CHAPITRE 3

APPLICATIONS DIFFÉRENTIABLES

5. Définitions et premières propriétés

[(1)] En guise d'introduction, commençons par un rappel sur les fonctions dérivables $\mathbb{R} \to \mathbb{R}$. Soient $a \in \mathbb{R}$ et $f : \mathbb{R} \to \mathbb{R}$ une fonction définie sur un intervalle ouvert I contenant a. On dit que f est *dérivable* en a si la limite

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}$$

existe ; dans ce cas elle est notée $f'(a)$ et l'on a $f(a+h) = f(a) + f'(a)h + h\varepsilon(h)$, où $\varepsilon : \mathbb{R} \to \mathbb{R}$ est une fonction *continue et nulle* en 0 (i.e. $\lim_{h\to 0} \varepsilon(h) = 0 = \varepsilon(0)$).

Dans ce cas, connaissant $f(a)$ on peut obtenir une assez bonne approximation de $f(x)$ au voisinage de a, i.e. de $f(a+h)$ pour h assez petit, en remplaçant $f(a+h)$ par la fonction « linéaire » $h \mapsto f(a)+f'(a)h$. Stricto sensu, c'est une application affine, mais comme la valeur en 0 est donnée par $f(a)$, « ce qui compte » est l'application linéaire $h \mapsto f(a+h) - f(a) = f'(a)h$. En résumé, on considère que $f : \mathbb{R} \to \mathbb{R}$ est une « bonne » fonction au voisinage d'un point a si elle possède une « *bonne approximation* » par une application linéaire $L_a : h \mapsto L_a(h)$, la notion de « bonne approximation » ayant le sens précis que la différence

$$f(a+h) - f(a) - L_a(h)$$

tend vers 0 plus vite que $|h|$, i.e. que $f(a+h) = f(a)+L_a(h)+|h|\varepsilon(h)$, où $\varepsilon : \mathbb{R} \to \mathbb{R}$ est une fonction *continue et nulle* en 0. C'est sous cette forme que la notion de fonction dérivable va s'étendre aux fonctions $\mathbb{R}^n \to \mathbb{R}^p$.

Notations 5.1. — Dans la suite, on munit $\mathbb{R}^n$ et $\mathbb{R}^p$ de la norme $\|\cdot\|_\infty$ (on pourrait aussi choisir la norme euclidienne $\|\cdot\|_2$).

On note $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$ l'espace vectoriel (de dimension np) des applications linéaires $\mathbb{R}^n \to \mathbb{R}^p$. On rappelle qu'une telle application est lipschitzienne donc continue. D'autre part, $\mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$ s'identifie à $M_{p,n}(\mathbb{R})$, l'espace de matrices à p lignes et n colonnes.

En particulier, si $n = 1$ alors $\mathscr{L}(\mathbb{R}, \mathbb{R}^p) = \mathbb{R}^p$ car une application linéaire $f : \mathbb{R} \to \mathbb{R}^p$ est déterminée par la donnée du vecteur $u = f(1)$, i.e. on a $f(t) = tu$ pour tout $t \in \mathbb{R}$. En termes

[(1)] Version du 20 mai 2016

matriciels, on a $\mathscr{L}(\mathbb{R},\mathbb{R}^p) = M_{p,1}(\mathbb{R})$, ce qui nous conduira dans la suite à représenter un élément $(y_1,\dots,y_p)$ de $\mathbb{R}^p$ par le vecteur **colonne** $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$.

Définition 5.2 **(Limites en** 0**).** — Soit $\varphi : \mathbb{R}^n \to \mathbb{R}^p$ une fonction définie sur une boule ouverte $B(0,R)$, sauf en 0, et soit $b \in \mathbb{R}^p$. On écrit

$$\lim_{\substack{h\to 0\\ h\neq 0}} \varphi(h) = b$$

si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour tout $h \neq 0$ vérifiant $\|h\| < \delta$ on ait $\|\varphi(h) - b\| < \varepsilon$. Dans ce cas, si l'on prolonge φ en 0 en posant $\varphi(0) = b$, on obtient une fonction qui est définie sur $B(0,R)$ et continue en 0. Si de plus $b = 0$, on dira que la fonction ainsi prolongée est « continue et nulle en 0 ».

D'autre part, si l'on note $\varphi_1,\dots,\varphi_p$ les composantes de φ et $b_1,\dots,b_p$ celles de b, alors comme $\|\phi(h) - b\| = \max_{i=1,\dots,p} |\varphi_i(h) - b_i|$, on voit que la condition précédente équivaut à dire que, pour tout $i = 1,\dots,p$,

$$\lim_{\substack{h\to 0\\ h\neq 0}} \varphi_i(h) = b_i.$$

(Q) ***Définition 5.3*** **(Fonctions dérivables** $\mathbb{R} \to \mathbb{R}^p$**).** — Soient $a \in \mathbb{R}$ et $f : \mathbb{R} \to \mathbb{R}^p$ une fonction définie sur un intervalle ouvert I contenant a. On « rappelle » que f est **dérivable** en a si la limite

$$\ell = \lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a)}{h}$$

existe dans $\mathbb{R}^p$; dans ce cas elle est notée $f'(a)$ et l'on a $f(a+h) = f(a) + f'(a)h + h\varepsilon(h)$, où $\varepsilon : \mathbb{R} \to \mathbb{R}^p$ est une fonction *continue et nulle* en 0 (i.e. $\lim_{h\to 0} \varepsilon(h) = 0 = \varepsilon(0)$).

Notant $f_1,\dots,f_p$ les composantes de f et $\ell_1,\dots,\ell_p$ celles de ℓ, on voit que la condition précédente équivaut à dire que, pour tout $i = 1,\dots,p$,

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{f(a+h) - f(a)}{h} = \ell_i$$

i.e. que f_i est dérivable en a de dérivée $f_i'(a) = \ell_i$. On obtient donc que f est dérivable en a ssi les f_i le sont, et dans ce cas le vecteur dérivé $f'(a)$ est

$$f'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_p'(a) \end{pmatrix}.$$

Notant $L_a(h)$ l'application linéaire $\mathbb{R} \to \mathbb{R}^p$, $h \mapsto hf'(a)$, on obtient donc que

$$f(a+h) - f(a) - L_a(h) = h\varepsilon(h),$$

où $\varepsilon : \mathbb{R} \to \mathbb{R}^p$ est une fonction *continue et nulle* en 0. C'est sous cette forme que la notion de fonction dérivable s'étend aux fonctions $\mathbb{R}^n \to \mathbb{R}^p$, avec $n > 1$.

Commençons par le lemme suivant.

Lemme 5.4. — *Soit* $L : \mathbb{R}^n \to \mathbb{R}^p$ *une application linéaire. Si* $\lim_{\substack{h\to 0\\ h\neq 0}} \dfrac{L(h)}{\|h\|} = 0$ *alors* $L = 0$.

Démonstration. — Notons S la sphère unité $S = S(0,1)$ et soit $x \in S$. Fixons $\varepsilon > 0$. Par hypothèse, il existe $\delta > 0$ tel que pour $\|h\| < 2\delta$ on ait

$$\|L(h)\| \leq \varepsilon \|h\|.$$

Posons $h = \delta x$, alors $\|h\| = \delta < 2\delta$ et donc

$$\|L(x)\| = \frac{1}{\delta} \|L(h)\| \leq \frac{1}{\delta} \varepsilon \|h\| = \varepsilon.$$

Comme ε est arbitraire, ceci entraîne $L(x) = 0$. Enfin, soit $y \in \mathbb{R}^n - \{0\}$ arbitraire. Alors $x = (1/\|y\|)y$ appartient à S donc $L(x) = 0$, et comme $y = \|y\| x$ on obtient $L(y) = 0$. $\square$

Terminologie 5.5. — Soient U un ouvert de $\mathbb{R}^n$, f une application $U \to \mathbb{R}^p$ et $a \in U$. Comme U est ouvert, il existe $r > 0$ tel que $B(a,r) \subset U$ et donc la fonction $h \mapsto f(a+h) - f(a)$ est définie pour $\|h\| < r$. On utilisera ceci de façon implicite dans la suite quand on dira que certaines fonctions de h sont définies pour h « assez petit ».

(Q) ***Définition 5.6.*** — Soient U un ouvert de $\mathbb{R}^n$ et f une application $U \to \mathbb{R}^p$.

(i) On dit que f est **différentiable** en un point a de U s'il existe une application linéaire $L_a \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$ telle que

$$(*) \qquad \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a) - L_a(h)}{\|h\|} = 0.$$

D'après le lemme précédent, L_a est uniquement déterminée si elle existe. Dans ce cas, L_a est notée $df(a)$ et est appelée *différentielle* ou *application linéaire tangente* de f en a.

