Chapitre 7

Fonctions réelles de plusieurs variables

7.1 Limite, continuité

On s'intéresse dans ce chapitre aux applications de $\mathbb{R}^n$ (ou d'une partie de $\mathbb{R}^n$) dans $\mathbb{R}^p$, avec $n, p \in \mathbb{N}^\star$. On va commencer par munir $\mathbb{R}^n$ (et $\mathbb{R}^p$) d'une "norme" particulière, appelée "norme euclidienne", qui jouera le rôle de la valeur absolue dans $\mathbb{R}$.

Définition 7.1 (Norme euclidienne) *Soit $m \in \mathbb{N}^\star$. Pour $x \in \mathbb{R}^m$, on note $x_1, \ldots, x_m$ les composantes de x et on pose :*

$$|x| = \left(\sum_{i=1}^{m} x_i^2\right)^{\frac{1}{2}}.$$

Si $x, y \in \mathbb{R}^m$, en notant $x_1, \ldots, x_m$ les composantes de x et $y_1, \ldots, y_m$ les composantes de y, on définit aussi le produit scalaire de x avec y, que nous noterons $x \cdot y$ par la formule :

$$x \cdot y = \sum_{i=1}^{m} x_i y_i.$$

Pour $m \in \mathbb{N}^\star$, l'application $x \mapsto |x|$ est une application de $\mathbb{R}^m$ dans $\mathbb{R}_+$ vérifiant les trois propriétés suivantes (qui font de cette application une norme) :

$$\text{(Non dégénérescence) pour tout } x \in \mathbb{R}^m,\ |x| = 0 \Leftrightarrow x = 0, \tag{7.1}$$

$$\text{(homogénéité positive) pour tout } x \in \mathbb{R}^m \text{ et tout } \lambda \in \mathbb{R},\ |\lambda x| = |\lambda||x|, \tag{7.2}$$

$$\text{(inégalité trinagulaire) pour tout } x, y \in \mathbb{R}^m,\ |x + y| \le |x| + |y|. \tag{7.3}$$

Remarque 7.1 *Soit $m \in \mathbb{N}^\star$. Si $x \mapsto \|x\|$ est une application de $\mathbb{R}^m$ dans $\mathbb{R}_+$ vérifiant les trois propriétés (7.1)-(7.3) (avec $\|x\|$, $\|\lambda x\|$, $\|x+y\|$ et $\|y\|$ au lieu de $|x|$, $|\lambda x|$, $|x+y|$ et $|y|$), on dit que cette application est une norme. Si $\|\cdot\|$ est une norme sur $\mathbb{R}^m$, on peut montrer qu'il existe $C_1, C_2 \in \mathbb{R}$ t.q. :*

$$x \in \mathbb{R}^m \Rightarrow C_1|x| \le \|x\| \le C_2|x|.$$

On dit que les normes $\|\cdot\|$ et $|\cdot|$ sont "équivalentes". Les propiétés d'une fonction f (de $\mathbb{R}^n$ dans $\mathbb{R}^p$) introduites dans ce chapitre (continuité, dérivabilité, caractère $C^1\ldots$) seraient alors les mêmes en remplaçant la norme euclidienne (sur $\mathbb{R}^n$ et $\mathbb{R}^p$) par d'autres normes. (Ce résultat d'équivalence des normes sur $\mathbb{R}^m$ est dû au fait que $\mathbb{R}^m$ est un espace vectoriel sur $\mathbb{R}$, de dimension finie. Sur un espace vectoriel de dimension infinie, les propriétés de f telles que continuité et dérivabilité dépendraient de la norme choisie.)

Une propriété intéressante de la norme euclidienne et du produit scalaire définis dans la définition 7.1 est l'inégalité de Cauchy-Schwarz que nous donnons maintenant (et qui est d'ailleurs utile pour démontrer la propriété (7.3))

Proposition 7.1 (Inégalité de Cauchy-Schwarz) *Soit $m \in \mathbb{N}^\star$ et $x, y \in \mathbb{R}^m$. On a alors*

$$|x \cdot y| \leq |x||y|. \tag{7.4}$$

On définit maintenant les notions de limite et de continuité pour une fonction de $\mathbb{R}^n$ dans $\mathbb{R}^p$ de manière similaire au cas $n = p = 1$.

Définition 7.2 (Boule ouverte) *Soit $m \in \mathbb{N}^\star$, $a \in \mathbb{R}^m$ et $\delta > 0$. La boule ouverte de centre a et de rayon δ est l'ensemble $B(a, \delta) = \{x \in \mathbb{R}^m,\ |x - a| < \delta\}$.*

Définition 7.3 (Limite un point de $\mathbb{R}$) *Soit $n, p \in \mathbb{N}^\star$, f une application de $D \subset \mathbb{R}^n$ dans $\mathbb{R}^p$ et $l \in \mathbb{R}^p$. Soit $a \in \mathbb{R}^n$. On suppose qu'il existe $\delta > 0$ t.q. $D \supset B(a, \delta) \setminus \{a\}$. On dit que $l = \lim_{x\to a} f(x)$ si pour tout $\varepsilon > 0$, il existe $\alpha > 0$ t.q. :*

$$x \in D,\ x \neq a,\ |x - a| \leq \alpha \ \Rightarrow\ |f(x) - l| \leq \varepsilon.$$

Comme dans le cas $n = p = 1$, la limite de f en a (sous les hypothèses de la définition 7.3), si elle existe, est unique. Il y a aussi une caractérisation séquentielle de la limite.

