Calcul différentiel

Thomas Letendre

Prépa Agreg – Automne 2023

Introduction

Cadre du cours. On va étudier le calcul différentiel pour des applications entre ouverts d'espaces de Banach. C'est un peu plus compliqué que le cadre du programme de l'Agrégation qui se restreint aux ouverts de $\mathbb{R}^n$. Cependant, c'est le cadre naturel pour cette théorie et cela permet de mettre en valeur le rôle fondamental de la complétude. La dimension finie n'est par contre pas centrale. Le fait de considérer des espaces de dimension infinie permet aussi de limiter la tentation d'avoir systématiquement recours à des écritures en coordonnées, ce qui est souvent moins clair que de raisonner sur les objets intrinsèques. On ne considérera que des espaces vectoriels réels.

Prérequis. Dans ces notes on suppose connues les bases de la topologie des espaces métriques (continuité, compacité, complétude, connexité, connexité par arcs,...), sur lesquelles on reviendra dans le cours de Topologie. On suppose aussi une certaine familiarité avec les applications linéaires continues entre espaces de Banach, y compris de dimensions infinies.

Philosophie du calcul différentiel. Les applications, même de $\mathbb{R}$ dans $\mathbb{R}$, sont des objets compliqués dont il est difficile d'avoir une bonne compréhension locale. À l'inverse, les applications linéaires sont bien comprises. L'analyse c'est difficile mais l'algèbre linéaire c'est simple.
Soient E et F deux Banachs et $f : E \to F$. Si f est différentiable en $a \in E$, sa différentielle est l'application linéaire *continue* $L : E \to F$ qui ressemble le plus à $x \mapsto f(a + x) - f(a)$, localement pour x petit. En d'autres termes $f(a + \cdot) \simeq f(a) + L$ et ceci est la meilleure approximation de f par une application affine (constante + linéaire) continue au voisinage de a. Le but du calcul différentiel est d'obtenir de l'information sur la fonction f à partir de la connaissance de sa différentielle. Cette information peut être quantitative, inégalité des accroissements finis par exemple, ou quantitative, théorème d'inversion locale par exemple.

Programme. Fonctions différentiables, inégalité des accroissements finis, différentielles d'ordres supérieurs, formules de Taylor, inversion locale, recherche d'extrema et étude des points critiques.

Références

[Gou08] X. Gourdon, *Les maths en tête : Analyse*, Ellipses, 2008.

[Hau07] B. Hauchecorne, *Les contre-exemples en mathématiques*, Ellipses, 2007.

[Lau11] F. Laudenbach, *Calcul différentiel et intégral*, Les éditons de l'École Polytechnique, 2011.

[Rou09] F. Rouvière, *Petit guide du calcul différentiel*, Cassini, 2009.

[SR08] J. St-Raymond, *Topologie calcul différentiel et variable complexe*, Calvage et Mounet, 2008.

Les deux références principales pour ce cours sont [Lau11, chap. II] et [SR08, chap. IX à XIV]. Il emprunte aussi ponctuellement à [Gou08] et [Hau07]. Enfin [Rou09] est une référence souvent appréciée des agrégatifs.

Table des matières

1 Rappels

1.1 Relations de comparaison

Soient X un espace métrique et $a \in X$. Soient $f : X \to F$ et $g : X \to G$, où F et G sont des espaces vectoriels normés.

Définition 1.1 (Grand O)**.** On dit que f est *dominée* par g (ou est grand O de g) sur X on note $f(x) = O\big(g(x)\big)$ s'il existe $C \geqslant 0$ tel que $\forall x \in X$, $\|f(x)\|_F \leqslant C\|g(x)\|_G$.
On dit que f est *dominée* par g (ou grand O de g) au voisinage de a et on note $f(x) \underset{x\to a}{=} O\big(g(x)\big)$ s'il existe un voisinage V de a et $C \geqslant 0$ tels que $\forall x \in V$, $\|f(x)\|_F \leqslant C\|g(x)\|_G$.

Définition 1.2 (Petit o)**.** On dit que f est *négligeable* devant g (ou petit o de g) quand x tend vers a et on note $f(x) \underset{x\to a}{=} o\big(g(x)\big)$ s'il existe $\varepsilon : X \to F$ telle que $f = \varepsilon\|g\|_G$ et $\varepsilon(x) \xrightarrow[x\to a]{} 0$.

Définition 1.3 (Équivalent)**.** Si $F = G$, on dit que f est *équivalent* à g lorsque x tend vers a et on note $f(x) \underset{x\to a}{\sim} g(x)$ si $f(x) - g(x) \underset{x\to a}{=} o(g(x))$.

On rappelle que $\underset{x\to a}{\sim}$ définit une relation d'équivalence sur les fonctions définies au voisinage de a. On laisse aux lecteurices le soin de réviser comment elle se comporte vis-à-vis des opérations algébriques, notamment son comportement pathologique vis-à-vis de l'addition.

Remarque 1.4.
- Ces définitions s'étendent immédiatement au cas où f et g sont définis sur $Y \subset X$ et $a \in \overline{Y}$. De même elles s'étendent au cas $a = \pm\infty$ avec $X \subset \mathbb{R}$ non borné.
- Définir les notions de o et O sans utiliser de quotient permet de leur donner un sens même lorsque g a des zéros dans tout voisinage de a.
- Les notions de o, O et $\sim$ dépendent a priori de $\|\cdot\|_F$ et $\|\cdot\|_G$. Néanmoins, elles sont conservées si on remplace ces normes par des normes équivalentes. En particulier, en dimension finie elle ne dépendent que de F et G.
- Les notations précédentes sont très mauvaises et font des choses horribles à l'égalité qui n'est plus ni transitive ni symétrique. Quand $x \to 0$ dans $\mathbb{R}$, on a $x^2 = o(x)$ et $x^3 = o(x)$. Mais on a plus de mal à écrire que $o(x) = x^2$, et personne ne déduirait de ce qui précède que $x^2 = x^3$.

1.2 Applications linéaires continues

Une référence pour cette section est [SR08, chap. VII.4]. Dans la suite on notera toutes les normes par $\|\cdot\|$ sauf s'il y a un risque d'ambiguïté.

Lemme 1.5. *Soient E et F deux espaces vectoriels normés et $L : E \to F$ une application linéaire, les énoncés suivants sont équivalents :*

1. *L est continue sur E,*
2. *L est continue en 0,*
3. *il existe $C \geqslant 0$ tel que $\forall x \in E$, $\|L(x)\| \leqslant C\|x\|$.*

Démonstration. En exercice. □

Lemme 1.6. *Si E est de dimension finie alors toute application linéaire $L : E \to F$ est continue.*

Démonstration. On définit une norme auxiliaire sur E par $x \mapsto \|x\|_E + \|L(x)\|_F$. Comme E est de dimension finie, cette norme auxiliaire est équivalente à $\|\cdot\|_E$. Il existe donc $C \geqslant 0$ telle que pour tout $x \in E$ $\|x\|_E + \|L(x)\|_F \leqslant C\|x\|_E$. En particulier, $\forall x \in E$, $\|L(x)\|_F \leqslant C\|x\|_E$. □

Définition 1.7. On note $\mathcal{L}(E,F)$ (resp. $\mathcal{L}(E)$) l'espace des applications linéaires *continues* de E dans F (resp. E). On définit sur $\mathcal{L}(E,F)$ la *norme d'opérateur* subordonnée à $\|\cdot\|_E$ et $\|\cdot\|_F$ par :

$$\forall L \in \mathcal{L}(E,F), \qquad \|L\|_{\mathcal{L}(E,F)} = \sup\left\{ \frac{\|L(x)\|_F}{\|x\|_E} \;\middle|\; x \neq 0 \right\} = \sup\{\|L(x)\|_F \mid \|x\|_E = 1\}.$$

C'est la constante optimale dans le point 3 du lemme 1.5, et le sup est un max si $\dim(E) < +\infty$.

Lemme 1.8. *Soient $L \in \mathcal{L}(E,F)$ et $K \in \mathcal{L}(F,G)$ alors $K \circ L \in \mathcal{L}(E,G)$ et $\|K \circ L\| \leqslant \|K\|\|L\|$.*

Démonstration. En exercice. □

Proposition 1.9. *Si F est de Banach alors $\mathcal{L}(E,F)$ est de Banach pour la norme de la définition 1.7. En particulier, $E' = \mathcal{L}(E,\mathbb{R})$ est de Banach pour la norme subordonnée à $\|\cdot\|_E$ et $|\cdot|$.*

Démonstration. En exercice. □

Notons $\mathrm{Iso}(E,F) = \big\{L \in \mathcal{L}(E,F) \bigm| L \text{ bijective et } L^{-1} \in \mathcal{L}(F,E)\big\}$ et $GL(E) = \mathrm{Iso}(E,E)$. Une raison pour laquelle l'algèbre linéaire (et donc le calcul différentiel) est plus agréable entre espaces de Banach est le résultat suivant, qui utilise la complétude sous la forme du théorème de Baire.

Théorème 1.10 (Isomorphisme de Banach)**.** *Soient E et F deux espaces de Banach. Si $L \in \mathcal{L}(E,F)$ est bijective alors $L^{-1} \in \mathcal{L}(F,E)$. En particulier* $\mathrm{Iso}(E,F) = \{L \in \mathcal{L}(E,F) \mid L \textit{ bijective}\}$.

