Notes de cours de l'ISIMA, première année
http://www.isima.fr/~leborgne

Fonctions de plusieurs variables et changements de variables

Gilles Leborgne

1er février 2017

Table des matières

1 Rappels sur les fonctions $\mathbb{R} \to \mathbb{R}$

1.1 Continuité

On considère un intervalle $I \subset \mathbb{R}$, i.e., I a l'une des formes suivantes : $]-\infty,\infty[(=\mathbb{R})$, $]-\infty,b[$, $]-\infty,b]$, $]a,\infty[$, $[a,\infty[$, $]a,b[$, $[a,b[$, $]a,b]$, $[a,b]$, où a et b sont deux réels tels que $a<b$ (dans les quatre derniers cas).

Définition 1.1 Soit $f : I \to \mathbb{R}$ et $x_0 \in I$. Alors f est continue en x_0 si :

$$\lim_{\substack{x\in I\\ x\to x_0}} f(x) = f(x_0) \quad \text{dans } \mathbb{R}. \tag{1.1}$$

Quand on écrit $x \to x_0$, on sous-entend bien sûr que $x \in I$, sinon $f(x)$ n'aurait pas de sens ; on ne précisera plus $x \in I$ dans la suite, i.e. on écrira simplement $\lim_{x\to x_0} f(x) = f(x_0)$.

On rappelle les notations :

$$\lim_{x\to a+} f(x) \stackrel{\text{déf}}{=} \lim_{\substack{x>a\\ x\to b}} f(x) \stackrel{\text{noté}}{=} f(a+),$$

et

$$\lim_{x\to b-} f(x) \stackrel{\text{déf}}{=} \lim_{\substack{x<b\\ x\to b}} f(x) \stackrel{\text{noté}}{=} f(b-).$$

Et la définition donnée par (1.1) contient la définition de la continuité à droite et à gauche : si $I=[a,b]$, f est continue à droite en a ssi :

$$f(a+) = f(a),$$

et f est continue à gauche en b ssi :

$$f(b-) = f(b).$$

La continuité en x_0 s'écrit aussi (avec la valeur absolue $|.|$ dans $\mathbb{R}$) :

$$\lim_{|x-x_0|\to 0} |f(x) - f(x_0)| = 0. \tag{1.2}$$

Et en termes de quantificateurs, f continue en x_0 s'écrit :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in I\ :\ |x - x_0| < \eta \implies |f(x) - f(x_0)| < \varepsilon,$$

qui s'énonce en "français" :

En x_0, pour tout $\varepsilon > 0$ (sous-entendu "aussi petit soit-il"), il existe $\eta > 0$ (qui dépend de ε et de x_0) tel que, dès que x vérifie $|x - x_0| < \eta$, on a $|f(x) - f(x_0)| < \varepsilon$, ou encore :

en x_0, pour tout $\varepsilon > 0$ (sous-entendu "aussi petit soit-il"), il existe $\eta > 0$ (qui dépend de ε et de x_0) tel que, dès que x est dans l'intervalle $]x_0-\eta, x_0+\eta[$, on a $f(x)$ dans l'intervalle $]f(x_0)-\varepsilon, f(x_0)+\varepsilon[$.

Définition 1.2 f est discontinue en x_0 ssi f n'est pas continue en x_0, i.e. ssi $\lim_{x\to x_0} f(x) \neq f(x_0)$. Ou encore si :

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x \in I \quad \text{t.q. :} \quad |x - x_0| < \eta \quad \text{et} \quad |f(x) - f(x_0)| \geq \varepsilon.$$

Définition 1.3 On dit que $f : I \to \mathbb{R}$ est continue dans I si f est continue en tout point x de I.

Cela s'écrit :

$$\forall x \in I,\ \forall \varepsilon > 0,\ \exists \eta > 0 :\ \forall \tilde{x} \in I,\ |\tilde{x} - x| < \eta \implies |f(\tilde{x}) - f(x)| < \varepsilon.$$

Définition 1.4 On appelle $C^0(I;\mathbb{R})$ l'ensemble des fonctions continues $f : I \to \mathbb{R}$, qu'on note simplement C^0 lorsqu'il n'y a pas de confusion possible.

Théorème 1.5 *(des valeurs intermédiaires.) Soit $f : I \to \mathbb{R}$ une fonction continue dans I, soit a et b deux points de I tels que $a < b$. Alors, si $f(a) < f(b)$ on a :*

$$\forall y_0 \in]f(a), f(b)[, \quad \exists x_0 \in]a, b[, \quad f(x_0) = y_0, \tag{1.3}$$

et si $f(a) > f(b)$, même résultat pour tout $y_0 \in]f(b), f(a)[$. I.e., si f est continue, alors toute valeur comprise entre $f(a)$ et $f(b)$ est atteinte.

En particulier, si f est continue sur I, alors $f(I)$ est un intervalle, à savoir l'un des intervalles $[m, M]$, $]m, M]$, $[m, M[$ ou $]m, M[$ où $m = \inf_{x\in I} f(x)$ et $M = \sup_{x\in I} f(x)$.

Preuve. Supposons $f(a) < f(b)$ (sinon on travaille avec la fonction $-f$).

Soit $y_0 \in]f(a), f(b)[$, soit $S = \{x \in]a, b[\ :\ f(x) < y_0\}$. Ayant $S \subset [a, b]$, on définit $c \in [a, b]$ comme étant la borne supérieur de S : si $x \in S$ alors $x \leq c$, et si $x > c$ alors $f(x) \geq y_0$.

On commence par remarquer que c vérifie $a<c<b$. En effet, montrons par exemple que $c \neq a$ (même raisonnement pour $c \neq b$) :

sinon $c = a$, et alors par définition de c, pour tout $x \in]a, b[$, on a $f(x) \geq y_0 > f(a)$. Mais alors $\lim_{x\to a+} f(x) \geq y_0 > f(a)$ et f est discontinue en a, contraire à l'hypothèse f continue en a. Donc $c = a$ est impossible : on a $c > a$.

D'où trois cas possibles :

1- Soit $f(c) = y_0$, et le théorème est démontré.

2- Soit $f(c) < y_0$, mais f étant continue en c, on a $\lim_{x\to c} f(x) = f(c)$ et par définition de c, $\lim_{x\to c+} f(x) \geq y_0 > f(c)$. C'est absurde. D'où $f(c) \geq y_0$.

3- Soit $f(c) > y_0$. Même démarche : c'est impossible.

Seul le cas 1- est possible, d'où (1.3). On en déduit que $f(I)$ est un intervalle : en effet, si y_1, y_2 sont deux point de $f(I)$ (i.e. il existe $x_1, x_2 \in I$ t.q. $f(x_1) = y_1$ et $f(x_2) = y2$, alors tout point de l'intervalle $[y_1, y_2]$ est dans $f(I)$, cf. (1.3). Et $f(I) \supset]m, M[$. ▪

Définition 1.6 On dit que $f : I \to \mathbb{R}$ est uniformément continue dans I si la continuité de f se mesure indépendamment de (uniformément en) x :

$$\forall \varepsilon > 0,\ \exists \eta_\varepsilon > 0 :\ \forall x, \tilde{x} \in I,\ |\tilde{x} - x| < \eta_\varepsilon \implies |f(\tilde{x}) - f(x)| < \varepsilon.$$

(On ne peut pas exprimer l'uniforme continuité en termes de limite.) Donc ici η ne dépend que de ε, pas de x : pour tout ε, il existe η tel que dès que x et $\tilde{x}$ vérifient $|\tilde{x}-x| < \eta_\varepsilon$, on a $|f(\tilde{x})-f(x)| < \varepsilon$.

Proposition 1.7 *Si f est uniformément continue sur I, elle est continue en tout point x de I. La réciproque est fausse, i.e. il existe un intervalle I de $\mathbb{R}$ et un fonction continue $f : I \to \mathbb{R}$ telle que f n'est pas uniformément continue.*

Preuve. On suppose f uniformément continue dans I. Soit $x \in I$, alors l'uniforme continuité donne en particulier : $\forall \varepsilon > 0,\ \exists \eta_\varepsilon > 0 :\ \forall \tilde{x} \in I,\ |\tilde{x} - x| < \eta_\varepsilon \implies |f(\tilde{x}) - f(x)| < \varepsilon$. I.e f est continue en x.

La réciproque est fausse : prendre $f : x \in]0,1] \to \frac{1}{x}$. Cette fonction est bien sûr continue en tout point de $]0,1]$, mais elle n'est pas uniformément continue dans $]0,1]$. En effet, $\exists \varepsilon > 0$, on prend $\varepsilon = 1$, tel que $\forall \eta > 0$, $\exists x > 0$ et $\exists y > 0$, à savoir $x = \min(\eta, 1)$ et $y = \frac{1}{2}x$, et on a à la fois $|x-y| < \eta$ et $|f(x) - f(y)| \geq \varepsilon$ (car $x - y = \min(\frac{\eta}{2}, 1)$ et $|f(x) - f(y)| = |\frac{1}{x} - \frac{2}{x}| = \frac{1}{x} = \max(\frac{1}{\eta}, 1) \geq 1 = \varepsilon$.

(Ou bien prendre $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$.) ∎

Exemple 1.8 Montrer que la fonction $f : x \in [0,1] \to f(x) = \sqrt{x} \in \mathbb{R}$ est uniformément continue (sans se servir de la proposition suivante).

Réponse. Regardons ce qui se passe au voisinage de $x = 0$, i.e. la continuité de $\sqrt{\ }$ en $x = 0$: on veut $|\sqrt{\tilde{x}} - \sqrt{0}| = \sqrt{\tilde{x}} < \varepsilon$ dès que $|\tilde{x} - 0| = \tilde{x} < \eta$. On prend $\eta = \varepsilon^2$, et on a $\tilde{x} < \eta$ donne bien $\sqrt{\tilde{x}} < \varepsilon$ (la fonction $\sqrt{\ }$ est continue à droite en 0 et on a caractérisé η en fonction de ε.

Montrons maintenant que si $|\tilde{x}-x| < \varepsilon^2$ alors $|\sqrt{\tilde{x}}-\sqrt{x}| < \varepsilon$. Pour ce, montrons que $|\sqrt{\tilde{x}}-\sqrt{x}|^2 \leq |\tilde{x}-x|$, i.e. que si $\tilde{x} \geq x$ alors $\tilde{x}+x-2\sqrt{\tilde{x}x} \leq \tilde{x}-x$, i.e. si $\tilde{x} \geq x$ alors $2x - 2\sqrt{\tilde{x}x} \leq 0$. C'est immédiat. Et si $\tilde{x} < x$ on a $|\tilde{x} - x| = x - \tilde{x}$ et de même $2\tilde{x} - 2\sqrt{\tilde{x}x} \leq 0$. ∎

Proposition 1.9 *Une fonction $f \in C^0([a,b],\mathbb{R})$ (continue sur un compact) est uniformément continue.*

Preuve. Supposons f non uniformément continue :

$$\exists \varepsilon > 0,\ \forall \eta > 0, \exists x \in [a,b],\ \exists \tilde{x} \in [a,b],\ |\tilde{x} - x| < \eta \text{ et } |f(\tilde{x}) - f(x)| \geq \varepsilon.$$

Soit donc un tel ε. Alors on fixe $n \in \mathbb{N}$ et on pose $\eta = \frac{1}{n}$. Puis on prend deux réels x_n et $\tilde{x}_n$ tels que $|\tilde{x}_n - x_n| < \eta$ et $|f(\tilde{x}_n) - f(x_n)| \geq \varepsilon$. On construit ainsi deux suites (x_n) et $(\tilde{x}_n)$ dans le compact $[a,b]$: on peut donc en extraire deux sous-suites convergentes qu'on notera encore, pour simplifier les notations, (x_n) et $(\tilde{x}_n)$.

Comme $|\tilde{x}_n - x_n| < \frac{1}{n}$, les deux sous-suites convergent vers une même valeur, et vérifient également $|f(\tilde{x}_n) - f(x_n)| \geq \varepsilon$. Comme f est continue, ceci est impossible. Donc f est bien uniformément continue. ∎

Proposition 1.10 *Si $f : [a,b] \to \mathbb{R}$ est continue, alors l'image $f([a,b])$ est compacte (bornée et fermée), et f atteint son maximum et son minimum :*

$$\exists x_M \in [a,b] : f(x_M) = \sup_{x\in[a,b]} f(x), \quad \exists x_m \in [a,b] : f(x_m) = \inf_{x\in[a,b]} f(x).$$

Preuve. Le théorème des valeurs intermédiaires indique que $f([a,b])$ est un intervalle, et il s'agit de montrer que cet intervalle est borné et fermé quand f est uniformément continue sur $[a,b]$ (avec la proposition 1.9).

Montrons que l'image $f([a,b])$ est bornée. L'intervalle $[a,b]$ étant compact, f y est uniformément continue, et donc à ε fixé, il existe η et donc il existe un nombre fini de points $x_0 < x_1 < ... < x_n$ tels que $x_{i+1} - x_i < \eta$ et $|f(x_{i+1}) - f(x_i)| < \varepsilon$ (à savoir $x_i = a + i\frac{2\eta}{b-a}$ pour $i = 1, ..., n$ où $n =$ partie entière de $\frac{b-a}{2\eta}$ points, quitte à prendre η suffisamment petit, i.e. tel que $\frac{b-a}{2\eta} \geq 1$).

Et donc pour tout $x \in [a,b]$, il existe $i \in [1,n]$ tel que : $|x - x_i| < \eta$ et $|f(x) - f(x_i)| < \varepsilon$. Et donc : $\sup\limits_{x\in[a,b]} f(x) \leq \max\limits_{i=1,...,n} (f(x_i)) + \varepsilon$, pour tout $x \in [a,b]$: et f est bien bornée (i.e. son image est bornée).

Montrons que $f([a,b])$ est fermé. Posons $M = \sup\limits_{x\in[a,b]} (|f(x)|)$, un tel sup existant dans $\mathbb{R}$ car f est bornée. Il s'agit de montrer qu'il existe $x_M \in [a,b]$ tel que $M = f(x_M)$:

par définition du sup, il existe une suite (x_n) dans $[a,b]$ telle que $f(x_n) \underset{n\to\infty}{\longrightarrow} M$. La suite (x_n) étant bornée dans le compact $[a,b]$, elle admet une sous suite convergente x_{n_k} dans $[a,b]$. Notons

$x_M = \lim_{k\to\infty} x_{n_k}$. Sachant f continue, on a $f(x_{n_k}) \to f(x_M)$, et on en déduit que $M = f(x_M)$, et donc que $M \in f([a,b])$. De même pour $m = \inf_{x\in[a,b]}(|f(x)|) \in f([a,b])$, ce qui prouve que $f([a,b])$ est un intervalle fermé. ▪

Remarque 1.11 En appliquant le théorème des accroissements finis (voir plus loin), on montre que si $f : I \to \mathbb{R}$ est continue dérivable de dérivée bornée sur I, alors f est uniformément continue sur I. ▪

1.2 Notations de Landau "o" et "O"

Définition 1.12 Soit $x_0 \in \mathbb{R}$, et soit I intervalle de $\mathbb{R}$ t.q. $x_0 \in I$. Étant données deux fonctions f et g de $\mathcal{F}(I;\mathbb{R})$, on dit que f est "petit o de g au voisinage de x_0", et on note "$f = o(g)$ au vois. de x_0", ssi

$$\forall \tilde\varepsilon > 0, \quad \exists \eta > 0, \quad \forall x \in I, \quad |x - x_0| < \eta \quad \Rightarrow \quad |f(x)| < \tilde\varepsilon |g(x)|. \tag{1.4}$$

On écrit également abusivement $f(x) = o(g(x))$ au vois. de x_0, et on dit que f est négligeable devant g au voisinage de x_0.

Exemple 1.13 Avec $g = 1_{\mathbb{R}}$ (la fonction constante valant 1 sur $\mathbb{R}$), on a "$f = o(1_{\mathbb{R}}) \stackrel{\text{noté}}{=} o(1)$ au vois. de x_0" ssi $f(x) \longrightarrow_{x\to x_0} 0$ (immédiat). ▪

Définition 1.14 Étant données deux fonctions f et g de $\mathcal{F}(\mathbb{R};\mathbb{R})$, on dit que f est "petit o de g au voisinage de $+\infty$", et on note "$f = o(g)$ au vois. de $+\infty$", ssi :

$$\forall \tilde\varepsilon > 0, \quad \exists M > 0, \quad \forall x > M \quad \Rightarrow \quad |f(x)| < \tilde\varepsilon |g(x)|. \tag{1.5}$$

On encore, $f = o(g)$ au voisinage de $+\infty$ ssi $F = o(G)$ au voisinage de 0 où on a posé $F(x) = f(\frac{1}{x})$ et $G(x) = g(\frac{1}{x})$.

Proposition 1.15 *Soit $x_0 \in I$, et soit I intervalle de $\mathbb{R}$ t.q. $x_0 \in I$. Étant données deux fonctions f et g de $\mathcal{F}(I;\mathbb{R})$, s'il existe une fonction $\varepsilon : \mathbb{R} \to \mathbb{R}$ telle que $\varepsilon(0) = 0$ et qui est continue en 0 (donc $\varepsilon(z) \underset{z\to 0}{\longrightarrow} 0$) et :*

$$f(x) = \varepsilon(x - x_0) g(x), \tag{1.6}$$

alors $f = o(g)$ au voisinage de x_0.

Preuve. Si une fonction ε vérifiant (1.6) existe, alors $|f(x)| \le |\varepsilon(x - x_0)||g(x)|$. Et donc, pour $\tilde\varepsilon > 0$ donné, choisissant $\eta > 0$ tel que $|z| < \eta$ implique $|\varepsilon(z)| < \tilde\varepsilon$ (un tel η existe par continuité de ε en 0), on obtient $|f(x)| \le \tilde\varepsilon |g(x)|$ dès que $|x - x_0| < \eta$, d'où (1.4). ▪

En particulier, si g ne s'annule pas dans un voisinage de x_0, on a f est "petit o de g au voisinage de x_0", et on note "$f = o(g)$ au vois. de x_0", ssi dans un voisinage de x_0 :

$$\frac{f(x)}{g(x)} = \varepsilon(x - x_0) \quad \text{i.e.} \quad \lim_{x\to x_0} \frac{f(x)}{g(x)} = 0.$$

On encore, si g ne s'annule pas dans un voisinage de x_0 sauf éventuellement en x_0 :

$$\lim_{\substack{x\to x_0 \\ x\neq x_0}} \frac{f(x)}{g(x)} = 0 \stackrel{\text{noté}}{=} \lim_{x\to x_0} \frac{f(x)}{g(x)},$$

la dernière notation pour alléger l'écriture (sachant que la division par 0 est interdite).

Notation. On note x^n la fonction $x \to x^n$ définie sur $\mathbb{R}$. Ainsi, si $g(x) = x^n$, si $f = o(g)$ au voisinage de x_0, on note $f = o(x^n)$ au voisinage de x_0.

Exemple 1.16 Toute fonction monôme $f(x) = x^n$ avec $n \ge 1$ est $o(1)$ au voisinage de $x_0 = 0$, où on a noté 1 la fonction constante $= 1$ sur $\mathbb{R}$, car $\frac{x^n}{1} \underset{x\to 0}{\longrightarrow} 0$. ▪

Exemple 1.17 La fonction $f(x) = x^n$ pour $n \ge 2$ est $o(x)$ au voisinage de 0 car $\frac{x^n}{x} \underset{x\to 0}{\longrightarrow} 0$: on note $f(x) = o(x)$ au voisinage de 0. ▪

Exemple 1.18 Si f est une fonction continue en x_0, alors $\dfrac{f(x)-f(x_0)}{1} \underset{x\to x_0}{\longrightarrow} 0$ et donc $f(x)-f(x_0) = o(1)$, i.e., $f(x) = f(x_0) + o(1)$ au voisinage de x_0 (développement limité de f à l'ordre 0 au voisinage de x_0).

Remarque 1.19 La notation de Landau n'est pas compatible avec l'addition, i.e. si $f_1 = o(g_1)$ et si $f_2 = o(g_2)$, la conclusion $f_1+f_2 = o(g_1+g_2)$ est fausse. Par exemple prendre $f_1(x) = f_2(x) = x^2$ et prendre $g_1(x) = x = -g_2(x)$: ici on a $g_1+g_2 = 0$ et $(f_1+f_2)(x) = 2x^2 \neq o(0) = 0$.

Proposition 1.20 *La notation de Landau est compatible avec la multiplication, i.e. si $f = o(g)$ au vois. de x_0 alors $fh = o(gh)$ au vois. de x_0 pour toute fonction h.*

Et la notation de Landau est transitive, i.e. si $f = o(g)$ et $g = o(h)$ au vois. de x_0 alors $f = o(h)$ au vois. de x_0.

Preuve. Si $f(x) = \varepsilon(x-x_0)g(x)$ alors $h(x)f(x) = \varepsilon(x-x_0)h(x)g(x)$.

Et si $f(x) = \varepsilon_1(x-x_0)g(x)$ et $g(x) = \varepsilon_2(x-x_0)h(x)$ alors $f(x) = (\varepsilon_1\varepsilon_2)(x-x_0)h(x)$, avec $\varepsilon_1\varepsilon_2$ qui tend vers 0 quand $x \to x_0$.

Définition 1.21 On définit la notion de 'comparable' ou de 'grand O' comme suit : étant données deux fonctions f et g, on dit que $f = O(g)$ au voisinage d'un point x_0 ssi :

$$\exists C > 0, \quad \exists \eta > 0, \quad \forall x \in I, \quad |x - x_0| < \eta, \quad |f(x)| \leq C\,|g(x)|.$$

et on note également $f(x) = O(g(x))$ au voisinage de x_0. On dit aussi que f est (au plus) du même ordre de grandeur que g au voisinage de x_0.

1.3 Dérivation

Définition 1.22 Soit $f : I \to \mathbb{R}$, et soit $x_0 \in I$. Quand la limite suivante existe dans $\mathbb{R}$:

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \quad \in \mathbb{R}, \tag{1.7}$$

i.e. quand il existe $c \in \mathbb{R}$ tel que $c = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$, on dit que f est dérivable en x_0, et on note $c = f'(x_0)$.

Donc, si f est dérivable en x_0, on note :

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0).$$

Quand on écrit $x \to x_0$, on sous-entend bien sûr que $x \in I$, sinon $f(x)$ n'aurait pas de sens. Cette définition contient donc la définition de la dérivée à droite (ou dérivée par la droite) correspondant au cas où $x_0 = a$ (et donc $I = [a, b[$ ou $[a, b]$ ou $[a, \infty[$) notée :

$$\lim_{x\to x_0+} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0+) \in \mathbb{R}, \quad \text{ou encore} \quad \lim_{h\to 0+} \frac{f(x_0+h)-f(x)}{h} = f'(x_0+) \tag{1.8}$$

avec la notation immédiate : $\displaystyle\lim_{x\to x_0+} = \lim_{\substack{x>x_0\\ x\to x_0}}$. De même pour la dérivée à gauche.

Une définition équivalente est :

$$\frac{f(x)-f(x_0)}{x-x_0} - f'(x_0) = o(1) \quad \text{au vois. de } x_0, \tag{1.9}$$

ou encore :

$$\frac{f(x)-f(x_0)}{x-x_0} = f'(x_0) + o(1) \quad \text{au vois. de } x_0, \tag{1.10}$$

ou encore :

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + o(x-x_0) \quad \text{au vois. de } x_0, \tag{1.11}$$

appelé développement limité de f à l'ordre 1 au voisinage de x_0.

La valeur $f'(x_0)$ est aussi appelée la pente de f en x_0 (= rapport côté opposé/côté adjacent) :

$$f'(x_0) = \text{pente en } x_0 = \lim_{x\to x_0} \frac{\Delta y}{\Delta x},$$

où on a noté $\Delta y = f(x)-f(x_0)$ et $\Delta x = x-x_0$.

On a le résultat immédiat :

Proposition 1.23 *Si f est dérivable en x_0, alors f est continue en x_0.*

Preuve. Par hypothèse, $f(x) = f(x_0) + (x - x_0)f'(x_0) + o(x - x_0)$, et donc, si $x \to x_0$ on a bien $f(x) \to f(x_0)$. ∎

Si les valeurs $f'(x_0)$ existent pour tous les $x_0 \in I$, où $I \subset \mathbb{R}$, on définit la fonction dérivée par :

$$f' : \begin{cases} I & \to \mathbb{R} \\ x & \mapsto f'(x). \end{cases}$$

Exemple 1.24 L'application affine $f(x) = ax + b$ est dérivable en tout point $x \in \mathbb{R}$, et a pour dérivée $f'(x) = a$ =constante pour tout $x \in \mathbb{R}$ (calcul immédiat). Et la représentation de f dans $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ par son graphe $\{(x, f(x)) : x \in \mathbb{R}\}$ est une droite de pente a passant par le point $(0, b)$. ∎

Exemple 1.25 L'application $f(x) = |x|$ est dérivable sur $\mathbb{R} - \{0\}$. En 0, la limite à droite est $+1$, et la limite à gauche est -1. Cette fonction n'est donc pas dérivable en 0. ∎

Exemple 1.26 L'application $f(x) = x^n$ pour $n \in \mathbb{N}$ est dérivable en tout point $x_0 \in \mathbb{R}$, et sa dérivée en x_0 est $f'(x_0) = nx_0^{n-1}$ (pour $n \geq 1$). En effet :

$$x^n - x_0^n = (x - x_0)(x^{n-1} + x^{n-2}x_0 + \ldots + xx_0^{n-2} + x_0^{n-1}) \tag{1.12}$$

et donc $\frac{x^n - x_0^n}{x - x_0}$ tend vers nx_0^{n-1} quand x tend vers x_0. Et par linéarité on déduit que toute fonction polynomiale est dérivable sur $\mathbb{R}$, et sa dérivée est immédiate à calculer. ∎

Exemple 1.27 L'application exponentielle $x \to exp(x) = e^x$ est définie comme étant la fonction qui est égale à sa dérivée en tout point $x \in \mathbb{R}$, i.e., pour tout $x \in \mathbb{R}$, on a $exp'(x) = exp(x)$. ∎

On note $C^1(I; \mathbb{R})$ l'ensemble des fonctions $f : I \to \mathbb{R}$ qui sont dérivables sur I de dérivée f' continue sur I (i.e. $f' \in C^0(I; \mathbb{R})$), et on note simplement C^1 lorsqu'il n'y a pas de confusion possible.

Remarque 1.28 Une fonction f peut être continue dans tout $\mathbb{R}$ et dérivable dans tout $\mathbb{R}$ sans que sa dérivée soit continue dans tout $\mathbb{R}$: par exemple, f définie sur $\mathbb{R}^*$ par $f(x) = x^2 \sin(\frac{1}{x})$ et en 0 par $f(0) = 0$ est continue sur $\mathbb{R}$.

Cette fonction est dérivable sur $\mathbb{R}^*$ de dérivée $f'(x) = 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x})$, et cette dérivée n'est pas prolongeable par continuité en 0. Pourtant $f'(0)$ existe et vaut 0 car $\frac{f(x)-f(0)}{x-0} = x \sin(\frac{1}{x})$ tend vers 0 quand $x \to 0$.

Cette fonction est donc dérivable en 0 sans que sa dérivée f' soit une fonction continue en 0. Les fonctions dérivables n'ont donc pas forcément leurs dérivées continues. ∎

La dérivation est l'opération qui consiste à calculer la dérivée. Etant donné que $\frac{a+b}{c} = \frac{a}{c} + \frac{b}{c}$ dans $\mathbb{R}$ dès que $c \neq 0$, on a immédiatement le théorème :

Théorème 1.29 *L'opération de dérivation est une opération linéaire, i.e., si f et g sont deux applications dérivables en x et si $\lambda \in \mathbb{R}$ alors :*

$$(f + \lambda g)'(x) = f'(x) + \lambda g'(x) \tag{1.13}$$

Définition 1.30 Si f est dérivable en x_0, alors la fonction affine définie par :

$$g_{x_0}(x) = f(x_0) + f'(x_0)(x - x_0), \tag{1.14}$$

i.e. par :

$$g_{x_0}(x) = ax + b \quad \text{avec} \quad a = f'(x_0) \quad \text{et} \quad b = f(x_0) - f'(x_0)x_0, \tag{1.15}$$

est appelée application affine tangente à f en x_0. Son graphe dans $\mathbb{R}^2$ est une droite tangente en x_0 au graphe de f.

Définition 1.31 Deux fonctions $f : \mathbb{R} \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ sont dites tangentes en un point x_0 ssi :

$$\lim_{x \to x_0} \frac{f(x) - g(x)}{x - x_0} = 0.$$

En particulier, si g est affine, alors g s'écrit $g(x) = a(x - x_0) + b$ où $g(x_0) = b$ et $g'(x_0) = a$, et si f est dérivable en x_0, alors $f(x) = f(x_0) + (x - x_0)f'(x_0) + o(x - x_0)$ au voisinage de x_0. Et on retrouve immédiatement que si g est tangente à f alors on doit avoir $g(x_0) = f(x_0)$ et $g'(x_0) = f'(x_0)$, et le graphe de g est la droite tangente au graphe de f en x_0.

1.4 Dérivation de produits et de composées

On rappelle que si f et g sont deux fonctions définies sur un même intervalle I, alors le produit fg est la fonction définie sur I par $fg(x) = f(x)g(x)$. Si de plus g ne s'annule pas sur I, alors $\frac{f}{g}$ est la fonction définie par $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$. Et si g est définie sur $f(I)$, alors $g \circ f$ est la fonction définie sur I par $(g \circ f)(x) = g(f(x))$.

Voici quelques règles usuelles de dérivation, basées sur les dérivées des fonctions affines $f(x) = a_0 + a_1 x$ et $g(y) = b_0 + b_1 y$ (on considère les développements limités au premier ordre, i.e. les parties affines de f et g au voisinage d'un point).

Dans ce cas $(fg)(x) = f(x)g(x) = (a_0 + a_1x)(b_0 + b_1x) = a_0b_0 + (a_0b_1 + a_1b_0)x + a_1b_1x^2$, d'où $(fg)'(x) = a_0b_1 + a_1b_0 + 2a_1b_1x = a_1(b_0 + b_1x) + (a_0 + a_1x)b_0 = (f'g + fg')(x)$.

Et $(g \circ f)(x) = g(a_0 + a_1x) = b_0 + b_1(a_0 + a_1x)$, d'où $(g \circ f)'(x) = b_1a_1 = g'(f(x))f'(x)$.

Cas général :

Théorème 1.32 *Soient f et g deux fonctions dérivables en $x \in \mathbb{R}$. Alors, le produit fg est dérivable en x, et si $g'(x) \neq 0$ alors $\frac{f}{g}$ est dérivable en x, et :*

$$\begin{aligned} &(i) \qquad (fg)'(x) = f'(x)g(x) + f(x)g'(x), \\ &(ii) \qquad \left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \end{aligned} \tag{1.16}$$

En particulier, $\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g(x)^2}$

Et si f est dérivable en x, et si g est dérivable en $y = f(x)$, alors $g \circ f$ est dérivable en x et :

$$(iii) \qquad (g \circ f)'(x) = g'(f(x))f'(x) \tag{1.17}$$

En particulier, si f est inversible au voisinage de x, si $f'(x) \neq 0$, et si f^{-1} est dérivable en $y = f(x)$, alors :

$$(iv) \qquad (f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}. \tag{1.18}$$

Preuve. Soit $x_0 \in \mathbb{R}$.

Démontrons (i), i.e., cherchons si la limite suivante existe :

$$\lim_{x \to x_0} \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \tag{1.19}$$

On a :

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = \frac{(f(x) - f(x_0))}{x - x_0} g(x) + f(x_0)\frac{(g(x) - g(x_0))}{x - x_0} \tag{1.20}$$

Mais g continue et dérivable en x_0 et f est dérivable en x_0 et donc :

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = (f'(x_0) + o(1))(g(x_0) + o(1)) + f(x_0)(g'(x_0) + o(1)). \tag{1.21}$$

L'opération passage à la limite étant linéaire, et la limite d'un produit étant égal aux produits des limites, on déduit l'existence de la limite et (i).

De même pour (ii) quand $f = 1$: on a

$$\frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} = \frac{g(x_0) - g(x)}{x - x_0} \frac{1}{g(x)g(x_0)} \tag{1.22}$$

d'où l'existence de la limite et (ii) quand $f = 1$. Puis (ii) découle de cette dernière formule et de (i) puisque $\frac{f}{g} = f\frac{1}{g}$.

Pour (iii), cherchons l'existence de la limite quand $x \to x_0$ de :

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(y) - g(y_0)}{x - x_0}, \tag{1.23}$$

où on a posé $y = f(x)$ et $y_0 = f(x_0)$. Comme g est dérivable en y_0, il vient :

$$g(y) = g(y_0) + g'(y_0)(y-y_0) + o(y-y_0) = g(y_0) + g'(y_0)(f(x)-f(x_0)) + o(f(x)-f(x_0)). \tag{1.24}$$

Et la dérivabilité de f en x_0 s'écrit $f(x)-f(x_0) = f'(x_0)(x-x_0) + o(x-x_0)$. D'où :

$$g(y) = g(y_0) + g'(y_0)[f'(x_0)(x-x_0) + o(x-x_0)] + o(f'(x_0)(x-x_0) + o(x-x_0)). \tag{1.25}$$

Et on obtient :

$$g(y) - g(y_0) = g'(y_0)f'(x_0)(x-x_0) + o(x-x_0), \tag{1.26}$$

i.e. la dérivée de $g \circ f$ en x_0 vaut $g'(y_0)f'(x_0)$.

Enfin pour (iv), si f^{-1} est dérivable en $y = f(x)$ avec f dérivable en x, de $(f \circ f^{-1})(y) = y$ on déduit :

$$f'(f^{-1}(y))(f^{-1})'(y) = 1,$$

i.e. (1.18). ▪

Remarque 1.33 Dans la dérivation de fonctions composées, il fallait comparer l'accroissement $g(f(x)) - g(f(x_0))$ avec l'accroissement $x - x_0$. Formellement, on aurait pu écrire :

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0} \longrightarrow g'(f(x_0))f'(x_0) \tag{1.27}$$

ce qui aurait donner directement le résutat. Cependant, il se peut que la fonction f oscille beaucoup autour de x_0, et on ne peut pas alors diviser par $f(x) - f(x_0)$ (exemple : $f(x) = x \sin(\frac{1}{x})$ avec $f(0) = 0$). C'est pourquoi on a composé les accroissements sans faire apparaître $f(x) - f(x_0)$ au dénominateur. ▪

Théorème 1.34 *(Leibniz) Si f et g sont n-fois dérivables en x alors fg est n-fois dérivable en x et :*

$$(fg)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x)g^{(n-k)}(x) \tag{1.28}$$

où les $\binom{n}{k} = \frac{n!}{k!(n-k)!} = C_n^k$ sont les coefficients binomiaux.

Preuve. Démonstration par récurrence, sachant que $\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$ (règle du triangle de Pascal). ▪

1.5 Théorème des accroissements finis dans $\mathbb{R}$

Théorème 1.35 *(Fermat) Si $f : [a, b] \to \mathbb{R}$ présente un extremum local (minimum ou maximum) en $x_0 \in]a, b[$ et si f est dérivable en x_0, alors $f'(x_0) = 0$.*

Preuve. Supposons f maximale en x_0 : $f(x) \leq f(x_0)$, donc $f'(x_0+) = \lim\limits_{\substack{x>x_0 \\ x\to x_0}} \dfrac{f(x) - f(x_0)}{x-x_0} \leq 0$ et $f'(x_0-) = \lim\limits_{\substack{x<x_0 \\ x\to x_0}} \dfrac{f(x) - f(x_0)}{x-x_0} \geq 0$. Comme f est dérivable en x_0 on a $f'(x_0+) = f'(x_0-) = f'(x_0)$. Donc $f'(x_0) = 0$. Et si f a un minimum en x_0, alors $-f$ est maximum en x_0 et $-f'(x_0) = 0$. ▪

Exercice 1.36 Montrer le théorème dit de Darboux : si f est dérivable sur $[a, b]$, si $f'(a) < f'(b)$ et si λ est tel que $f'(a) < \lambda < f'(b)$, alors il existe $\xi \in]a, b[$ tel que $\lambda = f'(\xi)$.

Réponse. L'inconnue du problème est un point ξ tel que $f'(\xi) = \lambda$, i.e. t.q. $f'(\xi) - \lambda = 0$. Soit $g(x) = f(x) - \lambda x$ de dérivée $g'(x) = f'(x) - \lambda$. On cherche donc un point ξ où g admet un extremum. g est continue sur $[a, b]$ et donc y atteint son maximum en un point ξ. g étant dérivable, le théorème de Fermat indique que $g'(\xi) = 0 = f'(\xi) - \lambda$. ▪

Théorème 1.37 *(Rolle) Si f est continue sur $[a,b]$, dérivable sur $]a,b[$, et si $f(a)=f(b)$ alors il existe $\xi \in]a,b[$ tel que $f'(\xi)=0$.*

Preuve. f étant continue sur $[a,b]$ y admet un extremum local, et f étant dérivable, on applique le théorème de Fermat. ∎

Théorème 1.38 *(des accroissements finis) Si $f:[a,b]\to\mathbb{R}$ est continue sur $[a,b]$ et dérivable sur $]a,b[$, alors :*

$$\exists c \in]a,b[, \quad \frac{f(b)-f(a)}{b-a} = f'(c). \tag{1.29}$$

I.e., il existe $c\in]a,b[$ tel que $f'(c)$ soit la pente moyenne. Ou encore :

$$\exists \theta \in]0,1[, \quad f(b)-f(a) = (b-a)f'(a+\theta(b-a)).$$

(Ce théorème est aussi parfois appelé théorème de Lagrange. Et le point $c=a+\theta(b-a)$ s'écrit aussi $c=(1-\theta)a+\theta b$, qui est l'écriture usuelle du barycentre de a et b.)

Preuve. La droite joignant a et b a pour équation $y=f(a)+\frac{f(b)-f(a)}{b-a}(x-a)$. On considère alors la fonction $g(x)=f(x)-f(a)-\frac{f(b)-f(a)}{b-a}(x-a)$, qui vérifie $g(b)=g(a)$, et on lui applique le théorème de Rolle : il existe $\xi\in]a,b[$ tel que $g'(\xi)=0=f'(\xi)-\frac{f(b)-f(a)}{b-a}$. ∎

Exercice 1.39 Montrer que si f est continue sur $[a,b]$ et dérivable sur $]a,b[$ et si $f'(x)=0$ pour tout x dans $]a,b[$, alors $f(x)$ est constante sur $[a,b]$.

Réponse : Le théorème des accroissements finis indique que, pour tout $y\in]a,b]$, il existe $c\in]a,y[$ tel que $f(y)-f(a)=f'(c)(y-a)=0$, d'où $f(y)=f(a)$ pour tout $y\in[a,b]$. ∎

Exercice 1.40 Montrer que si f est dérivable dans $]a,b[$ (où $a<b$), et si $m\le|f'(t)|\le M$ pour tout $t\in]a,b[$, alors $m\le\dfrac{|f(b)-f(a)|}{b-a}\le M$.

Réponse : Le théorème des accroissements finis indique que $f(b)-f(a)=f'(c)(b-a)$ pour un $c\in]a,b[$. D'où $|f(b)-f(a)|=(b-a)|f'(c)|$, d'où $m(b-a)\le|f(b)-f(a)|\le M(b-a)$. ∎

Théorème 1.41 *(accroissements finis généralisés) Si f et $g:[a,b]\to\mathbb{R}$ sont continues sur $[a,b]$ et dérivables sur $]a,b[$, alors :*

$$\exists c\in]a,b[, \quad (f(b)-f(a))g'(c)=(g(b)-g(a))f'(c).$$

En particulier, si $g(b)\neq g(a)$, il existe $c\in]a,b[$ tel que $\frac{f(b)-f(a)}{g(b)-g(a)}=\frac{f'(c)}{g'(c)}=\frac{\frac{f(b)-f(a)}{b-a}}{\frac{g(b)-g(a)}{b-a}}$.

Preuve. L'inconnue de ce problème est c. On pose :

$$F(x)=(f(b)-f(a))g(x)-(g(b)-g(a))f(x)$$

et on cherche s'il existe x tel que $F'(x)=0$. On a :

$$F(b)-F(a)=(f(b)-f(a))g(b)-(g(b)-g(a))f(b)-(f(b)-f(a))g(a)+(g(b)-g(a))f(a)=0,$$

et donc on peut appliquer le théorème de Rolle : $\exists c\in]a,b[$ t.q. $F'(c)=0$. Et $F'(c)=(f(b)-f(a))g'(c)-(g(b)-g(a))f'(c)$. ∎

Corollaire 1.42 *(Règle de l'Hôpital) Si f et $g:[a,b]\to\mathbb{R}$ sont continues sur $[a,b]$ et dérivables sur $]a,b[$, alors :*

$$\lim_{x\to a_+}\frac{f(x)-f(a)}{g(x)-g(a)}=\lim_{x\to a_+}\frac{f'(x)}{g'(x)}$$

dans le cas où ces limites existent. En particulier, si f et g sont dérivables à droite en a alors ces limites valent $\frac{f'(a)}{g'(a)}$.

Preuve. On applique le théorème précédent qui donne $(f(x)-f(a))g'(c_x) = (g(x)-g(a))f'(c_x)$ avec $a < c_x < x$, pour x quelconque dans $]a,b]$, d'où le résultat quand $x \to a$.
(C'est une autre façon d'écrire que $\dfrac{f(x)-f(a)}{g(x)-g(a)} = \dfrac{f(x)-f(a)}{x-a} \Big/ \dfrac{g(x)-g(a)}{x-a}$ quand ça a un sens, ou encore de dire que les comportements de f et g sont caractérisés par leur développement limité à l'ordre 1.) ■

Théorème 1.43 *(Théorème de Rolle Généralisé). Si f est une fonction $C^2([a,b])$ qui s'annule en les 3 points a, b et $c \in]a,b[$, alors il existe $\xi \in]a,b[$ tel que $f''(\xi) = 0$ (il y a alors un point ou la courbure est nulle).*

Et plus généralement, si $f \in C^n([a,b])$ s'annule en $n+1$ points distincts de $[a,b]$, alors il existe $\xi \in]a,b[$ tel que $f^{(n)}(\xi) = 0$.

Preuve. Le thoéorème de Rolle indique que f' s'annule en deux points distincts : une fois sur $]a,c[$ et une fois sur $]c,b[$, et donc $(f')'$ s'annule une fois sur $]a,b[$. Par récurrence, on généralise. ■

1.6 Primitives et intégrales

Définition 1.44 À $f : I \to \mathbb{R}$ fonction donnée, si le problème :
« trouver une fonction $F : I \to \mathbb{R}$ telle que $F'(x) = f(x)$ pour tout $x \in I$ »
a une solution F, on dit que la fonction F est une primitive de f sur I.

On note immédiatement que si F est une primitive de f, alors pour toute constante $c \in \mathbb{R}$, $F + c$ est aussi une primitive de f. Réciproquement :

Théorème 1.45 *Si F_1 et F_2 sont deux primitives de f alors il existe $c \in \mathbb{R}$ tel que $F_1 = F_2 + c$, i.e., deux primitives de f diffèrent au plus d'une constante.*

Preuve. La fonction $g(x) = F_1(x) - F_2(x)$ est telle que $g'(x) = 0$. Alors le théorème des accroissements finis indique que '$g =$ *constante*', voir exercice 1.39. ■

On peut alors définir :

Définition 1.46 On dit qu'une fonction f est intégrable sur I si elle admet une primitive sur I.

En fait, la définition d'une intégrale de f est plus générale : voir un cours sur les sommes de Riemann pour l'intégrale de Riemann, ou sur l'intégrale de Lebesgue pour l'intégrale de Lebesgue.

Définition 1.47 On appelle intégrale (définie) de f sur $I = [a,b]$ la différence $F(b) - F(a)$ (si elle existe), où F est une primitive de f sur I, et on note :

$$\int_a^b f(x)\,dx = F(b) - F(a) \overset{\text{noté}}{=} \int_a^b f. \tag{1.30}$$

Donc l'intégrale d'une fonction ne dépend que de la valeur d'une primitive aux extrémités de l'intervalle.

Et $\int_a^b f$ représente l'aire sous la courbe (voir les sommes de Riemann).

En particulier, si f est dérivable sur I, alors f' est intégrable sur I et :

$$\int_a^b f'(x)\,dx = f(b) - f(a). \tag{1.31}$$

Et on retrouve les formules de base de l'intégration :

Théorème 1.48 *L'intégration est une opération linéaire, i.e. "$\int_a^b (f + \lambda g) = \int_a^b f + \lambda \int_a^b g$", pour tout $\lambda \in \mathbb{R}$ et toutes fonctions intégrables f et g. Et pour tout $c \in [a,b]$, relation de Chasles :*

$$\int_a^b f = \int_a^c f + \int_c^b f. \tag{1.32}$$

Preuve. Si F et G sont des primitives de f et g alors immédiatement $(F+\lambda G)'(x) = F'(x)+\lambda G'(x)$. Puis $F(b) - F(a) = F(b) - F(c) + F(c) - F(a)$. ∎

On note, pour une fonction F donnée :

$$F(b) - F(a) \overset{\text{noté}}{=} [F]_a^b \overset{\text{noté}}{=} [F(x)]_a^b.$$

Théorème 1.49 *(Intégration par parties) Si f et g sont deux fonctions dérivables sur $I = [a, b]$ et si $f'g$ et fg' admettent des primitives sur I alors :*

$$\int_a^b f'(x)g(x)\,dx + \int_a^b f(x)g'(x)\,dx = [fg]_a^b \quad (= f(b)g(b) - f(a)g(a)), \tag{1.33}$$

soit :

$$\int_a^b f'(x)g(x)\,dx = -\int_a^b f(x)g'(x)\,dx + [f(x)g(x)]_a^b. \tag{1.34}$$

Preuve. fg est dérivable et $(fg)' = f'g + fg'$. ∎

Théorème 1.50 *(Changement de variables) Si f est dérivable sur $I = [a, b]$, si g est intégrable sur $f(I)$, et si $(g \circ f)f'$ est intégrable sur I alors :*

$$\int_{x=a}^b g(f(x))f'(x)\,dx = \int_{y=f(a)}^{f(b)} g(y)\,dy. \tag{1.35}$$

Preuve. Soit G une primitive de g. Alors avec la dérivation de fonctions composées, $(G \circ f)'(x) = g(f(x))f'(x)$, d'où :

$$\int_a^b (G \circ f)'(x)\,dx = [(G \circ f)(x)]_a^b = G(f(b)) - G(f(a)) = \int_{f(a)}^{f(b)} g(y)\,dy. \tag{1.36}$$

∎

Théorème 1.51 *(Théorème de la moyenne) Si f possède une primitive sur $[a, b]$ alors :*

$$\exists \xi \in [a, b], \quad \int_a^b f(x)\,dx = f(\xi)(b - a), \tag{1.37}$$

i.e. il existe $\xi \in [a, b]$ tel que l'aire sous le graphe de f entre a et b est égale à l'aire du rectangle de base $[a, b]$ et de hauteur $f(\xi)$.

Preuve. C'est une application du théorème des accroissements finis à F une primitive de f. ∎

Définition 1.52 On appelle valeur moyenne de f sur $[a, b]$ (ou hauteur moyenne de f entre a et b) le nombre $\bar{f} = \frac{1}{b-a}\int_a^b f(x)\,dx (= f(\xi))$.

Corollaire 1.53 *On suppose f et g intégrables sur $[a, b]$:*

(i) Si $f \geq 0$ sur $[a, b]$ alors $\int_a^b f(x)\,dx \geq 0$.

(ii) Si $f(x) \geq g(x)$ alors $\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx$.

(iii) Si $m \leq f(x) \leq M$ sur $[a, b]$, alors $m(b - a) \leq \int_a^b f(x)\,dx \leq M(b - a)$.

(iv) $|\int_a^b f(x)\,dx| \leq \int_a^b |f(x)|\,dx$.

(v) Si $f(x) \geq 0$ est continue sur $[a, b]$ et si $f \not\equiv 0$ alors $\int_a^b f(x)\,dx > 0$.

(vi) Si $g \geq 0$ sur $[a, b]$, si f, g et fg sont intégrables sur $[a, b]$ avec f continue sur $[a, b]$, alors (théorème généralisé de la moyenne) :

$$\exists \xi \in [a, b], \quad \int_a^b f(x)g(x)\,dx = f(\xi) \int_a^b g(x)\,dx. \tag{1.38}$$

1.7 Dérivation d'ordre supérieur

Définition 1.54 La fonction dérivée f' est dérivable en un point $x_0 \in I$ si la limite suivante existe dans $\mathbb{R}$ (notée dans ce cas $f''(x_0)$) :

$$\lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = f''(x_0) \in \mathbb{R} \quad \text{ou encore} \quad \lim_{h \to 0} \frac{f'(x_0 + h) - f'(x_0)}{h} = f''(x_0) \in \mathbb{R}. \tag{1.39}$$

Et on dit alors que f est deux fois dérivable en x_0.

Avec la notation 'petit o', cela s'écrit encore :

$$\frac{f'(x) - f'(x_0)}{x - x_0} = f''(x_0) + o(1), \quad \text{i.e.} \quad f'(x) = f'(x_0) + f''(x_0)(x - x_0) + o(x - x_0). \tag{1.40}$$

Exercice 1.55 Montrer que si $p : \mathbb{R} \to \mathbb{R}$ est un polynôme de degré 2 donné sous la forme $p(h) = a+bh+ch^2$, alors $a = p(0)$, $b = p'(0)$ et $c = \frac{p''(0)}{2}$, et donc que $p(h) = p(0)+h\,p'(0)+\frac{h^2}{2}p''(0)$.

Réponse. On a $p'(h) = b + 2ch$ et $p''(h) = 2c$, d'où $p'(0)$ et $p''(0)$.

Exercice 1.56 Montrer que si $p : \mathbb{R} \to \mathbb{R}$ est un polynôme de degré 2 donné sous la forme $p(x) = a + b(x-x_0) + c(x-x_0)^2$, alors on a $a = p(x_0)$, $b = p'(x_0)$ et $c = \frac{p''(x_0)}{2}$, et donc que $p(x) = p(x_0) + (x-x_0)\,p'(x_0) + \frac{(x-x_0)^2}{2}p''(x_0)$.

