Calcul Différentiel Avancé

par

Emmanuel Hebey

Université de Cergy-Pontoise

Année 2017-2018

1. Applications différentiables

On inverse les priorités. On ne cherche plus à calculer la différentielle mais bien à définir la différentiabilité. C'est elle qui vient en premier, et c'est la "bonne" façon de voir les choses.

Définition

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, $a \in \Omega$ un point de Ω, et $f : \Omega \to F$ une application de Ω dans F. On dit que f est différentiable au point a s'il existe une application linéaire continue $g \in L_c(E, F)$ telle que

$$f(x) - f(a) - g(x - a) = \|x - a\|_E \varepsilon(x) \qquad (\star)$$

pour tout $x \in \Omega$, où $\varepsilon : \Omega \to F$ est telle que $\varepsilon(a) = 0$ et $\varepsilon(x) \to 0$ (pour $\|\cdot\|_F$) lorsque $x \to a$ (pour $\|\cdot\|_E$). On dit que g est la différentielle de f au point a et on la note $Df(a)$.

Formellement, $(\star)$ signifie que

$$\forall \varepsilon > 0, \exists \eta > 0 \ / \ \forall x \in \Omega \backslash \{a\}, \ \|x-a\|_E < \eta$$
$$\Rightarrow \quad \|f(x) - f(a) - g(x-a)\|_F < \varepsilon \|x-a\|_E$$

Lemme

L'application linéaire g est unique.

Preuve : Si deux $g_1, g_2 \in L_c(E, F)$ vérifient $(\star)$, alors

$$g_1(x-a) - g_2(x-a) = \|x-a\|_E \varepsilon(x)$$

pour tout $x \in \Omega$, où $\varepsilon : \Omega \to F$ est telle que $\varepsilon(a) = 0$ et $\varepsilon(x) \to 0$ lorsque $x \to a$. Soit $h \in E$ un élément quelconque de E. Comme Ω est un ouvert, il existe $r_0 > 0$ tel que pour tout $t \in]-r_0, +r_0[$, le point $x = a + th$ est dans Ω. Bien sûr, $a + th \to a$ lorsque $t \to 0$. En posant $x = a + th$ dans l'égalité ci-dessus, avec $t > 0$, et par linéarité de g_1 et g_2, on obtient que

$$g_1(h) - g_2(h) = \|h\|_E \varepsilon_h(t)$$

Preuve suite et fin : pour tout $t > 0$ suffisamment petit, où $\varepsilon_h(t) = \varepsilon(a + th)$ est telle que $\varepsilon_h(t) \to 0$ lorsque $t \to 0$. En faisant effectivement tendre $t \to 0$, il suit que $g_2(h) = g_1(h)$ pour tout $h \in E$, et donc que $g_2 \equiv g_1$. On a donc bien unicité de g et on peut donc bien s'en servir pour définir $Df(a) = g$. CQFD.

Proposition

Si $E = \mathbb{R}^p$ et $F = \mathbb{R}^q$ alors on retrouve bien que

$$Df(a).(H) = \big(Df(a)^1.(H), \dots, Df(a)^q.(H)\big) \text{ avec}$$

$$Df(a)^j.(H) = \sum_{i=1}^{p} \frac{\partial f_j}{\partial x_i}(a) h_i$$

pour tout $H = (h_1, \dots, h_p) \in \mathbb{R}^p$, où les $f_1, \dots, f_q$ sont les fonctions composantes de f.

Preuve : pour faire simple supposons $q = 1$. On a

$$f(x) - f(a) - g(x - a) = \|x - a\|_E \varepsilon(x) \qquad (\star)$$

avec $\varepsilon(a) = 0$ et $\varepsilon(x) \to 0$ lorsque $x \to a$. Prenons successivement x sous la forme $x = x_t$ avec

$$x_t = (a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_p)\ .$$

Si $f_a^i(t) = f(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_p)$ est la $i$ème fonction partielle du Chapitre 2, alors $(\star)$ s'écrit aussi

$$f_a^i(t) - f_a^i(a_i) - \Phi_i(t - a_i) = |t - a_i| \varepsilon(t) \qquad (\star\star)$$

où $\varepsilon(a_i) = 0$ et $\varepsilon(t) \to 0$ lorsque $t \to a_i$, et la fonction

$$\Phi_i(t) = g(0, \dots, 0, t, 0 \dots, 0)$$

(le t à la $i$ème place) est une fonction linéaire de $\mathbb{R}$ dans lui-même. Les fonctions linéaires de $\mathbb{R}$ dans $\mathbb{R}$ s'écrivent toutes sous la forme $t \to \lambda t$ pour $\lambda \in \mathbb{R}$ un réel donné. Donc, en remplaçant, si on note $\Phi_i(t) = \lambda_i t$, on obtient avec $(\star\star)$ que

$$f_a^i(t) - f_a^i(a_i) - \lambda_i(t - a_i) = |t - a_i| \varepsilon(t)\ . \qquad (\star\star\star)$$

Preuve suite et fin : A partir de $(\star\star\star)$ on obtient que

$$\lim_{t \to a_i, t \neq a_i} \frac{f_a^i(t) - f_a^i(a_i)}{t - a_i} = \lambda_i$$

et donc

$$\lambda_i = (f_a^i)'(a_i) = \frac{\partial f}{\partial x_i}(a) \ .$$

En remarquant que $g(H) = \sum_{i=1}^{p} \Phi_i(h_i)$ pour tout $H = (h_1, \ldots, h_p)$, on voit que

$$g(H) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a) h_i$$

pour tout $H = (h_1, \ldots, h_p)$. D'où la proposition. CQFD.

Pourquoi demander dans $(\star)$ que $g \in L_c(E,F)$, et non pas tout simplement que $g \in L(E,F)$ sans continuité ? En raison du résultat suivant...

Théorème

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach. Soient Ω un ouvert de E, $a \in \Omega$ un point de Ω, et $f : \Omega \to F$ une application de Ω dans F. Si f est différentiable au point a, alors f est aussi continue au point a.

Preuve : On a

$$f(x) - f(a) - g(x-a) = \|x-a\|_E \varepsilon(x) \qquad (\star)$$

avec $\varepsilon(a) = 0$ et $\varepsilon(x) \to 0$ lorsque $x \to a$. Comme g est continue, $\lim_{x\to a} g(x-a) = 0$. Donc (comme aussi $\|x-a\|_E \varepsilon(x) \to 0$ lorsque $x \to a$), on voit avec $(\star)$ que

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

et f est bien continue en a. CQFD.

Définition

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, et soit Ω un ouvert de E. Soit $f : \Omega \to F$ une application de Ω dans F. On dit que f est différentiable sur Ω si f est différentiable en tout point a de Ω.

Si $f : \Omega \to F$ est différentiable sur Ω, alors on récupère une application

$$Df : \Omega \to L_c(E, F)$$

qui à $x \in \Omega$ associe $Df(x) \in L_c(E, F)$. Les applications Df et f sont de même type (ou même nature) : ce sont deux applications d'un espace de Banach dans un autre espace de Banach. On place ici sur $L_c(E, F)$ sa norme naturelle

$$\|\Phi\|_{L_c(E,F)} = \sup_{x \in E \setminus \{0\}} \frac{\|\Phi(x)\|_F}{\|x\|_E} ,$$

et on rappelle (cf. Chapitre 3) que $\big(L_c(E, F), \|\cdot\|_{L_c(E,F)}\big)$ est un Banach dès que $(F, \|\cdot\|_F)$ est un Banach.

La remarque que l'on vient de faire est d'importance. Elle permet de définir de façon naturelle les applications de classe C^1, les applications deux fois différentiables, les applications de classe C^2, les applications trois fois différentiables, etc. à partir des seules notions d'application différentiable et d'application continue entre Banach, comme c'était le cas pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Evidemment, si la théorie est claire et simple, dans la pratique les espaces se compliquent considérablement au fur et à mesure que l'on différencie :

$$Df : \Omega \to L_c(E, F) \ , \ D^2 f : \Omega \to L_c(E, L_c(E, F)) \ ,$$
$$D^3 f : \Omega \to L_c(E, L_c(E, L_c(E, F))) \ , \ \ldots$$

Les espaces d'arrivées seront à simplifier (par assimilation). Par exemple (cf. Chapitre 3), $L_c(E, L_c(E, F)) \approx L_c(E, E; F)$.

Définition

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, et soit Ω un ouvert de E. Soit $f : \Omega \to F$ une application différentiable de Ω dans F. On dit que f est de classe C^1 sur Ω si $Df : \Omega \to L_c(E, F)$ est continue sur Ω.

La continuité de DF en un point $a \in, \Omega$ s'écrit

$$\forall \varepsilon > 0, \exists \eta > 0 \ / \ \forall x \in \Omega, \|x - a\|_E < \eta \\ \Rightarrow \|Df(x) - Df(a)\|_{L_c(E,F)} < \varepsilon$$

et donc

$$\forall \varepsilon > 0, \exists \eta > 0 \ / \ \forall x \in \Omega, \|x - a\|_E < \eta \\ \Rightarrow \|Df(x).h - Df(a).h\|_F \leq \varepsilon \|h\|_E \text{ pour tout } h \in E \ .$$

Les deux formulations sont équivalentes en vertue de la définition même de la norme d'une application linéaire continue (et puisque $\varepsilon > 0$ est quelconque).