Proposition 7.2 (Caractérisation séquentielle de la limite) *Soit $n, p \in \mathbb{N}^\star$, f une application de $D \subset \mathbb{R}^n$ dans $\mathbb{R}^p$ et $l \in \mathbb{R}^p$. Soit $a \in \mathbb{R}^n$. On suppose qu'il existe $\delta > 0$ t.q. $D \supset B(a, \delta) \setminus \{a\}$. Alors, l est la limite en a de f si et seulement f transforme toute suite convergente vers a en suite convergente vers l, c'est-à-dire :*

$$(x_n)_{n\in\mathbb{N}} \subset D \setminus \{a\},\ \lim_{n\to+\infty} x_n = a \ \Rightarrow\ \lim_{n\to+\infty} f(x_n) = l.$$

Exemple 7.1 (Application linéaire) *Soit $n, p \in \mathbb{N}^\star$ et f une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Alors, f est continue. En fait, f est même lipschitzienne, c'est-à-dire qu'il existe $k \in \mathbb{R}$ t.q. $|f(x)| \leq k|x|$ pour tout $x \in \mathbb{R}^n$.*

7.2 Différentielle, dérivées partielles

Nous définissons d'abord les dérivées partielles d'une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$, puis la différentielle.

Définition 7.4 (Ouvert de $\mathbb{R}^n$) *Soit $n \in \mathbb{N}^\star$ et $\Omega \subset \mathbb{R}^n$. On dit que Ω est ouvert si pour tout $a \in \Omega$, il existe $\alpha > 0$ t.q. $B(a, \alpha) \subset \Omega$.*

Remarque 7.2 Soit Ω un ouvert de $\mathbb{R}^n$ $(n \geq 1)$ et a un point de Ω. D'après la définition 7.4, Il existe $\alpha > 0$ t.q. $B(a, \alpha) \subset \Omega$. Soit $j \in \{1, \ldots, n\}$. Si $x \in \mathbb{R}^n$ est t.q. $x_k = a_k$ pour $k \neq j$ et $|x_j - a_j| < \alpha$, on a donc $|x - a| < \alpha$ et donc $x \in \Omega$. Ceci va nous permettre de d'introduire la notion de "dérivées partielles" d'une application f de Ω dans $\mathbb{R}^p$ $(p \geq 1)$, consistant à fixer $(n-1)$ composantes de x et à considérer les applications qui à la composante restante de x associe l'une des composantes de f (on est ainsi ramené à étudier une application définie sur une partie de $\mathbb{R}$ dans $\mathbb{R}$).

Définition 7.5 *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Pour $x \in \mathbb{R}^n$, on note $x_1, \ldots, x_n$ les composantes de x et $f_1(x), \ldots, f_p(x)$ les composantes de $f(x)$ (de sorte que les f_i sont des applications de Ω dans $\mathbb{R}$).*

Soit $i \in \{1, \ldots, p\}$, $j \in \{1, \ldots, n\}$ et $a \in \Omega$. On note g l'application que à x_j associe $f_i(x)$, avec $x_k = a_k$ si $k \in \{1, \ldots, n\}$, $k \neq j$ (de sorte que g est une application définie sur un voisinage de a_j, dans $\mathbb{R}$, à valeurs dans $\mathbb{R}$, voir la remarque 7.2). Si g est dérivable au point a_j, on dit que f admet une dérivée partielle en a selon x_j et on pose :

$$\frac{\partial f_i}{\partial x_j}(a) = g'(a_j).$$

Nous donnons maintenant la définition de différentielle et nous faisons ensuite le lien entre différentielle et dérivées partielles.

Définition 7.6 *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Soit $a \in \Omega$. On dit que f est différentiable en a si il existe une application linéaire de R^n dans $\mathbb{R}^p$, notée $df(a)$, t.q.*

$$f(a+h) = f(a) + df(a)(h) + |h| g(h), \text{ pour } h \neq 0 \text{ t.q. } a + h \in \Omega,$$

avec $\lim_{h \to 0} g(h) = 0$ (et donc g continue en 0 si on pose $g(0) = 0$). L'application $df(a)$ est alors unique (voir la proposition 7.3).