Réponse. On a $p(x_0) = a$. Puis on a $p'(x) = b+2c(x-x_0)$ et $p''(x) = 2c$, d'où $p'(x_0) = b$ et $p''(x_0) = 2c$.

Définition 1.57 On définit les dérivées d'ordre n par récurrence : on note $f^{(0)} = f$, $f^{(1)} = f'$, et f est dérivable à l'ordre n en x_0 si $f^{(n-1)}$ est dérivable sur un voisinage de x_0.

Définition 1.58 On note $C^n(I;\mathbb{R})$ (ou plus simplement C^n), l'ensemble des fonctions f qui sont dérivables n-fois dans I et telles que $f^{(n)}$ est continue dans I.

1.8 Développement limité d'un polynôme

Soit le monôme défini par, pour $n \in \mathbb{N}^*$:

$$p(x) = x^n, \quad \forall x \in \mathbb{R}.$$

On s'intéresse à ce qui se passe au voisinage d'un point $x_0 \in \mathbb{R}$, ce qui revient à considérer les variations de $p(x_0+h)$ pour x_0 fixé et h variable (changement d'origine). On a, avec $\binom{n}{k} = \frac{n!}{k!(n-k)!}$:

$$p(x_0 + h) = (x_0 + h)^n = \sum_{k=0}^{n} \binom{n}{k} x_0^{n-k} h^k.$$

Et on remarque que, $p'(x_0) = nx_0^{n-1}$ et par récurrence, pour $0 \le k \le n$:

$$p^{(k)}(x_0) = n(n-1)...(n-k+1)x_0^{n-k} = \frac{n!}{(n-k)!}x_0^{n-k} = k!\binom{n}{k} x_0^{n-k},$$

d'où, pour $x, h \in \mathbb{R}$:

$$\begin{aligned} p(x_0 + h) &= \sum_{k=0}^{n} \frac{h^k}{k!} p^{(k)}(x_0) \\ &= p(x_0) + h\,p'(x_0) + \frac{h^2}{2}p''(x_0) + \ldots + \frac{h^n}{n!}p^{(n)}(x_0), \end{aligned}$$

expression du "développement" de $p(.)$ au voisinage de x_0 en fonction de ses dérivées en x_0, appelé développement limité en x_0.

Exemple 1.59 Pour $p(x) = x^2$, il vient $p(x_0 + h) = x_0^2 + 2x_0h + h^2 = p(x_0) + p'(x_0)h + \frac{p''(x_0)}{2}h^2$, de vérification immédiate car $p'(x_0) = 2x_0$ et $p''(x_0) = 2$.

La formule ci-dessus s'écrit également, pour tout $a, x \in \mathbb{R}$ (et ici $h = x-a$) :

$$p(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} p^{(k)}(a), \tag{1.41}$$

et c'est le développement limité de p en a. Et un polynôme étant une somme de monôme, cette formule est trivialement vraie pour tout polynôme.

1.9 Développements limités : formules de Taylor

Il s'agit d'approcher une fonction $f \in C^{n+1}(\mathbb{R})$ par un polynôme p_n de degré n au voisinage d'un point $a \in \mathbb{R}$:

$$f(x) = p_n(x) + R(x), \tag{1.42}$$

où $R(x)$ est le 'reste' (petit lorsque x est 'proche' de a). On va montrer que :

$$p_n(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a),$$

et que $R(x) = O((x-a)^{n+1})$ (ou encore : $R(x) = o((x-a)^n)$), et donc (alors immédiat) qu'en particulier $p_n^{(k)}(a) = f^{(k)}(a)$.

Le résultat (1.42) est déjà acquis lorsque f est un polynôme avec $R = 0$, cf. (1.41).

1.9.1 Définition

Définition 1.60 Une fonction $f : [a,b] \to \mathbb{R}$ admet une développement limité à l'ordre n au voisinage d'un point $x_0 \in]a,b[$ ssi il existe $n+1$ constantes $c_0, ..., c_n$ telles que :

$$f(x_0+h) = c_0 + c_1 h + ... + \frac{c_n}{n!} h^n + o(h^n) \quad (= \sum_{k=1}^{n} \frac{c_k}{k!} h^k + o(h^n)), \tag{1.43}$$

au voisinage de $h = 0$.

On va voir que si f est C^n alors $c_k = f^{(k)}(x_0)$ pour $k = 0, ..., n$.

1.9.2 Formule de Taylor avec reste intégral

Théorème 1.61 *Si f est dérivable k+1-fois sur $I = [a,b]$ alors au voisinage de a :*

$$\begin{aligned} f(x) &= f(a) + \ldots + f^{(k)}(a) \frac{(x-a)^k}{k!} + \int_a^x f^{(k+1)}(t) \frac{(x-t)^k}{k!} \, dt \\ &= f(a) + \ldots + f^{(k)}(a) \frac{(x{-}a)^k}{k!} + \frac{(x{-}a)^{k+1}}{k!} \int_0^1 f^{(k+1)}((x{-}a)u{+}a) \, (1{-}u)^k \, du. \end{aligned} \tag{1.44}$$

En particulier, si $0 \in [a,b]$, pour $x \in [a,b]$ alors au voisinage de 0 :

$$\begin{aligned} f(x) &= f(0) + \ldots + f^{(k)}(0) \frac{x^k}{k!} + \int_0^x f^{(k+1)}(t) \frac{(x-t)^k}{k!} \, dt \\ &= f(0) + \ldots + f^{(k)}(0) \frac{x^k}{k!} + \frac{x^{k+1}}{k!} \int_0^1 f^{(k+1)}(xu) \, (1{-}u)^k \, du. \end{aligned} \tag{1.45}$$

Preuve. On a par définition de l'intégrale $f(x) = f(a) + \int_a^x f'(t) \, dt$. La formule est donc vraie pour $k = 0$. Puis par intégration par parties (où à x fixé on pose $u'(t) = 1$, $v(t) = f'(t)$ puis $u(t) = t - x$ et $v'(t) = f''(t)$) :

$$\int_{t=a}^{x} f'(t) \, dt = [(t-x) f'(t)]_{t=a}^{x} - \int_{t=a}^{x} (t-x) f''(t) \, dt = (x-a) f'(a) + \int_{t=a}^{x} (x-t) f''(t) \, dt.$$

D'où au second ordre :

$$f(x) = f(a) + (x-a) f'(a) + \int_a^x (x-t) f''(t) \, dt.$$

Puis par intégration par parties successives (où à x fixé on pose $u'(t) = \frac{(x-t)^k}{k!}$, $v(t) = f^{(k+1)}(t)$ puis $u(t) = \frac{-(x-t)^{k+1}}{(k+1)!}$ et $v'(t) = f^{(k+2)}(t)$) :

$$\begin{aligned} \int_{t=a}^{x} f^{(k+1)}(t) \frac{(x-t)^k}{k!} \, dt &= [\frac{-(x-t)^{k+1}}{(k+1)!} f^{(k+1)}(t)]_a^x - \int_{t=a}^{x} f^{(k+2)}(t) \frac{-(x-t)^{k+1}}{(k+1)!} \, dt \\ &= \frac{(x-a)^{k+1}}{(k+1)!} f^{(k+1)}(a) + \int_{t=a}^{x} f^{(k+2)}(t) \frac{(x-t)^{k+1}}{(k+1)!} \, dt \end{aligned} \tag{1.46}$$

d'où le résultat par récurrence.

Puis on fait le changement de variable $u = \frac{t-a}{x-a} \in [0,1]$, soit $t = a{+}(x{-}a)u$ et $dt = (x{-}a)du$. Puis $x{-}t = (x{-}a)(1{-}u)$, d'où $(x{-}t)^k = (x{-}a)^k (1{-}u)^k$, d'où le résultat. ▪

1.9.3 Corollaire : formules de Taylor avec "o" et "O"

On déduit du théorème 1.61 précédent :

Théorème 1.62 *Si $f \in C^{k+1}([a,b])$ alors pour tout $x_0, x \in [a,b]$ il existe $\xi \in]a,b[$ tel que :*

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \ldots + f^{(k)}(x_0)\frac{(x-x_0)^k}{k!} + f^{(k+1)}(\xi)\frac{(x-x_0)^{k+1}}{(k+1)!}, \tag{1.47}$$

et (donc) :

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x-x_0) + \ldots + f^{(k+1)}(x_0)\frac{(x-x_0)^{k+1}}{(k+1)!} + o((x-x_0)^{k+1}) \\ &= f(x_0) + f'(x_0)(x-x_0) + \ldots + f^{(k)}(x_0)\frac{(x-x_0)^k}{k!} + O((x-x_0)^{k+1}). \end{aligned} \tag{1.48}$$

Preuve. On applique le théorème 1.51 de la valeur moyenne, d'où la première égalité.

Puis $f^{(k+1)}$ étant continue, on a $f^{(k+1)}(\xi) = f^{(k+1)}(x_0) + o(1)$, d'où la seconde égalité.

Et ayant $f \in C^{k+1}([a,b])$ on a $f^{(k+1)}$ borné sur $[a,b]$, et on déduit la troisième égalité à l'aide de la première. ∎

Remarque 1.63 Une application peut admettre un développement limité au 2nd ordre en un point sans que la différentielle seconde existe. Exemple : $f(x) = x^3 \sin(\frac{1}{x})$ avec $f(0) = 0$ (continue en 0) qui vérifie $f'(x) = 3x^2 \sin(\frac{1}{x}) - x\cos(\frac{1}{x})$ et $f'(0) = 0$ (continue en 0), et $f'(h) = o(h)$ au voisinage de 0. Et, appliquant le théorème des accroissements finis (1.29), on a au voisinage de $h = 0$:

$$f'(x_0+h) = o(h) \quad \Longrightarrow \quad f(x_0+h) = f(x_0) + o(h^2), \tag{1.49}$$

puisque $\frac{f(x_0+h)-f(x_0)}{h} = f'(c)$ pour un c entre x_0 et x_0+h.

Donc ici $f(h) = f(0) + o(h^2) = o(h^2)$, et f admet un développement limité à l'ordre 2 au voisinage de 0. Mais $\lim_{x\to 0} \frac{f'(x)-f'(0)}{x} = -\lim_{x\to 0}\cos(\frac{1}{x})$ n'existe pas (la valeur $f''(0)$ n'existe pas). ∎

2 Fonctions de plusieurs variables à valeurs scalaires

On considère $\mathbb{R}^n$ muni de sa base canonique $(\vec{e}_1, ..., \vec{e}_n)$ et de son produit scalaire canonique défini par :

$$(\vec{x}, \vec{y})_{\mathbb{R}^n} = \sum_{i=1}^n x_i y_i \quad (= x_1y_1 + ... + x_ny_n),$$

si $\vec{x} = \sum_{i=1}^n x_i\vec{e}_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ et $\vec{y} = \sum_{i=1}^n y_i\vec{e}_i = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, associé à la norme (la longueur donnée par formule de Pythagore) :

$$||\vec{x}||_{\mathbb{R}^n} = \sqrt{x_1^2 + ... + x_n^2}. \tag{2.1}$$

On se donne un ouvert $\Omega \subset \mathbb{R}^n$.

On s'intéresse aux fonctions de variables vectorielles à valeurs scalaires, i.e. aux fonctions de type :

$$f : \begin{cases} \Omega & \to \mathbb{R} \\ \vec{x} & \mapsto f(\vec{x}) \in \mathbb{R} \end{cases}$$

On rappelle que le graphe d'une telle fonction est le sous-ensemble de $\mathbb{R}^n \times \mathbb{R}$ défini par :

$$G(f) = \{(\vec{x}, z) \in \Omega \times \mathbb{R} : z = f(\vec{x})\}.$$

2.1 Continuité

Soit $f : \Omega \longrightarrow \mathbb{R}$ et $\vec{a} \in \Omega$. Alors f est continue en $\vec{a}$ si :

$$\lim_{\vec{x}\to\vec{a}} f(\vec{x}) = f(\vec{a}), \tag{2.2}$$

i.e. $\lim\limits_{||\vec{x}-\vec{a}||\to 0} |f(\vec{x}) - f(\vec{a})| = 0$, ou encore :

$$f(\vec{x}) = f(\vec{a}) + o(1) \quad \text{au vois. de } \vec{a},$$

développement limité de f à l'ordre 0.

En termes de quantificateurs, f continue en $\vec{a}$ s'écrit :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall \vec{x} \in \Omega\ :\ ||\vec{x} - \vec{a}|| < \eta \implies |f(\vec{x}) - f(\vec{a})| < \varepsilon.$$

(En français : en $\vec{a} \in \Omega$, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que, dès que $\vec{x}$ vérifie $||\vec{x} - \vec{a}|| < \eta$, on a $|f(\vec{x}) - f(\vec{a})| < \varepsilon$.)

Et f discontinue en $\vec{a}$ s'énonce :

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists \vec{x} \in \Omega\ :\ ||\vec{x} - \vec{a}|| < \eta \ \text{ et } \ |f(\vec{x}) - f(\vec{a})| \geq \varepsilon.$$

Définition 2.1 On appelle $C^0(\Omega, \mathbb{R})$ l'ensemble des fonctions $f : \Omega \to \mathbb{R}$ qui sont continues en tout point de Ω, et on note simplement C^0 lorsqu'il n'y a pas de confusion possible.

Remarque 2.2 Il n'est pas suffisant, pour montrer que f est continue en $\vec{a}$, de montrer que, pour tout $\vec{v}$, les fonctions $g_{\vec{v}} : t \to g(t) = f(\vec{a} + t\vec{v})$ sont continues en 0 avec $g_{\vec{v}}(0)$ un réel indépendant de $\vec{v}$. I.e., il est insuffisant de vérifier que f est continue dans toutes les directions $\vec{v}$ avec une limite $c = \lim\limits_{h\to 0} f(\vec{a} + h\vec{v})$ indépendante de la direction $\vec{v}$.

Prendre par exemple $f : \mathbb{R}^2 \to \mathbb{R}$ avec $f(0,0) = 0$ et pour $(x,y) \neq \vec{0}$:

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

Voir figure 2.1. Cette fonction est continue sur $\mathbb{R}^2 - \{\vec{0}\}$. En $\vec{0}$: on a $g_{\vec{v}}(h) = f(\vec{0} + h\vec{v}) \underset{h\to 0}{\longrightarrow} 0$ quelque soit le vecteur $\vec{v}$ (fixé). Mais f n'est pas continue en $\vec{0}$: on a $f(x, x^2) = \frac{1}{2}$. ▪

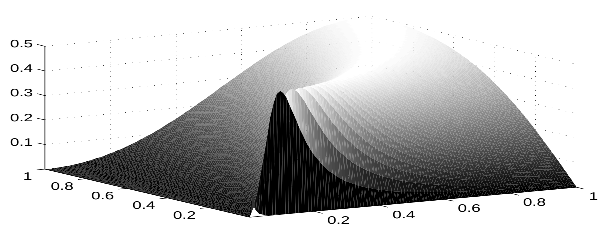

FIGURE 2.1 – Graphe de la fonction $f(x,y) = \dfrac{x^2 y}{x^4 + y^2}$.

Remarque 2.3 Pour montrer qu'une fonction est continue en un point $\vec{a}$, on peut passer en coordonnées polaires $\begin{pmatrix} x_1 = a_1 + r\cos\theta \\ x_2 = a_2 + r\sin\theta \end{pmatrix}$, et dire que $\vec{x} \to \vec{a}$, c'est dire que $||\vec{x} - \vec{a}|| = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2} = r \to 0$. ▪

2.2 Dérivation

2.2.1 Dérivée directionnelle

On s'intéresse à la dérivée d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ en un point $\vec{a} \in \mathbb{R}^n$ dans une direction $\vec{v} \in \mathbb{R}^n$. I.e. on restreint l'étude de f à la droite d'équation $t \to \vec{a} + t\vec{v}$ (droite dans $\mathbb{R}^n$), et on veut connaître la dérivée de f en $\vec{a}$ le long de cette droite. On pose :

$$g(t) = f(\vec{a} + t\vec{v}), \quad \forall t \in \mathbb{R},$$

i.e. on se ramène à l'étude de la fonction $g : \mathbb{R} \to \mathbb{R}$, et on veut regarder les variations $g'(t)$ de g en $t = 0$:

Définition 2.4 La dérivée de f en $\vec{a}$ dans la direction $\vec{v}$ est le réel :

$$g'(0) \overset{\text{noté}}{=} \partial_{\vec{v}} f(\vec{a}) \overset{\text{noté}}{=} \frac{\partial f}{\partial \vec{v}}(\vec{a}). \tag{2.3}$$

On a donc :

$$\frac{\partial f}{\partial \vec{v}}(\vec{a}) = \lim_{t\to 0} \frac{f(\vec{a}+t\vec{v}) - f(\vec{a})}{t}.$$

Définition 2.5 Lorsque $\vec{v} = \vec{e}_i$ (le i-ème vecteur de base), on note :

$$\partial_{\vec{e}_i} f(\vec{a}) \overset{\text{noté}}{=} \frac{\partial f}{\partial x_i}(\vec{a}) \overset{\text{noté}}{=} \partial_i f(\vec{a}), \tag{2.4}$$

et ces réels sont appelés les dérivées partielles de f en $\vec{a}$.

On a donc, par exemple pour $f : \mathbb{R}^2 \to \mathbb{R}$:

$$\frac{\partial f}{\partial x_1}(\vec{a}) = \lim_{t\to 0} \frac{f(a_1+t, a_2) - f(a_1, a_2)}{t},$$

où on a noté $\vec{a} = (a_1, a_2)$, ou encore avec des notations génériques, i.e. pour la dérivée en un point $\vec{x}$:

$$\frac{\partial f}{\partial x_1}(\vec{x}) = \lim_{h\to 0} \frac{f(x_1+h, x_2) - f(x_1, x_2)}{h}.$$

Et dans cette expression, x_2 joue le rôle d'une constante, la seule quantité variable étant x_1.

Remarque 2.6 Attention tout de même : si on note (y_1, y_2) les variables, i.e. on considère $f(y_1, y_2)$, la notation $\partial_1 f$ (qui n'est pas ambiguë) se note indifféremment $\frac{\partial f}{\partial x_1}$ ou $\frac{\partial f}{\partial y_1}$ (qui peut paraître ambiguë). ▪

On vient ainsi de définir les fonctions, pour $i = 1, 2$:

$$\partial_i f = \frac{\partial f}{\partial x_i} : \begin{cases} \Omega \to \mathbb{R} \\ \vec{x} \mapsto \partial_i f(\vec{x}) = \dfrac{\partial f}{\partial x_i}(\vec{x}) \end{cases}$$

quand elles ont un sens dans tout Ω.

Définition 2.7 Et une fonction qui est dérivable dans toutes les directions est appelée Fréchet dérivable ou bien Fréchet différentiable.

Exercice 2.8 Soit $f(x, y) = \sin(x^2 + y)$. Calculer $\partial_1 f$, $\partial_2 f$ et $\partial_{\vec{e}_1+\vec{e}_2} f$ en tout $\vec{x} = (x, y) \in \mathbb{R}^2$. ▪

Exercice 2.9 Soit $f(\vec{x}) = ||\vec{x}||_2 (= \sqrt{x^2+y^2})$. Calculer ses dérivées partielles en $\vec{x} \neq \vec{0}$. Est-ce que f a des dérivées partielles en $\vec{0}$? (On rappelle que $\sqrt{x^2} = |x|$.) ▪

Exercice 2.10 Soit $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$ si $\vec{x} \neq 0$ et $f(\vec{0}) = 0$. Calculer ses dérivées partielles en tout $\vec{x} \neq \vec{0}$, puis en $\vec{x} = \vec{0}$. Puis posant $\vec{v} = (v_1, v_2)$, calculer $\partial_{\vec{v}} f(\vec{x})$ en tout point $\vec{x}$. ▪

Exercice 2.11 Soit $f(x, y) = xy\frac{x^2-y^2}{x^2+y^2}$ si $\vec{x} \neq 0$ et $f(\vec{0}) = 0$. Montrer que f est continue en $\vec{0}$. Calculer ses dérivées partielles en tout $\vec{x} \neq \vec{0}$, puis en $\vec{x} = \vec{0}$.

Réponse. On passe en coordonnées polaires : $f(r\cos\theta, r\sin\theta) = r^2 \cos\theta \sin\theta(\cos^2\theta - \sin^2\theta) \longrightarrow_{r\to 0} 0$. D'où f est continue en $\vec{0}$. Et en $\vec{0}$ on a : $\lim_{h\to 0} \frac{f(\vec{0}+h\vec{e}_1)-f(\vec{0})}{h} = 0 = \frac{\partial f}{\partial x_1}(\vec{0})$. ▪

Exercice 2.12 Montrer que, pour $\lambda \in \mathbb{R}$ et $\lambda \neq 0$, on a :

$$\frac{\partial f}{\partial(\lambda\vec{v})}(\vec{a}) = \lambda \frac{\partial f}{\partial \vec{v}}(\vec{a})$$

Réponse : par définition, on a, posant $\vec{w} = \lambda\vec{v}$:

$$\begin{aligned} \frac{\partial f}{\partial \vec{w}}(\vec{a}) &= \lim_{h\to 0} \frac{f(\vec{a}+h\vec{w}) - f(\vec{a})}{h} = \lim_{h\to 0} \frac{f(\vec{a}+(h\lambda)\vec{v}) - f(\vec{a})}{h\lambda\dfrac{1}{\lambda}} \\ &= \lambda \lim_{k\to 0} \frac{f(\vec{a}+k\vec{v}) - f(\vec{a})}{k} = \lambda \frac{\partial f}{\partial \vec{v}}(\vec{a}), \end{aligned}$$

soit $\dfrac{\partial f}{\partial(\lambda\vec{v})}(\vec{a}) = \lambda \dfrac{\partial f}{\partial \vec{v}}(\vec{a})$ comme annoncé : donc attention aux changements d'unités (ex : mètre $\to$ yard). ▪

2.2.2 Fonction dérivable, différentielle, gradient

Pour simplifier les écritures, on traite le cas $\Omega \subset \mathbb{R}^2$, la généralisation au cas $\Omega \subset \mathbb{R}^n$ ne posant pas de problème.

Cependant, la dérivabilité dans toutes les directions n'est pas une notion assez forte : par exemple, la fonction définie sur $(\mathbb{R}^2)^*$ par $f(x_1,x_2) = \frac{x_1^2-x_2^2}{x_1^2+x_2^2}$ et par $f(\vec{0}) = 0$ est constante sur toutes les droites $t \to t\vec{v}$ privée du point $\vec{0}$ quelque soit la direction $\vec{v} \in \mathbb{R}^2$.

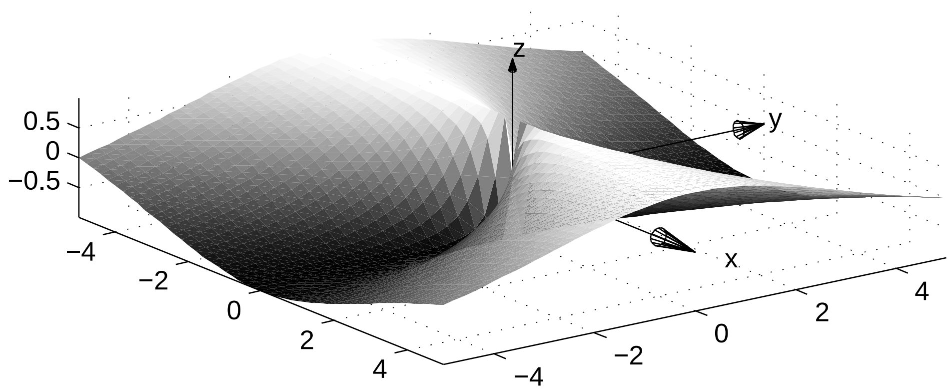

FIGURE 2.2 – Graphe de la fonction $f(x,y) = \dfrac{x^2-y^2}{x^2+y^2}$.

Cependant, en $\vec{a} = (0,0)$ la fonction f n'est même pas continue (voir figure 2.2).

D'où la définition suivante :

Définition 2.13 Soit $\vec{a} \in \Omega$. Une fonction $f : \Omega \to \mathbb{R}$ (à valeurs scalaires) est dite différentiable en $\vec{a}$, ou dérivable (au sens de Gateaux) en $\vec{a}$, ssi son graphe admet un plan tangent en $\vec{a}$, i.e. ssi il existe une application linéaire $\ell_{\vec{a}} : \vec{x} \in \mathbb{R}^n \to \ell_{\vec{a}}(\vec{x}) \in \mathbb{R}$ telle que :

$$f(\vec{x}) = f(\vec{a}) + \ell_{\vec{a}}(\vec{x}-\vec{a}) + o(||\vec{x}-\vec{a}||). \tag{2.5}$$

Et l'application linéaire $\ell_{\vec{a}}$ est appelée forme différentielle de f en $\vec{a}$, et notée $\ell_{\vec{a}} = df(\vec{a})$.

Localement au voisinage de $\vec{a}$, f est ainsi approximée par la forme linéaire $\ell_{\vec{a}}$, et l'expression (2.5) est le développement limité de f à l'ordre 1 au voisinage de $\vec{a}$.

Si on se place dans $\mathbb{R}^n$ muni de sa base canonique, une telle application linéaire s'écrit $\ell_{\vec{a}}(x_1,...,x_n) = A_1x_1 + ... + A_nx_n$, d'où la définition équivalente :

Définition 2.14 f est dite différentiable, ou dérivable (au sens de Gateaux), en $\vec{a} = (a_1,...,a_n)$ ssi :

$$\begin{aligned} &\exists A_1,...,A_n \in \mathbb{R} \quad \text{tels que, au voisinage de } \vec{x} = \vec{a} : \\ &\quad f(x_1,...,x_n) = f(a_1,...,a_n) + A_1(x_1-a_1) + ... + A_n(x_n-a_n) + o(||\vec{x}-\vec{a}||). \end{aligned} \tag{2.6}$$

Exemple 2.15 Pour $f : \mathbb{R}^2 \to \mathbb{R}$, on vient d'écrire, notant $z = f(x,y)$, que "$z = \alpha x + \beta y + \gamma + o(||\vec{x}-\vec{a}||)$" où $\alpha = A_1$, $\beta = A_2$ et $\gamma = f(\vec{a}) - A_1x - A_2y$, i.e. que le graphe de f admet en $\vec{a}$ un plan tangent à savoir le plan d'équation "$z = \alpha x + \beta y + \gamma$".

Autrement dit, f est différentiable en $\vec{a}$ ssi le graphe de f admet un plan tangent en $\vec{a}$. ▪

Proposition 2.16 *Si f est différentiable en $\vec{a} = (a_1,...,a_n)$, alors f est continue en $\vec{a}$, et on a $A_1 = \frac{\partial f}{\partial x_1}(\vec{a})$,..., $A_n = \frac{\partial f}{\partial x_n}(\vec{a})$, et donc :*

$$f(\vec{x}) = f(\vec{a}) + (x_1-a_1)\frac{\partial f}{\partial x_1}(\vec{a}) + ... + (x_n-a_n)\frac{\partial f}{\partial x_n}(\vec{a}) + o(||\vec{x}-\vec{a}||). \tag{2.7}$$

(Appelé développement limité de f au premier ordre au voisinage de $\vec{a}$.)

Preuve. La définition (2.6) donne $f(\vec{x}) = f(\vec{a}) + o(1)$, donc f est continue en $\vec{a}$.

Puis regardons par exemple A_1 (même démarche pour les A_i) : on prend $\vec{x} = \vec{a} + h\vec{e}_1 = (a_1{+}h, a_2, ...)$. Il vient :

$$f(a_1{+}h, a_2, ...) = f(a_1, a_2, ...) + A_1\ h + 0 + o(h).$$

D'où $\frac{f(a_1+h,a_2,...)-f(a_1,a_2,...)}{h} = A_1 + o(h)$, d'où $A_1 = \frac{\partial f}{\partial x_1}(\vec{a})$. ∎

Définition 2.17 Avec le produit scalaire canonique de $\mathbb{R}^n$, il existe un unique vecteur noté $\vec{\text{grad}} f(\vec{a})$, appelé gradient de f en $\vec{a}$, tel que $\ell_{\vec{a}}(\vec{v}) = (\vec{\text{grad}} f(\vec{a}), \vec{v})_n$ pour tout $\vec{v} \in \mathbb{R}^n$ (donné par le théorème de représentation de Riesz).

Et ici (pour le produit scalaire canonique) on a :

$$\vec{\text{grad}} f(\vec{a}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{a}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{a}) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} (\vec{a}) \overset{\text{noté}}{=} \vec{\nabla} f(\vec{a}). \tag{2.8}$$

Donc :

$$f(\vec{a} + h\vec{v}) - f(\vec{a}) = h\ (\vec{\text{grad}} f(\vec{a}), \vec{v})_{\mathbb{R}^n} + o(h). \tag{2.9}$$

Définition 2.18 On appelle matrice jacobienne de f en $\vec{a}$ la matrice ligne :

$$J_f(\vec{a}) = [\frac{\partial f}{\partial x_1}(\vec{a}) \quad \dots \quad \frac{\partial f}{\partial x_n}(\vec{a})].$$

(Matrice des composantes de la différentielles $df(\vec{a})$ dans la base duale de la base canonique, voir cours de géométrie différentielle).

En particulier, disposant du produit scalaire canonique, on a :

$$J_f(\vec{a}) = \vec{\text{grad}} f(\vec{a})^T. \tag{2.10}$$

Donc, si f est dérivable en $\vec{a}$, on a :

$$\begin{aligned} f(\vec{x}) &= f(\vec{a}) + (\vec{\text{grad}} f(\vec{a}), \vec{x} - \vec{a})_{\mathbb{R}^n} + o(||\vec{x} - \vec{a}||) \\ &= f(\vec{a}) + J_f(\vec{a}).(\vec{x} - \vec{a}) + o(||\vec{x} - \vec{a}||) \\ &= f(\vec{a}) + \vec{\text{grad}} f(\vec{a})^T.(\vec{x} - \vec{a}) + o(||\vec{x} - \vec{a}||). \end{aligned} \tag{2.11}$$

(Le "." étant la notation du produit matricielle.)

Proposition 2.19 *Si f est différentiable en $\vec{a}$, alors f est dérivable dans toutes les directions en $\vec{a}$, et on a :*

$$\frac{\partial f}{\partial \vec{v}}(\vec{a}) = (\vec{\text{grad}} f(\vec{a}), \vec{v})_{\mathbb{R}^n}. \tag{2.12}$$

Preuve. $f(\vec{a} + h\vec{v}) - f(\vec{a}) = A_1 h v_1 + ... + A_n h v_n + o(h) = h(\vec{\text{grad}} f(\vec{a}), \vec{v})_{\mathbb{R}^n} + o(h)$. ∎

Par contre, la réciproque est fausse : on peut avoir f dérivable dans toutes les directions en un point $\vec{x}$ sans que f soit différentiable en $\vec{x}$. Exemple $f(x,y) = x$ (de graphe un plan) si $x \neq 0$ et $f(0,y) = y$. Faire un dessin. Alors f est dérivable en $\vec{0}$ dans toutes les directions, mais n'a pas de plan tangent.

On a quand même un critère agréable dans le cas particulier suivant :

Théorème 2.20 *Si f possède des dérivées partielles (directionnelles) $\frac{\partial f}{\partial x_1}$,..., $\frac{\partial f}{\partial x_n}$ dans un voisinage de $\vec{a}$, et si ces dérivées partielles sont continues en $\vec{a}$, alors f est différentiable et C^1 en $\vec{a}$.*

Preuve. On traite le cas $\mathbb{R}^n = \mathbb{R}^2$, le cas général étant laissé au lecteur. On a :

$$f(x_1, x_2) - f(a_1, a_2) = (f(x_1, x_2) - f(a_1, x_2)) + (f(a_1, x_2) - f(a_1, a_2)). \tag{2.13}$$

Et le dernier terme donne immédiatement (fonction de la seule variable x_2) :

$$f(a_1, x_2) - f(a_1, a_2) = \frac{\partial f}{\partial x_2}(a_1, a_2)(x_2 - a_2) + o(|x_2 - a_2|). \tag{2.14}$$

Et le premier terme considéré à x_2 fixé s'écrit :

$$f(x_1, x_2) - f(a_1, x_2) = \frac{\partial f}{\partial x_1}(a_1, x_2)(x_1 - a_1) + o(|x_1 - a_1|), \tag{2.15}$$

ce qui a un sens car on a supposé que $\frac{\partial f}{\partial x_1}$ existait dans un voisinage de $\vec{a}$, et donc en (a_1, x_2) pour x_2 suffisamment proche de a_2. Et la dérivée partielle $\frac{\partial f}{\partial x_1}$ étant supposée continue en $\vec{a}$:

$$\frac{\partial f}{\partial x_1}(a_1, x_2) = \frac{\partial f}{\partial x_1}(a_1, a_2) + o(1), \tag{2.16}$$

et par conséquent, puisque par définition $(x_1 - a_1)o(1) = o(x_1 - a_1)$:

$$f(x_1, x_2) - f(a_1, x_2) = \frac{\partial f}{\partial x_1}(a_1, a_2)(x_1 - a_1) + o(|x_1 - a_1|), \tag{2.17}$$

d'où (2.13) donne bien (2.6) : f est dérivable en $\vec{a}$. Puis les dérivées partielles étant continues, grad$f$ l'est également, et f est C^1. ▪

Définition 2.21 On note $C^1(\Omega; \mathbb{R})$ le sous ensemble de $C^0(\Omega; \mathbb{R})$ formé des fonctions f qui sont dérivables en tout point de Ω et telles que les dérivées partielles $\frac{\partial f}{\partial x_i}$ sont continues en tout point de Ω pour tout $i = 1, ..., n$. C'est ensemble est également noté plus simplement C^1 si aucune confusion n'est possible.

Exercice 2.22 Soit f définie sur $\mathbb{R}^2$ par $f(x, y) = \dfrac{(xy)^2}{(x^2+y^2)^{\frac{3}{2}}}$ si $\vec{x} \neq \vec{0}$ et $f(0,0) = 0$. Montrer que f est continue en 0, calculer $\frac{\partial f}{\partial x_1}(\vec{x})$ ou $\frac{\partial f}{\partial x_2}(\vec{x})$ et en déduire que f n'est pas dérivable en 0. On pourra également poser $g(t) = f(t, t)$ pour montrer que g n'est pas dérivable en 0.

Réponse. f est une fraction rationnelle qui n'a qu'une racine en $(x, y) = \vec{0}$, et donc f est $C^\infty(\mathbb{R}^2 - \{\vec{0}\}; \mathbb{R})$. Regardons la continuité en $\vec{0}$. On passe en coordonnées polaires. $f(r\cos\theta, r\sin\theta) = r\cos^2\theta\sin^2\theta \underset{r\to 0}{\longrightarrow} 0$.

Regardons $\frac{\partial f}{\partial x_1}$: Si $\frac{\partial f}{\partial x_1}(\vec{0})$ existe, par définition on a :

$$\frac{\partial f}{\partial x_1}(\vec{0}) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0.$$

Donc la 1ère dérivée partielle existe. De même $\frac{\partial f}{\partial x_2}(\vec{0}) = 0$. Mais ça ne veut pas dire que f est dérivable. D'ailleurs f n'est pas dérivable : en effet si elle l'était, le vecteur gradient existerait et on aurait $\vec{\mathrm{grad}} f(\vec{0}) = \vec{0}$, d'où :

$$f(x, y) = f(0, 0) + (\vec{\mathrm{grad}} f(\vec{0}), \begin{pmatrix} x \\ y \end{pmatrix}) + o(||(x, y)||) = 0 + 0 + o(||(x, y)||).$$

En particulier on aurait $f(h, h) = o(h)$. Or on a $f(h, h) = \frac{h^4}{(2h^2)^{\frac{3}{2}}} = \frac{h^4}{\sqrt{8}|h|^3} = \frac{1}{\sqrt{8}}|h|$ qui n'est pas $o(h)$. Donc l'expression du gradient $\vec{\mathrm{grad}} f(\vec{0}) = \vec{0}$ est absurde, donc le gradient n'existe pas en $\vec{0}$, i.e. la fonction n'est pas dérivable en $\vec{0}$.

On peut aussi remarquer qu'on ne peut pas appliquer le théorème 2.20 : on a, quand $\vec{x} \neq \vec{0}$:

$$\frac{\partial f}{\partial x_1}(\vec{x}) = \frac{2xy^2(x^2+y^2)^{\frac{3}{2}} - 3x^3y^2(x^2+y^2)^{\frac{1}{2}}}{(x^2+y^2)^3} = xy^2 \frac{2y^2 - x^2}{(x^2+y^2)^{\frac{5}{2}}}.$$

En particulier, $\frac{\partial f}{\partial x_1}(0, h) = 0$ alors que $\frac{\partial f}{\partial x_1}(h, h) = \frac{1}{2^{\frac{5}{2}}}$ et donc que $\frac{\partial f}{\partial x_1}(0, h)$ et $\frac{\partial f}{\partial x_1}(h, h)$ n'ont pas la même limite quand $h \to 0$. Et donc $\frac{\partial f}{\partial x_1}$ n'est pas continue dans un voisinage de $\vec{0}$, et on ne peut pas appliquer le théorème. (Ici on a une valeur calculée $\frac{\partial f}{\partial x_1}(\vec{0}) = 0$ mais cette valeur ponctuelle n'est pas la prolongation par continuité de $\frac{\partial f}{\partial x_1}(\vec{x})$ quand $\vec{x} \to \vec{0}$: si on pouvait prolonger par continuité on aurait $0 = \frac{1}{2^{\frac{5}{2}}}$ par unicité de la limite, ce qui est absurde.) En particulier f n'est pas $C^1(\mathbb{R}^2;\mathbb{R})$.

Regardons maintenant directement $g(t) = f(t,t) = \frac{1}{2^{\frac{3}{2}}}\frac{t^4}{(t^2)^{\frac{3}{2}}} = \frac{1}{\sqrt{2}^3}\frac{t^4}{|t|^3} = \frac{1}{\sqrt{2}^3}|t|$: cette fonction n'est pas dérivable en 0, et donc $\frac{\partial f}{\partial(\vec{e}_1+\vec{e}_2)}$ n'est pas définie en $\vec{0}$. Donc f n'est pas dérivable en $\vec{0}$, sinon toutes les dérivées directionnelles existeraient, proposition 2.19, et en particulier $\frac{\partial f}{\partial(\vec{e}_1+\vec{e}_2)}$ existerait. ▪

Exercice 2.23 Montrer que si $f = \ell : \mathbb{R}^n \to \mathbb{R}$ est linéaire, alors en tout point $\vec{x} \in \mathbb{R}^n$ on a $d\ell(\vec{x}) = \ell$ (sa différentielle est indépendante du point $\vec{x}$ et vaut ℓ). Que vaut sa matrice jacobienne (matrice représentant $d\ell(\vec{x})$ dans les bases canoniques) ?

Réponse. On a $\ell(\vec{y})-\ell(\vec{x}) = \ell(\vec{y}-\vec{x})$ par linéarité, et donc la relation $\ell(\vec{y})-\ell(\vec{x}) = d\ell(\vec{x})(\vec{y}-\vec{x})+o(||\vec{y}-\vec{x}||)$ est trivialement satisfaite avec $d\ell(\vec{x}) = \ell$, le reste $o(||\vec{y}-\vec{x}||)$ étant identiquement nul (voir (2.5)).

Sa matrice jacobienne est la matrice représentant $d\ell(\vec{x})$ après avoir choisi les bases canoniques dans $\mathbb{R}^n$ et $\mathbb{R}^p$. C'est donc la matrice $[\ell]$ représentant l'application linéaire ℓ. Ainsi si $\ell(\vec{x}) = a_1x_1 + \ldots + a_nx_n$ on a $J_\ell(\vec{x}) = [a_1 \quad \ldots \quad a_n]$. ▪

Exercice 2.24 Soit $\mathrm{Tr} : A \in \mathbb{R}^{n^2} \to \mathrm{Tr}(A) \in \mathbb{R}$ l'application qui a toute matrice $A = [a_{ij}]$ associe sa trace $\mathrm{Tr}(A) = \sum_i a_{ii}$. Montrer qu'en chaque A la différentielle $d\mathrm{Tr}(A) = \mathrm{Tr}$. Montrer que sa matrice jacobienne est la matrice identité (après avoir choisi la base canonique dans $\mathbb{R}^{n^2}$).

Réponse. Approche se servant de l'exercice précédent : l'application Tr est linéaire. Donc $d\mathrm{Tr}(A) = \mathrm{Tr}$. On vérifie effectivement que $\mathrm{Tr}(B) = \mathrm{Tr}(A)+\mathrm{Tr}(B-A)$ donne trivialement $d\mathrm{Tr}(A) = \mathrm{Tr}$ fonction indépendante de A.

La base canonique de $\mathbb{R}^{n^2}$ est donnée par les matrices E_{ij} définies par : tous les termes sont nuls sauf le terme (i,j) qui vaut 1. Et Tr est connu dès qu'on connaît $\mathrm{Tr}(E_{ij})$ pour tout $i,j = 1,...,n$ (une application linéaire est connue dès qu'on connaît l'image de chacun de ses vecteurs de base). Ici $\mathrm{Tr}(E_{ij}) = \delta_{ij}$.

On a $\frac{\partial \mathrm{Tr}}{\partial E_{ij}}(A) = d\mathrm{Tr}(A)(E_{ij}) = \mathrm{Tr}(E_{ij}) = \delta_{ij}$ (vaut 1 si i=j et 0 sinon). Et donc la matrice jacobienne (matrice gradient) de Tr est la matrice $[\frac{\partial \mathrm{Tr}}{\partial E_{ij}}(A)] = [\delta_{ij}] = I=$ l'identité.

Remarque : pour une matrice $A = [a_{ij}]_{1\le i,j\le n} = \sum_{ij} a_{ij}E_{ij}$ on a

$\mathrm{Tr}(A) = \sum_{i,j=1}^n a_{ij}\mathrm{Tr}(E_{ij}) = \sum_{i,j=1}^n \delta_{ij}a_{ij} = I : A \quad (= \sum_i a_{ii})$

puisque $I = [\delta_{ij}]_{1\le i,j\le n}$, où ":" est la double contraction définie par $A : B = \sum_{i,j=1}^n a_{ij}b_{ij} = \mathrm{Tr}(AB^t)$. D'où $\mathrm{Tr} : \mathbb{R}^{n^2} \to \mathbb{R}$ est représentée par la matrice I dans la base (E_{ij}). D'où $d\mathrm{Tr}(A) = \mathrm{Tr} = I$. ▪

Exercice 2.25 On considère le déterminant $\det : A \in \mathbb{R}^{n^2} \to \det A \in \mathbb{R}$ (défini sur les matrices $n \times n$). Montrer que det est différentiable et que $d(\det A).B = \det A \ \mathrm{Tr}(A^{-1}B)$. Et comme la matrice des cofacteurs est définie par $\mathrm{cof}(A)^T = (\det A)A^{-1}$ on a donc également

$d(\det A).B = \mathrm{cof}(A) : B$ $(= \mathrm{Tr}(\mathrm{cof}(A)^T.B)$ avec la définition de la double contraction ci-dessus).

Réponse. Si H est une matrice dans un voisinage de la matrice nulle on a

$\det(I + hM) = 1 + h\mathrm{Tr}(M)+$ monomes de degré ≥ 2 en h,

par définition du déterminant. Or on a $\det(A+hB) = \det(A)\det(I+hA^{-1}B)$. On en déduit que $\det(A+hB) = \det(A)(1 + h\mathrm{Tr}(A^{-1}B)) + o(h)$. ▪

Exemple 2.26 En dimension infinie. Soit $Z(\varphi) = \int_\Omega f(\vec{x})\varphi(\vec{x})\,d\Omega$, où f et φ à valeurs réelles sont continues dans Ω ouvert de $\mathbb{R}^n$. On a alors Z dérivable sur $C^0(\Omega;\mathbb{R})$ de dérivée $dZ(\varphi) = Z$. En effet, Z est linéaire, et donc $Z(\varphi+h\psi) = Z(\varphi)+hZ(\psi) = Z(\varphi)+h\,dZ(\varphi).\psi+o(\psi)$ où $dZ(\varphi).\psi = Z(\psi)$. ▪

2.3 Interprétation géométrique du gradient

2.3.1 En terme de plus grande pente

Proposition 2.27 *Pour f fonction à valeurs scalaires, la direction donnée par le gradient $\vec{\mathrm{grad}}f(\vec{x})$ au point $\vec{x}$ est celle d'accroissement maximal de la fonction f à partir du point x. I.e., notant $\vec{n} = \frac{\vec{\mathrm{grad}}f(\vec{x})}{||\vec{\mathrm{grad}}f(\vec{x})||}$ le vecteur gradient normalisé en $\vec{x}$, on a :*

$$\forall \vec{v} \text{ t.q. } ||\vec{v}|| = 1, \quad \frac{\partial f}{\partial \vec{v}}(\vec{x}) \le \frac{\partial f}{\partial \vec{n}}(\vec{x}),$$

i.e. en termes de (limites de) taux d'accroissement :

$$\forall \vec{v} \text{ t.q. } ||\vec{v}|| = 1, \quad \lim_{h\to 0+} \frac{f(\vec{x}+h\vec{v})-f(\vec{x})}{h} \le \lim_{h\to 0+} \frac{f(\vec{x}+h\vec{n})-f(\vec{x})}{h}.$$

Preuve. Avec (2.9) il vient :

$$\frac{f(\vec{x}+h\vec{v})-f(\vec{x})}{h} = (\vec{\mathrm{grad}} f(\vec{x}), \vec{v})_{\mathbb{R}^2} + o(1).$$

Avec l'inégalité de Cauchy–Schwarz, on a $|(\vec{\mathrm{grad}} f(\vec{x}), \vec{v})_{\mathbb{R}^n}| \leq ||\vec{\mathrm{grad}} f(\vec{x})||\,||\vec{v}||$, et on a égalité lorsque $\vec{v} = \alpha\, \vec{\mathrm{grad}} f(\vec{x})$ pour un $\alpha \in \mathbb{R}$. Donc, pour $\vec{v}$ vecteur unitaire, le max du produit scalaire $(\vec{\mathrm{grad}} f(\vec{x}), \vec{v})_{\mathbb{R}^n}$ est atteint pour $\vec{v} = \frac{\vec{\mathrm{grad}} f(\vec{x})}{||\vec{\mathrm{grad}} f(\vec{x})||} = \vec{n}$. ▪

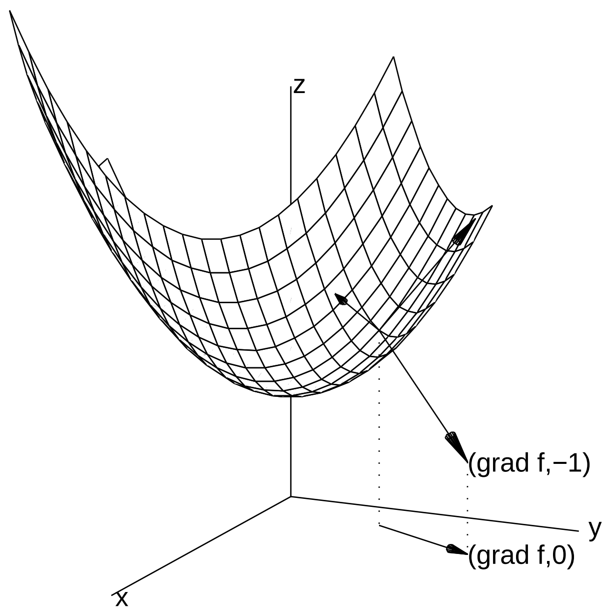

FIGURE 2.3 – Représentation du gradient sur le graphe d'une fonction.

2.3.2 En terme de normale au graphe

Définition 2.28 Une surface de $\mathbb{R}^3$ est un sous-ensemble de $\mathbb{R}^3$ de la forme $F(x,y,z)=0$ où F est une fonction de $\mathbb{R}^3$ dans $\mathbb{R}$ (relation implicite entre x, y et z).

Exemple 2.29 La sphère d'équation $x^2+y^2+z^2=1$. ▪

Exemple 2.30 Si $F(x,y,z) = z - f(x,y)$, alors on obtient la surface $z = f(x,y)$, surface explicite en z. ▪

Pour $f : \mathbb{R}^2 \to \mathbb{R}$, on visualise f à l'aide de son graphe :

$$G(f) = \{(x,y,z) \in \mathbb{R}^3 : z = f(x,y)\}.$$

Et $G(f)$ est une surface dans $\mathbb{R}^3$: posant $F(x,y,z) = z - f(x,y)$, le graphe $G(f)$ est la surface "$F(x,y,z)=0$".

Exemple 2.31 On peut représenter la sphère comme l'union des graphes des deux fonctions "$z = f(x,y) = \pm\sqrt{1-(x^2+y^2)}$" pour les x,y tels que $x^2+y^2 \leq 1$. ▪

Comme une fonction différentiable en un point $\vec{a}$ a son graphe tangent à un plan, pour connaître la normale en $\vec{a}$ à $G(f)$, il suffit de connaître la normale à un plan :

Cas d'un plan : une normale. S'il n'est pas vertical, un plan P dans $\mathbb{R}^3$ est le graphe $G(f)$ d'une fonction f du type :

$$z = f(x,y) = ax + by + c, \quad \forall x,y \in \mathbb{R}. \tag{2.18}$$

On a ici $F(x,y,z) = z - (ax+by+c)$.

