Chapitre 2

Différentielles d'ordre supérieur et formule de Taylor

Partant, d'une application $f : U \to F$ de classe C^1, où U est un ouvert de l'evn E, on considère l'application continue

$$\begin{array}{rccc} Df & : & U & \longrightarrow & \mathcal{L}(E,F) \\ & & x & \longmapsto & Df(x) \end{array}$$

On peut se demander si Df à son tour est différentiable, au quel cas la diddérentielle de Df serait une application : $D(Df) : U \longrightarrow \mathcal{L}(E, \mathcal{L}(E,F))$. Dans ce cas, on dit que f est deux fois différentiable et on note $D^2 f$ l'application $D(Df)$, qu'on appelle différentielle d'ordre 2 (ou différentielle seconde) de f sur U.
Si en plus $d^2 f$ est continue, on dit que f est de classe C^2 sur U.

Pour généraliser, on commence par définir par récurrence la suite d'evn :
$\mathcal{L}^1(E,F) = \mathcal{L}(E,F)$ et pour tout entier $r \geq 2$, $\mathcal{L}^r(E,F) = \mathcal{L}(E, \mathcal{L}^{r-1}(E,F))$.

Par récurrence sur l'entier naturel r, on dit qu'une application f est r fois différentible sur U si $D^{r-1}f$ est différentiable sur U. On note $D^r f(a)$ la différentielle d'ordre r de f en un point $a \in U$. dans ce cas $D^r f(a)$ est un élément de $\mathcal{L}^r(E,F)$.

2.0.29 DÉFINITION

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 $\mathcal{L}^2(E,F)$.
2. Par récurrence sur l'entier naturel r, on dit qu'une application f est de classe C^r sur U (r fois continûment différentiable) 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^∞ (ou indéfiniment différentiable) si f est de classe C^r pour tout $r \in \mathbb{N}$.

Notations : Pour tout entier, $r \in \mathbb{N}$, on note par $C^r(U,F)$ l'espace vectoriel des applications de classe C^r de U dans F et par $C^\infty(U,F)$ celui des applications de classe C^∞.

A noter que $C^0(U,F)$ est l'espace des fonctions continues, on a les inclusions

$$C^0(U,F) \supset C^1(U,F) \supset C^2(U,F) \supset C^3(U,F) \supset \ldots \supset C^r(U,F) \supset \ldots$$

et que

$$C^\infty(U,F) = \bigcap_{r\in\mathbb{N}} C^r(U,F).$$

2.0.30 EXEMPLE. (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 $B : E_1 \times E_2 \to F$ est bilinéaire continue, alors en tout point $(a,b) \in E_1 \times E_2$, $DB(a,b)$ est une application linéaire, par suite, $D^2 B$ est constante et $D^r B = D^{r-1} DB = 0$ si $r \geq 3$. Donc B est de classe C^∞.

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

2.0.31 DÉFINITION
On dit que $f : U \to F$ est r fois différentiable en un point $a \in U$, s'il existe un ouvert V tel que $a \in V \subset U$, tel que f soit $r-1$ fois différentiable sur V et $D^{r-1}(f|_V)$ soit différentiable en a.

2.0.2 Différentielles secondes (ou d'ordre 2)

Dérivées partielles secondes

Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}^p$ une application différentiable, alors les fonctions dérivées partielles sont bien définies, $\frac{\partial f}{\partial x_i} : U \to \mathbb{R}^p$ $1 \leq i \leq n$.

On peut alors définir (quand elles existent) leurs dérivées partielles. On les appellent les dérivées partielles secondes et on les notera par :

$D_j D_i f := D_j(D_i f)$ ou encore $\dfrac{\partial^2 f}{\partial x_j \partial x_i} := \dfrac{\partial}{\partial x_i}\left(\dfrac{\partial f}{\partial x_i}\right), 1 \leq i,j \leq n.$

2.0.32 REMARQUE
Sans conditions supplémentaires sur f, il n'y a pas en général d'égalité entre $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ et $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$.

2.0.33 EXEMPLE. Soit $f(x,y) = \dfrac{xy(x^2-y^2)}{x^2+y^2}$ si $(x,y) \neq (0,0)$ et $f(0,0) = 0$. On va montrer que $\dfrac{\partial^2 f}{\partial x \partial y}(0,0) \neq \dfrac{\partial^2 f}{\partial y \partial x}(0,0)$.