Remarque 7.3 On s'intéresse ici au cas particulier $p = 1$ et (pour simplifier) $n = 2$. Soit T une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$. On définit alors $a_1, a_2 \in \mathbb{R}$ par

$$\begin{aligned} &a_1 = T(h) \text{ si } h \text{ est le vecteur de composantes 1 et 0,} \\ &a_2 = T(h) \text{ si } h \text{ est le vecteur de composantes 0 et 1.} \end{aligned}$$

Avec ce choix de a_1 et a_2, on a, par linéarité de T,

$$T(h) = a_1 h_1 + a_2 h_2 \text{ si } h \text{ est le vecteur de composantes } h_1 \text{ et } h_2.$$

On note maintenant dx_1 l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $T(h) = h_1$ si h a pour composantes h_1 et h_2. De même, on note dx_2 l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $T(h) = h_2$ si h a pour composantes h_1 et h_2. On a alors $T(h) = a_1 dx_1(h) + a_2 dx_2(h)$ pour tout $h \in \mathbb{R}^2$, c'est-à-dire :

$$T = a_1 dx_1 + a_2 dx_2.$$

Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$, f une application de Ω dans $\mathbb{R}^p$ et $a \in \Omega$. Le fait que f admette des dérivées partielles en a n'implique pas que f soit différentiable en a (ni même que f soit continue en a, voir l'exercice 7.2). Par contre si f est différentiable en a, f admet alors des dérivées partielles en a, ceci est énoncé dans la proposition 7.3.

Proposition 7.3 (différentielle et dérivées partielles) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Soit $a \in \Omega$. On suppose que f est différentiable en a. Alors, f admet des dérivées partielles en a et, pour $i \in \{1, \ldots, p\}$, la i-ème composante de $df(a)(h)$, notée $(df(a)(h))_i$ vérifie :*

$$(df(a)(h))_i = \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(a) h_j, \text{ pour tout } h \in \mathbb{R}^n.$$

(En particulier, ceci montre l'unicité de la différentielle.)

Une réciproque partielle de la proposition 7.3 sera donnée dans la proposition 7.6.

Remarque 7.4 Sous les hypothèses de la proposition 7.3, dans le cas particulier $n = 2$ et $p = 1$ (dans ce cas, f n'a qu'une composante, notée f et non f_1), la remarque 7.3 donne

$$df(a) = \frac{\partial f}{\partial x_1}(a) dx_1 + \frac{\partial f}{\partial x_2}(a) dx_2.$$

Si f est différentiable en tout point de Ω, on définit donc une application $a \mapsto df(a)$, notée df, de Ω dans l'ensemble des applications linéaires de $\mathbb{R}^2$ dans $\mathbb{R}$ et on a

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2.$$

D'autres notations sont utilisées pour exprimer la différentielle au point a d'une application f de $\mathbb{R}^n$ dans $\mathbb{R}^p$:

- La matrice jacobienne. C'est une matrice de p lignes et n colonnes. Elle sera notée $M_f(a)$.
- Le gradient de f, si $p = 1$. C'est un vecteur de $\mathbb{R}^n$. Il sera noté $\nabla f(a)$.
- La dérivée de f, si $n = 1$. C'est un vecteur de $\mathbb{R}^p$. Il sera noté $f'(a)$.

Nous donnons maintenant les définitions de $M_f(a)$, $\nabla f(a)$ et $f'(a)$.

Définition 7.7 (Matrice jacobienne, gradient, dérivée) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Soit $a \in \Omega$. On suppose que f est différentiable en a.*

1. *(Matrice jacobienne) La matrice jacobienne de f en a, que nous noterons $M_f(a)$, est la matrice de p lignes et n colonnes dont le coefficient de la i-ème ligne et j-ème colonne est $\frac{\partial f_i}{\partial x_j}(a)$.*
2. *(Gradient) On suppose ici que $p = 1$. Le gradient de f en a, noté $\nabla f(a)$, est le vecteur de $\mathbb{R}^n$ dont la j-ème composante est $\frac{\partial f}{\partial x_j}(a)$.*
3. *(Dérivée) On suppose ici que $n = 1$. La dérivée de f en a, notée $f'(a)$, est le vecteur de $\mathbb{R}^p$ dont la i-ème composante est $f_i'(a)$. (Il n'y a pas ici de dérivée partielle car il n'y a qu'une seule variable.)*

Nous allons maintenant faire le lien entre différentielle, matrice jacobienne, gradient et dérivée. En pratique, on confond le vecteur x de $\mathbb{R}^n$, dont les composantes sont $x_1, \ldots, x_n$, avec la matrice à n lignes et 1 colonne formée avec $x_1, \ldots, x_n$, c'est-à-dire la matrice $\begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix}^t$ (appelée aussi "vecteur colonne"). Cette confusion nous permet d'écrire la remarque 7.5.

Remarque 7.5 (différentielle, matrice jacobienne, gradient et dérivée) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Soit $a \in \Omega$. On suppose que f est différentiable en a.*

1. *(Différentielle et matrice jacobienne) Pour tout $h \in \mathbb{R}^n$, on a $df(a)(h) = M_f(a)h$ (où $M_f(a)h$ désigne le produit de la matrice jacobienne avec la matrice à n lignes et 1 colonne que l'on confond avec le vecteur h).*

2. *(Différentielle, matrice jacobienne et gradient) Si $p = 1$, la matrice jacobienne est une matrice à 1 ligne et n colonnes. Le gradient de f en a est un vecteur de $\mathbb{R}^n$ que l'on confond donc avec une matrice à n lignes et 1 colonne, cette matrice est la transposée de la matrice jacobienne, c'est-à-dire $\nabla f(a) = M_f(a)^t$. On a donc :*

$$df(a)(h) = M_f(a)h = \nabla f(a) \cdot h, \text{ pour tout } h \in \mathbb{R}^n.$$