(ii) En posant $\varepsilon(h) = \dfrac{f(a+h) - f(a) - L_a(h)}{\|h\|}$ pour $h \neq 0$ et $\varepsilon(0) = 0$, $(*)$ équivaut à dire que pour h assez petit on a :

$$(\dagger) \qquad f(a+h) = f(a) + df(a)(h) + \|h\| \varepsilon(h),$$

avec ε continue et nulle en 0. On peut récrire cette égalité avec la notation $o()$:

$$(\ddagger) \qquad f(a+h) = f(a) + df(a)(h) + o(\|h\|),$$

où $o(\|h\|)$ désigne une fonction $\phi(h)$ telle que $\lim_{\substack{h \to 0 \\ h \neq 0}} \dfrac{\phi(h)}{\|h\|} = 0$. Avec l'une ou l'autre notation, ceci montre que si f est différentiable en a, elle admet en a un « développement de Taylor » à l'ordre 1, dont le terme linéaire est $df(a)(h)$.

(iii) On dit que f est **différentiable sur** U si elle est différentiable en tout point de U.

(Q) ***Exemples 5.7.*** — (1) Soit $f : \mathbb{R}^n \to \mathbb{R}^p$ linéaire. Alors f est différentiable sur $\mathbb{R}^n$ et pour tout $a \in \mathbb{R}^n$ on a $df(a) = f$. En effet, pour tout $h \in \mathbb{R}^n$ on a $f(a+h) - f(a) = f(h)$.

(2) L'application $Q : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto x \cdot x = (\|x\|_2)^2$ est différentiable sur $\mathbb{R}^n$. En effet, pour tout $a, h \in \mathbb{R}^n$ on a

$$Q(a+h) - Q(h) = a \cdot h + h \cdot a = 2a \cdot h$$

donc $dQ(a) : \mathbb{R}^n \to \mathbb{R}$ est la forme linéaire $h \mapsto 2a \cdot h$.

(3) L'application de multiplication $m : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto x_1 x_2$ est différentiable sur $\mathbb{R}^2$. En effet, pour $x = (x_1, x_2)$ et $h = (h_1, h_2)$ dans $\mathbb{R}^2$, on a

$$(x_1 + h_1)(x_2 + h_2) = x_1 x_2 + (x_2 h_1 + x_1 h_2) + h_1 h_2$$

et $|h_1 h_2| \leq \|h\|_\infty^2$, donc $dm(x)$ est la forme linéaire $\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \mapsto x_2 h_1 + x_1 h_2$, i.e. $dm(x)$ est donnée par la matrice ligne (x_2, x_1).

(4) Si $n = 1$, i.e. si I est un intervalle ouvert de $\mathbb{R}$ et $f = (f_1, \ldots, f_p)$ une application de I dans $\mathbb{R}^p$, alors f est différentiable en un point a de I ssi f est dérivable en a (i.e. chaque f_i l'est) et dans ce cas pour tout $h \in \mathbb{R}$ on a :

$$df(a)(h) = hf'(a) = h \begin{pmatrix} f_1'(a) \\ \vdots \\ f_p'(a) \end{pmatrix}.$$

En effet, on a vu en 5.3 que si f est dérivable en a elle y est différentiable et $df(a)$ est comme indiqué.

Réciproquement, supposons f différentiable en a. Comme tout $L \in \mathscr{L}(\mathbb{R}, \mathbb{R}^p)$ est de la forme $h \mapsto hu$ pour un certain $u \in \mathbb{R}^p$, ceci signifie qu'il existe $v \in \mathbb{R}^p$ tel que pour h assez petit on ait $f(a+h) - f(a) - hv = |h|\, \varepsilon(h)$ avec ε continue et nulle en 0, d'où

$$\left\| \frac{f(a+h) - f(a) - hv}{h} \right\| = \|\varepsilon(h)\|$$

pour $h \neq 0$, et donc

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h} = v.$$

Ceci montre que f est dérivable en a et $v = f'(a)$, et donc que $df(a)(h) = hv = hf'(a)$.

Remarque 5.8. — Si f est différentiable en a elle est continue en a, car $df(a)$ est continue (étant linéaire) et ε est continue en 0.

L'exemple (4) ci-dessus se généralise comme suit : En utilisant la norme $\|\cdot\|_\infty$ on obtient facilement la proposition suivante, qui généralise

(Q) ***Proposition 5.9.*** — *Soient U un ouvert de $\mathbb{R}^n$, $f = (f_1, \ldots, f_p)$ une application $U \to \mathbb{R}^p$ et $a \in U$. Alors f est différentiable en a ssi chaque f_i l'est, et dans ce cas on a*

$$df(a)(h) = \begin{pmatrix} df_1(a)(h) \\ \vdots \\ df_p(a)(h) \end{pmatrix}$$

pour tout $h \in \mathbb{R}^n$.

Démonstration. — Se donner une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ est « la même chose » que se donner p formes linéaires $L_i : \mathbb{R}^n \to \mathbb{R}$, i.e. on a

$$L(h) = \begin{pmatrix} L_1(h) \\ \vdots \\ L_p(h) \end{pmatrix}$$

pour tout $h \in \mathbb{R}^n$. (Du point de vue matriciel, on a $L \in M_{p,n}(\mathbb{R})$ et les L_i correspondent aux p lignes de cette matrice.) De même, la fonction $\varepsilon : \mathbb{R}^n \to \mathbb{R}^p$ qui apparaît dans la définition 5.6

s'écrit

$$\varepsilon(h) = \begin{pmatrix} \varepsilon_1(h) \\ \vdots \\ \varepsilon_p(h) \end{pmatrix}$$

et la condition que ε soit continue et nulle en 0 équivaut à dire que chaque ε_i l'est. On voit donc que f est différentiable en a ssi il existe des formes linéaires $L_1, \dots, L_p$ et des fonctions $\varepsilon_1, \dots, \varepsilon_p : \mathbb{R}^n \to \mathbb{R}$ continues et nulles en 0 telles que

$$f_i(a+h) = f_i(a) + L_i(a)(h) + \|h\| \, \varepsilon_i(h),$$

ce qui équivaut à dire que chaque f_i est différentiable en a et $L_i = df_i(a)$, et dans ce cas pour tout $h \in \mathbb{R}^n$ on a bien

$$df(a)(h) = \begin{pmatrix} L_1(h) \\ \vdots \\ L_p(h) \end{pmatrix} = \begin{pmatrix} df_1(a)(h) \\ \vdots \\ df_p(a)(h) \end{pmatrix}.$$

□

Avant de démontrer le théorème sur la composée d'applications différentiables, introduisons les dérivées partielles et la matrice jacobienne, qui donnerons un aspect plus concret à la notion de différentielle.

(Q) ***Définition 5.10* (Dérivée selon la direction v).** — Soit U un ouvert de $\mathbb{R}^n$, f une application $U \to \mathbb{R}$ et $a \in U$. Soit $v \in \mathbb{R}^n$ non nul. On dit que f admet en a une *dérivée partielle dans la direction v* si la fonction $\mathbb{R} \to \mathbb{R}$, $t \mapsto f(a+tv)$ est dérivable en $t = 0$, i.e. si la limite

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a+tv) - f(a)}{t}$$

existe, auquel cas elle est notée $\dfrac{\partial f}{\partial v}(a)$.

Remarque 5.11. — Intuitivement, la dérivée directionnelle $\dfrac{\partial f}{\partial v}(a)$ mesure les variation de f lorsqu'on se déplace au voisinage de a dans la direction v à la vitesse $\|v\|$. **Attention :** la terminologie est légèrement abusive, car cette dérivée dépend de v lui-même, et pas seulement de la direction $\mathbb{R}v$. En effet, si on remplace v par un multiple non nul $w = \lambda v$, alors

$$\frac{\partial f}{\partial w}(a) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a+\lambda t v) - f(a)}{t} = \lambda \lim_{\substack{\lambda t \to 0 \\ t \neq 0}} \frac{f(a+\lambda t v) - f(a)}{\lambda t} = \lambda \frac{\partial f}{\partial v}(a).$$

Ceci explique la remarque « intuitive » plus haut (en prenant pour « unité de vitesse » celle donnée par le vecteur unitaire $u = \dfrac{1}{\|v\|} v$).

Conservant les notations précédentes, on a en particulier :

(Q) ***Définition 5.12* (Dérivées partielles).** — Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. On note, si elle existe, $\dfrac{\partial f}{\partial x_i}(a)$ ou simplement $\partial_i f(a)$ la dérivée partielle de f en a dans la direction e_i, i.e.

$$\partial_i f(a) = \frac{\partial f}{\partial x_i}(a) = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a+te_i) - f(a)}{t} \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a_1, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) - f(a)}{t}$$

et l'on dit que c'est la dérivée partielle de f en a selon la i-ème variable.

Notation 5.13. — Si f admet des dérivées partielles $\dfrac{\partial f}{\partial x_i}(a)$ en tout point $a \in U$, on obtient ainsi pour tout $i = 1, \ldots, n$ une application :

$$\frac{\partial f}{\partial x_i} : U \to \mathbb{R}, \qquad a \mapsto \frac{\partial f}{\partial x_i}(a).$$

Avec la notation ci-dessus, il n'y a pas d'ambiguïté, mais souvent on écrit $a = (x_1, \ldots, x_n)$, d'où l'application

$$\frac{\partial f}{\partial x_i} : U \to \mathbb{R}, \qquad (x_1, \ldots, x_n) \mapsto \frac{\partial f}{\partial x_i}(x_1, \ldots, x_n).$$

On prendra garde que dans cette écriture, le x_i dans ∂x_i est un symbole pour désigner la dérivation selon le vecteur e_i, tandis que $(x_1, \ldots, x_n)$ est une « variable » qui décrit l'ouvert U.