On va commencer par déterminer les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$. Pour $(x,y) \neq (0,0)$, on obtient $\frac{\partial f}{\partial x}(x,y) = \dfrac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2}$ et $\frac{\partial f}{\partial y}(x,y) = \dfrac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2}$.
Au point $(0,0)$ en aura $\frac{\partial f}{\partial x}(0,0) = \lim_{x\to 0} \frac{f(x,0)=f(0,0)}{x-0} = \lim_{x\to 0} \frac{0}{x} = 0$ et $\frac{\partial f}{\partial y}(0,0) = \lim_{y\to 0} \frac{f(0,y)=f(0,0)}{y-0} = \lim_{y\to 0} \frac{0}{y} = 0$.

Ainsi, $\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$ et

$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$

On peut passer maintenant au calcul des dérivées partielles secondes en $(0,0)$.

Par définition $\dfrac{\partial^2 f}{\partial x \partial y}(0,0) = \lim\limits_{x\to 0} \dfrac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x-0} = \lim\limits_{x\to 0} \dfrac{x^5}{x^5} = 1$ et $\dfrac{\partial^2 f}{\partial y \partial x}(0,0) = \lim\limits_{y\to 0} \dfrac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y-0} = \lim\limits_{y\to 0} \dfrac{-y^5}{y^5} = -1$.
Ainsi, $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

Le théorème suivant donne une condition suffisante pour que les dérivées secondes mixtes soient égales.

2.0.34 THÉORÈME
Soit U un ouvert de $\mathbb{R}^n$. et $f : U \to \mathbb{R}^p$ une application de classe C^1 telle que les dérivées partielles secondes $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ et $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ soient continues dans un voisinage d'un point $a \in U$. Alors,

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

Démonstration: En utilisant une permutation des indices et le fait que le calcul des dérivées partielles ce fait sur chaque composante de f, on peut supposer sans perdre de généralité que $n = 2$ et $p = 1$.

On est donc dans la situation ou U est un ouvert de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ est de classe C^1 et dont les dérivées partielles secondes sont continues. On doit montrer que $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \dfrac{\partial^2 f}{\partial y \partial x}(y,x)$.

Par définition $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right)(x,y) = \lim\limits_{h\to 0} \dfrac{\frac{\partial f}{\partial y}(x+h,y) - \frac{\partial f}{\partial y}(x,y)}{h} = \lim\limits_{h\to 0} \dfrac{\lim_{k\to 0} \frac{f(x+h,y+k)-f(x+h,y)}{k} - \lim_{k\to 0} \frac{f(x,y+k)-f(x,y)}{k}}{h}$

$$= \lim_{h\to 0}\lim_{k\to 0} \frac{f(x+h,y+k) - f(x+h,y) - f(x,y+k) - f(x,y)}{hk}$$

On trouve de même que

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

Si on pose $A(h,k) = f(x+h,y+k) - f(x+h,y) - f(x,y+k) - f(x,y)$, alors $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \lim\limits_{h\to 0}\lim\limits_{k\to 0} \dfrac{A(h,k)}{hk}$ et $\dfrac{\partial^2 f}{\partial y \partial x}(x,y) = \lim\limits_{k\to 0}\lim\limits_{h\to 0} \dfrac{A(h,k)}{hk}$.

Ainsi, le résultat qu'on doit montrer est un résultat sur l'égalité après permutation des limites.

Posant $\phi(x) = f(x,y+k) - f(x,y)$. alors $A(h,k) = \phi(x+h) - \phi(x)$ et d'après le théorème des accroissement finis (pour les fonctions à valeurs dans $\mathbb{R}$)
il existe $\theta \in]0,1[$ tel que $\phi(x+h) - \phi(x) = \phi'(x+\theta h).h$,

d'où $A(h,k) = \phi'(x+\theta h).h = \left(\frac{\partial f}{\partial x}(x+\theta h, y+k) - \frac{\partial f}{\partial x}(x+\theta h, y)\right) h$

le théorème des accroissements finis donne l'existence d'un $theta' \in]0,1[$ tel que $\frac{\partial f}{\partial x}(x+\theta h, y+k) - \frac{\partial f}{\partial x}(x+\theta h, y) = \frac{\partial f}{\partial y}\left(\frac{\partial f}{\partial x}(x+\theta h, y+\theta' k)\right) k = \frac{\partial^2 f}{\partial y \partial x}(x+\theta h, y+\theta' k)k$
on a ainsi montrer que

$$A(h,k) = \phi(x+h) - \phi(x) = \frac{\partial^2 f}{\partial y \partial x}(x+\theta h, y+\theta' k)kh.$$

En posant maintenant, $\psi(y) = f(x+h,y) - f(x+h,y)$ on aura
$A(h,k) = \psi(y+k) - \psi(y)$.

