(a, r)$ on a que

$$B_<(a, r) \subseteqq \mathrm{int}(B_\leq(a, r)).$$

Supposons qu'il existe $b \in \mathrm{int}(B_\leq(a, r)) \setminus B_<(a, r)$. Alors $N(b - a) = r$ et (comme $\mathrm{int}(B_\leq(a, r))$ est ouverte) il existe $s > 0$ t.q.

$$B_<(b, s) \subseteqq \mathrm{int}(B_\leq(a, r)) \subseteqq B_\leq(a, r).$$

Soit $x = b + t(b - a)$ avec $t = \frac{s}{2r}$. Alors $N(b - x) = \frac{s}{2} < s$, i.e. $x \in B_<(b, s)$, et en même temps, $N(a - x) = (1 + t)r > r$, i.e. $x \notin B_\leq(a, r)$. Voilà une contradiction. □

1.3. **Les fermés.** Soit N une norme sur X.

Définition. Une partie A de X est dite *fermée* ou *un fermé* (p.r.à N) ssi le complementaire $X \setminus A$ est un ouvert (p.r. à N).

Notez que les parties $A = \emptyset$ et $A = X$ sont fermées (p.r.à n'importe quelle norme sur X).

Exemple. Soit $X = \mathbb{R}$ et soit N est la norme "valeure absolue". Alors (p.r. à N) $[0, 1]$ est fermée, mais $[0, 1[$, $]0, 1[$ et $]0, 1]$ ne le sont pas.

Proposition 1.12. *$B_{\leq}(a, r)$ est fermée*

Les boules élargies sont d'habitude appelées les *boules fermées.*

Démonstration. Il s'agit de montrer que $X \setminus B_{\leq}(a, r)$ est ouverte.
Soit $b \in X \setminus B_{\leq}(a, r)$ et soit $s = N(a - b) - r$. Alors $s > 0$ et

$$B_{<}(b, s) \subset X \setminus B_{\leq}(a, r).$$

En effet, si $x \in B_{<}(b, s)$, alors

$$N(a-x) = N(a-b+b-x) \geq N(a-b) - N(b-x) > N(a-b) - s = r,$$

i.e. $x \notin B_{\leq}(a, r)$.

□

Proposition 1.13.
(i) Une intersection (même non dénombrable) de fermés est fermée.
(ii) Une réunion finie de fermés est fermée.

Démonstration. Ceci suit de la Proposition (1.9) et les relations

$$X \setminus \bigcap_{\iota} A_{\iota} = \bigcup_{\iota} X \setminus A_{\iota}$$

et

$$X \setminus \bigcup_{\iota} A_{\iota} = \bigcap_{\iota} X \setminus A_{\iota}.$$

□

Notez que la réunion (non finie) $\bigcap_{s<r} B_{\leq}(a, s)$ n'est pas fermée.

1.3.1. *L'adhérence.*

Définition. Soit A une partie de X. *L'adhérence de A* (p.r.à N) est l'ensemble

$$\mathrm{adh}(A) = \bigcap_{\substack{B \supseteq A \\ B \text{ fermé}}} B,$$

i.e. c'est l'intersection de tous les fermés (p.r.à N) qui contiennent A.

Notez bien que l'adhérence dépend de la norme N qui n'est pas explicitée, mais sous-entendue, dans la notation $\mathrm{adh}(A)$.

Par Proposition (1.13) $adh(A)$ est lui-même un fermé, c'est donc *le plus petit fermé qui contient A.*

Exemple. Soit $X = \mathbb{R}$ et soit N est la norme "valeure absolue". Alors (p.r.à N) les ensembles $]0,1[$, $[0,1[$, $[0,1]$ et $]0,1]$ ont tous la même adhérence, à savoir $[0,1]$.

Proposition 1.14.

$$adh(B_<(a,r)) = B_\leq(a,r).$$

En particulier, $B_<(a,r)$ n'est pas un fermé.

Démonstration. Comme $B_\leq(a,r)$ est un fermé qui contient dans $B_<(a,r)$ on a que

$$\mathrm{adh}(B_<(a,r)) \subseteqq B_\leq(a,r).$$

Supposons qu'il existe $b \in B_\leq(a,r) \setminus \mathrm{adh}(B_<(a,r))$. Alors $N(b-a) = r$ et il existe $s > 0$ t.q.

$$B_<(b,s) \subseteqq X \setminus \mathrm{adh}(B_<(a,r)) \subseteqq X \setminus B_<(a,r).$$

Soit $x = b - t(b-a)$ avec $t = \frac{s}{2r}$. Alors $N(b-x) = \frac{s}{2} < s$, i.e. $x \in B_<(b,s)$, et en même temps, $N(a-x) = (1-t)r < r$, i.e. $x \in B_<(a,r)$. Voilà une contradiction.

$\square$

1.3.2. *Les parties ouvertes et fermées.* Il y a dans X des parties qui sont ouvertes mais pas fermées, comme les boules strictes.

Il y a dans X des parties qui sont fermées mais pas ouvertes, comme les boules élargies.

Il y a aussi dans X des parties qui ne sont ni ouvertes ni fermées, comme chaque partie A qui vérifie

$$B_<(a,r) \subsetneqq A \subsetneqq B_\leq(a,r).$$

Il y a finalement aussi des parties qui sont ouvertes et fermées, comme $\emptyset$ et X. Est-ce qu'il y en a d'autres ?

Théorème. *Dans un espace vectoriel normé (X,N) les seules parties qui sont à la fois ouvertes et fermées sont $\emptyset$ et X.*

On acceptera ce résultat qui sera démontré dans le cours de topologie.

1.4. **Convergence des suites dans un espace vectoriel normé.** Une *suite dans* l'espace vectoriel X n'est autre chose qu'une application $\alpha : \mathbb{N} \to X$. On note souvent cette suite

$$(\alpha(n))_{n\in\mathbb{N}}.$$

La suite α est *à valeurs dans* $A \subseteqq X$ ssi $\alpha(n) \in A$ pour chaque n.

Soit N une norme sur X.

Définition. Une suite α dans X *converge vers* $a \in X$ *(p.r.à* N*)* ssi

$$\lim_{n\to\infty} N(\alpha(n) - a) = 0.$$

On dit aussi que la suite α *converge vers* a *dans l'e.v.n.* (X, N).

Exemple. Dans $\mathbb{R}^2$, la suite

$$\alpha(n) = (1 + \frac{n}{n+1}, (-1)^n \frac{1}{n})$$

converge vers $(2, 0)$ (p.r.à la norme l^2). En effet

$$\|\alpha(n) - (2,0)\|_2 = \sqrt{(1 + \frac{n}{n+1} - 2)^2 + ((-1)^n \frac{1}{n} - 0)^2} \to 0, \quad n \to \infty.$$

Proposition 1.15. *Si la suite* α *converge vers* a *et vers* b *(p.r.à* N*) , alors* $a = b$*. On dit que* a *est la limite (p.r.à* N*) de la suite* α*.*

Démonstration.

$$N(a-b) = N(a-\alpha(n)+\alpha(n)-b) \le N(a-\alpha(n))+N(\alpha(n)-b) \to 0, \quad n \to \infty.$$

Donc, $N(a - b) = 0$ ce qui implique que $a - b = 0$, i.e. $a = b$. □

Proposition 1.16. *Si* α *et* β *sont deux suites dans* X *qui converge vers* a *et vers* b *(p.r.à* N*) respectivement, alors* $\alpha \pm \beta$ *convergent vers* $a \pm b$ *(p.r.à* N*) et, pour tout* $\lambda \in \mathbb{R}$*,* $\lambda\alpha$ *converge vers* λa *(p.r.à* N*).*

Démonstration.

$$N((a + b) - (\alpha(n) + \beta(n)) = N((a - \alpha(n)) + (b - \beta(n)))$$

$$\le N((a - \alpha(n))) + N(b - \beta(n)) \to 0, \quad n \to \infty.$$

Les preuves des autres énoncés sont similaires. □

1.4.1. *La norme de la convergence uniforme.* Soit I un intervalle dans $\mathbb{R}$ et soit $\mathcal{C}^0(I)$ l'espace de toutes les fonctions continues $f : I \to \mathbb{R}$. Alors

$$\|f\|_{\mathcal{C}^0} = \sup_{t\in I} |f(t)|$$

est une norme sur $\mathcal{C}^0(I)$.

Une suite α dans $\mathcal{C}^0(I)$ converge vers f p.r.à la norme $\| \ \|_{\mathcal{C}^0}$ ssi les fonctions $f_n = \alpha(n)$ converge uniformement (sur I) vers f. En effet

$$\lim_{n\to\infty} \|f - \alpha(n)\|_{\mathcal{C}^0} = 0 \iff \lim_{n\to\infty} \sup_{t\in I} |f(t) - f_n(t)| = 0.$$

Exemple. Dans $\mathcal{C}^0([0,1])$, la suite

$$\alpha = (f_n)_n, \quad f_n(t) = t^n$$

converge vers $f = 0$ p.r.à la norme $\| \ \|_{L^1}$. En effet

$$\|f_n - f\|_{L^1} = \int_0^1 t^n dt = \frac{1}{n+1} \to 0, \quad n \to \infty.$$

Mais elle ne converge pas vers $f = 0$ p.r.à la norme $\| \ \|_{\mathcal{C}^0}$ parce que la suite de fonctions $(f_n)_n$ ne converge pas uniformement vers 0 sur $[0,1]$.

1.4.2. *Convergence et adhérence.*

Lemme. *Si $a \in adh(A)$, alors*

$$B_<(a, \frac{1}{n}) \cap A \neq \emptyset \quad \forall n \geq 1.$$

Démonstration. Si, pour un $n \geq 1$, $B_<(a, \frac{1}{n}) \cap A = \emptyset$ alors $X \setminus B_<(a, \frac{1}{n})$ est un fermé qui contient A mais qui ne contient pas a. Comme adh(A) est le plus petit fermé qui contient A, alors adh(A) ne peut pas contenir a. $\square$

Proposition 1.17. *Soit A une partie de X et soit $a \in X$. Alors $a \in$ adh(A) (p.r.à N) ssi il existe une suite α à valeurs dans A et qui converge (p.r.à N) vers a.*

Démonstration. Si $a \notin$ adh(A), alors $a \in X \setminus$adh(A) qui est un ouvert. Il existe donc $r > 0$ t.q.

$$B_<(a, r) \subsetneqq X \setminus \text{adh}(A) \subsetneqq X \setminus A.$$

Ceci implique qu'aucune suite qui prend ses valeurs dans A peut converger vers a.

Si $a \in$ adh(A), alors (par le Lemme)

$$B_<(a, \frac{1}{n}) \cap A \neq \emptyset \quad \forall n \geq 1.$$

Soit $\alpha(n)$ un élément dans $B_<(a, \frac{1}{n}) \cap A$ pour chaque n. Alors la suite α prend ses valeurs dans A et converge (p.r.à N) à $a \notin A$. $\square$

Proposition 1.18. *Une partie A dans X est fermée (p.r.à N) ssi chaque suite à valeurs dans A et qui converge (p.r.à N) a sa limite dans A.*

Démonstration. Si A est fermée, alors $B = X \setminus A$ est ouverte et, donc, pour chaque $a \in B$ il existe $r_a > 0$ t.q.

$$B_<(a, r_a) \subsetneqq B,$$

i.e.

$$B_<(a, r_a) \cap A = \emptyset.$$

Aucune suite qui prend ses valeurs dans A peut donc converger vers un $a \notin A$.

Si A n'est pas fermée, alors il existe $a \in \mathrm{adh}(A) \setminus A$. Par la Proposition (1.17) il existe donc une une suite qui prend ses valeurs dans A et qui converge vers $a \notin A$. Alors la suite α prend ses valeurs dans A et converge (p.r.à N) à $a \notin A$. $\square$

Exemple. Dans $\mathbb{R}^2$, l'ensemble

$$A = \{(x_1, x_2) : \ x_1 \geq 1, \ x_2 \geq 0\}$$

est un fermé (p.r.à $\| \ \|_2$). En effet, soit $\alpha = (\alpha_1, \alpha_2)$ une suite à valeurs dans A qui converge vers $a = (a_1, a_2) \in \mathbb{R}^2$ (p.r.à $\| \ \|_2$). Alors

$$\|\alpha(n) - a\|_2 = \sqrt{(\alpha_1(n) - a_1)^2 + (\alpha_2(n) - a_2)^2} \to 0, \quad n \to \infty.$$

Ceci implique que

$$a_1 = \lim_{n\to\infty} \alpha_1(n) \quad \text{et} \quad a_2 = \lim_{n\to\infty} \alpha_2(n).$$

Comme α est à valeurs dans A on a $\alpha_1(n) \geq 1$ et $\alpha_2(n) \geq 0$ pour tout n. Donc $a_1 \geq 1$ et $a_2 \geq 0$, i.e. $a \in A$. A est donc fermé par Proposition (1.18).

L'adhérence (p.r.à $\| \ \|_2$) de

$$B = \{(x_1, x_2) : \ x_1 \geq 1, \ x_2 > 0\}$$

est $= A$. En effet, chaque $a = (a_1, 0)$, $a_1 \geq 1$, est la limite (p.r.à $\| \ \|_2$) de la suite

$$\alpha(n) = (a_1, \frac{1}{n+1}) \in A.$$

Par Proposition (1.17) on a $\mathrm{adh}B \supseteqq A \supseteqq B$, et, comme A est fermé, $\mathrm{adh}B = A$.

CM3

1.4.3. *Convergence et équivalence.*

Proposition 1.19. *Soit α une suite dans X qui converge vers $a \in X$ (p.r.à N). Si N' est une autre norme sur X qui est équivalente à N, alors α converge vers a (p.r.à N').*

Démonstration. Il existe $C > 0$ t.q.

$$N'(x) \leq CN(x) \quad \forall x \in X.$$

On a α converge vers a (p.r.à N) ssi

$$\lim_{n\to\infty} N(a - \alpha(n)) = 0 \implies \lim_{n\to\infty} N'(a - \alpha(n)) = 0$$

ssi α converge vers a (p.r.à N'). □

Dans un e.v. de dimension finie on peut donc parler de convergence *sans référence à une norme quelconque* parce que toutes les normes sont équivalentes. Par contre dans un e.v. de dimension ∞ il est essentiel de préciser de quelle norme on parle.

Proposition 1.20. *Soit $\alpha = (\alpha_1, \ldots, \alpha_m)$ une suite dans $\mathbb{R}^m$. Alors α converge vers $a = (a_1, \ldots, a_m) \in \mathbb{R}^m$ ssi*

$$\lim_{n\to\infty} \alpha_j(n) = a_j$$

pour chaque $j = 1, \ldots, m$.

Démonstration. On considère la norme l^∞. La suite α converge vers a (p.r.à $\| \ \|_\infty$) ssi

$$\lim_{n\to\infty} \|a - \alpha(n)\|_\infty \iff \lim_{n\to\infty} \max_i |a_i - \alpha_i(n)| = 0$$

ssi

$$\lim_{n\to\infty} |a_i - \alpha_i(n)| = 0 \quad \forall i.$$

□

1.5. **Continuité.** Soient X et Y deux espaces vectoriels, et soit

$$f : U \to Y$$

une fonction définie dans une partie $U \subset X$. On se donne deux normes, N_X sur X et N_Y sur Y.

Définition. La fonction f est *continue en* $a \in U$ (p.r.à N_X, N_Y) ssi pour chaque $\varepsilon > 0$ il existe $\delta > 0$ t.q.

$$N_X(x - a) < \delta,\ x \in U \implies N_Y(f(x) - f(a)) < \varepsilon.$$

On dit que f est *continue* (p.r.à N_X, N_Y) ssi f est continue en chaque $a \in U$ (p.r.à N_X, N_Y)

Notez que si $X = Y = \mathbb{R}$ et $N_X = N_Y$ est la norme “valeur absolue”, alors c'est la définition habituelle de continuité.

Notez aussi que chaque fonction constante est trivialement continue en chaque point $a \in U$.

Voici quelques exemples plus intéressants.

Exemple. Soit $X = \mathbb{R}$ avec N_X la norme “valeur absolue”, et soient u et $v \neq 0$ deux vecteurs dans Y. Alors la fonction

$$f(t) = u + tv, \quad t \in \mathbb{R},$$

est continue (p.r.à N_X, N_Y). En effet

$$N_Y(f(t) - f(t_0)) = N_Y((t - t_0)v) = |t - t_0|\, N_Y(v),$$

et ceci est $< \varepsilon$ dès que $|t - t_0| < \delta = \frac{\varepsilon}{N_Y(v)}$.

Exemple. Soit $X = \mathbb{R}^m$ avec $N_X = \| \ \|_2$. Alors, pour $j = 1, \ldots, m$, la fonction

$$\pi_j(x_1, \ldots, x_m) = x_j$$

– la *projection sur le j-ième coefficient* – est continue (p.r.à $\| \ \|_2, | \ |$). En effet

$$|\pi_j(x) - \pi_j(a)| = |x_j - a_j| \leq \|x - a\|_2$$

et ceci est $< \varepsilon$ dès que $\|x - a\|_2 < \delta = \varepsilon$.

Exemple. Soit $X = Y = \mathcal{M}_2$ muni de la norme

$$N(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = |a| + |b| + |c| + |d|\,.$$

Alors la fonction

$$f(A) = A^2$$

est continue (p.r.à N, N). En effet

$$N(f(B) - f(A)) = N(B^2 - A^2) = N(A(A - B) + (A - B)B)$$

$$\leq (N(A) + N(B))N(A - B) \leq 2N(A)N(A - B) + N(A - B)^2$$

et ceci est $< \varepsilon$ dès que $N(A - B) < \delta = \sqrt{N(A)^2 + \varepsilon} - N(A)$.

Proposition 1.21. *f est continue en $a \in U$ (p.r.à N_X, N_Y) ssi pour chaque suite α, à valeurs dans U, qui converge vers a (p.r.à N_X), la suite image $f \circ \alpha$, converge vers $f(a)$ (p.r.à N_Y).*

Démonstration. Soit α une suite à valeurs dans U qui converge vers $a \in U$ (p.r.à N_X), et soit $\varepsilon > 0$. Si f est continue en a, il existe $\delta = \delta_\varepsilon > 0$ t.q.

$$N_X(x - a) < \delta, \ x \in U \implies N_Y(f(x) - f(a)) < \varepsilon.$$

Comme α converge vers a il existe $N = N_\delta$ t.q.

$$N_X(\alpha(n) - a) < \delta \quad \forall n \geq N.$$

Donc

$$N_Y(f(\alpha(n)) - f(a)) < \varepsilon \quad \forall n \geq N,$$

i.e. la suite $f \circ \alpha$ converge vers $f(a) \in U$ (p.r.à N_Y).

Réciproquement, supposons que f n'est pas continue en a. Alors il existe $\varepsilon > 0$ t.q. pour tout $\delta > 0$ il existe x_δ t.q. $N_X(x_\delta) - a) \leq \delta$ et $N_Y(f(x_\delta)) - f(a)) > \varepsilon$. On choisit $\delta = \frac{1}{n+1}$ et pour chaque $n \in \mathbb{N}$ un tel x_n. Alors la suite $\alpha(n) = x_n$ converge vers a, tandis que la suite $f \circ \alpha$ ne converge pas vers $f(a)$.

□

Proposition 1.22. *Si $f, g : U \to Y$ sont continues en $a \in U$ (p.r.à N_X, N_Y), alors, pour chaque $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g : U \to Y$ est continue en $a \in U$ (p.r.à N_X, N_Y).*

Démonstration. Ceci suite des Propositions (1.21) et (1.16). □

1.5.1. *Composition.* Soit Z un troisième e.v. et soit $g : V \to Z$. Si $V \supset f(U)$ alors

$$g \circ f : U \to Z.$$

On se donne une norme N_Z sur Z.

Proposition 1.23. *Si f est continue en $a \in U$ (p.r.à N_X, N_Y), et g est continue en $b = f(a) \in V$ (p.r.à N_Y et N_Z), alors $g \circ f$ est continue en $a \in U$ (p.r.à N_X et N_Z).*

Démonstration. Soit α une suite à valeurs dans U qui converge vers a (p.r.à N_X).

Comme f est continue en a (p.r.à N_X, N_Y), on sait (par Proposition (1.21) que la suite image, $\beta = f \circ \alpha$, converge vers $b = f(a)$ (p.r.à N_Y). Comme g est continue en b (p.r.à N_Y, N_Z), on sait (par Proposition (1.21) que la suite image, $g \circ \beta$, converge vers $g(b)$) (p.r.à N_Z).

Donc, la suite $(g \circ f) \circ \alpha$ converge vers $(g \circ f)(a)$ ce qui implique (par Proposition (1.21) que $g \circ f$ est continue $a \in U$ (p.r.à N_X et N_Z). □

Corollaire 1.24. *Si f est continue en $a \in U$ (p.r.à N_X et N_Y), alors pour chaque $u \in X$,*

$$t \mapsto f(a + tu)$$

est continue en $0 \in \mathbb{R}$ (p.r.à $|\ \ |, N_X$).

Démonstration. Soit

$$V = \{t \in \mathbb{R}; a + tu \in U\}$$

et soit $g(t) = a + tu$: c'est une fonction continue en $0 \in V$ (p.r.à $|\ \ |$ et N_Y). Comme

$$f(a + tu) = f \circ g(t),$$

la Proposition (1.23) implique le résultat. □

Remarque. La réciproque de ce corollaire n'est pas vraie : voir exercice en TD.

1.5.2. *Equivalence.*

Proposition 1.25. *Soient N'_X et N'_Y deux autres normes équivalentes à N_X et à N_Y respectivement. Alors, f est continue en $a \in U$ p.r.à N_X, N_Y ssi elle est continue en $a \in U$ p.r.à N'_X, N'_Y.*

Démonstration. Ceci suit des Propositions (1.21) et (1.19).

□

Entre des e.v.'s de dimension finie on peut donc parler de continuité d'une fonction *sans référence à des normes quelconques* parce que toutes les normes sont équivalentes.

Proposition 1.26. *Soit $Y = \mathbb{R}^m$. Alors $f = (f_1, \ldots, f_m)$ est continue en $a \in U$ (p.r.à N_X, N_Y) ssi*

$$f_j : U \to \mathbb{R}$$

est continue en $a \in U$ (p.r.à $N_X, |\ \ |$) pour chaque $j = 1, \ldots, m$.

