1.6 Le théorème des accroissements finis

Rappelons le résultat classique pour les fonctions d'une variable réelle à valeurs dans $\mathbb{R}$.

Théorème 1.36 (Théorème des accroissements finis sur $\mathbb{R}$) *Soit f une fonction continue sur un intervalle $[a, b]$, à valeurs dans $\mathbb{R}$, dérivable sur $]a, b[$.*

Alors, il existe $c \in]a, b[$ tel que

$$f(b) - f(a) = f'(c)(b - a).$$

(où il existe $\theta \in]0, 1[$ tel que

$$f(b) - f(a) = f'(a + \theta(b - a))(b - a).$$

Une autre version (plus faible) de ce résultat est l'inégalité des accroissements finis :

Théorème 1.37 (l'inégalité des accroissements finis) *Soit f une fonction continue sur un intervalle $[a, b]$, à valeurs dans $\mathbb{R}$, dérivable sur $]a, b[$. Soit $M \geq 0$ telle que $|f'(x)| \leq M$, pour tout $x \in [a, b]$, alors $|f(y) - f(x)| \leq M|y - x|$ pour tout $x, y \in [a, b]$.*

On en déduit une première extension du théorème des accroissements finis pour les fonctions définies sur un ouvert d'un espace vectoriel normé E à valeurs dans $\mathbb{R}$.

Définition 1.6.1 *1) Soit E un espace vectoriel, $a, b \in E$. Le segment $[a, b]$ est le sous-ensemble de E défini par*

$$[a, b] = \{x \in E; \text{ il existe } t \in [0, 1] \text{ tel que } x = a + t(b - a)\}$$

2) Un sous-ensemble $U \subset E$ est dit **convexe** *si pour tout $a, b \in U$ le segment $[a, b] \subset U$.*

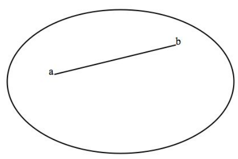

FIGURE 1 – convexe

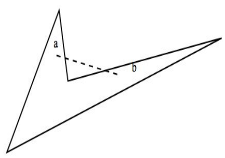

FIGURE 2 – non convexe

Théorème 1.38 (Théorème des accroissements finis à valeurs dans $\mathbb{R}$) .

Soit $f : U \longrightarrow \mathbb{R}$ une fonction différentiable dans l'ouvert $U \subset E$. Soit $a, b \in U$, si le segment $[a, b]$ est contenu dans U, il existe $\theta \in]0, 1[$ telle que

$$f(b) - f(a) = Df(a + \theta(b - a)).(b - a)$$

ce qui est équivalent à : il existe $c \in]a, b[$ tel que $f(b) - f(a) = Df(c).(b - a)$.

Démonstration 1.39 *On applique le théorème des accroissements finis à la fonction $g : [0, 1] \longrightarrow \mathbb{R}$, $t \mapsto g(t) = f(a + t(b - a))$. Alors, $g = f \circ A$ où $A : [0, 1] \longrightarrow \mathbb{R}^n$ est l'application définie par $A(t) = a + t(b - a)$.*
g est différentiable sur $[0, 1]$, comme composée de fonctions différentiables et

$$g'(t) = Df(A(t))(DA(t)) = Df(a + t(b - a)).(b - a)$$

Il existe donc $\theta \in]0, 1[$ tel que $g(1) - g(0) = g'(\theta)$,
qui se traduit par, il existe $\theta \in]0, 1[$ tel que

$$f(b) - f(a) = Df(a + \theta(b - a)).(b - a).$$

Remarque : L'exemple suivant montre que ce résultat est faux si f est à valeurs dans un espace vectoriel de dimension ≥ 2.

Exemple 1.6.2 *Soit $f : \mathbb{R} \longrightarrow \mathbb{R}^2$,$t \mapsto f(t) = (\cos(t), \sin(t))$. Sa différentielle au point t est $Df(t) = (-\sin(t), \cos(t))$,*

Par contre, $f(2\pi) - f(0)$ est le vecteur nulle et ne peut être égal à aucun des vecteurs de norme 2π, $(2\pi - 0)Df(\theta)$, avec $\theta \in \mathbb{R}$.

On va donner deux variantes du théorème précédent valables pour des fonctions à valeurs dans un espace vectoriel de dimension ≥ 2.

Théorème 1.40 (Théorème des accroissements finis à valeurs dans $\mathbb{R}^m$) .

Soit $f : U \longrightarrow \mathbb{R}^m$, $x \mapsto (f_1(x), \ldots, f_m(x))$ une fonction différentiable. Soit $a, b \in U$, si le segment $[a, b]$ est contenu dans U, il existe m points $c_1, \ldots, c_m \in]a, b[$ telle que pour tout $j \in \{1, \ldots, m\}$ on a

$$f_j(b) - f_j(a) = Df(c_j).(b - a)$$

Démonstration 1.41 *On applique 1.38 à chaque composante f_j de f.*

Une autre variante où l'égalité est remplacée par une inégalité sur les normes.

Théorème 1.42 (L'inégalité des accroissements finis) .

Soit $f : U \longrightarrow \mathbb{R}^m$ une application différentiable. U ouvert de $\mathbb{R}^n$.
Pour tout $a, b \in \mathbb{R}^n$ tels que le segment $[a, b] \subset U$, on a :

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

Démonstration 1.43 *On suppose $f(b)-f(a)\neq 0$, sinon l'inégalité est évidente. On pose $v=\frac{f(b)-f(a)}{\|f(b)-f(a)\|}$. On considére la forme linéaire continue ψ , définie par : pour tout $y\in F$, $\psi(y)=<y,v>$ où $<.,.>$ est le produit scalaire standard sur $\mathbb{R}^m$.*

D'après 1.38, il existe $\theta\in]0,1[$ telle que $\psi(f(b)-f(a))=\psi(f(b))-\psi(f(a))=D\psi\circ f(a+\theta(b-a)).(b-a)=\psi(Df(a+\theta_\psi(b-a))(b-a))$ s'écrit aussi $<f(b)-f(a),v>=<Df(a+\theta(b-a))(b-a),v>$ ou encore

$$\|f(b)-f(a)\|=<Df(a+\theta(b-a))(b-a),v>$$

et une application de l'inégalité de Cauchy-Schwarz nous donne :
$|<Df(a+\theta(b-a))(b-a),v>|\leq\|Df(a+\theta(b-a))(b-a)\|\|v\|\leq\|Df(a+\theta(b-a))(b-a)\|$ d'où $\|f(b)-f(a)\|\leq\|Df(a+\theta(b-a))(b-a)\|\leq\|Df(a+\theta(b-a))\|\|(b-a)\|$, finalement $\|f(b)-f(a)\|\leq\|(b-a)\|\left(\sup_{x\in]a,b[}\|Df(x)\|\right)$.
♣

Complément : Le cas général

Proposition 1.44 *Soient E et F deux espaces vectoriels normés et $f:U\longrightarrow F$ une application différentiable. U ouvert de E.*

Soient $a,b\in E$, tels que le segment $[a,b]\subset U$.

Alors, pour toute forme linéaire $\psi\in F'=\mathcal{L}(F,\mathbb{R})$, il existe $\theta_\psi\in]0,1[$ tel que

$$\psi\circ f(b)-\psi\circ f(a)=\psi\circ Df(a+\theta_\psi(b-a))(b-a)$$

ou encore $\psi(f(b)-f(a))=\psi(Df(a+\theta_\psi(b-a))(b-a))$.

Démonstration 1.45 *Soit $\psi\in F'$. On défini $g_\psi:[0,1]\longrightarrow\mathbb{R}$, $t\mapsto g_\psi(t)=\psi\circ f(a+th)$.*

Comme composée d'une application linéaire et d'une application différentiable, g_ψ est différentiable, et d'aprés 1.38, il existe donc $\theta_\psi\in]0,1[$ tel que $g_\psi(1)-g_\psi(0)=g'_\psi(\theta_\psi)$, c-à-d

$$\psi\circ f(b)-\psi\circ f(a)=\psi\circ Df(a+\theta_\psi(b-a))(b-a).♣$$

Comme conséquence on obtient :

Corollaire 1.46 (L'inégalité des accroissements finis à valeurs dans un espace euclidien) .