Un raisonnement analogue au précédent donne, l'existence d'un $theta_1 \in]0,1[$ tel que

$\psi(y+k) - \psi(y) = \psi'(y+\theta_1 k).k = \left(\frac{\partial f}{\partial y}(x+h, y+\theta_1 k) - \frac{\partial f}{\partial y}(x, y+\theta_1 k)\right) k$ puis l'existence d'un $\theta_1' \in]0,1[$ tel que

$\frac{\partial f}{\partial y}(x+h, y+\theta_1 k) - \frac{\partial f}{\partial y}(x, y+\theta_1 k) = \frac{\partial^2 f}{\partial y \partial x}(x\theta_1' h, y+\theta_1 k)h$ et donc

$$A(h,k) = \psi(y+k) - \psi(y) = \frac{\partial^2 f}{\partial x \partial y}(x+\theta_1' h, y+\theta_1 k)kh.$$

On a donc montrer l'existence de $\theta, \theta', \theta_1$ et $\theta_1' \in]0,1[$ tels que

$$\frac{A(h,k)}{hk} = \frac{\partial^2 f}{\partial y \partial x}(x+\theta h, y+\theta' k) = \frac{\partial^2 f}{\partial x \partial y}(x+\theta_1' h, y+\theta_1 k)$$

Comme, θh, $\theta_1' h$, $\theta' k$ et $\theta_1 k$ tendent vers 0 lorsque h et k tendent vers 0,
et l'hypothèse de continuité sur $\dfrac{\partial^2 f}{\partial x \partial y}$ et $\dfrac{\partial^2 f}{\partial y \partial x}(y,x)$
donne l'égalité des limites i.e. $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \dfrac{\partial^2 f}{\partial y \partial x}(y,x)$. ∎

Différentielle d'ordre 2

On va commencer par voir la différentielle seconde comme une application bilinéaire. Soit $f : U \to F$ 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 $h,k \in E$, $\left(D^2 f(a)h\right) k \in F$

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

On va donc, en utilisant cette manipulation, identifier $\mathcal{L}^2(E,F)$ avec l'espace des applications bilinéaires continues de $E \times E \to F$, noté $\mathcal{L}_2(E,F)$.

2.0.36 Théorème
L'application $\mu : \mathcal{L}^2(E,F) \to \mathcal{L}_2(E,F)$, $\varphi \mapsto \mu(\varphi) = B_\varphi$ telle que pour tout $(h,k) \in E \times E$, $B_\varphi(h,k) = \varphi(h)k$, est un isomorphisme isométrique.

Démonstration: L'application μ est évidemment linéaire, on va montrons d'abord qu'elle est surjective. Soit $B \in \mathcal{L}_2(E,F)$, On définit $\varphi_B : E \to \mathcal{L}(E,F)$, par $\varphi_B(h) = B(h,.)$, alors on a bien $\mu(\varphi_B) = B$. 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 μ 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 μ est une isométrie.■

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

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

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

2.0.38 Remarque (et notations). Si $E = \mathbb{R}^n$ et $\{e_1, \ldots, e_n\}$ sa base canonique, alors

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

2.0.39 Définition
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 (c'est un résultat plus général que 2.0.34, la condition de continuité des dérivées partielles secondes n'est pas requise)

2.0.40 THÉORÈME (THÉORÈME DE SYMÉTRIE DE SCHWARZ)
Soient $f : U \to F$ une application différentiable. Si f est deux fois différentiable en $a \in U$, $D^2 f(a)$ 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: 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 $\lim_{t\to 0} \frac{A}{t^2} = D^2f(a)(h,k)$, comme $A(t,h,k) = A(t,k,h)$, on aurait prouvé la symétrie de $D^2f(a)$. On doit alors montrer que $\lim_{t\to 0} \frac{A}{t^2} = D^2f(a)(h,k)$.

