Calcul différentiel

Laurent Charles

19 décembre 2018

Ces notes recouvrent la deuxième partie d'un cours enseigné à l'ENS Paris de 2016 à 2018.

Table des matières

Dans ce qui suit, on travaille avec des espaces de Banach. Cependant, on utilisera la complétude seulement pour définir l'intégrale et pour les théorèmes d'inversion locale et des fonctions implicites.

1 Prologue

1.1 Intégration des fonctions à valeurs dans un espace de Banach

Soient F un espace de Banach et $a, b \in \mathbb{R}$ tels que $a < b$.

Soit $f : [a, b] \to F$ une fonction continue et $\delta : \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0} \cup \{\infty\}$ un module de continuité uniforme de f, c'est à dire que δ est croissante, sa limite en 0 est 0 et pout tout $x, y \in [a, b]$, l'on a

$$\|f(x) - f(y)\| \leqslant \delta(|x - y|).$$

Par exemple, on peut choisir $\delta(\epsilon) = \sup_{|x-y|\leqslant\epsilon} \|f(x) - f(y)\|$.

Si $\sigma = \{x_0 = a < x_1 < \ldots < x_n = b\}$ est une subdivision de $[a, b]$, on note $|\sigma| = \max_{i=0,\ldots,n} |x_i - x_{i-1}|$ le pas de la subdivision et $R(\sigma, f)$ la somme de Riemann

$$R(\sigma, f) = \sum_{i=1}^{n} (x_i - x_{i-1}) f(x_{i-1}) \in F.$$

Lemme 1.1.1. *Pour tout subdivisions σ et σ' de $[a, b]$, l'on a*

$$\|R(\sigma, f) - R(\sigma', f)\| \leqslant (b - a)(\delta(|\sigma|) + \delta(|\sigma'|)).$$

Démonstration. Comme $(b - a)f(a) = \sum_{i=1}^{n} (x_i - x_{i-1}) f(a)$, on a que

$$\begin{aligned} \big\|R(\sigma, f) - (b - a)f(a)\big\| &= \left\| \sum_{i=1}^{n} (x_i - x_{i-1})(f(x_{i-1}) - f(a)) \right\| \\ &\leqslant \sum_{i=1}^{n} (x_i - x_{i-1}) \delta(b - a) = (b - a)\delta(b - a) \end{aligned} \tag{1}$$

Soit la subdivision $\sigma'' = \sigma' \cup \sigma$ de $[a.b]$. Notons $\sigma = \{x_0 < x_1 < \ldots < x_n\}$. Alors $\sigma_i := \sigma'' \cap [x_{i-1}, x_i]$ est une subdivision de $[x_{i-1}, x_i]$. On a que $R(\sigma'', f) = \sum_{i=1}^{n} R(\sigma_i, f)$ et en appliquant l'inégalité (1) à chaque σ_i, il vient que

$$\big\|R(\sigma'', f) - R(\sigma, f)\big\| \leqslant \sum_{i=1}^{n} \big\|R(\sigma_i, f) - (x_i - x_{i-1}) f(x_{i-1})\big\|$$
$$\leqslant \sum_{i=1}^{n} (x_i - x_{i-1}) \delta(|x_i - x_{i-1}|) \leqslant (b-a)\delta(|\sigma|)$$

Par le même argument, $\|R(\sigma'', f) - R(\sigma', f)\| \leqslant (b-a)\delta(|\sigma'|)$, ce qui termine la preuve. □

Conséquence 1.1.2. Il existe un unique $u \in F$ tel que pour toute subdivision σ de $[a, b]$, l'on ait $\|u - R(\sigma, f)\| \leqslant (b-a)\delta(|\sigma|)$.

Démonstration. Soit σ_n la subdivision de $[a, b]$ en n intervalles de longueurs égales. $|\sigma_n| = (b-a)/n$ tend vers 0 quand $n \to \infty$, donc $\delta(|\sigma_n|)$ tend aussi vers 0. Par conséquent, si le vecteur u existe, c'est la limite de $R(\sigma_n, f)$. Pour établir l'existence, on remarque que $(R(\sigma_n, f))$ est une suite de Cauchy par le lemme 1.1.1. F étant supposé complet, $(R(\sigma_n, f))$ converge vers une limite u. Par le lemme 1.1.1, l'on a $\|R(\sigma, f) - R(\sigma_n, f)\| \leqslant (b-a)(\delta(|\sigma|) + \delta(|\sigma_n|))$ et en faisant tendre n vers l'infini, il vient $\|R(\sigma, f) - u\| \leqslant (b-a)\delta(|\sigma|)$. □

L'intégrale $\int_a^b f$ est par définition le vecteur u. Ensuite pour tout $x, y \in [a, b]$, on pose $\int_x^y f := \int_x^y f|_{[x,y]}$ si $x < y$, $\int_x^y f := 0$ si $x = y$ et $\int_x^y f := -\int_y^x f$ si $y < x$.

Proposition 1.1.3.

1. *si $g \in \mathcal{C}([a, b], F)$ et $\lambda \in \mathbb{R}$, $\int_x^y (g + \lambda f) = \int_x^y g + \lambda \int_x^y f$ pour tout $x, y \in [a, b]$.*
2. *pour tout $x, y, z \in [a, b]$, $\int_x^z f = \int_x^y f + \int_y^z f$.*
3. $\big\|\int_a^b f\big\| \leqslant \int_a^b \|f\|$

Démonstration. On utilise que les sommes de Riemann $R(\sigma, f)$ sont linéaires en f, vérifient une identité de Chasles et $\|R(\sigma, f)\| \leqslant R(\sigma, \|f\|)$. □

Une dernière remarque est que pour toute application linéaire continue $L : F \to G$ avec G un deuxième espace de Banach, l'on a

$$\int_x^y L \circ f = L\Big(\int_x^y f\Big)$$

On en déduit que pour $F = \mathbb{R}^n$ et $f = (f_1, \ldots, f_n)$, l'intégrale $\int_x^y f$ est le vecteur de $\mathbb{R}^n$ de coordonnées $\int_x^y f_i$.

1.2 Accroissements finis

Soit I un intervalle ouvert, F un espace de Banach.

Definition 1.2.1. *Une application $f : I \to F$ est dérivable en $x \in I$, de dérivée $u \in F$, si*

$$\frac{f(x+\delta)-f(x)}{\delta} \to u \qquad \textit{quand } \delta \to 0.$$

On note $u = f'(x)$.

De manière équivalente, f est dérivable en x de dérivée u si $f(x+\delta) = f(x) + \delta u + \mathrm{o}(|\delta|)$ quand $\delta \to 0$. On remarquera que si f est dérivable en x, alors elle est continue en x.

Proposition 1.2.2. *Soit $g : I \to F$ une application continue et $x_0 \in I$. Alors la fonction*

$$f(x) = \int_{x_0}^{x} g(t)dt, \qquad \forall x \in I$$

est dérivable en tout point et $f'(x) = g(x)$.

Démonstration. Le taux d'accroissement de f en x est $\delta^{-1}\int_x^{x+\delta} g$ et

$$\left\| \frac{1}{\delta}\int_x^{x+\delta} g(t)dt - g(x) \right\| = \left\| \frac{1}{\delta}\int_x^{x+\delta} (g(t)-g(x))dt \right\| \leqslant \epsilon(|\delta|)$$

où $\epsilon(r) = \sup_{|t|\leqslant r} |g(x+t)-g(x)|$ est l'oscillation de g en x. □

L'égalité des accroissements finis ne se généralise pas aux fonctions à valeurs dans un espace vectoriel de dimension $\geqslant 2$. Par contre, on a le

Théorème 1.2.3 (accroissements finis)**.** *Si $f : I \to F$ est dérivable en tout point, alors pour tout $x, y \in I$, on a*

$$\|f(x) - f(y)\| \leqslant C|y-x|$$

avec $C = \sup\big\{\|f'(x+t(y-x))\| /\ t \in [0,1]\big\}$ éventuellement infini.