Démonstration. Voir [SR08, App. B]. □

Lemme 1.11 (Lemme de Neumann)**.** *Soient E un espace de Banach et $L \in \mathcal{L}(E)$. Si $\|L\| < 1$ alors $\mathrm{Id} - L \in GL(E)$ et $(\mathrm{Id} - L)^{-1} = \sum_{k \geqslant 0} L^k$.*

Démonstration. Pour tout $k \geqslant 0$, $\|L^k\| \leqslant \|L\|^k$. Comme $\|L\| < 1$, la série $\sum_{k\geqslant 0} L^k$ est normalement convergente, donc convergente dans le Banach $\mathcal{L}(E)$. Le résultat est alors conséquence des égalités $(\mathrm{Id} - L)\big(\sum_{k\geqslant 0} L^k\big) = \mathrm{Id} = \big(\sum_{k\geqslant 0} L^k\big)(\mathrm{Id} - L)$ et du théorème d'isomorphisme de Banach. □

Corollaire 1.12. *Si E est de Banach, alors le groupe $GL(E)$ est ouvert dans le Banach $\mathcal{L}(E)$.*

Démonstration. Soient $L \in GL(E)$ et $H \in \mathcal{L}(E)$, on écrit $L + H = L(\mathrm{Id} - K)$ où $K = -L^{-1}H$. Si $\|H\| < \frac{1}{\|L^{-1}\|}$ alors $\|K\| < 1$, donc $\mathrm{Id} - K \in GL(E)$ et $L + H \in GL(E)$. Donc $GL(E)$ contient la boule ouverte de centre L et de rayon $\frac{1}{\|L^{-1}\|}$. □

1.3 Applications multilinéaires continues

Soient $E_1, \dots, E_k$ et F des espaces vectoriels normés. Dans la suite on muni $E_1 \times \cdots \times E_k$, de la norme définie par $\|x\|_\infty = \max_{1\leqslant i\leqslant k} \|x_i\|_{E_i}$ pour tout $x = (x_1, \dots, x_k) \in E_1 \times \cdots \times E_k$.

Lemme 1.13. *Si $M : E_1 \times \cdots \times E_k \to F$ est k-linéaire, les énoncés suivants sont équivalents :*

1. *M est continue sur $E_1 \times \cdots \times E_k$,*
2. *M est continue en 0,*
3. *il existe $C \geqslant 0$ tel que $\forall (x_1, \dots, x_k) \in E_1 \times \cdots \times E_k$, $\|M(x_1, \dots, x_k)\|_F \leqslant C\|x_1\|_{E_1} \dots \|x_k\|_{E_k}$.*

Démonstration. En exercice, voir [SR08, chap. VII.6] dans le cas $k = 2$. □

Exemple 1.14. L'application $(K,L) \mapsto K \circ L$ de $\mathcal{L}(F,G) \times \mathcal{L}(E,F)$ vers $\mathcal{L}(E,G)$ est bilinéaire continue d'après le lemme 1.8.

Définition 1.15. On note $\mathcal{L}_k(E_1, \ldots, E_k, F)$ l'espace vectoriel des applications k-linéaires *continues* de $E_1 \times \cdots \times E_k$ dans F. On définit sur cet espace la *norme subordonnée* aux $\|\cdot\|_{E_i}$ et à $\|\cdot\|_F$ par :

$$\forall M \in \mathcal{L}_k(E_1, \ldots, E_k, F), \qquad \|M\| = \sup\left\{\|M(x_1, \ldots, x_k)\|_F \;\middle|\; \forall i \in [\![1, k]\!], \|x_i\|_{E_i} = 1\right\}.$$

On définit une application linéaire naturelle $\Phi : \mathcal{L}_k(E_1, \ldots, E_k, F) \to \mathcal{L}(E_k, \mathcal{L}_{k-1}(E_1, \ldots, E_{k-1}, F))$ en posant $\Phi(M) : x_k \mapsto M(\cdot, \ldots, \cdot, x_k)$, pour tout $M \in \mathcal{L}_k(E_1, \ldots, E_k, F)$. Les lecteurices vérifieront les détails, notamment que $\Phi(M) : E_k \to \mathcal{L}_{k-1}(E_1, \ldots, E_{k-1}, F)$ est bien linéaire et continue.

Lemme 1.16. *L'application linéaire Φ est une bijection isométrique, en particulier linéaire continue d'inverse linéaire et continu.*

Démonstration. Soit $\Psi : \mathcal{L}(E_k, \mathcal{L}_{k-1}(E_1, \ldots, E_{k-1}, F)) \to \mathcal{L}_k(E_1, \ldots, E_k, F)$ l'application linéaire définie par $\Psi(\mu) : (x_1, \ldots, x_k) \mapsto (\mu(x_k))(x_1, \ldots, x_{k-1})$. On vérifie à titre d'exercice que Φ et Ψ sont linéaires isométriques et inverse l'une de l'autre. Voir [SR08, chap. IX.1] dans le cas $k = 2$. $\square$

On obtient ainsi un isomorphisme isométrique $\mathcal{L}_k(E_1, \ldots, E_k, F) \simeq \mathcal{L}(E_k, \mathcal{L}_{k-1}(E_1, \ldots, E_{k-1}, F))$, d'où on déduit un isomorphisme isométrique $\mathcal{L}_k(E_1, \ldots, E_k, F) \simeq \mathcal{L}(E_k, \mathcal{L}(E_{k-1}, \ldots \mathcal{L}(E_1, F) \ldots))$ par récurrence. Ces identifications canoniques seront utiles lorsqu'on parlera de différentielle k-ième.

2 Applications différentiables

Dans toute la suite tous les espaces vectoriels considérés seront des espaces de Banach réels. On notera E et F deux tels espaces et $U \subset E$ un ouvert. On désigne par $\mathcal{L}(E, F)$ l'espace des applications linéaires continues de E dans F muni de sa norme d'opérateur.

2.1 Différentiabilité

Définition 2.1. Soit $f : U \to F$ et $x \in U$, on dit que f est *différentiable* en x s'il existe $L : E \to F$ linéaire *continue* telle que $f(x + h) \underset{h \to 0}{=} f(x) + L(h) + o(h)$, i.e. $\frac{\|f(x+h) - f(x) - L(h)\|}{\|h\|} \xrightarrow[h \to 0]{} 0$.

Remarque 2.2. La notion de fonction différentiable en x dépend des normes à équivalence près. En dimension finie cette notion est uniquement définie.

Lemme 2.3. *Si $f : U \to F$ est différentiable en $x \in U$ alors elle est continue en x.*

Démonstration. On a $f(x + h) = f(x) + L(h) + o(h) \xrightarrow[h \to 0]{} f(x)$, car L est supposée *continue* ! $\square$

La réciproque est bien sûr fausse, le contre-exemple classique étant $x \mapsto |x|$ de $\mathbb{R}$ dans $\mathbb{R}$ en 0.

Lemme 2.4. *Si f est différentiable en x, l'application $L \in \mathcal{L}(E, F)$ de la définition 2.1 est unique.*

Démonstration. Soient L et $L' \in \mathcal{L}(E, F)$ telles que $f(x+h) = f(x)+L(h)+o(h) = f(x)+L'(h)+o(h)$ au voisinage de 0. Alors $(L - L')(h) = o(h)$, d'où $t(L(v) - L'(v)) = (L - L')(tv) = o(tv) = o(t)$ lorsque $t \to 0$ à fixé. Donc $L(v) = L'(v)$ pour tout $v \in E$. $\square$

Définition 2.5.
- Si f est différentiable en x, l'application linéaire L de la définition 2.1 est appelée la *différentielle* en x de f et est notée $D_x f \in \mathcal{L}(E, F)$.
- On dit que f est *différentiable* sur U si elle est différentiable en tout point de U. Dans ce cas, l'application $Df : U \to \mathcal{L}(E, F)$ définie par $Df : x \mapsto D_x f$ est appelée la *différentielle* de f.
- Si de plus Df est continue, on dit que f est de classe $\mathcal{C}^1$ sur U. On note alors $f \in \mathcal{C}^1(U, F)$.

Exemple 2.6. • Si $L \in \mathcal{L}(E,F)$ alors $L(x+h) = L(x) + L(h)$ pour tout $x, h \in E$. Donc L est différentiable sur E et $DL : x \mapsto L$ est constante donc $L \in \mathcal{C}^1(E,F)$. Cela s'applique en particulier aux fonctions coordonnées dans $\mathbb{R}^n$.

- Soit $B \in L_2(E_1, E_2, F)$, pour tout $x = (x_1, x_2)$ et $h = (h_1, h_2) \in E_1 \times E_2$ on a

$$B(x+h) = B(x_1 + h_1, x_2 + h_2) = \underbrace{B(x_1, x_2)}_{=B(x)} + \underbrace{B(h_1, x_2) + B(x_1, h_2)}_{\text{linéaire continue en } h} + \underbrace{B(h_1, h_2)}_{=O(\|h_1\|\|h_2\|)=O(\|h\|_\infty^2)}.$$

Donc $B \in \mathcal{C}^1(E_1 \times E_2, F)$ avec $D_x B : h \mapsto B(h_1, x_2) + B(x_1, h_2)$ pour tout $x \in E_1 \times E_2$.

- Plus généralement, si $E_1, \ldots, E_k$ et F sont de Banach et $M \in \mathcal{L}_k(E_1, \ldots, E_k, F)$, alors M est $\mathcal{C}^1$ et pour tout $x = (x_1, \ldots, x_k) \in E_1 \times \cdots \times E_k$ on a

$$D_x M : (h_1, \ldots, h_k) \mapsto M(h_1, x_2, \ldots, x_k) + M(x_1, h_2, x_3, \ldots, x_k) + \cdots + M(x_1, \ldots, x_{k-1}, h_k).$$

- Soit E un Banach et $\mathrm{Inv} : L \mapsto L^{-1}$ de $GL(E)$ dans lui-même. D'après le lemme 1.11 on a :

$$(\mathrm{Id} + H)^{-1} = \mathrm{Id} - H + \sum_{k \geqslant 2} (-H)^k = \mathrm{Id} - H + H^2 \sum_{k \geqslant 0} (-H)^k.$$

Si $\|H\| \leqslant \frac{1}{2}$ alors $\left\|\sum_{k\geqslant 0}(-H)^k\right\| \leqslant \sum_{k \geqslant 0} \|H\|^k \leqslant 2$, et le dernier terme ci-dessus est $O(H^2)$ au voisinage de 0. Donc Inv est différentiable en Id avec $D_{\mathrm{Id}} \mathrm{Inv} : H \mapsto -H$. Soit maintenant $L \in GL(E)$, on a vu que pour H assez petit $L + H \in GL(E)$. Lorsque $H \to 0$ on a

$$\begin{aligned}(L+H)^{-1} &= (\mathrm{Id} + L^{-1}H)^{-1} L^{-1} = \big(\mathrm{Id} - L^{-1}H + o\big(L^{-1}H\big)\big) L^{-1} \\ &= L^{-1} - L^{-1} H L^{-1} + \underbrace{o\big(L^{-1} H L^{-1}\big)}_{=o(H)}.\end{aligned}$$

Comme $H \mapsto -L^{-1} H L^{-1}$ est continue de $\mathcal{L}(E)$ dans $\mathcal{L}(E)$, on en déduit que Inv est différentiable et $D_L \mathrm{Inv} : H \mapsto -L^{-1} H L^{-1}$ pour tout L. On peut ensuite vérifier que Inv est $\mathcal{C}^1$.