On a d'après 1.8.16
$$\begin{aligned}\|A(t,h,k) - t^2D^2f(a)(h,k)\| &= \|g(h) - g(0) - t^2D^2f(a)(h,k)\| \\ &\leq \|h\| \left(\sup_{x\in]0,h[} \|Dg(x) - t^2D^2f(a)(.,k)\| \right).\end{aligned} \tag{2.1}$$
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
$Df(a+t(x+k)) = Df(a) + tD^2f(a)(.,x+k) + \|t(x+k)\|\varepsilon_1(t(x+k))$ 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. Alors, 2.2 devient

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

■

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 $\lim_{t\to 0}\frac{A}{t^2} = D^2f(a)(h,k)$.■

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

Soit $f : U \to F$ 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 $x, y \in \mathbb{R}^n$.

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

2.0.43 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 $\frac{\partial f}{\partial x}(0,y) = \lim_{x\to 0} \frac{f(x,y)-f(0,y)}{x-0} = y\lim_{x\to 0} g(x,y)$, de même $\frac{\partial f}{\partial y}(x,0) = x\lim_{y\to 0} g(x,y)$,

d'où $\frac{\partial^2 f}{\partial y\partial x}(0,0) = \lim_{y\to 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y-0} = \lim_{y\to 0}\left(\lim_{x\to 0} g(x,y)\right)$.

Un calcul analogue montre que $\frac{\partial^2 f}{\partial x\partial y}(0,0) = \lim_{x\to 0}\left(\lim_{y\to 0} g(x,y)\right)$. Pour obtenir $\frac{\partial^2 f}{\partial y\partial x}(0,0) \neq \frac{\partial^2 f}{\partial y\partial x} f(0,0)$, il suffit de choisir g telle que $\lim_{y\to 0}\left(\lim_{x\to 0} g(x,y)\right) \neq \lim_{x\to 0}\left(\lim_{y\to 0} g(x,y)\right)$.

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 $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3\sin(\frac{1}{x})$ si $(x,y) \neq (0,0)$ et $f(0,0) = 0$; vérifie $\frac{\partial^2 f}{\partial x\partial y}(0,0) = \frac{\partial^2 f}{\partial y\partial x}(0,0) = 0$, mais n'est pas deux fois différentiable. (exercice)

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

On va commencer par voir la différentielle d'ordre r comme une application r-linéaire. Soit $f : U \to F$ de classe C^{r-1}. La différentielle d'ordre r en un point $a \in U$, $D^r f(a) \in \mathcal{L}^r(E,F) = \mathcal{L}\underbrace{(E,\ldots,(\mathcal{L}(E,\mathcal{L}(E}_{r},F))\ldots)$ d'où pour tout $(h_1,\ldots,h_r) \in E^r$,

$D^r f(a).h_1.h_2.\ldots.h_r = (D^r f(a)h_1)h_2)\ldots)h_r \in F$.

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, pour rendre les choses plus simples, en utilisant une généralisation de 2.0.36, identifier $\mathcal{L}^r(E,F)$ avec l'espace des applications r-linéaires continues de $E^r \to F$, noté $\mathcal{L}_r(E,F)$ (voir 1.5.1)

Ainsi la différentielle d'ordre r de f au point a sera vue comme une application r- linéaire continue.

L'espace $\mathcal{L}^r(E,F)$ est muni de la norme d'opérateur : $\|v\| = \sup_{x_i\neq 0} \frac{\|v.x_1.\ldots.x_r\|}{\|x_1\|.\ldots.\|x_r\|}$ et l'espace $\mathcal{L}_r(E,F)$ est muni de la norme : $\|L\| = \sup_{h_i\neq 0} \frac{\|L(h_1,\ldots,h_r)\|}{\|h_1\|.\ldots.\|h_r\|}$.

2.0.44 THÉORÈME

L'application $\mu : \mathcal{L}^r(E,F) \to \mathcal{L}_r(E,F), \phi \mapsto \mu(\phi) = L_\phi$ telle que pour tout $(h_1,\ldots,h_r) \in E\times\ldots\times E,\ \ L_\phi(h_1,h_2\ldots,h_r) = \phi.h_1.h_2.\ldots.h_r$, est un isomorphisme isométrique i.e. une bijection linéaire qui préserve la norme.

