Dans ce qui suit, U est un *ouvert* non vide de $\mathbb{R}^d$, $d \geq 1$. Une fonction

$$f: \begin{array}{ccc} U & \to & \mathbb{R}^p \\ x = (x_1, \ldots, x_d) & \mapsto & f(x) = (f_1(x), \ldots, f_p(x)) \end{array}$$

est dite *scalaire* si $p = 1$, à valeurs vectorielles sinon. Ses *composantes* sont les fonctions scalaires

$$f_i: \begin{array}{ccc} U & \to & \mathbb{R} \\ x = (x_1, \ldots, x_d) & \mapsto & f_i(x) \end{array}$$

pour $i \in \{1, \ldots, p\}$. Une *application partielle* de f est une fonction d'une variable, définie sur un ensemble

$$U_{a,j} := \{y \in \mathbb{R}\,;\ (a_1, \ldots, a_{j-1}, y, a_{j+1}, \ldots, a_d) \in U\}$$

où $a = (a_1, \ldots, a_d) \in U$ (noter que $U_{a,j}$ est un *voisinage* de a_j dans $\mathbb{R}$), par

$$\begin{array}{ccc} U_{a,j} & \to & \mathbb{R}^p \\ y & \mapsto & f(a_1, \ldots, a_{j-1}, y, a_{j+1}, \ldots, a_d) \end{array}$$

Fonctions continues Une fonction $f : U \to \mathbb{R}^p$ est *continue* si et seulement si ses composantes le sont. (La continuité des applications partielles ne suffit pas pour la continuité de f.)

Dérivées partielles Si l'application partielle

$$\begin{array}{ccc} U_{a,j} & \to & \mathbb{R}^p \\ y & \mapsto & f(a_1, \ldots, a_{j-1}, y, a_{j+1}, \ldots, a_d) \end{array}$$

est *dérivable* en a_j, on note $\dfrac{\partial f}{\partial x_j}(a)$ sa dérivée en a_j, appelée *dérivée partielle* de f par rapport à x_j au point a. Si f admet des dérivées partielles par rapport à toutes les variables x_j au point a, on définit $J_f(a)$ sa *matrice jacobienne* au point a comme la matrice à p lignes et d colonnes de coefficients

$$(J_f(a))_{ij} = \frac{\partial f_i}{\partial x_j}(a)\,,\ i \in \{1, \ldots, p\}\,,\ j \in \{1, \ldots, d\}\,.$$

Si f est scalaire ($p = 1$), on définit le *gradient* de f au point a par

$$\nabla f(a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_d}(a) \end{pmatrix} = {}^t(J_f(a))\,.$$

Si $d = p$, on définit la *divergence* de f au point a par

$$\nabla \cdot f(a) = \sum_{i=1}^{d} \frac{\partial f_i}{\partial x_i}(a) = \mathrm{tr}\,(J_f(a))\,.$$

Si $d = p = 3$, on définit le *rotationnel* de f au point a par

$$\nabla \wedge f(a) = \begin{pmatrix} \dfrac{\partial f_3}{\partial x_2}(a) - \dfrac{\partial f_2}{\partial x_3}(a) \\ \dfrac{\partial f_1}{\partial x_3}(a) - \dfrac{\partial f_3}{\partial x_1}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) - \dfrac{\partial f_1}{\partial x_2}(a) \end{pmatrix}.$$

NB: Il faut savoir *définir* les notions en *rouge* et démontrer les énoncés en bleu.

Dérivées directionnelles Étant donnés $a \in U$ et $h \in \mathbb{R}^d$, si la fonction d'une variable $t \mapsto f(a+th)$ est dérivable en zéro, on note

$$\mathrm{D}_h f(a) = \frac{\mathrm{d}}{\mathrm{d}t}(f(a+th))_{|t=0}$$

et on dit que f admet $\mathrm{D}_h f(a)$ pour *dérivée suivant le vecteur* h au point a. En particulier, si $(\mathrm{e}_1, \ldots, \mathrm{e}_d)$ désigne la base « canonique » de $\mathbb{R}^d$, les dérivées partielles par rapport à x_j sont les dérivées suivant les vecteurs e_j (si elles existent) :

$$\frac{\partial f}{\partial x_j}(a) = \frac{\mathrm{d}}{\mathrm{d}t}(f(a+t\mathrm{e}_j))_{|t=0} \;=\; \mathrm{D}_{\mathrm{e}_j} f(a)\,.$$