Démonstration. Ceci suit des Propositions (1.21) et (1.20). □

1.5.3. *Fonctions à valeurs dans $\mathbb{R}$.* Soit $Y = \mathbb{R}$ et $N_Y = |\ \ |$ la norme "valeur absolue".

Proposition 1.27. *Soient $f, g : U \to Y$ continues en $a \in U$ (p.r.à $N_X, |\ \ |$). Alors*
(i) fg est continue en $a \in U$ (p.r.à $N_X, |\ \ |$)
(ii) $\frac{f}{g}$ est continue en $a \in U$ (p.r.à $N_X, |\ \ |$), si $g(a) \neq 0$

Démonstration. Ceci suit de la Proposition (1.21) et les règles pour la convergence des suites numériques. □

Exemple. Soit $X = \mathbb{R}^3$. Alors la fonction

$$f(x_1, x_2, x_3) = x_1 x_2 x_3$$

est continue parce que

$$f = \pi_1 \pi_2 \pi_3$$

où π_j est la projection sur le j :ième coefficient. Aussi, la fonction

$$g(x_1, x_2, x_3) = \sin(x_1 x_2 x_3)$$

est continue parce qu'une composition de deux fonctions continues.

Corollaire 1.28. *Chaque fonction polynômiale en n variables est continue.*

Démonstration. Chaque fonction monomiale à est un produit de projections sur les coefficients. Chaque fonction polynomiale est une combinaison linéaire de fonction monomiale. □

1.5.4. *Continuité et topologie.*

Proposition 1.29. *Soit $f : X \to Y$. Alors sont équivalentes :*
(i) f est continue
(ii) $f^{-1}(V)$ est une partie ouverte dans X (p.r.à N_X) pour chaque partie ouverte V dans Y (p.r.à N_Y)
(iii) $f^{-1}(V)$ est une partie fermée dans X (p.r.à N_X) pour chaque partie fermeé V dans Y (p.r.à N_Y)

Démonstration. (i) implique (ii). Soit $a \in f^{-1}(V)$. Alors $b = f(a) \in V$ et il existe $\varepsilon > 0$ t.q.
$$B_<^{N_Y}(b, \varepsilon) \subset V.$$
Comme f est continue en a, il existe $\delta > 0$ t.q.
$$N_X(x - a) < \delta \implies N_Y(f(x) - f(a)) < \varepsilon,$$
i.e.
$$f(B_<^{N_X}(a, \delta)) \subset B_<^{N_Y}(b, \varepsilon) \subset V.$$
Donc
$$B_<^{N_X}(a, \delta) \subset f^{-1}(V).$$
(ii) implique (i). Soit $a \in X$ et $b = f(x) \in Y$. Pour tout $\varepsilon > 0$, $f^{-1}(B_<^{N_Y}(b, \varepsilon))$ est un ouvert dans X qui contient a. Il existe donc $\delta > 0$ t.q.
$$B_<^{N_X}(a, \delta) \subset f^{-1}(B_<^{N_Y}(b, \varepsilon)),$$
i.e.
$$N_X(x - a) < \delta \Longrightarrow N_Y(f(x) - f(a)) < \varepsilon,$$
i.e. f est continue en a.

(ii) est équivalente à (iii) p.q.
$$X \setminus f^{-1}(Y \setminus V) = f^{-1}(V).$$
□

Exemple. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue. Alors, l'ensemble
$$\{(x, y) \in \mathbb{R}^2 : 1 < f(x, y) < 2\}$$
est un ouvert, parce que $= f^{-1}(]1, 2[)$, et l'ensemble
$$\{(x, y) \in \mathbb{R}^2 : 1 \leq f(x, y) \leq 2\}$$

est un fermé, parce que $= f^{-1}([1,2])$.

Notez bien que pour l'ensemble $\{(x,y) \in \mathbb{R}^2 : 1 < f(x,y) \leq 2\}$ on ne peut rien dire en général.

CM4

1.6. **Fonctions linéaires.** Soit $L : X \to Y$ une fonction linéaire. On définе

$$|||L||| = \sup_{x \neq 0} \frac{N_Y(L(x))}{N_X(x)}.$$

Nota Bene. $|||L|||$ dépend des deux normes N_X et N_Y. Une notation sans ambiguité serait

$$|||L|||_{N_X,N_Y}.$$

$|||L|||$ est un objet important et le lemme suivant est souvent utile.

Lemme.

$$|||L||| = \sup_{N_X(x)=1} \frac{N_Y(L(x))}{N_X(x)}.$$

Démonstration. On a, pour chaque $x \neq 0$,

$$\frac{N_Y(L(x))}{N_X(x)} = N_Y(L(y)), \quad y = \frac{x}{N_X(x)}.$$

Comme $N_X(y) = 1$ on a donc

$$\{\frac{N_Y(L(x))}{N_X(x)} : x \neq 0\} = \{N_Y(L(x)) : N_X(x) = 1\}.$$

Comme les deux ensembles sont égaux, leur supremum sont aussi égaux. □

Exemple. Soit

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \in \mathcal{M}_2$$

et considèrons l'application $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ défine par

$$\begin{pmatrix} x \\ v \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \end{pmatrix}.$$

On munit $\mathbb{R}^2$ avec la norme l^2. Alors

$$|||L_A||| = \sup_{x^2+y^2=1} \|(x+2y,y)\|_2 = \sup_{x^2+y^2=1} \sqrt{(x+2y)^2+y^2}$$

qui est calculable – c'est égal à $\sqrt{3+\sqrt{8}}$.

1.6.1. *Fonctions linéaires continues.*

Proposition 1.30. *Pour une fonction linéaire $L : X \to Y$ sont équivalentes :*
(i) L est continue (p.r.à N_X, N_Y)
(ii) L est continue en 0 (p.r.à N_X, N_Y)
(iii) $|||L||| < \infty$

Démonstration. (i) est équivalente à (ii) p.q.

$$N_Y(L(x) - L(a)) = N_Y(L(x-a) - L(0)).$$

(iii) implique (ii) p.q.

$$N_Y(L(x) - L(0)) = N_Y(L(x)) \leq |||L||| N_X(x) = |||L||| N_X(x - 0).$$

(ii) implique (iii). Supposons $|||L||| = \infty$. Alors (par le Lemme précédent) il existe une suite $(x_n)_{n\in\mathbb{N}}$ dans X t.q. $N_X(x_n) = 1$ pour tout n et

$$s_n = N_Y(L(x_n)) \to \infty, \quad n \to \infty.$$

Soit $y_n = \frac{1}{\sqrt{s_n}} x_n$. Alors la suite $(y_n)_{n\in\mathbb{N}}$ converge vers 0 (p.r.à N_X) tandis que la suite $(L(y_n))_{n\in\mathbb{N}}$ ne converge pas vers 0 (p.r.à N_X). Par Proposition (1.21) il s'ensuit que L n'est pas continue en 0.

□

1.6.2. *Fonctions linéaires en dimension finie.*

Exemple. Soit $X = \mathcal{C}^0([0,1])$ et soit

$$L : f \mapsto f(1).$$

- c'est une application linéaire $X \to \mathbb{R}$. L n'est pas continue p.r.à la norme $\| \ \|_{L^1}$

Notez que dans cet exemple, X est de dimension ∞.

Proposition 1.31. *Si X est de dimension finie alors chaque application linéaire $L : X \to Y$ est continue (p.r.à n'importe quelles normes).*

Démonstration. Soit $u_1, \ldots, u_n$ une base pour X et soit $M : \mathbb{R}^n \to X$ la bijection linéaire

$$x = (x_1, \ldots, x_n) \mapsto u = x_1 u_1 + \cdots + x_n u_n.$$

Par Proposition (1.2) $N_X \circ M$ est une norme sur $\mathbb{R}^n$, et il existe donc (p.q. toutes les normes sur $\mathbb{R}^n$ sont équivalentes) une constante $C > 0$ t.q.

$$\|x\|_1 \leq C N_X \circ M(x), \quad x \in \mathbb{R}^n$$

Soit $u = M(x)$. Alors

$$N_Y(L(u)) = N_Y(x_1 L(u_1) + \cdots + x_n L(u_n)) \leq \|x\|_1 \max_j N_Y(L(u_j))$$

$$\leq C \max_j N_Y(L(u_j)) N_X \circ M(x) = C \max_j N_Y(L(u_j)) N_X(u).$$

Ceci montre que $|||L||| < \infty$. □

1.6.3. *La norme subordonnée.* On note

$$\mathcal{L}_c(X,Y) = \{L : X \to Y : \text{ linéaire et continue}\}$$

– c'est un espace vectoriel (qui à priori dépend de N_X, N_Y).

Proposition 1.32.

(i) $|||\ \ |||$ *est une norme sur* $\mathcal{L}_c(X,Y)$ *– c'est la norme subordonnée (à* N_X, N_Y*).*

(ii) Si $X = Y$ *alors*

$$|||K \circ L||| \leq |||K||| \cdot |||L||| \quad \forall K, L \in \mathcal{L}_c(X,Y)$$

– c'est une norme sous-multiplicative.

Démonstration. (i) On vérifie facilement les trois propriétés d'une norme.

(ii)

$$|||K \circ L||| = \sup_{x \neq 0} \frac{N_X(K(L(x)))}{N_X(x)} = \sup_{L(x) \neq 0} \frac{N_X(K(L(x)))}{N_X(x)}$$

$$= \sup_{L(x) \neq 0} \frac{N_X(K(L(x)))}{N_X(L(x))} \frac{N_X(L(x))}{N_X(x)} =$$

$$\leq \sup_{L(x) \neq 0} \frac{N_X(K(L(x)))}{N_X(L(x))} \cdot \sup_{L(x) \neq 0} \frac{N_X(L(x))}{N_X(x)} \leq |||K||| \cdot |||L|||.$$

□

1.7. **Complétude.** Soit (X, N) un e.v.n., et soit α une suite dans X. La série

$$\sum_{k \geq 0} \alpha(k)$$

s'appelle la *série de terme général* $(\alpha(k))_{k \in \mathbb{N}}$.

Définition. La série $\sum_{k \in \mathbb{N}} \alpha(k)$ *converge vers* $a \in X$ (p.r.à N) ssi la suite des sommes partielles

$$(\sum_{k=0}^{n} \alpha(k))_{n \in \mathbb{N}}$$

converge vers a (p.r.à N), i.e.

$$N(a - \sum_{k=0}^{n} \alpha(k)) \to 0 \quad n \to \infty.$$

La série est *convergente* ssi elle converge vers un vecteur dans X.

La série $\sum_{k\in\mathbb{N}} \alpha(k)$ est *absolument convergente* (p.r.à N) ssi la série

$$\sum_{k\in\mathbb{N}} N(\alpha(k)) < \infty.$$

L'exemple classique de

$$\sum_{k\geq 0} (-1)^k \frac{1}{k+1}$$

montre qu'une série peut être convergente sans être absolument convergente.

L'énoncé réciproque est vrai dans certain e.v.n. – par exemple dans $(\mathbb{R}, |\ |)$ – mais pas dans tous.

Définition. Un e.v.n. (X, N) est dit *complet* ssi chaque série absolument convergente (p.r.à N) est convergente (p.r.à N). Un *espace de Banach* est un e.v.n. complet.

1.7.1. *Dimension finie.*

Proposition 1.33. *Soit (X, N) un espace vectoriel normé. Si la dimension de X est finie, alors (X, N) est complet.*

Démonstration. On accepte ce résultat pour $(\mathbb{R}, |\ |)$.

C'est aussi vrai pour $(\mathbb{R}^n, \|\ \|_\infty)$. En effet, puisque chaque série $\sum_{k\in\mathbb{N}} \alpha(k)$ dans $(\mathbb{R}^n, \|\ \|_\infty)$ a n composantes $\sum_{k\in\mathbb{N}} \alpha_j(k)$, $j = 1, \ldots, n$ qui sont des séries dans $(\mathbb{R}, |\ |)$, on vérifie facilement que

$$\begin{aligned}
&\sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est absolument convergente} \\
\iff \quad &\sum_{k\in\mathbb{N}} \alpha_j(k) \quad \text{est absolument convergente pour chaque } j \\
\implies \quad &\sum_{k\in\mathbb{N}} \alpha_j(k) \quad \text{est convergente (vers } a_j\text{) pour chaque } j \\
\iff \quad &\sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est convergente (vers } a = (a_1, \ldots, a_n)\ ).
\end{aligned}$$

C'est aussi vrai pour $(\mathbb{R}^n, N)$ pour n'importe quelle norme N. En effet, puisque N est équivalente à $\|\ \|_\infty$, on vérifie facilement que

$$\begin{aligned}
&\sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est absolument convergente p.r. à } N \\
\iff \quad &\sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est absolument convergente p.r. à } \|\ \|_\infty
\end{aligned}$$

$$\Longrightarrow \quad \sum_{k\in\mathbb{N}} \alpha(k) \ \text{ est convergente (vers } a \text{) p.r. à } \| \ \|_\infty$$

$$\Longleftrightarrow \quad \sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est convergente (vers } a \text{) p.r. à } N.$$

C'est aussi vrai pour (X, N), $\dim X = n$, pour n'importe quelle norme N. En effet, soit $L : \mathbb{R}^n \to X$ une bijection linéaire et soit $\beta(k) = L^{-1}(\alpha(k))$. Comme $N \circ L$ est une norme sur $\mathbb{R}^n$, on vérifie facilement que

$$\sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est absolument convergente p.r. à } N$$

$$\Longleftrightarrow \quad \sum_{k\in\mathbb{N}} \beta(k) \quad \text{est absolument convergente p.r. à } N \circ L$$

$$\Longrightarrow \quad \sum_{k\in\mathbb{N}} \beta(k) \ \text{ est convergente (vers } b \text{) p.r. à } N \circ L$$

$$\Longleftrightarrow \quad \sum_{k\in\mathbb{N}} \alpha(k) \quad \text{est convergente (vers } a = L(b) \text{) p.r. à } N.$$

$\square$

1.7.2. *La norme de convergence uniforme.* Soit I un intervalle fermé et borné.

Proposition 1.34. *L'espace vectoriel normé* $(\mathcal{C}^0(I), \| \ \|_{\mathcal{C}^0})$ *est complet.*

Une série $\sum_{k\geq 0} f_k$, de fonctions continues $f_k : I \to \mathbb{R}$, qui converge absolument p.r. à la norme $\| \ \|_{\mathcal{C}^0}$ est dite *normalement convergente*

Démonstration. Soit $f_k : I \to \mathbb{R}$ des fonctions continues. Si

$$\sum_k \|f_k\|_{\mathcal{C}^0} < \infty,$$

alors

$$\sum_k |f_k(t)| < \infty, \quad \forall t \in I.$$

On définit alors

$$f(t) = \sum_k f_k(t) \quad \forall t \in I.$$

Il s'agit de montrer que f est une fonction continue et que la série $\sum_k f_k$ converge vers f p.r.à la norme $\| \ \|_{\mathcal{C}^0}$.

Soit $\varepsilon > 0$. Alors il existe $n = n_\varepsilon$ t.q. $\sum_{k\geq n} \|f_k\|_{\mathcal{C}^0} \leq \varepsilon$ pour tout $n \geq n_\varepsilon$. Maintenant

$$\left|f(t)-\sum_{k=0}^{n} f_k(t)\right| \le \left|f(t)-\sum_{k=0}^{n+m} f_k(t)\right| + \left|\sum_{k=0}^{n+m} f_k(t)-\sum_{k=0}^{n} f_k(t)\right|$$
$$\le \left|f(t)-\sum_{k=0}^{n+m} f_k(t)\right| + \sum_{k\ge n} \|f_k\|_{\mathcal{C}^0} \le \left|f(t)-\sum_{k=0}^{n+m} f_k(t)\right| + \varepsilon$$

pour tout $m \in \mathbb{N}$, tout $t \in I$ et tout $n \ge n_\varepsilon$. Donc, pour tout $t \in I$ et tout $n \ge n_\varepsilon$,

$$\left|f(t)-\sum_{k=0}^{n} f_k(t)\right| = \lim_{m\to\infty}\left|f(t)-\sum_{k=0}^{n} f_k(t)\right|$$
$$\le \lim_{m\to\infty}\left|f(t)-\sum_{k=0}^{n+m} f_k(t)\right| + \varepsilon = \varepsilon.$$

Ceci implique que

$$\sup_{t\in I}\left|f(t)-\sum_{k=0}^{n} f_k(t)\right| \le \varepsilon$$

pour tout $n \ge n_\varepsilon$, i.e. les fonctions $\sum_{k=0}^{n} f_k$ convergent uniformement sur I vers f. f est donc continue et

$$\left\|f-\sum_{k=0}^{n} f_k\right\|_{\mathcal{C}^0} = \sup_{t\in I}\left|f(t)-\sum_{k=0}^{n} f_k(t)\right| \to 0, \quad n\to\infty.$$

□

1.8. **Compacité.**

Définition. Une partie A de X est bornée (p.r.à N) ssi elle est contenue dans une boule $B^N_{\le}(0,r)$.

Notez bien que si N' est une autre norme équivalente à N, alors A est bornée (p.r.à N) ssi elle est bornée (p.r.à N'). Ceci est une conséquence du lemme de Proposition (1.10).

On acceptera le résultat suivant qui sera démontré dans le cours de topologie.

Théorème. *Soit X de dimension finie et soit U une partie bornée et fermée (et non-vide). Alors chaque fonction continue $f : U \to \mathbb{R}$ atteint son maximum et son minimum sur U.*

Ce résultat est faux en dimension ∞. Voici un exemple.

Exemple. On considère l'espace vectoriel

$$l^1(\mathbb{N}) = \{\alpha : \mathbb{N} \to \mathbb{R} : \underbrace{\sum_{n=0}^{\infty} |\alpha(n)|}_{\|\alpha\|_1} < \infty\}$$

muni de la norme $\| \ \|_1$ – c'est bien une norme.

Soit

$$U = \{\alpha \in l^1(\mathbb{N}) : \|\alpha\|_1 = 1\}$$

– c'est une partie trivialement bornée. Elle est aussi fermée p.q. $= B_{\leq}(0,1) \setminus B_{<}(0,1)$, où les boules ici sont définies avec la norme $\| \ \|_1$.

Soit $f : U \to \mathbb{R}$ la fonction

$$f(\alpha) = \sum_{n \in \mathbb{N}} \frac{1}{n+1} |\alpha(n)|$$

– c'est une fonction continue p.q.

$$|f(\beta) - f(\alpha)| \leq \sum_{n \in \mathbb{N}} \frac{1}{n+1} \big| |\alpha(n)| - |\beta(n)| \big| \leq$$

$$\leq \sum_{n \in \mathbb{N}} \frac{1}{n+1} |\alpha(n) - \beta(n)| \leq \|\alpha - \beta\|_1 .$$

De plus, si α_k est la suite qui est $= 1$ pour $n = k$ et $= 0$ pour $n \neq k$, alors $\alpha_k \in U$ et $f(\alpha_k) = \frac{1}{k+1}$. D'où

$$\inf_{\alpha \in U} f(\alpha) = 0.$$

Mais cette infimum n'est pas atteint p.q. $f(\alpha) > 0$ pour tout $\alpha \in U$.

2. Calcul Différentiel

CM5

Soit X un espace vectoriel, et soit $\| \ \|$ une norme sur X. Les notions topologiques comme "ouvert" ou "fermé" ou "continuité" vont se référer à cette norme-là, sans que cela soit explicitement dit.

On va considérer *un ouvert* U dans X, i.e. un ouvert p.r.à la norme $\| \ \|$ – cela dit pour la première et la dernière fois.

2.1. **Différentiabilité – fonctions à valeurs réelles.** Soit $f : U \to \mathbb{R}$ une fonction.

2.1.1. *Dérivées directionnelles.*

Définition. La fonction $f : U \to \mathbb{R}$ est *dérivable suivant le vecteur $v \in X$ en $a \in U$* ssi la fonction

$$t \mapsto f(a + tv) \in \mathbb{R}$$

est dérivable en $t = 0$. On note alors cette dérivée

$$\frac{\partial f}{\partial v}(a)$$

– c'est la *dérivée de f suivant le vecteur v en a.*

Exemple. Soit $f_1 : \mathbb{R}^2 \to \mathbb{R}$ la fonction $(x_1, x_2) = x_1^3 x_2^2$. Si $v = (v_1, v_2)$ et $a = (a_1, a_2)$, alors

$$f_1(a + tv) = (a_1 + tv_1)^3 (a_2 + tv_2)^2.$$

Cette fonction est dérivable en $t = 0$ et

$$\frac{\partial f_1}{\partial v}(a) = \frac{d}{dt} f(a + tv)_{/t=0} = 3a_1^2 a_2^2 v_1 + 2a_1^3 a_2 v_2.$$

Exemple. La fonction

$$f_2(x) = \begin{cases} \frac{x_1^6}{x_1^8 + (x_2 - x_1^2)^2} & x \neq (0,0) \\ 0 & x = (0,0) \end{cases}$$

est dérivable suivant chaque vecteur $v = (v_1, v_2)$ en $a = (0, 0)$. En effet, si $v_1 = 0$ $t \mapsto f_2(a + tv)$ est la fonction nulle, donc dérivable partout, et si $v_1 \neq 0$, alors la fonction

$$f_2(a + tv) = \begin{cases} \frac{t^4 v_1^6}{t^6 v_1^8 + (v_2 - tv_1^2)^2} & t \neq 0 \\ 0 & t = 0. \end{cases}$$

est dérivable en $t = 0$ avec dérivée $= 0$.

Exemple. La fonction

$$f_3(x) = \begin{cases} \frac{x_1^6}{x_1^8 + x_2^2} & x \neq (0,0) \\ 0 & x = (0,0) \end{cases}$$

n'est pas dérivable suivant le vecteur $v = (1, 0)$ en $a = (0, 0)$. En effet, la fonction

$$f_3(a + tv) = \begin{cases} \frac{t^4}{t^6} & t \neq 0 \\ 0 & t = 0. \end{cases}$$

n'est pas dérivable en $t = 0$.