Lemme 2.7. *Si $E = \mathbb{R}$ alors $f : U \to F$ est différentiable en $x \in U$ si et seulement si elle est dérivable en x. Si c'est le cas, on a $D_x f : h \mapsto f'(x)h$.*

Démonstration. Si f est dérivable en x alors $\frac{f(x+h)-f(x)}{h} \xrightarrow[h \to 0]{} f'(x) \in \mathbb{R}$, ce qui est équivalent à $f(x+h) = f(x) + f'(x)h + o(h)$. Donc f est différentiable en x et $D_x f : h \mapsto f'(x)h$. Inversement si f est différentiable en x, on a $f(x+h) = f(x) + D_x f \cdot h + o(h)$. Par linéarité de $D_x f : \mathbb{R} \to F$ on en déduit $f(x+h) = f(x) + h D_x f \cdot 1 + o(h)$. Donc f est dérivable en x de dérivée $f'(x) = D_x f \cdot 1$. □

2.2 Opérations sur les applications différentiables

Dans cette section on s'intéresse aux opérations qui préservent la classe des fonctions différentiables. C'est essentiel pour prouver qu'une fonction est différentiable si elle est construite à partir de fonctions plus simples que l'on sait déjà différentiables.

Lemme 2.8. *L'ensemble* $\mathrm{Diff}_x(U,F)$ *(resp.* $\mathrm{Diff}(U,F)$*, resp.* $\mathcal{C}^1(U,F)$*) des fonctions différentiables en $x \in U$ (resp. différentiables sur U, resp. de classe $\mathcal{C}^1$ sur U) est un sous-espace vectoriel de l'espace $\mathcal{F}(U,F)$ des fonctions de U dans F. De plus, les applications $D_x : \mathrm{Diff}_x(U,F) \to \mathcal{L}(E,F)$, $D : \mathrm{Diff}(U,F) \to \mathcal{F}(U, \mathcal{L}(E,F))$ et $D : \mathcal{C}^1(U,F) \to \mathcal{C}^0(U, \mathcal{L}(E,F))$ sont linéaires.*

Démonstration. En exercice. □

Lemme 2.9. *Soient E et $F_1, \dots, F_k$ des Banachs et $F = F_1 \times \dots \times F_k$ muni de $\|\cdot\|_\infty$. Soit $U \subset E$ ouvert et $f = (f_1, \dots, f_k) : U \to F$. Alors f est différentiable en $x \in U$ si et seulement si pour tout $i \in [\![1, k]\!]$, $f_i : U \to F_i$ est différentiable en x. Dans ce cas $D_x f : h \mapsto (D_x f_1 \cdot h, \dots, D_x f_k \cdot h)$.*

Démonstration. En exercice. □

Le résultat essentiel de cette section est la règle de la chaine ci-dessous. Le résultat est naturel si on écrit les développements limités naïvement, mais la preuve rigoureuse nécessite d'être très soigneux dans la gestion des termes d'erreur.

Proposition 2.10 (Règle de la chaine)**.** *Soient E, F et G trois espaces de Banach, soient $f : U \to V$ et $g : V \to G$ où $U \subset E$ et $V \subset F$ sont ouverts. Si f est différentiable en $x \in U$ et g est différentiable en $f(x) \in V$ alors $g \circ f$ est différentiable en x et $D_x(g \circ f) = D_{f(x)} g \circ D_x f$.*

Démonstration. La différentiabilité de f en x dit qu'il existe une fonction ε définie au voisinage de $0 \in E$ et à valeurs dans F telle que $\varepsilon(h) \xrightarrow[h \to 0]{} 0$ et

$$f(x+h) = f(x) + D_x f \cdot h + \|h\| \varepsilon(h). \tag{2.1}$$

De même, comme g est différentiable en $f(x)$, il existe une fonction η définie au voisinage de $0 \in F$ à valeurs dans G telle que $\eta(k) \xrightarrow[k \to 0]{} 0$ et

$$g(f(x)+k) = g(f(x)) + D_{f(x)} g \cdot k + \|k\| \eta(k). \tag{2.2}$$

D'après (2.1),

$$g \circ f(x+h) = g(f(x+h)) = g\left(f(x) + \underbrace{D_x f \cdot h + \|h\| \varepsilon(h)}_{=k(h)} \right).$$

On a $\|k(h)\| \leqslant \|h\| (\|D_x f\| + |\varepsilon(h)|) = O(h)$ au voisinage de 0. En particulier $k(h) \xrightarrow[h \to 0]{} 0$. On peut donc utiliser (2.2) dans l'égalité précédente, ce qui donne :

$$\begin{aligned} g \circ f(x+h) &= g(f(x)) + D_{f(x)} g \cdot k(h) + \|k(h)\| \eta(k(h)) \\ &= g \circ f(x) + \big(D_{f(x)} g \circ D_x f \big) \cdot h + \|h\| \underbrace{D_{f(x)} g \cdot \varepsilon(h)}_{\xrightarrow[h \to 0]{} 0} + O(h) \underbrace{\eta(\varepsilon(h))}_{\xrightarrow[h \to 0]{} 0} \\ &= g \circ f(x) + \big(D_{f(x)} g \circ D_x f \big) \cdot h + o(h). \end{aligned}$$

Comme $D_x f \in \mathcal{L}(E, F)$ et $D_{f(x)} g \in \mathcal{L}(F, G)$ on a bien $D_{f(x)} g \circ D_x f \in \mathcal{L}(E, G)$, d'où le résultat. □

Corollaire 2.11. *Si $f : U \to V$ et $g : V \to G$ sont $\mathcal{C}^1$ alors $g \circ f$ aussi.*

Démonstration. La prop. 2.10 donne $g \circ f$ différentiable sur U avec $D(g \circ f) : x \mapsto D_{f(x)} g \circ D_x f$.

$$\begin{tikzcd}
 & \mathcal{L}(E,F) & \\
U \arrow[ur, bend left, "Df"] \arrow[r, "(D_{f(\cdot)}g, D_f)"] \arrow[rr, bend right, "D(g\circ f)"'] \arrow[dr, "D_{f(\cdot)}g"'] \arrow[d, "f"'] & \mathcal{L}(F,G) \times \mathcal{L}(E,F) \arrow[r, "\circ"] & \mathcal{L}(E,F) \\
V \arrow[r, "Dg"'] & \mathcal{L}(F,G) &
\end{tikzcd}$$

Les applications correspondant à des flèches noires sont continues, par hypothèse ou par l'exemple 1.14. En composant des applications continues, les flèches bleues sont continues, donc la flèche rouge aussi. □

Corollaire 2.12. *Soient $I \subset \mathbb{R}$ intervalle ouvert de $U \subset E$ ouvert. Soient $\gamma : I \to U$ dérivable en $t \in I$ et $f : U \to F$ différentiable en $\gamma(t) \in U$, alors $f \circ \gamma$ est dérivable en t et $(f \circ \gamma)'(t) = D_{\gamma(t)} f \cdot \gamma'(t)$.*

La règle de la chaine permet de redémontrer la règle de Leibniz. Attention cependant, cette règle se prouve aussi plus élémentairement par un développement limité pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Corollaire 2.13 (Règle de Leibniz)**.** *Soient $I \subset \mathbb{R}$ un intervalle ouvert. Soient $f : I \to \mathbb{R}$ et $g : I \to \mathbb{R}$ dérivables en $t \in I$, alors fg est dérivable en t et $(fg)'(t) = f'(t)g(t) + f(t)g'(t)$.*

Démonstration. On applique la règle de la chaine à la composée de $(f, g) : I \to \mathbb{R}^2$ et de l'application bilinéaire $\times : \mathbb{R}^2 \to \mathbb{R}$. □

Exercice 2.14. Par la même stratégie, montrer le caractère $\mathcal{C}^1$ et calculer la différentielle de

- αf, où $\alpha : U \to \mathbb{R}$ et $f : U \to F$ sont $\mathcal{C}^1$;
- $\langle f , g \rangle$, où f et $g : U \to \mathbb{R}^n$ sont $\mathcal{C}^1$ et $\langle \cdot , \cdot \rangle$ est le produit scalaire euclidien ;
- $f_1 \times \cdots \times f_k$, où les $f_i : U \to \mathbb{R}$ sont $\mathcal{C}^1$.

Le dernier point appliqué aux applications coordonnées dans $\mathbb{R}^n$ montre que les monômes en n variables sont $\mathcal{C}^1$ sur $\mathbb{R}^n$. Il en est donc de même des éléments de $\mathbb{R}[X_1, \ldots, X_n]$.

Exercice 2.15. Montrer que $\det : M_n(\mathbb{R}) \to \mathbb{R}$ est de classe $\mathcal{C}^1$ et calculer sa différentielle.
Indication. Calculer d'abord la différentielle en Id, puis en $A \in GL_n(\mathbb{R})$ et enfin étendre à $M_n(\mathbb{R})$.