Différentiabilité Une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est *différentiable* en $x \in U$ s'il existe une application *linéaire* $u : \mathbb{R}^d \to \mathbb{R}^p$ telle que

$$\lim_{h \overset{\neq}{\to} 0} \frac{\|f(x+h) - f(x) - u(h)\|}{\|h\|} \;=\; 0\,.$$

Si c'est le cas, u est unique et on la note $\mathrm{d}f(x)$.
Si f est différentiable en a alors elle admet un dérivée suivant tout vecteur h au point a et

$$\boxed{\forall h \in \mathbb{R}^d\,,\ \mathrm{d}f(a)(h) = \mathrm{D}_h f(a) = \frac{\mathrm{d}}{\mathrm{d}t}(f(a+th))_{|t=0} \;=\; \sum_{j=1}^{d} h_j\, \frac{\partial f}{\partial x_j}(a)}\,.$$

L'existence de dérivées partielles, et même de dérivées directionnelles, ne suffit pas pour la différentiabilité.
Si f est fonction d'une variable ($d = 1$), sa différentiabilité en t équivaut à sa *dérivabilité* en t et

$$\forall k \in \mathbb{R}\,,\ \mathrm{d}f(t)(k) = k f'(t)\,, \quad f'(t) \;=\; \lim_{s \overset{\neq}{\to} 0} \frac{f(t+s) - f(t)}{s}\,.$$

La matrice de $\mathrm{d}f(x)$ dans les bases canoniques de $\mathbb{R}^d$ et $\mathbb{R}^p$ est la matrice jacobienne de f au point x.
La somme de deux fonctions f et g différentiables en x est différentiable en x et

$$\mathrm{d}(f+g)(x) = \mathrm{d}f(x) + \mathrm{d}g(x)\,.$$

Le produit d'une fonction f par une fonction scalaire λ est différentiable en x si les deux le sont et

$$\forall h \in \mathbb{R}^d\,, \quad \mathrm{d}(\lambda f)(x)(h) = \lambda(x)\, \mathrm{d}f(x)(h) + \mathrm{d}\lambda(x)(h)\; f(x)\,.$$

Une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est dite *différentiable* (tout court) si elle est différentiable en tout point, et sa *différentielle* est la fonction

$$\begin{array}{rccl} df: & U & \to & \mathscr{L}(\mathbb{R}^d;\mathbb{R}^p) \\ & x & \mapsto & \mathrm{d}f(x) \end{array}$$

où $\mathscr{L}(\mathbb{R}^d;\mathbb{R}^p)$ désigne l'espace des applications linéaires de $\mathbb{R}^d$ dans $\mathbb{R}^p$.

- ✓ Une fonction constante est différentiable et sa différentielle est identiquement l'application nulle.
- ✓ Une application linéaire est différentiable et sa différentielle est constante : pour $u \in \mathscr{L}(\mathbb{R}^d;\mathbb{R}^p)$,

$$\forall x \in \mathbb{R}^d\,, \forall h \in \mathbb{R}^d\,, \quad \mathrm{d}u(x)(h) = u(h)\,.$$

✓ Une application $\phi : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^p$ *bilinéaire* est différentiable et

$$\forall (x,y) \in \mathbb{R}^d \times \mathbb{R}^m\,,\ \forall (h,k) \in \mathbb{R}^d \times \mathbb{R}^m\,, \quad \mathrm{d}\phi(x,y)(h,k) = \phi(x,k) + \phi(h,y)\,.$$

Ceci s'applique par exemple, avec $d = m$ et $p = 1$, au *produit scalaire* $(x,y) \mapsto x \cdot y$.
Si $d = m$, l'application $q : x \in \mathbb{R}^d \mapsto q(x) := \phi(x,x)$ est différentiable et

$$\forall x \in \mathbb{R}^d\,,\ \forall h \mathbb{R}^d\,, \quad \mathrm{d}q(x)(h) = \phi(x,h) + \phi(h,x)\,.$$

Si $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est différentiable en $x \in U$, si $f(U) \subset V$, ouvert de $\mathbb{R}^p$ et $g : V \subset \mathbb{R}^p \to \mathbb{R}^m$ est différentiable en $f(x)$, alors $g \circ f : U \subset \mathbb{R}^d \to \mathbb{R}^m$ est différentiable en x et

$$\forall h \in \mathbb{R}^d\,,\ \mathrm{d}(g \circ f)(x)(h) = \mathrm{d}g(f(x))(\mathrm{d}f(x)(h))\,,$$

c'est-à-dire encore $\boxed{\mathrm{d}(g \circ f)(x) \,=\, \mathrm{d}g(f(x)) \circ \mathrm{d}f(x)}$, ce qui signifie aussi $\boxed{J_{g\circ f}(x) = J_g(f(x))\, J_f(x)}$.