Proposition

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, et $a \in \Omega$ un point de Ω. Soit $f : \Omega \to F$ une application de Ω dans F différentiable au point a de différentielle en ce point $Df(a) \in L_c(E, F)$. Si $\|\cdot\|'_E$ et $\|\cdot\|'_F$ sont deux autres normes sur E et F, respectivement équivalentes à $\|\cdot\|_E$ et $\|\cdot\|_F$, alors f est encore différentiable par rapport aux nouvelles normes et sa différentielle par rapport aux nouvelles normes est toujours $Df(a)$.

En d'autres termes, la notion de différentiabilité et de différentielle est inchangée (ou stable) par changement de normes équivalentes.

Preuve : Elle est très simple. On commence déjà par remarquer que $L_c(E, F)$ ne change pas par changement de normes équivalentes. Maintenant, comme $\|\cdot\|_E$ et $\|\cdot\|'_E$ sont équivalentes sur E, dire que $x \to a$ pour $\|\cdot\|_E$ équivaut à dire que $x \to a$ pour $\|\cdot\|'_E$.

Preuve suite et fin : De même, comme $\|\cdot\|_F$ et $\|\cdot\|'_F$ sont équivalentes sur F, dire que $\varepsilon(x) \to 0$ pour $\|\cdot\|_F$ équivaut à dire que $\varepsilon(x) \to 0$ pour $\|\cdot\|'_F$. Donc dire que l'on a

$$f(x) - f(a) - g(x - a) = \|x - a\|_E \varepsilon(x) \qquad (\star)$$

pour tout $x \in \Omega$, où $\varepsilon : \Omega \to F$ est telle que $\varepsilon(a) = 0$ et $\varepsilon(x) \to 0$ (pour $\|\cdot\|_F$) lorsque $x \to a$ (pour $\|\cdot\|_E$), équivaut à dire que l'on a

$$f(x) - f(a) - g(x - a) = \|x - a\|'_E \varepsilon(x) \qquad (\star')$$

pour tout $x \in \Omega$, où $\varepsilon : \Omega \to F$ est telle que $\varepsilon(a) = 0$ et $\varepsilon(x) \to 0$ (pour $\|\cdot\|'_F$) lorsque $x \to a$ (pour $\|\cdot\|'_E$). D'où la proposition. CQFD.

Théorème

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$, et $(G, \|\cdot\|_G)$ trois espaces de Banach, Ω un ouvert de E, Ω' un ouvert de F, et $a \in \Omega$ un point de Ω. Soient $f : \Omega \to F$ une application de Ω dans F, et $g : \Omega' \to G$ une application de Ω' dans G. On suppose que f est différentiable au point a, que $f(\Omega) \subset \Omega'$, et que g est différentiable au point $f(a)$. Alors $g \circ f : \Omega \to G$ est différentiable au point a de différentielle en ce point $D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$.

Preuve : On a

$$
\begin{aligned}
& g(y) - g\left(f(a)\right) - Dg\left(f(a)\right).\left(y - f(a)\right) = \|y - f(a)\|_F \tilde{\varepsilon}(y) \ , \\
& f(x) - f(a) - Df(a).(x - a) = \|x - a\|_E \varepsilon(x) \ ,
\end{aligned}
$$

où $\varepsilon(x) \to 0$ lorsque $x \to a$ et $\tilde{\varepsilon}(y) \to 0$ lorsque $y \to f(a)$. Posons $y = f(x)$ avec $x \to a$. Alors $y \to f(a)$ par continuité de f en a. Et donc

Preuve suite :

$$g(f(x)) - g(f(a)) - Dg(f(a)).(f(x) - f(a)) \\ = \|f(x) - f(a)\|_F \hat{\varepsilon}(x) \ ,$$

où $\hat{\varepsilon}(x) \to 0$ lorsque $x \to a$. On a

$$f(x) - f(a) = Df(a).(x - a) + \|x - a\|_E \varepsilon(x) \ .$$

Par continuité de $Df(a)$ (en tant qu'application linéaire donc), on en déduit que

$$\|f(x) - f(a)\|_F = O(\|x - a\|_E)$$

où par $O(\|x - a\|_E)$ on entend un terme dominé par $C\|x - a\|_E$ pour $C > 0$ une constante (c'est au moins vrai pour x au voisinage de a, et il n'y a que cela qui nous intéresse). Par continuité et linéarité de $Dg(f(a))$,

$$Dg(f(a)).(\|x - a\|_E \varepsilon(x)) = \|x - a\|_E \hat{\varepsilon}'(x)$$

où $\hat{\varepsilon}'(x) \to 0$ lorsque $x \to a$. De toutes ces relations on tire que

Preuve suite et fin :

$$\begin{aligned} & g\left(f(x)\right) - g\left(f(a)\right) - \left(Dg\left(f(a)\right) \circ Df(a)\right).(x-a) \\ & \qquad = \|f(x) - f(a)\|_F \overline{\varepsilon}(x) \ , \end{aligned}$$

où $\overline{\varepsilon}(x) \to 0$ lorsque $x \to a$. Or $Dg\left(f(a)\right) \circ Df(a) \in L_c(E, G)$, et donc $g \circ f$ est différentiable en a avec

$$D\left(g \circ f\right)(a) = Dg\left(f(a)\right) \circ Df(a) \ .$$

La proposition est démontrée. CQFD.

4. Différentielles d'applications particulières

On se propose ici de calculer des différentielles d'applications particulières. Déjà, de façon évidente, une application constante est différentiable en tout point et la différentielle d'une application constante est nulle en tout point. En d'autres termes, si $f : \Omega \to F$ est constante, Ω ouvert de E, alors f est différentieable en tout point a de Ω et $Df(a) \equiv 0$ est l'application linéaire nulle de E dans F. On calcule dans ce qui suit les différentielles des applications linéaires, et multilinéaires.

Théorème

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces de Banach, et $f \in L_c(E, F)$ une application linéaire continue de E dans F. Alors f est différentiable en tout point a de E de différentielle constante sur E donnée par $Df(a) = f$.

Preuve : De façon évidente, pour tout $x \in E$,

$$f(x) - f(a) - f(x-a) = 0$$

et on peut donc écrire a fortiori que

$$\|f(x) - f(a) - f(x-a)\|_F = \|x-a\|_E \varepsilon(x)$$

pour tout $x \in E$, où $\varepsilon(x) \to 0$ lorsque $x \to a$ (il suffit de prendre $\varepsilon \equiv 0$), ce qui prouve que f est différentiable au point a de différentielle en ce point $Df(a) = f$. CQFD.

L'application $Df : E \to L_c(E, F)$ est continue (puisque constante), et donc f est même de classe C^1 sur E (et ses dérivées successives sont nulles). On aborde maintenant la différentiabilité des applications bilinéaires.

Théorème

Soient $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$, *et* $(F, \|\cdot\|_F)$ *trois espaces de Banach, et* $f \in L_c(E_1, E_2; F)$ *une application bilinéaire continue de* $E_1 \times E_2$ *dans* F. *Alors* f *est différentiable en tout point* $a = (a_1, a_2)$ *de* $E_1 \times E_2$ *de différentielle en ce point donnée par*

$$Df(a).(h) = f(a_1, h_2) + f(h_1, a_2)$$

pour tout $h = (h_1, h_2)$ *dans* $E_1 \times E_2$.

Preuve : On place sur $E_1 \times E_2$ l'une des deux normes équivalentes $\|(x_1, x_2)\| = \|x_1\|_{E_1} + \|x_2\|_{E_2}$, ou $\|(x_1, x_2)\| = \sqrt{\|x_1\|_{E_1}^2 + \|x_2\|_{E_2}^2}$. Par exemple, on place la norme

$$\|(x_1, x_2)\|_{E_1 \times E_2} = \|x_1\|_{E_1} + \|x_2\|_{E_2} \ .$$

Soit $a = (a_1, a_2)$ un point fixé de $E_1 \times E_2$, et soit $g : E_1 \times E_2 \to F$ l'application définie par

Preuve suite :

$$g(h_1, h_2) = f(a_1, h_2) + f(h_1, a_2)$$

pour tout (h_1, h_2) dans $E_1 \times E_2$. Comme f est bilinéaire, g est une application linéaire de $E_1 \times E_2$ dans F. L'application g est de plus continue car f l'est. En effet,

$$\begin{aligned}
&\|g(h_1, h_2)\|_F \\
&\leq \|f(a_1, h_2)\|_F + \|f(h_1, a_2)\|_F \\
&\leq \|f\|_{L_c(E_1,E_2;F)}\|a_1\|_{E_1}\|h_2\|_{E_2} + \|f\|_{L_c(E_1,E_2;F)}\|h_1\|_{E_1}\|a_2\|_{E_2} \\
&\leq \|f\|_{L_c(E_1,E_2;F)}\left(\|a_1\|_{E_1} + \|a_2\|_{E_2}\right)\|(h_1, h_2)\|_{E_1 \times E_2}
\end{aligned}$$

et on obtient ainsi la continuité de l'application linéaire g. Par bilinéarité de f,

