

Ahmed Lesfari

L3
M1
M2

Fonctions de plusieurs variables réelles et complexes

Cours et exercices corrigés



Chapitre 1

Fonctions de plusieurs variables réelles

1.1 Calcul différentiel

Dans ce chapitre, on s'intéresse à l'étude des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$. Les résultats que l'on obtiendra s'appliquent aux espaces vectoriels normés E et F de dimension finie n et p respectivement. D'ailleurs, on se ramène dans ce cas par choix de base à des applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$. En outre certains de ces résultats s'appliquent aux espaces de dimension infinie, notamment les espaces de Banach.

1.1.1 Généralités

Soit Ω un sous-ensemble non-vide de $\mathbb{R}^n$ et soit f une fonction de Ω dans $\mathbb{R}^p$ qui à $x = (x_1, \dots, x_n) \in \Omega$, on associe l'élément de $\mathbb{R}^p$,

$$f(x) = f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_p(x_1, \dots, x_n)),$$

où $f_1, \dots, f_p$ sont des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$. On parle de fonctions de $n \geq 2$ variables à valeurs réelles ($\mathbb{R}$) ou vectorielles ($\mathbb{R}^p, p \geq 2$).

Définition 1.1 On appelle ensemble de définition de la fonction f , noté D_f , l'ensemble défini par

$$D_f = \{x \in \Omega : f(x) \text{ existe}\},$$

et graphe de la fonction f l'ensemble $\{(x, f(x)), x \in \Omega\} \subset \Omega \times \mathbb{R}^p \times \mathbb{R}^{n+p}$.

Exemple 1.2 *Le domaine de définition de la fonction*

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad (x, y) \longmapsto \ln(1 - x^2 - y^2),$$

est le disque de centre $(0, 0)$ et de rayon 1, car pour que cette fonction soit définie il faut et il suffit que $1 - x^2 - y^2 > 0$, c.-à-d., $x^2 + y^2 < 1$.

1.1.2 Limite-continuité-différentiabilité

Soient une application f d'un ouvert Ω de $\mathbb{R}^n$ dans $\mathbb{R}^p$, $a = (a_1, \dots, a_n)$ un point de Ω ou sur sa frontière et $l = (l_1, \dots, l_p) \in \mathbb{R}^p$.

Définition 1.3 *On dit que f admet pour limite l lorsque x tend vers a dans Ω si,*

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall x \in \Omega, \|x - a\| \leq \delta \implies \|f(x) - l\| \leq \varepsilon.$$

On écrit $l = \lim_{x \rightarrow a} f(x)$, et on note que si la limite existe, elle est unique.

Dans cette définition, $\|x - a\|$ désigne la norme de $x - a$ pour une norme quelconque sur $\mathbb{R}^n$ et $\|f(x) - l\|$ désigne la norme de $f(x) - l$ pour une norme quelconque sur $\mathbb{R}^p$ (rappelons que comme les espaces $\mathbb{R}^n$ et $\mathbb{R}^p$ sont de dimensions finies, alors toutes les normes sont équivalentes sur ces espaces et sur tout espace vectoriel normé de dimension finie. Dans certains problèmes, on précisera en cas de besoin de calcul explicite telle ou telle norme à utiliser, par exemple $\|\cdot\|_1$, $\|\cdot\|_2$ ou $\|\cdot\|_\infty$).

Une fonction à valeurs dans $\mathbb{R}^p$, c'est aussi p fonctions à valeurs réelles. Comme on peut toujours calculer les limites composante par composante, la définition ci-dessus revient à dire que f admet pour limite l lorsque x tend vers a dans Ω si chaque fonction coordonnée $f_k : \Omega \longrightarrow \mathbb{R}$, admet l_k pour limite.

Exemple 1.4 *Montrons en utilisant la définition que :*

$$\lim_{(x,y) \rightarrow (0,0)} x^2 \sin \frac{1}{x^2 + y^2} = 0.$$

En effet, posons $u = x^2 + y^2$. Comme $\lim_{u \rightarrow 0} u \sin \frac{1}{u} = 0$, c.-à-d.,

$$\forall \varepsilon > 0, \exists \eta > 0 : \forall u \in \mathbb{R}^*, |u| \leq \eta \implies \left| u \sin \frac{1}{u} \right| \leq \varepsilon.$$

Il suffit de choisir $\delta \leq \sqrt{\eta}$. D'où,

$$\|(x, y)\|_2 \leq \delta \implies |x^2 + y^2| \leq \eta \implies \left| (x^2 + y^2) \sin \frac{1}{x^2 + y^2} \right| \leq \varepsilon,$$

et le résultat découle du fait que l'on a

$$\left| x^2 \sin \frac{1}{x^2 + y^2} \right| \leq \left| (x^2 + y^2) \sin \frac{1}{x^2 + y^2} \right| \leq \varepsilon.$$

Exemple 1.5 Montrons que : $\lim_{\substack{(x,y) \rightarrow (0,0) \\ (x,y) \neq (0,0)}} f(x,y)$ n'existe pas où

$$f(x,y) = \frac{\sin(x+y)}{|x| + |y|}.$$

En effet, sur l'ensemble $\{(x,0) : x > 0\}$, on a

$$\lim_{(x,y) \rightarrow (0,0)} f(x,y) = \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{\sin x}{x} = 1,$$

et sur l'ensemble $\{(x,0) : x < 0\}$, on a

$$\lim_{(x,y) \rightarrow (0,0)} f(x,y) = - \lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{\sin x}{x} = -1.$$

Définition 1.6 On dit que la fonction f est continue en $a \in \Omega$ si elle définie en a et si $\lim_{x \rightarrow a} f(x) = f(a)$, autrement dit, si

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall x \in \Omega, \|x - a\| \leq \delta \implies \|f(x) - f(a)\| \leq \varepsilon.$$

On dit que f est continue sur un sous-ensemble Ω de D_f si elle est continue en tout point de Ω .

Remarque 1.7 Souvent lors de la résolution des exercices, pour prouver la continuité d'une fonction de plusieurs variables en un point a de son domaine, on majore $|f(x) - f(a)|$ par une expression mieux connue, tendant vers 0 avec $\|x - a\|$. En revanche, pour prouver la discontinuité d'une fonction de plusieurs variables en un point a de son domaine, on prouve que pour x tendant vers a le long d'un chemin particulier, $f(x)$ ne tend pas vers $f(a)$. Quelques majorations utiles lors de la résolution des exercices :

$$|x| \leq \sqrt{x^2 + y^2}, \quad |y| \leq \sqrt{x^2 + y^2},$$

$$|xy| \leq \frac{1}{2}(x^2 + y^2), \quad |xy| \leq (|x| + |y|)(|x| + |y|).$$

Soit $a = (a_1, \dots, a_n) \in \Omega \subset \mathbb{R}^n$, $f : \Omega \longrightarrow \mathbb{R}^p$. Considérons l'application

$$f^i : \Omega_i \longrightarrow \mathbb{R}^p, \quad x_i \longmapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n),$$

la $i^{\text{ème}}$ fonction partielle de f au point a où

$$\Omega_i = \{x_i \in \mathbb{R} : (a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n) \in \Omega\}, \quad 1 \leq i \leq n.$$

Proposition 1.8 *Si f est continue en a , alors chaque f^i est continue en a_i .*

Démonstration : Comme f est continue, alors par définition, on a

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall x \in \Omega, \|x - a\| \leq \delta \implies \|f(x) - f(a)\| \leq \varepsilon.$$

Dès lors, $x = (a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ et $|x_i - a_i| \leq \delta$ impliquent $\|x - a\| \leq \delta$ et par conséquent,

$$\begin{aligned} & \|f^i(x_i) - f^i(a_i)\| = \\ & \|f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n)\| \leq \varepsilon, \end{aligned}$$

et le résultat en découle. ■

Exemple 1.9 *Montrons que la réciproque de la proposition précédente est fausse en général. Soit f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par*

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

(Ici, on a $x_1 = x$, $x_2 = y$, $a_1 = a_2 = 0$). La 1^{re} fonction partielle de f au point $a = (0, 0)$ est $x \mapsto f(x, 0) = 0$ et il en est de même de la 2^{ème} fonction partielle $y \mapsto f(0, y) = 0$. Ces fonctions sont évidemment continues, et pourtant la fonction f n'est pas continue en $a = (0, 0)$. En effet, posons $y = kx$, $k \neq 0$, d'où

$$f(x, kx) = \begin{cases} \frac{k}{1 + k^2}, & \text{si } x \neq 0 \\ 0, & \text{si } x = 0 \end{cases}$$

ce qui montre que la fonction f n'est pas continue en 0, donc elle n'est pas continue en $a = (0, 0)$. (Par exemple pour $k = 1$, on a pour tout $x \in \mathbb{R}^*$, $f(x, x) = \frac{1}{2}$, et par hypothèse $f(0, 0) = 0$). Plus précisément, la fonction f n'a pas de limite quand $(x, y) \rightarrow (0, 0)$. Comme nous l'avons déjà signalé, pour que la limite de f existe, elle ne doit pas dépendre de la manière dont $(x, y) \rightarrow (0, 0)$, ce qui n'est pas le cas ici ou autrement dit cela dépend de la limite s'il existe de $\frac{y}{x}$.

Définition 1.10 Soit $u \in \mathbb{R}^n$ avec $u \neq 0$. La dérivée directionnelle de f en a dans la direction u est la dérivée en 0 de la fonction d'une seule variable réelle $t \mapsto f(a + tu)$, c.-à-d., la limite si elle existe,

$$\lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{f(a + tu) - f(a)}{t},$$

et on la note $\frac{\partial f}{\partial u}(a)$ ou $\partial_u f(a)$ ou encore $Df(a, u)$.

Soit

$$a_1, \dots, a_n \in \mathbb{R}, \quad a = a_1 e_1 + \dots + a_n e_n \in \Omega \subset \mathbb{R}^n,$$

où $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$. La dérivée partielle de f au point a par rapport à la $i^{\text{ème}}$ variable, si elle existe, s'obtient en prenant pour u le $i^{\text{ème}}$ vecteur $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ de la base canonique $(e_1, \dots, e_n)$. On la note $\frac{\partial f}{\partial x_i}(a)$ ou $f'_i(a)$. D'après la définition des dérivées directionnelles, on a

Définition 1.11 La $i^{\text{ème}}$ dérivée partielle première de f en a , si elle existe, est définie par

$$\begin{aligned} \frac{\partial f}{\partial x_i}(a) &= \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a_1, \dots, a_i, \dots, a_n)}{t}, \\ &= \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{f(a + t e_i) - f(a)}{t}. \end{aligned}$$

Au fond, on calcule la dérivée au point a_i de la $i^{\text{ème}}$ fonction partielle de f au point a : $x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$. En d'autres termes, pour calculer la dérivée partielle $\frac{\partial f}{\partial x_i}(a)$, on fixe toutes les variables, sauf la $i^{\text{ème}}$, et on dérive la fonction d'une variable ainsi obtenue.

Remarque 1.12 L'existence des dérivées partielles en un point n'assure pas la continuité de la fonction en ce point, ni l'existence des dérivées directionnelles en ce point. De même, l'existence de toutes les dérivées directionnelles en un point n'implique pas la continuité de la fonction en ce point.

Exemple 1.13 On reprend l'exemple 1.9. On a

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{f(t, 0) - f(0, 0)}{t} = 0,$$

et on a également $\frac{\partial f}{\partial y}(0,0) = 0$. Pour $(x,y) \neq 0$, on a

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

On constate que les dérivées partielles 1^{re} existent au point $(0,0)$ mais la fonction f (voir exemple 1.9) n'est pas continue en ce point.

Définition 1.14 On dit que f est différentiable en a s'il existe une application linéaire $L : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, telle que :

$$\forall h \in \mathbb{R}^n, \quad f(a+h) = f(a) + L(h) + \varepsilon(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon(h)}{\|h\|} = 0,$$

(on écrit parfois, $\varepsilon(h) = o(\|h\|)$). De façon équivalente (où $x = a + h$), s'il existe une application linéaire $L : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, telle que :

$$f(x) = f(a) + L(x-a) + \varepsilon(x-a), \quad \lim_{\substack{x \rightarrow a \\ x \in \Omega \setminus \{a\}}} \frac{\varepsilon(x-a)}{\|x-a\|} = 0.$$

L'application L si elle existe, est unique et s'appelle la différentielle de f au point a . On la note $df(a)$ ou df_a .

Remarques 1.15 a) Dans la littérature, on utilise aussi l'écriture suivante :

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

(où $L(h) = df(a)(h)$) pour tout h de norme assez petite, ou encore

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

pour tout h de norme assez petite. Cette expression n'est autre que le développement limité de f à l'ordre 1 au voisinage de a .

b) Notons que $f(a+h)$ a un sens pour $\|h\|$ suffisamment petit car Ω est un ouvert. L'application L est définie sur un ensemble contenant la boule $B(0,r)$ où $r > 0$ de sorte que $B(a,r) \subset \Omega$, on la prolonge ensuite à $\mathbb{R}^n$ par linéarité.

c) Si les applications f et g de Ω dans $\mathbb{R}^p$ sont différentiables en a , alors pour tout $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ est différentiable en a et on a

$$d(\alpha f + \beta g)(a) = \alpha df(a) + \beta dg(a).$$

d) La fonction f à valeurs dans $\mathbb{R}^p$ est différentiable en a si et seulement si ses composantes f_j sont différentiables en a et on a $df(a) = (df_1(a), \dots, df_p(a))$.

e) L'espace vectoriel $\mathbb{R}^p$ étant de dimension finie, alors l'application L est continue et par conséquent, on peut appliquer la définition ci-dessus aux cas des applications définies sur des espaces vectoriels normés et à valeurs dans des espaces vectoriels normés; il suffit d'imposer dans la définition ci-dessus que L soit une application linéaire continue.