✓ La *norme euclidienne* $N : x = (x_1, \ldots, x_d) \mapsto \|x\|_2 = \sqrt{\sum_{i=1}^d x_i^2}$ est différentiable sur $\mathbb{R}^d \backslash \{0\}$: quels que soient $x \in \mathbb{R}^d \backslash \{0\}$ et $h \in \mathbb{R}^d$, $\mathrm{d}N(x)(h) \,=\, \dfrac{(x \cdot h)}{\|x\|_2}$.

Intégration des fonctions à valeurs vectorielles Soient $a, b \in \mathbb{R}$ avec $a < b$ et

$$\begin{array}{rccl} f: & [a,b] & \to & \mathbb{R}^p \\ & t & \mapsto & f(t) = (f_1(t), \ldots, f_p(t)) \end{array}$$

continue. Son intégrale sur le segment $[a,b]$ est le vecteur de $\mathbb{R}^p$ défini par :

$$\int_a^b f(t)\mathrm{d}t \,=\, \left(\int_a^b f_1(t)\mathrm{d}t, \ldots, \int_a^b f_p(t)\mathrm{d}t \right)\,.$$

C'est la limite des *sommes de Riemann* de f sur $[a,b]$, c'est-à-dire la limite lorsque n tend vers $+\infty$ de

$$\sum_{i=1}^n (x_i - x_{i-1})\, f(y_i)$$

où $(x_i)_{0\le i\le n}$ est une *subdivision* de $[a,b]$ et $y_i \in [x_{i-1}, x_i]$ pour tout $i \in \{1, \ldots, n\}$. (Noter que les points x_i de la subdivision ainsi que les points y_i dépendent de n : on ne le fait pas apparaître pour ne pas alourdir les notations.)

L'application $f \mapsto \int_a^b f(t)\mathrm{d}t$ est linéaire. On définit $\int_b^a f(t)\mathrm{d}t \,=\, -\, \int_a^b f(t)\mathrm{d}t$.
Si $f : I \to \mathbb{R}^p$ est continue sur l'*intervalle* I, quels que soient a, b, $c \in I$ on a
$\displaystyle\int_a^b f(t)\mathrm{d}t \,=\, \int_a^c f(t)\mathrm{d}t \,+\, \int_c^b f(t)\mathrm{d}t$ (*relation de Chasles*).

De plus, la fonction $x \mapsto \int_a^x f(t)\mathrm{d}t$ est dérivable et

$$\frac{\mathrm{d}}{\mathrm{d}x} \int_a^x f(t)\mathrm{d}t \,=\, f(x)\,.$$

Plus généralement, si α et β sont des fonctions dérivables d'une variable et à valeurs dans I on a

$$\frac{\mathrm{d}}{\mathrm{d}u} \int_{\alpha(u)}^{\beta(u)} f(t)\mathrm{d}t \,=\, \beta'(u)\, f(\beta(u)) \,-\, \alpha'(u) f(\alpha(u))\,.$$

Quelle que soit la norme $\|\cdot\|$ choisie dans $\mathbb{R}^p$, on a

$$\left\| \int_a^b f(t)\mathrm{d}t \right\| \leq \left| \int_a^b \|f(t)\|\mathrm{d}t \right| .$$

Si $g : I \to \mathbb{R}^p$ est une autre fonction continue et $\|\cdot\|_2$ est la norme euclidienne on a

$$\left| \int_a^b f(t) \cdot g(t)\mathrm{d}t \right| \leq \left| \int_a^b \|f(t)\|_2^2\mathrm{d}t \right|^{1/2} \left| \int_a^b \|g(t)\|_2^2\mathrm{d}t \right|^{1/2} \quad (\textit{inégalité de Cauchy-Schwarz}).$$

Théorème des accroissements finis Si $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ est continue et dérivable sur $]a, b[$ il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$. C'est la *formule des accroissements finis*, fausse pour les fonctions à valeurs vectorielles.