2.1.2. *Dérivées partielles.* Soit $X = \mathbb{R}^n$. Si $a = (a_1, \ldots, a_n) \in U$ et si l'on fixe tous les variables $(x_1, \ldots, x_n)$ sauf le premier

$$x_2 = a_2, \ldots, x_n = a_n$$

on obtient une fonction d'une variable

$$x_1 \mapsto f(x_1, a_2, \ldots, a_n).$$

Si cette fonction est dérivable en $x_1 = a_1$ on dit que f est *dérivable p.r.à x_1 en a_1* et l'on note

$$\frac{\partial f}{\partial x_1}(a) = \frac{d}{dx_1} f(x_1, a_2, \ldots, a_n)/_{x_1 = a_1}.$$

– on appelle cette dérivée la *dérivée partielle de f p.r.à x_1 en a.*

Idem pour les autres variables.

On va dire que f est *partiellement dérivable en a* ssi f a des dérivées partielles p.r.à toutes les variables $x_1, \ldots, x_n$ en a.

Exemple. La fonction $f_4(x_1, x_2) = x_1^3 x_2^2$ est partiellement dérivable en chaque $a = (a_1, a_2) \in \mathbb{R}^2$.

En effet, la fonction

$$x_1 \mapsto f_4(x_1, a_2) = x_1^3 a_2^2$$

est dérivable avec la dérivée $3x_1^2 a_2^2$. On a donc $\frac{\partial 4}{\partial x_1}(a) = 3a_1^2 a_2^2$.

De plus, la fonction

$$x_2 \mapsto f_4(a_1, x_2) = a_1^3 x_2^2$$

est dérivable avec la dérivée $2a_1^3 a_2$. On a donc $\frac{\partial f_4}{\partial x_2}(a) = 2a_1^3 a_2$.

On se convainc facilement que les dérives partielles sont des dérivés directionelles et que

$$\frac{\partial f}{\partial x_j}(a) = \frac{\partial f}{\partial e_j}(a), \quad j = 1, \ldots, n,$$

où $\{e_1, \ldots, e_n\}$ est la base canonique de $\mathbb{R}^n$.

En particulier, la fonction f_2 (ci-dessus) est partiellement dérivable en $a = (0, 0)$ et $\frac{\partial f_2}{\partial x_1}(a) = \frac{\partial f_2}{\partial x_1}(a) = 0$.

La fonction f_3 (ci-dessus) n'est pas partiellement dérivable en $a = (0, 0)$. La dérivée partielle p.r.à x_2 en a existe – $\frac{\partial f_3}{\partial x_2}(a) = 0$ – mais la dérivée partielle p.r.à x_1 en a n'existe pas.

Notez que

$$f(x_1, x_2) = f_2(x_1, x_2 + x_1^2)$$

est une composition de fonctions qui sont partiellement dérivables [1] en chaque point, mais f elle-même n'a pas cette propriété. C'est pour cette

1. même dérivables suivant tout vecteur

raison qu'on a cherché une autre généralisation de dérivabilité avec de meilleurs propriétés “structurales”.

Gateaux-différentiabilité. Une fonction $f : U \to \mathbb{R}$ est dite *Gateaux-différentiabie en $a \in U$* ssi f est dérivable suivant chaque vecteur $v \in X$ et la fonction

$$X \ni v \mapsto \frac{\partial f}{\partial v}(0)$$

est linéaire et continue.

La fonction f_2 est Gateaux-différentiabie en $(0,0)$ ce qui montre que Gateaux-différentiabilité n'implique pas la continuité. L'exemple

$$f(x_1, x_2) = f_2(x_1, x_2 + x_1^2)$$

montre aussi que cette propriété n'est pas préservé sous compositions, p.q. les fonctions

$$(x_1, x_2) \mapsto x_1 \quad \text{et} \quad (x_1, x_2) \mapsto x_2 + x_1^2$$

sont toutes les deux Gateaux-différentiables partout.

2.1.3. *Différentiabilité.*

Définition. La fonction $f : U \to \mathbb{R}$ est *différentiable en $a \in U$* ssi il existe une application linéaire et continue $L : X \to \mathbb{R}$ t.q.

$$(*) \quad \frac{|f(x+a) - f(a) - L(x)|}{\|x\|} \to 0, \quad \|x\| \to 0.$$

f est *différentiable* ssi elle est différentiable en chaque $a \in U$.

Notez que la continuité de L est la continuité p.r.à la norme $\| \ \|$. Rappelons que, si X est de dimension finie alors chaque application linéaire $L : X \to \mathbb{R}$ est continue mais ceci n'est pas le cas si X est de dimension ∞.

Notez aussi que la différentiabilité fait intervenir la norme $\| \ \|$. C'est donc d'une *différentiabilité p.r. à* $\| \ \|$ dont on parle ici.

Proposition 2.1. *Si f est différentiable en $a \in U$, alors l'application linéaire et continue $L : X \to \mathbb{R}$ en $(*)$ est unique. On l'appelle la différentielle de f en a, notée Df_a ou $Df(a)$.*

Démonstration. Soient deux applications linéaires et continues L_1 et $L_2 : X \to \mathbb{R}$ dont chacune vérifie la condition $(*)$, i.e.

$$\frac{|f(x+a) - f(a) - L_j(x)|}{\|x\|} \to 0, \quad \|x\| \to 0$$

pour $j = 1, 2$. Comme

$$|(L_1 - L_2)(x)| \leq |f(x+a) - f(a) - L_1(x)| + |f(x+a) - f(a) - L_2(x)|$$

on a
$$\frac{|(L_1 - L_2)(x)|}{\|x\|} \to 0, \quad \|x\| \to 0.$$
Alors
$$\frac{|(L_1 - L_2)(x)|}{\|x\|} = \frac{|(L_1 - L_2)(\lambda x)|}{\|\lambda x\|} \to 0, \quad \lambda \to 0,$$
ce qui implique que $(L_1 - L_2)(x) = 0$, i.e. $L_1(x) = L_2(x)$. □

Si f est une fonction constante alors f est différentiable et Df_a est l'application linéaire nulle :
$$Df_a : x \mapsto 0, \quad x \in X.$$

Si $f = L$ est une fonction linéaire et continue, alors f est différentiable et $Df_a = L$ pour chaque a.

Si $X = \mathbb{R}$ et $\| \ \|$ est la valeur absolue, alors f est différentiable $a \in U \subset \mathbb{R}$ ssi f est dérivable en a et
$$Df_a : x \mapsto f'(a)x, \quad x \in \mathbb{R}.$$

Exemple. Soit $X = \mathbb{R}^2$ et $\| \ \|$ la norme l^2. La fonction
$$f(x,y) = \begin{cases} \frac{x^3y}{x^2+y^2} & x \neq (0,0) \\ 0 & x = (0,0) \end{cases}$$
est différentiable en $(0,0)$ et $Df_{(0,0)}$ est l'application linéaire nulle sur $\mathbb{R}^2$. En effet
$$|f((x,y) + (0,0)) - f(0,0)| = \left| \frac{x^3y}{x^2+y^2} \right| \leq$$
$$\leq \frac{|x|^3 |y|}{x^2+y^2} \leq \frac{(\sqrt{x^2+y^2})^4}{x^2+y^2} = \|(x,y)\|^2 .$$

Exemple. Soit $X = \mathcal{M}_2$. On munit X de la norme
$$\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\| = |a| + |b| + |c| + |d|$$
qui est, rappelons-le, sous-multiplicative :
$$\|AB\| \leq \|A\| \|B\| .$$

La fonction $f(A) = Tr(A^2)$ est différentiable en chaque $A \in \mathcal{M}_2$ et
$$Df_A : B \mapsto 2Tr(AB).$$

En effet
$$f(A+B) - f(A) = Tr(AB + BA + B^2) = 2Tr(AB) + Tr(B^2)$$
et
$$L : B \to 2Tr(AB)$$

est une applicaton linéaire et, donc, continue (p.q. $\dim X = 4 < \infty$). De plus

$$\frac{|f(A+B)-f(A)-L(B)|}{\|B\|} = \frac{|Tr(B^2)|}{\|B\|} \leq \frac{\|B^2\|}{\|B\|} \leq \|B\| \to 0, \quad \|B\| \to 0.$$

2.1.4. *Continuité et dérivées directionnelles.*

Proposition 2.2. *Soit f différentiable en $a \in U$. Alors*
(i) f est continue en a
(ii) f est dérivable suivant tout vecteur $v \in X$ en a et

$$\frac{\partial f}{\partial v}(a) = Df_a(v).$$

Démonstration. (i) On a

$$|f(x+a)-f(a)| \leq |f(x+a)-f(a)-Df_a(x)| + |Df_a(x)| \to 0, \quad \|x\| \to 0.$$

(ii)

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \frac{f(a+tv)-f(a)}{t}$$

$$= \lim_{t\to 0} \Big(\frac{f(a+tv)-f(a)-Df_a(tv)}{t} + \frac{Df_a(tv)}{t}\Big)$$

$$= \Big(\lim_{t\to 0} \frac{f(a+tv)-f(a)-Df_a(tv)}{\|tv\|}\Big)\Big(\frac{|t|}{t}\Big)\|v\| + Df_a(v) = Df_a(v)$$

$\square$

Exemple. La fonction

$$f(x) = \begin{cases} \frac{x_1^6}{x_1^8+(x_2-x_1^2)^2} & x \neq (0,0) \\ 0 & x = (0,0) \end{cases}$$

est dérivable suivant chaque vecteur $v = (v_1, v_2)$ en $a = (0,0)$, mais elle n'est pas différentiable en $(0,0)$. En effet, elle n'est même pas continue en $(0,0)$.

Corollaire 2.3. *Soit $X = \mathbb{R}^n$. Si $f : U \to \mathbb{R}$ est différentiable en $a \in U$, alors*

$$Df_a(v) = \frac{\partial f}{\partial x_1}(a)v_1 + \cdots + \frac{\partial f}{\partial x_n}(a)v_n$$

pour tout $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$.

On peut donc écrire $Df_a(v)$ comme le produit scalaire (euclidien)

$$Df_a(v) = \langle \underbrace{(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a))}_{\nabla f(a)}, (x_1, \dots, x_n) \rangle$$

ou comme le produit matriciel

$$Df_a(v) = \underbrace{\begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \dots & \frac{\partial f}{\partial x_n}(a) \end{pmatrix}}_{\mathrm{Jac} f(a)} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Le vecteur $\nabla f(a)$ s'appelle le *gradient* de f en a, et la matrice $\mathrm{Jac} f(a)$ s'appelle la *jacobienne* de f en a.

Nota Bene. Df_a n'est ni un vecteur ni une matrice, mais une *fonction linéaire.*

Démonstration. Soit $\{e_1, \dots, e_n\}$ la base canonique de $\mathbb{R}^n$. Alors

$$v = (v_1, \dots, v_n) = v_1 e_1 + \dots + v_n e_n$$

et, comme Df_a est linéaire,

$$Df_a(v) = v_1 Df_a(e_1) + \dots v_n Df_a(e_n).$$

Par définition

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

pour chaque j. □

2.1.5. *Linéarité et produit.* Une notation habituelle et souvent pratique est *"le petit o"*. On note $g(x) = o(\|x\|^n)$ pour une fonction g définie au voisinage de $x = 0$ t.q.

$$\frac{g(x)}{\|x\|^n} \to 0, \quad \|x\| \to 0.$$

Avec cette notation on peut donc formuler $(*)$ comme

$$|f(x+a) - f(a) - L(x)| = o(\|x\|).$$

Proposition 2.4. *Si $f, g : U \to \mathbb{R}$ sont différentiables en $a \in U$, alors*

(i) pour tous $\lambda, \mu \in \mathbb{R}$, la fonction $\lambda f + \mu g$ est différentiable en $a \in U$ et

$$D(\lambda f + \mu g)_a = \lambda Df_a + \mu Dg_a$$

(ii) la fonction fg est différentiable en $a \in U$ et

$$D(fg)_a = g(a) Df_a + f(a) Dg_a$$

Démonstration. (i) est une vérification facile.

(ii) On a

$$(fg)(x+a) - (fg)(a) - g(a)Df_a(x) - f(a)Dg_a(x) =$$
$$= g(x+a)\big(f(x+a) - f(a) - Df_a(x)\big) + f(a)\big(g(x+a) - g(a) - Dg_a(x)\big) +$$
$$+\big(g(x+a) - g(a)\big)Df_a(x)$$

ce qui est un $o(\|x\|)$. En effet

$$\frac{g(x+a)\big(f(x+a) - f(a) - Df_a(x)\big)}{\|x\|} \to g(a)\cdot 0 = 0, \quad \|x\| \to 0$$

p.q. f est différentiable en a et g est continue en a,

$$\frac{f(a)\big(g(x+a) - g(a) - Dg_a(x)\big)}{\|x\|} \to f(a)\cdot 0 = 0, \quad \|x\| \to 0$$

p.q. g est différentiable en a,

$$\frac{\big|\big(g(x+a) - g(a)\big)Df_a(x)\big|}{\|x\|} \le |g(x+a) - g(a)|\, |||Df_a||| \to 0, \quad \|x\| \to 0$$

p.q g est continue en a et $|||Df_a||| < \infty$.

□

Corollaire 2.5. *Chaque fonction polynômiale en n variables est différentiable.*

Démonstration. Chaque monôme est un produit de projections sur les coordonnées qui sont des applications linéaires, donc différentiables. Chaque polynôme est une combinaison linéaire de monômes.

□

2.1.6. *Composition.*

Proposition 2.6. *Si $f : U \to \mathbb{R}$ sont différentiables en $a \in U$, et si $\varphi : \mathbb{R} \to \mathbb{R}$ est dérivable en $b = f(a) \in \mathbb{R}$, alors $\varphi \circ f$ est différentiable en $a \in U$ et*

$$D(\varphi\circ f)_a = \varphi'(b)Df_a.$$

Démonstration. On a, pour $x \neq 0$,

$$\frac{f(x+a) - f(a)}{\|x\|} = \frac{f(x+a) - f(a) - Df_a(x)}{\|x\|} + \frac{Df_a(x)}{\|x\|}$$

et, comme f est différentiable en a, il existe $r > 0$ t.q.

$$\left|\frac{f(x+a) - f(a)}{\|x\|}\right| \le 1 + |||Df_a|||$$

si $\|x\| < r$.

Soit $y = f(x+a) \neq f(a) = b$. Alors

$$\|x\| \to 0 \quad \Longrightarrow \quad |y-b| \to 0 \quad \Longrightarrow \quad \frac{\varphi(y)-\varphi(b)-\varphi'(b)(y-b)}{y-b} \to 0$$

p.q. f est continue en a et φ est differentiable en b.

D'où

$$\frac{\varphi(f(x+a))-\varphi(f(a))-\varphi'(f(a))(f(x+a)-f(a))}{\|x\|}$$

$$= \frac{\varphi(y)-\varphi(b)-\varphi'(b)(y-b)}{y-b}\frac{f(x+a)-f(a)}{\|x\|} \to 0 \quad \|x\| \to 0,$$

i.e.

$$\varphi(f(x+a))-\varphi(f(a))-\varphi'(f(a))(f(x+a)-f(a)) = o(\|x\|).$$

D'autre part

$$\varphi'(b)\big(f(a+x)-f(a)-Df_a(x))\big) = o(\|x\|)$$

p.q. f est différentiable en a.

Comme

$$\varphi\circ f(x+a)-\varphi\circ f(a)-\varphi'(b)Df_a(x) =$$

$$= \varphi(f(x+a))-\varphi(f(a))-\varphi'(b)(f(a+x)-f(a))+ \\ + \varphi'(b)\big(f(a+x)-f(a)-Df_a(x))\big)$$

on voit que

$$\varphi\circ f(x+a)-\varphi\circ f(a)-\varphi'(b)Df_a(x) = o(\|x\|).$$

□

Corollaire 2.7. *Si $f, g : U \to \mathbb{R}$ sont différentiables en $a \in U$ et si $g(a) \neq 0$ alors la fonction $\frac{f}{g}$ est différentiable en $a \in U$ et*

$$D(\frac{f}{g})_a = \frac{1}{g(a)^2}(g(a)Df_a - f(a)Dg_a).$$

Démonstration. Pour $f = 1$ c'est une conséquence de Proposition (2.6) Ensuite $\frac{f}{g} = f\frac{1}{g}$ et l'on applique Proposition (2.4)

□

Exemple. La fonction $f : \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \sin(\log(\frac{x^2+1}{y^2+1})),$$

est différentiable.

2.1.7. *Equivalence.* La notion de différentiabilité et la différentielle sont définies avec une norme sur X, et dépendent donc, à priori, de cette norme.

Proposition 2.8. *Soit N une autre norme sur X qui est équivalente à $\| \|$. S'il existe une application linéaire $L : X \to \mathbb{R}$, continue (p.r.à $\| \|, | |$), t.q.*

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

alors L est continue (p.r.à $N, | |$) et

$$|f(x+a) - f(a) - L(x)| = o(N(x)).$$

Démonstration. Il existe une constante $C \geq 1$ t.q. $\frac{1}{C}\|x\| \leq N(x) \leq C\|x\|$, $x \in X$. Alors

$$\|x\| \to 0 \quad \Longleftrightarrow \quad N(x) \to 0$$

et

$$\frac{|f(x+a) - f(a) - L(x)|}{N(x)} \leq C \frac{|f(x+a) - f(a) - L(x)|}{\|x\|} \to 0, \quad N(x) \to 0.$$

Le fait que L est continue (p.r.à $N(x), | |$) est donné dans Proposition (1.25). □

Si X est de dimension finie on peut donc parler de la différentiabilité d'une fonction *sans référence à des normes quelconques* parce que toutes les normes sont équivalentes.

2.1.8. *Théorème des accroissements finis.* Soient $a, b \in X$. On note CM6

$$[a, b] = \{(1-t)a + tb : \ 0 \leq t \leq 1\}$$

– le *segment entre a et b.*

Définition. Une partie U de X est dite *convexe* ssi pour tous $a, b \in X$, le segment

$$[a, b] \subset X.$$

Notez que chaque boule $B = B_<^{\| \|}(c, r)$ est convexe. En effet, si $a, b \in B$, alors, pour $0 \leq t \leq 1$,

$$\|ta + (1-t)b - c\| = \|t(a-c) + (1-t)(b-c)\| \leq$$
$$\leq t\|a-c\| + (1-t)\|b-c\| < tr + (1-t)r = r.$$

Le même est vrai pour les boules élargies.

Proposition 2.9. *Soit $f : U \to \mathbb{R}$ une fonction différentiable.*

(i) Si $[a, b] \subset U$, alors il existe $c \in [a, b]$ t.q.

$$f(b) - f(a) = Df_c(b-a).$$

(ii) Si U est convexe, alors

$$|f(a) - f(b)| \leq \sup_{c \in U} |||Df_c||| \, \|b - a\| \quad \forall a, b \in U,$$

où $|||Df_c|||$ est la norme subordonnée de l'application Df_c.

Démonstration. (i) Soit $g : [0, 1] \to \mathbb{R}$ la fonction $g(t) = f((1 - t)a + tb) = f(a + t(b - a))$. On a

$$g'(t) = \frac{\partial f}{\partial (b - a)}(a + t(b - a)) = Df_{a+t(b-a)}(b - a), \quad \forall t \in [0, 1].$$

Par le théorème des accroissements finis habituel, il existe $0 < s < 1$ t.q. $g(1) - g(0) = g'(s)$, ce qui donne

$$f(b) - f(a) = Df_c(b - a), \quad c = a + s(b - a).$$

(ii) Comme U est convexe on a donc, par (i),

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

Par la définition de la norme subordonnée on a

$$|Df_c(b - a)| \leq |||Df_c||| \, \|b - a\| \, .$$

□

Corollaire 2.10. *Soit $f : U \to \mathbb{R}$ une fonction différentiable. Si U est convexe et*

$$Df_a = 0 \quad \forall a \in U,$$

alors f est une fonction constante.

Démonstration. Par la proposition on voit que $f(a) = f(b)$ pour tous $a, b \in U$. □

2.2. **Fonctions à valeurs réelles de classe $\mathcal{C}^1$.** Nous allons nous restreindre au cas d'une fonction $f : U \to \mathbb{R}$, où U est un ouvert de $X = \mathbb{R}^n$.

Définition. La fonction $f : U \to \mathbb{R}$ est *de classe $\mathcal{C}^1$* ssi f est partiellement dérivable en chaque point de U et les fonctions

$$\frac{\partial f}{\partial x_j} : U \to \mathbb{R}$$

sont continues, $j = 1, \ldots, n$.

Exemple. La fonction

$$f(x,y) = x(y+1) + \log(x^2+1) + \sin^2(y), \quad (x,y) \in \mathbb{R}^2,$$

est de classe $\mathcal{C}^1$. En effet,

$$\frac{\partial f}{\partial x}(x,y) = y + 1 + \frac{2x}{x^2+1} \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = x + 2\sin y \cos y$$

sont des fonctions continues.

La proposition suivante montre que la notion de classe $\mathcal{C}^1$ exclut des fonctions "perverses" comme

$$f(x) = \begin{cases} \frac{x_1^6}{x_1^8 + (x_2 - x_1^2)^2} & x \neq (0,0) \\ 0 & x = (0,0) \end{cases}$$

qui est partiellement dérivable en $(0,0)$ sans être différentiables en $(0,0)$.

Proposition 2.11. *Sont équivalentes pour une fonction $f : U \to \mathbb{R}$:*
(i) f est de classe $\mathcal{C}^1$
(ii) f est différentiable et la fonction

$$U \ni a \mapsto Df_a \in \mathcal{L}_c(X, \mathbb{R})$$

est continue p.r.à la norme subordonnée sur $\mathcal{L}_c(X, \mathbb{R})$, i.e. la norme

$$|||L||| = \sup_{x \neq 0} \frac{|L(x)|}{\|x\|}.$$

Démonstration. (ii) implique (i) par Proposition (2.2) qui dit que

$$\frac{\partial f}{\partial x_j}(x) = Df_x(e_j).$$

Alors

$$\left| \frac{\partial f}{\partial x_j}(y) - \frac{\partial f}{\partial x_j}(x) \right| \leq |||Df_y - Df_x||| \to 0, \quad y \to x$$

par la continuité de Df_x.