Démonstration. On peut supposer C fini, sinon il n'y a rien à montrer. Soit $M > C$ et $J = \{t \in [0,1]/\ \|f(x+t(y-x))\| \leqslant Mt|y-x|\}$. Nous allons montrer que $1 \in J$. De là, on conclut en faisant tendre M vers C.

J est fermé, $0 \in J$ donc J admet un plus grand élément s (sa borne supérieure). Comme f est dérivable en $z := x + s(y-x)$, $f(z+\delta) = f(z) +$

$\delta(f'(z) + \mathrm{o}(1))$ quand $\delta \to 0$. Lorsque δ est assez petit, $\|f'(z) + \mathrm{o}(1)\| \leqslant M$ car $M > C$, donc

$$\|f(z+\delta) - f(z)\| \leqslant M|\delta| \tag{2}$$

On a alors pour tout $t \in [s, 1]$ suffisamment proche de s :

$$\begin{aligned}
&\|f(x + t(y-x)) - f(x)\| \\
&\leqslant \|f(x + t(y-x)) - f(x + s(y-x))\| + \|f(x + s(y-x)) - f(x)\| \\
&\leqslant M(t-s)|y-x| + Ms|y-x| = Mt|y-x|
\end{aligned}$$

la première majoration venant de (2) avec $\delta = (t-s)(y-x)$, la seconde car $s \in J$. Donc $t \in J$. Conclusion : $s = 1$ □

Conséquence 1.2.4.
1. Si $f : I \to F$ est dérivable avec $f' = 0$ en tout point, alors f est constante.
2. Si $f : I \to F$ est de classe $\mathcal{C}^1$ et $x_0 \in I$, alors

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt. \tag{3}$$

On remarquera que (3) implique à nouveau l'inégalité des accroissements finis, sous l'hypothèse plus forte que f est de classe $\mathcal{C}^1$,

2 Dérivée

Soient E et F deux espaces de Banach. On note $\mathcal{L}(E, F)$ l'espace des applications linéaires continues de E dans F. Rappelons que $\mathcal{L}(E, F)$ a une norme naturelle

$$\|L\|_{\mathcal{L}(E,F)} = \sup\left\{\|L(u)\|_F / \ \|u\|_E \leqslant 1\right\}$$

qui en fait un espace de Banach. On se donne aussi un ouvert U de E et on considérera des applications $f : U \to F$.

2.1 Définitions

Si V est un voisinage de l'origine dans E, $f : V \to F$ est une application et $h : V \to \mathbb{R}_{\geqslant 0}$ est une fonction qui ne s'annule pas en dehors de l'origine, on note $f(u) = \mathrm{o}(h(u))$ quand $u \to 0$ pour dire que $h(u)^{-1} f(u) \to 0$ quand $u \to 0$. Comme on considérera uniquement des o en l'origine, on omettra parfois de préciser $u \to 0$.

Definition 2.1.1. *Une application $f : U \to F$ est dérivable en $x \in U$ si il existe $L \in \mathcal{L}(E,F)$ tel que*

$$f(x+u) = f(x) + L(u) + \mathrm{o}(\|u\|) \tag{4}$$

quand $u \to 0$.

Si f est dérivable en x, $f(x+u) = f(x) + \mathrm{o}(1)$, c'est-à-dire que f est continue en x. On peut interpréter (4) comme le fait que les variations $f(x+u) - f(x)$ de f en x sont au premier ordre linéaire en u. L'application linéaire L est uniquement déterminée par f. Pour s'en convaincre, il suffit de vérifier que pour tout $M \in \mathcal{L}(E,F)$, $M(u) = \mathrm{o}(\|u\|)$ entraine $M = 0$. L'application L s'appelle la dérivée de f en x. On la note $f'(x)$.

Le fait d'être dérivable en x et la dérivée $f'(x)$ ne dépendent pas du choix des normes de E et F à équivalence près. Remarquons aussi que si f et g sont deux applications de U dans F dérivables en x, alors pour tout $\lambda \in \mathbb{R}$, $f + \lambda g$ est dérivable en x de dérivée $f'(x) + \lambda g'(x)$.

Lorsque $E = \mathbb{R}$, on dispose de deux définitions pour la dérivabilité : 1.2.1 et 2.1.1. Elles sont équivalentes. Dans un cas, la dérivée est un vecteur de F, dans l'autre cas, une application linéaire de $\mathbb{R}$ dans F. En fait, F et $\mathcal{L}(\mathbb{R}, F)$ sont naturellement isomorphes comme espaces vectoriels normés par l'application qui envoie $L \in \mathcal{L}(\mathbb{R}, F)$ sur $L(1) \in F$. Et cet isomorphisme envoie l'application linéaire dérivée sur le vecteur dérivé.

Si f est dérivable en x et $u \in E$, on appelle $f'(x)(u)$ la dérivée de f en x dans la direction u. Remarquons que $f'(x)(u)$ est le vecteur dérivée en 0 de l'application h définie au voisinage de l'origine dans $\mathbb{R}$ par $h(t) = f(x+tu)$.

Pour $E = \mathbb{R}^n$, on peut définir les dérivées partielles d'une applications f de U dans F : la i-ième dérivée partielle $\partial_i f(x) \in F$ de f en x s'obtient en dérivant par rapport à la i-ième coordonnée de $x = (x_1, \ldots, x_n)$, les autres coordonnées étant fixes :

$$\partial_i f(x) := \lim_{t \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) - f(x)}{t}. \tag{5}$$

Comme cas particulier du paragraphe précédent, si f est dérivable en x, alors elle admet des dérivées partielles en x données par $\partial_i f(x) = f'(x)(e_i)$ où l'on note (e_i) la base canonique de $\mathbb{R}^n$. La réciproque est fausse. Par exemple, la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y) = \frac{xy}{x^4 + y^4} \text{ si } (x,y) \neq (0,0) \qquad \text{et} \qquad f(0,0) = 0$$

admet des dérivées partielles en tout point (remarquer que la restriction de f à l'axe des abscisses ou des ordonnée est identiquement nulle). Par contre

f n'est pas dérivable en l'origine, car elle n'est pas bornée au voisinage de 0 donc pas même continue en 0.

Proposition 2.1.2. *$f : U \to E$ est dérivable en tout point $x \in U$ avec $f'(x) = 0$ si et seulement si f est localement constante.*

Démonstration. Pour le sens direct, on se donne $x \in U$ et $r > 0$ tel que $B(x, r) \subset U$. Supposons f dérivable en tout point. Alors pout tout u de norme $< r$, l'application $t \to f(x + tu)$ est dérivable sur $[0, 1]$ de dérivée $f'(x + tu)(u)$. En effet,

$$\begin{aligned} f(x + (t + \delta)u) &= f(x + tu) + f'(x + tu)(\delta u) + \mathrm{o}(\|\delta u\|) \\ &= f(x + tu) + \delta f'(x + tu)(u) + \mathrm{o}(|\delta|). \end{aligned}$$

Par conséquent, $f' = 0$ entraine que $f(x) = f(x + u)$, cf. conséquence 1.2.4 des accroissements finis. $\square$

Pour clore ce chapitre, donnons la définition des dérivées supérieures.

Definition 2.1.3. *$f : U \to F$ est continuement dérivable si elle est dérivable en tout point et l'application $f' : U \to \mathcal{L}(E, F)$, $x \to f'(x)$ est continue. On note $\mathcal{C}^1(U, F)$ l'espace des applications continuement dérivables de U dans F. La classe $\mathcal{C}^k$ est définie par récurrence sur k : f est de classe $\mathcal{C}^{k+1}$ si f est dérivable sur U et sa dérivée f' est de classe $\mathcal{C}^k$.*

On se convainc sans peine que $\mathcal{C}^k(U, E)$ est un sous-espace vectoriel de $\mathcal{C}^0(U, F)$.

2.2 Applications multilinéaires et composition

Comme on l'a vu, une application constante est dérivable de dérivée identiquement nulle. Une telle application est donc de classe $\mathcal{C}^\infty$. Une application linéaire $f \in \mathcal{L}(E, F)$ est dérivable en tout point $x \in E$, et sa dérivée est $f'(x) = f$. En effet, $f(x+u) = f(x)+f(u)$. Autrement dit, l'application dérivée f' est constante égale à f. On en déduit que f est de classe $\mathcal{C}^\infty$.