Un point du plan affine "$z = ax + by + c$" vérifie $\begin{pmatrix} x \\ y \\ z-c \end{pmatrix} . \begin{pmatrix} -a \\ -b \\ 1 \end{pmatrix} = 0$, et un point du plan vectoriel "$z = ax + by$" associé vérifie $z - ax - by = 0$, i.e. $\begin{pmatrix} x \\ y \\ z \end{pmatrix} . \begin{pmatrix} -a \\ -b \\ 1 \end{pmatrix} = 0$. Donc tout point $\vec{x} = (x, y, z)$ du plan est orthogonal au vecteur $(-a, -b, 1)$, i.e. est orthogonal aux vecteurs unitaires :

$$\vec{n}(\vec{x}) = \pm \frac{1}{\sqrt{a^2 + b^2 + 1}} \begin{pmatrix} -a \\ -b \\ 1 \end{pmatrix}.$$

Comme $f(\vec{x}) = f(x, y) = ax + by - c$ avec $a = \frac{\partial f}{\partial x}(\vec{x})$ et $b = \frac{\partial f}{\partial y}(\vec{x})$, on a également, en un point $\vec{x}$ du plan :

$$\vec{n}(\vec{x}) \ // \ \begin{pmatrix} -\frac{\partial f}{\partial x_1}(\vec{x}) \\ -\frac{\partial f}{\partial x_2}(\vec{x}) \\ 1 \end{pmatrix} = \begin{pmatrix} -\left(\vec{\text{grad}} f(\vec{x})\right) \\ 1 \end{pmatrix}. \tag{2.19}$$

Avec $F(x, y, z) = z - ax - by - c$, le graphe représente la surface d'équation $F(x, y, z) = 0$, et un calcul immédiat donne également :

$$\vec{n}(\vec{x}) \ // \ \begin{pmatrix} -\left(\vec{\text{grad}} f(\vec{x})\right) \\ 1 \end{pmatrix} = \vec{\text{grad}} F(\vec{x}), \tag{2.20}$$

i.e. $\vec{n}$ est parallèle à $\vec{\text{grad}} F$.

Cas d'une fonction différentiable en $\vec{a}$: une normale $\vec{n}(\vec{a})$. Pour $f(x, y)$ fonction dérivable en (a_1, a_2), son plan tangent en (a_1, a_2) a pour équation :

$$g(x, y) = f(\vec{a}) + \vec{\text{grad}} f(\vec{a})^T . \begin{pmatrix} x - a_1 \\ y - a_2 \end{pmatrix} = A_0 + A_1 x + A_2 y, \tag{2.21}$$

où $A_1 = \frac{\partial f}{\partial x_1}(\vec{a})$, $A_1 = \frac{\partial f}{\partial x_2}(\vec{a})$ et $A_0 = f(\vec{a}) - \vec{\text{grad}} f(\vec{a})^T . \vec{a}$. Ce plan tangent est représenté par le graphe :

$$z = A_0 + A_1 x + A_2 y. \tag{2.22}$$

D'où une normale à ce plan tangent en $\vec{a}$ donnée par :

$$\vec{n}(\vec{a}) \ // \ \begin{pmatrix} -A_1 \\ -A_2 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{\partial f}{\partial x_1}(\vec{a}) \\ -\frac{\partial f}{\partial x_2}(\vec{a}) \\ 1 \end{pmatrix}.$$

Et si on pose :

$$F(x, y, z) = z - (A_0 + A_1 x + A_2 y), \tag{2.23}$$

le plan tangent est la surface plane d'équation $F(x, y, z) = 0$: et une normale à ce plan en $\vec{a} = (a_1, a_2)$ est donnée par :

$$\vec{n}(\vec{a}) = \begin{pmatrix} \left(-\vec{\text{grad}} f(\vec{a})\right) \\ 1 \end{pmatrix} = \vec{\text{grad}} F(\vec{a}, f(\vec{a})). \tag{2.24}$$

Remarque 2.32 Un point du plan a pour coordonnées $\vec{r}(x, y) = \begin{pmatrix} x \\ y \\ z = ax + by + c \end{pmatrix} \in \mathbb{R}^3$, pour tout $\vec{x} = (x, y) \in \mathbb{R}^2$, et deux vecteurs tangents en un point $\vec{a} \in \mathbb{R}^2$ sont donnés par :

$$\frac{\partial \vec{r}}{\partial x}(\vec{a}) = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x_1}(\vec{a}) \end{pmatrix}, \quad \frac{\partial \vec{r}}{\partial y}(\vec{a}) = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial x_2}(\vec{a}) \end{pmatrix}. \tag{2.25}$$

On vérifie effectivement que pour un point $\vec{r} \in P$ alors tout point de type, pour $\alpha, \beta \in \mathbb{R}$: $\vec{X} = \vec{r} + \alpha \frac{\partial \vec{r}}{\partial x} + \beta \frac{\partial \vec{r}}{\partial y} = \begin{pmatrix} X = x + \alpha \\ Y = y + \beta \\ Z = ax + by + c + \alpha a + \beta b \end{pmatrix}$ est dans P, car on vérifie bien que $Z = aX + bY + c$.

Et un calcul immédiat redonne :

$$\vec{n} \, // \, \frac{\partial \vec{r}}{\partial x} \wedge \frac{\partial \vec{r}}{\partial y}. \tag{2.26}$$

■

2.4 Règles de dérivation

Théorème 2.33 *1- Si f et g fonctions $\mathbb{R}^n \to \mathbb{R}$ sont dérivables dans la direction $\vec{e}_i$ en $\vec{a}$, alors le produit fg est dérivable dans la direction $\vec{e}_i$ en $\vec{a}$ et :*

$$\forall i = 1, ..., n, \quad \frac{\partial (fg)}{\partial x_i}(\vec{a}) = \frac{\partial f}{\partial x_i}(\vec{a}) g(\vec{a}) + f(\vec{a}) \frac{\partial g}{\partial x_i}(\vec{a}) \quad \in \mathbb{R}. \tag{2.27}$$

Ces n équations s'écrivent encore de manière condensée (système de n équations) :

$$\vec{\text{grad}}(fg) = (\vec{\text{grad}} f) g + f (\vec{\text{grad}} g) \quad \in \mathcal{F}(\mathbb{R}^n; \mathbb{R}^n). \tag{2.28}$$

2- Et si $f : \mathbb{R}^n \to \mathbb{R}$ est dérivable en $\vec{a} \in \mathbb{R}^n$, et si $g : \mathbb{R} \to \mathbb{R}$ est dérivable en $f(\vec{a}) \in \mathbb{R}$, alors $g \circ f : \mathbb{R}^n \to \mathbb{R}$ est dérivable en $\vec{a} \in \mathbb{R}^n$, et :

$$\forall i = 1, ..., n, \quad \frac{\partial (g \circ f)}{\partial x_i}(\vec{a}) = g'(f(\vec{a})) \frac{\partial f}{\partial x_i}(\vec{a}) \quad \in \mathbb{R}, \tag{2.29}$$

soit de manière condensée (système de n équations) :

$$\vec{\text{grad}}(g \circ f) = (g' \circ f) \, \vec{\text{grad}} f \quad \in \mathcal{F}(\mathbb{R}^n; \mathbb{R}^n). \tag{2.30}$$

Attention : $(g' \circ f)(\vec{a})$ est un réel valant $g'(f(\vec{a})) \in \mathbb{R}$. Et donc "$(g' \circ f)(\vec{a}) \, \vec{\text{grad}} f(\vec{a})$" est le vecteur : résultat du produit du réel $(g' \circ f)(\vec{a}) \in \mathbb{R}$ par le vecteur $\vec{\text{grad}} f(\vec{a}) \in \mathbb{R}^n$.

Preuve. Pour démontrer les dérivations de produit et de changement de variables, on applique le théorème 1.32 en se ramenant au cas de fonctions d'une seule variable : on pose $f_i(x) = f(x_1, ..., x_{i-1}, x, x_i, ..., x_n)$ fonction de la seule variable x, les x_j pour $j = 1, ..., i-1$ et $j = i+1, ..., n$ étant fixés. Même notation pour $g_i(x)$. Et $\frac{\partial (fg)}{\partial x_i} = (f_i g_i)'$, d'où le résultat. Puis pour la composée avec $g : \mathbb{R} \to \mathbb{R}$ on a $\frac{\partial (g \circ f)}{\partial x_i} = (g \circ f_i)'$, d'où le résultat. ■

2.5 Théorème des accroissements finis

(En anglais = Mean Value Theorem.)

Théorème 2.34 *Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ une fonction à valeurs scalaires différentiable dans Ω ouvert de $\mathbb{R}^n$. Soient $\vec{x}, \vec{y} \in \mathbb{R}^n$ tels que le segment $[\vec{x}, \vec{y}] = \{\vec{z} = t(\vec{y} - \vec{x}) + \vec{x} : t \in [0, 1]\}$ soit dans Ω. Alors il existe $\vec{c} \in [\vec{x}, \vec{y}]$ tel que :*

$$f(\vec{y}) - f(\vec{x}) = \vec{\text{grad}} f(\vec{c}).(\vec{y} - \vec{x}). \tag{2.31}$$

Preuve. Soit $g(t) = f(t(\vec{y} - \vec{x}) + \vec{x})$. On a $g : [0, 1] \to \mathbb{R}$ avec g différentiable, d'où il existe $t_0 \in]0, 1[$ tel que $g(1) - g(0) = g'(t_0)(1-0) = g'(t_0)$. Avec $g'(t) = \vec{\text{grad}} f(t(\vec{y} - \vec{x}) + \vec{x}).\vec{y} - \vec{x}$, et $g(1) = f(\vec{y})$ et $g(0) = f(\vec{x})$, d'où le résultat avec $\vec{c} = t_0(\vec{y} - \vec{x}) + \vec{x}$. ■

2.6 Dérivées d'ordres supérieurs, hessien, théorème de Schwarz

Soit $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ une fonction à valeurs scalaires. On a défini (quand cela avait un sens) les fonctions $f_{,i} = \frac{\partial f}{\partial x_i} : \Omega \to \mathbb{R}$ comme étant les dérivées directionnelles dans la direction $\vec{e}_i$.

Pour chacune des fonctions $f_{,i}$ on peut également considérer les dérivées directionnelles en un point $\vec{a}$:

$$\frac{\partial f_{,i}}{\partial \vec{v}}(\vec{a}) = \lim_{h \to 0} \frac{f_{,i}(\vec{a} + h\vec{v}) - f_{,i}(\vec{a})}{h}$$

qui a un sens dès que la limite existe.

En particulier pour les $\vec{v} = \vec{e}_j$, on peut définir ainsi n^2 valeurs $\frac{\partial f_{,i}}{\partial \vec{e}_j}(\vec{a})$, pour $1 \leq i,j \leq n$ (à condition que les n^2 limites existent). On note :

$$\frac{\partial f_{,i}}{\partial \vec{e}_j}(\vec{a}) = \frac{\partial f_{,i}}{\partial x_j}(\vec{a}) = \frac{\partial \frac{\partial f}{\partial x_i}}{\partial x_j}(\vec{a}) = \frac{\partial^2 f}{\partial x_j x_i}(\vec{a}).$$

Notation. On note $C^2(\Omega;\mathbb{R})$ l'ensemble des fonctions

$$C^2(\Omega;\mathbb{R}) = \{f \in C^1(\Omega;\mathbb{R}) : \forall 1 \leq i,j \leq n,\ \frac{\partial^2 f}{\partial x_j x_i} \in C^0(\Omega;\mathbb{R})\}.$$

Théorème 2.35 *(Théorème de Schwarz)* **et définition**. *Si f, les $\frac{\partial f}{\partial x_i}$ pour $1{\leq}i{\leq}n$ et les $\frac{\partial^2 f}{\partial x_i x_j}$ pour $1{\leq}i,j{\leq}n$ existent et sont continues en $\vec{a}$ (en particulier quand f est C^2 dans un voisinage de $\vec{a}$), alors :*

$$\frac{\partial^2 f}{\partial x_i x_j}(\vec{a}) = \frac{\partial^2 f}{\partial x_j x_i}(\vec{a}), \quad \forall 1{\leq}i,j{\leq}n.$$

Autrement dit, la matrice hessienne en $\vec{a}$ définie par :

$$H(\vec{a}) = [H_{ij}(\vec{a})]_{1\leq i,j\leq n} = [\frac{\partial^2 f}{\partial x_i x_j}(\vec{a})]_{1\leq i,j\leq n} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\vec{a}) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\vec{a}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\vec{a}) \end{pmatrix}$$

est symétrique en $\vec{a}$ (le hessien est le déterminant de la matrice hessienne).

Preuve. Il suffit de faire la démonstration dans $\mathbb{R}^2$ et prendre $x_i{=}x_1$ et $x_j{=}x_2$ (les autres coordonnées restent fixées et jouent le rôle de paramètres).

Notons donc $f : (x,y) \in \mathbb{R}^2 \to f(x,y) \in \mathbb{R}$.

On souhaite montrer que $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \dfrac{\partial^2 f}{\partial y \partial x}(x,y)$ aux points (x,y) où $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ sont continues. Un calcul direct donne :

$$\begin{aligned}\frac{\partial^2 f}{\partial x \partial y}(x,y) \overset{\text{déf}}{=} \frac{\partial \frac{\partial f}{\partial y}}{\partial x}(x,y) &= \frac{1}{h}\Big(\frac{\partial f}{\partial y}(x{+}h,y) - \frac{\partial f}{\partial y}(x,y)\Big) + o(1) \\ &= \frac{1}{h}\Big(\frac{1}{k}\big(f(x{+}h,y{+}k) - f(x{+}h,y)\big){+}o(1) - \frac{1}{k}\big(f(x,y{+}k) - f(x,y)\big){+}o(1)\Big) + o(1)\end{aligned}$$

Mais on est coincé : en développant, on a un terme en $o(\dfrac{1}{h})$ qui n'est pas contrôlé quand $h \to 0$. Il faut donc modifier cette approche de manière à ne pas avoir de terme en $o(1)$ relatif à k.

L'idée est de partir à l'envers : à (x,y) fixé, on regarde vers quoi tend l'accroissement :

$$Q(h,k) = \frac{1}{h}\Big(\frac{1}{k}\big(f(x{+}h,y{+}k) - f(x{+}h,y)\big) - \frac{1}{k}\big(f(x,y{+}k) - f(x,y)\big)\Big)$$

(donc, sans les termes en $o(1)$) quand $h,k \to 0$: on va montrer que $Q : (h,k) \to Q(h,k)$ est continue en 0 et que $Q(0,0) = \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$.

On pose $g_{y,k}(u) = f(u,y{+}k) - f(u,y)$ d'où :

$$Q(h,k) = \frac{1}{hk}(g_{y,k}(x+h) - g_{y,k}(x)) = \frac{1}{k}\frac{g_{y,k}(x+h) - g_{y,k}(x)}{h},$$

et le théorème de accroissements finis donne :

$$\exists \theta \in]0,1[, \quad Q(h,k) = \frac{1}{k}g'_{y,k}(x{+}\theta h) = \frac{1}{k}(\frac{\partial f}{\partial x}(x{+}\theta h, y{+}k) - \frac{\partial f}{\partial x}(x+\theta h, y)).$$

On applique encore une fois le théorème des accroissements finis pour obtenir :

$$\exists \theta \in]0,1[, \quad \exists \eta \in]0,1[, \quad Q(h,k) = \frac{\partial^2 f}{\partial y \partial x}(x{+}\theta h, y{+}\eta k)$$

La dérivée partielle $\frac{\partial^2 f}{\partial y \partial x}$ étant supposée continue en $\vec{x}$, on en déduit que Q est continue en 0 et $Q(0,0) = \frac{\partial^2 f}{\partial y \partial x}(x,y)$.

Puis on remarque que :

$$Q(h,k)=\frac{1}{k}\Big(\frac{1}{h}\big(f(x{+}h,y{+}k)-f(x,y{+}k)\big)-\frac{1}{h}\big(f(x{+}h,y)-f(x,y)\big)\Big),$$

expression symétrique de la précédente. D'où un calcul similaire au précédent donne $Q(0,0)=\frac{\partial^2 f}{\partial x\partial y}(x,y)$. D'où le théorème. ▪

Exemple 2.36 Soit $f(x,y)=\cos(xy)$. Vérifiez que $\frac{\partial^2 f}{\partial x\partial y}(\vec a){=}{-}a_1a_2\cos(a_1a_2){-}\sin(a_1a_2){=}\frac{\partial^2 f}{\partial y\partial x}(\vec a)$ en tout point $\vec a=(a_1,a_2)\in\mathbb{R}^2$. ▪

Exemple 2.37 Soit $f(x,y)=xy\dfrac{x^2-y^2}{x^2+y^2}$ et $f(0,0)=0$. Vérifier que f est continue en 0, calculer $\frac{\partial f}{\partial x}(x,y)$ et $\frac{\partial f}{\partial y}(x,y)$. Ces dérivées partielles sont-elles continues en $(0,0)$?

Puis montrer que $\frac{\partial^2 f}{\partial x\partial y}(x,0){=}1$ et $\frac{\partial^2 f}{\partial y\partial x}(0,y){=}{-}1$. En déduire que f n'est pas dans $C^2(\mathbb{R}^2;\mathbb{R})$.

Réponse. La continuité de f ne pose pas de problème. Par exemple, $|x^2-y^2|\le x^2+y^2$, d'où $0\le|f(x,y)|\le|xy|$ qui tend bien vers 0 avec (x,y).

Regardons si les dérivées partielles en $\vec 0$ ont un sens. On a :

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

d'où $\frac{\partial f}{\partial x}(\vec 0)=0$ et $\frac{\partial f}{\partial y}(\vec 0)=0$ sont bien définis (les limites quand $h\to 0$ existent et valent 0).

Puis pour $(x,y)\ne 0$, ayant $f(x,y)=\frac{x^3y-xy^3}{x^2+y^2}$:

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

et donc :

$$\frac{\partial f}{\partial x}(r\cos\theta,r\sin\theta)=O(r)\underset{r\to0}{\longrightarrow}0.$$

D'où la continuité de $\frac{\partial f}{\partial x}$ dans $\mathbb{R}^2$. Idem pour $\frac{\partial f}{\partial y}$.

Puis on a $\frac{\partial f}{\partial x}(0,y)=-y$ pour $y\ne0$, d'où $\lim_{k\to0}\frac{\frac{\partial f}{\partial x}(0,k)-\frac{\partial f}{\partial x}(0,0)}{k}=-1=\frac{\partial^2 f}{\partial y\partial x}(0,y)$ pour $y\ne0$, et si on prolongeait par continuité en 0 on obtiendrait $\frac{\partial^2 f}{\partial y\partial x}(0,0)=-1$. De même $\frac{\partial f}{\partial y}(x,0)=x$, puis $\frac{\partial^2 f}{\partial x\partial y}(x,0){=}1$ pour $x\ne0$, et $\frac{\partial^2 f}{\partial x\partial y}(0,0)=1$. D'où $\frac{\partial^2 f}{\partial x\partial y}(0,0)\ne\frac{\partial^2 f}{\partial y\partial x}(0,0)$ et f n'est pas C^2 en $(0,0)$. ▪

Remarque 2.38 Attention, il n'est pas suffisant de montrer que les fonctions, à h fixé $Q_h:k\to Q(h,k)$ et à k fixé $Q_k:h\to Q(h,k)$ sont continues en 0 pour avoir la continuité de Q en 0.

Voici par exemple une mauvaise démonstration du théorème :

Le théorème des accroissements finis nous indique que, à k fixé, il existe $\beta_k\in]0,1[$ et $\alpha_k\in]0,1[$ tels que :

$$\begin{cases}\dfrac{1}{k}\big(f(x{+}h,y{+}k)-f(x{+}h,y)\big){=}\dfrac{\partial f}{\partial y}(x{+}h,y{+}\beta_k k),\quad\text{et}\\[2mm] \dfrac{1}{k}\big(f(x,y{+}k)-f(x,y)\big){=}\dfrac{\partial f}{\partial y}(x,y{+}\alpha_k k).\end{cases}$$

On en déduit :

$$Q(h,k)=\frac{1}{h}\Big(\frac{\partial f}{\partial y}(x{+}h,y{+}\beta_k k)-\frac{\partial f}{\partial y}(x,y{+}\alpha_k k)\Big).$$

Mais $\frac{\partial f}{\partial y}$ est supposée continue dans un voisinage de (x,y), on en déduit que (quitte à prendre h assez petit), à $h{\ne}0$ fixé, la fonction $k\to Q(h,k)$ est continue au voisinage de $k=0$. Ayant $\beta_k k\to0$ et $\alpha_k k\to0$ quand $k\to0$, on obtient :

$$Q(h,k)\underset{k\to0}{\longrightarrow}Q(h,0)=\frac{1}{h}\Big(\frac{\partial f}{\partial y}(x{+}h,y)-\frac{\partial f}{\partial y}(x,y)\Big)$$

Et par définition de la dérivée en x, il vient $Q(h,0)=\dfrac{\partial\frac{\partial f}{\partial y}}{\partial x}(x,y)+o(1)$ au voisinage de $h=0$ (avec $h{\ne}0$).

D'où la fonction $h \to Q(h,0)$ est prolongeable continue au voisinage de $h=0$ par

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

Mais malheureusement, cela ne veut pas dire que Q est continue au voisinage de $(0,0)$, mais seulement que dans la direction $\vec{v}=(1,0)$ la fonction $g(t)=Q(t\vec{v})$ est prolongeable par continuité. D'ailleurs, ici le but est de montrer que si on prolonge $k \to Q(0,k)$ pour $k=0$, on obtient la même limite... ∎

Remarque 2.39 On peut démontrer qu'il suffit de supposer que $\frac{\partial^2 f}{\partial y \partial x}$ soit continue en (x,y) pour que le théorème de Schwarz soit vrai (donc sans avoir à supposer que $\frac{\partial^2 f}{\partial x \partial y}$ est continue).

En effet, l'hypothèse $\frac{\partial f}{\partial y}$ continue permet d'avoir :

$$Q(h,k) \underset{k\to 0}{\longrightarrow} \frac{1}{h}\Big(\frac{\partial f}{\partial y}(x+h,y) - \frac{\partial f}{\partial y}(x,y)\Big) \tag{2.32}$$

Et l'hypothèse d'existence de $\frac{\partial^2 f}{\partial y \partial x}$ permet d'avoir :

$$\exists \theta \in]0,1[,\ \exists \eta \in]0,1[,\ Q(h,k) = \frac{\partial^2 f}{\partial y \partial x}(x+\theta h, y+\eta k)$$

Et $\frac{\partial^2 f}{\partial y \partial x}$ étant continue, à ε fixé, dès que h et k sont assez petit, on a :

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) - \varepsilon < \frac{\partial^2 f}{\partial y \partial x}(x+\theta h, y+\eta k) = Q(h,k) < \frac{\partial^2 f}{\partial y \partial x}(x,y) + \varepsilon$$

D'où avec (2.32) :

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) - \varepsilon \le \frac{1}{h}\Big(\frac{\partial f}{\partial y}(x+h,y) - \frac{\partial f}{\partial y}(x,y)\Big) \le \frac{\partial^2 f}{\partial y \partial x}(x,y) + \varepsilon.$$

Et quand $h \to 0$ on a $\frac{1}{h}\Big(\frac{\partial f}{\partial y}(x+h,y) - \frac{\partial f}{\partial y}(x,y)\Big) \to \frac{\partial^2 f}{\partial x \partial y}(x,y)$. D'où l'égalité des dérivées partielles. ∎

2.7 * Lever les ambiguïtés des notations : introduction de la notation ∂_i

On a un problème de notations pour la dérivée de fonctions composées.

2.7.1 Premier problème à résoudre

Soit $f : (x,y) \to f(x,y) \in \mathbb{R}$ et $g(x,y) \to g(x,y) \in \mathbb{R}$ deux fonctions dans $C^1(\mathbb{R}^2;\mathbb{R})$ telles que :

$$f(x,y) = g(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}). \tag{2.33}$$

(Formule de changement de base orthonormée $(\vec{e}_1, \vec{e}_2)$ transformée en $\frac{1}{\sqrt{2}}(\vec{e}_1+\vec{e}_2) + \frac{1}{\sqrt{2}}(\vec{e}_1-\vec{e}_2)$.)

Question : calculer $\frac{\partial f}{\partial x}$ en fonction de $\frac{\partial g}{\partial x}$ et de $\frac{\partial g}{\partial y}$.

Réponse 1. On suppose ici que $\frac{\partial g}{\partial x}$ (resp. $\frac{\partial g}{\partial y}$) désigne la dérivée par rapport à la première variable de g (resp. à la seconde variable de g). Dans ce cas :

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \frac{\partial g}{\partial x}(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}})\frac{\partial \frac{x+y}{\sqrt{2}}}{\partial x}(x,y) + \frac{\partial g}{\partial y}(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}})\frac{\partial \frac{x-y}{\sqrt{2}}}{\partial x}(x,y) \\ &= \frac{1}{\sqrt{2}}\frac{\partial g}{\partial x}(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}) + \frac{1}{\sqrt{2}}\frac{\partial g}{\partial y}(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}).\end{aligned} \tag{2.34}$$

Remarque : au vu de la question on a été obligé de supposé que la notation $\frac{\partial g}{\partial x}$ était la dérivation par rapport à la première variable de g. Mais suivant le contexte du problème, cette supposition n'est pas nécessairement pertinente.

Réponse 2. (Souvent utilisée en physique.) On pose $X(x,y) = \frac{x+y}{\sqrt{2}}$ et $Y(x,y) = \frac{x-y}{\sqrt{2}}$. Donc $f(x,y) = g(X(x,y), Y(x,y))$, et en “respectant” le nom des variables :

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \frac{\partial g}{\partial X}(X(x,y),Y(x,y))\frac{\partial X}{\partial x}(x,y) + \frac{\partial g}{\partial y}(X(x,y),Y(x,y))\frac{\partial Y}{\partial x}(x,y)\\ &= \frac{1}{\sqrt{2}}\frac{\partial g}{\partial X}(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}) + \frac{1}{\sqrt{2}}\frac{\partial g}{\partial Y}(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}).\end{aligned}$$

Réponse 3. Paragraphe suivant.

2.7.2 Second problème à résoudre.

Exercice 2.40 Soit $f : (x,y) \to f(x,y) \in \mathbb{R}$ et $h(x,y) \to h(x,y) \in \mathbb{R}$ deux fonctions dans $C^2(\mathbb{R}^2;\mathbb{R})$ telles que :

$$h(x,y) = f(y,x) \tag{2.35}$$

Exprimer $\frac{\partial^2 h}{\partial x \partial y}$ à l'aide de $\frac{\partial^2 f}{\partial x \partial y}$, et vérifier le résultat quand :

$$f(x,y) = x^3y^2, \qquad \text{et donc} \qquad h(x,y) = y^3x^2 = x^2y^3. \tag{2.36}$$

Réponse. 1. Calcul générique : $\frac{\partial h}{\partial x}(x,y) = \frac{\partial f}{\partial x}(y,x)$, donc $\frac{\partial^2 h}{\partial y \partial x}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(y,x)$.

Vérification : $h(x,y) = x^2y^3$ donne $\frac{\partial h}{\partial x}(x,y) = 2xy^3$, d'où $\frac{\partial^2 h}{\partial y \partial x}(x,y) = 6xy^2$.

Puis : $f(y,x) = y^3x^2 = x^2y^3$ donne $\frac{\partial f}{\partial x}(y,x) = 2y^3x$ d'où $\frac{\partial^2 f}{\partial y \partial x}(y,x) = 6y^2x$.

Il faut donc faire attention au nom des variables, et cet exemple est simple...

Réponse. 2. (Souvent utilisée en physique.) Soit $X(x,y) = y$ et $Y(x,y) = x$ les fonctions inversant le nom des variables. On a $f(x,y) = h(X(x,y), Y(x,y))$.

En utilisant le nom des variables pour les dérivations, la dérivation dans la direction x s'écrit :

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \frac{\partial h}{\partial X}(X(x,y),Y(x,y))\frac{\partial X}{\partial x}(x,y) + \frac{\partial h}{\partial Y}(X(x,y),Y(x,y))\frac{\partial Y}{\partial x}(x,y)\\ &= 0 + \frac{\partial h}{\partial Y}(X(x,y),Y(x,y)) = \frac{\partial h}{\partial Y}(y,x),\end{aligned}$$

D'où en dérivant dans la direction y :

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x}(x,y) &= \frac{\partial^2 h}{\partial X \partial Y}(X(x,y),Y(x,y))\frac{\partial X}{\partial y}(x,y) + \frac{\partial^2 h}{\partial Y \partial Y}(X(x,y),Y(x,y))\frac{\partial Y}{\partial y}(x,y)\\ &= \frac{\partial^2 h}{\partial X \partial Y}(X(x,y),Y(x,y)) + 0 = \frac{\partial^2 h}{\partial X \partial Y}(y,x).\end{aligned}$$

Réponse. 3. Voir paragraphe suivant. ▪

2.7.3 Lever les ambiguïtés : notation ∂_i

Pour lever les ambiguïtés, on décide de ne pas utiliser les noms de variables pour les dérivées partielles : par exemple on n'utilise pas $\frac{\partial}{\partial x}$ mais ∂_1 la dérivée par rapport à la première variable.

Ainsi : soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction différentiable en $\vec{x}$. Le développement limité (2.7) s'écrit :

$$\varphi(\vec{x}) = \varphi(\vec{x}_0) + \partial_1 f(\vec{x})(x_1 - x_{01}) + ... + \partial_1 f(\vec{x})(x_n - x_{0n}) + o(||\vec{x} - \vec{x}_0||), \tag{2.37}$$

de même que :

$$\varphi(\vec{y}) = \varphi(\vec{y}_0) + \partial_1 f(\vec{y})(x_1 - x_{01}) + ... + \partial_1 f(\vec{y})(x_n - x_{0n}) + o(||\vec{y} - \vec{y}_0||). \tag{2.38}$$

La notation ∂_1 n'est pas attachée au nom de la variable, ici x ou y (contrairement à $\partial_1 =^{\text{noté}} \frac{\partial f}{\partial x}$ par exemple).

Réponse 3 du paragraphe 2.7.1. On reformule la question comme : relativement à (2.33), calculer $\partial_1 f$ en fonction de $\partial_1 g$ et de $\partial_2 g$. Par dérivation de fonctions composées on a :

$$\begin{aligned}\partial_1 f(x,y) &= \partial_1 g(X(x,y),Y(x,y))\partial_1 X(x,y) + \partial_2 g(X(x,y),Y(x,y))\partial_1 Y(x,y)\\ &= \frac{1}{\sqrt{2}}\partial_1 g(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}) + \frac{1}{\sqrt{2}}\partial_2 g(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}).\end{aligned}$$

Exercice 2.41 Suite du paragraphe 2.7.1, exercice (2.40).

Réponse. 3.

Sans utiliser le nom des variables pour les dérivations, la dérivation dans la direction 1 s'écrit, avec $X(x,y)=y$ et $Y(x,y)=x$:

$$\begin{aligned}\partial_1 f(x,y) &= \partial_1 h(X(x,y),Y(x,y))\partial_1 X(x,y)+\partial_2 h(X(x,y),Y(x,y))\partial_1 Y(x,y)\\ &= 0+\partial_2 h(X(x,y),Y(x,y))=\partial_2 h(y,x).\end{aligned}$$

D'où, en dérivant dans la direction 2 :

$$\begin{aligned}\partial_{21} f(x,y) &= \partial_{12} h(X(x,y),Y(x,y))\partial_2 X(x,y)+\partial_{22} h(X(x,y),Y(x,y))\partial_2 Y(x,y)\\ &= \partial_{12} h(X(x,y),Y(x,y))=\partial_{12} h(y,x).\end{aligned}$$

2.8 Formule de Taylor dans $\mathbb{R}^n$

On s'intéresse aux fonctions $f:\mathbb{R}^n\to\mathbb{R}$ (fonctions à valeurs scalaires).

La démarche est la même que pour la formule de Taylor dans $\mathbb{R}$. On souhaite par exemple qu'un polynôme :

$$p(x,y)=a+bx+cy+dx^2+exy+fy^2+...$$

soit donné par son développement au voisinage d'un point $\vec{x}=(x,y)$. L'identification des coefficients donne immédiatement :

$$\begin{aligned}p(\vec{x}) &= p(\vec{0})+\frac{\partial p}{\partial x}(\vec{0})\,x+\frac{\partial p}{\partial y}(\vec{0})\,y+\frac{1}{2!}\Big(\frac{\partial^2 p}{\partial x^2}(\vec{0})\,x^2+2\frac{\partial^2 p}{\partial x\partial y}(\vec{0})\,xy+\frac{\partial^2 p}{\partial y^2}(\vec{0})\,y^2\Big)\\ &\quad+\frac{1}{3!}\Big(\frac{\partial^3 p}{\partial x^3}(\vec{0})\,x^3+3\frac{\partial^3 p}{\partial x^2\partial y}(\vec{0})\,x^2y+2\frac{\partial^3 p}{\partial x\partial y^2}(\vec{0})\,xy^2+\frac{\partial^3 p}{\partial y^3}(\vec{0})\,y^3\Big)+...\\ &= p(\vec{0})+\sum_{1\le i\le 2}\frac{\partial p}{\partial x_i}(\vec{0})\,x_i+\frac{1}{2!}\sum_{k=0}^{2}C_2^k x_1^k x_2^{2-k}\frac{\partial^2 p}{\partial x_1^k\partial x_2^{2-k}}(\vec{0})\\ &\quad+\frac{1}{3!}\sum_{k=0}^{3}C_3^k x_1^k x_2^{3-k}\frac{\partial^3 p}{\partial x_1^k\partial x_2^{3-k}}(\vec{0})+...\end{aligned}$$

On le démontre directement : $p(\vec{0})=a$, puis $\frac{\partial p}{\partial x}(\vec{0})=b$..., puis $\frac{\partial p^2}{\partial x^2}(\vec{0})=2d$,...

Remarque 2.42 Notation abrégée :

$$p(\vec{x})=p(\vec{0})+\frac{1}{1!}dp(\vec{0}).\vec{x}+\frac{1}{2!}d^2p(\vec{0})(\vec{x},\vec{x})+\frac{1}{3!}d^3p(\vec{0})(\vec{x},\vec{x},\vec{x})+... \tag{2.39}$$

On se limitera dans ce cours au développement au second ordre. Ce qui s'écrit alors, avec H_p la matrice Hessienne de p :

$$p(\vec{x})=p(\vec{0})+\vec{\mathrm{grad}}p(\vec{0})^t.\vec{x}+\frac{1}{2}\vec{x}^t.H_p(\vec{0}).\vec{x}+o(||\vec{x}||^2)$$

Et on souhaite qu'une fonction quelconque soit bien approximé par un tel polynôme. On a :

Théorème 2.43 *(Formule de Taylor) Si $f:\mathbb{R}^n\to\mathbb{R}$ est dérivable k+1-fois au voisinage de $\vec{0}$, alors (avec H_f la matrice hessienne de f) :*

$$\begin{aligned}f(\vec{x}) &= f(\vec{0})+\vec{\mathrm{grad}}f(\vec{0})^t.\vec{x}+\frac{1}{2}\vec{x}^t.H_f(\vec{0}).\vec{x}+\dots\\ &\quad+\sum_{\substack{i_1,\dots,i_n\in\mathbb{N}\\ i_1+\dots+i_n=k}}\frac{x_1^{i_1}\dots x_n^{i_n}}{i_1!\dots i_n!}\frac{\partial^k f}{\partial x_{i_1}^{i_1}\dots\partial x_{i_n}^{i_n}}(\vec{0})\;+R_k(\vec{x},\vec{0})\end{aligned} \tag{2.40}$$

où :

$$R_k(\vec{x},\vec{0})=\frac{1}{k!}\int_0^1 g^{k+1}(t)\,(1-t)^k\,dt=O(||\vec{x}||^{k+1})=o(||\vec{x}||^k), \tag{2.41}$$

où on a posé $g(t)=f(t\vec{x})$.

Preuve. Soit $g(t) = f(t\vec{x}) = f(tx_1, \dots, tx_n)$ pour $t \in [0,1]$, i.e. on se place sur le segment $[\vec{0}, \vec{x}]$. Et on a $g'(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(t\vec{x})x_i$, puis $g''(t) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(t\vec{x})x_i x_j$, puis... Et on applique la formule de Taylor avec reste intégrale (1.45) à la fonction g entre 0 et 1, i.e.

$g(1) = g(0) + \dots + \frac{t^k}{k!}g^{(k)}(0) + \frac{1}{k!}\int_{u=0}^1 g^{(k+1)}(u)(1-u)^k\,du.$

D'où $g(1) = f(\vec{x}) =$ le résultat annoncé. ∎

Au voisinage d'un point $\vec{a}$:

Corollaire 2.44 *(Formule de Taylor) Si f est dérivable k+1-fois au voisinage de $\vec{a}$, alors :*

$$\begin{aligned} f(\vec{x}) = f(\vec{a}) &+ \vec{\operatorname{grad}} f(\vec{a})^t.(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^t.H_f(\vec{a}).(\vec{x}-\vec{a}) + \dots \\ &+ \frac{1}{k!}\sum_{\substack{i_1,\dots,i_n \in \mathbb{N} \\ i_1+\dots+i_n=k}} \frac{k!}{i_1!\dots i_n!}\frac{\partial^k f}{\partial x_{i_1}\dots\partial x_{i_n}}(\vec{a})\,(x_{i_1}-a_{i_1})\dots(x_{i_n}-a_{i_n}) + R_k(\vec{x},\vec{a}) \end{aligned} \tag{2.42}$$

où :

$$R_k(\vec{x},\vec{a}) = \frac{1}{k!}\int_0^1 g^{(k+1)}(t)\;(1-t)^k\,dt = O(||\vec{x}-\vec{a}||^{k+1} = o(||\vec{x}-\vec{a}||^k), \tag{2.43}$$

où on a posé $g(t) = f(\vec{a}+t(\vec{x}-\vec{a}))$.

Preuve. Soit $g(t) = f(\vec{a} + t(\vec{x}-\vec{a}))$ pour $t \in [-1,1]$, i.e. on se place sur la droite passant par $\vec{x}$ et $\vec{a}$. Et on applique la formule de Taylor avec reste intégrale (1.44) à la fonction g. ∎

Cas particulier du développement à l'ordre 2 d'une fonction supposée C^3, notant H la matrice Hessienne de f (symétrique car f est C^2, théorème de Schwarz 2.35) :

$$f(\vec{x}) = f(\vec{a}) + \vec{\operatorname{grad}} f(\vec{a}).(\vec{x}-\vec{a}) + \frac{1}{2}(\vec{x}-\vec{a})^t.H_f(\vec{a}).(\vec{x}-\vec{a}) + 0(||\vec{x}-\vec{a}||^3).$$

Exercice 2.45 Faire la démonstration de la formule de Taylor (2.42) à l'aide de la formule de Taylor (2.40) et de la translation $\vec{x} \to \vec{x}-\vec{a}$ (changement d'origine).

Réponse. On fait un changement d'origine. On pose $f(\vec{x}) = h(\vec{x}-\vec{a})$, ce qui fait que $f(\vec{a}) = h(\vec{0})$ et qu'on peut se servir du développement de Taylor de h. On a donc $h(\vec{x}) = f(\vec{x}+\vec{a}) = f(\vec{y}) = f(\vec{y}(\vec{x}))$ où $\vec{y}(\vec{x}) = \vec{x} + \vec{a} = \begin{pmatrix} y_1 = x_1+a_1 \\ y_2 = x_2+a_2 \end{pmatrix}$, et le développement de h en $\vec{0}$ est donné par (2.40), d'où celui de f en $\vec{a}$, sachant que par exemple :

$$\frac{\partial h}{\partial x_1}(\vec{x}) = \frac{\partial f}{\partial y_1}(\vec{y}(\vec{x})).\frac{\partial y_1}{\partial x_1}(\vec{x}) + \frac{\partial f}{\partial y_2}(\vec{y}(\vec{x})).\frac{\partial y_2}{\partial x_1}(\vec{x}) = \frac{\partial f}{\partial y_1}(\vec{x}+\vec{a}).1 + \frac{\partial f}{\partial y_2}(\vec{x}+\vec{a}).0 = \frac{\partial f}{\partial y_1}(\vec{x}+\vec{a})$$

Et donc $\frac{\partial h}{\partial x_1}(\vec{0}) = \frac{\partial f}{\partial y_1}(\vec{a})$. Et on fait de même pour les autres dérivées de f. Ainsi, de la formule de Taylor au 1er ordre :

$$h(\vec{x}) = h(\vec{0}) + \frac{\partial h}{\partial x_1}(\vec{0})\,x_1 + \frac{\partial h}{\partial x_2}(\vec{0})\,x_2 + o(||\vec{x}||),$$

est réécrite comme :

$$f(\vec{y}) = f(\vec{a}) + \frac{\partial f}{\partial y_1}(\vec{a})\,(y_1-a_1) + \frac{\partial f}{\partial y_2}(\vec{a})\,(y_2-a_2) + o(||\vec{y}-\vec{a}||).$$

∎

3 Fonctions de plusieurs variables à valeurs vectorielles

3.1 Introduction et continuité

Les fonctions à valeurs vectorielles sont de type, pour Ω ouvert de $\mathbb{R}^n$:

$$\vec{f} : \begin{cases} \Omega \subset \mathbb{R}^n & \to \mathbb{R}^m \\ \vec{x} & \mapsto \vec{f}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix}, \end{cases}$$

où les m fonctions $f_i : \Omega \to \mathbb{R}$ sont à valeurs scalaires, pour $i = 1, ..., m$. Les $f_i(\vec{x})$ sont les m composantes de $\vec{f}(\vec{x})$, pour $\vec{x} \in \Omega$.

Proposition 3.1 *La fonction $\vec{f}$ est continue en un point $\vec{x}$ ssi les n fonctions f_i, $i = 1, ..., n$, sont continues en $\vec{x}$.*

Preuve. $\vec{f}$ est continue en $\vec{x}$ ssi $\lim_{\vec{v}\to 0} ||\vec{f}(\vec{x}+\vec{v}) - \vec{f}(\vec{x})|| = 0$. Et $||\vec{f}(\vec{x}+\vec{v}) - \vec{f}(\vec{x})|| = \left(\sum_{i=1}^n (f_i(\vec{x}+\vec{v}) - f_i(\vec{x}))^2\right)^{\frac{1}{2}}$. D'où le résultat. ▪

On traitera souvent le cas $m{=}2, 3$, i.e. $\mathbb{R}^m{=}\mathbb{R}^2$ ou $\mathbb{R}^3$ dans ce cours.

Exemple 3.2 Soit une fonction à valeurs vectorielles :

$$\vec{f} : \begin{cases} \qquad\qquad \mathbb{R}^2 & \to \mathbb{R}^2 \\ \vec{x} = (x_1, x_2) & \mapsto \vec{f}(\vec{x}) = \vec{f}(x_1, x_2) = \begin{pmatrix} y_1 = f_1(\vec{x}) = f_1(x_1, x_2) = x_1 \cos x_2 \\ y_2 = f_2(\vec{x}) = f_2(x_1, x_2) = x_1 \sin x_2 \end{pmatrix}, \end{cases}$$

où donc f_1 et f_2 sont deux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$ (à valeurs scalaires). Et il est immédiat que f_1 et f_2 sont continues, et donc $\vec{f}$ est continue.

Pour cette fonction $\vec{f}$, les notations usuelles sont données par $x_1 = r$ et $x_2 = \theta$ lorsque $x_1 > 0$ et $x_2 \in \mathbb{R}$, et $y_1 = x$ et $y_2 = y$. Et la fonction $\vec{f}$ ci-dessus exprime une position (x, y) en fonction des coordonnées polaires (r, θ). ▪

Notation de Landau. On a avec la notation de Landau, si $\vec{f}$ est continue en $\vec{x}$:

$$\vec{f}(\vec{y}) - \vec{f}(\vec{x}) = o(1) \quad \text{au voisinage de } \vec{x},$$

au sens où $f_i(\vec{y}) - f_i(\vec{x}) = o(1)$ pour tout $i = 1, ..., n$.

3.2 Dérivations et matrice jacobienne

On a vu qu'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ (à valeurs scalaires) est dérivable en $\vec{a} \in \mathbb{R}^n$ s'il existe une application linéaire $\ell_{\vec{a}} : \vec{x} = (x_1, ..., x_n) \in \mathbb{R}^n \to \ell_{\vec{a}}(\vec{x}) = \ell_{\vec{a}}(x_1, ..., x_n) = A_1 x_1 + ... + A_n x_n$ telle que, voir (2.5) :

$$f(x_1, ..., x_n) = f(a_1, , ..., a_n) + \ell_{\vec{a}}(x_1{-}a_1, ..., x_n{-}a_n) + o(||\vec{x}{-}\vec{a}||).$$

Et l'application linéaire $\ell_{\vec{a}}$ est appelée forme différentielle de f en $\vec{a}$, et notée $\ell_{\vec{a}} = df(\vec{a})$ (localement au voisinage de $\vec{a}$, f est ainsi approximée par la forme linéaire $\ell_{\vec{a}}$).

D'où :

Définition 3.3 On dit qu'une fonction $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ (à valeurs vectorielles) est dérivable en $\vec{a} \in \mathbb{R}^n$ s'il existe une application linéaire $L_{\vec{a}} : \mathbb{R}^n \to \mathbb{R}^m$ telle que :

$$\vec{f}(\vec{x}) = \vec{f}(\vec{a}) + L_{\vec{a}}(\vec{x} - \vec{a}) + o(||\vec{x}{-}\vec{a}||) \quad \in \mathbb{R}^m. \tag{3.1}$$

Et l'application linéaire $L_{\vec{a}}$ est appelée la différentielle de $\vec{f}$ en $\vec{a}$, et notée $L_{\vec{a}} = d\vec{f}(\vec{a})$.

Ainsi, localement au voisinage de $\vec{a}$, on peut commencer l'étude de $\vec{f}$ par celle de l'application linéaire $L_{\vec{a}}$ (dans un premier temps en tout cas).

Proposition 3.4 *L'application $\vec{f}$ est dérivable en $\vec{a}$ ssi les applications composantes f_i sont dérivables en $\vec{a}$, et on a :*

$$d\vec{f}(\vec{a})(\vec{v}) = \begin{pmatrix} df_1(\vec{a})(\vec{v}) \\ \vdots \\ df_m(\vec{a})(\vec{v}) \end{pmatrix},$$

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

Preuve. Immédiat : écrire (3.1) comme m lignes (composantes dans $\mathbb{R}^m$), et poser $\vec{x} = \vec{v} + \vec{a}$. ▪

Proposition 3.5 *On prend les bases canoniques de $\mathbb{R}^n$ et de $\mathbb{R}^m$. Alors l'application linéaire $d\vec{f}(\vec{a})$ peut-être représentée par une matrice $[d\vec{f}(\vec{a})]$, appelée matrice jacobienne de $\vec{f}$ en $\vec{a}$.*

La matrice $[d\vec{f}(\vec{a})]$ a sa i-ème ligne donnée par les $(\frac{\partial f_i}{\partial x_j}(\vec{a}))_{j=1,\dots,n}$, et a sa j-ème colonne donnée par les coordonnées de $d\vec{f}(\vec{a})(\vec{e}_j) = \frac{\partial \vec{f}}{\partial x_j}(\vec{a})$ dans la base canonique de $\mathbb{R}^m$:

$$[d\vec{f}(\vec{a})] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{pmatrix} = \begin{pmatrix} (\vec{\text{grad}} f_1(\vec{a})^T) \\ \vdots \\ (\vec{\text{grad}} f_n(\vec{a})^T) \end{pmatrix}. \tag{3.2}$$

Preuve. On a effectivement $df_i(\vec{a})(\vec{v}) = \vec{\text{grad}} f_i(\vec{a})^T.\vec{v}$, pour tout $\vec{v} \in \mathbb{R}^n$ et tout $i = 1, ..., n$. Et la ligne i est donnée par les n termes $\frac{\partial f_i}{\partial x_j}(\vec{a})$ pour $j = 1, ..., n$.

Puis $d\vec{f}(\vec{a})(\vec{e}_j) = \begin{pmatrix} df_1(\vec{a})(\vec{e}_j) \\ \vdots \\ df_m(\vec{a})(\vec{e}_j) \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_j}(\vec{a}) \\ \vdots \\ \frac{\partial f_m}{\partial x_j}(\vec{a}) \end{pmatrix} = \frac{\partial \vec{f}}{\partial x_j}(\vec{a})$, d'où l'interprétation des colonnes. ▪

On a donc, pour la représentation dans les bases canoniques, avec les notations du calcul matriciel :

$$\vec{f}(\vec{x}) = \vec{f}(\vec{a}) + d\vec{f}(\vec{a}).(\vec{x} - \vec{a}) + o(||\vec{x} - \vec{a}||). \tag{3.3}$$

Exercice 3.6 Soit $\vec{f}(x, y) = \begin{pmatrix} x \\ y^2 \\ xy^2 \end{pmatrix}$ définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}^3$. Écrire la matrice jacobienne de $\vec{f}$ en un point $\vec{x}$ (matrice $3 * 2$). ▪

Exercice 3.7 Soit $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$ définie sur $\mathbb{R}^2 - \{\vec{0}\}$ à valeurs dans $\mathbb{R}$. Écrire la matrice jacobienne de f en un point $\vec{x} \neq \vec{0}$ (matrice $1 * 2$). ▪

Exercice 3.8 Soit $\vec{f}(\rho, \theta) = \begin{pmatrix} \rho\cos\theta \\ \rho\sin\theta \end{pmatrix}$ définie sur $U = \mathbb{R}_+^* \times]-\pi, \pi[$. Écrire la matrice jacobienne de $\vec{f}$ en un point $\vec{u} = (\rho, \theta) \in U$ (matrice $2 * 2$). ▪

Exercice 3.9 Soit $\vec{f}(\rho, \theta, \varphi) = \begin{pmatrix} \rho\cos\theta\cos\varphi \\ \rho\sin\theta\cos\varphi \\ \rho\sin\varphi \end{pmatrix}$ définie sur $U = \mathbb{R}_+^* \times]-\pi, \pi[\times]-\frac{\pi}{2}, \frac{\pi}{2}[$. Écrire la matrice jacobienne de $\vec{f}$ en un point $\vec{u} = (\rho, \theta, \varphi) \in U$ (matrice $3 * 3$). ▪

Exercice 3.10 Montrer que la matrice jacobienne de l'application $I : \vec{x} \in \mathbb{R}^n \to \vec{x} \in \mathbb{R}^n$ est la matrice identité. Et montrer que toute application linéaire $L : \mathbb{R}^n \to \mathbb{R}^m$ a pour matrice jacobienne la matrice constante $[L]$ représentant L dans la base canonique.