(i) implique (ii). On va donner la preuve pour $n = 2$. Considérons d'abord le point $(0,0)$.

Soit

$$I(x,y) = f(x,y) - f(0,0) - \frac{\partial f}{\partial x}(0,0)x - \frac{\partial f}{\partial y}(0,0)y =$$

$$\int_0^y \big(\frac{\partial f}{\partial y}(x,t) - \frac{\partial f}{\partial y}(0,0)\big)dt + \int_0^x \big(\frac{\partial f}{\partial x}(t,0) - \frac{\partial f}{\partial x}(0,0)\big)dt.$$

Soit $\varepsilon > 0$. Alors il existe $\delta = \delta_\varepsilon > 0$ t.q.

$$\|(x,y)\|_1 < \delta \quad \Longrightarrow \quad \begin{cases} \left|\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(0,0)\right| < \varepsilon \\ \left|\frac{\partial f}{\partial y}(x,y) - \frac{\partial f}{\partial y}(0,0)\right| < \varepsilon \end{cases}$$

– ici on utilise la continuité des dérivées partielles.

Alors

$$\|(x,y)\|_1 < \delta \quad \Longrightarrow \quad |I(x,y)| \le \varepsilon\,|y| + \varepsilon\,|x| = \varepsilon\,\|(x,y)\|_1\,.$$

Donc

$$I(x,y) = o(\|(x,y)\|_1),$$

i.e. f est différentiable en $(0,0)$. De plus

$$Df_{(0,0)}(\xi,\eta) = \frac{\partial f}{\partial x}(0,0)\xi - \frac{\partial f}{\partial y}(0,0)\eta \quad \forall(\xi,\eta) \in \mathbb{R}^2.$$

On démontre de la même façon que f est différentiable en chaque point $(x,y) \in U$ et que

$$Df_{(x,y)}(\xi,\eta) = \frac{\partial f}{\partial x}(x,y)\xi - \frac{\partial f}{\partial y}(x,y)\eta \quad \forall(\xi,\eta) \in \mathbb{R}^2.$$

Cette formule implique la continuité de $Df_{(x,y)}$. En effet

$$\left|Df_{(x',y')}(\xi,\eta) - Df_{(x,y)}(\xi,\eta)\right| \le$$

$$\le \Big(\left|\frac{\partial f}{\partial x}(x',y') - \frac{\partial f}{\partial x}(x,y)\right| + \left|\frac{\partial f}{\partial y}(x',y') - \frac{\partial f}{\partial y}(x,y)\right| \Big)\, \|(\xi,\eta)\|_1$$

Donc

$$\left|\!\left|\!\left|Df_{(x',y')} - Df_{(x,y)}\right|\!\right|\!\right| \le \Big(\left|\frac{\partial f}{\partial x}(x',y') - \frac{\partial f}{\partial x}(x,y)\right| + \left|\frac{\partial f}{\partial y}(x',y') - \frac{\partial f}{\partial y}(x,y)\right| \Big)$$

et ceci $\to 0$ quand $(x',y') \to (x,y)$ par la continuité des dérivées partielles. □

Exemple. La fonction

$$f(x,y) = x(y+1) + \log(x^2+1) + \sin^2(y), \quad (x,y) \in \mathbb{R}^2,$$

est différentiable p.q. de classe $\mathcal{C}^1$.

Proposition 2.12. *Si $f, g : U \to \mathbb{R}$ sont de classe $\mathcal{C}^1$, alors*

(i) pour tous $\lambda, \mu \in \mathbb{R}$, la fonction $\lambda f + \mu g$ est de classe $\mathcal{C}^1$.

(ii) la fonction fg est de classe $\mathcal{C}^1$

(iii) Si g ne s'annule pas, alors la fonction $\frac{f}{g}$ est de classe $\mathcal{C}^1$

(iv) $\varphi \circ g$ est de classe $\mathcal{C}^1$ pour chaque fonction $\varphi : V \to \mathbb{R}$ de classe $\mathcal{C}^1$, où V est un ouvert qui contient $f(U)$.

Démonstration. Ceci suit des règles de dérivation, qui sont valables aussi pour les dérivations partielles, et des propriétés de continuité. □

Corollaire 2.13. *Chaque fonction polynômiale en n variables est de classe $\mathcal{C}^1$.*

Démonstration. Chaque monôme est un produit de projections sur les coordonnées qui sont des applications linéaires, donc de classe $\mathcal{C}^1$. Chaque polynôme est une combinaison linéaire de monômes.

□

2.2.1. *Classe $\mathcal{C}^1$ générale.* Dans un espace vectoriel normé général X on peut aussi définir la notion de classe $\mathcal{C}^1$.

Définition. La fonction $f : U \to \mathbb{R}$ est *de classe* $\mathcal{C}^1$ ssi f est différentiable et la fonction

$$U \ni a \mapsto Df_a \in \mathcal{L}_c(X, \mathbb{R})$$

est continue p.r.à la norme subordonnée sur $\mathcal{L}_c(X, \mathbb{R})$, i.e. la norme

$$|||L||| = \sup_{x \neq 0} \frac{|L(x)|}{\|x\|}.$$

Exemple. Soit $X = \mathcal{M}_2$. On munit X de la norme

$$\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\| = |a| + |b| + |c| + |d| \,.$$

La fonction $f(A) = Tr(A^2)$ est différentiable et

$$Df_A : B \mapsto 2Tr(AB).$$

Alors

$$\big(Df_{A+C} - Df_A\big)(B) = 2Tr((A + C)B) - 2Tr(AB) = 2Tr(CB).$$

Donc

$$|||Df_{A+C} - Df_A||| = \sup_{b \neq 0} \frac{\big|\big(Df_{A+C} - Df_A\big)(B)\big|}{\|B\|} =$$

$$= \sup_{B \neq 0} 2 \frac{|Tr(CB)|}{\|B\|} \leq 2 \|C\| \,.$$

2.3. **Limites et intégration p.r.à un paramètre.** Soit I un intervalle dans $\mathbb{R}$ et soient CM7

$$-\infty \leq \alpha = \inf I < \sup I = \beta \leq \infty.$$

On a donc que l'intérieure de I

$$\mathrm{int}(I) =]\alpha, \beta[$$

mais I lui-même n'est pas forcément ouvert.

2.3.1. *Intégrabilité.*

Définition. Une fonction continue $f : I \to \mathbb{R}$ est dite *intégrable* ssi

$$\int_\alpha^\beta |f(t)|\, dt < \infty,$$

i.e. il existe une constante $C > 0$ t.q.

$$\int_\gamma^\delta |f(t)|\, dt < C$$

pour tout $\alpha < \gamma < \delta < \beta$.

Noter que si l'intervalle I est fermé et borné, alors chaque fonction continue sur I est intégrable. Mais pour d'autres intervalles ce n'est pas le cas.

Exemple. $I =]0,1]$. La fonction $f(t) = \frac{1}{\sqrt{t}}$ est intégrable, mais la fonction $g(t) = \frac{1}{t}$ ne l'est pas.

Noter que si $|f| \leq |g|$ et g est intégrable, alors f l'est aussi.

Proposition 2.14. *Si $f : I \to \mathbb{R}$ est intégrable, alors l'intégrale (de Riemann) généralisée*

$$\int_\alpha^\beta f(t)dt$$

existe.

Démonstration. Soient deux suites $\alpha_n \searrow \alpha$ et $\beta_n \nearrow \beta$. Alors

$$\int_{\beta_n}^\beta |f(t)|\, dt + \int_\alpha^{\alpha_n} |f(t)|\, dt \to 0, \quad n \to \infty.$$

Comme

$$\left| \int_{\alpha_{n+m}}^{\beta_{n+m}} f(t)dt - \int_{\alpha_n}^{\beta_n} f(t)dt \right| \leq \int_{\beta_n}^\beta |f(t)|\, dt + \int_\alpha^{\alpha_n} |f(t)|\, dt,$$

la suite

$$\Big(\int_{\alpha_n}^{\beta_n} f(t)dt \Big)_n$$

est une suite de Cauchy. Elle converge donc, et sa limite est (par définition) l'intégrale de Riemann généralisée. □

2.3.2. *Convergence dominée.* Soit $(f_n)_n$ une suite de fonctions $I \to \mathbb{R}$ qui converge simplement vers $f : I \to \mathbb{R}$. On dit que la convergence est *dominée par* $g : I \to \mathbb{R}$ ssi

$$|f_n(t)| \leq g(t) \quad \forall t \in I,\ n \in \mathbb{N}.$$

Le *théorème de convergence dominée* dit

Théorème. *Soit $(f_n)_n$ une suite de fonctions continues $I \to \mathbb{R}$ qui converge simplement vers $f : I \to \mathbb{R}$ et dont la convergence est dominée par une fonction $g : I \to \mathbb{R}$.*

Si f est continue et g est intégrable, alors

$$\int_\alpha^\beta f_n(t)dt \to \int_\alpha^\beta f(t)dt, \quad n \to \infty.$$

Nous allons accepter ce résultat.

Remarque. Ce resultat, ici formulé pour des fonctions continues est aussi valable pour les fonctions intégrable au sens de Riemann, et même dans un cadre beaucoup plus large. La démonstration du théorème de convergence dominée (dans le cadre de l'intégrale de Lebesgue) se trouve dans tous les exposés sur la théorie de l'intégration de Lebesgue. L'identification de l'intégrale de Lebesgue avec l'intégrale de Riemann pour des fonctions continue (et même intégrables au sens de Riemann) see trouve par exemple dans W. Rudin : Principles of Mathematical Analysis (thm 11.33, 3rd edition).

Exemple. Soit

$$f_n(t) = \frac{\sqrt{t}}{1 + n^2t^2}, \quad t \in I = [0, 1].$$

La suite $(f_n)_n$ converge simplement vers $f = 0$ sur I. Comme

$$|f_n(t)| \leq \sqrt{t}$$

la convergence est dominée par la fonction $g(t) = \sqrt{t}$ qui est intégrable sur I. Donc

$$\lim_{n\to\infty} \int_0^1 f_n(t)dt = \int_0^1 f(t)dt = 0.$$

Soit

$$g_n(t) = \frac{n^{\frac{3}{2}}\sqrt{t}}{1 + n^2t^2}, \quad t \in I = [0, 1].$$

La suite $(g_n)_n$ converge simplement vers $f = 0$ sur I, mais

$$\lim_{n\to\infty} \int_0^1 g_n(t)dt = \int_0^\infty \frac{\sqrt{t}}{1 + t^2}dt \neq 0 = \int_0^1 f(t)dt.$$

La convergence de $(g_n)_n$ vers f n'est donc dominée par aucune fonction intégrable.

2.3.3. *Intégration p.r.à un paramètre.* L'exemple classique d'une intégrale p.r.à un paramètre qui n'est pas continue est

$$\int_0^\infty \frac{x}{1+x^2t^2}dt = \begin{cases} \frac{\pi}{2} & x>0 \\ 0 & x=0 \\ -\frac{\pi}{2} & x<0. \end{cases}$$

Mais sous des conditions assez générales ceci n'a pas lieu.

Proposition 2.15. *Soit $f: U\times I \to \mathbb{R}$, où U est une partie de $\mathbb{R}^n$. Si*

(i) f est continue et

(ii) il existe $g: I\to\mathbb{R}$ intégrable t.q.

$$|f(x,t)| \le g(t) \quad \forall (x,t)\in U\times I$$

alors la fonction

$$U \ni x \mapsto h(x) = \int_\alpha^\beta f(x,t)dt$$

est continue.

Démonstration. Soit $x\in U$ fixé et soit $(x_n)_n$ une suite dans U qui converge vers x. Alors, pour chaque $t\in I$,

$$f(x_n,t)\to f(x,t) \quad n\to\infty,$$

par (i). De plus, par (ii), cette convergence est dominée par la fonction intégrable g. Par le théorème de convergence dominée on a donc que $h(x_n)$ converge vers $h(x)$ quand $n\to\infty$, i.e. h est continue en x.

□

Exemple. Soient $U=\mathbb{R}^2$, $I=[0,\infty[$ et

$$f(x,y,t) = \frac{\sin(xt+yt^2)}{1+t^2} \quad (x,y,t)\in\mathbb{R}^2\times[0,\infty[[.$$

f est continue et

$$|f(x,y,t)| \le \frac{1}{1+t^2} = g(t).$$

Comme $g: I\to\mathbb{R}$ est intégrable, la fonction

$$h(x,y) = \int_0^\infty \frac{\sin(xt+yt^2)}{1+t^2}dt$$

est continue sur $U=\mathbb{R}^2$.

Proposition 2.16. *Soit $f: U\times I\to\mathbb{R}$, où U est un ouvert de $\mathbb{R}^n$. Si*

(i) $f : U \times int(I) \to \mathbb{R}$ *est de classe* $\mathcal{C}^1$, *et*
(ii) il existe $g : I \to \mathbb{R}$ *intégrable t.q.*

$$|f(x,t)| \quad et \quad \left|\frac{\partial f}{\partial x_j}(x,t)\right| \leq g(t) \quad \forall (x,t) \in U \times I,\ j = 1, \dots, n$$

alors la fonction

$$U \ni x \mapsto h(x) = \int_a^b f(x,t)dt$$

est de classe $\mathcal{C}^1$ *et*

$$\frac{\partial h}{\partial x_j}(x) = \int_\alpha^\beta \frac{\partial f}{\partial x_j}(x,t)dt, \quad \forall x \in U,\ j = 1, \dots, n.$$

Démonstration. Nous allons traiter $j = 1$, les autres j étant pareils.

Fixons $x = (x_1, x') \in U$, et soit $(\varepsilon_n)_n$ une suite qui tend vers 0. Alors

$$\frac{h(x_1 + \varepsilon_n, x') - (x_1, x')}{\varepsilon_n} - \int_\alpha^\beta \frac{\partial f}{\partial x_1}(x_1, x', t)dt =$$

$$= \int_\alpha^\beta \Big(\underbrace{\frac{f(x_1 + \varepsilon_n, x', t) - f(x_1, x', t)}{\varepsilon_n} - \frac{\partial f}{\partial x_1}(x_1, x', t)}_{F_n(t)} \Big) dt$$

Par (i) $F_n \to 0$ (simplement) quand $n \to \infty$. Par le thm des accroissements finis

$$\left|\frac{f(x_1 + \varepsilon_n, x', t) - f(x_1, x', t)}{\varepsilon_n}\right| \leq \sup_y \left|\frac{\partial f}{\partial x_1}(y, x', t)\right|.$$

Donc, par (ii),

$$|F_n(t)| \leq 2g(t) \quad \forall t, n,$$

i.e. la convergence de $(F_n)_n$ est dominée par $2g$.

Par le théorème de convergence dominée on conclut que

$$\lim_{n\to\infty} \frac{h(x_1 + \varepsilon_n, x') - h(x_1, x')}{\varepsilon_n} = \int_\alpha^\beta \frac{\partial f}{\partial x_1}(x_1, x', t)dt.$$

Ceci qui montre que la dérivée partielle de h p.r.à x_1 existe en chaque point $x = (x_1, x') \in U$, et est

$$= \int_\alpha^\beta \frac{\partial f}{\partial x_j}(x,t)dt.$$

Enfin, la proposition précédente donne qu'elle est continue en x.

□

Exemple. La fonction
$$h(x,y)=\int_0^\infty \frac{\sin(xt+yt^2)}{1+t^2}dt$$
n'est pas de classe $\mathcal{C}^1$ sur $U=\mathbb{R}^2$ p.q.
$$\frac{h(\varepsilon,0)-h(0,0)}{\varepsilon}=\frac{1}{\varepsilon}\int_0^\infty \frac{\sin(\varepsilon t)}{1+t^2}dt=\int_0^\infty \frac{\sin(t)}{\varepsilon^2+t^2}dt$$
et
$$\int_0^\infty \frac{\sin(t)}{\varepsilon^2+t^2}dt\to\infty,\quad \varepsilon\to 0.$$
En effet,
$$\int_0^\infty \frac{\sin(t)}{\varepsilon^2+t^2}dt=\int_1^\infty \frac{\sin(t)}{\varepsilon^2+t^2}dt+\int_0^1 \frac{\sin(t)-t}{\varepsilon^2+t^2}dt+\int_0^1 \frac{t}{\varepsilon^2+t^2}dt$$
et (par convergence dominée)
$$\lim_{\varepsilon\to 0}\int_0^\infty \frac{\sin(t)}{\varepsilon^2+t^2}dt=\int_1^\infty \frac{\sin(t)}{t^2}dt+\int_0^1 \frac{\sin(t)-t}{t^2}dt+\lim_{\varepsilon\to 0}\int_0^1 \frac{t}{\varepsilon^2+t^2}dt$$
et
$$\int_0^1 \frac{t}{\varepsilon^2+t^2}dt=\frac{1}{2}\big(\log(1+\varepsilon^2)-\log(\varepsilon^2)\big)\to\infty,\quad \varepsilon\to 0.$$

2.3.4. *Limites.*

Proposition 2.17. *Soit $f_n : U\to\mathbb{R}$ de classe $\mathcal{C}^1$ et telles que la suite $(f_n)_n$ converge simplement vers $f:U\to\mathbb{R}$ et les suites $(\frac{\partial f_n}{\partial x_j})_n$ convergent simplement vers $g_j:U\to\mathbb{R}$, $j=1,\dots,n$.*

Si

(i) g_j est continue, $j=1,\dots,n$, et

(ii) il existe $h:U\to\mathbb{R}$ continue t.q.
$$\left|\frac{\partial f_n}{\partial x_j}(x)\right|\le h(x)\quad \forall x\in U,\ j=1,\dots,n,$$
alors f est de classe $\mathcal{C}^1$ et
$$\frac{\partial f}{\partial x_j}=g_j\quad j=1,\dots,n.$$

Démonstration. Nous allons traiter $j=1$, les autres j étant pareils.

Soient (a,x') et (b,x') deux points fixées dans U, $a<b$, t.q. le segment $[(a,x'),(b,x')]\subset U$. Alors
$$f_n(b,x')-f_n(a,x')=\int_\alpha^\beta \frac{\partial f_n}{\partial x_1}(t,x')dt.$$

Par le théorème de convergence dominée on a

$$\int_\alpha^\beta \frac{\partial f_n}{\partial x_1}(t,x')dt \to \int_\alpha^\beta g_1(t,x')dt, \quad n\to\infty.$$

Comme $f_n(b,x') \to f(b,x')$ et $f_n(a,x') \to f(a,x')$ quand $n\to\infty$ on conclut que

$$f(b,x') - f(a,x') = \int_\alpha^\beta g_1(t,x')dt.$$

Comme $t \mapsto g_1(t,x')$ est continue, cette formule montre que $\frac{\partial f}{\partial x_1}(a,x')$ existe et est égale à $g_1(a,x')$.

□

2.4. **Fonctions à valeurs réelles de classe $\mathcal{C}^2$.** Nous allons nous restreindre au cas d'une fonction $f : U \to \mathbb{R}$, où U est un ouvert de $\mathbb{R}^n$. CM8

Si f est de classe $\mathcal{C}^1$, elle a des dérivées partielles

$$\frac{\partial f}{\partial x_j} : U \to \mathbb{R}, \quad j = 1,\dots,n$$

qui sont des fonctions continues. Il se peut que ces fonctions soient dérivables (partiellement) et alors on note

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

– c'est la *dérivée partielle d'ordre deux p.r.à x_j et à x_i de f en x.*

Exemple. La fonction

$$f(x,y) = x(y+1) + \log(x^2+1) + \sin^2(y), \quad (x,y)\in\mathbb{R}^2,$$

est de classe $\mathcal{C}^1$ et ses dérivées partielles sont

$$\frac{\partial f}{\partial x}(x,y) = y+1+\frac{2x}{x^2+1} \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = x + 2\sin y\cos y.$$

Chacune de ces deux dérivées sont de nouveaux dérivables (partiellement) :

$$\frac{\partial^2 f}{\partial x\partial x}(x,y) = \frac{2-2x^2}{(x^2+1)^2}, \quad \frac{\partial^2 f}{\partial y\partial y}(x,y) = 2\cos^2 y - 2\sin^2 y$$

et

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

Dans cet exemple les dérivées partielles d'ordre deux sont même des fonctions continues.

Définition. f est *de classe* $\mathcal{C}^2$ ssi f est de classe $\mathcal{C}^1$ et toutes les fonctions

$$\frac{\partial f}{\partial x_j} : U \to \mathbb{R}, \quad j = 1, \dots, n,$$

sont de classe $\mathcal{C}^1$.

Proposition 2.18. *Si $f, g : U \to \mathbb{R}$ sont de classe $\mathcal{C}^2$ alors*
(i) $\lambda f + \mu g$ est de classe $\mathcal{C}^2$ pour toutes λ, μ
(ii) fg est de classe $\mathcal{C}^2$
(iii) $\frac{f}{g}$ est de classe $\mathcal{C}^2$ si g ne s'annule pas
(iii) $\varphi \circ f$ est de classe $\mathcal{C}^2$ pour chaque fonction $\varphi : V \to \mathbb{R}$ de classe $\mathcal{C}^2$ où V est un ouvert qui contient $f(U)$.