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

où $h_1 = x_1 - a_1$ et $h_2 = x_2 - a_2$. Par suite

Preuve suite et fin : pour tout $(x_1, x_2) \in E_1 \times E_2$,

$$\begin{aligned}
&\|f(x_1, x_2) - f(a_1, a_2) - f(a_1, x_2 - a_2) - f(x_1 - a_1, a_2)\|_F \\
&\leq \|f\|_{L_c(E_1, E_2; F)} \|x_1 - a_1\|_{E_1} \|x_2 - a_2\|_{E_2} \\
&\leq \frac{1}{2} \|f\|_{L_c(E_1, E_2; F)} \|(x_1 - a_1, x_2 - a_2)\|^2_{E_1 \times E_2} ,
\end{aligned}$$

et on a montré que pour tout $(x_1, x_2) \in E_1 \times E_2$,

$$\begin{aligned}
&f(x_1, x_2) - f(a_1, a_2) - f(a_1, x_2 - a_2) - f(x_1 - a_1, a_2) \\
&= \|(x_1 - a_1, x_2 - a_2)\|_{E_1 \times E_2} \varepsilon(x_1, x_2) ,
\end{aligned}$$

où $\varepsilon(x_1, x_2) \to 0$ lorsque $(x_1, x_2) \to (a_1, a_2)$. En particulier, f est différentiable au point $a = (a_1, a_2)$ de différentielle en ce point

$$Df(a).(h) = f(a_1, h_2) + f(h_1, a_2)$$

pour tout $h = (h_1, h_2)$ dans $E_1 \times E_2$. CQFD.

La proposition que l'on vient de démontrer se généralise aux applications multilinéaires d'ordre quelconque.

Théorème

Soient $(E_1, \|\cdot\|_{E_1})$,. . .,$(E_n, \|\cdot\|_{E_n})$, $(F, \|\cdot\|_F)$ *des espaces de Banach. Soit* $f \in L_c(E_1, \dots, E_n; F)$ *une application n-linéaire continue de* $E_1 \times \cdots \times E_n$ *dans* F*. Alors* f *est différentiable en tout point* $a = (a_1, \dots, a_n)$ *de* $E_1 \times \cdots \times E_n$*, de différentielle donnée par*

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

pour tout $h = (h_1, \dots, h_n)$ *dans* $E_1 \times \cdots \times E_n$*.*

5. Applications à valeurs dans un produit d'espaces de Banach

On considère $(E, \|\cdot\|_E)$, $(F_1, \|\cdot\|_{F_1})$,..., $(F_k, \|\cdot\|_{F_k})$ des espaces de Banach. On place sur le produit $F = F_1 \times \cdots \times F_k$ l'une des deux normes produits définies comme auparavant par

$$\|(y_1, \ldots, y_k)\|_{F_1 \times \cdots \times F_k} = \sum_{i=1}^{k} \|y_i\|_{F_i} \ , \text{ ou}$$

$$\|(y_1, \ldots, y_k)\|_{F_1 \times \cdots \times F_k} = \sqrt{\sum_{i=1}^{k} \|y_i\|_{F_i}^2} \ .$$

Etant donnés Ω un ouvert de E, et $f : \Omega \to F$, avec $F = F_1 \times \cdots \times F_k$, une application de Ω dans le produit $F_1 \times \cdots \times F_k$, on note $f_i : \Omega \to F_i$ les applications définies par $f = (f_1, \ldots, f_k)$. Donc $f(x) = (f_1(x), \ldots, f_k(x))$ pour tout $x \in \Omega$.

Proposition

L'application $f : \Omega \to F$ est différentiable en un point a de Ω si et seulement si les applications f_i le sont. De plus, pour tout $x \in E$,

$$Df(a).(x) = (Df_1(a).(x), \ldots, Df_k(a).(x))$$

où les $Df(a) \in L_c(E, F)$ et $Df_i(a) \in L_c(E, F_i)$ sont les différentielles de f et f_i en a.

Preuve : Plaçons sur F la norme

$$\|(y_1, \ldots, y_k)\|_{F_1 \times \cdots \times F_k} = \sum_{i=1}^{k} \|y_i\|_{F_i} \ .$$

Un raisonnement avec l'autre norme donnerait le même résultat puisque différentiabilité et différentielle sont des notions invariantes par changement de normes équivalentes. Supposons que f est différentiable en a. Il existe alors $g \in L_c(E, F)$ telle que

Preuve suite et fin :

$$\|f(x) - f(a) - g(x-a)\|_F = \|x-a\|_E \varepsilon(x) \qquad (\star)$$

pour tout $x \in \Omega$, où $\varepsilon : \Omega \to \mathbb{R}^+$ est telle que $\varepsilon(x) \to 0$ lorsque $x \to a$. Clairement g s'écrit sous la forme $g = (g_1, \dots, g_k)$, où les $g_i : E \to F_i$ sont linéaires continues de E dans F_i. Comme

$$\|f(x) - f(a) - g(x-a)\|_F = \sum_{i=1}^{k} \|f_i(x) - f_i(a) - g_i(x-a)\|_{F_i} \ ,$$

on déduit de $(\star)$ que pour tout i,

$$\|f_i(x) - f_i(a) - g_i(x-a)\|_{F_i} = \|x-a\|_E \varepsilon(x) \ .$$

Ainsi f_i est différentiable au point a, et $g_i = Df_i(a)$. A l'inverse, on obtient tout aussi facilement que si les f_i sont différentiables au point a, alors f l'est aussi avec

$$Df(a).(x) = (Df_1(a).(x), \dots, Df_k(a).(x)) \ .$$

D'où la proposition. CQFD.

Une application bien utile de cette proposition et de la proposition portant sur la différentiabilité des applications bilinéaires est donnée par le corollaire suivant.

Corollaire

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ *deux espaces de Banach. Soit* $\Phi : E \times E \to F$ *une application bilinéaire continue de* $E \times E$ *dans* F*, donc* $\Phi \in L_c(E, E; F)$*. L'application* $f : E \to F$ *définie par* $f(x) = \Phi(x, x)$ *est différentiable sur* E *de différentielle en tout point* a *de* E *donnée par* $Df(a).(x) = \Phi(a, x) + \Phi(x, a)$ *pour tout* $x \in E$*. En particulier,* $Df(a).(x) = 2\Phi(a, x)$ *pour tout* $x \in E$ *si* Φ *est symétrique.*

Preuve : On écrit que f est la composée de Φ et de l'application $F : E \to E \times E$ définie par $F(x) = (x, x)$. On vérifie facilement que F est différentiable sur E de différentielle en a donnée par

$$DF(a).(x) = (x, x) \ .$$

Preuve suite et fin : Par différentiabilité des fonctions composées, et différentiabilité des applications bilinéaires, on en déduit que

$$Df(a).(x) = D\Phi(F(a)).\,(DF(a).(x)) = D\Phi(a,a).(x,x) = \Phi(a,x)+\Phi(x,a$$

pour tout $x \in E$, et donc le corollaire. CQFD.

A titre d'application immédiate, notons $L^2(\mathbb{R})$ l'espace de Lebesgue des fonctions mesurables de $\mathbb{R}$ dans $\mathbb{R}$ qui sont de carré intégrables, à savoir qui sont telles que $\int_R u^2 dx < +\infty$. La norme

$$\|u\|_{L^2} = \sqrt{\int_{\mathbb{R}} u^2 dx}$$

fait de $L^2(\mathbb{R})$ un espace de Banach. L'application $f : L^2(\mathbb{R}) \to \mathbb{R}$ définie par $u \to \int_{\mathbb{R}} u^2 dx$ est alors différentiable sur $L^2(\mathbb{R})$ de différentielle

$$Df(u).(v) = 2\int_{\mathbb{R}} uvdx$$

En effet, $(u,v) \to \int_{\mathbb{R}} uvdx$ est une forme bilinéaire symétrique continue (par Cauchy-Schwarz) sur $L^2(\mathbb{R})$.

6. Pseudo produits d'applications différentiables

Soient $(E, \|\cdot\|_E)$, $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$, et $(F, \|\cdot\|_F)$ quatre espaces de Banach. Soient de plus Ω un ouvert de E, $u_1 : \Omega \to E_1$ et $u_2 : \Omega \to E_2$ deux applications de Ω dans E_1 et E_2, et $f : E_1 \times E_2 \to F$ une application bilinéaire continue de $E_1 \times E_2$ dans F. On définit le pseudo-produit u de u_1 et u_2 comme étant l'application $u : \Omega \to F$ donnée par

$$u(x) = f\left(u_1(x), u_2(x)\right) \ .$$

On parle de pseudo-produit par analogie avec le cas réel où $u_1 : \mathbb{R} \to \mathbb{R}$, $u_2 : \mathbb{R} \to \mathbb{R}$, et $u(x) = u_1(x)u_2(x)$ (à savoir $f(x, y) = xy$ de $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$). On démontre la proposition suivante.

Proposition

Soient $(E, \|\cdot\|_E)$, $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$, et $(F, \|\cdot\|_F)$ quatre espaces de Banach. Soient de plus Ω un ouvert de E, $u_1 : \Omega \to E_1$ et $u_2 : \Omega \to E_2$ deux applications de Ω dans E_1 et E_2, et $f : E_1 \times E_2 \to F$ une application bilinéaire continue de $E_1 \times E_2$ dans F. Si u_1, u_2 sont différentiables en un point a, alors $u : \Omega \to F$ est aussi différentiable en a et

$$Du(a).(x) = f\left(Du_1(a).(x), u_2(a)\right) + f\left(u_1(a), Du_2(a).(x)\right)$$

pour tout $x \in E$.