1.1.3 Propriétés diverses des fonctions différentiables

Proposition 1.16 *Si f est différentiable en a , alors f est continue en a .*

Démonstration : Par hypothèse f est différentiable en a , c.-à-d.,

$$\forall h \in \mathbb{R}^n, \quad f(a+h) = f(a) + L(h) + \varepsilon(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon(h)}{\|h\|} = 0.$$

On a $L(h) = df(a)(h)$. Comme L est linéaire, elle est continue (car $\mathbb{R}^n$ est de dimension finie) et donc $\lim_{h \rightarrow 0} L(h) = \lim_{h \rightarrow 0} df(a)(h) = 0$. D'autre part,

$$\lim_{h \rightarrow 0} \varepsilon(h) = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \|h\| \cdot \frac{\varepsilon(h)}{\|h\|} = \lim_{h \rightarrow 0} \|h\| \cdot \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon(h)}{\|h\|} = 0,$$

donc

$$\lim_{h \rightarrow 0} f(a+h) = \lim_{h \rightarrow 0} (f(a) + L(h) + \varepsilon(h)) = f(a).$$

Par conséquent, la fonction f est continue en a . ■

Remarque 1.17 *Si f est différentiable en a , alors f est dérivable suivant tout vecteur u de $\mathbb{R}^n$ et $\frac{\partial f}{\partial u}(a) = df(a)u$, pour chaque u de $\mathbb{R}^n$. Cette formule implique que si L est l'application linéaire intervenant dans la définition de la différentiabilité de f en a , alors $L(u) = \frac{\partial f}{\partial u}(a)$ pour tout $u \in \mathbb{R}^n$, d'où l'unicité de L .*

Proposition 1.18 *Si f est différentiable au point a , alors les dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ existent et on a*

$$df = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n.$$

Démonstration : Soient $h \in \mathbb{R}$ et $(e_1, \dots, e_n)$ une base canonique de $\mathbb{R}^n$. Par hypothèse, la fonction f est différentiable au point a , d'où

$$f(a + he_i) = f(a) + df(a)(he_i) + \varepsilon(he_i), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon(he_i)}{\|he_i\|} = 0.$$

Donc les dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ existent et on a $\frac{\partial f}{\partial x_i}(a) = df(a)(e_i)$. Comme $h = h_1 e_1 + \cdots + h_n e_n$, alors

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

Autrement dit, on a

$$\forall(h_1, \dots, h_n) \in \mathbb{R}^n, \quad df(a)(h_1, \dots, h_n) = \sum_{i=1}^n h_i df(a)(e_i) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a).$$

En général, on utilise (par abus de notations) dx_i en tant qu'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$ telle que : pour tout $h \in \mathbb{R}^n$, $dx_i(h) = h_i$. Dès lors, avec cette notation, on écrit

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

d'où

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

(Notons que $(dx_1, \dots, dx_n)$ n'est autre que la base duale de $(e_1, \dots, e_n)$). ■

Définition 1.19 On dit que f est continûment différentiable ou de classe $\mathcal{C}^1$ sur Ω lorsque les dérivées partielles $\frac{\partial f}{\partial x_i}$ de f existent et sont continues sur Ω .

Exemple 1.20 Montrons que la fonction définie par

$$f(x, y) = \begin{cases} xy \frac{x^4 - y^4}{x^4 + y^4}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

est de classe $\mathcal{C}^1$. La fonction f est continue et différentiable sur $\mathbb{R}^2$. En effet, f est continue et différentiable sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Au point $(0, 0)$, la fonction f est continue car

$$|f(x, y)| \leq |xy|, \quad \lim_{(x, y) \rightarrow (0, 0)} f(x, y) = 0.$$

En outre, elle est différentiable en ce point car

$$\frac{\partial f}{\partial x}(x, y) = \frac{y(x^8 - y^8 + 8x^4 y^4)}{(x^4 + y^4)^2}, \quad (x, y) \neq (0, 0)$$

$$\frac{\partial f}{\partial y}(x, y) = \frac{x(x^8 - y^8 - 8x^4 y^4)}{(x^4 + y^4)^2}, \quad (x, y) \neq (0, 0)$$

et

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

Les deux dérivées partielles $\frac{\partial f}{\partial x}(x, y)$ et $\frac{\partial f}{\partial y}(x, y)$ sont continues en $(0, 0)$ car

$$\left| \frac{\partial f}{\partial x}(x, y) \right| = \left| \frac{y(x^8 - y^8 + 8x^4y^4)}{(x^4 + y^4)^2} \right| \leq \left| \frac{y \cdot 4((x^4 + y^4)^2)}{(x^4 + y^4)^2} \right| = 4|y|,$$

$$\left| \frac{\partial f}{\partial y}(x, y) \right| = \left| \frac{x(x^8 - y^8 - 8x^4y^4)}{(x^4 + y^4)^2} \right| \leq \left| \frac{x \cdot 4((x^4 + y^4)^2)}{(x^4 + y^4)^2} \right| = 4|x|,$$

d'où

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\partial f}{\partial x}(x, y) = 0 = \frac{\partial f}{\partial x}(0, 0), \quad \lim_{(x,y) \rightarrow (0,0)} \frac{\partial f}{\partial y}(x, y) = 0 = \frac{\partial f}{\partial y}(0, 0).$$

Dès lors, la fonction f est de classe $\mathcal{C}^1$.

Proposition 1.21 Si (i) l'une des dérivées partielles $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$, de la fonction f existe au point $a = (a_1, \dots, a_n)$ et si (ii) les $n-1$ autres dérivées partielles existent dans un voisinage de a et sont continues en a , alors la fonction f est différentiable en a .

Démonstration : Sans perte de généralité, on suppose que c'est la dérivée partielle $\frac{\partial f}{\partial x_n}$ qui existe au point $a = (a_1, \dots, a_n)$ et soit $B(a, r)$, $r > 0$, une boule ouverte centrée en a , contenue dans Ω (voisinage de a) et telle que les $n-1$ autres dérivées partielles existent sur $B(a, r)$ et sont continues en a . Nous allons montrer que sous les hypothèses de l'énoncé,

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - h_1 \frac{\partial f}{\partial x_1}(a) - \dots - h_n \frac{\partial f}{\partial x_n}(a)}{\|h\|} = 0.$$

On écrit $f(a+h) - f(a)$ sous la forme,

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2 + h_2, \dots, a_n + h_n) - f(a_1, a_2, \dots, a_n), \\ &= f(a_1 + h_1, a_2 + h_2, \dots, a_n + h_n) - f(a_1, a_2 + h_2, \dots, a_n + h_n) \\ &\quad + f(a_1, a_2 + h_2, \dots, a_n + h_n) - f(a_1, a_2, a_3 + h_3, \dots, a_n + h_n) \\ &\quad + \dots + f(a_1, \dots, a_{n-2}, a_{n-1} + h_{n-1}, a_n + h_n) - f(a_1, \dots, a_{n-1}, a_n + h_n) \\ &\quad + f(a_1, \dots, a_{n-1}, a_n + h_n) - f(a_1, \dots, a_n). \end{aligned}$$

D'après le théorème des accroissements finis, il existe $\theta_k \in]0, 1[$, ($1 \leq k \leq n-1$) tels que :

$$\begin{aligned} & f(a_1 + h_1, a_2 + h_2, \dots, a_n + h_n) - f(a_1, a_2 + h_2, \dots, a_n + h_n) \\ &= h_1 \frac{\partial f}{\partial x_1}(a_1 + \theta_1 h_1, a_2 + h_2, \dots, a_n + h_n), \\ & f(a_1, a_2 + h_2, \dots, a_n + h_n) - f(a_1, a_2, a_3 + h_3, \dots, a_n + h_n) \\ &= h_2 \frac{\partial f}{\partial x_2}(a_1, a_2 + \theta_2 h_2, a_3 + h_3, \dots, a_n + h_n), \end{aligned}$$

et ainsi de suite, jusqu'à l'ordre $n-1$,

$$\begin{aligned} & f(a_1, \dots, a_{n-2}, a_{n-1} + h_{n-1}, a_n + h_n) - f(a_1, \dots, a_{n-1}, a_n + h_n) \\ &= h_{n-1} \frac{\partial f}{\partial x_{n-1}}(a_1, \dots, a_{n-2}, a_{n-1} + \theta_{n-1} h_{n-1}, a_n + h_n). \end{aligned}$$

On obtient alors pour $h_n \neq 0$,

$$\begin{aligned} f(a + h) - f(a) - \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) = \\ & h_1 \left(\frac{\partial f}{\partial x_1}(a_1 + \theta_1 h_1, a_2 + h_2, \dots, a_n + h_n) - \frac{\partial f}{\partial x_1}(a_1, \dots, a_n) \right) \\ & + h_2 \left(\frac{\partial f}{\partial x_2}(a_1, a_2 + \theta_2 h_2, a_3 + h_3, \dots, a_n + h_n) - \frac{\partial f}{\partial x_2}(a_1, \dots, a_n) \right) \\ & \vdots \\ & + h_{n-1} \left(\frac{\partial f}{\partial x_{n-1}}(a_1, \dots, a_{n-2}, a_{n-1} + \theta_{n-1} h_{n-1}, a_n + h_n) - \frac{\partial f}{\partial x_{n-1}}(a_1, \dots, a_n) \right) \\ & + h_n \left(\frac{f(a_1, \dots, a_{n-1}, a_n + h_n) - f(a_1, \dots, a_n)}{h_n} - \frac{\partial f}{\partial x_n}(a_1, \dots, a_n) \right). \end{aligned}$$

Les $n-1$ termes entre parenthèses tendent vers zéro du fait de la continuité de $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_{n-1}}$. Le dernier terme entre parenthèse tend aussi vers zéro en vertu de la définition de $\frac{\partial f}{\partial x_n}(a_1, \dots, a_n)$, ce qui montre que la fonction f est différentiable en a . ■

Remarque 1.22 On remarque évidemment au niveau de la preuve ci-dessus (ainsi que celle proposée dans [21]) que dans le cas d'une fonction $f(x, y)$ à deux variables la condition suffisante de différentiabilité de cette fonction consiste à

utiliser la continuité d'une des dérivées partielles mais pas les deux ! La continuité de toutes les dérivées partielles intervient seulement pour montrer que f est de classe $\mathcal{C}^1$. Par ailleurs (bien que cela n'a pas beaucoup d'importance étant donné que la plus part des exercices que l'on rencontre à ce niveau concernent des fonctions à deux variables !), on voit aussi au niveau de la solution que cet énoncé se généralise facilement pour le cas d'une fonction à plusieurs variables mais en affaiblissant la condition de continuité des dérivées partielles en demandant seulement la continuité par rapport à certaines variables.

Nous allons reprendre explicitement la preuve ci-dessus dans le cas particulier d'une fonction $f(x, y)$ à deux variables, en dimension deux le résultat devient :

Proposition 1.23 *si l'une des dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ de la fonction f existe au point $a = (a_1, a_2)$ et si l'autre dérivée partielle existe dans un voisinage de a et que de plus cette dernière est continue en a , alors la fonction f est différentiable en a .*

Démonstration : Sans perte de généralité, on suppose que c'est la dérivée partielle $\frac{\partial f}{\partial y}$ qui existe au point $a = (a_1, a_2)$ et soit $B(a, r)$, $r > 0$, une boule ouverte centrée en a , contenue dans Ω (voisinage de a) et telle que l'autre dérivée partielle $\frac{\partial f}{\partial x}$ existe sur $B(a, r)$ et est continue en a . Nous allons montrer que sous les hypothèses de l'énoncé,

$$\lim_{h \rightarrow 0} \frac{f(a + h) - f(a) - h_1 \frac{\partial f}{\partial x}(a) - h_2 \frac{\partial f}{\partial y}(a)}{\|h\|} = 0.$$

On a

$$\begin{aligned} f(a + h) - f(a) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) \\ &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) + f(a_1, a_2 + h_2) - f(a_1, a_2). \end{aligned}$$

En utilisant le théorème des accroissements finis, alors on peut trouver θ_1, θ_2 appartenant à $]0, 1[$, tels que :

$$\begin{aligned} f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) &= h_1 \frac{\partial f}{\partial x_1}(a_1 + \theta_1 h_1, a_2 + h_2), \\ f(a_1, a_2 + h_2) - f(a_1, a_2) &= h_2 \frac{\partial f}{\partial x_2}(a_1, a_2 + \theta_2 h_2). \end{aligned}$$

Dès lors, on obtient pour $h_2 \neq 0$ l'expression suivante :

$$\begin{aligned} f(a+h) - f(a) - h_1 \frac{\partial f}{\partial x}(a) - h_2 \frac{\partial f}{\partial y}(a) \\ = h_1 \left(\frac{\partial f}{\partial x}(a_1 + \theta_1 h_1, a_2 + h_2) - \frac{\partial f}{\partial x}(a_1, a_2) \right) \\ + h_2 \left(\frac{\partial f}{\partial y}(a_1, a_2 + \theta_2 h_2) - \frac{\partial f}{\partial y}(a_1, a_2) \right). \end{aligned}$$

Le premier terme entre parenthèses tend vers zéro du fait de la continuité de $\frac{\partial f}{\partial x}$. Le dernier terme entre parenthèses tend aussi vers zéro en vertu de la définition de $\frac{\partial f}{\partial y}(a_1, a_2)$, ce qui montre que la fonction f est différentiable au point a . ■

Proposition 1.24 *Si les dérivées partielles $\frac{\partial f}{\partial x_i}$, ($1 \leq i \leq n$) de la fonction f existent dans un voisinage du point a et sont continues en a , alors f est différentiable en a . Autrement dit, si f est de classe C^1 sur Ω , alors f est différentiable sur Ω .*

Démonstration : C'est essentiellement la même preuve que pour la proposition 1.21. Plus précisément, on obtient l'expression :

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

Le second terme tend vers 0 quand $h \rightarrow 0$ et donc le résultat. ■

La proposition 1.21 donne une condition suffisante sur les dérivées partielles de la fonction f pour que celle-ci soit différentiable en a , ce qui montre (voir aussi remarque 1.22) que la proposition 1.24 n'est pas optimal, au sens que ses hypothèses sont un peu trop fortes.