Démonstration. Ceci suit de Proposition (2.12), des règles habituelles de dérivation et des résultats correspondants pour la continuité. □

Corollaire 2.19. *Chaque fonction polynômiale en n variables est de classe $\mathcal{C}^2$.*

Démonstration. Chaque monôme est un produit de projections sur les coordonnées qui sont des applications linéaires, donc de classe $\mathcal{C}^2$. Chaque polynôme est une combinaison linéaire de monômes. □

2.4.1. *Le théorème de Schwarz.*

Proposition 2.20. *(Le théorème de Schwarz) Si $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^2$ et $a \in U$, alors*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a) \quad \forall i, j = 1, \dots, n.$$

Démonstration. Soit d'abord $n = 2$. Supposons que $a = 0$ et soit, pour $xy \neq 0$,

$$\Phi(x, y) = \frac{f(x, y) - f(0, y) - f(x, 0) + f(0, 0)}{xy} =$$

$$= \frac{1}{xy} \int_0^x \frac{\partial f}{\partial x}(t, y) - \frac{\partial f}{\partial x}(t, 0) dt = \frac{1}{y} \int_0^1 \frac{\partial f}{\partial x}(sx, y) - \frac{\partial f}{\partial x}(sx, 0) ds =$$

$$= \int_0^1 \frac{1}{y} \Big(\int_0^y \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(sx, t) dt \Big) ds = \int_0^1 \Big(\int_0^1 \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(sx, ry) dr \Big) ds.$$

Alors

$$\left| \Phi(x, y) - \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(0, 0) \right| \leq \int_0^1 \Big(\int_0^1 \big(\frac{\partial}{\partial y} \frac{\partial f}{\partial x}(sx, ry) - \frac{\partial}{\partial y} \frac{\partial f}{\partial x}(0, 0) \big) dr \Big) ds$$

Soit $\varepsilon > 0$. Comme $\frac{\partial}{\partial y}\frac{\partial f}{\partial x}$ est continue, il existe $\delta > 0$ t.q.

$$\|(x,y)\| < \delta \quad \Longrightarrow \quad \left|\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x,y) - \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(0,0)\right| < \varepsilon.$$

Donc

$$\|(x,y)\| < \delta \quad \Longrightarrow \quad \left|\Phi(x,y) - \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(0,0)\right| < \int_0^1\int_0^1 \varepsilon dr ds = \varepsilon,$$

i.e.

$$\Phi(x,y) \to \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(0,0), \quad \|(x,y)\| \to 0.$$

Mais

$$\Phi(x,y) = \frac{f(x,y) - f(x,0) - f(0,y) + f(0,0)}{xy} =$$
$$= \frac{1}{xy}\int_0^y \frac{\partial f}{\partial y}(x,t) - \frac{\partial f}{\partial y}(0,t)dt$$

et on obtient (de la même façon) que

$$\Phi(x,y) \to \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(0,0), \quad \|(x,y)\| \to 0.$$

D'où

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

Le cas $a \neq (0,0)$ est ramené au cas $a = 0$ en considérant la fonction $g(x) = f(a+x)$.

Le cas $n \geq 2$ est similaire au cas $n = 2$, mais avec des notations plus compliquées.

$\square$

Corollaire 2.21. *Si f est de classe $\mathcal{C}^2$ alors pour tout $x \in U$*

$$\mathrm{Hess} f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

est une matrice symétrique.

$\mathrm{Hess} f(x)$ s'appelle la matrice *hessienne de f en x.*

Exemple. Si $f(x,y) = x(y+1) + \log(x^2+1) + \sin^2 y$ alors

$$\mathrm{Hess} f(x) = \begin{pmatrix} \frac{2-2x^2}{(x^2+1)^2} & 1 \\ 1 & 2\cos^2 y - 2\sin^2 y \end{pmatrix}$$

Remarque. Le théorème de Schwarz reste valable sous des conditions plus faible – à savoir que f est de classe $\mathcal{C}^1$ avec des dérivées partielles qui sont différentiables en a. Mais alors la démonstration est différente (voir H. Cartan : Calcul différentiel).

Remarque. Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction deux fois dérivables, mais pas de classe $\mathcal{C}^2$, et soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction

$$f(x, y) = g(x), \quad (x, y) \in \mathbb{R}^2.$$

Alors f n'est pas de classe $\mathcal{C}^2$, mais toutes les dérivées partielles d'ordre deux existent et

$$\frac{\partial^2 f}{\partial x \partial x}(x, y) = g''(x), \quad \frac{\partial^2 f}{\partial y \partial y}(x, y) = 0,$$

et

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

La fonction g peut se construire de façon suivant. Si $\varphi(t) = \int_0^t \sin(\frac{1}{s})ds$, alors φ est dérivable et

$$\varphi'(t) = \begin{cases} \sin(\frac{1}{t}) & t \neq 0 \\ 0 & t = 0. \end{cases}$$

On prend $g(x) = \int_0^x \varphi(t)dt$.

2.4.2. *Les polynômes de Taylor de degré* 1 *et* 2*.*

Définition. Soit $a \in U$. Si f est de classe $\mathcal{C}^1$ alors le polynôme

$$T_a^1 f(x) = f(a) + \mathrm{Jac}(a) \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix}$$

$$= f(a) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a)(x_j - a_j)$$

s'appelle le *polynôme de Taylor de degré* 1 *de* f *en* a.

Si f est de classe $\mathcal{C}^2$ alors le polynôme

$$T_a^2 f(x) = T_a^1 f(x) + \frac{1}{2} \begin{pmatrix} x_1 - a_1 & \dots & x_n - a_n \end{pmatrix} \mathrm{Hess} f(a) \begin{pmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{pmatrix}$$

$$= T_a^1 f(x) + \frac{1}{2} \sum_{i,j=1^n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j)$$

s'appelle le *polynôme de Taylor de degré* 2 *de* f *en* a.

Exemple. Si $f(x,y) = x(y+1) + \log(x^2+1) + \sin^2 y$ et $a = (0,0)$ alors $f(a) = 0$ et

$$\mathrm{Jac}f(a) = \begin{pmatrix} 1 & 0 \end{pmatrix} \quad \text{et} \quad \mathrm{Hess}f(a) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

et

$$T_a^1 f(x,y) = x \quad \text{et} \quad T_a^2 f(x,y) = x + x^2 + xy + y^2.$$

Si f est de classe $\mathcal{C}^1$ alors, par la définition de différentiabilité

$$\left| f(x) - T_a^1 f(x) \right| = o(\|x-a\|)$$

mais si f de classe $\mathcal{C}^2$ on a des propriétés plus fortes.

Proposition 2.22. *(Formule de Taylor avec reste intégral.) Si f est de classe $\mathcal{C}^2$ alors*

$$f(x) - T_a^1 f(x) = \int_0^1 (1-t) \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a + t(x-a))(x_i - a_i)(x_j - a_j) dt$$

pour chaque $[x,a] \subset U$.

Démonstration. Soit $a = 0$. On a

$$I = \int_0^1 (1-t) \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(tx) x_i x_j dt =$$

$$= \sum_i x_i \int_0^1 (1-t) \sum_j \frac{\partial g_i}{\partial x_j}(tx) x_j dt = \sum_i x_i \int_0^1 (1-t) \frac{d}{dt} g_i(tx) dt$$

où $g_i = \frac{\partial f}{\partial x_i}$.

Comme

$$\int_0^1 (1-t) \frac{d}{dt} g_i(tx) dt = -g_i(0) + \int_0^1 g_i(tx) dt$$

on obtient

$$I = -\sum_i g_i(0) x_i + \sum_i \int_0^1 g_i(tx) x_i dt =$$

$$= -\sum_i \frac{\partial f}{\partial x_i}(0) x_i + \sum_i \int_0^1 \frac{\partial f}{\partial x_i}(tx) x_i dt =$$

$$= -\sum_i \frac{\partial f}{\partial x_i}(0) x_i + \int_0^1 \frac{d}{dt} f(tx) dt =$$

$$= -\sum_i \frac{\partial f}{\partial x_i}(0) x_i + f(x) - f(0).$$

Le cas $a \neq 0$ est similaire. $\square$

Corollaire 2.23. *(Formule de Taylor avec reste Taylor-Lagrange.) Si f est de classe $\mathcal{C}^2$ et U est convexe, alors*

$$\left|f(x) - T_a^1 f(x)\right| \le \frac{C}{2} \|x-a\|_1^2,$$

où

$$C = \sup_{1\le i,j\le n} \sup_{x\in U} \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right|.$$

Démonstration.

$$\left|\int_0^1 (1-t)\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+t(x-a))(x_i-a_i)(x_j-a_j)dt\right| \le$$

$$\le C\sum_{i,j} |x_i-a_i|\,|x_j-a_j| \int_0^1 (1-t)dt = \frac{C}{2}\|x-a\|_1^2.$$

□

Corollaire 2.24. *(Formule de Taylor avec reste Taylor-Young.) Si f est de classe $\mathcal{C}^2$ alors*

$$\left|f(x) - T_a^2 f(x)\right| = o(\|x-a\|^2).$$

Démonstration. Soit $a=0$. Alors

$$\left|f(x) - T_a^2 f(x)\right| = \left|f(x) - T_a^1 f(x) - \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}(0)x_i x_j\right| =$$

$$= \left|\int_0^1 (1-t)\sum_{i,j} \Big(\frac{\partial^2 f}{\partial x_i \partial x_j}(tx) - \frac{\partial^2 f}{\partial x_i \partial x_j}(0)x_i x_j dt\right| \le$$

$$\le \int_0^1 (1-t)\sum_{i,j} \underbrace{\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(tx) - \frac{\partial^2 f}{\partial x_i \partial x_j}(0)\right|}_{g_{i,j}(tx)} |x_i x_j|\, dt$$

Soit $\varepsilon > 0$. Comme f est de classe $\mathcal{C}^2$ il existe $\delta > 0$ t.q.

$$\|x\| < \delta \quad \Longrightarrow \quad g_{i,j}(x) < \varepsilon, \quad \forall i,j.$$

Donc

$$\left|f(x) - T_a^2 f(x)\right| \le \int_0^1 (1-t)\sum_{i,j} \varepsilon\,|x_i x_j| = \frac{\varepsilon}{2}\|x\|_1^2 \le C\varepsilon\|x\|^2$$

si $\|x\| < \delta$. Ceci montre que $|f(x) - T_a^2 f(x)| = o(\|x\|^2)$.

Le cas $a \neq 0$ est similaire. □

Exemple. Soit $f(x,y) = x\sin(x+y)+y^2\cos y$. On s'intéresse à la limite de

$$\frac{f(x,y)}{(x^2+y^2)^\alpha}$$

quand $\|(x,y)\| \to 0$ pour différentes valeurs de $0 < \alpha \leq 1$. On écrit

$$\frac{f(x,y)}{(x^2+y^2)^\alpha} = \frac{f(x,y) - T^2_{(0,0)}f(x,y)}{(x^2+y^2)^\alpha} + \frac{T^2_{(0,0)}f(x,y)}{(x^2+y^2)^\alpha}.$$

Comme f est de classe $\mathcal{C}^2$, Taylor-Young nous dit que

$$\frac{f(x,y) - T^2_{(0,0)}f(x,y)}{(x^2+y^2)^\alpha} \to 0, \quad \|(x,y)\| \to 0.$$

Il suffit donc d'étudier le deuxième terme.

$$T^2_{(0,0)}f(x,y) = f(0,0) + \frac{\partial f}{\partial x}(0,0)x + \frac{\partial f}{\partial y}(0,0)+$$

$$+\frac{1}{2}\big(\frac{\partial^2 f}{\partial x \partial x}(0,0)x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0,0)xy + \frac{\partial^2 f}{\partial y \partial y}(0,0)y^2\big) = xy + y^2$$

et

$$\frac{xy+y^2}{(x^2+y^2)^\alpha} \to 0, \quad \|(x,y)\| \to 0$$

si $\alpha < 1$. Pour $\alpha = 1$ la limite n'existe pas.

2.4.3. *Meilleure approximation polynômiale.*

Proposition 2.25. *Soit $P : \mathbb{R}^n \to \mathbb{R}$ un polynôme de degré $\leq d$. Si*

$$P(x) = o(\|x\|)^d),$$

alors $P = 0$.

Démonstration. Soit P un monôme de degré d. Alors, pour tout x et tout $\lambda \in \mathbb{R}$,

$$\frac{|P(x)|}{\|x\|^d} = \frac{|P(\lambda x)|}{\|\lambda x\|^d} \to 0, \lambda \to 0,$$

ce qui implique que P est le monôme nul.

Ceci reste toujours vrai si P est un monôme de degré $< d$, et chaque polynôme de degré $\leq d$ est une combinaison linéaire de monôme de degré $\leq d$. $\square$

Il suit de cette proposition que $T^1_a f$ est *la meilleure approximation de $f(x)$ au voisinage de $x = a$ par un polynôme de degré un* p.q. c'est l'unique polynôme de degré un qui vérifie

$$\left|f(x) - T^1_a f(x)\right| = o(\|x-a\|).$$

De même, $T_a^2 f$ est *la meilleure approximation de $f(x)$ au voisinage de $x = a$ par un polynôme de degré deux* p.q. c'est l'unique polynôme de degré deux qui vérifie

$$\left|f(x) - T_a^2 f(x)\right| = o(\|x-a\|^2).$$

Exemple. $P(x,y) = y + x^2 + xy + y^2$ est l'unique polynôme de degré deux qui vérifie

$$\left|x(y+1) + \log(x^2+1) + \sin^2 y - P(x,y)\right| = o(\|(x,y)\|^2).$$

2.5. **Extrema locaux.** Soit U une partie (pas nécessairement ouverte) de $\mathbb{R}^n$ et soit $f : U \to \mathbb{R}$.

Définition. $a \in U$ est un *point d'extremum de f* ssi

$$f(x) \leq f(a), \quad \forall x \in U \quad \text{— un } \textit{point de maximum}$$

ou

$$f(x) \geq f(a), \quad \forall x \in U \quad \text{— un } \textit{point de minimum}.$$

On dit qu'un point d'extremum a de f est *strict* ssi $f(x) \neq f(a)$ pour tous $x \neq a$.

Si f est constante, alors chaque point $a \in U$ est à la fois point de maximum et de minimum de f, mais il n'est jamais strict

Définition. $a \in U$ est un *point d'extremum local de f* ssi il existe un ouvert $V \subset \mathbb{R}^n$ contenant a t.q. a est un point d'extremum de la restriction $f/_{V\cap U} : V \cap U \to \mathbb{R}$.

Un point d'extremum local est ou un *point de maximum local* ou un *point de minimum local.*

Exemple. La fonction $f(x) = x(x^2-1)$, $x \in \mathbb{R}$, a deux points d'extrema locaux, à savoir $x = -\frac{1}{\sqrt{3}}$ qui est un point de maximum local strict $x = \frac{1}{\sqrt{3}}$ qui est un point de minimum local strict.

Définition. Un point $a \in int(U)$ est un *point critique* de f ssi f est différentiable en a et $Df_a = 0$

Exemple. La fonction $f(x) = x(x^2-1)$, $x \in \mathbb{R}$, a deux points critiques, à savoir $\pm\frac{1}{\sqrt{3}}$. En effet

$$f'(x) = 3x^2 - 1 = 0$$

ssi $x = \pm\frac{1}{\sqrt{3}}$

Exemple. La fonction $f(x,y)=x^2+2xy+y^3$, $(x,y)\in\mathbb{R}^2$, a deux points critiques à savoir $(0,0)$ et $(-\frac{2}{3},\frac{2}{3})$. En effet

$$\begin{cases} \frac{\partial f}{\partial x}(x,y)=2x+2y=0\\ \frac{\partial f}{\partial y}(x,y)=2x+3y^2=0\end{cases}$$

ssi $(x,y)=(0,0)$ ou $(x,y)=(-\frac{2}{3},\frac{2}{3})$.

Proposition 2.26. *Soit $a\in int(U)$ un point d'extremum local. Si f est différentiable en a, alors a est un point critique de f.*

Exemple. La fonction $f(x)=x^3$, $x\in\mathbb{R}$, a un point critique, à savoir $x=0$, mais aucun point d'extremum local.

Exemple. Pour la $f(x,y)=x^2+2xy+y^3$, $x\in\mathbb{R}$, les seules points qui peuvent être des extrema locaux sont $(0,0)$ et $(-\frac{2}{3},\frac{2}{3})$.

Démonstration. Soit $n=2$. Si (a,b) est un point de minimum local, alors $f(x,y)\geq f(a,b)$ pour $\|(x-a,y-b)\|_2<r$. Alors, pour $|x|<r$,

$$\frac{f(x+a,b)-f(a,b)}{x}\begin{cases}\geq 0 & x>0\\ \leq 0 & x<0\end{cases}$$

ce qui implique que

$$\frac{\partial f}{\partial x}(a,b)=\lim_{x\to 0}\frac{f(x+a,b)-f(a,b)}{x}=0.$$

Si a est un point de maximum local, alors $f(x,y)\leq f(a,b)$ pour $\|(x-a,y-b)\|_2<r$. Alors, pour $|x|<r$,

$$\frac{f(x+a,b)-f(a,b)}{x}\begin{cases}\leq 0 & x>0\\ \geq 0 & x<0\end{cases}$$

ce qui implique que

$$\frac{\partial f}{\partial x}(a,b)=\lim_{x\to 0}\frac{f(x+a,b)-f(a,b)}{x}=0.$$

On a donc $\frac{\partial f}{\partial x}(a,b)=0$, et de la même façon $\frac{\partial f}{\partial y}(a,b)=0$.

Le cas $n\geq 3$ est similaire. □

2.5.1. *Matrices symétriques.* Soit A une matrice symétrique $n\times n$. CM9

Définition. On dit que A est *positive* ssi toutes les valeurs propres de A sont ≥ 0.

On dit que A est *définie positive* ssi toutes les valeurs propres de A sont >0.

On dit que A est *négative/définie négative* ssi $-A$ est positive/définie positive.

Une matrice qui n'est ni positive, ni négative a (au moins) une valeur propre >0 et (au moins) une valeur propre <0.

Soit q_A la forme quadratique associée à A :

$$q_A(x) = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Lemma 2.27.

(i) A est positive ssi

$$q_A(x) \geq 0 \quad \forall x \in \mathbb{R}^n$$

(ii) A est définie positive ssi

$$q_A(x) > 0 \quad \forall x \in \mathbb{R}^n,\ x \neq 0;$$

Notez que si A est définie positive alors $\sqrt{q_A}$ est une norme euclidienne. C'est la norme associée au produit scalaire

$$(x,y) \mapsto \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Démonstration. La matrice symétrique A est diagonalisable dans une base orthonormée, i.e. il existe un matrice orthogonale C t.q.

$$D = C^{-1}AC$$

est diagonale avec $\lambda_1, \dots, \lambda_n$ – les valeurs propres de A – sur la diagonale. On observe que

$$\sum_j \lambda_j x_j^2 = q_D(x) = q_A(Cx).$$

On a donc A positive ssi chaque $\lambda_j \geq 0$ ssi $q_A(Cx) \geq 0$ pour chaque $x \in \mathbb{R}^n$ ssi (parce que C est inversible) $q_A(x) \geq 0$ pour chaque $x \in \mathbb{R}^n$.

Le cas "définie negative" est pareil. □

2.5.2. *Caractérisation des points critiques.*

Proposition 2.28. *Soient U ouvert et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^2$. Soit $a \in U$ un point critique de f.*

(i) Si $\mathrm{Hess} f(a)$ est définie positive, alors a est un point de minimum local de f et, en plus, strict.

(ii) Si $\mathrm{Hess} f(a)$ est définie négative, alors a est un point de maximum local de f et, en plus, strict.

(iii) Si $\mathrm{Hess} f(a)$ n'est ni positive ni négative, alors a n'est pas un point d'extremum local de f.

Exemple. Les points critiques de la fonction $f(x,y) = x^2 + 2xy + y^3$, $(x,y) \in \mathbb{R}^2$, sont $(0,0)$ et $(-\frac{2}{3}, \frac{2}{3})$.

La matrice

$$\text{Hess}f(-\frac{2}{3}, \frac{2}{3}) = \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix}$$

a deux valeurs propres strictement positives (le déterminant et la trace sont > 0). Le point $(-\frac{2}{3}, \frac{2}{3})$ est donc un minimum local strict de f.

La matrice

$$\text{Hess}f(0,0) = \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix}$$

a deux valeurs propres de signe différent (le déterminant est < 0). Le point $(0,0)$ n'est donc pas un point d'extremum local de f.

Notez que ni le cas $\text{Hess}f(a)$ *positive mais pas définie positive*, ni le cas $\text{Hess}f(a)$ *négative mais pas définie négative* sont traités par cette proposition.

Exemple. Les deux fonctions $f_\pm(x) = x^2 \pm y^4$, $(x,y) \in \mathbb{R}^2$, n'ont qu'un point critique, à savoir $(0,0)$. Ce point est un minimum local strict pour f_+ mais il n'est pas un point d'extremum local pour f_-. Pourtant les deux fonctions ont la même matrice hessienne en $(0,0)$ – elle est postive, mais pas définie positive.

Démonstration. Supposons d'abord que $a = 0$ et $f(a) = 0$. Si l'on note $A = \frac{1}{2}\text{Hess}f(a)$ alors $T_a^2 f(x) = q_A(x)$ et la formule de Taylor dit que

$$(*) \quad |f(x) - q_A(x)| = o(\|x\|^2).$$

(i) Dans ce cas $\sqrt{q_A}$ est une norme sur $\mathbb{R}^n$ et, comme elle est équivalente à $\| \; \|$, $(*)$ implique que

$$|f(x) - q_A(x)| = o(q_A(x)).$$

Il existe donc un $\delta > 0$ t.q.

$$\|x\| < \delta \quad \Longrightarrow |f(x) - q_A(x)| \leq \frac{1}{2} q_A(x).$$

En particulier

$$f(x) \geq \frac{1}{2} q_A(x), \quad \|x\| < \delta,$$

i.e. $a = 0$ est un point de minimum local strict de f.

(ii) De (i) il suit que $a = 0$ est un point de minimum local strict de $-f$, i.e. $a = 0$ est un point de maximum local strict de f.

(iii) Alors il existe deux vecteurs $u, v \in \mathbb{R}^n$ t.q. $q_A(u) > 0$ et $q_A(v) < 0$. $(*)$ implique que

$$|f(tu) - q_A(tu)| = o(\|tu\|^2) = o(t^2).$$

Il existe donc un $\delta > 0$ t.q.

$$|t| < \delta \quad \Longrightarrow |f(tu) - q_A(tu)| \leq \frac{1}{2} q_A(u) t^2.$$

En particulier

$$f(tu) \geq \frac{1}{2} q_A(tu) = \frac{1}{2} q_A(u) t^2, \quad |t| < \delta,$$

i.e. $a = 0$ n'est pas un point de maximum local de f.

Aussi $(*)$ implique que

$$|f(tv) - q_A(tv)| = o(\|tv\|^2) = o(t^2).$$

Il existe donc un $\delta > 0$ t.q.

$$|t| < \delta \quad \Longrightarrow |f(tv) - q_A(tv)| \leq \frac{1}{2}(-q_A(v)) t^2.$$

En particulier

$$f(tv) \leq \frac{1}{2} q_A(tv) = \frac{1}{2} q_A(v) t^2, \quad |t| < \delta,$$

i.e. $a = 0$ n'est pas un point de minimum local de f.