Soit E un evn et F un espace euclidien munit du produit scalaire $<.,.>$.

Soit $f:U\longrightarrow F$ une application différentiable. U ouvert de E.

Pour tout $a,b\in E$ tels que le segment $[a,b]\subset U$, on a :

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

Démonstration 1.47 *On suppose $f(b)-f(a)\neq 0$, sinon l'inégalité est évidente. On pose $v=\frac{f(b)-f(a)}{\|f(b)-f(a)\|}$. On considére la forme linéaire continue ψ , définie par : pour tout $y\in F$, $\psi(y)=<y,v>$ où $<.,.>$ est le produit scalaire.*

Alors, l'équation $\psi(f(b)-f(a))=\psi(Df(a+\theta_\psi(b-a))(b-a))$ s'écrit $<f(b)-f(a),v>=<Df(a+\theta_\psi(b-a))(b-a),v>$ ou encore

$$\|f(b)-f(a)\|=<Df(a+\theta_\psi(b-a))(b-a),v>$$

et l'inégalité de Cauchy-Schwarz nous donne $|<Df(a+\theta_\psi(b-a))(b-a),v>|\leq\|Df(a+\theta_\psi(b-a))(b-a)\|\|v\|\leq\|Df(a+\theta_\psi(b-a))(b-a)\|$ d'où

$\|f(b)-f(a)\|\leq\|Df(a+\theta_\psi(b-a))(b-a)\|\leq\|Df(a+\theta_\psi(b-a))\|\|(b-a)\|$, a fortiori $\|f(b)-f(a)\|\leq\|(b-a)\|\left(\sup_{x\in]a,b[}\|Df(x)\|\right)$. ♣

et le cas général

Théorème 1.48 (L'inégalité des accroissements finis (cas général)) .

Soient E et F deux evn et $f:U\longrightarrow F$ une application différentiable. U ouvert de E.

Pour tout $a,b\in E$ tels que le segment $[a,b]\subset U$, on a :

$$\|f(b)-f(a)\|\leq\|b-a\|\left(\sup_{x\in]a,b[}\|Df(x)\|\right). \tag{1.5}$$

Démonstration 1.49 *Soit $a, b \in E$, $a \neq b$, tels que le segment $[a, b] \subset U$. On pose $M = \sup\limits_{x \in]a,b[} \|Df(x)\|$. On peut supposer $M < +\infty$ sinon il n'y a rien à démontrer.*

On va montrer que pour tout $\varepsilon > 0$ on a

$$\|f(b) - f(a)\| \leq (M + \varepsilon)\|(b - a)\|. \tag{1.6}$$

Comme ε est arbitraire, on obtient l'inégalité 1.5, en faisant tendre ε vers 0.

On pose

$$E = \{\theta \in [0, 1],\ \text{pour tout } 0 \leq t \leq \theta,\ \|f(a + t(b - a)) - f(a)\| \leq t(M + \varepsilon)\|(b - a)\|\}.$$

L'ensemble E n'est pas vide, puisqu'il contient 0, donc admet un plus grand élément. On note $\xi = \sup E$.

Comme f est continue, $\xi \in E$ d'où $E = [0, \xi]$.

On doit maintenant montrer que $\xi = 1$.

On suppose $\xi < 1$. D'après la définition de la différentiabilité de f au point $a + \xi(b - a)$, il existe $\delta > 0$, avec $\xi + \delta < 1$ et pour tout $t \in [\xi, \xi + \delta[$

$$\frac{\|f(a + t(b - a)) - f(a + \xi(b - a)) - Df(a + \xi(b - a))\left((t - \xi)(b - a)\right)\|}{\|(t - \xi)(b - a)\|} \leq \varepsilon \tag{1.7}$$

ceci entraîne

$$\|f(a + t(b - a) - f(a + \xi(b - a))\| \leq \|\|Df(a + \xi(b - a))\|(t - \xi)(b - a)\| + \varepsilon\|(t - \xi)(b - a)\|.$$

ou encore

$$\|f(a + t(b - a) - f(a + \xi(b - a))\| \leq (M + \varepsilon)(t - \xi)\|(b - a)\|.$$

D'où

$$\begin{aligned}\|f(a + t(b - a)) - f(a)\| &\leq \|f(a + t(b - a)) - f(a + \xi(b - a))\| + \|f(a + \xi(b - a)) - f(a)\| \\ &\leq (t - \xi)(M + \varepsilon)\|(b - a)\| + \xi(M + \varepsilon)\|(b - a)\| \\ &= t(M + \varepsilon)\|(b - a)\|.\end{aligned}$$

Comme $t > \xi$, ceci contredit la définition de ξ. Donc $\xi = 1$. ♣

Remarque : Pour terminer signalons une autre démonstration de ce théorème : D'après la proposition 1.44, pour tout $\psi \in F'$ il existe $\theta_\psi \in]0, 1[$ tel :

$\psi\left(f(b) - f(a)\right) = \psi\left(Df(a + \theta_\psi(b - a))(b - a)\right)$, d'où

$$\begin{aligned}|\psi\left(f(b) - f(a)\right)| = |\psi\left(Df(a + \theta_\psi(b - a))(b - a)\right)| &\leq \|\psi\|.\|Df(a + \theta_\psi(b - a))(b - a)\| \\ &\leq \|\psi\|.\|Df(a + \theta_\psi(b - a))\|\|(b - a)\| \\ &\leq \|\psi\|\|(b - a)\|\left(\sup_{x \in]a,b[}\|Df(x)\|\right).\end{aligned}$$

Une conséquence du théorème de Hahn-Banach (hors programme du cours), dit que

$$\|f(b) - f(a)\| = \sup\{|\psi\left(f(b) - f(a)\right)|; \|\psi\| \leq 1\}$$

D'où

$$\|f(b) - f(a)\| \leq \|(b - a)\|\left(\sup_{x \in]a,b[}\|Df(x)\|\right). \clubsuit$$

1.7 Quelques applications du Théorème des accroissements finis

Corollaire 1.50 *Soient E et F deux evn. $f : U \longrightarrow F$ une application différentiable. U ouvert de E.*

Soit $T \in \mathcal{L}(E, F)$. Pour tout $a, b \in E$ tels que le segment $[a, b] \subset U$, on a :

$$\|f(b) - f(a) - T(b-a)\| \leq \|b-a\| \left(\sup_{x \in]a,b[} \|Df(x) - T\| \right). \tag{1.8}$$

En particulier, si $T = Df(a)$ on aura

$$\|f(b) - f(a) - Df(a)(b-a)\| \leq \|b-a\| \left(\sup_{x \in]a,b[} \|Df(x) - Df(a)\| \right). \tag{1.9}$$

Démonstration 1.51 *Résulte de l'inégalité des accroissement finis appliquée à $f - T$ et de la linéarité de T qui entraîne $DT(x) = T$.*

Définition 1.7.1 (application lipschitzienne) *Soient E et F deux espaces vectoriels normés et $K \geq 0$.*

Une application $f : U \longrightarrow F$ est dite K-lipschitzienne (ou lipschitzienne de rapport K) si pour tout $x, y \in U$ on a

$$\|f(x) - f(y)\| \leq K\|x - y\|.$$

Corollaire 1.52 *Soient E et F deux des espaces vectoriels normés, U ouvert convexe de E.*

Soit $f : U \longrightarrow F$ une application différentiable, on suppose qu'il existe $K \geq 0$ tel que $\|Df(x)\| \leq K$ pour tout $x \in U$. Alors f est K-lipschitzienne .