2.3 Dérivées directionnelles et dérivées partielles

Définition 2.16. Soient $f : U \to F$, $x \in U$ et $v \in E$. Si $t \mapsto f(x + tv)$ est dérivable en 0, sa dérivée en 0 est appelée la *dérivée directionnelle* de f en x dans la direction v. On la notera $\partial_v f(x)$ (attention notation non standard).

Définition 2.17. Si $E = \mathbb{R}^n$, la i-ième dérivée partielle de $f : U \to F$ en $x \in U$ est la dérivée en $t = 0$ de $t \mapsto f(x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n)$, si elle existe. On la note $\partial_i f(x)$. Si $(e_1, \ldots, e_n)$ est la base canonique de $\mathbb{R}^n$ on a donc $\partial_i f(x) = \partial_{e_i} f(x)$.

Lemme 2.18. *Si $f : U \to F$ est différentiable en $x \in U$ alors elle admet des dérivées directionnelles en x dans toutes les directions, et $\partial_v f(x) = D_x f \cdot v$ pour tout $v \in E$.*
En particulier, si $E = \mathbb{R}^n$ elle admet des dérivées partielles en x et $D_x f : h \mapsto \sum_{i=1}^n h_i \partial_i f(x)$.

Démonstration. On applique le cor. 2.12 avec $\gamma : t \mapsto x + tv$, dérivable en 0 telle que $\gamma'(0) = v$. □

À l'inverse, l'existence de dérivées directionnelle ne suffit pas à garantir le différentiabilité de f. Par exemple, dans $\mathbb{R}^2$ l'indicatrice de l'ensemble $\left\{(x_1, x_2) \in \mathbb{R}^2 \,\middle|\, 0 < x_2 < x_1\right\}$ admet une dérivée directionnelle nulle dans toutes les directions en $(0, 0)$, mais elle n'y est même pas continue !
Les dérivées partielles et directionnelles sont cependant très utiles pour les calculs. De plus, même si elles ne suffisent pas à prouver la différentiabilité, elles donnent l'unique candidat pour être $D_x f$.

Exemple 2.19. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x_1, x_2) = 0$ si $x_2 \neq 0$ et $f(x_1, x_2) = x_1$ sinon. Cette fonction admet des dérivées directionnelles dans toutes les directions en 0. Soit $v = (v_1, v_2)$, si $v_2 \neq 0$ alors $\partial_v f(0) = 0 = f(v)$ et si $v_2 = 0$ alors $\partial_v f(0) = v_1 = f(v)$. Si f était différentiable en 0 sa différentielle serait donc f, mais cette application n'est pas linéaire.

Définition 2.20. Soit $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m$ différentiable en x, alors la matrice de $D_x f$ dans la base canonique de $\mathbb{R}^n$ et $\mathbb{R}^m$ est la matrice *jacobienne* $\left(\partial_j f_i(x)\right)_{1 \leqslant i \leqslant m, 1 \leqslant j \leqslant n}$ de f en x.

Pour des applications explicites de $\mathbb{R}^n$ dans $\mathbb{R}^m$, le calcul de la différentielle revient souvent au calcul de la jacobienne. La composition dans la règle de la chaine se ramène alors à un produit matriciel.

3 Accroissements finis

Dans cette section, on s'intéresse au théorème des accroissements et à ses conséquences, qui sont nombreuses et profondes. Notons tout de suite que le résultat important est l'inégalité des accroissements finis. L'égalité des accroissements finis est plus anecdotique, et n'est valable que pour les fonctions à valeurs dans $\mathbb{R}$.
Encore une fois, dans toute cette section E et F sont des espaces de Banach et $U \subset E$ est ouvert. On notera aussi souvent a et b deux réels avec $a \leqslant b$.

3.1 Inégalité des accroissements finis

On commence avec une version forte qui sera utile pour prouver la formule de Taylor–Lagrange, voir la section 4.3.

Théorème 3.1 (Inégalité des accroissements finis forte)**.** *Soient $f : [a,b] \to F$ et $g : [a,b] \to \mathbb{R}$ deux fonctions continues sur $[a,b]$ et dérivables sur $]a,b[$. On suppose que $\forall x \in]a,b[$, $\|f'(x)\| \leqslant g'(x)$. Alors $\|f(b) - f(a)\| \leqslant g(b) - g(a)$.*

Démonstration. Soit $\varepsilon > 0$, on note $\varphi_\varepsilon : x \mapsto \|f(x) - f(a)\| - (g(x) - g(a)) - \varepsilon(x - a)$ qui est continue de $[a,b]$ dans $\mathbb{R}$. Soit $I_\varepsilon = \varphi_\varepsilon^{-1}(]-\infty,\varepsilon]) = \left\{x \in [a,b] \;\middle|\; \|f(x) - f(a)\| \leqslant (g(x) - g(a)) + \varepsilon(x - a) + \varepsilon\right\}$, on va montrer que $b \in I_\varepsilon$.
On a $a \in I_\varepsilon \subset [a,b]$, donc I_ε est non vide et majoré, donc il admet une borne sup que l'on note $x \in [a,b]$. L'ensemble I_ε est fermé dans $[a,b]$ car $]-\infty,\varepsilon]$ est fermé et φ_ε est continue. Donc $x \in I_\varepsilon$.
Comme $\varphi_\varepsilon(a) = 0$ et φ_ε est continue, pour $\delta > 0$ assez petit $\varphi_\varepsilon(a + \delta) \leqslant \varepsilon$. Donc $a < a + \delta \leqslant x$.
Par l'absurde on suppose que $x < b$, de sorte que $x \in]a,b[$. Pour $h > 0$ assez petit on a alors

$$\begin{aligned}\frac{\|f(x+h) - f(x)\|}{h} - \frac{g(x+h) - g(x)}{h} &= \left\|\frac{f(x+h) - f(x)}{h}\right\| - \frac{g(x+h) - g(x)}{h} \\ &\xrightarrow[h \to 0]{} \|f'(x)\| - g'(x) \leqslant 0.\end{aligned}$$

Il existe donc $\delta > 0$ tel que pour tout $h \in [0,\delta]$, $\frac{\|f(x+h)-f(x)\|}{h} - \frac{g(x+h)-g(x)}{h} \leqslant \varepsilon$. Donc

$$\begin{aligned}\varphi_\varepsilon(x+h) &= \|f(x+h) - f(a)\| - (g(x+h) - g(a)) - \varepsilon(x + h - a) \\ &\leqslant \underbrace{\|f(x) - f(a)\| - (g(x) - g(a)) - \varepsilon(x - a)}_{= \varphi_\varepsilon(x) \leqslant \varepsilon} + \underbrace{\|f(x+h) - f(x)\| - (g(x+h) - g(x)) - \varepsilon h}_{\leqslant 0}.\end{aligned}$$

Donc $[x, x+\delta] \subset I_\varepsilon$, ce qui contredit la définition de x. Donc $x = b$.
Pour tout $\varepsilon > 0$, on a montré que $\|f(b) - f(a)\| \leqslant (g(b) - g(a)) + \varepsilon(b - a) + \varepsilon$. En faisant $\varepsilon \to 0$, on obtient donc $\|f(b) - f(a)\| \leqslant g(b) - g(a)$. □

Remarque 3.2. Le résultat reste vrai en supposant f et g seulement dérivables à droite et $\|f_d'\| \leqslant g_d'$.

Corollaire 3.3 (Inégalité des accroissements finis classique)**.** *Si $f : [a,b] \to F$ est continue sur $[a,b]$, dérivable sur $]a,b[$ et $\|f'(x)\| \leqslant M$ pour tout $x \in]a,b[$, alors $\|f(b) - f(a)\| \leqslant M(b - a)$.*

Démonstration. On applique le théorème 3.1 avec $g : x \mapsto Mx$. □

Si on interprète f comme une position en fonction du temps et f' comme la vitesse associée, le corollaire 3.3 correspond à l'intuition physique suivante. Si on a marché pendant 1h à moins de 5 km/h alors on a parcouru moins de 5km.

Corollaire 3.4. *Soient $U \subset E$ un ouvert convexe et $f : U \to F$ différentiable telle que $\|D_x f\| \leqslant M$ pour tout $x \in U$. Pour tout x et $y \in U$ on a $\|f(y) - f(x)\| \leqslant M\|y - x\|$.*

Démonstration. Soient x et $y \in U$, on note $\gamma : [0,1] \to U$ définie par $\gamma : t \mapsto x + t(y - x)$. La fonction $f \circ \gamma : [0,1] \to F$ est continue sur $[0,1]$, dérivable sur $]0,1[$ et, pour tout $t \in]0,1[$, $\|(f \circ \gamma)'(t)\| = \|D_{\gamma(t)} f \cdot \gamma'(t)\| \leqslant \|D_{\gamma(t)} f\| \|\gamma'(t)\| \leqslant M\|y - x\|$. D'après le corollaire 3.3 on a donc $\|f(y) - f(x)\| = \|f(\gamma(1)) - f(\gamma(0))\| \leqslant M\|y - x\|$. $\square$

Plus généralement, si U n'est pas convexe mais qu'il existe $\gamma : [0,1] \to U$ un chemin joignant x à y comme ci-dessus, la même preuve montre que $\|f(y) - f(x)\| \leqslant \sup_{t \in]0,1[} \|(f \circ \gamma)'(t)\|$. En particulier, si $[x,y] \subset U$ on a toujours $\|f(y) - f(x)\| \leqslant M\|y - x\|$. Faire des dessins.