3. *(Différentielle, matrice jacobienne et dérivée) Si $n = 1$, la matrice jacobienne est une matrice à n lignes et 1 colonne. La dérivée f en a est un vecteur de $\mathbb{R}^p$ que l'on confond donc avec une matrice à p lignes et 1 colonne. Cette matrice est la matrice jacobienne, c'est-à-dire $f'(a) = M_f(a)$. On a donc :*

$$df(a)(h) = M_f(a)h = hf'(a), \text{ pour tout } h \in \mathbb{R}.$$

Remarque 7.6 (base canonique et représentation d'une application linéaire)
Soit $n \geq 1$. Pour $i \in \{1, \ldots, n\}$, on note e_i l'élément de $\mathbb{R}^n$ dont toutes les composantes sont nulles sauf la i-ème, cette i-ème composante valant 1. On rappelle que la famille $\{e_1, \ldots, e_n\}$ est une base de $\mathbb{R}^n$ (et s'appelle "base canonique de $\mathbb{R}^n$"). Si $x \in \mathbb{R}^n$ a pour composantes $x_1, \ldots, x_n$, la décomposition de x dans cette base est (avec les opérations usuelles dans $\mathbb{R}^n$ d'addition et de multiplication pas un nombre réel) :

$$x = \sum_{i=1}^{n} x_i e_i.$$

Soit maintenant $n, p \in \mathbb{N}^\star$ et T une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Pour tout $x \in \mathbb{R}^n$ on a donc, en notant $x_1, \ldots, x_n$ les composantes de x et $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{R}^n$:

$$T(x) = \sum_{i=1}^{n} x_i T(e_i).$$

Pour tout $i \in \{1, \ldots, n\}$, le vecteur $T(e_i)$ est un vecteur de $\mathbb{R}^p$ (on le confond donc avec une matrice à p lignes et 1 colonne). On note A la matrice (à p lignes et n colonnes) dont, pour tout $i \in \{1, \ldots, n\}$, la i-ème colonne est formée de $T(e_i)$. On a alors $T(x) = Ax$ (le vecteur x de $\mathbb{R}^n$ est confondu avec une matrice à n lignes et 1 colonne). La matrice A est la matrice qui représente l'application linéaire T dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$.

Proposition 7.4 (Différentielle d'une fonction composée) *Soit $n, p, q \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$, O un ouvert de $\mathbb{R}^p$, f une application de Ω dans $\mathbb{R}^p$ et g une application de O dans $\mathbb{R}^q$. Soit $a \in \Omega$. On suppose que f est différentiable en a, $f(a) \in O$ et g différentiable en $f(a)$. Alors $g \circ f$ (qui est bien définie au voisinage de a) est différentiable en a et $d(g \circ f)(a) = dg(f(a)) \circ df(a)$, ou encore, avec les matrices jacobiennes, $M_{g \circ f}(a) = M_g(f(a))M_f(a)$.*

Exemple 7.2 On considére ici une application de $\mathbb{R}$ de $\mathbb{R}$ mais définie "en passant par $\mathbb{R}^2$". c'est-à-dire que γ est une application de $\mathbb{R}$ dans $\mathbb{R}^2$ et φ est une application de $\mathbb{R}^2$ dans $\mathbb{R}$. On suppose que γ et φ sont de classe C^1 et on considère la fonction $f = \varphi \circ \gamma$, de sorte que f est une application de $\mathbb{R}$ dans $\mathbb{R}$ de classe C^1. Pour tout $t \in \mathbb{R}$, la proposition 7.4 donne

$$f'(t) = d\varphi(\gamma(t))(\gamma'(t)) = M_\varphi(\gamma(t))\gamma'(t).$$

Noter que $d\varphi(\gamma(t))$ est une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ (représentée par la matrice $M_\varphi(\gamma(t))$ de 1 ligne et 2 colonnes) et que $\gamma'(t)$ est un vecteur de $\mathbb{R}^2$. En notant γ_1 et γ_2 les deux composantes de γ, on a aussi

$$f'(t) = \frac{\partial \varphi}{\partial x_1}(\gamma(t))\gamma_1'(t) + \frac{\partial \varphi}{\partial x_1}(\gamma(t))\gamma_2'(t).$$

On note $\gamma(0) = a$ et $\gamma(1) = b$. On a donc $f(0) = \varphi(a)$ et $f(1) = \varphi(b)$. Comme f est de classe C^1, le chapitre 5 (voir la remarque 5.8) donne

$$\varphi(b) - \varphi(a) = f(1) - f(0) = \int_0^1 f'(t) = \int_0^1 d\varphi(\gamma(t))(\gamma'(t))dt.$$

Le terme de droite, c'est-à-dire $\int_0^1 d\varphi(\gamma(t))(\gamma'(t))dt$, s'appelle "intégrale de la différentielle de φ le long du chemin $\{\gamma(t),\, t \in [0,1]\}$" (ce "chemin" est un "arc de courbe" dans $\mathbb{R}^2$). Nous venons de démontrer que cette intégrale ne dépend donc que des extrémités du chemin c'est-à-dire de a et b (et, bien sûr, de φ) et non de l'ensemble du chemin (pourvu que celui ci soit de classe C^1, cette hypothèse peut être légèrement affaiblie).