Lors de la résolution des exercices sur l'étude de la différentiabilité d'une fonction f en un point a , plusieurs méthodes peuvent être utilisées. Si f n'est pas continue en a , alors f n'est pas différentiable en a . Rappelons aussi (proposition 1.23) que dans le cas d'une fonction à deux variables la condition suffisante de différentiabilité de cette fonction consiste à utiliser la continuité d'une des dérivées partielles mais pas nécessairement les deux. L'organigramme suivant résume la démarche à suivre lors de l'étude de la différentiabilité d'une fonction f en un point a .

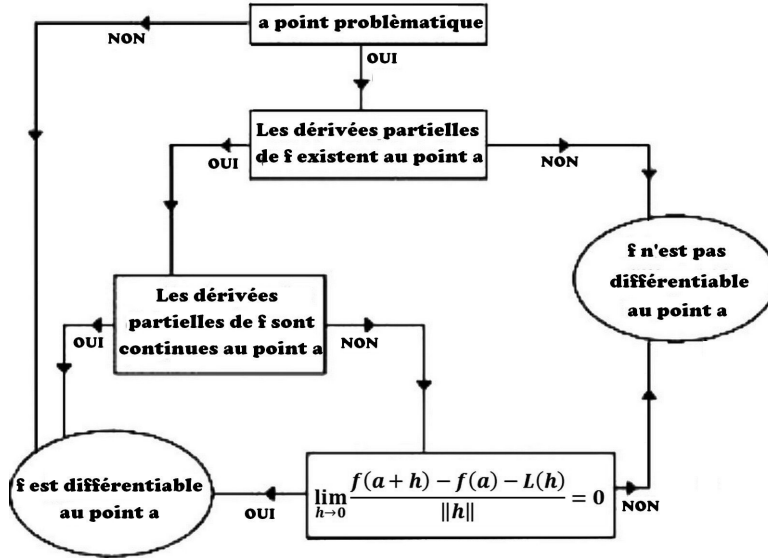


Figure 1 – Organigramme (différentiabilité)

Proposition 1.25 Soient $f, g : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$, des fonctions différentiables au point $a \in \Omega$. Alors, $f + g$ est différentiable en a et on a

$$d(f + g)(a) = df(a) + dg(a).$$

Démonstration : Par hypothèse, les fonctions f et g sont différentiables, donc on a par définition

$$\forall h \in \mathbb{R}^n, \quad f(a + h) = f(a) + L(h) + \varepsilon_1(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = 0,$$

$$\forall h \in \mathbb{R}^n, \quad g(a + h) = g(a) + M(h) + \varepsilon_2(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_2(h)}{\|h\|} = 0,$$

où L et M désignent des applications linéaires de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}^p$. Dès lors,

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{(f + g)(a + h) - (f + g)(a) - (L + M)(h)}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{f(a + h) - f(a) - L(h) + g(a + h) - g(a) - M(h)}{\|h\|} \\ &= \lim_{h \rightarrow 0} \frac{\varepsilon_1(h) + \varepsilon_2(h)}{\|h\|} = 0, \end{aligned}$$

ce qui montre que $f + g$ est différentiable en a . Plus précisément, en posant $\varepsilon_3 \equiv \varepsilon_1 + \varepsilon_2$, $N \equiv L + M$ et les théorèmes sur les limites, on obtient

$$(f + g)(a + h) = (f + g)(a) + N(h) + \varepsilon_3(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_3(h)}{\|h\|} = 0.$$

Comme la fonction $f + g$ est linéaire, on en déduit que $f + g$ est différentiable en a avec $d(f + g)(a) = df(a) + dg(a)$. ■

Proposition 1.26 *Soient $f, g : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, des fonctions différentiables au point $a \in \Omega$. Alors, fh est différentiable en a et on a*

$$d(fh)(a) = h(a)df(a) + f(a)dh(a).$$

Démonstration : Par hypothèse, les fonctions f et g sont différentiables, donc on a par définition

$$\forall h \in \mathbb{R}^n, \quad f(a + h) = f(a) + L(h) + \varepsilon_1(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = 0,$$

$$\forall h \in \mathbb{R}^n, \quad g(a + h) = f(a) + M(h) + \varepsilon_2(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_2(h)}{\|h\|} = 0,$$

où L et M désignent des applications linéaires de $\Omega \subset \mathbb{R}^n$ dans $\mathbb{R}$. On écrit,

$$f(a + h)g(a + h) - f(a)g(a) = f(a + h)[g(a + h) - g(a)] + g(a)[f(a + h) - f(a)].$$

D'où,

$$f(a + h)g(a + h) - f(a)g(a) = f(a + h)[M(h) + \varepsilon_2(h)] + g(a)[L(h) + \varepsilon_1(h)].$$

Dès lors,

$$\begin{aligned} & \frac{f(a + h)g(a + h) - f(a)g(a) - f(a)M(h) - g(a)L(h)}{\|h\|} \\ &= \frac{[f(a + h) - f(a)]M(h) + f(a + h)\varepsilon_2(h) + g(a)\varepsilon_1(h)}{\|h\|}, \end{aligned}$$

et cette expression tend vers 0 quand $h \rightarrow 0$ car f est continue en a (d'après la proposition 1.16) et par hypothèse

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_2(h)}{\|h\|} = 0,$$

d'où le résultat. ■

Proposition 1.27 Soient $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$ et $g : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, des fonctions différentiables au point $a \in \Omega$. Si $g(a) \neq 0$, alors $\frac{f}{g}$ est différentiable en a et

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

Démonstration : Comme $\frac{f}{g} = f \cdot \frac{1}{g}$, il suffit de prouver que $\frac{1}{g}$ est différentiable en a , avec

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

Le reste découle de la proposition précédente relative au produit. On vérifie sans peine qu'il existe $\delta > 0$ tel que : pour tout x appartenant au domaine de définition de g , on a $g(x) \neq 0$. En effet, par hypothèse la fonction g est différentiable en a , donc elle est continue en a , c.-à-d.,

$$\forall \varepsilon > 0, \exists \delta > 0 : \|x - a\| \leq \delta \implies |g(x) - g(a)| \leq \varepsilon.$$

Comme $g(a) \neq 0$ (par hypothèse), on choisit $\varepsilon = \frac{|g(a)|}{2}$, donc on a

$$|g(x) - g(a)| \leq \frac{|g(a)|}{2} \implies -\frac{|g(a)|}{2} \leq |g(x)| - |g(a)| \leq \frac{|g(a)|}{2}.$$

Par conséquent,

$$g(x) \geq \frac{|g(a)|}{2} > 0.$$

Par hypothèse, la fonction g est différentiable, donc on a par définition

$$\forall h \in \mathbb{R}^n, \quad g(a+h) = g(a) + L(h) + \varepsilon_1(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = 0,$$

où $L = dg(a)(h)$. On a

$$\begin{aligned} \frac{1}{g(a+h)} - \frac{1}{g(a)} - d\left(\frac{1}{g(a)}\right)(h) &= \frac{1}{g(a+h)} - \frac{1}{g(a)} + \frac{dg(a)}{g^2(a)}(h), \\ &= \frac{g(a)[g(a) - g(a+h)] + g(a+h)dg(a)(h)}{g(a+h)g^2(a)}, \\ &= \frac{g(a)[-dg(a)(h) - \varepsilon(h)] + [g(a) + dg(a)(h) + \varepsilon(h)]dg(a)(h)}{g(a+h)g^2(a)}, \\ &= \frac{-g(a)\varepsilon(h) + dg(a)(h)dg(a)(h) + \varepsilon(h)dg(a)(h)}{g(a+h)g^2(a)}, \\ &= \frac{dg(a)(h)dg(a)(h) + [dg(a)(h) - g(a)]\varepsilon(h)}{g(a+h)g^2(a)}. \end{aligned}$$

Comme dans la preuve de la proposition précédente, on divise cette expression par $\|h\|$ tout en tenant compte du fait que $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon(h)}{\|h\|} = 0$, on montre qu'elle tend vers 0 quand $h \rightarrow 0$. La preuve s'achève. ■

Nous allons maintenant démontrer la formule de dérivation des fonctions composées (appelée parfois règle de la chaîne ou de dérivation en chaîne).

Théorème 1.28 Soient $a \in \mathbb{R}^n$, Ω un voisinage de a et $f : \Omega \longrightarrow \mathbb{R}^p$. Posons $b = f(a)$ et soit Δ un voisinage de b et $g : \Delta \longrightarrow \mathbb{R}^q$. On suppose que f est différentiable en a et g est différentiable en b . Alors $g \circ f$ est différentiable en a et on a,

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

Démonstration : Par hypothèse la fonction f est différentiable en a , donc elle est continue en a (proposition 1.16). On en déduit que $f^{-1}(\Delta)$ est un voisinage en a et $g \circ f$ est définie sur ce voisinage. Par définition, on a

$$\forall h \in \mathbb{R}^n, \quad f(a + h) = f(a) + df(a)(h) + \varepsilon_1(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = 0,$$

avec $df(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^p$ une application linéaire continue, et

$$\forall k \in \mathbb{R}^p, \quad g(a + k) = g(a) + dg(b)(k) + \varepsilon_2(k), \quad \lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{\varepsilon_2(k)}{\|k\|} = 0,$$

avec $dg(b) : \mathbb{R}^p \longrightarrow \mathbb{R}^q$ une application linéaire continue. Posons $b = f(a)$ et choisissons $k = f(a + h) - f(a) = df(a)(h) + \varepsilon_1(h)$, alors k est une fonction de h qui tend vers 0 quand $h \rightarrow 0$. On compose g avec f et on développe comme suit,

$$\begin{aligned} g \circ f(a + h) &= g(f(a + h)) = g(b + k) = g(b) + dg(b)(k) + \varepsilon_2(k), \\ &= g(b) + dg(b)(df(a)(h) + \varepsilon_1(h)) + \varepsilon_2(k), \\ &= g(f(a)) + dg(b)(df(a)(h)) + dg(b)(\varepsilon_1(h)) + \varepsilon_2(k), \\ &= g \circ f(a) + dg(f(a)) \circ df(a)(h) + dg(b)(\varepsilon_1(h)) + \varepsilon_2(k). \end{aligned}$$

Notons que l'application $dg(f(a)) \circ df(a)$ est linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^q$ comme composée de deux applications linéaires. Pour démontrer la proposition, il suffit de prouver que :

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{dg(b)(\varepsilon_1(h)) + \varepsilon_2(k)}{\|h\|} = 0.$$

(*) Comme $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = 0$ (par hypothèse), alors on a (par continuité de l'application linéaire $dg(b)$),

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{dg(b)(\varepsilon_1(h))}{\|h\|} = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} dg(b) \cdot \frac{\varepsilon_1(h)}{\|h\|} = dg(b) \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = dg(b)(0) = 0.$$

(**) Reste à montrer que $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_2(k)}{\|h\|} = 0$. Soit $\varphi : \Omega \rightarrow \mathbb{R}^q$, la fonction définie par

$$\varphi(k) = \begin{cases} \frac{\varepsilon_2(k)}{\|k\|}, & \text{si } k \neq 0 \\ 0, & \text{si } k = 0 \end{cases}$$

On a évidemment pour $h \neq 0$,

$$\frac{\varepsilon_2(k)}{\|h\|} = \varphi(k) \cdot \frac{\|k\|}{\|h\|},$$

car cela résulte immédiatement de la définition de $\varphi(k)$ pour $k \neq 0$, tandis que pour $k = 0$ les deux membres de cette expression sont nuls. Montrons que $\frac{\|k\|}{\|h\|}$ est borné sur un voisinage suffisamment petit de a , tandis que $\varphi(k)$ tend vers 0 quand $k \rightarrow 0$.

