Chapitre 17. Fonctions de plusieurs variables

Plan du chapitre

Nous allons étudier des fonctions ayant n variables réelles, $n \in \mathbb{N}^*$, à valeurs dans un $\mathbb{R}$-espace de dimension finie p non nulle. Le cas $n = p = 1$ est le programme de maths sup du premier semestre et le cas $n = 1$ et $p \geqslant 1$ est le chapitre 8 : « fonctions vectorielles ».
Le cas $n = 2$ et $p = 1$ (fonctions numériques de deux variables) a été étudié en fin de maths sup.
Le programme officiel de maths spé est ambitieux sur le sujet des fonctions de plusieurs variables, mais dans la pratique des problèmes de concours, assez souvent, seuls les cas $n = 2$ et $p = 1$ ou $n = 2$ et $p = 2$ apparaissent effectivement. Il s'agit donc au sortir de ce chapitre de maîtriser au moins les fonctions de deux variables.

1 Fonctions composantes, fonctions coordonnées, fonctions partielles

• Le modèle de base d'une fonction à n variables réelles et p composantes, $(n, p) \in (\mathbb{N}^*)^2$, est une fonction d'une partie D de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ de la forme

$$f : \begin{array}{ccc} D \subset \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ (x_1, \ldots, x_n) & \longmapsto & f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_p(x_1, \ldots, x_n)) \end{array}.$$

On doit faire attention à la notation avec une seule parenthèse : $f(x_1, \ldots, x_n)$, pour spécifier qu'il y a n variables. Si on met deux parenthèses : $f((x_1, \ldots, x_n))$, il y a une seule variable qui est un n-uplet.

Les fonctions f_1, ..., f_p, sont les **fonctions composantes** de la fonction f. Ce sont toujours des fonctions de n variables, mais à valeurs dans $\mathbb{R}$.

Par exemple, la fonction $f : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (r, \theta) & \mapsto & (r\cos\theta, r\sin\theta) \end{array}$ (passage des coordonnées polaires aux coordonnées cartésiennes) a deux fonctions composantes, les fonctions $f_1 : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (r, \theta) & \mapsto & r\cos\theta \end{array}$ et $f_2 : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (r, \theta) & \mapsto & r\sin\theta \end{array}$.

• On peut considérer plus généralement une fonction f d'une partie D d'un $\mathbb{R}$-espace E de dimension finie non nulle n dans un $\mathbb{R}$-espace vectoriel F de dimension finie non nulle p. Dans ce cas, pour décrire les valeurs de f, on peut utiliser une base $\mathscr{B} = (e_1, \ldots, e_p)$ de E et une base $\mathscr{B}'$:

$$\forall x \in D,\ f(x) = f_1(x)\, e'_1 + \ldots + f_p(x)\, e'_p.$$

Les fonctions f_1, ..., f_p, sont les **fonctions coordonnées dans la base** $\mathscr{B}$ de la fonction f. Ce sont des fonctions d'une partie de E vers $\mathbb{R}$. On note que les fonctions composantes d'une fonction f de $\mathbb{R}^n$ dans $\mathbb{R}^p$ ne sont autre que les fonctions coordonnées de f dans la base canonique de $\mathbb{R}^p$.

Si la base $\mathscr{B}$ de E est fixée, le programme officiel accepte l'abus de notation $f(x) = f(x_1, \ldots, x_n)$ à la place de $f(x) = f(x_1e_1 + \ldots + x_ne_n)$. De même, à l'arrivée, si la base $\mathscr{B}'$ de F est fixée, on peut écrire par abus de notation $f(x) = (f_1(x), \ldots, f_p(x))$ au lieu de $f(x) = f_1(x)e'_1 + \ldots + f_p(x)e'_p$.

Pour ces raisons, par la suite, nous ne détaillerons le plus souvent que le cas particulier des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$, fonctions à n variables et p composantes.

• La notion de fonctions composantes ou plus généralement de fonctions coordonnées dans une base est une première tentative de simplifier l'abord d'une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Mais le problème central demeure : il y a toujours n variables qui varient. L'idée qui vient est alors de n'en laisser varier qu'une seule dans un premier temps, en fixant les autres. Plus précisément, soit f une fonction d'une partie D de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et soit $a = (a_1, \ldots, a_n)$ un élément donné de D.

Pour $i \in [\![1, n]\!]$, la **i-ème application partielle** de f en a est la fonction d'une seule variable réelle

$$f_{x_i,a} : t \mapsto f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n).$$

En indice, on a placé la variable numéro i (qu'on laisse varier) et on a placé $a = (a_1, \ldots, a_n)$ pour indiquer comment les autres variables sont fixées.

Par exemple, soit $f : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x, y) & \mapsto & ye^{x^2+y^2} \end{array}$. La première application partielle de f en $(1, 0)$ est l'application définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R},\ f_{x,(1,0)}(x) = f(x, 0) = 0$$

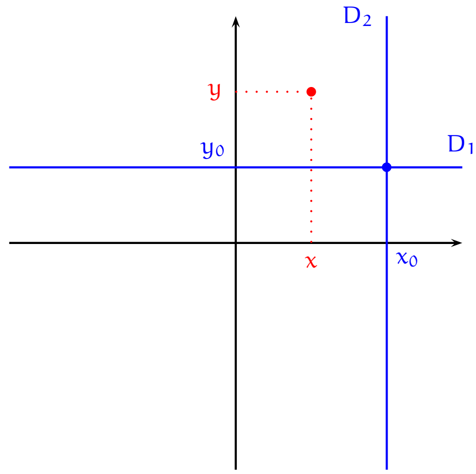
et la deuxième application partielle de f en $(1, 0)$ est l'application définie sur $\mathbb{R}$ par

$$\forall y \in \mathbb{R},\ f_{y,(1,0)}(y) = f(1, y) = ye^{y^2+1}.$$

Plus généralement, les deux applications partielles de f en $(x_0, y_0) \in \mathbb{R}^2$ sont les applications définies sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R},\ f_{x,(x_0,y_0)}(x) = f(x, y_0) = y_0 e^{x^2+y_0^2} \text{ et } \forall y \in \mathbb{R},\ f_{y,(x_0,y_0)}(y) = f(x_0, y) = y e^{x_0^2+y^2}.$$

On note que la seule connaissance des deux applications partielles en (x_0, y_0) ne suffit pas à reconstituer l'application f puisqu'on ne calcule jamais l'image par f d'un couple (x, y) tel que $x \neq x_0$ et $y \neq y_0$ c'est-à-dire l'image d'un couple (x, y) n'appartenant pas à $D_1 \cup D_2$:



2 Limites. Continuité

Le cours général sur les notions de limites et de continuité a été effectué dans le chapitre « Topologie ». Il s'agit maintenant de se concentrer sur l'aspect technique.

2.1 Limite en un point

Soit f une fonction définie sur une partie D de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$ et soit a un point adhérent à D. Deux situations se présentent : la fonction f a une limite en a ou la fonction f n'a pas de limite en a. Nous allons détailler ces deux situations d'un point de vue technique sur deux exemples pour des fonctions de deux variables à valeurs dans $\mathbb{R}$ (on rappelle que dans le cas général, f a une limite en a si et seulement si chacune de ses composantes a une limite en a et dans ce cas, $\lim\limits_{x\to a} f(x) = \left(\lim\limits_{x\to a} f_1(x), \ldots, \lim\limits_{x\to a} f_p(x)\right)$).

• On commence par un exemple où la fonction f a effectivement une limite. On met en œuvre les techniques usuelles de l'analyse comme par exemple l'utilisation d'inégalités.

Exemple. Pour $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, on pose $f(x, y) = \dfrac{x^2y^2}{x^2 + y^2}$. La fonction f est définie sur $D = \mathbb{R}^2 \setminus \{(0, 0)\}$ à valeurs dans $\mathbb{R}$ et $(0, 0)$ est adhérent à D. On a $\lim\limits_{(x,y)\to(0,0)} x^2y^2 = 0$ et $\lim\limits_{(x,y)\to(0,0)} \left(x^2 + y^2\right) = 0$. Il y a donc un problème.

Pour $(x, y) \in \mathbb{R}^2$, $(|x| - |y|)^2 \geqslant 0$ et donc $|xy| \leqslant \dfrac{1}{2}\left(x^2 + y^2\right)$ (on doit mémoriser l'inégalité précédente qui est fréquemment utilisée en pratique). Par suite, pour $(x, y) \in D$,

$$|f(x, y)| = |xy|\frac{|xy|}{x^2 + y^2} \leqslant \frac{1}{2}|xy|.$$

Maintenant, $\lim\limits_{(x,y)\to(0,0)} \dfrac{1}{2}|xy| = 0$ (car la fonction $g\ :\ (x, y) \mapsto \dfrac{1}{2}|xy|$ est continue sur $\mathbb{R}^2$ et donc en $(0, 0)$) et donc $\lim\limits_{(x,y)\to(0,0)} f(x, y) = 0$. La fonction f a une limite en $(0, 0)$ et cette limite est égale à 0.

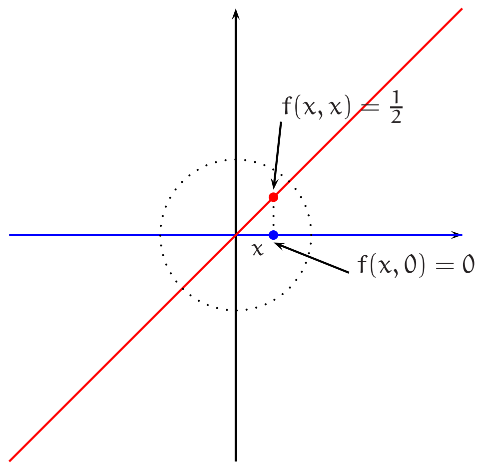
Ici, on a pu conclure grâce à une majoration simple de $|f(x, y)|$. Plus généralement, si f est une fonction de $D \subset \mathbb{R}^n$ dans $\mathbb{R}$, pour prouver que $f(x)$ tend vers ℓ quand x tend vers a (où x et a sont des éléments de $\mathbb{R}^n$ et $\ell \in \mathbb{R}$), on essaie de majorer $|f(x) - \ell|$ par une expression tendant vers 0 quand $x = (x_1, \ldots, x_n)$ tend vers $a = (a_1, \ldots, a_n)$. Si f est à valeurs dans $\mathbb{R}^p$, c'est $\|f(x) - \ell\|$ qu'on cherche à majorer (où $\|\ \|$ est une norme donnée sur $\mathbb{R}^p$).

• Passons maintenant à un exemple où la fonction f n'a pas de limite.

Dans ce cas, dans la quasi-totalité des situations pratiques, l'outil de base est la notion de limite suivant un sous-ensemble : si une fonction f a une limite $\ell = (\ell_1, \ldots, \ell_n)$ quand $x = (x_1, \ldots, x_n)$ tend vers $a = (a_1, \ldots, a_n)$ adhérent à D, alors f a encore pour limite ℓ quand x tend vers a en restant dans un sous-ensemble D_1 de D tel que a est adhérent à D_1.

Ce résultat est utilisé de la façon suivante : si on trouve deux sous-ensembles D_1 et D_2 de D tels que f ait une limite ℓ_1 quand x tend vers a en restant dans D_1 et une limite ℓ_2 quand x tend vers a en restant dans D_2 avec $\ell_1 \neq \ell_2$ ou si on trouve un sous-ensemble D_1 de D tel que f n'a pas de limite dans $\mathbb{R}^p$ quand x tend vers a en restant dans D_1, alors f n'a pas de limite en a.

Exemple. Pour $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$, on pose $f(x,y) = \dfrac{xy}{x^2+y^2}$. La fonction f est définie sur $D = \mathbb{R}^2 \setminus \{(0,0)\}$ et $(0,0)$ est adhérent à D.
Quand x tend vers 0, le couple $(x,0)$ tend vers le couple $(0,0)$. Pour $x \in \mathbb{R}\setminus\{0\}$, $f(x,0) = \dfrac{x \times 0}{x^2+0^2} = 0$. Donc, $\lim\limits_{x\to 0} f(x,0) = 0$ (si f a une limite quand (x,y) tend vers $(0,0)$, cette limite ne peut être que 0).
Quand x tend vers 0, le couple (x,x) tend vers le couple $(0,0)$. Pour $x \in \mathbb{R}\setminus\{0\}$, $f(x,x) = \dfrac{x \times x}{x^2+x^2} = \dfrac{1}{2}$. Donc, $\lim\limits_{x\to 0} f(x,x) = \dfrac{1}{2}$ (si f a une limite quand (x,y) tend vers $(0,0)$, cette limite ne peut être que $\dfrac{1}{2}$).
Puisque $\dfrac{1}{2} \neq 0$, la fonction f n'a pas de limite en $(0,0)$. Pour l'établir, on a constaté que dans tout voisinage de l'origine, il y a des couples (x,y) tels que $f(x,y) = 0$ et des couples (x,y) tels que $f(x,y) = \dfrac{1}{2}$. En particulier, en tendant vers $(0,0)$ en suivant des chemins différents (la droite d'équation $y = 0$ ou la droite d'équation $y = x$), on obtient des limites différentes ce qui empêche l'existence d'une limite en $(0,0)$.



2.2 Continuité. Continuité partielle

On redit que la notion de continuité a été étudiée dans le chapitre « Topologie ». De nouveau, on se concentre sur l'aspect technique.

Soit f la fonction définie sur $\mathbb{R}^2$ par :

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} \dfrac{x^2y^2}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) = 0 \end{cases}.$$

La fonction f est définie sur $\mathbb{R}^2$. De plus, pour $(x,y) \neq (0,0)$,

$$|f(x,y)| = \frac{x^2y^2}{x^2+y^2} = |xy|\frac{|xy|}{x^2+y^2} \leqslant \frac{|xy|}{2}.$$

Puisque $\lim\limits_{(x,y)\to(0,0)} \dfrac{1}{2}|xy| = 0$, on en déduit que $\lim\limits_{\substack{(x,y)\to(0,0)\\(x,y)\neq(0,0)}} f(x,y) = 0 = f(0,0)$ et donc que f est continue en $(0,0)$.

De manière générale, pour établir la continuité d'une fonction f en un point $a = (a_1, \ldots, a_n)$, ou bien on dispose d'un théorème général (combinaison linéaire, produit ...), ou bien on montre directement que $\lim\limits_{\substack{x\to a\\x\neq a}} f(x) = f(a)$.

Exercice 1. Soit f la fonction définie sur $\mathbb{R}^2$ par :

$$\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} \dfrac{xy\left(x^2-y^2\right)}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) = 0 \end{cases}.$$

Etudier la continuité de f sur $\mathbb{R}^2$.

Solution 1. La fonction f est continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$ en tant que quotient de fonctions continues sur $\mathbb{R}^2 \setminus \{(0,0)\}$ dont le dénominateur ne s'annule pas sur $\mathbb{R}^2 \setminus \{(0,0)\}$ (ou plus simplement en tant que fraction rationnelle définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$).

Pour $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$,

$$|f(x,y)| = \frac{|xy|\left|x^2-y^2\right|}{x^2+y^2} \leqslant \frac{|xy|\left(\left|x^2\right|+\left|y^2\right|\right)}{x^2+y^2} = \frac{|xy|\left(x^2+y^2\right)}{x^2+y^2} = |xy|.$$

Puisque $\lim\limits_{(x,y)\to(0,0)} |xy| = 0$, on en déduit que $\lim\limits_{\substack{(x,y)\to(0,0)\\(x,y)\neq(0,0)}} f(x,y) = 0 = f(0,0)$. Ceci montre que f est continue en $(0,0)$.

En résumé, f est continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$ et en $(0,0)$ et finalement, f est continue sur $\mathbb{R}^2$.

Considérons un autre exemple. Pour $(x,y) \in \mathbb{R}^2$, on pose $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) = 0 \end{cases}$. Les deux applications partielles de f en $(0,0)$ sont les fonctions définies sur $\mathbb{R}$ par : $\forall x \in \mathbb{R},\ f(x,0) = \begin{cases} 0 \text{ si } x \neq 0 \\ 0 \text{ si } x = 0 \end{cases} = 0$ et $\forall y \in \mathbb{R},\ f(0,y) = 0$.

Les deux applications partielles en $(0,0)$ sont nulles et donc continues sur $\mathbb{R}$. En particulier, les deux applications partielles en $(0,0)$ sont continues en 0 et 0 respectivement. Mais ceci ne suffit absolument pas pour affirmer que la fonction f est continue en $(0,0)$. De fait, on a vu plus que la fonction f n'a pas de limite en $(0,0)$ et en particulier, n'est pas continue en $(0,0)$. On a l'habitude de dire sur le sujet que

la continuité partielle n'entraine pas la continuité.

Ainsi, une solution du genre : « la fonction $(x,y) \mapsto \dfrac{xy}{x^2+y^2}$ est continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$ car à y fixé, la fonction de x est continue et à x fixé, la fonction de y est continue » est une solution totalement fausse.

Par contre, si une application f est continue en un point $a = (a_1, \ldots, a_n)$, alors bien sûr, pour tout $i \in [\![1,n]\!]$, la i-ème application partielle de f en a est continue en a_i. En résumé,

$$\begin{array}{c} \text{Continuité} \\ \Big\Downarrow \quad \not\Big\Uparrow \\ \text{Continuité partielle} \end{array}$$

3 Dérivées partielles d'ordre 1

3.1 Dérivées partielles d'ordre 1 en un point. Fonctions dérivées partielles d'ordre 1

3.1.1 Dérivées partielles d'ordre 1 en un point

DÉFINITION 1. Soit f une fonction définie sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Soit a un point de Ω. Soit $i \in [\![1, n]\!]$.

On dit que f admet en a une dérivée partielle par rapport à sa i-ème variable si et seulement si la i-ème application partielle de f en a est dérivable en a_i. Dans ce cas, la dérivée de la i-ème application partielle de f en a s'appelle la dérivée partielle de la fonction f par rapport à sa i-ème variable en a et se note $\dfrac{\partial f}{\partial x_i}(a)$ (ou aussi $\partial_i f(a)$) :

$$\forall i \in [\![1, n]\!],\ \frac{\partial f}{\partial x_i}(a) = \lim_{t \to a_i} \frac{1}{t - a_i} \left(f\left(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n\right) - f\left(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n\right)\right),$$

ou aussi

$$\forall i \in [\![1, n]\!],\ \frac{\partial f}{\partial x_i}(a) = \lim_{h \to 0} \frac{1}{h} \left(f\left(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n\right) - f\left(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n\right)\right).$$

⇛ **Commentaire**.

◇ *∂ est le « d rond » par opposition au « d droit ». Si on note simplement f_i la i-ème application partielle de f en a, alors*

$$\frac{\partial f}{\partial x_i}(a) = f_i'\left(a_i\right) = \frac{df_i}{dt}\left(a_i\right).$$

Ainsi, dériver partiellement une fonction de plusieurs variables en un point de $\mathbb{R}^n$ s'effectue en dérivant une certaine fonction d'une seule variable en un réel.

◇ *Dans la définition, f est définie sur un ouvert non vide Ω de $\mathbb{R}^n$. Ceci assure le fait que si a est un point du domaine de définition de f, f est bien définie sur un voisinage de a. Mais bien sûr, on peut élargir la définition dans certains cas à des ensembles non ouverts en s'adaptant.* ❏

◇ *La définition 1 se place dans le cadre restreint des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Si plus généralement, f est une application d'une partie d'un $\mathbb{R}$-espace E de dimension finie non nulle n dans un $\mathbb{R}$-espace F de dimension finie non nulle p, la notion de dérivée partielle dépend du choix d'une base de E. Plus précisément, soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base de E. Un élément x de E s'écrit de manière unique sous la forme $x = x_1e_1 + \ldots + x_ne_n$ où $(x_1, \ldots, x_n) \in \mathbb{R}^n$. La dérivée partielle de f en $a = a_1e_1 + \ldots + a_ne_n$ par rapport à la i-ème coordonnée dans la base $\mathscr{B}$ qui s'appelle encore la **i-ème dérivée partielle dans la base $\mathscr{B}$ de f en a** est*

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \to a_i} \frac{1}{t - a_i} \left(f\left(a_1e_1 + \ldots + te_i + \ldots + a_ne_n\right) - f\left(a_1e_1 + \ldots + a_ie_i + \ldots + a_ne_n\right)\right),$$

ou plus simplement

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \to a_i} \frac{1}{t - a_i} \left(f\left(a + (t - a)e_i\right) - f(a)\right) = \lim_{h \to 0} \frac{1}{h} \left(f\left(a + he_i\right) - f(a)\right).$$

Si la base $\mathscr{B}$ est fixée, on peut identifier $f\left(x_1e_1 + \ldots + x_ne_n\right)$ et $f\left(x_1, \ldots, x_n\right)$ et utiliser la définition 1.

Analysons maintenant le lien entre l'existence de dérivées partielles et la continuité. Reprenons l'exemple de la fonction f définie sur $\mathbb{R}^2$ par : $\forall (x, y) \in \mathbb{R}^2$, $f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \\ 0 \text{ si } (x, y) = 0 \end{cases}$. On a déjà vu que les deux applications partielles de f en $(0, 0)$ étaient continues en 0 et 0 respectivement mais que f n'était pas continue en $(0, 0)$. Etudions maintenant l'existence de dérivées partielles en $(0, 0)$.

Pour $x \neq 0$, $\dfrac{f(x, 0) - f(0, 0)}{x - 0} = \dfrac{0 - 0}{x - 0} = 0$ et donc $\lim\limits_{x \to 0} \dfrac{f(x, 0) - f(0, 0)}{x - 0} = 0$. f admet donc une dérivée partielle par rapport à sa première variable en $(0, 0)$ et $\dfrac{\partial f}{\partial x}(0, 0) = 0$. De même, f admet une dérivée partielle par rapport à sa deuxième variable en $(0, 0)$ et $\dfrac{\partial f}{\partial y}(0, 0) = 0$.

On note que f n'est pas continue en $(0, 0)$ et donc que

l'existence de dérivées partielles en un point n'entraine pas la continuité en ce point.

Par contre, on sait qu'une fonction d'une seule variable, dérivable en un certain réel, est automatiquement continue en ce réel et donc l'existence de dérivées partielles entraine la continuité partielle. On peut résumer ces implications par le graphique ci-dessous qui va se compléter au fur et à mesure du chapitre (toute implication non écrite étant fausse) :

$$\begin{tikzcd}
\text{Existence de dérivées partielles} \arrow[dr, Rightarrow] & & \text{Continuité} \arrow[dl, Rightarrow] \\
& \text{Continuité partielle} &
\end{tikzcd}$$

❏

Les théorèmes généraux sur les fonctions d'une seule variable fournissent immédiatement

Théorème 1. Soient f et g deux fonctions définies sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Soient $a \in \Omega$ et $i \in [\![1, n]\!]$.

Si f et g admettent une i-ème dérivée partielle en a, alors pour tout $(\lambda, \mu) \in \mathbb{K}^2$, $\lambda f + \mu g$ admet une i-ème dérivée partielle en a et

$$\frac{\partial(\lambda f + \mu g)}{\partial x_i}(a) = \lambda \frac{\partial f}{\partial x_i}(a) + \mu \frac{\partial g}{\partial x_i}(a).$$

Théorème 2. Soient f et g deux fonctions définies sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Soient $a \in \Omega$ et $i \in [\![1, n]\!]$.

Si f et g admettent une i-ème dérivée partielle en a, alors $f \times g$ admet une i-ème dérivée partielle en a et

$$\frac{\partial(f \times g)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) \times g(a) + f(a) \times \frac{\partial g}{\partial x_i}(a).$$

Théorème 3. Soient f et g deux fonctions définies sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Soient $a \in \Omega$ et $i \in [\![1, n]\!]$.

Si f et g admettent une i-ème dérivée partielle en a et si $g(a) \neq 0$, alors $\dfrac{f}{g}$ admet une i-ème dérivée partielle en a et

$$\frac{\partial \left(\dfrac{f}{g}\right)}{\partial x_i}(a) = \frac{\dfrac{\partial f}{\partial x_i}(a) \times g(a) - f(a) \times \dfrac{\partial g}{\partial x_i}(a)}{(g(a))^2}.$$

3.1.2 Fonctions dérivées partielles d'ordre 1

Définition 2. Soit f une fonction définie sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$.

Pour $i \in [\![1, n]\!]$, f admet une i-ème dérivée partielle sur Ω si et seulement si f admet une i-ème dérivée partielle en chaque point a de Ω.

Dans ce cas, on peut définir la i-ème fonction dérivée partielle sur Ω notée $\dfrac{\partial f}{\partial x_i}$: c'est une fonction de n variables, définie sur Ω à valeurs dans $\mathbb{R}^p$.

