Dans ce qui suit, d et p sont des entiers naturels non nuls. Une fonction

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

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

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

pour $i \in \{1, \ldots, p\}$. À tout élément $a = (a_1, \ldots, a_d)$ de D, à tout entier $j \in \{1, \ldots, d\}$, on associe une *application partielle* de f, définie comme la fonction d'une variable

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

où

$$D_{a,j} := \{t \in \mathbb{R}\,;\; (a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_d) \in D\}\,.$$

(Dans les cas $j = 1$ et $j = d$, les écritures ci-dessus sont à remplacer respectivement par $(t, a_2, \ldots, a_d)$ et $(a_1, \ldots, a_{d-1}, t)$.) Noter que $D_{a,j}$ contient a_j, et que $D_{a,j}$ est ouvert si D est ouvert.

Fonctions continues Une fonction $f : D \subset \mathbb{R}^d \to \mathbb{R}^p$ est *continue* en $a \in D$ si

$$\boxed{\forall \varepsilon > 0\,,\ \exists \eta > 0\,; \forall x \in D\,,\ \|x - a\|_{\mathbb{R}^d} \leqslant \eta \implies \|f(x) - f(a)\|_{\mathbb{R}^p} \leqslant \varepsilon}$$

Cette définition ne dépend pas des normes choisies sur $\mathbb{R}^d$ et sur $\mathbb{R}^p$. La *somme* de deux fonctions continues est continue. Le *produit* d'une fonction continue par une fonction scalaire continue est continu. La *composée* de deux fonctions continues est continue.
Une fonction $f : D \subset \mathbb{R}^d \to \mathbb{R}^p$ est continue en $a \in D$ si et seulement si, pour tout suite $(x_n)_{n\in\mathbb{N}} \in D^{\mathbb{N}}$ convergeant vers a, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$.
Une fonction vectorielle est continue si et seulement si ses composantes le sont. Si une fonction de plusieurs variables est continue alors toutes ses applications partielles sont continues. Cette condition nécessaire n'est pas suffisante.
Toute *norme* sur $\mathbb{R}^d$ définit une fonction scalaire continue. Toute application linéaire $\mathbb{R}^d \to \mathbb{R}^p$ est continue.

Dérivées partielles Désormais, on considère une fonction f définie sur un ouvert U de $\mathbb{R}^d$ à valeurs dans $\mathbb{R}^p$. Si l'application partielle

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

est *dérivable* en a_j, on note $\partial_j f(a)$ sa dérivée en a_j, appelée *dérivée partielle* de f par rapport à sa j-ème variable au point a. On note souvent aussi

$$\partial_j f(a) = \frac{\partial f}{\partial x_j}(a)\,.$$

Si f admet des dérivées partielles par rapport à toutes ses variables 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

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

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

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

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_d f(a) \end{pmatrix} = {}^t(J_f(a)).$$

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

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

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

$$\nabla \wedge f(a) = \begin{pmatrix} \partial_2 f_3(a) - \partial_3 f_2(a) \\ \partial_3 f_1(a) - \partial_1 f_3(a) \\ \partial_1 f_2(a) - \partial_2 f_1(a) \end{pmatrix}.$$

Dérivées directionnelles Étant donnés $a \in U$ et $h \in \mathbb{R}^d$, si la fonction d'une variable $\theta \mapsto f(a+\theta h)$ est dérivable en zéro, on note

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

et on dit que f admet $\mathrm{D}_h f(a)$ pour *dérivée suivant le vecteur* h au point a. En particulier, si $(\mathrm{e}_1, \dots, \mathrm{e}_d)$ désigne la base « canonique » de $\mathbb{R}^d$, f admet une dérivée partielle par rapport à sa j-ème variable si et seulement si elle admet une dérivée suivant le vecteur e_j, et si c'est le cas :

$$\partial_j f(a) = \mathrm{D}_{\mathrm{e}_j} f(a).$$

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 $\mathrm{d}f(x)$. 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.
Si f est différentiable en a alors elle est continue en a, elle admet une dérivée suivant tout vecteur h au point a et

$$\boxed{\forall h \in \mathbb{R}^d,\ \mathrm{d}f(a)(h) = \mathrm{D}_h f(a)}.$$

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

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

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 $\mathbb{R}^p$.