Ceci se généralise aux applications bilinéaires et même multilinéaires continues. Soient $E_1, \ldots, E_n$ des espaces de Banach. Le produit de ces espaces admet plusieurs normes équivalentes

$$\|u\|_\infty = \max_{i=1,\ldots,n} \|u_i\|_{E_i}, \qquad \|u\|_p = \left(\sum_{i=1}^n \|u_i\|_{E_i}^p\right)^{1/p}, \quad p \in [1, \infty).$$

Les plus simples pour travailler sont $\|\cdot\|_\infty$ ou $\|\cdot\|_1$.

Proposition 2.2.1. *Soit $f : E_1 \times \ldots \times E_n \to F$ une application multilinéaire. Alors les assertions suivantes sont équivalentes*

1. *f est continue*
2. *il existe $C > 0$ tel que $\|f(u_1, \ldots, u_n)\| \leqslant C\|u_1\| \ldots \|u_n\|$ pour tout $(u_1, \ldots, u_n) \in E_1 \times \ldots \times E_n$,*
3. *$\|f\| := \sup\big\{\|f(u_1, \ldots, u_n)\| / \ \|u_1\|, \ldots, \|u_n\| \leqslant 1\big\}$ est fini.*

De plus, si f est continue, alors f est de classe $\mathcal{C}^\infty$ et sa dérivée est donnée par $f'(x_1, \ldots, x_n)(u_1, \ldots, u_m) = f(u_1, x_2, \ldots, x_n) + f(x_1, u_2, x_3, \ldots, x_n) + \ldots + f(x_1, \ldots, x_{n-1}, u_n)$.

Remarquons aussi que $\|f\|$ est la plus petite constante C vérifiant la deuxième assertion. On peut aussi vérifier que $\|f\|$ définit bien une norme sur l'espace vectoriel des applications multilinéaires continues de $E_1 \times \ldots \times E_n$ dans F.

Démonstration. Si l'on munit $E_1 \times \ldots \times E_n$ de la norme infini, la boule unité est le produit $B_1 \times \ldots \times B_n$ des boules unitée des E_i. Si f est continue, elle est alors bornée au voisinage de l'origine, et donc il existe $r > 0$ tel que $f(r(B_1 \times \ldots \times B_n))$ soit contenue dans la boule unité de F. Ceci implique la deuxième assertion avec $C = 1/r$. Ceci implique à son tour que pour $x = (x_1, \ldots, x_n)$ et $u = (u_1, \ldots, u_n)$ de norme $\leqslant 1$, l'on a

$$\begin{aligned} f(x+u) = f(x) + f(u_1, x_2, \ldots, x_n) + f(x_1, u_2, x_3, \ldots, x_n) \\ + \ldots + f(x_1, \ldots, x_{n-1}, u_n) + \mathcal{O}(\|u\|^2) \end{aligned}$$

Ainsi f est dérivable donc continue. La dérivée étant la somme de n applications $(n-1)$-multilinéaires, l'on en déduit par une récurrence sur n que f est de classe $\mathcal{C}^\infty$. □

Proposition 2.2.2. *Si $f : U \to F$ est une application dérivable en $x \in U$, V un ouvert de F contenant $f(U)$ et $g : V \to G$ est dérivable en $f(x)$ avec G un espace de Banach, alors $g \circ f$ est dérivable en x et*

$$(g \circ f)'(x) = g'(f(x)) \circ f'(x). \tag{6}$$

De plus, si f et g sont de classe $\mathcal{C}^k$, $g \circ f$ l'est également.

Démonstration. Si f est dérivable en x, alors $f(x+u) - f(x) = v(u)$ avec

$$v(u) = f'(x)(u) + \|u\|\alpha(u) \qquad \text{où } \alpha(u) = \mathrm{o}(1).$$

Si g est dérivable en $f(x)$, alors

$$g(f(x)+v) = g(f(x)) + g'(f(x))(v) + \|v\|\beta(v) \qquad \text{avec } \beta(v) = \mathrm{o}(1).$$

Par conséquent

$$\begin{aligned} g(f(x+u)) = g(f(x)+v(u)) = g(f(x)) + g'(f(x))(f'(x)(u)) \\ + \|u\| g'(f(x))(\alpha(u)) + \|v(u)\|\beta(v(u)). \end{aligned}$$

On remarque que $g'(f(x))(\alpha(u)) = \mathrm{o}(1)$. De plus, $v(u) = \mathcal{O}(\|u\|)$ au voisinage de l'origine, donc $\|v(u)\|\beta(v(u)) = \mathrm{o}(\|u\|)$. Ceci montre que $g \circ f$ est dérivable en x avec la dérivée donnée par (6).

Supposons maintenant f dérivable sur U et g sur V. Alors d'après ce qui précède, $g \circ f$ est dérivable sur U et $(g \circ f)' = \varphi \circ \psi$ où $\psi : U \to \mathcal{L}(F,G) \times \mathcal{L}(E,F)$ envoie x sur $(g'(f(x)), f'(x))$ et $\varphi : \mathcal{L}(F,G) \times \mathcal{L}(E,F) \to \mathcal{L}(E,G)$ envoie (L,M) sur $L \circ M$. Remarquons que φ est bilinéaire continue donc de classe $\mathcal{C}^\infty$. On en déduit que par récurrence sur k que si f et g sont de classe $\mathcal{C}^k$, alors $g \circ f$ est de classe $\mathcal{C}^k$. Tout d'abord, f' et g' continue entraine que ψ est continue donc $\varphi \circ \psi = (g \circ f)'$ l'est aussi. Ensuite si f et g sont de classe $\mathcal{C}^{k+1}$, alors par hypothèse de récurrence $g' \circ f$ est de classe $\mathcal{C}^k$, donc ψ est de classe $\mathcal{C}^k$ par la proposition 2.3.1, donc $\varphi \circ \psi$ de classe $\mathcal{C}^k$ à nouveau par hypothèse de récurrence, autrement dit $g \circ f$ est de classe $\mathcal{C}^{k+1}$. $\square$

On notera les cas particuliers suivants.

1. Si I est un ouvert de $\mathbb{R}$, $h : I \to E$ admet pour vecteur dérivé u en $t \in I$, $h(I) \subset U$ et $f : U \to F$ est dérivable en $h(t)$, alors $f \circ h : I \to F$ est dérivable en t avec pour vecteur dérivée $f'(h(t))(u)$.
2. Si $f : U \to F$ est dérivable en x et $L \in \mathcal{L}(F,G)$, alors $L \circ f : U \to G$ est dérivable en x avec $(L \circ f)'(x) = L \circ f'(x)$.
3. Si $f_1 : U \to F_1$ et $f_2 \to F_2$ sont toutes deux dérivables en x et $B : F_1 \times F_2 \to G$ est une application bilinéaire continue, alors $g(x) = B(f_1(x), f_2(x))$ est dérivable en x et
$$g'(x)(u) = B(f_1'(x)(u), f_2(x)) + B(f_1(x), f_2'(x)(u)).$$

2.3 Espace de départ ou d'arrivée produit

Supposons $F = F_1 \times \ldots \times F_p$ avec les F_j des espaces de Banach. Alors F est lui-même un espace de Banach pour la norme $\|v\|_F = \max_{j=1,\ldots,p} \|v_j\|_{F_j}$.