Les théorèmes locaux 1, 2 et 3 du paragraphe précédent ont aussi leurs versions globales :

Théorème 4. Soient f et g deux fonctions définies sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans dans $\mathbb{R}^p$. Soit $i \in [\![1, n]\!]$.

Si f et g admettent une i-ème dérivée partielle sur Ω, alors pour tout $(\lambda, \mu) \in \mathbb{K}^2$, $\lambda f + \mu g$ admet une i-ème dérivée partielle sur Ω et

$$\frac{\partial(\lambda f + \mu g)}{\partial x_i} = \lambda \frac{\partial f}{\partial x_i} + \mu \frac{\partial g}{\partial x_i}.$$

Théorème 5. Soient f et g deux fonctions définies sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Soiit $i \in [\![1, n]\!]$.

Si f et g admettent une i-ème dérivée partielle sur Ω, alors $f \times g$ admet une i-ème dérivée partielle sur Ω et

$$\frac{\partial(f \times g)}{\partial x_i} = \frac{\partial f}{\partial x_i} \times g + f \times \frac{\partial g}{\partial x_i}.$$

Théorème 6. Soient f et g deux fonctions définies sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}$. Soit $i \in [\![1, n]\!]$.

Si f et g admettent une i-ème dérivée partielle sur Ω et si g ne s'annule pas sur Ω, alors $\dfrac{f}{g}$ admet une i-ème dérivée partielle sur Ω et

$$\frac{\partial\left(\dfrac{f}{g}\right)}{\partial x_i} = \frac{\dfrac{\partial f}{\partial x_i} \times g - f \times \dfrac{\partial g}{\partial x_i}}{g^2}.$$

Par exemple, si pour tout (x, y) de $\mathbb{R}^2$, $f(x, y) = xe^{x^2+y^2}$, alors pour tout $(x, y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x, y) = e^{x^2+y^2} + x \times 2xe^{x^2+y^2} = \left(2x^2 + 1\right) e^{x^2+y^2}$$

et

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

Dans la notation $\dfrac{\partial f}{\partial x}(x, y)$, il est important de comprendre que la lettre x utilisée à deux endroits différents de l'expression ne désigne pas la même chose. Dans le quotient $\dfrac{\partial f}{\partial x}$, la lettre x désigne la variable par rapport à laquelle on a dérivé alors que dans (x, y), la lettre x, entre autres, désigne ce en quoi on a évalué. Si on évalue en (x_0, y_0), on obtient $\dfrac{\partial f}{\partial x}(x_0, y_0)$ et non pas $\dfrac{\partial f}{\partial x_0}(x_0, y_0)$.

Exercice 2. Soit f la fonction définie sur $\mathbb{R}^2$ par :

$$\forall (x, y) \in \mathbb{R}^2,\ f(x, y) = \begin{cases} \dfrac{xy\left(x^2 - y^2\right)}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \\ 0 \text{ si } (x, y) = 0 \end{cases}.$$

Etudier l'existence des dérivées partielles de f sur $\mathbb{R}^2$ et les déterminer.

Solution 2. On a déjà vu dans l'exercice 1, page 5, que f est continue sur $\mathbb{R}^2$.

Dérivées partielles sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

f admet sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ une dérivée partielle par rapport à sa 1ère variable x en tant que quotient de fonctions admettant une dérivée partielle par rapport à leur 1ère variable x dont le dénominateur ne s'annule pas sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. De plus, pour $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$,

$$\frac{\partial f}{\partial x}(x, y) = y\frac{\left(3x^2 - y^2\right)\left(x^2 + y^2\right) - \left(x^3 - xy^2\right)(2x)}{\left(x^2 + y^2\right)^2} = \frac{y\left(x^4 + 4x^2y^2 - y^4\right)}{\left(x^2 + y^2\right)^2}.$$

De même, pour $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$,

$$\frac{\partial f}{\partial y}(x, y) = \frac{x\left(x^4 - 4x^2y^2 - y^4\right)}{\left(x^2 + y^2\right)^2},$$

(obtenu en échangeant les rôles de x et y et en changeant de signe car $f(x, y) = -\dfrac{yx\left(y^2 - x^2\right)}{y^2 + x^2} = -f(y, x)$).

Dérivées partielles en $(0, 0)$. Pour $x \neq 0$,

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

et donc $\displaystyle\lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = 0$. Par suite, f admet une dérivée partielle par rapport à sa première variable en $(0, 0)$ et $\dfrac{\partial f}{\partial x}(0, 0) = 0$. De même, f admet une dérivée partielle par rapport à sa deuxième variable en $(0, 0)$ et $\dfrac{\partial f}{\partial y}(0, 0) = 0$.

En résumé, f admet sur $\mathbb{R}^2$ des dérivées partielles par rapport à chacune de ses deux variables et de plus

$$\forall (x,y) \in \mathbb{R}^2,\ \frac{\partial f}{\partial x}(x,y) = \begin{cases} \dfrac{y\left(x^4+4x^2y^2-y^4\right)}{\left(x^2+y^2\right)^2} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) = (0,0) \end{cases}$$

et

$$\forall (x,y) \in \mathbb{R}^2,\ \frac{\partial f}{\partial y}(x,y) = \begin{cases} \dfrac{x\left(x^4-4x^2y^2-y^4\right)}{\left(x^2+y^2\right)^2} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) = (0,0) \end{cases}.$$

3.2 Dérivée suivant un vecteur

Pour analyser l'existence et la valeur de la i-ème dérivée partielle en $a = (a_1, \dots, a_n)$, on s'est intéressé à

$$\lim_{t \to a_i} \frac{1}{t - a_i} \left(f\left(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n\right) - f\left(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n\right)\right)$$

qui peut aussi s'écrire

$$\lim_{t \to 0} \frac{1}{t} \left(f\left(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n\right) - f\left(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n\right)\right).$$

En notant $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$, on a donc

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \to 0} \frac{1}{t} \left(f\left(a + te_i\right) - f(a)\right).$$

On dit alors qu'on a dérivé la fonction f en a suivant le vecteur e_i. On généralise cette notion :

DÉFINITION 3. Soit f une fonction définie sur un ouvert non vide Ω de $\mathbb{R}^n$ à valeurs dans $\mathbb{R}^p$. Soit a un point de Ω. Soit v un vecteur **non nul** de $\mathbb{R}^n$ donné.

f est **dérivable en a suivant le vecteur** v (ou dans la direction de v) si et seulement si la fonction d'une variable réelle $t \mapsto \dfrac{1}{t}(f(a+tv) - f(a))$ a une limite quand t tend vers 0 ce qui équivaut au fait que la fonction $t \mapsto f(a+tv)$ est dérivable en 0. Dans ce cas, cette limite s'appelle la **dérivée de f en a suivant le vecteur** v et se note $D_v f(a)$:

$$D_v f(a) = \lim_{t \to 0} \frac{1}{t}(f(a+tv) - f(a)).$$

v étant toujours fixé, la fonction $D_v f\ :\ a \mapsto D_v f(a)$ est la **fonction dérivée suivant le vecteur** v.

⇛ **Commentaire**. *En particulier, si $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$,*

$$\frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a).$$

Ainsi, les dérivées partielles premières de f en a ne sont autres que les dérivées de f en a suivant chacun des vecteurs de la base canonique de $\mathbb{R}^n$. Il est donc clair que si f est dérivable suivant tout vecteur en a, alors en particulier f admet des dérivées partielles par rapport à chacune de ses variables en a.

Fournissons maintenant un exemple de fonction admettant des dérivées partielles en un point sans être dérivable suivant tout vecteur en ce point et un exemple de fonction dérivable suivant tout vecteur en un point mais qui n'est pas continue en ce point.

Exemple 1. Soit f la fonction définie sur $\mathbb{R}^2$ par : $\forall (x,y) \in \mathbb{R}^2,\ f(x,y) = \begin{cases} 1 \text{ si } x \neq 0 \text{ et } y \neq 0 \\ 0 \text{ si } x = 0 \text{ ou } y = 0 \end{cases}$. Pour $x \neq 0$, $\dfrac{f(x,0) - f(0,0)}{x - 0} = 0$ puis $\displaystyle\lim_{x \to 0} \frac{f(x,0) - f(0,0)}{x - 0} = 0$. Donc, $\dfrac{\partial f}{\partial x}(0,0)$ existe et $\dfrac{\partial f}{\partial x}(0,0) = 0$. De même, $\dfrac{\partial f}{\partial y}(0,0)$ existe et $\dfrac{\partial f}{\partial y}(0,0) = 0$.

Soit $v = (1,1) \neq (0,0)$. Pour $t \neq 0$, $\dfrac{1}{t}(f(tv) - f(0)) = \dfrac{1}{t}f(t,t) = \dfrac{1}{t}$, expression qui n'a pas de limité quand t tend vers $0 = (0,0)$. f n'est donc pas dérivable en $(0,0)$ suivant le vecteur $v = (1,1)$.

Ainsi, f admet des dérivées partielles par rapport à chacune de ses deux variables en $(0,0)$ mais n'est pas dérivable suivant tout vecteur en $(0,0)$ et plus généralement n'est dérivable en aucun vecteur $v = (a,b)$ avec $a \neq 0$ et $b \neq 0$). ❏

Exemple 2. Soit f la fonction définie sur $\mathbb{R}^2$ par : $\forall(x,y)\in\mathbb{R}^2$, $f(x,y)=\begin{cases}1 \text{ si } y=x^2 \text{ et } x>0\\ 0 \text{ sinon}\end{cases}$.

On note que $f(x,y)\neq 0 \Rightarrow x>0$ et $y>0$.

Vérifions que f n'est pas continue en $(0,0)$. Quand x tend vers 0, $(x,0)$ tend vers $(0,0)$ et $f(x,0)=0$ tend vers 0. Quand x tend vers 0 en restant strictement positif, $\left(x,x^2\right)$ tend vers $(0,0)$ et $f\left(x,x^2\right)=1$ tend vers $1\neq 0$. Donc, f n'est pas continue en $(0,0)$.

Vérifions maintenant que f est dérivable suivant tout vecteur en $(0,0)$. Soit $v=(\alpha,\beta)\neq(0,0)$.

- Si $\alpha=0$ (et donc $\beta\neq 0$), pour $t\in\mathbb{R}\setminus\{0\}$,

$$\frac{1}{t}(f((0,0)+t(\alpha,\beta))-f(0,0))=\frac{1}{t}f(0,t\beta)=0 \text{ (car } t\beta\neq 0=0^2\text{)}$$

  et donc $\lim\limits_{t\to 0}\frac{1}{t}(f((0,0)+t(\alpha,\beta))-f(0,0))=0$.

- Si $\alpha\neq 0$ et $\alpha\beta\leqslant 0$, alors pour $t\neq 0$, $t\alpha$ et $t\beta$ ne sont pas tous deux strictement positifs et donc $\frac{1}{t}(f(((0,0)+t(\alpha,\beta))-f(0,0))=0$ puis $\lim\limits_{t\to 0}\frac{1}{t}(f((0,0)+t(\alpha,\beta))-f(0,0))=0$.
- Si par exemple, $\alpha>0$ et $\beta>0$, pour $t<0$, $\frac{1}{t}(f((0,0)+t(\alpha,\beta))-f(0,0))=0$ et pour $t>0$,

$$f((0,0)+t(\alpha,\beta))=1\Leftrightarrow t\beta=t^2\alpha^2\Leftrightarrow t=\frac{\beta}{\alpha^2}.$$

  Par suite, pour $t\in\left]0,\frac{\beta}{\alpha^2}\right[$, $\frac{1}{t}(f((0,0)+t(\alpha,\beta))-f(0,0))=0$. On en déduit encore une fois que $\lim\limits_{t\to 0}\frac{1}{t}(f((0,0)+t(\alpha,\beta))-f(0,0))=0$.

Ainsi, f est dérivable suivant tout vecteur en $(0,0)$ (et pour tout vecteur $v\neq 0$, $D_v(f)(0,0)=0$) mais f n'est pas continue en $(0,0)$. ❑

En résumé, (toute implication non écrite étant fausse)

$$\begin{tikzcd}
\text{Dérivabilité suivant tout vecteur en } a \arrow[d, Rightarrow] & & \\
\text{Existence de dérivées partielles en } a \arrow[dr, Rightarrow] & & \text{Continuité en } a \arrow[dl, Rightarrow] \\
& \text{Continuité partielle en } a &
\end{tikzcd}$$

❑

3.3 Fonction différentiable. Différentielle

3.3.1 Fonctions différentiables en un point. Différentielle d'une fonction en un point

Il s'agit maintenant d'approcher la différence finie (à opposer à différence infinitésimale) $f(a+h)-f(a)$ à l'ordre 1. Si f est une fonction définie sur un intervalle ouvert I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$ et a est un point de I, on sait déjà le faire quand f est dérivable en a :

$$f(a+h)=f(a)+hf'(a)+o(h)$$

où $\lim\limits_{h\to 0}\frac{o(h)}{h}=0$. L'application $L : h\mapsto hf'(a)$ est la meilleure approximation à l'ordre 1 de la différence $f(a+h)-f(a)$ quand h tend vers 0 ou encore l'application $L : h\mapsto hf'(a)$ est une **application linéaire** de $\mathbb{R}$ dans $\mathbb{R}$ telle que l'expression $f(a+h)-f(a)-L(h)$ soit négligeable devant h quand h tend vers 0.

On généralise :

Définition 4. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. On note $\|\ \|$ une norme donnée sur E.
Soit f une application d'un ouvert non vide Ω de E vers F. Soit $a \in \Omega$.

f est **différentiable en** a si et seulement si il existe une application linéaire L de E vers F telle que l'expression $\dfrac{1}{\|h\|}(f(a+h) - f(a) - L(h))$ tende vers 0 quand $h \in E$ tend vers 0.

Il revient au même de dire que f est **différentiable en** a si et seulement si il existe une application linéaire L de E vers F telle que, pour h au voisinage de $0 \in E$,

$$f(a+h) = f(a) + L(h) + o(h)$$

où $h \mapsto o(h)$ est une fonction de $h \in E$, définie sur un voisinage de 0, à valeurs dans F et vérifiant $\displaystyle\lim_{h\to 0} \frac{1}{\|h\|} o(h) = 0$.

Il revient au même de dire que f est **différentiable en** a si et seulement si il existe une application linéaire L de E vers F et une application ε définie sur un voisinage de $0 \in E$ telle que, pour h au voisinage de 0,

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h)$$

où ε est une fonction de $h \in E$, définie sur un voisinage de 0 et vérifiant $\displaystyle\lim_{h\to 0} \varepsilon(h) = 0$

Théorème 7. Si L existe, alors L est unique.

Démonstration. On suppose qu'il existe deux applications linéaires L_1 et L_2 de E vers F telles que

$$f(a+h) = f(a) + L_1(h) + o(h) = f(a) + L_2(h) + o(h).$$

Alors, $(L_1 - L_2)(h) = o(h)$ ou encore $\displaystyle\lim_{h\to 0} \frac{1}{\|h\|}(L_1 - L_2)(h) = 0$. Soit u un vecteur non nul de E.
Pour t réel strictement positif, on pose $h = tu$.

$$\frac{1}{\|h\|}(L_1 - L_2)(h) = \frac{1}{\|tu\|}(L_1 - L_2)(tu) = \frac{1}{\|u\|}(L_1 - L_2)(u).$$

Quand le réel t tend vers 0 par valeurs supérieures, le vecteur $h = tu$ tend vers $0 \in E$. On en déduit que

$$0 = \lim_{t\to 0} \frac{1}{\|t\|}(L_1 - L_2)(tu) = \frac{1}{\|u\|}(L_1 - L_2)(u)$$

et donc $(L_1 - L_2)(u) = 0$. Ainsi, pour tout vecteur non nul u, $L_1(u) = L_2(u)$ ce qui reste vrai quand $u = 0$ et donc $L_1 = L_2$.

❏

Définition 5. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle.
Soit f une application d'un ouvert non vide Ω de E vers F, différentiable en un élément $a \in \Omega$.

La **différentielle** de f en a est l'unique application linéaire L de E vers F telle que $f(a+h) \underset{h\to 0}{=} f(a) + L(h) + o(h)$.

On note $df(a)$ ou df_a la différentielle de f en a. df_a est un élément de $\mathscr{L}(E, F)$ vérifiant

$$f(a+h) \underset{h\to 0}{=} f(a) + df_a(h) + o(h).$$

La différentielle de f en a s'appelle aussi l'**application linéaire tangente** à f en a.

On a donc essentiellement deux notations pour la différentielle de f en a : df_a et $df(a)$. Si on évalue en h, on doit écrire $df_a(h)$ ou $(df(a))(h)$ ou $df(a).h$. Le programme officiel prévoit d'utiliser plutôt cette dernière notation : $df(a).h$. Le point signifie qu'on a évalué l'application linéaire $df(a)$ en h :

$$df(a).h = (df(a))(h) = df_a(h).$$

Exemple. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien. On note $\|\ \|$ la norme euclidienne associée au produit scalaire $\langle\ ,\ \rangle$. Pour tout x de E, on pose $f(x) = \|x\|^2$. f est une application de E dans $\mathbb{R}$.

Soit $a \in E$. Pour tout h de E,

$$f(a+h) - f(a) = \|a+h\|^2 - \|a\|^2 = 2\langle a, h\rangle + \|h\|^2.$$

L'application $h \mapsto 2\langle a, h\rangle$ est une application linéaire de E vers $\mathbb{R}$ (c'est-à-dire une forme linéaire sur E) et d'autre part, $\|h\|^2 \underset{h\to 0}{=} o(h)$ car $\displaystyle\lim_{h\to 0} \frac{1}{\|h\|}\|h\|^2 = \lim_{h\to 0} \|h\| = 0$. Donc, f est différentiable en a et

$$\forall h \in E,\ df_a(h) = 2\langle a, h\rangle.$$

Un résultat immédiat est :

Théorème 8. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle.
Soit f une application d'un ouvert non vide Ω de E vers F.

Si f est constante sur Ω, alors pour tout $a \in \Omega$, f est différentiable en a et $df_a = 0$.

On détaille maintenant le cas particulier des fonctions d'une seule variable.

Théorème 9. Soit f une application d'un intervalle ouvert non vide I de $\mathbb{R}$ à valeurs dans un $\mathbb{R}$-espace vectoriel E de dimension finie non nulle. Soit $a \in I$.

f est différentiable en a si et seulement si f est dérivable en a. De plus, dans ce cas,

$$\forall h \in \mathbb{R},\ df_a(h) = hf'(a).$$

Remarque. On note que dans ce cas, $f'(a) = df_a(1)$ ou $f'(a) = df(a).1$. ❏

Démonstration . On sait que f est dérivable en a si et seulement si f admet en a un développement limité d'ordre 1 et que dans ce cas, ce développement limité s'écrit

$$f(a+h) \underset{h\to 0}{=} f(a) + hf'(a) + o(h),$$

ce qui établit le résultat puisque l'application $h \mapsto hf'(a)$ est linéaire.

❏

Pour les fonctions d'une seule variable, on sait que si f est dérivable en a, alors f est continue en a (et on sait que la réciproque est fausse). Ce théorème se généralise de la façon suivante :

Théorème 10. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle.
Soit f une application d'un ouvert non vide Ω de E vers F. Soit $a \in \Omega$.

Si f est différentiable en a, alors f est continue en a (et la réciproque de cette implication est fausse).

Démonstration. Puisque f est différentiable en a, $f(a+h) \underset{h\to 0}{=} f(a) + df_a(h) + o(h)$. Maintenant, puisque E est de dimension finie, on sait que l'application linéaire df_a est continue sur E (voir chapitre « Topologie »). Donc, $\lim\limits_{\substack{h\to 0\\ h\neq 0}} df_a(h) = df_a(0) = 0$. On en déduit que

$$\lim_{\substack{h\to 0\\ h\neq 0}} f(a+h) = \lim_{\substack{h\to 0\\ h\neq 0}} (f(a) + df_a(h) + o(h)) = f(a),$$

ce qui démontre la continuité de f en a.

Exercice 3. Soit $f : \begin{array}{ccc} GL_n(\mathbb{R}) & \to & \mathscr{M}_n(\mathbb{R}) \\ A & \mapsto & A^{-1} \end{array}$.

1) Montrer que $GL_n(\mathbb{R})$ est un ouvert de $\mathbb{R}$.

2) Montrer que f est continue sur $GL_n(\mathbb{R})$.

3) Montrer que f est différentiable en tout point de $GL_n(\mathbb{R})$ et déterminer la différentielle de f en tout point de $GL_n(\mathbb{R})$.

Solution 3. On munit $\mathscr{M}_n(\mathbb{R})$ d'une norme sous-multiplicative notée $\|\ \|$.

1) $GL_n(\mathbb{R})$ est un ouvert de $\mathscr{M}_n(\mathbb{R})$ en tant qu'image réciproque de l'ouvert $\mathbb{R}\setminus\{0\}$ de $\mathbb{R}$ par l'application $A \mapsto \det(A)$ qui est continue sur $\mathscr{M}_n(\mathbb{R})$.

2) On sait que pour tout $A \in GL_n(\mathbb{R})$, $A^{-1} = \dfrac{1}{\det(A)}(\mathrm{com}(A))^T$.

Soit $(i,j) \in [\![1,n]\!]^2$. L'application qui à $A \in \mathscr{M}_n(\mathbb{R})$ associe la matrice extraite de A obtenue en supprimant sa ligne i et sa colonne j, est continue sur $\mathscr{M}_n(\mathbb{R})$ car linéaire sur $\mathscr{M}_n(\mathbb{R})$ puis l'application $A \mapsto A_{i,j}$ (où $A_{i,j}$ est le cofacteur de $a_{i,j}$) est continue par continuité du déterminant. On en déduit que l'application $A \mapsto \mathrm{com}(A)$ est continue sur $\mathscr{M}_n(\mathbb{R})$ car chacune de ses n^2 composantes l'est.

La transposition est continue sur $\mathscr{M}_n(\mathbb{R})$ car linéaire sur $\mathscr{M}_n(\mathbb{R})$ et donc l'application $A \mapsto (\text{com}(A))^T$ est continue dans $\mathscr{M}_n(\mathbb{R})$. D'autre part, l'application $A \mapsto \dfrac{1}{\det(A)}$ est continue sur $GL_n(\mathbb{R})$ en tant qu'inverse d'une fonction continue sur $GL_n(\mathbb{R})$ à valeurs dans $\mathbb{R}$, ne s'annulant pas sur $GL_n(\mathbb{R})$.

Finalement, l'application $f : A \mapsto A^{-1} = \dfrac{1}{\det(A)}(\text{com}(A))^T$ est continue sur $GL_n(\mathbb{R})$.

3) Soit $A \in GL_n(\mathbb{R})$. Puisque $GL_n(\mathbb{R})$ est un ouvert de $\mathscr{M}_n(\mathbb{R})$, il existe un voisinage $\mathcal{V}$ de 0 dans $\mathscr{M}_n(\mathbb{R})$ tel que, pour tout $H \in \mathcal{V}$, $A + H \in GL_n(\mathbb{R})$. Pour $H \in \mathcal{V}$,

$$(A+H)^{-1} - A^{-1} = (A+H)^{-1}\left(I_n - (A+H)A^{-1}\right) = -(A+H)^{-1}HA^{-1}$$

puis

$$\begin{aligned}(A+H)^{-1} - A^{-1} + A^{-1}HA^{-1} &= \left(-(A+H)^{-1} + A^{-1}\right)HA^{-1} = (A+H)^{-1}\left(-I_n + (A+H)A^{-1}\right)HA^{-1} \\ &= (A+H)^{-1}HA^{-1}HA^{-1}.\end{aligned}$$

Par suite, pour $H \in \mathcal{V} \setminus \{0\}$,

$$\frac{1}{\|H\|}\left\|(A+H)^{-1} - A^{-1} + A^{-1}HA^{-1}\right\| = \frac{1}{\|H\|}\left\|(A+H)^{-1}HA^{-1}HA^{-1}\right\| \leqslant \left\|(A+H)^{-1}\right\|\left\|A^{-1}\right\|^2\|H\|.$$

Cette dernière expression tend vers 0 quand H tend vers 0 (car la fonction f est continue en A de sorte que $(A+H)^{-1}$ tend vers A^{-1} quand H tend vers 0 puis $\left\|(A+H)^{-1}\right\|$ tend vers $\left\|A^{-1}\right\|$ quand H tend vers 0 par continuité de la norme). Ainsi, $(A+H)^{-1} - A^{-1} + A^{-1}HA^{-1} \underset{H \to 0}{=} o(H)$ ou encore

$$f(A+H) \underset{H \to 0}{=} f(A) - A^{-1}HA^{-1} + o(H).$$

Maintenant, l'application $H \mapsto -A^{-1}HA^{-1}$ est linéaire et on a donc montré que f est différentiable en A et

$$\forall H \in \mathscr{M}_n(\mathbb{R}),\ df_A(H) = -A^{-1}HA^{-1}.$$

3.3.2 Lien avec la dérivée suivant un vecteur

Théorème 11. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit f une application définie sur Ω, ouvert non vide de E, à valeurs dans F. Soit $a \in \Omega$.