Réponse. L'application identité est un cas particulier d'une application linéaire. Soit $L : \mathbb{R}^n \to \mathbb{R}^m$ une application linéaire donnée. Soient L_i pour $i = 1, ..., m$ ses composantes, i.e. pour $\vec{x} \in \mathbb{R}^n$ on pose $L(\vec{x}) = \begin{pmatrix} L_1(\vec{x}) \\ \vdots \\ L_n(\vec{x}) \end{pmatrix}$. Chaque L_i est linéaire, i.e. de la forme $L_i(\vec{x}) = a_{i1}x_1 + \ldots + a_{in}x_n$, d'où $L(\vec{x}) = A.\vec{x}$ où $A = [a_{ij}]_{\substack{1 \le i \le m \\ 1 \le j \le n}} = [L]$. Comme L est linéaire, on a $L(\vec{y}) = L(\vec{x}) + L(\vec{y} - \vec{x}) = L(\vec{x}) + L(\vec{y} - \vec{x}) + o(||\vec{y} - \vec{x}||)$. Donc la différentielle de L en $\vec{x}$ est l'application linéaire $dL(\vec{x}) = L$, et la matrice jacobienne de L en $\vec{x}$ est la matrice de $dL(\vec{x})$ i.e. est la matrice $[L] = A$ (indépendante du point $\vec{x}$). ▪

3.3 Composition et dérivations : produit des jacobiennes

On considère une fonction $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ à valeurs vectorielles :

$$\vec{f} : \begin{cases} \mathbb{R}^n & \to \mathbb{R}^m \\ \vec{x} = (x_1, ..., x_m) & \mapsto \vec{f}(\vec{x}) = \vec{f}(x_1, ..., x_m) = \begin{pmatrix} f_1(x_1, ..., x_m) \\ \vdots \\ f_m(x_1, ..., x_m) \end{pmatrix}, \end{cases}$$

et soit une fonction $\vec{g}$ à valeurs vectorielles :

$$\vec{g} : \begin{cases} \mathbb{R}^m \to \mathbb{R}^p \\ \vec{y} = (y_1, ..., y_m) \mapsto \vec{g}(\vec{y}) = \vec{g}(y_1, ..., y_m) = \begin{pmatrix} g_1(x_1, ..., x_m) \\ \vdots \\ g_p(x_1, ..., x_m) \end{pmatrix}. \end{cases}$$

On considère alors la fonction composée :

$$\vec{g} \circ \vec{f} : \begin{cases} \mathbb{R}^n \to \mathbb{R}^p \\ \vec{x} = (x_1, ..., x_n) \mapsto \vec{g}(\vec{y}) = \vec{g}(\vec{f}(\vec{x})) = \begin{pmatrix} (g_1 \circ \vec{f})(\vec{x})) \\ \vdots \\ (g_p \circ \vec{f})(\vec{x})) \end{pmatrix} = \begin{pmatrix} g_1(f_1(\vec{x}), ..., f_m(\vec{x})) \\ \vdots \\ g_p(f_1(\vec{x}), ..., f_m(\vec{x})) \end{pmatrix}. \end{cases}$$

On a alors :

Théorème 3.11 *Si $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ est différentiable en $\vec{a} \in \mathbb{R}^n$ et si $g : \mathbb{R}^m \to \mathbb{R}$ est différentiable en $\vec{b} = \vec{f}(\vec{a}) \in \mathbb{R}^m$, alors la fonction $(g \circ \vec{f}) : \mathbb{R}^n \to \mathbb{R}$ est différentiable en $\vec{a}$, et on a l'égalité entre applications linéaires :*

$$d(g \circ \vec{f})(\vec{a}) = dg(\vec{b}) \circ d\vec{f}(\vec{a}). \tag{3.4}$$

Et dans les bases canoniques, on obtient pour les matrices jacobiennes :

$$[d(g \circ \vec{f})(\vec{a})] = [dg(\vec{b})].[d\vec{f}(\vec{a})] \tag{3.5}$$

produit des matrices jacobiennes (aux points où cela a un sens), i.e. pour tout $j = 1, ..., n$:

$$\frac{\partial(g \circ \vec{f})}{\partial x_j}(\vec{a}) = \sum_{k=1}^{m} \frac{\partial g}{\partial y_k}(\vec{b}) \frac{\partial f_k}{\partial x_j}(\vec{a}). \tag{3.6}$$

(La dernière formule est également écrite, si on veut éviter la référence au nom des variables y_i pour $g(y_1, ..., y_n)$ et x_i pour $\vec{f}(x_1, ..., x_n)$: $\partial_j(g \circ \vec{f})(\vec{a}) = \sum_{k=1}^{m} (\partial_k g)(\vec{b})(\partial_j f_k)(\vec{a})$.)

Preuve. La démonstration est similaire à celle du théorème 1.32 : pour $\vec{x} \in \mathbb{R}^n$ on a $(g \circ \vec{f})(\vec{x})) = g(f_1(\vec{x}), ..., f_m(\vec{x}))$, et pour $n = 2 = m$ pour simplifier les écritures (cas général laissé au lecteur) :

$$\begin{aligned} \frac{\partial(g \circ \vec{f})}{\partial x_1}(\vec{x}) &= \lim_{h \to 0} \frac{(g \circ \vec{f})(\vec{x}+h\vec{e}_1) - (g \circ \vec{f})(\vec{x})}{h} \\ &= \lim_{h \to 0} \frac{g(f_1(\vec{x}+h\vec{e}_1), f_2(\vec{x}+h\vec{e}_j)) - g(f_1(\vec{x}), f_2(\vec{x}))}{h}, \end{aligned}$$

avec :

$$\begin{aligned} &\frac{g(f_1(\vec{x}+h\vec{e}_1), f_2(\vec{x}+h\vec{e}_j)) - g(f_1(\vec{x}), f_2(\vec{x}))}{h} \\ &= \frac{g(f_1(\vec{x}+h\vec{e}_1), f_2(\vec{x}+h\vec{e}_j)) - g(f_1(\vec{x}+h\vec{e}_1), f_2(\vec{x}))}{h} + \frac{g(f_1(\vec{x}+h\vec{e}_1), f_2(\vec{x})) - g(f_1(\vec{x}), f_2(\vec{x}))}{h} \end{aligned}$$

qui tend quand $h \to 0$ vers $\dfrac{\partial g}{\partial y_2}(\vec{y}) \dfrac{\partial f_2}{\partial x_1}(\vec{x}) + \dfrac{\partial g}{\partial y_1}(\vec{y}) \dfrac{\partial f_1}{\partial x_1}(\vec{x})$ où $\vec{y} = \vec{f}(\vec{x}) \in \mathbb{R}^m$ (même démarche que pour la démonstration 1.32).

Cela s'écrit aussi :

$$\begin{pmatrix} \frac{\partial(g \circ \vec{f})}{\partial x_1}(\vec{a}) & \frac{\partial(g \circ \vec{f})}{\partial x_2}(\vec{a}) \end{pmatrix} = \begin{pmatrix} \frac{\partial g}{\partial y_1}(\vec{b}) & \frac{\partial g}{\partial y_2}(\vec{b}) \end{pmatrix} . \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_1} \end{pmatrix}, \tag{3.7}$$

soit $[d(g \circ \vec{f})(\vec{a})] = [dg(\vec{b})].[d\vec{f}(\vec{a})]$. ∎

Remarque 3.12 Autre preuve : $\vec{f}(\vec{x}+h\vec{e}_i) = \vec{f}(\vec{x}) + h\, d\vec{f}(\vec{x}).\vec{e}_i + o(h)$, d'où $g(\vec{f}(\vec{x}+h\vec{e}_i)) = g(\vec{f}(\vec{x})) + dg(\vec{f}(\vec{x})).(h\, d\vec{f}(\vec{x}).\vec{e}_i + o(h)) + o(\vec{f}(\vec{x}+h\vec{e}_i) - \vec{f}(\vec{x}))$, d'où $g(\vec{f}(\vec{x}+h\vec{e}_i)) = g(\vec{f}(\vec{x})) + h\, dg(\vec{f}(\vec{x})).\frac{\partial \vec{f}}{\partial x_i}(\vec{x}).\vec{e}_i + o(h)$. ∎

Corollaire 3.13 *Si $\vec{f}:\mathbb{R}^n\to\mathbb{R}^m$ est différentiable en $\vec{a}\in\mathbb{R}^n$ et si $\vec{g}:\mathbb{R}^m\to\mathbb{R}^p$ est différentiable en $\vec{b}=\vec{f}(\vec{a})\in\mathbb{R}^m$, alors la fonction $(\vec{g}\circ\vec{f})$ est différentiable en $\vec{a}$, et on a l'égalité entre applications linéaires :*

$$d(\vec{g}\circ\vec{f})(\vec{a})=d\vec{g}(\vec{b})\circ d\vec{f}(\vec{a}). \tag{3.8}$$

Et dans les bases canoniques, on obtient pour les matrices jacobiennes :

$$[d(\vec{g}\circ\vec{f})(\vec{a})]=[d\vec{g}(\vec{b})].[d\vec{f}(\vec{a})], \tag{3.9}$$

produit des matrices jacobiennes (aux points où cela a un sens), i.e. :

$$\frac{\partial(g_i\circ\vec{f})}{\partial x_j}(\vec{a})=\sum_{k=1}^{m}\frac{\partial g_i}{\partial y_k}(\vec{b})\frac{\partial f_k}{\partial x_j}(\vec{a}), \tag{3.10}$$

pour tout $i=1,...,p$ et $j=1,...,n$.

Preuve. On s'intéresse aux p composantes $g_i\circ\vec{f}$ pour $i=1,...,p$ de $\vec{g}\circ\vec{f}$: théorème précédent, et donc en écrivant les n lignes pour $i=1,...,n$, on obtient (3.9) comme annoncé. ∎

Remarque 3.14 La formule (3.9) n'est autre que la formule de composition des applications linéaires : si $L_1:\mathbb{R}^n\to\mathbb{R}^m$ est une application linéaire, représentée dans la base canonique par la matrice $[L_1]$ de dimension $m*n$, et si $L_2:\mathbb{R}^m\to\mathbb{R}^p$ est une application linéaire, représentée dans la base canonique par la matrice $[L_2]$ de dimension $p*m$, alors $L_2\circ L_1:\mathbb{R}^n\to\mathbb{R}^p$ est une application linéaire, représentée dans la base canonique par la matrice produit $[L_2].[L_1]$ de dimension $p*n$.

Et dans le théorème précédent, $L_1=d\vec{f}(\vec{a})$ et $L_2=d\vec{g}(\vec{b})$ où $\vec{b}=\vec{f}(\vec{a})$, de matrices $[L_1]=[d\vec{f}(\vec{a})]$ et $[L_2]=[d\vec{g}(\vec{b})]$. Et on a $d(\vec{g}\circ\vec{f})(\vec{a})=d\vec{g}(\vec{b})\circ d\vec{f}(\vec{a})=L_2\circ L_1$, représenté par la matrice $[L_2].[L_1]=[d\vec{g}(\vec{b})].[d\vec{f}(\vec{a})]$. ∎

Proposition 3.15 *Si $\vec{f}:\mathbb{R}^n\to\mathbb{R}^n$ est dérivable en $\vec{a}$, inversible d'inverse $\vec{f}^{-1}$ dérivable en $\vec{b}=\vec{f}(\vec{a})$, alors $\det([d\vec{f}(\vec{a})])\neq 0$ et :*

$$[d(\vec{f}^{-1})(\vec{b})]=\Big([d\vec{f}(\vec{a})]\Big)^{-1}. \tag{3.11}$$

Preuve. On a $[d(\vec{f}^{-1}\circ\vec{f})(\vec{a})]=I$ où I est la matrice identité, i.e. $[d(\vec{f}^{-1})(\vec{b})].[d\vec{f}(\vec{a})]=I$. ∎

Exercice 3.16 Soit $\vec{\psi}(\rho,\theta)=\begin{pmatrix}\rho\cos\theta\\ \rho\sin\theta\end{pmatrix}$ définie sur $\mathbb{R}_+^*\times]-\pi,\pi[$ à valeurs dans $\mathbb{R}^2$, et soit $f(x,y)=\sqrt{x^2+y^2}$ définie sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$.

Calculer $[d\vec{\psi}]$, $[df]$, $[d(f\circ\vec{\psi})]$ après avoir calculé $f\circ\vec{\psi}$, et $[df(\vec{\psi}(\rho,\theta))].[d\vec{\psi}(\rho,\theta)]$. ∎

Exercice 3.17 Soit $F(x,y,z)=(x+y^2,xy^2z)$ et $G(u,v)=(u^2+v,uv,e^v)$. Donner les ensembles de départ et d'arrivée de F et de G, calculer dF, dG, $dF(G(u,v)).dG(u,v)$, et $d(F\circ G)(u,v)$ après avoir donné $F\circ G$. ∎

Exercice 3.18 Soit $f(x,y)=\sin(x^2-y^2)$ et soit $\vec{g}(x,y)=(x+y,x-y)$. Donner les espaces de départ et d'arrivée de f et de $\vec{g}$. Calculer $f\circ\vec{g}$, puis $d(f\circ\vec{g})$. Calculer df, puis $d\vec{g}$, puis $df(\vec{g}(\vec{x})).d\vec{g}(\vec{x})$.

Réponse. On a $f:\mathbb{R}^2\to\mathbb{R}$ et $\vec{g}:\mathbb{R}^2\to\mathbb{R}^2$. On a $f\circ\vec{g}:\mathbb{R}^2\to\mathbb{R}$ définie par $(f\circ\vec{g})(x,y)=f(\vec{g}(x,y))=f(g_1(x,y),g_2(x,y))=\sin((x+y)^2-(x-y)^2)=\sin(4xy)$.

D'où $[d(f\circ\vec{g})(\vec{x})]=(\,4y\sin(4xy)\quad,\quad 4x\sin(4xy)\,)$ (matrice ligne).

Puis on change les notations pour éviter toute ambiguïté. On note $f(X,Y)=\sin(X^2-Y^2)$, et on garde la notation $\vec{g}(x,y)=\begin{pmatrix}g_1(\vec{x})=x+y\\ g_2(\vec{x})=x-y\end{pmatrix}$. Ainsi on notera $f(\vec{g}(x,y))=f(X,Y)$ où $X=g_1(\vec{x})=x+y$ et $Y=g_2(\vec{x})=x-y$.

Calcul immédiat :
$[d\vec{g}(x,y)]=\begin{pmatrix}1&1\\1&-1\end{pmatrix}$ (matrice constante) et
$[df(X,Y)]=(\,2X\sin(X^2-Y^2)\quad,\quad -2Y\sin(X^2-Y^2)\,)$ (matrice ligne).

On a besoin de $df(\vec{g}(\vec{x}))$, i.e. $df(X,Y)$ où $X=g_1(\vec{x})=x+y$ et $Y=g_2(\vec{x})=x-y$:

$$\begin{aligned}[df(\vec{g}(\vec{x}))]&=(\,2(x+y)\sin((x+y)^2-(x-y)^2)\quad,\quad -2(x-y)\sin(x+y)^2-(x-y)^2)\,)\\&=(\,2(x+y)\sin(4xy)\quad,\quad -2(x-y)\sin(4xy)\,)\end{aligned}$$

D'où le produit $[df(\vec{g}(\vec{x}))].[d\vec{g}(x,y)]=(\,4y\sin(4xy)\quad,\quad 4x\sin(4xy)\,)$, déjà obtenu directement. ∎

Exercice 3.19 Soit $\vec{\psi}(\rho,\theta,\varphi)=\begin{pmatrix}\rho\cos\theta\cos\varphi\\ \rho\sin\theta\cos\varphi\\ \rho\sin\varphi\end{pmatrix}$ définie sur $U=\mathbb{R}_+^*\times]-\pi,\pi[\times]-\frac{\pi}{2},\frac{\pi}{2}[$. Et soit $G(x,y,z)=\begin{pmatrix}x^2+y^2+z^2\\ \frac{\sqrt{x^2+y^2}}{z}\end{pmatrix}$ définie que $\mathbb{R}\times\mathbb{R}\times\mathbb{R}^*$. On notera $\vec{u}=(\rho,\theta,\varphi)$ et $\vec{x}=(x,y,z)$.

Calculer $d\vec{\psi}(\vec{u})$, $dG(\vec{x})$ puis le produit $dG(\vec{\psi}(\vec{u})).d\vec{\psi}(\vec{u})$.

Calculer $(G\circ\vec{\psi})(\vec{u})$ et $d(G\circ\vec{\psi})(\vec{u})$. Vérifier le résultat.

Exercice 3.20 Soit $f:\mathbb{R}^2\to\mathbb{R}$ une fonction telle que $f(x,y)=-f(y,x)$.

Montrer que $\frac{\partial f}{\partial y}(x,y)=-\frac{\partial f}{\partial x}(x,y)$.

Réponse. On pose $g(x,y)=-f(y,x)$, i.e. $g(x,y)=-(f\circ\vec{\varphi})(x,y)$ où $\vec{\varphi}(x,y)=\begin{pmatrix}\varphi_1(x,y)=y\\ \varphi_2(x,y)=x\end{pmatrix}$. On obtient :

$$\frac{\partial g}{\partial x}(x,y)=-\left(\frac{\partial f}{\partial x}(\vec{\varphi}(x,y)).\frac{\partial\varphi_1}{\partial x}(x,y)+\frac{\partial f}{\partial y}(\vec{\varphi}(x,y)).\frac{\partial\varphi_2}{\partial x}(x,y)\right)=-\frac{\partial f}{\partial y}(\vec{\varphi}(x,y)).$$

Comme on a aussi $g(x,y)=f(x,y)$, on obtient $\frac{\partial g}{\partial x}(x,y)=-\frac{\partial f}{\partial y}(\vec{\varphi}(x,y))$.

Exercice 3.21 Soit $g:\mathbb{R}^m\to\mathbb{R}$ une fonction dérivable, et soit $h:\mathbb{R}\to\mathbb{R}$ la fonction définie par $h(t)=g(t,t,...,t)$. Calculer $h'(t)$. Donner le résultat lorsque m=2 et $g(x,y)=x\sin y$.

Réponse. On note $\vec{y}=(y_1,...,y_m)\in\mathbb{R}^m$. On a $g(\vec{y})=g(y_1,...,y_m)\in\mathbb{R}$, et on veut que "$y_i=t$" : on pose donc $f_i(t)=t$ pour $i=1,...,m$, puis $\vec{f}:\mathbb{R}\to\mathbb{R}^m$ la fonction définie par $\vec{f}(t)=\begin{pmatrix}f_1(t)\\ \vdots\\ f_m(t)\end{pmatrix}$. Et la fonction h est définie par $h=g\circ\vec{f}$. On a donc :

$$h'(t)=\frac{\partial g}{\partial y_1}(\vec{f}(t))\frac{\partial f_1}{\partial t}(t)+...+\frac{\partial g}{\partial y_m}(\vec{f}(t))\frac{\partial f_m}{\partial t}(t).$$

Comme $f_i(t)=t$ pour tout $i=1,...,m$ et tout $t\in\mathbb{R}$ (et donc $f_i'(t)=1$), on en déduit $h'(t)=\sum_{i=1}^m\frac{\partial g}{\partial y_1}(t,t,...,t)$. Pour $g(x,y)=x\sin y$, on obtient $h'(t)=\frac{\partial g}{\partial x}(t,t)+\frac{\partial g}{\partial y}(t,t)=\sin t+t\cos t$. Comme ici $h(t)=t\sin t$, on vérifie directement le résultat obtenu.

Exercice 3.22 Soit $g:\mathbb{R}^m\to\mathbb{R}$ une fonction dérivable, soit $\vec{x}\in\mathbb{R}^m$, et soit $h:\mathbb{R}\to\mathbb{R}$ la fonction définie par $h(t)=g(t\vec{x})$. Calculer $h'(t)$. Donner le résultat lorsque m=2, $\vec{x}=(2,3)$ et $g(u,v)=u\sin v$.

Réponse. Même démarche que précédemment, avec $f_i(t)=tx_i$ pour tout $i=1,...,m$ et tout $t\in\mathbb{R}$, d'où $h'(t)=\sum_{i=1}^m x_i\frac{\partial g}{\partial y_1}(t\vec{x})$. Pour $g(x,y)=x\sin y$ et $\vec{x}=(2,3)$, on obtient $h'(t)=2\frac{\partial g}{\partial x}(2t,3t)+3\frac{\partial g}{\partial y}(2t,3t)=2\sin(3t)+6t\cos(3t)$. Comme ici $h(t)=2t\sin(3t)$, on vérifie directement le résultat obtenu.

Exercice 3.23 Soit $f:\mathbb{R}^2\to\mathbb{R}$ une fonction C^1. On pose :

$$h(x)=\int_0^x f(x,t)\,dt.$$

Que vaut $h'(x)$?

Réponse. On pose $g(x,y)=\int_0^y f(x,t)\,dt$. On a $h(x)=g(x,x)$, i.e. $h(x)=(g\circ\vec{\varphi})(x)$ où $\vec{\varphi}(x)=\begin{pmatrix}x\\ x\end{pmatrix}$. D'où $h'(x)=\frac{\partial g}{\partial x}(\vec{\varphi}(x))\frac{\partial\varphi_1}{\partial x}(x)+\frac{\partial g}{\partial y}(\vec{\varphi}(x))\frac{\partial\varphi_2}{\partial x}(x)$, i.e. $h'(x)=\int_0^x\frac{\partial f}{\partial x}(x,t)\,dt+f(x,x)$.

(Autre méthode : faire le changement de variable $u=\frac{t}{x}$ quand $x\neq 0$.)

4 Théorème des accroissements finis : cas de $\mathbb{R}^n$

On commence par regarder le cas des fonctions d'une variable réelle à valeurs vectorielles, i.e. des fonctions de type $\vec{\varphi}:[a,b]\to\mathbb{R}^n$. Une telle fonction a son graphe représenté par une trajectoire dans $\mathbb{R}\times\mathbb{R}^n=\mathbb{R}^{n+1}$.

Pour simplifier, on prendra $n=2$ et on notera $\vec{\varphi}(t)=\begin{pmatrix}\varphi_1(t)\\ \varphi_2(t)\end{pmatrix}$, où donc les fonctions φ_1 et $\varphi_2:\mathbb{R}\to\mathbb{R}$.

Dans ce cas le théorème des accroissements finis 1.38 page 10 n'est plus applicable.

Exemple 4.1 Soit $\vec\varphi : t\in[0,2\pi]\to\vec\varphi(t)=\begin{pmatrix}\cos(t)=\varphi_1(t)\\ \sin(t)=\varphi_2(t)\end{pmatrix}\in\mathbb{R}^2$.

Sa dérivée est $\vec\varphi\,'(t)=\begin{pmatrix}-\sin(t)\\ \cos(t)\end{pmatrix}\in\mathbb{R}^2$, et on a $\vec\varphi(0)-\vec\varphi(2\pi)=0$ ne peut pas être égale à $(2\pi-0)\vec\varphi\,'(c)$ pour un $c\in]0,2\pi[$, car $\vec\varphi\,'(c)\neq 0$ pour tout c (en effet $||\vec\varphi\,'(c)||=(\sin^2c+\cos^2c)^{\frac12}=1$).

Continuons sur l'interprétation du graphe : dans le cas des fonctions $f:I\to\mathbb{R}$ dérivables, le théorème des accroissements finis 1.38 indique qu'il existe c tel que l'accroissement fini $\frac{f(b)-f(a)}{b-a}$ vaut $f'(c)$.

Le théorème des accroissements finis 1.38 est bien sûr applicable à φ_1 et à φ_2 : il existe un réel c_1 pour φ_1 et un réel c_2 pour φ_2 tels que $-\sin c_1=\varphi_1'(c_1)=\frac{\varphi_1(2\pi)-\varphi_1(0)}{2\pi-0}=0$, et $\cos c_2=\varphi_2'(c_1)=\frac{\varphi_2(2\pi)-\varphi_2(0)}{2\pi-0}=0$. Mais $c_1\neq c_2$: ici $c_1=\pi \bmod(\pi)$ est toujours différent de $c_2=\frac{\pi}{2} \bmod(\pi)$.

On a cependant une inégalité (accroissements finis 'affaiblis') pour $E=\mathbb{R}^n$. Rappelons que si f et g de $[a,b]\to\mathbb{R}$ sont continues sur $[a,b]$ et dérivables sur $]a,b[$ et si $|f'|\leq g'$, alors $|\int_a^b f'(t)dt|\leq\int_a^b g'(t)dt$, i.e. $|f(b)-f(a)|\leq g(b)-g(a)$.

Et ce résultat est conservé pour E espace de Banach (en particulier pour $E=\mathbb{R}^n$) :

Théorème 4.2 *(accroissement finis) Soient deux réels a et b, et E un espace de Banach dont la norme est notée $||.||_E$.*

Soient $\vec\varphi:[a,b]\to E$ et $g:[a,b]\to\mathbb{R}$ deux applications continues sur $[a,b]$ et différentiables sur $]a,b[$ telles que :

$$\forall t\in]a,b[,\quad ||\vec\varphi\,'(t)||_E\leq g'(t). \tag{4.1}$$

Alors :

$$||\vec\varphi(b)-\vec\varphi(a)||_E\leq g(b)-g(a) \tag{4.2}$$

(l'intégration passe sous la norme), i.e. l'accroissement de $\vec\varphi$ est majoré par celui de g.

La dernière inégalité s'écrit aussi $||\int_a^b\vec\varphi'(t)dt||_E\leq\int_a^b g't)dt$, ce qui est une généralisation du théorème des accroissements finis dans le cas des fonctions à valeurs scalaires.

Preuve. On va en fait montrer que :

$$\forall\varepsilon>0,\quad\forall t\in[a,b],\quad ||\vec\varphi(t)-\vec\varphi(a)||_E\leq g(t)-g(a)+\varepsilon(t-a)+\varepsilon.$$

(On a ajouté une 'petite' fonction affine.) On fera alors tendre ε vers 0 pour avoir le résultat.

Soit donc $\varepsilon>0$ donné. On pose pour tout $t\in[a,b]$:

$$F(t)=||\vec\varphi(t)-\vec\varphi(a)||_E-\big(g(t)-g(a)+\varepsilon(t-a)+\varepsilon\big),$$

où donc $F(a)=-\varepsilon$, et on va montrer que :

$$U=\{t\in[a,b]:F(t)>0\}=\emptyset$$

ensemble vide, d'où le résultat.

1- U est ouvert car l'application F est continue sur $[a,b]$ et $U=F^{-1}(]0,\infty[)$.

Puis supposons $U\neq\emptyset$.

2- Soit $c=\inf\{t\in U\}$. Alors :

21- $c>a$: en effet, l'application F est continue sur $[a,b]$ et $F(a)=-\varepsilon$ donc $F(t)<0$ pour t suffisamment proche de a.

22- $c\notin U$: en effet, sinon $U=[c,...)$ et U ne serait pas ouvert.

23- $c<b$: sinon $c=b$, donc $U=\{b\}$ et U est fermé.

3- On vient d'obtenir $c\in]a,b[$ et donc, par hypothèse, $||\vec\varphi'(c)||_E\leq g'(c)$. Et donc, ε étant donné, il existe $\eta>0$ tel que $]c,c+\eta[\subset]a,b[$ et, par définition de la dérivée :

$$\forall t\in]c,c+\eta[,\quad \frac{||\vec\varphi(t)-\vec\varphi(c)||_E}{t-c}-\frac{\varepsilon}{2}\leq||\vec\varphi'(c)||_E\quad\text{et}\quad g'(c)\leq\frac{g(t)-g(c)}{t-c}+\frac{\varepsilon}{2}.$$

On en déduit :

$$\forall t\in]c,c+\eta[,\quad \frac{||\vec\varphi(t)-\vec\varphi(c)||_E}{t-c}\leq\frac{g(t)-g(c)}{t-c}+\varepsilon$$

soit :

$$\forall t \in]c, c+\eta[, \quad ||\vec{\varphi}(t) - \vec{\varphi}(c)||_E \le g(t) - g(c) + \varepsilon(t-c)$$

Mais on a aussi $c \notin U$ et donc $F(c) \le 0$, i.e. :

$$||\vec{\varphi}(c) - \vec{\varphi}(a)||_E \le g(c) - g(a) + \varepsilon(c-a) + \varepsilon$$

d'où en sommant (et inégalité triangulaire) :

$$\forall t \in]c, c+\eta[, \quad ||\vec{\varphi}(t) - \vec{\varphi}(a)||_E \le g(t) - g(a) + \varepsilon(t-a) + \varepsilon$$

Mais alors pour tout $t \in [c, c+\frac{\eta}{2}]$ on a $F(t) \le 0$ et donc $c + \frac{\eta}{2} \le \inf\{t \in U\} = c$. Contradiction. Donc $U = \emptyset$. ∎

Exercice 4.3 Montrer que, pour $\vec{\varphi} : [a,b] \to E$, si $\vec{\varphi}' = 0$ alors $\vec{\varphi}$ =cste.

Réponse. Ayant $\vec{\varphi}' = 0$, on a $||\vec{\varphi}'(t)|| \le 0$ pour tout $t{\in}[a,b]$, et donc $||\vec{\varphi}(t) - \vec{\varphi}(a)||_E \le 0$ pour tout $t{\in}[a,b]$, i.e. $\vec{\varphi}(t) = \vec{\varphi}(a)$ pour tout $t{\in}[a,b]$. ∎

Exercice 4.4 Montrer que, pour $\vec{\varphi} : [a,b] \to E$, s'il existe deux réels $m(\ge 0)$ et $M(\ge 0)$ tels que $m \le ||\vec{\varphi}'(t)|| \le M$ pour tout $t{\in}[a,b]$, alors $||\vec{\varphi}(b) - \vec{\varphi}(a)|| \le M(b-a)$, mais que le résultat $||\vec{\varphi}(b) - \vec{\varphi}(a)|| \ge m(b-a)$ est faux en général.

Réponse. On pose $g(t) = Mt$ pour tout $t{\in}[a,b]$. Par hypothèse $||\vec{\varphi}'(t)|| \le g'(t) = M$ pour tout $t{\in}[a,b]$. D'où $||\vec{\varphi}(b) - \vec{\varphi}(a)|| \le M(b-a)$.

Par contre, pour $\vec{\varphi}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ sur $\mathbb{R}$, on a $m = 1$ car $||\vec{\varphi}'(t)|| = 1$ pour tout t, mais $\vec{\varphi}(2\pi) - \vec{\varphi}(0) = 0$ et $0 \not\ge m(2\pi{-}0)$.

À comparer avec l'exercice 1.40. ∎

D'où :

Théorème 4.5 *Soit $\vec{\varphi} : [a,b] \to E$, avec $\vec{\varphi} \in C^1([a,b];E)$. Alors :*

$$||\vec{\varphi}(b) - \vec{\varphi}(a)||_E \le k\ (b-a) \quad \text{où} \quad k = \sup_{t\in[a,b]} ||\vec{\varphi}'(t)||_E. \tag{4.3}$$

Preuve. On prend $g(t) = kt$ où $k = \sup_{t\in[a,b]} ||\vec{\varphi}'(t)||_E$ existe car $\vec{\varphi}' \in C^0([a,b];E)$. Et on a $||\vec{\varphi}'(t)||_E \le g'(t)$ pour tout t, d'où le résultat. ∎

D'où on déduit pour les fonctions de variables vectorielles :

Théorème 4.6 *(Théorème des accroissements finis.) Soit Ω un ouvert de $\mathbb{R}^n$ et $\vec{f} \in C^1(\Omega;E)$. Pour $\vec{a}, \vec{b} \in \mathbb{R}^n$, on note $[\vec{a},\vec{b}]$ le segment $[\vec{a},\vec{b}] = \{\vec{x}(t) = t(\vec{b}{-}\vec{a}) + \vec{a},\ t \in [0,1]\}$. Alors, pour $\vec{a}, \vec{b} \in \Omega$, si $[\vec{a},\vec{b}] \in \Omega$, on a :*

$$||\vec{f}(\vec{b}) - \vec{f}(\vec{a})||_E \le k\ ||\vec{b} - \vec{a}|| \quad \text{où} \quad k = \sup_{\vec{x}\in[\vec{a},\vec{b}]} ||d\vec{f}(\vec{x})||_{L(\mathbb{R}^n;E)}. \tag{4.4}$$

(On rappelle que $||d\vec{f}(\vec{x})||_{L(\mathbb{R}^n;E)} = \sup_{\vec{y}\in\mathbb{R}^n} \frac{||d\vec{f}(\vec{x})(\vec{y})||_E}{||\vec{y}||_{\mathbb{R}^n}}$ et en particulier si $E = \mathbb{R}^p$ et $[d\vec{f}(\vec{x})]$ la matrice jacobienne de $\vec{f}$ en $\vec{x}$, alors $||d\vec{f}(\vec{x})||_{L(\mathbb{R}^n;\mathbb{R}^p)} = \sup_{\vec{y}\in\mathbb{R}^n} \frac{||[d\vec{f}(\vec{x})].\vec{y}||_{\mathbb{R}^p}}{||\vec{y}||_{\mathbb{R}^n}}$.)

Preuve. Soit $\vec{\psi}(t) = \vec{f}(t(\vec{b}{-}\vec{a}) + \vec{a})$ pour $t \in [0,1]$. On a $\vec{\psi} \in C^1([0,1];E)$ avec $\vec{\psi}'(t) = d\vec{f}(\vec{x}).(\vec{b} - \vec{a})$ quand $x = t(\vec{b}{-}\vec{a}) + \vec{a}$, et le théorème précédent donne $||\vec{\psi}(1) - \vec{\psi}(0)||_E \le k_{\vec{\psi}}$ où $k_{\vec{\psi}} = \sup_{t\in[a,b]} ||\vec{\psi}'(t)||_E \le \sup_{\vec{x}\in[\vec{a},\vec{b}]} ||d\vec{f}(\vec{x})||_{L(\mathbb{R}^n;E)} ||\vec{b} - \vec{a}||_E$. ∎

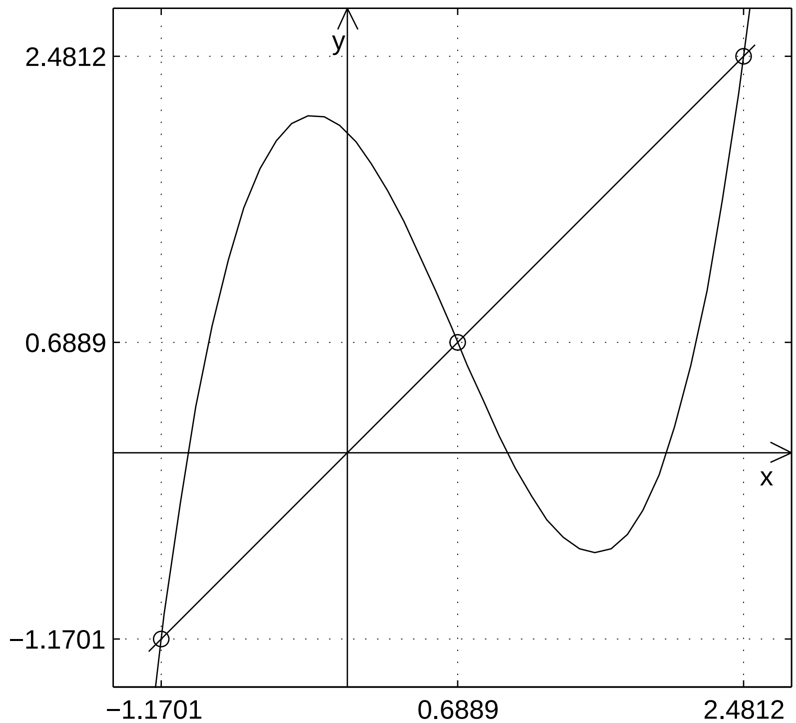

FIGURE 5.1 – Les 3 points fixes de la fonction $f(x) = (x+1)(x-1)(x-2) = x^3 - 2x^2 - x + 2$.

5 Théorème du point fixe de Banach

Soit E un espace de Banach (espace vectoriel normé et complet pour la norme). On prendra par exemple $E = \mathbb{R}^n$ avec la norme euclidienne.

Définition 5.1 Si $f : \Omega \to E$ et $a \in E$ vérifient :

$$a = f(a),$$

a est appelé point fixe de f.

Dans le cas d'une fonction f de $\mathbb{R} \to \mathbb{R}$, un point fixe est l'intersection de la droite "$y = x$" avec la fonction "$y = f(x)$" : un tel x vérifie $x = y = f(x)$ comme souhaité.

Définition 5.2 S'il existe $k > 0$ tel que :

$$\forall x, y \in U, \quad ||f(x) - f(y)||_E \leq k||x - y||_E,$$

alors f est appelée lipschitzienne de rapport k. Et si $k < 1$ alors f est appelée une contraction.

Il est immédiat que si f est lipschitzienne alors f est continue, et même uniformément continue.

On commence par le cas $\Omega = E$:

Théorème 5.3 *Si E est un espace de Banach, si $f : E \to E$ est une contraction, alors f possède un et un seul point fixe :*

$$\exists! a \in E \text{ tel que } a = f(a).$$

Preuve. Soit $b \in E$ un point quelconque, on pose $b_0 = b$. Et soit $b_1 = f(b)$, $b_2 = f^2(b) = f(b_1)$, et pour $n \geq 1$, $b_n = f(b_{n-1}) = f^n(b)$. Alors $(b_n)_{\mathbb{N}}$ est une suite de Cauchy. En effet, on a, pour

$p, n \in \mathbb{N}$, sachant $k < 1$:

$$\begin{aligned}||b_p - b_0||_E \le ||b_p - b_{p-1}||_E + ... + ||b_1 - b_0||_E \le (k^p + ... + k + 1)||b_1 - b_0||_E \\ \le \frac{1-k^{p+1}}{1-k}||b_1 - b_0||_E \le \frac{1}{1-k}||b_1 - b_0||_E,\end{aligned}$$

et donc :

$$\begin{aligned}||b_{n+p} - b_n||_E \le k||b_{n+p-1} - b_{n-1}||_E \le ... \le k^n||b_p - b_0||_E \\ \le k^n\frac{1}{1-k}||b_1 - b_0||_E \underset{n,p\to\infty}{\longrightarrow} 0.\end{aligned}$$

La suite (b_n) est donc de Cauchy dans E complet : elle converge. Soit a sa limite, $a = \lim_n(b_n)$.

Et f étant continue (car Lipschitz), on a $f(a) = f(\lim_n(b_n)) = \lim_n(f(b_n)) = \lim_n(b_{n+1}) = \lim_n(b_n) = a$, et donc a est point fixe.

Enfin si a' est un autre point fixe, i.e. $a' = f(a')$, alors $||a - a'|| = ||f(a) - f(a')|| \le k||a - a'||$ et donc $||a - a'|| = 0$ car $k < 1$. ∎

Si f est uniquement définie sur $\Omega \subset E$ à valeurs dans E, il faut s'assurer dans le théorème précédent que pour $b \in \Omega$ donné, $b_1 = f(b) \in \Omega$. C'est l'objet des hypothèses supplémentaires dans le théorème suivant :

Théorème 5.4 *Si $f : \Omega \to E$ est k-lipschitzienne avec $k < 1$ (et f est une contraction), et s'il existe un point $b \in \Omega$ et un réel $M \in \mathbb{R}$ tels que $d(b, E-\Omega) > M$ et tel que $d(b, f(b)) < M(1-k)$ (i.e. b est suffisamment loin de sa frontière), alors :*

$$\exists! a \in E \text{ tel que } a = f(a),$$

i.e. il existe un unique point fixe de f dans E.

Preuve. Démonstration similaire. ∎

Exercice 5.5 Soit $F : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F(x, y) = \begin{pmatrix}\cos\frac{x+y}{2}\\ \sin\frac{x+y}{2}\end{pmatrix}$. On note $\vec{x} = (x, y)$.

1- Montrer qu'il existe B boule fermée de $\mathbb{R}^2$ telle que $F(B) \subset B$.

2- Montrer que l'application linéaire $dF(\vec{x})$ (représentée par la matrice jacobienne de F en $\vec{x}$ dans la base canonique) vérifie $||df(\vec{x})|| \le \frac{1}{\sqrt{2}}$ pour $\vec{x} \in B$. (On rappelle que si L est une application linéaire, alors $||L|| = \sup_{\vec{x}\neq 0}\frac{||L(\vec{x})||_2}{||\vec{x}||_2}$).

3- En déduire que F admet un unique point fixe dans B.

Réponse. 1- On a $||F(\vec{x})|| = 1$ pour tout $\vec{x}$ (l'image de F est dans le cercle unité) ; donc si $||\vec{x}|| \le r$, i.e. si $\vec{x} \in \bar{B}(\vec{0}, r)$, on a $F(\vec{x}) \in \bar{B}(\vec{0}, 1)$, et $B(\vec{0}, 1) \subset B(\vec{0}, r)$ ssi $r \ge 1$; et donc, toute boule B de centre 0 et de rayon ≥ 1 vérifie $F(B) \subset B$. On prend par exemple $B = \bar{B}(\vec{0}, 1)$.

2- $[df(\vec{x})] = \begin{pmatrix}-\frac{1}{2}\sin\frac{x+y}{2} & -\frac{1}{2}\sin\frac{x+y}{2}\\ \frac{1}{2}\cos\frac{x+y}{2} & \frac{1}{2}\cos\frac{x+y}{2}\end{pmatrix}$, d'où $df(\vec{x}).\vec{v} = \frac{v_1+v_2}{2}\begin{pmatrix}-\sin\frac{x+y}{2}\\ \cos\frac{x+y}{2}\end{pmatrix}$, d'où $||df(\vec{x}).\vec{v}|| = \frac{|v_1+v_2|}{2}$. D'où $\sup_{\vec{v}}\frac{||df(\vec{x}).\vec{v}||}{||\vec{v}||} = \frac{1}{\sqrt{2}}$. En effet, $\frac{(v_1+v_2)^2}{v_1^2+v_2^2} = \frac{v_1^2+v_2^2+2v_1v_2}{v_1^2+v_2^2}$ qui est maximum pour $v_1 = v_2$ avec pour valeur max 2.

3- D'où F est une contraction dans la boule unité fermée, et donc F admet un unique point fixe dans cette boule (d'ailleurs sur le cercle unité image de F) : il existe un unique point (x, y) de $\mathbb{R}^2$ t.q. $x = \cos\frac{x+y}{2}$ et $y = \sin\frac{x+y}{2}$. ∎

6 Théorème d'inversion locale

Soit une fonction inversible :

$$f : \begin{cases}[a, b] \subset \mathbb{R} & \to [c, d],\\ \quad x & \mapsto y = f(x),\end{cases} \qquad f^{-1} : \begin{cases}[c, d] \subset \mathbb{R} & \to [a, b],\\ \quad y & \mapsto f^{-1}(y) = x, .\end{cases}$$

On rappelle que le graphe de f^{-1} se déduit du graphe de f par symétrie autour de la droite $y = x$, voir figure 6.1.

En effet, on commence par remarquer que : si le graphe de f est représenté dans le repère $(0, x, y)$, alors celui de f^{-1} est représenté dans le repère $(0, y, x)$: pour f, c'est le segment $[a, b]$ qui est en abscisse et le segment $[c,d]$ qui est en ordonnées, alors que pour f^{-1}, c'est le segment $[c,d]$ qui est en abscisse et le segment $[a,b]$ qui est en ordonnées.

Puis, dans le repère $(0,x,y)$ soit un point $P=(x,y{=}f(x))$ du graphe de f. Montrons que $Q=(q_1,q_2)=(y{=}f(x),x)=$ le symétrique de P par rapport à la première diagonale "$y=x$" appartient au graphe de f^{-1}, i.e. que le point Q dans le repère $(0,y,x)$ vérifie : $f^{-1}(q_1)=q_2$. C'est immédiat car $f^{-1}(q_1)=f^{-1}(f(x))=x=q_2$.

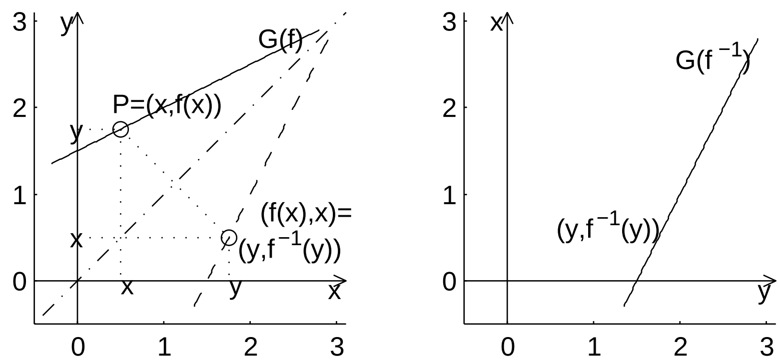

FIGURE 6.1 – Construction des fonctions réciproques : par symétrie par rapport à la bissectrice du premier cadran $y=x$.

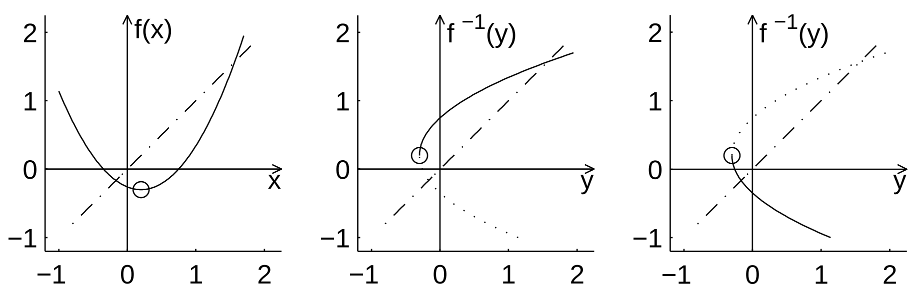

FIGURE 6.2 – La fonction $f(x)=y=(x-0.2)^2-.3$, et ses fonctions réciproques : la première $f^{-1}(y)=x=\sqrt{y+0.3}$ (branche supérieure) définie pour $y\geq-0.3$ et la seconde $f^{-1}(y)=x=-\sqrt{y+0.3}$ (branche inférieure) définie pour $y\geq-0.3$.

Théorème 6.1 *On suppose $f\in C^1([a,b])$ (où $a<b$), et on se donne un point $c\in]a,b[$ tel que $f'(c)\neq0$. Alors il existe $r>0$ tel que $]c{-}r,c{+}r[\subset[a,b]$ pour lequel f est bijective et dont l'inverse est C^1, avec de plus :*

$$\forall y\in f(]c{-}r,c{+}r[),\quad \textit{posant } x=f^{-1}(y),\quad (f^{-1})'(y)=\frac{1}{f'(x)}.$$

Un dessin aide à s'en convaincre : la fonction f^{-1} est la fonction miroir de f par rapport à l'axe défini par la droite "$y=x$", et la pente pour la fonction f^{-1} est l'inverse de la pente de la fonction f (à condition que cette pente ne soit pas nulle).

Preuve. Par hypothèse on a $f'(c)\neq0$. Quitte à considérer $-f$, supposons $f'(c)>0$.

f' continue et $f'(c)>0$ impliquent qu'il existe $r>0$ et un intervalle $]c{-}r,c{+}r[$ sur lequel $f'(x)>\delta$ pour $\delta=\frac{f'(c)}{2}$. On s'intéresse à la restriction f_r de f à l'intervalle $]c{-}r,c{+}r[$, restriction qu'on note également (abusivement) f.

Sur l'intervalle $]c{-}r,c{+}r[$, f est injective : sinon avec le théorème de Rolle, il existerait un point de cet intervalle où f' s'annule, contraire à l'hypothèse $\delta>0$.

Donc f est bijective de $]c{-}r,c{+}r[$ dans son image $f(]c{-}r,c{+}r[)=J$. Et on peut considérer son inverse $f^{-1}:J\to]c{-}r,c{+}r[$.

Montrons que f^{-1} est continue dans J. Soit $y_0\in J$, avec donc $y_0=f(x_0)$ pour un unique $x_0\in]c{-}r,c{+}r[$. Et pour $h\in\mathbb{R}$ suffisamment petit pour que $y=y_0+h\in J$, et y est associé à un

$x \in]c-r, c+r[$ tel que $y = f(x)$. Le théorème de accroissements finis indique qu'il existe ξ entre x_0 et x tel que $f(x) - f(x_0) = (x - x_0)f'(\xi)$, d'où :

$$|f^{-1}(y) - f^{-1}(y_0)| = |x - x_0| = \frac{|f(x) - f(x_0)|}{|f'(\xi)|} \le \frac{|f(x) - f(x_0)|}{\delta} = \frac{y - y_0}{\delta},$$

et donc $|f^{-1}(y) - f^{-1}(y_0)| \to 0$ quand $y \to y_0$: f^{-1} est bien continue en y_0, pour tout $y_0 \in J$: f^{-1} est continue dans J.

Montrons que f^{-1} est dérivable en $y_0 \in J$. On a :

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{x - x_0}{(x - x_0)f'(\xi_x)}$$

pour un ξ_x entre x_0 et x, le dénominateur ne s'annulant pas car $f'(\xi_x) > \frac{\delta}{2} > 0$. D'où :

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} \underset{y \to y_0}{\longrightarrow} \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))} \overset{\text{noté}}{=} (f^{-1})'(y_0).$$

Enfin, montrons que la fonction $(f^{-1})'$ ainsi définie sur J est continue. Les fonctions f' et f^{-1} sont continues, f' ne s'annule pas dans $]c-r, c+r[$, et $X \to \frac{1}{X}$ est continue là où $X \neq 0$. Donc $(f^{-1})'$ est bien continue. ∎

Et en n dimensions on a :

Théorème 6.2 *Soit $\vec{f} : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction continue qui est C^1. Et soit $\vec{c} \in \Omega$ un point tel que $\det([d\vec{f}(\vec{c})]) \neq 0$ (avec $[d\vec{f}(\vec{c})] = [\frac{\partial f_i}{\partial x_j}(\vec{c})]_{1 \le i,j \le n}$ la matrice jacobienne de $\vec{f}$ en $\vec{c}$ et $\det([d\vec{f}(\vec{c})])$ son jacobien). Alors il existe $U \subset \Omega$, U ouvert et $U \ni c$, pour lequel la restriction de $\vec{f}$ à U : $\vec{f}_{|U} : \left\{ \begin{array}{ll} U & \to V = f(U) \\ \vec{x} & \mapsto \vec{y} = \vec{f}(\vec{x}) \end{array} \right\}$ est un difféomorphisme (bijective C^1 d'inverse C^1). Et pour tout $\vec{y} \in V$, notant $\vec{x} = \vec{f}^{-1}(\vec{y})$:*

$$[d(\vec{f}^{-1})(\vec{y})] = [d\vec{f}(\vec{x})]^{-1},$$

i.e. la matrice jacobienne de l'inverse est l'inverse de la matrice jacobienne, i.e., la matrice $[d\vec{f}(\vec{x})]$ représentant l'application linéaire tangente à $\vec{f}$ en $\vec{x}$ a pour inverse la matrice $[d\vec{f}(\vec{x})]^{-1} = d(\vec{f}^{-1})(\vec{y})$ représentant l'application linéaire tangente à $\vec{f}^{-1}$ en $\vec{y} = \vec{f}(\vec{x})$.

