CHAPITRE II. — APPLICATIONS DIFFERENTIABLES

Dorénavant, tous nos espaces vectoriels sont de dimension finie. On note $(\mathrm{E}, ||.||_{\mathrm{E}})$ et $(\mathrm{F}, ||.||_{\mathrm{F}})$ des espaces vectoriels normés sur $\mathbb{R}$, U un ouvert non vide de E et $f : \mathrm{U} \to \mathrm{F}$ une application. Par un choix de bases pour E et F, on se ramène à $f : \mathbb{R}^p \to \mathbb{R}^q$.

3 Définitions, exemples et premières propriétés

Définition 3.1 (différentiabilité). — *Une application $f : \mathrm{U} \to \mathrm{F}$ est différentiable au point $a \in \mathrm{U}$ s'il existe une application linéaire $\mathrm{L} \in \mathcal{L}(\mathrm{E}, \mathrm{F})$ et une application $\varepsilon : \mathrm{U} \to \mathrm{F}$ telles que*

$$f(x) = f(a) + \mathrm{L}(x - a) + \|x - a\|_{\mathrm{E}}\varepsilon(x)$$

et où $\lim_{x\to a} \varepsilon(x) = 0$.

Autrement dit, si on se fixe $r > 0$ de sorte que $\mathrm{B}(a, r) \subset \mathrm{U}$ et h avec $\|h\|_{\mathrm{E}} \leq r$, alors

$$f(a + h) = f(a) + \mathrm{L}(h) + \|h\|_{\mathrm{E}}\rho(h)$$

où $\rho : \mathrm{B}(0, r) \to \mathrm{F}$ vérifie $\lim_{h\to 0} \rho(h) = 0$.

Une autre formulation consiste à écrire que, pour toute marge d'erreur $\varepsilon > 0$, il existe un degré de précision $\delta > 0$ telle que, si $\|h\|_{\mathrm{E}} \leq \delta$ alors $\|f(a + h) - f(a) - \mathrm{L}(h)\|_{\mathrm{F}} \leq \varepsilon\|h\|_{\mathrm{E}}$.

Lemme 3.2. — *Si f est différentiable en a alors l'application linéaire L est unique et s'appelle la différentielle de f au point a, ou la dérivée. On la note $\mathrm{L} = \mathrm{D}_a f$ ou $f'(a)$.*

DÉMONSTRATION. — Supposons que l'on ait deux applications linéaires L_1 et L_2 de sorte qu'il existe $\varepsilon_1, \varepsilon_2$ telles que $\lim_0 \varepsilon_1 = \lim_0 \varepsilon_2 = 0$ et

$$\begin{cases} f(a + h) = f(a) + \mathrm{L}_1(h) + \|h\|_{\mathrm{E}}\varepsilon_1(h) \\ f(a + h) = f(a) + \mathrm{L}_2(h) + \|h\|_{\mathrm{E}}\varepsilon_2(h) \end{cases}$$

En faisant la différence, on obtient

$$\|(\mathrm{L}_1 - \mathrm{L}_2)(h)\|_{\mathrm{F}} = \|h\|_{\mathrm{E}} \times \|\varepsilon_1(h) - \varepsilon_2(h)\|_{\mathrm{F}}$$

pour tout h. D'après le corollaire 2.11, cela implique $\mathrm{L}_1 = \mathrm{L}_2$. ■

Définition 3.3. — *On dit que $f : \mathrm{U} \to \mathrm{F}$ est différentiable sur U si f est différentiable en tout point de U. On obtient alors $\mathrm{D}f : \mathrm{U} \to \mathcal{L}(\mathrm{E}, \mathrm{F})$. On dit que f est continûment différentiable (ou $\mathcal{C}^1$) au point $a \in \mathrm{U}$ si f est différentiable au voisinage de a et $\mathrm{D}f$ est continue au point a. On dit que f est de classe $\mathcal{C}^1$ dans U si f est continûment différentiable en chaque point de U.*

3.1 Exemples

1. *Si f est constante, alors f est différentiable, même $\mathcal{C}^1$, et $\mathrm{D}_a f = 0$.*

En effet, on a $f(a + h) = f(a) + 0 + 0$.

2. *Si f est la restriction d'une application linéaire, alors f est de classe $\mathcal{C}^1$ et $\mathrm{D}_a f = f$ pour tout $a \in \mathrm{U}$.*

En effet, on a $f(a + h) = f(a) + f(h) + 0$ avec f linéaire. Le corollaire 2.9 montre que $\mathrm{D}f = f$ est continue.

3. *Si $f : \mathrm{E}_1 \times \mathrm{E}_2 \to \mathrm{F}$ est bilinéaire alors f est différentiable et $\mathrm{D}_{(a_1,a_2)}f(h, k) = f(a_1, k) + f(h, a_2)$.*

On munit $\mathrm{E} = \mathrm{E}_1 \times \mathrm{E}_2$ de la norme $\|(a, b)\|_{\mathrm{E}} = \max\{\|a\|_{\mathrm{E}_1}, \|b\|_{\mathrm{E}_2}\}$. On écrit

$$f(a_1 + h, a_2 + k) = f(a_1, a_2) + f(h, a_2) + f(a_1, k) + f(h, k)\,.$$

On vérifie que $\mathrm{L} : (h, k) \mapsto f(h, a_2) + f(a_1, k)$ est bien linéaire et $\|f(h, k)\| \leq \|f\|\|h\|_{\mathrm{E}_1}\|k\|_{\mathrm{E}_2} \leq \|(h, k)\|_{\mathrm{E}}^2 \times \|f\|$ d'après la proposition 2.14.

4. *Si* $(\mathrm{E}, \|\cdot\|)$ *est un espace euclidien, alors* $\mathrm{N} : x \mapsto \|x\|^2$ *est différentiable et* $\mathrm{D}_x\mathrm{N}(h) = 2\langle x, h\rangle$.

Comme ci-dessus, on a $\mathrm{N}(x+h) = \mathrm{N}(x) + 2\langle x, h\rangle + \|h\|^2$.

5. *Soit* E *un evn; l'application* $f : u \in \mathrm{GL(E)} \mapsto u^{-1}$ *est différentiable et* $\mathrm{D}_u f(h) = -u^{-1} \circ h \circ u^{-1}$ *pour tout* $u \in \mathrm{GL(E)}$.

On munit $\mathcal{L}(\mathrm{E})$ de la norme d'opérateur et on rappelle que GL(E) est un ouvert de $\mathcal{L}(\mathrm{E})$. Commençons par étudier la différentiabilité de f en l'identité Id. On étudie donc $(\mathrm{Id}+h)^{-1}$, $h \in \mathcal{L}(\mathrm{E})$. En analogie avec la formule $(1+x)^{-1} = \sum_{n\geq 0}(-1)^n x^n$, on s'intéresse à la suite d'applications linéaires $u_n = \sum_{0\leq k\leq n}(-1)^k h^k$ où $h^k = h \circ \ldots \circ h$ k fois. Si $\|h\| < 1$, alors cette série est normalement convergente puisque

$$\sum_{0\leq k\leq n} \|(-1)^k h^k\| \leq \sum_{0\leq k\leq n} \|h\|^k \leq \frac{1}{1-\|h\|}\,.$$

Donc (u_n) est convergente car $\mathcal{L}(\mathrm{E})$ est de dimension finie et complet. Vérifions maintenant que $(1+h)^{-1} = \sum_{k\geq 0}(-1)^k h^k$: pour cela, il suffit de montrer que $\lim(\mathrm{Id}+h)\circ u_n = \mathrm{Id}$. Or, on a

$$(\mathrm{Id}+h)\circ u_n = \sum_{0\leq k\leq n}(-1)^k h^k + \sum_{0\leq k\leq n}(-1)^k h^{k+1} = \mathrm{Id} + (-1)^n h^{n+1}\,.$$

Comme $\|(-1)^n h^{n+1}\| \leq \|h\|^{n+1}$ et $\|h\| < 1$, on a bien montré que $(\mathrm{Id}+h)$ est inversible et on a calculé $(\mathrm{Id}+h)^{-1}$. L'application $h \mapsto -h$ est linéaire et $\sum_{k\geq 2}(-1)^k h^k = \|h\|\varepsilon(h)$ avec $\lim_0 \varepsilon = 0$ puisque $\varepsilon(h) = (h^2/\|h\|)\sum_{k\geq 0}(-1)^k h^k = (h^2/\|h\|)(\mathrm{Id}+h)^{-1}$.