Preuve : On écrit que u est la composée de f et de l'application $F : \Omega \to E_1 \times E_2$ définie par $F(x) = (u_1(x), u_2(x))$. Clairement F est différentiable en a si u_1 et u_2 sont différentiables en a, et sa différentielle en a est donnée par

$$DF(a).(x) = (Du_1(a).(x), Du_2(a).(x))$$

pour tout $x \in E$. Par différentiabilité des fonctions composées, et

Preuve suite et fin : différentiabilité des applications bilinéaires, on en déduit que

$$\begin{aligned} Du(a).(x) &= Df(F(a)).(DF(a).(x)) \\ &= Df\,(u_1(a), u_2(a))\,.\,(Du_1(a).(x), Du_2(a).(x)) \\ &= f\,(Du_1(a).(x), u_2(a)) + f\,(u_1(a), Du_2(a).(x)) \end{aligned}$$

pour tout $x \in E$. La proposition est démontrée. CQFD.

A noter : on retrouve, dans le cas réel, la formule $(uv)' = u'v + uv'$.

On considère ici le cas où l'espace de départ est un produit d'espaces de Banach. Soient donc $(E_1, \|\cdot\|_{E_1})$,...,$(E_k, \|\cdot\|_{E_k})$, et $(F, \|\cdot\|_F)$ des espaces de Banach. On place sur $E = E_1 \times \cdots \times E_k$ l'une des deux normes produits équivalentes

$$\|(x_1, \ldots, x_k)\|_{E_1 \times \cdots \times E_k} = \sum_{i=1}^{k} \|x_i\|_{E_i} \ , \ \text{ou}$$

$$\|(x_1, \ldots, x_k)\|_{E_1 \times \cdots \times E_k} = \sqrt{\sum_{i=1}^{k} \|x_i\|_{E_i}^2} \ .$$

On place par exemple la première de ces deux normes. Soit Ω un ouvert de E, $a \in \Omega$ un point de Ω, et $f : \Omega \to F$ une application. On note $a = (a_1, \ldots, a_k)$ où les $a_i \in E_i$. Comme Ω est un ouvert, il existe $r_i = r_i(a)$, $r_i > 0$, tel que pour tout $x_i \in B_{a_i}(r_i)$, on ait $(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_k) \in \Omega$, où $B_{a_i}(r_i)$ est la boule de centre a_i et rayon r_i dans E_i. En effet, comme $a \in \Omega$ et Ω est un ouvert, il existe $r > 0$ tel que $B_a(r) \subset \Omega$, et il reste à remarquer que $(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_k) \in B_a(r)$ pour tout $x_i \in B_{a_i}(r)$ de sorte que l'on peut prendre $r_i = r$.

Cette remarque faite, on peut maintenant construire k applications partielles f_i de f en a, définies au voisinage de a_i par

$$f_a^i(x) = f\left(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_k\right) \ .$$

Donc $f_a^i : \Omega_i \to F$ où $\Omega_i \subset E_i$ est un voisinage de a_i dans E_i, à savoir Ω_i est tel que $B_{a_i}(r_i) \subset \Omega_i$ pour un certain $r_i > 0$.

Proposition

Si $f : \Omega \to F$ est différentiable au point $a \in \Omega$, les applications partielles $f_a^i : \Omega_i \to F$ sont différentiables au point a_i pour tout $i = 1, \dots, k$, et

$$Df(a).(x) = \sum_{i=1}^{k} Df_a^i(a_i).(x_i)$$

pour tout $x = (x_1, \dots, x_k)$ dans E.

Les $Df_a^i(a_i)$ sont souvent notés $\frac{\partial f}{\partial x_i}(a)$. Attention, ce ne sont pas des réels (comme vous en avez l'habitude) mais des applications linéaires continues :

$$\frac{\partial f}{\partial x_i}(a) \in L_c(E_i, F)$$

pour tout $i = 1, \ldots, k$. La relation du théorème s'écrit alors

$$Df(a).(x) = \sum_{i=1}^{k} \frac{\partial f}{\partial x_i}(a).(x_i)$$

pour tout $x = (x_1, \ldots, x_k)$ dans E. Les dérivées partielles d'une application sont liées à la structure d'espace produit au départ.

On démontre la proposition dans ce qui suit.

Preuve : L'application $\theta_i : E_i \to E$ qui à x associe

$$\theta_i(x) = (a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_k)$$

est clairement différentiable au point a_i de différentielle en ce point l'application $D\theta_i(a_i) \in L_c(E_i, E)$ donnée par

$$D\theta_i(a_i).(x) = (0, \ldots, 0, x, 0, \ldots, 0)$$

pour tout $x \in E_i$. On pourra par exemple appliquer ce qui a été dit dans la section précédente. Comme $f_a^i = f \circ \theta_i$, on obtient que f_a^i est bien différentiable en a_i. Sa différentielle vaut $Df_a^i(a_i) = Df(a) \circ D\theta_i(a_i)$, et donc

$$Df(a).(0, \ldots, 0, x_i, 0, \ldots, 0) = Df_a^i(a_i).(x_i)$$

pour tout $x_i \in E_i$. En écrivant que

$$Df(a).(x_1, \ldots, x_k) = \sum_{i=1}^{k} Df(a).(0, \ldots, 0, x_i, 0, \ldots, 0) = \sum_{i=1}^{k} Df_a^i(a_i).(x_i)\,,$$

on obtient la proposition. CQFD.

8. Applications de $\mathbb{R}$ dans un Banach

Si $f : \mathbb{R} \to \mathbb{R}$ est dérivable en un point x, alors on récupère deux notions de dérivées en x, la dérivée $f'(x)$ et la différentielle $Df(x)$, toutes deux reliées par

$$Df(x).(h) = f'(x)h$$

pour tout $h \in \mathbb{R}$. La remarque s'étend au cas où l'espace d'arrivé est un Banach (mais l'espace de départ est $\mathbb{R}$). Supposons que f soit une fonction de $\mathbb{R}$, ou d'un intervalle ouvert de $\mathbb{R}$, à valeurs dans un espace de Banach $(F, \|\cdot\|_F)$. A savoir $f : \mathbb{R} \to F$. Alors là encore on peut définir une dérivée $f'(x)$, avec $f'(x) \in F$, et une différentielle $Df(x) \in L_c(\mathbb{R}, F)$. Et là encore, les deux sont reliées par la formule

$$Df(x).(h) = hf'(x)$$

pour tout $h \in \mathbb{R}$ (le membre de droite fait sens comme un réel que multiplie un vecteur). On a $\|Df(x)\|_{L_c(\mathbb{R},F)} = \|f'(x)\|_F$. Plus précisément, on définit $f'(x)$ par

$$f'(x) = \lim_{y \to x, y \neq x} \frac{1}{y-x} \left(f(y) - f(x)\right)$$

et on vérifie que

(1) si $f'(x)$ existe, alors f est différentiable en x et $Df(x).(h) = hf'(x)$ pour tout h,

(2) si f est différentiable en x, alors $f'(x)$ existe et $Df(x).(h) = hf'(x)$ pour tout h.

Pour le voir on commence par remarquer qu'une application linéaire de $\mathbb{R}$ dans un espace vectoriel F est forcément du type $h \to hX$ pour X un vecteur de F (tout simplement : $X = f(1)$). Si f est différentiable en a alors il existe $X \in F$ tel que

$$f(x) - f(a) - (x-a)X = |x-a|\varepsilon(x) \ ,$$

où $\varepsilon(x) \to 0$ lorsque $x \to a$. En divisant par $x - a$ et en faisant tendre $x \to a$, on obtient que

$$\lim_{x \to a, x \neq a} \frac{f(x) - f(a)}{x-a} = X \ .$$

Donc $X = f'(a)$ et (2) est démontré. A l'inverse, si

$$\lim_{x \to a, x \neq a} \frac{f(x) - f(a)}{x - a} = f'(a) \ ,$$

alors

$$\lim_{x \to a, x \neq a} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = 0 \ ,$$

et donc

$$f(x) - f(a) - (x - a)f'(a) = |x - a|\varepsilon(x) \ ,$$

où $\varepsilon(x) \to 0$ lorsque $x \to a$. D'où (1).

Lemme

Soient $(F, \|\cdot\|_F)$ un espace de Banach, $f : [a,b] \to F$, et $g : [a,b] \to \mathbb{R}$ deux applications continues d'un intervalle $[a,b]$ de $\mathbb{R}$ respectivement dans F et $\mathbb{R}$. On suppose que f et g sont toutes deux différentiables sur $]a,b[$ et que

$$\|Df(x)\|_{L_c(\mathbb{R},F)} \leq g'(x)$$

pour tout $x \in]a,b[$, où $g'(x)$ est la dérivée de g, et $\|\cdot\|_{L_c(\mathbb{R},F)}$ est la norme de $L_c(\mathbb{R},F)$. Alors $\|f(b)-f(a)\|_F \leq g(b)-g(a)$.