- On a

$$\frac{\|k\|}{\|h\|} = \frac{\|df(a)(h) + \varepsilon_1(h)\|}{\|h\|} \leq \frac{\|df(a)(h)\|}{\|h\|} + \frac{\|\varepsilon_1(h)\|}{\|h\|}.$$

Par hypothèse f est différentiable en a et donc de la définition de ε_1 , on a $\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon_1(h)}{\|h\|} = 0$ (il existe donc $\eta > 0$ tel que par exemple $h \neq 0$ et $\|h\| \leq \eta$ implique $\frac{\|\varepsilon_1(h)\|}{\|h\|} \leq 1$, et donc $\frac{\|\varepsilon_1(h)\|}{\|h\|}$ est borné). Concernant l'autre terme, en

tenant compte de la remarque 1.29 ci-dessous appliquée à l'application linéaire $df(a)$, il existe $C > 0$ tel que : pour tout $h \in \mathbb{R}^n$, $h \neq 0$, $\|df(a)(h)\| \leq C\|h\|$ et donc $\frac{\|df(a)(h)\|}{\|h\|}$ est borné. (Sinon, on peut le montrer d'une autre façon en

notant que l'ensemble $\left\{ \frac{h}{\|h\|} : h \neq 0 \right\}$ est inclus dans une sphère de rayon unité et comme celle-ci est un compact (fermé borné), alors l'image de cet ensemble par la fonction continue $df(a)$ (elle est linéaire) est borné. Donc $\frac{\|k\|}{\|h\|}$ est borné dans un voisinage assez petit.

- Montrons que $\lim_{k \rightarrow 0} \varphi(k) = 0$. Comme $\lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{\varepsilon_2(k)}{\|k\|} = 0$, c.-à-d.,

$$\forall \varepsilon > 0, \exists \delta > 0, \forall y \text{ si } k \neq 0 \text{ et } \|k\| \leq \delta \implies \frac{\varepsilon_2(k)}{\|k\|} \leq \varepsilon.$$

Pour $\varepsilon > 0$ arbitrairement fixé, choisissons un tel $\delta > 0$. Par hypothèse, la fonction f est différentiable en a , donc continue en a (proposition 1.16) et dès lors, pour ce $\delta > 0$, on a $\exists \eta > 0, \forall h (\|h\| \leq \eta \implies \|k\| \leq \delta)$. Pour le $\varepsilon > 0$ fixé, on a donc un $\eta > 0$ tel que : $\|h\| \leq \eta \implies \|\varphi(h)\| \leq \varepsilon$ car si $k = 0$, alors $\varphi(k) = 0$, et si $k \neq 0$, on déduit du fait (ci-dessus) qu'on a

$$\|k\| \leq \delta \implies \|\varphi(k)\| = \frac{\varepsilon_2(k)}{\|k\|} \leq \varepsilon,$$

et on en déduit que : $\lim_{k \rightarrow 0} \varphi(k) = 0$. Finalement,

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{dg(b)(\varepsilon_1(h)) + \varepsilon_2(k)}{\|h\|} = 0,$$

et on obtient,

$$gof(a+h) = gof(a) + dg(f(a))odf(a)(h) + \varepsilon(h), \quad \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\varepsilon(h)}{\|h\|} = 0.$$

Par conséquent, gof est différentiable en a et on a $d(gof)(a) = dg(f(a))odf(a)$. Ceci achève la démonstration. ■

Remarque 1.29 Soit $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, une application linéaire. Alors, il existe $C > 0$ tel que : pour tout $x \in \mathbb{R}^n$, $x \neq 0$, $\|\psi(x)\| \leq C\|x\|$. En effet, soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Pour tout $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, on a $x = \sum_{i=1}^n x_i e_i$. D'où,

$$\|\psi(x)\| = \left\| \psi\left(\sum_{i=1}^n x_i e_i\right) \right\| = \sum_{i=1}^n |x_i| \cdot \|\psi(e_i)\| = \|x\| \cdot \|\psi(e_i)\| \leq C\|x\|,$$

où $C = \sup_{1 \leq i \leq n} \|\psi(e_i)\|$.

Considérons l'application $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^p$, $x \longmapsto y = f(x)$. On a

$$y_1 = f_1(x_1, \dots, x_n), \quad y_2 = f_2(x_1, \dots, x_n), \dots, y_p = f_p(x_1, \dots, x_n).$$

Définition 1.30 Soit $a \in \Omega$ et supposons que les dérivées partielles $\frac{\partial f_i}{\partial x_j}(a)$, $1 \leq i \leq p$, $1 \leq j \leq n$, existent. On appelle matrice jacobienne de f en a , la matrice d'ordre $p \times n$ suivante :

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \dots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$

Si $p = 1$, $J_f(a)$ se réduit à un vecteur de $\mathbb{R}^n$,

$$\nabla f \equiv \text{grad } f = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right), \quad (f \equiv f_1),$$

appelé gradient de f . Si $n = p$, le déterminant de la matrice $J_f(a)$ s'appelle jacobien de f en a et on écrit

$$\det J_f(a) = \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)}(a).$$

Corollaire 1.31 On a $J_{g \circ f}(a) = J_g(f(a)) \cdot J_f(a)$.

Démonstration : En effet, exprimée (matriciellement) en terme de matrice jacobienne, la proposition précédente donne le résultat. Plus précisément, d'après la proposition précédente, on a

$$d(g \circ f)(a) = dg(b) \circ df(a), \quad b = f(a).$$

Cette règle de composition correspond au produit des matrices jacobienes :

$$\begin{aligned} & \begin{pmatrix} \frac{\partial(g_1 \circ f)}{\partial x_1} & \dots & \frac{\partial(g_1 \circ f)}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial(g_q \circ f)}{\partial x_1} & \dots & \frac{\partial(g_q \circ f)}{\partial x_n} \end{pmatrix} (a) \\ &= \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \dots & \frac{\partial g_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial g_q}{\partial y_1} & \dots & \frac{\partial g_q}{\partial y_n} \end{pmatrix} (f(a)) \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1} & \dots & \frac{\partial f_p}{\partial x_n} \end{pmatrix}, \end{aligned}$$

et la preuve s'achève. ■

Corollaire 1.32 Soient $a \in \mathbb{R}^n$, Ω un voisinage de a et $f : \Omega \longrightarrow \mathbb{R}^p$. Posons $b = f(a)$ et soit Δ un voisinage de b et $g : \Delta \longrightarrow \mathbb{R}$. On suppose que f est différentiable en a et g est différentiable en b . Alors, on a (règle de dérivation en chaîne), pour tout $i \in \{1, \dots, n\}$,

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{j=1}^p \frac{\partial g}{\partial y_j}((f(a))) \cdot \frac{\partial f_j}{\partial x_i}(a).$$

Démonstration : En effet, c'est un cas particulier de la proposition précédente avec $q = 1$. D'après cette proposition, on a

$$d(g \circ f)(a) = dg(b) \circ df(a), \quad b = f(a).$$

et cette règle de composition correspond au niveau des dérivées partielles à

$$\frac{\partial(g \circ f)}{\partial x_i}(a) = \sum_{j=1}^p \frac{\partial g}{\partial y_j}((f(a))) \cdot \frac{\partial f_j}{\partial x_i}(a), \quad i \in \{1, \dots, n\},$$

d'où le résultat. ■

Supposons de plus que f est bijective avec $g = f^{-1}$. D'où $\det J_{g \circ f}(a) = 1$, et par conséquent,

$$\frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} = \frac{1}{\frac{\partial(x_1, \dots, x_n)}{\partial(f_1, \dots, f_n)}}.$$

Cette formule est très utile (notamment lors du calcul des intégrales multiples) car elle permet souvent d'éviter l'inversion explicite d'une fonction.

Proposition 1.33 (théorème des accroissements finis). Soient Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \longrightarrow \mathbb{R}$ une application, $a \in \Omega$, $h \in \mathbb{R}^n$ tels que le segment $[a, a + h] = \{a + th : 0 \leq t \leq 1\}$ soit inclus dans Ω . On suppose que f est différentiable sur Ω . Alors, il existe un réel $\theta \in]0, 1[$ tel que :

$$f(a + h) - f(a) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a + \theta h), \quad h = (h_1, \dots, h_n) \in \mathbb{R}^n$$

Démonstration : On considère l'application F définie par

$$F : [0, 1] \xrightarrow{g} [a, a + h] \subset \Omega \xrightarrow{f} \mathbb{R},$$

$$t \longmapsto a + th = (a_1 + th_1, \dots, a_n + th_n) \longmapsto f(a + th).$$

Cette fonction est différentiable sur $[0, 1]$ en vertu de la dérivée de fonction composée et

$$\forall t \in [0, 1], \quad F'(t) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a + th),$$

où $(h_1, \dots, h_n)$ sont des coordonnées de h dans la base canonique de $\mathbb{R}^n$. D'après la formule des accroissements finis,

$$\exists \theta \in]0, 1[\quad F(1) - F(0) = F'(\theta).$$

Or $F(1) = f(a + h)$ et $F(0) = f(a)$, donc

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

où $\theta \in]0, 1[$ et $h = \sum_{i=1}^n h_i e_i$. ■.

Remarque 1.34 *Le théorème des accroissements finis n'est plus vrai pour les fonctions à valeurs vectorielles (en particulier à valeurs complexes). Pour s'en convaincre, on considère deux exemples simples. La fonction $f : \mathbb{R} \rightarrow \mathbb{R}^2$, $x \mapsto f(x) = (x^2, x^3)$ est dérivable avec $f'(x) = (2x, 3x^2)$, mais $\nexists c \in]0, 1[$ tel que : $f(1) - f(0) = f'(c)(1 - 0)$, c.-à-d., $(1, 1) - (0, 0) = (2c, 3c^2)$ car explicitement $2c = 1 \Rightarrow c = 1/2$ et $3c^2 = 1 \Rightarrow 3 \cdot (1/4) \neq 1$. De même, la fonction $g : \mathbb{R} \rightarrow \mathbb{C}$, $x \mapsto g(x) = e^{ix} = \cos x + i \sin x$, est dérivable avec $g'(x) = ie^{ix}$, mais $\nexists c \in]0, 2\pi[$ tel que : $g(2\pi) - g(0) = g'(c)(2\pi - 0)$, c.-à-d., $1 - 1 = ie^{ic} \cdot 2\pi$ car $0 \neq e^{ic}$.*

Proposition 1.35 *(Inégalité des accroissements finis). Soient Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \rightarrow \mathbb{R}^p$ une application, $a \in \Omega$, $h \in \mathbb{R}^n$ tels que le segment $[a, a + h]$ soit inclus dans Ω . On suppose que f est continue sur $[a, a + h]$, différentiable sur $]a, a + h[$ et que : $\exists M > 0, \forall x \in]a, a + h[, \|df(x)\| \leq M$. Alors,*

$$\|f(a + h) - f(a)\| \leq M\|h\|.$$

(En posant $b = a + h$, cette inégalité s'écrit $\|f(b) - f(a)\| \leq M\|b - a\|$).

Démonstration : Considérons la fonction

$$g : [0, 1] \rightarrow \mathbb{R}^p, \quad t \mapsto g(t) = f(a + th).$$

Elle est continue sur $[0, 1]$, différentiable sur $]0, 1[$ et on a $g'(t) = df(a + th)(h)$, d'après la formule de dérivation de fonctions composées. Dès lors,

$$\|g'(t)\| = \|df(a + th)(h)\| \leq M\|h\|,$$

par hypothèse. Par ailleurs, il existe $t \in]0, 1[$ tel que :

$$\|g(t) - g(0)\| \leq \|g'(t)\|(1 - 0),$$

et par conséquent,

$$\|g(1) - g(0)\| = \|f(a + h) - f(a)\| \leq M\|h\|,$$

ce qui achève la démonstration. ■

Soient $a \in \mathbb{R}^n$, Ω un voisinage de a et $f : \Omega \longrightarrow \mathbb{R}^p$. Rappelons que f est de classe $\mathcal{C}^1$ au voisinage de a , si les dérivées partielles $\frac{\partial f}{\partial x_i}$ existent au point $a \in \Omega$ et sont continues en a . Si ces dérivées partielles possèdent elles-mêmes des dérivées partielles, on les appellent dérivées partielles secondes et on note

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

Si ces dérivées partielles secondes existent au voisinage de a et sont continues en a , on dit que f est de classe $\mathcal{C}^2$ au voisinage de a . On définit ainsi par récurrence les dérivées partielles $k^{\text{èmes}}$ et la notion de fonction de classe $\mathcal{C}^k$. On dit enfin qu'une fonction est de classe $\mathcal{C}^\infty$ au voisinage de a si toutes ses dérivées partielles, de tous les ordres, existent au voisinage de a et sont continues en a .