Le calcul de $\dfrac{\partial f}{\partial x_i}(a)$ consiste donc à ne dériver l'expression de f que par rapport à la variable x_i.

Exemples 5.14. — (1) Si $f : \mathbb{R}^3 \to \mathbb{R}$ est définie par $f(x, y, z) = -2x \cos y$, on a :

$$\frac{\partial f}{\partial x}(x, y, z) = -2\cos y, \quad \frac{\partial f}{\partial y}(x, y, z) = 2x \sin y, \quad \frac{\partial f}{\partial z}(x, y, z) = 0.$$

(2) Soit $f : \mathbb{R}_+^* \times \mathbb{R} \to \mathbb{R}$ définie par $f(x, y) = \mathrm{Arctan}(y/x)$. Comme la dérivée de $\mathrm{Arctan}(u)$ est $\dfrac{1}{1+u^2}$, on a :

$$(\dagger) \qquad \frac{\partial f}{\partial x}(x, y) = \frac{1}{1 + (y/x)^2}\,\frac{-y}{x^2} = \frac{-y}{x^2 + y^2}, \qquad \frac{\partial f}{\partial y}(x, y) = \frac{1}{1 + (y/x)^2}\,\frac{1}{x} = \frac{x}{x^2 + y^2}.$$

Lemme 5.15. — *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^p$. Si f est différentiable en $a \in U$ alors,*
(Q) *pour tout $v \in \mathbb{R}^n - \{0\}$, f admet en a une dérivée dans la direction v et l'on a*

$$\frac{\partial f}{\partial v}(a) = df(a)(v).$$

En particulier, f admet en a des dérivées partielles $\dfrac{\partial f}{\partial x_i}(a)$ et pour tout $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ on a :

$$df(a)(h) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) h_j.$$

Démonstration. — La fonction $\phi : t \mapsto f(a + tv)$ est la composée de f et de la fonction dérivable $u : t \mapsto a + tv$, de dérivée $u'(t) = v$. D'après le théorème 5.17, ϕ est dérivable en $t = 0$ de vecteur dérivé $df(u(0))(u'(0)) = df(a)(v)$, d'où $df(a)(v) = \dfrac{\partial f}{\partial v}(a)$.

En particulier, pour $v = e_j$ on obtient $df(a)(e_j) = \dfrac{\partial f}{\partial e_j}(a) = \dfrac{\partial f}{\partial x_j}(a)$. Enfin, comme $df(a)$ est linéaire, pour $h = (h_1, \ldots, h_n) = \sum_j h_j e_j$ on obtient :

$$df(a)(h) = \sum_j h_j df(a)(e_j) = \sum_j \frac{\partial f}{\partial x_j}(a)\, h_j.$$

□

Ainsi, une fonction f différentiable en un point a admet des dérivées partielles. On verra dans la section suivante que la réciproque est fausse en général, mais que si f admet sur un ouvert U des dérivées partielles qui sont *continues*, alors f est différentiable sur U.

On a vu plus haut (5.15) que si $f : \mathbb{R}^n \to \mathbb{R}$ est différentiable en a, alors $df(a)$ est la **forme linéaire** $\mathbb{R}^n \to \mathbb{R}$ donnée par

$$df(a)(h) = \frac{\partial f}{\partial x_1}(a)h_1 + \cdots + \frac{\partial f}{\partial x_n}(a)h_n$$

i.e. $df(a)$ est donnée par la matrice **ligne** :

$$(\star) \qquad \Big(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a)\Big).$$

Définition 5.16 **(Matrice jacobienne).** — Soit maintenant $f = (f_1, \dots, f_p)$ une fonction $\mathbb{R}^n \to \mathbb{R}^p$. D'après la proposition 5.9, f est différentiable en un point a ssi chaque f_i l'est, et dans ce cas pour tout $h \in \mathbb{R}^n$ on a :

$$df(a)(h) = \begin{pmatrix} df_1(a)(h) \\ \vdots \\ df_p(a) \end{pmatrix}.$$

(Q) Tenant compte de l'expression pour chaque $df_i(a)(h)$ donnée en $(\star)$ plus haut, on obtient que la matrice de $df(a)$ est la matrice

$$Df(a) = \Big(\frac{\partial f_i}{\partial x_j}\Big)_{\substack{1\le i\le p \\ 1\le j\le n}} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \cdots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

dont la i-ème ligne est donnée par $df_i(a)$. Cette matrice est appelée **matrice jacobienne** de f en a. Lorsque $p = n$, $Df(a)$ est une matrice carrée et on note $J_f(a)$ son déterminant, qu'on appelle le (déterminant) **jacobien** de f en a.

Revenant au cas n, p arbitraires, rappelons que pour toute application linéaire $u : \mathbb{R}^n \to \mathbb{R}^p$, de matrice $A \in M_{p,n}(\mathbb{R})$, l'image par u d'un vecteur $x = (x_1, \dots, x_n)$ s'obtient en appliquant A au vecteur **colonne** $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, i.e. l'élément $u(x)$ de $\mathbb{R}^p$ est donné par le vecteur colonne $AX \in \mathbb{R}^p$.

(Q) ***Théorème 5.17.*** — *Soient U, V des ouverts de $\mathbb{R}^n$ et $\mathbb{R}^p$ et $f : U \to V$ et $g : V \to \mathbb{R}^q$ des applications.*

(i) *Si f est différentiable en a et g en $f(a)$, alors $g \circ f$ est différentiable en a et l'on a :*

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

(ii) *Si f est différentiable sur U et g sur V, alors $g \circ f$ est différentiable sur U.*

(iii) *En particulier, si $n = 1$ et $U = I$ est un intervalle ouvert de $\mathbb{R}$, l'application $g \circ f : I \to \mathbb{R}^p$ est dérivable et pour tout $t \in I$ on a :*

$$(*) \qquad (g \circ f)'(t) = dg(f(t))(f'(t)).$$

Démonstration. — (i) Posons $b = f(a)$. Par hypothèse, il existe des fonctions $\eta : \mathbb{R}^n \to \mathbb{R}^p$ et $\mu : \mathbb{R}^p \to \mathbb{R}^q$, continues et nulles en 0, telles que pour $h \in \mathbb{R}^n$ et $h' \in \mathbb{R}^p$ assez petits, on ait :

$$f(a+h) = b + df(a)(h) + \|h\|\,\eta(h)$$
$$g(b+h') = g(b) + dg(b)(h') + \|h'\|\,\mu(h').$$

Pour h assez petit, posons

$$k(h) = f(a+h) - f(a) = df(a)(h) + \|h\|\,\eta(h).$$

Alors pour $h \neq 0$ on a

$$\frac{g(f(a+h)) - g(b) - dg(b)(df(a)(h))}{\|h\|} = dg(b)(\eta(h)) + \frac{\|k(h)\|}{\|h\|}\mu(k(h)).$$

Montrons que le membre de droite tend vers 0 quand $h \neq 0$ tend vers 0. Pour le premier terme c'est clair, car $dg(b)$ est continue (car linéaire) et η est continue et nulle en 0.

Notons $\psi(h)$ le second terme. Comme η est continue et nulle en 0, il existe $\delta_0 > 0$ tel que $\|\eta(h)\| < 1$ si $\|h\| < \delta_0$. Comme $df(a)$ est L-lipschitzienne, posant $C = L+1$, on obtient que pour tout h tel que $\|h\| < \delta_0$, on a

$$\|k(h)\| \leq C\,\|h\| \qquad \text{et donc} \qquad \|\psi(h)\| \leq C\,\|\mu(k(h))\|.$$

Comme $\mu \circ k$ est continue et nulle en 0, il en résulte que $\lim_{\substack{h\to 0\\ h\neq 0}} \psi(h) = 0$. Ceci prouve (i) et (ii).

Déduisons-en le cas particulier (iii). D'après 5.7 (4), une application $\phi : I \to \mathbb{R}^q$ est différentiable en un point t ssi elle est dérivable en t et dans ce cas $d\phi(t)$ est l'application linéaire $h \mapsto h\phi'(t)$.

Ici, on sait d'après (i), que $g \circ f$ est différentiable en t, de différentielle $dg(f(t)) \circ df(t)$. Or $df(t)$ est l'application linéaire $\mathbb{R} \to \mathbb{R}^n$, $h \mapsto hf'(t)$ et donc $d(g \circ f)(t)$ est l'application linéaire $\mathbb{R} \to \mathbb{R}^n$, $h \mapsto hdg(f(t))(f'(t))$. Il en résulte que $(g \circ f)'(t) = dg(f(t))(f'(t))$. □

Remarque 5.18. — La définition de la différentiabilité et le théorème précédent illustrent un principe général en mathématiques : il a fallu travailler un peu pour établir la définition (i.e. montrer que $df(a)$ est unique si elle existe) puis pour démontrer le théorème, mais ce travail ayant été fait une fois pour toutes, on dispose d'une notion qui est facile à manipuler, comme le montre la jolie formule $d(g \circ f)(a) = dg(f(a)) \circ df(a)$, dont on verra plus bas la traduction en termes de produit de matrices.