Remarque : Dans le résultat précédent l'hypothèse de convexité est essentielle. l'exemple suivant montre que le résultat n'est pas nécessairement vrai même si on suppose que U est connexe.
En effet, si $U = \{(x, y) \in \mathbb{R}^2; x > 0 \,\text{et}\, 1 < x^2 + y^2 < 2\}$ une demi-couronne ouverte et $f : U \longrightarrow \mathbb{R}$ définie par : $(x, y) \mapsto f(x, y) = \arctan(\frac{y}{x})$.
Alors $Df(x, y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$, donc $\|Df(x, y)\| = \frac{1}{\sqrt{x^2+y^2}} < 1$. Mais, $\lim_{n \longrightarrow +\infty} |f(\frac{1}{n}, 1) - f(\frac{1}{n}, -1)| = \pi$ et $\|(\frac{1}{n}, 1) - (\frac{1}{n}, -1)\| = 2$, montre que f n'est pas 1-lipschitzienne dans l'ouvert connexe U bien que $\sup_{(x,y) x \in U} \|Df(x, y)\| < 1$.

Définition 1.7.2 *Une application $f : U \longrightarrow F$ est dite localement constante si pour tout $a \in U$, il existe $r > 0$ tel que $B(a, r) \subset U$ et $f(x) = f(a)$, pour tout $x \in B(a, r)$.*

Exercice 1.7.3 .

Montrer que si U est un ouvert connexe, alors les conditions suivantes sont équivalentes :

1. *f est localement constante*
2. *f est constante*

Indication : Soit $a \in U$ fixé, montrer que $A := \{x \in U; f(x) = f(a)\}$ est un ouvert, fermé non vide de U.

Remarque : La connexité de U est essentielle. En effet $f :]0,1[\cup]2,3[\longrightarrow \mathbb{R}$ définie par $f(x) = 0$ si $x \in]0,1[$ et $f(x) = 1$ si $x \in]2,3[$, est localement constante mais n'est pas constante.

Corollaire 1.53 *Soient E et F deux evn, U ouvert connexe de E.*

Soit $f : U \longrightarrow F$ une application différentiable telle que $Df(x) = 0$ pour tout $x \in U$. Alors f est constante .

Démonstration 1.54 *Soit $a \in U$. Il existe $r > 0$ tel que $B(a,r) \subset U$. d'après le corollaire précédent, appliqué au convexe $B(a,r)$, on a $\|f(x) - f(a)\| = 0$ pour tout $x \in B(a,r)$. Donc la restriction de f à toute boule contenue dans U est constante autrement dit f est localement constante dans U.*

Comme U est un ouvert connexe, f y est alors constante, d'après l'exercice précédent.

1.8 Un critère de Différentiabilité (condition suffisante)

Théorème 1.55 *Soit $f : U \longrightarrow \mathbb{R}^m$, $x \mapsto (f_1(x),\dots,f_m(x))$, U ouvert de $\mathbb{R}^m$ et $a \in U$.*

Si les dérivées partielles de f existent et sont continues en a, sauf eventuellement une qui n'existe qu'en a. Alors, f est différentiable en a.

Démonstration 1.56 *On va supposer, pour éviter des notations encombrantes, (et ceci sans perdre de généralité) que $n = 2$ et $m = 1$.*

On suppose $\mathbb{R}^2$, munit de la norme $\|(x_1,x_2).\|_\infty = \max\{|x_1|,|x_2|\}$.

On a par hypothèse : $\frac{\partial f}{\partial x_1}$ est une fonction continue en $a = (a_1,a_2)$ et $\frac{\partial f}{\partial x_2}(a)$ existe.

On doit montrer :

$$\lim_{(h_1,h_2)\longrightarrow 0} \frac{|f(a_1+h_1,a_2+h_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1 - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2|}{\|(h_1,h_2)\|} = 0$$

On écrit

$f(a_1+h_1,a_2+h_2)-f(a_1,a_2) = (f(a_1+h_1,a_2+h_2) - f(a_1,a_2+h_2))+(f(a_1,a_2+h_2) - f(a_1,a_2))$

En utilisant l'inégalité triangulaire on obtient

$$\frac{|f(a_1+h_1,a_2+h_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1 - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2|}{\|(h_1,h_2)\|}$$

$$\leq \frac{|f(a_1+h_1,a_2+h_2) - f(a_1,a_2+h_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1|}{|h_1|} + \frac{|f(a_1,a_2+h_2) - f(a_1,a_2+h_2) - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2|}{|h_2|}$$

Alors :

1. *d'après le théorème des accroissements finis, il existe $c \in]a_1, a_1+h_1[$ tel que $f(a_1+h_1,a_2+h_2)-f(a_1,a_2+h_2) = \frac{\partial f}{\partial x_1}(c,a_2).h_1$ et $\lim_{(h_1,h_2)\longrightarrow 0} \frac{\partial f}{\partial x_1}(c,a_2) = \frac{\partial f}{\partial x_1}(a_1,a_2)$ par continuité de $\frac{\partial f}{\partial x_1}$ en a, d'où* $\lim\limits_{(h_1,h_2)\longrightarrow 0} \frac{|f(a_1+h_1,a_2+h_2) - f(a_1,a_2+h_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1|}{|h_1|} = \lim\limits_{(h_1,h_2)\longrightarrow 0} \left|\frac{\partial f}{\partial x_1}(c,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2)\right| = 0.$
2. *$\frac{\partial f}{\partial x_2}(a)$ existe est équivalent à $\lim_{h_2\longrightarrow 0} \frac{|f(a_1,a_2+h_2)-f(a_1,a_2+h_2)-\frac{\partial f}{\partial x_2}(a_1,a_2).h_2|}{|h_2|} = 0$*

Ainsi, $\lim_{(h_1,h_2)\longrightarrow 0} \frac{|f(a_1+h_1,a_2+h_2)-f(a_1,a_2)-\frac{\partial f}{\partial x_1}(a_1,a_2).h_1-\frac{\partial f}{\partial x_2}(a_1,a_2).h_2|}{\|(h_1,h_2)\|} = 0.$♣

Corollaire 1.57 *Soit* $f : U \longrightarrow \mathbb{R}^m$, $x \mapsto (f_1(x),\dots,f_m(x))$, U *ouvert de* $\mathbb{R}^m$ *et* $a \in U$.

Si les dérivées partielles de f *existent et sont continues en* a *alors,* f *est différentiable en* a.

Exemple 1.8.1 *Soit* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *définie par :*

$$f(x,y) = \begin{cases} \dfrac{x^4}{(x^2+y^2)} + y & \textit{si } (x,y) \neq (0,0) \\ 0 & \textit{si } (x,y) = (0,0) \end{cases}$$

on veut montrer que f *est différentiable en* $(0,0)$.

On commence par vérifier que les dérivées partielles existent en ce point.

1. $\lim_{x\longrightarrow 0} \frac{f(x,0)-f(0,0)}{x} = \lim_{x\longrightarrow 0} \frac{x^2}{x} = 0$ *d'où* $\frac{\partial f}{\partial x}(0,0) = 0$.
2. $\lim_{y\longrightarrow 0} \frac{f(0,y)-f(0,0)}{y} = \lim_{y\longrightarrow 0} \frac{y}{y} = 1$ *d'où* $\frac{\partial f}{\partial y}(0,0) = 1$.

Maintenant, il suffit de montrer que l'une des dérivées partielles est continue en $(0,0)$.

On a pour $(x,y) \neq (0,0)$, $\frac{\partial f}{\partial y}(x,y) = \frac{\partial}{\partial y} f(x,y) = \frac{-2yx^4}{(x^2+y^2)^2} + 1$. *On doit montrer que* $\lim_{(x,y)\longrightarrow(0,0)} \frac{\partial f}{\partial y}(x,y) = \frac{\partial f}{\partial y}(0,0) = 1$.