Le cas général est réduit au cas précédent en considérant la fonction $g(x) = f(x+a) - f(a)$. En effet $\mathrm{Hess}g(0) = \mathrm{Hess}f(a)$, et 0 est un point de maximum/minimum local (strict) de g ssi a est un point de maximum/minimum local (strict) de f. □

2.6. **Fonctions de classe $\mathcal{C}^k$ à variables réelles.** Nous allons nous restreindre au cas d'une fonction $f : U \to \mathbb{R}$, où U est un ouvert de $X = \mathbb{R}^n$.

Définition. La fonction $f : U \to \mathbb{R}$ est *de classe* $\mathcal{C}^k$, $k \geq 2$, ssi toutes les dérivées partielles de f existent et les fonctions

$$\frac{\partial f}{\partial x_j} : U \to \mathbb{R}$$

sont de classe $\mathcal{C}^{k-1}$, $j = 1, \ldots, n$.

f est *de classe* $\mathcal{C}^\infty$ ssi elle est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.

Si f est de classe $\mathcal{C}^k$ alors on note

$$\frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_2} \partial x_{i_1}}(x) = \frac{\partial}{\partial x_{i_k}} \ldots \frac{\partial}{\partial x_{i_2}} \frac{\partial f}{\partial x_{i_1}}(x)$$

– le *k-ième dérivée partielle p.r.à* $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$. C'est la fonction obtenue en dérivant f k fois : d'abord p.r.à x_{i_1}, ensuite p.r.à x_{i_2}, ensuite p.r.à x_{i_3} etc. etc.

2.6.1. *Règles.*

Proposition 2.29. *Si $f, g : U \to \mathbb{R}$ sont de classe $\mathcal{C}^k$, alors*
(i) pour tous $\lambda, \mu \in \mathbb{R}$, la fonction $\lambda f + \mu g$ est de classe $\mathcal{C}^k$.
(ii) la fonction fg est de classe $\mathcal{C}^k$
(iii) si g ne s'annule pas, alors la fonction $\frac{f}{g}$ est de classe $\mathcal{C}^k$,
(iv) $\varphi \circ f$ est de classe $\mathcal{C}^k$ pour chaque fonction $\varphi : V \to \mathbb{R}$ de classe $\mathcal{C}^k$, où V est un ouvert qui contient $f(U)$.

Démonstration. Ceci suit des règles de dérivation, qui sont valables aussi pour les dérivations partielles, et des propriétés de continuité. □

Corollaire 2.30. *Chaque fonction polynômiale en n variables est de classe $\mathcal{C}^\infty$.*

Démonstration. Chaque monôme est un produit de projections sur les coordonnées qui sont des applications linéaires, donc de classe $\mathcal{C}^1$. Chaque polynôme est une combinaison linéaire de monômes.

□

Exemple. La fonction $f : \mathbb{R}^3 \to \mathbb{R}$,

$$f(x, y, z) = xy^2 + \sin(yz^2)$$

est de classe $\mathcal{C}^\infty$.

2.6.2. *Le théorème de Schwarz.*

Proposition 2.31. *(Le théorème de Schwarz) Si $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^k$, alors*

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_2} \partial x_{i_1}}(x), \quad x \in U,$$

ne dépend pas de l'ordre des dérivations.

Démonstration. induction sur k
soit $j = (j_k, \dots, j_1)$ une permutation de $i = (i_k, \dots, i_1)$
si $j_k = i_k$ alors

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(x) = \frac{\partial}{\partial x_{i_k}} \frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \dots \partial x_{i_1}}(x) =$$

$$= \frac{\partial}{\partial x_{j_k}} \frac{\partial^{k-1} f}{\partial x_{j_{k-1}} \dots \partial x_{j_1}}(x)$$

si $j_k \neq i_k$ alors on peut supposer que $j_k = i_{k-1}$ et que $i_k = j_{k-1}$. Alors

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(x) = \frac{\partial}{\partial x_{i_k}} \frac{\partial}{\partial x_{i_{k-1}}} \frac{\partial^{k-2} f}{\partial x_{i_{k-2}} \dots \partial x_{1_1}}(x) =$$

$$= \frac{\partial}{\partial x_{j_k}} \frac{\partial}{\partial x_{j_{k-1}}} \frac{\partial^{k-2} f}{\partial x_{j_{k-2}} \dots \partial x_{j_1}}(x).$$

$\square$

2.6.3. *Les polynômes de Taylor de degré k.*

Définition. Soit $a \in U$. Si f est de classe $\mathcal{C}^k$, $k \geq 3$ alors le polynôme

$$T_a^k f(x) = T_a^{k-1} f(x) +$$
$$+ \frac{1}{k!} \sum_{i_1, i_2, \dots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_2} \partial x_{i_1}}(a)(x_{i_k} - a_{i_k}) \dots (x_{i_1} - a_{i_1})$$

s'appelle le *polynôme de Taylor de degré k de f en a.*

Cette une formule lourde qui est un peu accessible quand $n = 1$ et 2. Pour $n = 1$ elle s'écrit :

$$\begin{aligned} T_a^k f(x) = & f(a) + \frac{1}{1!} f'(a)(x-a) + \frac{1}{2!} f''(a)(x-a)^2 + \\ & + \dots + \frac{1}{k!} f^{(k)}(a)(x-a)^k. \end{aligned}$$

Pour $n = 2$ elle s'écrit :

$$\begin{aligned} T_{(a,b)}^k f(x,y) = & f(a) + \\ & + \frac{1}{1!} \sum_{j=0}^{1} \frac{\partial f}{\partial x^{1-j} \partial y^j}(a,b)(x-a)^{1-j}(y-b)^j + \\ & + \frac{1}{2!} \sum_{j=0}^{2} \binom{2}{j} \frac{\partial f}{\partial x^{2-j} \partial y^j}(a,b)(x-a)^{2-j}(y-b)^j + \\ & + \dots + \\ & \frac{1}{k!} \sum_{j=0}^{k} \binom{k}{j} \frac{\partial^k f}{\partial x^{k-j} \partial y^j}(a,b)(x-a)^{k-j}(y-b)^j. \end{aligned}$$

Proposition 2.32. *(Formule de Taylor avec reste intégrale.) Si f est de classe $\mathcal{C}^k$ alors, pour chaque $[x,a] \subset U$,*

$$f(x) - T_a^{k-1} f(x) =$$
$$= \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \sum_{i_1, i_2, \dots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_2} \partial x_{i_1}}(a + t(x-a))(x_{i_k} - a_{i_k}) \dots (x_{i_1} - a_{i_1}) dt.$$

Démonstration. Soit $a = 0$. On va procéder par induction sur k. On sait que c'est vrai pour $k = 2$ ot l'on suppose vraie pour k. On a

$$I = \int_0^1 \frac{(1-t)^k}{k!} \sum_{i_1,i_2,\ldots,i_k,j=1}^{n} \frac{\partial^{k+1} f}{\partial x_j \partial x_{i_k} \ldots \partial x_{i_2} \partial x_{i_1}}(tx) x_j x_{i_k} \ldots x_{i_1} dt$$

$$= \sum_{i_1,i_2,\ldots,i_k=1}^{n} x_{i_k} \ldots x_{i_1} \int_0^1 \frac{(1-t)^k}{k!} \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}(tx) x_j dt =$$

$$= \sum_{i_1,i_2,\ldots,i_k=1}^{n} x_{i_k} \ldots x_{i_1} \int_0^1 \frac{(1-t)^k}{k!} \frac{dg_i}{dt}(tx) dt.$$

où

$$g_i = \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_2} \partial x_{i_1}}, \quad i = (i_k, \ldots, i_1).$$

Comme

$$\int_0^1 \frac{(1-t)^k}{k!} \frac{dg_i}{dt}(tx) dt = -\frac{1}{k!} g_i(0) + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} g_i(tx) dt$$

on obtient

$$I = -\frac{1}{k!} \sum_{i_1,i_2,\ldots,i_k=1}^{n} g_i(0) x_{i_k} \ldots x_{i_1} + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} \sum_{i_1,i_2,\ldots,i_k=1}^{n} g_i(tx) x_{i_k} \ldots x_{i_1} dt =$$

$$= f(x) - T_a^{k-1} f(x) - \frac{1}{k!} \sum_{i_1,i_2,\ldots,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_{k+1}} \ldots \partial x_{i_2} \partial x_{i_1}}(0) x_{i_k} \ldots x_{i_1}$$

et ceci est $f(x) - T_a^k f(x)$.

Le cas $a \neq 0$ est similaire. □

Corollaire 2.33. *(Formule de Taylor avec reste Taylor-Lagrange.) Si f est de classe $\mathcal{C}^k$ et U est convexe, alors*

$$\left| f(x) - T_a^{k-1} f(x) \right| \leq \frac{C}{k!} (\|x - a\|_1)^k,$$

où

$$C = \sup_{1 \leq i_1, i_2, \ldots, i_k \leq n} \sup_{x \in U} \left| \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_2} \partial x_{i_1}}(x) \right|.$$

La démonstration est comme celle déjà faite dans le cas $\mathcal{C}^2$.

Corollaire 2.34. *(Formule de Taylor avec reste Taylor-Young.) Si f est de classe $\mathcal{C}^k$ alors*

$$\left| f(x) - T_a^k f(x) \right| = o(\|x - a\|^k).$$

La démonstration est comme celle déjà faite dans le cas $\mathcal{C}^2$.

2.7. **Différentiabilité – fonctions à valeurs vectorielles.** Nous allons considerer deux espaces vectoriels X et Y munis de deux normes qu'on notera $\| \|_X$ et $\| \|_Y$. Les notions topologiques comme "ouvert" ou "fermé" ou "continuité" vont se référer à cettes normes-là – sans que cela soit explicitement dit. CM10

Soit $f : U \to Y$, où U est *un ouvert* de $(X, \| \|_X)$.

Définition. La fonction $f : U \to Y$ est *différentiable en* $a \in U$ ssi il existe une application linéaire et continue $L : X \to Y$ t.q.

$$(**) \quad \frac{\|f(x+a) - f(a) - L(x)\|_Y}{\|x\|_X} \to 0, \quad \|x\|_X \to 0.$$

f est *différentiable* ssi elle est différentiable en chaque $a \in U$.

En termes du petit o, la condition $(**)$ se formule

$$\|f(x+a) - f(a) - L(x)\|_Y = o(\|x\|_X).$$

Notez que la différentiabilité fait intervenir les deux norme $\| \|_X$ et $\| \|_Y$. C'est donc d'une *différentiabilité p.r. à* $\| \|_X$ *et* $\| \|_Y$ dont on parle ici.

Proposition 2.35. *Si f est différentiable en $a \in U$, alors la fonction linéaire et continue $L : X \to Y$ en $(*)$ est unique. On l'appelle a différentielle de f en a, notée Df_a ou $Df(a)$.*

Démonstration. La preuve est la même que dans le cas $Y = \mathbb{R}$ (déjà faite) remplaçant simplement la valeur absolue $| \, |$ par la norme $\| \|_Y$. □

Si f est une fonction constante alors f est différentiable et Df_a est l'application linéaire nulle :

$$Df_a : x \mapsto 0, \quad x \in X.$$

Si $f = L$ est une fonction linéaire et continue $X \to Y$, alors f est différentiable et $Df_a = L$, $a \in U$.

2.7.1. *Deux exemples.* Soit $X = Y = \mathcal{M}_n$, les deux espaces munis d'une même norme $\| \|$ qu'on va supposer sous-multiplicative. On rappelle que chaque fonction linéaire $L : X \to Y$ est continue p.q. $\dim X = n^2 < \infty$.

Exemple. La fonction $f(A) = A^2$, $A \in \mathcal{M}_n$, est différentiable et

$$Df_A : B \mapsto AB + BA, \quad A \in \mathcal{M}_n.$$

En effet
$$f(A+B) - f(A) = AB + BA + B^2$$
où l'on voit la fonction
$$L : B \to AB + BA$$
qui est linéaire et continue. De plus
$$\frac{\|f(A+B) - f(A) - L(B)\|}{\|B\|} = \frac{\|B^2\|}{\|B\|} \leq \|B\| \to 0, \quad \|B\| \to 0.$$

On démontre de la même façon que la fonction $f(A) = A^3$, $A \in \mathcal{M}_n$, est différentiable et
$$Df_A : B \mapsto A^2B + ABA + BA^2, \quad A \in \mathcal{M}_n.$$

Exemple. On rappelle que
$$Gl_n = \{A \in \mathcal{M}_n : \ A \text{ inversible}\}$$
est un ouvert de $\mathcal{M}_n$.

La fonction $f(A) = A^{-1}$, $A \in U = Gl_n$, est différentiable et
$$Df_A : B \mapsto -A^{-1}BA^{-1}, \quad A \in U.$$

Si $A \in U$ et $A + B \in U$, alors
$$f(A+B) - f(A) = (I + C)^{-1}A^{-1} - A^{-1}, \quad C = A^{-1}B,$$
et si $\|C\| < 1$ ceci est
$$= -A^{-1}BA^{-1} + \Big(\sum_{n\geq 2}(-C)^n\Big)A^{-1}.$$
Ici l'on voit la fonction
$$L : B \mapsto -A^{-1}BA^{-1}$$
qui est linéaire et continue. De plus
$$\frac{\|f(A+B) - f(A) - L(B)\|}{\|B\|} \leq$$
$$\leq \Big(\sum_{n\geq 2}\|C\|^n\Big)\frac{\|A^{-1}\|}{\|B\|} \leq \frac{\|A^{-1}\|^3}{1 - \|C\|}\|B\| \to 0, \quad \|B\| \to 0.$$

2.7.2. *Continuité et dérivées directionnelles.* On dit que f est *dérivable suivant le vecteur* $v \in X$ en a ssi il existe $u \in Y$ t.q.

$$\frac{1}{t}\big(f(a+tv) - f(a)\big) \to u, \quad t \to 0.$$

Si c'est le cas, le vecteur u est unique et s'appelle la *dérivée de f suivant le vecteur v en a* — notée $\frac{\partial f}{\partial v}(a)$.

Proposition 2.36. *Soit f différentiable en $a \in U$. Alors*
(i) f est continue en a
(ii) f est dérivables suivant tout vecteurs $v \in X$ en a et

$$\frac{\partial f}{\partial v}(a) = Df_a(v).$$

Démonstration. (i) On a

$$\|f(x+a) - f(a)\|_Y \leq \|f(x+a) - f(a) - Df_a(x)\|_Y + \|Df_a(x)\|_Y \to 0, \quad \|x\|_X \to 0.$$

(ii)

$$\frac{\partial f}{\partial v}(a) - Df_a(v) = \lim_{t\to 0} \frac{1}{t}(f(a+tv) - f(a) - Df_a(tv))$$

et

$$\frac{\|f(a+tv) - f(a) - Df_a(tv)\|_Y}{|t|} = \frac{\|f(a+tv) - f(a) - Df_a(tv)\|_Y}{\|tv\|_X}\|v\|_X$$

et ceci $\to 0$ quand $t \to 0$.

□

Soit $Y = \mathbb{R}^m$ et notons les vecteurs dans Y sous formes colonnes :

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}.$$

Corollaire 2.37.
(i) f est différentiable en a ssi chaque composante f_j est différentiable en a.
(ii) Si f est différentiable en a, alors

$$Df_a(x) = \begin{pmatrix} D(f_1)_a \\ D(f_2)_a \\ \vdots \\ D(f_m)_a \end{pmatrix}$$

(iii) Si f est différentiable en a et si $X = \mathbb{R}^n$, alors

$$Df_a(x) = \underbrace{\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \dots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}}_{\mathrm{Jac} f(a)} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

La matrice $\mathrm{Jac}\, f(a)$ s'appelle la matrice *jacobienne de f en a.*

Nota Bene. Df_a n'est pas une matrice, mais une *fonction linéaire.*

Exemple. Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ la fonction

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x^2 y + \cos(xz) \\ xyz \end{pmatrix}$$

Alors f est différentiable p.q. chaque composante est différentiable et

$$\mathrm{Jac} f(x,y,z) = \begin{pmatrix} 2xy - \sin(xz)z & x^2 & -\sin(xz)x \\ yz & xz & xy \end{pmatrix}.$$

Démonstration. (i)+(ii) Soient

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \quad \text{et} \quad L = \begin{pmatrix} L_1 \\ \vdots \\ L_m \end{pmatrix}.$$

Alors

$$\|f(x+a) - f(a) - L(x)\|_Y = o(\|x\|_X)$$

ssi

$$\|f(x+a) - f(a) - L(x)\|_\infty = o(\|x\|_X)$$

ssi

$$\max_j \big(|f_j(x+a) - f_j(a) - L_j(x)| \big) = o(\|x\|_X)$$

ssi

$$|f_j(x+a) - f_j(a) - L_j(x)| = o(\|x\|_X), \quad \forall j = 1, \dots, m.$$

(iii) suit de (ii) et de Corollaire 2.3. □

2.7.3. *Linéarité.*

Proposition 2.38. *Si $f, g : U \to Y$ sont différentiables en $a \in U$, alors, pour tous $\lambda, \mu \in X$, la fonction $\lambda f + \mu g$ est différentiable en $a \in U$ et*

$$D(\lambda f + \mu g)_a = \lambda D f_a + \mu D g_a.$$

Une verification facile.

2.7.4. *Composition.* Soit Z un troisième e.v. muni d'une norme $\| \|_Z$. Soit $g : V \to Z$, où V est un ouvert dans Y qui contient $f(U)$. Alors la fonction

$$g \circ f : U \to Z$$

est bien définie.

Proposition 2.39. *Si f est différentiable en $a \in U$ et si g est différentiable en $b = f(a) \in V$, alors $g \circ f$ est différentiable en a et*

$$D(g \circ f)_a = D g_b \circ D f_a.$$

En particulier, si $X = \mathbb{R}^n$, $Y = \mathbb{R}^k$ et $Z = \mathbb{R}^k$ alors

$$Jac(g \circ f)(a) = Jac\, g(b) \cdot Jac\, f(a).$$

Exemple. Soient $f : \mathbb{R}^3 \to \mathbb{R}^2$ et $g : \mathbb{R}^2 \to \mathbb{R}$ les fonctions

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x^2 y + \cos(xz) \\ xyz \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} x \\ y \end{pmatrix} \mapsto xy.$$

Alors $g \circ f : \mathbb{R}^3 \to \mathbb{R}$ est la fonction

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto (x^2 y + \cos(xz)) xyz$$

et

$$\text{Jac}(g \circ f)(x, y, z) = \text{Jac} g(f(x, y, z)) \text{Jac} f(x, y, z) =$$

$$= \begin{pmatrix} xyz & x^2 y + \cos(xz) \end{pmatrix} \cdot \begin{pmatrix} 2xy - \sin(xz) z & x^2 & \sin(xz) x \\ yz & xz & xy \end{pmatrix}$$

Démonstration. La preuve est la même que dans le cas $Y = \mathbb{R}$ (déjà faite) remplaçant simplement la valeur absolue $| \, |$ par la norme $\| \|_Y$. $\square$

2.7.5. *Théorème des accroissements finis.* Soit $f : \mathbb{R} \to \mathbb{R}^2$ la fonction

$$x \mapsto \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = \begin{pmatrix} x(1-x) \\ x^3 \end{pmatrix}.$$

Alors

$$f_1(1) - f_1(0) = 0 = f_1'(x)(1-0) \qquad \Longleftrightarrow \qquad x = \frac{1}{2}$$

et

$$f_2(1) - f_2(0) = 1 = f_2'(x)(1-0) \qquad \Longleftrightarrow \qquad x = \pm\frac{1}{\sqrt{3}},$$

i.e. il n'est pas possible de trouver $x \in \mathbb{R}$ t.q.

$$f(1) - f(0) = f'(x)(1-0).$$

Le théorème des accroissements finis dans sa forme d'égalité n'est donc pas valable pour des fonctions à valeurs vectorielles.

Proposition 2.40. *Soit $f : U \to Y$ une fonction différentiable.*

(i) Si $[a,b] \subset U$, alors il existe $c \in [a,b]$ t.q.

$$\|f(a) - f(b)\|_Y \leq |||Df_c|||\, \|b-a\|_X\,.$$

(ii) Si U est convexe, alors

$$\|f(a) - f(b)\|_Y \leq \sup_{c \in U} |||Df_c|||\, \|b-a\|_X \quad \forall a, b \in U.$$

— ici $|||Df_c|||$ est la norme subordonnée de l'application Df_c.

Exemple. Soit $X = Y = \mathcal{M}_2$, les deux espaces munis de la même norme

$$\left\| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right\| = |a| + |b| + |c| + |d|\,.$$

Soit

$$U = \{A \in \mathcal{M}_2 : \|A\| < 1\}.$$

Alors, pour toutes $A, B \in U$,

$$\left\|B^2 - A^2\right\| \leq 2\,\|B - A\|\,.$$

En effet, $f(A) = A^2$ est différentiable et

$$Df_A : C \mapsto AC + CA,$$

d'où

$$|||Df_A||| \leq 2\,\|A\|$$

qui est ≤ 2 si $A \in U$. Comme U est convex (p.q. U est une boule pour $\|\ \|$), le thm des accroissements finis donne alors

$$\|f(B) - f(A)\| \leq 2\,\|B - A\|\,.$$

Démonstration. (ii) est une conséquence de (i). Il s'agit donc de montrer (i).

Soit $L : Y \to \mathbb{R}$ une fonction linéaire et continue, et soit $f_L = L \circ f : U \to \mathbb{R}$. Alors f_L est différentiable et $D(f_L)_a = L \circ Df_a$.