Proposition 2.3.1. *Soient $f : U \to F$ et $f_1, \ldots, f_p$ les composantes de f. Alors f vérifie $\mathcal{P}$ si et seulement tous les f_j vérifient $\mathcal{P}$ avec $\mathcal{P} =$ « est continue en x », « est dérivable en x » ou « est de classe $\mathcal{C}^k$ ». De plus lorsque f est dérivable en x,*

$$f'(x)(h) = (f_1'(x)(h), \ldots, f_p'(x)(h)).$$

Démonstration. Tout est facile à vérifier. On fera notamment une récurrence sur k en utilisant que $\mathcal{L}(E, F)$ est isomorphe à $\mathcal{L}(E, F_1) \times \ldots \times \mathcal{L}(E, F_p)$ par l'application qui envoie L sur $(L_1, \ldots, L_p)$ où $L(u) = (L_1(u), \ldots, L_p(u))$. De plus $\|L\|_{\mathcal{L}(E,F)} = \max_{j=1,\ldots,p} \|L_j\|$ si la norme de F est définie comme précédemment. □

Supposons que $E = E_1 \times \ldots \times E_n$ est un espace de Banach produit. Soit $f : U \to F$ avec U un ouvert de E. On appelle i-ième dérivée partielle de f en $x \in U$ la dérivée en x_i de l'application $y \to f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$. On la note $D_i f(x)$. Elle appartient à $\mathcal{L}(E_i, F)$. Si la i-ième dérivée partielle existe en tout point, elle définit une application $D_i f : U \to \mathcal{L}(E_i, F)$.

Bien entendu, si f est dérivable en $x \in U$, elle admet des dérivées partielles en x et pour tout i, l'on a

$$D_i f(x)(h) = f'(x)(0, \ldots, 0, h, 0, \ldots, 0)$$

où le h est en i-ième position. La réciproque est fausse comme on l'a déjà vu, cf. paragraphe avant la proposition 2.1.2. Par contre, si l'on suppose les dérivées partielles continues, on peut remonter.

Théorème 2.3.2. *Si $f : U \to F$ admet des dérivées partielles en tout point qui sont continues en $x \in U$, alors f est dérivable en x.*

Pour la preuve et d'autres preuves plus tard, on utilisera la conséquence suivante de l'inégalité des accroissements finis.

Lemme 2.3.3. *Soit $h : U \to F$ dérivable en tout point. Soient $x \in U$ et $u \in E$ tel que le segment $[x, x+u]$ soit contenu dans U. Pour tout $L \in \mathcal{L}(E, F)$,*

$$\|h(x+u) - h(x) - L(u)\| \leqslant \|u\| \sup_{t \in [0,1]} \|h'(x+tu) - L\|. \tag{7}$$

Démonstration. On applique le théorème 1.2.3 à la fonction $t \to h(x+tu) - tL(u)$ □

Démonstration du théorème 2.3.2. Pour simplifier, on suppose $E = E_1 \times E_2$. On applique à deux reprises la conséquence (7) des accroissements finis :

$$\|f(x_1+u_1, x_2+u_2) - f(x_1+u_1, x_2) - D_2 f(x_1, x_2)(u_2)\| \leqslant C_2 \|u_2\| \quad (8)$$

avec $C_2 = \sup_{t\in[0,1]} \|D_2 f(x_1+u_1, x_2+tu_2) - D_2 f(x_1, x_2)\|$. Et

$$\|f(x_1+u_1, x_2) - f(x_1, x_2) - D_1 f(x_1, x_2)(u_1)\| \leqslant C_1 \|u_1\| \quad (9)$$

avec $C_1 = \sup_{t\in[0,1]} \|D_1 f(x_1+tu_1, x_2) - D_1 f(x_1, x_2)\|$. En sommant (8) et (9) et en notant $x = (x_1, x_2)$ et $u = (u_1, u_2)$, il vient

$$\|f(x+u) - f(x) - D_1 f(x)(u_1) - D_2 f(x)(u_2)\| \leqslant C_1(x,u)\|u_1\| + C_2(x,u)\|u_2\|$$

avec $C_1(x,u)$ et $C_2(x,u)$ les constantes précédentes. Pour conclure, $D_1 f$ et $D_2 f$ continues en x entraine que $C_1(x,u)$ et $C_2(x,u)$ tendent vers 0 lorsque $u \to 0$. $\square$

Proposition 2.3.4. *Soit $k \in \mathbb{N}$. Alors $f : U \to F$ est de classe $\mathcal{C}^{k+1}$ si et seulement si elle admet des dérivées partielles $D_1 f$, ..., $D_n f$ en tout point qui sont de classe $\mathcal{C}^k$.*

Démonstration. On a un isomorphisme d'espaces vectoriels

$$\mathcal{L}(E,F) \simeq \mathcal{L}(E_1,F) \times \ldots \times \mathcal{L}(E_n,F)$$

qui envoie L sur $(L_1, \ldots, L_n)$ si $L(h) = L_1(h_1) + \ldots + L_n(h_n)$. Si de plus la norme de E est définie par $\|h\| = \max_{i=1,\ldots,n} \|h_i\|_{E_i}$, alors les normes correspondantes d'applications linéaires sont

$$\|L\|_{\mathcal{L}(E,F)} = \sum_{i=1}^{n} \|L_i\|_{\mathcal{L}(E_i,F)}.$$

Et inversement, si on définit la norme de h comme la somme des $\|h_i\|$, la norme de L est le maximum des $\|L_i\|$. En fait cela n'a aucune importance pour la suite, car pour un produit fini, les normes définis comme un maximum ou une somme sont équivalentes.

D'après le théorème 2.3.2, on peut supposer pour la preuve que f est dérivable et $f' = (D_1 f, \ldots, D_n f)$. On a alors $f \in \mathcal{C}^{k+1}(U,F)$ ssi $f' \in \mathcal{C}^k(U,F)$ ssi pour tout i, $D_i f \in \mathcal{C}^k(U, \mathcal{L}(E_i, F)$ d'après la proposition 2.3.1 sur les espaces d'arrivées produits. $\square$

Si $E = \mathbb{R}^n$ et $i = 1, \ldots, n$, la dérivée partielle $D_i f(x) \in \mathcal{L}(\mathbb{R}, F)$ s'identifie au vecteur dérivée partielle $\partial_i f(x) \in F$ défini en (5). Si $i_1, \ldots, i_\ell \in \{1, \ldots, n\}$, on appelle $\partial_{i_1} \ldots \partial_{i_\ell} f(x)$ une dérivée partielle d'ordre ℓ.

Proposition 2.3.5. *Si $E = \mathbb{R}^n$, alors f est de classe $\mathcal{C}^k$ ssi toutes les dérivées partielles de f d'ordre $\leqslant k$ existent et sont continues. Si de plus $F = \mathbb{R}^p$ et $f = (f_1, \ldots, f_p)$, alors f est de classe $\mathcal{C}^k$ ssi pour tout $j = 1, \ldots, p$, toutes les dérivées partielles de f_j d'ordre $\leqslant k$ existent et sont continues*

Démonstration. La première assertion s'obtient par applications répétées de la proposition 2.3.4. On passe à la seconde assertion par la proposition 2.3.1. □

3 Dérivées supérieures

3.1 Le lemme de Schwarz

Théorème 3.1.1. *Soit $f : U \to E$ dérivable en tout point et $x \in U$ tel que f admette une dérivée seconde en x. Alors pour tout $u, v \in E$, on a*

$$f''(x)(u)(v) = f''(x)(v)(u).$$

La preuve est basée sur le lemme suivant.

Lemme 3.1.2. *Sous les mêmes hypothèses, l'on a*

$$f(x+u+v)-f(x+u)-f(x+v)+f(x) = f''(x)(v)(u)+\epsilon_x(u,v)\|u\|(\|u\|+\|v\|)$$

où $\epsilon_x(u, v) = \mathrm{o}(1)$ quand (u, v) tend vers 0.