On a $\lim_{(x,y)\longrightarrow(0,0)} \left|\frac{-2yx^4}{(x^2+y^2)^2}\right| \leq \lim_{(x,y)\longrightarrow(0,0)} \left|2y\frac{x^4}{(x^2+y^2)^2}\right| \leq 2|y| = 0$.

Ainsi $\lim_{(x,y)\longrightarrow(0,0)} \frac{\partial f}{\partial y}(x,y) = \lim_{(x,y)\longrightarrow(0,0)} \frac{-2yx^4}{(x^2+y^2)^2} + 1 = 1 = \frac{\partial f}{\partial y}(0,0)$. ♣

Remarque : La différentiabilité n'entraîne pas n'entraîne la continuité d'aucune dérivée partielle. Un exemple est donné par l'exercice suivant

Exercice 1.8.2 *La fonction* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *définie par :*

$$f(x,y) = \begin{cases} (x^2+y^2)\sin\left(\dfrac{1}{x^2+y^2}\right) & \textit{si } (x,y) \neq (0,0) \\ 0 & \textit{si } (x,y) = (0,0) \end{cases}$$

est différentiable en $(0,0)$*, mais* $\frac{\partial f}{\partial x}$ *et* $\frac{\partial f}{\partial y}$ *ne sont pas continues en* $(0,0)$.

Exercice 1.8.3 *La fonction* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *définie par :*

$$f(x,y) = \begin{cases} yx^2\sin\left(\dfrac{1}{x^2}\right) & \textit{si } x \neq 0 \\ 0 & \textit{si } x = 0 \end{cases}$$

Soit $y_0 \neq 0$.

Montrer que $\frac{\partial f}{\partial x}$ *n'est pas continue en* $(0,y_0)$ *mais que* $\frac{\partial f}{\partial y}$ *est continue en* $(0,y_0)$.

Soit U un ouvert de $\mathbb{R}^n$ et $f : U \longrightarrow \mathbb{R}^m$ une application. On dit que f est de **classe** C^1 **(** ou **continûment différentiable) sur** U si elle est différentiable en tout point de U et si l'application :$Df : U \longrightarrow \mathcal{L}(\mathbb{R}^n;\mathbb{R}^m)$, $\quad a \mapsto Df(a)$ est continue.

Le résultat suivant donne une condition nécessaire et suffisante pour qu'une application soit de classe C^1.

Théorème 1.58 *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \longrightarrow \mathbb{R}^m$ une application.*

Alors f est de classe C^1 si et seulement si ses dérivées partielles sont continues i.e. pour tout $j \in \{1, \ldots, n\}$, $\frac{\partial f}{\partial x_j}$ existe et est continue.

Démonstration 1.59 *“ $\Longrightarrow''$ Soit $a \in U$. Si f est de classe C^1, alors $\lim_{x \longrightarrow a} Df(x) - Df(a) = 0$, ainsi $\lim_{x \longrightarrow a} \frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(a) = Df(x)e_j - Df(a)e_j = (Df(x) - Df(a))e_j = 0.e_j = 0$, d'où la continuiteé de $\frac{\partial f}{\partial x_j}$ en a.*

“ $\Longleftarrow''$ Si toutes les dérivées partielles sont continues en a alors, d'après le corollaire 1.57, f est différentiable et $Df(x).h = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)h_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)\pi_j(h)$, où $\pi_j : \mathbb{R}^n \longrightarrow \mathbb{R}$ est la projection $\pi_j(h) = h_j$.

Ainsi $Df = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \circ \pi_j$ est une somme finie d'applications continues, donc est continue, c-à-d que f est C^1.♣

Complément : Soient $E = E_1 \times \ldots \times E_n$ et F des espaces vectoriels normés , $U \subset E$ un ouvert et $f : U \longrightarrow F$ une application.

On a vu précédemment que si f est différentiable en tout point $a \in U$, alors $Df(a) = (D_1f(a), \ldots, D_nf(a))$ et chaque composante $D_if(a) \in \mathcal{L}(E, F)$, mais que la réciproque est en générale fausse.

Le résultat suivant donne le lien entre continuité des dérivées partielles et continuité de la différentielle.

Théorème 1.60 *Soient $E = E_1 \times \ldots \times E_n$ et F des espaces vectoriels normés , $U \subset E$ un ouvert et $f : U \longrightarrow F$ une application.*

Alors f est de classe C^1 si et seulement si et $j \in \{1, \ldots, n\}$, D_jf existe et est continue.

Démonstration 1.61 *” $\Longrightarrow$ ” Si Df est continue il en est de même de ses composantes D_if, donc les dérivées partielles sont continues.*

” $\Longleftarrow$ ” Supposons que pour tout $j \in \{1, \ldots, n\}$, D_jf existe et est continue.

Pour alléger les notatiosn on prendra $n = 2$, la même technique marche pour $n \geq 3$.

Soit $a = (a_1, a_2) \in U$ et $(h_1, h_2) \in E = E_1 \times E_2$, on écrit

$$\begin{aligned} f(a+h) - f(a) &= f(a_1+h_1, a_2+h_2) - f(a_1, a_2) \\ &= f(a_1+h_1, a_2+h_2) - f(a_1+h_1, a_2) + f(a_1+h_1, a_2) - f(a_1, a_2) \end{aligned}$$

Soit $\varepsilon > 0$. Par hypothèse, il existe $\delta > 0$ tel que pour $\|h_1\| < \delta$ et $\|h_2\| < \delta$ on a

$$\|D_1f(a_1+h_1, a_2+h_2) - D_1f(a_1, a_2)\| < \varepsilon \quad \text{et} \quad \|D_2f(a_1+h_1, a_2+h_2) - D_2f(a_1, a_2)\| < \varepsilon.$$

D'après le corollaire 1.50, on a

$$\begin{aligned} &\|f(a_1+h_1, a_2+h_2) - f(a_1+h_1, a_2) - D_2f(a_1, a_2)h_2\| \\ &\leq \|h_2\| \sup_{s \in]a_2, a_2+h_2[} \|D_2f(a_1+h_1, s) - D_2f(a_1, a_2)\|. \end{aligned}$$

et

$$\begin{aligned} &\|f(a_1+h_1, a_2) - f(a_1, a_2) - D_1f(a_1, a_2)h_1\| \\ &\leq \|h_1\| \sup_{t \in]a_1, a_1+h_1[} \|D_1f(t, a_2) - D_1f(a_1, a_2)\|. \end{aligned}$$

D'où

$$\|f(a+h) - f(a) - (D_1f(a), D_2f(a)) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \| \leq \varepsilon(\|h_1\| + \|h_1\|)$$

Ainsi f est différentiable et de différentielle $Df(a) = (D_1f(a), D_2f(a))$, comme D_1f et D_2f sont continues, il en est de même de Df.♣

2 Différentielles d'ordre supérieur et formule de Taylor

Partant d'une application $f : U \longrightarrow F$ de classe C^1, on considère l'application continue

$$Df : U \longrightarrow \mathcal{L}(E, F)$$
$$x \mapsto Df(x)$$

Définition 2.0.4 *1. On dit que f est de classe C^2, si l'application Df est de classe C^1. La différentielle seconde au point $a \in U$, notée $D^2 f(a)$ est un élément de $\mathcal{L}(E, \mathcal{L}(E, F))$.*

2. Par induction sur l'entier naturel r, on définit une application f est de classe C^r si sa différentielle Df est de classe C^{r-1}. On note $D^r f(a)$ la différentielle d'ordre r de f en un point $a \in U$.

3. On dit que f est de classe C^∞ si f est de classe C^r pour tout $r \in \mathbb{N}$.

Exemple 2.0.5 *(a) Si $f \in \mathcal{L}(E, F)$, on a vu que $Df = f$ est constante, d'où $D^2 f = 0$, ceci entraîne que pour tout $r \geq 2$, $D^r f = 0$. Donc f est de classe C^∞.*

(b) Si $B : E_1 \times E_2 \longrightarrow F$ est bilinéaire continue, alors en tout point $(a, b) \in E_1 \times E_2$, $DB(a, b)$ est une application linéaire, par suite $D^r B = D^{r-1} DB = 0$ si $r \geq 3$. Donc B est de classe C^∞.