Démonstration: Voir 2.0.36 ∎

2.0.46 DÉFINITION
Une application r-linéaire $u \in \mathcal{L}_r(E,F)$ est dite *symétrique* si pour tout $(h_1,\dots,h_r) \in E^r$ et toute permutation $\sigma \in \mathcal{S}(r)$ i.e. toute bijection
$\sigma : \{1,\dots,r\} \to \{1,\dots,r\}$ on a :

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

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

2.0.47 THÉORÈME (THÉORÈME DE SYMÉTRIE DE SCHWARZ)
Soient $f : U \to F$ une application r fois différentiable . Alors $D^r f(a)$ est une application r-linéaire symétrique i.e.pour tout $(h_1,\dots,h_r) \in E^r$ et toute permutation $\sigma \in \mathcal{S}(r)$ on a

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

2.0.48 EXEMPLE. Si $f : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto f(x,y)$ 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 $\frac{\partial^3 f}{\partial x \partial y^2}(a) = \frac{\partial^3 f}{\partial y \partial x \partial y}(a) = \frac{\partial^3 f}{\partial y^2 \partial x}(a)$

2.0.4 Dérivées partielles d'ordre supérieur

Soit $f : U \to \mathbb{R}^p$, U ouvert de $\mathbb{R}^n$ et $\{e_1,\dots,e_n\}$ la base canonique.
Pour $r \in \mathbb{N}^*$, des indices $i_1,\dots,i_r \in \{1,\dots,n\}$ on appelle le vecteur

$$\frac{\partial^r f(x)}{\partial x_{i_1},\dots,\partial x_{i_r}} := D_{e_{i_1}}\dots D_{e_{i_r}} f(x)$$

une dérivée partielle d'ordre r de f en x.

2.0.49 THÉORÈME
Soit $f : U \to \mathbb{R}^p$, U ouvert de $\mathbb{R}^n$ et $r \in \mathbb{N}^*$.

1. $f \in C^r(U,\mathbb{R}^p)$ si et seulement si f a toute ses dérivées partielles d'ordre k continues pour tout $0 \le k \le r$.
2. Pour $f \in C^r(U,\mathbb{R}^p)$, on a

$$\frac{\partial^r f(x)}{\partial x_{i_1},\dots,\partial x_{i_r}} = \frac{\partial^r f(x)}{\partial x_{i_{\sigma(1)}},\dots,\partial x_{i_{\sigma(r)}}}$$

pour toute permutation $\sigma \in \mathcal{S}(r)$, i.e. les dérivées partielles ne dépendent pas de l'ordre de dérivation.

Démonstration: On observe que pour $1 \leq k \leq r$ et $h_1, \ldots, h_k \in \mathbb{R}^n$ on a :

$$D^k f(x)(h_1, \ldots, h_k) = D(\ldots(D(Df(x)h_1)h_2)\ldots)h_k$$

pour tout $x \in U$. En particulier,

$$D^k f(x)(e_{i_1}, \ldots, e_{i_k}) = D_{e_{i_1}} \ldots D_{e_{i_k}} f(x) = \frac{\partial^k f(x)}{\partial x_{i_1}, \ldots, \partial x_{i_k}} \tag{2.4}$$

Par conséquent, pour $h_i = \sum_{j=1}^n h_i^j e_j$, en utilisant la multilinéarité de $D^k f(x)$ on a

$$D^k f(x)(h_1, \ldots, h_k) = \sum_{i_1,\ldots,i_k=1}^{n} D_{e_{i_1}} \ldots D_{e_{i_k}} f(x) h_1^{i_1} \ldots h_k^{i_k} = \sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial^k f(x)}{\partial x_{i_1}, \ldots, \partial x_{i_k}} h_1^{i_1} \ldots h_k^{i_k} \tag{2.5}$$

■

1. La formule 2.4 montre que si $f \in C^r(U, \mathbb{R}^p)$, alors ses dérivées partielles d'ordre $k \leq r$ sont continues. Inversement, on suppose que f a toutes ses dérivées partielles d'ordre $k \leq r$ sont continues.
   On sait que pour $r = 1$, il y a équivalence entre dérivées partielles continues et $f \in C^1(U, \mathbb{R}^p)$. On suppose, par récurrence que le résultat est vrai à l'ordre $r-1$. D'après la formule 2.5, on voit alors que les dérivées partielles $D_j(D^{r-1}f)$ existent et sont continues pour tout $i \in \{1, \ldots, n\}$, d'où $D(D^{r-1}f)$ existe et est continue. On a ainsi montrer que $D^r f = D(D^{r-1}f)$ existe et est continue i.e. $f \in C^r(U, \mathbb{R}^p)$.
2. Conséquence du théorème de symétrie de Schwarz.■