Proposition 1.36 (lemme de Schwarz). Si $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ existent dans un voisinage Ω d'un point $a = (a_1, \dots, a_n)$ et sont continues en ce point, alors

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

Démonstration : Notons d'abord que l'on peut travailler composante par composante. Ce n'est donc pas une restriction réelle de supposer f à valeurs réelles. D'autre part on peut se limiter à la considération de deux variables à la fois. Nous supposons donc $n = 2$ et nous écrirons (x, y) au lieu de (x_1, x_2) et (a, b) au lieu de (a_1, a_2) . L'existence des dérivées partielles $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ dans Ω implique celle de $\frac{\partial f}{\partial y}$ et $\frac{\partial f}{\partial x}$. Soit (h, k) assez petit pour que le rectangle de sommets opposés (a, b) et $(a + h, b + k)$ soit dans Ω . Posons

$$F(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b),$$

et montrons que $F(h, k)$ a une limite lorsque $(h, k) \rightarrow (0, 0)$. Nous allons calculer cette limite de deux manières différentes ; cela conduira au résultat. Posons $\varphi_1(t) = f(t, b+k) - f(t, b)$, d'où $F(h, k) = \varphi_1(a+h) - \varphi_1(a)$. Les fonctions $t \mapsto f(t, b+k)$ et $t \mapsto f(t, b)$ sont dérivables, donc la fonction $t \mapsto \varphi_1(t)$ est aussi dérivable et on a

$$\varphi_1'(t) = \frac{\partial f}{\partial x}(t, b+k) - \frac{\partial f}{\partial x}(t, b).$$

D'après le théorème des accroissements finis appliqué à φ_1 , il existe un réel $\theta_1 \in]0, 1[$ tel que :

$$F(h, k) = h\varphi_1'(a + \theta_1 h) = h \left(\frac{\partial f}{\partial x}(a + \theta_1 h, b+k) - \frac{\partial f}{\partial x}(a + \theta_1 h, b) \right).$$

Comme la fonction $t \mapsto \frac{\partial f}{\partial x}(a + \theta_1 h, t)$ est dérivable au voisinage de b , de dérivée $\frac{\partial^2 f}{\partial y \partial x}(a + \theta_1 h, t)$, le théorème des accroissements finis permet de trouver un réel $\tau_1 \in]0, 1[$ tel que :

$$F(h, k) = hk \frac{\partial^2 f}{\partial y \partial x}(a + \theta_1 h, b + \tau_1 k).$$

Or $\lim_{(h,k) \rightarrow (0,0)} (a + \theta_1 h, b + \tau_1 k) = (a, b)$, donc en tenant compte de la continuité de $\frac{\partial^2 f}{\partial y \partial x}$, on obtient

$$\lim_{(h,k) \rightarrow (0,0)} \frac{F(h, k)}{hk} = \frac{\partial^2 f}{\partial y \partial x}(a, b).$$

La suite de la preuve consiste à reprendre le raisonnement précédent en posant au départ $\varphi_2(s) = f(a+h, s) - f(a, s)$, d'où $F(h, k) = \varphi_2(b+k) - \varphi_2(b)$. On obtiendrait

$$F(h, k) = hk \frac{\partial^2 f}{\partial x \partial y}(a + \tau_2 h, b + \theta_2 k),$$

où θ_2 et τ_2 sont deux réels strictement compris entre 0 et 1. La continuité de $\frac{\partial^2 f}{\partial x \partial y}$ entraîne :

$$\lim_{(h,k) \rightarrow (0,0)} \frac{F(h, k)}{hk} = \frac{\partial^2 f}{\partial x \partial y}(a, b).$$

L'unicité de la limite implique alors l'égalité entre les dérivées $\frac{\partial^2 f}{\partial y \partial x}$ et $\frac{\partial^2 f}{\partial x \partial y}$ au point (a, b) , ce qui achève la démonstration. ■

Remarque 1.37 La contraposée du lemme de Schwarz permet de montrer que certaines fonctions ne sont pas de classe $\mathcal{C}^2$. A titre d'exemple, on reprend la fonction définie par (voir exemple 1.20),

$$f(x, y) = \begin{cases} xy \frac{x^4 - y^4}{x^4 + y^4}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

et on montre qu'elle n'est pas de classe $\mathcal{C}^2$. On a montré dans l'exemple 1.20, que cette fonction est continue, différentiable sur $\mathbb{R}^2$ et de classe $\mathcal{C}^1$. Pour prouver qu'elle n'est pas de classe $\mathcal{C}^2$, il suffit d'utiliser la contraposée du lemme de Schwarz car

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) (0, 0) = \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial y}(h, 0) - \frac{\partial f}{\partial y}(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{h}{h} = 1,$$

et

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) (0, 0) = \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, h) - \frac{\partial f}{\partial x}(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{-h}{h} = -1,$$

d'où,

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

On constate d'ailleurs que même si les dérivées partielles $\frac{\partial^2 f}{\partial y \partial x}$ et $\frac{\partial^2 f}{\partial x \partial y}$ en un point existent, elles ne sont pas nécessairement égales.

Proposition 1.38 Soit f une fonction de deux variables à valeurs réelles, définie dans un voisinage Ω du point $(a, b) \in \mathbb{R}^2$. On suppose que les conditions suivantes sont satisfaites :

(i) les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial^2 f}{\partial y \partial x}$ existent dans Ω .

(ii) $\frac{\partial^2 f}{\partial y \partial x}$ est continue sur Ω .

(iii) $\frac{\partial f}{\partial y}(x, b)$ existe pour x dans un voisinage de a .

Alors, $\frac{\partial^2 f}{\partial x \partial y}$ existe en (a, b) et on a

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

Démonstration : Comme dans la preuve de la proposition précédente, on pose

$$F(h, k) = f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b).$$

Dès lors

$$\begin{aligned} \lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{F(h, k)}{k} &= \lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{f(a + h, b + k) - f(a + h, b)}{k} \\ &\quad - \lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{f(a, b + k) - f(a, b)}{k}, \\ &= \frac{\partial f}{\partial y}(a + h, b) - \frac{\partial f}{\partial y}(a, b), \end{aligned} \tag{1.1}$$

et cette limite existe par hypothèse. D'après la proposition précédente, on sait que :

$$F(h, k) = h \left(\frac{\partial f}{\partial x}(a + \theta_1 h, b + k) - \frac{\partial f}{\partial x}(a + \theta_1 h, b) \right), \quad \theta_1 \in]0, 1[$$

En appliquant la formule des accroissements finis à $\frac{\partial f}{\partial x}$ en la deuxième variable, on obtient

$$F(h, k) = hk \frac{\partial^2 f}{\partial y \partial x}(a + \theta_1 h, b + \tau_1 k), \quad \tau_1 \in]0, 1[$$

Par hypothèse $\frac{\partial^2 f}{\partial y \partial x}$ existe et est continue sur Ω , donc

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{1}{h} \lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{F(h, k)}{k} = \frac{\partial^2 f}{\partial y \partial x}(a, b),$$

et d'après (1.1), on en déduit que la limite

$$\begin{aligned} \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{1}{h} \lim_{\substack{k \rightarrow 0 \\ k \neq 0}} \frac{F(h, k)}{k} &= \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{\frac{\partial f}{\partial y}(a + h, b) - \frac{\partial f}{\partial y}(a, b)}{h}, \\ &= \frac{\partial^2 f}{\partial x \partial y}(a, b), \end{aligned}$$

existe et est égale à $\frac{\partial^2 f}{\partial y \partial x}(a, b)$, en vertu de l'unicité de la limite. ■

Proposition 1.39 (*Formule de Taylor*). Soient $a \in \mathbb{R}^n$, Ω un voisinage de a , $f : \Omega \rightarrow \mathbb{R}$ et $h \in \mathbb{R}^n$ tel que $a + th \in \Omega$ pour tout $t \in [0, 1]$ (c.-à-d., tel que le segment $[a, a + h]$, soit contenu dans Ω). On suppose que $f \in \mathcal{C}^{r+1}$ sur Ω . Alors, il existe $\theta \in]0, 1[$ tel que :

$$\begin{aligned} f(a + h) &= f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i + \frac{1}{2} \sum_{i_1, i_2=1}^n \frac{\partial^2 f}{\partial x_{i_2} \partial x_{i_1}}(a) h_{i_1} h_{i_2} + \cdots \\ &+ \frac{1}{r!} \sum_{i_1, \dots, i_r=1}^n \frac{\partial^r f}{\partial x_{i_r} \dots \partial x_{i_1}}(a) h_{i_1} \dots h_{i_r} \\ &+ \frac{1}{(r+1)!} \sum_{i_1, \dots, i_{r+1}=1}^n \frac{\partial^{r+1} f}{\partial x_{i_{r+1}} \dots \partial x_{i_1}}(a + \theta h) h_{i_1} \dots h_{i_{r+1}}. \end{aligned}$$

Démonstration : L'idée de la preuve consiste à utiliser un raisonnement par récurrence. Considérons la fonction $\varphi : [0, 1] \rightarrow \mathbb{R}$, $t \mapsto f(a + th)$. Alors, il existe $\theta \in]0, 1[$ tel que l'on ait :

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2!} g''(0) + \cdots + \frac{1}{r!} g^{(r)}(0) + \frac{1}{r!} g^{(r)}(0) \frac{1}{(r+1)!} g^{(r+1)}(\theta).$$

Notons que : $\varphi(0) = f(a)$ et $\varphi(1) = f(a + h)$. (Notons que pour une fonction $g : \Omega \rightarrow \mathbb{R}$ de classe $\mathcal{C}^1$, on a d'après la règle de différentiation de fonctions composées appliquée à $[0, 1] \rightarrow \Omega \rightarrow \mathbb{R}$, $t \mapsto a + th \mapsto g(a + th)$, l'expression :

$$\frac{d}{dt} g(a + th) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a + th) \cdot h_j.$$

On calcul les dérivées successives de la fonction φ . On a

$$\varphi'(t) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(a + th) \cdot h_j.$$

Pour j fixé, on peut dériver $t \mapsto \frac{\partial f}{\partial x_j}(a + th)$, d'où

$$\varphi''(t) = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(a + th) \cdot h_i h_j.$$

Montrons par récurrence sur k , $1 \leq k \leq r$, que l'on a

$$\varphi^{(k)}(t) = \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a + th) \cdot h_{i_1} \dots h_{i_k}.$$

Acceptant par hypothèse de récurrence cette expression, on donc

$$\begin{aligned}
\varphi^{(k+1)}(t) &= \frac{d}{dt} \left(\sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a+th) \cdot h_{i_1} \dots h_{i_k} \right), \\
&= \sum_{i_1, \dots, i_k=1}^n \frac{d}{dt} \left(\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(a+th) \right) h_{i_1} \dots h_{i_k}, \\
&= \sum_{i_1, \dots, i_k=1}^n \left(\sum_{i=1}^n \frac{\partial^{k+1} f}{\partial x_i \partial x_{i_k} \dots \partial x_{i_1}}(a+th) h_i \right) h_{i_1} \dots h_{i_k}, \\
&= \sum_{i_1, \dots, i_k, i_{k+1}=1}^n \frac{\partial^{k+1} f}{\partial x_{k+1} \partial x_{i_k} \dots \partial x_{i_1}}(a+th) h_{i_1} \dots h_{i_k} h_{i_{k+1}}, \quad (i_{k+1} \equiv i),
\end{aligned}$$

d'où le résultat annoncé. ■

Remarques 1.40 a) En pratique on utilise parfois d'autres notations. Le terme d'indice r dans la proposition précédente, se note souvent (en utilisant une puissance symbolique d'ordre r),

$$\frac{1}{r!} \sum_{i_1, \dots, i_r=1}^n \frac{\partial^r f}{\partial x_{i_r} \dots \partial x_{i_1}}(a) h_{i_1} \dots h_{i_r} \equiv \frac{1}{r!} \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^r f(a).$$

Par exemple, pour $n = 2$ et $r = 3$,

$$\frac{1}{3!} \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} \right)^3 f(a),$$

est une notation commode pour

$$\frac{1}{3!} \left(h_1^3 \frac{\partial^3 f}{\partial x_1^3}(a) + 3h_1^2 h_2 \frac{\partial^3 f}{\partial x_1^2 \partial x_2}(a) + 3h_1 h_2^2 \frac{\partial^3 f}{\partial x_1 \partial x_2^2}(a) + h_2^3 \frac{\partial^3 f}{\partial x_2^3}(a) \right).$$

b) Pour le cas d'une fonction de deux variables, on écrit parfois la formule de Taylor (d'ordre 2) en utilisant les notations de Monge :

$$p = \frac{\partial f}{\partial x}, \quad q = \frac{\partial f}{\partial y}, \quad r = \frac{\partial^2 f}{\partial x^2}, \quad s = \frac{\partial^2 f}{\partial x \partial y}, \quad t = \frac{\partial^2 f}{\partial y^2}.$$

Par exemple, on a

$$\begin{aligned}
f(a+h, b+k) &= f(a, b) + (hp + kq)(a, b) + \frac{1}{2!} (h^2 r + 2hks + k^2 t)(a, b) \\
&\quad + \frac{1}{3!} (hp + kq)^3 (a + \theta h, b + \theta k).
\end{aligned}$$

Exemple 1.41 La formule ci-dessus peut aussi être utile lors de la détermination des valeurs approchées de certains nombres. Par exemple pour déterminer la valeur approchée de $(0.95)^{2.01}$, on pose $f(x, y = y^x = e^{x \ln y})$. Avec les notations de Monge, on a

$$p = \frac{\partial f}{\partial x} = y^x \ln y, \quad q = \frac{\partial f}{\partial y} = xy^{x-1}, \quad r = \frac{\partial^2 f}{\partial x^2} = y^x (\ln y)^2,$$

$$s = \frac{\partial^2 f}{\partial x \partial y} = y^{x-1} + xy^{x-1} \ln y, \quad t = \frac{\partial^2 f}{\partial y^2} = x(x-1)y^{x-2}.$$

D'où,

$$(0.95)^{2.01} = f(2 + 0.01, 1 - 0.05) = f(a + h, b + k),$$

$$\approx f(a, b) + (hp + kq)(a, b) + \frac{1}{2!}(h^2r + 2hks + k^2t)(a, b),$$

où $a = 2$, $b = 1$, $h = 0.01$, $k = -0.05$, $f(a, b) = 1$, $(hp + kq)(a, b) = -0.01$ et $(h^2r + 2hks + k^2t)(a, b) = 0.004$. Par conséquent,

$$(0.95)^{2.01} \approx 1 - 0.1 + \frac{1}{2}(0.004) = 0.902.$$

Définition 1.42 On dit qu'une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$, est homogène de degré α si, pour $\lambda \in \mathbb{R}_+^*$ et tout $x \in \mathbb{R}$, on ait $f(\lambda x) = \lambda^\alpha f(x)$.