Si f est différentiable en a, alors f est dérivable suivant tout vecteur en a et dans ce cas, pour tout $v \in E \setminus \{0\}$,

$$D_v f(a) = df_a(v) \text{ (ou } D_v f(a) = df(a).v).$$

Démonstration . Soit $v \in E \setminus \{0\}$. Puisque f est différentiable en a, quand le vecteur h tend vers 0

$$f(a+h) = f(a) + df_a(h) + o(h).$$

En particulier, quand le réel t tend vers 0,

$$f(a+tv) = f(a) + df_a(tv) + o(t) = f(a) + t\,df_a(v) + o(t)$$

et donc

$$\frac{1}{t}(f(a+tv) - f(a)) - df_a(v) \underset{t \to 0}{=} o(1).$$

Ceci montre que f est dérivable suivant le vecteur v et que $D_v f(a) = df_a(v)$.

❏

La liste des différentes implications s'est allongée (toute implication non écrite étant fausse) :

$$
\begin{tikzcd}
 & \text{Différentiabilité en } a \arrow[dl, Rightarrow] \arrow[ddr, Rightarrow] & \\
\text{Dérivabilité suivant tout vecteur en } a \arrow[d, Rightarrow] & & \\
\text{Existence de dérivées partielles en } a \arrow[dr, Rightarrow] & & \text{Continuité en } a \arrow[dl, Rightarrow] \\
 & \text{Continuité partielle en } a &
\end{tikzcd}
$$

3.3.3 Lien avec les dérivées partielles. Expression de la différentielle en un point

On applique le théorème 11 au cas particulier où le vecteur v est l'un des vecteurs de la base canonique de $\mathbb{R}^n$ et on obtient immédiatement :

Théorème 12. Soit f une application d'un ouvert non vide Ω de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Soit $a \in \Omega$.

Si f est différentiable en a, alors f admet des dérivées partielles par rapport à chacune de ses variables en a et de plus,

$$\forall i \in [\![1, n]\!],\ \frac{\partial f}{\partial x_i}(a) = df_a\left(e_i\right).$$

Le théorème précédent se généralise bien sûr aux dérivées partielles dans une base $(e_1, \ldots, e_n)$ d'un $\mathbb{R}$-espace E de dimension n pour une application f de E vers un $\mathbb{R}$-espace F de dimension p.

Avec ce résultat, on est maintenant en mesure d'exprimer la différentielle d'une fonction f en un point a à partir des dérivées partielles de f en a :

Théorème 13. Soit f une application d'un ouvert non vide Ω de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Soit $a \in \Omega$. Si f est différentiable en a, alors

$$\forall h = (h_1, \ldots, h_n) \in \mathbb{R}^n,\ df_a(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a).$$

Démonstration. On note $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. Puisque $df_a \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$, pour $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$,

$$df_a(h) = df_a\left(\sum_{i=1}^{n} h_i e_i\right) = \sum_{i=1}^{n} h_i df_a\left(e_i\right) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a),$$

(d'après le théorème 12).

❏

Plus généralement, si f est une application d'un ouvert non vide Ω d'un $\mathbb{R}$-espace E de dimension finie $n \in \mathbb{N}^*$ à valeurs dans un $\mathbb{R}$-espace F de dimension finie $p \in \mathbb{N}^*$ et si $\mathscr{B} = (e_1, \ldots, e_n)$ est une base donnée de E, alors pour tout $h = \sum_{i=1}^{n} h_i e_i \in E$, où $(h_1, \ldots, h_n) \in \mathbb{R}^n$,

$$df_a(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a),$$

où les $\frac{\partial f}{\partial x_i}(a)$, $i \in [\![1, n]\!]$, sont les dérivées partielles de f dans la base $\mathscr{B}$ en a.

3.3.4 Différentiabilité et différentielle d'une application linéaire

Théorème 14. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit $f \in \mathscr{L}(E, F)$.

Pour tout $a \in E$, f est différentiable en a et $df_a = f$.

Démonstration . Puisque f est linéaire, l'application $h \mapsto f(a+h) - f(a) - f(h)$ est l'application nulle sur E. En particulier,

$$f(a+h) \underset{h \to 0}{=} f(a) + f(h) + o(h),$$

avec f linéaire de E vers F. Ceci montre que f est différentiable en a et que $df_a = f$.

❏

On analyse maintenant un cas particulier important du théorème précédent. Notons $\mathscr{B} = (e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Pour $i \in [\![1, n]\!]$, notons e_i^* la i-ème forme coordonnée dans la base $\mathscr{B}$. Pour $i \in [\![1, n]\!]$, e_i^* est définie par les égalités : $\forall j \in [\![1, n]\!]$, $e_i^*(e_j) = \delta_{i,j}$ ou plus généralement

$$\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n,\ e_i^*(x) = x_i.$$

Puisque e_i^* est une forme linéaire sur $\mathbb{R}^n$, sa différentielle en tout point (c'est-à-dire la différentielle en tout point de l'application $x \mapsto x_i$) est elle-même. Il est donc cohérent de noter

$$e_i^* = dx_i.$$

dx_i est ainsi une forme linéaire sur $\mathbb{R}^n$ et plus précisément dx_i est la i-ème forme coordonnée dans la base $\mathscr{B}$. Par définition, pour tout $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, $dx_i(h) = h_i$.

3.3.5 Différentielle d'une fonction sur un ouvert

Définition 6. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit f une application d'un ouvert Ω de E vers F.

f est **différentiable sur** Ω si et seulement si f est différentiable en chaque point de Ω. Dans ce cas, la différentielle de f sur Ω est l'application, notée df, qui à chaque a de Ω associe df_a :

$$df :\ \begin{array}{ccc} \Omega & \to & \mathscr{L}(E, F) \\ a & \mapsto & df_a \end{array}.$$

L'ensemble des fonctions différentiables sur Ω à valeurs dans F se note $D^1(\Omega, F)$.

df est donc une application de Ω dans $\mathscr{L}(E, F)$, l'ensemble des applications linéaires de E vers F. C'est un objet assez compliqué ...

Avec les notations du paragraphe précédent, on peut donner une expression de la différentielle de f sur Ω à l'aide des fonctions dérivées partielles :

Théorème 15. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit f une application d'un ouvert Ω de E vers F. Si f est différentiable sur Ω, alors

$$df = \sum_{i=1}^{n} dx_i \frac{\partial f}{\partial x_i}.$$

Si on évalue en $a \in \Omega$ les deux membres de l'égalité précédente, on obtient

$$df_a = \sum_{i=1}^{n} dx_i \frac{\partial f}{\partial x_i}(a),$$

et si on évalue les deux membres de l'égalité précédente en $h = \sum_{i=1}^{n} h_i e_i$, on obtient

$$df_a(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a).$$

On note que l'ordre des objets $dx_i \dfrac{\partial f}{\partial x_i}(a)$ ne peut être modifié car dx_i est une forme linéaire et $\dfrac{\partial f}{\partial x_i}(a)$ est un vecteur élément de F. Si f est à valeurs dans $\mathbb{R}$ (cas où F est de dimension 1), alors chaque $\dfrac{\partial f}{\partial x_i}(a)$ est un nombre réel et l'ordre peut être inversé :

$$df_a = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) dx_i$$

ou

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i.$$

Exemple. Pour $(x, y) \in \mathbb{R}^2$, on pose $f(x, y) = xe^{x^2+y^2}$. Alors,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = \left(2x^2+1\right) e^{x^2+y^2} dx + 2xye^{x^2+y^2} dy,$$

puis, pour tout $(x_0, y_0) \in \mathbb{R}^2$,

$$df_{(x_0,y_0)} = \frac{\partial f}{\partial x}(x_0, y_0)\, dx + \frac{\partial f}{\partial y}(x_0, y_0)\, dy = \left(2x_0^2+1\right) e^{x_0^2+y_0^2} dx + 2x_0y_0e^{x_0^2+y_0^2} dy,$$

puis, pour tout $(x_0, y_0) \in \mathbb{R}^2$ et tout $(h, k) \in \mathbb{R}^2$,

$$df_{(x_0,y_0)}(h, k) = \left(\left(2x_0^2+1\right) h + 2x_0y_0k\right) e^{x_0^2+y_0^2}.$$

❑

3.3.6 Matrice jacobienne

Définition 7. (matrice jacobienne)

• Soit f une application d'un ouvert non vide Ω de $\mathbb{R}^n$ dans $\mathbb{R}^p$, différentiable en $a \in \Omega$.

La **matrice jacobienne** de f en a, notée $J(f, a)$ est la matrice de l'application linéaire df_a relativement aux bases canoniques $\mathscr{B}_n$ et $\mathscr{B}_p$ de $\mathbb{R}^n$ et $\mathbb{R}^p$ respectivement :

$$J(f, a) = \text{Mat}_{\mathscr{B}_n, \mathscr{B}_p}\left(df_a\right).$$

C'est un élément de $\mathscr{M}_{p,n}(\mathbb{R})$.

• Plus généralement, soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit f une application d'un ouvert non vide Ω de E vers F, différentiable en $a \in \Omega$. Soient $\mathscr{B}$ une base de E et $\mathscr{B}'$ une base de F.

La **matrice jacobienne** de f en a relativement aux bases $\mathscr{B}$ et $\mathscr{B}'$ est la matrice de l'application linéaire df_a relativement aux bases $\mathscr{B}$ et $\mathscr{B}'$.

Théorème 16. Soit f une application d'un ouvert non vide Ω de $\mathbb{R}^n$ dans $\mathbb{R}^p$, différentiable en $a \in \Omega$. On note f_1, ..., f_p, les applications composantes de f : $\forall (x_1, \ldots, x_n) \in \Omega$, $f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_p(x_1, \ldots, x_n))$. Alors,

$$J(f, a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{1 \leqslant i \leqslant p,\ 1 \leqslant j \leqslant n}.$$

Démonstration. On note $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. Soit $j \in [\![1, n]\!]$.
La j-ème colonne de $J(f, a)$ est le vecteur colonne dont les composantes sont les coordonnées de $df_a(e_j)$ dans la base canonique de $\mathbb{R}^p$. D'après le théorème 12, $df_a(e_j) = \dfrac{\partial f}{\partial x_j}(a)$ et donc la j-ème colonne de $J(f, a)$ est $\begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_i}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_p}{\partial x_j}(a) \end{pmatrix}$.

❑

Exemple. On munit le plan $\mathbb{R}^2$ de sa structure euclidienne et de son orientation canonique. On note $\left(\overrightarrow{u}, \overrightarrow{v}\right)$ la base canonique (orthonormée directe) de $\mathbb{R}^2$. Pour $\theta \in \mathbb{R}$, on pose $\overrightarrow{u_\theta} = (\cos\theta)\overrightarrow{u} + (\sin\theta)\overrightarrow{v}$ (de sorte que $\overrightarrow{u} = \overrightarrow{u_0}$ et $\overrightarrow{v} = \overrightarrow{u_{\frac{\pi}{2}}}$). Soit $M = (x, y)$ un point de $\mathbb{R}^2$. On appelle coordonnées polaires de M tout couple $[r, \theta]$ de réels tels que $\overrightarrow{OM} = r\overrightarrow{u_\theta}$. On a donc

$$x\overrightarrow{u}+y\overrightarrow{v}=\overrightarrow{OM}=(r\cos\theta)\overrightarrow{u}+(r\sin\theta)\overrightarrow{v}.$$

Le « passage des coordonnées polaires aux coordonnées cartésiennes » s'écrit donc $\left\{\begin{array}{l} x=r\cos\theta \\ y=r\sin\theta \end{array}\right.$.

Pour $(r,\theta)\in\mathbb{R}^2$, on pose $f(r,\theta)=(r\cos\theta,r\sin\theta)$. La matrice jacobienne de f en un point (r,θ) de $\mathbb{R}^2$ est

$$\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

En première colonne, on a dérivé par rapport à r et en deuxième colonne, on a dérivé par rapport à θ. ❑

3.3.7 Opérations sur les différentielles

Théorème 17. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soient f et g deux applications de Ω, ouvert non vide de E, vers F.

1) Soit $a\in\Omega$. Si f et g sont différentiables en a, alors pour tout $(\lambda,\mu)\in\mathbb{R}^2$, $\lambda f+\mu g$ est différentiable en a et

$$d(\lambda f+\mu g)_a=\lambda df_a+\mu dg_a.$$

L'ensemble des fonctions différentiables en a à valeurs dans F est un $\mathbb{R}$-espace vectoriel.

2) Si f et g sont différentiables sur Ω, alors pour tout $(\lambda,\mu)\in\mathbb{R}^2$, $\lambda f+\mu g$ est différentiable sur Ω et

$$d(\lambda f+\mu g)=\lambda df+\mu dg.$$

L'ensemble $D^1(\Omega,F)$ des fonctions différentiables sur Ω à valeurs dans F est un $\mathbb{R}$-espace vectoriel.

Démonstration .

$$\begin{aligned}(\lambda f+\mu g)(a+h)-(\lambda f+\mu g)(a)&=\lambda(f(a+h)-f(a))+\mu(g(a+h)-g(a))\\ &\underset{h\to 0}{=}\lambda\left(df_a(h)+o(h)\right)+\mu\left(dg_a(h)+o(h)\right)\\ &\underset{h\to 0}{=}\left(\lambda df_a+\mu dg_a\right)(h)+o(h).\end{aligned}$$

De plus, $\lambda df_a+\mu dg_a$ est une application linéaire de E vers F et on a donc montré que $\lambda f+\mu g$ est différentiable en a et que de plus, $d(\lambda f+\mu g)_a=\lambda df_a+\mu dg_a$.

❑

Théorème 18. Soient E, F, G et H quatre $\mathbb{R}$-espaces vectoriels de dimension finie non nulle et soit Ω un ouvert non vide de E. Soient f une application de Ω dans F et g une application de Ω dans G. Soit enfin B une application de $F\times G$ dans H, bilinéaire sur $F\times G$.

1) Soit $a\in\Omega$. Si f et g sont différentiables en a, alors $B(f,g)$ est différentiable en a et

$$d(B(f,g))_a=B(df_a,g(a))+B(f(a),dg_a).$$

2) Si f et g sont différentiables sur Ω, alors $B(f,g)$ est différentiable sur Ω et

$$d(B(f,g))=B(df,g)+B(f,dg).$$

⇛ **Commentaire** . *Le théorème 18 se généralise par récurrence à des applications multilinéaires* $M\ :\ E_1\times\ldots\times E_p$ *:*

$$d\left(M\left(f_1,\ldots,f_p\right)\right)_a=\sum_{i=1}^{p}M\left(f_1(a),\ldots,f_{i-1}(a),\left(df_i\right)_a,f_{i+1}(a),\ldots,f_p(a)\right).$$

Démonstration . On note $\mathscr{B}=(e_1,\ldots,e_p)$ une base de F et $\mathscr{B}'=(e'_1,\ldots,e'_q)$ une base de G. On note $\|\ \|_F$ et $\|\ \|_G$ les normes infini respectivement associées. On note encore $\|\ \|_H$ une norme sur H.

• Montrons d'abord qu'il existe $K\in\mathbb{R}^{+*}$ tel que $\forall(x,y)\in F\times G$, $\|B(x,y)\|_H\leqslant K\|x\|_F\|y\|_G$.

Pour $x=\displaystyle\sum_{i=1}^{p}x_ie_i$ et $y=\displaystyle\sum_{j=1}^{q}y_je'_j$,

$$\begin{aligned}\|B(x,y)\|_H = \left\|\sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} x_iy_jB\left(e_i,e'_j\right)\right\|_H &\leqslant \sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}} |x_i|\,|y_j|\,\left\|B\left(e_i,e'_j\right)\right\|_H\\ &\leqslant \left(\sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}\left\|B\left(e_i,e'_j\right)\right\|_H\right)\|x\|_F\|y\|_G.\end{aligned}$$

Le réel $K = \sum_{\substack{1\leqslant i\leqslant p\\1\leqslant j\leqslant q}}\left\|B\left(e_i,e'_j\right)\right\|_H$ convient et est ainsi dorénavant fixé.

• Pour h au voisinage de 0,

$$\begin{aligned}B(f,g)(a+h)-B(f,g)(a) &= B(f(a+h),g(a+h))-B(f(a),g(a+h))+B(f(a),g(a+h))-B(f(a),g(a))\\ &= B(f(a+h)-f(a),g(a+h))+B(f(a),g(a+h)-g(a))\end{aligned}$$

puis

$$\begin{aligned}&B(f,g)(a+h)-B(f,g)(a)-B\left(df_a(h),g(a)\right)-B\left(f(a),dg_a(h)\right)\\ &\qquad= B(f(a+h)-f(a),g(a))-B\left(df_a(h),g(a)\right)+B(f(a),g(a+h)-g(a))-B\left(f(a),dg_a(h)\right)\\ &\qquad+B(f(a+h)-f(a),g(a+h)-g(a))\\ &\qquad= B\left(f(a+h)-f(a)-df_a(h),g(a)\right)+B(f(a),g(a+h)-g(a)-dg_a(h))\\ &\qquad+B(f(a+h)-f(a),g(a+h)-g(a)).\end{aligned}$$

Par suite, pour h au voisinage de 0 et non nul

$$\begin{aligned}&\left\|\frac{1}{\|h\|_E}\left(B(f,g)(a+h)-B(f,g)(a)-B\left(df_a(h),g(a)\right)-B\left(f(a),dg_a(h)\right)\right)\right\|_H\\ &\qquad\leqslant K\left(\frac{1}{\|h\|_E}\left\|f(a+h)-f(a)-df_a(h)\right\|_F\|g(a)\|_G\right.\\ &\qquad\left.+\|f(a)\|_F\frac{1}{\|h\|_E}\left\|g(a+h)-g(a)-dg_a(h)\right\|_G+\frac{1}{\|h\|_E}\left\|f(a+h)-f(a)\right\|_F\left\|g(a+h)-g(a)\right\|_G\right).\end{aligned}$$

Puisque f et g sont différentiables en a, chacune des deux expressions $\frac{1}{\|h\|_E}\left\|f(a+h)-f(a)-df_a(h)\right\|_F$ et $\frac{1}{\|h\|_E}\left\|g(a+h)-g(a)-dg_a(h)\right\|_G$ tend vers 0 quand h tend vers 0. D'autre part, quand h tend vers 0,

$$\frac{1}{\|h\|_E}\left\|f(a+h)-f(a)\right\|_F\left\|g(a+h)-g(a)\right\|_G = \left\|df_a\left(\frac{h}{\|h\|_E}\right)+\frac{o(h)}{\|h\|_E}\right\|_F\left\|dg_a(h)+o(h)\right\|_G.$$

df_a est une application linéaire de E vers F et E est de dimension finie. Donc, df_a est continue sur E puis on sait qu'il existe $M\in\mathbb{R}$ tel que $\forall x\in E$, $\|df_a(x)\|_F\leqslant M\|x\|_E$. Par suite, pour h non nul au voisinage de 0, $\left\|df_a\left(\frac{h}{\|h\|_E}\right)\right\|_F\leqslant M\left\|\frac{h}{\|h\|_E}\right\|_E = M$. Ainsi, l'expression $df_a\left(\frac{h}{\|h\|_E}\right)$ est bornée au voisinage de 0. Puisque $\frac{o(h)}{\|h\|_E}$ tend vers 0 quand h tend vers 0, l'expression $df_a\left(\frac{h}{\|h\|_E}\right)+\frac{o(h)}{\|h\|_E}$ est bornée au voisinage de 0. D'autre part, $dg_a(h)+o(h)$ tend vers 0 quand h tend vers 0 par continuité de dg_a en 0. Ceci montre que $\frac{1}{\|h\|_E}\left\|f(a+h)-f(a)\right\|_F\left\|g(a+h)-g(a)\right\|_G$ tend vers 0 quand h tend vers 0.

Finalement, l'expression $\frac{1}{\|h\|_E}\left(B(f,g)(a+h)-B(f,g)(a)-B\left(df_a(h),g(a)\right)-B\left(f(a),dg_a(h)\right)\right)$ tend vers 0 quand h tend vers 0 ou encore

$$B(f,g)(a+h)\underset{h\to0}{=}B(f,g)(a)+B\left(df_a(h),g(a)\right)+B\left(f(a),dg_a(h)\right)+o(h).$$

Puisque l'application $h\mapsto B\left(df_a(h),g(a)\right)+B\left(f(a),dg_a(h)\right)$ est linéaire, on a montré que $B(f,g)$ est différentiable en a et que

$$d(B(f,g))_a = B(df_a,g(a))+B(f(a),dg_a).$$

❏

Un cas particulier du théorème 18 est (en appliquant à $B\ :\ \begin{array}{ccc}\mathbb{R}^2 & \rightarrow & \mathbb{R}\\ (x,y) & \mapsto & x\times y\end{array}$) :

Théorème 19. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie non nulle. Soient f et g deux applications de Ω, ouvert non vide de E, vers $\mathbb{R}$.

1) Soit $a \in \Omega$. Si f et g sont différentiables en a, alors $f \times g$ est différentiable en a et

$$d(f \times g)_a = g(a)df_a + f(a)dg_a.$$

2) Si f et g sont différentiables sur Ω, alors $f \times g$ est différentiable sur Ω et

$$d(f \times g) = g\,df + f\,dg.$$

Théorème 20. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie non nulle. Soient f et g deux applications de Ω, ouvert non vide de E, vers $\mathbb{R}$.

1) Soit $a \in \Omega$.

a) Si f est différentiable en a et si $f(a) \neq 0$, alors $\dfrac{1}{f}$ est différentiable en a et

$$d\left(\frac{1}{f}\right)_a = -\frac{df_a}{(f(a))^2}.$$

b) Si f et g sont différentiables en a et si $g(a) \neq 0$, alors $\dfrac{f}{g}$ est différentiable en a et

$$d\left(\frac{f}{g}\right)_a = \frac{g(a)df_a - f(a)dg_a}{(g(a))^2}.$$

2)

a) Si f est différentiable sur Ω et si f ne s'annule pas sur Ω, alors $\dfrac{1}{f}$ est différentiable sur Ω et

$$d\left(\frac{1}{f}\right) = -\frac{1}{f^2}df.$$

b) Si f et g sont différentiables sur Ω et si g ne s'annule pas sur Ω, alors$\dfrac{f}{g}$ est différentiable sur Ω et

$$d\left(\frac{f}{g}\right) = \frac{1}{g^2}(g\,df - f\,dg).$$

Démonstration . On montre 1) a) et b).

a) f est différentiable en a et en particulier continue en a. Puisque f ne s'annule pas en a, f ne s'annule pas au voisinage de a.

Pour h au voisinage de 0, $\dfrac{1}{f(a+h)} - \dfrac{1}{f(a)} = -\dfrac{f(a+h) - f(a)}{f(a)f(a+h)}$ et donc,

$$\frac{1}{f(a+h)} - \frac{1}{f(a)} \underset{h \to 0}{=} -\frac{df_a(h) + o(h)}{f(a)f(a+h)} \underset{h \to 0}{=} -\frac{df_a(h) + o(h)}{(f(a))^2 + o(1)} \underset{h \to 0}{=} -\frac{1}{(f(a))^2}\left(df_a(h) + o(h)\right)(1 + o(1))$$

(car f est différentiable en a et en particulier continue en a de sorte que $f(a)f(a+h)$ tend vers $(f(a))^2 \neq 0$ quand h tend vers 0). Maintenant, puisque df_a est continue sur E car linéaire sur un espace de dimension finie, on sait qu'il existe $M \in \mathbb{R}^+$ tel que, pour tout $h \in E$, $\|df_a(h)\| \leqslant M\|h\|$. On en déduit que $df_a(h)o(1) \underset{h \to 0}{=} o(h)$ et finalement que,

$$\frac{1}{f(a+h)} - \frac{1}{f(a)} \underset{h \to 0}{=} -\frac{df_a(h)}{(f(a))^2} + o(h).$$

Puisque l'application $h \mapsto -\dfrac{df_a(h)}{(f(a))^2}$ est linéaire, on a montré que $\dfrac{1}{f}$ est différentiable en a et que $d\left(\dfrac{1}{f}\right)_a = -\dfrac{df_a}{(f(a))^2}$.

b) On applique a) et le théorème 19. On obtient la différentiabilité en a de $\dfrac{f}{g} = f \times \dfrac{1}{g}$ et de plus,

$$d\left(\frac{f}{g}\right)_a = \frac{1}{g(a)}df_a + f(a)\frac{-dg_a}{(g(a))^2} = \frac{g(a)df_a - f(a)dg_a}{(g(a))^2}.$$

❏

Théorème 21. Soient E, F et G trois $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soient f une application d'un ouvert non vide Ω de E vers F et g une application d'un ouvert Ω' de F vers G telles que $f(\Omega) \subset \Omega'$.