(c) Plus généralement, pour tout entier $n \geq 1$, toute application n-ilinéaire continue est de classe C^∞ et ses différentielles d'ordre $\geq n + 1$ sont nulles.

2.1 Différentielles secondes

On va commencer par voir la différentielle seconde comme une application bilinéaire.

Soit $f : U \longrightarrow F$ de classe C^1.

La différentielle seconde en un point $a \in U$, $D^2 f(a) \in \mathcal{L}(E, \mathcal{L}(E, F))$ c-à-d que pour tout $h \in E$, $D^2 f(a)h \in \mathcal{L}(E, F)$ ou encore que pour tout $h, k \in E$, $\big(D^2 f(a)h\big)\, k \in F$

On voit alors que l'application $(h, k) \mapsto \big(D^2 f(a)h\big)\, k$ est bilinéaire.

On va donc, en utilisant cette manipulation, identifier $\mathcal{L}(E, \mathcal{L}(E, F))$ avec l'espace des applications bilinéaires de $E \times E \longrightarrow F$, noté $\mathcal{L}_2(E, F)$.

Théorème 2.1 *L'application $\Phi : \mathcal{L}(E, \mathcal{L}(E, F)) \longrightarrow \mathcal{L}_2(E, F)$, $\varphi \mapsto \Phi(\varphi) = B_\varphi$ telle que pour tout $(h, k) \in E \times E$, $B_\varphi(h, k) = \varphi(h)k$, est un isomorphisme isométrique.*

Démonstration 2.2 *L'application Φ est évidemment linéaire, on va montrons d'abord qu'elle est surjective. Soit $B \in \mathcal{L}_2(E, F)$, On définit $\varphi_B : E \longrightarrow \mathcal{L}(E, F)$, par $\varphi_B(h) = B(h, .)$, alors on a bien $\Phi(\varphi_B) = B$. Pour l'injectivité, il suffit de remarquer que si $B_\varphi = 0$, alors pour tout $(h, k) \in E \times E$, $\varphi(h)k = B_\varphi(h, k) = 0$, donc pour tout $h \in E$, $\varphi(h) = 0$, par suite $\varphi = 0$. Donc Φ est un isomorphisme linéaire.*

De plus on a,

$$\|\varphi_B\| = \sup_{h \neq 0} \frac{\|\varphi_B(h)\|}{\|h\|} = \sup_{h \neq 0} \frac{\|B(h, .)\|}{\|h\|} = \sup_{h \neq 0} \sup_{k \neq 0} \frac{\|B(h, k)\|}{\|h\|.\|k\|} = \|B\|$$

Donc, l'application Φ est une isométrie.

Maintenant, la différentielle seconde $D^2f(a)$ vue comme une application bilinéaire, est l'application

$$\begin{gathered} E \times E \longrightarrow F \\ (h,k) \mapsto D^2f(a)(h,k) = D_h\left(D_k f\right)(a) \end{gathered}$$

où $D_h f$ est la dérivée directionnelle dans la direction h.

Remarque : [et notations] Si $E = \mathbb{R}^n$ et $(e_1, \dots, e_n)$ sa base canonique, alors

$$D^2f(a)(e_j, e_k) = D_{e_j}\left(D_{e_k} f\right)(a) = D_j\left(D_k f\right)(a) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right)(a) = \frac{\partial^2 f}{\partial x_j \partial x_k}(a).$$

Définition 2.1.1 *Une application bilinéaire $u \in \mathcal{L}_2(E,F)$ est dite symétrique si pour tout $(h,k) \in E \times E$ on a :*

$$u(h,k) = u(k,h).$$

Le résultat suivant exprime le symétrie de la différentielle seconde.

Théorème 2.3 (Théorème de symétrie de Schwarz) *Soient $f : U \longrightarrow F$ une application différentiable. Si f est deux fois différentiable en $a \in U$, $D^2f(a)$ est une application bilinéaire symétrique c-à-d pour tout $(h,k) \in E \times E$*

$$D^2f(a)(h,k) = D^2f(a)(k,h).$$

Démonstration 2.4 *Soit $(h,k) \in E \times E$ fixé.*

Puisque U est un ouvert, il existe $r > 0$ tel que pour tout $\lambda, \mu \in \mathbb{R}$, $|\lambda| \leq r$ et $|\mu| \leq r$ on ait : $a + \lambda h + \mu k \in U$.

Posons pour tout $t \in [-r,r]$,

$$A(t,h,k) = f(a + t(h+k)) - f(a+th) - f(a+tk) + f(a)$$

et $g(v) = f(a + t(v+k)) - f(a+tv) - f(a+tk)$

Alors $A = g(h) - g(0)$

Supposons que l'on ait pu montrer que $\lim_{t \longrightarrow 0} \dfrac{A}{t^2} = D^2f(a)(h,k)$, comme $A(t,h,k) = A(t,k,h)$, on aurait prouvé la symétrie de $D^2f(a)$. On doit alors montrer que $\lim_{t \longrightarrow 0} \dfrac{A}{t^2} = D^2f(a)(h,k)$.

On a d'après 1.50

$$\begin{aligned} \|A(t,h,k) - t^2D^2f(a)(h,k)\| &= \|g(h) - g(0) - t^2D^2f(a)(h,k)\| \\ &\leq \|h\| \left(\sup_{x \in]0,h[} \|Dg(x) - t^2D^2f(a)(.,k)\| \right). \end{aligned} \tag{2.1}$$

car l'application $h \mapsto t^2D^2f(a)(h,k)$ est linéaire.En différentiant g, on trouve $Dg(x) = tDf(a + t(x+k)) - tDf(a+tx)$ d'où

$$\|A(t,h,k) - t^2D^2f(a)(h,k)\| \leq \|h\| \left(\sup_{x \in]0,h[} \|tDf(a + t(x+k)) - tDf(a+tx) - t^2D^2f(a)(.,k)\| \right). \tag{2.2}$$

D'autre part, f deux fois différentiable en a, entraîne que $Df(a+t(x+k)) = Df(a)+tD^2f(a)(.,x+k)+\|t(x+k)\|\varepsilon_1(t(x+k))$ et $Df(a+tx) = Df(a)+tD^2f(a)(.,x)+\|tx\|\varepsilon_2(tx)$ avec la limite quand t tends vers 0 de ε_1 et ε_2 égale à 0. 2.2 devient

$$\begin{aligned}\|A(t,h,k) - t^2D^2f(a)(h,k)\| &\leq \|h\|(\sup_{x\in]0,h[} \|t\left[Df(a) + tD^2f(a)(.,x+k) + \|t(x+k)\|\varepsilon_1(t(x+k))\right] \\ &\quad - t\left[Df(a) + tD^2f(a)(.,x) + \|tx\|\varepsilon_2(tx)\right] - t^2D^2f(a)(.,k)\|) \\ &\leq t^2\|h\| \left(\sup_{x\in]0,h[} \|\|(x+k)\|\varepsilon_1(t(x+k)) - \|x\|\varepsilon_2(tx)\|\right) \end{aligned} \tag{2.3}$$

Finalement,

$$\|\frac{A(t,h,k)}{t^2} - D^2f(a)(h,k)\| \leq \|h\| \left(\sup_{x\in]0,h[} \|(x+k)\|\|\varepsilon_1(t(x+k))\| + \|x\|\|\varepsilon_2(tx)\| \right)$$

d'où $\lim_{t\longrightarrow 0} \frac{A}{t^2} = D^2f(a)(h,k)$.♣

Cas particulier $E = \mathbb{R}^n$ et $F = \mathbb{R}$.

Soit $f : U \longrightarrow F$ deux fois différentiable en $a \in U$, alors $D^2f(a)$ est une forme bilinéaire symétrique.