Inégalité des accroissements finis pour les fonctions d'une variable : Soit $f : I \subset \mathbb{R} \to \mathbb{R}^p$ dérivable sur l'intervalle ouvert U et soient $a, b \in I$, $a < b$. S'il existe $M \in \mathbb{R}^+$ tel que

$$\forall t \in [a, b], \quad \|f'(t)\| \leq M ,$$

alors $\|f(b) - f(a)\| \leq M|b - a|$.

Si la fonction f ci-dessus est *de classe $\mathscr{C}^1$*, c'est-à-dire que $f' : I \to \mathbb{R}^p$ est continue, l'inégalité des accroissements finis se déduit de la formule

$$f(b) - f(a) = \int_a^b f'(t)\mathrm{d}t = \int_0^1 f'(a + \theta(b - a))(b - a)\mathrm{d}\theta .$$

Inégalité des accroissements finis pour les fonctions de plusieurs variables : Soit $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ différentiable sur l'ouvert U et soient $x, y \in U$ tels que $[x, y] := \{x + \theta(y - x) ;\ \theta \in [0, 1]\} \subset U$. S'il existe $M \in \mathbb{R}^+$ tel que

$$\forall \theta \in [0, 1], \quad ||| \mathrm{d}f(x + \theta(y - x)) ||| \leq M ,$$

alors $\|f(x) - f(y)\| \leq M\|x - y\|$.

Une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est dite *de classe $\mathscr{C}^1$* si elle est différentiable et si sa différentielle $df : U \to \mathscr{L}(\mathbb{R}^d; \mathbb{R}^p)$ est continue ($\mathscr{L}(\mathbb{R}^d; \mathbb{R}^p)$ étant un espace vectoriel normé de dimension finie, la *norme subordonnée* $|||\cdot|||$ est *équivalente* à toutes les autres), ou de façon équivalente, si elle admet des dérivées partielles $\dfrac{\partial f}{\partial x_i} : U \to \mathbb{R}^p$ continues ($i \in \{1, \dots d\}$).

Théorème du point fixe de Picard Soit $f : F \subset \mathbb{R}^d \to \mathbb{R}^d$ telle que $f(F) \subset F$, où F est un *fermé*. Si f est *contractante* (c'est-à-dire telle que : $\exists k \in [0, 1[, \forall x, y \in A , \|f(x) - f(y)\| \leq k\|x - y\|$) alors il existe un unique $x \in F$ tel que $f(x) = x$.
(Ce théorème est faux pour les fonctions *Lipschitziennes* de rapport $k \geq 1$.)

Difféomorphismes Une fonction $f : U \subset \mathbb{R}^d \to V \subset \mathbb{R}^p$ est un *difféomorphisme* de l'ouvert U sur l'ouvert V si elle différentiable, *bijective* et sa *réciproque* est différentiable. C'est un *$\mathscr{C}^1$-difféomorphisme* si elle de plus de classe $\mathscr{C}^1$ et sa réciproque aussi. Si c'est le cas, on a nécessairement $d = p$ et $\mathrm{d}f(x)$ est un *isomorphisme* de $\mathbb{R}^d$ quel que soit $x \in U$, la différentielle de la fonction réciproque f^{-1} étant donnée par $\boxed{\mathrm{d}(f^{-1})(y) = (\mathrm{d}f(f^{-1}(y)))^{-1}}$ quel que soit $y \in V$.

NB: Il faut savoir *définir* les notions en *rouge* et démontrer les énoncés en bleu.