2.0.5 Polynômes et puissances symboliques d'une différentielle

2.0.51 DÉFINITION (POLYNÔME GÉNÉRALISÉ)

1. Soient E et F deux espaces vectoriels normés et $r \in \mathbb{N}^*$.
   Une application $f : E \to F$ est un **polynôme k-homogène** si il existe une application k-linéaire continue et symétrique $g : E^k \to F$ telle que pour tout $x \in E$, $f(x) = g(x \ldots, x)$.
2. Une application $P : E \to F$ est dite **polynomiale de degré** n, si elle s'écrit : $P(x) = \sum_0^n a_k(x, \ldots, x)$, pour tout $x \in E$ où pour tout $k \in \{0, \ldots, r\}$, a_k est un polynôme k-homogène et $a_k \not\equiv 0$.

2.0.52 PROPOSITION

Soit $f : E \to F$ un polynôme k-homogène application tel que pour tout $x \in E$, $f(x) = g(x, \ldots, x)$ où g est une application k-linéaire symétrique et continue.

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

Démonstration: On sait déja voir 2.0.30, que f est de classe C^∞ et que $D^r f = 0$ si $r > k$.
On va démontrer par récurrence sur r la formule pour $1 \leq r \leq k$.
Si $r = 0$, on a $D^r f(a) = f(a) = g(a, \ldots, a)$.
Supposons que $D^r f$ est l'application r-linéaire, telle que $D^r f(x) = \frac{k!}{(k-r)!} g(\underbrace{x, \ldots, x}_{k-r}, \underbrace{\ldots}_{r})$.

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

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

Par multilinéarité de g on a

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

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

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

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

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

Finalement, $D^{r+1} f(a) h = \frac{k!}{(k-r)!} (k-r) g(x, \ldots, x, h, \ldots) = \frac{k!}{(k-r-1)!} g(x, \ldots, x, h, \ldots)$

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

2.0.54 THÉORÈME

Soient E, F deux espaces vectoriels normés et P de E dans F une application polynomiale de degré n. Alors, $P(x) = \displaystyle\sum_{k=0}^{n} \frac{1}{k!} D^k P(0)(x, \ldots, x)$.

Démonstration: Comme $P(x) = \sum_{k=0}^n a_k(x, \ldots, x)$ et que d'après la proposition précédente 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 $P(x) = \sum_{k=0}^n a_k(x, \ldots, x) = \sum_{k=0}^n \frac{1}{k!} D^k P(0)(x, \ldots, x)$.$\blacksquare$

2.0.56 COROLLAIRE
Soient $P = \sum_{k=0}^{n} a_k$ et $Q = \sum_{k=0}^{n} b_k$ deux applications polynomiales de E dans F de degré n. Alors $P = Q$ si et seulement si pour tout $k \in \{0, \ldots, n\}$, $a_k = b_k$.

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

2.0.58 DÉFINITION
Soient E et F des evn. $a \in U$ ouvert de E. Soit $f : U \to F$ une application r fois différentiable en a. L'application $\phi : E \to F$ définie par $\phi(h) = D^r f(a)(h, \ldots, h)$ est appelée la **puissance symbolique d'ordre r de la différentielle de f en a**.

Notation : On notera $D^r f(a)(h, \ldots, h)$ par $D^r f(a).h^r$.

2.0.59 REMARQUE
Soit $f : U \to F$ une application r fois différentiable en $a \in U$.
On suppose que $[a, a + h] \subset U$ et on définit $g : [0, 1] \to F$ par $g(t) = f(a + th)$.
A l'aide de la règle de composition et d'une récurrence sur r on obtient la formule

$$\boxed{g^{(k)}(t) = \frac{d^k g}{dt^k}(t) = D^k f(a + th)h^k, \quad 1 \leq k \leq r.}$$

En effet, si $g^{(k-1)}(t) = D^{k-1} f(a + th) \underbrace{(h, \ldots, h)}_{k-1}$ alors,

$g^{(k)}(t) = (g^{(k-1)})'(t) = (D^{k-1} f(a + th)(h, \ldots, h))' = D(D^{k-1} f(a + th)(h, \ldots, h))h$
$= D^k f(a + th).h^k$.

En particulier $g^{(k)}(0) = D^k f(a)h^k, \quad 1 \leq k \leq r$.