Comme une forme bilinéaire symétrique B sur $\mathbb{R}^n$ est représentée, dans la base canonique de $\mathbb{R}^n$, par une matrice symétrique $A \in \mathcal{M}_n(\mathbb{R})$ c-à-d

$$B(x,y) =< Ax, y >=< x, Ay >= (x_1,\ldots,x_n)A \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} a_{ii}x_iy_i + 2\sum_{1\leq i<j\leq n} a_{ij}x_iy_j,$$

pour tout $x, y \in \mathbb{R}^n$.

Définition 2.1.2 (matrice hessienne) *La matrice symétrique qui représente $D^2f(a)$, dans la base canonique de $\mathbb{R}^n$, est appelée matrice* **hessienne** *de f au point a et est notée $H_f(a)$.*

Les coefficients de la matrice $H_f(a)$ sont les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_j \partial x_k}(a)$, $1 \leq k, j \leq n$,

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \ldots & \ldots & \ldots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \ldots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

et pour tout $1 \leq k, j \leq n$, $\boxed{\frac{\partial^2 f}{\partial x_j \partial x_k}(a) = \frac{\partial^2 f}{\partial x_k \partial x_j}(a)}$.

Remarque :

1. Si f n'est pas deux fois différentiable, la matrice hessienne n'est pas nécessairement symétrique.

   Par exemple, si $f(x,y) = xy.g(x,y)$, alors $D_1f(0,y) = \lim_{x\longrightarrow 0} \frac{f(x,y) - f(0,y)}{x} = y \lim_{x\longrightarrow 0} g(x,y)$, de même $D_2f(x,0) = x \lim_{y\longrightarrow 0} g(x,y)$, d'où $D_2D_1f(0,0) =$

$\lim\limits_{y \longrightarrow 0} \dfrac{D_1 f(0,y) - D_1 f(0,0)}{y} = \lim\limits_{y \longrightarrow 0} \left(\lim\limits_{x \longrightarrow 0} g(x,y) \right)$ de même $D_1 D_2 f(0,0) = \lim\limits_{x \longrightarrow 0} \left(\lim\limits_{y \longrightarrow 0} g(x,y) \right)$.

Pour obtenir $D_2 D_1 f(0,0) \neq D_1 D_2 f(0,0)$, il suffit de choisir g telle que $\lim\limits_{y \longrightarrow 0} \left(\lim\limits_{x \longrightarrow 0} g(x,y) \right) \neq \lim\limits_{x \longrightarrow 0} \left(\lim\limits_{y \longrightarrow 0} g(x,y) \right)$. Par exemple $g(x,y) = \frac{x^2-y^2}{x^2+y^2}$ si $(x,y) \neq (0,0)$ et $g(0,0) = 1$ convient.

2. D'autre part, la symétrie de la hessienne n'entraîne pas que f est deux fois différentiable. En effet, la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x,y) = x^3 \sin(\frac{1}{x})$ si $(x,y) \neq (0,0)$ et $f(0,0) = 0$
   satisfait $D_2 D_1 f(0,0) = D_1 D_2 f(0,0) = 0$, mais n'est pas deux fois différentiable. (à vérifier !)

2.2 Généralisation à l'ordre $r \geq 2$

On va commencer par voir la différentielle d'ordre r comme une application r-linéaire.

Soit $f : U \longrightarrow F$ de classe C^{r-1}.

La différentielle d'ordre r en un point $a \in U$, $D^r f(a) \in \mathcal{L}\,(\underbrace{E, \ldots, (\mathcal{L}(E}_{r}, \mathcal{L}(E,F))\ldots)$

d'où pour tout $(h_1, \ldots, h_r) \in E$, $(D^r f(a) h_1) h_2) \ldots) h_r \in F$.

On voit alors que l'application $(h_1, \ldots, h_r) \mapsto (D^r f(a) h_1) h_2) \ldots) h_r$ est r-linéaire.

On va donc, en utilisant une généralisation de 4.3, identifier $\mathcal{L}\,(\underbrace{E, \ldots, (\mathcal{L}(E}_{r}, \mathcal{L}(E,F))\ldots)$ avec l'espace des applications r-linéaires de $E^r \longrightarrow F$, noté $\mathcal{L}_r(E,F)$.

Théorème 2.5 *L'application* $\Phi : \mathcal{L}\,(\underbrace{E, \ldots, (\mathcal{L}(E}_{r}, \mathcal{L}(E,F))\ldots) \longrightarrow \mathcal{L}_r(E,F)$, $\phi \mapsto \Phi(\phi) = B_\phi$ *telle que pour tout* $(h_1, \ldots, h_r) \in E \times \ldots \times E$, $B_\phi(h_1, \ldots, h_r) = \phi(h_1) \ldots) h_r$, *est un isomorphisme isométrique.*

Définition 2.2.1 *Une application r-linéaire* $u \in \mathcal{L}_r(E,F)$ *est dite symétrique si pour tout* $(h_1, \ldots, h_r) \in E^r$ *et toute bijection* $\sigma : \{1, \ldots, r\} \longrightarrow \{1, \ldots, r\}$ *on a*

$$u(h_{\sigma(1)} \ldots, h_{\sigma(r)}) = u(h_1, \ldots, h_r)$$

Le résultat suivant exprime le symétrie de la différentielle d'ordre r.

Théorème 2.6 (Théorème de symétrie de Schwarz) *Soient* $f : U \longrightarrow F$ *une application r différentiable . Alors* $D^r f(a)$ *est une application* r*-linéaire symétrique i.e.pour tout* $(h_1, \ldots, h_r) \in E^r$

$$D^r f(a)(h_{\sigma(1)} \ldots, h_{\sigma(r)}) = D^r f(a)(h_1, \ldots, h_r).$$

Exemple 2.2.2 *Si* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, $(x,y) \mapsto f(x,y)$ *est* 3 *fois différentiable en un point* $a \in \mathbb{R}^2$, *alors :*
$\frac{\partial^3 f}{\partial x^2 \partial y}(a) = \frac{\partial^3 f}{\partial x \partial y \partial x}(a) = \frac{\partial^3 f}{\partial y \partial x^2}(a)$ *et* $\frac{\partial^3 f}{\partial x \partial y^2}(a) = \frac{\partial^3 f}{\partial y \partial x \partial y}(a) = \frac{\partial^3 f}{\partial y^2 \partial x}(a)$

Exemple 2.2.3 *Soit $f : E \longrightarrow F$ une application telle que pour tout $x \in E$ $f(x) = g(x,\dots,x)$ où g est une application k-linéaire symétrique.*

Alors, $D^r f(x) = \begin{cases} \frac{k!}{(k-r)!} g(\underbrace{x,\dots,x}_{k-r},\underbrace{\dots}_{r}) & \text{si } 0 \le r \le k \\ 0 & \text{si } r > k \end{cases}$.

On sait déja voir 2.0.5, que f est de classe C^∞ et que $D^r f = 0$ si $r > k$.
On va démontrer par induction sur r la formule pour $1 \le r \le k$.
Si $r = 0$, on a $D^r f(a) = f(a) = g(a,\dots,a)$.
Supposons que $D^r f$ est l'application r-linéaire, telle que $D^r f(x) = \frac{k!}{(k-r)!} g(\underbrace{x,\dots,x}_{k-r},\underbrace{\dots}_{r})$.

Soit $h \in E$, alors $D^{r+1} f(x) \underbrace{(h,\dots,h)}_{r+1} = \lim_{t\longrightarrow 0} \frac{D^r f(x+th) - D^r f(x)}{t}$

$= \lim_{t\longrightarrow 0} \frac{k!}{(k-r)!} \frac{g(\underbrace{x+th,\dots,x+th}_{k-r},\underbrace{\dots}_{r}) - g(\underbrace{x,\dots,x}_{k-r},\underbrace{\dots}_{r})}{t}$

Par multilinéarité de g on a

$D^r f(x+th) - D^r f(x) = \frac{k!}{(k-r)!} \left(g(\underbrace{x+th,\dots,x+th}_{k-r},\underbrace{\dots}_{r}) - g(\underbrace{x,\dots,x}_{k-r},\underbrace{\dots}_{r}) \right)$

$= \frac{k!}{(k-r)!} t \left(\sum_{1\le i\le k-r} g(x,\dots,x,h,x\dots,x,\dots) \right) + O(t^2))$

D'où $D^{r+1} f(x) h = \frac{k!}{(k-r)!} (\sum_{1\le i\le k-r} g(x,\dots,x,\underbrace{h}_{i^e\ place},x\dots,x,\dots))$.