Un isomorphisme de $\mathbb{R}^d$ est un $\mathscr{C}^1$-difféomorphisme de $\mathbb{R}^d$ sur $\mathbb{R}^d$.
COORDONNÉES POLAIRES La fonction suivante est un $\mathscr{C}^1$-difféomorphisme :

$$\begin{array}{ccc} \mathbb{R}^{+*}\times]-\frac{\pi}{2},\frac{\pi}{2}[& \to & \mathbb{R}^2\backslash(\mathbb{R}^-\times\{0\}) \\ (r,\theta) & \mapsto & (x,y)=(r\cos\theta, r\sin\theta)\,. \end{array}$$

Théorème d'inversion locale Soient U un ouvert de $\mathbb{R}^d$, une fonction $f : U \to \mathbb{R}^d$ de classe $\mathscr{C}^1$ et $a \in U$ tels que $\mathrm{d}f(a)$ soit un isomorphisme de $\mathbb{R}^d$. Alors il existe un ouvert $U_a \subset U$ contenant a et un ouvert V_b contenant $b = f(a)$ tels que $f : U_a \to V_b$ soit un $\mathscr{C}^1$-difféomorphisme.

Théorème d'inversion globale Si $f : U \to \mathbb{R}^d$ est de classe $\mathscr{C}^1$ sur l'ouvert U et

- pour tout $x \in U$, $\mathrm{d}f(x)$ est un isomorphisme de $\mathbb{R}^d$,
- la fonction f est *injective*,

alors f est un $\mathscr{C}^1$-difféomorphisme de U sur $f(U)$.

Théorème des fonctions implicites Soient W un ouvert de $\mathbb{R}^d$, V un ouvert de $\mathbb{R}^p$, $U = W \times V$, et $f : U \to \mathbb{R}^d$ de classe $\mathscr{C}^1$. Pour tout $(x, y) \in U$, on note $\mathrm{d}_2 f(x, y)$ la différentielle au point y de l'application partielle

$$\begin{array}{ccc} V & \to & \mathbb{R}^p \\ v & \mapsto & f(x,v)\,. \end{array}$$

Si $(a, b) \in U$ est tel que $f(a, b) = 0$ et $\mathrm{d}_2 f(a, b)$ est un isomorphisme de $\mathbb{R}^p$, alors il existe un ouvert $U_{(a,b)} \subset U$ contenant (a, b), un ouvert $W_a \subset W$ contenant a et une fonction $\varphi : W_a \to \mathbb{R}^p$ de classe $\mathscr{C}^1$ tels que

$$((x,y)\in U_{(a,b)}\,,\ f(x,y)=0)\ \Leftrightarrow\ (x\in W_a\,,\ y=\varphi(x))\,.$$

Si $f : U = W \times V \to \mathbb{R}^d$ et $\varphi : W \to \mathbb{R}^p$ sont de classe $\mathscr{C}^1$, telles que $\mathrm{d}_2 f(x, y)$ est un isomorphisme de $\mathbb{R}^p$ quel que soit $(x, y) \in U$, et $f(x, \varphi(x)) = 0$ quel que soit $x \in W$, alors

$$\forall x\in W\,,\ \forall h\in\mathbb{R}^d\,,\ \mathrm{d}\varphi(x)(h)=-(\mathrm{d}_2 f(x,\varphi(x)))^{-1}(\mathrm{d}_1 f(x,\varphi(x))(h))\,,$$

où $\mathrm{d}_1 f(x, y)$ désigne la différentielle au point x de l'application partielle

$$\begin{array}{ccc} W & \to & \mathbb{R}^p \\ w & \mapsto & f(w,y)\,. \end{array}$$

Différentielle seconde Soit $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ où U est un ouvert. On dit que f est *deux fois différentiable* en $a \in U$ s'il existe un voisinage $U_a \subset U$ de a où f est différentiable et si la fonction

$$\begin{array}{cccc} \mathrm{d}f: & U_a & \to & \mathscr{L}(\mathbb{R}^d;\mathbb{R}^p) \\ & x & \mapsto & \mathrm{d}f(x) \end{array}$$

est différentiable. Si c'est le cas on définit l'application bilinéaire $\mathrm{d}^2 f(a) : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^p$ par

$$\mathrm{d}^2 f(a)(h,k)=(\mathrm{d}(\mathrm{d}f))(a)(h)(k)\,,\quad \forall(h,k)\in\mathbb{R}^d\times\mathbb{R}^d\,.$$

Autrement dit, $\mathrm{d}^2 f(a)(h, k) = \mathrm{d}g_k(a)(h)$ où la fonction $g_k : \mathbb{R}^d \to \mathbb{R}^p$ est définie, à $k \in \mathbb{R}^d$ fixé, par $g_k(x) = \mathrm{d}f(x)(k)$.
Théorème de Schwarz : si f est deux fois différentiable en a alors $\mathrm{d}^2 f(a)$ est *symétrique*, c'est-à-dire

$$\mathrm{d}^2 f(a)(h,k)=\mathrm{d}^2 f(a)(k,h)\,,\quad \forall(h,k)\in\mathbb{R}^d\times\mathbb{R}^d\,.$$