Preuve. Montrons que $\vec{f}$ est inversible dans un voisinage de $\vec{c}$.

Ayant $\det([d\vec{f}(\vec{c})]) \neq 0$, la matrice $[d\vec{f}(\vec{c})]$ est inversible, et l'application linéaire $df(\vec{c})$ est inversible. Soit $L = (d\vec{f}(\vec{c}))^{-1}$ l'application linéaire inverse, avec donc $L.d\vec{f}(\vec{c}) = I$ identité. Ayant $\vec{f} \in C^1$, on a $d\vec{f} \in C^0$, et donc il existe $\rho > 0$ tel que :

$$\forall \vec{x} \in B(\vec{c}, \rho), \quad |||d\vec{f}(\vec{x}) - d\vec{f}(\vec{c})||| \le \frac{1}{2|||L|||}, \tag{6.1}$$

où $|||L|||$ est la norme de L et $B(\vec{c}, \rho)$ est la boule centrée en $\vec{c}$ de rayon ρ. Puis soit $r = \frac{\rho}{2|||L||}$ et $\vec{y} \in B(\vec{f}(\vec{c}), r)$. On pose :

$$G_{\vec{y}} : \left\{ \begin{array}{ll} B(\vec{c}, \rho) & \to \mathbb{R}^n \\ \vec{x} & \mapsto G_{\vec{y}}(\vec{x}) = \vec{x} - L(\vec{f}(\vec{x}) - \vec{y}). \end{array} \right.$$

Montrons que $G_{\vec{y}}$ est contractante. On a $dG_{\vec{y}}(\vec{x}) = I - L.d\vec{f}(\vec{x}) = L(d\vec{f}(\vec{c}) - d\vec{f}(\vec{x}))$ d'où $|||dG_{\vec{y}}(\vec{x})||| \le |||L||| \frac{1}{2|||L|||} = \frac{1}{2}$ dans $B(\vec{c}, \rho)$.

D'où étant donné deux points $x_1, x_2 \in B(\vec{c}, \rho)$, le théorème des accroissements finis 4.2 (son corollaire 4.6) donne :

$$||G_{\vec{y}}(\vec{x}_1) - G_{\vec{y}}(\vec{x}_2)|| \le \sup_{\vec{x} \in [\vec{x}_1, \vec{x}_2]} (|||dG_{\vec{y}}(\vec{x})|||)||\vec{x}_1 - \vec{x}_2|| \le \frac{1}{2}||\vec{x}_1 - \vec{x}_2||. \tag{6.2}$$

D'où on a $G_{\vec{y}}(B(\vec{c}, \rho)) \subset B(\vec{c}, \rho)$ car pour $\vec{x} \in B(\vec{c}, \rho)$ et $\vec{y} \in B(\vec{f}(\vec{c}), r)$:

$$||G_{\vec{y}}(\vec{x}) - \vec{c}|| \le ||G_{\vec{y}}(\vec{x}) - G_{\vec{y}}(\vec{c})|| + ||G_{\vec{y}}(\vec{c}) - \vec{c}|| \le \frac{1}{2}||\vec{x} - \vec{c}|| + ||L(\vec{y} - \vec{f}(\vec{c}))|| \le \frac{\rho}{2} + |||L|||\, r \le \rho.$$

D'où $G_{\vec{y}}$ est contractante dans $B(\vec{c}, \rho)$ de rapport $\frac{1}{2}$.

Donc $G_{\vec{y}}$ a un unique point fixe $\vec{x}$: $G_{\vec{y}}(\vec{x}) = \vec{x}$, i.e. à $\vec{y} \in B(\vec{f}(\vec{c}), r)$ fixé, il existe un unique point $\vec{x} \in B(\vec{c}, \rho)$ t.q. $\vec{y} = \vec{f}(\vec{x})$, et donc $\vec{f}$ est bijective de $f^{-1}(B(\vec{f}(\vec{c}), r))$ dans $B(\vec{f}(\vec{c}), r)$.

On note $V = B(\vec{f}(\vec{c}), r)$ (qui est ouvert) et $U = \vec{f}^{-1}(B(\vec{f}(\vec{c}), r))$, et on a obtenu que f est bijective de U dans V. Et comme $\vec{f}$ est continue, et U est l'image inverse de V, on a U ouvert qui contient $\vec{c}$.

Montrons que $\vec{f}^{-1} : V \to U$ est continue sur V. Montrons que, pour tout $\vec{x}_1, \vec{x}_2 \in U$:

$$||\vec{f}(\vec{x}_2) - \vec{f}(\vec{x}_1)|| \geq \frac{1}{2|||L|||}||\vec{x}_1 - \vec{x}_2||. \tag{6.3}$$

On a :

$$||G_{\vec{y}}(\vec{x}_1) - G_{\vec{y}}(\vec{x}_2)|| = ||(\vec{x}_1 - \vec{x}_2) - L(\vec{f}(\vec{x}_1) - \vec{f}(\vec{x}_2))|| \geq ||\vec{x}_1 - \vec{x}_2|| - |||L|||\,||\vec{f}(\vec{x}_1) - \vec{f}(\vec{x}_2)||,$$

d'où avec (6.2) :

$$|||L|||\,||\vec{f}(\vec{x}_2) - \vec{f}(\vec{x}_1)|| \geq \frac{1}{2}||\vec{x}_1 - \vec{x}_2||.$$

D'où $||\vec{f}^{-1}(\vec{y}_1) - \vec{f}^{-1}(\vec{y}_2)|| \leq 2|||L|||\,||\vec{y}_1 - \vec{y}_2||$, et $\vec{f}$ est lipschitzienne sur V donc continue.

Montrons que $\vec{f}^{-1}$ est dérivable sur V, i.e., pour $\vec{y} \in V$, il existe une application linéaire $L_{\vec{y}} : \mathbb{R}^n \to \mathbb{R}^n$ telle que $\vec{f}^{-1}(\vec{y}_\varepsilon) - \vec{f}^{-1}(\vec{y}) = L_{\vec{y}}(\vec{y}_\varepsilon - \vec{y}) + o(\vec{y}_\varepsilon - \vec{y})$. Montrons que $L_{\vec{y}} = (d\vec{f}(\vec{x}))^{-1}$ pour $\vec{x} = f^{-1}(\vec{y})$. On a $d\vec{f}(\vec{x}).(\vec{x}_\varepsilon - \vec{x}) = \vec{f}(\vec{x}_\varepsilon) - \vec{f}(\vec{x}) + o(\vec{x}_\varepsilon - \vec{x})$, d'où :

$$\vec{x}_\varepsilon - \vec{x} = d\vec{f}(\vec{x})^{-1}.(\vec{f}(\vec{x}_\varepsilon) - \vec{f}(\vec{x})) + o(\frac{1}{2|||L|||}(\vec{x}_\varepsilon - \vec{x})).$$

En effet, on a (6.1) et donc $|||d\vec{f}(\vec{c})||| - |||d\vec{f}(\vec{x})||| \leq \frac{1}{2|||L|||}$, donc $|||d\vec{f}(\vec{x})||| \geq |||d\vec{f}(\vec{c})||| - \frac{1}{2|||L|||}$, et comme $I = L.d\vec{f}(\vec{c})$, on a $1 \leq |||L|||\,|||d\vec{f}(\vec{c})|||$, d'où $|||d\vec{f}(\vec{c})||| \geq \frac{1}{|||L|||}$ et $|||d\vec{f}(\vec{x})||| \geq \frac{1}{2|||L|||}$.

Donc : $\vec{x}_\varepsilon - \vec{x} = d\vec{f}(\vec{x})^{-1}.(\vec{f}(\vec{x}_\varepsilon) - \vec{f}(\vec{x})) + o((\vec{x}_\varepsilon - \vec{x}))$. Et $\vec{f}^{-1}$ est continue sur le compact $[\vec{y}, \vec{y}_\varepsilon]$ et donc $||\vec{x}_\varepsilon - \vec{x}|| \leq M||\vec{y}_\varepsilon - \vec{y}||$ où M est la borne supérieure de f^{-1} sur $\overline{B(\vec{f}(\vec{c}), \frac{d}{2})}$ dès que $\vec{y}_\varepsilon \in B(\vec{f}(\vec{c}), \frac{d}{2})$ où $d = d(\vec{y}, \partial B(f(\vec{c}), r))$ est la distance de $\vec{y}$ au bord de la boule $V = B(f(\vec{c}), r)$. Et donc $o((\vec{x}_\varepsilon - \vec{x})) = o((\vec{y}_\varepsilon - \vec{y}))$ au voisinage de $\vec{y}$, et donc $\vec{f}^{-1}(\vec{y}_\varepsilon) - \vec{f}^{-1}(\vec{y}) = d\vec{f}(\vec{x})^{-1}(\vec{y}_\varepsilon - \vec{y}) + o(\vec{y}_\varepsilon - \vec{y})$ au voisinage de $\vec{y}$. Donc $\vec{f}^{-1}$ est dérivable en tout $\vec{y} \in V$ et $d\vec{f}^{-1}(\vec{y}) = d\vec{f}(\vec{x})^{-1}$.

Montrons que $d\vec{f}^{-1}$ est continue sur V. On a $d\vec{f}^{-1}(\vec{y}) = (d\vec{f}(\vec{x}))^{-1} = (d\vec{f} \circ \vec{f}^{-1})(\vec{y}))^{-1}$ avec $d\vec{f}$ et $\vec{f}^{-1}$ continues et avec l'application linéaire inverse $M \to M^{-1}$ qui est continue en M lorsque l'application linéaire M est inversible. ▪

7 Théorème des fonctions implicites

Une relation implicite (un lien, une liaison) entre deux variables x et y est une équation de type $f(x, y) = 0$ où f est une fonction donnée. Si cette équation peut se mettre sous la forme $y = g(x)$, où g est une fonction donnée, on a "explicité" la relation "implicite" $f(x, y) = 0$. Des conditions simples pour permettant d'expliciter sont données par le théorème des fonctions implicites.

7.1 Cas des fonction $\mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}$

On se donne une fonction :

$$f : \begin{cases} \Omega \subset \mathbb{R}^n & \to \mathbb{R}, \\ (x_1, ..., x_n) & \mapsto z = f(x_1, ..., x_n). \end{cases}$$

Le graphe de cette fonction est le sous-ensemble de $\mathbb{R}^{n+1}$ défini par :

$$G_f = \{(x_1, ..., x_n, z) : z = f(x_1, ..., x_n)\} \quad \subset \mathbb{R}^{n+1}.$$

Et on s'intéresse à l'intersection de G_f avec le plan $z = c$.

7.1.1 Courbe de niveau

Définition 7.1 Soit $c \in \mathbb{R}$. On appelle surface de niveau de f de valeur c l'ensemble :

$$\Gamma_c = \{(x_1, ..., x_n) \in \Omega : f(x_1, ..., x_n) = c\}$$

Si $\mathbb{R}^n = \mathbb{R}^2$, la surface de niveau est appelée courbe de niveau : c'est l'ensemble des (x, y) tels que $f(x, y) = z$. Et si $\mathbb{R}^n = \mathbb{R}$, la courbe de niveau est au plus un ensemble de points (et au plus un point unique si f est bijective).

Sur une surface de niveau, les coordonnées $(x_1, ..., x_n)$ ne sont pas indépendantes : elles sont liées par l'équation $f(x_1, ..., x_n) = c$, ce lien étant implicite. La question est alors de savoir si de ces n variables dépendantes on peut extraire $n-1$ variables indépendantes. Si c'est le cas, on pourra (au moins localement) exprimer l'une d'elles en fonction des autres, par exemple, on pourra exprimer x_n =fonction de $(x_1, ..., x_{n-1})$. Le lien est alors explicite : on a explicité x_n en fonction des autres variables.

Le théorème des fonctions implicites donne une condition suffisante pour qu'une relation explicite existe.

Exemple 7.2 Soit $f(x, y) = ax+by$ avec $(a, b) \neq (0, 0)$, et soit $z_0 = c \in \mathbb{R}$ donné. Alors l'ensemble des (x, y) qui satisfont à $ax + by = c$ est une droite. Et les (x, y) qui sont sur cette droite peuvent également s'exprimer sous la forme $y = g(x)$ où $g(x) = \dfrac{c - ax}{b}$ existe à condition que $b \neq 0$. Noter que $b = \frac{\partial f}{\partial y}(x, y)$.

Les (x, y) qui sont sur cette droite peuvent également s'exprimer sous la forme $x = h(y)$ où $h(y) = \dfrac{c - by}{a}$ existe à condition que $a \neq 0$. ▪

Localement, toute fonction $f \in C^1(\mathbb{R}^2; \mathbb{R})$ au voisinage d'un point (x_0, y_0) admet un développement limité :

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

Et f est approximée par son plan tangent en (x_0, y_0). On en déduit que, dès que $B = \frac{\partial f}{\partial y}(x_0, y_0) \neq 0$, le long d'une courbe de niveau $f(x, y) = c$:

$$y = y_0 - \frac{A}{B}(x - x_0) + o(x-x_0, y-y_0) \overset{\text{noté}}{=} g(x),$$

où on a noté $A = \frac{\partial f}{\partial x}(x_0, y_0)$, i.e. on a explicité y en fonction de x sur la courbe de niveau. Et en particulier, on a $g'(x_0) = -\frac{A}{B} = -\frac{\frac{\partial f}{\partial x}(x_0,y_0)}{\frac{\partial f}{\partial y}(x_0,y_0)}$.

Exemple 7.3 Soit $f(x, y) = xe^y$, et soit $c \neq 0$ donné. On regarde l'ensemble des (x, y) tels que $xe^y = c$. On obtient la relation explicite $x = h(y) = ce^{-y}$, et quand cela a un sens la relation $y = g(x) = \log(\frac{|c|}{|x|})$: cette relation n'est pas valable pour tout x dans $\mathbb{R}$ (problème au voisinage de $x = 0$), et une telle relation ne peut être établie que localement dans $]-\infty, 0[$ où dans $]0, \infty[$.

Et si on exprime f en fonction de son développement limité :

$$f(x, y) = x_0e^{y_0} + (x-x_0)(e^{y_0}) + (y-y_0)(x_0e^{y_0}) + o(x-x_0, y-y_0)$$

le plan tangent a une équation de la forme $z = Ax + By + C$ et $B = x_0e^{y_0}$ est nul si $x_0 = 0$, auquel cas on ne peut pas expliciter y en fonction de x au voisinage d'un point $(0, y_0)$ lorsqu'on considèrera l'équation $f(x, y) = 0$. Par contre, $A = e^{y_0} \neq 0$ et on peut toujours expliciter x en fonction de y comme on l'a vu pour résoudre l'équation $f(x, y) = c$: à savoir $x = ce^{-y}$. ▪

7.1.2 Cas des fonction $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$

On commence par écrire et démontrer ce théorème pour les fonctions $\mathbb{R}^2 \to \mathbb{R}$:

Théorème 7.4 *Soit Ω ouvert de $\mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction $C^1(\Omega, \mathbb{R})$. Soient $(x_0, y_0) \in \Omega$ et $c \in \mathbb{R}$ tels que :*

$$f(x_0, y_0) = c \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) \neq 0.$$

Alors il existe un voisinage ouvert $V_1 \times V_2$ de (x_0, y_0), et il existe une fonction dérivable $g : V_1 \to V_2$ telle que :

$$\forall (x, y) \in V_1 \times V_2, \quad f(x, y) = c \quad \iff \quad \forall (x, y) \in V_1 \times V_2, \quad y = g(x).$$

De plus, la dérivée de g en x_0 est donnée par :

$$g'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}. \tag{7.1}$$

Preuve. Si g existe et est dérivable alors on a : $\forall x \in V_1$, $f(x, g(x)) = c$. D'où par dérivation :

$$\forall x \in V_1, \text{ posant } y = g(x), \ \frac{\partial f}{\partial x}(x, y) + \frac{\partial f}{\partial y}(x, y) g'(x) = 0,$$

d'où la relation (7.1).

Montrons que g existe et est dérivable. On suppose $c = 0$, sinon on remplace f par la fonction $(x, y) \to f(x, y) - c$. Et on suppose que $\frac{\partial f}{\partial y}(x_0, y_0) = m > 0$ sinon on considère $-f$.

Ayant $\frac{\partial f}{\partial y}$ continue, il vient :

$$\exists a > 0, \ \exists b > 0, \quad \forall x, y \in [x_0{-}a, x_0{+}a] \times [y_0{-}b, y_0{+}b], \quad \frac{\partial f}{\partial y}(x, y) \geq \frac{m}{2}.$$

On va alors démontrer le résultat suivant, sachant $f(x_0, y_0) = 0 (= c)$:

$$\exists \alpha > 0, \ \alpha \leq a, \quad \forall x \in [x_0{-}\alpha, x_0{+}\alpha], \quad f(x, y_0 - b) < 0 < f(x, y_0 + b). \tag{7.2}$$

(Il suffira de prendre x suffisamment proche de x_0, f et $\frac{\partial f}{\partial x}$ étant continues.)

Pour cela : on sait que dans le voisinage $[x_0{-}a, x_0{+}a] \times [y_0{-}b, y_0{+}b]$, on a, $\frac{\partial f}{\partial x}$ étant continu :

$$\exists M > 0, \quad \forall x, y \in [x_0{-}a, x_0{+}a] \times [y_0{-}b, y_0{+}b], \quad \frac{\partial f}{\partial x}(x, y) \leq M.$$

Et le théorème des valeurs intérmédiaires donne :

$$\forall x \in [x_0{-}a, x_0{+}a], \quad \exists \xi_x \in (x_0, x), \quad f(x, y_0) = f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(\xi_x, y_0).$$

(La notation $\xi_x \in (x_0, x)$ signifie ici que si $x_0 \leq x$ alors $\xi_x \in [x_0, x]$ et si $x \leq x_0$ alors $\xi_x \in [x, x_0]$.) Et donc, sachant $f(x_0, y_0) = 0$:

$$\forall \alpha, \ 0 < \alpha < a, \quad \forall x \in [x_0{-}\alpha, x_0{+}\alpha], \quad |f(x, y_0)| \leq \alpha M.$$

De même,

$$\exists \xi_y \in [y_0, y_0{+}b], \quad f(x, y_0 + b) - f(x, y_0) = b \frac{\partial f}{\partial y}(x_0, \xi_y) \geq b \frac{m}{2}.$$

On en déduit, en procédent de même avec $y_0 - b$:

$$\begin{cases} f(x, y_0 + b) = [f(x, y_0 + b) - f(x, y_0)] + f(x, y_0) > b \dfrac{m}{2} - \alpha M, \\[2ex] f(x, y_0 - b) = [f(x, y_0 - b) - f(x, y_0)] + f(x, y_0) < -b \dfrac{m}{2} + \alpha M, \end{cases}$$

Et il suffit de prendre $\alpha = \frac{bm}{4M}$ pour obtenir (7.2).

Et de (7.2), à x fixé dans $[x_0{-}\alpha, x_0{+}\alpha]$, par continuité de f, on déduit qu'il existe un point $y \in [y_0{-}b, y_0{+}b]$ tel que $f(x, y) = 0$. Et comme la fonction $y \in [y_0{-}b, y_0{+}b] \to f(x, y)$ est strictement croissante ($\frac{\partial f}{\partial y}(x, y) > 0$), ce point y est unique : on pose $y = g(x)$.

Il reste à montrer que la fonction g ainsi définie est dérivable. On prend $y = g(x)$ et k tel que $y + k = g(x + h)$, et on s'intéresse à $\frac{g(x+h) - g(x)}{h}$ qui vaut donc $\frac{(y+k) - y}{h} = \frac{k}{h}$.

Le théorème des accroissements finis appliqué à la fonction $t \in \mathbb{R} \to f(x+th, y+tk) \in \mathbb{R}$ (on se place sur le segment de droite qui va de (x,y) à $(x+h, y+k)$) donne :

$$\exists(\xi,\eta) \in [x, x{+}h]\times[y, y{+}k], \quad f(x{+}h, y{+}k) - f(x,y) = h\frac{\partial f}{\partial x}(\xi,\eta) + k\frac{\partial f}{\partial y}(\xi,\eta).$$

Mais $y = g(x)$ et $y + k = g(x+h)$ donnent $f(x{+}h, y{+}k) = f(x,y) = 0$ (on est sur la courbe de niveau), d'où :

$$\frac{k}{h} = -\frac{\frac{\partial f}{\partial x}(\xi,\eta)}{\frac{\partial f}{\partial y}(\xi,\eta)}.$$

Et donc :

$$\lim_{h\to 0} \frac{g(x+h)-g(x)}{h} = -\frac{\frac{\partial f}{\partial x}(x,y)}{\frac{\partial f}{\partial y}(x,y)} \in \mathbb{R}.$$

Ce qui conclut la démonstration. ▪

Exercice 7.5 Soit $f(x,y) = x^2 + y^2$ définie sur $\mathbb{R}^2$. On s'intéresse à la courbe de niveau R avec $R > 0$, i.e. à l'ensemble des (x,y) t.q. $f(x,y) = R$, et en particulier au voisinage d'un point (x_0, y_0) sur cette courbe. Expliciter cette courbe au voisinage d'un point tel que $y_0 \neq 0$ sous la forme $y = g(x)$, puis calculer $g'(x_0)$ et $g''(x_0)$. Puis donner le développement limité de $h : x \to \sqrt{R - x^2}$ à l'ordre 2 au voisinage de x_0.

Réponse. On a $\frac{\partial f}{\partial y}(x,y) = 2y$ est non nul pour $y \neq 0$. On commence par considérer un point (x,y) tel que $y \neq 0$ (faire un dessin du cercle de rayon R). D'où il existe un voisinage $V_1 \times V_2$ de (x_0, y_0) et une fonction $g : V_1 \to V_2$ telle que $f(x,y) = R$ équivaut à $y = g(x)$ pour tout $(x,y) \in V_1 \times V_2$, avec $g(x_0) = y_0$. On a $f(x, g(x)) = R$ pour tout $x \in V_1$, d'où $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial x} g'(x) = 0$, i.e. $2x + 2g(x)\,g'(x) = 0$, d'où d'une part $g'(x_0) = -\frac{x_0}{y_0}$ et d'autre part $2 + 2g'(x)^2 + 2g(x)g''(x) = 0$, d'où $1 + \frac{x_0^2}{y_0^2} + y_0 g''(x_0) = 0$, d'où $g''(x_0) = -\frac{R^2}{y_0^3}$.

Par ailleurs, $h(x) = (R - x^2)^{\frac{1}{2}}$, $h(x_0) = y_0$, $h'(x) = -x(R - x^2)^{-\frac{1}{2}}$, $h'(x_0) = -\frac{x_0}{y_0}$, $h''(x) = -(R - x)^{-\frac{1}{2}} - x^2(R - x^2)^{-\frac{3}{2}}$, $h''(x_0) = -\frac{1}{y_0} - \frac{x_0^2}{y_0^3} = -\frac{R^2}{y_0^3}$: donc g et h ont même développement à l'ordre 2 (d'ailleurs à tout ordre puisqu'ici on connaît la solution analytique à savoir $g = h$). ▪

Exercice 7.6 On pose $f(x,y) = \sin x - e^y + e^{-y}$. Montrer que l'équation "$\sin x - e^y + e^{-y} = 0$" est explicitable au voisinage de 0 par une fonction $y = \varphi(x)$ où $\varphi \in C^1(\mathbb{R};\mathbb{R})$. Calculer $\varphi'(0)$.

Réponse. Si φ existe, on a $f(0, \varphi(0)) = 0$, i.e. $-e^{\varphi(0)} + e^{-\varphi(0)} = 0 = 2\sinh(\varphi(0))$ (sinus hyperbolique), et donc $\varphi(0) = 0$. On a $\frac{\partial f}{\partial y}(x,y) = -e^y - e^{-y} = -2\cosh(y)$ (cosinus hyperbolique), et donc au point $(x{=}0, y{=}\varphi(0){=}0)$ on a $\frac{\partial f}{\partial y}(0,0) = -2 \neq 0$. Donc on peut appliquer le théorème des fonctions implicites : il existe une fonction $\varphi : \mathbb{R} \to \mathbb{R}$ définie au voisinage de $x{=}0$ telle que $\varphi(x) = y \Leftrightarrow f(x,y) = 0$. On a donc $f(x, \varphi(x)) = 0$, donc $\frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x))\varphi'(x) = 0 = \cos x - (e^{\varphi(x)}y + e^{-\varphi(x)})\varphi'(x)$, d'où $\varphi'(0) = \frac{1}{2}$. ▪

Exercice 7.7 Même question pour l'équation $x^3 + y^3 - 3xy - 1 = 0$, où $\varphi(0) = 1$. Et donner un développement limité de φ à l'ordre 3 au voisinage de 0.

Réponse. On pose $f(x,y) = x^3 + y^3 - 3xy - 1$. On a $\frac{\partial f}{\partial y}(x,y) = 3y^2 - 3x$. En particulier, $\frac{\partial f}{\partial y}(0,1) = 3 \neq 0$. On peut donc expliciter f au voisinage de $(0,1)$: il existe une fonction $\varphi : V_1 \to V_2$ avec V_1 voisinage de 0 et V_2 voisinage de 1 telle que $y = \varphi(x) \Leftrightarrow f(x,y) = 0$. Pour $x \in V_1$ on a donc $f(x, \varphi(x)) = 0$, d'où $\frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x))\varphi'(x) = 0$. Soit $2x^2 - 3\varphi(x) + (3\varphi(x)^2 - 3x)\varphi'(x) = 0$. D'où, avec $\varphi(0) = 1$, on obtient $-3 + 3\varphi'(0) = 0$, d'où $\varphi'(0) = 1$. Puis $4x - 3\varphi'(x) + (6\varphi'(x)\varphi(x) - 3)\varphi'(x) + (3\varphi(x)^2 - 3x)\varphi''(x) = 0$. D'où $-3 + (6-3) + 3\varphi''(0) = 0$, d'où $\varphi''(0) = 0$. D'où $4 - 3\varphi''(x) + (6\varphi''(x)\varphi(x) + 6\varphi'(x)^2)\varphi'(x) + (6\varphi'(x)\varphi(x) - 3)\varphi''(x) + (6\varphi'(x)\varphi(x) - 3)\varphi''(x) + (3\varphi(x)^2 - 3x)\varphi'''(x) = 0$. D'où $4 + 6 + 3\varphi'''(0) = 0$, d'où $\varphi'''(0) = -10/3$. ▪

Exercice 7.8 Même question pour l'équation $1 - ye^x + xe^y = 0$. Et donner un développement limité de φ à l'ordre 2 au voisinage de 0. ▪

7.1.3 Cas des fonction $\mathbb{R}^{n-1}\times\mathbb{R}\to\mathbb{R}$

Théorème 7.9 *Soit Ω ouvert de $\mathbb{R}^n$ et $f:\Omega\to\mathbb{R}$ une fonction $C^1(\Omega,\mathbb{R})$. Pour $(x_1,...,x_n)\in\mathbb{R}^n$, on note $\vec{x}=(x_1,...,x_{n-1})$ et $y=x_n$. Soit $(\vec{x}_0,y_0)\in\Omega\subset\mathbb{R}^{n-1}\times\mathbb{R}$ et soit $c\in\mathbb{R}$ tels que :*

$$f(\vec{x}_0,y_0)=c \quad \text{et} \quad \frac{\partial f}{\partial x_n}(\vec{x}_0,y_0)\neq 0.$$

Alors il existe un voisinage ouvert $V_1\times V_2$ de $(\vec{x}_0,y_0)$, et il existe une fonction dérivable $g:V_1\to V_2$ telle que :

$$\forall(\vec{x},y)\in V_1\times V_2,\quad f(\vec{x},y)=c \quad\Longleftrightarrow\quad \forall(\vec{x},y)\in V_1\times V_2,\quad y=g(\vec{x}).$$

De plus, ayant $f(\vec{x},g(\vec{x}))=c$ sur la ligne de niveau, les dérivées partielles de g en $\vec{x}_0$ sont données par :

$$\frac{\partial g}{\partial x_i}(\vec{x})=-\frac{\frac{\partial f}{\partial x_i}(\vec{x},g(\vec{x}))}{\frac{\partial f}{\partial x_n}(\vec{x},g(\vec{x}))}. \tag{7.3}$$

Preuve. Exercice : adapter la démonstration précédente en considérant $\vec{x}=(x_1,...,x_{n-1})$ et $y=x_n$. En particulier, pour $(\vec{x},y)\in V_1\times V_2$ on a $y=g(\vec{x})$ et $f(\vec{x},g(\vec{x}))=c$, d'où $df(\vec{x},g(\vec{x}))=0$, i.e. pour tout $i=1,...,n-1$ on a : $\frac{\partial f}{\partial x_i}(\vec{x},g(\vec{x}))+\frac{\partial f}{\partial x_n}(\vec{x},g(\vec{x}))\frac{\partial g}{\partial x_i}(\vec{x})=0$. ∎

Exercice 7.10 Soit la fonction $f(x,y,z)=x^2+4y^2-3z^2$. On s'intéresse au voisinage du point $(3,0,1)$ et donc à la courbe de niveau $=6$ puisque $f(3,0,1)=6$. Montrer qu'au voisinage du point $(3,0,1)$, l'équation implicite $f(x,y,z)=6$ peut être explicitée sous la forme $z=g(x,y)$, et calculer $[dg(x,y)]$.

Réponse. On a $\frac{\partial f}{\partial z}(3,0,1)=-6\neq 0$. Donc il existe une fonction $g:V_1\subset\mathbb{R}^2\to V_2\subset\mathbb{R}$, avec V_1 voisinage de $(3,0)$ et v_2 voisinage de 1, telle que la courbe de niveau $f(x,y,z)=6$ au voisinage du point $(3,0,1)$ s'écrit $z=g(x,y)$. Ayant $f(x,y,g(x,y))=6$, posant $\vec{X}=(x,y,g(x,y))$, on déduit en dérivant en x que $\frac{\partial f}{\partial x}(\vec{X})+\frac{\partial f}{\partial z}(\vec{X})\frac{\partial g}{\partial x}(x,y)=0$, d'où $\frac{\partial g}{\partial x}(x,y)=-\frac{\frac{\partial f}{\partial x}(\vec{x})}{\frac{\partial f}{\partial z}(\vec{x})}$, d'où $\frac{\partial g}{\partial x}(3,0)=1$. Et de même avec $f(x,y,g(x,y))=6$, en dérivant en y, on a $\frac{\partial f}{\partial y}(\vec{X})+\frac{\partial f}{\partial z}(\vec{X})\frac{\partial g}{\partial y}(x,y)=0$, d'où $\frac{\partial g}{\partial y}(3,0)=0$. ∎

7.2 Cas des fonction $\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}^m$

7.2.1 Développement limités, fonction $\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$

Pour $f\in C^1(\mathbb{R}^n\times\mathbb{R}^m;\mathbb{R})$, on peut écrire son développement limité au premier ordre comme, pour $(\vec{x},\vec{y})\in\mathbb{R}^n\times\mathbb{R}^m$, $h\in\mathbb{R}$, $(\vec{v},\vec{w})\in\mathbb{R}^n\times\mathbb{R}^m$:

$$\begin{aligned} f((\vec{x},\vec{y})+h(\vec{v},\vec{w})) &= f(\vec{x},\vec{y})+h\,df(\vec{x},\vec{y}).(\vec{v},\vec{w})+o(h)\\ &= f(\vec{x},\vec{y})+h\,D_1f(\vec{x},\vec{y}).\vec{v}+h\,D_2f(\vec{x},\vec{y}).\vec{w}+o(h)\end{aligned}$$

où l'application linéaire $df(\vec{x},\vec{y}):\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ est écrite :

$$df(\vec{x},\vec{y}).(\vec{v},\vec{w})=D_1\vec{f}(\vec{x},\vec{y}).\vec{v}+D_2\vec{f}(\vec{x},\vec{y}).\vec{w}.$$

En d'autres termes, si à $\vec{y}$ fixé on pose $f_{\vec{y}}(\vec{x})=f(\vec{x},\vec{y})$, on a $D_1f(\vec{x},\vec{y})=df_{\vec{y}}(\vec{x})$; et si à $\vec{x}$ fixé on pose $f_{\vec{x}}(\vec{y})=f(\vec{x},\vec{y})$, on a $D_2f(\vec{x},\vec{y})=df_{\vec{x}}(\vec{y})$.

Exemple 7.11 $\mathbb{R}^3=\mathbb{R}\times\mathbb{R}^2$, i.e. $n=1$ et $m=2$.
On a $df(x,y,z).(v_1,v_2,v_3)=\frac{\partial f}{\partial x}(x,y,z)v_1+\frac{\partial f}{\partial y}(x,y,z)v_2+\frac{\partial f}{\partial z}(x,y,z)v_3$. Posons $\vec{x}=x$ et $\vec{y}=(y,z)$. Et on note $\vec{v}=v_1$ et $\vec{w}=(v_2,v_3)$. On a alors $D_1f(x,(y,z)).v_1=\frac{\partial f}{\partial x}(x,y,z)v_1$ et $D_2f(x,(y,z)).(v_2,v_3)=\frac{\partial f}{\partial y}(x,y,z)v_2+\frac{\partial f}{\partial z}(x,y,z)v_3$. Et on a $df(\vec{x},\vec{y}).(\vec{v},\vec{w})=D_1f(\vec{x},\vec{y}).\vec{v}+D_2f(\vec{x},\vec{y}).\vec{w}$. ∎

Exemple 7.12 Avec $\mathbb{R}^3=\mathbb{R}^2\times\mathbb{R}$, i.e. $n=2$ et $m=1$.
On a $df(x,y,z).(v_1,v_2,v_3)=\frac{\partial f}{\partial x}(x,y,z)v_1+\frac{\partial f}{\partial y}(x,y,z)v_2+\frac{\partial f}{\partial z}(x,y,z)v_3$. Posons $\vec{x}=(x,y)$ et $\vec{y}=z$. Et on note $\vec{v}=(v_1,v_2)$ et $\vec{w}=v_3$. On a alors $D_1f((x,y),z)).(v_1,v_2)=\frac{\partial f}{\partial x}(x,y,z)v_1+\frac{\partial f}{\partial y}(x,y,z)v_2$ et $D_2f((x,y),z).v_3=\frac{\partial f}{\partial z}(x,y,z)v_3$.
Et on a $df(\vec{x},\vec{y}).(\vec{v},\vec{w})=D_1f(\vec{x},\vec{y}).\vec{v}+D_2f(\vec{x},\vec{y}).\vec{w}$. ∎

7.2.2 Développement limités, fonction $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$

Et pour $\vec{f} \in C^1(\mathbb{R}^n \times \mathbb{R}^m; \mathbb{R}^p)$, on peut écrire son développement limité au premier ordre comme, pour $(\vec{x}, \vec{y}) \in \mathbb{R}^n \times \mathbb{R}^m$, $h \in \mathbb{R}$, $(\vec{v}, \vec{w}) \in \mathbb{R}^n \times \mathbb{R}^m$:

$$\vec{f}((\vec{x}, \vec{y}) + h(\vec{v}, \vec{w})) = \vec{f}(\vec{x}, \vec{y}) + h\, D_1\vec{f}(\vec{x}, \vec{y}).\vec{v} + h\, D_2\vec{f}(\vec{x}, \vec{y}).\vec{w} + o(h),$$

où l'application linéaire $d\vec{f}(\vec{x}, \vec{y}) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ est écrite $d\vec{f}(\vec{x}, \vec{y}).(\vec{v}, \vec{w}) = D_1\vec{f}(\vec{x}, \vec{y}).\vec{v} + D_2\vec{f}(\vec{x}, \vec{y}).\vec{w}$. (Dans une base de $\mathbb{R}^p$ on a ainsi p équations.) On a donc :

$$\vec{f}(\vec{x}, \vec{y}) = \vec{f}(\vec{x}_0, \vec{y}_0) + A.(\vec{x}-\vec{x}_0) + B.(\vec{y}-\vec{y}_0) + o(||(\vec{x}-\vec{x}_0), (\vec{y}-\vec{y}_0)||,$$

où $A : \mathbb{R}^n \to \mathbb{R}^p$ et $B : \mathbb{R}^m \to \mathbb{R}^p$ sont les applications linéaires $A = D_1\vec{f}(\vec{x}_0, \vec{y}_0)$ et $B = D_2\vec{f}(\vec{x}_0, \vec{y}_0)$.

7.2.3 Application : fonction $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$

En particulier, quand $m = p$, l'application linéaire $B = D_2 f$ est un endomorphisme de $\mathbb{R}^m$, et s'il est inversible, on a localement :

$$\vec{y} = \vec{y}_0 + B^{-1}.(\vec{f}(\vec{x}, \vec{y}) - \vec{f}(\vec{x}_0, \vec{y}_0) - A.(\vec{x}-\vec{x}_0)) + o(||(\vec{x}-\vec{x}_0), (\vec{y}-\vec{y}_0)||).$$

Et le long d'une courbe de niveau $\vec{f}(\vec{x}, \vec{y}) = \vec{c}$, on a ainsi explicité $\vec{y}$ en fonction de $\vec{x}$ au voisinage du point $(\vec{x}_0, \vec{y}_0)$:

$$\vec{y} = \vec{y}_0 - B^{-1}.A.(\vec{x}-\vec{x}_0) + o(||(\vec{x}-\vec{x}_0), (\vec{y}-\vec{y}_0)||) \overset{\text{noté}}{=} \vec{g}(\vec{x}).$$

Le théorème s'énonce 7.4 se généralise alors comme (avec ici $D_2 f(\vec{x}, \vec{y})$ endomorphisme inversible) :

Théorème 7.13 *Soit Ω ouvert de $\mathbb{R}^n \times \mathbb{R}^m$ et $\vec{f} : \Omega \to \mathbb{R}^m$ une fonction $C^1(\Omega, \mathbb{R}^m)$. Soient $(\vec{x}_0, \vec{y}_0) \in \Omega$ et $\vec{c} \in \mathbb{R}^m$ tels que :*

$$\vec{f}(\vec{x}_0, \vec{y}_0) = \vec{c} \quad \text{et} \quad D_2\vec{f}(\vec{x}_0, \vec{y}_0) \text{ isomorphisme de } \mathbb{R}^m,$$

où $D_2\vec{f}(\vec{x}_0, \vec{y}_0)$ désigne la différentielle de $\vec{f}$ par rapport à la variable $\vec{y}$ au point $(\vec{x}_0, \vec{y}_0)$, i.e. la différentielle $d\vec{f}_{\vec{x}_0}(\vec{y}_0) : \mathbb{R}^m \to \mathbb{R}^m$ de l'application $\vec{f}_{\vec{x}_0} : \vec{y} \to \vec{f}_{\vec{x}_0}(\vec{y}) = f(\vec{x}_0, \vec{y})$.

Alors il existe un voisinage ouvert $V_1 \times V_2$ de $(\vec{x}_0, \vec{y}_0)$, et il existe une fonction dérivable $\vec{g} : V_1 \to V_2$ telle que :

$$\forall (\vec{x}, \vec{y}) \in V_1 \times V_2, \quad \vec{f}(\vec{x}, \vec{y}) = \vec{c} \quad \Longleftrightarrow \quad \vec{y} = \vec{g}(\vec{x}).$$

Et donc localement, la "surface de niveau" $\vec{f}(\vec{x}_0, \vec{y}_0) = \vec{c}$ est représentable par une fonction $\vec{y} = \vec{g}(\vec{x})$. Et avec $\vec{f}(\vec{x}, g(\vec{x})) = 0$ on obtient :

$$D_1\vec{f}(\vec{x}, g(\vec{x})) + D_2\vec{f}(\vec{x}, g(\vec{x})) \circ d\vec{g}(\vec{x}) = 0. \tag{7.4}$$

Ou encore, en termes de matrices jacobiennes :

$$[D_1\vec{f}(\vec{x}, g(\vec{x}))] + [D_2\vec{f}(\vec{x}, g(\vec{x}))].[d\vec{g}(\vec{x})] = 0. \tag{7.5}$$

Preuve. On regarde ici le cas $n = m = p = 1$, la démarche dans le cas général étant la même. On va appliquer le théorème d'inversion locale à la fonction, pour $\Omega \subset \mathbb{R} \times \mathbb{R}$:

$$\vec{h} : \begin{cases} \quad \Omega & \to \mathbb{R} \times \mathbb{R}, \\ (x, y) & \mapsto \vec{h}(x, y) = \begin{pmatrix} x \\ f(x, y) \end{pmatrix}. \end{cases}$$

La matrice jacobienne de $\vec{h}$ est $d\vec{h} = \begin{pmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{pmatrix}$ de déterminant (le jacobien) en (x_0, y_0) qui vaut $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Donc $\vec{h}$ est (localement) inversible d'un voisinage $V \subset \Omega$ du point (x_0, y_0) à valeurs dans $W = \vec{h}(V)(\subset \mathbb{R}^2)$.

On pose alors :

$W_1 = \{s \in \mathbb{R} : (s,c) \in W\}$ (section de W), et

$g : x \in W_1 \to g(x) = y$ où y t.q. $\vec{h}(x,y) = (x,c)$ existe, est unique, car h inversible.

(On rappelle que $c = f(x,y)$.) La fonction g est C^1 car h^{-1} l'est, et par construction, $\vec{h}(x,g(x)) = (x,c)$ et donc $f(x,g(x)) = c$. D'où le résultat.

On renvoie à Avez [2] pour les détails. ▪

Exercice 7.14 Soit $\vec{f} : \left\{ \begin{array}{cl} \mathbb{R}^3 & \to \mathbb{R}^2 \\ (x,y,z) & \mapsto \begin{pmatrix} f_1(x,y,z) = x + e^{y+z} - 1 \\ f_2(x,y,z) = x^2 + y^2 + z^2 - x - 2y - z \end{pmatrix} \end{array} \right\}$. À x fixé, on pose $\vec{g}(y,z) = \vec{f}(x,y,z)$. Calculer la matrice jacobienne de $\vec{g}$. Montrer qu'il existe un intervalle ouvert A contenant 0, un voisinage ouvert B contenant $(0,0)$ et une fonction $\vec{\varphi} : x \in A \to \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}$ de classe C^1 tels que $\vec{\varphi}(0) = (0,0)$ et $\vec{f}(x, \varphi_1(x), \varphi_2(x)) = 0$ pour tout $x \in A$. Puis calculer $\vec{\varphi}'(0)$.

Réponse. $\vec{f}$ est trivialement C^1 ainsi que $\vec{g}$. On a $[dg(y,z)] = \begin{pmatrix} e^{y+z} & e^{y+z} \\ 2y-2 & 2z-1 \end{pmatrix}$, d'où $[d\vec{g}(0,0)] = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}$ de detérminant $\neq 0$. On peut donc appliquer le théorème des fonctions implicites : ayant $\vec{f}(0,(0,0)) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, il existe A un intervalle ouvert contenant 0, B un intervalle ouvert contenant $(0,0)$ et une application $\vec{\varphi} : A \to B$ tels que $(y,z) = \vec{\varphi}(x)$ équivaut à $\vec{f}(x,(y,z)) = 0$ pour tout $(x,(y,z)) \in A \times B$, avec donc $\vec{\varphi}(0) = (0,0)$. Et $\vec{f}$ et $\vec{\varphi}$ vérifient $\vec{f}(x, \vec{\varphi}(x)) = 0$, i.e. $\left\{ \begin{array}{l} f_1(x,\varphi_1(x),\varphi_2(x)) = 0 \\ f_2(x,\varphi_1(x),\varphi_2(x)) = 0 \end{array} \right.$, pour tout $x \in A$.

On en déduit en dérivant en x que, notant $T = (x, \varphi_1(x), \varphi_2(x))$, $\frac{\partial f_i}{\partial x}(T) + \frac{\partial f_i}{\partial y}(T)\varphi_1'(x) + \frac{\partial f_i}{\partial z}(T)\varphi_2'(x) = 0$, pour $i = 1,2$, et donc pour $x = 0$ et $T = \vec{0}$: $1 + \varphi_1'(0) + \varphi_2'(0) = 0$ et $-1 - 2\varphi_1'(0) - \varphi_2'(0) = 0$. Et donc $\varphi_1'(0) = 0$ et $\varphi_2'(0) = -1$, i.e. $\vec{\varphi}'(0) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. ▪

Exercice 7.15 Soit $\vec{f} : \mathbb{R}^3 \to \mathbb{R}^2$ donnée par : $\vec{f}(t,x,y) = \begin{pmatrix} f_1(t,x,y) = x^2 + t^2 + 2y - 2t \\ f_2(t,x,y) = y^2 - x \end{pmatrix}$. Montrer qu'il existe un intervalle ouvert $I \subset \mathbb{R}$ contenant 0, un voisinage ouvert J de $(0,0)$ dans $\mathbb{R}^2$ et une application $\vec{\varphi} : I \to J$, où on note $\vec{\varphi}(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \end{pmatrix}$, tels que : $\vec{\varphi}(0) = \vec{0}$, et pour tout $t \in I$, $\vec{f}(t, \varphi_1(t), \varphi_2(t)) = \vec{0}$. Calculer $\vec{\varphi}\,'(0)$.

Réponse. Ici $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R}^2$. On pose $t = t$ et $\vec{y} = (x,y)$. 0n a $[D_{\vec{y}} f(t,\vec{y})] = \begin{pmatrix} 2x & 2 \\ -1 & 2y \end{pmatrix}$, ici indépendant de t. En particulier en $(t,\vec{y}) = (0,\vec{0})$, on a $\det([D_2 f(0,\vec{)}]) = 2 \neq 0$, et on peut appliquer le théorème des fonctions implicites. D'où l'existence de I, J et $\vec{\varphi} : t \in I \to (x,y) \in J$ tels que $f(t, \varphi_1(t), \varphi_2(t)) = 0$. Par dérivation de fonctions composées on trouve $\varphi_1'(0)$ et $\varphi_2'(0)$... ▪

Exercice 7.16 Soit $F : \mathbb{R}^3 \to \mathbb{R}^2$ donnée par $(x,y,z) \mapsto \begin{pmatrix} f_1(x,y,z) = x^2 - y^2 + z^2 - 1 \\ f_2(x,y,z) = xyz - 1 \end{pmatrix}$. Montrer qu'il existe un intervalle ouvert A contenant 1 et un voisinage ouvert B de $(1,1)$ dans $\mathbb{R}^2$ et une application $\Phi : A \to B$ tels que $\Phi(1) = (1,1)$ et que pour tout $x \in A$ on ait $F(x, \varphi_1(x), \varphi_2(x)) = 0$. Donner la dérivée $\Phi'(1) = (\varphi_1'(1), \varphi_2'(1))$ de Φ en 1. ▪

8 Changement de variables

On notera $(\vec{e}_1, ..., \vec{e}_n)$ la base canonique de $\mathbb{R}^n$. On rappelle que la base canonique est la base construite sur le produit cartésien $\mathbb{R}^n = \mathbb{R} \times ... \times \mathbb{R}$ (n-fois) à partir de la base "1" de $\mathbb{R}$ (élément neutre de la multiplication dans $\mathbb{R}$), i.e. c'est la base formée des vecteurs $\vec{e}_i = (0, ..., 0, 1, 0, ...0)$ où seule la i-ème composante est non nulle et vaut 1.

Ainsi : $\vec{x} \in \mathbb{R}^n$ équivaut à :

il existe $(x_1, ..., x_n) \in \mathbb{R} \times ... \times \mathbb{R}$ tel que $\vec{x} = (x_1, ..., x_n)$, ce qui s'écrit :

$\vec{x} = x_1(1, 0, ..., 0) + ... + x_n(0, ..., 0, 1)$, ou encore :

$\vec{x} = x_1\vec{e}_1 + ... + x_n\vec{e}_n$ (par définition de la base canonique).

8.1 Définitions et résultats

On se donne deux ouverts $U \subset \mathbb{R}^n$ et $V \subset \mathbb{R}^n$.

8.1.1 Changement de variables

Définition 8.1 Un changement de variables est une fonction $\vec{\psi} : U \to V$ qui est un difféomorphisme, i.e. une fonction $\vec{\psi}$ qui est bijective et C^1 telle que $\vec{\psi}^{-1} : V \to U$ soit également C^1.

U est appelé espace des paramètres, et V l'espace géométrique.

Les paramètres $(u_1, ..., u_n) \in U$ sont appelés coordonnées curvilignes, et les $(x_1, ..., x_n) \in V$ les coordonnées cartésiennes.