***Remarque 5.19* (Traduction matricielle)**. — Le théorème précédent s'écrit en termes matriciels comme suit. Considérons des ouverts $U \subset \mathbb{R}^n$ et $V \subset \mathbb{R}^p$ et des applications différentiables $f : U \to V$ et $g : V \to \mathbb{R}^q$. Pour tout $a \in U$, soit $A = Df(a) \in M_{p,n}(\mathbb{R})$ la matrice jacobienne de f en a et $B = Dg(f(a)) \in M_{q,p}(\mathbb{R})$ celle de g en $b = f(a)$. Alors la matrice jacobienne de $g \circ f$ en a est

$$D(g \circ f)(a) = BA$$

puisque $d(g \circ f)(a) = dg(f(a)) \circ f(a)$.

***Remarque 5.20* (Attention !)**. — Contrairement aux fonctions d'une seule variable, où l'on peut écrire $(g \circ f)'(a) = g'(f(a))f'(a) = f'(a)g'(f(a))$ (puisque le produit dans $\mathbb{R}$ est commutatif), l'ordre d'apparition des différentielles dans la formule $d(g \circ f)(a) = dg(f(a)) \circ f(a)$ est extrêmement important. En effet, $df(a)$ va de $\mathbb{R}^n \to \mathbb{R}^p$ et $dg(f(a))$ va de $\mathbb{R}^p$ dans $\mathbb{R}^q$, donc on ne peut même pas les composer dans le « mauvais sens » si $n \neq q$. Et même si $n = p = q$, la composition dans le « mauvais sens » ne donne pas le bon résultat, puisque la multiplication dans $M_n(\mathbb{R})$ n'est pas commutative.

Remarque 5.21. — Écrivons la différentielle d'une composée $\mathbb{R}^n \xrightarrow{f} \mathbb{R}^p \xrightarrow{g} \mathbb{R}^q$ en termes de dérivées partielles. Posons $A = Df(a)$ et $B = Dg(f(a))$, alors $D(g \circ f) = BA$. Donc pour tout $j = 1, \ldots, n$ et $i = 1, \ldots, q$, on a

$$(BA)_{ij} = \sum_{k=1}^{p} B_{ik}A_{kj}.$$

Si l'on note $(u_1, \ldots, u_p)$ les coordonnées sur $\mathbb{R}^p$, alors on a

$$B_{ik} = \frac{\partial g_i}{\partial u_k}(f(a)) \qquad \text{et} \qquad A_{kj} = \frac{\partial f_k}{\partial x_j}(a)$$

(Q) et donc l'égalité précédente donne :

$$(\dagger) \qquad \frac{\partial (g \circ f)_i}{\partial x_j}(a) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial u_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a).$$

Le calcul consiste donc à : dériver g_i par rapport à la variable u_k et évaluer le résultat en $f(a)$, puis multiplier par la dérivée de f_k par rapport à la variable x_j évaluée en a, puis sommer par rapport à k.

(Q) ***Proposition 5.22.*** — *Soient U un ouvert de $\mathbb{R}^n$, f, g deux applications différentiables $U \to \mathbb{R}^p$, et $a \in U$.*

(i) *Pour tout $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ est différentiable en a et $d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a)$.*

(ii) *Si $p = 1$, alors fg est différentiable en a et $d(fg)(a) = f(a)dg(a) + g(a)df(a)$.*

Démonstration. — (i) est laissé en exercice ; prouvons (ii). D'après la proposition 5.9, l'application $F : \mathbb{R}^n \to \mathbb{R}^2$, $x \mapsto (f(x), g(x))$ est différentiable en a, et pour tout $h \in \mathbb{R}^n$ on a

$$dF(a)(h) = \big(df(a)(h), dg(a)(h)\big).$$

D'autre part, d'après l'exemple 5.7 (3), l'application $m : \mathbb{R}^2 \to \mathbb{R}$, $(x_1, x_2) \mapsto x_1 x_2$, est différentiable sur $\mathbb{R}^2$ et sa différentielle en $F(a) = (f(a), g(a))$ est la forme linéaire $(h_1, h_2) \mapsto g(a)h_1 + f(a)h_2$.

D'après le théorème 5.17, l'application $fg = m \circ F$ est donc différentiable en a, de différentielle $d(fg)(a) = dm(F(a)) \circ dF(a)$, i.e. pour tout $h \in \mathbb{R}^n$ on a

$$d(fg)(a)(h) = dm(F(a))\big(df(a)(h), dg(a)(h)\big) = g(a)df(a)(h) + f(a)dg(a)(h)$$

i.e. $d(fg)(a)$ est la forme linéaire $g(a)df(a) + f(a)dg(a)$. □

Démontrons maintenant deux résultats qui utilisent la notion de connexité ou de convexité.

***Corollaire 5.23* (du théorème 5.17).** — *Soient $U \neq \varnothing$ un ouvert connexe de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^p$ différentiable. Si $df(a) = 0$ pour tout $a \in U$, alors f est constante sur U.*

Démonstration. — Soient $a, x \in U$. Ils peuvent être joints par une ligne brisée $L = (a = x_0; \ldots; x_p = x)$ et il suffit de montrer que $f(x_i) = f(x_{i+1})$ pour $i = 0, \ldots, p-1$. On est ainsi ramené au cas où le segment $[a, x]$ est contenu dans U. L'application $\gamma : \mathbb{R} \to \mathbb{R}^n$, $t \mapsto a + t(x - a)$ est dérivable et $\gamma([0,1]) = [a, x]$. Comme U est ouvert, $\gamma^{-1}(U)$ est un intervalle ouvert I contenant $[0, 1]$. D'après le point (iii) du théorème précédent, $f \circ \gamma : I \to \mathbb{R}^p$ est dérivable, de dérivée nulle, donc constante. Il en résulte $f(x) = f(a)$, donc f est constante. □

Bien entendu, il est nécessaire de supposer U connexe. Sinon prendre $U = \mathbb{R}^*$ et $f(x) = 1$ si $x > 0$, $f(x) = -1$ si $x < 0$.

Notation 5.24. — Soit $\phi : \mathbb{R}^n \to \mathbb{R}^p$ une application linéaire. On note

$$\|\phi\| = \sup_{x \neq 0} \frac{\|\phi(x)\|}{\|x\|}.$$

la constante de Lipschitz de ϕ. On dit que c'est la norme d'opérateur (ou norme matricielle) *subordonnée* aux normes choisies sur $\mathbb{R}^n$ et $\mathbb{R}^p$.

***Théorème 5.25* (Inégalité des accroissements finis).** — *Soit $U \subset \mathbb{R}^n$ un ouvert* <u>convexe</u> *et $f : U \to \mathbb{R}^p$ différentiable. Supposons que $M = \sup_{z \in [a,b]} \|df(z)\|$ soit $< +\infty$. Alors, pour tout $a, b \in U$, on a*

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

Démonstration. — La démonstration pour une norme quelconque est assez technique. Contentons-nous de le prouver pour la norme $\|\cdot\|_\infty$ ou $\|\cdot\|_2$ sur $\mathbb{R}^p$. Posons $f = (f_1, \ldots, f_p)$.

Le cas de la norme $\|\cdot\|_\infty$ est facile : il suffit de montrer que

$$|f_i(b) - f_i(a)\| \leq M \, \|b - a\|$$

pour tout i, ce qui nous ramène au cas $p = 1$ et l'on peut omettre l'indice i. On définit les applications $u : [0,1] \to [a,b]$ et $g : [0,1] \to \mathbb{R}$ par $u(t) = a + t(b-a)$ et $g(t) = f(u(t))$. Alors g est continue sur $[0,1]$ et dérivable sur $]0,1[$, de dérivée en t

$$g'(t) = df(u(t))(u'(t)) = t \, df(u(t))(b - a).$$

D'après le théorème des accroissements finis en dimension 1, il existe $t_0 \in \,]0,1[$ tel que

$$f(b) - f(a) = g(1) - g(0) = g'(t_0)(1 - 0) = g'(t_0)$$

et d'après la définition de M on a

$$|g'(t_0)| = t_0 \, \|df(u(t_0))(b-a)\| \leq t_0 \, M \, \|b - a\| \leq M \, \|b - a\|$$

d'où le résultat.