Si f et g sont deux fonctions homogènes de degré α , alors $f + g$ est homogène de degré α . Si f est homogène de degré α et g est homogène de degré β , alors fg est homogène de degré $\alpha + \beta$ et $\frac{f}{g}$ est homogène de degré $\alpha - \beta$. Si f est homogène de degré α et $s \in \mathbb{R}^*$, alors la fonction f^s est homogène de degré α^s . Si f est différentiable et homogène de degré α , alors les dérivées partielles de f sont homogènes de degré $\alpha - 1$.

Proposition 1.43 Si f est différentiable en x et homogène de degré α , alors on a la formule d'Euler :

$$\sum_{k=1}^n x_k \frac{\partial f}{\partial x_k}(x) = \alpha f(x).$$

Démonstration : En effet, on a $f(\lambda x) = \lambda^\alpha f(x)$. En dérivant cette égalité par rapport à λ , on obtient

$$\sum_{k=1}^n \frac{\partial f}{\partial x_k}(\lambda x) \cdot x_k = \alpha \lambda^{\alpha-1} f(x).$$

Prenons $\lambda = 1$, d'où

$$\sum_{k=1}^n x_k \frac{\partial f}{\partial x_k}(x) = \alpha f(x),$$

ce qui achève la démonstration. ■

Soient $\Omega \subset \mathbb{R}^n$, un ouvert, $f : \Omega \longrightarrow \mathbb{R}^p$ et $a \in \Omega$. On dit qu'un polynôme $P : \mathbb{R}^n \longrightarrow \mathbb{R}^p$, de degré $\leq n$ est un développement limité de f à l'ordre n au point a , si

$$\|f(a+x) - P(x)\| = o(\|x\|^n).$$

(Cela revient à dire que la fonction $x \longmapsto f(a+x)$ est " n -tangente" à P à l'origine). Si P existe, alors il est unique. Dans le cas où f est n -fois différentiable au point a , la formule de Taylor exprime précisément que f admet un développement limité P à l'ordre n au point a .

1.1.4 Inversion locale et fonctions implicites

Considérons le système linéaire suivant de n équations à n inconnues,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n, \end{aligned}$$

ou sous forme condensée, $\sum_{j=1}^n a_{ij}x_j = b_i$, $1 \leq i \leq n$, ou encore sous forme

matricielle, $Ax = b$, avec $A = (a_{ij})$ une matrice $n \times n$, $b = (b_i)$ une matrice $n \times 1$ et $x = (x_j)$ une matrice $n \times 1$. La matrice A définit une transformation linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ pour laquelle on utilise aussi la notation A :

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad (x_1, \dots, x_n) \longmapsto (y_1, \dots, y_n), \quad y_i = \sum_{j=1}^n a_{ij}x_j, \quad 1 \leq i \leq n.$$

On sait que le système précédent a une solution unique si et seulement si la matrice A est inversible, donc si $\det A \neq 0$. Cela revient à dire que l'application A est bijective. La question que l'on se pose est qu'en est-il si on considère un système de n équations non-linéaires à n inconnues :

$$\begin{aligned} f_1(x_1, \dots, x_n) &= b_1, \\ f_2(x_1, \dots, x_n) &= b_2, \\ &\vdots \\ f_n(x_1, \dots, x_n) &= b_n, \end{aligned}$$

ou sous forme condensée $f(x) = b$? A la place de A interviendra le jacobien de f , c.-à-d.,

$$\det J_f(x_0) = \frac{\partial(f_1, \dots, f_n)}{\partial(x_1, \dots, x_n)} = \det \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{1 \leq i, j \leq n}.$$

Supposons que f soit de classe $\mathcal{C}^1$ au voisinage d'un point x_0 et que $f(x_0) = b_0$. L'équation $f(x) = b$ peut alors s'écrire $f(x) - f(x_0) = b - b_0$, ou encore sous la forme $df(x_0).(x - x_0) + \varepsilon(x - x_0) = b - b_0$. Une approximation linéaire de l'équation proposée est donc fournie par $df(x_0).(x - x_0) = b - b_0$. Comme la fonction linéaire $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^n$, est représentée par la matrice jacobienne $J_f(x_0) = \left(\frac{\partial f_j}{\partial x_i} \right)_{1 \leq i, j \leq n}$, on peut écrire l'équation linéaire approchée sous la forme $J_f(x_0).(x - x_0) = b - b_0$. L'équation approchée possède donc pour chaque b une solution unique en x si et seulement si la matrice jacobienne $J_f(x_0)$ est inversible et donc si et seulement si le jacobien de f en x_0 (c.-à-d. $\det J_f(x_0)$) n'est pas nul. Pourrait-on espérer que l'équation $f(x) = b$, elle-même possède une solution unique si c'est le cas de l'équation linéaire approchée (autrement dit si l'application linéaire approchant f est inversible)? Posée globalement, la question appelle une réponse négative. Mais localement, la réponse est positive.

Précisons cela en examinant un exemple simple dans le cas $n = 1$. Soit $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, une fonction de classe $\mathcal{C}^1$ sur l'ouvert I (suggérée par le graphe de la figure suivante) tout en supposant que $f' \neq 0$ en un point $x_0 \in I$.

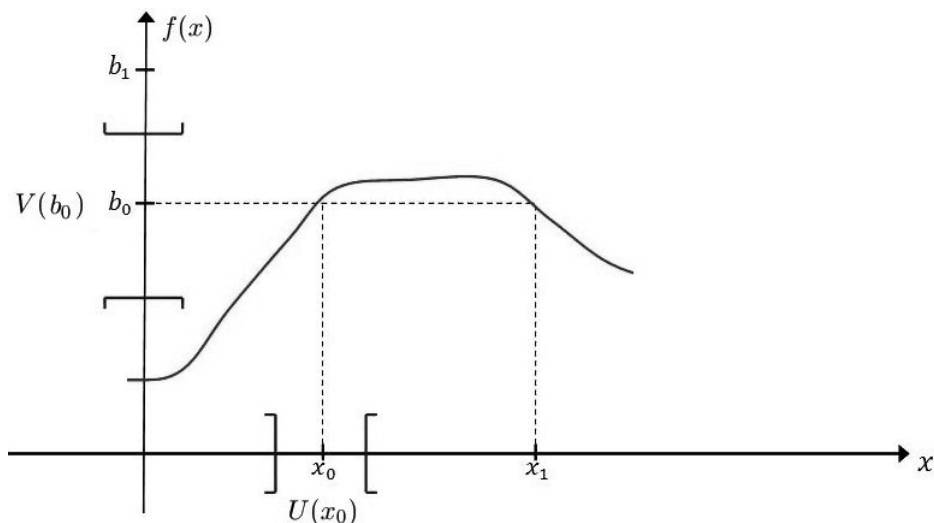


Figure 2 – Courbe

Pour fixer les idées, on suppose que $f'(x_0) > 0$, donc f' reste > 0 dans un voisinage $U(x_0)$ de x_0 , d'où f est strictement croissante et par conséquent

injective sur ce voisinage. L'image de ce voisinage est un ouvert $V(b_0)$ contenant $f(x_0)$. On a donc trouvé un voisinage ouvert $U(x_0)$ de x_0 et un voisinage ouvert $V(b_0)$ de $f(x_0)$ tels que : $f : U(x_0) \longrightarrow V(b_0)$ soit bijective. La matrice jacobienne de f en x_0 se réduit ici au seul nombre réel $\frac{df}{dx}(x_0)$ qui n'est pas nul (noter que la tangente à la courbe en (x_0, b_0) n'est pas horizontale). On peut donc en résumé, faire les observations suivantes :

(i) négativement : l'équation $f(x) = b$ n'a pas nécessairement de solution pour aucun réel ; c'est le cas par exemple de l'équation $f(x) = b_1$. De plus si l'équation $f(x) = b$ a une solution, celle-ci n'est pas nécessairement unique. Par exemple, l'équation $f(x) = b_0$ admet les deux solutions x_0 et x_1 .

(ii) positivement : si on considère les voisinages $U(x_0)$ de x_0 et $V(b_0)$ de b_0 , la fonction $f : U(x_0) \longrightarrow V(b_0)$ est bijective : pour tout $b \in V(b_0)$, il existe un et un seul $x \in U(x_0)$ tel que $f(x) = b$. En outre, la fonction réciproque $f^{-1} : V(b_0) \longrightarrow U(x_0)$ est encore différentiable et de classe $\mathcal{C}^1$.

C'est là un fait général comme le montre le théorème suivant (avec un léger changement de notations) :

Théorème 1.44 (*d'inversion locale*). Soit Ω un ouvert de $\mathbb{R}^n$ et $f : \Omega \longrightarrow \mathbb{R}^n$, une fonction de classe $\mathcal{C}^1$. Soit $a \in \Omega$, $b = f(a)$. Supposons que $df(a)$ soit inversible, autrement dit que

$$\det J_f(a) = \det \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \dots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \dots & \frac{\partial f_n}{\partial x_n}(a) \end{pmatrix} \neq 0,$$

(c.-à-d., la matrice jacobienne $\left(\frac{\partial f_i}{\partial x_j}\right)(a)$ est régulière). Alors, il existe un voisinage $U(a)$ de a et un voisinage $V(b)$ de b tels que la restriction de f à $U(a)$ soit une bijection de $U(a)$ sur $V(b)$. En outre, la réciproque $f^{-1} : V(b) \longrightarrow U(a)$, est de classe $\mathcal{C}^1$. (Si f est de classe $\mathcal{C}^k$, $k \in \mathbb{N}^*$, alors f^{-1} est de classe $\mathcal{C}^k$).

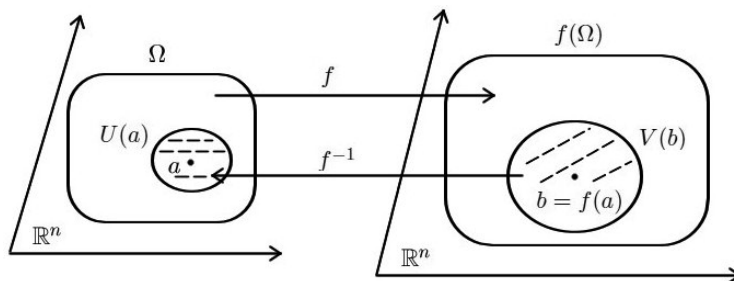


Figure 3 – Inversion locale

Démonstration : La preuve se fera en plusieurs étapes.

a) Posons

$$g(x) = (df(a))^{-1}(f(a+x) - f(a)), \quad x \in \Omega \setminus \{a\} = \{y - a : y \in \Omega\}.$$

Alors par ce changement de f , on se ramène à inverser $g : \Omega \setminus \{a\} \longrightarrow \mathbb{R}^n$ au voisinage de 0 avec $g(0) = 0$ et

$$dg(0) = (df(a))^{-1}(df(a)) = \mathbb{1} (\equiv id_{\mathbb{R}^n}).$$

Si g est un difféomorphisme local, il en est de même de

$$f(x) = df(a)(g(x) - a) + f(a).$$

En fait, par une telle transformation et composition, on se ramène au cas où $a = 0$, $b = f(a) = 0$ et le théorème en question s'appliquera à la fonction $h(x) = f(a+x) - f(a)$, car $h(0) = 0$ et $dh(0) = d(f(a))$. Dès lors, h admettra en 0 un inverse local g et donc admettra en a un inverse local $y \mapsto g(y - f(a)) + a$. on a

$$dg(0) = d(df(a))^{-1}(0)odf(a) = (df(a))^{-1}odf(a) = \mathbb{1}.$$

Supposons que $df(0) = \mathbb{1}$ car si le théorème est vrai dans ce cas, alors il s'appliquera à la fonction $(df(0))^{-1}of$ (où on compose $f : \Omega \longrightarrow \mathbb{R}^n$ avec $(df(0))^{-1} \longrightarrow \mathbb{R}^n$), puisque évidemment,

$$d((df(a))^{-1}of)(0) = d(df(0))^{-1}(f(0))odf(0) = (df(0))^{-1}odf(0) = \mathbb{1}.$$