On utilise parfois la notion de dérivée dans une direction donnée, on donne ci après la définition de cette dérivée (en se limitant au cas des fonctions à valeurs dans $\mathbb{R}$).

Définition 7.8 (Dérivée directionnelle) *Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$. Soit $a \in \Omega$ et $x \in \mathbb{R}^n$, $x \neq 0$. Pour $t \in \mathbb{R}$ t.q. $a + tx \in \Omega$, on pose $g(t) = f(a + xt)$ (comme Ω est ouvert, la fonction g est donc définie sur un voisinage de 0). On dit que f admet une dérivée en a dans la direction x si la fonction $t \mapsto \frac{g(t)-g(0)}{t}$ admet une limite à droite en 0. On note $f_x'(a)$ cette limite, c'est-à-dire :*

$$f_x'(a) = \lim_{t \to 0, t > 0} \frac{f(a + tx) - f(a)}{t}.$$

Avec les notations de la définition 7.8, il est assez facile de voir que si f est différentiable en a, alors, pour tout $x \in \mathbb{R}^n$, $x \neq 0$, f admet une dérivée en a dans la direction x et cette dérivée est $f_x'(a) = df(a)(x)$. Ceci est une conséquence de la proposition 7.4 (voir l'exercice 7.3). Par contre l'existence de la dérivée directionnelle en a dans toutes les directions (c'est-à-dire pour tout $x \neq 0$) ne donne par la différentiabilité de f en a (exercice 7.4).

Définition 7.9 (Norme d'une application linéaire) *Soit $n, p \in \mathbb{N}^\star$ et f une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$. On définit alors la norme de f par la formule suivante :*

$$\|f\| = \sup_{x \in \mathbb{R}^n, x \neq 0} \frac{|f(x)|}{|x|}.$$

(Cette application est bien une norme sur l'ensemble des applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}^p$, qui est un espace vectoriel sur $\mathbb{R}$.)

Si f une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$, on peut remarquer que

$$|f(x)| \le \|f\||x| \;\; \text{pour tout } x \in \mathbb{R}^n. \tag{7.5}$$

Soit f continue de $\mathbb{R}$ dans $\mathbb{R}$, dérivable en tout point de $\mathbb{R}$. On rappelle (théorème 3.2) que si $a, b \in \mathbb{R}$, $a \neq b$, il existe $c \in]a, b[$ t.q. $f(b) - f(a) = f'(c)(b - a)$. Ce résultat peut être mis en défaut pour une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$ dès que $p > 1$, comme le montre l'exemple suivant :

On prend ici $n = 1$, $p = 2$ et, pour tout $x \in \mathbb{R}$, $f(x)$ est le vecteur dont les composantes sont $\cos(x)$ et $\sin(x)$, c'est-à-dire $f(x) = (\cos(x), \sin(x))^t$. L'application f est de classe C^1 sur $\mathbb{R}$ et la matrice jacobienne de f au point x est $M_f(x) = (-\sin(x), \cos(x))$. On prend maintenant $a = 0$ et $b = 2\pi$ et on remarque que, pour tout $c \in \mathbb{R}$ on $f(b) - f(a) \neq df(c)(b - a)$. En effet, on a $f(b) - f(a) = 0$ et $df(c)(b - a) = M_f(c)(b - a) = 2\pi(-\sin(c), \cos(c))^t \neq 0$ car $\sin(c)$ et $\cos(c)$ ne s'annule jamais tous les deux pour une même valeur de c.

On va toutefois donner une généralisation convenable du théorème des accroissements finis (théorème 7.1). Pour cela, si $n \ge 1$ et $a, b \in \mathbb{R}^n$, on note $]a, b[= \{ta + (1 - t)b,\ t \in]0, 1[\}$ et $[a, b] = \{ta + (1 - t)b,\ t \in [0, 1]\}$

Théorème 7.1 (Théorème des Accroissements Finis, cas vectoriel) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application continue de Ω dans $\mathbb{R}^p$. Soit $a, b \in \Omega$ t.q. $[a, b] \subset \Omega$, $a \neq b$. On suppose f différentiable en tout point de $]a, b[$. On a alors :*

$$|f(b) - f(a)| \le \left(\sup_{c \in]a,b[} \|df(c)\| \right) |b - a|.$$

Démonstration : On donne ici une démonstration consistant à se ramener au théorème des accroissements finis pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$ (théorème 3.2). Mais, cette démonstration est limitée au choix de la norme euclidienne dans l'espace $\mathbb{R}^p$ (choix que nous avons fait pour tout ce chapitre) alors que le résultat est vrai pour tout choix de norme sur $\mathbb{R}^n$ et $\mathbb{R}^p$ (la démonstration doit alors être légèrement modifiée).