Du coup, on vient d'établir $\mathrm{D}_{\mathrm{Id}} f(h) = -h$.

Si $u \in \mathrm{GL(E)}$, alors il existe $\delta > 0$ telle que $\mathrm{B}(u,\delta) \subset \mathrm{GL(E)}$. Prenons $h \in \mathcal{L}(\mathrm{E})$, avec $\|h\| < \delta$. On a $(u+h)^{-1} = (\mathrm{Id}+u^{-1}h)^{-1}u^{-1}$ donc

$$\begin{aligned}(u+h)^{-1} &= \left(\mathrm{Id} - u^{-1}h + \|h\|\frac{(u^{-1}h)^2}{\|h\|}(\mathrm{Id}+u^{-1}h)^{-1}\right)u^{-1}\\ &= u^{-1} - u^{-1}hu^{-1} + \|h\|\frac{(u^{-1}h)^2}{\|h\|}(\mathrm{Id}+u^{-1}h)^{-1}u^{-1}\end{aligned}$$

donc on en déduit que f est différentiable au point u et $\mathrm{D}_u f(h) = -u^{-1}hu^{-1}$.

3.2 Remarques

1. Si $p = q = 1$, on retrouve bien la définition standard: $f(a+h) = f(a) + f'(a)h + h\varepsilon(h)$. Plus généralement, si $f : \mathbb{R} \to \mathrm{F}$ est différentiable au point a, alors la dérivée est de la forme $h \in \mathbb{R} \mapsto hv$ où $v \in \mathrm{F}$ est un vecteur. Du coup, on peut identifier $\mathrm{D}_a f$ avec v.

2. La différentiabilité d'une application et la différentielle ne dépend pas des normes choisies car elles sont toutes équivalentes: si $\|\cdot\|_{\mathrm{E}}$ et $\|\cdot\|'_{\mathrm{E}}$ (resp. $\|\cdot\|_{\mathrm{F}}$ et $\|\cdot\|'_{\mathrm{F}}$) sont des normes sur E (resp. sur F), alors il existe $\mathrm{C} \geq 1$ telle que $(1/\mathrm{C})\|\cdot\|_{\mathrm{E}} \leq \|\cdot\|'_{\mathrm{E}} \leq \mathrm{C}\|\cdot\|_{\mathrm{E}}$ et $(1/\mathrm{C})\|\cdot\|_{\mathrm{F}} \leq \|\cdot\|'_{\mathrm{F}} \leq \mathrm{C}\|\cdot\|_{\mathrm{F}}$. Donc, si $f : (\mathrm{U}, \|\cdot\|_{\mathrm{E}}) \to (\mathrm{F}, \|\cdot\|_{\mathrm{F}})$ est différentiable alors $f : (\mathrm{U}, \|\cdot\|'_{\mathrm{E}}) \to (\mathrm{F}, \|\cdot\|'_{\mathrm{F}})$ est aussi différentiable. En effet, on écrit $f(a+h) = f(a) + \mathrm{D}_a f(h) + \|h\|_{\mathrm{E}}\varepsilon_h(h)$ avec $\lim_{h\to 0}\|\varepsilon(h)\|_{\mathrm{F}} = 0$ et on doit vérifier

$$\lim_{h\to 0}\left\|\frac{\|h\|_{\mathrm{E}}}{\|h\|'_{\mathrm{E}}}\varepsilon(h)\right\|'_{\mathrm{F}} = 0, \text{ qui découle de } \left\|\frac{\|h\|_{\mathrm{E}}}{\|h\|'_{\mathrm{E}}}\varepsilon(h)\right\|'_{\mathrm{F}} \leq \mathrm{C}\|\varepsilon(h)\|'_{\mathrm{F}} \leq \mathrm{C}^2\|\varepsilon(h)\|_{\mathrm{F}}.$$

3. On a l'interprétation géométrique suivante: Si $\mathrm{E} = \mathbb{R}^p$ et $\mathrm{F} = \mathbb{R}^q$, on représente le graphe Γ de f dans $(\mathbb{R}^{p+q}, \|\cdot\|_\infty)$ par l'ensemble des couples $(x, f(x))$. Si f est différentiable au point a, on note $\mathrm{T}_a\Gamma$ l'espace affine décrit par $\{(x, f(a) + \mathrm{D}_a f(x-a)),\ x \in \mathbb{R}^p\}$. Cet espace est tangent au point $(a, f(a))$ à Γ au sens suivant:

$$\lim_{x\to a}\frac{d((x,f(x)), \mathrm{T}_a\Gamma)}{d((x,f(x)),(a,f(a)))} = 0\,.$$

En effet,

$$d((x,f(x)),\mathrm{T}_a\Gamma) \leq d((x,f(x)),(x,f(a)+\mathrm{D}_af(x-a))) \leq \|f(x)-f(a)-\mathrm{D}_af(x-a)\| \leq \|x-a\|\|\varepsilon(x-a)\|$$

et $d((x,f(x)),(a,f(a))) = \max\{\|x-a\|, \|f(x)-f(a)\|\} \geq \|x-a\|$, donc

$$\frac{d((x,f(x)),\mathrm{T}_a\Gamma)}{d((x,f(x)),(a,f(a)))} \leq \frac{\|x-a\|\|\varepsilon(x-a)\|}{\|x-a\|} \leq \|\varepsilon(x-a)\| .$$

4. La notion de dérivabilité et celle de dérivée au point a, comme la notion de limite au point a dont elles sont issues, sont des notions *locales* (il en est de même de la notion de continuité au point a). *Localité* signifie que si deux fonctions f et g coïncident sur un voisinage de a et que f est différentiable en a, alors g aussi et $\mathrm{D}_af = \mathrm{D}_ag$.

3.3 Propriétés

On donne des outils pour vérifier qu'une transformation est différentiable et pour donner des méthodes générales de calcul. On établit ensuite deux propriétés simples découlant de la différentiabilité.