L'isomorphisme linéaire $(df(0))^{-1}$ étant une bijection $\mathcal{C}^\infty$, prouver le théorème pour f est équivalent à le prouver pour $(df(0))^{-1}of$. Sans restreindre la généralité, on est donc ramené à traiter le cas $f : \Omega \longrightarrow \mathbb{R}^n$ avec $a = 0 \in \Omega$, $f(a) = 0$ et $df(0) = \mathbb{1}$, ce qu'on supposera désormais dans la suite. On posera donc $h(x) = f(x) - x$, d'où $h(0) = 0$, $dh(0) = df(0) - \mathbb{1} = 0$. Comme dh est continue en 0, alors il existe $r > 0$ tel que la boule fermée $\overline{B}(0, r)$ de centre 0 et de rayon r est contenue dans Ω et pour tout $x \in \overline{B}(0, r)$, $\|dh(x)\| \leq 1/2$. En appliquant l'inégalité des accroissements finis, on obtient pour tout $x \in \overline{B}(0, r)$, $\|h(x)\| \leq \frac{1}{2}\|x\|$. Par conséquent, l'image de $\overline{B}(0, r)$ par l'application h est contenue dans $\overline{B}(0, r/2)$.

b) Il s'agit maintenant de construire deux voisinages ouverts $U(0)$ et $V(0)$ de 0 tels que $g : U(0) \longrightarrow V(0)$ soit bijective. Nous allons montrer que pour $y \in \overline{B}(0, r/2)$, il existe un point fixe unique $x \in \overline{B}(0, r)$ tel que : $g(x) = y$. (L'image de $\overline{B}(0, r)$ par g est égale à $\overline{B}(0, r/2)$). Pour cela, on se propose de déterminer un point fixe de la fonction définie sur $\overline{B}(0, r)$ par

$$h_y(x) = x + y - g(x) = y - h(x),$$

où $y \in \overline{B}(0, r/2)$ (fixé). Si x est un point fixe de h_y , on a $x = y + x - g(x)$ ou $y = g(x)$. Donc x est un point fixe de h_y si et seulement si x est solution de l'équation $g(x) = y$. Le problème consiste à trouver une boule sur laquelle h_y soit contractante pour y dans un certain voisinage de 0. On a pour tout $x \in \overline{B}(0, r)$, $h(x) \in \overline{B}(0, r/2)$ et

$$\|h_y\| = \|x + y - g(x)\| = \|y - h(x)\| \leq \|y\| + \|h(x)\| \leq \frac{r}{2} + \frac{r}{2} = r,$$

et dès lors, $h_y(x) \in \overline{B}(0, r)$. Ainsi h_y envoie la boule fermée $\overline{B}(0, r)$ dans elle-même. On a $dh_y(x) = dh(x)$ et $\|dh_y(x)\| \leq 1/2$, pour tout $x \in \overline{B}(0, r)$. (Il est important que le côté droit soit plus petit que 1 mais il n'est pas nécessaire qu'il soit en particulier égal à $1/2$. Cela implique que pour chaque $x \in U(0)$, $dg(x)$ est inversible). Pour $x, \tilde{x} \in \overline{B}(0, r)$, on a alors d'après l'inégalité des accroissements finis :

$$\|h_y(x) - h_y(\tilde{x})\| = \|h(x) - h(\tilde{x})\| \leq \|x - \tilde{x}\| \sup_{\xi \in [x, \tilde{x}]} \|dh(\xi)\| \leq \frac{\|x - \tilde{x}\|}{2}.$$

Cela prouve que pour $y \in B(0, r/2)$, h_y est une contraction de $\overline{B}(0, r)$ dans $\overline{B}(0, r)$. D'après le théorème du point fixe, il existe un unique $x \in B(0, r)$ tel que : $h_y(x) = x$, c.-à-d., l'équation $g(x) = y$ possède une solution unique dans $B(0, r)$. Dès lors, $\exists! x \in B(0, r)$ tel que :

$$h_y(x) = x \iff g(x) = y,$$

ce qui définit une application inverse de g de $\overline{B}(0, r/2)$ dans $\overline{B}(0, r)$. On pose $U(0) \equiv g^{-1}(B(0, r/2))$ et $V(0) \equiv B(0, r/2)$, et on note $g^{-1} : V(0) \rightarrow U(0)$.

c) Étudions la continuité de g^{-1} . Soient $x = g^{-1}(y)$ et $\tilde{x} = g^{-1}(\tilde{y})$ avec $x, \tilde{x} \in B(0, r)$ et $y, \tilde{y} \in V(0)$. On a

$$\begin{aligned} \|x - \tilde{x}\| &= \|x - g(x) + g(x) - \tilde{x} + g(\tilde{x}) - g(\tilde{x})\|, \\ &\leq \|g(x) - g(\tilde{x})\| + \|h(x) - h(\tilde{x})\|, \\ &\leq \|y - \tilde{y}\| + \frac{\|x - \tilde{x}\|}{2}, \end{aligned}$$

donc

$$\|x - \tilde{x}\| \leq 2\|y - \tilde{y}\|,$$

ce qui prouve que g^{-1} est continue sur $V(0)$. (g^{-1} est lipschitzienne).

d) Étudions maintenant la différentiabilité de g^{-1} . Soient $y \in V(0)$ et $z \in \Omega$ tel que : $y + u \in V(0)$. On note comme ci-dessus $x = g^{-1}(y) \in U(0)$ et

$x + v = g^{-1}(y + u) \in U(0)$. On a

$$\begin{aligned} g^{-1}(y + u) - g^{-1}(y + u)g^{-1}(y) - (dg_x)^{-1}(u) \\ = x + v - x - (dg_x)^{-1}(g(x + v) - g(x)), \quad (x = g^{-1}(y)), \\ = (dg_x)^{-1}(g(x + v) - g(x) - dg_x(v)). \end{aligned}$$

Comme dg_x est inversible (en fait en diminuant r si nécessaire, on peut supposer que pour tout $x \in B(0, r)$, dg_x est un isomorphisme). Dès lors, pour chaque $x \in \overline{B}(0, r)$, il existe une constante $C > 0$ tel que : $\|(dg_x)^{-1}\| \leq C$. D'après c), g^{-1} est lipschitzienne, donc $\|v\| \leq 2\|u\|$ et dès lors,

$$\begin{aligned} \frac{\|g^{-1}(y + u) - g^{-1}(y)g^{-1}(y) - (dg_x)^{-1}(u)\|}{\|u\|} \\ \leq \|(dg_x)^{-1}\| \frac{\|g(x + v) - g(x) - dg_x(v)\|}{\|u\|}, \quad (x = g^{-1}(y)), \\ \leq 2C \frac{\|g(x + v) - g(x) - dg_x(v)\|}{\|v\|}. \end{aligned}$$

Lorsque $v \rightarrow 0$, le dernier facteur ci-dessus tend vers 0 par la différentiabilité de g . Cela montre que g^{-1} est différentiable en y et que pour chaque $y \in V(0)$ avec $x = g^{-1}(y)$,

$$dg^{-1}(y) = (dg(x))^{-1}.$$

e) Comme g est de classe $\mathcal{C}^r$, alors dg est de classe $\mathcal{C}^{r-1}$. On a vu que $dg^{-1}(y) = (dg(x))^{-1}$ et comme l'opération inverse sur les matrices inversibles est lisse, cela établit que dg^{-1} est de classe $\mathcal{C}^{r-1}$. Par conséquent, g^{-1} est de classe $\mathcal{C}^r$. Ceci achève la démonstration. ■

Définition 1.45 Une bijection f d'un ouvert Ω de $\mathbb{R}^n$ sur un ouvert $f(\Omega)$ de $\mathbb{R}^n$ qui est de classe $\mathcal{C}^k$, $k \in \mathbb{N}^*$, ainsi que sa réciproque f^{-1} s'appelle un *difféomorphisme de classe $\mathcal{C}^k$* (ou simplement *$\mathcal{C}^k$ -difféomorphisme*).

Remarques 1.46 a) Sous les mêmes hypothèses du théorème précédent (avec f de classe $\mathcal{C}^k$, $k \in \mathbb{N}^*$), f est un $\mathcal{C}^k$ -difféomorphisme de $U(a)$ sur $V(b)$.

b) Il faut bien noter que le théorème précédent est local (ce qui justifie son appellation). Même si $df(x)$ est inversible pour tout x , l'inverse global de f n'existe pas nécessairement. Pour s'en convaincre, considérons la fonction,

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad (x, y) \longmapsto (e^x \cos y, e^x \sin y).$$

Le jacobien

$$\det J_f(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^x \neq 0,$$

ne s'annule jamais. Donc cette fonction a une matrice jacobienne partout inversible et pourtant f n'a pas de réciproque globale car f n'est pas bijective ; en effet f est localement injective mais elle n'est pas injective car $f(x, y) = f(x, y + 2\pi)$.

c) Le théorème précédent signifie qu'une fonction est inversible au voisinage d'un point en lequel sa différentielle est inversible. La condition $\det J_f(a) \neq 0$ est équivalente à $df(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ est un isomorphisme (c.-à-d., ici bijective). Une question qui se pose souvent est comment déterminer la différentielle. Or ce théorème ne permet pas toujours de calculer explicitement cette fonction pour tout $y \in V$, mais on peut procéder sans la connaissance explicite de f^{-1} comme suit. On notera que sous les hypothèses du théorème, puisque $f : U \longrightarrow V$ et $f^{-1} : V \longrightarrow U$ sont de classe $\mathcal{C}^1$ donc différentiables, on a par la formule des fonctions composées, $df(x) \circ df^{-1}(y) = df^{-1}(y) \circ df(x) = I$ pour $y = f(x)$ avec $x \in U$ (I étant l'application identité dans $\mathbb{R}^n$) ; par conséquent $df^{-1}(y) = (df(x))^{-1}$ où $(df(x))^{-1}$ désigne la réciproque de l'isomorphisme $df(x)$. Autrement dit, la matrice jacobienne de f^{-1} en y est l'inverse de la matrice jacobienne de f en x .

d) Le raisonnement fait dans la remarque b) concerne évidemment aussi le cas du calcul des dérivées partielles de la réciproque $\phi \equiv f^{-1} : V \longrightarrow U$, sans passer par la détermination explicite de la fonction ϕ (car comme nous l'avons signalé ci-dessus le théorème ne permet pas toujours de la calculer explicitement). Comme $\phi \circ f = \text{id}_U$, alors la matrice jacobienne de $\phi \circ f$ est la matrice unité (que l'on note aussi $I = (\delta_{ij})$ où δ_{ij} est le symbole de Kronecker qui vaut 1 si $i = j$ et 0 sinon). Dès lors, $J_\phi(y) \cdot J_f(x) = I$ ou

$$\sum_{k=1}^n \frac{\partial \phi_i}{\partial y_k}(y) \cdot \frac{\partial f_k}{\partial x_j}(x) = \delta_{ij}.$$

La résolution de ce système d'équations linéaires donne une information sur les $\frac{\partial \phi_i}{\partial y_k}$. Par exemple, si $y = f(x) = e^x + 3 \cos x \sin x$, alors on a

$$\frac{df}{dx} = e^x - 3 \sin^2 x + 3 \cos^2 x \implies \frac{d\phi}{dy}(0) \cdot \frac{df}{dx}(0) = 1 \implies \frac{d\phi}{dy}(0) = \frac{1}{4}.$$

Corollaire 1.47 Soient $\Omega \subset \mathbb{R}^n$ un ouvert et $f : \Omega \longrightarrow \mathbb{R}^n$ une fonction de classe $\mathcal{C}^1$. On suppose que $df(x)$ est un isomorphisme pour tout $x \in \Omega$ (ou ce qui est équivalent $\det J_f(x) \neq 0$).

- a) Si $\Delta \subset \Omega$ est un ouvert, alors $f(\Delta) \subset \mathbb{R}^n$ est un ouvert.
- b) La fonction f est injective au voisinage de chaque point de Ω .
- c) La fonction f peut ne pas être injective sur Ω tout entier même si Ω est connexe (et un tel phénomène ne peut pas se produire lorsque $n = 1$).