Par Proposition (2.9) il existe $c \in [a, b]$ t.q.

$$f_L(b) - f_L(a) = D(f_L)_c(b-a) = L(Df_c(b-a)),$$

ce qui implique

$$|L(f(b) - f(a))| \leq |||L||| \cdot \|Df_c(b-a)\|_Y \leq |||L||| \cdot |||Df_c||| \cdot \|b-a\|_X .$$

Supposons que $f(b) \neq f(a)$ (sinon, il n'y a rien a démontrer) et soit

$$v = \frac{1}{\|f(b) - f(a)\|_Y}(f(b) - f(a))$$

– v est un vecteur unitaire dans Y. Soit maintenant une fonction linéaire t.q.

$$(*) \quad L(v) = 1 \quad \text{et} \quad |||L||| = 1.$$

Alors

$$\|f(b) - f(a)\|_Y = L(f(b) - f(a)) \leq |||L||| \cdot |||Df_c||| \cdot \|b-a\|_X = |||Df_c||| \cdot \|b-a\|_X .$$

. □

Remarque. Dans la démonstration nous avon utilisé l'existence d'une fonction linéaire $L : X \to \mathbb{R}$ vérifiant $(*)$. L'existence d'une telle fonction est facile de montrer si $(X, \| \ \|_X)$ est une espace euclidien (et même un espace préhilbertien) mais bien plus difficile en générale. C'est une conséquence d'un célèbre résultat en analyse fonctionnelle : *le théorème de Hahn-Banach.*

Corollaire 2.41. *Soit $f : U \to Y$ une fonction différentiable. Si U est convexe et*

$$Df_a = 0 \quad \forall a \in U,$$

alors f est une fonction constante.

Démonstration. Par la proposition on voit que $f(a) = f(b)$ pour tous $a, b \in U$. □

2.7.6. *Equivalence.* La notion de différentiabilité et la différentielle sont définies avec des deux normes $\| \ \|_X$ et $\| \ \|_Y$ sur X, et Y. Mais

Proposition 2.42. *Soient N_X N_Y deux autres normes sur X et Y qui sont équivalentes à $\| \ \|_X$ et $\| \ \|_Y$. S'il existe une application linéaire $L : X \to Y$, continue (p.r.à $\| \ \|_X , \| \ \|_Y$), t.q.*

$$\|f(x+a) - f(a) - L(x)\|_Y = o(\|x\|_X),$$

alors L est continue (p.r.à N_X, N_Y) et

$$N_Y(f(x+a) - f(a) - L(x)) = o(N_X(x)).$$

Démonstration. Voir la preuve de Proposition (2.8) □

Si X et Y sont de dimension finie on peut donc parler de la différentiabilité d'une fonction *sans référence à des normes quelconques* parce que toutes les normes (sur X et sur Y respectivement) sont équivalentes.

2.7.7. *La deuxième différentielle.*

Définition. La fonction $f : U \to Y$ est *de classe* $\mathcal{C}^1$ ssi
(i) f est différentiable, et
(ii) la fonction

$$U \ni a \mapsto Df_a \in \mathcal{L}_c(X, Y)$$

est continue p.r.à la norme subordonnée sur $\mathcal{L}_c(X, Y)$, i.e. la norme

$$|||L||| = \sup_{x \neq 0} \frac{\|L(x)\|_Y}{\|x\|_X}.$$

Exemple. Soit $X = Y = \mathcal{M}_n$ munis de la même norme $\| \ \|$ qu'on suppose d'être sous-multiplicative, et soit $f : \mathcal{M}_n \to \mathcal{M}_n$ la fonction $f(A) = A^2$. Elle est de classe $\mathcal{C}^1$.

En effet f est differentiable et

$$Df_A : B \mapsto AB + BA.$$

De plus, pour tout B et C

$$\big(Df_{A+C} - Df_A\big)(B) = Df_{A+C}(B) - Df_A(B) =$$

$$= \big((A+C)B - B(A+C)\big) - \big(AB + BA\big) = CB - BC.$$

D'où

$$\big\|\big(Df_{A+C} - Df_A\big)(B)\big\| \leq 2\,\|B\|\,\|C\|$$

et donc

$$|||Df_{A+C} - Df_A||| \leq 2\,\|C\| \to 0, \quad \|C\| \to 0.$$

La fonction

$$\mathcal{M}_n \ni A \mapsto Df_A \in \mathcal{L}_c(\mathcal{M}_n, \mathcal{M}_n)$$

est donc continue.

Exemple. Soit $X = Y = \mathcal{M}_n$ munis de la même norme $\| \ \|$ qu'on suppose d'être sous-multiplicative. La fonction $f(A) = A^{-1}$, $A \in U = Gl_n$, est de classe $\mathcal{C}^1$.

En effet, f est différentiable et

$$Df_A : B \mapsto -A^{-1}BA^{-1}, \quad A \in U.$$

De plus, pour tout B et tout C t.q. $A + C \in U$,

$$\big(Df_{A+C} - Df_A\big)(B) = -(A+C)^{-1}B(A+C)^{-1} + A^{-1}BA^{-1} =$$

$$= -(I-E)^{-1}A^{-1}B(I-E)^{-1}A^{-1} + A^{-1}BA^{-1}, \quad E = -A^{-1}C.$$

Si $\|E\| < 1$ alors

$$\big(Df_{A+C} - Df_A\big)(B) = -\sum_{m,n\geq 0} E^m A^{-1}BE^n A^{-1} + A^{-1}BA^{-1}$$

et

$$\big\|\big(Df_{A+C} - Df_A\big)(B)\big\| \leq \sum_{m,n\geq 1} \|E\|^{m+n} \left\|A^{-1}\right\|^2 \|B\| =$$

$$\leq \big((\frac{1}{1-\|E\|})^2 - 1\big) \left\|A^{-1}\right\|^2 \|B\| .$$

Donc

$$|||Df_{A+C} - Df_A||| \leq \big((\frac{1}{1-\|E\|})^2 - 1\big) \left\|A^{-1}\right\|^2 \to 0, \quad \|E\| \to 0.$$

D'ù

$$|||Df_{A+C} - Df_A||| \to 0, \quad \|C\| \to 0$$

p.q.

$$\|E\| \leq \|C\| \left\|A^{-1}\right\| \to 0, \quad \|C\| \to 0.$$

La fonction

$$\mathcal{M}_n \ni A \mapsto Df_A \in \mathcal{L}_c(\mathcal{M}_n, \mathcal{M}_n)$$

est donc continue.

Définition. La fonction $f : U \to Y$ est *deux fois différentiable* en a ssi

(i) f est de classe $\mathcal{C}^1$ sur un ouvert $U' \subset U$ contenant a, et

(ii) la fonction

$$U' \ni a \mapsto Df_a \in \mathcal{L}_c(X, Y)$$

est différentiable en a p.r.à la norme subordonnée sur $\mathcal{L}_c(X, Y)$, i.e. la norme

$$|||L||| = \sup_{x\neq 0} \frac{\|L(x)\|_Y}{\|x\|_X}.$$

Si f est deux fois différentiable en a, alors

$$D(Df)_a \in \mathcal{L}_c(X, \mathcal{L}_c(X,Y))$$

est la *deuxième différentielle* de f en a.

$D(Df)_a$ est un objet délicat, mais on peut l'identifier à une fonction bi-linéaire, bien plus comprehensible.

2.7.8. *Fonctions bi-linéaires.* Pour simplifier un peu nous allons supposer que X est de dimension finie. On note

$$\mathcal{L}_c^2(X,Y) = \{b : X \times X \to Y : \text{ bi-linéaire}\}$$

– c'est un espace vectoriel.[2]

Il y a une construction qui permet associer à chaque $L \in \mathcal{L}_c(X, \mathcal{L}_c(X,Y))$ une fonction dans $\mathcal{L}_c^2(X,Y)$:

si $L \in \mathcal{L}_c(X, \mathcal{L}_c(X,Y))$, alors $L(u) \in \mathcal{L}_c(X,Y)$ pour chaque $u \in X$ et on définit

$$b_L : X \times X \ni (u,v) \mapsto L(u)(v) \in Y$$

– c'est une fonction bi-linéaire.

Cette construction nous permet d'identifier la deuxième différentiel $D(Df)_a$ avec la fonction bi-linéaire

$$D^2 f_a = b_{D(Df)_a} \in \mathcal{L}_c^2(X,Y).$$

Exemple. Soit $X = \mathbb{R}^2$ et $Y = \mathbb{R}$. Une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^2$, par exemple

$$f(x,y) = x^2 y + \sin(xy),$$

est deux fois différentiable en chaque point $(a,b) \in \mathbb{R}^2$ et l'on peut montrer que

$$D^2 f_{(a,b)} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$$

est la fonction

$$D^2 f_{(a,b)} : \big((x,y),(\xi,\eta)\big) \mapsto \begin{pmatrix} x & y \end{pmatrix} \mathrm{Hess} f(a,b) \begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

Exemple. Soit $X = Y = \mathcal{M}_n$ et soit $f : \mathcal{M}_2 \to \mathcal{M}_2$ la fonction $f(A) = A^3$.

Elle est différentiable en chaque point $A \in \mathcal{M}_n$ et

$$Df_A : \mathcal{M}_n \to \mathcal{M}_n$$

est la fonction

$$Df_A : B \mapsto A^2 B + ABA + BA^2.$$

2. si X est de dimension ∞ il faut distinguer les fonctions bi-linéaire et continues des fonctions bi-linéaire

On peut montrer que f est deux fois différentiable en chaque point $A \in \mathcal{M}_n$ et

$$D^2 f_A : \mathcal{M}_n \times \mathcal{M}_n \to \mathcal{M}_n$$

est la fonction

$$D^2 f_A : (B, C) \mapsto ABC + BAC + ACB + BCA + CAB + CBA.$$

Remarque. Dans les deux exemples ci-dessus la deuxième différentielle est une application bi-linéaire *symétrique*. Ce n'est pas par hasard : la deuxième différentielle $D^2 f_a$ est toujours une fonction bi-linéaire symétrique.

3. Le théorème d'inversion locale et le théorème des fonctions implicites

CM11

Soit une fonction $f : U \to \mathbb{R}^m$ où U est un ouvert de $\mathbb{R}^n$. On dira que f est de classe $\mathcal{C}^k$ ssi chaque composante $f_1, \ldots, f_m$ est de classe $\mathcal{C}^k$.

Dans toute cette section nous allons supposer, sauf si précisé autrement, que f est (au moins) de classe $\mathcal{C}^1$.

3.1. **Difféomorphismes.**

Définition. Une fonction $f : U \to \mathbb{R}^n$ de classe $\mathcal{C}^1$ est un *difféomorphisme local* ssi

$$Df_a : \mathbb{R}^n \to \mathbb{R}^n$$

est injective pour chaque $a \in U$.

Rappelons que Df_a est injective ssi la matrice jacobienne $\mathrm{Jac} f(a)$ est inversible.

Exemple. Soit

$$f(x) = x^3.$$

C'est une fonction de classe $\mathcal{C}^1$ et $f'(x) = 3x^2 \neq 0$ ssi $x \neq 0$.

La fonction $f : \mathbb{R}^* \to \mathbb{R}$ est donc un difféomorphisme local, mais la fonction $f : \mathbb{R} \to \mathbb{R}$ ne l'est pas.

Exemple. Soit

$$f(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

C'est une fonction de classe $\mathcal{C}^1$ et

$$\mathrm{Jac} f(x, y) = \begin{pmatrix} 2x & 2y \\ 2y & 2x \end{pmatrix}$$

est inversible ssi $(x, y) \neq (0, 0)$.

La fonction $f : \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}^2$ est donc un difféomorphisme local, mais la fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ ne l'est pas.

3.1.1. *Injectivité locale.* Soit $\| \ \|$ une norme sur $\mathbb{R}^n$.

Proposition 3.1. *Soit $f : U \to \mathbb{R}^n$ un difféomorphisme local. Pour tout $a \in U$ il existe $C \geq 0$ et $r > 0$ t.q.*

$$\frac{1}{C} \|x - x'\| \leq \|f(x) - f(x')\| \leq C \|x - x'\|, \quad x, x' \in B_{\leq}(a, r).$$

Démonstration. Soit $B = B_{\leq}(a, s) \subset U$. Par le théorème des accroissements fini on a

$$\|f(x) - f(x')\| \leq \sup_{y \in B} |||Df_y||| \cdot \|x - x'\|, \quad x, x' \in B.$$

Comme f est de classe $\mathcal{C}^1$, la fonction

$$B \ni y \mapsto |||Df_y|||$$

est continue et, donc (par le théorème de la section **??**), bornée. Donc

$$\|f(x) - f(x')\| \leq \alpha \|x - x'\|, \quad x, x' \in B,$$

où

$$\alpha = \sup_{y \in B} |||Df_y|||.$$

Soit $A = \mathrm{Jac} f(a)$ – c'est une matrice inversible. Alors $x \mapsto \|Ax\|$ est une norme sur $\mathbb{R}^n$ et il existe $\beta > 0$ t.q.

$$\frac{1}{\beta} \|x\| \leq \|Ax\| \leq \beta \|x\|, \quad x \in \mathbb{R}^n.$$

Soit $g(x) = f(x) - T_a^1 f(x)$. Comme f est de classe $\mathcal{C}^1$, il existe $t > 0$ t.q.

$$|||Dg_x||| = |||Df_x - Df_a||| \leq \frac{1}{2\beta}, \quad x \in B_{\leq}(a, t),$$

et, par le théorème des accroissements fini on a

$$\|g(x) - g(x')\| \leq \frac{1}{2\beta}, \|x - x'\|, \quad x, x' \in B_{\leq}(a, t)$$

Maintenant,

$$\|f(x) - f(x')\| = \|g(x) - g(x') + A(x - x')\| \geq$$

$$\geq \|A(x - x')\| - \|g(x) - g(x')\| \geq \frac{1}{\beta} \|x - x'\| - \|g(x) - g(x')\|$$

et ceci est

$$\geq \frac{1}{\beta} \|x - x'\| - \frac{1}{2\beta} \|x - x'\| = \frac{1}{2\beta} \|x - a\|, \quad x, x' \in B_{\leq}(a, t).$$

On a donc
$$\frac{1}{C}\|x-x'\| \le \|f(x)-f(x')\| \le C\|x-x'\|, \quad x,x' \in B_{\le}(a,r)$$
pour $C = \max(2\beta, \alpha)$ et $r = \min(s,t)$.

$\square$

Corollaire 3.2. *Soit $f : U \to \mathbb{R}^n$ un difféomorphisme local. Alors, pour tout $a \in U$ il existe un ouvert $U' \ni a$ t.q.*
$$f/_{U'} : U' \to \mathbb{R}^n$$
soit injective – on dit que f est localement injective.

Exemple. Soit
$$f(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$
Pour tout $(a,b) \neq (0,0)$, il existe un ouvert $U' \ni ((a,b)$ t.q. $f/_{U'}$ est injective (p.q. $f : \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}^2$ est un difféomorphisme local).

Par contre ceci n'est pas vrai si $(a,b) = (0,0)$ p.q.
$$f(x,-x) = f(-x,x) \quad x \in \mathbb{R}.$$

3.1.2. *Un théorème de point fixe.* Soit $\| \ \|$ une norme sur $\mathbb{R}^n$.

Proposition 3.3. *Soit $B = B_{\le}(0,r)$ et soit $g : B \to B$ t.q.*
$$\|g(x) - g(x')\| \le \frac{1}{2}\|x - x'\|, \quad x,x' \in B.$$
Alors il existe un unique $x \in B$ t.q.
$$g(x) = x.$$

Démonstration. Soit $x_0 \in B$. On définit une suite $(x_n)_n$ dans B, de façon inductive, par
$$x_{n+1} = g(x_n), \quad n \ge 0.$$

Supposons que cette suite converge et soit x sa limite. Comme B est fermée et comme g est continue on a alors
$$g(x) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

Pourquoi la suite $(x_n)_n$ converge-t-elle ? On a, pour tout $n \ge 1$,
$$\|x_{n+1} - x_n\| = \|g(x_n) - g(x_{n-1})\| \le \frac{1}{2}\|x_n - x_{n-1}\|$$
et, par induction sur n,
$$\|x_{n+1} - x_n\| \le \frac{1}{2^n}\|x_1 - x_0\|.$$

Ceci montre que la série $\sum_{n\geq 0}(x_{n+1}-x_n)$ est absolument convergente, donc convergente. Si y est sa somme, alors

$$y=\sum_{n\geq 0}(x_{n+1}-x_n)=\lim_{k\to\infty}\sum_{n=0}^{k}(x_{n+1}-x_n))=$$

$$=\lim_{k\to\infty}(x_{k+1}-x_0)=\lim_{k\to\infty}x_{k+1}-x_0$$

d'où

$$\lim_{k\to\infty}x_{k+1}=y+x_0=x.$$

Ceci montre l'existence de x, et l'unicite est facile. En effet, si $g(x)=x$ et $g(x')=x'$ alors

$$\|x-x'\|=\|g(x)-g(x')\|\leq\frac{1}{2}\|x-x'\|$$

ce qui implique que $\|x-x'\|=0$, i.e. $x=x'$. □

Remarque. Le même résultat est vrai si $B=B_{\leq}(0,r)$ est une boule dans un espace de Banach $(X,\|\ \|)$ quelconque. La preuve est la même.

Remarque. Le même résultat est vrai si $g:B\to B$ vérifie

$$\|g(x)-g(x')\|\leq\alpha\|x-x'\|,\quad x,x'\in B$$

pour un $\alpha\in]0,1[$ – une telle fonction s'appelle une *contraction*. La preuve est similaire.

3.1.3. *Applications ouvertes.* Une fonction $f:U\to\mathbb{R}^n$ s'appelle une *application ouverte* ssi $f(V)$ est un ouvert de $\mathbb{R}^n$ pour chaque V ouvert $\subset U$.

Exemple. La fonction $f:\mathbb{R}\to\mathbb{R}$ définie par

$$f(x)=x^2,\quad x\in\mathbb{R}$$

n'est pas ouverte – on a, par exemple, $f(]-1,+1[)=[0,1[$. f est de classe $\mathcal{C}^1$, mais elle n'est pas un difféomorphisme local p.q. $f'(0)=0$.

Proposition 3.4. *Un difféomorphisme local $f:U\to\mathbb{R}^n$ est une application ouverte.*

Démonstration. Il suffit de montrer que $f(U)$ est ouvert. (Pourquoi ?)

Soit $a\in U$. On va montrer qu'il existe $r_a>0$ t.q. $B_<(f(a),r_a)\subset f(U)$. Cela sera suffisant p.q. alors

$$f(U)=\bigcup_{a\in U}B_<(f(a),r_a).$$

Cas particulier. On suppose d'abord que $a = 0$, que $f(0) = 0$ et que $Df_0 = \mathrm{id}$ – la fonction d'identité. Comme f est differentiable en 0 on a

$$\|f(x) - x\| = \|f(x) - f(0) - Df_0)x\| = o(\|x\|)$$

ce qui implique qu'il existe $r_1 > 0$ t.q.

$$\|f(x) - x\| \leq \frac{1}{2} \|x\|, \quad \|x\| \leq r_1.$$

Comme f est de classe $\mathcal{C}^1$ il existe $r_2 > 0$ t.q.

$$|||Df_x - Df_0||| \leq \frac{1}{2}, \quad \|x\| \leq r_2.$$

Soit $r = \min(r_1, r_2)$ et fixons $y \in B_{\leq}(0, \frac{r}{2})$. Soit

$$g(x) = y - (f(x) - x).$$

On observe que

$$\|g(x)\| \leq r, \quad \|x\| \leq r,$$

et que

$$|||Dg_x||| = |||Df_x - Df_0||| \leq \frac{1}{2}, \quad \|x\| \leq r.$$

Par le théorème des accroissements finis on a

$$\|g(x) - g(x')\| \leq \sup_{\|z\| \leq r} |||Dg_z||| \cdot \|x - x'\| \leq \frac{1}{2} \|x - x'\|, \quad \|x\|, \|x'\| \leq r.$$

Par Proposition 3.3 il existe donc un (unique) $\|x\| \leq r$ t.q. $g(x) = x$, i.e. $f(x) = y$. Donc

$$B_{\leq}(0, \frac{r}{2}) \subset f(U).$$

Cas général. Soit $b = f(a)$ et rappelons que Df_a est injective, donc bijective. On définie

$$h(x) = (Df_a)^{-1}(x) - a, \quad k(y) = y - b$$

et

$$\tilde{f}(x) = k \circ f \circ h$$

– ce sont des fonctions de classe $\mathcal{C}^1$.

Alors $\tilde{f}$ est définie sur $\tilde{U} = h^{-1}(U)$,

$$\tilde{f}(0) = f(a) - b = 0$$

et

$$D\tilde{f}_0 = \mathrm{id} \circ Df_a \circ (Df_a)^{-1} = \mathrm{id}.$$

On applique le cas particulier à $\tilde{f}$. On a donc

$$B_{\leq}(0, \frac{r}{2}) \subset \tilde{f}(\tilde{U}) = k(f(U)) = f(U) - b,$$

i.e.

$$B_{\leq}(b, \frac{r}{2}) \subset f(U).$$

$\square$

Remarque. Il existe une version beaucoup plus forte de ce résultat. C'est le *théorème de l'invariance de domaine* (de Brouwer) qui dit qu'une fonction continue et localement injective est une application ouverte.

3.1.4. *L'inverse.*

Proposition 3.5. *Soit $f : U \to \mathbb{R}^n$ un difféomorphisme local. Si f est injective, alors la fonction inverse*

$$f^{-1} : V = f(U) \to U$$

est un difféomorphisme local et, pour tout $x \in U$,

$$D(f^{-1})_y = (Df_x)^{-1}, \quad y = f(x).$$

De plus, si f est de classe $\mathcal{C}^k$ $(k \geq 1)$ alors f^{-1} l'est aussi.