Dérivées partielles d'ordre supérieur Si $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est deux fois différentiable alors ses toutes ses dérivées partielles $\frac{\partial f}{\partial x_i}$ admettent des dérivées partielles, que l'on note $\frac{\partial^2 f}{\partial x_j \partial x_i}$. On a de plus

$$\mathrm{d}^2 f(a)(h,k) = \sum_{i=1}^{d}\sum_{j=1}^{d} h_j k_i \frac{\partial^2 f}{\partial x_j \partial x_i}(a)\,, \quad \forall (h,k) \in \mathbb{R}^d \times \mathbb{R}^d\,,$$

et d'après le théorème de Schwarz

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}\,,$$

de sorte que

$$\mathrm{d}^2 f(a)(h,h) = \sum_{i=1}^{d} h_i^2 \frac{\partial^2 f}{\partial x_i^2}(a) \,+\, 2\sum_{i<j} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,, \quad \forall h \in \mathbb{R}^d\,,$$

où le symbole $\sum_{i<j}$ représente la *somme double* sur tous les indices $i, j \in \{1, \dots, d\}$ tels que $i < j$, et

$$\frac{\partial^2 f}{\partial x_i^2} := \frac{\partial^2 f}{\partial x_i \partial x_i}\,.$$

Si f est scalaire ($p = 1$) on définit $\mathrm{Hess} f(a)$ la *Hessienne* de f au point a comme la matrice (symétrique) de coefficients $\frac{\partial^2 f}{\partial x_i \partial x_j}$.
Tant qu'elles existent on définit les *dérivées partielles d'ordre n* par récurrence : si $j_1, \dots j_n \in \{1, \dots, d\}$ sont tels que la fonction $\dfrac{\partial^n f}{\partial x_{j_n} \dots \partial x_{j_1}}$ admet une dérivée partielle par rapport à $j_{n+1} \in \{1, \dots, d\}$, on note $\dfrac{\partial^{n+1} f}{\partial x_{j_{n+1}} x_{j_n} \dots \partial x_{j_1}}$ cette dérivée partielle. D'après le théorème de Schwarz on a

$$\frac{\partial^n f}{\partial x_{j_n} \dots \partial x_{j_1}} = \frac{\partial^n f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}\,,$$

où pour tout $i \in \{1, \dots, d\}$, α_i est le nombre d'entiers j_k valant i. Noter que $\alpha_1 + \cdots + \alpha_d = n$ puisqu'il y a n entiers j_k. Lorsque f admet des dérivées partielles jusqu'à l'ordre n continues, on dit qu'elle est de *classe $\mathscr{C}^n$*. Si c'est le cas on note

$$\mathrm{d}^n f(x)(h, \dots, h) = \sum_{\alpha_1 + \cdots + \alpha_d = n} \frac{n!}{\alpha_1! \cdots \alpha_d!}\, h_1^{\alpha_1} \cdots h_d^{\alpha_d} \frac{\partial^n f}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}\,.$$

Ci-dessus le symbole $\sum_{\alpha_1 + \cdots + \alpha_d = n}$ représente la somme sur tous les d-uplets $(\alpha_1, \dots, \alpha_d)$ d'entiers naturels α_i tels que $\alpha_1 + \cdots + \alpha_d = n$ (noter que ces entiers sont forcément inférieurs à n). On a

$$\mathrm{d}^n f(x)(h, \dots, h) = \frac{\mathrm{d}^n}{\mathrm{d}t^n}(f(x + th))_{|t=0}\,.$$

Formules de Taylor

Formule de Taylor avec reste intégral : Si $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est de classe $\mathscr{C}^{n+1}$ sur l'ouvert *convexe* [a] U, alors quel soit $x \in U$ et quel que soit $h \in \mathbb{R}^d$ tel que $x + h \in U$,

$$f(x+h) = f(x) + \sum_{p=1}^{n} \frac{1}{p\,!}\, \mathrm{d}^p f(x)(h, \dots, h) + \int_0^1 \frac{(1-t)^n}{n\,!}\, \mathrm{d}^{n+1} f(x + th)(h, \dots, h)\, \mathrm{d}t\,.$$

a. La convexité de U signifie que, quels que soient x et y dans U, le *segment* $[x, y] := \{z = \theta x + (1-\theta) y\,;\ \theta \in [0,1]\}$ est inclus dans U.