Démonstration : a) Soit $y \in f(\Delta)$ ou $\Delta \subset \Omega$ est un ouvert. Dès lors, $\exists x \in \Delta$ tel que $y = f(x)$. D'après le théorème d'inversion locale (appliqué au point x), il existe un voisinage de y entièrement contenu dans l'image $f(\Delta)$, d'où le résultat.

b) Découle immédiatement du théorème d'inversion locale.

c) En effet, il suffit de considérer l'exemple étudié dans la remarque 1.46, b), c.-à-d., $f(x, y) = (e^x \cos y, e^x \sin y)$. On a montré que f est localement injective mais f n'est pas injective puisque $f(x, y) = f(x, y + 2\pi)$. Pour le cas $n = 1$, on a déjà expliqué cela pour la fonction suggérée par le graphe de la figure 2. ■

Si on suppose de plus dans le théorème d'inversion locale (voir aussi le corollaire précédent) que la fonction f est injective sur l'ouvert Ω , alors l'image $f(\Omega)$ est un ouvert de $\mathbb{R}^n$ et f est un $\mathcal{C}^1$ -difféomorphisme de Ω sur $f(\Omega)$; c'est le *théorème d'inversion globale*. (On peut remplacer $\mathcal{C}^1$ par $\mathcal{C}^k$, $k \in \mathbb{N}^*$ si f est de classe $\mathcal{C}^k$). C'est une conséquence du théorème d'inversion locale (il suffit de montrer que f^{-1} est de classe $\mathcal{C}^k$). L'hypothèse d'injectivité dans ce théorème est essentielle, sans cela le théorème est faux. En général, un difféomorphisme local n'est pas un difféomorphisme global. Pour s'en convaincre, considérons la fonction $f(x, y) = (e^x \cos y, e^x \sin x)$. On a $\det J_f(x) = e^x \neq 0$ pour tout $(x, y) \in \mathbb{R}^2$. C'est donc un difféomorphisme local (en vertu du théorème d'inversion locale), toutefois la fonction f n'est pas injective car $f(x, y + 2\pi) = f(x, y)$ et dès lors ne peut être un difféomorphisme global. Un autre exemple est fournie par la fonction $f : \Omega = \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$, $(x, y) \mapsto (x^2 - y^2, 2xy)$, qui est de classe $\mathcal{C}^1$ et que ses différentielles sont inversibles. On a $\det J_f(x) = 4(x^2 + y^2) \neq 0$ si $(x, y) \neq (0, 0)$, $f(x, y) = f(-x, -y)$ pour tout $(x, y) \in \mathbb{R}^2$, la fonction f n'est pas injective au voisinage de $(0, 0)$ et elle n'est pas un $\mathcal{C}^1$ -difféomorphisme de $\mathbb{R}^2$ dans $\mathbb{R}^2$.

Corollaire 1.48 *Sous les hypothèses de la proposition précédente, pour tout voisinage ouvert $W \subset U$ de x , alors $f(W)$ est un voisinage de ouvert de $f(x)$ et en outre, f est bijective de W sur $f(W)$ avec une réciproque de classe $\mathcal{C}^1$ ($\mathcal{C}^k$ si f l'est).*

Démonstration : Il suffit de vérifier que $f(W)$ est un ouvert, la suite découle directement du théorème d'inversion locale. Comme $f^{-1} \circ f = I$ (identité sur U), alors d'après la formule de dérivation des fonctions composées, on a

$$df^{-1}(f(x)) \circ df(x) = I,$$

pour tout $x \in U$. Dès lors, $df(x)$ est un isomorphisme pour tout $x \in U$ (la condition $\det J_f(x) \neq 0$ dans la proposition est équivalent à $df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ est un isomorphisme). La conclusion suit donc le corollaire précédent, point a) et la preuve s'achève. ■

Remarque 1.49 Soient $(x, y) \in \mathbb{R}^n \times \mathbb{R}^p$ et W un voisinage de (x, y) dans $\mathbb{R}^n \times \mathbb{R}^p$. Alors il existe U un voisinage de x dans $\mathbb{R}^n$ et V un voisinage de y dans $\mathbb{R}^p$ tels que : $U \times V \subset W$. Par ailleurs, si U est un voisinage ouvert de $(x, 0)$ dans $\mathbb{R}^n \times \mathbb{R}^p$ alors U défini par $U \times \{0\} = U \cap (\mathbb{R}^n \times \{0\})$ est un voisinage ouvert de x dans $\mathbb{R}^n$.

Le théorème des fonctions implicites suivant permet de résoudre une équation de la forme $g(x, y) = 0$ en exprimant une des variables en fonction des autres, comme par exemple $y = f(x)$ où f est une fonction connue implicitement. De manière plus précise, le théorème en question s'énonce comme suit,

Théorème 1.50 (des fonctions implicites). Soit Ω un ouvert de $\mathbb{R}^n \times \mathbb{R}^p$ et $g : \Omega \rightarrow \mathbb{R}^p$ une fonction de classe $\mathcal{C}^1$. Soit $(a, b) \in \Omega$. Supposons que :

(i) $g(a, b) = 0$.

(ii) la matrice $\left(\frac{\partial g_i}{\partial y_j}(a, b) \right)_{1 \leq i, j \leq p}$ est inversible.

Alors, il existe un voisinage U de a dans $\mathbb{R}^n$ et un voisinage V de b dans $\mathbb{R}^p$, avec $(a, b) \in U \times V \subset \Omega$, tels qu'il existe une fonction unique $f : U \rightarrow V$, avec

(i)' $b = f(a)$.

(ii)' $g(x, f(x)) = 0$ pour tout $x \in U$.

La fonction f est de classe $\mathcal{C}^1$ et de plus, si g est de classe $\mathcal{C}^k$ ($k \geq 1$), f est de classe $\mathcal{C}^k$.

Démonstration : Soit

$$G : \Omega \rightarrow \mathbb{R}^n \times \mathbb{R}^p, \quad (x, y) \mapsto (x, g(x, y)),$$

la fonction de classe $\mathcal{C}^1$ ($\mathcal{C}^k$ si g l'est) et dont la matrice jacobienne est

$$\begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & & & & \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_n} & \frac{\partial g_1}{\partial y_1} & \dots & \frac{\partial g_1}{\partial y_p} \\ \vdots & & & & & & \\ \frac{\partial g_p}{\partial x_1} & \frac{\partial g_p}{\partial x_2} & \dots & \frac{\partial g_p}{\partial x_n} & \frac{\partial g_p}{\partial y_1} & \dots & \frac{\partial g_p}{\partial y_p} \end{pmatrix} = \begin{pmatrix} I_n & O \\ \left(\frac{\partial g_i}{\partial x_k} \right)_{\substack{1 \leq i \leq p \\ 1 \leq k \leq n}} & \left(\frac{\partial g_i}{\partial y_j} \right)_{1 \leq i, j \leq p} \end{pmatrix},$$

où I_n est la matrice unité d'ordre n et O est la matrice nulle à n lignes et p colonnes. La matrice jacobienne de G en (a, b) est

$$\begin{pmatrix} I_n & O \\ \left(\frac{\partial g_i}{\partial x_k}(a, b) \right)_{\substack{1 \leq i \leq p \\ 1 \leq k \leq n}} & \left(\frac{\partial g_i}{\partial y_j}(a, b) \right)_{1 \leq i, j \leq p} \end{pmatrix}.$$

On note, par suite de l'hypothèse (ii), que cette matrice jacobienne est inversible et donc le théorème d'inversion locale s'applique à G . Il existe donc un voisinage ouvert W de (a, b) dans $\mathbb{R}^n \times \mathbb{R}^p$ tel que $G(W)$ est un voisinage ouvert de $G(a, b) = (a, g(a, b)) = (a, 0)$ (d'après (i)) dans $\mathbb{R}^n \times \mathbb{R}^p$ et G est une bijection $\mathcal{C}^1$ de W sur $G(W)$ dont la réciproque $G^{-1} : G(W) \rightarrow W$ est de classe $\mathcal{C}^1$ ($\mathcal{C}^k$ si g l'est). D'après le corollaire 1.48 et la remarque 1.49, on peut choisir W de la forme $U' \times V$ où U' est un voisinage ouvert de a dans $\mathbb{R}^n$ et V est un voisinage ouvert de b dans $\mathbb{R}^p$. La fonction G^{-1} est nécessairement de la forme, $G^{-1} : G(U' \times V) \rightarrow U' \times V$, $(x, t) \mapsto (x, \varphi(x, t))$ où $(x, \varphi(x, t))$ est l'unique élément de $U' \times V$ tel que : $G((x, \varphi(x, t))) = (x, t)$, c.-à-d., $\varphi(x, t)$ est l'unique élément de V tel que : $g(x, \varphi(x, t)) = t$. En outre, $\varphi : G(U' \times V) \rightarrow V$ est de classe $\mathcal{C}^1$ ($\mathcal{C}^k$ si g l'est). Par ailleurs (remarque 1.49), il existe un voisinage ouvert $U \subset U'$ de a dans $\mathbb{R}^n$ tel que : $U \times \{0\} \subset G(U' \times V)$. Posons $f(x) = \varphi(x, 0)$, $x \in U$. Autrement dit, pour x tel que $(x, 0) \in V_1$, on pose $G^{-1}(x, 0) = (x, f(x))$. Notons que f est définie et de classe $\mathcal{C}^1$ dans un voisinage de a (car G^{-1} est définie et de classe $\mathcal{C}^1$ dans un voisinage de $G(a, b) = (a, 0)$). On a $(a, f(a)) = G^{-1}(a, 0) = G^{-1}(G(a, b)) = (a, b)$ et donc $f(a) = b$. On a aussi $(x, 0) = G(G^{-1}(x, 0)) = G(x, f(x)) = (x, g(x, f(x)))$ et donc $f(x)$ est l'unique élément de V tel que : $g(x, f(x)) = 0$. La figure ci-dessous illustre cette situation.

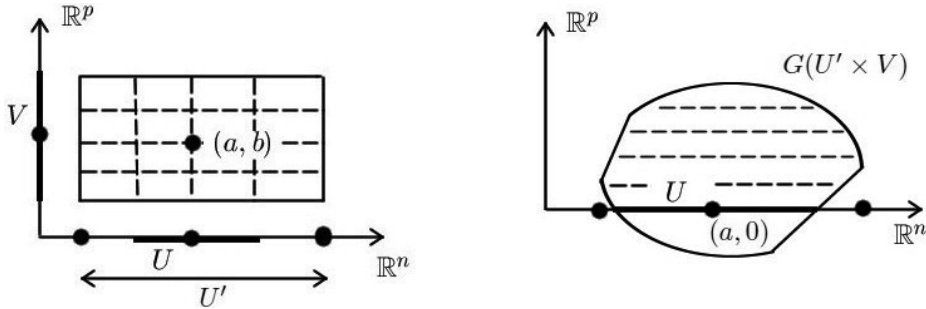


Figure 4 – Fonction implicite

Enfin, si g est de classe $\mathcal{C}^k$, G est de classe $\mathcal{C}^k$ donc G^{-1} est de classe $\mathcal{C}^k$ et f est bien de classe $\mathcal{C}^k$. ■

Exemple 1.51 Montrons que l'on peut représenter la courbe d'équation :

$$g(x, y) = y^2 - 2x^3 - x^2 = 0,$$

par une équation $x = f(y)$, au voisinage du point $(1, \sqrt{3})$ et déterminons $f'(\sqrt{3})$. En effet, on a

$$\frac{\partial g}{\partial x} = -6x^2 - 2x \implies \frac{\partial g}{\partial x}(1, \sqrt{3}) = -8 \neq 0,$$

et le résultat découle immédiatement du théorème des fonctions implicites. En outre, on a $0 = g(x, y) = g(f(y), y)$, $u = f(y)$, $v = y$, d'où

$$0 = \frac{\partial g}{\partial y} = \frac{\partial g}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial g}{\partial v} \frac{\partial v}{\partial y} = \frac{\partial g}{\partial x} f'(y) + \frac{\partial g}{\partial y} = (-6x^2 - 2x)f'(y) + 2y.$$

Pour $(x, y) = (1, \sqrt{3})$, on a $-8f'(\sqrt{3}) + 2\sqrt{3} = 0$, d'où $f'(\sqrt{3}) = \frac{\sqrt{3}}{4}$.

1.1.5 Recherche d'extremum

Soient Ω un ouvert de $\mathbb{R}^n$, $f : \Omega \longrightarrow \mathbb{R}$ et $a \in \Omega$.

Définition 1.52 a) On dit que f possède en a un maximum (resp. minimum) local ou relatif s'il existe un voisinage $\mathcal{V}$ de a inclus dans Ω tel que :

$$\forall x \in \mathcal{V}, \quad f(x) \leq f(a) \quad (\text{resp. } f(x) \geq f(a)).$$

b) On dit que f possède en a un maximum (resp. minimum) global si :

$$\forall x \in \Omega, \quad f(x) \leq f(a) \quad (\text{resp. } f(x) \geq f(a)).$$

c) Un extremum est un maximum ou un minimum.

Proposition 1.53 (Condition nécessaire) : Si f est différentiable en a et présente un extremum en a , alors $df(a) = 0$. (On dit dans ce cas que le point a est un point stationnaire ou encore point critique).

Démonstration : Si f possède un extremum en a , il existe une boule ouverte $B(a, r) \subset \Omega$ telle que : $\forall x \in B(a, r)$, $f(x) \leq f(a)$, (ou $f(x) \geq f(a)$). Soit $h \in \mathbb{R}^n$, $h \neq 0$. Pour tout $t \in \mathbb{R}$ tel que : $|t| \leq \frac{r}{\|h\|}$, on a $a + th \in B(a, r)$. La fonction $g(t) = f(a + th)$ est définie sur $\left] -\frac{r}{\|h\|}, \frac{r}{\|h\|} \right]$, elle possède un extremum en $t = 0$ et y est dérivable. Dès lors, $g'(0) = 0$. Or $g'(t) = f'(a + th)h$, donc $f'(a)h = 0$ pour tout $h \in \mathbb{R}^n$; l'application $f'(a) : \mathbb{R}^n \longrightarrow \mathbb{R}$ est nulle. ■

Définition 1.54 Soit $f : \Omega \longrightarrow \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. On appelle matrice hessienne (ou tout simplement hessienne) de f en a , la matrice suivante :

$$\mathcal{H}(f, a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix} = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \right)_{1 \leq i, j \leq n}.$$