1) Soit $a \in \Omega$. Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et de plus,

$$d(g \circ f)_a = dg_{f(a)} \circ df_a.$$

2) Si f est différentiable sur Ω et g est différentiable sur Ω', alors $g \circ f$ est différentiable sur Ω et de plus,

$$d(g \circ f) = (dg \circ f) \circ df.$$

Démonstration . On note $\| \|_E$ une norme donnée dans E, $\| \|_F$ une norme donnée dans F.

f est différentiable en a et donc il existe une fonction ε_1, définie sur un voisinage de 0 dans E, telle que

$$f(a+h) \underset{h \to 0}{=} f(a) + df_a(h) + \|h\|_E \varepsilon_1(h) \text{ et } \lim_{h \to 0} \varepsilon_1(h) = 0.$$

De même, g est différentiable en $f(a)$ et donc il existe une fonction ε_2, définie sur un voisinage de 0 dans F, telle que

$$g(f(a)+k) \underset{k \to 0}{=} g(f(a)) + dg_{f(a)}(k) + \|k\|_F \varepsilon_2(k) \text{ et } \lim_{k \to 0} \varepsilon_2(k) = 0.$$

Par suite, puisque $k = df_a(h) + \|h\|_E \varepsilon_1(h) \underset{h \to 0}{\to} 0$ (par continuité de df_a en 0),

$$\begin{aligned}
g(f(a+h)) &\underset{h \to 0}{=} g\left[f(a) + (df_a(h) + \|h\|_E \varepsilon_1(h))\right] \\
&\underset{h \to 0}{=} g(f(a)) + dg_{f(a)}\left[df_a(h) + \|h\|_E \varepsilon_1(h)\right] + \|df_a(h) + \|h\|_E \varepsilon_1(h)\|_F \, \varepsilon_2\left(df_a(h) + \|h\|_E \varepsilon_1(h)\right) \\
&\underset{h \to 0}{=} g(f(a)) + dg_{f(a)}\left(df_a(h)\right) + \|h\|_E dg_{f(a)}\left(\varepsilon_1(h)\right) + \|df_a(h) + \|h\|_E \varepsilon_1(h)\|_F \, \varepsilon_2\left(df_a(h) + \|h\|_E \varepsilon_1(h)\right).
\end{aligned}$$

Ensuite, puisque $dg_{f(a)}\left(\varepsilon_1(h)\right) \underset{h \to 0}{\to} 0$ (par continuité de $dg_{f(a)}$ en 0), on a $\|h\|_E dg_{f(a)}\left(\varepsilon_1(h)\right) \underset{h \to 0}{=} o(h)$. Ensuite,

$$\frac{1}{\|h\|_E} \|df_a(h) + \|h\|_E \varepsilon_1(h)\|_F \, \varepsilon_2\left(df_a(h) + \|h\|_E \varepsilon_1(h)\right) = \left\| df_a\left(\frac{1}{\|h\|_E} h\right) + \varepsilon_1(h) \right\|_F \varepsilon_2\left(df_a(h) + \|h\|_E \varepsilon_1(h)\right).$$

Puisque E est de dimension finie et que df_a est linéaire sur E, on sait qu'il existe $K \in \mathbb{R}^{+*}$ tel que

$$\forall h \in E, \ \|df_a(h)\|_F \leqslant K \|h\|_E$$

et donc $\forall h \in E \setminus \{0\}$, $\left\| df_a\left(\frac{1}{\|h\|_E} h\right) \right\|_F \leqslant K$.

Ainsi, l'expression $df_a\left(\frac{1}{\|h\|_E} h\right)$ est bornée et donc l'expression $df_a\left(\frac{1}{\|h\|_E} h\right) + \varepsilon_1(h)$ est bornée sur un voisinage de 0. On en déduit que $\left\| df_a\left(\frac{1}{\|h\|_E} h\right) + \varepsilon_1(h) \right\|_F \varepsilon_2\left(df_a(h) + \|h\|_E \varepsilon_1(h)\right) \underset{h \to 0}{\to} 0$ et finalement que

$$\|df_a(h) + \|h\|_E \varepsilon_1(h)\|_F \, \varepsilon_2\left(df_a(h) + \|h\|_E \varepsilon_1(h)\right) \underset{h \to 0}{=} o(h).$$

En résumé,

$$g(f(a+h)) \underset{h \to 0}{=} g(f(a)) + \left(dg_{f(a)} \circ df_a\right)(h) + o(h).$$

Puisque $dg_{f(a)} \circ df_a \in \mathscr{L}(E, G)$, on a montré que $g \circ f$ est différentiable en a et que $d(g \circ f)_a = dg_{f(a)} \circ df_a$.

❏

Une conséquence est que (des bases étant fixées une bonne fois pour toutes) la matrice jacobienne de $g \circ f$ en a est le produit de la matrice jacobienne de g en $f(a)$ par de la matrice jacobienne de f en a :

$$J(g \circ f, a) = J(g, f(a)) \times J(f, a).$$

Analysons maintenant les conséquences de cette formule sur les dérivées partielles de fonctions composées.

Soient

$$f : \begin{array}{ccc} \Omega \subset \mathbb{R}^n & \to & \mathbb{R}^p \\ (x_1, \dots, x_n) & \mapsto & \left(f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)\right) \end{array}$$

et

$$g : \begin{array}{ccc} \Omega' \subset \mathbb{R}^p & \to & \mathbb{R}^q \\ (y_1, \dots, y_p) & \mapsto & \left(g_1(y_1, \dots, y_p), \dots, g_q(y_1, \dots, y_p)\right) \end{array}$$

telles que $f(\Omega) \subset \Omega'$. $J(f, a)$ est la matrice de format (p, n) dont le coefficient ligne i, colonne j, $(i, j) \in [\![1, p]\!] \times [\![1, n]\!]$ est $\dfrac{\partial f_i}{\partial x_j}(a)$ et $J(g, f(a))$ est la matrice de format (q, p) dont le coefficient ligne i, colonne j, $(i, j) \in [\![1, q]\!] \times [\![1, p]\!]$ est $\dfrac{\partial g_i}{\partial y_j}(f(a))$.

Puisque $J(g \circ f, a) = J(g, (f(a)) \times J(f, a)$, on obtient (règle de la chaîne)

$$\forall (i, j) \in [\![1, q]\!] \times [\![1, n]\!],\ \frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial y_k}(f(a)) \times \frac{\partial f_k}{\partial x_j}(a) \quad (*).$$

Mettons en œuvre cette formule pénible sur un exemple. On reprend l'exemple du passage des coordonnées cartésiennes aux coordonnées polaires que l'on détaille d'abord en tant que tel avant de donner un exemple de dérivée partielle de fonction composée.

• Pour $(r, \theta) \in \mathbb{R}^2$, on pose $\varphi(r, \theta) = (r \cos \theta, r \sin \theta) = (x, y)$. φ est une fonction de deux variables r et θ et à deux composantes $\varphi_1 : (r, \theta) \mapsto r \cos \theta$ et $\varphi_2 : (r, \theta) \mapsto r \sin \theta$.

On a vu que la matrice jacobienne de φ en un point (r_0, θ_0) de $\mathbb{R}^2$ est $\begin{pmatrix} \cos(\theta_0) & -r_0 \sin(\theta_0) \\ \sin(\theta_0) & r_0 \cos(\theta_0) \end{pmatrix}$. Ceci s'écrit de manière abrégée :

$$\boxed{\begin{array}{ll} \dfrac{\partial x}{\partial r} = \cos(\theta) & \dfrac{\partial x}{\partial \theta} = -r \sin(\theta) \\[2ex] \dfrac{\partial y}{\partial r} = \sin(\theta) & \dfrac{\partial y}{\partial \theta} = r \cos(\theta) \end{array}}$$

• Soit maintenant f une fonction différentiable sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$. f a deux variables x et y et une composante. On passe en polaires en posant

$$\forall (x, y) \in \mathbb{R}^2,\ f(x, y) = f(r \cos \theta, r \sin \theta) = g(r, \theta)$$

ou encore on pose $g = f \circ \varphi$. g est alors une fonction de deux variables r et θ et à une composante. La formule de dérivation partielle des fonctions composées $(*)$ fournit pour tout $(r_0, \theta_0) \in \mathbb{R}^2$

$$\frac{\partial g}{\partial r}(r_0, \theta_0) = \frac{\partial f}{\partial x}(r_0 \cos(\theta_0), r_0 \sin(\theta_0)) \times \frac{\partial \varphi_1}{\partial r}(r_0 \cos(\theta_0), r_0 \sin(\theta_0))(r_0, \theta_0) + \frac{\partial f}{\partial y} \times \frac{\partial \varphi_2}{\partial r}(r_0, \theta_0)$$

ou encore de manière plus abrégée (en un point distinct de l'origine) :

$$\frac{\partial g}{\partial r} = \frac{\partial f}{\partial x} \times \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \times \frac{\partial y}{\partial r} = \cos(\theta) \frac{\partial f}{\partial x} + \sin(\theta) \frac{\partial f}{\partial y} = \frac{1}{r}\left(x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y}\right),$$

et de même

$$\frac{\partial g}{\partial \theta} = \frac{\partial f}{\partial x} \times \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \times \frac{\partial y}{\partial \theta} = -r \sin(\theta) \frac{\partial f}{\partial x} + r \cos(\theta) \frac{\partial f}{\partial y} = -y \frac{\partial f}{\partial x} + x \frac{\partial f}{\partial y},$$

• Maintenant, passer en polaires consiste plutôt à faire l'inverse, c'est-à-dire à exprimer $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ en fonction de $\dfrac{\partial g}{\partial r}$ et $\dfrac{\partial g}{\partial \theta}$. Ceci s'obtient en dérivant partiellement l'égalité $f = g \circ \varphi^{-1}$ (en restreignant les ensembles de départ et d'arrivée de φ à $]0, +\infty[\times[0, 2\pi[$ et $\mathbb{R}^2 \setminus \{0\}$ respectivement, on obtient une bijection que l'on note encore φ). Dans le calcul, nous avons besoin d'inverser les formules $\dfrac{\partial x}{\partial r} = \ldots$ sous la forme $\dfrac{\partial r}{\partial x} = \ldots$

Nous avons plusieurs méthodes pour y parvenir. La première d'entre elles est d'expliciter r et θ en fonction de x et y. C'est assez pénible car il faut discuter suivant la position de (x, y) dans le plan. Par exemple, si $x > 0$, on a $r = \sqrt{x^2 + y^2}$ et $\theta = \mathrm{Arctan}\left(\dfrac{y}{x}\right)$.

Une autre idée bien meilleure est d'utiliser l'égalité $\varphi \circ \varphi^{-1} = \mathrm{Id}$ qui une fois différenciée, fournit de manière très condensée (mais très imprécise) $d\varphi \circ d\varphi^{-1} = \mathrm{Id}$. Une conséquence est que l'inverse de la matrice jacobienne de φ en un point $(r_0, \theta_0) \in]0, +\infty[\times[0, 2\pi[$ à savoir

$$\begin{pmatrix} \cos(\theta_0) & -r_0 \sin(\theta_0) \\ \sin(\theta_0) & r_0 \cos(\theta_0) \end{pmatrix}$$

est la matrice jacobienne de φ^{-1} en $(x_0, y_0) = \varphi(r_0, \theta_0)$. On obtient

$$J\left(\varphi^{-1}, (x_0, y_0)\right) = \frac{1}{r_0}\begin{pmatrix} r_0 \cos(\theta_0) & r_0 \sin(\theta_0) \\ -\sin(\theta_0) & \cos(\theta_0) \end{pmatrix} = \begin{pmatrix} \cos(\theta_0) & \sin(\theta_0) \\ -\dfrac{\sin(\theta_0)}{r_0} & \dfrac{\cos(\theta_0)}{r_0} \end{pmatrix}.$$

On a ainsi obtenu (et on mémorisera) :

$$\boxed{\begin{array}{cc} \dfrac{\partial r}{\partial x} = \cos(\theta) & \dfrac{\partial r}{\partial y} = \sin(\theta) \\[2ex] \dfrac{\partial \theta}{\partial x} = -\dfrac{\sin(\theta)}{r} & \dfrac{\partial \theta}{\partial y} = \dfrac{\cos(\theta)}{r} \end{array}}$$

On note au passage le comportement des dérivées partielles. On a obtenu $\dfrac{\partial x}{\partial r} = \cos(\theta) = \dfrac{\partial r}{\partial x}$ et donc

$$\frac{\partial r}{\partial x} \neq \frac{1}{\dfrac{\partial x}{\partial r}},$$

alors que pour les « d droits », $\dfrac{\mathrm{d}x}{\mathrm{d}y} = \dfrac{1}{\dfrac{\mathrm{d}y}{\mathrm{d}x}}$ (dérivée d'une réciproque). De même, $\dfrac{\mathrm{d}z}{\mathrm{d}x} = \dfrac{\mathrm{d}z}{\mathrm{d}y} \times \dfrac{\mathrm{d}y}{\mathrm{d}x}$ (dérivée d'une composée) alors que $\dfrac{\partial g}{\partial r} \times \dfrac{\partial r}{\partial x} \neq \dfrac{\partial g}{\partial x}$.

• On peut maintenant exprimer les dérivées partielles de f par rapport à x et y en fonction des dérivées partielles de g par rapport à r et θ :

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial r} \times \frac{\partial r}{\partial x} + \frac{\partial g}{\partial \theta} \times \frac{\partial \theta}{\partial x} = \cos(\theta)\frac{\partial g}{\partial r} - \frac{\sin(\theta)}{r}\frac{\partial g}{\partial \theta}$$

et

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial r} \times \frac{\partial r}{\partial y} + \frac{\partial g}{\partial \theta} \times \frac{\partial \theta}{\partial y} = \sin(\theta)\frac{\partial g}{\partial r} + \frac{\cos(\theta)}{r}\frac{\partial g}{\partial \theta}$$

3.4 Fonctions de classe C^1

-

Le programme officiel adopte la définition suivante :

DÉFINITION 8. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit Ω un ouvert non vide de E. Soit f une fonction définie sur Ω à valeurs dans F.

f est de **classe** C^1 **sur** Ω si et seulement si f est différentiable sur Ω et de plus l'application $\mathrm{d}f\ :\ a \mapsto \mathrm{d}f_a$ est continue sur Ω (à valeurs dans $\mathscr{L}(E, F)$).

L'ensemble des fonctions de classe C^1 sur Ω à valeurs dans F se note $C^1(\Omega, F)$.

Par définition, on a $C^1(\Omega, F) \subset D^1(\Omega, F)$ et on obtient le graphique définitif (toute implication non écrite étant fausse) :

$$\begin{tikzcd}
 & f \text{ de classe } C^1 \text{ sur } \Omega \arrow[d, Rightarrow] & \\
 & \text{Différentiabilité en } a \in \Omega \arrow[dl, Rightarrow] \arrow[ddr, Rightarrow] & \\
\begin{array}{c}\text{Dérivabilité suivant}\\ \text{tout vecteur en } a\end{array} \arrow[d, Rightarrow] & & \\
\begin{array}{c}\text{Existence de}\\ \text{dérivées partielles en } a\end{array} \arrow[dr, Rightarrow] & & \text{Continuité en } a \arrow[dl, Rightarrow] \\
 & \text{Continuité partielle en } a &
\end{tikzcd}$$

Dans la pratique, ce n'est pas la définition 8 qui est mise en œuvre car :

Théorème 22. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit Ω un ouvert non vide de E. Soit f une fonction définie sur Ω à valeurs dans F.

f est de classe C^1 sur Ω si et seulement si f admet sur Ω des dérivées partielles premières sur Ω et chacune des fonctions dérivées partielles est continue sur Ω.

Démonstration . Pour simplifier les notations, nous supposerons que $E = \mathbb{R}^n = \Omega$ et $F = \mathbb{R}^p$. On munit $\mathbb{R}^n$ et $\mathbb{R}^p$ de la norme infini que l'on note $\| \ \|$ dans les deux cas.

• Supposons que f admette sur $\mathbb{R}^n$ des dérivées partielles par rapport à chacune de ses variables et que ces dérivées partielles soient des fonctions continues sur $\mathbb{R}^n$.
Soit $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Vérifions que f est différentiable en a. L'idée est de ne faire varier qu'une variable à la fois pour disposer de l'égalité des accroissements finis : pour $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$,

$$\begin{aligned}
f(a+h) - f(a) &= f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1, \ldots, a_n) \\
&= (f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1, a_2 + h_2 \ldots, a_n + h_n)) \\
&+ (f(a_1, a_2 + h_2, \ldots, a_n + h_n) - f(a_1, a_2, a_3 + h_3, \ldots, a_n + h_n)) \\
&+ \ldots \\
&+ (f(a_1, \ldots, a_{i-1}, a_i + h_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - f(a_1, \ldots, a_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n)) \ \ (= D_i) \\
&+ \ldots \\
&+ (f(a_1, \ldots, a_{n-1}, a_n + h_n) - f(a_1, \ldots, a_{n-1}, a_n)).
\end{aligned}$$

(si $i = 1$, par convention, $D_i = f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1, \ldots, a_n)$). Pour $i \in [\![1, n-1]\!]$ posons

$$\forall x \in \mathbb{R},\ g_i(x) = f(a_1, \ldots, a_{i-1}, x, a_{i+1} + h_{i+1}, \ldots, a_n + h_n).$$

de sorte que $D_i = g_i(a_i + h_i) - g_i(a_i)$. g_i est une fonction définie sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$ (si $\Omega \subsetneq \mathbb{R}^n$, g_i est définie au moins sur un intervalle ouvert contenant a_i obtenu en projetant Ω sur $\mathrm{Vect}(e_i)$), dérivable sur $\mathbb{R}$. D'après le théorème des accroissements finis, il existe un réel c_i compris entre a_i et $a_i + h_i$ (fonction de a et de h) tel que

$$D_i = g_i(a_i + h_i) - g_i(a_i) = h_i g_i'(c_i) = h_i \frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n).$$

On en déduit que $f(a+h) - f(a) = \displaystyle\sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n)$ puis que

$$\begin{aligned}
\left\| f(a+h) - f(a) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) \right\| &= \left\| \sum_{i=1}^{n} h_i \left(\frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) \right) \right\| \\
&\leqslant \sum_{i=1}^{n} |h_i| \left\| \frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) \right\| \\
&\leqslant \|h\| \sum_{i=1}^{n} \left\| \frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) \right\|.
\end{aligned}$$

Maintenant, pour chaque $i \in [\![1, n]\!]$, c_i est compris entre a_i et $a_i + h_i$ et donc c_i tend vers a_i quand h_i tend vers 0 puis $(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n)$ tend vers $(a_1, \ldots, a_n)$ quand h tend vers 0. Puisque par hypothèse, les dérivées partielles de f sont continues sur Ω et donc en a, l'expression
$\displaystyle\sum_{i=1}^{n} \left\| \frac{\partial f}{\partial x_i}(a_1, \ldots, a_{i-1}, c_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a_1, \ldots, a_n) \right\|$ tend vers 0 quand h tend vers 0.
Finalement, $\displaystyle\frac{1}{\|h\|}\left(f(a+h) - f(a) - \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) \right)$ tend vers 0 quand h tend vers 0 ou encore

$$f(a+h) \underset{h \to 0}{=} f(a) + \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a) + o(h).$$

Puisque la fonction $(h_1, \ldots, h_n) \mapsto \displaystyle\sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a)$ est linéaire, on a montré que f est différentiable en a et que

$$\forall a \in \mathbb{R}^n,\ \forall h \in \mathbb{R}^n,\ df_a(h) = \sum_{i=1}^{n} h_i \frac{\partial f}{\partial x_i}(a).$$

Enfin, puisque les fonctions $\dfrac{\partial f}{\partial x_i}$, $1 \leqslant i \leqslant n$, sont continues sur $\mathbb{R}^n$, il en est de même de $df\ :\ a \mapsto df_a = \displaystyle\sum_{i=1}^{n} dx_i \frac{\partial f}{\partial x_i}(a)$ et donc f est de classe C^1 sur $\mathbb{R}^n$.

• Réciproquement, supposons f de classe C^1 sur $\mathbb{R}^n$. En particulier, f est différentiable en chaque $a \in \mathbb{R}^n$ et donc admet des dérivées partielles par rapport à chacune des ses variables en tout $a \in \mathbb{R}^n$.

On munit $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$ de la norme :

$$\forall \varphi \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^p),\ |||\varphi||| = \text{Sup}\left\{\frac{\|\varphi(x)\|}{\|x\|},\ x \in \mathbb{R}^n \setminus \{0\}\right\},$$

(norme subordonnée). On sait que pour tout $x \in \mathbb{R}^n$, $\|\varphi(x)\| \leqslant |||\varphi||| \|x\|$.
Soit $i \in [\![1, n]\!]$. Soit $a \in \mathbb{R}^n$. Pour h élément de $\mathbb{R}^n$ donné, en notant $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$,

$$\left\|\frac{\partial f}{\partial x_i}(a+h) - \frac{\partial f}{\partial x_i}(a)\right\| = \|(df_{a+h} - df_a)(e_i)\| \leqslant |||df_{a+h} - df_a|||\ \|e_i\|.$$

Par hypothèse, df est continue sur $\mathbb{R}^n$ et donc $|||df_{a+h} - df_a(h)|||\ \|e_i\|$ tend vers 0 quand h tend vers 0. Il en est de même que $\left\|\frac{\partial f}{\partial x_i}(a+h) - \frac{\partial f}{\partial x_i}(a)\right\|$ ce qui démontre la continuité de $\frac{\partial f}{\partial x_i}$ en a.

On a montré que f admet des dérivées partielles par rapport à chacune de ses variables en chaque point a de $\mathbb{R}^n$ et que ses dérivées partielles sont continues sur $\mathbb{R}^n$.

❑

Exercice 4. Soit f la fonction définie sur $\mathbb{R}^2$ par :

$$\forall (x, y) \in \mathbb{R}^2,\ f(x, y) = \begin{cases} \dfrac{xy\left(x^2 - y^2\right)}{x^2 + y^2} \text{ si } (x, y) \neq (0, 0) \\ 0 \text{ si } (x, y) = 0 \end{cases}.$$

Montrer que f est de classe C^1 sur $\mathbb{R}^2$.

Rappel : on a déjà montré que f est continue sur $\mathbb{R}^2$ (exercice n° 1, page 5) et admet sur $\mathbb{R}^2$ des dérivées partielles par rapport à chacune de ses deux variables (exercice n° 2, page 8) définies par : $\forall (x, y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} \dfrac{y\left(x^4 + 4x^2y^2 - y^4\right)}{\left(x^2 + y^2\right)^2} \text{ si } (x, y) \neq (0, 0) \\ 0 \text{ si } (x, y) = 0 \end{cases} \quad \text{et} \quad \frac{\partial f}{\partial y}(x, y) = \begin{cases} \dfrac{x\left(x^4 - 4x^2y^2 - y^4\right)}{\left(x^2 + y^2\right)^2} \text{ si } (x, y) \neq (0, 0) \\ 0 \text{ si } (x, y) = 0 \end{cases}.$$

Solution 4. Les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ en tant que quotient de fonctions continues sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ dont le dénominateur ne s'annule pas sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Pour $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$,

$$\begin{aligned} \left|\frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(0, 0)\right| &= \frac{|y|\left|x^4 + 4x^2y^2 - y^4\right|}{\left(x^2 + y^2\right)^2} \leqslant \frac{|y|\left(x^4 + 4x^2y^2 + y^4\right)}{\left(x^2 + y^2\right)^2} \\ &\leqslant \frac{|y|\left(2x^4 + 4x^2y^2 + 2y^4\right)}{\left(x^2 + y^2\right)^2} = 2|y|. \end{aligned}$$

$2|y|$ tend vers 0 quand (x, y) tend vers $(0, 0)$ et donc $\displaystyle\lim_{\substack{(x,y)\to(0,0) \\ (x,y)\neq(0,0)}} \frac{\partial f}{\partial x}(x, y) = 0 = f(0, 0)$. Ainsi, la fonction $\frac{\partial f}{\partial x}$ est continue en $(0, 0)$ et finalement $\frac{\partial f}{\partial x}$ est continue sur $\mathbb{R}^2$. On montre de même que la fonction $\frac{\partial f}{\partial y}$ est continue sur $\mathbb{R}^2$.