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 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 *forme bilinéaire* $\phi : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ 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* $(x,y) \mapsto \langle x \cdot y\rangle$.
Si $d = m$, l'application $q : x \in \mathbb{R}^d \mapsto q(x) := \phi(x,x)$ est différentiable et

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

En particulier, la *forme quadratique* $q : x \in \mathbb{R}^d \mapsto q(x) := \langle x \cdot x\rangle$ est différentiable et

$$\forall x \in \mathbb{R}^d\,,\ \forall h \in \mathbb{R}^d\,, \quad \mathrm{d}q(x)(h) = 2\langle x \cdot h\rangle\,.$$

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

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 $\boxed{J_{g\circ f}(x) = J_g(f(x))\,J_f(x)}$.

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

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

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

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

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 $(x_i)_{0\leqslant i\leqslant n}$ est une *subdivision* de $[a,b]$ et $y_i \in [x_{i-1},x_i]$ pour tout $i \in \{1,\dots,n\}$. (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

$$\int_a^b f(t)\mathrm{d}t = \int_a^c f(t)\mathrm{d}t + \int_c^b f(t)\mathrm{d}t \text{ (\textit{relation de Chasles})}.$$

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

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

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

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

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 I et soient $a, b \in I$. S'il existe $M \in \mathbb{R}^+$ tel que

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

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

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

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

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

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

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. (En général, on munit $\mathscr{L}(\mathbb{R}^d;\mathbb{R}^p)$ de la *norme subordonnée* $|||\cdot|||$ à celles de $\mathbb{R}^d$ et $\mathbb{R}^p$. Cependant, comme $\mathscr{L}(\mathbb{R}^d;\mathbb{R}^p)$ est un espace vectoriel normé de dimension finie, cette norme est *équivalente* à toutes les autres.) Une fonction de classe $\mathscr{C}^1$ admet des dérivées partielles continues par rapport à toutes ses variables. On démontre la réciproque grâce au théorème des accroissements finis : si $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ admet des dérivées partielles continues par rapport à toutes ses variables, $\partial_i f : \ U \ \to \ \mathbb{R}^p \ \ (i \in \{1, \dots d\})$, alors elle est de classe $\mathscr{C}^1$.

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 : $\exists \alpha \in]0,1[\,,\ \forall x, y \in F\,,\ \|f(x) - f(y)\| \leqslant \alpha \|x - y\|$) alors il existe un unique $x \in F$ tel que $f(x) = x$.
(Ce théorème est faux pour les fonctions *Lipschitziennes* de rapport $k \geqslant 1$.)

Difféomorphismes Une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ est un *difféomorphisme* de l'ouvert U sur l'ouvert $V \subset \mathbb{R}^p$ 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 $y \in V$.
Un isomorphisme de $\mathbb{R}^d$ est un $\mathscr{C}^1$-difféomorphisme de $\mathbb{R}^d$ sur $\mathbb{R}^d$.
La fonction tan est un $\mathscr{C}^1$-difféomorphisme de $]-\pi/2, \pi/2[$ sur $\mathbb{R}$, de réciproque arctan. (La fonction $x \in \mathbb{R} \mapsto x^3$ n'est pas un difféomorphisme.)
COORDONNÉES POLAIRES La fonction suivante est un $\mathscr{C}^1$-difféomorphisme :

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

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 $\mathbb{R}^d$ (c'est-à-dire $J_f(a) \in \mathrm{GL}_d(\mathbb{R})$). Alors il existe un ouvert $U_a \subset U$ contenant a et un ouvert V_b contenant $b = f(a)$ tels que la restriction $f_{|U_a}$ de f à U_a soit un $\mathscr{C}^1$-difféomorphisme de U_a sur V_b.

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 $\mathbb{R}^p$, V un ouvert de $\mathbb{R}^q$, $U = W \times V \subset \mathbb{R}^{p+q}$, et $f : U \to \mathbb{R}^q$ de classe $\mathscr{C}^1$. Soit $(a,b) \in U$ tel que $f(a,b) = 0_{\mathbb{R}^q}$. On suppose l'application partielle

$$\begin{array}{rccl} g: & V & \to & \mathbb{R}^q \\ & y & \mapsto & f(a,y) \end{array}$$

telle que $\mathrm{d}g(b)$ soit un isomorphisme de $\mathbb{R}^q$ (c'est-à-dire $J_g(b) \in \mathrm{GL}_q(\mathbb{R})$). Alors il existe un ouvert $U_{(a,b)} \subset U$ contenant (a,b), un ouvert $W_a \subset W$ contenant a et une fonction $\varphi : W_a \to \mathbb{R}^q$ de classe $\mathscr{C}^1$ tels que