Preuve du Théorème 3.1.1. D'aprés le lemme 3.1.2, l'on a

$$f(x + tu + tv) - f(x + tu) - f(x + tv) + f(x) = t^2 f''(x)(v)(u) + \mathrm{o}(t^2)$$

et le terme de gauche est symmétrique en u, v. □

Preuve du lemme 3.1.2. On introduit la différence $\Delta_u f(x) = f(x+u)-f(x)$. On remarque que

$$\Delta_u \Delta_v f(x) = f(x + u + v) - f(x + u) - f(x + v) + f(x).$$

En appliquant la conséquence (7) de l'inégalité des accroissements finis à la fonction $h = \Delta_v f$ et $L = f''(x)(v)$, il vient

$$\|\Delta_u \Delta_v f(x) - f''(x)(v)(u)\| \leqslant \|u\| \sup_{t\in[0,1]} \|\Delta_v g(x + tu) - g'(x)(v)\|$$

où l'on a noté $g = f'$ de sorte que $(\Delta_v f)' = \Delta_v g$ et $L = g'(x)(v)$. Par hypothèse, g est dérivable en x, donc $g(x+w) = g(x) + g'(x)(w) + \epsilon(w)\|w\|$. Alors

$$\begin{aligned}\Delta_v g(x+tu) - g'(x)(v) &= g(x+tu+v) - g(x+tu) - g'(x)(v)\\ &= g(x) + g'(x)(tu+v) + \epsilon(tu+v)\|tu+v\|\\ &\quad - g(x) - g'(x)(tu) - \epsilon(tu)\|tu\| - g'(x)(v)\\ &= \epsilon(tu+v)\|tu+v\| - \epsilon(tu)\|tu\|\end{aligned}$$

Pour $t \in [0,1]$, $\|tu+v\| \leqslant \|u\| + \|v\|$ et $\|tu\| \leqslant \|u\| \leqslant \|u\| + \|v\|$. Enfin $\sup_{t\in[0,1]} \|\epsilon(tu+v)\|$ et $\sup_{t\in[0,1]} \|\epsilon(tu)\|$ sont tous deux des $\mathrm{o}(1)$ quand (u,v) tend vers 0. $\square$

Lorsque f est de classe $\mathcal{C}^2$, on a que

$$\Delta_u\Delta_v f(x) = f''(x)(v)(u) + \mathrm{o}(\|u\|\|v\|).$$

Pour le voir, il suffit d'appliquer à deux reprises la conséquence (7) des accroissements finis. Plus généralement si f est de classe $\mathcal{C}^k$, l'on a

$$\Delta_{u_k}\ldots\Delta_{u_1} f(x) = f^{(k)}(x)(u_1)\ldots(u_k) + \mathrm{o}(\|u_1\|\ldots\|u_k\|).$$

lorsque $(u_1,\ldots,u_k) \to 0$ dans E^k.

3.2 Interlude algébrique

Si $f : U \to F$ admet une dérivée k-ième en x, celle-ci appartient à $\mathcal{L}^k(E,F) := \mathcal{L}(E,\mathcal{L}(E,\ldots,\mathcal{L}(E,F)\ldots))$, où on a répété k fois le E.

$\mathcal{L}^k(E,F)$ est naturellement isomorphe comme espace vectoriel normé à l'espace $\mathcal{L}_k(E,F)$ des applications k-multilinéaires de E^k dans F qui sont continues : on envoie $\varphi \in \mathcal{L}^k(E,F)$ sur $\psi \in \mathcal{L}_k(E,F)$ donné par

$$\psi(u_1,\ldots,u_k) = \varphi(u_1)(u_2)\ldots(u_k).$$

La norme sur $\mathcal{L}_k(E,F)$ est définie comme dans la proposition 2.2.1 par

$$\|\psi\| = \sup\{\|\psi(u_1,\ldots,u_k)\| / \ \|u_1\| = \ldots = \|u_k\| = 1\}.$$

Pour vérifier qu'il s'agit bien d'un isomorphisme, il semble plus prudent de faire une récurrence et montrer que $\mathcal{L}(E,\mathcal{L}_k(E,F)) \simeq \mathcal{L}_{k+1}(E,F)$ par l'application qui envoie φ sur $\psi(u_0,\ldots,u_k) = \varphi(u_0)(u_1,\ldots,u_k)$.

Proposition 3.2.1. *Soit $f : U \to F$ qui admet une dérivée k-ième $f^{(k)}(x) \in \mathcal{L}_k(E, F)$ en $x \in U$. Alors cette dérivée est symétrique, c'est à dire que pour toute permutation σ de $\{1, \ldots, k\}$, l'on a que*

$$f^{(k)}(x)(u_{\sigma(1)}, \ldots, u_{\sigma(k)}) = f^{(k)}(x)(u_1, \ldots, u_k).$$

Démonstration. On le montre par récurrence sur k. Si $k = 2$, c'est le théorème 3.1.1. Supposons que f admette une dérivée $(k+1)$-ième en x. Pour toute permutation σ de $\{1, \ldots, k\}$, introduisons l'application

$$L_\sigma : \mathcal{L}_k(E, F) \to \mathcal{L}_k(E, F), \qquad L_\sigma(\psi)(u_1, \ldots, u_k) = \psi(u_{\sigma(1)}, \ldots, u_{\sigma(k)})$$

Par hypothèse de récurrence, $L_\sigma \circ f^{(k)} = f^{(k)}$. Remarquons que L_σ est linéaire continue. En dérivant la relation précédente, il vient que $L_\sigma f^{(k+1)}(x)(u_0) = f^{(k+1)}(x)(u_0)$, autrement dit

$$f^{(k+1)}(x)(u_0, u_{\sigma(1)}, \ldots, u_{\sigma(k)}) = f^{(k+1)}(x)(u_0, u_1, \ldots, u_k).$$

Par ailleurs, le théorème 3.1.1 implique que

$$f^{(k+1)}(x)(u_0, u_1, u_2, \ldots, u_k) = f^{(k+1)}(x)(u_1, u_0, u_2, \ldots, u_k).$$

Pour conclure, le groupe des permutations de $\{0, \ldots, k\}$ est engendré par la transposition $(0, 1)$ et les permutations de $\{1, \ldots, k\}$. $\square$

3.3 Interprétation géométrique

Soit $f : U \to F$ qui admet une dérivée k-ième en $x \in U$.

Proposition 3.3.1. *Pour tout $u_1, \ldots, u_k \in E$, l'on a que*

$$f^{(k)}(x)(u_1, \ldots, u_k) = \frac{\partial^k \varphi}{\partial t_1 \ldots \partial t_k}(0, \ldots, 0)$$

où φ est la fonction définie au voisinage de l'origine dans $\mathbb{R}^k$ par $\varphi(t_1, \ldots, t_k) = f(x + t_1 u_1 + \ldots + t_k u_k)$.

Démonstration. On fait une récurrence sur k. Soit $\varphi(t_0, \ldots, t_k) = f(x + t_0 u_0 + t_1 u_1 + \ldots + t_k u_k)$. Par hypothèse de récurrence,

$$\frac{\partial^k \varphi}{\partial t_1 \ldots \partial t_k}(t_0, 0, \ldots, 0) = f^{(k)}(x + t_0 u_0)(u_1, \ldots, u_k).$$

Pour calculer ensuite la dérivé par rapport à t_0 du terme de droite, on le considère comme la composition des applications $t_0 \to x + t_0 u_0$, $f^{(k)}$ et $\mathcal{L}_k(E, F) \to F$, $\varphi \to \varphi(u_1, \ldots, u_k)$ qui est linéaire continue. $\square$

On en déduit pour $E = \mathbb{R}^n$ et (e_i) la base canonique de $\mathbb{R}^n$ que

$$\partial_{i_1} \dots \partial_{i_k} f(x) = f^{(k)}(x)(e_{i_1}, \dots, e_{i_k}). \tag{10}$$

Donc par la proposition 3.2.1, $\partial_{i_1} \dots \partial_{i_k} f(x)$ est symétrique en $(i_!, \dots, i_k)$. Pour cette raison, on a que

$$\partial_{i_1} \dots \partial_{i_k} f(x) = \partial^\alpha f(x)$$

où $\alpha \in \mathbb{N}^n$ est le multi-indice donné par $\alpha(j) = \sharp\{\ell/\ i_\ell = j\}$ pour tout $j \in \{1, \dots, n\}$ et

$$\partial^\alpha f(x) := \partial_1^{\alpha(1)} \dots \partial_n^{\alpha(n)} f(x). \tag{11}$$

3.4 Formule de Taylor

On donne un énoncé sous des hypothèses assez fortes, à savoir $f : U \to F$ de classe $\mathcal{C}^{k+1}$.