1. <u>**Formule développée pour les fonctions de n variables :**</u> $E = \mathbb{R}^n$ et $F = \mathbb{R}$.
   Si $\{e_1, \ldots, e_n\}$ est la base canonique de $\mathbb{R}^n$, alors pour $h = \sum_{i=1}^{n} h_i e_i$ on a

   $$\boxed{D^r f(a).h^r = D^r f(a)(h, \ldots, h) = \sum_{1 \leq i_1, \ldots, i_r \leq n} h_{i_1} \ldots h_{i_r} \frac{\partial^r f}{\partial x_{i_1} \ldots \partial x_{i_r}}(a)}$$

   où la somme est étendue à toutes les suites $(i_1 \ldots, i_r)$ de $\{1, \ldots, n\}$.
   Cette écriture ne tient pas compte des symétries (Théorème de Schwarz).

Ce qui importe le plus est le nombre de fois où l'indice i_s apparaît dans cette formule .

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

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

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

2.0.60 LEMME

$$m(\alpha) = \frac{r!}{\alpha!} = \frac{r!}{\alpha_1! \dots \alpha_n!}.$$

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

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

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

2.0.62 COROLLAIRE (GÉNÉRALISATION DE LA FORMULE DU BINÔME)
pour tout entier $r \ge 1$ et $h_1, \dots, h_r \in \mathbb{R}$ on a

$$(h_1 + \dots + h_n)^r = \sum_{\alpha_1 + \dots + \alpha_n = r} \frac{r!}{\alpha_1! \dots \alpha_n!} h_1^{\alpha_1} \dots h_n^{\alpha_n} = \sum_{|\alpha|=r} \frac{r!}{\alpha!} h^\alpha.$$

la somme est étendue à tout multi-indice $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ tel que $|\alpha| = \sum_{i=1}^n \alpha_i = r$.

2.0.63 COROLLAIRE

1. Avec cette notation on a :

$$\boxed{\frac{1}{r!} D^r f(a).h^r = \frac{1}{r!} D^r f(a)(h, \dots, h) = \sum_{|\alpha|=r} \frac{D^\alpha f(a)}{\alpha!} h^\alpha}$$

   Cette dernière formule tient bien compte des symétries.

2. Si P est le polynôme défini par $P(h) = \sum_{r=0}^{n} \frac{1}{r!} D^r f(a).h^r$ alors sa différentielle

$$\boxed{DP(h) = \sum_{r=1}^{n} \frac{1}{(r-1)!} D^r f(a).h^{r-1}}$$

2.0.6 Formules de Taylor

2.0.64 THÉORÈME (FORMULE DE TAYLOR-YOUNG)
Soient $U \subset E$, un ouvert et f une application de U dans F.

Si f est r fois différentiables au point $a \in U$, alors elle admet un développement de Taylor-Young à l'ordre r au point a c-à-d : qu' il existe une fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}^p$, avec $\lim_{h \to 0} \varepsilon(h) = 0$ telle que :

$$\boxed{f(a+h) = f(a) + \sum_{k=1}^{r} \frac{1}{k!} D^k f(a).h^k + \|h\|^r \varepsilon(h)}$$

Démonstration: Il s'agit de montrer par récurrence l'assertion $\mathcal{P}_r$ suivante :
$\mathcal{P}_r$: " l'application, dite reste d'ordre r, $R_r(h) = f(a+h) - f(a) - \sum_{k=1}^{r} \frac{1}{k!} D^k f(a).h^k$ vérifie : $\lim_{h \to 0} \frac{R_r(h)}{\|h\|^r} = 0.$"
La définition de la différentielle d'une fonction en un point montre que $\mathcal{P}_1$ est vraie. Soit un entier $r \geq 2$. Nous supposons, que $\mathcal{P}_{r-1}$ est vraie et montrons $\mathcal{P}_r$.

Soit f r fois différentiable en a et associons-lui le polynôme $P(h) = \sum_{r=0}^{n} \frac{1}{r!} D^r f(a).h^r$. Alors $DP(h) = \sum_{r=1}^{n} \frac{1}{(r-1)!} D^r f(a).h^{r-1}$.

On constate que $DR_r(h) = Df(a+h) - DP(h)$ est le reste d'ordre $r-1$ pour l'application Df.

Mais, d'après $\mathcal{P}_{r-1}$, $\lim_{h \to 0} \frac{DR_r(h)}{\|h\|^{r-1}} = 0$, i.e. pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que $\|k\| < \delta$ entraîne $\|DR_r(k)\| < \epsilon \|k\|^{r-1}$.