Exemple 8.2 (Coordonnées polaires.) Soient $(u, v) = (\rho, \theta) \in U = \mathbb{R}_+^* \times]-\pi, \pi[$, et $\vec{\psi}(\rho, \theta) = \begin{pmatrix} \psi_1(\rho, \theta) = \rho \cos\theta =^{\text{noté}} x(\rho, \theta) \\ \psi_2(\rho, \theta) = \rho \sin\theta =^{\text{noté}} y(\rho, \theta) \end{pmatrix}$, où $V = \mathrm{Im}\vec{\psi} = \mathbb{R}^2 - \{(x, 0) : x \leq 0\}$ (= $\mathbb{R}^2$ privé du demi-axe $x \leq 0$). L'espace des paramètres est l'espace des (ρ, θ), et l'espace géométrique est l'espace des (x, y). ▪

Remarque 8.3 L'application $f :]-1, 1[\to]-1, 1[$ définie par $f(u) = u^3$ est C^1, bijective d'inverse $f^{-1}(x) = x^{\frac{1}{3}}$ qui est C^0 sur $]-1, 1[$ mais n'est pas C^1 en 0 (pente $+\infty$ en 0). La proposition suivante rappellera qu'il faut $f'(u) \neq 0$ pour tout u (pas de pente horizontale pour f) pour que : $f \in C^1$ et f bijective donne $f^{-1} \in C^1$ (théorème d'inversion locale 6.1). ▪

Notons $\vec{u} = (u_1, ..., u_n) \in U$ et $\vec{x} = (x_1, ..., x_n) \in V$. Puis :

$$\vec{\psi}(\vec{u}) = \sum_{i=1}^{n} \psi_i(\vec{u})\ \vec{e}_i = \begin{pmatrix} \psi_1(\vec{u}) \\ \vdots \\ \psi_n(\vec{u}) \end{pmatrix} \overset{\text{noté}}{=} \begin{pmatrix} x_1(\vec{u}) \\ \vdots \\ x_n(\vec{u}) \end{pmatrix} \overset{\text{noté}}{=} \vec{x}(\vec{u}) \in V,$$

où les n fonctions $\psi_i : \vec{u} \in U \to \psi_i(\vec{u}) \in \mathbb{R}$ sont données de telle sorte que $\vec{\psi}$ est un changement de variables (= un difféomorphisme).

Proposition 8.4 *1- Si $\vec{\psi}$ est un changement de variable alors $\det(J_\psi(\vec{u})) \neq 0$ pour tout $\vec{u} \in U$.*

2- Si $\vec{\psi} \in C^1(U, V)$, si $\vec{\psi}$ est bijective et si $\vec{\psi}$ vérifie $\det(J_\psi(\vec{u})) \neq 0$ pour tout $u \in U$ alors $\vec{\psi}^{-1} \in C^1(V, U)$.

3- Si $\vec{\psi} \in C^1(U, V)$ et $\vec{\psi}$ vérifie $\det(J_\psi(\vec{u})) \neq 0$ pour tout $\vec{u} \in U$, alors localement au voisinage de chaque point $\vec{u}$, $\vec{\psi}$ définit un changement de variables, i.e. $\forall \vec{u} \in U$, $\exists r > 0$, $\vec{\psi} : B(\vec{u}, r) \to \mathrm{Im}(B(\vec{u}, r))$ définit un changement de variables, où $B(\vec{u}, r)$ est la boule ouverte de centre $\vec{u}$ et de rayon r.

Preuve. 1- Soit $\vec{\psi}$ un difféomorphisme, alors $\vec{\psi} \in C^1$ ainsi que $\vec{\psi}^{-1}$ et $\vec{\psi}^{-1} \circ \vec{\psi} = I$, et $\det(J_{\vec{\psi}^{-1}}(x, u)). \det(J_{\vec{\psi}}(u, v)) = 1$, donc les déterminants sont non nuls (on a appliqué la proposition 3.15).

2- C'est le théorème d'inversion locale 6.2 lorsque $\vec{\psi}$ est bijective.

3- Supposons que $\vec{\psi} \in C^1$ et $\det(J_{\vec{\psi}}(\vec{u})) \neq 0$ pour tout $\vec{u} \in U$. Alors on applique le théorème d'inversion locale 6.2 : il existe $r > 0$ tel que $\vec{\psi}$ est inversible sur $B(\vec{u}, r)$. ▪

Exercice 8.5 Soit $G : \mathbb{R}^3 \to \mathbb{R}^3$ application définie par $G(u, v, w) = \begin{pmatrix} u + v + w^2 \\ v + w + u^2 \\ w + u + v^2 \end{pmatrix}$. Montrer que G définit un difféomorphisme dans un voisinage de $\vec{0}$ (on dit que G définit un difféomorphisme local au voisinage de $\vec{0}$).

Réponse. G est C^1 sur $\mathbb{R}^2$ donc en $\vec{0}$, donc J_G est C^0 sur $\mathbb{R}^2$ donc en $\vec{0}$. Et $\det(J_G(\vec{0})) = 2$, donc $\det(J_G(\vec{0}))>0$ dans un voisinage de $\vec{0}$. Donc on peut appliquer le théorème d'inversion locale 6.2 ou son expression donnée par la proposition 8.4. On pourra également montrer que pour $|u|, |v|, |w| < \frac{1}{4}$, on a $\det(J_G(\vec{u})) > 0$. ▪

8.1.2 Coordonnées polaires

On se donne un répère $(O,\vec{e}_1,\vec{e}_2)$ dans l'espace affine $\mathbb{R}^2$. La base $(\vec{e}_1,\vec{e}_2)$ est la base canonique, et on dispose également du produit scalaire canonique, i.e. de la forme bilinéaire $g:\mathbb{R}^2\to\mathbb{R}$ définie positive qui est définie sur les vecteurs de base par $g(\vec{e}_i,\vec{e}_j)=\delta_{ij}$ pour $j=1,2$. Ce produit scalaire canonique est usuellement noté $g(\vec{v},\vec{w})=(\vec{v},\vec{w})$, et permet de définir les angles à l'aide du cosinus.

Définition 8.6 Les coordonnées polaires d'un point $\vec{x}\in\mathbb{R}^2$ sont par définition, la distance $\rho=||\vec{x}||$ par rapport à l'origine, et l'angle $\theta=\widehat{(\vec{e}_1,\vec{x})}$ (entre $\vec{e}_1$ et $\vec{x}$).

Un point $\vec{x}=(x,y)$ est donc repéré par sa distance ρ à l'origine et son angle θ par rapport au demi axe des abscisses positives, et sera noté "$\vec{x}=\vec{x}(\rho,\theta)=(x(\rho,\theta),y(\rho,\theta))$".

On notera dans la suite $\vec{u}=(\rho,\theta)$ et on applique les définitions du paragraphe 8.1 précédent.

La fonction permettant de passer des coordonnées polaires aux coordonnées cartésiennes est par exemple (à une distance ρ et un angle θ on associe un point $\vec{x}=\vec{\psi}(\rho,\theta)$) :

$$\vec{\psi}:\begin{cases} U=]0,\infty[\times]-\pi,\pi[\;\longrightarrow\; V=\mathbb{R}^2-\{0\} \\ \vec{u}=\begin{pmatrix}\rho\\ \theta\end{pmatrix} \longmapsto \vec{\psi}(\rho,\theta)=\begin{pmatrix}\psi_1(\rho,\theta)\overset{\text{déf}}{=}\rho\cos\theta\overset{\text{noté}}{=}x(\rho,\theta)\\ \psi_2(\rho,\theta)\overset{\text{déf}}{=}\rho\sin\theta\overset{\text{noté}}{=}y(\rho,\theta)\end{pmatrix}\overset{\text{noté}}{=}\vec{x}(\rho,\theta),\end{cases}\tag{8.1}$$

où U est $\mathbb{R}^2$ privé du demi-axe des x négatifs. Cette application est bijective, d'application inverse :

$$\vec{\psi}^{-1}:\begin{cases} V=\mathbb{R}^2-\{0\} \;\longrightarrow\; U=]0,\infty[\times]-\pi,\pi[\\ \begin{pmatrix}x\\ y\end{pmatrix} \longmapsto \vec{\psi}^{-1}(x,y)=\begin{pmatrix}(\psi^{-1})_1(x,y)\overset{\text{déf}}{=}\sqrt{x^2+y^2}\overset{\text{noté}}{=}\rho(x,y)\\ (\psi^{-1})_2(x,y)\overset{\text{déf}}{=}2\arctan\dfrac{y}{x+\sqrt{x^2+y^2}}\overset{\text{noté}}{=}\theta(x,y)\end{pmatrix}\end{cases}\tag{8.2}$$

La fonction $\vec{\psi}$ définit un difféomorphisme : on a exhibé son inverse, et on compose des fonctions C^1 dans des ouverts.

Remarque 8.7 Si on souhaite récupérer un difféomorphisme dans un voisinage du demi-axe des x négatifs, il suffit de considérer de considérer $\tilde{\vec{\psi}}$ où $U=]0,\infty[\times]0,2\pi[$ (i.e. avec $\theta\in]0,2\pi[$), qui dans ce cas exclut le demi-axe des x positifs.

De manière générale, la fonction $\vec{\psi}$ étant périodique de période 2π, on peut tout aussi bien considérer $\vec{\psi}$ avec $U=]0,\infty[\times]\theta_0,\theta_0+2\pi[$ où θ_0 est un réel quelconque. ▪

La formule donnant $\theta=2\arctan\frac{y}{x+\sqrt{x^2+y^2}}$ appelle quelques commentaires :

1- la fonction tangente : $\theta\to\tan\theta=\frac{\sin\theta}{\cos\theta}$ est bijective de $]-\frac{\pi}{2},\frac{\pi}{2}[\to]-\infty,\infty[$ (le cosinus s'annule pour $\theta=\pm\frac{\pi}{2}$). Or on voudrait avoir une relation donnant θ pour $\theta\in]-\pi,\pi[$, i.e. pour $\frac{\theta}{2}\in]-\frac{\pi}{2},\frac{\pi}{2}[$: la variable adaptée à notre problème est donc $\frac{\theta}{2}$. Et on a :

$$\tan(\frac{\theta}{2})=\frac{\sin\frac{\theta}{2}}{\cos\frac{\theta}{2}}=\frac{\sin\frac{\theta}{2}\cos\frac{\theta}{2}}{\cos^2\frac{\theta}{2}}=\frac{\frac12\sin(\theta)}{\frac12(\cos(\theta)+1)}=\frac{\frac{y}{\rho}}{\frac{x}{\rho}+1}=\frac{y}{x+\sqrt{x^2+y^2}}.$$

D'où le résultat annoncé.

2- La fonction $\theta(x,y)=2\arctan\frac{y}{x+\sqrt{x^2+y^2}}$ n'est pas très utilisée. On préfère souvent utiliser la formule usuelle $\theta=\arctan\frac{y}{x}$ (qui dit que : $\tan\theta=\frac{\text{côté opposé}}{\text{côté adjacent}}$).

Mais comme arctan : $]-\infty,+\infty[\to]-\frac{\pi}{2},\frac{\pi}{2}[$ ne peut donner que des valeurs de $\theta\in]-\frac{\pi}{2},\frac{\pi}{2}[$ (correspondant à $x>0$), on distingue les cas suivants :

$$\theta=\begin{cases}\arctan\dfrac{y}{x}, & \text{si } x>0,\\ \pi+\arctan\dfrac{y}{x}, & \text{si } x<0,\quad (\arctan\dfrac{y}{x}=\arctan\dfrac{-y}{-x}\text{ donne le }\theta\text{ du point opposé }(-x,-y)).\\ \dfrac{\pi}{2}, & \text{si } x=0 \text{ et } y>0,\\ -\dfrac{\pi}{2}, & \text{si } x=0\text{ et } y<0.\end{cases}$$

Exercice 8.8 Montrer que la matrice jacobienne de $\vec{\psi}$ en (ρ,θ) est donnée par $J_{\vec{\psi}}(\rho,\theta) = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$, avec $\det(J_{\vec{\psi}}(\rho,\theta)) = \rho$, et que $J_{\vec{\psi}^{-1}}(x,y) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{\rho}\sin\theta & \frac{1}{\rho}\cos\theta \end{pmatrix}$, où $\rho = \rho(x,y)$ et $\theta = \theta(x,y)$, avec $\det(J_{\vec{\psi}^{-1}}(x,y)) = \frac{1}{\rho(x,y)}$. ▪

Exercice 8.9 Soit $f(x,y) = xy\frac{x^2-y^2}{x^2+y^2}$ et $f(0,0) = 0$. Utiliser le changement de variables des coordonnées polaires pour montrer que f est $C^1(\mathbb{R}^2;\mathbb{R})$.

Réponse. Réponse. f est trivialement C^1 dans $\mathbb{R}^2 - \{\vec{0}\}$. On s'intéresse donc au voisinage de $\vec{0}$.

On définit $g(\rho,\theta) = f(x,y)$, i.e. $g = f\circ\vec{\psi}$ avec $(x,y) = \vec{\psi}(\rho,\theta)$ où $\vec{\psi}$ définit le changement de coordonnées polaire, avec de plus $g(0,0) = 0 = f(0,0)$ (a priori le changement polaire n'est pas défini pour $\rho = 0$). Donc :

$$g(\rho,\theta) = f(\rho\cos\theta,\rho\sin\theta) = \rho^2\cos\theta\sin\theta(\cos^2\theta - \sin^2\theta) = \frac{1}{4}\rho^2\sin(4\theta).$$

Il est immédiat qu'à θ fixé on a $g(\rho,\theta) \to 0$ quand $\rho \to 0_+$, i.e. que $f(x,y) \to 0$ quand $(x,y) \to \vec{0}$ dans la direction fixée donnée par θ. Cette direction θ étant qcq, on a f continue en 0.

Par dérivation :

$$\frac{\partial g}{\partial\rho}(\rho,\theta) = \frac{1}{2}\rho\sin(4\theta), \quad \frac{\partial g}{\partial\theta}(\rho,\theta) = \rho^2\cos(4\theta).$$

D'où, avec $f = g\circ\vec{\psi}^{-1}$ et donc $J_f(x,y) = J_g(\rho,\theta).J_{\vec{\psi}^{-1}}(x,y)$, on a :

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y) = \dfrac{1}{2}\rho\cos\theta\sin(4\theta) - \rho\sin\theta\cos(4\theta), \\ \dfrac{\partial f}{\partial y}(x,y) = \dfrac{1}{2}\rho\sin\theta\sin(4\theta) - \rho\cos\theta\cos(4\theta), \end{cases}$$

où $\rho = \rho(x,y)$ et $\theta = \theta(x,y)$. Et quand $\rho \to 0$ à θ fixé (qcq), on a $\frac{\partial f}{\partial x}(x,y)$ qui tend vers 0 donc si $\frac{\partial f}{\partial x}(0,0)$ existe et vaut 0, on a $\frac{\partial f}{\partial x}$ qui est une fonction continue en $\vec{0}$. Et on vérifie immédiatement que $\frac{f(h,0)-f(0,0)}{h} = 0 \to 0$ quand $h \to 0$ donc $\frac{\partial f}{\partial x}(0,0)$ existe et vaut 0. De même pour $\frac{\partial f}{\partial y}$. Donc les dérivées partielles sont continues sur $\mathbb{R}^2$, donc $f \in C^1(\mathbb{R}^2;\mathbb{R})$, voir théorème 2.20. ▪

8.1.3 Coordonnées cylindriques

Ce sont les coordonnées polaires plongées dans $\mathbb{R}^3$: il ne se passe rien en z :

$$\vec{\psi} : \begin{cases} U =]0,\infty[\times]-\pi,\pi[\times\mathbb{R} \longrightarrow V = \mathbb{R}^2-\{0\}\times\mathbb{R} \\ \vec{u} = \begin{pmatrix} \rho \\ \theta \\ z \end{pmatrix} \mapsto \vec{\psi}(\rho,\theta,z) = \begin{pmatrix} \psi_1(\rho,\theta,z) \overset{\text{déf}}{=} \rho\cos\theta \overset{\text{noté}}{=} x(\rho,\theta) \\ \psi_2(\rho,\theta,z) \overset{\text{déf}}{=} \rho\sin\theta \overset{\text{noté}}{=} y(\rho,\theta) \\ \psi_3(z) \overset{\text{déf}}{=} z \end{pmatrix} \overset{\text{noté}}{=} \vec{x}(\rho,\theta,z), \end{cases} \tag{8.3}$$

Les variables (ρ,θ) sont totalement découplées de la variable z, et traiter des coordonnées cylindriques n'est techniquement pas autre chose que traiter des coordonnées polaires, où on ajoute un composante verticale (en z) qui est l'identité : $\varphi_3(z) = z$.

8.1.4 Coordonnées sphériques

On se place dans $\mathbb{R}^3$. Les coordonnées sphériques sont basées sur les coordonnées polaires : un vecteur $\vec{x} \in \mathbb{R}^3$ s'écrit

$$\vec{x} = \vec{\psi}(\rho,\theta,\varphi) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \rho\cos\theta\cos\varphi \\ \rho\sin\theta\cos\varphi \\ \rho\sin\varphi \end{pmatrix} = \begin{pmatrix} \rho\cos\theta \\ \rho\sin\theta \\ 0 \end{pmatrix}\cos\varphi + \begin{pmatrix} 0 \\ 0 \\ \rho \end{pmatrix}\sin\varphi. \tag{8.4}$$

Ce pour :

$$\rho \in]0,\infty[, \quad \theta \in]-\pi,\pi[, \quad \varphi \in]-\frac{\pi}{2},\frac{\pi}{2}[.$$

Sur la sphère terrestre, θ définit la longitude (sur les parallèles) et φ la latitude (sur les méridiens).

Et on a immédiatement :

$$\rho = \sqrt{x^2+y^2+z^2}, \qquad \theta = 2\arctan\frac{y}{x+\sqrt{x^2+y^2}}, \qquad \varphi = \arctan\frac{z}{\sqrt{x^2+y^2}}. \tag{8.5}$$

(Ou plus simplement $\theta = \arctan\dfrac{y}{x}$, si $x > 0$, voir § précédent.)

Remarque 8.10 Au lieu de $\vec{x} = \vec{\psi}(\rho,\theta,\varphi) = \begin{pmatrix} x = \rho\cos\theta\cos\varphi \\ y = \rho\sin\theta\cos\varphi \\ z = \rho\sin\varphi \end{pmatrix}$, les mécaniciens considèrent également :

$$\vec{x} = \tilde{\vec{\psi}}(\rho,\theta,\tilde{\varphi}) = \begin{pmatrix} \tilde{x} = \rho\cos\theta\sin\tilde{\varphi} \\ \tilde{y} = \rho\sin\theta\sin\tilde{\varphi} \\ \tilde{z} = \rho\cos\tilde{\varphi} \end{pmatrix}, \rho \in]0,\infty[, \quad \theta \in]-\pi,\pi[, \quad \tilde{\varphi} \in]0,\pi[, \tag{8.6}$$

et l'angle $\tilde{\varphi}$ mesure l'angle entre le pôle nord et un parallèle. Si c'est cette formule que vous comptez utiliser, comme $\varphi = \pi/2 - \tilde{\varphi}$, dans la suite du cours on remplacera φ par sa valeur en $\tilde{\varphi}$: en effet, $\cos\varphi = \sin\tilde{\varphi}$, $\sin\varphi = \cos\tilde{\varphi}$.

Mais attention aux signes quand on dérive : $\frac{\partial\vec{\psi}}{\partial\varphi}(\rho,\theta,\varphi) = -\frac{\partial\tilde{\vec{\psi}}}{\partial\tilde{\varphi}}(\rho,\theta,\tilde{\varphi})$:
$\frac{\partial\tilde{x}}{\partial\tilde{\varphi}} = \rho\cos\theta\cos\tilde{\varphi}$ $(= +\rho\cos\theta\sin\varphi)$, alors que $\frac{\partial x}{\partial\varphi} = -\rho\cos\theta\sin\varphi$,
$\frac{\partial\tilde{y}}{\partial\tilde{\varphi}} = \rho\sin\theta\cos\tilde{\varphi}$ $(= +\rho\sin\theta\sin\varphi)$, alors que $\frac{\partial y}{\partial\varphi} = -\rho\sin\theta\sin\varphi$, et
$\frac{\partial\tilde{z}}{\partial\tilde{\varphi}} = -\rho\cos\tilde{\varphi}$ $(= -\rho\sin\varphi)$, alors que $\frac{\partial z}{\partial\varphi} = +\rho\sin\varphi$:
tous les signes de dérivation en $\tilde{\varphi}$ sont inversés par rapport aux dérivations en φ. ▪

Remarque 8.11 Pour θ, voir remarque 8.7. ▪

Exemple 8.12 Montrer que pour les coordonnées sphériques, les matrices jacobiennes sont données par :

$$J_{\vec{\psi}}(\rho,\theta,\varphi) = \begin{pmatrix} \cos\theta\cos\varphi & -\rho\sin\theta\cos\varphi & -\rho\cos\theta\sin\varphi \\ \sin\theta\cos\varphi & \rho\cos\theta\cos\varphi & -\rho\sin\theta\sin\varphi \\ \sin\varphi & 0 & \rho\cos\varphi \end{pmatrix}, \tag{8.7}$$

et que les vecteurs colonnes $J_{\vec{\psi}}(\vec{u}).\vec{e}_j$ constituant la jacobienne sont orthogonaux et ont pour norme 1, $\rho\cos\varphi$ et ρ.

Montrer que le jacobien (déterminant de la jacobienne) est donné par :

$$\det(J_{\vec{\psi}}(\vec{u})) = \rho^2\cos\varphi.$$

Et :

$$J_{\vec{\psi}}^{-1}(x,y,z) = \begin{pmatrix} \cos\theta\cos\varphi & \sin\theta\cos\varphi & \sin\varphi \\ -\frac{\sin\theta}{\rho\cos\varphi} & \frac{\cos\theta}{\rho\cos\varphi} & 0 \\ -\frac{\sin\varphi\cos\theta}{\rho} & -\frac{\sin\varphi\sin\theta}{\rho} & \frac{\cos\varphi}{\rho} \end{pmatrix}. \tag{8.8}$$

On rappelle que le déterminant donne le volume formé par les vecteurs colonnes de la matrice, et qu'en particulier, quand ces colonnes sont des vecteurs orthogonaux le volume est donné par le produit des normes des vecteurs colonnes.

On rappelle que pour une matrice dont les colonnes sont orthogonales, la matrice inverse est obtenue à partir de la matrice transposée en divisant les lignes de la matrice transposée par la norme au carré. En effet, $\begin{pmatrix} \frac{1}{||\vec{v}_1||^2}\vec{v}_1^T \\ \vdots \\ \frac{1}{||\vec{v}_n||^2}\vec{v}_n^T \end{pmatrix} . \Big(\begin{pmatrix} \vec{v}_1 \end{pmatrix} \ \ ... \begin{pmatrix} \vec{v}_n \end{pmatrix} \Big) = [\frac{\vec{v}_i^t.\vec{v}_j}{||\vec{v}_i||^2}] = Id$ lorsque $\vec{v}_i \perp \vec{v}_j$ pour tout $i \neq j$.

On rappelle aussi que lorsque les colonnes sont orthogonales, il n'y a aucune raison pour que les lignes le soit, à moins que les vecteurs n'aient même norme. Exemple $\begin{pmatrix} 2 & -2 \\ 4 & 1 \end{pmatrix}$ dont les colonnes sont orthogonales, mais pas les lignes, ou encore la matrice $J_{\vec{\psi}}$ dont les colonnes sont orthogonales et pas les lignes. ▪

Exercice 8.13 Montrer que pour $\tilde{\vec{\psi}}$ défini en (8.6), on a :

$$J_{\tilde{\vec{\psi}}}(\rho,\theta,\varphi) = \begin{pmatrix} \cos\theta\sin\tilde{\varphi} & -\rho\sin\theta\sin\tilde{\varphi} & \rho\cos\theta\cos\tilde{\varphi} \\ \sin\theta\sin\tilde{\varphi} & \rho\cos\theta\sin\tilde{\varphi} & \rho\sin\theta\cos\tilde{\varphi} \\ \cos\tilde{\varphi} & 0 & -\rho\sin\tilde{\varphi} \end{pmatrix}, \qquad \tilde{\varphi} \in]0,\pi[,$$

de déterminant (jacobien) $-\rho^2\sin\tilde{\varphi}$. ▪

8.2 Exemples

Exemple 8.14 En utilisant les coordonnées polaires, déterminer les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$ solutions de $(x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y})(x,y) = \frac{\sqrt{x^2+y^2}}{1+x^2+y^2}$.

Réponse. On pose $g(r,\theta) = f(x,y) = (f\circ\vec{\psi})(r,\theta)$ où $\vec{\psi}(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$. On note $\vec{x} = (x,y)$ et $\vec{u} = (r,\theta)$. On a $J_{\vec{\psi}}(\vec{u}) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$ d'où $J_{\vec{\psi}^{-1}}(\vec{x}) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{r}\sin\theta & \frac{1}{r}\cos\theta \end{pmatrix}(\vec{x}) = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix}(\vec{x})$. (Calcul de $\frac{\partial\theta}{\partial x}(\vec{x})$ et $\frac{\partial\theta}{\partial y}(\vec{x})$ plus agréable qu'à l'aide des formules (8.2).

D'où $\frac{\partial f}{\partial x}(\vec{x}) = (\frac{\partial g}{\partial r}(\vec{u})\frac{\partial r}{\partial x}(\vec{x}) + \frac{\partial g}{\partial \theta}(\vec{u})\frac{\partial \theta}{\partial x}(\vec{x}) = \frac{\partial g}{\partial r}(\vec{u})\cos\theta(\vec{x}) - \frac{\partial g}{\partial \theta}(\vec{u})\frac{\sin\theta}{r}(\vec{x})$ et de même $\frac{\partial f}{\partial y}(\vec{x}) = (\frac{\partial g}{\partial r}(\vec{u})\frac{\partial r}{\partial y}(\vec{x}) + \frac{\partial g}{\partial \theta}(\vec{u})\frac{\partial \theta}{\partial y}(\vec{x}) = \frac{\partial g}{\partial r}(\vec{u})\sin\theta(\vec{x}) + \frac{\partial g}{\partial \theta}(\vec{u})\frac{\cos\theta}{r}(\vec{x})$.

D'où $(x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y})(\vec{x}) = r\frac{\partial g}{\partial r}(\vec{u})$, d'où $\frac{\partial g}{\partial r}(r,\theta) = \frac{1}{1+r^2}$, d'où $g(r,\theta) = \arctan(r) + h(\theta)$ où $h \in C^1(\mathbb{R}^2;\mathbb{R})$ est quelconque, d'où $f(x,y) = \arctan(r(x,y)) + h(\theta(x,y))$. ▪

Exemple 8.15 Soit $f \in C^2(\mathbb{R}^2;\mathbb{R})$ et soit $\Box F = \frac{pa^2 f}{\partial x^2} - \frac{pa^2 f}{\partial y^2}$ le d'alembertien. Soit $\vec{\psi} : (x,y) \to (u = \frac{x+y}{2}, v = \frac{x-y}{2})$.

Montrer que $\vec{\psi} : \mathbb{R}^2 \to \mathbb{R}^2$ est un changement de variables. On pose $g(u,v) = f(x,y)$, i.e. $g = f\circ\vec{\psi}^{-1}$.

Montrer que $\Box f(x,y) = \frac{\partial^2 g}{\partial u\partial v}(u,v)$, puis donner les fonctions f solutions de $\Box f = 0$.

Réponse. $\vec{\psi}^{-1}(u,v) = \begin{pmatrix} x = u+v \\ y = u-v \end{pmatrix}$, et $\vec{\psi}$ et $\vec{\psi}^{-1}$ sont trivialement C^∞. En particulier, ayant $f \in C^2$ on obtient $g \in C^2$.

Puis $f_1(\vec{x}) =^{\text{noté}} \frac{\partial f}{\partial x}(\vec{x}) = \frac{\partial g}{\partial u}(\vec{u})\frac{1}{2} + \frac{\partial g}{\partial v}(\vec{u})\frac{1}{2} =^{\text{noté}} g_1(\vec{u})$. Puis $\frac{\partial f_1}{\partial x}(\vec{x}) = \frac{\partial g_1}{\partial u}(\vec{u})\frac{1}{2} + \frac{\partial g_1}{\partial v}(\vec{u})\frac{1}{2}$, d'où $\frac{pa^2 f}{\partial x^2}(\vec{x}) = \frac{1}{4}(\frac{pa^2 g}{\partial u^2} + 2\frac{pa^2 g}{\partial u\partial v} + \frac{pa^2 g}{\partial v^2})(\vec{u})$. Même démarche pour $\frac{pa^2 f}{\partial y^2}(\vec{x})$. D'où $\Box f(\vec{x}) = \frac{\partial^2 g}{\partial u\partial v}(u,v)$.

Et $\frac{\partial^2 g}{\partial u\partial v}(u,v) = 0$ donne $\frac{\partial g}{\partial v}(u,v) = \alpha(v)$ où α est une fonction C^1 quelconque. D'où $g(u,v) = \beta(v) + \gamma(u)$ où β est γ sont des fonctions C^2 quelconques (avec ici β une primitive de α). ▪

Exemple 8.16 Soit $\vec{\psi} : (u,v) \in\to (x = u, y = uv) \in \mathbb{R}^2$. Calculer une jacobienne de $\vec{\psi}$.

Montrer que $\vec{\psi} \in C^1(\mathbb{R}_+^* \times \mathbb{R}; \mathbb{R}_+^* \times \mathbb{R})$ est un difféomorphisme, et calculer $\vec{\psi}^{-1}$.

Soit f une fonction telle que $x^2\frac{\partial^2 f}{\partial x^2}(\vec{x}) + xy\frac{\partial^2 f}{\partial x\partial y}(\vec{x}) + y^2\frac{\partial^2 f}{\partial y^2}(\vec{x}) = 0$. Poser $g(u,v) = f(x,y)$, déterminer l'équation satisfaite par g, résoudre cette équation, en déduire f. ▪

8.3 Formules de changement de base

8.3.1 Rappel : changement de vecteurs

On se donne une base $(\vec{a}) = (\vec{a}_j)_{j=1,\dots,n}$ de $\mathbb{R}^n$.

Soit $L : \mathbb{R}^n \to \mathbb{R}^n$ un endomorphisme (une application linéaire d'un espace dans lui-même). Une application linéaire est entièrement définie par la donnée des images $L(\vec{a}_j)$ des vecteurs de base. Notons $(L_{ij})_{i=1,\dots,n}$ les composantes (réelles) de chaque $L(\vec{a}_j)$ dans la base $(\vec{a}_i)$:

$$\forall j = 1,\dots,n, \qquad L(\vec{a}_j) = \sum_{i=1}^n L_{ij}\vec{a}_i = \begin{pmatrix} L_{1j} \\ \vdots \\ L_{nj} \end{pmatrix}_{|(\vec{a})}. \tag{8.9}$$

On peut alors représenter L par sa matrice $[L] = [L_{ij}]_{\substack{1\le i\le n \\ 1\le j\le n}}$ dans la base $(\vec{a}_j)$, la colonne j de la matrice $[L]$ donnant les composantes de $L(\vec{a}_j)$ dans la base $(\vec{a})$.

Notons :

$$\forall j = 1,\dots,n, \qquad \vec{b}_j = L(\vec{a}_j) = \begin{pmatrix} L_{1j} \\ \vdots \\ L_{nj} \end{pmatrix}_{(|\vec{a})} = [L].\begin{pmatrix} 0 \\ \vdots \\ 1 \leftarrow \text{ligne } j \\ \vdots \\ 0 \end{pmatrix}. \tag{8.10}$$

Et L linéaire est ainsi représentée dans la base $(\vec{a}_i)$ par la matrice :

$$[L] \overset{\text{notée}}{=} [L_{ij}]_{\substack{1\le i\le n\\1\le j\le n}} \overset{\text{notée}}{=} [L]_{ij,\,\substack{1\le i\le n\\1\le j\le n}} = \begin{pmatrix} \begin{pmatrix} \vec{b}_1 \end{pmatrix} & \dots & \begin{pmatrix} \vec{b}_n \end{pmatrix} \end{pmatrix}_{|(\vec{a})},$$

(avec i indice de ligne et j indice de colonne) la colonne j de la matrice étant donnée par les composantes de $\vec{b}_j = L(\vec{a}_j)$ dans la base $(\vec{a})$.

On suppose L bijective, et donc $(\vec{b}) = (\vec{b}_i)_{i=1,...,n}$ est une base, d'application inverse L^{-1} linéaire définie sur les vecteurs de base par :

$$\forall j = 1,...,n, \qquad L^{-1}(\vec{b}_j) = \sum_{i=1}^{n} M_{ij}\vec{b}_i = \begin{pmatrix} M_{1j} \\ \vdots \\ M_{nj} \end{pmatrix}_{|(\vec{b})}, \tag{8.11}$$

où on a noté $(M_{ij})_{i=1,...,n}$ les composantes (réelles) de chaque $\vec{a}_j = L^{-1}(\vec{b}_j)$ dans la base $(\vec{b})$. Ainsi $[M] = [L^{-1}]$ est la matrice représentant L^{-1} dans la base $(\vec{b})$. Et donc, les coordonnées de $\vec{a}_j = L^{-1}(\vec{b}_j)$ dans la base $(\vec{b}_i)$ sont stockés dans la colonne j de $[M]$.

Proposition 8.17 *On a :*

$$[L^{-1}] = [L]^{-1} \qquad (\textit{i.e.} : [M] = [L]^{-1}).$$

Preuve. En effet, on a $L^{-1}(L(\vec{a}_j)) = \vec{a}_j$ par définition de l'inverse, et $L^{-1}(L(\vec{a}_j))$ est calculé à l'aide du calcul matriciel :

$\vec{a}_j = L^{-1}(L(\vec{a}_j)) = L^{-1}(\vec{b}_j) = \sum_{i=1}^n M_{ij}\vec{b}_i = \sum_{i=1}^n M_{ij}\sum_{k=1}^n L_{ki}\vec{a}_k = \sum_{k=1}^n(\sum_{i=1}^n L_{ki}M_{ij})\vec{a}_k$
d'où $(\sum_{i=1}^n L_{ki}M_{ij}) = \delta_{kj}$, i.e. $([L][M])_{kj} = \delta_{kj}$, d'où $[L][M] = I$ (matrice identité), i.e. $[L]_{|(\vec{a}_i)}.[L^{-1}]_{|(\vec{b}_i)} = I$. ▪

Remarque 8.18 L étant linéaire, elle est différentiable en tout point $\vec{x} \in \mathbb{R}^n$, et sa différentielle $dL(\vec{x})$ en $\vec{x}$ n'est autre que L. En effet, on a $L(\vec{x}+\vec{v}) = L(\vec{x}) + L(\vec{v})$ et $L(\vec{x}+\vec{v}) = L(\vec{x}) + dL(\vec{x})(\vec{v}) + o(||\vec{v}||)$ pour tout $\vec{v} \in \mathbb{R}^n$. Et $dL(\vec{x})$ étant linéaire, on a bien $dL(\vec{x}) = L$. Et ce quelque soit $\vec{x} \in \mathbb{R}^n$.

Et donc la matrice jacobienne $J_L(\vec{x})$ de L (application linéaire) n'est autre que la matrice constante $[L]$. Et donc (8.11) s'écrit aussi :

$$\vec{b}_j = \sum_{i=1}^{n} dL(\vec{x})(\vec{a}_i) = \sum_{i=1}^{n} [J_L(\vec{x})]_{ij}\vec{a}_i.$$

Dans la suite, on aura $L = d\vec{\psi}(\vec{x})$ application linéaire tangente d'une application $\vec{\psi}$ non linéaire de changement de coordonnées, et on aura $\vec{b}_j = \sum_{i=1}^n [J_{\vec{\psi}}(\vec{x})]_{ij}\vec{a}_i$. ▪

Exemple 8.19 Soit $L : \mathbb{R}^2 \to \mathbb{R}^2$ linéaire définie par :

$$L(\vec{e}_1) = \vec{e}_1 + \vec{e}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_{|(\vec{e})}, \qquad L(\vec{e}_2) = \vec{e}_2 - \vec{e}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}_{|(\vec{e})}.$$

Alors $[L] = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ est la matrice représentant L dans la base $(\vec{e}_1, \vec{e}_2)$.

Sa matrice inverse est $[L]^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$.

On note $\vec{b}_1 = L(\vec{e}_1) = \vec{e}_1 + \vec{e}_2$ et $\vec{b}_2 = L(\vec{e}_2) = \vec{e}_2 - \vec{e}_1$, et on a :

$$L^{-1}(\vec{b}_1) = \vec{e}_1 = \frac{\vec{b}_1 - \vec{b}_2}{2} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \end{pmatrix}_{|(\vec{b})}, \qquad L^{-1}(\vec{b}_2) = \vec{e}_2 = \frac{\vec{b}_1 + \vec{b}_2}{2} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}_{|(\vec{b})}.$$

D'où $[L^{-1}] = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ représente l'application linéaire L^{-1} dans la base $(\vec{b}_1, \vec{b}_2)$ et on a bien $[L]^{-1} = [L^{-1}]$. ▪

Exemple 8.20 On se place dans $\mathbb{R}^2$. On se fixe un réel $\theta \in \mathbb{R}$. On pose :

$$\vec{e}_\rho = \cos\theta\vec{e}_1 + \sin\theta\vec{e}_2 = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}_{|(\vec{e}_1,\vec{e}_2)}, \qquad \vec{e}_\theta = -\sin\theta\vec{e}_1 + \cos\theta\vec{e}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}_{|(\vec{e}_1,\vec{e}_2)}.$$

Géométriquement, $\vec{e}_\rho$ et $\vec{e}_\theta$ sont obtenus par rotation d'angle $+\theta$ à partir de $\vec{e}_1$ et de $\vec{e}_2$.

Et on définit l'application linéaire $L : \mathbb{R}^2 \to \mathbb{R}^2$ à l'aide de $L(\vec{e}_1) = \vec{e}_\rho$ et $L(\vec{e}_2) = \vec{e}_\theta$. Dans la base $(\vec{e}_1, \vec{e}_2)$, la matrice représentant L est donc $[L] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, de matrice inverse $[L]^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. Et $[L]^{-1}$ est la matrice représentant L^{-1} dans la base $(\vec{e}_\rho, \vec{e}_\theta)$, dont les colonnes donnent les composantes de $\vec{e}_1$ et $\vec{e}_2$ dans cette base :

$$\vec{e}_1 = \cos\theta\vec{e}_\rho - \sin\theta\vec{e}_\theta = \begin{pmatrix} \cos\theta \\ -\sin\theta \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta)}, \qquad \vec{e}_2 = \sin\theta\vec{e}_\rho + \cos\theta\vec{e}_\theta = \begin{pmatrix} \sin\theta \\ \cos\theta \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta)}.$$

Géométriquement, $\vec{e}_1$ et $\vec{e}_2$ sont obtenus par rotation d'angle $-\theta$ à partir de $\vec{e}_\rho$ et de $\vec{e}_\theta$. ▪

Exercice 8.21 Reprendre exemple 8.20 en remplaçant $(\vec{e}_\rho, \vec{e}_\theta)$ par la base $(\vec{b}_1, \vec{b}_2)$ où $\vec{b}_1 = \vec{e}_\rho$ et $\vec{b}_2 = \rho\vec{e}_\theta$ (la base $(\vec{b}_1, \vec{b}_2)$ est la base du système de coordonnées polaires).

Réponse. $[L] = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$, d'où $[L^{-1}] = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{\rho}\sin\theta & \frac{1}{\rho}\cos\theta \end{pmatrix}$, d'où :

$$\vec{e}_1 = \cos\theta\vec{b}_1 - \frac{1}{\rho}\sin\theta\vec{b}_2 = \begin{pmatrix} \cos\theta \\ -\frac{1}{\rho}\sin\theta \end{pmatrix}_{|(\vec{b}_1,\vec{b}_2)}, \qquad \vec{e}_2 = \sin\theta\vec{b}_1 + \frac{1}{\rho}\cos\theta\vec{b}_2 = \begin{pmatrix} \sin\theta \\ \frac{1}{\rho}\cos\theta \end{pmatrix}_{|(\vec{b}_1,\vec{b}_2)}.$$

Remarquer que les dimensions sont respectées dans les expressions de $\vec{e}_1$ et de $\vec{e}_2$ car $\vec{b}_2$ à la dimension de ρ. ▪

8.3.2 Rappel : changement de coordonnées

On se donne deux bases $(\vec{a}_i)_{i=1,\dots,n}$ et $(\vec{b}_i)_{i=1,\dots,n}$, et on définit l'application linéaire $L : \mathbb{R}^n \to \mathbb{R}^n$ par $L(\vec{a}_j) = \vec{b}_j$ pour tout $j = 1, \dots, n$.

Soit $\vec{v} \in \mathbb{R}^n$, et soient $(\alpha_i)_{i=1,\dots,n} \in \mathbb{R}^n$ ses composantes dans la base $(\vec{a}_i)$, et $(\beta_i)_{i=1,\dots,n} \in \mathbb{R}^n$ ses composantes dans la base $(\vec{b}_i)$:

$$\vec{v} = \sum_{i=1}^n \alpha_i\vec{a}_i = \sum_{i=1}^n \beta_i\vec{b}_i. \tag{8.12}$$

Exprimons la relation (matricielle) entre les composantes (α_i) et (β_i).

Proposition 8.22 *On a :*

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = [L^{-1}].\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \text{et} \quad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = [L].\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. \tag{8.13}$$

(lorsque $\vec{b}_j = L(\vec{a}_j)$.)

Preuve. On a $\vec{v} = \sum_j \beta_j\vec{b}_j = \sum_j \beta_j(L(\vec{a}_j)) = \sum_j \beta_j(\sum_k L_{kj}\vec{a}_k) = \sum_k(\sum_j L_{kj}\beta_j)\vec{a}_k$. D'où $\alpha_k = \sum_j L_{kj}\beta_j =$ ligne k de L multipliée par colonne $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$, comme annoncé. ▪

Exercice 8.23 Poursuivre l'examble 8.20 : si $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 = w_1\vec{e}_\rho + w_2\vec{e}_\theta$, exprimer w_1 et w_2 en fonction de v_1 et de v_2.

Réponse. $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} . \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, et donc $\begin{pmatrix} w_1 = \cos\theta v_1 + \sin\theta v_2 \\ w_2 = -\sin\theta v_1 + \cos\theta v_2 \end{pmatrix}$. ▪

Exercice 8.24 Poursuivre l'exercice 8.21 : si $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 = w_1\vec{b}_1 + w_2\vec{b}_2$, exprimer w_1 et w_2 en fonction de v_1 et de v_2.

Réponse. $\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{\rho}\sin\theta & \frac{1}{\rho}\cos\theta \end{pmatrix} . \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$, et donc $\begin{pmatrix} w_1 = \cos\theta v_1 + \sin\theta v_2 \\ w_2 = -\frac{1}{\rho}\sin\theta v_1 + \frac{1}{\rho}\cos\theta v_2 \end{pmatrix}$. Et on vérifie les dimensions (w_2 est en $\frac{1}{\rho}$ et $\vec{b}_2$ en ρ). ▪

Exercice 8.25 Soit $A = [G]_{|(\vec{a}_i)}$ la matrice d'une application linéaire $G : \mathbb{R}^n \to \mathbb{R}^n$ dans une base $(\vec{a}_i)$ donnée, et soit $B = [G]_{|(\vec{b}_i)}$ la matrice de l'application linéaire $G : \mathbb{R}^n \to \mathbb{R}^n$ dans une base $(\vec{b}_i)$ donnée. Montrer que $B = P^{-1}AP$, où P est la matrice de passage, i.e. la matrice de l'application linéaire L définie par $L(\vec{a}_j) = \vec{b}_j$.

Réponse. On a $P = [L]_{|(\vec{a}_i)}$ la matrice de L dans la base $(\vec{a}_i)$ et $P^{-1} = [L^{-1}]_{|(\vec{b}_i)}$ la matrice de L^{-1} dans la base $(\vec{b}_i)$, i.e. $L(\vec{a}_j) = \sum_i P_{ij}\vec{a}_i$ et $L^{-1}(\vec{b}_j) = \sum_i (P^{-1})_{ij}\vec{b}_i$.

On a $G(\vec{a}_j) = \sum_i A_{ij}\vec{a}_i$ et $G(\vec{b}_j) = \sum_i B_{ij}\vec{b}_i$ par définition des matrices A et B, d'où :

$$\begin{aligned}\sum_i A_{ij}\vec{a}_i = G(\vec{a}_j) = G(L^{-1}(\vec{b}_j)) = G(\sum_k P^{-1}_{kj}\vec{b}_k) = \sum_k P^{-1}_{kj} G(\vec{b}_k) = \sum_{k\ell} P^{-1}_{kj} B_{\ell k}\vec{b}_\ell \\ = \sum_{k\ell} P^{-1}_{kj} B_{\ell k} L(\vec{a}_\ell) = \sum_{ik\ell} P^{-1}_{kj} B_{\ell k} P_{i\ell}\vec{a}_i.\end{aligned}$$

D'où $A_{ij} = \sum_{k\ell} P^{-1}_{kj} B_{\ell k} P_{i\ell} = [PBP^{-1}]_{ij}$ pour tout i, j. ▪

9 * Complément

9.1 Matrices jacobiennes et interprétation

Soit $\vec{\psi} : U \to V$ un changement de variables. Les matrices jacobiennes de $\vec{\psi}$ en $\vec{u}$ et de $\vec{\psi}^{-1}$ en $\vec{x}$ sont :

$$J_{\vec{\psi}}(\vec{u}) = [\frac{\partial \psi_i}{\partial u_j}(\vec{u})]_{\substack{i=1:n \\ j=1:n}} = [\frac{\partial x_i}{\partial u_j}(\vec{u})]_{\substack{i=1:n \\ j=1:n}}, \qquad J_{\vec{\psi}^{-1}}(\vec{x}) = [\frac{\partial (\psi^{-1})_i}{\partial u_j}(\vec{x})]_{\substack{i=1:n \\ j=1:n}} = [\frac{\partial u_i}{\partial x_j}(\vec{x})]_{\substack{i=1:n \\ j=1:n}}.$$

Les matrices jacobiennes $J_{\vec{\psi}}(\vec{u})$ et $J_{\vec{\psi}^{-1}}(\vec{x})$ représentent les matrices des applications linéaires $d\vec{\psi}(\vec{u})$ et $d\vec{\psi}^{-1}(\vec{x})$.

Question : dans quelles bases ?

Réponse.

Les paramètres $u_1, \dots, u_n$ sont supposés indépendants : on impose donc, dans l'espace des paramètres $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R} \supset U$, de prendre comme base la base canonique $(\vec{E}_1, \dots, \vec{E}_n)$ de $\mathbb{R}^n$, base définie par $\vec{E}_i = (0, \dots, 0, 1, 0, \dots, 0)$ avec le 1 en i-ème position et 0 ailleurs, pour $i = 1, ..., n$. (On pourra penser aux coordonnées polaires $(\rho, \theta) = \vec{u}$ qui sont indépendantes et on écrira $\vec{u} = \rho\vec{E}_1 + \theta\vec{E}_2$ représenté par $[\vec{u}] = \begin{pmatrix} \rho \\ \theta \end{pmatrix}$ dans la base $(\vec{E}_1, \vec{E}_2)$.)

Soit $f : U \to \mathbb{R}$ C^1 à valeurs scalaires ; son développement limité au premier ordre est :

$$f(\vec{v}) = f(\vec{u}) + L_{\vec{u}}(\vec{v} - \vec{u}) + o(||\vec{v} - \vec{u}||) \quad \in \mathbb{R},$$

où $L_{\vec{u}} = df(\vec{u}) : \mathbb{R}^n \to \mathbb{R}$ est la forme linéaire définissant le plan tangent au graphe de f au point $\vec{u}$, i.e. le (hyper-)plan d'équation : $\vec{v} \in \mathbb{R}^n \to z = f(\vec{u}) + L_{\vec{u}}(\vec{v} - \vec{u}) \in \mathbb{R}$.

Puis $L_{\vec{u}} : \mathbb{R}^n \to \mathbb{R}$, forme linéaire, est définie à l'aide des images $L_{\vec{u}}(\vec{E}_i)$ des vecteurs de base $\vec{E}_i$. Et, par définition :

$$\frac{\partial f}{\partial u_i}(\vec{u}) = L_{\vec{u}}(\vec{E}_i) = df(\vec{u})(\vec{E}_i) = \lim_{h \to 0} \frac{f(\vec{u} + h\vec{E}_i) - f(\vec{u})}{h}$$

est la dérivée de f le long du i-ème vecteur de base canonique $\vec{E}_i$ de $\mathbb{R}^n$ espace des paramètres. Ainsi, tout vecteur $\vec{v} \in \mathbb{R}^n$, $\vec{v} = \sum_i v_i\vec{E}_i$, est transformé par $L_{\vec{u}}$ en $L_{\vec{u}}(\vec{v}) = \sum_i v_i\, L_{\vec{u}}(\vec{E}_i) = \sum_i v_i\, \frac{\partial f}{\partial u_i}(\vec{u})$.

Et $J_f(\vec{u}) = (\frac{\partial f}{\partial u_1}(\vec{u}) \dots \frac{\partial f}{\partial u_n}(\vec{u}))$ est la matrice ligne qui stocke les $\frac{\partial f}{\partial u_1}(\vec{u})$ (qui sont les images des vecteurs de base $\vec{E}_i$ de l'espace des paramètres par $df(\vec{u}) = L_{\vec{u}}$). Et donc, si $\vec{v} = \sum v_i \vec{E}_i$ est représenté par la matrice $[\vec{v}] = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ (représentation de $\vec{v}$ dans la base $(\vec{E}_i)$), on a :

$$J_f(\vec{u}).[\vec{v}] = (\frac{\partial f}{\partial u_1}(\vec{u}) \dots \frac{\partial f}{\partial u_n}(\vec{u})) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_i \frac{\partial f}{\partial u_i}(\vec{u})\, v_i = L_{\vec{u}}(\vec{v}) \quad \in \mathbb{R}.$$

Maintenant, si $(\vec{e}_i)$ est une base donnée de $\mathbb{R}^n$ espace géométrique, pour U ouvert de $\mathbb{R}^n$ (espace des paramètres) et V ouvert de $\mathbb{R}^n$ (espace géométrique), on a :

$$\begin{aligned} \vec{\psi} : U &\to V \\ \vec{u} &\to \vec{\psi}(\vec{u}) = \sum_{i=1}^n \psi_i \vec{e}_i = \begin{pmatrix} \psi_1(\vec{u}) \\ \vdots \\ \psi_n(\vec{u}) \end{pmatrix}_{(\vec{e}_i)} , \end{aligned}$$

chaque $\psi_i : U \to \mathbb{R}$ étant C^1 ; chaque $d\psi_i(\vec{u})$ est une application linéaire de matrice jacobienne $J_{\psi_i}(\vec{u}) = (\frac{\partial \psi_i}{\partial u_1}(\vec{u}) \dots \frac{\partial \psi_i}{\partial u_n}(\vec{u}))$ (matrice ligne). Et $J_{\vec{\psi}}(\vec{u})$ est la matrice constituée des lignes $J_{\psi_i}(\vec{u})$.