Considérons maintenant la norme euclidienne $\|\cdot\|_2$. Si $f(b) = f(a)$ il n'y a rien à montrer, donc on peut supposer $f(b) \neq f(a)$. Posons $v = f(b) - f(a)$ et définissons les applications $u : [0,1] \to [a,b]$ et $g : [0,1] \to \mathbb{R}$ par $u(t) = a + t(b-a)$ et

$$g(t) = v \cdot f(u(t)) = \sum_{i=1}^{p} v_i f_i(u(t)),$$

i.e. g est la composée de $f \circ u$ et de l'application linéaire $x \mapsto v \cdot x$. Alors g est continue sur $[0,1]$ et dérivable sur $]0,1[$, de dérivée en t

$$g'(t) = v \cdot (f \circ u)'(t) = v \cdot df(u(t))(u'(t)) = v \cdot t \, df(u(t))(b - a).$$

D'après le théorème des accroissements finis en dimension 1, il existe $t_0 \in \,]0,1[$ tel que

$$v \cdot f(b) - v \cdot f(a) = g(1) - g(0) = g'(t_0)(1 - 0) = g'(t_0),$$

soit $\|v\|^2 = v \cdot t_0 \, df(u(t_0))(b-a)$. Par conséquent, d'après l'inégalité de Cauchy-Schwarz et la définition de M, on a :

$$\|v\|^2 \leq \|v\| \, \|t_0 \, df(u(t_0))(b-a)\| \leq \|v\| \, t_0 \, M \, \|b - a\|$$

et en divisant par $\|v\| > 0$ on obtient

$$\|f(b) - f(a)\| \leq t_0 \, M \, \|b - a\| \leq M \, \|b - a\|.$$

□

Terminons cette section avec la définition du *gradient* de f.

***Définition 5.26* (gradient).** — Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une application différentiable. On a vu que pour tout $a \in U$, $df(a)$ est une *forme linéaire* sur $\mathbb{R}^n$. Comme il est psychologiquement plus facile (et plus visuel) de travailler avec des vecteurs que des formes linéaires, on fait ce qui suit.

(1) On munit $\mathbb{R}^n$ du produit scalaire standard $x \cdot y = \sum_i x_i y_i$ et l'on munit désormais, sauf mention du contraire, $\mathbb{R}^n$ de la norme euclidienne $\|\cdot\|_2$, qu'on notera simplement $\|\cdot\|$.

(2) Le produit scalaire induit un isomorphisme θ entre $\mathbb{R}^n$ et son dual, donné par $\theta(x)(h) = x \cdot h$ pour tout $x, h \in \mathbb{R}^n$. Par conséquent, pour toute forme linéaire $\phi : \mathbb{R}^n \to \mathbb{R}$, il existe un unique vecteur $u = u_\phi$ tel que $\phi(h) = u \cdot h$ pour tout $h \in \mathbb{R}^n$.

(Q) (3) Le vecteur colonne de $\mathbb{R}^n$ correspondant à la forme linéaire $df(a)$ est noté $\nabla f(a)$ et appelé le **gradient** de f en a ; on a donc :

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_n f(a) \end{pmatrix} \qquad \text{et} \qquad df(a)(h) = \nabla f(a) \cdot h = \sum_{i=1}^{n} \partial_i f(a)\, h_i$$

pour tout $h = (h_1, \dots, h_n)$.

6. Applications de classe C^1

Signalons qu'une fonction dont toutes les dérivées partielles sont partout définies n'est pas forcément différentiable :

Exemple 6.1. — Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie, pour tout $(x, y) \in \mathbb{R}^2$, par

$$f(x,y) = \begin{cases} \dfrac{y^2}{x} & \text{si } x \neq 0, \\ 0 & \text{si } x = 0. \end{cases}$$

Alors, pour tout vecteur $v_\mu = (1, \mu)$ avec $\mu \in \mathbb{R}$, et tout $t \neq 0$, on a

$$\frac{f(0 + tv_\mu) - f(0)}{t} = \frac{t^2\mu^2}{t^2} = \mu^2,$$

donc f est dérivable en 0 dans la direction v_μ. De plus, pour $v = (0,1) = e_2$ et $t \neq 0$, on a $f(0 + te_2) = 0 = f(0)$ donc $(\partial f / \partial x_2)(0) = 0$. Donc f admet en 0 des dérivées dans toutes les directions. Mais f n'est pas différentiable en 0 car elle n'y est pas continue. En effet, la suite $u^k = (1/k^2, 1/k)$ pour $k \in \mathbb{N}^*$ converge vers 0 mais $f(u^k) = 1$ pour tout k.

La situation est meilleure lorsque les dérivées partielles sont continues.

***Définition 6.2* (Fonctions de classe C^1).** — Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ une application. On dit que f est de classe C^1 sur U si f admet des dérivées partielles $\partial_j f$ pour
(Q) $j = 1, \dots, n$ et si celles-ci sont **continues** sur U.

De même, pour $f = (f_1, \dots, f_p) : U \to \mathbb{R}^p$ on dit que f est de classe C^1 si chaque f_i l'est : ceci équivaut à dire que l'application

$$\Phi : U \to M_{p,n}(\mathbb{R}), \qquad a \mapsto \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial f_p}{\partial x_1}(a) & \cdots & \dfrac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

est **continue**.

On notera $\mathscr{C}^1(U,\mathbb{R}^p)$ l'ensemble des fonctions de classe C^1 sur U à valeurs dans $\mathbb{R}^p$.

(Q) ***Théorème 6.3.*** — *Soient $U \subset \mathbb{R}^n$ un ouvert et $f \in \mathscr{C}^1(U,\mathbb{R}^p)$. Alors f est différentiable sur U et l'application*

$$Df : U \to M_{p,n}(\mathbb{R}), \qquad a \mapsto Df(a)$$

est une application continue.

Démonstration. — D'abord, en considérant les composantes $f_1, \dots, f_p$ de f, il suffit de traiter le le cas $p = 1$. Faisons alors la démonstration pour $p = 1$ et $n = 3$, ce qui est suffisant pour bien comprendre l'idée. On utilise la norme $\|\cdot\|_\infty$. Soit $a \in U$. Quitte à faire le changement de variable $x' = a + x$, on peut supposer $a = 0$, ce qui va permettre d'alléger l'écriture. Fixons $\varepsilon > 0$.

Soit $r > 0$ tel que $B(a,r) \subset U$. Pour $j = 1,2,3$, comme $\partial_j f$ est continue en $a = 0$, il existe $\delta_j \in]0,r[$ tel que $|\partial_j(x) - \partial_j(0)| < \varepsilon/3$ si $\|x\| < \delta_j$. Posons $\delta = \min_{i=1,2,3} \delta_i$, alors pour tout $x \in B(0,\delta)$ on a

$$(*) \qquad |\partial_j(x) - \partial_j(0)| < \frac{\varepsilon}{3}.$$

Soit $h = (h_1,h_2,h_3) \in B(0,\delta)$, alors $(h_1,h_2,0)$ et $(h_1,0,0)$ sont aussi dans $B(0,\delta)$. Comme $\partial_3 f(h_1,h_2,0)$ existe, on peut écrire

$$(3) \qquad f(h_1,h_2,h_3) = f(h_1,h_2,0) + \partial_3 f(h_1,h_2,0)h_3 + |h_3|\eta_3(h),$$

où η_3 est continue et nulle en 0. Puis, comme $\partial_3 f(h_1,h_2,0)$ existe, on peut écrire

$$(2) \qquad f(h_1,h_2,0) = f(h_1,0,0) + \partial_2 f(h_1,0,0)h_2 + |h_2|\eta_2(h),$$

où η_2 est continue et nulle en 0. Puis, comme $\partial_1 f(h_1,0,0)$ existe, on peut écrire

$$(1) \qquad f(h_1,0,0) = f(0) + \partial_1 f(0)h_1 + |h_1|\eta_1(h),$$

où η_1 est continue et nulle en 0. De plus, d'après $(*)$ on a

$$|\partial_3 f(h_1,h_2,0)h_3 + \partial_2 f(h_1,0,0) + \partial_1 f(0,0,0)h_1 - \sum_{j=1}^{3} \partial_j(0)h_j| \le \frac{\varepsilon}{3} \sum_{j=1}^{3} |h_j| \le \varepsilon\|h\|.$$

En posant $\eta(h) = \sum_{j=1}^3 |\eta_j(h)|$, on en déduit :

$$(\dagger) \qquad |f(h_1,h_2,h_3) - f(0) - \sum_{j=1}^{3} \partial_j(0)h_j| \le \|h\|\,(\varepsilon + \eta(h)).$$

Comme chaque η_j est continue et nulle en 0, ainsi que l'application $\mathbb{R} \to \mathbb{R}_+$, $x \mapsto |x|$, il en est de même de η et donc il existe $\delta_0 > 0$ tel que $0 \le \eta(h) < \varepsilon$ si $\|h\| < \delta_0$. Posons $\delta_1 = \min(\delta,\delta_0)$. Alors pour tout h tel que $\|h\| < \delta_1$, on a :

$$|f(h_1,h_2,h_3) - f(0) - \sum_{j=1}^{3} \partial_j(0)h_j| \le \|h\|\,2\varepsilon.$$

Ceci prouve que f est différentiable en $a = 0$, de différentielle la forme linéaire

$$(h_1,h_2,h_3) \mapsto \sum_{j=1}^{3} \frac{\partial f}{\partial x_j}(a)h_j.$$

De plus, l'application $df : U \to \mathscr{L}(\mathbb{R}^3, \mathbb{R}) = M_{1,3}(\mathbb{R})$ (matrices à une ligne et 3 colonnes) est donnée par

$$a \mapsto \big(\partial_1 f(a), \partial_2 f(a), \partial_3 f(a)\big)$$

donc est continue. De même, pour p et n arbitraires, l'application $Df : U \to M_{p,n}(\mathbb{R})$ associe à a la matrice $\left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$ dont la composante d'indice (i,j) est $\dfrac{\partial f_i}{\partial x_j}$ qui est continue sur U, donc Df est bien continue. Ceci achève la preuve du théorème. □

***Remarque 6.4* (Attention !).** — Il faut se garder de croire que l'application $U \to M_{p,n}(\mathbb{R})$, $a \mapsto Df(a)$ est linéaire : en effet, chaque coefficient d'indice (i,j) de la matrice $Df(a)$ est donné par $\dfrac{\partial f_i}{\partial x_j}(a)$, qui en général n'est pas une fonction linéaire de a et peut être arbitrairement compliqué. Par exemple, pour $U = \mathbb{R}$ et $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^n$, c'est l'application $U \to M_{1,1}(\mathbb{R})$ qui à tout $a \in U$ associe la matrice (na^{n-1}).