3.3.1 Propriétés de stabilité

Proposition 3.4. — *Soient $\lambda, \mu \in \mathbb{R}$. Si $f, g : \mathrm{U} \to \mathrm{F}$ sont différentiables au point a, alors $\lambda f + \mu g$ est aussi différentiable au point a et $\mathrm{D}_a(\lambda f + \mu g) = \lambda \mathrm{D}_af + \mu \mathrm{D}_ag$. Si f et g sont de classes $\mathcal{C}^1$, alors $\lambda f + \mu g$ aussi.*

Démonstration. — On écrit

$$\begin{aligned}
(\lambda f + \mu g)(a+h) &= \lambda f(a+h) + \mu g(a+h) \\
&= \lambda f(a) + \lambda \mathrm{D}_af(h) + \lambda\|h\|\varepsilon_f(h) + \mu g(a) + \mu \mathrm{D}_ag(h) + \mu\|h\|\varepsilon_g(h) \\
&= (\lambda f + \mu g)(a) + (\lambda \mathrm{D}_af + \mu \mathrm{D}_ag)(h) + \|h\|(\lambda\varepsilon_f(h) + \mu\varepsilon_g(h)) .
\end{aligned}$$

∎

Proposition 3.5. — *Soient $\mathrm{E}, \mathrm{F}, \mathrm{G}$ trois evn, $\mathrm{U} \subset \mathrm{E}$ et $\mathrm{V} \subset \mathrm{F}$ des ouverts. Si $f : \mathrm{U} \to \mathrm{F}$ et $g : \mathrm{V} \to \mathrm{G}$ sont des applications différentiables au point $a \in \mathrm{U}$ et au point $f(a) \in \mathrm{V}$ alors $g \circ f$ est différentiable et $\mathrm{D}_a(g \circ f) = \mathrm{D}_{f(a)}g \circ \mathrm{D}_af$. Si f, g sont de classes $\mathcal{C}^1$ alors $g \circ f$ est aussi de classe $\mathcal{C}^1$ sur $\mathrm{U} \cap f^{-1}(\mathrm{V})$.*

Démonstration. — On écrit

$$f(a+h) = f(a) + (\mathrm{D}_af(h) + \|h\|\varepsilon_f(h))$$

de sorte que

$$\begin{aligned}
(g \circ f)(a+h) &= (g \circ f)(a) + \mathrm{D}_{f(a)}g(\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)) + \|\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)\|\varepsilon_g(\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)) \\
&= (g \circ f)(a) + (\mathrm{D}_{f(a)}g \circ \mathrm{D}_af)(h) + \|h\|\mathrm{D}_{f(a)}g(\varepsilon_f(h)) + \|\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)\|\varepsilon_g(\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)) .
\end{aligned}$$

On vérifie que $h \mapsto (\mathrm{D}_{f(a)}g \circ \mathrm{D}_af)(h)$ est dans $\mathcal{L}(\mathrm{E}, \mathrm{G})$ et

$$\lim_{h \to 0} \frac{\|h\|\mathrm{D}_{f(a)}g(\varepsilon_f(h)) + \|\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)\|\varepsilon_g(\mathrm{D}_af(h) + \|h\|\varepsilon_f(h))}{\|h\|} = 0$$

en utilisant le fait que $\|\mathrm{D}_af(h)\| \leq \|\mathrm{D}_af\| \cdot \|h\|$. En effet, on en déduit que

$$\begin{cases}
\|\mathrm{D}_{f(a)}g(\varepsilon_f(h))\| \leq \|\mathrm{D}_{f(a)}g\| \cdot \|\varepsilon_f(h)\| \to 0 \text{ quand } h \to 0 , \\
\lim_{h \to 0} \varepsilon_g(\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)) = 0 \\
\|\mathrm{D}_af(h) + \|h\|\varepsilon_f(h)\| \leq (\|\mathrm{D}_af\| + \|\varepsilon_f(h)\|) \cdot \|h\| .
\end{cases}$$

∎

Proposition 3.6. — *Soient* $\mathrm{E}, \mathrm{F}_1, \ldots, \mathrm{F}_k$ *des evn et* $f : \mathrm{U} \to \mathrm{F}_1 \times \ldots \times \mathrm{F}_k$ *une fonction définie par* $x \mapsto f(x) = (f_1(x), \ldots, f_k(x))$. *Soit* $a \in \mathrm{U}$; *la fonction* f *est différentiable au point* a *si et seulement si toutes les composantes* f_j *sont différentiables au point* a *et*

$$\mathrm{D}_a f(h) = \big(\mathrm{D}_a f_1(h), \ldots, \mathrm{D}_a f_k(h)\big)\ .$$

De même, f *est de classe* $\mathcal{C}^1$ *si et seulement si tous les* f_j *le sont.*

En particulier, une fonction $f : \mathrm{U} \to \mathbb{R}^n$ *est différentiable au point* $a \in \mathrm{U}$ *si et seulement si chaque composante est différentiable.*

Remarque 3.7. — *Cette proposition nous dit que les problèmes de différentiabilité d'une fonction se ramène à ceux de fonctions à valeurs réelles.*

Démonstration. — Comme les projections $\pi_j : \mathrm{F}_1 \times \ldots \times \mathrm{F}_k \to \mathrm{F}_j$ sont linéaires, elles sont différentiables et $\mathrm{D}\pi_j = \pi_j$. Or $f_j = \pi_j \circ f$ donc si f st différentiable au point a, il est en de même pour les composantes et $\mathrm{D}_a f_j = \pi_j \circ \mathrm{D}_a f$.

Réciproquement, on suppose chaque composante différentiable. On écrit

$$\begin{aligned} f(a+h) &= \big(f_1(a+h), \ldots, f_k(a+h)\big) \\ &= f(a) + \big(\mathrm{D}_a f_1(h) + \|h\|\varepsilon_1(h), \ldots, \mathrm{D}_a f_k(h) + \|h\|\varepsilon_k(h)\big) \\ &= f(a) + \big(\mathrm{D}_a f_1(h), \ldots, \mathrm{D}_a f_k(h)\big) + \|h\| \big(\varepsilon_1(h), \ldots, \varepsilon_k(h)\big)\ . \end{aligned}$$

■

Proposition 3.8. — *On suppose que* $f : \mathrm{U} \to \mathrm{F}$ *est un homéomorphisme sur* $f(\mathrm{U})$, *que* f *est différentiable et que* $\mathrm{D}_a f$ *est un isomorphisme. Alors* f^{-1} *est différentiable au point* $f(a)$ *et* $\mathrm{D}_{f(a)} f^{-1} = (\mathrm{D}_a f)^{-1}$.

Démonstration. — On écrit $g = f^{-1}$ et $b = f(a)$. On cherche à estimer $g(b+h) - g(b)$. La différentiabilité de f nous dit

$$\begin{aligned} h = (b+h) - b &= f(g(b+h)) - f(a) \\ &= \mathrm{D}_a f(g(b+h) - g(b)) + \|g(b+h) - a\|\varepsilon(g(b+h) - a)\,. \end{aligned}$$

Du coup, on a

$$\mathrm{D}_a f(g(b+h) - g(b)) = h - \|g(b+h) - a\|\varepsilon(g(b+h) - a)\,.$$

Nous obtenons en appliquant $(\mathrm{D}_a f)^{-1}$

$$g(b+h) - g(b) = (\mathrm{D}_a f)^{-1}(h) - \|g(b+h) - a\|(\mathrm{D}_a f)^{-1}(\varepsilon(g(b+h) - a))\,. \tag{3.1}$$

Donc, pour conclure, il suffit de montrer que

$$\lim_{h \to 0} \frac{\|g(b+h) - a\|(\mathrm{D}_a f)^{-1}(\varepsilon(g(b+h) - a))}{\|h\|} = 0\,.$$

A partir de (3.1), on obtient

$$\begin{aligned} \|g(b+h) - a\| &= \|(\mathrm{D}_a f)^{-1}(h - \|g(b+h) - a\|\varepsilon(g(b+h) - a))\| \\ &\leq \|(\mathrm{D}_a f)^{-1}\| \cdot (\|h\| + \|g(b+h) - a\| \cdot \|\varepsilon(g(b+h) - a)\|)\,. \end{aligned}$$