En résumé, f admet sur $\mathbb{R}^2$ des dérivées partielles par rapport à chacune de ses variables et ces dérivées partielles sont continues sur $\mathbb{R}^2$. On a montré que f est de classe C^1 sur $\mathbb{R}^2$.

Sinon, on a immédiatement les théorèmes généraux usuels :

Théorème 23. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit Ω un ouvert non vide de E.

f est de classe C^1 sur Ω si et seulement si ses applications coordonnées dans une base de F donnée sont de classe C^1 sur Ω.

En particulier, si $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$, f est de classe C^1 sur Ω si et seulement si ses applications composantes le sont.

Théorème 24. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit Ω un ouvert non vide de E.

Pour tout $(f,g) \in \left(C^1(\Omega,F)\right)^2$ et tout $(\lambda,\mu) \in \mathbb{R}^2$, $\lambda f + \mu g \in C^1(\Omega,F)$.

$C^1(\Omega,F)$ est un sous-espace vectoriel de $D^1(\Omega,F)$.

Théorème 25. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie non nulle. Soit Ω un ouvert non vide de E.

Pour tout $(f,g) \in \left(C^1(\Omega,\mathbb{R})\right)^2$, $f \times g \in C^1(\Omega,\mathbb{R})$.

Théorème 26. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie non nulle. Soit Ω un ouvert non vide de E.

Pour tout $(f,g) \in \left(C^1(\Omega,\mathbb{R})\right)^2$ tel que g ne s'annule pas sur Ω, $\dfrac{f}{g} \in C^1(\Omega,\mathbb{R})$.

Théorème 27. Soient E, F et G trois $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit Ω un ouvert non vide de E et Ω' un ouvert non vide de F. Soient f une application de Ω vers F et g une application de Ω' vers G telles que $f(\Omega) \subset \Omega'$.

Si f est de classe C^1 sur Ω et g est de classe C^1 sur Ω', alors $g \circ f$ est de classe C^1 sur Ω.

3.5 Dérivation et intégration le long d'un arc

Théorème 28. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit f une application d'un ouvert non vide Ω de E vers F, différentiable sur Ω. Soit $\gamma\ :\ t \mapsto \gamma(t)$ une application définie et dérivable sur un intervalle I de $\mathbb{R}$ à valeurs dans E tel que $\forall t \in I$, $\gamma(t) \in \Omega$. Alors

$$\forall t \in I,\ (f\circ\gamma)'(t) = df_{\gamma(t)}(\gamma'(t)).$$

En particulier, si a est un point de Ω, si v est un vecteur non nul de E et si pour tout réel t appartenant à un intervalle ouvert I contenant 0, $\gamma(t) = a + tv$, alors

$$\forall t \in I,\ (f\circ\gamma)'(t) = df_{a+tv}(v) = D_v f(a+tv).$$

⇛ **Commentaire** . *Une application $\gamma\ :\ t \mapsto \gamma(t)$ d'un intervalle I de $\mathbb{R}$ à valeurs dans E s'appelle un arc paramétré de E. L'ensemble des points $(t,\gamma(t))$, $t \in I$, est une courbe (paramétrée) tracée dans l'ouvert Ω.*

Démonstration . γ est dérivable sur I et donc différentiable sur I. De plus, $\gamma(I) \subset \Omega$ et f est différentiable sur Ω. Donc, $f\circ\gamma$ est différentiable sur I ou encore $f\circ\gamma$ est dérivable sur I. De plus, pour $t \in I$ et $h \in \mathbb{R}$,

$$h(f\circ\gamma)'(t) = d(f\circ\gamma)_t(h) = df_{\gamma(t)}\left(d\gamma_t(h)\right) = df_{\gamma(t)}(h\gamma'(t)) = hdf_{\gamma(t)}(\gamma'(t))$$

et on obtient le résultat en prenant $h = 1$.

❏

⇛ **Commentaire** . *Ce qui précède n'est pas dans la mentalité du programme officiel. La formule $d(g\circ f)_a = dg_{f(a)} \circ df_a$ $(*)$ (et sa longue démonstration) n'est pas à proprement parler au programme et nous sommes donc censés montrer directement le théorème 28 sans connaître $(*)$.*

La formule $(*)$ a permis entre autre de découvrir la règle de la chaine (voir page 20 et 21) donnant les dérivées partielles de fonctions composées. Mais, on peut arriver à montrer cette règle de la chaine uniquement avec les formules du théorème 28 (ce que nous ne ferons pas). C'est ce que voulait le programme officiel.

Théorème 29. Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit f une application d'un ouvert non vide Ω de E vers F, de classe C^1 sur Ω.
Soit $(a,b) \in \Omega^2$. Soit $\gamma\ :\ t \mapsto \gamma(t)$ un arc paramétré défini et de classe C^1 sur $[0,1]$ à valeurs dans E tel que
- $\gamma(0) = a$ et $\gamma(1) = b$
- $\forall t \in I$, $\gamma(t) \in \Omega$.

Alors

$$f(b) - f(a) = \int_0^1 df_{\gamma(t)}(\gamma'(t))\ dt.$$

Démonstration . L'application $(f \circ \gamma)'$ est continue sur le segment $[0, 1]$. D'après le théorème précédent,

$$f(b) - f(a) = f(\gamma(1)) - f(\gamma(0)) = \int_0^1 (f \circ \gamma)'(t)\, dt = \int_0^1 df_{\gamma(t)}(\gamma'(t))\, dt.$$

❑

3.6 Cas particulier des fonctions numériques

(Une fonction numérique est une fonction à valeurs dans $\mathbb{R}$).

3.6.1 Egalité des accroissements finis. Caractérisation des fonctions constantes

Théorème 30. (Egalité des accroissements finis pour les fonctions numériques)

Soit f une application d'un ouvert non vide Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$, différentiable sur Ω.

Pour tout $(a, b) \in \Omega^2$ tel que $[a, b] \subset \Omega$, il existe $\lambda \in [0, 1]$ tel que $f(b) - f(a) = df_{(1-\lambda)a+\lambda b}(b - a)$.

⇛ **Commentaire** . *La condition* $[a, b] \subset \Omega$ *est assurée pour tout* $(a, b) \in \Omega^2$ *si* Ω *est convexe.*

Démonstration . Soit $(a, b) \in \Omega^2$ tel que $[a, b] \subset \Omega$. Donc, pour $t \in [0, 1]$, $(1 - t)a + tb \in \Omega$. Pour $t \in [0, 1]$, on pose $g(t) = f((1 - t)a + tb)$. Puisque f est de classe C^1 sur Ω, g est une application de $[0, 1]$ dans $\mathbb{R}$, de classe C^1 sur $[0, 1]$. D'après le théorème des accroissements finis, il existe $\lambda \in [0, 1]$ tel que $g(1) - g(0) = g'(\lambda)$.

On note γ l'application de $[0, 1]$ dans Ω définie par : $\forall t \in [0, 1]$, $\gamma(t) = (1 - t)a + tb$. On a alors $g = f \circ \gamma$ puis, pour $t \in [0, 1]$,

$$g'(\lambda) = (f \circ \gamma)'(\lambda) = df_{\gamma(\lambda)}(\gamma'(\lambda)) = df_{(1-\lambda)a+\lambda b}(b - a).$$

L'égalité $g(1) - g(0) = g'(\lambda)$ s'écrit alors

$$f(b) - f(a) = df_{(1-\lambda)a+\lambda b}(b - a).$$

❑

Théorème 31.

1) Soit f une application d'un ouvert non vide connexe par arcs Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$, de classe C^1 sur Ω.

f est constante sur Ω si et seulement si $df = 0$.

2) Soit f une application d'un ouvert non vide connexe par arcs Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans un $\mathbb{R}$-espace vectoriel F de dimension finie non nulle, de classe C^1 sur Ω.

f est constante sur Ω si et seulement si $df = 0$.

Démonstration . On sait que si f est constante sur un ouvert Ω quelconque, alors $df = 0$. On établit la réciproque dans le cas particulier d'un ouvert non vide convexe et on admet le cas plus général des ouverts non vides connexes par arcs (comme le prévoit le programme officiel).

1) Réciproquement, soit f une application de classe C^1 sur Ω à valeurs dans $\mathbb{R}$ telle que $df = 0$. Soit $(a, b) \in \Omega^2$. Puisque Ω est convexe, $[a, b] \subset \Omega$. D'après le théorème précédent, il existe $\lambda \in [0, 1]$ tel que $f(b) - f(a) = df_{(1-\lambda)a+\lambda b}(b - a) = 0$ car $df_{(1-\lambda)a+\lambda b} = 0$.

Donc, $\forall (a, b) \in \Omega^2$, $f(a) = f(b)$ ou encore f est constante sur Ω.

2) On applique le 1) à chacune des fonctions coordonnées de f dans une base donnée de F.

❑

3.6.2 Gradient

On rappelle un résultat issu du chapitre « Espaces euclidiens ». Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien.
Soit $u \in E$. L'application $\varphi_u\ :\ x \mapsto \langle u, x\rangle$ est une forme linéaire sur E.
Réciproquement, pour toute forme linéaire φ sur E, il existe un vecteur u et un seul tel que $\forall x \in E$, $\varphi(x) = \langle u, x\rangle$.

Définition 9. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien de dimension finie non nulle. Soit f une application définie et différentiable sur un ouvert non vide Ω de E à valeurs dans $\mathbb{R}$. Soit $a \in \Omega$.

Le **vecteur gradient** de f en a est l'unique vecteur u de E tel que $\forall h \in E$, $df_a(h) = \langle u, h\rangle$. Il est noté $\nabla f(a)$ (∇ se lit nabla) ou $\overrightarrow{\text{grad}}(f)_a$. Donc,

$$\forall h \in E,\ df_a(h) = \langle \nabla f(a), h\rangle.$$

⇛ **Commentaire** . *La forme ∇ est la forme Δ renversée. Δf est le laplacien de f, utilisé en physique. La forme ∇ a été choisie pour rendre certaines formules facilement mémorisable. Le nom « nabla » fait référence à une harpe grecque qui avait approximativement cette forme.*

On donne maintenant les coordonnées du vecteur gradient dans une base orthonormée de l'espace euclidien $(E, \langle\ ,\ \rangle)$.

Théorème 32. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien de dimension finie non nulle. Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base **orthonormée** de cet espace.
Soit f une application définie et différentiable sur un ouvert non vide Ω de E à valeurs dans $\mathbb{R}$. Soit $a \in \Omega$. Alors,

$$\nabla f(a) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) e_i,$$

où les $\dfrac{\partial f}{\partial x_i}(a)$ sont les dérivées partielles de f en a dans la base $\mathscr{B}$.

Démonstration . Posons $\nabla f(a) = \displaystyle\sum_{i=1}^{n} x_i e_i$ où $(x_1, \ldots, x_n) \in \mathbb{R}^n$. Puisque la base $\mathscr{B}$ est orthonormée, on sait que pour $i \in [\![1, n]\!]$,

$$x_i = \langle \nabla f(a), e_i\rangle = df_a(e_i) = \frac{\partial f}{\partial x_i}(a).$$

❏

Théorème 33. Si $\nabla f(a) \neq 0$, ∇f est positivement colinéaire au vecteur unitaire selon lequel la dérivée de f en a est maximale.

Démonstration . Soit v un vecteur unitaire. D'après l'inégalité de Cauchy-Schwarz,

$$D_v f(a) = df_a(v) = \langle \nabla f, \rangle \leqslant |\langle \nabla f, v\rangle| \leqslant \|\nabla f\|\|v\| = \|\nabla f\|.$$

De plus, si $\nabla f \neq 0$, on sait qu'on a l'égalité si et seulement si v est positivement colinéaire au vecteur ∇f ce qui équivaut à $v = \dfrac{1}{\|\nabla f\|}\nabla f$.
Ainsi, $\text{Max}\{D_v f(a),\ v \text{ unitaire}\}$ existe et $\text{Max}\{D_v f(a),\ v \text{ unitaire}\} = \|\nabla f\|$, le maximum étant atteint pour $v_0 = \dfrac{1}{\|\nabla f\|}\nabla f$. Mais alors $\nabla f = \|\nabla f\| v_0$ est positivement colinéaire au vecteur unitaire selon lequel la dérivée de f en a est maximale.

❏

Essayons de donner une interprétation plus concrète du gradient en dimension 3 comme dirigeant à tout instant une « ligne de plus grande pente » sur une surface d'équation $z = f(x, y)$.

Soit $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction numérique de deux variables. La représentation graphique de la fonction f est l'ensemble d'équation $z = f(x, y)$. Quand f « est simple », cette représentation graphique est interprétée comme une surface. Aucun cours sur les surfaces n'est prévu dans le programme officiel à l'exception de la détermination du plan tangent en un point de cette surface ce qui sera fait plus loin.

Soit donc (S) la surface d'équation $z = f(x, y)$. Soit (x_0, y_0, z_0) un point de (S) (donc $z_0 = f(x_0, y_0)$). Soit $a = (x_0, y_0)$ (a est la projection sur le plan (xOy) du point (x_0, y_0, z_0)). On va chercher dans quelle direction et quel sens démarrer, à partir du point (x_0, y_0, z_0), pour maximiser la croissance de $z = f(x, y)$. On se donne donc un vecteur unitaire $v = (\alpha, \beta)$ du plan (xOy) et on cherche à maximiser en t_0 la croissance de $z(t) = f(a + tv) = f(x_0 + t\alpha, y_0 + t\beta)$ ou encore on cherche v maximisant la dérivée en t_0 de la fonction $t \mapsto z(t) = f(a + tv)$. Mais

$$z'(t_0) = D_v f(a) = df_a(v) = \langle \nabla f(a), v\rangle.$$

D'après le théorème 33, si $\nabla f(a) \neq 0$, $z'(t_0)$ est maximal si et seulement si v est colinéaire à $\nabla f(a)$ et de même sens que $\nabla f(a)$. Le gradient de f en a indique donc la direction horizontale à prendre pour se déplacer sur (S) suivant une ligne de plus grande pente.

Si (S) est la modélisation d'une montagne assez lisse (comme un sommet du massif central ou des Vosges mais pas un sommet des Alpes ou des Pyrénées), le vecteur gradient de f indique à tout instant la projection horizontale de la direction d'un torrent dévalant la montagne.

3.7 Vecteurs tangents à une partie

DÉFINITION 10. Soit E un $\mathbb{R}$-espace vectoriel de dimension finie non nulle. Soit X une partie non vide de E. Soient $x \in X$ et v un vecteur de E.

v est tangent à X en x s'il existe $\varepsilon > 0$ et un arc γ défini sur $]-\varepsilon, \varepsilon[$, dérivable en 0, à valeurs dans X, tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

⇛ **Commentaire**.

◇ *Les conditions « $\gamma(0) = x$ et γ est à valeurs dans X » peuvent se réénoncer de manière plus imagée sous la forme « l'arc γ « est » un arc tracé sur l'ensemble X et passant par le point x ». Les tangentes en chaque point d'un arc tracé sur X (la tangente à l'arc γ en $\gamma(0) = x$ est dirigée par $\gamma'(0) = v$) sont par définition des tangentes à X.*

◇ *Le programme officiel n'impose pas à un vecteur tangent v d'être non nul, ce qui paraît douteux, le vecteur nul « ne dirigeant rien » et en particulier ne dirigeant pas une tangente.*

Notation. L'ensemble des vecteurs tangents à X en x se note T_xX.

Exemple. Prenons pour X un sous-espace affine $\mathscr{F}$ de E. Notons F la direction de $\mathscr{F}$. Soit $x \in \mathscr{F}$.

Soit $v \in F$. L'application $\gamma : t \mapsto x + tv$ est une application définie et de classe C^1 sur $\mathbb{R}$, à valeurs dans $\mathscr{F}$ (par définition d'un sous-espace affine), vérifiant $\gamma(0) = x$ et $\gamma'(0) = v$ (et plus généralement, pour tout $t \in \mathbb{R}$, $\gamma'(t) = v$). v est donc un vecteur tangent à $\mathscr{F}$ ou encore $v \in T_x\mathscr{F}$.

Inversement, soit v un élément de $T_x\mathscr{F}$. Il existe $\varepsilon > 0$ puis $\gamma :]-\varepsilon, \varepsilon[\to E$, telle que $\gamma(0) = x$, pour tout $t \in]-\varepsilon, \varepsilon[$, $\gamma(t) \in \mathscr{F}$ et $\gamma'(0) = v$. Pour tout $t \in]-\varepsilon, \varepsilon[\setminus\{0\}$, $\dfrac{1}{t}(\gamma(t) - \gamma(0))$ est un élément de F. En particulier, pour n suffisamment grand, $t_n = \dfrac{1}{n} \in]-\varepsilon, \varepsilon[\setminus\{0\}$ puis $x_n = \dfrac{1}{t_n}(\gamma(t_n) - \gamma(0)) \in F$.

Puisque E est de dimension finie et que F est un sous-espace vectoriel de E, on sait que F est un fermé de E. Puisque γ est dérivable en 0,

$$v = \gamma'(0) = \lim_{t \to 0} \frac{1}{t}(\gamma(t) - \gamma(0)) = \lim_{n \to +\infty} \frac{1}{t_n}(\gamma(t_n) - \gamma(0)) = \lim_{n \to +\infty} x_n.$$

Ainsi, v est la limite d'une suite convergente d'éléments de F et donc $v \in F$ puisque F est fermé. On a montré que

$$\forall x \in \mathscr{F},\ T_x\mathscr{F} = F.$$

L'ensemble des vecteurs tangents à un sous-espace affine est donc l'ensemble des vecteurs de sa direction. ❏

On suppose maintenant que X est un sous-ensemble de E d'équation $f(x) = 0$ où f est une fonction différentiable sur un certain ouvert Ω de E (X est donc un sous-ensemble de Ω). On va vérifier que le vecteur gradient de f en un point x de X est orthogonal à tout vecteur tangent à X en x.

Théorème 34. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien de dimension finie non nulle n. Soit f une application définie et différentiable sur un ouvert Ω d'un espace eucilidien $(E, \langle\ ,\ \rangle)$, à valeurs dans $\mathbb{R}$.

Soit $X = \{x \in \Omega /\ f(x) = 0\}$. Soit $x \in X$. S'il existe un arc γ passant par x et tracé sur X comme dans la définition 10, alors $\nabla f(x)$ est orthogonal à $\gamma'(0)$.

DÉMONSTRATION. Avec les notations de la définition 10, pour tout $t \in]-\varepsilon, \varepsilon[$, $f(\gamma(t)) = 0$. En dérivant cette égalité, on obtient (voir théorème 28, page 25)

$$\forall t \in]-\varepsilon, \varepsilon[,\ df_{\gamma(t)}\left(\gamma'(t)\right) = 0.$$

En particulier,

$$0 = df_{\gamma(0)}\left(\gamma'(0)\right) = df_x(v) = \langle \nabla f(x), v \rangle.$$

Le vecteur gradient $\nabla f(x)$ est effectivement orthogonal à un vecteur v tangent à X en x. ❏

Analysons le cas de la dimension 3. On suppose que l'ensemble d'équation $z = f(x, y)$ est une « vraie surface » que l'on note (S). L'équation $z = f(x, y)$ s'écrit encore $g(x, y, z) = 0$ où $g(x, y, z) = z - f(x, y)$. Soit (x_0, y_0, z_0) un point de (S). Le vecteur gradient de g en (x_0, y_0, z_0) est orthogonal à toute tangente en (x_0, y_0, z_0) à (S). On en déduit que le vecteur gradient de g en (x_0, y_0, z_0) est un vecteur normal au plan tangent à (S) en (x_0, y_0, z_0) (le programme officiel ne prévoit aucun cours sur les surfaces de $\mathbb{R}^3$ et nous ne définirons donc pas précisément la notion de surface et encore moins le plan tangent à une surface en un point de cette surface). Puisque $\nabla g_{(x_0,y_0,z_0)} = \begin{pmatrix} -\dfrac{\partial f}{\partial x}(x_0, y_0, z_0) \\ -\dfrac{\partial f}{\partial y}(x_0, y_0, z_0) \\ 1 \end{pmatrix}$, on obtient

Théorème 35. Soit f une application d'un ouvert non vide Ω de $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$, différentiable sur Ω.

Soit (S) l'ensemble d'équation $z = f(x, y)$. Soit (x_0, y_0, z_0) un point de (S). Dans un repère orthonormé de $\mathbb{R}^3$, une équation du plan tangent à (S) en (x_0, y_0, z_0) est :

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0, z_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0, z_0)(y - y_0).$$

Exemple. Soit $R > 0$. Pour $(x, y) \in \mathbb{R}^2$ tel que $x^2 + y^2 \leqslant R^2$, posons $f(x, y) = \sqrt{R^2 - x^2 - y^2}$ et notons (S) l'ensemble d'équation $z = \sqrt{R^2 - x^2 - y^2}$ dans un certain repère orthonormé de l'espace ((S) est une demi-sphère de centre O et de rayon R). Soit (x_0, y_0, z_0) un point de (S). Une équation du plan tangent à (S) en (x_0, y_0, z_0) est

$$z - z_0 = -\frac{x_0}{\sqrt{R^2 - x_0^2 - y_0^2}}(x - x_0) - \frac{x_0}{\sqrt{R^2 - x_0^2 - y_0^2}}(y - y_0),$$

ou encore

$$z - z_0 = -\frac{x_0}{z_0}(x - x_0) - \frac{x_0}{z_0}(y - y_0),$$

ou enfin

$$x_0(x - x_0) + y_0(y - y_0) + z_0(z - z_0) = 0.$$

Un vecteur normal au plan tangent à (S) en $M_0(x_0, y_0, z_0)$ est le vecteur de coordonnées (x_0, y_0, z_0) c'est-à-dire le vecteur $\overrightarrow{OM_0}$. Ainsi, le plan tangent à une sphère en un point est perpendiculaire au rayon correspondant. ❏

On énonce maintenant un théorème fournissant un procédé pour déterminer T_xX et dont le démonstration est totalement hors programme.

Théorème 36. Soit $(E, \| \ \|)$ un $\mathbb{R}$-espace vectoriel normé de dimension finie n non nulle.
Soit g une application d'un ouvert non vide Ω de E à valeurs dans $\mathbb{R}$, de classe C^1 sur Ω.

Soit X une partie non vide de Ω puis x un élément de X. On suppose que $dg_x \neq 0$. Alors

$$T_xX = \mathrm{Ker}\,(dg_x).$$

On note en particulier que T_xX est un sous-espace vectoriel de E. Plus précisément, dg_x étant une application linéaire non nulle de E vers $\mathbb{R}$ ou encore dg_x étant une forme linéaire non nulle, $T_xX = \mathrm{Ker}\,(dg_x)$ est un hyperplan de E.

Si de plus, E est muni d'un produit scalaire, le résultat s'énonce plus simplement :

Théorème 37. Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien de dimension finie n non nulle.
Soit g une application d'un ouvert non vide Ω de E à valeurs dans $\mathbb{R}$, de classe C^1 sur Ω.

Soit X une partie non vide de Ω puis x un élément de X. On suppose que $\nabla g(x) \neq 0$. Alors

$$T_xX = (\nabla g(x))^{\perp}.$$

T_xX est donc l'hyperplan de vecteur normal $\nabla g(x)$.