Le théorème 6.3 est fort utile en pratique. En effet, en dehors de certains cas particuliers, il est généralement difficile de montrer directement qu'une fonction f est différentiable sur un ouvert U. Une possibilité consiste donc à montrer que f admet des dérivées partielles sur U et que celles-ci continues sont sur U. Le théorème précédent assure alors la différentiabilité de f sur U.

Exemple 6.5. — Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$f(x,y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

La fonction f est bien définie, continue et admet des dérivées partielles sur $\mathbb{R}^2 \setminus \{(0,0)\}$. De plus, pour tout $(x,y) \neq (0,0)$, on a

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

donc $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues $\mathbb{R}^2 - \{(0,0)\}$. Par ailleurs, le calcul des dérivées partielles en $(0,0)$ donne

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

Comme $|x^4 - y^4 \pm 4x^2y^2| \leq x^4 + y^4 + 4x^2y^2 \leq 2(x^2+y^2)^2$, on a pour $(x,y) \neq (0,0)$:

$$\left|\frac{\partial f}{\partial x}(x,y)\right| \leq 2|y| \leq 2\|(x,y)\|, \qquad \left|\frac{\partial f}{\partial y}(x,y)\right| \leq 2|x| \leq 2\|(x,y)\|$$

donc $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues en $(0,0)$. Ceci assure que f est différentiable sur $\mathbb{R}^2$.

Signalons au passage la proposition suivante. Comme les dérivées partielles s'obtiennent en dérivant par rapport à une seule variable, elles jouissent des mêmes propriétés que celles connues pour les fonctions d'une seule variable, la démonstration est omise.

Proposition 6.6. — *Soient $f, g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions admettant en a une dérivée partielle par rapport à la variable x_i.*

(i) *Pour tout $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ admet en a une dérivée partielle par rapport à la variable x_i et*

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

(ii) *fg admet une dérivée partielle en a par rapport à la variable x_i et*

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

(iii) *Si $f(a) \neq 0$ alors $1/f$ admet une dérivée partielle en a par rapport à la variable x_i et*

$$\frac{\partial(1/f)}{\partial x_i}(a) = \frac{-1}{f(a)^2}\frac{\partial f}{\partial x_i}(a).$$

Terminons cette section avec un exemple important d'utilisation des différentielles, la recherche d'extrema. (Ceci aurait pu figurer dans la section précédente.)

***Définitions 6.7* (Extrema locaux).** — Soient U un ouvert de $\mathbb{R}^n$, f une application $U \to \mathbb{R}$ et $a \in U$.

(i) On dit que f admet en a un minimum (resp. maximum) *local* s'il existe $r > 0$ tel que $B(a, r) \subset U$ et $f(a) \leq f(x)$ (resp. $f(a) \geq f(x)$) pour tout $x \in B(a, r)$.

(ii) On dit que f admet en a un minimum (resp. maximum) global si $f(a) \leq f(x)$ (resp. $f(a) \geq f(x)$) pour tout $x \in U$.

(iii) On utilisera le mot « extremum »[(2)] (local ou global) pour désigner sans distinction un maximum ou un minimum (local ou global).

(iv) Il est clair qu'un extremum global est *a fortiori* un extremum local. Mais il se peut que f admette des extrema locaux mais aucun extremum global. Par exemple, si $P(x)$ est un polynôme de degré 3 ayant trois racines réelles, par exemple $P(x) = x(x^2 - 1)$, alors P admet un maximum local et un mininum local (en les points où le polynôme dérivé $P'(x)$ s'annule) mais aucun extremum global.

Définition 6.8. — Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une application différentiable. On dit que $a \in U$ est un **point critique** de f si $df(a) = 0$.

(Q) ***Proposition 6.9.*** — *Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une application différentiable. Si f admet un extremum local en $a \in U$, alors a est un point critique de f (i.e. $df(a) = 0$).*

Démonstration. — Supposons par exemple que f ait un minimal local en a. Il existe $r > 0$ tel que $f(a + h) \geq f(a)$ pour tout $h \in \mathbb{R}^n$ vérifiant $\|h\| < 2r$. Fixons un tel h, posons $I =]-1, 1[$ et considérons la fonction $g : I \to \mathbb{R}$ définie par $g(t) = f(a + th)$. Alors g est dérivable sur I et

$$g'(t) = df(a + th)(h)$$

pour tout $t \in I$; en particulier $g'(0) = df(a)(h)$. D'autre part, comme g a un minimum local en 0, on a $g'(0) = 0$. Rappelons la démonstration : pour tout $t \in I$ on a $g(t) - g(0) \geq 0$ donc

$$\lim_{t \to 0^-} \frac{g(t) - g(0)}{t} \leq 0 \leq \lim_{t \to 0^+} \frac{g(t) - g(0)}{t}$$

et donc $g'(0)$, qui est la valeur commune des deux limites, est nul. On a donc $df(a)(h) = 0$ pour tout h vérifiant $\|h\| < 2r$. Comme on l'a déjà vu, ceci entraîne que $L = df(a)$ est nulle : en effet, pour $h \neq 0$ arbitraire, posons $\lambda = \|h\|/r$ et $h' = (1/\lambda)h$, alors $\|h'\| = r < 2r$ donc $L(h') = 0$, et comme $h = \lambda h'$ on a $L(h) = \lambda L(h') = 0$. □

Être un point critique est donc une condition nécessaire pour être un extremum. Elle n'est cependant pas suffisante au regard des exemples suivants.

---

(2) pluriel *extrema*

Exemples 6.10. — (1) Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par $f(x, y) = x^2 - y^2$ pour tout $(x, y) \in \mathbb{R}^2$. Elle est différentiable puisque c'est un polynôme. Pour tout $(x, y) \in \mathbb{R}^2$, $df(x, y)$ est la forme linéaire donnée par la matrice ligne :

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

qui s'annule en $(0, 0)$ (et en ce point uniquement). Donc $(0, 0)$ est un point critique de f. Mais, $f(0, 0) = 0$ et pour tout $\varepsilon \neq 0$ on a

$$f(\varepsilon, 0) = \varepsilon^2 > f(0, 0) > -\varepsilon^2 = f(0, \varepsilon)$$

donc f n'a pas en a de maximum ou minimum local.

(2) Un autre exemple est donné par $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$. La dérivée $f'(x) = 3x^2$ s'annule en 0 mais f n'a pas en 0 un extremum local car $f(x) > f(0) > f(-x)$ pour tout $x > 0$.

Pour avoir plus d'information afin de décider si un point critique a de f est un extremum local, on va supposer dans la section suivante que f admet des dérivées partielles d'ordre 2 et considérer la matrice hessienne de f en a.

7. Dérivées partielles d'ordre deux, matrice hessienne, application aux extrema

Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ différentiable et a un point critique de f (i.e. $df(a) = 0$). Une condition *suffisante* pour que f ait en a un extremum local sera donnée en termes de la matrice des dérivées partielles d'ordre 2 de f en a, que nous allons maintenant introduire. (La notion de dérivées partielles d'ordre k existe aussi pour $k \geq 3$, mais nous nous limiterons ici au cas $k = 2$.)

***Définition 7.1* (matrice hessienne).** — Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ de classe C^1.

(i) Soit $a \in U$. On dit que f admet en a des dérivées partielles d'ordre 2 si les fonctions $\dfrac{\partial f}{\partial x_1}, \dots, \dfrac{\partial f}{\partial x_n}$ admettent en a des dérivées partielles. On note alors pour tout $i, j = 1, \dots, n$,

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

la i-ème dérivée partielle de la j-ème dérivée partielle de f en a.

(Q) (ii) La matrice formée de ces dérivées partielles d'ordre 2 en a est notée

$$D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{1 \leq i,j \leq n}$$

et appelée **matrice hessienne** de f au point a.

(iii) On dit que f est de classe C^2 sur U si les dérivées partielles $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ sont définies et continues en tout point de U. Dans ce cas, chaque application $\dfrac{\partial f}{\partial x_i}$ est de classe C^1 sur U, donc *différentiable* sur U, d'après le théorème 6.3.