Or, si h est assez petit, alors la continuité de g nous permet d'affirmer $\|\varepsilon(g(b+h) - a)\| \leq 1/(2\|(\mathrm{D}_a f)^{-1}\|)$ donc

$$\|g(b+h) - a\| \leq 2\|(\mathrm{D}_a f)^{-1}\| \cdot \|h\|\,.$$

Du coup,

$$\lim_{h \to 0} \frac{\|g(b+h) - a\|\varepsilon(g(b+h) - a)}{\|h\|} = \lim_{h \to 0} \varepsilon(g(b+h) - g(b)) = 0\,.$$

Ceci montre que (3.1) nous donne la différentielle de f^{-1}. ■

Proposition 3.9. — *Soient $f, g : \mathrm{U} \to \mathbb{R}$ deux applications différentiables.*

1. *La fonction fg est différentiable au point a et $\mathrm{D}_a(fg) = g(a)\mathrm{D}_a f + f(a)\mathrm{D}_a g$.*

2. *Si $g(a) \neq 0$, alors f/g est différentiable au point a et $\mathrm{D}_a(f/g) = (1/g(a)^2)[g(a)\mathrm{D}_a f - f(a)\mathrm{D}_a g]$.*

Démonstration. — On écrit

$$\begin{aligned}(fg)(a+h) &= (f(a) + \mathrm{D}_a f(h) + \|h\|\varepsilon_f(h))(g(a) + \mathrm{D}_a g(h) + \|h\|\varepsilon_g(h)) \\ &= (fg)(a) + f(a)\mathrm{D}_a g(h) + g(a)\mathrm{D}_a f(h) + \|h\|[(f(a) + \mathrm{D}_a f(h))\varepsilon_g(h)) + (g(a) + \mathrm{D}_a g(h))\varepsilon_f(h))]\,.\end{aligned}$$

Pour f/g, on applique la formule du produit et la formule de composition à $x \mapsto (1/x)$ et g. ■

Exemple 3.10. — *Les fonctions polynomiales et les fractions rationnelles sont différentiables là où elles sont définies. En particulier, le déterminant d'une matrice carrée est différentiable.*

3.3.2 Autres propriétés élémentaires

Proposition 3.11. — *Une application $f : \mathrm{U} \to \mathrm{F}$ qui est différentiable au point a est aussi continue au point a.*

Démonstration. — En effet, on a

$$\|f(x) - f(a)\| \leq (\|\mathrm{D}_a f\| + \varepsilon(x-a)) \cdot \|x - a\|\,.$$

■

Proposition 3.12. — *Si $f : \mathrm{U} \to \mathbb{R}$ admet un extremum local au point $a \in \mathrm{U}$ et si f est différentiable au point a alors $\mathrm{D}_a f = 0$.*

Démonstration. — Sans perte de généralité, on peut supposer que a est un maximum local. Du coup, il existe un voisinage $\mathrm{V} \subset \mathrm{U}$ de a tel que $f(x) \leq f(a)$ pour tout $x \in \mathrm{V}$. Si $(a \pm h) \in \mathrm{V}$, $h \neq 0$, alors, comme $f(a+h) = f(a) + \mathrm{D}_a f(h) + \|h\|\varepsilon(h)$, on en déduit $\mathrm{D}_a f(h) + \|h\|\varepsilon(h) \leq 0$, impliquant ainsi

$$\limsup_{h \to 0} \frac{\mathrm{D}_a f(h)}{\|h\|} \leq 0\,.$$

En considérant $x = a - h$, on obtient de même $\mathrm{D}_a f(h) \geq \|h\|\varepsilon(-h)$, impliquant donc

$$\liminf_{h \to 0} \frac{\mathrm{D}_a f(h)}{\|h\|} \geq 0 \quad \text{puis} \quad \lim_{h \to 0} \frac{\mathrm{D}_a f(h)}{\|h\|} = 0\,.$$

Par conséquent, $\mathrm{D}_a f = 0$. ■

3.4 Dérivées directionnelles, dérivées partielles et matrice jacobienne

Soit U un ouvert d'un evn E et $f : \mathrm{U} \to \mathrm{F}$ une application à valeurs dans un evn F. Si $v \in \mathrm{E}$, alors la *dérivée au point $a \in \mathrm{U}$ dans la direction v* est définie, lorsqu'elle existe, par la limite

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

et on la note $\partial_v f(a)$, ou $\dfrac{\partial f}{\partial v}(a)$. C'est un vecteur de F.

Si on identifie E à $\mathbb{R}^p$ et que l'on note $\mathcal{B} = (\mathrm{e}_1, \ldots, \mathrm{e}_p)$ la base canonique, alors la dérivée dans la direction e_j s'appelle la *j-ième dérivée partielle de f*. On la note $\partial_j f(a)$, $f'_{x_j}(a)$ ou $\dfrac{\partial f}{\partial x_j}(a)$.

Lemme 3.13. — *Si f est différentiable au point a, alors toutes les dérivées directionnelles existent au point a et on a $\partial_v f(a) = \mathrm{D}_a f(v)$.*

DÉMONSTRATION. — En effet, on a $f(a+tv) - f(a) = t\mathrm{D}_a f(v) + |t| \cdot \|v\| \varepsilon(tv)$ donc

$$\lim_{t\to 0} \frac{f(a+tv)-f(a)}{t} = \lim_{t\to 0} \left(\mathrm{D}_a f(v) + \frac{|t|\cdot\|v\|}{t}\varepsilon(tv) \right) = \mathrm{D}_a f(v)\,.$$

■

Remarque 3.14. — *L'existence de dérivées partielles, voire de dérivées dans toutes les directions, n'impliquent pas la différentiabilité, ni même la continuité. En effet, prenons par exemple $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3/y$ si $y \neq 0$ et $f(x,0) = 0$ sinon. On a $\partial_1 f(0,0) = \partial_2 f(0,0) = 0$ car $f(x,0) = f(0,y) \equiv 0$ et plus généralement, pour $a, b \neq 0$,*

$$\partial_{(a,b)} f(0,0) = \lim_{t\to 0} \frac{(ta)^3/(tb)}{t} = 0\,.$$

Or $f(x,x^4) = 1/x$ qui n'admet pas de limite en 0.

Nous verrons plus loin que l'existence de dérivées partielles continues au voisinage du point a implique la différentiabilité au point a.

Identifions maintenant F à $\mathbb{R}^q$ et considérons les composantes de $f = (f_1, \dots, f_q)$. Si f est différentiable au point $a \in \mathbb{R}^p$, alors la *matrice jacobienne* de f au point a est la matrice de terme général $(\partial_j f_i(a))_{1\le i\le q, 1\le j\le p} \in \mathcal{M}_{q,p}(\mathbb{R})$. Il s'agit donc de la représentation matricielle de $\mathrm{D}_a f$ dans les bases canoniques. Le *jacobien de* f est le déterminant de la matrice jacobienne.