Démonstration. On sait que pour tout $h \in E$, $dg_x(h) = \langle \nabla(x), h\rangle$ et donc

$$h \in T_xX \Leftrightarrow dg_x(h) = 0 \Leftrightarrow \langle \nabla(x), h\rangle = 0 \Leftrightarrow h \in (\nabla g(x))^{\perp}.$$

❏

Analysons par exemple, le cas d'une sphère d'un espace euclidien. On se place dans $E = \mathbb{R}^n$ muni du produit scalaire canonique. On se donne $a = (a_1, \dots, a_n) \in E$ et $R > 0$. On considère $\mathscr{S}$ la sphère de centre a et de rayon R. Dans la base canonique de $\mathbb{R}^n$, une équation de $\mathscr{S}$ est

$$(x_1 - a_1)^2 + \dots + (x_n - a_n)^2 = R^2 \ (E)$$

ou encore $f(x_1,\dots,x_n)=0$ où, pour tout $x=(x_1,\dots,x_n)\in\mathbb{R}^n$, $f(x_1,\dots,x_n)=(x_1-a_1)^2+\dots+(x_n-a_n)^2-R^2$. L'hyperplan tangent à la sphère $\mathscr{S}$ en un point $b=(b_1,\dots,b_n)$ de cette sphère est l'hyperplan affine passant par b de vecteur normal $\nabla f(b)=(2(b_1-a_1),\dots,2(b_n-a_n))\neq 0$. Un autre vecteur normal est $\dfrac{1}{2}\nabla f(b)=(b_1-a_1,\dots,b_n-a_n)=b-a$ et donc, une équation de l'hyperplan tangent à $\mathscr{S}$ en b est

$$(b_1-a_1)(x_1-b_1)+\dots+(b_n-a_n)(x_n-b_n)=0\quad (E').$$

Maintenant, puisque $b\in\mathscr{S}$,

$$\begin{aligned}(b_1-a_1)(x_1-b_1)&+\dots+(b_n-a_n)(x_n-b_n)\\ &=(b_1-a_1)(x_1-a_1+a_1-b_1)+\dots+(b_n-a_n)(x_n-a_n+a_n-b_n)\\ &=(b_1-a_1)(x_1-a_1)+\dots+(b_n-a_n)(x_n-a_n)-R^2\end{aligned}$$

et donc, l'équation (E') peut se réécrire

$$(b_1-a_1)(x_1-a_1)+\dots+(b_n-a_n)(x_n-a_n)=R^2\quad (E').$$

Le remplacement mécanique d'un terme $(x_i-a_i)^2$ par $(b_i-a_i)(x_i-a_i)$ pour passer d'une équation de la sphère à une équation de l'hyperplan tangent en un point de cette sphère, rentre dans le cadre plus général d'une règle appelée « règle de dédoublement des termes », règle valable pour toute surface du second degré.

4 Dérivées partielles d'ordre supérieur

4.1 Fonctions de classe C^k

Nous allons maintenant dériver les dérivées partielles. On suppose que $f\;:\;\Omega\subset\mathbb{R}^n\to\mathbb{R}^p$ admet des dérivées partielles $\dfrac{\partial f}{\partial x_i}$, $1\leqslant i\leqslant n$, sur Ω. Soit $i\in[\![1,n]\!]$. Si la fonction $\dfrac{\partial f}{\partial xi}$ admet sur Ω une dérivée partielle par rapport à sa j-ème variable, $j\in[\![1,n]\!]$, on la note $\dfrac{\partial^2 f}{\partial x_j\partial x_i}$ (en faisant attention à l'ordre des variables) au lieu de $\dfrac{\partial}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)$. Dans le cas particulier où $j=i$, on écrit $\dfrac{\partial^2 f}{\partial x_i^2}$. On peut ainsi calculer n^2 dérivées partielles secondes.

Dans la notation $\dfrac{\partial^2 f}{\partial x_j\partial x_i}$, la dernière dérivation effectuée est écrite à gauche et la première est écrite à droite. Il existe une autre notation, plus dangereuse mais plus condensée : $\dfrac{\partial^2 f}{\partial x_j\partial x_i}=f''_{x_i x_j}$. Dans ce cas, la première dérivation est écrite à gauche et la deuxième à droite.

L'ordre des dérivations est essentiel comme le montre l'exercice suivant :

Exercice 5. Soit f la fonction définie sur $\mathbb{R}^2$ par : $\forall(x,y)\in\mathbb{R}^2$, $f(x,y)=\left\{\begin{array}{l} y^2\sin\left(\dfrac{x}{y}\right)\ \text{si}\ y\neq 0\\ 0\ \text{si}\ y=0\end{array}\right.$.

Vérifier que $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)$ et $\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$ existent et sont différents.

Solution 5. On pose $D=\{(x,0),\ x\in\mathbb{R}\}$ puis $\Omega=\mathbb{R}^2\setminus D$.

• f est définie sur $\mathbb{R}^2$.

• f est de classe C^1 sur Ω en vertu de théorèmes généraux et pour $(x,y)\in\Omega$,

$$\frac{\partial f}{\partial x}(x,y)=y\cos\left(\frac{x}{y}\right)\ \text{et}\ \frac{\partial f}{\partial y}(x,y)=2y\sin\left(\frac{x}{y}\right)-x\cos\left(\frac{x}{y}\right).$$

• Etudions l'existence et la valeur éventuelle de $\dfrac{\partial f}{\partial x}(x_0,0)$, x_0 réel donné. Pour $x\neq x_0$,

$$\frac{f(x,0)-f(x_0,0)}{x-x_0}=\frac{0-0}{x-x_0}=0.$$

Donc $\dfrac{f(x,x_0)-f(x_0,0)}{x-x_0}$ tend vers 0 quand x tend vers x_0. On en déduit que $\dfrac{\partial f}{\partial x}(x_0,0)$ existe et $\dfrac{\partial f}{\partial x}(x_0,0)=0$. Finalement, la fonction $\dfrac{\partial f}{\partial x}$ est définie sur $\mathbb{R}^2$ par

$$\forall (x,y)\in\mathbb{R}^2,\ \frac{\partial f}{\partial x}(x,y)=\left\{\begin{array}{l}y\cos\left(\dfrac{x}{y}\right)\ \text{si}\ y\neq 0\\ 0\ \text{si}\ y=0\end{array}\right..$$

• Etudions l'existence et la valeur éventuelle de $\dfrac{\partial f}{\partial y}(x_0,0)$, x_0 réel donné. Pour $y\neq 0$,

$$\frac{f(x_0,y)-f(x_0,0)}{y-0}=\frac{y^2\sin\left(\dfrac{x_0}{y}\right)}{y}=y\sin\left(\frac{x_0}{y}\right).$$

On en déduit que $\left|\dfrac{f(x_0,y)-f(x_0,0)}{y-0}\right|\leqslant |y|$ puis que $\dfrac{f(x_0,y)-f(x_0,0)}{y-0}$ tend vers 0 quand y tend vers 0. Par suite, $\dfrac{\partial f}{\partial y}(x_0,0)$ existe et $\dfrac{\partial f}{\partial y}(x_0,0)=0$. Finalement, la fonction $\dfrac{\partial f}{\partial y}$ est définie sur $\mathbb{R}^2$ par

$$\forall (x,y)\in\mathbb{R}^2,\ \frac{\partial f}{\partial y}(x,y)=\left\{\begin{array}{l}2y\sin\left(\dfrac{x}{y}\right)-x\cos\left(\dfrac{x}{y}\right)\ \text{si}\ y\neq 0\\ 0\ \text{si}\ y=0\end{array}\right..$$

• Etudions l'existence et la valeur éventuelle de $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)$. Pour $x\neq 0$,

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

Donc $\dfrac{\dfrac{\partial f}{\partial y}(x,0)-\dfrac{\partial f}{\partial y}(0,0)}{x-0}$ tend vers 0 quand x tend vers 0. On en déduit que $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)$ existe et $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)=0$.

• Etudions l'existence et la valeur éventuelle de $\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$. Pour $y\neq 0$,

$$\frac{\dfrac{\partial f}{\partial x}(0,y)-\dfrac{\partial f}{\partial x}(0,0)}{y-0}=\frac{y\cos\left(\dfrac{0}{y}\right)}{y}=1.$$

Donc $\dfrac{\dfrac{\partial f}{\partial x}(0,y)-\dfrac{\partial f}{\partial x}(0,0)}{y-0}$ tend vers 1 quand y tend vers 0. On en déduit que $\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$ existe et $\dfrac{\partial^2 f}{\partial y\partial x}(0,0)=1$. On a montré que $\dfrac{\partial^2 f}{\partial x\partial y}(0,0)$ et $\dfrac{\partial^2 f}{\partial y\partial x}(0,0)$ existent et sont différents.

On définit ensuite par récurrence les dérivées partielles d'ordre $k\geqslant 2$ par $\dfrac{\partial^k f}{\partial x_{i_k}\ldots\partial x_{i_1}}=\dfrac{\partial}{\partial x_{i_k}}\left(\dfrac{\partial^{k-1} f}{\partial x_{i_{k-1}}\ldots\partial x_{i_1}}\right)$ (en cas d'existence) et on peut poser la définition suivante :

Définition 11. Soit f une application d'un ouvert non vide Ω d'un espace de dimension finie non nulle E vers un espace de dimension finie non nulle F.

Soit $k\geqslant 1$. f est de classe C^k sur Ω si et seulement si f admet sur Ω des dérivées partielles jusqu'à l'ordre k par rapport à tout i-uplet de variables, $1\leqslant i\leqslant k$, et les dérivées partielles k-èmes sont continues sur Ω. On note $C^k(\Omega,F)$ l'ensemble des fonctions de classe C^k sur Ω à valeurs dans F.

f est de classe C^∞ sur Ω si et seulement si, pour tout $k\in\mathbb{N}^*$, f est de classe C^k sur Ω. On note $C^\infty(\Omega,F)$ l'ensemble des fonctions de classe C^∞ sur Ω à valeurs dans F.

On peut démontrer et on admettra que les théorèmes généraux usuels (combinaisons linéaires, produits, quotients, composées ...) restent valables pour les fonctions de classe C^k ou C^∞. En particulier,

Théorème 38. Soient E et F deux $\mathbb{R}$-espaces de dimensions finies non nulles. Soit Ω un ouvert non vide de E.

$\forall k \in \mathbb{N}^*$, $C^k(\Omega, F)$ est un $\mathbb{R}$-espace vectoriel.

$C^\infty(\Omega, F)$ est un $\mathbb{R}$-espace vectoriel.

$\forall k \in \mathbb{N}^*$, $C^\infty(\Omega, F) \underset{\neq}{\subset} \ldots C^{k+1}(\Omega, F) \underset{\neq}{\subset} C^k(\Omega, F) \underset{\neq}{\subset} \ldots \underset{\neq}{\subset} C^0(\Omega, F)$.

et

Théorème 39.

Un polynôme à n variables réelles est de classe C^∞ sur $\mathbb{R}^n$.

Une fraction rationnelle à n variables réelles est de classe C^∞ sur son domaine de définition.

4.2 Théorème de Schwarz

Théorème 40. (théorème de Schwarz)

Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie non nulle. Soit Ω un ouvert non vide de E. Soit f une application de Ω vers F. Soit $k \geqslant 2$.

Si f est de classe C^k sur Ω, alors $\forall (i_1, \ldots, i_k) \in [\![1, n]\!]^k$, $\forall \sigma \in \mathscr{S}_k$,

$$\frac{\partial^k f}{\partial x_{\sigma(i_k)} \ldots \partial x_{\sigma(i_k)}} = \frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}.$$

Démonstration . (le programme officiel précise que cette démonstration n'est pas exigible.)

Puisque toute permutation est produit de transpositions, il suffit de prouver le théorème de Schwarz quand σ est une transposition. Ceci ramène au cas des fonctions de deux variables que l'on dérive partiellement 2 fois. De même, si on démontre le résultat pour chacune des fonctions coordonnées, le théorème de Schwarz sera démontré.

On suppose donc dorénavant que f est une application d'un ouvert non vide Ω de $\mathbb{R}^2$ vers $\mathbb{R}$, de classe C^2 sur Ω, et on montre que $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$.

Soit $a = (a_1, a_2) \in \Omega$. On va établir que

$$\frac{\partial^2 f}{\partial y \partial x}(a) = \lim_{\substack{(h_1, h_2) \to 0 \\ h_1 h_2 \neq 0}} \frac{f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1, a_2)}{h_1 h_2} \quad (*).$$

Posons $u(h) = u(h_1, h_2) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1, a_2)$. Puisque Ω est ouvert et que $a \in \Omega$, u est définie sur un voisinage V de $(0, 0)$.

Soit $h = (h_1, h_2) \in V$ fixé. D'après l'égalité des accroissements finis appliqué à la fonction $v : t \mapsto u(t, h_2)$, il existe un réel c_1 (fonction de h_1 et h_2) compris entre 0 et h_1 tel que

$$u(h_1, h_2) = v(h_1) - v(0) = h_1 v'(c_1) = h_1 \left(\frac{\partial f}{\partial x}(a_1 + c_1, a_2 + h_2) - \frac{\partial f}{\partial x}(a_1 + c_1, a_2) \right)$$

De même, l'égalité des accroissements finis appliqué à la fonction $w : t \mapsto u(a_1 + c_1, t)$ fournit un réel c_2 (fonction de h_1 et h_2) compris entre 0 et h_2 tel que

$$u(h_1, h_2) = h_1 \left(\frac{\partial f}{\partial x}(a_1 + c_1, a_2 + h_2) - \frac{\partial f}{\partial x}(a_1 + c_1, a_2) \right) = h_1 h_2 \frac{\partial^2 f}{\partial y \partial x}(a_1 + c_1, a_2 + c_2).$$

Par suite, pour $(h_1, h_2) \in V$ tel que $h_1 h_2 \neq 0$,

$$\frac{u(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial y \partial x}(a_1 + c_1, a_2 + c_2).$$

Puisque c_1 est compris entre 0 et h_1 et c_2 est compris entre 0 et h_2, le couple $(a_1 + c_1, a_2 + c_2)$ tend vers le couple $a = (a_1, a_2)$ quand le couple $h = (h_1, h_2)$ tend vers le couple $(0, 0)$. Puisque par hypothèse, $\dfrac{\partial^2 f}{\partial y \partial x}$ est continue sur Ω et en particulier en a, $\dfrac{\partial^2 f}{\partial y \partial x}(a_1 + c_1, a_2 + c_2)$ tend vers $\dfrac{\partial^2 f}{\partial y \partial x}(a_1, a_2)$ quand (h_1, h_2) tend vers $(0, 0)$. Ceci démontre $(*)$.

Mais alors, par symétrie des rôles,

$$
\begin{aligned}
\frac{\partial^2 f}{\partial y \partial x}(a) &= \lim_{\substack{(h_1,h_2)\to 0\\ h_1 h_2 \neq 0}} \frac{f(a_1+h_1,a_2+h_2)-f(a_1,a_2+h_2)-f(a_1+h_1,a_2)+f(a_1,a_2)}{h_1 h_2}\\
&= \lim_{\substack{(h_1,h_2)\to 0\\ h_1 h_2 \neq 0}} \frac{f(a_1+h_1,a_2+h_2)-f(a_1+h_1,a_2)-f(a_1,a_2+h_2)+f(a_1,a_2)}{h_1 h_2}\\
&= \frac{\partial^2 f}{\partial x \partial y}(a).
\end{aligned}
$$

❏

Exercice 6. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une application de classe C^2 sur $\mathbb{R}^2$. Le laplacien de f est la fonction Δf définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$ définie par :

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

Calculer Δf en polaires (ne se poser aucun problème théorique, les problèmes de calcul se suffisent à eux mêmes).

Solution 6. Pour $(x,y) \in \mathbb{R}^2$, on pose $f(x,y) = f(r\cos\theta, r\sin\theta) = g(r,\theta)$ ou encore si pour $(r,\theta) \in \mathbb{R}^2$, $\varphi(r,\theta) = (r\cos\theta, r\sin\theta)$, on pose $g = f \circ \varphi$. Il s'agit de calculer Δf en fonction des différentes dérivées partielles de g.

On reprend les calculs laissés en attente à la page 22 : $\dfrac{\partial f}{\partial x} = \cos\theta\dfrac{\partial g}{\partial r} - \dfrac{\sin\theta}{r}\dfrac{\partial g}{\partial \theta}$ et $\dfrac{\partial f}{\partial y} = \sin\theta\dfrac{\partial g}{\partial r} + \dfrac{\cos\theta}{r}\dfrac{\partial g}{\partial \theta}$.

On rappelle aussi que $\dfrac{\partial r}{\partial x} = \cos\theta$, $\dfrac{\partial r}{\partial y} = \sin\theta$, $\dfrac{\partial \theta}{\partial x} = -\dfrac{\sin\theta}{r}$ et $\dfrac{\partial \theta}{\partial y} = \dfrac{\cos\theta}{r}$. D'après le théorème de Schwarz, $\dfrac{\partial^2 g}{\partial r \partial \theta} = \dfrac{\partial^2 g}{\partial \theta \partial r}$ et donc

$$
\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\cos\theta\frac{\partial g}{\partial r} - \frac{\sin\theta}{r}\frac{\partial g}{\partial \theta}\right)\\
&= (-\sin\theta)\frac{\partial \theta}{\partial x}\frac{\partial g}{\partial r} + \cos\theta\left(\frac{\partial r}{\partial x}\frac{\partial^2 g}{\partial r^2} + \frac{\partial \theta}{\partial x}\frac{\partial^2 g}{\partial r \partial \theta}\right) - \cos\theta\frac{\partial \theta}{\partial x}\frac{1}{r}\frac{\partial g}{\partial \theta} - \sin\theta\left(-\frac{1}{r^2}\right)\frac{\partial r}{\partial x}\frac{\partial g}{\partial \theta} - \frac{\sin\theta}{r}\left(\frac{\partial r}{\partial x}\frac{\partial^2 g}{\partial r \partial \theta} + \frac{\partial \theta}{\partial x}\frac{\partial^2 g}{\partial \theta^2}\right)\\
&= \frac{\sin^2\theta}{r}\frac{\partial g}{\partial r} + \cos^2\theta\frac{\partial^2 g}{\partial r^2} - \frac{\cos\theta\sin\theta}{r}\frac{\partial^2 g}{\partial r \partial \theta} + \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial g}{\partial \theta} - \frac{\sin\theta}{r}\left(\cos\theta\frac{\partial^2 g}{\partial r \partial \theta} - \frac{\sin\theta}{r}\frac{\partial^2 g}{\partial \theta^2}\right)\\
&= \frac{\sin^2\theta}{r}\frac{\partial g}{\partial r} + \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial g}{\partial \theta} + \cos^2\theta\frac{\partial^2 g}{\partial r^2} - \frac{2\cos\theta\sin\theta}{r}\frac{\partial^2 g}{\partial r \partial \theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 g}{\partial \theta^2}
\end{aligned}
$$

puis

$$
\begin{aligned}
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\sin\theta\frac{\partial g}{\partial r} + \frac{\cos\theta}{r}\frac{\partial g}{\partial \theta}\right)\\
&= \cos\theta\frac{\partial \theta}{\partial y}\frac{\partial g}{\partial r} + \sin\theta\left(\frac{\partial r}{\partial y}\frac{\partial^2 g}{\partial r^2} + \frac{\partial \theta}{\partial y}\frac{\partial^2 g}{\partial r \partial \theta}\right) + (-\sin\theta)\frac{\partial \theta}{\partial y}\frac{1}{r}\frac{\partial g}{\partial \theta} + \cos\theta\left(-\frac{1}{r^2}\right)\frac{\partial r}{\partial y}\frac{\partial g}{\partial \theta} + \frac{\cos\theta}{r}\left(\frac{\partial r}{\partial y}\frac{\partial^2 g}{\partial r \partial \theta} + \frac{\partial \theta}{\partial y}\frac{\partial^2 g}{\partial \theta^2}\right)\\
&= \frac{\cos^2\theta}{r}\frac{\partial g}{\partial r} + \sin^2\theta\frac{\partial^2 g}{\partial r^2} + \frac{\cos\theta\sin\theta}{r}\frac{\partial^2 g}{\partial r \partial \theta} - \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial g}{\partial \theta} + \frac{\cos\theta}{r}\left(\sin\theta\frac{\partial^2 g}{\partial r \partial \theta} + \frac{\cos\theta}{r}\frac{\partial^2 g}{\partial \theta^2}\right)\\
&= \frac{\cos^2\theta}{r}\frac{\partial g}{\partial r} - \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial g}{\partial \theta} + \sin^2\theta\frac{\partial^2 g}{\partial r^2} + \frac{2\cos\theta\sin\theta}{r}\frac{\partial^2 g}{\partial r \partial \theta} + \frac{\cos^2\theta}{r^2}\frac{\partial^2 g}{\partial \theta^2}
\end{aligned}
$$

et finalement

$$\Delta f = \frac{\partial^2 g}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2} + \frac{1}{r}\frac{\partial g}{\partial r}.$$

⇛ **Commentaire**. *Nous vous laissons le soin de calculer le laplacien du passage en sphérique en dimension* 3.

5 Optimisation

5.1 Extrema des fonctions numériques

DÉFINITION 12. Soit f une application d'un sous-ensemble non vide D d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$. Soit $a \in D$.

f admet un minimum (global) en a si et seulement si $\forall x \in D,\ f(x) \geqslant f(a)$.
f admet un maximum (global) en a si et seulement si $\forall x \in D,\ f(x) \leqslant f(a)$.

f admet un minimum local en a si et seulement si il existe un voisinage V de a dans E tel que $\forall x \in D \cap V,\ f(x) \geqslant f(a)$.
f admet un maximum local en a si et seulement si il existe un voisinage V de a dans E tel que $\forall x \in D \cap V,\ f(x) \leqslant f(a)$.

5.2 Etude à l'ordre 1

Rappelons tout d'abord quelques erreurs classiques dans la mémorisation d'un certain nombre de théorèmes de maths sup.

• Le « théorème » : « soit $f\ :\ I \to \mathbb{R}$, I intervalle de $\mathbb{R}$, et $x_0 \in I$. f admet un extremum local en x_0 si et seulement si $f'(x_0) = 0$ » est faux. Déjà, il manque l'hypothèse de dérivabilité. Par exemple, la fonction $x \mapsto |x-1|$ admet un minimum en 1 mais n'est pas dérivable en 1.

• Améliorons. Le « théorème » : « soit $f\ :\ I \to \mathbb{R}$, dérivable sur I intervalle de $\mathbb{R}$, et $x_0 \in I$. Si f admet un extremum en x_0 alors $f'(x_0) = 0$ » est faux. La fonction $f\ :\ \begin{array}{ccc} [0,1] & \to & \mathbb{R} \\ x & \mapsto & x^2 \end{array}$ admet un maximum en 1 et pourtant $f'(1) = 2 \neq 0$. Ici, il manque l'hypothèse « I ouvert ».

• Le « théorème » : « soit $f\ :\ I \to \mathbb{R}$, dérivable sur I intervalle de $\mathbb{R}$, et $x_0 \in I$. Si $f'(x_0) = 0$ alors f admet un extremum local en x_0 » est faux. Soit $f\ :\ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$. $f'(0) = 0$ et pourtant la fonction cube n'admet en 0 ni un minimum local, ni un maximum local.

• Deux « théorèmes vrais » sont : « soit $f\ :\ I \to \mathbb{R}$, dérivable sur I intervalle ouvert de $\mathbb{R}$, et $x_0 \in I$. Si f admet un extremum local en x_0 alors $f'(x_0) = 0$ » ou aussi « soit $f\ :\ I \to \mathbb{R}$, dérivable sur I intervalle de $\mathbb{R}$, et x_0 un point intérieur à I. Si f admet un extremum local en x_0 alors $f'(x_0) = 0$ ».

• Un autre « théorème vrai » est : « soit $f\ :\ I \to \mathbb{R}$, dérivable sur I intervalle ouvert de $\mathbb{R}$, et $x_0 \in I$. Si f' s'annule en x_0 en changeant de signe, alors f admet un extremum local en x_0 ».

Maintenant, nous généralisons :

Théorème 41. Soit f une application d'un **ouvert** non vide Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$, **différentiable** sur Ω. Soit $a \in \Omega$.

Si f admet un extremum local en a, alors $df_a = 0$.

DÉMONSTRATION . On fixe une base $\mathscr{B} = (e_1, \dots, e_n)$ de E. Soit $a \in \Omega$. On suppose que f admet en a un extremum local en $a = a_1e_1 + \ldots + a_ne_n$.

Soit $i \in [\![1, n]\!]$. Soit $g_i\ :\ t \mapsto f(a_1e_1 + \ldots + a_{i-1}e_{i-1} + te_i + a_{i+1}e_{i+1} + \ldots + a_ne_n)$ la i-ème application partielle de f dans la base $\mathscr{B}$ en a. Le fait que Ω soit ouvert assure le fait que g_i est définie (au moins) sur un intervalle ouvert non vide I contenant a_i.

g_i est dérivable sur l'intervalle ouvert I (car f est différentiable sur Ω) et g_i admet un extremum local en a_i (car f admet un extremum local en a). Donc, $g_i'(a_i) = 0$ ou encore $\dfrac{\partial f}{\partial x_i}(a) = 0$.

Puisque toutes les dérivées partielles en a sont nulles, on a encore $df_a = 0$.

❑

DÉFINITION 13. Soit f une application d'un ouvert non vide Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$, différentiable sur Ω.

Un point $a \in \Omega$ en lequel $df_a = 0$ s'appelle un **point critique de** f.

Un point critique de f est donc un point en lequel toutes les dérivées partielles s'annulent car $df_a = 0 \Leftrightarrow \forall i \in [\![1, n]\!],\ \dfrac{\partial f}{\partial x_i}(a) = 0$.