Démonstration. $(k = 1, 2)$ Par Proposition (3.4) l'ensemble $V = f(U)$ est un ouvert. Soit $g = f^{-1} : V \to U \subset \mathbb{R}^n$ et soit $b = f(a) \in V$.

g est différentiable en b et $Dg_b = (Df_a)^{-1}$. En effet, si $y = f(x) \in V$, alors

$$\begin{aligned} &\left\| g(y) - g(b) - (Df_a)^{-1}(y - b) \right\| = \\ &= \left\| (Df_a)^{-1}\big(Df_a(x - a) - f(x) + f(a)\big) \right\| \leq \\ &\leq \left\| (Df_a)^{-1} \right\| \|f(x) - f(a) - Df_a(x - a)\| = o(\|x - a\|). \end{aligned}$$

Par Proposition (3.1) on a

$$\frac{1}{C} \|x - a\| \leq \|f(x) - f(a)\| \leq C \|x - \| , \quad x \in B_{\leq}(a, r)$$

pour un $C > 0$ et un $r > 0$, ce qui implique que

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

Donc

$$\left\| g(y) - g(b) - (Df_a)^{-1}(y - b) \right\| = o(\|y - b\|).$$

g est de classe $\mathcal{C}^1$. En effet, comme $f(g(y)) = y$, $y \in V$ on a, par Proposition 2.39,

$$\mathrm{Jac} f(g(y)) \mathrm{Jac} g(y) = I, \quad y \in V,$$

i.e.

$$\mathrm{Jac} g(y) = (\mathrm{Jac} f(g(y)))^{-1}, \quad y \in V.$$

La fonction $G : y \mapsto \mathrm{Jac}g(y)$ est donc la composition de $F : y \mapsto \mathrm{Jac}f(g(y))$ et de

$$\Phi : Gl_n \ni A \mapsto A^{-1} \in \mathcal{M}_n$$

La fonction F est continue p.q. f est de classe $\mathcal{C}^1$ et g est continue (p.q. différentiable). Comme la fonction Φ est continue (p.q. différentiable – voir section 2.7.1), la fonction G est aussi continue. D'où g de classe $\mathcal{C}^1$.

g est de classe $\mathcal{C}^2$ si f est de classe $\mathcal{C}^2$. En effet, la fonction F est de classe $\mathcal{C}^1$ p.q. f est de classe $\mathcal{C}^2$ et g est de classe $\mathcal{C}^1$. Comme la fonction Φ est de classe $\mathcal{C}^1$ (voir section 2.7.7), la fonction G est aussi de classe $\mathcal{C}^1$. D'où g de classe $\mathcal{C}^2$. $\square$

Exemple. La fonction $f :]0, \infty[\times \mathbb{R} \to \mathbb{R}^2$,

$$f(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix},$$

est un difféomorphisme local (c'est dèjà vérifié) et elle est injective. En effet

$$\begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} = (u, v)$$

implique (p.q. $x > 0$) que

$$x^2 - (\frac{v}{2x})^2 = u \quad \iff \quad x^4 - (\frac{v}{2})^2 = ux^2$$

$$\iff \quad x^2 = \frac{u}{2} + \frac{1}{2}\sqrt{u^2 + v^2} \quad \iff \quad x = \sqrt{\frac{u}{2} + \frac{1}{2}\sqrt{u^2 + v^2}}.$$

Son inverse

$$f^{-1} : V = f(U) \to \mathbb{R}^n$$

est donc un difféomorphisme local, et de classe $\mathcal{C}^k$ (p.q. f en est).

Exemple. La fonction $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = x^3, \quad x \in \mathbb{R},$$

est de classe $\mathcal{C}^1$ et injective mais elle n'est pas difféomorphisme local....et son inverse n'est pas de classe $\mathcal{C}^1$ p.q. non-dérivable en 0.

3.1.5. *Difféomorphismes.*

Définition. Soient U et V deux ouverts de $\mathbb{R}^n$. Une fonction $f : U \to V$ de classe $\mathcal{C}^1$ est une *difféomorphisme* ssi

(i) f est bijective, et
(ii) $f^{-1} : V \to U$ est différentiable.

Exemple. L'application $f :]0,\infty[\times\mathbb{R} \to \mathbb{R}^2 \setminus \{(0,0)\}$,

$$f(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix},$$

est un difféomorphisme. En effet

$$\begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} = (u,v)$$

$$\iff \quad x = \sqrt{\frac{u}{2} + \frac{1}{2}\sqrt{u^2+v^2}} \text{ et } y = \frac{v}{2x}.$$

Proposition 3.6. *Un difféomorphisme $f : U \to V$ est un difféomorphisme local.*

En particulier, l'inverse f^{-1} est de classe de classe $\mathcal{C}^k$ $(k \geq 1)$ si f l'est.

Démonstration. On a

$$f^{-1} \circ f(x) = x, \quad \forall x \in U.$$

Par la différentielle d'une composition on a

$$Df_y^{-1} \circ Df_x = \mathrm{id}, \quad y = f(x),$$

ce qui implique que Df_x est injective. $\square$

Proposition 3.7. *(Le théorème d'inversion local.) Soit $f : U \to \mathbb{R}^n$ de classe $\mathcal{C}^1$, et soit $a \in U$. Si Df_a est inversible, alors il existent deux ouverts $U' \ni a$ et $V' \ni f(a)$ t.q.*

$$f/_{U'} : U' \to V'$$

est un difféomorphisme.

Démonstration. Comme f est de classe $\mathcal{C}^1$ on peut, en restreignant U, supposer que Df_a est inversible pour chaque $a \in U$, i.e. que f est un difféomorphisme local.

Par Corollaire 3.2 il existe un ouvert $U' \ni a$ t.q. $f/_{U'}$ soit injective. Par Proposition 3.4, $V' = f(U')$ est un ouvert et, par Proposition 3.5, l'inverse est de classe $\mathcal{C}^1$. $\square$

Le charactère $\mathcal{C}^1$ ici est important.

Exemple. Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction

$$f(x) = \int_0^x \sin(\frac{1}{t})dt + x, \quad x \neq 0$$

et $f(0) = 0$. Alors f est différentiable,

$$f'(x) = \sin(\frac{1}{x}) + 1, \quad x \neq 0$$

et $f'(0) = 1$.

Comme $f'(x) > 0$ sauf pour $x = \frac{2}{-\pi+4n\pi}$ on voit que f est strictement croissante, donc injective. Mais l'inverse f^{-1} n'est pas différentiable au voisinage de 0 p.q. $f'(x) = 0$ pour $x = \frac{2}{-\pi+4n\pi}$.

Remarque. Une fonction $f : U \to \mathbb{R}^n$ différentiable est un *difféomorphisme local* ssi pour chaque $a \in U$ il existent deux ouverts $U' \ni a$ et $V' \ni f(a)$ t.q.

$$f/_{U'} : U' \to V'$$

est bijective et $f^{-1} : V' \to U'$ est différentiable.

Le théorème d'inversion locale montre que cette définition est la même que celle que nous avons choisie quand f est de classe $\mathcal{C}^1$.

CM12

3.2. **Le théorème des fonctions implicites.** Soit $F : U \times V \to \mathbb{R}^m$, où U est un ouvert de $\mathbb{R}^n$ et V est un ouvert de $\mathbb{R}^m$. On va supposer que F est de classe $\mathcal{C}^1$ (au moins).

On s'intéresse aux équations

$$F(x, y) = 0, \quad (x, y) \in U \times V$$

et on se demande si ces équations déterminent (de façon "implicite") une fonction $f : U \to V$ de classe $\mathcal{C}^1$. C'est à dire, existe-t-il une fonction de classe $\mathcal{C}^1$ t.q.

$$F(x, y) = 0 \quad \Longleftrightarrow \quad y = f(x)$$

pour tout $(x, y) \in U \times V$? [3]

Exemple. Soit $F : \mathbb{R}^2 \to \mathbb{R}$,

$$F(x, y) = -x^2 + y - 1.$$

Alors

$$F(x, y) = 0 \quad \Longleftrightarrow \quad y = 1 + x^2$$

et l'équation détermine donc la fonction

$$y = f(x) = 1 + x^2.$$

Mais la situation est rarement si simple et notre question n'admet pas une bonne réponse en général. Une question moins ambitieuse est si, pour une solution $F(a, b) = 0$ donnée, on peut trouver un ouvert $U' \times V'$ contenant (a, b) t.q les équations

$$F(x, y) = 0, \quad (x, y) \in U' \times V'$$

3. on peut, bien sur, aussi s'intéresser à des fonctions continues, mais ici c'est la classe $\mathcal{C}^1$, et $\mathcal{C}^k$, dont il s'agit

détermine une fonction $f : U' \to V'$ de classe $\mathcal{C}^1$. On dit alors que les équations $F(x, y) = 0$ déterminent, *localement sur* $U' \times V'$, une fonction de classe $\mathcal{C}^1$.

Exemple. Soit $F : \mathbb{R}^2 \to \mathbb{R}$,

$$F(x, y) = -x^2 + y^2 - 1.$$

Alors

$$F(x, y) = 0 \quad \iff \quad y^2 = 1 + x^2 \quad \iff \quad y = \pm\sqrt{1 + x^2},$$

et l'on voit que l'équation $F(x, y) = 0$ détermine localement sur $\mathbb{R}\times]0, \infty[$ la fonction

$$y = f(x) = \sqrt{1 + x^2},$$

et localement sur $\mathbb{R}\times] - \infty, 0[$ la fonction

$$y = f(x) = -\sqrt{1 + x^2}.$$

Quand $y = 0$ il n'y a pas de solutions, et ces deux fonctions d'écrivent dont toutes les solutions de $F(x, y) = 0$

Exemple. Soit $F : \mathbb{R}^2 \to \mathbb{R}$,

$$F(x, y) = x^2 + y^2 - 1.$$

Alors

$$F(x, y) = 0 \quad \iff \quad y^2 = 1 - x^2 \quad \iff \quad y = \pm\sqrt{1 - x^2},$$

et l'on voit que l'équation $F(x, y) = 0$ détermine localement sur $] - 1, 1[\times]0, \infty[$ la fonction

$$y = f(x) = \sqrt{1 - x^2},$$

et localement sur $] - 1, 1[\times] - \infty, 0[$ la fonction

$$y = f(x) = -\sqrt{1 - x^2}.$$

Il y a deux autres solutions de $F(x, y) = 0$, à savoir $(1, 0)$ et $(-1, 0)$. Proche de $(1, 0)$ l'équation $F(x, y) = 0$ ne détermine pas y comme une fonction de x, mais x comme une fonction de y :

$$F(x, y) = 0 \quad \iff \quad x = f(y) = \sqrt{1 - y^2}$$

pour tout $(x, y) \in \mathbb{R}\times] - 1, 1[$. Idem pour le point $(-1, 0)$.

Exemple. Soit $F : \mathbb{R}^2 \to \mathbb{R}$,

$$F(x, y) = -x^2 + y^2.$$

Alors

$$F(x, y) = 0 \quad \iff \quad y = \pm x.$$

Cette équation détermine localement sur $\mathbb{R}^*\times]0,\infty[$ la fonction

$$y = f(x) = x$$

et localement sur $\mathbb{R}^*\times]-\infty,0[$ la fonction

$$y = f(x) = -x.$$

Il y a aussi la solution $(0,0)$, mais au voisinage de ce point les solutions de $F(x,y)=0$ ne détermine aucune fonction, ni de x ni de y. (Notez que $(0,0)$ est un point critique de la fonction F.)

Exemple. Et pour la fonction $F:\mathbb{R}^2\to\mathbb{R}$,

$$F(x,y) = y^3 + 4x^4y^2 + 2xy + x^2 - 1,$$

que peut-on dire ?

3.2.1. *Le théorème des fonctions implicites.* Soit

$$F : U\times V \to \mathbb{R}^m$$

une fonction de classe $\mathcal{C}^k$, $k\geq 1$. On va decomposer la matrice jacobienne $\mathrm{Jac}F(x,y)$ en deux parties :

$$\mathrm{Jac}_xF(x,y) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(x,y) & \cdots & \frac{\partial F_1}{\partial x_n}(x,y) \\ \vdots & \vdots & \vdots \\ \frac{\partial F_m}{\partial x_1}(x,y) & \cdots & \frac{\partial F_m}{\partial x_n}(x,y) \end{pmatrix}$$

et

$$\mathrm{Jac}_yF(x,y) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(x,y) & \cdots & \frac{\partial F_1}{\partial y_m}(x,y) \\ \vdots & \vdots & \vdots \\ \frac{\partial F_m}{\partial y_1}(x,y) & \cdots & \frac{\partial F_m}{\partial y_m}(x,y) \end{pmatrix}.$$

Proposition 3.8. *Soit*

$$F : U\times V \to \mathbb{R}^m$$

une fonction de classe $\mathcal{C}^k$, $k\geq 1$, *et soit* $(a,b)\in U\times V$ *t.q.* $F(a,b)=0$.

Si $\mathrm{Jac}_yF(a,b)$ *est inversible, alors*

(I) il existent deux ouverts $a\in U'\subset U$ *et* $b\in V'\subset V$ *et une fonction de classe* $\mathcal{C}^k$,

$$f : U'\to V',\quad f(a)=b,$$

t.q., pour tout $(x,y)\in U'\times V'$,

$$F(x,y)=0 \iff y=f(x),$$

i.e. les équations $F(x, y = 0)$ déterminent, localement sur $U' \times V'$, une fonction f de classe $\mathcal{C}^k$.

(II)

$$Jacf(a) = -Jac_y F(a,b)^{-1} Jac_x F(a,b).$$

Exemple. La fonction $F : \mathbb{R}^2 \to \mathbb{R}$,

$$F(x,y) = y^3 + 4x^4y^2 + 2xy + x^2 - 1,$$

est de classe $\mathcal{C}^k$ et $F(0,1) = 0$. On a

$$\mathrm{Jac}_y F(x,y) = \big(\tfrac{\partial F}{\partial y}(x,y)\big) = \big(2x + 8x^4y + 3y^2\big)$$

et, donc,

$$\mathrm{Jac}_y F(0,1) = \big(3\big)$$

qui est inversible – c'est une matrice 1×1.

Par le TFI avec $(a,b) = (0,1)$ il existent un ouvert $U' \ni a$ et une fonction de classe $\mathcal{C}^k$

$$f : U' \to \mathbb{R}, \quad f(0) = 1,$$

t.q.

$$F(x, f(x)) = 0, \quad x \in U',$$

i.e.

$$(*) \quad f(x)^3 + x^4 f(x)^2 + 2xf(x) + x^2 - 1, \quad x \in U'.$$

La dérivée $f'(0)$ se calcul par (II) :

$$f'(0) = -\frac{1}{\frac{\partial F}{\partial y}(0,1)} \frac{\partial F}{\partial x}(0,1) = -\frac{2}{3}.$$

Les dérivées supérieure de f en 0 peuvent se calculer à partir de l'équation $(*)$.

Exemple. La fonction $F : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$, $x \in \mathbb{R}$, $y = (y_1, y_2) \in \mathbb{R}^2$,

$$F(x,y) = \begin{pmatrix} x^2 + xy_1y_2 + y_1^2y_2^3 - 1 \\ x^3 + y_1^3y_2^2 - 1 \end{pmatrix}$$

est de classe $\mathcal{C}^\infty$ et $F(0,1,1) = 0$.

On a

$$\mathrm{Jac}_y F(x,y) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(x,y) & \frac{\partial F_1}{\partial y_2}(x,y) \\ \frac{\partial F_2}{\partial y_1}(x,y) & \frac{\partial F_2}{\partial y_2}(x,y) \end{pmatrix} = \begin{pmatrix} xy_2 + 2y_1y_2^3 & xy_1 + 3y_1^2y_2^2 \\ 3y_1^2y_2^2 & 2y_1^3y_2 \end{pmatrix}$$

et, donc,

$$\mathrm{Jac}_y F(0,1,1) = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix},$$

qui est une matrice inversible – le déterminant est $= -5$.

Par le TFI avec $(a,b)=(0,1,1)$ il existent un ouvert $U'\ni 0$ et une fonction de classe $\mathcal{C}^k$

$$f=(f_1,f_2):U'\to\mathbb{R}^2,\quad f(0)=(1,1)$$

t.q.

$$F(x,f_1(x),f_2(x))=0,\quad x\in U'.$$

La dérivée de f_1 et de f_2 en 0 se calculent par (II) :

$$\begin{pmatrix} f_1'(0)\\ f_2'(0)\end{pmatrix}=-\begin{pmatrix}2&3\\3&2\end{pmatrix}^{-1}\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}\frac{2}{5}\\ -\frac{3}{5}\end{pmatrix}.$$

Démonstration. La fonction

$$G:U\times V\ni(x,y)\mapsto(x,F(x,y))\ni\mathbb{R}^n\times\mathbb{R}^m$$

est de classe $\mathcal{C}^k$, $G(a,b)=(a,0)$ et

$$\mathrm{Jac}G(x,y)=\begin{pmatrix} I & 0\\ \mathrm{Jac}_xF(x,y) & \mathrm{Jac}_yF(x,y)\end{pmatrix}.$$

On note que $\mathrm{Jac}G(a,b)$ est inversible p.q.

$$\det\mathrm{Jac}G(a,b)=\det\mathrm{Jac}_yF(a,b)\neq 0.$$

Le théorème d'inversion locale implique donc qu'il existe un ouvert $U'\times V'\subset U\times V$ contenant (a,b) t.q.

$$G/_{U'\times V'}:U'\times V'\to G(U'\times V')=W$$

est une $\mathcal{C}^k$-difféomorphisme. Noter que l'image W est de la forme

$$W=\bigcup_{x\in U'}W_x,\quad W_x\subset\mathbb{R}^m.$$

Soit

$$H=(H_1,H_2):W\to U'\times V'$$

son inverse, i.e.

$$G(H(x,y))=(x,y)\quad(x,y)\in W,$$

i.e.

$$(H_1(x,y),F(H_1(x,y),H_2(x,y))=(x,y)\quad(x,y)\in W.$$

D'où

$$H_1(x,y)=x\quad(x,y)\in W$$

et

$$F(x,H_2(x,y))=y\quad(x,y)\in W.$$

Donc,

$$F(x,y)=0\quad\iff\quad y=H_2(x,0)$$

pour tout $(x, y) \in U' \times V'$. Et la fonction
$$f : U' \ni x \mapsto H_2(x, 0)$$
est de classe $\mathcal{C}^k$. Ceci donne (I).

Par composition on sait que, pour tout $(x, y) \in W$,
$$\mathrm{Jac}G(H(x, y)) \cdot \mathrm{Jac}H(x, y) = I.$$
Alors
$$\begin{pmatrix} I & 0 \\ \mathrm{Jac}_x F(a, b) & \mathrm{Jac}_y F(a, b) \end{pmatrix} \cdot \begin{pmatrix} I & 0 \\ \mathrm{Jac}_x H_2(a, 0) & \mathrm{Jac}_y H_2(a, 0) \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}.$$
D'où
$$\mathrm{Jac}_x F(a, b) + \mathrm{Jac}_y F(a, b) \mathrm{Jac}_x H_2(a, 0) = 0,$$
et alors
$$\mathrm{Jac}f(a) = \mathrm{Jac}_x H_2(a, 0) = -\big(\mathrm{Jac}_x F(a, b)\big)^{-1} \mathrm{Jac}_x F(a, b).$$
Ceci donne (II).

□

3.2.2. *Application.*

Proposition 3.9. *Soit*
$$F(x, y) = y^m + \sum_{j=0}^{m-1} p_j(x) y^j$$
un polynôme de degrée k en y dont les coefficients
$$p_0, \ldots, p_{m-1} : U \to \mathbb{R}$$
sont des fonctions de classe $\mathcal{C}^k$ sur $U \subset \mathbb{R}^n$

Soit $a \in U$ et supposons que les m racines du polynôme $F(a, y)$ sont simples et réelles.

Alors il existent un ouvert $U' \subset U$ contenant a et des fonctions de classe $\mathcal{C}^k$,
$$f_1, \ldots, f_m : U' \to \mathbb{R},$$
t.q. $f_1(x), \ldots, f_m(x)$ sont les racines du polynôme $F(x, y)$ pour chaque $x \in U'$.

Exemple. Soit
$$F(x, y) = y^3 + 4x^4y^2 + 2xy + x^2 - 1.$$
Les trois racines de
$$F(1, y) = y^3 + 4y^2 + 2y = y(y^2 + 4y + 2)$$
simples et réelles.

Il existe donc un ouvert U' contenant 1 t.q. toutes les racines du polynôme

$$F(x,y) = y^3 + 4x^4y^2 + 2xy + x^2 - 1, \quad x \in U',$$

sont des fonctions de classe $\mathcal{C}^k$ sur U'.

Démonstration. Soient $b_1, \ldots, b_m$ les m racines de $F(a,y)$. Alors

$$F(a,y) = \prod_{j=1}^{m} (y - b_j)$$

et

$$\frac{\partial F}{\partial y}(a,y) = \sum_{i=1}^{m} \prod_{j \neq i} (y - b_j).$$

Comme $F(a, b_1) = 0$ et

$$\frac{\partial F}{\partial y}(a,b_1) = \prod_{j \neq 1} (b_1 - b_j) \neq 0$$

le TFI montre qu'il existent un ouvert U_1 contenant a et une fonction de classe $\mathcal{C}^k$

$$f_1 : U_1 \to \mathbb{R}, \quad f_1(a) = b_1,$$

t.q.

$$F(x, f_1(x)) = 0, \quad x \in U_1.$$

Idem pour $b_2, \ldots, b_m$, et les fonctions $f_2, \ldots, f_m$ sont définies sur $U = \cap U_j$.

□

Remarque. La proposition n'est pas vrai quand il y a des racines doubles. Voir par exemple $F(x,y) = y^2 - x$ qui a une racine double quand $x = 0$ et dont les racines sont $\pm\sqrt{x}$ pour $x \geq 0$. Par contre, la proposition reste vrai pour des racines complexes simples.

Corollaire 3.10. *Soit $M : U \to \mathcal{M}_m$ classe $\mathcal{C}^k$, U un ouvert dans $\mathbb{R}^n$*

Soit $a \in U$ et supposons que les m valeurs propres de la matrice $M(a)$ sont simples et réelles.

Alors il existent un ouvert $U' \subset U$ contenant a et des fonctions de classe $\mathcal{C}^k$,

$$f_1, \ldots, f_m : U' \to \mathbb{R},$$

t.q. $f_(x), \ldots, f_m(x)$ sont les valeurs propres de $M(x)$ pour chaque $x \in U'$.

Démonstration. On applique la proposition à

$$F(x,y) = -\det\big(M(x) - yI\big).$$

□

Institut de Mathématiques de Jussieu-Paris Rive Gauche, Université de Paris - Campus des Grands Moulins, Bâtiment Sophie Germain, Boite Courrier 7012, 8 Place Aurélie Nemours, 75205 Paris Cedex 13

Email address: `hakan.eliasson@imj-prg.fr`