Preuve : Soit $\varepsilon > 0$ quelconque. On va montrer que pour tout $x \in [a,b]$,

$$\|f(x)-f(a)\|_F \leq g(x)-g(a)+\varepsilon(x-a)+\varepsilon \ . \qquad (\star)$$

Comme $\varepsilon > 0$ est quelconque, il s'ensuivra en posant $x = b$ et en faisant $\varepsilon \to 0$ que l'on a bien que $\|f(b)-f(a)\|_F \leq g(b)-g(a)$.

Preuve suite : Notons $\Phi : [a, b] \to \mathbb{R}$ la fonction

$$\Phi(x) = \|f(x) - f(a)\|_F - g(x) + g(a) - \varepsilon(x - a) \ .$$

On veut montrer que $\Phi(x) \leq \varepsilon$ pour tout $x \in [a, b]$. Il est déjà clair, puisque f et g le sont, que la fonction Φ est continue sur $[a, b]$. Comme $\Phi(a) = 0$, et donc $\Phi(a) < \varepsilon$, il suit de la continuité de Φ que $\Phi(x) \leq \varepsilon$ pour $x \geq a$ voisin de a. Maintenant, soit $\Phi(x) \leq \varepsilon$ pour tout $x \in [a, b]$, auquel cas on a l'inégalité que l'on voulait, soit, en raisonnant par l'absurde, on peut supposer qu'il existe $d \in]a, b[$ tel que $\Phi(d) > \varepsilon$. Notons

$$I = \left\{ x \in [a, b] \ / \ \Phi(t) \leq \varepsilon \text{ pour tout } t \in [a, x] \right\} \ .$$

En raison de ce que l'on vient de dire, I est un intervalle du type $I = [a, c]$ (fermé à droite par continuité de Φ) avec $a < c$ et $c < b$ (car $c \leq d$). Toujours par continuité,

$$\Phi(c) = \varepsilon \ .$$

Comme $c \in]a, b[$, on a par hypothèse que $\|Df(c)\|_{L_c(\mathbb{R},F)} \leq g'(c)$.

Preuve suite : Par définition de $Df(c)$,

$$\forall \varepsilon' > 0 \ , \ \exists \eta > 0 \ / \ \forall x \in]a, b[\ , \ |x - c| < \eta$$
$$\Rightarrow \ \|f(x) - f(c) - Df(c).(x - c)\|_F \leq \varepsilon' |x - c| \ .$$

En posant $\varepsilon' = \varepsilon/2$, et par inégalité triangulaire, on récupère que

$$\frac{\|f(x) - f(c)\|_F}{|x - c|} - \frac{\varepsilon}{2} \leq \|Df(c)\|_{L_c(\mathbb{R},F)}$$

pour $x \in]c, c + \eta[$. De même, quitte à diminuer $\eta > 0$, on va aussi pouvoir montrer que

$$g'(c) \leq \frac{g(x) - g(c)}{x - c} + \frac{\varepsilon}{2}$$

pour $x \in]c, c + \eta[$. L'inégalité $\|Df(c)\|_{L_c(\mathbb{R},F)} \leq g'(c)$ du théorème donne alors que

$$\|f(x) - f(c)\|_F \leq g(x) - g(c) + \varepsilon(x - c)$$

pour tout $c \leq x < c + \eta$. Comme $\Phi(c) = \varepsilon$, on a que

$$\|f(c) - f(a)\|_F - g(c) + g(a) - \varepsilon(c - a) = \varepsilon \ .$$

Preuve suite et fin : Il suit de ces deux dernières relations que

$$\begin{aligned}\|f(x)-f(a)\|_F &\leq \|f(x)-f(c)\|_F + \|f(c)-f(a)\|_F \\ &\leq g(x)-g(a)+\varepsilon(x-a)+\varepsilon\end{aligned}$$

pour $c \leq x < c+\eta$, et donc qu'il y a des $x > c$ qui sont tels que $x \in I$, ce qui est impossible puisque $I = [a, c]$. Par l'absurde on a donc montré $\Phi(x) \leq \varepsilon$ pour tout $x \in [a, b]$. Le lemme suit facilement. CQFD.

Soit $(E, \|\cdot\|_E)$ un espace de Banach, et soit Ω un ouvert de E. Etant donnés $x, y \in \Omega$ deux points dans Ω, on dit que le ségment $[x, y]$ est entièrement contenu dans Ω si pour tout $t \in [0, 1]$,

$$tx + (1-t)y \in \Omega \ .$$

On démontre la version vectorielle du théorème des accroissements finis.

Théorème (Théorème des accroissements finis.)

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, et $f : \Omega \to F$ une application différentiable dans Ω. Soient aussi $x, y \in \Omega$ deux points de Ω avec la propriété que le ségment $[x, y]$ est entièrement contenu dans Ω. Alors

$$\|f(y) - f(x)\|_F \leq \Lambda \|y - x\|_E \ ,$$

où $\Lambda = \sup_{t\in[0,1]} \|Df(tx + (1-t)y)\|_{L_c(E,F)}$ peut être infini.

Une situation où l'on sait que Λ est fini est la situation où f est de classe C^1, car alors Λ est la borne supérieure d'une fonction continue, la fonction $t \to \|Df(tx + (1-t)y)\|_{L_c(E,F)}$, sur le compact $[0, 1]$, et une fonction continue sur un compact est bornée (et atteint ses bornes).

Preuve : On note h la fonction de t définie par

$$h(t) = f\left(tx + (1-t)y\right) \ .$$

Alors h est définie et différentiable sur un intervalle du type $]-\varepsilon, 1+\varepsilon[$ (car $[x, y]$ est contenu dans Ω, et Ω est un ouvert). Sa différentielle en un point t est donnée par le théorème de composition en remarquant que h est la composée de f avec $L(t) = tx + (1-t)y$. Sa différentielle en t vaut alors $Dh(t) = Df(L(t)) \circ DL(t)$, et sa dérivée en t (on rappelle que $h'(t) = Dh(t).(1)$) vaut donc

$$h'(t) = Df(L(t)).(x-y) \ .$$

En particulier,

$$\begin{aligned} \|Dh(t)\|_{L_c(\mathbb{R},F)} &= \|h'(t)\|_F \\ &\leq \Lambda \|y-x\|_E = g'(t) \end{aligned}$$

pour tout $t \in [0,1]$, où $g : \mathbb{R} \to \mathbb{R}$ est la fonction définie par

$$g(t) = \Lambda \|y-x\|_E t \ .$$

Preuve suite et fin : Avec le lemme il suit que

$$\|h(1) - h(0)\|_F \leq g(1) - g(0) \ ,$$

ce qui est l'inégalité du théorème des accroissements finis. CQFD.

Pour rappel, une remarque importante sur le théorème des accroissments finis (cf. Chapitre 2) est que l'égalité du cas réel ne se prolonge pas au cas vectoriel. Plus précisemment, on ne pourra en général pas écrire que

$$f(y) - f(x) = Df(c).(y - x)$$

pour un ceratin $c \in]x, y[$ dès lors que f est à valeurs vectorielles.

Par définition, un ouvert Ω est dit convexe si pour tous $x, y \in \Omega$, le ségment $[x, y]$ est contenu dans Ω (par exemple une boule dans un espace vectoriel normé est convexe). Par ailleurs, on rappelle qu'une application $f : X \to Y$ entres espaces métriques (X, d_X) et (Y, d_Y) est dite lipschitzienne sur X s'il existe un réel $K \geq 0$ tel que

$$d_Y\left(f(x), f(y)\right) \leq K d_X(x, y)$$

pour tous $x, y \in X$. Une fonction lipschitzienne est toujours uniformément continue.

Corollaire

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert convexe de E, et $f : \Omega \to F$ une application différentiable dans Ω. On suppose qu'il existe $K \geq 0$ tel que pour tout $x \in \Omega$, $\|Df(x)\|_{L_c(E,F)} \leq K$. Alors

$$\|f(y) - f(x)\|_F \leq K\|y - x\|_E$$

pour tous $x, y \in \Omega$. En particulier, f est lipschitzienne sur Ω.

Preuve : Elle est immédiate à partir du théorème des accroissements finis. CQFD.

Toujours par définition, un ouvert Ω est dit connexe s'il n'est pas réunion de deux ouverts disjoints non vides. Donc Ω ouvert est connexe s'il n'existe pas Ω_1 et Ω_2 deux ouverts non vides tels que $\Omega = \Omega_1 \cup \Omega_2$ et $\Omega_1 \cap \Omega_2 = \emptyset$. Une propriété remarquable des espaces vectoriels normés est la suivante : "Un ouvert Ω d'un espace vectoriel normé est connexe si et seulement si pour tous $x, y \in \Omega$ il existe $\gamma : [0,1] \to \Omega$ un chemin continu qui ait x et y pour extrémités, à savoir qui soit tel que $\gamma(0) = x$ et $\gamma(1) = y$." On dit encore qu'un ouvert d'un espace vectoriel normé est connexe si et seulement si il est connexe par arc.

Corollaire

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert connexe de E, et $f : \Omega \to F$ une application différentiable dans Ω. Si $Df(x) \equiv 0$ pour tout $x \in \Omega$, alors f est constante sur Ω.