Formule de Taylor-Young : Si $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est de classe $\mathscr{C}^n$ au voisinage de $x \in U$ alors

$$f(x+h) \;=\; f(x) \;+\; \sum_{p=1}^{n} \frac{1}{p\,!}\, \mathrm{d}^p f(x)(h,\ldots,h) \;+\; o(\|h\|^n)$$

lorsque h tend vers zéro.

Extrema On dit qu'une fonction $f : D \subset \mathbb{R}^d \to \mathbb{R}$ a un *minimum global* (resp. *maximum global*) en $a \in D$ si $f(x) \geq f(a)$ (resp. $f(x) \leq f(a)$) quel que soit $x \in D$. Elle a un *minimum local* (resp. *maximum local*) en $a \in D$ s'il existe un voisinage V de a tel que $f(x) \geq f(a)$ (resp. $f(x) \leq f(a)$) quel que soit $x \in D \cap V$. Un *extremum* (global ou local) est soit un minimum soit un maximum (global ou local). Il est dit *strict* si l'inégalité ($f(x) \geq f(a)$ pour un minimum, resp. $f(x) \leq f(a)$ pour un maximum) est stricte quel que soit $x \neq a$ (c'est-à-dire que $f(x) > f(a)$ pour un minimum strict, resp. $f(x) < f(a)$ pour un maximum strict).

Extrema sur un compact Si K est *compact* et $f : K \subset \mathbb{R}^d \to \mathbb{R}$ est continue, alors elle admet un minimum global et un maximum global.

Extrema sur un ouvert Soit U un ouvert de $\mathbb{R}^d$ et $f : U \to \mathbb{R}^d$ différentiable.

- *Condition nécessaire* : si f a un extremum local en $a \in U$ alors $\mathrm{d}f(a) = 0$.
- *Condition suffisante* : si f est deux fois différentiable, $\mathrm{d}f(a) = 0$, et il existe $\alpha > 0$ tel que $\mathrm{d}f^2(a)(h,h) \geq \alpha\|h\|^2$ quel que soit $h \in \mathbb{R}^d$, alors f admet un minimum local strict en a.

La condition $(\exists\alpha > 0\,;\ \forall h \in \mathbb{R}^d\,,\ \mathrm{d}f^2(a)(h,h) \geq \alpha\|h\|^2)$ est satisfaite dès que $\boxed{\forall h \in \mathbb{R}^d\backslash\{0\}\,,\ \mathrm{d}f^2(a)(h,h) \;>\; 0}$.
Cette condition revient, si $d = 2$ et seulement dans ce cas, à $\boxed{\mathrm{tr}(\mathrm{Hess} f(a)) > 0\,,\ \mathrm{dét}(\mathrm{Hess} f(a)) > 0}$.
Attention, en cas d'inégalité large ($\mathrm{d}f^2(a)(h,h) \geq 0$), on ne peut pas conclure.

Extrema liés (optimisation sous contrainte d'égalités)

Théorème des multiplicateurs de Lagrange : Soit U un ouvert de $\mathbb{R}^d$, $f : U \to \mathbb{R}^d$ et $g : U \to \mathbb{R}^p$ de classe $\mathscr{C}^1$. On suppose que $a \in U$ est tel que $g(a) = 0$, et $(\mathrm{d}g_1(a), \ldots, \mathrm{d}g_p(a))$ forme une *famille libre* de $\mathscr{L}(\mathbb{R}^d; \mathbb{R})$ (où $g_1, \ldots, g_p$ désignent les composantes de g). Si la fonction f admet un minimum local en a sous la contrainte $g(x) = 0$ (c'est-à-dire que la *restriction* de f à $D := \{x \in U\,;\ g(x) = 0\}$ admet un minimum local en a), alors il existe des nombres réels $\lambda_1, \ldots, \lambda_p$ tels que $\mathrm{d}f(a) = \sum_{i=1}^{p} \lambda_i\, \mathrm{d}g_i(a)$.