Par symétrie de g on a $g(x,\dots,x,h,x,\dots,x,\dots) = g(\underbrace{x,\dots,x}_{k-r-1},h,\underbrace{\dots}_{r})$,

par suite $\displaystyle\sum_{1\le i\le k-r} g(x,\dots,x,h,x\dots,x,\dots) = (k-r) g(\underbrace{x,\dots,x}_{k-r-1},h,\underbrace{\dots}_{r})$.

Finalement, $D^{r+1} f(a) h = \frac{k!}{(k-r)!} (k-r) g(x,\dots,x,h,\dots) = \frac{k!}{(k-r-1)!} g(x,\dots,x,h,\dots)$ *Donc* $D^{r+1} f(a) = \frac{k!}{(k-r-1)!} g(\underbrace{x,\dots,x}_{k-r-1},\underbrace{\dots}_{r+1})$.

2.3 Formules de Taylor

Théorème 2.7 (Formule de Taylor avec reste intégral) *Soient U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}$ une application de classe C^{p+1}, $a \in U$ et $h \in \mathbb{R}^n$ tels que le segment $[a, a+h] \subset U$. Alors*

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{1}{k!} D^k f(a)(h,\dots,h) + \frac{1}{p!} \int_0^1 (1-t)^p D^{p+1} f(a+th)(h,\dots,h) dt$$

Démonstration 2.8 *On applique le lemme suivant à la fonction $g : [0,1] \longrightarrow \mathbb{R}$, $t \mapsto g(t) = f(a+th)$.*

Lemme 2.3.1 *Soit I un intervalle ouvert de $\mathbb{R}$, contenant $[0,1]$ et $g : I \longrightarrow \mathbb{R}$ une fonction de classe C^{p+1}. Alors :*

$$g(1) = g(0) + \sum_{k=1}^{p} \frac{1}{k!} g^{(k)}(0) + \frac{1}{p!} \int_0^1 (1-t)^p g^{(p+1)}(t) dt.$$

Démonstration 2.9 (du lemme) *On utilise une recurrence sur l'entier p. Le lemme est vrai pour $p = 0$, par le "théorème fondamentale de l'analyse" à savoir $g(1) - g(0) = \int_0^1 g'(t) dt$. Supposons le lemme vrai à l'ordre $p-1$, donc $g(1) = \sum_{k=0}^{p-1} \frac{1}{k!} g^{(k)}(0) + \frac{1}{(p-1)!} \int_0^1 (1-t)^{p-1} g^{(p)}(t) dt$. En utilisant $\left(\frac{(1-t)^p}{p!}\right)' = -\frac{(1-t)^{p-1}}{(p-1)!}$ et une intégration par parties on obtient*

$$\frac{1}{(p-1)!}\int_0^1 (1-t)^{p-1} g^{(p)}(t)dt = \frac{1}{p!}g^{(p)}(0) + \frac{1}{k!}\int_0^1 (1-t)^p g^{(p+1)}(t)dt.$$

D'où

$$\begin{aligned} g(1) &= \sum_{k=0}^{p-1} \frac{1}{k!}g^{(k)}(0) + \frac{1}{(p-1)!}\int_0^1 (1-t)^{p-1} g^{(p)}(t)dt \\ &= \sum_{r=0}^{p-1} \frac{1}{k!}g^{(k)}(0) + \frac{1}{k!}g^{(p)}(0) + \frac{1}{p!}\int_0^1 (1-t)^p g^{(p+1)}(t)dt.\clubsuit \end{aligned}$$

Théorème 2.10 (Formule de Taylor-Lagrange) .

Soient U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}$ une application $p+1$ fois différentiable, $a \in U$ et $h \in \mathbb{R}^n$ tels que le segment $[a, a+h] \subset U$. Alors il existe $\theta \in]0,1[$ tel que :

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{1}{r!}D^k f(a)(h, \ldots, h) + \frac{1}{(p+1)!}D^{p+1}f(a+\theta h)$$

Démonstration 2.11 *On applique le lemme suivant à la fonction $g : [0,1] \longrightarrow \mathbb{R}$, $t \mapsto g(t) = f(a+th)$.*

Lemme 2.3.2 *Soit I un intervalle ouvert de $\mathbb{R}$, contenant $[0,1]$ et $g : I \longrightarrow \mathbb{R}$ une fonction $p+1$ dérivable. Alors il existe $\theta \in]0,1[$ tel que*

$$g(1) = g(0) + \sum_{k=1}^{p} \frac{1}{k!}g^{(k)}(0) + \frac{1}{(p+1)!}g^{(p+1)}(\theta).$$

Remarque : Ce théorème n'est pas vrai pour les fonctions vectorielles c-àd à valeurs dans $\mathbb{R}^m$, $m \geq 2$ (voir la remarque 1.6)

Complément :

Définition 2.3.3 (polynôme généralisé) *Soient E et F deux espaces vectoriels normés et $k \in \mathbb{N}$.*

Une application $P : E \longrightarrow F$ est dite polynomiale de degré k, si elle s'écrit : $P(x) = \sum_{0\leq r\leq k} a_r(x, \ldots, x)$, où pour tout $r \in \{0, \ldots, k\}$, a_r est une une application r-linéaire symétrique de $E^r \longrightarrow F$ et $a_k \not\equiv 0$.

Théorème 2.12 *Soient E, F deux espaces vectoriels normés et P de E dans F une application polynomiale de degré n.*

Alors, $P(x) = \sum_{k=0}^{n} \frac{1}{k!} D^k P(0)(x, \ldots, x)$.

Démonstration 2.13 *Comme $P(x) = \sum_{k=0}^{n} a_k(x, \ldots, x)$ et que d'après 2.0.5 pour tout $k \in \{0, \ldots, n\}$, $D^r a_k(x, \ldots, x) =$*

$$\begin{cases} \frac{k!}{(k-r)!} a_k(\underbrace{x, \ldots, x}_{k-r}, \underbrace{\ldots}_{r}) & \text{si } 0 \leq r \leq k \\ 0 & \text{si } r > k \end{cases}$$

on a alors $D^k P(0) = k! a_k$. D'où $P(x) = \sum_{k=0}^{n} a_k(x, \ldots, x) = \sum_{k=0}^{n} \frac{1}{k!} D^k P(0)(x, \ldots, x)$.$\clubsuit$

Corollaire 2.14 *Soient $P = \sum_{r=0}^{n} a_r$ et $Q = \sum_{r=0}^{n} b_r$ deux applications polynomiales de E dans F de degré n. Alors $P = Q$ si et seulement si pour tout $r \in \{0, \ldots, n\}$, $a_r = b_r$.*

Démonstration 2.15 $P = Q \Leftrightarrow P - Q = 0 \Leftrightarrow$ *pour tout* $r \in \{0, \ldots, n\}$, $D^r(P-Q)(0) = 0 \Leftrightarrow$ *pour tout* $r \in \{0, \ldots, n\}$, $a_r = b_r$. $\clubsuit$

Théorème 2.16 (Formule de Taylor-Young) *Soient E, F deux espaces vectoriels normés, U un ouvert de E, f une application de U dans F et a, b deux points de U tels que le segment $[a, b] \subset U$.*