Corollaire 3.5. *Soient $U \subset E$ un ouvert connexe et $f : U \to F$ différentiable telle que $D_x f = 0$ pour tout $x \in U$, alors f est constante.*

Démonstration. Soient $x_0 \in U$ et $\Omega = \{x \in U \mid f(x) = f(x_0)\}$. L'ensemble Ω est non vide et fermé, par continuité de f. Soit $x \in \Omega$, il existe $R > 0$ tel que $B(x,R) \subset U$. D'après le cor. 3.4 appliqué à l'ouvert convexe $B(x,R)$, la fonction f est constante sur $B(x,R)$, donc $B(x,R) \subset \Omega$. Donc Ω est ouvert. Par connexité de U, on a $\Omega = U$. Donc f est constante. $\square$

3.2 Théorème fondamental de l'analyse

Dans cette section on fait le lien entre l'inégalité des accroissements finis et le théorème fondamental de l'analyse, qui affirme que $f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t$ si f est de classe $\mathcal{C}^1$. Il est important de se rappeler que le théorème fondamental de l'analyse est conséquence de l'inégalité des accroissements finis, et pas l'inverse !

Soit E un espace de Banach, on va vouloir intégrer des fonctions continues d'un segment $[a,b] \subset \mathbb{R}$ à valeurs dans E. Si $E = \mathbb{R}$, on utilise l'intégrale de Lebesgue. Si E est de dimension finie, on peut choisir une base et intégrer composante à composante par la théorie de Lebesgue. En revanche cette méthode ne marche plus en dimension infinie. En dimension infinie, on sait cependant construire l'intégrale de Riemann, le prix à payer étant qu'on ne sait intégrer que des fonctions continues (par morceaux). La construction de l'intégrale de Riemann pour les fonctions à valeurs dans un Banach E est détaillée dans [Gou08, chap. 3]. Elle utilise de façon essentielle la complétude de E.

Lemme 3.6. *Soient $f \in \mathcal{C}^0([a,b], E)$ et $F : x \mapsto \int_a^x f(t)\,\mathrm{d}t$. La fonction F est $\mathcal{C}^1$ sur $[a,b]$ et $F' = f$.*

Démonstration. En exercice, en utilisant la relation de Chasles et l'uniforme continuité de f. $\square$

Corollaire 3.7 (Théorème fondamental de l'analyse)**.** *Si $f : [a,b] \to E$ est de classe $\mathcal{C}^1$ alors $f(b) - f(a) = \int_a^b f'(t)\,\mathrm{d}t$.*

Démonstration. D'après le lemme 3.6 la fonction $\varphi : x \mapsto f(x) - f(a) - \int_a^x f'(t)\,\mathrm{d}t$ est $\mathcal{C}^1$ sur $[a,b]$ de dérivée nulle. Par l'inégalité des accroissement finis (cor. 3.3), la fonction φ est constante. Comme $\varphi(a) = 0$ elle est nulle. Donc $\varphi(b) = 0$. $\square$

Corollaire 3.8 (Intégration par parties)**.** *Si f et g sont $\mathcal{C}^1$ de $[a,b]$ dans $\mathbb{R}$, alors*

$$\int_a^b f(t)g'(t)\,\mathrm{d}t = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)\,\mathrm{d}t.$$

Démonstration. Comme fg est $\mathcal{C}^1$, on peut lui appliquer le théorème fondamental de l'analyse. Grâce à la règle de Leibniz on obtient :

$$[f(t)g(t)]_a^b = \int_a^b (fg)'(t)\,\mathrm{d}t = \int_a^b f'(t)g(t)\,\mathrm{d}t + \int_a^b f(t)g'(t)\,\mathrm{d}t.$$ $\square$

3.3 Le retour des dérivées partielles

Grâce à l'inégalité des accroissements finis, on sait aussi donner un critère sur les dérivées partielles d'une fonction pour qu'elle soit différentiable. En pratique, c'est souvent comme ça qu'on prouve la différentiabilité de fonctions explicites.

Proposition 3.9. *Soient $U \subset \mathbb{R}^n$ un ouvert, $x \in U$ et $f : U \to F$. Si f admet des dérivées partielles sur U, et si pour tout $i \in [\![1, n]\!]$ la fonction $\partial_i f$ est continue en x (par rapport à toutes les variables) alors f est différentiable en x.*

Démonstration. Si on savait déjà que f est différentiable en x, alors sa différentielle en x serait l'application linéaire continue $L : (h_1, \dots, h_n) \mapsto \sum_{i=1}^n h_i \partial_i f(x)$. On connait donc l'unique candidat pour être $D_x f$ et il faut vérifier qu'il convient.
Dans cette preuve, on note $B(z, R)$ la boule fermée de centre z et de rayon R dans $\mathbb{R}^n$ pour $\|\cdot\|_\infty$. Soit $h \in \mathbb{R}^n$ tel que $B(x, \|h\|_\infty) \subset U$ et $y = x + h$, on a :

$$\begin{aligned}
\|f(x+h) - f(x) - L(h)\| =& \left\| f(y_1, \dots, y_n) - f(x_1, \dots, x_n) - \sum_{i=1}^n h_i \partial_i f(x) \right\| \\
\leqslant& \|f(y_1, y_2, \dots, y_n) - f(x_1, y_2, \dots, y_n) - h_1 \partial_1 f(x)\| \\
&+ \|f(x_1, y_2, y_3, \dots, y_n) - f(x_1, x_2, y_3, \dots, y_n) - h_2 \partial_2 f(x)\| \\
&+ \dots + \|f(x_1, \dots, x_{n-1}, y_n) - f(x_1, \dots, x_{n-1}, x_n) - h_n \partial_n f(x)\|,
\end{aligned}$$

où tous les points intermédiaires sont dans $B(x, \|h\|_\infty) \subset U$. Le i-ème terme de la somme est $\|g_i(h_i) - g_i(0)\|$ où on a noté $g_i : t \mapsto f(x_1, \dots, x_{i-1}, x_i + t, y_{i+1}, \dots, y_n) - t\partial_i f(x)$. Par hypothèse g_i est dérivable sur un voisinage de $[0, h_i]$, et $g_i'(t) = \partial_i f(x_1, \dots, x_{i-1}, x_i + t, y_{i+1}, \dots, y_n) - \partial_i f(x)$ pour tout $t \in [0, h_i]$.
Fixons $\varepsilon > 0$. Par continuité en x des $\partial_i f$, il existe $\delta > 0$ tel que $\|\partial_i f(x + k) - \partial_i f(x)\| \leqslant \varepsilon$ pour tout $i \in [\![1, n]\!]$ et $k \in B(0, \delta)$. Si $\|h\|_\infty \leqslant \delta$, alors pour tout $i \in [\![1, n]\!]$, on a $\sup_{t \in [0, h_i]} \|g_i'(t)\| \leqslant \varepsilon$. Donc, par l'inégalité des accroissements finis, $\|g_i(h_i) - g_i(0)\| \leqslant \varepsilon |h_i|$. Ainsi, dès que $\|h\|_\infty \leqslant \delta$, on a $\|f(x+h) - f(x) - L(h)\| \leqslant \varepsilon \sum_{i=1}^n |h_i| \leqslant n\varepsilon \|h\|_\infty$. Donc $\|f(x+h) - f(x) - L(h)\| \underset{h \to 0}{=} o(h)$. □

Corollaire 3.10. *Si f admet des dérivées partielles continues sur U alors f est $\mathcal{C}^1$.*

Démonstration. La proposition 3.9 montre que f est différentiable sur U. Pour tout $x \in U$, on a $D_x f : (h_1, \dots, h_n) \mapsto \sum_{i=1}^n h_i \partial_i f(x)$. Soient $x, y \in U$ et $h \in \mathbb{R}^n$, on a

$$\|D_y f \cdot h - D_x f \cdot h\| \leqslant \sum_{i=1}^n |h_i| \|\partial_i f(y) - \partial_i f(x)\| \leqslant \|h\| \sum_{i=1}^n \|\partial_i f(y) - \partial_i f(x)\|.$$

Donc $\|D_y f - D_x f\| \leqslant \sum_{i=1}^n \|\partial_i f(y) - \partial_i f(x)\| \xrightarrow[y \to x]{} 0$ et Df est continue de U vers $\mathcal{L}(\mathbb{R}^n, F)$. □

3.4 Suite de fonctions différentiables

Proposition 3.11. *Soit U un ouvert convexe borné et soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions différentiables de U dans F. On suppose que :*

- *il existe $x_0 \in U$ tel que $(f_n(x_0))_{n \in \mathbb{N}}$ converge ;*
- *il existe $g : U \to \mathcal{L}(E, F)$ telle que $Df_n \xrightarrow[n \to +\infty]{} g$ uniformément.*

Alors $(f_n)_{n \in \mathbb{N}}$ converge uniformément sur U vers une fonction f différentiable et $Df = g$.