En particulier, si $f : \mathrm{U} \to \mathbb{R}$ alors

$$\mathrm{D}_a f(h_1, \dots, h_p) = \sum_{j=1}^{p} \partial_j f(a) h_j\,.$$

De même, si V est un ouvert de $\mathbb{R}^q$ contenant $f(a)$ et $g : \mathrm{V} \to \mathbb{R}^m$ est une application différentiable au point $f(a)$ alors la formule de composition conduit à, pour $1 \le i \le m$ et $1 \le j \le p$,

$$\partial_j (g\circ f)_i(a) = \sum_{k=1}^{q} \partial_k g_i(f(a)) \partial_j f_k(a) = \sum_{k=1}^{q} \frac{\partial g_i}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a)\,.$$

Plus généralement, si E se décompose en $\mathrm{E}_1 \times \ldots \times \mathrm{E}_k$, alors on peut considérer les différentielles partielles par rapport à chaque sous-espace E_j. Plus précisément, on se fixe $a = (a_1, \dots, a_k) \in \mathrm{U}$ et on considère $\iota_j : \mathrm{E}_j \to \mathrm{E}$ défini par $\iota_j(x) = (a_1, \dots, a_{j-1}, x, a_{j+1}, \dots, a_k)$ avec les conventions d'usage si $j \in \{1, k\}$. On définit $\partial_j f_a = \mathrm{D}_a(f \circ \iota_j)$ lorsqu'elle est bien définie.

Proposition 3.15. — *Si f est différentiable au point a, alors f admet des dérivées partielles par rapport à chaque sous-espace et*

$$\mathrm{D}_a f : (h_1, \dots, h_k) \mapsto \sum_{j=1}^{k} \partial_j f_a(h_j)\,.$$

DÉMONSTRATION. — Les applications ι_j sont affines, donc $f\circ\iota_j$ est différentiable pour tout $j \in \{1, \dots, k\}$ et, pour $h_j \in \mathrm{E}_j$,

$$\mathrm{D}_a(f\circ\iota_j)(h_j) = \mathrm{D}_a f \begin{pmatrix} 0 \\ \vdots \\ 0 \\ h_j \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{impliquant ainsi} \quad \mathrm{D}_a f \begin{pmatrix} h_1 \\ \vdots \\ h_k \end{pmatrix} = \sum_{j=1}^{k} \partial_j f_a(h_j)\,.$$

■

3.5 Gradient

Dans ce paragraphe, on suppose que E est un espace euclidien, U un ouvert de E. On rappelle que, dans un espace euclidien, pour toute forme linéaire $L \in E^*$, il existe un vecteur $v \in E$ tel que $L(x) = \langle v, x\rangle$ pour tout $x \in E$. Si $f : U \to \mathbb{R}$ est une application différentiable en un point $a \in U$, on définit le *gradient* de f au point a comme le vecteur $grad_a f = \nabla f(a)$ qui vérifie, pour tout $x \in E$,

$$D_a f(h) = \langle \nabla f(a), h\rangle\,.$$

Si $E = \mathbb{R}^p$, alors

$$\nabla f(a) = \begin{pmatrix} \partial_1 f(a) \\ \vdots \\ \partial_p f(a) \end{pmatrix}.$$

On a les interprétations suivantes du gradient.

- Si on considère le graphe $\Gamma = \{(x, f(x)) \in U \times \mathbb{R},\ x \in U\}$ alors $\begin{pmatrix} \nabla f(a) \\ (-1) \end{pmatrix}$ est orthogonal à $T_a\Gamma$ car
$$\left\langle \begin{pmatrix} \nabla f(a) \\ (-1) \end{pmatrix}, \begin{pmatrix} x \\ D_a f(x) \end{pmatrix} \right\rangle = \langle \nabla f(a), x\rangle - D_a f(x) = 0\,.$$
- Le gradient $\nabla f(a)$ est orthogonal à l'ensemble de niveau $\mathcal{N} = \{x \in E,\ f(x) = f(a)\}$. En effet, si $x \in \mathcal{N}$ est proche de a, alors $f(x) = f(a) + D_a f(x-a) + \|x-a\|\varepsilon(x-a)$ donc $D_a f(x-a) = -\|x-a\|\varepsilon(x-a)$ et
$$\lim_{x\to a} \left\langle \nabla f(a), \frac{x-a}{\|x-a\|} \right\rangle = 0\,.$$
- Le gradient $\nabla f(a)$ pointe vers la plus grande pente de f en a au sens suivant: on a $f(x) = f(a) + \langle \nabla f(a), x\rangle + \|x-a\|\varepsilon(x-a)$. En première approximation, $f(x)$ est maximale si $\langle \nabla f(a), x\rangle$ l'est. Or le produit scalaire est maximal si x est positivement colinéaire à $\nabla f(a)$.

4 Théorème de la moyenne et applications

On rappelle les énoncés pour les fonctions numériques avant de passer aux versions vectorielles.

Théorème 4.1 (Rolle). — *Soient $a, b \in \mathbb{R}$ tels que $a < b$ et $f : [a, b] \to \mathbb{R}$ une application continue, dérivable sur $]a, b[$. Si $f(a) = f(b)$ alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.*

Démonstration. — Comme $[a, b]$ est compact, la fonction f atteint ses bornes. Si elle est constante, n'importe quel point $c \in]a, b[$ convient; sinon, comme $f(a) = f(b)$ un des extrema de f, que l'on notera c, n'est ni a ni b. Donc $c \in]a, b[$ et $f'(c) = 0$ d'après la proposition 3.12. ■

Théorème 4.2 (des accroissements finis). — *Soient $a, b \in \mathbb{R}$ tels que $a < b$ et $f : [a, b] \to \mathbb{R}$ une application continue, dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b-a)$.*

Démonstration. — On applique le théorème de Rolle à la fonction auxiliaire $g : [a, b] \to \mathbb{R}$ définie par $g(t) = f(t)(b-a) - (f(b) - f(a))t$. ■

Théorème 4.3 (de la moyenne). — *Soient $a, b \in \mathbb{R}$ tels que $a < b$, F un evn, $f : [a, b] \to F$ et $g : [a, b] \to \mathbb{R}$ deux applications continues, dérivables sur $]a, b[$. Si, pour tout $t \in]a, b[$, on a $\|f'(t)\| \le g'(t)$ alors*
$$\|f(b) - f(a)\| \le g(b) - g(a)\,.$$

Démonstration. — Il suffit de montrer que, pour tout $\varepsilon > 0$, on a $\|f(b) - f(a)\| \le g(b) - g(a) + \varepsilon(b-a)$. Une fois cette assertion établie, comme ε est arbitraire, on obtiendra $\|f(b) - f(a)\| \le g(b) - g(a)$ et donc le théorème.

A cet effet, on considère $\mathrm{A} = \{t \in [a,b] \ \|f(t) - f(a)\| \leq g(t) - g(a) + \varepsilon(t-a)\}$ pour $\varepsilon > 0$ fixé et on remarque que $\mathrm{A} \neq \varnothing$ puisque $a \in \mathrm{A}$ (car $0 \leq 0$). Notons $s = \sup \mathrm{A}$. Par définition, il existe une suite croissante (s_n) contenue dans A qui tend vers s. Pour tout n, on a donc

$$\|f(s_n) - f(a)\| \leq g(s_n) - g(a) + \varepsilon(s_n - a)\,.$$

Comme f et g sont continues, on obtient à la limite $\|f(s) - f(a)\| \leq g(s) - g(a) + \varepsilon(s-a)$ montrant ainsi $s \in \mathrm{A}$.