Passons maintenant à l'utilisation du théorème 41. Il fournit une **condition nécessaire** d'existence d'un extremum local et malheureusement pas une **condition suffisante**. Donc, si on recherche les extrema locaux d'une fonction numérique (différentiable sur un ouvert), on commence par déterminer les points critiques de f puis, pour chacun des points obtenus, on se débrouille sans théorème, « à la main », pour savoir si oui ou non, on est en présence d'un extremum local.

Exercice 7. Déterminer les extrema (locaux ou globaux) de la fonction f définie sur $\mathbb{R}^2$ par : $\forall(x,y)\in\mathbb{R}^2$, $f(x,y) = -2(x-y)^2+x^4+y^4$.

Solution 7. La fonction f est de classe C^1 sur $\mathbb{R}^2$ en tant que polynôme à plusieurs variables et $\mathbb{R}^2$ est un ouvert de $\mathbb{R}^2$. Donc, si f admet un extremum local en un point $(x_0,y_0)\in\mathbb{R}^2$, (x_0,y_0) est un point critique de f. Soit $(x,y)\in\mathbb{R}^2$.

$$df_{(x,y)} = 0 \Leftrightarrow \left\{\begin{array}{l}\dfrac{\partial f}{\partial x}(x,y)=0\\ \dfrac{\partial f}{\partial y}(x,y)=0\end{array}\right. \Leftrightarrow \left\{\begin{array}{l}-4(x-y)+4x^3=0\\ 4(x-y)+4y^3=0\end{array}\right. \Leftrightarrow \left\{\begin{array}{l}x^3+y^3=0\\ -(x-y)+x^3=0\end{array}\right. \Leftrightarrow \left\{\begin{array}{l}y=-x\\ x^3-2x=0\end{array}\right.$$
$$\Leftrightarrow (x,y)\in\left\{(0,0),\left(\sqrt{2},-\sqrt{2}\right),\left(-\sqrt{2},\sqrt{2}\right)\right\}.$$

• Etude en $\left(\sqrt{2},-\sqrt{2}\right)$. $f\left(\sqrt{2},-\sqrt{2}\right)=-8$ puis, pour tout $(x,y)\in\mathbb{R}^2$,

$$\begin{aligned}f(x,y)-f\left(\sqrt{2},-\sqrt{2}\right) &= x^4+y^4-2x^2-2y^2+4xy+8 \geqslant x^4+y^4-2x^2-2y^2-2\left(x^2+y^2\right)+8\\ &= x^4-4x^2+y^4-4y^2+8=\left(x^2-2\right)^2+\left(y^2-2\right)^2\geqslant 0.\end{aligned}$$

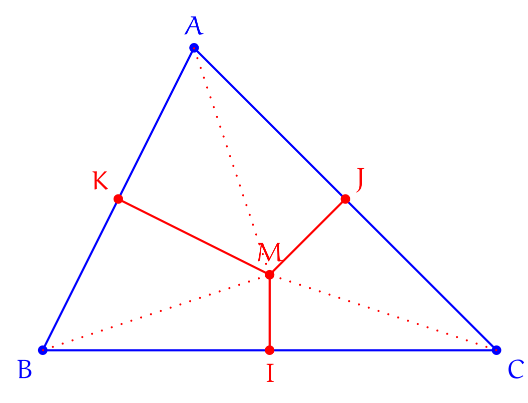
et donc f admet un minimum global en $\left(\sqrt{2},-\sqrt{2}\right)$ égal à -8.

• Etude en $\left(-\sqrt{2},\sqrt{2}\right)$. Pour tout $(x,y)\in\mathbb{R}^2$, $f(-x,-y)=f(x,y)$ et donc f admet aussi un minimum global en $\left(-\sqrt{2},\sqrt{2}\right)$ égal à -8.

• Etude en $(0,0)$. $f(0,0)=0$. Pour $x\neq 0$, $f(x,x)=2x^4>0$ et donc f prend des valeurs strictement supérieures à $f(0,0)$ dans tout voisinage de $(0,0)$. Pour $x\in\left]-\sqrt{2},\sqrt{2}\right[\setminus\{0\}$, $f(x,0)=x^4-2x^2=x^2(x^2-2)<0$ et f prend des valeurs strictement inférieures à $f(0,0)$ dans tout voisinage de $(0,0)$. Finalement, f n'admet pas d'extremum local en $(0,0)$.

Exercice 8. Maximum du produit des distances d'un point M, intérieur à un triangle ABC, aux côtés de ce triangle.

Solution 8. On pose $BC=a$, $CA=b$ et $AB=c$ et on note $\mathscr{A}$ l'aire du triangle ABC. Soit M un point intérieur au triangle ABC. On note I, J et K les projetés orthogonaux de M sur les droites (BC), (CA) et (AB) respectivement.



On pose $u=$ aire de MBC, $v=$ aire de MCA et $w=$ aire de MAB. On a

$$d(M,(BC))\times d(M,(CA))\times d(M,(AB)) = MI\times MJ\times MK = \frac{2u}{a}\times\frac{2v}{b}\times\frac{2w}{c}=\frac{8}{abc}uv(\mathscr{A}-u-v).$$

Il s'agit alors de trouver le maximum de la fonction $f\ :\ (u,v)\mapsto uv(\mathscr{A}-u-v)$ sur le domaine

$$T=\left\{(u,v)\in\mathbb{R}^2/\ u\geqslant 0,\ v\geqslant 0 \text{ et } u+v\leqslant\mathscr{A}\right\}.$$

T est un compact de $\mathbb{R}^2$. En effet :

- $\forall(u,v)\in T^2$, $\|(u,v)\|_1=u+v\leqslant\mathscr{A}$ et donc T est bornée.
- Les applications $\varphi_1\ :\ (u,v)\mapsto u$, $\varphi_2\ :\ (u,v)\mapsto v$ et $\varphi_3\ :\ (u,v)\mapsto u+v$ sont continues sur $\mathbb{R}^2$ en tant que formes linéaires sur un espace de dimension finie. Donc les ensembles $P_1=\{(u,v)\in\mathbb{R}^2/\ u\geqslant 0\}=\varphi_1^{-1}([0,+\infty[)$, $P_2=\{(u,v)\in\mathbb{R}^2/\ v\geqslant 0\}=\varphi_2^{-1}([0,+\infty[)$ et $P_3=\{(u,v)\in\mathbb{R}^2/\ u+v\leqslant\mathscr{A}\}=\varphi_3^{-1}(]-\infty,\mathscr{A}])$ sont des fermés de $\mathbb{R}^2$

en tant qu'images réciproques de fermés par des applications continues. On en déduit que $T = P_1 \cap P_2 \cap P_3$ est un fermé de $\mathbb{R}^2$ en tant qu'intersection de fermés de $\mathbb{R}^2$.

Puisque T est un fermé borné de $\mathbb{R}^2$, T est un compact de $\mathbb{R}^2$ puisque $\mathbb{R}^2$ est de dimension finie et d'après le théorème de Borel-Lebesgue.

f est continue sur le compact T à valeurs dans $\mathbb{R}$ en tant que polynôme à plusieurs variables et donc f admet un maximum sur T.

Pour tout (u, v) appartenant à la frontière de T, on a $f(u, v) = 0$. Comme f est strictement positive sur $\overset{\circ}{T} = \{(u, v) \in \mathbb{R}^2 /\ u > 0,\ v > 0 \text{ et } u + v < 0\}$, f admet son maximum dans $\overset{\circ}{T}$. Puisque f est de classe C^1 sur $\overset{\circ}{T}$ qui est un ouvert de $\mathbb{R}^2$, si f admet un maximum en $(u_0, v_0) \in \overset{\circ}{T}$, (u_0, v_0) est nécessairement un point critique de f. Soit $(u, v) \in \overset{\circ}{T}$.

$$\left\{\begin{array}{l} \dfrac{\partial f}{\partial u}(u, v) = 0 \\ \dfrac{\partial f}{\partial v}(u, v) = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} v(\mathscr{A} - 2u - v) = 0 \\ u(\mathscr{A} - u - 2v) = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} 2u + v = \mathscr{A} \\ u + 2v = \mathscr{A} \end{array}\right. \Leftrightarrow u = v = \frac{\mathscr{A}}{3}.$$

Puisque f admet un point critique et un seul à savoir $(u_0, v_0) = \left(\dfrac{\mathscr{A}}{3}, \dfrac{\mathscr{A}}{3}\right)$, f admet son maximum en ce point et ce maximum vaut $f(u_0, v_0) = \dfrac{\mathscr{A}^3}{27}$. Le maximum du produit des distances d'un point M intérieur au triangle ABC aux cotés de ce triangle est donc $\dfrac{8\mathscr{A}^3}{27abc}$.

On étudie maintenant une situation plus générale. On dispose d'une fonction numérique f différentiable sur un ouvert Ω. On cherche alors une condition nécessaire pour que la restriction de f à une partie X de Ω admette un extremum local en x élément donné de X. Un tel problème est un problème d'optimisation sous contrainte. On commence par établir :

Théorème 42. Soit f une application d'un **ouvert** non vide Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$, **différentiable** sur Ω. Soient X une partie non vide non vide de Ω puis x un élément donné de X.

Si f_X admet un extremum local en x, alors df_x s'annule en tout vecteur tangent à X en x ou encore $T_xX \subset \operatorname{Ker}(df_x)$.

Démonstration. Supposons par exemple que f_X admet un minimum local en x. Soit v un vecteur non nul tangent à X en x. Il existe $\varepsilon > 0$ puis $\gamma\ :\ ]-\varepsilon, \varepsilon[\to X$, dérivable en 0, tel que $\gamma(0) = x$ et $\gamma'(0) = v$.

Par hypothèse, pour tout $t \in]-\varepsilon, \varepsilon[$, $f(\gamma(t)) \geqslant f(x) = f(\gamma(0))$. Donc, la fonction $g\ :\ t \mapsto f(\gamma(t))$ admet un minimum local en 0. Puisque g est dérivable sur l'intervalle ouvert $]-\varepsilon, \varepsilon[$, $g'(0) = 0$ ou encore $df_{\gamma(0)}(\gamma'(0)) = 0$ ou enfin $df_x(v) = 0$.

❏

On en déduit le « théorème des extrema liés » appelé aussi le « théorème d'optimisation sous contrainte » :

Théorème 43. Soient f et g deux fonctions définies et de classe C^1 sur Ω ouvert non vide d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$. Soit $X = g^{-1}(\{0\}) = \{x \in \Omega /\ g(x) = 0\}$ l'ensemble des zéros de la fonction g. On suppose X non vide.

Soit $x \in X$. Si $dg_x \neq 0$ et si f_X admet un extremum local en x, alors df_x est colinéaire à dg_x.

Démonstration. D'après le théorème 36, page 29, $\operatorname{Ker}(dg_x) = T_xX$. D'après le théorème précédent, $\operatorname{Ker}(df_x) \subset T_xX = \operatorname{Ker}(dg_x)$.

Si $df_x = 0$, alors df_x est colinéaire à dg_x. Dorénavant, $df_x \neq 0$. $\operatorname{Ker}(df_x)$ et $\operatorname{Ker}(dg_x)$ sont deux hyperplans de E tels que $\operatorname{Ker}(df_x) \subset \operatorname{Ker}(dg_x)$. Par égalité des dimensions, on en déduit que $\operatorname{Ker}(df_x) = \operatorname{Ker}(dg_x)$. On note H cet hyperplan.

Soient e un vecteur non dans $\operatorname{Ker}(dg_x)$ (e existe car $\operatorname{Ker}(dg_x) \subsetneq E$) puis $D = \operatorname{Vect}(e)$. On sait que $E = H \oplus D$. Soit $\lambda = \dfrac{df_x(e)}{dg_x(e)}$ de sorte que $df_x(e) = \lambda dg_x(e)$. On a alors $(df_x)_{/D} = \lambda (dg_x)_{/D}$ (car ces deux formes linéaires coïncident sur une base de D) et $(df_x)_{/H} = \lambda (dg_x)_{/H}$ (car $(df_x)_{/H}$ et $(dg_x)_{/H}$ sont nulles). Les deux formes linéaires df_x et λdg_x coïncident sur deux sous-espaces supplémentaires de E et on en déduit que $df_x = \lambda dg_x$.

❏

⇛ **Commentaire**. *Si* $(E, \langle\ ,\ \rangle)$ *est euclidien, la conclusion du théorème précédent peut s'écrire sous la forme :* $\nabla f(x)$ *est colinéaire à* $\nabla g(x)$.

Voyons maintenant à travers un exercice comment nous pouvons utiliser ce théorème dans la pratique.

Exercice 9. Trouver le maximum de $x^2+y^2+z^2-2xy+2xz-2yz$ sur $\mathscr{S}=\left\{(x,y,z)\in\mathbb{R}^3/\ x^2+y^2+z^2=1\right\}$.

Solution 9. Pour $(x,y,z)\in\mathbb{R}^3$, posons $f(x,y,z)=x^2+y^2+z^2-2xy+2xz-2yz$ et $g(x,y,z)=x^2+y^2+z^2-1$. On note que pour tout $(x,y,z)\in\mathbb{R}^3$, $f(-x,-y,-z)=f(x,y,z)$.

f est continue sur $\mathbb{R}^3$, en tant que polynôme, à valeurs dans $\mathbb{R}$ et $\mathscr{S}$ est un compact de $\mathbb{R}^3$ (sphère unité pour la norme euclidienne). Donc, f admet un maximum en (au moins) un point (x_0,y_0,z_0) de $\mathscr{S}$.

f et g sont de classe C^1 sur $\mathbb{R}^3$ qui est un ouvert de $\mathbb{R}^3$ et $\mathscr{S}=g^{-1}(\{0\})$. De plus, pour tout $(x,y,z)\in\mathscr{S}$, $\nabla g(x,y,z)=\begin{pmatrix}2x\\2y\\2z\end{pmatrix}\neq 0$ (car $x^2+y^2+z^2=1\neq 0$). Donc, d'après le théorème des extremas liés, en un point (x,y,z) de $\mathscr{S}$ en lequel f atteint son maximum, il existe $\lambda\in\mathbb{R}$ tel que $\nabla f(x,y,z)=\lambda\nabla g(x,y,z)$.

Cette dernière égalité s'écrit encore $\left\{\begin{array}{l}2x-2y+2z=2\lambda x\\2y-2x-2z=2\lambda y\\2z+2x-2y=2\lambda z\end{array}\right.$ c'est-à-dire $\left\{\begin{array}{l}(1-\lambda)x-y+z=0\\-x+(1-\lambda)y-z=0\\x-y+(1-\lambda)z=0\end{array}\right.$ (S).

Si $\det(S)\neq 0$, (S) admet l'unique solution $(x,y,z)=(0,0,0)$. Mais $(0,0,0)$ n'est pas un point de $\mathscr{S}$ et donc, en (x_0,y_0,z_0), on a nécessairement $\det(S)=0$. Calculons donc $\det(S)$. En développant suivant la première colonne, on obtient

$$\begin{aligned}\det(S)&=\begin{vmatrix}1-\lambda&-1&1\\-1&1-\lambda&-1\\1&-1&1-\lambda\end{vmatrix}=(1-\lambda)\left((1-\lambda)^2-1\right)+(\lambda)+(\lambda)=(1-\lambda)(-\lambda)(2-\lambda)+2\lambda\\&=-\lambda((\lambda-1)(\lambda-2)-2)=\lambda^2(\lambda-3).\end{aligned}$$

Donc, $\lambda\in\{0,3\}$.

• Si $\lambda=0$, $(S)\Leftrightarrow\left\{\begin{array}{l}x-y+z=0\\-x+y-z=0\\x-y+z=0\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}z=-x+y\\-x+y-(-x+y)=0\\x-y+(-x+y)=0\end{array}\right.\Leftrightarrow z=-x+y.$
On a alors $f(x,y,z)=f(x,y,-x+y)=x^2+y^2+(-x+y)^2-2xy+2x(-x+y)-2y(-x+y)=0$. Puisque $f(1,0,0)=1>0$, le maximum de f sur $\mathscr{S}$ n'est pas 0.

• Les points en lesquels f atteint son maximum sur $\mathscr{S}$ sont nécessairement obtenus dans le cas $\lambda=3$. Dans ce cas, $(S)\Leftrightarrow\left\{\begin{array}{l}-2x-y+z=0\\-x-2y-z=0\\x-y-2z=0\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}z=2x+y\\-x-2y-(2x+y)=0\\x-y-2(2x+y)=0\end{array}\right.\Leftrightarrow\left\{\begin{array}{l}y=-x\\z=x\end{array}\right.$.

De plus, la condition $(x,y,z)\in\mathscr{S}$ fournit $x^2+(-x)^2+x^2=1$ puis $x=\pm\dfrac{1}{\sqrt{3}}$.

On a obtenu les deux points $\pm\left(\dfrac{1}{\sqrt{3}},-\dfrac{1}{\sqrt{3}},\dfrac{1}{\sqrt{3}}\right)$. Enfin, $f\left(\pm\left(\dfrac{1}{\sqrt{3}},-\dfrac{1}{\sqrt{3}},\dfrac{1}{\sqrt{3}}\right)\right)=3\times\dfrac{1}{3}+2\times\left(\dfrac{1}{3}+\dfrac{1}{3}-\dfrac{1}{3}\right)=\dfrac{5}{3}$.

Le maximum de f sur $\mathscr{S}$ est donc égal à $\dfrac{5}{3}$.

5.3 Etude à l'ordre 2

Le paragraphe précédent nous a fourni une condition nécessaire pour avoir un extremum : si f est de classe C^1 sur un ouvert Ω à valeurs dans $\mathbb{R}$ et si f admet un extremum local en a, alors $df_a=0$. On cherche maintenant une condition suffisante. Comme dans le cas, des fonctions réelles d'une variable, un développement limité d'ordre 2 peut suffire pour conclure.

DÉFINITION 14. Soit f une application d'un ouvert non vide Ω d'un $\mathbb{R}$-espace E de dimension finie non nulle à valeurs dans $\mathbb{R}$, de classe C^2 sur Ω.

La **matrice hessienne** de f en un élément x de Ω est

$$H_f(x)=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right)_{1\leqslant i,j\leqslant n}.$$

⇛ **Commentaire**. *Puisque f est de classe C^2 sur Ω, le théorème de* SCHWARZ *permet d'affirmer que pour tout x de Ω, pour tout $(i,j)\in[\![1,n]\!]^2$, $\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x)=\dfrac{\partial^2 f}{\partial x_j\partial x_i}(x)$. Par suite, pour tout $x\in\Omega$, $H_f(x)$ est une matrice symétrique réelle.*

La matrice hessienne va nous fournir une expression condensée du développement limité à l'ordre 2 d'une fonction en un point.

Théorème 44. (Formule de TAYLOR-YOUNG à l'ordre 2).

Soit f une fonction de classe C^2 sur Ω ouvert non vide d'un $\mathbb{R}$-espace vectoriel euclidien $(E, \langle\ ,\ \rangle$ à valeurs dans $\mathbb{R}$. Alors,

$$f(x+h) \underset{h\to 0}{=} f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}\langle H_f(x)h, h\rangle + o\left(\|h\|^2\right)$$

ce qui s'écrit aussi

$$f(x+h) \underset{h\to 0}{=} f(x) + \nabla f(x)^T h + \frac{1}{2}h^T H_f(x)h + o\left(\|h\|^2\right).$$

Plus explicitement, en se plaçant dans le cas particulier $E = \mathbb{R}^n$ muni du produit scalaire canonique (pour alléger les notations)

$$\begin{aligned} f(x+h) &= f(x_1+h_1, \dots, x_n+h_n) \\ &\underset{(h_1,\dots h_n)\to(0,\dots,0)}{=} f(x) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{1\leqslant i,j\leqslant n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + o\left(\|h\|^2\right). \end{aligned}$$

(On rappelle que si $A = (a_{i,j})_{1\leqslant i,j\leqslant n} \in \mathscr{M}_n(\mathbb{R})$ et $X = (x_i)_{1\leqslant i\leqslant n}$, alors $X^T AX = \displaystyle\sum_{1\leqslant i,j\leqslant n} a_{i,j}x_i x_j$).

Le programme officiel prévoit que la démonstration de ce théorème n'est pas exigible. Néanmoins :

DÉMONSTRATION . On se place dans $E = \mathbb{R}^n$ muni du produit scalaire canonique noté $\langle\ ,\ \rangle$. On note $\|\ \|$ la norme associée. Soit $x = (x_1, \dots, x_n) \in \Omega$. Puisque Ω est ouvert, il existe $r > 0$ tel que $B_o(x, r) \subset \Omega$.

Soit $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ tel que $\|h\| < \frac{r}{3}$. Pour $t \in I = \left]-\frac{1}{2}, \frac{3}{2}\right[$, posons $\varphi(t) = f(x+th)$. Pour tout réel t de I, $\|(x+th) - x\| = |t|\|h\| \leqslant \frac{3}{2} \times \frac{r}{3} = \frac{r}{2} < r$. La fonction $t \mapsto x + th$ est de classe C^2 sur I à valeurs dans $B_o(x, r)$ et donc à valeurs dans Ω et f est de classe C^2 sur Ω. Puisque I et Ω sont ouverts, on en déduit que la fonction φ est de classe C^2 sur I et en particulier sur $[0, 1]$.

La fonction φ est de classe C^2 sur $[0, 1]$ à valeurs dans $\mathbb{R}$. D'après la formule de TAYLOR-LAPLACE

$$f(x+h) - f(x) = \varphi(1) - \varphi(0) = (1-0)\varphi'(0) + \int_0^1 (1-t)\varphi''(t)\ dt\ (*).$$

Pour tout réel $t \in [0, 1]$, $\varphi'(t) = \displaystyle\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x+th)$ et en particulier, $\varphi'(0) = \displaystyle\sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x)$. Ensuite, pour tout $t \in I$, $\varphi''(t) = \displaystyle\sum_{1\leqslant i,j\leqslant n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x+th)$ et en particulier, $\varphi''(0) = \displaystyle\sum_{1\leqslant i,j\leqslant n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$. L'égalité $(*)$ s'écrit alors :

$$f(x+h) - f(x) - \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(x) = \int_0^1 (1-t)\varphi''(t)\ dt$$

puis, en tenant compte de l'égalité $\displaystyle\int_0^1 (1-t)\varphi''(0)\ dt = \varphi''(0)\int_0^1 (1-t)\ dt = \frac{1}{2}\varphi''(0)$,

$$f(x+h) - f(x) - \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)h_i - \frac{1}{2}\sum_{1\leqslant i,j\leqslant n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \int_0^1 (1-t)(\varphi''(t) - \varphi''(0))\ dt\ (**).$$

et donc

$$\left| f(x+h) - f(x) - \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)h_i - \frac{1}{2}\sum_{1\leqslant i,j\leqslant n} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right| \leqslant \int_0^1 (1-t)|\varphi''(t) - \varphi''(0)|\ dt \leqslant \int_0^1 |\varphi''(t) - \varphi''(0)|\ dt$$

avec pour tout $t \in [0, 1]$,

$$\begin{aligned}|\varphi''(t)-\varphi''(0)| &= \left|\sum_{1\leqslant i,j\leqslant n} h_ih_j\left(\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right)\right|\\ &\leqslant \sum_{1\leqslant i,j\leqslant n}|h_i|\,|h_j|\left|\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right| \leqslant \|h\|_\infty^2\sum_{1\leqslant i,j\leqslant n}\left|\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\\ &\leqslant \|h\|^2\sum_{1\leqslant i,j\leqslant n}\left|\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|.\end{aligned}$$

Ainsi, pour tout h au voisinage de 0 et non nul,

$$\frac{1}{\|h\|^2}\left|f(x+h)-f(x)-\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x)h_i-\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right| \leqslant \sum_{1\leqslant i,j\leqslant n}\int_0^1\left|\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\,dt.$$

Soit alors $\varepsilon>0$. Pour $(i,j)\in[\![1,n]\!]^2$, la fonction $\dfrac{\partial^2 f}{\partial x_i\partial x_j}$ est continue en x et donc il existe $\alpha_{i,j}>0$ tel que, pour tout $x'\in\Omega$, si $\|x'-x\|\leqslant\alpha$, alors $\left|\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x')-\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\leqslant\dfrac{\varepsilon}{n^2}$.

Soit $\alpha=\operatorname{Min}\left\{\dfrac{r}{2},\operatorname{Min}\left\{\alpha_{i,j},\ (i,j)\in[\![1,n]\!]^2\right\}\right\}>0$. Pour $h\in E$ tel que $\|h\|\leqslant\alpha$, on a tout d'abord $\|(x+h)-x\|=\|h\|<r$ et donc $x+h\in\Omega$.
Soit $(i,j)\in[\![1,n]\!]^2$. Pour tout $t\in[0,1]$, $\|(x+th)-x\|=t\|h\|\leqslant\alpha\leqslant\alpha_{i,j}$ et donc $\left|\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\leqslant\dfrac{\varepsilon}{n^2}$ puis en intégrant, $\displaystyle\int_0^1\left|\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\,dt\leqslant\frac{\varepsilon}{n^2}$.
Mais alors, $\displaystyle\sum_{1\leqslant i,j\leqslant n}\int_0^1\left|\frac{\partial^2 f}{\partial x_i\partial x_j}(x+th)-\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\,dt\leqslant n^2\frac{\varepsilon}{n^2}=\varepsilon$.