Pour $t \in [0, 1]$, on a $tb + (1 - t)a \in \Omega$ et on pose $\varphi(t) = (f(tb + (1 - t)a) - f(a)) \cdot (f(b - f(a))$. La fonction φ est donc continue de $[0, 1]$ dans $\mathbb{R}$. Par dérivation de fonctions composées, elle est aussi dérivable sur $]0, 1[$ et on a, pour tout $t \in]0, 1[$,

$$\varphi'(t) = df(tb + (1 - t)a)(b - a) \cdot (f(b) - f(a)).$$

On peut alors utiliser le théorème 3.2 sur la fonction φ. On obtient qu'il existe $s \in]0, 1[$ t.q.

$$\varphi(1) - \varphi(0) = \varphi'(s) = df(sb + (1 - s)a)(b - a) \cdot (f(b) - f(a)).$$

On remarque maintenant que $\varphi(1) = |f(b) - f(a)|^2$ et $\varphi(0) = 0$. En posant $y = sb + (1 - s)a$ et en utilisant l'inégalité de Cauchy-Schwarz (inégalité (7.4)) on a donc

$$|f(b) - f(a)|^2 = df(y)(b - a) \cdot (b - a) \le |df(y)(b - a)||f(b) - f(a)|.$$

On en déduit, avec (7.5)

$$|f(b) - f(a)| \le \|df(y)\||b - a| \le \left(\sup_{c \in]a,b[} \|df(c)\| \right) |b - a|.$$

Ce qui termine cette démonstration. ■

Définition 7.10 (Fonctions de classe C^k) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$.*

1. *On dit que f est classe C^0 sur Ω si f est continue sur Ω, c'est-à-dire continue en tout point de Ω. On note alors $f \in C(\Omega, \mathbb{R}^p)$ ou $f \in C^0(\Omega, \mathbb{R}^p)$.*
2. *On dit que f est classe C^1 sur Ω si f est différentiable en tout point de Ω et que, pour tout $(i, j) \in \{1, \ldots, p\} \times \{1, \ldots, n\}$, la fonction $\frac{\partial f_i}{\partial x_j}$ est continue sur Ω. On note alors $f \in C^1(\Omega, \mathbb{R}^p)$.*
3. *Pour $k \geq 2$, on dit que f est de classe C^k (et on écrit $f \in C^k(\Omega, \mathbb{R}^p)$) si f est différentiable en tout point de Ω et les fonctions $\frac{\partial f_i}{\partial x_j}$ sont de classe C^{k-1} sur Ω (pour tout $(i, j) \in \{1, \ldots, p\} \times \{1, \ldots, n\}$).*
4. *On dit que f est classe C^∞ sur Ω si f est de classe C^k pour tout $k \in \mathbb{N}$.*

Proposition 7.5 (C^k est un e.v.) *Soit $n, p \in \mathbb{N}^\star$, $k \in \mathbb{N} \cup \{+\infty\}$ et Ω un ouvert de $\mathbb{R}^n$. Alors, $C^k(\Omega, \mathbb{R}^p)$ est un espace vectoriel sur $\mathbb{R}$.*

Proposition 7.6 (Existence de la différentielle) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. On suppose que f admet des dérivées partielles en tout point de Ω.*

1. *Soit $a \in \Omega$. On suppose que les dérivées partielles de f sont continues en a. Alors, f est différentiable en a.*
2. *On suppose que $f \in C^1(\Omega, \mathbb{R}^p)$. Alors , f est différentiable en tout point de Ω.*

Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Si f est de classe C^2, les dérivées partielles de $\frac{\partial f_i}{\partial x_j}$ admettent elle-mêmes des dérivées partielles. On note, pour $a \in \Omega$, $i \in \{1, \ldots, p\}$ et $(j, k) \in \{1, \ldots, n\}^2$:

$$\frac{\partial^2 f_i}{\partial x_k \partial x_j}(a) = \frac{\partial(\frac{\partial f_i}{\partial x_j})}{\partial x_k}(a).$$

Le théorème 7.2 compare $\frac{\partial^2 f_i}{\partial x_k \partial x_j}(a)$ et $\frac{\partial^2 f_i}{\partial x_j \partial x_k}(a)$.

Théorème 7.2 (Dérivées croisées) *Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. On suppose f de classe C^2. On a alors $\frac{\partial^2 f_i}{\partial x_k \partial x_j}(a) = \frac{\partial^2 f_i}{\partial x_j \partial x_k}(a)$ pour tout $a \in \Omega$, tout $i \in \{1, \ldots, p\}$ et tout $(j, k) \in \{1, \ldots, n\}^2$.*

Remarque 7.7 Voici une conséquence intéressante du théorème 7.2. Soit Ω un ouvert de $\mathbb{R}^2$ et g_1, g_2 deux applications de Ω dans $\mathbb{R}$, de classe C^1. On se demande si il existe une application f de Ω dans $\mathbb{R}$, différentiable et t.q., pour tout $a \in \Omega$, $\frac{\partial f}{\partial x_1}(a) = g_1(a)$ et $\frac{\partial f}{\partial x_2}(a) = g_2(a)$ (dans ce cas, on a $df = g_1 dx_1 + g_2 dx_2$). Une condition nécessaire pour l'existence de f est donnée par le théorème 7.2, il faut que, pour tout $a \in \Omega$, $\frac{\partial g_1}{\partial x_2}(a) = \frac{\partial g_2}{\partial x_1}(a)$. Si $\Omega = \mathbb{R}^2$ ou $\Omega = (\mathbb{R}_+^\star)^2$ (qui est une cas intéressant en thermodynamique) ou, plus généralement, si Ω est "étoilé par rapport à un point", cette condition est suffisante (voir l'exercice 7.6).