Proposition 3.4.1. *Soit $x \in U$ et $u \in E$ tel que le segment $[x, x+u]$ soit contenu dans U. Alors*

$$f(x+u) = \sum_{\ell=0}^{k} \frac{1}{\ell!} f^{(\ell)}(x)(u, \dots, u) + R_k(x,u)$$

avec le reste $R_k(x,u)$ qui vérifie

$$R_k(x,u)\| \leqslant \frac{\|u\|^{k+1}}{(k+1)!} \sup_{t\in[0,1]} \|f^{(k+1)}(x+tu)\|.$$

Démonstration. Si $\epsilon > 0$ et $h :]-\epsilon, 1+\epsilon[\to F$ est une application de classe $\mathcal{C}^{k+1}$, on rappelle la formule bien connue

$$h(1) = \sum_{\ell=0}^{k} \frac{h^{(\ell)}(0)}{\ell!} + R_k \qquad \text{où} \qquad R_k = \frac{1}{k!} \int_0^1 h^{(k+1)}(t)(1-t)^k dt.$$

Ceci se démontre par récurrence sur k à coup d'intégrations par partie, qui sont bien justifiées d'après (3). Notons que

$$\|R_k\| \leqslant \left(\sup_{[0,1]} \|h^{(k+1)}\|\right) \frac{1}{k!} \int_0^1 (1-t)^k\, dt = \left(\sup_{[0,1]} \|h^{(k+1)}\|\right) \frac{1}{(k+1)!}.$$

On applique ceci à la fonction $h(t) = f(x+tu)$. Sa dérivée ℓ-ième est

$$h^{(\ell)}(t) = f^{(\ell)}(x+tu)(u, \ldots, u). \tag{12}$$

Cette dernière formule s'établit par récurrence sur ℓ. On dérive (12) en considérant la composition de $t \to x+tu$, $f^{(\ell)}$ et $\mathcal{L}_k(E,F) \to F$, $\varphi \to \varphi(u,\ldots,u)$, cette dernière application est linéaire continue. □

Dans $\mathbb{R}^n$, les termes de la formule de Taylor se réécrivent comme suit.

Proposition 3.4.2. *Si $E = \mathbb{R}^n$ et $u = (u_1, \ldots, u_n)$, alors*

$$\frac{1}{\ell!} f^{(\ell)}(x)(u, \ldots, u) = \frac{1}{\ell!} \sum_{i_1, \ldots, i_\ell = 1}^{n} (\partial_{i_1} \ldots \partial_{i_k} f)(x) u_{i_1} \ldots u_{i_\ell} \tag{13}$$

$$= \sum_{\alpha \in \mathbb{N}^n,\ |\alpha| = \ell} \frac{1}{\alpha!} \partial^\alpha f(x) u^\alpha \tag{14}$$

où $|\alpha| = \alpha(1) + \ldots + \alpha(n)$, $\partial^\alpha f(x)$ définie comme en (11) et $u^\alpha = u_1^{\alpha(1)} \ldots u_n^{\alpha(n)}$.

Démonstration. Pour la première égalité, on remplace chaque u par $\sum u_i e_i$, on développe et on utilise (10). Pour obtenir la seconde égalité, il faut compter pour chaque multi-indice α les ℓ-uplets $(i_1, \ldots, i_\ell)$ tels que $\alpha(j) = \sharp\{\ell / i_\ell = j\}$ pour tout j. Pour se faire, considérer l'action du groupe symétrique S_ℓ sur $\{1, \ldots, \ell\}$, chaque α correspond à une orbite, d'isotropie $S_{\alpha(1)} \times \ldots \times S_{\alpha(k)}$, donc de cardinal $\ell!/\alpha!$. □

Une dernière remarque : pour établir la formule de Taylor lorsque $E = \mathbb{R}^n$ avec les termes sous la forme (13), au lieu de passer par (12), on peut calculer les dérivées successives de $h(t) = f(x+tu)$ en utilisant que $h'(t) = \sum_i u_i \partial_i f(x+tu)$ et en réitérant.

4 Inversion locale et fonctions implicites

4.1 La fonction inverse des applications linéaires

Notons $\mathcal{L}(E) = \mathcal{L}(E,E)$. Un élément de $\mathcal{L}(E)$ est dit inversible s'il admet un inverse dans $\mathcal{L}(E)$.

Proposition 4.1.1. *L'ensemble U des éléments inversibles de $\mathcal{L}(E)$ est un ouvert. L'application $\varphi : U \to \mathcal{L}(E)$ qui envoie L sur L^{-1} est de classe $\mathcal{C}^\infty$ et*

$$\varphi'(L)(u) = -L^{-1} u L^{-1} \tag{15}$$

pour tout $L \in U$, $u \in \mathcal{L}(E)$.

Pour la preuve, on utilise essentiellement que

$$\|LM\| \leqslant \|L\| \, \|M\|, \qquad \forall L, M \in \mathcal{L}(E).$$

Démonstration. Si $u \in \mathcal{L}(E)$ est de norme < 1, alors la série

$$f(u) = \sum_{n=0}^{\infty} (-1)^n u^n$$

converge normalement. De plus, $(\mathrm{id}_E + u) f(u) = f(u)(\mathrm{id}_E + u) = \mathrm{id}_E$. Donc U contient la boule ouverte $B(\mathrm{id}_E, 1)$ et $\varphi(\mathrm{id}_{\mathrm{E}} + u) = f(u)$. Enfin,

$$f(u) = \mathrm{id}_E - u + \mathcal{O}(\|u\|^2) \tag{16}$$

donc φ est dérivable en id_E et $\varphi'(\mathrm{id}_E)(u) = -u$.

Soit maintenant $L \in \mathcal{L}(E)$ inversible. Alors pour tout $u \in \mathcal{L}(E)$,

$$L + u = L(\mathrm{id}_E + L^{-1}u). \tag{17}$$

On en déduit que $B(L, R_L) \subset U$ avec $R_L = \|L^{-1}\|^{-1}$. En effet $\|u\| < R_L$ entraine $\|L^{-1}u\| < 1$, donc $\mathrm{id}_E + L^{-1}u$ est inversible et par (17), $L + u$ est inversible. Donc U est un ouvert. De plus, (17) implique

$$\begin{aligned} (L+u)^{-1} &= (\mathrm{id}_E + L^{-1}u)^{-1} L^{-1} = f(L^{-1}u) L^{-1} \\ &= (\mathrm{id}_E - L^{-1}u + \mathcal{O}(\|L^{-1}u\|^2)) L^{-1} = L^{-1} - L^{-1}uL^{-1} + \mathcal{O}(\|u\|^2) \end{aligned}$$

où l'on a utilisé (16) et le $\mathcal{O}$ est uniforme pour u dans un voisinage de l'origine. Nous venons de montrer que φ est dérivable sur U et que la dérivée est bien donnée par (15).

Pour finir nous montrons par récurrence sur k que φ est de classe $\mathcal{C}^k$. Nous avons $\varphi' = B \circ \psi$ avec $\psi : U \to \mathcal{L}(E) \times \mathcal{L}(E)$, $L \to (L^{-1}, L^{-1})$ et $B : \mathcal{L}(E) \times \mathcal{L}(E) \to \mathcal{L}(\mathcal{L}(E))$ telle que $B(M, N)(u) = -MuN$. On remarque que B est bilinéaire continue donc de classe $\mathcal{C}^\infty$. D'autre part, φ est dérivable donc continue, donc ψ est continue, donc φ' est continue, donc φ est de classe $\mathcal{C}^1$. Supposons φ de classe $\mathcal{C}^k$, alors ψ est aussi $\mathcal{C}^k$, ainsi que $B \circ \psi$, donc φ est de classe $\mathcal{C}^{k+1}$. $\square$

4.2 Inversion locale

Definition 4.2.1. *Soient E et F des espaces de Banach, U et V des ouverts de E et F respectivement et k un entier $\geqslant 1$. Une application $f : U \to V$ est un difféomorphisme de classe $\mathcal{C}^k$ si elle est bijective, de classe $\mathcal{C}^k$ ainsi que son inverse.*

Soit $f : U \to V$ un difféomorphisme de classe $\mathcal{C}^k$. Alors

1. pour tout ouvert $U' \subset U$, f se restreint en un difféomorphisme de classe $\mathcal{C}^k$ de U' dans $V' = f(U')$.
2. pour tout $x \in U$, $f'(x) : E \to F$ est inversible d'inverse $(f^{-1})'(f(x)) : F \to E$. Pour le voir, il suffit de dériver les deux relations $\mathrm{id}_U = f^{-1} \circ f$ et $\mathrm{id}_V = f \circ f^{-1}$.