Pour conclure, on doit montrer $s = b$. Supposons donc $s < b$. Par définition de la dérivée au point s, on peut trouver $\delta > 0$ telle que $s + \delta < b$ et, pour $t \in]s, s+\delta]$, on a

$$\|f(t) - f(s)\| \leq \|\mathrm{D}_s f\| \cdot (t-s) + (t-s)\frac{\varepsilon}{2}$$

et

$$g'(s)(t-s) \leq g(t) - g(s) + \frac{\varepsilon}{2}(t-s)$$

de sorte que

$$\begin{aligned} \|f(t) - f(s)\| &\leq \|\mathrm{D}_s f\| \cdot (t-s) + (t-s)\frac{\varepsilon}{2} \\ &\leq g'(s) \cdot (t-s) + (t-s)\frac{\varepsilon}{2} \\ &\leq g(t) - g(s) + \varepsilon(t-s)\,. \end{aligned}$$

Or comme $s \in \mathrm{A}$, on a, pour $t \in]s, s+\delta[$,

$$\begin{aligned} \|f(t) - f(a)\| &\leq \|f(t) - f(s)\| + \|f(s) - f(a)\| \\ &\leq g(t) - g(s) + \varepsilon(t-s) + g(s) - g(a) + \varepsilon(s-a) \\ &\leq g(t) - g(a) + \varepsilon(t-a) \end{aligned}$$

ce qui montre $t \in \mathrm{A}$ et contredit le fait que $s = \sup \mathrm{A}$. Du coup, $s = b$. ■

Soit E un evn. Si $a, b \in \mathrm{E}$, le *segment* $[a,b]$ est $\{(1-t)a + bt,\ t \in [0,1]\}$. Un sous-ensemble X de E est *convexe* si, pour tous $x, y \in \mathrm{X}$, le segment $[x,y]$ est contenu dans X. Par exemple une boule, ouverte ou fermée, est convexe.

Théorème 4.4 (inégalité des accroissements finis). — *Soient $k \geq 0$, E, F deux evn, U un ouvert convexe et $f : \mathrm{U} \to \mathrm{F}$ une application différentiable. Si, pour tout $a \in \mathrm{U}$, on a $\|\mathrm{D}_a f\| \leq k$, alors, pour tous $x, y \in \mathrm{U}$, on a*

$$\|f(x) - f(y)\| \leq k\|x - y\|\,.$$

Démonstration. — On applique le théorème de la moyenne à l'application auxiliaire $g : [0,1] \to \mathrm{F}$ définie par $g(t) = f((1-t)x + ty)$: on a $g'(t) = \mathrm{D}_{(1-t)x+ty} f(y-x)$ donc $\|g'(t)\| \leq k\|y-x\|$ et

$$\|g(1) - g(0)\| = \|f(y) - f(x)\| \leq k\|y - x\|\,.$$

■

4.1 Applications de dérivée nulle

Un ouvert U est *connexe par arcs* si pour tous $x, y \in \mathrm{U}$, il existe une fonction continue $c : [0,1] \to \mathrm{U}$, autrement dit un *chemin*, telle que $c(0) = x$ et $c(1) = y$. Un ouvert convexe est en particulier connexe par arcs.

Proposition 4.5. — *Soient $\mathrm{U} \subset \mathrm{E}$ un ouvert connexe par arcs et $f : \mathrm{U} \to \mathrm{F}$ une application différentiable. Si, pour tout $a \in \mathrm{U}$, on a $\mathrm{D}_a f = 0$ alors f est constante.*

La fonction qui prend la valeur 0 sur $]0,1[$ et 1 sur $]2,3[$ est différentiable, de dérivée identiquement nulle, mais non constante. Ce n'est pas une contradiction puisque $]0,1[\cup]2,3[$ n'est pas connexe par arcs.

Démonstration. — Soient $x, y \in \mathrm{U}$ et $c : [0,1] \to \mathrm{U}$ un chemin reliant x à y. Notons $\mathrm{A} = \{t \in [0,1],\ (f \circ c)(t) = f(x)\}$. Cet ensemble est non vide car il contient $\{0\}$. Notons $s = \sup \mathrm{A}$ et montrons que $s = 1$ par l'absurde. Si $s < 1$, notons $a = c(s)$ et $r > 0$ de sorte que $\mathrm{B}(a, r) \subset \mathrm{U}$. La continuité de c nous donne l'existence de $\delta > 0$, $\delta \leq \min\{s, 1-s\}$ telle que $c(]s-\delta, s+\delta[) \subset \mathrm{B}(a, r)$.

L'inégalité des accroissements finis affirme que, pour $b \in \mathrm{B}(a, r)$, on a $\|f(b) - f(a)\| \leq \sup \|\mathrm{D}_x f\| \|b - a\| = 0$. Par définition de s, on peut choisir $t \in \mathrm{A} \cap]s-\delta, s]$ de sorte que $f(a) = f(c(s)) = f(x)$ car $t \in \mathrm{A}$; par ailleurs, si $t = s + \delta/2$, alors $f(c(t + \delta/2)) = f(a) = f(x)$, ce qui contredirait la définition de s. Donc $s = 1$ et $f(x) = f(y)$. ■

4.2 Caractérisation des fonctions de classe $\mathcal{C}^1$

On fournit une sorte de réciproque au lemme 3.13 qui donne un moyen assez simple d'établir en pratique la différentiabilité de fonctions.

Théorème 4.6. — *Une application $f : \mathrm{U}(\subset \mathbb{R}^n) \to \mathrm{F}$ est de classe $\mathcal{C}^1$ si et seulement si toutes les dérivées partielles de f existent et sont continues. Dans ce cas, on a*

$$\mathrm{D}_a f(h_1, \ldots, h_n) = \sum_{j=1}^{n} h_j \partial_j f(a)\,.$$

Dans cette écriture, on rappelle que $\partial_j f(a)$ désigne chaque fois des vecteurs de F.

Démonstration. — Le lemme 3.13 implique que si f est de classe $\mathcal{C}^1$, alors les dérivées partielles de f existent et sont continues, comme composées des fonctions $\mathrm{D}f$ et des évaluations sur les vecteurs de base $\mathrm{L} \in \mathcal{L}(\mathbb{R}^n, \mathrm{F}) \mapsto \mathrm{L}(\mathrm{e}_j)$. Passons maintenant à la réciproque.

On munit $\mathbb{R}^n$ de la norme $\|\cdot\|_\infty$. Montrons la différentiabilité de f au point $a \in \mathrm{U}$. Notons que l'on a déjà un candidat pour la différentielle. On suppose $\mathrm{B}(a, \mathrm{R}) \subset \mathrm{U}$. On considère $h_1, \ldots, h_n \in \mathbb{R}$ avec $|h_j| \leq \mathrm{R}$ pour chaque $j \in \{1, \ldots, n\}$ et on notera $h = (h_1, \ldots, h_n)$.

Afin d'exploiter les hypothèses, on introduit les points suivants

$$\begin{cases} x_j = (a_1 + h_1, \ldots, a_{j-1} + h_{j-1}, a_j, \ldots, a_n) & \text{pour } j \in \{1, \ldots, n\} \\ x_{n+1} = a + h & \\ u_j = (0, \ldots, 0, h_j, 0, \ldots, 0) = h_j \mathrm{e}_j & \text{pour } j \in \{1, \ldots, n\}\,. \end{cases}$$

Il vient

$$f(a+h) - f(a) = \sum_{j=1}^{n} f(x_{j+1}) - f(x_j) = \sum_{j=1}^{n} f(x_j + u_j) - f(x_j)$$

On obtient ainsi

$$\begin{aligned} \left\| f(a+h) - f(a) - \sum_{j=1}^{n} h_j \partial_j f(a) \right\| &= \left\| \sum_{j=1}^{n} [f(x_{j+1}) - f(x_j) - h_j \partial_j f(x_j)] + \sum_{j=1}^{n} (\partial_j f(x_j) - \partial_j f(a)) h_j \right\| \\ &\leq \sum_{j=1}^{n} \|f(x_{j+1}) - f(x_j) - h_j \partial_j f(x_j)\| + \|h\|_\infty \sum_{j=1}^{n} \|\partial_j f(x_j) - \partial_j f(a)\|\,. \end{aligned}$$