Preuve : Il y a un peu plus à dire que pour le corollaire précédent. Les boules étant convexes, si f est telle que $Df(x) \equiv 0$ pour tout $x \in \Omega$, alors f est localement constante sur Ω en vertue du corollaire précédent. A savoir : pour tout $x \in \Omega$, $\exists r > 0$ tel que $f = Cste$ sur $B_x(r)$. Soient maintenant x et y deux points dans Ω. Comme Ω est connexe, il existe $\gamma : [0,1] \to \Omega$ un chemin continu tel que $\gamma(0) = x$ et $\gamma(1) = y$. L'application $f \circ \gamma : [0,1] \to F$ est continue sur $[0,1]$, comme composée d'applications continues, et localement constante sur $[0,1]$ puisque f est localement constante sur Ω. Notons $g = f \circ \gamma$, et soit

$$I = \left\{ t \in [0,1] \ / \ \forall \tau \in [0,t] \ , \ g(\tau) = g(0) \right\} \ .$$

Comme g est localement constante au voisinage de 0 (f est localement constante au voisinage de x) on peut écrire que I n'est pas vide. Alors I est un intervalle du type $I = [0, t_0]$. Or $t_0 < 1$ est impossible car alors f devrait être constante sur un intervalle du type $]t_0 - \varepsilon, t_0 + \varepsilon[$ pour un certain $\varepsilon > 0$, et donc on devrait avoir $t_0 \geq t_0 + \varepsilon$, ce qui est impossible. Donc $t_0 = 1$ et $f(x) = f(y)$. Or x et y sont quelconques dans Ω. CQFD.

On revient à la notion de dérivée partielle et on considère donc le cas où l'espace de départ est un produit d'espaces de Banach. Soient $(E_1, \|\cdot\|_{E_1})$,...,$(E_k, \|\cdot\|_{E_k})$, et $(F, \|\cdot\|_F)$ des espaces de Banach. On place sur $E = E_1 \times \cdots \times E_k$ l'une des normes produits équivalentes habituelles. On démontre alors le théorème suivant.

Théorème

Soient $(E_1, \|\cdot\|_{E_1})$,...,$(E_k, \|\cdot\|_{E_k})$, et $(F, \|\cdot\|_F)$ des espaces de Banach, Ω un ouvert de $E_1 \times \cdots \times E_k$, et $f : \Omega \to F$ une application. Alors f est de classe C^1 sur Ω si et seulement si les dérivées partielles $\frac{\partial f}{\partial x_i}$ de f existent pour tout i et sont continues sur Ω.

Preuve : Pour simplifier on démontre le théorème dans le cas où $k = 2$. On change alors les notations pour considérer trois espaces de Banach $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$, et $(G, \|\cdot\|_G)$, Ω un ouvert de $E \times F$, et $f : \Omega \to G$. On note (x, y) la variable de $E \times F$.

Preuve suite : Supposons pour commencer que f soit de classe C^1 sur Ω. Alors ses dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ existent en tout point de Ω et on a que

$$Df(x,y).(h,k) = \frac{\partial f}{\partial x}(x,y).(h) + \frac{\partial f}{\partial y}(x,y).(k)$$

pour tous $(x,y) \in \Omega$ et tous $(h,k) \in E \times F$. On a déjà démontré ce résultat. Reste maintenant à montrer que les dérivées partielles sont bien C^1 si f l'est. On a

$$\frac{\partial f}{\partial x}(x,y).(h) = Df(x,y).(h,0) \ , \ \text{et}$$
$$\frac{\partial f}{\partial y}(x,y).(k) = Df(x,y).(0,k)$$

pour tous $(x,y) \in \Omega$, tous $h \in E$, et tous $k \in F$.

Preuve suite : On en déduit que

$$
\begin{aligned}
&\|\frac{\partial f}{\partial x}(x,y)-\frac{\partial f}{\partial x}(a,b)\|_{L_c(E,G)} \\
&= \sup_{h\in E, h\neq 0} \frac{\|\frac{\partial f}{\partial x}(x,y).(h)-\frac{\partial f}{\partial x}(a,b).(h)\|_G}{\|h\|_E} \\
&= \sup_{h\in E, h\neq 0} \frac{\|\left(Df(x,y)-Df(a,b)\right).(h,0)\|_G}{\|h\|_E} \\
&\leq \|Df(x,y)-Df(a,b)\|_{L_c(E\times F,G)}
\end{aligned}
$$

pour tous (a,b) et (x,y) dans Ω. En particulier, la continuité de $\frac{\partial f}{\partial x}$ en (a,b) suit de celle de Df en (a,b). Un raisonnement analogue peut être fait pour $\frac{\partial f}{\partial y}$ et on voit donc que si f est de classe C^1 sur Ω alors $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont aussi de classe C^1 sur Ω.

Preuve suite : Réciproquement, on suppose maintenant que $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues sur Ω. Etant donné $(a, b) \in \Omega$, on écrit que

$$\begin{aligned} & f(x,y) - f(a,b) - \frac{\partial f}{\partial x}(a,b).(x-a) - \frac{\partial f}{\partial y}(a,b).(y-b) \\ & = f(x,y) - f(a,y) - \frac{\partial f}{\partial x}(a,b).(x-a) \\ & \quad + f(a,y) - f(a,b) - \frac{\partial f}{\partial y}(a,b).(y-b) \ . \end{aligned}$$

Pour (x, y) suffisamment proche de (a, b) les points (a, b), (a, y), et (x, y) vont être dans une boule $B_{(a,b)}(r)$, avec $r > 0$ petit de sorte que la boule soit contenue dans Ω. Soit $\varepsilon > 0$. Comme les dérivées partielles de f sont continues, quitte à choisir $r > 0$ suffisamment petit on va pouvoir écrire que

$$\begin{aligned} & \|\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(a,b)\|_{L_c(E,G)} < \frac{\varepsilon}{2} \ , \ \text{et} \\ & \|\frac{\partial f}{\partial y}(a,y) - \frac{\partial f}{\partial y}(a,b)\|_{L_c(F,G)} < \frac{\varepsilon}{2} \end{aligned}$$

Preuve suite : pour tous $(x, y) \in B_{(a,b)}(r)$. On applique maintenant le théorème des accroissements finis aux applications

$$g_1 : x \to f(x, y) - \frac{\partial f}{\partial x}(a, b).(x - a) \ , \ \text{et}$$
$$g_2 : y \to f(a, y) - \frac{\partial f}{\partial y}(a, b).(y - b) \ .$$

Ces applications sont différentiables de différentielles respectives

$$Dg_1(x) = \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(a, b) \ , \ \text{et}$$
$$Dg_2(y) = \frac{\partial f}{\partial y}(a, y) - \frac{\partial f}{\partial y}(a, b) \ .$$

Le théorème des accroissements finis appliqué à g_1 et g_2 donne alors que

$$\|g_1(x) - g_1(a)\|_G \leq \frac{\varepsilon}{2}\|x - a\|_E \ \text{ et } \ \|g_2(y) - g_2(b)\|_G \leq \frac{\varepsilon}{2}\|y - b\|_F$$

soit encore que

Preuve suite :

$$
\begin{aligned}
&\|f(x,y) - f(a,b) - \frac{\partial f}{\partial x}(a,b).(x-a) - \frac{\partial f}{\partial y}(a,b).(y-b)\|_G \\
&\leq \frac{\varepsilon}{2}\|x-a\|_E + \frac{\varepsilon}{2}\|y-b\|_F \\
&\leq \varepsilon\|(x,y)-(a,b)\|_{E\times F}
\end{aligned}
$$

pour tous $(x,y) \in B_{(a,b)}(r)$. En d'autres termes on a montré que f est différentiable en tout point (a,b) de Ω de différentielle en ce point l'application donnée par

$$
Df(a,b).(h,k) = \frac{\partial f}{\partial x}(a,b).(h) + \frac{\partial f}{\partial y}(a,b).(k)
$$

pour tous $(h,k) \in E \times F$.

Preuve suite et fin : Reste à montrer que $Df : \Omega \to L_c(E \times F, G)$ est continue. En écrivant que

$$
\begin{aligned}
& \|Df(x,y) - Df(a,b)\|_{L_c(E\times F,G)} \\
& = \sup_{(h,k)\neq(0,0)} \frac{\|Df(x,y).(h,k) - Df(a,b).(h,k)\|_G}{\|h\|_E + \|k\|_F} \\
& \leq \sup_{(h,k)\neq(0,0)} \frac{\|\frac{\partial f}{\partial x}(x,y).h - \frac{\partial f}{\partial x}(a,b).h\|_G}{\|h\|_E + \|k\|_F} \\
& \quad + sup_{(h,k)\neq(0,0)} \frac{\|\frac{\partial f}{\partial y}(x,y).k - \frac{\partial f}{\partial y}(a,b).k\|_G}{\|h\|_E + \|k\|_F} \\
& \leq \|\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(a,b)\|_{L_c(E,G)} + \|\frac{\partial f}{\partial y}(x,y) - \frac{\partial f}{\partial y}(a,b)\|_{L_c(F,G)}
\end{aligned}
$$

on voit que la continuité de Df en (a,b) suit de la continuité des dérivées partielles de f en (a,b). CQFD.