Une autre remarque facile est que la composition de deux difféomorphismes de classe $\mathcal{C}^k$ est un difféomorphisme de classe $\mathcal{C}^k$.

Théorème 4.2.2. *Soient U un ouvert de E et $f : U \to F$ de classe $\mathcal{C}^k$ avec $k \geqslant 1$. On suppose qu'il existe $x_0 \in U$ tel que $f'(x_0)$ admette un inverse dans $\mathcal{L}(F, E)$. Alors il existe des ouverts U' et V de E et F respectivement tels que $x_0 \in U'$, $U' \subset U$, $f(U') = V$ et f se restreigne en un difféomorphisme de classe $\mathcal{C}^k$ de U' dans V.*

Considérons $f : U \to F$ de classe $\mathcal{C}^1$ telle que pour tout $x \in U$, $f'(x)$ est inversible. Alors d'après le théorème 4.2.2, f est une application ouverte. Donc $V = f(U)$ est ouvert dans F et f est bien entendu surjective de U dans V. Toujours d'après le théorème 4.2.2, f est localement injective. Cependant, il se peut que f ne soit pas injective. Par exemple, on prendra $U =]0, 1[\cup]2, 3[$ et $f(x) = x$ si $x < 1$, $x - 2$ sinon. Pour un exemple avec U connexe, on pourra considérer $U = \mathbb{C} \setminus \{0\}$ et $f(z) = z^2$.

Démonstration. 1. On se ramène à $E = F$, $x_0 = 0$, $f(x_0) = 0$ et $f'(x_0) = \mathrm{id}_E$. Pour ce faire, utiliser que les translations et isomorphismes des espaces de Banach sont des difféomorphismes de classe $\mathcal{C}^\infty$. On peut alors passer de f à $x \to f(x_0 + x) - f(x_0)$ qui est la composée de $x \to x_0 + x$, f et $y \to y - f(x_0)$. On peut ensuite passer de f à $L \circ f$ où $L = (f^{-1})'(x_0) : F \to E$.

2. Introduisons $g(x) = x - f(x)$. Alors $g'(0) = 0$ donc il existe $r > 0$ tel que $\overline{B}_r \subset U$ et pour tout $x \in \overline{B}_r$, $\|g'(x)\| \leqslant 1/2$. Ici on note B_r la boule ouverte $\{x \in E / \|x\| < r\}$ et $\overline{B}_r$ son adhérence. Par le théorème des accroissements finis, il vient que pour tout $x_1, x_2 \in \overline{B}_r$,

$$\|g(x_1) - g(x_2)\| \leqslant \frac{1}{2}\|x_1 - x_2\| \tag{18}$$

Nous en déduisons que

$$\|x_1 - x_2\| \leqslant 2\|f(x_1) - f(x_2)\| \tag{19}$$

En effet, $x = f(x) + g(x)$, donc $\|x_1 - x_2\| \leqslant \|f(x_1) - f(x_2)\| + \|g(x_1) - g(x_2)\| \leqslant \|f(x_1) - f(x_2)\| + \frac{1}{2}\|x_1 - x_2\|$ d'après (18) et on retranche ensuite $\frac{1}{2}\|x_1 - x_2\|$ des deux côtés.

3. Montrons que pour tout $y \in B_{r/2}$, il existe un unique $x \in B_r$ tel que $f(x) = y$. D'après (18), g envoie $\overline{B}_r$ dans $\overline{B}_{r/2}$. Donc g_y définie par

$$g_y(x) = y + x - f(x) = y + g(x)$$

envoie $\overline{B}_r$ dans lui-même. De plus, pour tout x_1, x_2 dans $\overline{B}_r$, $\|g_y(x_1) - g_y(x_2)\| = \|g(x_1) - g(x_2)\| \leqslant \frac{1}{2}\|x_1 - x_2\|$ d'après (18). Donc le théorème du point fixe s'applique à la restriction de g_y à $\overline{B}_r$ et il existe un unique $x \in \overline{B}_r$ tel que $g_y(x) = x$, c'est à dire $f(x) = y$. D'après (19), x appartient à B_r.

4. Soit $V = B_{r/2}$ et $U' = f^{-1}(B_{r/2}) \cap B_r$. D'après l'étape précédente, f se restreint en une bijection de U' dans V. De plus U' et V sont clairement ouverts et contiennent tous deux l'origine. Enfin, d'après (19), l'inverse $f^{-1} : V \to U$ est continue.

5. Il reste à montrer que f^{-1} est de classe $\mathcal{C}^k$. On commence par la dérivabilité. Soit $x \in U'$ et $y = f(x)$. Comme $x \in B_r$, $\|g'(x)\| \leqslant 1/2 < 1$, et donc $f'(x) = \mathrm{id}_E - g'(x)$ est inversible dans $\mathcal{L}(E)$. Montrons que f^{-1} est dérivable en y de dérivée $L := (f'(x))^{-1}$. Pour tout $v \in E$ suffisamment petit de sorte que $y + v \in V$, on note $u(v) = f^{-1}(y+v) - f^{-1}(y)$. Nous avons que $y + v = f(f^{-1}(y+v)) = f(x + u(v))$. Donc

$$v = f(x + u(v)) - f(x) = f'(x)(u(v)) + \epsilon(u(v))\|u(v)\|$$

où $\epsilon(u) = \mathrm{o}(1)$ en l'origine. En appliquant L, il vient

$$u(v) = L(v) - \|u(v)\| L(\epsilon(u(v)).$$

D'après (19), $\|u(v)\| = \|f^{-1}(y+v) - f^{-1}(y)\| \leqslant 2\|v\|$. Donc $\|u(v)\| L(\epsilon(u(v)) = \mathrm{o}(\|v\|)$ quand $v \to 0$. Nous avons donc montré que f^{-1} est dérivable en y.

Comme $(f^{-1})' = I \circ f' \circ f^{-1}$ où I est l'inversion de $\mathcal{L}(E)$, $(f^{-1})'$ est continue donc f^{-1} de classe $\mathcal{C}^1$. En utilisant cette même relation et le fait que I est de classe $\mathcal{C}^\infty$, on montre par récurrence que f de classe $\mathcal{C}^k$ entraine f^{-1} de classe $\mathcal{C}^k$. □

4.3 Fonctions implicites

On se donne E, F et G espaces de Banach, W un ouvert de $E \times F$ et $f : W \to G$ une application dérivable en $(x, y) \in W$. Alors la dérivée se décompose en somme des dérivées partielles

$$f'(x,y)(u,v) = D_1 f(x,y)(u) + D_2 f(x,y)(v), \qquad (u,v) \in E \times F$$

avec $D_1 f(x,y) \in \mathcal{L}(E,G)$ et $D_2 f(x,y) \in \mathcal{L}(F,G)$.

Théorème 4.3.1. *On suppose f de classe $\mathcal{C}^k$ avec $k \geqslant 1$. Soit $(x_0, y_0) \in W$ tel que $f(x_0, y_0) = 0$ et $D_2 f(x_0, y_0)$ soit inversible d'inverse continue. Alors il existe un voisinage ouvert U de x_0, un voisinage ouvert V de y_0 et une application $g : U \to V$ de classe $\mathcal{C}^k$ tels que $U \times V \subset W$, $g(x_0) = y_0$ et*

$$\forall (x, y) \in U \times V, \qquad f(x, y) = 0 \Leftrightarrow y = g(x). \tag{20}$$

On interprète généralement x comme un paramètre, y une inconnue et f une équation. En particulier pour $E = \mathbb{R}^k$, $F = G = \mathbb{R}^n$, on est ramené à des paramètres, inconnues et équations scalaires : $y = (y_1, \dots, y_k)$, $x = (x_1, \dots, x_n)$ et $f = (f_1, \dots, f_n)$. On a autant d'équations que d'inconnues, par contre le nombre de paramètres est arbitraire.