Comme :

$$d\vec{\psi}(\vec{u}) = \sum_{i=1}^n d\psi_i(\vec{u})\vec{e}_i = \begin{pmatrix} d\psi_1(\vec{u}) \\ \vdots \\ d\psi_n(\vec{u}) \end{pmatrix}_{(\vec{e}_i)} \quad \in L(\mathbb{R}^n, \mathbb{R}^n),$$

(chaque $d\psi_i(\vec{u})$ est une forme linéaire $\in L(\mathbb{R}^n, \mathbb{R})$) on a :

$$d\vec{\psi}(\vec{u})(\vec{E}_j) = \sum_{i=1}^n d\psi_i(\vec{u})(\vec{E}_j)\vec{e}_i = \begin{pmatrix} d\psi_1(\vec{u})(\vec{E}_j) \\ \vdots \\ d\psi_n(\vec{u})(\vec{E}_j) \end{pmatrix}_{|(\vec{e}_i)} = \begin{pmatrix} \frac{\partial \psi_1}{\partial u_j}(\vec{u}) \\ \vdots \\ \frac{\partial \psi_n}{\partial u_j}(\vec{u}) \end{pmatrix} \in \mathbb{R}^n.$$

D'où $J_{\vec{\psi}}(\vec{u})$ est la matrice représentant l'application linéaire $d\vec{\psi}(\vec{u})$ dans les bases $(\vec{E}_i)$ (de l'espace de départ des paramètres) et $(\vec{e}_i)$ (de l'espace d'arrivée géométrique). En particulier, la j-ème colonne de $J_{\vec{\psi}}(\vec{u})$ est donnée par les composantes de $d\vec{\psi}(\vec{u})(\vec{E}_j) = \sum_{i=1}^n d\psi_i(\vec{u})(\vec{E}_j)\vec{e}_i = \sum_{i=1}^n \frac{\partial \psi_i}{\partial u_j}(\vec{u})\vec{e}_i$, i.e. par les $(\frac{\partial \psi_i}{\partial u_j}(\vec{u}))_{i=1,\dots,n}$. Et donc, pour tout vecteur $\vec{v} \in \mathbb{R}^n$, avec $\vec{v} = \sum_i v_i \vec{E}_i$, on a :

$$d\vec{\psi}(\vec{u})(\vec{v}) = \sum_j v_j\, d\vec{\psi}(\vec{u})(\vec{E}_j) = \sum_j \begin{pmatrix} \frac{\partial \psi_1}{\partial u_j}(\vec{u}) \\ \vdots \\ \frac{\partial \psi_n}{\partial u_j}(\vec{u}) \end{pmatrix} v_j = \begin{pmatrix} \frac{\partial \psi_1}{\partial u_1}(\vec{u})v_1 + \dots + \frac{\partial \psi_1}{\partial u_n}(\vec{u})v_n \\ \vdots \\ \frac{\partial \psi_n}{\partial u_1}(\vec{u})v_1 + \dots + \frac{\partial \psi_n}{\partial u_n}(\vec{u})v_n \end{pmatrix} = J_{\vec{\psi}}(\vec{u}).[\vec{v}].$$

Interprétation des colonnes

Pour $\vec{x} = \vec{\psi}(\vec{u})$, on définit au point $\vec{x}$ de l'espace géométrique le vecteur, dessiné dans l'espace géométrique $\mathrm{Im}(d\vec{\psi}(\vec{u}))$:

$$\vec{b}_j(\vec{x}) = d\vec{\psi}(\vec{u})(\vec{E}_j) \quad \text{i.e.} \quad \vec{b}_j(\vec{x}) = \sum_{i=1}^n \frac{\partial \psi_i}{\partial u_j}(\vec{u})\vec{e}_i. \tag{9.1}$$

Comme $d\vec{\psi}(\vec{u}) \in L(\mathbb{R}^n \supset U, \mathbb{R}^n \supset V)$, on a $\mathrm{Im}(d\vec{\psi}(\vec{u})) \subset \mathbb{R}^n$ espace géométrique, et $\vec{b}_j(\vec{x}) \in \mathrm{Im}(d\vec{\psi}(\vec{u}))$ est bien un vecteur de l'espace géométrique.

Et les composantes de $\vec{b}_j(\vec{x})$ dans la base $(\vec{e}_i)$ sont stockées dans la colonne j de $J_{\vec{\psi}}(\vec{u})$:

$$[\vec{b}_j(\vec{x})]_{(\vec{e}_i)} = J_{\vec{\psi}}(\vec{u}).[\vec{E}_j]_{(\vec{E}_i)} = \begin{pmatrix} \frac{\partial \psi_1}{\partial u_j}(\vec{u}) \\ \vdots \\ \frac{\partial \psi_n}{\partial u_j}(\vec{u}) \end{pmatrix}_{(\vec{e}_j)} , \tag{9.2}$$

Donc quand $\vec{x} = \vec{\psi}(\vec{u})$, les $\vec{b}_j(\vec{x})$ sont les images des $\vec{E}_j$ par l'application linéaire tangente $d\vec{\psi}(\vec{u})$, et les composantes de $\vec{b}_j(\vec{x})$ dans la base $(\vec{e}_i)$ sont stockées dans la colonne j de $J_{\vec{\psi}}(\vec{u})$.

9.2 Système de coordonnées, lignes de coordonnées

On s'intéresse ici à l'interprétation géométrique des vecteurs $\vec{b}_j(\vec{x})$ définis en (9.1).

Un changement de variables $\vec{\psi} : U \to V$ permet de définir pour chaque $\vec{u} \in U$ espace des paramètres, au voisinage de chaque $\vec{x} = \psi(\vec{u})$ espace géométrique, n courbes $(\vec{c}_j)_{j=1,...,n}$ appelées lignes de coordonnées en $\vec{x}$ (à dessiner dans l'espace géométrique) définies par :

$$\vec{c}_j : \begin{cases}]-\varepsilon_j, \varepsilon_j[& \to V \\ \qquad t & \mapsto \vec{c}_j(t) = \vec{\psi}(\vec{u} + t\vec{E}_i) = \vec{\psi}(u_1, ..., u_{j-1}, u_j + t, u_{j+1}, ..., u_n) \in \mathbb{R}^n, \end{cases} \tag{9.3}$$

où $\varepsilon_j > 0$ est choisi tel que les $(u_1, ..., u_{j-1}, u_j + t, u_{j+1}, ..., u_n) \in U$ pour tout $t \in]-\varepsilon_j, \varepsilon_j[$. En particulier $\vec{c}_j(0) = \vec{x}$, et $\{\vec{c}_j(t) : t \in]-\varepsilon_j, \varepsilon_j[\}$ est une courbe qui passe par $\vec{x}$. Les images $\mathrm{Im}(\vec{c}_j)$ sont les images des segments de droite $]\vec{u} - \varepsilon_j\vec{E}_j, \vec{u} + \varepsilon_j\vec{E}_j[$ par $\vec{c}_j$.

Définition 9.1 Ces n courbes $(\vec{c}_j)_{j=1,...,n}$ attachées au système de coordonnées forment ce qui est appelé un système de coordonnées en $\vec{x}$.

Donc par définition des $\vec{c}_j$ on a $\vec{x} = \vec{\psi}(\vec{u}) = \vec{c}_j(0)$ pour tout $j = 1, ..., n$ et :

$$\vec{c}_j(t) = \vec{\psi}(\vec{u} + t\vec{E}_j) \qquad (= \sum_{i=1}^n \psi_i(\vec{u} + t\vec{E}_j)\, \vec{e}_i = \begin{pmatrix} \psi_1(\vec{u} + t\vec{E}_j) \\ \vdots \\ \psi_n(\vec{u} + t\vec{E}_j) \end{pmatrix}_{|(\vec{e}_i)}). \tag{9.4}$$

Si on note $\vec{\tau}_j(t) = \vec{u} + t\vec{E}_j = (u_1, ..., u_{j-1}, u_j + t, u_{j+1}, ..., u_n)$ (translation de t le long du j-ème vecteur de base canonique dans l'espace des paramètres), alors on a $\vec{c}_j = \vec{\psi} \circ \vec{\tau}_j$, i.e. :

$$\vec{c}_j(t) = (\vec{\psi} \circ \vec{\tau}_j)(t) \qquad (= \vec{\psi}(\vec{\tau}_j(t)) = \sum_{i=1}^n \psi_i(\vec{\tau}_j(t))\, \vec{e}_i). \tag{9.5}$$

Exemple 9.2 Pour les coordonnées polaires où $\vec{u} = (\rho, \theta)$ et $\vec{x} = \vec{\psi}(\rho, \theta) = \begin{pmatrix} x_1 = \rho\cos\theta \\ x_2 = \rho\sin\theta \end{pmatrix}$, on a $\vec{c}_1(t) = \begin{pmatrix} (\rho + t)\cos\theta \\ (\rho + t)\sin\theta \end{pmatrix}$ et $\vec{c}_1$ définit (i.e. a pour image) un segment de droite "radial", et $\vec{c}_2(t) = \begin{pmatrix} \rho\cos(\theta + t) \\ \rho\sin(\theta + t) \end{pmatrix}$ et $\vec{c}_2$ définit (i.e. a pour image) un arc de cercle. Cela dans un voisinage du point $\vec{x} = \vec{\psi}(\vec{u})$ où $\vec{u} = (\rho, \theta)$ est fixé. Notez que $\vec{c}_1$ a un sens pour $t > -\rho$, et $\vec{c}_2$ pour $\theta + t \in]-\pi, \pi[$, par exemple. ∴

Exemple 9.3 Pour les coordonnées sphériques où $\vec{u} = (\rho, \theta, \varphi)$, dans un voisinage du point $\vec{x} = \vec{\psi}(\vec{u})$ où $\vec{u} = (\rho, \theta, \varphi)$ est fixé, la courbe $\vec{c}_1(t) = \begin{pmatrix} (\rho + t)\cos\theta\cos\varphi \\ (\rho + t)\sin\theta\cos\varphi \\ (\rho + t)\sin\varphi \end{pmatrix}$ définit un segment de droite "radial", la courbe $\vec{c}_2(t) = \begin{pmatrix} \rho\cos(\theta + t)\cos\varphi \\ \rho\sin(\theta + t)\cos\varphi \\ \rho\sin\varphi \end{pmatrix}$ définit un parallèle, et la courbe $\vec{c}_3(t) = \begin{pmatrix} \rho\cos\theta\cos(\varphi + t) \\ \rho\sin\theta\cos(\varphi + t) \\ \rho\sin(\varphi + t) \end{pmatrix}$ définit un méridien. ∴

9.3 Base du système de coordonnées

Proposition 9.4 et définition. *Soit un changement de variables $\vec{\psi}$, et soient, définis pour chaque $\vec{u} \in U$ avec $\vec{x} = \vec{\psi}(\vec{u})$, les n vecteurs $\vec{b}_j(\vec{x})$ définis en (9.1), pour $j = 1, ..., n$. On a :*

$$\vec{b}_j(\vec{x}) = \frac{d\vec{c}_j}{dt}(t{=}0) \qquad (= \frac{\partial\vec{\psi}}{\partial u_j}(\vec{u}) = \sum_{i=1}^n \frac{\partial\psi_i}{\partial u_j}(\vec{u})\vec{e}_i = \begin{pmatrix} \frac{\partial\psi_1}{\partial u_j}(\vec{u}) \\ \vdots \\ \frac{\partial\psi_n}{\partial u_j}(\vec{u}) \end{pmatrix}_{|(\vec{e}_i)}), \tag{9.6}$$

i.e. les $\vec{b}_j(\vec{x})$ sont les vecteurs tangents aux lignes de coordonnées au point $\vec{x} = \vec{\psi}(\vec{u})$.

Ces vecteurs forment une base $(\vec{b}_1(\vec{x}), ..., \vec{b}_n(\vec{x}))$ appelée base du système de coordonnées en $\vec{x} = \vec{\psi}(\vec{u})$ (ce sont les vecteurs de base "naturellement" attachées au système de coordonnées).

Preuve. Avec (9.5) on a $\frac{d\vec{c}_j}{dt}(t) = d\vec{\psi}(\vec{r}_j(t)).\frac{d\vec{r}_j}{dt}(t)$, et donc pour $t=0$ on obtient $\frac{d\vec{c}_j}{dt}(0) = d\vec{\psi}(\vec{u}).\vec{E}_j$, i.e. $\frac{d\vec{c}_j}{dt}(0) = \vec{b}_j(\vec{x})$, i.e. (9.6). (Et on rappelle que par définition $d\vec{\psi}(\vec{u}).\vec{E}_j = \frac{\partial\vec{\psi}}{\partial u_j}(\vec{u}) = \begin{pmatrix} \frac{\partial\psi_1}{\partial u_j}(\vec{u}) \\ \vdots \\ \frac{\partial\psi_n}{\partial u_j}(\vec{u}) \end{pmatrix}_{|(\vec{e}_i)}$.)

Comme $\vec{\psi}$ définit un changement de variables, sa matrice jacobienne est inversible, et donc ses colonnes sont indépendantes, i.e. les vecteurs $\frac{\partial\vec{\psi}}{\partial u_j}(\vec{u}) = \vec{b}_j(\vec{x})$ sont indépendants (toujours avec $\vec{x} = \vec{\psi}(\vec{u})$). Étant au nombre de n, ils forment une base. ∎

Définition 9.5 Lorsqu'en chaque point $\vec{x} \in V$ les vecteurs $b_i(\vec{x})$ sont 2 à 2 orthogonaux, pour $i = 1, ..., n$, le système de coordonnées est dit orthogonal.

Exemple 9.6 Pour les coordonnées polaires, en un point $\vec{x} = \begin{pmatrix} \rho\cos\theta \\ \rho\sin\theta \end{pmatrix}$, les deux vecteurs de base de coordonnées sont $\vec{b}_1(\vec{x}) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ et $\vec{b}_2(\vec{x}) = \begin{pmatrix} -\rho\sin\theta \\ \rho\cos\theta \end{pmatrix}$ (composantes données dans la base canonique $(\vec{e}_1, \vec{e}_2)$, avec $\rho = \rho(\vec{x})$ et $\theta = \theta(\vec{x})$).

Attention : cette base $(\vec{b}_1(\vec{x}), \vec{b}_2(\vec{x}))$ est orthogonale, mais non orthonormée. La base usuelle est obtenue en normant, i.e. c'est la base orthonormée $(\vec{e}_\rho, \vec{e}_\theta)$ donnée par :

$$\vec{e}_\rho(\vec{x}) = \vec{b}_1(\vec{x}) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}_{|(\vec{e}_1,\vec{e}_2)}, \qquad \vec{e}_\theta(\vec{x}) = \frac{1}{\rho}\vec{b}_2(\vec{x}) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}_{|(\vec{e}_1,\vec{e}_2)}.$$

∎

Exemple 9.7 Pour les coordonnées sphériques, en un point $\vec{x}$, les trois vecteurs de base de coordonnées sont $\vec{b}_1(\vec{x}) = \begin{pmatrix} \cos\theta\cos\varphi \\ \sin\theta\cos\varphi \\ \sin\varphi \end{pmatrix}$, $\vec{b}_2(\vec{x}) = \begin{pmatrix} -\rho\sin\theta\cos\varphi \\ \rho\cos\theta\cos\varphi \\ 0 \end{pmatrix}$ et $\vec{b}_3(\vec{x}) = \begin{pmatrix} -\rho\cos\theta\sin\varphi \\ -\rho\sin\theta\sin\varphi \\ \rho\cos\varphi \end{pmatrix}$ (composantes données dans la base canonique $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ avec $\rho = \rho(\vec{x})$, $\theta = \theta(\vec{x})$ et $\varphi = \varphi(\vec{x})$).

Attention : cette base $(\vec{b}_1(\vec{x}), \vec{b}_2(\vec{x}), \vec{b}_3(\vec{x}))$ est orthogonale, mais non orthormée. La base usuelle est obtenue en normant, i.e. c'est la base orthonormée $(\vec{e}_\rho, \vec{e}_\theta, \vec{e}_\varphi)$ donnée par, en un point $\vec{x}$:

$$\vec{e}_\rho = \vec{b}_1 = \begin{pmatrix} \cos\theta\cos\varphi \\ \sin\theta\cos\varphi \\ \sin\varphi \end{pmatrix}, \quad \vec{e}_\theta = \frac{1}{\rho\cos\varphi}\vec{b}_2 = \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix}, \quad \vec{e}_\varphi = \frac{1}{\rho}\vec{b}_3 = \begin{pmatrix} -\cos\theta\sin\varphi \\ -\sin\theta\sin\varphi \\ \cos\varphi \end{pmatrix}.$$

∎

9.4 Les dérivées dans les nouvelles coordonnées

Exprimer une fonction $f : \vec{x} \in V \to f(\vec{x}) \in \mathbb{R}$ en fonction des paramètres $\vec{u} \in U$, $\vec{u} = \vec{\psi}^{-1}(\vec{x})$, consiste à définir la fonction $\tilde{f} : U \to \mathbb{R}$ par :

$$\tilde{f} = f \circ \vec{\psi}, \qquad \text{i.e.} \qquad \tilde{f}(\vec{u}) = f(\vec{x}) \qquad \text{quand} \qquad \vec{x} = \vec{\psi}(\vec{u}). \tag{9.7}$$

Et on a donc également :

$$f = \tilde{f} \circ \vec{\psi}^{-1}.$$

Exemple 9.8 Si $f(x, y) = e^{\sqrt{x^2+y^2}}$ on prend les coordonnées polaires et on a $f(x, y) = \tilde{f}(\rho, \theta)$ avec $\tilde{f}(\rho, \theta) = e^\rho$, i.e. $f = \tilde{f} \circ \vec{\psi}^{-1}$, et on écrit $f(x, y) = e^\rho$, avec sous-entendu $\rho = \rho(x, y)$. ∎

Problème : f est défini en un point $\vec{x}$, et $\frac{\partial f}{\partial x_i}$ a un sens au point $\vec{x}$. Et comme $\vec{x} = \vec{\psi}(\vec{u})$, se pose la question de dériver f par rapport à u_i, ce qui a priori n'a pas de sens car f est définie dans V espace géométrique (pour un point $\vec{x}$) et n'est pas défini dans U espace des paramètres (pour un point $\vec{u}$).

Définition 9.9 Soit $f : \vec{x} \in V \to f(\vec{x}) \in \mathbb{R}$ une application $C^1(V;\mathbb{R})$, et soit $\vec{u} = \vec{\psi}^{-1}(\vec{x}) \in U$.

Dériver f en $\vec{x} \in V$ par rapport à u_j signifie par définition : dériver $\tilde{f} = f \circ \vec{\psi}$ au point $\vec{u}$ par rapport à u_j, i.e. pour $\vec{x} \in V$ avec $\vec{x} = \vec{\psi}(\vec{u})$:

$$\frac{\partial f}{\partial u_j}(\vec{x}) \overset{\text{déf}}{=} \frac{\partial \tilde{f}}{\partial u_j}(\vec{u}). \tag{9.8}$$

Proposition 9.10 *On a :*

$$\frac{\partial f}{\partial u_j}(\vec{x}) = df(\vec{x}).\vec{b}_j(\vec{x}). \tag{9.9}$$

C'est la dérivée de f en $\vec{x}$ le long du vecteur $\vec{b}_j(\vec{x})$ (vecteur tangent à la j-ème ligne de coordonnées).

Preuve. On a $\tilde{f} = f \circ \vec{\psi}$, d'où :

$$\frac{\partial f}{\partial u_j}(\vec{x}) \overset{\text{déf}}{=} \frac{\partial (f \circ \vec{\psi})}{\partial u_j}(\vec{u}) = df(\vec{x}).\frac{\partial \vec{\psi}}{\partial u_j}(\vec{u}) = df(\vec{x}).\vec{b}_j(\vec{x}). \tag{9.10}$$

▪

Exemple 9.11 Soit $f(x,y) = e^{\sqrt{x^2+y^2}} = e^\rho$ (et les coordonnées polaires). Alors on a $\tilde{f}(\rho,\theta) = e^\rho$, d'où $\frac{\partial f}{\partial \rho}(\vec{x}) =^{\text{déf}} \frac{\partial \tilde{f}}{\partial \rho}(\vec{u}) = e^\rho$ et $\frac{\partial f}{\partial \theta}(\vec{x}) =^{\text{déf}} \frac{\partial \tilde{f}}{\partial \theta}(\vec{u}) = 0$.

Résultat retrouvé avec la dérivation des fonctions composées : par exemple $\frac{\partial (f\circ\vec{\psi})}{\partial \rho}(\vec{u}) = \frac{\partial f}{\partial x}(\vec{x})\frac{\partial \psi_1}{\partial \rho}(\vec{u}) + \frac{\partial f}{\partial y}(\vec{x})\frac{\partial \psi_2}{\partial \rho}(\vec{u}) = \cos\theta e^\rho.\cos\theta + \sin\theta e^\rho.\sin\theta = e^\rho$, car $\psi_1(\rho,\theta) = \rho\cos\theta$ et $\psi_2(\rho,\theta) = \rho\sin\theta$.

On encore avec $\frac{\partial f}{\partial \rho}(\vec{x}) = df(\vec{x}).\vec{b}_1(\vec{x}) = (\frac{\partial f}{\partial x}\ \frac{\partial f}{\partial y})\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$, idem pour $\frac{\partial f}{\partial \theta}(\vec{x})$. ▪

Et, pour tout i=1, ..., n, si $\vec{x} = \vec{\psi}(\vec{u}) \in V$, pour $f \in C^1(V;\mathbb{R})$, les formules (3.9) de dérivation de la composée $\tilde{f} = f \circ \vec{\psi}$ donnent : "$J_{\tilde{f}} = J_f.J_{\vec{\psi}}$", i.e. avec $\vec{x} = \vec{\psi}(\vec{u})$:

$$\Big(\frac{\partial \tilde{f}}{\partial u_1}(\vec{u})\ \dots\ \frac{\partial \tilde{f}}{\partial u_n}(\vec{u})\Big) = \Big(\frac{\partial f}{\partial x_1}(\vec{x})\ \dots\ \frac{\partial f}{\partial x_n}(\vec{x})\Big).J_{\vec{\psi}}(\vec{u}), \tag{9.11}$$

soit en transposant :

$$\begin{pmatrix}\frac{\partial \tilde{f}}{\partial u_1}(\vec{u}) \\ \vdots \\ \frac{\partial \tilde{f}}{\partial u_n}(\vec{u})\end{pmatrix} = J_{\vec{\psi}}(\vec{u})^T.\begin{pmatrix}\frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x})\end{pmatrix}. \tag{9.12}$$

Ou bien, si on veut exprimer les $\frac{\partial f}{\partial x_i}(\vec{x})$ en fonction des $\frac{\partial \tilde{f}}{\partial u_j}(\vec{u})$:

$$\begin{pmatrix}\frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x})\end{pmatrix} = (J_{\vec{\psi}}(\vec{u}))^{-T}.\begin{pmatrix}\frac{\partial \tilde{f}}{\partial u_1}(\vec{u}) \\ \vdots \\ \frac{\partial \tilde{f}}{\partial u_n}(\vec{u})\end{pmatrix}. \tag{9.13}$$

Pour une matrice M, la notation M^{-T} signifie : $M^{-T} = (M^{-1})^T$ (donc $= (M^T)^{-1}$).

Remarque 9.12 Donc si f est défini en $\vec{x} = \vec{\psi}(\vec{u}) \in V$ et si au point $\vec{x}$ on veut dériver f par rapport à u_i, c'est la fonction $\tilde{f} = f \circ \vec{\psi}$ qu'on veut dériver par rapport à u_i, et ce au point $\vec{u} = \vec{\psi}^{-1}(\vec{x})$. ▪

Exemple 9.13 Soit $f(x,y) = xy \in \mathbb{R}$, et on prend les coordonnées polaires $(x,y) = \vec{\psi}(\rho,\theta) = (\rho\cos\theta, \rho\sin\theta)$. On a donc $\tilde{f}(\rho,\theta) = \rho^2\cos\theta\sin\theta$, et donc : $\displaystyle\frac{\partial f}{\partial \rho}(x,y) \overset{\text{déf}}{=} \frac{\partial \tilde{f}}{\partial \rho}(\rho,\theta) = 2\rho\cos\theta\sin\theta$ où $\rho = \rho(x,y)$ et $\theta = \theta(x,y)$.

Ce calcul est un cas particulier du calcul : $\frac{\partial (f\circ\vec{\psi})}{\partial \rho}(\rho,\theta) = \frac{\partial f}{\partial x}(x,y)\frac{\partial \psi_1}{\partial \rho}(\rho,\theta) + \frac{\partial f}{\partial y}(x,y)\frac{\partial \psi_2}{\partial \rho}(\rho,\theta)$ qui ici $= y\cos\theta + x\sin\theta = (\rho\sin\theta)\cos\theta + (\rho\cos\theta)\sin\theta = 2\rho\cos\theta\sin\theta$. ▪

Exemple 9.14 Pour les coordonnées polaires, $J_{\vec\psi}(\rho,\theta) = \begin{pmatrix}\cos\theta & -\rho\sin\theta\\ \sin\theta & \rho\cos\theta\end{pmatrix}$. Et pour $f(x,y)$ donné, on a avec (9.12) :

$$\begin{cases} \dfrac{\partial\tilde f}{\partial\rho}(\rho,\theta) = \dfrac{\partial(f\circ\vec\psi)}{\partial\rho}(\rho,\theta) = \cos\theta\dfrac{\partial f}{\partial x}(x,y)+\sin\theta\dfrac{\partial f}{\partial y}(x,y) \quad (= df(\vec x).\vec b_1(\vec x)),\\ \dfrac{\partial\tilde f}{\partial\theta}(\rho,\theta) = \dfrac{\partial(f\circ\vec\psi)}{\partial\theta}(\rho,\theta) = -\rho\sin\theta\dfrac{\partial f}{\partial x}(x,y)+\rho\cos\theta\dfrac{\partial f}{\partial y}(x,y) \quad (= df(\vec x).\vec b_2(\vec x)),\end{cases} \tag{9.14}$$

et on retrouve bien (9.12) : $\begin{pmatrix}\frac{\partial\tilde f}{\partial\rho}(\rho,\theta)\\ \frac{\partial\tilde f}{\partial\theta}(\rho,\theta)\end{pmatrix} = J_{\vec\psi}(\rho,\theta)^T.\begin{pmatrix}\frac{\partial f}{\partial x}(x,y)\\ \frac{\partial f}{\partial y}(x,y)\end{pmatrix}$.

9.5 Les dérivées secondes dans les nouvelles coordonnées et le laplacien

Si $f\in C^2(V;\mathbb{R})$, si $\tilde f(\vec u) =^{\text{déf}} f(\vec x)$ quand $\vec x=\vec\psi(\vec u)$, alors on définit :

$$\frac{\partial^2 f}{\partial u_i\partial u_j}(\vec x) \stackrel{\text{déf}}{=} \frac{\partial^2\tilde f}{\partial u_i\partial u_j}(\vec u) \quad \Big(= \frac{\partial^2(f\circ\vec\psi)}{\partial u_i\partial u_j}(\vec u)\Big). \tag{9.15}$$

Vérifions la cohérence de ces notations : on a :

$$\frac{\partial^2\tilde f}{\partial u_i\partial u_j}(\vec u) = \frac{\partial\frac{\partial\tilde f}{\partial u_j}}{\partial u_i}(\vec u) \stackrel{\text{noté}}{=} \frac{\partial\frac{\partial f}{\partial u_j}}{\partial u_i}(\vec x) \stackrel{\text{noté}}{=} \frac{\partial^2 f}{\partial u_i u_j}(\vec x).$$

Exercice 9.15 Soit $f\in C^2(\mathbb{R}^2;\mathbb{R})$. Calculer les dérivées de f en coordonnées polaires, à l'aide des $\frac{\partial\tilde f}{\partial u_i}$ et des $\frac{\partial^2\tilde f}{\partial u_i\partial u_j}$ pour $u_i=\rho$ ou θ. Montrer que le laplacien en coordonnées polaires, défini en un point $\vec x$ par $\Delta f(\vec x) = \frac{\partial^2 f}{\partial x^2}(\vec x)+\frac{\partial^2 f}{\partial y^2}(\vec x)$, vérifie :

$$\Delta f(\vec x) = \frac{\partial^2\tilde f}{\partial\rho^2}(\vec u)+\frac{1}{\rho}\frac{\partial\tilde f}{\partial\rho}(\vec u)+\frac{1}{\rho^2}\frac{\partial^2\tilde f}{\partial\theta^2}(\vec u)$$

Réponse. On considère la fonction $\tilde f(\rho,\theta)=(f\circ\vec\psi)(\rho,\theta)=f(\rho\cos\theta,\rho\sin\theta)$. Les formules (9.13) donnent, sachant $(J_{\vec\psi}(\rho,\theta)^T)^{-1} = \begin{pmatrix}\cos\theta & \frac{-1}{\rho}\sin\theta\\ \sin\theta & \frac{1}{\rho}\cos\theta\end{pmatrix}$:

$$\begin{pmatrix}g_1(\vec x)\\ g_2(\vec x)\end{pmatrix} = \begin{pmatrix}\frac{\partial f}{\partial x}(\vec x)\\ \frac{\partial f}{\partial y}(\vec x)\end{pmatrix} = (J_{\vec\psi}(\rho,\theta)^T)^{-1}.\begin{pmatrix}\frac{\partial\tilde f}{\partial\rho}(\vec u)\\ \frac{\partial\tilde f}{\partial\theta}(\vec u)\end{pmatrix} = \begin{pmatrix}\cos\theta\frac{\partial\tilde f}{\partial\rho}(\vec u)-\frac{1}{\rho}\sin\theta\frac{\partial\tilde f}{\partial\theta}(\vec u)\\ \sin\theta\frac{\partial\tilde f}{\partial\rho}(\vec u)+\frac{1}{\rho}\cos\theta\frac{\partial\tilde f}{\partial\theta}(\vec u)\end{pmatrix} \stackrel{\text{noté}}{=} \begin{pmatrix}\tilde g_1(\vec u)\\ \tilde g_2(\vec u)\end{pmatrix}.$$

De même :

$$\begin{pmatrix}\frac{\partial\frac{\partial f}{\partial x}}{\partial x}\\ \frac{\partial\frac{\partial f}{\partial x}}{\partial y}\end{pmatrix} = \begin{pmatrix}\frac{\partial g_1}{\partial x}(\vec x)\\ \frac{\partial g_1}{\partial y}(\vec x)\end{pmatrix} = (J_{\vec\psi}(\rho,\theta)^T)^{-1}.\begin{pmatrix}\frac{\partial\tilde g_1}{\partial\rho}(\vec u)\\ \frac{\partial\tilde g_1}{\partial\theta}(\vec u)\end{pmatrix} = \begin{pmatrix}\cos\theta\frac{\partial\tilde g_1}{\partial\rho}(\vec u)-\frac{1}{\rho}\sin\theta\frac{\partial\tilde g_1}{\partial\theta}(\vec u)\\ \sin\theta\frac{\partial\tilde g_1}{\partial\rho}(\vec u)+\frac{1}{\rho}\cos\theta\frac{\partial\tilde g_1}{\partial\theta}(\vec u)\end{pmatrix},$$

et :

$$\begin{pmatrix}\frac{\partial\frac{\partial f}{\partial y}}{\partial x}\\ \frac{\partial\frac{\partial f}{\partial y}}{\partial y}\end{pmatrix} = \begin{pmatrix}\frac{\partial g_2}{\partial x}(\vec x)\\ \frac{\partial g_2}{\partial y}(\vec x)\end{pmatrix} = \begin{pmatrix}---\\ \sin\theta\frac{\partial\tilde g_2}{\partial\rho}(\vec u)+\frac{1}{\rho}\cos\theta\frac{\partial\tilde g_2}{\partial\theta}(\vec u)\end{pmatrix}.$$

Et on a :

$$\begin{cases}\dfrac{\partial\tilde g_1}{\partial\rho}(\vec u) = \cos\theta\dfrac{\partial^2\tilde f}{\partial\rho^2}(\vec u)+\dfrac{1}{\rho^2}\sin\theta\dfrac{\partial\tilde f}{\partial\theta}(\vec u)-\dfrac{1}{\rho}\sin\theta\dfrac{\partial^2\tilde f}{\partial\rho\partial\theta}(\vec u),\\ \dfrac{\partial\tilde g_1}{\partial\theta}(\vec u) = -\sin\theta\dfrac{\partial\tilde f}{\partial\rho}(\vec u)+\cos\theta\dfrac{\partial^2\tilde f}{\partial\rho\partial\theta}(\vec u)-\dfrac{1}{\rho}\cos\theta\dfrac{\partial\tilde f}{\partial\theta}(\vec u)-\dfrac{1}{\rho}\sin\theta\dfrac{\partial^2\tilde f}{\partial\theta^2}(\vec u),\\ \dfrac{\partial\tilde g_2}{\partial\rho}(\vec u) = \sin\theta\dfrac{\partial^2\tilde f}{\partial\rho^2}(\vec u)-\dfrac{1}{\rho^2}\cos\theta\dfrac{\partial\tilde f}{\partial\theta}(\vec u)+\dfrac{1}{\rho}\cos\theta\dfrac{\partial^2\tilde f}{\partial\rho\partial\theta}(\vec u),\\ \dfrac{\partial\tilde g_2}{\partial\theta}(\vec u) = \cos\theta\dfrac{\partial\tilde f}{\partial\rho}(\vec u)+\sin\theta\dfrac{\partial^2\tilde f}{\partial\rho\partial\theta}(\vec u)-\dfrac{1}{\rho}\sin\theta\dfrac{\partial\tilde f}{\partial\theta}(\vec u)-\dfrac{1}{\rho}\sin\theta\dfrac{\partial^2\tilde f}{\partial\rho\partial\theta}(\vec u)+\dfrac{1}{\rho}\cos\theta\dfrac{\partial^2\tilde f}{\partial\theta^2}(\vec u).\end{cases}$$

D'où :

$$\begin{cases} \dfrac{\partial^2 f}{\partial x^2}(\vec{x}) & = \cos\theta^2 \dfrac{\partial^2 \tilde{f}}{\partial \rho^2} + \dfrac{1}{\rho^2}\sin\theta^2 \dfrac{\partial^2 \tilde{f}}{\partial \theta^2} - \dfrac{2}{\rho}\sin\theta\cos\theta \dfrac{\partial^2 \tilde{f}}{\partial\rho\partial\theta} + \dfrac{1}{\rho}\sin\theta^2 \dfrac{\partial \tilde{f}}{\partial \rho} + \dfrac{2}{\rho^2}\sin\theta\cos\theta \dfrac{\partial \tilde{f}}{\partial \theta}, \\ \dfrac{\partial^2 f}{\partial x\partial y}(\vec{x}) & = \cos\theta\sin\theta \dfrac{\partial^2 \tilde{f}}{\partial \rho^2} - \dfrac{1}{\rho^2}\sin\theta\cos\theta \dfrac{\partial^2 \tilde{f}}{\partial \theta^2} + \dfrac{1}{\rho}(\cos\theta^2 - \sin\theta^2) \dfrac{\partial^2 \tilde{f}}{\partial\rho\partial\theta} \\ & \quad - \dfrac{1}{\rho}\sin\theta\cos\theta \dfrac{\partial \tilde{f}}{\partial \rho} - \dfrac{1}{\rho^2}(\cos\theta^2 - \sin\theta^2) \dfrac{\partial^2 \tilde{f}}{\partial \theta^2}, \\ \dfrac{\partial^2 f}{\partial y^2}(\vec{x}) & = \sin\theta^2 \dfrac{\partial^2 \tilde{f}}{\partial \rho^2} + \dfrac{1}{\rho^2}\cos\theta^2 \dfrac{\partial^2 \tilde{f}}{\partial \theta^2} + \dfrac{2}{\rho}\sin\theta\cos\theta \dfrac{\partial^2 \tilde{f}}{\partial\rho\partial\theta} + \dfrac{1}{\rho}\cos\theta^2 \dfrac{\partial \tilde{f}}{\partial \rho} - \dfrac{2}{\rho^2}\sin\theta\cos\theta \dfrac{\partial \tilde{f}}{\partial \theta}. \end{cases}$$

D'où en particulier :

$$\frac{\partial^2 f}{\partial x^2}(\vec{x}) + \frac{\partial^2 f}{\partial y^2}(\vec{x}) = \frac{\partial^2 \tilde{f}}{\partial \rho^2}(\vec{u}) + \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \rho}(\vec{u}) + \frac{1}{\rho^2}\frac{\partial^2 \tilde{f}}{\partial \theta^2}(\vec{u})$$

ce qui est l'expression cherchée.

Remarque 9.16 À quoi ça sert ?

Beaucoup de problème physiques consistent à trouver une fonction f telle que $\Delta f = g$ où g est une fonction donnée. La fonction f s'exprime naturellement en fonction de x et y. Mais la fonction g ne dépend parfois que de ρ, et alors pour des raisons de symétries, on cherche une fonction f qui ne dépende également que de ρ (la distance à l'origine), et pas de "l'angle θ".

Il est alors naturel de chercher f sous la forme $f(x,y) = \tilde{f}(\rho,\theta) = \tilde{f}(\rho)$ qui ne dépend pas de θ, et on obtient immédiatement $\Delta \tilde{f} = \frac{\partial^2 \tilde{f}}{\partial \rho^2}(\rho) + \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \rho}(\rho) = g$, équation différentielle en ρ facile à intégrer (on pourra poser $\frac{\partial \tilde{f}}{\partial \rho}(\rho) = h(\log(\rho)) = h(w)$). On trouve $\tilde{f}(\rho)$, et comme $\tilde{f}(\rho) = f(x,y)$, on a trouvé f.

Et lors d'une résolution numérique approchée, on a ainsi transformé un problème à deux dimensions (x et y) en un problème en une dimension (ρ), beaucoup moins coûteux à résoudre.

Exercice 9.17 Montrer qu'en coordonnées sphériques :

$$\Delta f = \frac{\partial^2 \tilde{f}}{\partial \rho^2} + \frac{1}{\rho^2\cos\varphi^2}\frac{\partial^2 \tilde{f}}{\partial \theta^2} + \frac{1}{\rho^2}\frac{\partial^2 \tilde{f}}{\partial \varphi^2} + \frac{2}{\rho}\frac{\partial \tilde{f}}{\partial \rho} + \frac{\tan\varphi}{\rho^2}\frac{\partial \tilde{f}}{\partial \varphi}$$

Exercice 9.18 Soit $f \in C^2(V;\mathbb{R})$ où $V = \mathbb{R}^2 - \{(x,y) : y = 0,\ x \leq 0\}$, et soit le changement de coordonnées polaires $\vec{\psi}(\rho,\theta) = (\rho\cos\theta, \rho\sin\theta)$. Calculer les dérivées secondes $\frac{\partial^2 \tilde{f}}{\partial u_i \partial u_j}$ pour $u_i = \rho$ ou θ en fonction des $\frac{\partial f}{\partial x_i}$ et des $\frac{\partial^2 f}{\partial x_i \partial x_j}$ pour $x_i = x$ ou y.

Réponse. On a déjà (9.14) :

$$\begin{cases} \dfrac{\partial \tilde{f}}{\partial \rho}(\vec{u}) = \cos\theta \dfrac{\partial f}{\partial x}(\vec{x}) + \sin\theta \dfrac{\partial f}{\partial y}(\vec{x}) = \cos\theta\ h_1(\vec{u}) + \sin\theta\ h_2(\vec{u}), \\ \dfrac{\partial \tilde{f}}{\partial \theta}(\vec{u}) = -\rho\sin\theta \dfrac{\partial f}{\partial x}(\vec{x}) + \rho\cos\theta \dfrac{\partial f}{\partial y}(\vec{x}) = -\rho\sin\theta\ h_1(\vec{u}) + \rho\cos\theta\ h_2(\vec{u}). \end{cases}$$

où on a posé $h_1(\vec{u}) = (\frac{\partial f}{\partial x} \circ \vec{\psi})(\vec{u}) = \frac{\partial f}{\partial x}(\vec{x})$ et $h_2(\vec{u}) = (\frac{\partial f}{\partial y} \circ \vec{\psi})(\vec{u}) = \frac{\partial f}{\partial y}(\vec{x})$. D'où :

$$\begin{cases} \dfrac{\partial \frac{\partial \tilde{f}}{\partial \rho}}{\partial \rho}(\vec{u}) = \cos\theta \dfrac{\partial h_1}{\partial \rho}(\vec{u}) + \sin\theta \dfrac{\partial h_2}{\partial \rho}(\vec{u}), \\ \dfrac{\partial \frac{\partial \tilde{f}}{\partial \rho}}{\partial \theta}(\vec{u}) = -\sin\theta\ h_1(\vec{u}) + \cos\theta \dfrac{\partial h_1}{\partial \theta}(\vec{u}) + \cos\theta\ h_2(\vec{u}) + \sin\theta \dfrac{\partial h_2}{\partial \theta}(\vec{u}), \\ \dfrac{\partial \frac{\partial \tilde{f}}{\partial \theta}}{\partial \theta}(\vec{u}) = -\rho\cos\theta\ h_1(\vec{u}) - \rho\sin\theta \dfrac{\partial h_1}{\partial \theta}(\vec{u}) - \rho\sin\theta\ h_2(\vec{u}) + \rho\cos\theta \dfrac{\partial h_2}{\partial \theta}(\vec{u}). \end{cases}$$

Et :

$$\frac{\partial h_i}{\partial u_j}(\vec{u}) = \frac{\partial \frac{\partial f}{\partial x_i}}{\partial x}(\vec{x})\frac{\partial \psi_1}{\partial u_j}(\vec{u}) + \frac{\partial \frac{\partial f}{\partial x_i}}{\partial y}(\vec{x})\frac{\partial \psi_2}{\partial u_j}(\vec{u}),$$

i.e. :

$$\begin{cases} \dfrac{\partial h_1}{\partial \rho}(\vec{u}) = \dfrac{\partial^2 f}{\partial x^2}\cos\theta + \dfrac{\partial^2 f}{\partial x \partial y}\sin\theta, \\ \dfrac{\partial h_1}{\partial \theta}(\vec{u}) = -\dfrac{\partial^2 f}{\partial x^2}\rho\sin\theta + \dfrac{\partial^2 f}{\partial x \partial y}\rho\cos\theta, \\ \dfrac{\partial h_2}{\partial \rho}(\vec{u}) = \dfrac{\partial^2 f}{\partial x \partial y}\cos\theta + \dfrac{\partial^2 f}{\partial y^2}\sin\theta, \\ \dfrac{\partial h_2}{\partial \theta}(\vec{u}) = -\dfrac{\partial^2 f}{\partial x \partial y}\rho\sin\theta + \dfrac{\partial^2 f}{\partial y^2}\rho\cos\theta. \end{cases}$$

Finalement, on obtient :

$$\begin{cases} \dfrac{\partial^2 \tilde{f}}{\partial \rho^2}(\vec{u}) = \cos^2\theta\dfrac{\partial^2 f}{\partial x^2}(\vec{x}) + 2\cos\theta\sin\theta\dfrac{\partial^2 f}{\partial x \partial y}(\vec{x}) + \sin^2\theta\dfrac{\partial^2 f}{\partial y^2}(\vec{x}), \\ \dfrac{\partial^2 \tilde{f}}{\partial \rho \partial \theta}(\vec{u}) = -\sin\theta\,\dfrac{\partial f}{\partial x}(\vec{x}) + \cos\theta\,\dfrac{\partial f}{\partial y}(\vec{x}) \\ \qquad -\rho\cos\theta\sin\theta\dfrac{\partial^2 f}{\partial x^2}(\vec{x}) + \rho(\cos^2\theta - \sin^2\theta)\dfrac{\partial^2 f}{\partial x \partial y}(\vec{x}) + \rho\cos\theta\sin\theta\dfrac{\partial^2 f}{\partial y^2}(\vec{x}), \\ \dfrac{\partial^2 \tilde{f}}{\partial \theta^2}(\vec{u}) = -\rho\cos\theta\,\dfrac{\partial f}{\partial x}(\vec{x}) - \rho\sin\theta\dfrac{\partial f}{\partial y}(\vec{x}) \\ \qquad +\rho^2\sin^2\theta\dfrac{\partial^2 f}{\partial x^2}(\vec{x}) - 2\rho^2\sin\theta\cos\theta\dfrac{\partial^2 f}{\partial x \partial y}(\vec{x}) + \rho^2\cos^2\theta\dfrac{\partial^2 f}{\partial y^2}(\vec{x}). \end{cases}$$

Ce sont les dérivées cherchées, notées abusivement $\frac{\partial^2 f}{\partial \rho^2}$, $\frac{\partial^2 f}{\partial \rho \theta}$ et $\frac{\partial^2 f}{\partial \theta^2}$. Notez que cela peut se mettre sous la forme matricielle (non usitée) :

$$\begin{pmatrix} \frac{\partial^2 \tilde{f}}{\partial \rho^2}(\vec{u}) \\ \frac{\partial^2 \tilde{f}}{\partial \rho \theta}(\vec{u}) \\ \frac{\partial^2 \tilde{f}}{\partial \theta^2}(\vec{u}) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ -\sin & 0 & \cos \\ -\rho\cos & 0 & -\rho\sin \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x}(\vec{x}) \\ 0 \\ \frac{\partial f}{\partial y}(\vec{x}) \end{pmatrix} + \begin{pmatrix} cos^2 & 2\cos\sin & \sin^2 \\ -\rho\cos\sin & \rho(\cos^2 - \sin^2) & \rho\cos\sin \\ \rho^2\sin^2 & -2\rho^2\cos\sin & \rho^2\cos^2 \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(\vec{x}) \\ \frac{\partial^2 f}{\partial x \partial y}(\vec{x}) \\ \frac{\partial^2 f}{\partial y^2}(\vec{x}) \end{pmatrix}.$$

9.6 Formules de changement de base

9.6.1 Formules de changement de variables dans la base du système de coordonnées

Soit $\vec{\psi} : U \to V$ un changement de variables. On peut appliquer directement (8.13), lorsque $\vec{b}_j(\vec{x}) = d\vec{\psi}(\vec{u})(\vec{E}_j)$, toujours avec $(\vec{E}_i)_{i=1:n}$ la base canonique de $\mathbb{R}^2$ (pour l'espace des paramètres) et $(\vec{e}_i)_{i=1:n}$ une base donnée de $\mathbb{R}^2$ (pour l'espace géométrique).

Ici on se place dans l'espace géométrique, et on veut faire le changement de base $(\vec{e}_i) \to (\vec{b}_i(\vec{x}))$ en un point $\vec{x} = \vec{\psi}(\vec{u})$ donné :

Proposition 9.19 *Pour un changement de variables $\vec{\psi} : U \to V$ donné, pour $\vec{u} \in U$ fixé, et $(\vec{b}_i(\vec{x}))_{i=1,\dots,n}$ la base du système de coordonnées au point $\vec{x} = \vec{\psi}(\vec{u}) \in V$, i.e. $\vec{b}_j(\vec{x}) = d\vec{\psi}(\vec{u}).\vec{E}_j$:*
si $\vec{v} = \sum_{i=1}^n \alpha_i \vec{e}_i = \sum_{i=1}^n \beta_i \vec{b}_i(\vec{x})$ est un vecteur donné de l'espace géométrique, alors :

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = J_{\vec{\psi}^{-1}}(\vec{x}).\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \qquad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = J_{\vec{\psi}}(\vec{u}).\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}. \tag{9.16}$$

Preuve. On applique directement (8.13).

Exemple 9.20 Pour les coordonnées polaires, en un point $\vec{x} = \vec{\psi}(\vec{u})$ fixé et la base $(\vec{b}_1(\vec{x}), \vec{b}_2(\vec{x}))$ du système de coordonnées, si $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 = w_1\vec{b}_1(\vec{x}) + w_2\vec{b}_2(\vec{x})$, alors

$$\begin{cases} w_1 = \cos\theta v_1 + \sin\theta v_2, \\ w_2 = \dfrac{1}{\rho}(-\sin\theta v_1 + \cos\theta v_2). \end{cases}$$

Exemple 9.21 Et si $f : V \to \mathbb{R}$ est une fonction donnée :

$$\vec{\text{grad}} f(\vec{x}) = \begin{pmatrix} \frac{\partial f}{\partial x}(\vec{x}) \\ \frac{\partial f}{\partial y}(\vec{x}) \end{pmatrix}_{|(\vec{e}_1,\vec{e}_2)},$$

i.e. $\vec{\text{grad}} f(\vec{x}) = \frac{\partial f}{\partial x}(\vec{x})\vec{e}_1(\vec{x}) + \frac{\partial f}{\partial y}(\vec{x})\vec{e}_2)$. Et on a, avec les coordonnées polaires :

$$\vec{\text{grad}} f(\vec{x}) = \begin{pmatrix} \cos\theta\frac{\partial f}{\partial x}(\vec{x}) + \sin\theta\frac{\partial f}{\partial y}(\vec{x}) \\ \frac{1}{\rho}(-\sin\theta\frac{\partial f}{\partial x}(\vec{x}) + \cos\theta\frac{\partial f}{\partial y}(\vec{x})) \end{pmatrix}_{|(\vec{b}_1,\vec{b}_2)}$$

i.e. $\vec{\text{grad}} f(\vec{x}) = (\cos\theta\frac{\partial f}{\partial x}(\vec{x}) + \sin\theta\frac{\partial f}{\partial y}(\vec{x}))\vec{b}_1(\vec{x}) + \frac{1}{\rho}(-\sin\theta\frac{\partial f}{\partial x}(\vec{x}) + \cos\theta\frac{\partial f}{\partial y}(\vec{x}))\vec{b}_2(\vec{x})$.