Enfin, on termine ce paragraphe en définissant le jacobien d'une application dfférentiable de $\mathbb{R}^n$ dans $\mathbb{R}^n$ (utile, par exemple, pour les calculs d'intégrales en utilisant des changements de variables).

Définition 7.11 (Jacobien) *Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^n$. Soit $a \in \mathbb{R}^n$. on suppose f différentiable en a. le jacobien de f est a, noté $J(a)$, est la valeur absolue du déterminant de la matrice jacobienne de f en a, c'est-à-dire :*

$$J(a) = |det(M_f(a))|.$$

7.3 Recherche d'un extremum

On s'intéresse dans cette section, à la recherche d'un point où une fonction f, définie sur un ouvert de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}$, atteint son maximum ou son minimum.

Définition 7.12 (Point critique, minima) *Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$. Soit $a \in \Omega$.*

1. *On dit que a est un point critique de f si f est différentiable en a et $\nabla f(a) = 0$.*
2. *On dit que f atteint un minimum local en a si il existe $\delta > 0$ t.q. $f(a) \le f(x)$ pour tout $x \in B(a, \delta)$.*

Proposition 7.7 (Condition nécessaire de minimalité) *Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$. Soit $a \in \Omega$. On suppose que f atteint un minimum local en a. Alors, si f différentiable en a, on a $\nabla f(a) = 0$.*

Pour obtenir une condition nécessaire plus précise que celle donnée dans la proposition 7.7 (et aussi pour obtenir une condition suffisante), on suppose dans le proposition 7.8 que f est de classe C^2 sur Ω et on introduit la matrice hessienne de f.

Définition 7.13 (Matrice hessienne) *Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$, de classe C^2. Soit $a \in \Omega$. On appelle matrice hessienne de f en a la matrice à n lignes et n colonnes dont le terme à la i-ieme ligne et j-ieme colonne est $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$. On note $H(a)$ cette matrice (c'est une matrice symétrique, grâce au théorème 7.2).*

Proposition 7.8 (Condition nécessaire et condition suffisante de minimalité)
Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$, de classe C^2. Soit $a \in \Omega$.

1. *(Condition nécessaire) On suppose que f atteint un minimum local en a. On a alors $\nabla f(a) = 0$ et $H(a)$ (matrice hessienne de f en a) est (symétrique) semi-définie positive (c'est-à-dire $H(a)h \cdot h \ge 0$ pour tout $h \in \mathbb{R}^n$).*
2. *(Condition suffisante) On suppose que $\nabla f(a) = 0$ et que $H(a)$ est (symétrique) définie positive (c'est-à-dire $H(a)h \cdot h > 0$ pour tout $h \in \mathbb{R}^n$, $h \ne 0$). Alors, f atteint un minimum local en a.*

7.4 Exercices

Exercice 7.1 (Différentiable implique continu)
Soit $n, p \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}^p$. Soit $a \in \Omega$. On suppose que f est différentiable en a. Montrer que f est continue.

Exercice 7.2 (Existence des dérivées partielles n'implique pas continuité)
Soit f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par :

$$f(x) = \frac{x_1 x_2}{x_1^2 + x_2^2} \text{ pour } x = (x_1, x_2)^t \in \mathbb{R}^2,\, x \ne 0,$$

$$f(0) = 0.$$

Montrer que f admet des dérivées partielles en 0 mais que f n'est pas continue en 0 (et donc que f n'est pas différentiable en 0).

Exercice 7.3 (Différentiable implique existence des dérivées directionnelles)

Soit $n \in \mathbb{N}^\star$, Ω un ouvert de $\mathbb{R}^n$ et f une application de Ω dans $\mathbb{R}$. Soit $a \in \Omega$ et $x \in \mathbb{R}^n$, $x \neq 0$. On suppose que f est différentiable en a, montrer que f est dérivable en a dans la direction x et que $f'_x(a) = df(a)(x)$.

Exercice 7.4 (Existence des dérivées directionnelles n'implique pas continuité)

Pour $x = (x_1, x_2)^t \in \mathbb{R}^2$, $x \neq 0$, il existe un et un seul couple $(\rho, \theta) \in \mathbb{R}_+^\star \times [0, 2\pi[$ t.q. $x_1 = \rho\cos(\theta)$ et $x_2 = \rho\sin(\theta)$. On pose alors $f(x) = \frac{\rho^2}{2\pi-\theta}$. On pose aussi $f(0) = 0$ (de sorte que f est définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$).

1. Soit $x \in \mathbb{R}^2$, $x \neq 0$. Montrer que f admet une dérivée en 0 dans la direction x et calculer $f'_x(0)$.

2. Montrer qu'il existe une application T, linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$, t.q., pour tout $x \in \mathbb{R}^2$, $x \neq 0$, $f'_x(0) = T(x)$.