On introduit maintenant les fonctions auxiliaires

$$g_j : t \in [0,1] \mapsto f(x_j + t u_j) - t h_j \partial_j f(x_j)$$

auxquelles on souhaite appliquer le théorème de la moyenne. Vérifions donc que g_j est différentiable. Le second terme est différentiable car linéaire en t. Pour le premier, posons $y_j = x_j + t u_j$ et prenons $\alpha \in \mathbb{R}$. En se souvenant que $u_j = h_j \mathrm{e}_j$, on obtient

$$\lim_{\alpha \to 0} \frac{f(y_j + \alpha u_j) - f(y_j)}{\alpha} = \lim_{\alpha \to 0} \frac{f(y_j + (\alpha h_j)\mathrm{e}_j) - f(y_j)}{\alpha h_j} h_j = \partial_j f(y_j) h_j\,.$$

Du coup, g_j est bien dérivable et

$$g_j'(t) = \partial_j f(x_j + tu_j)h_j - h_j\partial_j f(x_j)\,.$$

Le théorème de la moyenne donne $\|g_j(1) - g_j(0)\| \leq \sup_{t\in[0,1]} \|g_j'(t)\|$ soit

$$\|f(x_{j+1}) - f(x_j) - h_j\partial_j f(x_j)\| \leq \sup_{t\in[0,1]} \|\partial_j f(x_j + tu_j) - \partial_j f(x_j)\| \cdot |h_j|\,.$$

On se fixe $\varepsilon > 0$. Par continuité des dérivées partielles, on peut trouver $r \in]0, \mathrm{R}[$ de sorte que, pour tous $x, y \in \mathrm{B}(a, r)$, on ait $\|\partial_j f(x) - \partial_j f(y)\| \leq \varepsilon/(2n)$. Du coup, il vient

$$\left\| f(a+h) - f(a) - \sum_{j=1}^{n} h_j\partial_j f(a) \right\| \leq \sum_{j=1}^{n} (\varepsilon/2n)|h_j| + \|h\|_\infty \sum_{j=1}^{n} \varepsilon/(2n) \leq \|h\|_\infty \varepsilon\,.$$

Ceci montre la différentiabilité au point a et identifie sa différentielle. Il reste à vérifier la continuité de $\mathrm{D}f$. Prenons $a, b \in \mathrm{U}$; pour tout $h \in \mathbb{R}^n$ tel que $\|h\|_\infty$, on a

$$\|(\mathrm{D}_a f - \mathrm{D}_b f)(h)\|_{\mathrm{F}} \leq \sum_{j=1}^{n} |h_j| \cdot \|\partial_j f(a) - \partial_j f(b)\|_{\mathrm{F}} \leq \sum_{j=1}^{n} \|\partial_j f(a) - \partial_j f(b)\|_{\mathrm{F}}$$

donc

$$\|\mathrm{D}_a f - \mathrm{D}_b f\| \leq \sum_{j=1}^{n} \|\partial_j f(a) - \partial_j f(b)\|_{\mathrm{F}}$$

et la continuité de $\mathrm{D}f$ découle de celle des dérivées partielles. ■

Le même argument conduit au résultat plus général suivant qui améliore la proposition 3.15.

Théorème 4.7. — *Soient* $\mathrm{E}_1, \ldots, \mathrm{E}_k, \mathrm{F}$ *des espaces de dimension finie,* $f : \mathrm{U} \to \mathrm{F}$ *une application définie sur un ouvert de* $\mathrm{E} \overset{def.}{=} \mathrm{E}_1 \times \ldots \times \mathrm{E}_k$*. L'application* f *est de classe* $\mathcal{C}^1$ *sur* U *si et seulement si* f *admet des différentielles partielles continues* $\partial_j f$*,* $j = 1, \ldots, k$*, le long de chaque sous-espace* E_j*. Dans ce cas,*

$$\mathrm{D}_a f(h) = \sum \partial_j f_a(h_j)\,.$$

4.3 Limite de suites de fonctions différentiables

On s'intéresse maintenant aux limites de fonctions différentiables.

Théorème 4.8 (d'inversion des limites et des dérivées). — *Soient* $\mathrm{U} \subset \mathrm{E}$ *un ouvert connexe par arcs et* $(f_n)_{n\geq 0}$ *une suite d'applications différentiables de* U *dans* F *telles que*

1. *il existe* $x_0 \in \mathrm{U}$ *tel que la suite* $(f_n(x_0))_{n\geq 0}$ *est convergente.*
2. *pour tout* $a \in \mathrm{U}$*, il existe* $r > 0$ *tel que la suite* $(\mathrm{D}f_n)_n$ *est uniformément convergente sur* $\mathrm{B}(a, r)$*.*

Alors la suite $(f_n)_n$ *est uniformément convergente sur chaque boule* $\mathrm{B}(a, r)$ *vers une fonction* f*, différentiable sur* U*, de différentielle la limite des différentielles* $(\mathrm{D}f_n)_n$*:*

$$\lim_n \mathrm{D}f_n = \mathrm{D}(\lim_n f_n)\,.$$

Remarque 4.9. — *Les deux conditions sont nécessaires. Prenons par exemple la suite de fonctions constantes à valeurs réelles* n*: les différentielles sont indentiquement nulles donc uniformément convergentes, mais la suite est divergente; ceci montre la nécessité de la première condition. Prenons sur* $]-1, 1[$*, la suite* $(f_n)_n$ *définies par* $x \mapsto \sqrt{x^2 + 1/(n+1)}$*. On a convergence uniforme vers* $|\cdot|$*; on peut vérifier que la suite de dérivées tend simplement vers* 1 *pour* $x > 0$*,* (-1) *pour* $x < 0$ *et* 0 *pour* $x = 0$*.*

Démonstration. — Prenons tout d'abord $a \in \mathrm{U}$ et $r > 0$ de sorte que **2.** soit vérifié. L'inégalité des accroissements finis appliquée à $f_p - f_q$ implique, pour $x, y \in \mathrm{B}(a, r)$,

$$\|(f_p(x) - f_q(x)) - (f_p(y) - f_q(y))\| \leq \sup_{z \in \mathrm{B}(a,r)} \|\mathrm{D}_z f_p - \mathrm{D}_z f_q\| \cdot \|x - y\| . \tag{4.1}$$

Comme on a convergence uniforme des dérivées, la convergence en un seul point de $\mathrm{B}(a, r)$ implique la convergence uniforme de (f_n) sur la boule. Justifions cette assertion: supposons que $(f_n(x))_n$ soit convergente. Alors $(f_n(x))_n$ est une suite de Cauchy, tout comme $(\mathrm{D}f_n)_n$ est une suite de Cauchy uniforme sur $\mathrm{B}(a, r)$. Prenons $\varepsilon > 0$, et fixons n_0 assez grand de sorte que si, $p, q \geq n_0$, alors

$$\|f_p(x) - f_q(x)\| \leq \varepsilon/2 \quad \text{ et pour tout } z \in \mathrm{B}(a, r),\ \|\mathrm{D}_z f_p - \mathrm{D}_z f_q\| \leq \varepsilon/(2r) .$$

On en déduit, avec (4.1), pour $p, q \geq n_0$ et tout $y \in \mathrm{B}(a, r)$,

$$\|f_p(y) - f_q(y)\| \leq \|f_p(x) - f_q(x)\| + \|(f_p(x) - f_q(x)) - (f_p(y) - f_q(y))\| \leq \varepsilon/2 + (\varepsilon/2r)\|x - y\| \leq \varepsilon$$

puisque $\|x - y\| \leq \|x - a\| + \|a - y\| < 2r$. Donc la suite $(f_n)_n$ est de Cauchy uniforme sur $\mathrm{B}(a, r)$, impliquant la convergence uniforme de $(f_n)_n$ sur $\mathrm{B}(a, r)$.