Si on se place dans la base normalisée $\vec{e}_\rho = \vec{b}_1$ et $\vec{e}_\theta = \frac{1}{\rho}\vec{b}_2$ on obtient :

$$\vec{\text{grad}} f(\vec{x}) = \begin{pmatrix} \cos\theta\frac{\partial f}{\partial x}(\vec{x}) + \sin\theta\frac{\partial f}{\partial y}(\vec{x}) \\ -\sin\theta\frac{\partial f}{\partial x}(\vec{x}) + \cos\theta\frac{\partial f}{\partial y}(\vec{x}) \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta)},$$

i.e. $\vec{\text{grad}} f(\vec{x}) = (\cos\theta\frac{\partial f}{\partial x}(\vec{x}) + \sin\theta\frac{\partial f}{\partial y}(\vec{x}))\vec{e}_\rho(\vec{x}) + (-\sin\theta\frac{\partial f}{\partial x}(\vec{x}) + \cos\theta\frac{\partial f}{\partial y}(\vec{x}))\vec{e}_\theta(\vec{x})$. ∎

Exercice 9.22 Le faire pour les coordonnées sphériques. ∎

9.6.2 Le gradient dans la base de coordonnées

Le gradient $\vec{\text{grad}} f(\vec{x})$ est un vecteur qui peut s'exprimer dans différentes bases. Dans la base $(\vec{e}_i)$, et dans la base du système de coordonnées, on pose, en un point $\vec{x} = \vec{\psi}(\vec{u})$:

$$\vec{\text{grad}} f(\vec{x}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{x})\vec{e}_i = \sum_{i=1}^n \beta_i(\vec{u})\vec{b}_i(\vec{x}). \tag{9.17}$$

Exemple 9.23 Avec les coordonnées polaires on a, quand $\vec{x} = \vec{\psi}(\rho,\theta)$:

$$\begin{cases} \beta_1(\rho,\theta) = \cos\theta\dfrac{\partial f}{\partial x}(\vec{x}) + \sin\theta\dfrac{\partial f}{\partial y}(\vec{x}), \\ \beta_2(\rho,\theta) = -\dfrac{\sin\theta}{\rho}\dfrac{\partial f}{\partial x}(\vec{x}) + \dfrac{\cos\theta}{\rho}\dfrac{\partial f}{\partial y}(\vec{x}), \end{cases}$$

d'après l'exemple 9.21. ∎

L'application de (9.16), et (9.13) donnent :

Proposition 9.24 *Soit $f : V \to \mathbb{R}$ une fonction C^1 donnée, et soit $\tilde{f} : U \to \mathbb{R}$ la fonction définie par $\tilde{f}(\vec{u}) = f(\vec{x})$, i.e. $\tilde{f} = f \circ \vec{\psi}^{-1}$. Les coordonnées du vecteur $\vec{\text{grad}} f(\vec{x})$ dans la base $(\vec{b}_i(\vec{x}))_{i=1,\dots,n}$ du système de coordonnées au point $\vec{x}$ sont données à l'aide des $\frac{\partial \tilde{f}}{\partial u_j}(\vec{u})$ par, quand $\vec{x} = \vec{\psi}(\vec{u})$:*

$$\begin{pmatrix} \beta_1(\vec{u}) \\ \vdots \\ \beta_n(\vec{u}) \end{pmatrix} = J_{\vec{\psi}}(\vec{u})^{-1}.J_{\vec{\psi}}^{-T}(\vec{u}).\begin{pmatrix} \frac{\partial \tilde{f}}{\partial u_1}(\vec{u}) \\ \vdots \\ \frac{\partial \tilde{f}}{\partial u_n}(\vec{u}) \end{pmatrix} \tag{9.18}$$

Preuve. (9.16) donne $\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = J_{\vec{\psi}}(\vec{u})^{-1}.\begin{pmatrix} \frac{\partial f}{\partial x_1}(\vec{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\vec{x}) \end{pmatrix}$, puis on applique (9.13). ∎

Exercice 9.25 Calculer le gradient en coordonnées polaires. Appliquer le résultat à $f(x,y) = e^{\sqrt{x^2+y^2}}$.

Réponse. Soit $f(x,y)$ donnée et soit $\tilde{f}(\rho,\theta) = f(x,y)$. Dans l'application proposée on aura $\tilde{f}(\rho,\theta) = e^\rho$. Calculer le gradient en coordonnées polaires veut dire calculer les composantes du gradient dans la base

$(\vec{b}_1, \vec{b}_2)$ du système de coordonnées polaires en fonction de $\tilde{f}$. On a $(J_{\vec{\psi}}(\vec{u}))^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{\rho}\sin\theta & \frac{1}{\rho}\cos\theta \end{pmatrix}$, puis :

$$J(\vec{u})^{-1}.J(\vec{u})^{-T} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\rho^2} \end{pmatrix}, \tag{9.19}$$

d'où avec (9.17) et (9.18) :

$$\begin{aligned} \vec{\text{grad}} f(\vec{x}) &= \frac{\partial \tilde{f}}{\partial \rho}(\vec{u})\vec{b}_1 + \frac{1}{\rho^2}\frac{\partial \tilde{f}}{\partial \theta}(\vec{u})\vec{b}_2 = \begin{pmatrix} \frac{\partial \tilde{f}}{\partial \rho}(\vec{u}) \\ \frac{1}{\rho^2}\frac{\partial \tilde{f}}{\partial \theta}(\vec{u}) \end{pmatrix}_{|(\vec{b}_1,\vec{b}_2)} \\ &= \frac{\partial \tilde{f}}{\partial \rho}(\vec{u})\vec{e}_\rho + \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \theta}(\vec{u})\vec{e}_\theta = \begin{pmatrix} \frac{\partial \tilde{f}}{\partial \rho}(\vec{u}) \\ \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \theta}(\vec{u}) \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta)}, \end{aligned}$$

où $(\vec{e}_\rho = \vec{b}_1, \vec{e}_\theta = \frac{\vec{b}_2}{\rho})$ est la base normalisée, la dernière expression étant la plus utilisée (expression dans la base normalisée $(\vec{e}_\rho, \vec{e}_\theta)$).

Dans l'application proposé $f(x,y) = e^\rho$ on obtient $\vec{\text{grad}} f(\vec{x}) = \begin{pmatrix} e^\rho \\ 0 \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta)}$, i.e. $\vec{\text{grad}} f(x,y) = \begin{pmatrix} e^{\sqrt{x^2+y^2}} \\ 0 \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta)}$. ▪

Remarque 9.26 Si M est une matrice dont les colonnes sont 2 à 2 orthogonales et de norme λ_i, i.e. une matrice vérifiant $M^T.M = \text{diag}(\lambda_i^2)$, alors $M^{-1} = \text{diag}(\frac{1}{\lambda_i^2}).M^T$. En effet, $M.(\text{diag}(\frac{1}{\lambda_i^2}).M^T) = (M.\text{diag}(\frac{1}{\lambda_i}).(M.\text{diag}(\frac{1}{\lambda_i}))^T = I$ car les matrices entre parenthèses sont unitaires (leurs colonnes sont orthonormées), et $(\text{diag}(\frac{1}{\lambda_i^2}).M^T).M = \text{diag}(\frac{1}{\lambda_i^2}).(M^T.M) = \text{diag}(\frac{1}{\lambda_i^2}).\text{diag}(\lambda_i^2) = I$. Et donc $M^T = \text{diag}(\lambda_i^2).M^{-1}$.

D'où, pour un système de coordonnées orthogonales : $J_{\vec{\psi}}^T.J_{\vec{\psi}} = \text{diag}(\lambda_i^2)$: on retrouve (9.19).

Noter que quand les colonnes sont orthogonales il n'y a aucune raison pour avoir les lignes orthogonales. Par exemple $M = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$ a ses colonnes orthogonales (et on a $M^T.M = \text{diag}(1,\rho^2)$), mais ses lignes ne sont pas orthogonales (sauf si $\rho = 1$ auquel cas la matrice est unitaire). Ici on a $M^tM = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix}$ (les colonnes de M sont orthogonales), et on a $MM^T = \begin{pmatrix} \cos^2\theta + \rho^2\sin^2\theta & (1-\rho^2)\cos\theta\sin\theta \\ (1-\rho^2)\cos\theta\sin\theta & \sin^2\theta + \rho^2\cos^2\theta \end{pmatrix}$ non diagonale quand $\rho \neq 1$ (les lignes de M ne sont pas orthogonales). ▪

Exercice 9.27 Soient $(\vec{b}_i(\vec{x}))$ les vecteurs de base en $\vec{x}$ d'un système de coordonnées orthogonal (i.e. tels que les $\vec{b}_i(\vec{x})$ sont 2 à 2 othogonaux). Soient $\lambda_i(\vec{x}) = ||\vec{b}_i(\vec{x})||$, et soient :

$$\vec{e}_{u_i}(\vec{x}) = \frac{\vec{b}_i(\vec{x})}{||\vec{b}_i(\vec{x})||} = \frac{\vec{b}_i(\vec{x})}{\lambda_i(\vec{x})},$$

les vecteurs normalisés. Montrer que : $\vec{\text{grad}} f(\vec{x}) = \displaystyle\sum_{i=1}^n \frac{1}{\lambda_i(\vec{u})}\frac{\partial \tilde{f}}{\partial u_i}(\vec{u})\, \vec{e}_{u_i}(\vec{x})$ quand $\tilde{f} = f \circ \vec{\psi}^{-1}$.

Réponse. Le système de coordonnées étant orthogonal, on a $\vec{b}_i^T.\vec{b}_j = \lambda_i^2\delta_{ij}$, et donc

$$(J_{\vec{\psi}}(\vec{u}))^T.J_{\vec{\psi}}(\vec{u}) = \text{diag}(\lambda_i^2),$$

où $\text{diag}(\lambda_i^2)$ est la matrice diagonale de coeff les λ_i^2 sur la diagonale. Et donc $(J_{\vec{\psi}}(\vec{u}))^{-1}.J_{\vec{\psi}}(\vec{u})^{-T} = \text{diag}(\frac{1}{\lambda_i^2})$, et avec (9.18) :

$$\vec{\text{grad}} f(\vec{x}) = \begin{pmatrix} \frac{1}{\lambda_1^2}\frac{\partial \tilde{f}}{\partial u_1}(\vec{u}) \\ \vdots \\ \frac{1}{\lambda_n^2}\frac{\partial \tilde{f}}{\partial u_n}(\vec{u}) \end{pmatrix}_{|(\vec{b}_i)} = \begin{pmatrix} \frac{1}{\lambda_1}\frac{\partial \tilde{f}}{\partial u_1}(\vec{u}) \\ \vdots \\ \frac{1}{\lambda_n}\frac{\partial \tilde{f}}{\partial u_n}(\vec{u}) \end{pmatrix}_{|(\vec{e}_{u_i})},$$

qui est le résultat annoncé. ▪

Exercice 9.28 Appliquer l'exercice précédent au cas des coordonnées polaires et sphériques. On montrera en particulier qu'en sphérique :

$$\vec{\mathrm{grad}} f(\vec{x}) = \begin{pmatrix} \frac{\partial \tilde{f}}{\partial \rho}(\vec{u}) \\ \frac{1}{\rho\cos\varphi}\frac{\partial \tilde{f}}{\partial \theta}(\vec{u}) \\ \frac{1}{\rho}\frac{\partial \tilde{f}}{\partial \varphi}(\vec{u}) \end{pmatrix}_{|(\vec{e}_\rho,\vec{e}_\theta,\vec{e}_\varphi)}$$

▪

9.6.3 Divergence et rotationnel dans les nouvelles coordonnées

Le paragraphe précédent traitait des fonctions à valeurs scalaires. On s'intéresse ici à des fonctions à valeurs vectorielles $\vec{f} : \mathbb{R}^n \to \mathbb{R}^n$:

$$\vec{f} : \begin{cases} \mathbb{R}^n \to \mathbb{R}^n \\ \vec{x} = \sum_{i=1}^n x_i \vec{e}_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{|(\vec{e}_i)} \mapsto \vec{f}(\vec{x}) = \sum_{i=1}^n f_i(\vec{x})\vec{e}_i = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix}_{|(\vec{e}_i)} \end{cases}.$$

Soit $\vec{\psi} : \vec{u} \in U \to \vec{x} \in V$ un changement de variables. Pour chaque point $\vec{u} \in U$, on pose $\vec{x} = \vec{\psi}(\vec{u})$, et on dispose de la base $(\vec{b}_i(\vec{x}))_{i=1,\dots,n}$ du système de coordonnées, avec "$\vec{b}_i = J_{\vec{\psi}}.\vec{E}_i$" pour tout i.

Le vecteur $\vec{f}(\vec{x}) = \sum_{i=1}^n f_i(\vec{x})\vec{e}_i$ s'écrit également $\vec{f}(\vec{x}) = \sum_{i=1}^n \beta_i(\vec{u})\vec{b}_i(\vec{x})$, i.e.

$$\vec{f}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix}_{|(\vec{e}_i)} = \begin{pmatrix} \beta_1(\vec{u}) \\ \vdots \\ \beta_n(\vec{u}) \end{pmatrix}_{|(\vec{b}_i)}.$$

On définit $\tilde{\vec{f}} : U \to \mathbb{R}^n$ par $\tilde{\vec{f}}(\vec{u}) = \vec{f}(\vec{x})$, i.e. $\tilde{\vec{f}} = \vec{f} \circ \vec{\psi}^{-1}$, i.e.

$$f_i(\vec{x}) = \tilde{f}_i(\vec{u}) \quad \forall i = 1, \dots, n.$$

On a donc posé, quand $\vec{x} = \vec{\psi}(\vec{u})$:

$$\vec{f}(\vec{x}) = \tilde{\vec{f}}(\vec{u}) = \sum_{i=1}^n f_i(\vec{x})\vec{e}_i = \sum_{i=1}^n \tilde{f}_i(\vec{u})\vec{e}_i = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix}_{|(\vec{e}_i)} = \begin{pmatrix} \tilde{f}_1(\vec{u}) \\ \vdots \\ \tilde{f}_n(\vec{u}) \end{pmatrix}_{|(\vec{e}_i)},$$

Et les formules de changement de coordonnées donnent :

$$\begin{pmatrix} \tilde{f}_1(\vec{u}) \\ \vdots \\ \tilde{f}_n(\vec{u}) \end{pmatrix} = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{pmatrix} = (J_{\vec{\psi}}(\vec{u})).\begin{pmatrix} \beta_1(\vec{u}) \\ \vdots \\ \beta_n(\vec{u}) \end{pmatrix} = \begin{pmatrix} \sum_k \frac{\partial \psi_1}{\partial u_k}(\vec{u})\beta_k(\vec{u}) \\ \vdots \\ \sum_k \frac{\partial \psi_n}{\partial u_k}(\vec{u})\beta_k(\vec{u}) \end{pmatrix}. \tag{9.20}$$

Et les $\tilde{f}_i$ vérifient, pour tout $i, j = 1, \dots, n$:

$$\frac{\partial \tilde{f}_i}{\partial u_j} = \sum_{k=1}^n \frac{\partial^2 \psi_i}{\partial u_j \partial u_k}\beta_k + \sum_{k=1}^n \frac{\partial \psi_i}{\partial u_k}\frac{\partial \beta_k}{\partial u_j} = (H_{\psi_i}.\vec{\beta})_j + (J_{\vec{\psi}}.J_{\vec{\beta}})_{ij}$$

où H_i est la matrice Hessienne (symétrique) de ψ_i, et $J_{\vec{\psi}}$ et $J_{\vec{\beta}}$ les matrices jacobiennes de $\vec{\psi}$ et $\vec{\beta}$.

D'où la matrice jacobienne $J_{\tilde{\vec{f}}}$ de $\tilde{\vec{f}}$:

$$[J_{\tilde{\vec{f}}}]_{ij} = [(H_{\psi_i}.\vec{\beta})_j] + [J_{\vec{\psi}}.J_{\vec{\beta}}]_{ij}. \tag{9.21}$$

D'où

Proposition 9.29 *On a :*

$$\begin{aligned} \frac{\partial f_i}{\partial x_j} = [J_{\vec f}]_{ij} &= \sum_k \Big([(H_{\psi_i}.\vec\beta)_k] + [J_{\vec\psi}.J_{\vec\beta}]_{ik} \Big) [J_{\vec\psi^{-1}}]_{kj} \\ &= \sum_k \sum_\ell \Big(\frac{\partial^2 \psi_i}{\partial u_k \partial u_\ell} \beta_\ell + \frac{\partial \psi_i}{\partial u_\ell} \frac{\partial \beta_\ell}{\partial u_k} \Big) \frac{\partial u_k}{\partial x_j} \end{aligned} \tag{9.22}$$

On a ainsi toutes les dérivées $\frac{\partial f_i}{\partial x_j}(\vec x)$ en fonction des $\beta_\ell(\vec u)$ et des $\frac{\partial \beta_\ell}{\partial u_k}(\vec u)$.

Preuve. En effet, $\vec f = \tilde{\vec f} \circ \vec\psi^{-1}$ donne $J_{\vec f} = J_{\tilde{\vec f}}.J_{\vec\psi^{-1}}$, et donc $[J_{\vec f}]_{ij} = \sum_k [J_{\tilde{\vec f}}]_{ik} [J_{\vec\psi^{-1}}]_{kj}$. ▪

Exercice 9.30 Calculer $J_{\vec f}$ et $\mathrm{div} \vec f$ en coordonnées polaires.

Réponse. On pose $\vec u = (\rho, \theta)$. On a $\psi_1(\vec u) = \rho \cos\theta$ et $\psi_2(\vec u) = \rho \sin\theta$. On a $J_{\vec\psi}(\vec u) = \begin{pmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{pmatrix}$, et $\vec b_1(\vec u) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ et $\vec b_2 = \begin{pmatrix} -\rho\sin\theta \\ \rho\cos\theta \end{pmatrix}$.

Puis $H(\psi_1)(\vec u) = \begin{pmatrix} 0 & -\sin\theta \\ -\sin\theta & -\rho\cos\theta \end{pmatrix}$ et $H(\psi_2)(\vec u) = \begin{pmatrix} 0 & \cos\theta \\ \cos\theta & -\rho\sin\theta \end{pmatrix}$, d'où

$$[(H_{\psi_i}.\vec\beta)_j] = \begin{pmatrix} -\sin\theta\beta_2, & -\sin\theta\beta_1 - \rho\cos\theta\beta_2 \\ \cos\theta\beta_2, & \cos\theta\beta_1 - \rho\sin\theta\beta_2 \end{pmatrix}.$$

Puis $(J_{\vec\psi}(\vec u))^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\frac{1}{\rho}\sin\theta & \frac{1}{\rho}\cos\theta \end{pmatrix}$. D'où :

$$[(H_{\psi_i}.\vec\beta)_j].(J_{\vec\psi})^{-1} = \begin{pmatrix} \frac{1}{\rho}\sin\theta^2\beta_1, & -\frac{1}{\rho}\cos\theta\sin\theta\beta_1 - \beta_2 \\ -\frac{1}{\rho}\cos\theta\sin\theta\beta_1 + \beta_2, & \frac{1}{\rho}\cos\theta^2\beta_1 \end{pmatrix}$$

Et :

$$\begin{aligned} (J_{\vec\psi}.J_{\vec\beta}.(J_{\vec\psi})^{-1})_{11} &= \cos\theta^2 \frac{\partial\beta_1}{\partial\rho} - \rho\cos\theta\sin\theta \frac{\partial\beta_2}{\partial\rho} - \frac{1}{\rho}\cos\theta\sin\theta\frac{\partial\beta_1}{\partial\theta} + \sin\theta^2\frac{\partial\beta_2}{\partial\theta} \\ (J_{\vec\psi}.J_{\vec\beta}.(J_{\vec\psi})^{-1})_{12} &= \cos\theta\sin\theta\frac{\partial\beta_1}{\partial\rho} - \rho\sin\theta^2\frac{\partial\beta_2}{\partial\rho} + \frac{1}{\rho}\cos\theta^2\frac{\partial\beta_1}{\partial\theta} - \cos\theta\sin\theta\frac{\partial\beta_2}{\partial\theta} \\ (J_{\vec\psi}.J_{\vec\beta}.(J_{\vec\psi})^{-1})_{21} &= \cos\theta\sin\theta\frac{\partial\beta_1}{\partial\rho} + \rho\cos\theta^2\frac{\partial\beta_2}{\partial\rho} - \frac{1}{\rho}\sin\theta^2\frac{\partial\beta_1}{\partial\theta} - \cos\theta\sin\theta\frac{\partial\beta_2}{\partial\theta} \\ (J_{\vec\psi}.J_{\vec\beta}.(J_{\vec\psi})^{-1})_{22} &= \sin\theta^2\frac{\partial\beta_1}{\partial\rho} + \rho\cos\theta\sin\theta\frac{\partial\beta_2}{\partial\rho} + \frac{1}{\rho}\cos\theta\sin\theta\frac{\partial\beta_1}{\partial\theta} + \cos\theta^2\frac{\partial\beta_2}{\partial\theta} \end{aligned}$$

D'où :

$$\mathrm{div}\vec f(\vec x) = \frac{1}{\rho}\beta_1 + \frac{\partial\beta_1}{\partial\rho} + \frac{\partial\beta_2}{\partial\theta}$$

Attention : le calcul a été fait dans la base $(\vec b_i)$ du système de coordonnées. Si on prend la base usuelle $\vec e_\rho = \vec b_1$ et $\vec e_\theta = \frac{1}{\rho}\vec b_2$, alors $\vec f(\vec x) = f_1(\vec x)\vec e_1 + f_2(\vec x)\vec e_1 = f_\rho(\vec u)\vec e_\rho(\vec u) + f_\theta(\vec u)\vec e_\theta(\vec u)$ où $f_\rho = \beta_1$ et $f_\theta = \rho\beta_2$, et alors :

$$\mathrm{div}\vec f(\vec x) = \frac{1}{\rho} f_\rho(\vec u) + \frac{\partial f_\rho}{\partial\rho}(\vec u) + \frac{1}{\rho}\frac{\partial f_\theta}{\partial\theta}(\vec u),$$

expression usuelle de la divergence en coordonnées polaires. ▪

Proposition 9.31 *Soient $(\vec b_i(\vec x))$ les vecteurs de base d'un système de coordonnées orthogonal $\vec\psi$ (i.e. système de coordonnées tel que les $\vec b_i(\vec x)$ sont 2 à 2 othogonaux pour tout $\vec x$). On pose $\lambda_j(\vec x) = ||\vec b_j(\vec x)|| = \left(\sum_k \frac{\partial\psi_k}{\partial u_j}(\vec u)^2\right)^{\frac{1}{2}}$ pour tout $j = 1, ..., n$, où $\vec x = \vec\psi(\vec u)$.*

On décompose $\vec f$ sur les bases $(\vec e_i)$ et $(\vec b_i)$ comme $\vec f(\vec x) = \sum_{i=1}^n f_i(\vec x)\vec e_i = \sum_{i=1}^n \beta_i(\vec u)\vec b_i(\vec x)$ lorsque $\vec x = \vec\psi(\vec u)$. Alors :

$$\begin{aligned} \mathrm{div}\vec f &= \sum_{i=1}^n \beta_i \Big(\sum_{k=1}^n \frac{1}{\lambda_k}\frac{\partial\lambda_k}{\partial u_i} \Big) + \sum_{i=1}^n \frac{\partial\beta_i}{\partial u_i} \\ &= \frac{1}{\prod_{k=1}^n \lambda_k} \sum_{i=1}^n \Big(\frac{\partial(\beta_i(\prod_{k=1}^n \lambda_k))}{\partial u_i} \Big) \end{aligned} \tag{9.23}$$

où $\mathrm{div}\vec f$ est pris en $\vec x$ et le membre de droite est pris en $\vec u = \vec\psi^{-1}(\vec x)$.

Preuve. On dispose de (9.22) et donc on a $\mathrm{div}\vec{f}(\vec{x}) = \sum_{i,k,\ell} \frac{\partial^2\psi_i}{\partial u_k \partial u_\ell}\beta_\ell \frac{\partial u_k}{\partial x_i} + \frac{\partial \psi_i}{\partial u_\ell}\frac{\partial \beta_\ell}{\partial u_k}\frac{\partial u_k}{\partial x_i}$.

On a $\lambda_k(\vec{x}) = ||\vec{b}_k(\vec{x})|| = (\sum_j \frac{\partial\psi_j}{\partial u_k}(\vec{u})^2)^{\frac12}$, d'où :

$$\frac{\partial\lambda_k}{\partial u_i} = \frac12 (\sum_j (\frac{\partial\psi_j}{\partial u_k})^2)^{-\frac12} \sum_j 2 \frac{\partial^2\psi_j}{\partial u_i u_k}\frac{\partial\psi_j}{\partial u_k} = \frac{1}{\lambda_k}\sum_j \frac{\partial^2\psi_j}{\partial u_i u_k}\frac{\partial\psi_j}{\partial u_k},$$

d'où :

$$\sum_{i=1}^n \beta_i \Big(\sum_{k=1}^n \frac{1}{\lambda_k}\frac{\partial\lambda_k}{\partial u_i}\Big) = \sum_{i,j,k=1}^n \beta_i \frac{1}{\lambda_k^2}\frac{\partial^2\psi_j}{\partial u_i u_k}\frac{\partial\psi_j}{\partial u_k}.$$

Comme $J_{\vec{\psi}}^T.J_{\vec{\psi}} = \mathrm{diag}(\lambda_k^2)$ par hypothèse d'orthogonalité du système de coordonnées, et donc que $(J_{\vec{\psi}})^{-1} = \mathrm{diag}(\frac{1}{\lambda_k^2})J_{\vec{\psi}}^T = [\frac{1}{\lambda_i^2}\frac{\partial\psi_j}{\partial u_i}]_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ (voir remarque 9.26), on obtient :

$$\sum_{i=1}^n \beta_i \Big(\sum_{k=1}^n \frac{1}{\lambda_k}\frac{\partial\lambda_k}{\partial u_i}\Big) = \sum_{i,k,j=1}^n \beta_i [(J_{\vec{\psi}})^{-1}]_{kj}\frac{\partial^2\psi_j}{\partial u_i u_k} = \sum_{i,k,j=1}^n \beta_i \frac{\partial u_k}{\partial x_j}\frac{\partial^2\psi_j}{\partial u_i u_k},$$

ce qui est le premier terme cherché.

Puis :

$$\sum_{i,j,k}\frac{\partial\psi_i}{\partial u_j}\frac{\partial\beta_j}{\partial u_k}\frac{\partial\psi_k^{-1}}{\partial x_i} = \sum_{i,j,k}\frac{\partial\psi_k^{-1}}{\partial x_i}\frac{\partial\psi_i}{\partial u_j}\frac{\partial\beta_j}{\partial u_k},$$

avec $\sum_i \frac{\partial\psi_k^{-1}}{\partial x_i}\frac{\partial\psi_i}{\partial u_j} = [(J_{\vec{\psi}})^{-1}.J_{\vec{\psi}}]_{kj} = [\delta_{kj}]$, on obtient :

$$\sum_{i,j,k}\frac{\partial\psi_i}{\partial u_j}\frac{\partial\beta_j}{\partial u_k}\frac{\partial\psi_k^{-1}}{\partial x_i} = \sum_{j,k}[\delta_{kj}]\frac{\partial\beta_j}{\partial u_k} = \sum_j \frac{\partial\beta_j}{\partial u_j},$$

ce qui est le second terme cherché. ∎

Corollaire 9.32 *Soit $(\vec{e}_i)$ la base orthonormée associée au système orthogonal $(\vec{b}_i(\vec{x}))$: on a $\vec{e}_{u_i} = \frac{\vec{b}_i}{\lambda_i}$, avec $\lambda_i = ||\vec{b}_i||$. Si $\vec{f}(\vec{x}) = \sum_i f_i(\vec{x})\vec{e}_i = \sum_i f_{u_i}(\vec{u})\vec{e}_{u_i}(\vec{u})$ où $\vec{x} = \vec{\psi}(\vec{u})$, alors :*

$$\begin{aligned}\mathrm{div}\vec{f}(\vec{x}) &= \frac{1}{\prod_{k=1,\dots,n}\lambda_k}\sum_{i=1}^n \Big(\frac{\partial(f_{u_i}(\prod_{\substack{k=1,\dots,n\\ k\neq i}}\lambda_k))}{\partial u_i}\Big)(\vec{u})\\ &= \sum_{i=1}^n \frac{f_{u_i}}{\lambda_i}\Big(\sum_{k=1}^n\frac{1}{\lambda_k}\frac{\partial\lambda_k}{\partial u_i}\Big) - \sum_{i=1}^n \frac{f_{u_i}}{\lambda_i^2}\frac{\partial\lambda_i}{\partial u_i} + \sum_{i=1}^n\frac{1}{\lambda_i}\frac{\partial f_{u_i}}{\partial u_i}.\end{aligned}\tag{9.24}$$

Preuve. On a $\beta_i = \frac{f_{u_i}}{\lambda_i}$. D'où le résultat. ∎

Exercice 9.33 Retrouver l'expression de la divergence en coordonnées polaires. ∎

Exercice 9.34 Montrer que pour les coordonnées sphériques (8.4) on a :

$$\mathrm{div}\vec{f}(\vec{x}) = (\frac{\partial f_\rho}{\partial\rho} + \frac{1}{\rho\cos\varphi}\frac{\partial f_\theta}{\partial\theta} + \frac1\rho\frac{\partial f_\varphi}{\partial\varphi} + 2\frac{f_\rho}{\rho} - \tan\varphi\frac{f_\varphi}{\rho})(\rho,\theta,\varphi),$$

où on a posé $\vec{f}(\vec{x}) = f_\rho(\vec{u})\vec{e}_\rho(\vec{u}) + f_\theta(\vec{u})\vec{e}_\theta(\vec{u}) + f_\varphi(\vec{u})\vec{e}_\varphi(\vec{u})$. ∎

Exercice 9.35 En déduire que pour les coordonnées sphériques (8.6) (souvent utilisées par les physiciens) :

$$\mathrm{div}\vec{f}(\vec{x}) = (\frac{\partial f_\rho}{\partial\rho} + \frac{1}{\rho\sin\tilde\varphi}\frac{\partial f_\theta}{\partial\theta} + \frac1\rho\frac{\partial f_{\tilde\varphi}}{\partial\tilde\varphi} + 2\frac{f_\rho}{\rho} + \mathrm{cotan}\tilde\varphi\frac{f_{\tilde\varphi}}{\rho})(\rho,\theta,\tilde\varphi),$$

où on a posé $\vec{f}(\vec{x}) = f_\rho(\vec{u})\vec{e}_\rho(\vec{u}) + f_\theta(\vec{u})\vec{e}_\theta(\vec{u}) + f_{\tilde\varphi}(\vec{u})\vec{e}_{\tilde\varphi}(\vec{u})$.

Indication : remarquer que $\vec{e}_{\tilde\varphi}(\rho,\theta,\tilde\varphi) = -\vec{e}_\varphi(\rho,\theta,\frac\pi2-\varphi) = -\vec{e}_\varphi(\rho,\theta,\tilde\varphi(\varphi))$ et donc que :
$f_{\tilde\varphi}(\rho,\theta,\tilde\varphi) = -f_\varphi(\rho,\theta,\frac\pi2-\varphi) = -f_\varphi(\rho,\theta,\tilde\varphi(\varphi))$.
Puis que, sachant $\frac{\partial\tilde\varphi}{\partial\varphi} = -1$:
$\frac{\partial f_{\tilde\varphi}}{\partial\tilde\varphi}(\rho,\theta,\tilde\varphi) = \frac{\partial f_\varphi}{\partial\varphi}(\rho,\theta,\tilde\varphi)$. ∎

Définition 9.36 Et si $n = 3$ (donc dans $\mathbb{R}^3$), le rotationnel de $\vec{f} \in C^1(\mathbb{R}^3;\mathbb{R}^3)$ en $\vec{x}$ est définie par, dans la base canonique de $\mathbb{R}^3$:

$$(\vec{\mathrm{rot}}\vec{f})(\vec{x}) = \begin{pmatrix} \frac{\partial f_3}{\partial x_2}(\vec{x}) - \frac{\partial f_2}{\partial x_3}(\vec{x}) \\ \frac{\partial f_1}{\partial x_3}(\vec{x}) - \frac{\partial f_3}{\partial x_1}(\vec{x}) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}) - \frac{\partial f_1}{\partial x_2}(\vec{x}) \end{pmatrix} \quad \in \mathbb{R}^3 \tag{9.25}$$

(expression vectorielle, donc dont les coordonnées dépendent d'une base.)

On peut également noté, avec la convention d'indice $f_4 = f_1$, $f_5 = f_2$, $x_4 = x_1$ et $x_5 = x_2$:

$$\begin{aligned}(\vec{\mathrm{rot}}\vec{f})(\vec{x}) &= \sum_{i=1}^{3}(\frac{\partial f_{i+2}}{\partial x_{i+1}}(\vec{x}) - \frac{\partial f_{i+1}}{\partial x_{i+2}}(\vec{x}))\ \vec{e}_i \\ &\overset{\text{noté}}{=} (\nabla \wedge \vec{f})(\vec{x}) \overset{\text{noté}}{=} \det(\begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{pmatrix})(\vec{x}).\end{aligned}$$

Exercice 9.37 Calculer le rotationnel en coordonnées polaires. ▟

Exercice 9.38 Soient $(\vec{b}_i(\vec{x}))$ les vecteurs de base d'un système de coordonnées orthogonal dans $\mathbb{R}^3$ (les $\vec{b}_i(\vec{u})$ sont 2 à 2 othogonaux pour tout $\vec{u} \in U$). Soient $\lambda_i(\vec{x}) = ||\vec{b}_i(\vec{x})||$, et soient $\vec{e}_{u_i}(\vec{x}) = \frac{\vec{b}_i(\vec{x})}{\lambda_i(\vec{x})}$ les vecteurs orthonormaux de base.

On pose $\vec{f}(\vec{x}) = f_1(\vec{x})\vec{e}_1 + ... + f_3(\vec{x}) = f_{u_1}(\vec{u})\vec{e}_{u_1}(\vec{u}) + ... + f_{u_3}(\vec{u})\vec{e}_{u_3}(\vec{u})$. Montrer que, dans $\mathbb{R}^3$ et la base $(\vec{e}_{u_i})$:

$$\vec{\mathrm{rot}}\vec{f} = \frac{1}{\lambda_1\lambda_2\lambda_3} \det \begin{pmatrix} \lambda_1\vec{e}_{u_1} & \frac{\partial}{\partial u_1} & \lambda_1\beta_1 \\ \lambda_2\vec{e}_{u_2} & \frac{\partial}{\partial u_2} & \lambda_2\beta_2 \\ \lambda_3\vec{e}_{u_3} & \frac{\partial}{\partial u_3} & \lambda_3\beta_3 \end{pmatrix}$$

▟

10 Annexe : rappels de formules trigonométriques

La fonction $t \to \sin t$ s'annule aux points $t = k\pi$ pour $k \in \mathbb{Z}$, et la fonction $t \to \cos t$ s'annule aux points $t = \frac{\pi}{2} + k\pi$ pour $k \in \mathbb{Z}$. On écrit les formules génériques, à considérer là où elles sont bien définies.

1. $\sin^2 t + \cos^2 t = 1$.
2. $\sec t \stackrel{\text{déf}}{=} \dfrac{1}{\cos t}$, $\qquad \tan t = \dfrac{\sin t}{\cos t}$.
3. $\sin(-t) = -\sin t$, $\qquad \cos(-t) = \cos t$.
4. $\sin(\dfrac{\pi}{2} - t) = \cos t$, $\qquad \cos(\dfrac{\pi}{2} - t) = \sin t$.
5. $\sin(\dfrac{\pi}{2} + t) = \cos t$, $\qquad \cos(\dfrac{\pi}{2} + t) = -\sin t$.
6. $\sin(t \pm s) = \sin t \cos s \pm \cos t \sin s$, $\qquad \cos(t \pm s) = \cos t \cos s \mp \sin t \sin s$.
7. $\sin(2t) = 2 \sin t \cos t$, $\qquad \cos(2t) = \cos^2 t - \sin^2 t = 2\cos^2 t - 1 = 1 - 2\sin^2 t$.
8. $\tan(2t) = \frac{2\tan t}{1 - \tan^2 t}$.
9. $\sin \dfrac{t}{2} = \pm\sqrt{\dfrac{1 - \cos t}{2}}$, $\qquad \cos \dfrac{t}{2} = \pm\sqrt{\dfrac{1 + \cos t}{2}}$.
10. $(\sin t - \cos t)^2 = 1 - \sin(2t)$.
11. $\tan(s + t) = \dfrac{\tan s + \tan t}{1 - \tan s \tan t}$.
12. $\sin t \sin s = -\dfrac{1}{2}(\cos(t+s) - \cos(t-s))$, $\qquad \cos t \cos s = \dfrac{1}{2}(\cos(t+s) + \cos(t-s))$.
13. $\sin t \cos s = \dfrac{1}{2}(\sin(t-s) + \sin(t+s))$, $\qquad \cos t \sin s = \dfrac{1}{2}(\sin(t+s) - \sin(t-s))$.
14. $\sin t + \sin s = 2\sin(\dfrac{t+s}{2}) \cos \dfrac{t-s}{2})$, $\qquad \sin t - \sin s = 2\cos(\dfrac{t+s}{2}) \sin \dfrac{t-s}{2})$.
15. $\cos t + \cos s = 2\cos(\dfrac{t+s}{2}) \cos \dfrac{t-s}{2})$, $\qquad \cos t - \cos s = -2\sin(\dfrac{t+s}{2}) \sin \dfrac{t-s}{2})$.
16. Pour $\sin t + \cos s$, écrire que $\cos s = \sin(\dfrac{\pi}{2} - s)$ par exemple.
17. $e^{it} = \cos t + i \sin t$, $\qquad \cos t = \dfrac{e^{it} + e^{-it}}{2}$, $\qquad \sin t = \dfrac{e^{it} - e^{-it}}{2i}$.
18. $\cos(3t) = 4\cos^3 t - 3\cos t$, $\qquad \sin(3t) = 3\sin t - 4\sin^3(t)$
    (parties réelles et imaginaires de $e^{3it} = (e^{it})^3$).
19. $\text{ch} t = \dfrac{e^t + e^{-t}}{2}$, $\qquad \text{sh} t = \dfrac{e^t - e^{-t}}{2}$.
20. $\text{ch}^2 t - \text{sh}^2 t = 1$, $\qquad \text{ch}(2t) = \text{ch}^2 t + \text{sh}^2 t$, $\qquad \text{sh}(2t) = 2\text{ch} t\, \text{sh} t$.
21. $4\text{ch}^3 t = \text{ch}3t + 3cht$, $\qquad 4\text{sh}^3 t = \text{sh}3t - 3\text{sh}t$.

11 Annexe : quelques intégrales

1. $\displaystyle\int_0^\infty x^n e^{-x}\, dx = n! = \Gamma(n+1)$.
2. $\displaystyle\int_{-\infty}^\infty e^{-ax^2}\, dx = \frac{\sqrt{\pi}}{a}$.
3. $\displaystyle\int_0^1 x^m (1-x)^n\, dx = \frac{m!n!}{(m+n+1)!}$.
4. $\displaystyle\int_0^{\frac{\pi}{2}} \sin^n x\, dx = \int_0^{\frac{\pi}{2}} \cos^n x\, dx = \begin{cases} \dfrac{\pi}{2} \dfrac{1.3....(n-1)}{2.4....n} & \text{si } n \text{ pair}, \\ \dfrac{2.4....(n-1)}{3.5...n} & \text{si } n \text{ impair}, \end{cases}$
5. $\displaystyle\int_0^\infty \frac{\sin^2 x}{x^2}\, dx = \int_0^\infty \frac{\sin x}{x}\, dx = \frac{\pi}{2}$.

12 Annexe : quelques dérivées (et primitives) usuelles

Voici une liste de dérivées (ce qui réciproquement donne également les primitives) :

1. si f est une fonction constante, elle est dérivable de dérivée nulle.
2. $(x^n)' = nx^{n-1}$, pour tout $n = 1, 2, \dots$ (preuve en considérant la dérivée du produit xx^{n-1} et par récurrence).
3. $(x^\alpha)' = \alpha x^{\alpha-1}$, pour tout $\alpha \neq 0$, $x > 0$, $\alpha \in \mathbb{R}$ (preuve par la composée $\exp(\alpha \log x)$).
4. $(e^{\alpha x})' = \alpha e^x$ (par définition de la fonction exp pour $\alpha \in \mathbb{R}$).
5. $(\ln \alpha x)' = \frac{\alpha}{x}$ pour $x > 0$ et $\alpha > 0$ (par dérivation de $(\exp \circ \ln)(y) = y$).
6. $(x\ln x - x)' = \ln x$ pour $x > 0$ (par dérivation de $(\ln \circ \exp)(y) = y$).
7. $(\sin x)' = \cos x$ et $(\cos x)' = -\sin x$ (preuve par $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$ et $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$).
8. $(\mathrm{sh}x)' = \mathrm{ch}x$ et $(\mathrm{ch}x)' = \mathrm{sh}x$ (preuve par $\mathrm{sh}x = \frac{1}{2}(e^x - e^{-x})$ et $\mathrm{ch}x = \frac{1}{2}(e^x + e^{-x})$).
9. $(\tan x)' = \dfrac{1}{\cos^2 x} = 1 + \tan^2 x$, pour $x \in]\frac{-\pi}{2}, \frac{\pi}{2}[$ (par dérivation du quotient $\tan x = \frac{\sin x}{\cos x}$).
10. $(\mathrm{cotan}x)' = -\dfrac{1}{\sin^2 x}$, pour $x \in]0, \pi[$ (par dérivation du quotient $\mathrm{cotan}x = \frac{\cos x}{\sin x}$).
11. $(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$, pour $x \in]-1, 1[$ (puisque $\sin(\arcsin\theta) = \theta$).
12. $(\arccos x)' = \dfrac{-1}{\sqrt{1-x^2}}$, pour $x \in]-1, 1[$ (puisque $\cos(\arccos\theta) = \theta$).
13. $(\arctan x)' = \dfrac{1}{1+x^2}$ (puisque $\tan(\arctan) = I$).
14. $(\arctan \dfrac{x}{\alpha})' = \dfrac{\alpha}{x^2+\alpha^2}$, pour $\alpha \neq 0$.
15. $(\ln(|\dfrac{1}{\cos x} + \tan x|))' = \dfrac{1}{\cos x}$.
16. $(\ln(|\dfrac{1}{\sin x} - \mathrm{cotan}x|))' = \dfrac{1}{\sin x} = -(\ln(|\dfrac{1}{\sin x} + \mathrm{cotan}x|))'$.
17. $(\dfrac{\tan x}{\cos x} + \ln|\dfrac{1}{\cos x} + \tan x|)' = \dfrac{2}{\cos^3 x}$.
18. $\ln(\dfrac{|x+a|}{|x-a|})' = \dfrac{2a}{a^2-x^2} = 2\tanh^{-1}(\dfrac{x}{a})'$, pour $x \neq a$.
19. $\ln(|x + \sqrt{x^2+a}|)' = \dfrac{1}{\sqrt{x^2+a}}$.
20. $\big(x\sqrt{x^2+a}\big)' = \dfrac{2x^2+a}{\sqrt{x^2+a}}$.
21. D'où : $\big(x\sqrt{x^2+a}\big)' + a\ln(|x + \sqrt{x^2+a}|)' = 2\sqrt{x^2+a}$.
22. si $b^2 - 4ac > 0$: $\ln(\dfrac{|2ax + b - \sqrt{b^2-4ac}|}{|2ax + b + \sqrt{b^2-4ac}|})' = \dfrac{\sqrt{b^2-4ac}}{ax^2+bx+c}$.
23. si $b^2 - 4ac < 0$: $\arctan(\dfrac{2ax+b}{\sqrt{4ac-b^2}})' = \dfrac{1}{2}\dfrac{\sqrt{4ac-b^2}}{ax^2+bx+c}$.
24. si $b^2 - 4ac = 0$: $(\dfrac{2}{2ax+b})' = \dfrac{1}{ax^2+bx+c}$.
25. si $a > 0$: $\ln(|2ax + b + 2\sqrt{a}\sqrt{ax^2+bx+c}|)' = \dfrac{\sqrt{a}}{\sqrt{ax^2+bx+c}}$.
26. $\mathrm{argsh}'(x) = \frac{1}{\sqrt{1+x^2}}$, et $\mathrm{argch}'(x) = \frac{1}{\sqrt{x^2-1}}$.
27. $(\mathrm{argsh}(x) + x\sqrt{1+x^2})' = 2\sqrt{1+x^2}$.
28. Une primitive de $\tan^2 x$ est $\tan x - x + c$
    (car $\tan^2 x = \dfrac{1}{\cos^2 x} - 1$).
29. Une primitive de $\dfrac{1}{1+\sin x}$ est $\tan x - \dfrac{1}{\cos x} + c$
    (on multiplie par $1 - \sin x$, i.e. on calcule la primitive de $\dfrac{1-\sin x}{1-\sin x^2} = \dfrac{1-\sin x}{\cos^2 x}$).

30. Primitives de $\dfrac{1}{1+a\sin x+b\cos x+c\tan x}$: faire le changement de variable $u=\tan\dfrac{x}{2}$, qui s'exprime également $\sin x=\dfrac{2u}{1+u^2}$, $\cos x=\dfrac{1-u^2}{1+u^2}$, et qui donne $dx=\dfrac{2\,du}{1+u^2}$. Et on obtient la primitive d'une fraction rationnelle.
31. N.B. : les fonctions $\cos(x^2)$ et $\dfrac{1}{\sqrt{b-\sin^2 x}}$ (par exemple) n'ont pas de primitive élémentaire (la dernière primitive est une "intégrale elliptique"). On calcule les intégrales correspondantes de manière approchée (intégration numérique).

13 Annexe : racines des polynômes de degré 3 et 4

Pour les racinies d'un polynôme de degré 2 : $ax^2+bx+c=0$, on rappelle que $x=\frac{-b\pm\sqrt{\Delta}}{2a}$ où $\Delta=b^2-4ac$ (racines éventuellement complexes).

On vérifie immédiatement que $(x-\frac{-b+\sqrt{b^2-4ac}}{2a})(x-\frac{-b-\sqrt{b^2-4ac}}{2a})=x^2+\frac{b}{a}x+\frac{c}{a}$, puisque c'est $((x+\frac{b}{2a})-\frac{\sqrt{b^2-4ac}}{2a})((x+\frac{b}{2a})+\frac{\sqrt{b^2-4ac}}{2a})$.

Pour les racines d'un polynôme de degré 3 (Girolamo Cardano (1501-1576)) :

$$ax^3+bx^2+cx+d=0,$$

on cherche une racine r (puis on décompose sous la forme $a(x-r)(x^2+\alpha x+\beta)=0$). On commence par mettre l'équation sous la forme $y^3+Ay=B$. Il suffit pour cela de poser $x=y-\frac{3b}{a}$, ce qui donne :

$$y^3+Ay=B,\qquad A=\frac{1}{a}(c-\frac{b^2}{3a}),\quad B=-\frac{1}{a}(d+\frac{2b^3}{27a^2}-\frac{bc}{3a}).$$

Puis on cherche y sous la forme $y=s-t$ où s et t sont solution du système :

$$\begin{cases} 3st=A,\\ s^3-t^3=B.\end{cases}$$

On vérifie immédiatement qu'un tel $y=s-t$ est solution de $(s-t)^3+A(s-t)=B$, avec A et B donnés ci-dessus.

Substituant $s=\frac{A}{3t}$ dans la 2ème équation, on obtient : $t^6+Bt^3-\frac{A^3}{27}=0$, soit $T^2+BT-\frac{A^3}{27}=0$ où $T=t^3$. D'où $t^3=T=\frac{-B\pm\sqrt{B^2-\frac{4A^3}{27}}}{2}$. On garde par exemple $t=(\frac{-B\pm\sqrt{B^2-\frac{4A^3}{27}}}{2})^{\frac{1}{3}}$, d'où s, d'où $y=s-t$, d'où x.

Le résultat final est :

$$x=(q+\sqrt{q^2+(r-p)^2})^{\frac{1}{3}}+(q+\sqrt{q^2-(r-p)^2})^{\frac{1}{3}}+p,$$

où $p=\frac{-b}{3a}$, $q=p^3+\frac{bc-3ad}{6a^2}$ et $r=\frac{c}{3a}$. (Il est bien sûr hors de question d'apprendre ces formules par coeur.)

Références :
http ://www.sosmath.com/algebra/factor/fac11/fac11.html
http ://atlas.math.vanderbilt.edu/~schectex/courses/cubic/

Pour les formules donnant une racine d'un polynôme de degré 4, voir par exemple :
http ://www.sosmath.com/algebra/factor/fac12/fac12.html

Références

[1] Arnaudiès J.M., Fraysse H. : Cours de mathématiques – 2, Analyse, classes préparatoires. Dunod 1988.

[2] Avez A. : *Calcul différentiel*. Masson, 1983.

[3] Ayres F., Mendelson E. : *Calcul différentiel et intégral, cours et problèmes.* Série Schaum, McGraw-Hill, 1993.

[4] Cartan H. : *Cours de calcul différentiel*. Hermann méthodes, 1977.

[5] Courant R., John F. : Introduction to Calculus and Analysis. Volume 1, 2. Springer-Verlag, 1989.

[6] Couzeix M., Mignot A. : *Analyse numérique des équations différentielles*. Masson, 1992.

[7] Geffroy J. : *Équations différentielles*. Puf, 1983.

[8] Germain P. : *Mécaniques des milieus continus*, cours de l'Ecole Polytechique 1982. Edition de l'Ecole Polytechnique, 1982.

[9] Kartachev A., Rojdestvenski B. : Analyse mathématique. Éditions Mir, 1988.

[10] Lang S. : Calculus of Several Variables. Springer-Verlag, Third Edition,1994.

[11] O'Neil P. : *Advanced Engineering Mathematics*. PWS, 1995.

[12] Reinhard : *équations différentielles*. Dunod.

[13] Schatzman M. : *Analyse numérique, cours et exercices pour la licence*. InterEdition, 1991.

[14] Strang G. : Calculus. Wellesley Cambridge Press 1991.

[15] http ://www.les-mathematiques.net