On a montré que

$$\forall\varepsilon>0,\ \exists\alpha>0/\ \forall h\in E,\ \left(0<\|h\|\leqslant\alpha\Rightarrow x+h\in\Omega\text{ et }\frac{1}{\|h\|^2}\left|f(x+h)-f(x)-\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x)h_i-\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right|\leqslant\varepsilon\right).$$

Par suite, $\displaystyle\lim_{h\to 0}\frac{1}{\|h\|^2}\left(f(x+h)-f(x)-\sum_{i=1}^n\frac{\partial f}{\partial x_i}(x)h_i-\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(x)\right)=0$ ou encore

$$f(x+h)\underset{h\to 0}{=}f(x)+\sum_{i=1}^n h_i\frac{\partial f}{\partial x_i}(x)+\frac{1}{2}\sum_{1\leqslant i,j\leqslant n}h_ih_j\frac{\partial^2 f}{\partial x_i\partial x_j}(x)+o\left(\|h\|^2\right).$$

❏

On applique maintenant la formule de Taylor-Young à l'ordre 2 à l'étude des extrema d'une fonction numérique de classe C^2. On commence par affiner le théorème 41, page 34, avec une condition nécessaire plus forte. Dans ce qui suit, $E=\mathbb{R}^n$ est muni du produit scalaire canonique.

Théorème 45. Soit f une fonction définie sur un ouvert non vide Ω de $\mathbb{R}^n$, de classe C^2 sur Ω.

Si f admet un minimum local (resp. maximum local) en un point x de Ω, alors $\nabla f(x)=0$ (ou encore $df_x=0$ ou encore x est un point critique de f) et de plus $H_f(x)\in\mathscr{S}_n^+(\mathbb{R})$ (resp. $\mathscr{S}_n^-(\mathbb{R})$).

Démonstration. Supposons que f admette un minimum local $x\in\Omega$. Donc, pour tout $h\in E$ au voisinage de 0, $f(x+h)-f(x)\geqslant 0$. Puisque f est de classe C^1 (car de classe C^2) sur Ω, d'après le théorème 41, on a $df_x=0$ ou encore $\nabla f(x)=0$. D'après le théorème 44,

$$f(x+h)-f(x)\underset{h\to 0}{=}\frac{1}{2}h^T H_f(x)h+o\left(\|h\|^2\right).$$

On a déjà dit que $H_f(x)$ est une matrice symétrique réelle. On sait alors que le polynôme caractéristique de $H_f(x)$ est scindé sur $\mathbb{R}$ puis $H_f(x)\in\mathscr{S}_n^+(\mathbb{R})$ si et seulement si toutes les valeurs propres de $H_f(x)$ sont des réels positifs.

Soient donc $\lambda\in\mathbb{R}$ puis e un vecteur propre associé. Pour t réel au voisinage de 0, le vecteur $h=te$ est tel que $x+h$ est dans Ω puis quand le réel t tend vers le réels 0, le vecteur h tend vers le vecteur nul. Mais alors

$$(f(x+te)-f(x))\underset{t\to 0}{=}\frac{1}{2}(te)^T H_f(x)(te)+o\left(t^2\right),$$

puis $\frac{1}{t^2}(f(x+te)-f(x)) \underset{t\to 0}{=} \frac{1}{2}e^T H_f(x)e + o(1) \underset{t\to 0}{=} \frac{1}{2}\lambda\|e\|^2 + o(1)$ (car $H_f(x)e = \lambda e$). Puisque l'expression $\frac{1}{t^2}(f(x+te)-f(x))$ est positive quand t est au voisinage de 0, par passage à la limite quand t tend vers 0, on obtient $\frac{\lambda}{2}\|e\|^2 \geqslant 0$. En tenant compte de $\frac{1}{2}\|e\|^2 > 0$ car $e \neq 0$, on obtient finalement $\lambda \geqslant 0$.

On a montré que si f admet un minimum local en x, alors $\mathrm{Sp}(H_f(x)) \subset \mathbb{R}^+$ et donc $H_f(x) \in \mathscr{S}_n^+(\mathbb{R})$.

Si f a un maximum local en x, on applique ce qui précède à la fonction $-f$. La matrice $H_{-f}(x) = -H_f(x)$ est symétrique réelle positive et donc la matrice $H_f(x)$ est symétrique réelle négative.

❏

On donne maintenant une condition suffisante pour l'existence d'un extremum local.

Théorème 46. Soit f une fonction définie sur un ouvert non vide Ω de $\mathbb{R}^n$, de classe C^2 sur Ω. Soit $x \in \Omega$.

Si $\nabla f(x) = 0$ (ou encore si x est un point critique de f) et si de plus $H_f(x) \in \mathscr{S}_n^{++}(\mathbb{R})$ (resp. $\mathscr{S}_n^{--}(\mathbb{R})$), alors f admet en x un minimum local (resp. maximum local).

Démonstration. On munit toujours $E = \mathbb{R}^n$ du produit scalaire canonique et on note $\|\ \|$ la norme euclidienne associée.

Supposons que $\nabla f(x) = 0$ et que $H_f(x) \in \mathscr{S}_n^{++}(\mathbb{R})$. On a toujours (en tenant compte de $\nabla f(x) = 0$)

$$f(x+h)-f(x) \underset{h\to 0}{=} \frac{1}{2}h^T H_f(x)h + o\left(\|h\|^2\right)\ (*).$$

L'application $\begin{array}{ccc}(\mathbb{R}^n)^2 & \to & \mathbb{R}\\ (u,v) & \mapsto & \frac{1}{2}u^T H_f(x)v\end{array}$ (en identifiant un n-uplet de réels et un vecteur colonne) est une forme bilnéaire, symétrique (car $H_f(x)$ est symétrique), définie, positive (car $H_f(x) \in \mathscr{S}_n^{++}(\mathbb{R})$). Donc, l'application $\begin{array}{ccc}(\mathbb{R}^n)^2 & \to & \mathbb{R}\\ (u,v) & \mapsto & \frac{1}{2}u^T H_f(x)v\end{array}$ est un produit scalaire sur $\mathbb{R}^n$. On note N la norme associée. $(*)$ s'écrit alors

$$\frac{1}{\|h\|^2}(f(x+h)-f(x)) \underset{h\to 0,\ h\neq 0}{=} \left(\frac{N(h)}{\|h\|}\right)^2 + \varepsilon(h),$$

où ε est une fonction définie sur un voisinage de 0 telle que $\lim\limits_{h\to 0}\varepsilon(h) = 0$. Puisque $E = \mathbb{R}^n$ est de dimension finie, la norme N est équivalente à la norme $\|\ \|$ et donc, il existe deux réels α et β strictement positifs tels que $\alpha\|\ \| \leqslant N \leqslant \beta\|\ \|$. Pour tout $h \neq 0$, $\left(\frac{N(h)}{\|h\|}\right)^2 \geqslant \alpha^2 (> 0)$. D'autre part, puisque $\lim\limits_{h\to 0}\varepsilon(h) = 0$, il existe un voisinage V de 0 tel que, pour $h \in V$, $\varepsilon(h) \geqslant -\frac{\alpha^2}{2}$. Pour h au voisinage de 0, on a donc

$$\frac{1}{\|h\|^2}(f(x+h)-f(x)) \geqslant \alpha^2 - \frac{\alpha^2}{2} = \frac{\alpha^2}{2} \geqslant 0$$

et donc $f(x+h)-f(x) \geqslant 0$. Ceci montre que f admet un minimum local en x.

❏

On va maintenant détailler le cas particulier des fonctions de deux variables (cas $n = 2$). Dans ce cas, la matrice hessienne de f en $a = (x,y) \in \mathbb{R}^2$ s'écrit

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(a) & \frac{\partial^2 f}{\partial x\partial y}(a)\\ \frac{\partial^2 f}{\partial x\partial y}(a) & \frac{\partial^2 f}{\partial y^2}(a)\end{pmatrix}.$$

Pour alléger les notations, on définit les notations (dites notations de Monge) :

$$\boxed{\begin{array}{c} p = \frac{\partial f}{\partial x} \quad q = \frac{\partial f}{\partial y}\\ r = \frac{\partial^2 f}{\partial x^2},\quad s = \frac{\partial^2 f}{\partial x\partial y} \quad \text{et} \quad t = \frac{\partial^2 f}{\partial y^2}.\end{array}}$$

(p, q, r, s et t sont des fonctions de deux variables). La matrice hessienne de f s'écrit alors de manière plus condensée $H_f = \begin{pmatrix} r & s\\ s & t\end{pmatrix}$. Les deux valeurs propres réelles λ et μ de $H_f(a)$ vérifient

$$\lambda\mu = \det(H_f(a)) = rt - s^2 \text{ et } \lambda + \mu = \mathrm{Tr}(H_f(a)) = r + t.$$

Ces deux valeurs propres sont non nulles et de même signe si et seulement si $\lambda\mu = rt - s^2 > 0$. Plus précisément, ces deux valeurs propres sont strictement positives (resp. strictement négatives) si et seulement si $rt - s^2 > 0$ et $r + t > 0$ (resp. $r + t < 0$). On obtient alors le théorème :

Théorème 47. Soit f une fonction définie sur un ouvert non vide Ω de $\mathbb{R}^2$, de classe C^2 sur Ω. Soit $a = (x, y) \in \Omega$.

Si f admet un extremum local en a, alors $p(a) = q(a) = 0$.
Réciproquement, si $p = q = 0$, $rt - s^2 > 0$ et $r + t > 0$, f admet un minimum local en a et si $p = q = 0$, $rt - s^2 > 0$ et $r + t < 0$, f admet un maximum local en a.

Revenons alors sur l'exercice n° 7 de la page 35. Nous avons étudié les extrema locaux de la fonction $f : (x, y) \mapsto -2(x - y)^2 + x^4 + y^4$. Pour tout $(x, y) \in \mathbb{R}^2$,

$$p(x, y) = \frac{\partial f}{\partial x}(x, y) = -4(x - y) + 4x^3 \text{ et } q(x, y) = \frac{\partial f}{\partial y}(x, y) = 4(x - y) + 4y^3.$$

puis

$$r(x, y) = \frac{\partial^2 f}{\partial x^2}(x, y) = -4 + 12x^2,\ s(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y) = 4 \text{ et } t(x, y) = \frac{\partial^2 f}{\partial y^2}(x, y) = -4 + 12y^2.$$

Mais alors $\left(rt - s^2\right)(x, y) = 16\left[\left(3x^2 - 1\right)\left(3y^2 - 1\right) - 1\right]$ et $(r + t)(x, y) = 4\left(3x^2 + 3y^2 - 2\right)$.

Nous avions obtenu trois points critiques, les points $A = \left(\sqrt{2}, -\sqrt{2}\right)$, $A' = \left(-\sqrt{2}, \sqrt{2}\right) = -A$ et $O = (0, 0)$.

$\left(rt - s^2\right)\left(\sqrt{2}, -\sqrt{2}\right) = 16 \times 24 > 0$. Donc, f a un extremum local en A. De plus, $(r + t)\left(\sqrt{2}, -\sqrt{2}\right) = 4 \times 10 > 0$ et donc f a un minimum local en A. Le théorème 47 nous a donc apporté un outil supplémentaire efficace.

Par contre, $\left(rt - s^2\right)(0, 0) = 0$ et on ne peut pas améliorer la solution déjà fournie page 35.

6 Exemples de résolutions d'équations aux dérivées partielles

Une équation aux dérivées partielles est, comme son nom l'indique, une équation dont l'inconnue est une fonction de plusieurs variables dont les dérivées partielles premières, secondes ... vérifient une certaine égalité.

Par exemple, en physique, l'**équation de la diffusion thermique** s'écrit $\dfrac{\partial T}{\partial t} = \dfrac{k}{\rho c}\dfrac{\partial^2 T}{\partial x^2}$ ou l'**équation généralisée de la chaleur** est $\dfrac{\partial T}{\partial t} - \dfrac{k}{\rho c}\Delta T = \dfrac{S}{\rho C_P}$.

L'**équation des cordes vibrantes** est quant à elle $\dfrac{1}{c^2}\dfrac{\partial^2 u}{\partial t^2} = \dfrac{\partial^2 u}{\partial x^2}$.

Nous allons analyser quelques équations simples et nous contenter d'une sensibilisation à quelques problèmes techniques.

• « L »'équation de référence est $\dfrac{\partial f}{\partial y} = 0$ où f est une fonction de classe C^1 sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$. Pour x_0 réel donné, la fonction $g : y \mapsto f(x_0, y)$ a une dérivée nulle sur $\mathbb{R}$ ($g'(y) = \dfrac{\partial f}{\partial y}(x_0, y)$). Pour x_0 réel donné, la fonction $y \mapsto f(x_0, y)$ est donc constante quand y varie. Maintenant, la valeur de cette constante est fonction de x et donc il existe une fonction φ de $\mathbb{R}$ dans $\mathbb{R}$, nécessairement de classe C^1 sur $\mathbb{R}$, telle que

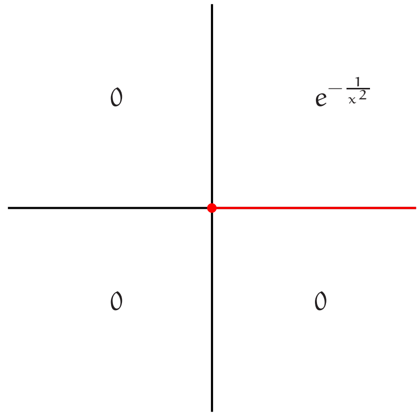
$$\forall (x, y) \in \mathbb{R}^2,\ f(x, y) = \varphi(x).$$

Réciproquement, une telle fonction convient.

• Les choses se compliquent un peu si on résout sur un ouvert Ω de $\mathbb{R}^2$ qui n'est pas $\mathbb{R}^2$. Reprenons l'équation $\dfrac{\partial f}{\partial y} = 0$ où f est une fonction de classe C^1 sur $\mathbb{R}^2 \setminus \Delta$ où $\Delta = \{(x, 0),\ x \geqslant 0\}$.

Soit f_0 la fonction définie sur $\mathbb{R}^2 \setminus \Delta$ par

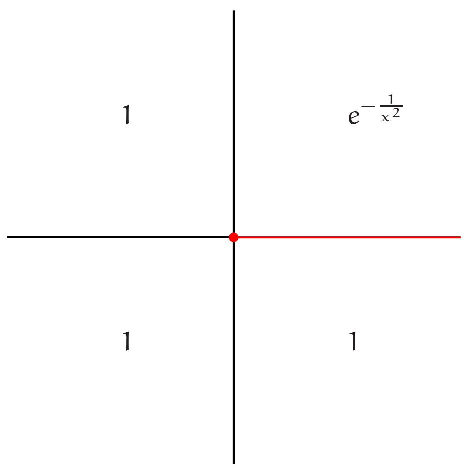
$$\forall (x, y) \in \mathbb{R}^2 \setminus \Delta,\ f_0(x, y) = \begin{cases} e^{-\frac{1}{x^2}} \text{ si } x > 0 \text{ et } y > 0 \\ 0 \text{ sinon} \end{cases}.$$



On montre facilement que f_0 est solution de l'équation $\dfrac{\partial f}{\partial y} = 0$ sur $\mathbb{R}^2 \setminus \Delta$ mais on doit noter que par exemple la fonction $y \mapsto f_0(1, y)$ n'est constante sur $]-\infty, 0[\cup]0, +\infty[$. Le problème vient du fait que $]-\infty, 0[\cup]0, +\infty[$ n'est pas un intervalle.

Plus généralement, si f est une solution de l'équation et si x_0 est un réel positif donné, la fonction $y \mapsto f(x_0, y)$ est définie et de classe C^1 sur $]-\infty, 0[\cup]0, +\infty[$. Il existe donc nécessairement deux fonctions g_1 et g_2 de classe C^1 sur $[0, +\infty[$ telles que $\forall (x, y) \in [0, +\infty[\times]0, +\infty[$, $f(x, y) = g_1(x)$ et $\forall (x, y) \in [0, +\infty[\times]-\infty, 0[$, $f(x, y) = g_2(x)$. Mais il s'agit ensuite qu'une fois la résolution achevée, les fonctions se « recollent » correctement. C'est le cas de la fonction f_0 citée plus haut et ce n'est pas le cas de la fonction f_1 définie sur $\mathbb{R}^2 \setminus \Delta$ par

$$\forall (x, y) \in \mathbb{R}^2 \setminus \Delta,\ f_1(x, y) = \left\{\begin{array}{l} e^{-\frac{1}{x^2}} \text{ si } x > 0 \text{ et } y > 0 \\ 1 \text{ sinon} \end{array}\right. .$$



La fonction f_1 ci-dessus n'est pas de classe C^1 sur $\mathbb{R}^2 \setminus \Delta$ en raison du mauvais recollement sur la demi-droite $]Oy) = \{(0, y),\ y > 0\}$. Nous ne détaillerons pas davantage la résolution.

• On s'intéresse maintenant à l'équation $2\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y}$ où f est une fonction de classe C^1 sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$. On va résoudre cette équation grâce aux changement de variables : $u = x + y$ et $v = x + 2y$.

Comme toujours, **changer de variables, c'est en fait changer de fonction inconnue**. Tout d'abord

$$\left\{\begin{array}{l} u = x + y \\ v = x + 2y \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} x = 2u - v \\ y = -u + v \end{array}\right. .$$

Pour $(x, y) \in \mathbb{R}^2$, on pose donc

$$f(x, y) = f(2u - v, -u + v) = g(u, v),$$

ou encore, si pour $(u, v) \in \mathbb{R}^2$, on pose $\varphi(u, v) = (2u - v, -u + v)$, φ est une bijection de $\mathbb{R}^2$ sur lui-même (φ est un automorphisme de $\mathbb{R}^2$) et $g = f \circ \varphi$ ou aussi $f = g \circ \varphi^{-1}$. De plus, φ et φ^{-1} sont de classe C^1 sur $\mathbb{R}^2$ et $\mathbb{R}^2$ respectivement et donc, f est de classe C^1 sur $\mathbb{R}^2$ si et seulement si g est de classe C^1 sur $\mathbb{R}^2$.

La formule de dérivation partielle d'une fonction composée (règle de la chaîne) fournit

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial g}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial g}{\partial v} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}$$

et

$$\frac{\partial f}{\partial y}=\frac{\partial u}{\partial y}\frac{\partial g}{\partial u}+\frac{\partial v}{\partial y}\frac{\partial g}{\partial v}=\frac{\partial g}{\partial u}+2\frac{\partial g}{\partial v},$$

puis

$$2\frac{\partial f}{\partial x}-\frac{\partial f}{\partial y}=\frac{\partial g}{\partial u}.$$

Par suite,

$$\begin{aligned}2\frac{\partial f}{\partial x}=\frac{\partial f}{\partial y}&\Leftrightarrow\frac{\partial g}{\partial u}=0\\&\Leftrightarrow\exists\varphi\in C^1(\mathbb{R},\mathbb{R})\,/\,\forall(u,v)\in\mathbb{R}^2,\ g(u,v)=\varphi(v)\\&\Leftrightarrow\exists\varphi\in C^1(\mathbb{R},\mathbb{R})\,/\,\forall(x,y)\in\mathbb{R}^2,\ f(x,y)=\varphi(x+2y).\end{aligned}$$

Les solutions sur $\mathbb{R}^2$ de l'équation proposée sont les fonctions de la forme $(x,y)\mapsto\varphi(x+y)$ où φ est une fonction de classe C^1 sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Par exemple, la fonction $f\ :\ \cos(x+2y)+e^{3x+6y+2}+4$ est une solution.

• En dernier exemple, on veut résoudre l'équation (E)

$$x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}=\sqrt{x^2+y^2}$$

sur $P=]0,+\infty[\times\mathbb{R}=\left\{(x,y)\in\mathbb{R}^2/\ x>0\right\}$ en passant en polaires.

Soit $\varphi\ :\ \begin{array}{ccc}]0,+\infty[\times\left]-\frac{\pi}{2},\frac{\pi}{2}\right[&\to&P\\(r,\theta)&\mapsto&(r\cos\theta,r\sin\theta)\end{array}$. φ est une bijection de $]0,+\infty[\times\left]-\dfrac{\pi}{2},\dfrac{\pi}{2}\right[$ sur P, de classe C^1 sur P et sa réciproque est $\varphi^{-1}\ :\ \begin{array}{ccc}P&\to&]0,+\infty[\times\left]-\frac{\pi}{2},\frac{\pi}{2}\right[\\(x,y)&\mapsto&\left(\sqrt{x^2+y^2},\mathrm{Arctan}\left(\frac{y}{x}\right)\right)\end{array}$ qui est de classe C^1 sur P.

Le changement de variables s'écrit explicitement $\left\{\begin{array}{l}x=r\cos\theta\\y=r\sin\theta\end{array}\right.$. On pose $g=f\circ\varphi$ ou encore $\forall(x,y)\in\mathbb{R}^2$,

$$f(x,y)=f(r\cos\theta,r\sin\theta)=g(r,\theta).$$

On rappelle que (voir page 21 et 22) $\dfrac{\partial r}{\partial x}=\cos\theta$, $\dfrac{\partial r}{\partial y}=\sin\theta$, $\dfrac{\partial\theta}{\partial x}=-\dfrac{\sin\theta}{r}$ et $\dfrac{\partial\theta}{\partial y}=\dfrac{\cos\theta}{r}$. On a donc

$$\frac{\partial f}{\partial x}=\frac{\partial r}{\partial x}\frac{\partial g}{\partial r}+\frac{\partial\theta}{\partial x}\frac{\partial g}{\partial\theta}=\cos\theta\frac{\partial g}{\partial r}-\frac{\sin\theta}{r}\frac{\partial g}{\partial\theta}$$

et

$$\frac{\partial f}{\partial y}=\frac{\partial r}{\partial y}\frac{\partial g}{\partial r}+\frac{\partial\theta}{\partial y}\frac{\partial g}{\partial\theta}=\sin\theta\frac{\partial g}{\partial r}+\frac{\cos\theta}{r}\frac{\partial g}{\partial\theta},$$

et donc

$$x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}=r\cos\theta\left(\cos\theta\frac{\partial g}{\partial r}-\frac{\sin\theta}{r}\frac{\partial g}{\partial\theta}\right)+r\sin\theta\left(\sin\theta\frac{\partial g}{\partial r}+\frac{\cos\theta}{r}\frac{\partial g}{\partial\theta}\right)=r\frac{\partial g}{\partial r}.$$

On en déduit que

$$\begin{aligned}f\text{ solution de }(E)\text{ sur }P&\Leftrightarrow\forall(r,\theta)\in]0,+\infty[\times\left]-\frac{\pi}{2},\frac{\pi}{2}\right[,\ r\frac{\partial g}{\partial r}(r,\theta)=r\\&\Leftrightarrow\forall(r,\theta)\in]0,+\infty[\times\left]-\frac{\pi}{2},\frac{\pi}{2}\right[,\ \frac{\partial g}{\partial r}(r,\theta)=1\\&\Leftrightarrow\exists\varphi\in C^1\left(\left]-\frac{\pi}{2},\frac{\pi}{2}\right[,\mathbb{R}\right)/\ \forall(r,\theta)\in]0,+\infty[\times\left]-\frac{\pi}{2},\frac{\pi}{2}\right[,\ g(r,\theta)=r+\varphi(\theta)\\&\Leftrightarrow\exists\varphi\in C^1\left(\left]-\frac{\pi}{2},\frac{\pi}{2}\right[,\mathbb{R}\right)/\ \forall(x,y)\in P,\ f(x,y)=\sqrt{x^2+y^2}+\varphi\left(\mathrm{Arctan}\left(\frac{y}{x}\right)\right)\\&\Leftrightarrow\exists\psi\in C^1(\mathbb{R},\mathbb{R})\,/\,\forall(x,y)\in P,\ f(x,y)=\sqrt{x^2+y^2}+\psi\left(\frac{y}{x}\right).\end{aligned}$$

Les solutions de (E) sur P sont les fonctions de la forme $(x,y)\mapsto\sqrt{x^2+y^2}+\psi\left(\dfrac{y}{x}\right)$ où ψ est une fonction de classe C^1 sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$.