Démonstration. Soient x et $y \in U$, pour tout m et $n \in \mathbb{N}$ on a :

$$\begin{aligned} \|(f_n(x) - f_n(y)) - (f_m(x) - f_m(y))\| &= \|(f_n(x) - f_m(x)) - (f_n(y) - f_m(y))\| \\ &\leqslant \|Df_n - Df_m\|_\infty \|x - y\| \end{aligned} \tag{3.1}$$

en appliquant le théorème des accroissements finis à $f_n - f_m$ différentiable sur U.
Comme U est borné, il existe $R > 0$ tel que $U \subset B(x_0, R)$. L'équation (3.1) avec $y = x_0$ donne

$$\|(f_n - f_n(x_0)) - (f_m - f_m(x_0))\|_\infty \leqslant \|Df_n - Df_m\|_\infty R$$

pour tout $m, n \in \mathbb{N}$. Comme $(Df_n)_{n\in\mathbb{N}}$ converge uniformément elle est de Cauchy uniforme. Donc $(f_n - f_n(x_0))_{n\in\mathbb{N}}$ est de Cauchy uniforme. Comme F est complet, cette suite de fonction converge uniformément. Donc f_n converge uniformément sur U vers une limite notée f.
Soit $x \in U$, pour tout h assez petit et tout $m \in \mathbb{N}$,

$$\begin{aligned} \|f(x+h) - f(x) - g(x) \cdot h\| \leqslant &\|(f(x+h) - f(x)) - (f_m(x+h) - f_m(x))\| \\ &+ \|f_m(x+h) - f_m(x) - D_x f_m \cdot h\| + \|D_x f_m \cdot h - g(x) \cdot h\| \\ \leqslant &2\|g - Df_m\|_\infty \|h\| + \|f_m(x+h) - f_m(x) - D_x f_m \cdot h\|, \end{aligned}$$

où on a utilisé (3.1) avec $y = x + h$ et $n \to +\infty$ pour contrôler le premier terme à droite de la première inégalité. Soit $\varepsilon > 0$, il existe m tel que $\|Df_m - g\|_\infty \leqslant \varepsilon$. Par différentiabilité en x de f_m, pour h assez petit, $\|f_m(x+h) - f_m(x) - D_x f_m \cdot h\| \leqslant \varepsilon\|h\|$. Donc pour tout h assez petit $\|f(x+h) - f(x) - g(x) \cdot h\| \leqslant 3\varepsilon\|h\|$. Donc $\|f(x+h) - f(x) - g(x) \cdot h\| = o(h)$. Donc f est différentiable en x et $D_x f = g(x)$. □

On peut énoncer des variantes de ce résultat sur des ouverts connexes, en remplaçant les convergences par des convergences uniformes sur toute partie bornée. Il y a bien sûr des résultats analogues pour les séries de fonctions différentiables, qu'on laisse aux lecteurices le soin d'énoncer.

Exercice 3.12. Soit E un Banach et $\exp : \mathcal{L}(E) \to GL(E)$ définie par $\exp : L \mapsto \sum_{k\geqslant 0} \frac{L^k}{k!}$. Montrer que exp est $\mathcal{C}^1$ et calculer sa différentielle.

4 Différentielles d'ordres supérieurs

Dans cette section, on s'intéresse aux différentielles d'ordres supérieurs d'une fonction, et aux formules de Taylor que ces différentielles itérées permettent des formuler.

4.1 Différentielle k-ième

Soient E et F deux Banachs et $U \subset E$ un ouvert. Dans suite on notera $\mathcal{L}_k(E, F)$ au lieu de $\mathcal{L}_k(E, \dots, E, F)$ pour tout $k \in \mathbb{N}^*$.

Définition 4.1. Soient $f : U \to F$ et $x \in U$, on dit que f est 2 fois différentiable en x si Df est bien définie au voisinage de x et est différentiable en x. Si c'est le cas, pour tout $v_1, v_2 \in E$ on note $D_x^2 f(v_1, v_2) = (D_x(Df) \cdot v_2) \cdot v_1$.
On dit que f est $k+1$ fois différentiable en x si $D^k f$ est bien définie au voisinage de x et est différentiable en x. On note alors $D_x^{(k+1)} f : (v_1, \dots, v_{k+1}) \mapsto \big(D_x\big(D^k f\big) \cdot v_{k+1}\big) \cdot (v_1, \dots, v_k)$.

On a $D_x(Df) \in \mathcal{L}(E, \mathcal{L}(E, F))$ et $D_x^2 f \in \mathcal{L}_2(E, E, F)$ est l'application bilinéaire qui lui correspond par l'isomorphisme isométrique du lemme 1.16. De même $D_x(D^k f) \in \mathcal{L}(E, \mathcal{L}_k(E, F))$ correspond à $D_x^{k+1} f \in \mathcal{L}_{k+1}(E, F)$.

Définition 4.2. On dit que $f : U \to F$ est k-fois différentiable (resp. $\mathcal{C}^k$) si $D^k f : U \to \mathcal{L}_k(E, F)$ est bien définie (resp. continue) sur U. On dit que f est $\mathcal{C}^\infty$ si elle est $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.

Comme lorsque $k = 1$, ces notions sont stables par combinaison linéaire et D^k est linéaire par rapport à f. Si $F = F_1 \times \cdots \times F_p$ la k-différentiabilité se vérifie composante par composante.

Lemme 4.3. *Si $f : U \to V$ et $g : V \to G$ sont de classe $\mathcal{C}^k$ alors $g \circ f$ aussi.*

Démonstration. La preuve se fait par récurrence. On peut initialiser à $k = 0$ qui est connu. L'étape de récurrence se fait exactement comme dans la preuve du cas $\mathcal{C}^1$, voir corollaire 2.11. □

Exemple 4.4. Les fonctions linéaires ou multilinéaires continues sont $\mathcal{C}^\infty$, de même que les fonctions polynomiales. Les fonctions $\mathrm{Inv} : GL(E) \to GL(E)$ et $\det : M_n(\mathbb{R}) \to \mathbb{R}$ sont $\mathcal{C}^\infty$.

Exercice 4.5. Soit $M \in \mathcal{L}_k(E, F)$ une application k-linéaire symétrique continue. Calculer les différentielles à tout ordre de la fonction $f : x \mapsto \frac{1}{k!} M(x, x, \ldots, x)$ de E dans F.

Indication. Commencer par les cas $k = 1$ et $k = 2$ puis généraliser.

4.2 Dérivées directionnelles et partielles itérées

Lemme 4.6. *Si $f : U \to F$ est $k + 1$ fois différentiable en x, alors pour tout $v_1, \ldots, v_{k+1} \in E$ on a*

$$D_x^{k+1} f \cdot (v_1, \ldots, v_{k+1}) = \partial_{v_{k+1}}(\cdots \partial_{v_2}(\partial_{v_1} f) \cdots)(x).$$

Démonstration. Soit $\mathrm{ev}_v : \mathcal{L}_k(E, F) \to F$ l'application linéaire définie par $\mathrm{ev}_v : M \mapsto M(v_1, \ldots, v_k)$. On calcule

$$\begin{aligned}
\partial_{v_{k+1}}\Big(D^k f(v_1, \ldots, v_k)\Big)(x) &= D_x\Big(\mathrm{ev}_v \circ D^k f\Big) \cdot v_{k+1} = \Big(\mathrm{ev}_v \circ D_x(D^k f)\Big) \cdot v_{k+1} \\
&= \mathrm{ev}_v\Big(D_x(D^k f) \cdot v_{k+1}\Big) = \Big(D_x(D^k f) \cdot v_{k+1}\Big) \cdot (v_1, \ldots, v_k) \\
&= D_x^{k+1} f(v_1, \ldots, v_{k+1}).
\end{aligned}$$

Par récurrence $D_x^{k+1} f(v_1, \ldots, v_{k+1}) = \partial_{v_{k+1}}\big(D^k f(v_1, \ldots, v_k)\big)(x) = \partial_{v_{k+1}}(\cdots \partial_{v_2}(\partial_{v_1} f) \cdots)(x)$. □

Théorème 4.7 (Schwarz). *Si $f : U \to F$ est deux fois différentiable en x alors, pour tout $u, v \in E$, $D_x^2 f(u, v) = D_x^2 f(v, u)$.*

Corollaire 4.8. *Si $f : U \to F$ est k-fois différentiable en x alors $D_x^k f$ est symétrique, au sens où pour tout $v_1, \ldots, v_k \in E$ et tout $\sigma \in \mathfrak{S}_k$ on a $D_x^k f(v_1, \ldots, v_k) = D_x^k f(v_{\sigma(1)}, \ldots, v_{\sigma(1)})$.*

Démonstration. Il suffit de le montrer pour une transposition du type $(i, i + 1)$ puisqu'elles engendrent $\mathfrak{S}_k$. Grâce au lemme 4.6, on peut alors se ramener au cas $k = 2$, qui est le théorème 4.7. □

Le corollaire 4.8 dit que $D_x^k f$ est symétrique dès qu'elle est bien définie. Si f est k-différentiable, alors cela montre que les dérivées directionnelles commutent. Lorsque $E = \mathbb{R}^n$, on a en particulier que les dérivées partielles d'ordre au plus k existent et commutent. Cela justifie les notations du type $\partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n}$ avec $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ et $|\alpha| = \alpha_1 + \cdots + \alpha_n \leqslant k$ puisque l'ordre dans lequel on calcule les dérivées partielles n'est pas important.

Démonstration du thm. 4.7. On pose $S : (u,v) \mapsto f(x+u+v) - f(x+u) - f(x+v) + f(x)$, qui est symétrique en (u,v). Le point essentiel est de prouver que $S(u,v) - D_x^2 f(u,v) = o\big(\|(u,v)\|_\infty^2\big)$ lorsque $(u,v) \to 0$. On en déduit alors que $D_x^2 f(u,v) - D_x^2 f(v,u) = o\big(\|(u,v)\|_\infty^2\big)$. Donc si $(u,v) \neq 0$

$$\frac{D_x^2 f(u,v) - D_x^2 f(v,u)}{\|(u,v)\|_\infty^2} = \frac{D_x^2 f(tu,tv) - D_x^2 f(tv,tu)}{\|(tu,tv)\|_\infty^2} \xrightarrow[t \to 0]{} 0$$

et donc $D_x^2 f(u,v) = D_x^2 f(v,u)$.
Comme f est deux fois différentiable en x, il existe η définie au voisinage de $0 \in E$ à valeurs dans $\mathbb{R}$ telle que $\eta(h) \xrightarrow[h \to 0]{} 0$ et $\|D_{x+h} f - D_x f - D_x(Df) \cdot h\| = \|h\|\eta(h)$. Pour tout u et $v \in E$ assez petits,

$$\|S(u,v) - D_x^2 f(u,v)\| \leqslant \underbrace{\|S(u,v) - D_{x+v} f \cdot u + D_x f \cdot u\|}_{=A} + \underbrace{\|D_{x+v} f \cdot u - D_x f \cdot u - (D_x(Df) \cdot v) \cdot u\|}_{=B}.$$