3. Montrer que f n'est pas continue en 0 (et donc n'est pas différentiable en 0).

Exercice 7.5 (Mécanique, Analyse et Algèbre)

Soient x et y deux fonctions de $\mathbb{R}$ dans $\mathbb{R}^2$, de classe C^1. On désigne par $\mathcal{M}_2(\mathbb{R})$ l'ensemble des matrices 2×2 à coefficients dans $\mathbb{R}$. On note $M(t) \in \mathcal{M}_2(\mathbb{R})$ la matrice dont les deux lignes sont formées par les vecteurs (lignes) $x(t)$ et $y(t)$, $M_x(t)$ la matrice dont les deux lignes sont formées par les vecteurs (lignes) $x'(t)$ et $y(t)$ et $M_y(t)$ la matrice dont les deux lignes sont formées par les vecteurs (lignes) $x(t)$ et $y'(t)$.

1. Soit f une fonction bilinéaire de $\mathbb{R}^2 \times \mathbb{R}^2$ dans $\mathbb{R}$ (c'est-à-dire que, pour tout $u \in \mathbb{R}^2$, les applications $z \mapsto f(z, u)$ et $z \mapsto f(u, z)$ sont des applications linéaires de $\mathbb{R}^2$ dans $\mathbb{R}$). On rappelle qu'une application bilinéaire est nécessairement continue. Pour tout réel t, on pose $g(t) = f(x(t), y(t))$. Montrer que g est une application de $\mathbb{R}$ dans $\mathbb{R}$, de classe C^1, et que
$$g'(t) = f(x'(t), y(t)) + f(x(t), y'(t))$$
pour tout $t \in \mathbb{R}$. [On pourra, par exemple, écrire $f(x(t+h), y(t+h)) - f(x(t), y(t)) = f(x(t+h) - x(t), y(t+h)) + f(x(t), y(t+h) - y(t))$.]

2. En utilisant la question 1 et l'application $N \mapsto \det(N)$ (de $\mathcal{M}_2(\mathbb{R})$ dans $\mathbb{R}$), montrer que $\varphi'(t) = \det(M_x(t)) + \det(M_y(t))$, avec $\varphi(t) = \det(M(t))$.

3. On suppose que pour tout $t \in \mathbb{R}$, on suppose qu'il existe une matrice $A(t) = \begin{bmatrix} a_{1,1}(t) & a_{1,2}(t) \\ a_{2,1}(t) & a_{2,2}(t) \end{bmatrix}$ t.q., pour tout $t \in \mathbb{R}$, $M'(t) = A(t)M(t)$. Montrer que $M_x(t) = A_x(t)M(t)$ et $M_y(t) = A_y(t)M(t)$ avec
$$A_x = \begin{bmatrix} a_{1,1}(t) & a_{1,2}(t) \\ 0 & 1 \end{bmatrix},\ A_y = \begin{bmatrix} 1 & 0 \\ a_{2,1}(t) & a_{2,2}(t) \end{bmatrix}.$$

4. Montrer que $\det(M_x(t)) = a_{1,1}(t)\det(M(t))$ et $\det(M_y(t)) = a_{2,2}(t)\det(M(t))$. En déduire que $\varphi'(t) = \mathrm{trace}(A(t))\varphi(t)$ pour tout $t \in \mathbb{R}$ (avec $\varphi(t) = \det(M(t))$).

NB: Les courageux pourront montrer que le résultat de la question 4 (c'est-à-dire $\varphi'(t) = \mathrm{trace}(A(t))\varphi(t)$ pour tout $t \in \mathbb{R}$) est encore vrai si la matrice M(t) est formée de n vecteurs lignes $x_1(t), \ldots, x_n(t)$ de $\mathbb{R}^n$ et que $M'(t) = A(t)M(t)$ avec $n \geq 1$ et $x_1, \ldots, x_n$ fonctions de $\mathbb{R}$ dans $\mathbb{R}^n$, de classe C^1.

Exercice 7.6 (CNS pour une forme différentielle exacte)

Soit Ω un ouvert de $\mathbb{R}^2$ et g_1, g_2 deux applications de Ω dans $\mathbb{R}$, de classe C^1 et t.q., pour tout $a \in \Omega$, $\frac{\partial g_1}{\partial x_2}(a) = \frac{\partial g_2}{\partial x_1}(a)$. On suppose aussi qu'il existe $b \in \Omega$ t.q. pour tout $a \in \Omega$, l'intervalle $[b, a]$ est inclus dans Ω (on dit alors que Ω est étoilé par rapport à b). On note b_1 et b_2 les composantes de b. Pour $a = (a_1, a_2)^t \in \Omega$, on pose

$$f(a) = (a_1 - b_1) \int_0^1 g_1(ta + (1-t)b)dt + (a_2 - b_2) \int_0^1 g_2(ta + (1-t)b)dt.$$

Montrer que f est une application différentiable de Ω dans $\mathbb{R}$ et que, pour tout $a \in \Omega$, $\frac{\partial f}{\partial x_1}(a) = g_1(a)$ et $\frac{\partial f}{\partial x_2}(a) = g_2(a)$. (On a donc $df = g_1 dx_1 + g_2 dx_2$.)