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

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

De plus, si l'on note $\mathrm{d}_2 f(x,y)$ la différentielle au point y de l'application partielle

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

il existe un ouvert $\widetilde{U}_{(a,b)} \subset U_{(a,b)}$ contenant (a,b) tel que pour tout $(x,y) \in \widetilde{U}_{(a,b)}$, $\mathrm{d}_2 f(x,y)$ soit un isomorphisme de $\mathbb{R}^q$. Alors, pour tout $x \in \widetilde{W}_a := \{x \in W_a\,;\ (x,\varphi(x)) \in \widetilde{U}_{(a,b)}\}$,

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

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

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

Formules de Taylor pour les fonctions vectorielles d'une variable Soit $g : I \subset \mathbb{R} \to \mathbb{R}^p$, où I est un intervalle de $\mathbb{R}$.

Formule de Taylor avec reste intégral : On suppose g de classe $\mathscr{C}^{n+1}$. Quels que soient a, $b \in I$,

$$g(b) \;=\; g(a) \;+\; \sum_{k=1}^{n} \frac{g^{(k)}(a)}{k\,!}\,(b-a)^k \;+\; \int_0^1 \frac{(1-t)^n}{n\,!}\; g^{(n+1)}(a+t(b-a))\,(b-a)^{n+1}\,\mathrm{d}t\,.$$

Formule de Taylor-Lagrange : On suppose que g est $(n+1)$ dérivable et il existe $M \geqslant 0$ tel que $\|g^{(n+1)}(t)\| \leqslant M$ pour tout $t \in I$. Quels que soient a, $b \in I$,

$$\left\| g(b) \;-\; g(a) \;-\; \sum_{k=1}^{n} \frac{g^{(k)}(a)}{k\,!}\,(b-a)^k \right\| \leqslant \frac{M}{(n+1)\,!}\,|b-a|^{n+1}\,.$$

Formule de Taylor-Young : On suppose que g est n fois dérivable en $a \in I$. Alors

$$g(a+h) \underset{h\to 0}{=} g(a) \;+\; \sum_{k=1}^{n} \frac{g^{(k)}(a)}{k\,!}\,h^k \;+\; o(h^n)\,.$$

La formule de Taylor avec reste intégral se démontre en observant que, si $\varphi : [0,1] \to \mathbb{R}^p$ est $(n+1)$ fois dérivable, on a

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\varphi(t) \;+\; \sum_{p=1}^{n} \frac{(1-t)^k}{p\,!}\;\varphi^{(k)}(t)\Big) \;=\; \frac{(1-t)^n}{n\,!}\;\varphi^{(n+1)}(t)\,,\ t \in [0,1]\,.$$

Formules de Taylor pour les fonctions vectorielles de plusieurs variables Soit une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$, où U est un ouvert *convexe* de $\mathbb{R}^d$. Pour a, $b \in U$ fixés, lorsque la fonction d'une variable $t \mapsto f(a+t(b-a))$ est n fois dérivable (avec $n \in \mathbb{N}^*$), on note

$$\mathrm{d}^n f(x)(b-a)^{[n]} \;=\; \frac{\mathrm{d}^n}{\mathrm{d}t^n}(f(a+t(b-a)))_{|t=0}\,.$$

Dans ce qui suit, on suppose pour simplifier que la fonction d'une variable $t \mapsto f(a+t(b-a))$ est de classe $\mathscr{C}^{n+1}$.

Formule de Taylor avec reste intégral :

$$f(b) \;=\; f(a) \,+\, \sum_{k=1}^{n} \frac{1}{k\,!}\, \mathrm{d}^k f(a)(b-a)^{[k]} \,+\, \int_0^1 \frac{(1-t)^n}{n\,!}\, \mathrm{d}^{n+1} f(a+t(b-a))(b-a)^{[n+1]}\, \mathrm{d}t\,.$$

Formule de Taylor-Young :

$$f(a+h) \underset{h\to 0_{\mathbb{R}^d}}{=} f(a) \,+\, \sum_{k=1}^{n} \frac{1}{k\,!}\, \mathrm{d}^k f(a)(b-a)^{[k]} \,+\, o(\|h\|^n)\,.$$

On définit par récurrence le fait que f soit n fois différentiable, ou encore de classe $\mathscr{C}^n$: pour un entier $n \geqslant 2$, une fonction $f : U \subset \mathbb{R}^d \to \mathbb{R}^p$ (avec U ouvert)

- est *n fois différentiable* si elle est différentiable et si $\mathrm{d}f$ est $(n-1)$ fois différentiable ;
- est de *classe $\mathscr{C}^n$* si elle est différentiable et si $\mathrm{d}f$ est de classe $\mathscr{C}^{n-1}$.