11. Le théorème d'inversion locale

Par définition on appellera voisinage ouvert d'un point a (dans un espace de Banach E) tout ouvert contenant a. On dira par ailleurs qu'une application $f : U \to V$ entre deux ouverts d'espaces de Banach est un C^1-difféomorphisme de U sur V si :
(i) f est de classe C^1 sur U,
(ii) f est bijective de U sur V,
(iii) f^{-1} est de classe C^1 sur V.
On note $\mathrm{Isom}(E, F)$ le sous espace de $L(E, F)$ constitué des isomorphismes de E sur F.

Théorème (Théorème d'inversion locale)

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, et $f : \Omega \to F$ une application de classe C^1. Soit $a \in \Omega$. On suppose que $Df(a) \in \mathrm{Isom}(E, F)$ est un isomorphisme de E sur F. Alors il existe un voisinage ouvert U de a, $U \subset \Omega$, et un voisinage ouvert V de $f(a)$ pour lesquels f est un C^1-difféomorphisme de U sur V.

Le théorème d'inversion locale est long à démontrer. Pour ne pas perdre trop de temps on va se limiter à commenter le théorème, sans le démontrer.

Une première remarque est que si l'application $f : U \to V$ est un C^1-difféomorphisme de U sur V, alors pour tout $x \in U$ on a $Df(x) \in \mathrm{Isom}(E, F)$ et, de plus,

$$Df(x)^{-1} = Df^{-1}\left(f(x)\right) \ .$$

En effet, $f \circ f^{-1}(y) = y$ pour tout $y \in V$ tandis que $f^{-1} \circ f(x) = x$ pour tout $x \in U$. On différencie ces relations et on obtient alors avec la formule de composition que

$$Df^{-1}\left(f(x)\right) \circ Df(x) = Id_E \ \text{ et } \ Df\left(f^{-1}(y)\right) \circ Df^{-1}(y) = Id_F \ .$$

En posant $y = f(x)$, on trouve la relation annoncée. On peut alors voir le théorème d'inversion locale comme une réciproque locale de cette propriété : si $Df(x)$ est un isomorphisme pour un x, alors f est un C^1-difféomorphisme au voisinage de x.

Une seconde remarque concerne le cas de la dimension finie. Dans ce cas on rappelle qu'il convient déjà de demander que les deux espaces aient même dimensions (car il n'y a pas d'isomorphisme entres espaces de dimensions différentes). Et on rappelle que lorsque $E = F = \mathbb{R}^n$ (tout espace vectoriel de dimension n est un $\mathbb{R}^n$ par choix d'une base) la condition $Df(x) \in \mathrm{Isom}(\mathbb{R}^n, \mathbb{R}^n)$ équivaut à la non nullité du jacobien de f en x (voir le Chapitre 2).

On énonce et démontre maintenant une version "globale" du théorème d'inversion locale.

Corollaire

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, et $f : \Omega \to F$ une application de classe C^1. Alors f est un C^1-difféomorphisme de Ω sur $f(\Omega)$ si et seulement si f est injective sur Ω et $Df(x) \in Isom(E, F)$ pour tout $x \in \Omega$.

Preuve : Si f est un C^1-difféomorphisme de Ω sur un ouvert $\Omega' = f(\Omega)$ de F alors clairement f est injective sur Ω (car bijectif implique injectif), et on a (cf. ci-dessus) automatiquement que $Df(x) \in \mathrm{Isom}(E, F)$ pour tout $x \in \Omega$.

Réciproquement, supposons que f soit injective sur Ω et que $Df(x) \in \mathrm{Isom}(E, F)$ pour tout $x \in \Omega$. Comme "bijectif équivaut à injectif et surjectif", il est clair que f réalise une bijection de Ω sur $f(\Omega)$. Par le théorème d'inversion locale, pour tout $x \in \Omega$ il existe U un voisinage ouvert de x dans Ω, et V un voisinage ouvert de $f(x)$ dans $f(\Omega)$ tels que f réalise un C^1-difféomorphisme de U sur V. Une première conséquence est que tout $y = f(x)$ dans $f(\Omega)$ possède un voisinage ouvert $V \subset f(\Omega)$. Il s'ensuit que $f(\Omega)$ est donc bien un ouvert.

Preuve suite et fin : Une seconde conséquence est que tout $y = f(x)$ dans $f(\Omega)$ possède un voisinage ouvert $V \subset f(\Omega)$ sur lequel la bijection réciproque $f^{-1} : f(\Omega) \to \Omega$ est de classe C^1 (il n'y a qu'une application réciproque f^{-1}). Et être de classe C^1 au voisinage de tout point d'un ouvert ou être de classe C^1 sur l'ouvert sont deux notions identiques. Il s'ensuit que f^{-1} est de classe C^1 sur $f(\Omega)$ et donc que f est bien un C^1-difféomorphisme de Ω sur un ouvert $\Omega' = f(\Omega)$ de F. CQFD.

Note

La classe de différentiabilité s'adapte au théorème : on peut remplacer C^1 par n'importe quel C^p du moment que $p \geq 1$.

12. Le théorème des fonctions implicites

Le théorème des fonctions implicites est plus délicat à comprendre que le théorème d'inversion locale (les deux théorèmes sont pourtant équivalents l'un à l'autre). En gros, avec le théorème des fonctions implicites, on veut dire que si $f : E \times F \to G$ est une application avec des hypothèses "convenables", alors une équation du type $f(x, y) = 0$ va définir (implicitement, d'où le nom du théorème) une fonction $g : E \to F$ telle que

$$f(x, y) = 0 \Leftrightarrow y = g(x) \ .$$

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach. On place sur $E \times F$ une des normes produit usuelles.

Théorème (Théorème des fonctions implicites)

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$, *et* $(G, \|\cdot\|_G)$ *trois espaces de Banach. Soit* Ω *un ouvert de* $E \times F$, *et* $f : \Omega \to G$ *une application de classe* C^1. *Soit* $(a, b) \in \Omega$ *tel que* $f(a, b) = 0$. *On suppose que*

$$\frac{\partial f}{\partial y}(a, b) \in Isom(F, G)\ .$$

Alors il existe U *un voisinage ouvert de* (a, b) *contenu dans* Ω, *il existe un voisinage* V *de* a *dans* E, *et il existe* $g : V \to F$ *de classe* C^1 *tels que si* $(x, y) \in U$ *vérifie* $f(x, y) = 0$ *alors* $x \in V$ *et* $y = g(x)$. *Réciproquement si* $x \in V$ *et* $y = g(x)$, *alors* $(x, y) \in U$ *et* $f(x, y) = 0$.

On a supposé ici que $f(a, b) = 0$. Si ce n'est pas le cas il suffit de considérer l'application $\tilde{f} = f - f(a, b)$. Par ailleurs on a bien sûr que $b = g(a)$ puisque $(a, b) \in U$.

Preuve : On ramène la preuve du théorème des fonctions implicites au théorème d'inversion locale. On considère $\tilde{f} : \Omega \to E \times G$ l'application définie par

$$\tilde{f}(x,y) = (x, f(x,y)) \ .$$

Cette application est clairement de classe C^1 sur Ω, et sa différentielle en un point (x,y) de Ω est donnée par

$$D\tilde{f}(x,y).(h,k) = \left(h, \frac{\partial f}{\partial x}(x,y).(h) + \frac{\partial f}{\partial y}(x,y).(k)\right) \ .$$

Clairement, $D\tilde{f}(a,b) \in \mathrm{Isom}(E \times F, E \times G)$ dès lors que $\frac{\partial f}{\partial y}(a,b) \in \mathrm{Isom}(F,G)$. Par exemple,

$$\begin{aligned} & D\tilde{f}(a,b).(h,k) = (h',k') \\ \Leftrightarrow \quad & h = h' \text{ et} \\ & \quad k = \frac{\partial f}{\partial y}(a,b)^{-1}\left(k' - \frac{\partial f}{\partial x}(a,b).h'\right) \ , \end{aligned}$$

de sorte que $D\tilde{f}(a,b)$ est bien inversible de $E \times F$ sur $E \times G$.

Preuve suite : Pour le caractère continue de l'inverse on peut soit appliquer le théorème de Banach, soit le vérifier diréctement sur la formule ci-dessus. Sachant que $\tilde{f}$ est de classe C^1 et que $D\tilde{f}(a,b) \in \mathrm{Isom}(E \times F, E \times G)$, on peut appliquer le théorème d'inversion locale à $\tilde{f}$. On obtient l'existence d'un voisinage ouvert U de (a,b) dans $E \times F$, et d'un voisinage ouvert $W = \tilde{f}(U)$ de $\tilde{f}(a,b) = (a, f(a,b)) = (a,0)$ dans $E \times G$ pour lesquels $\tilde{f}$ réalise un C^1-difféomorphisme de U sur W. En raison de la forme de $\tilde{f}$, le difféomorphisme réciproque est forcément du type

$$\tilde{f}^{-1}(x,z) = (x, \tilde{g}(x,z)) \ ,$$

où donc $\tilde{g} : W \to F$ est de c lasse C^1. Clairement on a équivalence entre les deux propositions :

(1) $(x,y) \in U$ et $f(x,y) = z$,

(2) $(x,z) \in W$ et $y = \tilde{g}(x,z)$.