Si f est p fois différentiables dans U, alors elle admet un développement de Taylor-Young à l'ordre p en tout point $a \in U$ c-à-d : pour tout $a \in U$, il existe une fonction $\varepsilon : E \longrightarrow F$, avec $\lim_{h \longrightarrow 0} \varepsilon(h) = 0$ tels que :

$$f(a+h) = f(a) + \sum_{k=1}^{p} \frac{1}{k!} D^k f(a)(h, \ldots, h) + \|h\|^p \varepsilon(h)$$

Démonstration 2.17 *Pour simplifier les notation , on suppose que $a = 0$. On va prouver ce théorème par récurrence sur p.*
Notons par $P(h)$ le polynôme $\sum_{k=0}^{p} \frac{1}{k!} D^k f(a)(h, \ldots, h)$.
Pour $p = 0$, le théorème est vrai par définition de la continuité de f en 0. (Pour $p = 1$, le théorème est vrai par définition de la différentiabilité de f en 0.) Supposons, le théorème est vrai jusqu'a l'ordre $p-1$ et montrons le pour l'ordre p. Comme f est p fois différentiable en 0, Df est $p-1$ différentiable en 0, et par hypothèse de récurrence $Df(h) = DP(h) + \|h\|^{p-1} \varepsilon_1(h)$ avec $\lim_{h \longrightarrow 0} \varepsilon_1(h) = 0$. On peut appliquer l'inégalité des accroissements finis à $f - P$ sur $[0, h]$, pour obtenir

$$\|f(h) - P(h)\| \leq \|h\| \left(\sup_{x \in]0,h[} \|Df(x) - DP(x)\| \right) \leq \|h\|^p \|\varepsilon(h)\|. \tag{2.4}$$

où $\|\varepsilon(h)\| \leq \sup_{x \in]0,h[} \|\varepsilon_1(x)\|$, d'où $\lim_{h \longrightarrow 0} \varepsilon(h) = 0$. ♣

Remarque :

1. Ce théorème est une généralisation du développement de Taylor-Lagrange pour les fonctions d'une variable réelle, comme l'inégalité des accroissements finis est une généralisation du théorème des accroissements finis.
2. La réciproque de ce théorème est fausse pour $n \geq 2$.

   En effet, la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ définie par $f(x) = \begin{cases} x^3 \sin(\frac{1}{x}) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$

   admet un développement de Taylor-Young à l'ordre 2 en 0, avec $P \equiv 0$, mais n'y est pas 2 fois différentiable. (à vérifier !)

Exemple 2.3.4 *Soit $f : B((0,0), 1) \longrightarrow \mathbb{R}$ définie par $f(x, y) = \frac{e^{\sin(\sqrt{1+x} - \sqrt{1+y})}}{2+x-y}$.*

On se propose de déterminer le développement de Taylor-Young de f à l'ordre 2 en $(0, 0)$.

$$f(x, y) = f(0, 0) + J_f(0, 0) \begin{pmatrix} x \\ y \end{pmatrix} + \frac{1}{2}(x, y) H_f(0, 0) \begin{pmatrix} x \\ y \end{pmatrix} + (x^2 + y^2)\varepsilon(x, y)$$

avec $\lim_{(x,y) \longrightarrow (0,0)} \varepsilon(x, y) = 0$.

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

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

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

$\frac{\partial^2 f}{\partial x^2}(0,0) = \lim_{x \longrightarrow 0} \frac{\frac{\partial f}{\partial x}(x,0) - \frac{\partial f}{\partial x}(0,0)}{x} = 0$
$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{y \longrightarrow 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y} = -\frac{1}{8}$
$\frac{\partial^2 f}{\partial y^2}(0,0) = \lim_{y \longrightarrow 0} \frac{\frac{\partial f}{\partial y}(0,y) - \frac{\partial f}{\partial y}(0,0)}{y} = \frac{1}{4}$
d'où $f(x,y) = \frac{1}{2} - \frac{1}{8}xy + \frac{1}{8}y^2 + (x^2 + y^2)\varepsilon(x,y)$

Remarque :

1. Soit $f : U \longrightarrow F$ une application r différentiable en $a \in U$. On suppose $[a, a + h] \subset U$ et on définit $g : [0,1] \longrightarrow F$ par $g(t) = f(a + th)$.
   à l'aide de la règle de composition et d'une récurrence sur r on obtient
   Formule : $\boxed{g^{(r)}(t) = \frac{d^r g}{dt^r}(t) = D^r f(a + th)(h, \dots, h), \quad 1 \leq k \leq r.}$
   En effet, si $g^{(r-1)}(t) = D^{r-1} f(a + th) \underbrace{(h, \dots, h)}_{r-1}$ alors,
   $g^{(r)}(t) = (g^{(r-1)})'(t) = (D^{r-1} f(a+th)(h, \dots, h))' = D(D^{r-1} f(a+th)(h, \dots, h))h$
   $= D^r f(a + th) \underbrace{(h, \dots, h)}_{r}.$
   En particulier $g^{(r)}(0) = D^r f(a)(h, \dots, h), \quad 1 \leq k \leq r.$
2. **Formule développée :** $E = \mathbb{R}^n$ et $F = \mathbb{R}$.
   Si $(e_1, \dots, e_n)$ est une base de $\mathbb{R}^n$. Alors $h = \sum_{i=1}^n h_i e_i$ et
   $$\boxed{D^r f(a)(h, \dots, h) = \sum_{1 \leq i_1, \dots, i_r \leq n} h_{i_1} \dots h_{i_r} \frac{\partial^r f}{x_{i_1 \dots x_{i_r}}}(a)}$$
   où la somme est étendue à toutes les suites $(i_1 \dots, i_r)$ de $\{1, \dots, n\}$.
   Cette écriture ne tient pas compte des symétries (Théorème de schwarz).

Ce qui importe est le nombre de fois où l'indice i_s apparaît.

Pour chaque suite $(i_1, \dots, i_r)$, on note α_i le nombre d'indices $i_s = i$, pour tout $1 \leq i \leq n$. On associe alors *le multi-indice* $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ et sa *longueur* $|\alpha| = \sum_{i=1}^n \alpha_i = r$.

Notations : On note $\alpha! := \alpha_1! \dots \alpha_n!$ pour $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$,
$h^\alpha := h_1^{\alpha_1} \dots h_n^{\alpha_n}$ pour $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ et $D^\alpha := D_1^{\alpha_1} \dots D_n^{\alpha_n}$.

Pour un multi-indice donné $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, on note $m(\alpha)$ le nombre de suites $(i_1, \dots, i_r)$ qui lui sont associées. Alors $m(\alpha) = \frac{r!}{\alpha!}$.

En effet, $m(\alpha)$ est déterminé par l'identité

$$\sum_{|\alpha|=r} m(\alpha) h^\alpha = \sum_{1 \leq i_1, \dots, i_r \leq n} h_{i_1} \dots h_{i_r} = (h_1 + \dots + h_n)^r.$$

Alors, $m(\alpha)$ est égal au choix de α_1 éléments parmis r, puis de α_2 parmis $r - \alpha_1$, $\dots$, et enfin de α_n parmis $r - (\alpha_1 + \dots + \alpha_{n-1})$.

D'où $m(\alpha) = \binom{r}{\alpha_1}\binom{r-\alpha_1}{\alpha_2} \cdots \binom{r-(\alpha_1+\dots+\alpha_{n-1})}{\alpha_n} = \frac{r!}{\alpha_1! \dots \alpha_n!} = \frac{r!}{\alpha!}.\clubsuit$

Remarque : Avec cette notation on a : $\frac{1}{r!} D^r f(a)(h, \dots, h) = \sum_{|\alpha|=r} \frac{h^\alpha}{\alpha!} D^\alpha f(a)$

et par suite $P(h) = \sum_{r=0}^{n} \frac{1}{r!} D^r f(a)(h, \dots, h) = \sum_{0 \leq |\alpha| \leq n} \frac{h^\alpha}{\alpha!} D^\alpha f(a)$

Cette dernière formule tient compte des symétries.