Soit $h \in E$ tel que $\|h\| < \delta$. Pour tout $k \in [0, h]$, vérifie $\|k\| \leq \|k\|^{r-1} < \delta$ et alors $\|DR_r(k)\| < \epsilon \|k\|^{r-1} \leq \epsilon \|h\|^{r-1}$.

D'après le théorème des accroissements finis appliquée dans le convexe $[0, h]$ on aura $\|R_r(h)\| = \|R_r(h) - R_r(0)\| < \epsilon \|k\|^{r-1}.(\|h\| - 0) = \epsilon \|h\|^r$. On a ainsi montré que pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que $\|h\| < \delta$ entraîne $\|R_r(h)\| < \epsilon \|k\|^r$ i.e. $\lim_{h \to 0} \frac{R_r(h)}{\|h\|^r} = 0$ et donc $\mathcal{P}_r$ est vérifiée. ■

2.0.66 REMARQUE. Soit $f \in C^r(U, \mathbb{R})$, alors f admet un développement de Taylor-Young en tout point $a \in U$. En particulier si $E = \mathbb{R}^n$ on a

1. Si $r = 1$, on a

$$\begin{aligned} f(x) = f(a) + Df(a)(x-a) + \|x-a\|.\varepsilon(x) &= f(a) + \langle \nabla f(a), (x-a) \rangle + \|x-a\|.\varepsilon(x) \\ &= f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).x_i + \|x-a\|.\varepsilon(x) \end{aligned}$$

avec $\lim_{x \to a} \varepsilon(x) = 0$.
(on utilise la formule de Taylor-Young à l'ordre 1, en posant $h = x - a$.)

2. Si $r = 2$, on a

$$f(x) = f(a) + Df(a)(x-a) + \frac{1}{2}D^2 f(a)(x-a, x-a) + \|x-a\|^2.\varepsilon(x)$$
$$= f(a) + \langle \nabla f(a), (x-a) \rangle + \frac{1}{2}\langle H_f(a)(x-a), (x-a) \rangle + \|x-a\|^2.\varepsilon(x)$$
$$= f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a).x_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a).x_i x_j + \|x-a\|^2.\varepsilon(x)$$

avec $\lim_{x\to a} \varepsilon(x) = 0$.

2.0.67 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 $f : \mathbb{R} \to \mathbb{R}$ 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 !)

2.0.68 EXEMPLE. Soit $f : B((0,0),1) \to \mathbb{R}$ 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 $a = (0,0)$.

$$f(x,y) = f(0,0) + \langle \nabla f(0,0), \begin{pmatrix} x \\ y \end{pmatrix} \rangle + \frac{1}{2}\langle H_f(0,0)\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \rangle + \left\| \begin{pmatrix} x \\ y \end{pmatrix} \right\|_2^2 .\varepsilon(x)$$

$$= f(0,0) + \frac{\partial f}{\partial x}(0,0)x + \frac{\partial f}{\partial y}(0,0)y + +\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(0,0)x^2 + 2\frac{\partial^2 f}{\partial x \partial y}(0,0)xy + \frac{\partial^2 f}{\partial y^2}(0,0)y^2\right) + (x^2+y^2).\varepsilon(x,y)$$

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

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

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

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

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

$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{y\to 0} \frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y} = -\frac{1}{8}$

$\frac{\partial^2 f}{\partial y^2}(0,0) = \lim_{y\to 0} \frac{\frac{\partial f}{\partial y}(0,y) - \frac{\partial f}{\partial y}(0,0)}{y} = \frac{1}{4}$.

Ainsi, $f(x,y) = \frac{1}{2} - \frac{1}{8}xy + \frac{1}{8}y^2 + (x^2+y^2).\varepsilon(x,y)$

2.0.7 Annexe : d'autres formules de Taylor

2.0.69 THÉORÈME (FORMULE DE TAYLOR AVEC RESTE INTÉGRAL)
Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ 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

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

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

2.0.71 LEMME
Soit I un intervalle ouvert de $\mathbb{R}$, contenant $[0,1]$ et $g : I \to \mathbb{R}$ 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: (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

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

2.0.73 THÉORÈME (FORMULE DE TAYLOR-LAGRANGE)
Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ 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 :

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

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

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

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

2.0.76 REMARQUE. Ce théorème n'est pas vrai pour les fonctions vectorielles c-à-d à valeurs dans $\mathbb{R}^m$, $m \geq 2$ (voir la remarque 1.8.6)