Attention ! On peut montrer que la fonction f de l'exemple 6.5 vérifie $\dfrac{\partial^2 f}{\partial x \partial y} = 1 \neq \dfrac{\partial^2 f}{\partial y \partial x} = -1$. Mais un résultat fondamental est le théorème de Schwarz ci-dessous qui affirme que ceci ne peut se produire si les dérivées partielles secondes sont continues.

(Q) ***Théorème 7.2* (de Schwarz).** — *Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ une application de classe C^1 qui admet des dérivées partielles secondes sur U qui sont continues en un point a de U. Alors pour tout $i, j = 1, \dots, n$, on a :*

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

Par conséquent, la matrice hessienne $D^2 f(a)$ est **symétrique**.

Démonstration. — Fixons deux indices $i \neq j$ et posons

$$\Delta(s,t) = f(a + se_i + te_j) - f(a + se_i) - f(a + te_j) + f(a).$$

On va montrer que $\lim_{\substack{t\to 0\\ t\neq 0}} \dfrac{\Delta(t,t)}{t^2} = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$. Comme i, j jouent des rôles symétriques dans la définition de $\Delta(t,t)$, cette limite sera aussi égale à $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$, d'où l'égalité voulue.

Comme U est ouvert, il existe un intervalle ouvert I contenant 0 tel que Δ soit définie sur le carré ouvert $I \times I$. Fixons $t \neq 0$ dans I. D'abord, pour $s \in I$, posant $u(s) = a + se_i$, $v(s) = u(s) + te_j$ et $g(s) = f(v(s)) - f(u(s))$, on a

$$\Delta(t,t) = g(t) - g(0). \tag{1}$$

Comme f est différentiable et $u'(s) = e_i = v'(s)$, alors g est dérivable sur I, de dérivée

$$g'(s) = df(v(s))(v'(s)) - d(u(s))(u'(s)) = \partial_i f(u(s) + te_j) - \partial_i f(u(s)).$$

D'après le théorème des accroissements finis, il existe $\theta_1 \in [0,1]$ tel que

$$\Delta(t,t) = g(t) - g(0) = tg'(\theta_1 t) = t\big(\partial_i f(a + \theta_1 te_i + te_j) - \partial_i f(a + \theta_1 te_i)\big). \tag{2}$$

Posons maintenant $w(s) = a + \theta_1 te_i + se_j$ et $\psi(s) = \partial_i f(w(s))$. Alors ψ est dérivable sur I de dérivée

$$\psi'(s) = \partial_j \partial_i f(a + \theta_1 te_i + se_j)$$

et donc, d'après le théorème des accroissements finis il existe $\theta_2 \in [0,1]$ tel que $\psi(t) - \psi(0) = t\psi'(\theta_2)$. Comme $\Delta(t,t) = t(\psi(t) - \psi(0))$, on obtient :

$$\frac{\Delta(t,t)}{t^2} = \partial_j \partial_i f(a + \theta_1 te_i + \theta_2 te_j). \tag{3}$$

Soit maintenant $\varepsilon > 0$. Comme $\partial_j \partial_i f$ est continue en a, il existe $\delta > 0$ tel que, pour tout $h \in \mathbb{R}^n$ vérifiant $\|h\| < \delta$, on ait

$$|\partial_j \partial_i f(a + h) - \partial_j \partial_i f(a)| < \varepsilon.$$

Si $|t| < \delta$ alors, comme $\theta_1, \theta_2 \in [0,1]$, cette condition est vérifiée par le vecteur $h = \theta_1 te_i + \theta_2 te_j$, et l'on a donc

$$\left| \frac{\Delta(t,t)}{t^2} - \partial_j \partial_i f(a) \right| < \varepsilon$$

pour tout $t \neq 0$ vérifiant $|t| < \delta$. Ceci prouve le résultat désiré. □

Pour une fonction de classe C^2, on a la formule de Taylor à l'ordre 2 suivante :

(Q) ***Théorème 7.3* (Formule de Taylor à l'ordre 2).** — *Soit $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}$ de classe C^2 sur U. Pour tout $a \in U$, on a pour $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$ assez petit,*

$$\begin{aligned} f(a+h) &= f(a) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) h_j + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o(\|h\|^2) \\ &= f(a) + df(a)(h) + \frac{1}{2} Q(h) + o(\|h\|^2), \end{aligned}$$

où $Q(h) = h \cdot D^2 f(a) h$ est la forme quadratique associée à la matrice symétrique $D^2 f(a)$.

Démonstration. — On donnera une démonstration complète dans une section en appendice à la fin de ce chapitre. Contentons-nous de démontrer ici un résultat un peu plus faible, à savoir le cas où h varie dans une direction fixée. Soit $a \in U$. Comme U est ouvert, il contient une boule ouverte $B(a,r)$. Fixons $h \in B(0,r)$ et posons $I =]-1,1[$. Alors la fonction $g : I \to \mathbb{R}$ définie par $g(t) = f(a+th)$ est dérivable, de dérivée

$$g'(t) = df(a+th)(h) = \sum_{j=1}^{n} \partial_j f(a+th) h_j.$$

Comme f est de classe C^2, chaque $\partial_j f$ est de classe C^1 donc différentiable, donc $g'(t)$ est dérivable (et même de classe C^1) de dérivée :

$$g''(t) = \sum_{j=1}^{n} h_j d(\partial_j f)(a+th) = \sum_{j=1}^{n} h_j \sum_{i=1}^{n} \partial_i \partial_j (a+th) h_i = \sum_{i,j=1}^{n} \partial_i \partial_j (a+th) h_i h_j.$$

Comme $g : I \to \mathbb{R}$ est deux fois dérivable (et même de classe C^2), la formule de Taylor donne

$$g(t) = g(0) + t g'(0) + \frac{t^2}{2} g''(0) + o(t^2)$$

au voisinage de $t = 0$, d'où le résultat annoncé. □

(Q) ***Théorème 7.4.*** — *Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ de classe C^2 et $a \in U$ un point critique de f.*

(i) *Si f admet un minimum local en a, alors les valeurs propres de la matrice hessienne $D^2 f(a)$ sont toutes ≥ 0.*

(ii) *Réciproquement, si ces valeurs propres sont toutes > 0, alors f admet un minimum local en a.*

(iii) *De même, si f admet un maximum local en a, alors les valeurs propres de la matrice hessienne $D^2 f(a)$ sont toutes ≤ 0. Réciproquement, si ces valeurs propres sont toutes < 0, alors f f admet un maximum local en a.*

Démonstration. — Dans cette démonstration, on munit $\mathbb{R}^n$ de la norme euclidienne $\|\cdot\|_2$. D'après le théorème de Schwarz, la matrice hessienne $S = D^2 f(a) \in M_n(\mathbb{R})$ est **symétrique**. D'après un théorème d'algèbre linéaire, on sait donc qu'elle est *diagonalisable dans une base orthonormée* ; il existe donc des réels $\lambda_1 \leq \cdots \leq \lambda_n$ et une base orthonormée $(v_1, \ldots, v_n)$ de $\mathbb{R}^n$ tels que $S v_i = \lambda_i v_i$ pour $i = 1, \ldots, n$.

Comme $df(a) = 0$, la formule de Taylor donne pour tout t assez petit :

$$f(a + tv_i) = f(a) + \frac{t^2}{2} v_i \cdot Sv_i + t^2 \eta(t) = f(a) + \frac{t^2}{2}(\lambda_i + 2\eta(t))$$

avec $\lim_{t\to 0} \eta(t) = 0$.

Si f admet un minimum local en a, on a donc $\lambda_i \geq -2\eta(t)$ pour tout t assez petit, d'où par passage à la limite $\lambda_i \geq 0$. Ceci prouve (i).

Supposons maintenant $\lambda_1 > 0$. Pour $h = \sum_{i=1}^{n} h_i v_i$ assez petit, la formule de Taylor donne

$$f(a + h) = f(a) + \frac{1}{2} h \cdot Sh + \|h\|^2 \eta(h) = f(a) + \frac{1}{2} \sum_{i=1}^{n} \lambda_i h_i^2 + \|h\|^2 \eta(h)$$

avec $\lim_{t\to 0} \eta(t) = 0$, et comme par hypothèse $\lambda_n \geq \cdots \geq \lambda_1 > 0$, on obtient

$$f(a + h) \geq f(a) + \|h\|^2 \Big(\frac{\lambda_1}{2} + \eta(h) \Big)$$

d'où $f(a + h) \geq f(a)$ pour tout h tel que $|\eta(h)| \leq \lambda_1/2$. Ceci prouve que f admet un minimum local en a, d'où (ii). Et bien entendu, le cas (iii) se traite de façon analogue (ou en changeant f en $-f$). □

(Q) ***Remarque 7.5***. — En dimension $n = 2$, il est facile de connaître le signe des valeurs propres λ_1 et λ_2 de $S = D^2 f(a)$. En effet, on a $\text{dét}(S) = \lambda_1 \lambda_2$ et $\text{Tr}(S) = \lambda_1 + \lambda_2$. Si $\text{dét}(S) = 0$ alors l'une des valeurs propres est nulle, et si $\text{dét}(S) < 0$ elles sont de signe opposé ; dans ces deux cas on ne peut pas conclure. Par contre, si $\text{dét}(S) > 0$ alors les deux valeurs propres sont non nulles et de même signe, ce qui implique que a est un extremum local ; plus précisément :

(i) Si $\text{dét}(S) > 0$ et $\text{Tr}(S) > 0$ alors $\lambda_1, \lambda_2 > 0$ et a est un minimum local.
(ii) Si $\text{dét}(S) > 0$ et $\text{Tr}(S) < 0$ alors $\lambda_1, \lambda_2 < 0$ et a est un maximum local.