En particulier, une fonction deux fois différentiable admet des *dérivées partielles secondes*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} \;:=\; \partial_i(\partial_j f)\,.$$

Lorsque $i = j$ on note

$$\frac{\partial^2 f}{\partial x_i^2} \;:=\; \partial_i(\partial_i f)\,.$$

Théorème de Schwarz : si f est deux fois différentiable, quels que soient i et $j \in \{1, \dots, d\}$,

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

Par suite,

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

On note encore $\mathrm{d}^2 f(a)(h,h) = \mathrm{d}^2 f(a) h^{[2]}$. Lorsque f est scalaire ($p = 1$), on définit $\mathrm{Hess} f(a)$ la matrice *hessienne* de f au point a comme la matrice (symétrique) de coefficients $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$. Ainsi, en notant H le vecteur colonne de composantes $(h_1, \dots, h_d)$, on a

$$\mathrm{d}^2 f(a)(h,h) \;=\; {}^{\mathrm{t}}H\; \mathrm{Hess} f(a)\; H\,.$$

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 $f(x) \geqslant f(a)$ (resp. $f(x) \leqslant f(a)$) 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 $f(x) \geqslant f(a)$ (resp. $f(x) \leqslant f(a)$) 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 ($f(x) \geqslant f(a)$ pour un minimum, resp. $f(x) \leqslant f(a)$ 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 $f : U \subset \mathbb{R}^d \to \mathbb{R}$ deux fois différentiable.

- *Conditions nécessaires* :
  - Si f a un extremum local en $a \in U$ alors $\mathrm{d}f(a) = 0$ (dans $\mathscr{L}(\mathbb{R}^d; \mathbb{R})$), ce qui équivaut à $\nabla f(a) = 0$ (dans $\mathbb{R}^d$). Dans ce cas, on dit que a est un *point critique* de f.
  - Si f a un minimum local en $a \in U$ alors $\mathrm{d}f(a) = 0$ *et* $\mathrm{d}f^2(a)(h,h) \geqslant 0$ pour tout $h \in \mathbb{R}^d$.
  - Si f a un maximum local en $a \in U$ alors $\mathrm{d}f(a) = 0$ *et* $\mathrm{d}f^2(a)(h,h) \leqslant 0$ pour tout $h \in \mathbb{R}^d$.
- *Conditions suffisantes* :
  - Si $\mathrm{d}f(a) = 0$ *et* $\mathrm{d}f^2(a)(h,h) > 0$ quel que soit $h \in \mathbb{R}^d \backslash \{0\}$, alors f admet un minimum local strict en a.
  - Si $\mathrm{d}f(a) = 0$ *et* $\mathrm{d}f^2(a)(h,h) < 0$ quel que soit $h \in \mathbb{R}^d \backslash \{0\}$, alors f admet un maximum local strict en a.

Nature des points critiques pour les fonctions de deux variables Soit U un ouvert de $\mathbb{R}^2$, $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ deux fois différentiable, et a un point critique de f (c'est-à-dire que $\nabla f(a) = 0$).

- Si $\mathrm{dét}(\mathrm{Hess} f(a)) > 0$ et $\mathrm{tr}(\mathrm{Hess} f(a)) > 0$, la fonction f admet un minimum local strict en a.
- Si $\mathrm{dét}(\mathrm{Hess} f(a)) > 0$ et $\mathrm{tr}(\mathrm{Hess} f(a)) < 0$, la fonction f admet un maximum local strict en a.
- Si $\mathrm{dét}(\mathrm{Hess} f(a)) < 0$, on dit que la fonction f admet un *point selle* en a.
- Si $\mathrm{dét}(\mathrm{Hess} f(a)) = 0$, la nature du point critique a est indéterminée.

Extrema liés (optimisation sous contrainte d'égalité)

Théorème du *multiplicateur de Lagrange*
Soit U un ouvert de $\mathbb{R}^d$, $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ de classe $\mathscr{C}^1$. On suppose que $a \in U$ est tel que $g(a) = 0$ et $\nabla g(a) \neq 0$. Si la fonction f admet un extremum 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 extremum local en a, alors il existe un réel λ tel que $\nabla f(a) = \lambda \nabla g(a)$.