Si f est linéaire continue de $E \times F$ dans G, l'hypothèse est que $f(0, \cdot) : F \to G$ est inversible. Si l'on note λ son inverse, alors

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

avec $g(x) = -\lambda(f(x, 0))$.

Tentative de preuve : on considére x comme un paramètre et on applique le théorème d'inversion locale à $h_x(y) = f(x, y)$ de sorte à pouvoir écrire $h_x(y) = 0$ ssi $y = h_x^{-1}(0)$. Comme $h_x'(y) = D_2 f(x, y)$, l'hypothèse $D_2 f(x_0, y_0)$ inversible nous garantit que $h_x'(y)$ est inversible lorsque (x, y) décrit un voisinage de (x_0, y_0). On en déduit que h_x réalise un difféomorphisme d'un voisinage ouvert U_x de x_0 sur un ouvert V_x. Le problème est que rien ne nous assure que $0 \in V_x$ lorsque $x \neq x_0$. Pour régler ceci, nous allons appliquer le théorème d'inversion locale sur $E \times F$.

Démonstration. Quitte à remplacer f par $\lambda \circ f$ où $\lambda : G \to F$ est l'inverse de $D_2 f(x_0, y_0)$, on a que $F = G$ et $D_2 f(x_0, y_0) = \mathrm{id}_F$. Soit $\varphi : W \to E \times F$ qui envoie (x, y) sur $(x, f(x, y))$. Alors φ est de classe $\mathcal{C}^k$ et sa dérivée en (x, y) est l'application de $E \times F \simeq E \oplus F$ dans lui-même

$$\begin{bmatrix} \mathrm{id}_E & 0 \\ A & \mathrm{id}_F \end{bmatrix} \qquad \text{avec } A = D_1 f(x_0, y_0).$$

Cette dérivée est inversible d'inverse $\begin{bmatrix} \mathrm{id}_E & 0 \\ -A & \mathrm{id}_F \end{bmatrix}$. Donc par le théoréme d'inversion locale, il existe U' voisinage ouvert de x_0 et V voisinage ouvert de y_0 tels que $\varphi(U' \times V)$ soit ouvert et φ se restreint en un difféomorphisme de $U' \times V$ sur $\varphi(U' \times V)$. Remarquons que l'inverse de φ est de la forme

$$\forall (x, y) \in \varphi(U' \times V), \qquad \varphi^{-1}(x, y) = (x, h(x, y)) \quad \text{avec } h(x, y) \in V.$$

En effet, $\varphi^{-1}(x,y) = (x',y')$ ssi $(x',y') = \varphi(x,y) = (x, f(x,y))$ et donc $x' = x$.

Alors, pour tout $(x,y) \in U' \times V$ tel que $(x,0) \in \varphi(U' \times V)$, on a

$$f(x,y) = 0 \Leftrightarrow \varphi(x,y) = (x,0) \Leftrightarrow (x,y) = \varphi^{-1}(x,0) \Leftrightarrow y = h(x,0).$$

On pose donc $g(x) = h(x,0)$ et $U := \{x \in U' / \ (x,0) \in \varphi(U' \times V)\}$ de sorte à avoir l'équivalence (20). On a bien que $x_0 \in U$ et $g(x_0) = y_0$ car $x_0 \in U$ et $\varphi(x_0,y_0) = (x_0,0)$. U est un ouvert car $U = U' \cap S^{-1}(\varphi(U' \times V))$ où $S : E \to E \times F$ envoie x sur $(x,0)$. Enfin g est de classe $\mathcal{C}^k$ car c'est la composée de $S : U \to U \times F$, $\varphi^{-1} : \varphi(U' \times V) \to E \times F$ et la projection $E \times F \to F$. $\square$

5 Fonction exponentielle

Comme au chapitre 4.1 sur la fonction inverse, on note $\mathcal{L}(E) = \mathcal{L}(E,E)$ et l'on rappelle que pour tout $A, B \in \mathcal{L}(E)$, $\|AB\| \leqslant \|A\|\|B\|$. Par conséquent, $\|A^n\| \leqslant \|A\|^n$ pour tout n et l'on en déduit que la série $\sum A^n/n!$ converge normalement. Sa somme est par définition l'exponentielle de A

$$\exp(A) := \sum_{n=0}^{\infty} \frac{1}{n!} A^n.$$

Les propriétés suivantes se démontrent sans peine :

1. $\|\exp(A)\| \leqslant \exp(\|A\|)$
2. $\exp(0) = \mathrm{id}_E$
3. si $AB = BA$, alors $\exp(A+B) = \exp(A)\exp(B)$.
4. $\exp(A)$ est inversible d'inverse $\exp(-A)$.
5. si $P \in \mathcal{L}(E)$ est inversible, alors $P\exp(A)P^{-1} = \exp(PAP^{-1})$.

Nous allons établir que l'application exponentielle est de classe $\mathcal{C}^\infty$. Pour se faire, nous utiliserons le résultat suivant.

Théorème 5.0.2. *Soit U un ouvert de E et (f_j) une suite de $\mathcal{C}^1(U,F)$. On suppose que (f_j) admet une limite ponctuelle f et que (f_j') converge localement uniformément vers g. Alors f est de classe $\mathcal{C}^1$ et $f' = g$.*

(f_j') converge localement uniformément vers g signifie que pour tout $x \in U$, il existe $\epsilon > 0$ tel que $B(x,\epsilon) \subset U$ et $\sup_{B(x,\epsilon)} \|f_j' - g\|$ converge vers 0. Ceci implique que g est continue.

Démonstration. Pour tout $x \in U$ et $u \in E$ tels que $[x, x+u] \subset U$, l'on a par la conséquence (7) des accroissements finis que

$$\|f_j(x+u) - f_j(x) - f'_j(x)(u)\| \leqslant \|u\| \sup_{t\in[0,1]} \|f'_j(x+tu) - f'_j(x)\|$$

Si maintenant la convergence de f'_j vers g est uniforme sur $[x, x+u]$, il vient

$$\|f(x+u) - f(x) - g(x)(u)\| \leqslant \|u\| \sup_{t\in[0,1]} \|g(x+tu) - g(x)\|$$

et l'on conclut en utilisant que g est continue. □

Proposition 5.0.3. *L'application* exp *est de classe* $\mathcal{C}^\infty$*.*

Démonstration. La fonction $h_n(A) = A^n/n!$ est de classe $\mathcal{C}^\infty$. De plus

$$h'_n(A)(B) = \frac{1}{n!}\big(BA^{n-1} + ABA^{n-2} + A^2BA^{n-3} + \ldots + A^{n-1}B\big)$$

et par conséquent $\|h'_n(A)\| \leqslant \|A\|^{n-1}/(n-1)!$. On en déduit que la série des dérivées $\sum h'_n$ converge normalement uniformément sur toute partie bornée de $\mathcal{L}(E)$. Par le théorème 5.0.2, exp est alors de classe $\mathcal{C}^1$ et admet pour dérivée $\sum h'_n$.

On montre de même que l'exponentielle est de classe $\mathcal{C}^k$. En effet, on a que

$$\|h_n^{(k)}(A)\| \leqslant \|A\|^{n-k}/(n-k)! \qquad \text{si } n \geqslant k,$$

et $h_n^{(k)} = 0$ si $n < k$. Pour s'en convaincre, remarquer que $h_n^{(k)}(A)(B_1, \ldots, B_k)$ est la somme de $n!/(n-k)!$ monômes en les variables A, B_1, …, B_k. Et tous ces monômes sont majorés en norme par $\|A\|^{n-k}\|B_1\| \ldots \|B_k\|/n!$. □

Notons $\exp'$ la fonction dérivée de exp. D'après la première partie de la preuve, $\exp'(A)(B) = \sum_n h'_n(A)(B)$. D'où

$$AB = BA \qquad \Rightarrow \qquad \exp'(A)(B) = B\exp(A) = \exp(A)B.$$

En particulier, la dérivée de l'application $t \to \exp(tA)$ est $A\exp(tA)$.