8. Équations aux dérivées partielles

De nombreux phénomènes mécaniques, physiques, biologiques ou économiques se modélisent à l'aide d'équations dont l'inconnue est une fonction de plusieurs variables (on peut penser à la variable temporelle t et à la variable d'espace x), et qui font intervenir les dérivées partielles de la fonction. De telles équations s'appellent *équations aux dérivées partielles*. Dans ce chapitre, nous présentons quelques exemples d'équations aux dérivées partielles ainsi que des méthodes de résolution dans des cas très particuliers.

8.1. Équation de transport. — Soit $a \in \mathbb{R}^n - \{0\}$ et $u_0 : \mathbb{R}^n \to \mathbb{R}$ une fonction de classe C^1 donnée. L'**équation de transport** consiste à chercher une fonction $u : \mathbb{R}^n \times [0, +\infty[\to \mathbb{R}$ telle que

$$\begin{cases} \dfrac{\partial u}{\partial t}(x,t) + a \cdot \nabla_x u(x,t) = 0, & \text{pour tout } (x,t) \in \mathbb{R}^n \times]0, +\infty[, \\ u(x,0) = u_0(x), & \text{pour tout } x \in \mathbb{R}^n. \end{cases}$$

Pour résoudre cette équation, on utilise la méthode des (hypersurfaces) *caractéristiques* que nous présentons pour $n = 1$. Dans ce cas, $a \in \mathbb{R}^*$ et l'équation de transport devient

$$\frac{\partial u}{\partial t}(x,t) + a \frac{\partial u}{\partial x}(x,t) = 0, \quad \text{pour tout } (x,t) \in \mathbb{R} \times]0, +\infty[.$$

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

$$\Phi(x,t) = \begin{pmatrix} x - at \\ x + at \end{pmatrix}.$$

On vérifie aisément que Φ est une application linéaire dont le déterminant de la matrice est donné par $2a \neq 0$. Par conséquent, Φ est inversible et

$$\begin{pmatrix} y \\ z \end{pmatrix} = \Phi(x,t) \quad \Longleftrightarrow \quad \Phi^{-1}(y,z) = \begin{pmatrix} \frac{z+y}{2} \\ \frac{z-y}{2a} \end{pmatrix}.$$

Par conséquent, Φ réalise un $\mathscr{C}^1$-difféomorphisme de $\mathbb{R}^2$ dans $\mathbb{R}^2$. Posons $v(y,z) = u \circ \Phi^{-1}(y,z)$ pour tout $(y,z) \in \mathbb{R}^2$. La formule de dérivation des fonctions composées montre que

$$\frac{\partial v}{\partial z}(y,z) = \nabla u(\Phi^{-1}(y,z)) \cdot \frac{\partial \Phi^{-1}}{\partial z}(y,z) = \nabla u(\Phi^{-1}(y,z)) \cdot \begin{pmatrix} 1/2 \\ 1/2a \end{pmatrix} = \frac{1}{2a}\left(a\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t}\right)(\Phi^{-1}(y,z)).$$

Ainsi, u est solution de l'équation de transport si et seulement $\frac{\partial v}{\partial z}(y,z) = 0$. Autrement dit, si et seulement s'il existe une fonction $f : \mathbb{R} \to \mathbb{R}$ telle que $v(y,z) = f(y)$ pour tout $(y,z) \in \mathbb{R}^2$. Par conséquent, les solutions de l'équation de transport sont les fonctions u de la forme $u(x,t) = v(\Phi(x,t)) = f(x - at)$ pour tout $(x,t) \in \mathbb{R} \times [0,+\infty[$. Pour déterminer la fonction f, on utilise la condition initiale qui assure que pour tout $x \in \mathbb{R}$, on a $u(x,0) = u_0(x)$ et donc $f(x) = u_0(x)$. Ceci montre que la solution de l'équation de transport est donnée par

$$u(x,t) = u_0(x - at), \qquad \text{pour tout } (x,t) \in \mathbb{R} \times [0,+\infty[.$$

8.2. Équation des ondes. — L'**équation des ondes** régit la propagation d'une onde électromagnétique. Étant donné un réel $a \neq 0$ et des fonctions $u_0, u_1 : \mathbb{R}^n \to \mathbb{R}$ de classe C^2 et C^1 respectivement, elle s'écrit

$$\begin{cases} \frac{\partial^2 u}{\partial t^2}(x,t) - a^2 \Delta_x u(x,t) = 0, & \text{pour tout } (x,t) \in \mathbb{R}^n \times]0,+\infty[, \\ u(x,0) = u_0(x), & \text{pour tout } x \in \mathbb{R}^n, \\ \frac{\partial u}{\partial t}(x,0) = u_1(x), & \text{pour tout } x \in \mathbb{R}^n. \end{cases}$$

Tout comme l'équation de transport, elle peut se résoudre grâce à la méthode des caractéristiques. De nouveau, en dimension $n = 1$, l'équation des ondes devient

$$\frac{\partial^2 u}{\partial t^2}(x,t) - a^2 \frac{\partial^2 u}{\partial x^2}(x,t) = 0, \qquad \text{pour tout } (x,t) \in \mathbb{R} \times]0,+\infty[.$$

Avec les mêmes notations qu'à la sous-section précédente, on calcule

$$\frac{\partial^2 v}{\partial y \partial z}(y,z) = \nabla\left(\frac{1}{2}\frac{\partial u}{\partial x} + \frac{1}{2a}\frac{\partial u}{\partial t}\right)(\Phi^{-1}(y,z)) \cdot \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2a} \end{pmatrix} = \frac{1}{4a^2}\left(\frac{\partial^2 u}{\partial t^2}(x,t) - a^2 \frac{\partial^2 u}{\partial x^2}\right)(\Phi^{-1}(y,z)).$$

Ainsi, u est solution de l'équation des ondes si et seulement $\frac{\partial^2 v}{\partial y \partial z}(y,z) = 0$. Il existe donc une fonction $g : \mathbb{R} \to \mathbb{R}$ telle que $\frac{\partial v}{\partial z}(y,z) = g(z)$. En intégrant cette nouvelle équation différentielle, on trouve que $v(y,z) = F(y) + G(z)$, pour des fonctions $F, G : \mathbb{R} \to \mathbb{R}$ et donc

$$u(x,t) = F(x - at) + G(x + at), \quad \text{pour tout } (x,t) \in \mathbb{R} \times [0,+\infty[.$$

De nouveau, on utilise les conditions initiales pour déterminer F et G. En effet, pour tout $x \in \mathbb{R}$,

$$u_0(x) = u(x,0) = F(x) + G(x), \quad u_1(x) = \frac{\partial u}{\partial t}(x,0) = -aF'(x) + aG'(x).$$

On peut résoudre facilement ce système différentiel en dérivant la première équation, en la multipliant par a, puis en l'additionnant et la soustrayant à la seconde équation. Il vient alors que

$$G'(x) = \frac{au_0'(x) + u_1(x)}{2a}, \quad F'(x) = \frac{au_0'(x) - u_1(x)}{2a},$$

soit

$$G(x) = \frac{1}{2}u_0(x) + \frac{1}{2a}\int_0^x u_1(y)\,dy + \beta, \quad F(x) = \frac{1}{2}u_0(x) - \frac{1}{2a}\int_0^x u_1(y)\,dy + \alpha,$$

où α et $\beta \in \mathbb{R}$ sont des constantes. Par conséquent, pour tout $(x,t) \in \mathbb{R} \times [0,+\infty[$

$$u(x,t) = \frac{1}{2}u_0(x+at) + \frac{1}{2a}\int_0^{x+at} u_1(y)\,dy + \frac{1}{2}u_0(x-at) - \frac{1}{2a}\int_0^{x-at} u_1(y)\,dy + c,$$

où $c = \alpha + \beta$. En utilisant de nouveau la condition initiale $u(x,0) = u_0(x)$, on trouve que $c = 0$ et donc que la solution de l'équation des ondes est donnée par la fonction

$$u(x,t) = \frac{1}{2}u_0(x+at) + \frac{1}{2a}\int_0^{x+at} u_1(y)\,dy + \frac{1}{2}u_0(x-at) - \frac{1}{2a}\int_0^{x-at} u_1(y)\,dy$$

pour tout $(x,t) \in \mathbb{R} \times [0,+\infty[$.

TABLE DES MATIÈRES