Soit $x \in \mathrm{U}$ et prenons un chemin $c : [0, 1] \to \mathrm{U}$ qui relie x_0 à x. Notons A l'ensemble des $t \in [0, 1]$ pour lesquels on a convergence de la suite $(f_n(c(t)))_n$ et $s = \sup \mathrm{A}$. L'analyse précédente montre que $s = 1$. En effet, cet ensemble est non vide car il contient 0. Si $s < 1$, notons $a = c(s)$ et $r > 0$ de sorte que l'on ait convergence uniforme des dérivées sur $\mathrm{B}(a, r)$. La continuité de c nous donne l'existence de $\delta > 0$, $\delta \leq \min\{s, 1 - s\}$, telle que $c(]s - \delta, s + \delta[) \subset \mathrm{B}(a, r)$. Par définition de s, il existe $t \in \mathrm{A} \cap]s - \delta, s]$. Donc on a convergence uniforme de (f_n) sur $\mathrm{B}(a, r)$, en particulier au point $t = s + \delta/2$, contredisant la définition de s. Comme $s = 1$, on trouve t assez proche de 1 pour que (4.1) implique la convergence au point x et sur la boule $\mathrm{B}(x, r_x)$ correspondante.

Notons f la limite de $(f_n)_n$ et L celle de $(\mathrm{D}_a f_n)_n$. Fixons-nous $\mathrm{B}(a, r)$ où ces convergences sont uniformes. Nous allons les exploiter pour établir la différentiabilité de f:

$$\begin{aligned}
\|f(a + h) - f(a) - \mathrm{L}(h)\| &= \|(f(a + h) - f_n(a + h)) - (f(a) - f_n(a)) \\
&\quad + (f_n(a + h) - f_n(a) - \mathrm{D}_a f_n(h)) + (\mathrm{L} - \mathrm{D}_a f_n)(h)\| \\
&\leq \|(f(a + h) - f_n(a + h)) - (f(a) - f_n(a))\| \\
&\quad + \|f_n(a + h) - f_n(a) - \mathrm{D}_a f_n(h)\| + \|\mathrm{L} - \mathrm{D}_a f_n\| \cdot \|h\| .
\end{aligned}$$

Fixons-nous $\varepsilon > 0$. Pour p, q assez grands, on a $\|\mathrm{D}_x f_p - \mathrm{D}_x f_q\| \leq \varepsilon/3$ pour tout $x \in \mathrm{B}(a, r)$. D'après (4.1), on obtient en passant à la limite $p \to \infty$ et en prenant $q = n$ assez grand,

$$\begin{cases} \|(f(a + h) - f_n(a + h)) - (f(a) - f_n(a))\| \leq \varepsilon\|h\|/3 , \\ \\ \|\mathrm{L} - \mathrm{D}_a f_n\| \cdot \|h\| \leq \varepsilon\|h\|/3 . \end{cases}$$

On fixe maintenant n assez grand, et on utilise la différentiabilité de f_n pour choisir h assez petit de sorte que

$$\|f_n(a + h) - f_n(a) - \mathrm{D}_a f_n(h)\| \leq \varepsilon\|h\|/3 .$$

En rassemblant ces trois estimées, on conclut

$$\|f(a + h) - f(a) - \mathrm{L}(h)\| \leq \varepsilon\|h\| .$$

■

En prenant des sommes partielles, on obtient

Théorème 4.10 (d'inversion des séries et des dérivées). — *Soient* $\mathrm{U} \subset \mathrm{E}$ *un ouvert connexe par arcs et* $(f_n)_{n\geq 0}$ *une suite d'applications différentiables de* U *dans* F *telles que*

1. *il existe* $x_0 \in \mathrm{U}$ *tel que la série de terme général* $(f_n(x_0))_{n\geq 0}$ *est convergente.*

2. *pour tout* $a \in \mathrm{U}$, *il existe* $r > 0$ *tel que la série de terme général* $(\mathrm{D}f_n)_n$ *est uniformément convergente sur* $\mathrm{B}(a, r)$.

Alors la série de terme général $(f_n)_n$ *est uniformément convergente sur chaque boule* $\mathrm{B}(a, r)$ *vers une fonction* f, *différentiable sur* U, *de différentielle la somme des différentielles* $(\mathrm{D}f_n)_n$:

$$\sum_n \mathrm{D}f_n = \mathrm{D}\left(\sum_n f_n\right).$$

On en déduit l'application suivante:

Proposition 4.11. — *Soit* $\sum_{n\geq 0} a_n z^n$ *une série entière de rayon de convergence* $\mathrm{R} > 0$. *La fonction* $f : \mathcal{L}(\mathbb{R}^n) \to \mathcal{L}(\mathbb{R}^n)$ *définie par* $f(\mathrm{A}) = \sum_{n\geq 0} a_n \mathrm{A}^n$ *est bien définie et différentiable, de classe* $\mathcal{C}^1$ *sur* $\mathrm{B}_{\mathcal{L}(\mathbb{R}^n)}(0, \mathrm{R})$. *De plus, on a*

$$\mathrm{D_A}f(\mathrm{H}) = \sum_{n\geq 0} a_n(\mathrm{HA}^{n-1} + \mathrm{AHA}^{n-2} + \ldots + \mathrm{A}^{n-1}\mathrm{H}).$$

Démonstration. — Si $\|\mathrm{A}\| < \mathrm{R}$ alors $\sum_{n\geq 0} \|a_n \mathrm{A}^n\| \leq \sum |a_n| \|\mathrm{A}\|^n$ qui est finie puisque $\|\mathrm{A}\| < \mathrm{R}$. Notons que $\mathrm{A} \mapsto \mathrm{A}^n$ est différentiable de dérivée $\mathrm{H} \mapsto \mathrm{HA}^{n-1} + \mathrm{AHA}^{n-2} + \ldots + \mathrm{A}^{n-1}\mathrm{H}$. Par conséquent, on obtient

$$\sum \|a_n(\mathrm{HA}^{n-1} + \mathrm{AHA}^{n-2} + \ldots + \mathrm{A}^{n-1}\mathrm{H})\| \leq \sum n|a_n| \|\mathrm{A}\|^{n-1}$$

qui est donc uniformément convergente sur $\mathrm{B}(0, \mathrm{R}')$ pour tout $\mathrm{R}' < \mathrm{R}$. Du coup le théorème précédent s'applique.

Enfin, l'expression de $\mathrm{D_A}f$ dépend continûment de A. ■

Cette proposition s'applique par exemple à

$$\begin{cases} (\mathrm{I}+\mathrm{A})^{-1} & = & \displaystyle\sum_{n\geq 0}(-1)^n \mathrm{A}^n, & \|\mathrm{A}\| < 1 \\ \exp \mathrm{A} & = & \displaystyle\sum_{n\geq 0} \frac{\mathrm{A}^n}{n!}, & \text{A quelconque,} \\ \mathrm{Log}(\mathrm{I}+\mathrm{A}) & = & \displaystyle\sum_{n\geq 1} \frac{(-1)^{n+1}\mathrm{A}^n}{n}, & \|\mathrm{A}\| < 1\,. \end{cases}$$