D'une part, $B \leqslant \|u\| \|D_{x+v} f - D_x f - D_x(Df) \cdot v\| = \|u\| \|v\| \eta(v) \leqslant \|(u,v)\|_\infty^2 \eta(v) = o\big(\|(u,v)\|_\infty^2\big)$.
D'autre part,

$$A = \|f(x+u+v) - f(x+u) - f(x+v) + f(x) - D_{x+v} f \cdot u + D_x f \cdot u\| = \|\alpha(u) - \alpha(0)\|$$

où $\alpha : h \mapsto \underbrace{f(x+v+h) - f(x+h)}_{\text{différentiable}} \underbrace{- D_{x+v} f \cdot h + D_x f \cdot h}_{\text{linéaire}}$ est bien définie et différentiable sur un voisinage de $[0,u] \subset U$. On va appliquer l'inégalité des accroissements finis à α. Pour tout $h \in [0,u]$,

$$\begin{aligned} D_h \alpha &= D_{x+v+h} f - D_{x+h} f - D_{x+v} f + D_x f \\ &= \underbrace{\big(D_{x+v+h} f - D_x f - D_x(Df) \cdot (v+h)\big)}_{\|\cdot\| = \|v+h\| \eta(v+h)} \\ &\quad - \underbrace{\big(D_{x+h} f - D_x f - D_x(Df) \cdot h\big)}_{\|\cdot\| = \|h\| \eta(h)} - \underbrace{\big(D_{x+v} f - D_x f - D_x(Df) \cdot v\big)}_{\|\cdot\| = \|v\| \eta(v)} \end{aligned}$$

Donc $\|D_h \alpha\| \leqslant \|v+h\| \eta(v+h) + \|v\| \eta v + \|h\| \eta(h)$.
Soit $\varepsilon > 0$ il existe $\delta > 0$ tel que si $\|h\| \leqslant 2\delta$ alors $\eta(h) \leqslant \varepsilon$. Si $\|(u,v)\|_\infty \leqslant \delta$, alors $\|u\| \leqslant \delta$ et $\|v\| \leqslant \delta$. Donc pour tout $h \in [0,u]$ on a $\|D_h \alpha\| \leqslant \varepsilon(\|v+h\| + \|v\| + \|h\|) \leqslant 4\varepsilon \|(u,v)\|_\infty$. L'inégalité des accroissements finies montre alors que

$$A = \|\alpha(u) - \alpha(0)\| \leqslant 4\varepsilon \|(u,v)\|_\infty \|h\| \leqslant 4\varepsilon \|(u,v)\|_\infty^2.$$

Donc $A = o\big(\|(u,v)\|_\infty^2\big)$, donc $S(u,v) - D_x^2 f(u,v) = o\big(\|(u,v)\|_\infty^2\big)$. □

En pratique, on montre souvent le caractère $\mathcal{C}^k$ d'une application grâce à la proposition suivante.

Proposition 4.9. *On suppose que $U \subset \mathbb{R}^n$. Soit $f : U \to F$, si pour tout $\alpha \in \mathbb{N}^n$ tel que $|\alpha| = k$ la dérivée partielle $\partial^\alpha f$ est bien définie et continue sur U alors f est $\mathcal{C}^k$.*

Démonstration. Par récurrence à partir du cas $\mathcal{C}^1$ traité dans le corollaire 3.10. □

Définition 4.10. Soient $U \subset \mathbb{R}^n$ et $f : U \to \mathbb{R}$. Si f est 2-différentiable en x, sa matrice *Hessienne* est $H_x(f) = \big(\partial_i \partial_j f(x)\big)_{1 \leqslant i,j \leqslant n}$. C'est la matrice de la forme bilinéaire symétrique $D_x^2 f$ dans la base canonique de $\mathbb{R}^n$.

4.3 Les formules de Taylor

Soient E et F deux espaces de Banach et soit $f : U \to F$ où $U \subset E$ est ouvert.

Définition 4.11. Si $x \in U$ est f est k-différentiable en x, on notera $T_x^k f : E \to F$ l'application

$$T_x^k f : h \longmapsto \sum_{p=0}^{k} \frac{1}{p!} D_x^p f(h, h, \dots, h) = f(x) + D_x f(h) + \frac{1}{2} D_x^2 f(h, h) + \dots + \frac{1}{k!} D_x^k f(h, \dots, h).$$

On dit que $T_x^k f$ est le *développement de Taylor* de f à l'ordre k en x (notation non standard). Si $E = \mathbb{R}^n$ et $F = \mathbb{R}$, on parlera du *polynôme de Taylor* de f à l'ordre k en x.

Remarque 4.12. Attention, si $E \neq \mathbb{R}^n$, qu'est-ce qu'un polynôme défini sur E ?

Exercice 4.13. Si $E = \mathbb{R}^n$ et f est k-différentiable en x, montrer que $\forall h = (h_1, \dots, h_n) \in \mathbb{R}^n$

$$\frac{1}{k!} D_x^k f(h, \dots, h) = \sum_{\alpha_1 + \dots + \alpha_n = k} \frac{1}{\alpha_1! \dots \alpha_n!} \partial_1^{\alpha_1} \dots \partial_n^{\alpha_n} f(x) h_1^{\alpha_1} \dots h_n^{\alpha_n} = \sum_{|\alpha| = k} \frac{1}{\alpha!} \partial^\alpha f(x) h^\alpha.$$

En particulier, si $F = \mathbb{R}$ on a $T_x^k f = \sum_{|\alpha| \leqslant k} \frac{1}{\alpha!} \partial^\alpha f(x) X_1^{\alpha_1} \dots X_n^{\alpha_n} \in \mathbb{R}[X_1, \dots, X_n]$.

Les formules de Taylor permettent d'estimer l'erreur que l'on fait si on remplace la fonction f par son développement de Taylor au voisinage de x, sous diverses hypothèses de régularité.

Théorème 4.14 (Formule de Taylor avec reste intégral)**.** *On suppose que f est $\mathcal{C}^{k+1}$ sur U. Soient x et h tels que $[x, x+h] \subset U$ alors $f(x+h) = T_x^k f(h) + \int_0^1 \frac{(1-s)^k}{k!} D_{x+sh}^{k+1} f(h, \dots, h) \, \mathrm{d}s$.*

Démonstration. Commençons par le cas $E = \mathbb{R}$, $x = 0$ et $h = 1$ que l'on prouve par récurrence. Dans ce cas, $D_y^p f(h, \dots, h) = f^{(p)}(y)$ pour tout y et p. On doit donc montrer que

$$f(1) = \sum_{p=0}^{k} \frac{f^{(p)}(0)}{p!} + \int_0^1 \frac{(1-s)^k}{k!} f^{(k+1)}(s) \, \mathrm{d}s$$

Pour $k = 0$, le théorème fondamental de l'analyse nous dit que $f(1) = f(0) + \int_0^1 f'(s) \, \mathrm{d}s$. L'étape de récurrence, se fait par intégration par parties. Si f est $\mathcal{C}^{k+1}$ alors

$$\int_0^1 \frac{(1-s)^{k-1}}{(k-1)!} f^{(k)}(s) \, \mathrm{d}s = \frac{1}{k!} f^{(k)}(0) + \int_0^1 \frac{(1-s)^k}{k!} f^{(k+1)}(s) \, \mathrm{d}s.$$

Pour le cas général, on applique le cas précédent à $\varphi : t \mapsto f(x + th)$. $\square$

Théorème 4.15 (Formule de Taylor–Lagrange)**.** *On suppose que f est $(k+1)$ fois différentiable sur U et qu'il existe $M \geqslant 0$ tel que pour tout $x \in U$, $\|D_x^{k+1} f\| \leqslant M$. Soient x et h tels que $[x, x+h] \subset U$, alors $\|f(x+h) - T_x^k f(h)\| \leqslant \frac{M}{(k+1)!} \|h\|^{k+1}$.*

Démonstration. Commençons par le cas $E = \mathbb{R}$, $x = 0$ et $h = 1$. Dans ce cas,

$$\left\| f(x+h) - T_x^k f(h) \right\| = \left\| f(1) - \sum_{p=0}^{k} \frac{1}{p!} f^{(p)}(0) \right\| = \|\Phi(1) - \Phi(0)\|,$$

où $\Phi : t \mapsto \sum_{p=0}^{k} \frac{(1-t)^p}{p!} f^{(p)}(t)$. La fonction Φ est dérivable et, pour tout $t \in [0,1]$,

$$\Phi'(t) = \sum_{p=0}^{k} \frac{(1-t)^p}{p!} f^{(p+1)}(t) - \sum_{p=1}^{k} \frac{(1-t)^{p-1}}{(p-1)!} f^{(p)}(t) = \frac{(1-t)^k}{k!} f^{(k+1)}(t),$$

de sorte que $\|\Phi'(t)\| \leqslant \frac{(1-t)^k}{k!} M$ pour tout $t \in [0,1]$. Le majorant est la dérivée de $g : t \mapsto -\frac{(1-t)^{k+1}}{(k+1)!} M$, donc par l'inégalité des accroissements finis version forte, $\|\Phi(1) - \Phi(0)\| \leqslant g(1) - g(0) = \frac{1}{(k+1)!} M$. Cela conclut la preuve du cas particulier.
Pour le cas général, on applique le cas précédent à $\varphi : t \mapsto f(x+th)$. □

Théorème 4.16 (Formule de Taylor–Young)**.** *On suppose que f est k fois différentiable en $x \in U$, alors $f(x+h) = T_x^k f(h) + o(\|h\|^k)$ lorsque $h \to 0$.*

Remarque 4.17. Attention, cette fois on ne peut pas se ramener au cas de dimension 1 et l'appliquer radialement. Si on a $o(\|h\|^k)$ radialement dans toutes les directions, on ne connait pas la dépendance en la direction de la fonction cachée dans le o. C'est trop peu d'information pour recoller les estimations en un $o(\|h\|^k)$ global.