Preuve suite et fin : Pour le voir il suffit de remarquer que $f(x,y) = z$ équivaut à $\tilde{f}(x,y) = (x,z)$, puis d'appliquer $\tilde{f}^{-1}$ à cette équation pour dire que cela équivaut à son tour à $\tilde{f}^{-1}(x,z) = (x,y)$, et donc à $y = \tilde{g}(x,z)$. On pose maintenant $z = 0$ dans (1) et (2). Alors les deux propositions suivantes deviennent équivalentes :

(3) $(x,y) \in U$ et $f(x,y) = 0$,

(4) $(x,0) \in W$ et $y = \tilde{g}(x,0)$.

Reste à remarquer si $V = \{x \in E \ / \ (x,0) \in W\}$, alors V est un voisinage ouvert de a dans E, et que si $g(x) = \tilde{g}(x,0)$, alors g est de classe C^1 dans V. Les propositions (3) et (4) donnent alors $(x,y) \in U$ et $f(x,y) = 0$ si et seulement si $x \in V$ et $y = g(x)$, ce qui est exactement ce qu'affirme le théorème des fonctions implicites. CQFD.

On vient de voir que le théorème des fonctions implicites est une conséquence du théorème d'inversion locale. La réciproque est vraie. En quelques mots, soit $f : E \to F$ telle que

$$Df(a) \in \mathrm{Isom}(E, F) \ .$$

On considère l'application $\Phi : F \times E \to F$ définie par

$$\Phi(y, x) = f(x) - y \ .$$

Si $b = f(a)$, alors $\Phi(b, a) = 0$ et

$$\frac{\partial \Phi}{\partial x}(b, a) = Df(a)$$

est un isomorphisme de E sur F. On déduit alors du théorème des fonctions implicites qu'il existe $g : F \to E$ (en ne faisant pas attention au domaine) telle que $\Phi\big(y, g(y)\big) = 0$, et donc telle que $f\big(g(y)\big) = y$. En gros, g est l'inverse de f et on a récupéré le théorème d'inversion locale.

Une remarque sur le théorème des fonctions implicites est que $g(a) = b$. Une remarque supplémentaire est qu'en différentiant l'équation $f\big(x, g(x)\big) = 0$ on obtient la valeur de $Dg(a)$. Plus précisemment,

$$\frac{\partial f}{\partial x}(a, b) + \frac{\partial f}{\partial y}(a, b) \circ Dg(a) = 0$$

et donc

$$Dg(a) = -\frac{\partial f}{\partial y}(a, b)^{-1} \circ \frac{\partial f}{\partial x}(a, b) \ .$$

Une autre remarque est que g est unique (au moins au voisinage de a). Si on a deux applications g_1 et g_2 qui vérifient le théorème des fonctions implicites, alors $g_1 = g_2$ au voisinage de a. Une dernière remarque est que là encore le théorème s'adapte à n'importe quelle classe de différenciation : on peut remplacer C^1 par C^p dès lors que $p \geq 1$.

Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, et soit Ω un ouvert de E. Soit aussi $f : \Omega \to F$ une application différentiable sur Ω. On dit que f est deux fois différentiable en un point $a \in \Omega$ si l'application $Df : \Omega \to L_c(E, F)$ est différentiable au point a. On note $D^2f(a)$ la différentielle de cette application en a. Alors $D^2f(a) \in L_c\big(E, L_c(E, F)\big)$. Cela étant dit, on a vu au chapitre 3, qu'il y a une isométrie naturelle entre cet espace $L_c\big(E, L_c(E, F)\big)$ et l'espace $L_c(E, E; F)$ des applications bilinéaires continues de $E \times E$ dans F. On regardera alors $D^2f(a)$ comme étant un élément de cet espace $L_c(E, E; F)$. Donc

$$D^2f(a) \in L_c(E, E; F) \ ,$$

ce qui revient à poser que $D^2f(a).(h, k) = \big(D^2f(a).(h)\big).(k)$ pour tout $(h, k) \in E \times E$. Il y a un problème de convention : on aurait aussi pu poser $D^2f(a).(h, k) = \big(D^2f(a).(k)\big).(h)$. Ce problème est réglé par le lemme de Schwarz.

Théorème (Lemme de Schwarz)

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, et $f : \Omega \to F$ une application différentiable sur Ω et deux fois différentiable en un point $a \in \Omega$. Alors $D^2 f(a)$, regardée comme application de $L_c(E, E; F)$, est symétrique. A savoir $D^2 f(a).(h, k) = D^2 f(a).(k, h)$ pour tout $(h, k) \in E \times E$.

Preuve : On démontre très brièvement ce théorème. On note $\Phi : E \times E \to F$ l'application

$$\Phi(h, k) = f(a + h + k) - f(a + h) - f(a + k) + f(a) \ .$$

Clairement Φ est symétrique en h et k. Pour ne pas alourdir la preuve on admet (ce n'est pas trivial) que

$$\|D^2 f(a).(h, k) - \Phi(h, k)\|_F = o\left(\|h\|_E^2 + \|k\|_E^2\right) \ ,$$

au sens où $o\left(\|h\|_E^2 + \|k\|_E^2\right) = \left(\|h\|_E^2 + \|k\|_E^2\right)\varepsilon(h, k)$ avec $\varepsilon(h, k) \to 0$ lorsque $(h, k) \to (0, 0)$.

Preuve suite : Par suite, avec cette relation,

$$\|D^2 f(a).(h,k) - D^2 f(a).(k,h)\|_F = o\left(\|h\|_E^2 + \|k\|_E^2\right) \ ,$$

et on va maintenant exploiter la non homogénéité de cette relation. On fixe $h, k \in E$ et on considère les vecteurs λh et λk, où $\lambda > 0$ est quelconque. Alors

$$\begin{aligned} &\|D^2 f(a).(\lambda h, \lambda k) - D^2 f(a).(\lambda k, \lambda h)\|_F \\ &= \lambda^2 \|D^2 f(a).(h,k) - D^2 f(a).(k,h)\|_F \\ &= \lambda^2 \left(\|h\|_E^2 + \|k\|_E^2\right) \varepsilon(\lambda) \ , \end{aligned}$$

où $\varepsilon(\lambda) \to 0$ lorsque $\lambda \to 0$. En particulier,

$$\|D^2 f(a).(h,k) - D^2 f(a).(k,h)\|_F = \left(\|h\|_E^2 + \|k\|_E^2\right) \varepsilon(\lambda)$$

pour tout $\lambda > 0$.

Preuve suite et fin : On fait alors tendre $\lambda \to 0$, et on récupère que

$$D^2 f(a).(h,k) = D^2 f(a).(k,h) \ .$$

Comme h et k sont quelconques, le théorème est démontré.
CQFD.

On définit la différentiabilité d'ordre n par récurrence avec une formule de récurrence du type $D^n = DD^{n-1}$. Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux espaces de Banach, et soit Ω un ouvert de E. Soit aussi $f : \Omega \to F$ une application différentiable sur Ω. On veut définir la différentielle $D^n f(a)$ d'ordre n de f en a comme application multilinéaire de $L_c(E^n; F)$. A savoir on veut

$$D^n f(a) \in L_c(E^n; F) \ ,$$

où $L_c(E^n; F)$ est l'espace des applications n-linéaires continues de $E \times \cdots \times E$ (n fois) dans F. Pour $n = 2$ on a déjà vu comment procéder. Pour $n = 3$, lorsque f est deux fois différentiables sur Ω, on regarde $D^2 f$ comme application

$$D^2 f : \Omega \to L_c(E, E; F) \ .$$

Par suite, $D^3 f(a) = DD^2 f(a)$ est une application

$$D^3 f(a) \in L_c\big(E, L_c(E, E; F)\big) \ .$$

Or on va facilement pouvoir assimiler $L_c\big(E, L_c(E, E; F)\big)$ à $L_c(E, E, E; F)$ en posant

$$\Phi(x, y, z) = \Phi(x).(y, z) \ .$$

Et on procède ainsi de suite par récurrence pour passer à $n = 4$, $n = 5$, etc. L'extension du lemme de Schwarz donne : si f est n fois différentiable en un point $a \in \Omega$, alors $D^n f(a)$ est symétrique. A savoir

$$D^n f(a).\big(h_1, \ldots, h_n\big) = D^n f(a).\big(h_{\sigma(1)}, \ldots, h_{\sigma(n)}\big)$$

pour tous $h_1, \ldots, h_n \in E$, et toute permutation σ de $\{1, \ldots, n\}$.

14. Formule de Taylor asymptotique

Théorème (Formule de Taylor asymptotique)

Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ deux espaces de Banach, Ω un ouvert de E, $f : \Omega \to F$ une application $(n-1)$-fois différentiable sur Ω et n-fois différentiable en un point $a \in \Omega$. Alors

$$f(a+h) = f(a) + Df(a).(h) + \cdots + \frac{1}{n!} D^n f(a).(h^n) + \|h\|_E^n \varepsilon(h)$$

pour tout h suffisamment petit, où $\varepsilon(h) \to 0$ lorsque $h \to 0$, et où la notation $D^n f(a).(h^n)$ tient pour
$D^n f(a).(h^n) = D^n f(a).(h, \ldots, h)$ (n-fois).

Si $n = 1$, on a juste écrit ici la définition de la différentiabilité au point a.