Démonstration du thm. 4.16. On fait une preuve par récurrence sur $k \in \mathbb{N}^*$. Pour $k = 1$ on a $f(x+h) = f(x) + D_x f \cdot h + o(\|h\|)$ par définition de la différentiabilité en x. Soit $k \geqslant 2$, supposons le résultat vrai à l'ordre $k-1$. Soit $f : U \to F$ qui est k-fois différentiable en x.
On note $\varphi : h \mapsto f(x+h) - T_x^k f(h)$, qui est bien définie au voisinage de 0 et k fois différentiable en 0. En particulier, φ est différentiable au voisinage de 0. En utilisant la symétrie et la multilinéarité des $D_x^p f$ (voir exercice 4.5), pour tout h assez petit,

$$D_h\varphi \cdot v = D_{x+h} f \cdot v - \sum_{p=1}^{k} \frac{D_x^p f(v,h,\ldots,h)}{(p-1)!} = D_{x+h} f \cdot v - \underbrace{\sum_{p=0}^{k-1} \frac{1}{p!} (D_x^p(D_f)(h,\ldots,h)) \cdot v}_{=T_x^{k-1}(D_f)(h)\cdot v},$$

donc $\|D_h\varphi\| = \big\|D_{x+h} f - T_x^{k-1}(D_f)(h)\big\| = o\big(\|h\|^{k-1}\big)$ par hypothèse de récurrence.
Soit $\varepsilon > 0$, il existe $\delta > 0$ tel que pour tout $h \in E$ si $\|h\| < \delta$ alors $\|D_h\varphi\| \leqslant \varepsilon \|h\|^{k-1}$. Soit $h \in E$, si $\|h\| < \delta$ on applique l'inégalité des accroissements finis à φ entre 0 et h :

$$\|\varphi(h)\| = \|\varphi(h) - \varphi(0)\| \leqslant \|h\| \sup_{v\in[0,h]} \|D_v\varphi\| \leqslant \|h\| \sup_{v\in[0,h]} \varepsilon\|v\|^{k-1} \leqslant \varepsilon \|h\|^k.$$

Donc $\frac{\|\varphi(h)\|}{\|h\|^k} \xrightarrow[h\to 0]{} 0$, c'est-à-dire $\varphi(h) = f(x+h) - T_x^k f(h) = o\big(\|h\|^k\big)$. □

Corollaire 4.18. *Soient $U \subset \mathbb{R}^n$ et $f : U \to \mathbb{R}$. Si f est 2 fois différentiable en x, $D_x f = 0$ et $D_x^2 f$ est définie-positive alors x est un minimum local de f.*

Démonstration. Par Taylor–Young à l'ordre 2, on obtient $f(x+h) = f(x) + D_x^2 f(h,h) + o(\|h\|^2)$. Pour tout $h \in \mathbb{R}^n$ on a $D_x^2 f(h,h) \geqslant \lambda \|h\|^2$ où $\lambda > 0$ est la plus petite valeur propre de la Hessienne de f en x. Donc $f(x+h) \geqslant f(x) + \|h\|^2(\lambda + o(1))$ est strictement supérieur à $f(x)$ pour h petit. □

5 Inversion locale

Dans cette section on considère des ouverts $U \subset E$ et $V \subset F$, où E et F sont des espaces de Banach. Quand on parle de régularité $\mathcal{C}^k$, on suppose implicitement que $k \geqslant 1$.

5.1 Homéomorphismes et difféomorphismes

Définition 5.1. Soit $f : U \to V$ une bijection on dit que f est un *homéomorphisme* si f et f^{-1} sont continues. On dit que f est un *difféomorphisme* (resp. $\mathcal{C}^k$*-difféomorphisme*) si f et f^{-1} sont différentiables (resp. $\mathcal{C}^k$).

Lemme 5.2. *Soit $f : U \to V$ une bijection. Si f est différentiable en x et f^{-1} est différentiable en $f(x)$ alors $D_x f \in \operatorname{Iso}(E, F)$ et $(D_x f)^{-1} = D_{f(x)}(f^{-1})$.*

Démonstration. On applique la règle de la chaine à $f^{-1} \circ f = \mathrm{Id}_U$ et $f \circ f^{-1} = \mathrm{Id}_V$, ce qui prouve que $D_{f(x)}(f^{-1}) \circ D_x f = \mathrm{Id}_E$ et $D_x f \circ D_{f(x)}(f^{-1}) = \mathrm{Id}_F$. □

Corollaire 5.3. *Si f est $\mathcal{C}^k$ et f^{-1} est différentiable, alors f^{-1} est $\mathcal{C}^k$.*

Démonstration. D'après le lemme 5.2 on a $D(f^{-1}) : y \mapsto \big(D_{f^{-1}(y)} f\big)^{-1} = \operatorname{Inv} \circ Df \circ f^{-1}(y)$. On sait que $\operatorname{Inv} : \operatorname{Iso}(E, F) \to \operatorname{Iso}(F, E)$ est $\mathcal{C}^\infty$ (au moins pour $E = F$ et on peut s'y ramener). Comme Df est $\mathcal{C}^{k-1}$ et f^{-1} est continue, on prouve par récurrence que $D(f^{-1})$ est $\mathcal{C}^r$ pour tout $r \in [\![0, k-1]\!]$. □

Remarque 5.4. Si $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$, s'il existe $f : U \to V$ un difféomorphisme alors $m = n$, car $D_x f$ est un isomorphisme pour tout $x \in U$. C'est encore vrai si on suppose seulement qu'il existe un homéomorphisme $f : U \to V$, mais c'est beaucoup plus dur.

5.2 Le théorème d'inversion locale

Proposition 5.5. *Soit $f : U \to V$ un homéomorphisme. Si f est différentiable en $a \in U$ (resp. $\mathcal{C}^k$ sur U) et $D_a f \in \operatorname{Iso}(E, F)$ (resp. pour tout $x \in U$, $D_x f \in \operatorname{Iso}(E, F)$) alors f^{-1} est différentiable en $f(a)$ (resp. $\mathcal{C}^k$ sur V).*

Démonstration. Le cas $\mathcal{C}^k$ est une conséquence du cas différentiable et du corollaire 5.3 ci-dessus. Prouvons donc le cas différentiable. D'après le lemme 5.2, le seul candidat pour être la différentielle en $f(a)$ de f^{-1} est $(D_a f)^{-1}$, qui est continue par le théorème d'isomorphisme de Banach (thm. 1.10). On doit montrer qu'il convient, c'est-à-dire que $f^{-1}(f(a) + h) = a + (D_a f)^{-1} \cdot h + o(h)$.
Quitte à remplacer f par $x \mapsto f(a + x) - f(a)$ (homéomorphisme d'inverse $y \mapsto f^{-1}(f(a) + y) - a$), on peut supposer que $a = 0$ et $f(a) = 0$. Quitte à remplacer ce nouveau f par $(D_0 f)^{-1} \circ f$, on peut supposer que $F = E$ et $D_0 f = \mathrm{Id}$. On se ramène ainsi à montrer que si f est un homéomorphisme différentiable en 0 tel que $f(0) = 0$ et $D_0 f = \mathrm{Id}$ alors $f^{-1}(y) = y + o(y)$, i.e. $f^{-1}(y) \underset{y \to 0}{\sim} y$.
L'hypothèse de différentiabilité sur f montre que $\frac{f(x) - x}{\|x\|} \xrightarrow[x \to 0]{} 0$. Par continuité de f^{-1} en 0, on a aussi $f^{-1}(y) \xrightarrow[y \to 0]{} 0$ et donc $\frac{f(f^{-1}(y)) - f^{-1}(y)}{\|f^{-1}(y)\|} \xrightarrow[y \to 0]{} 0$. D'où $y - f^{-1}(y) = o\big(f^{-1}(y)\big)$. C'est-à-dire $f^{-1}(y) \underset{y \to 0}{\sim} y$. □

Remarque 5.6. C'est faux si on ne suppose pas $D_a f$ inversible. Considérer par exemple $x \mapsto x^3$ en 0.

Théorème 5.7 (Inversion locale)**.** *Soient $f : U \to F$ une fonction $\mathcal{C}^k$ et $a \in U$ tel que $D_a f$ soit un isomorphisme, alors il existe des ouverts $U' \subset U$ et $V \subset F$ voisinages de a et $f(a)$ respectivement et tels que $f_{|U'} : U' \to V$ soit un $\mathcal{C}^k$-difféomorphisme.*

Remarque 5.8. D'après le théorème d'isomorphisme de Banach, il suffit de vérifier que $D_a f$ est bijective. Lorsque $E = F = \mathbb{R}^n$, la condition $D_a f$ inversible se vérifie souvent en montrant que le déterminant de la matrice jacobienne est non nul.

Démonstration. Comme dans la preuve de la proposition 5.5, on peut se ramener au cas où $E = F$ avec $a = 0 = f(a)$ et $D_a f = \mathrm{Id}$. Comme Df est continue et $GL(E)$ est ouvert, on peut de plus supposer que $D_x f \in GL(E)$ pour tout $x \in U$.
Il suffit de prouver qu'il existe des voisinages $U' \subset U$ et V de 0 tels que $f_{|U'} : U' \to V$ soit un homéomorphisme. Si c'est la cas, f^{-1} est $\mathcal{C}^k$ sur V par la proposition 5.5. Le point essentiel est donc de montrer que f est un homéomorphisme local, et en fait c'est la bijectivité locale qui est difficile.
Bijectivité locale. Soit $y \in E$, on veut montrer que si $\|y\|$ est assez petit alors l'équation $f(x) = y$ admet une unique solution proche de 0. Posons $T_y : x \mapsto y - f(x) + x$, on a $f(x) = y$ si et seulement si x est un point fixe de T_y. Notre but est de montrer que T_y est contractante pour appliquer le théorème du point fixe de Picard.
Par continuité de Df en 0, il existe $R > 0$ tel que $\forall x \in \overline{B(0,R)}$, $\|D_x T_y\| = \|D_x f - \mathrm{Id}\| \leqslant \frac{1}{2}$ et $\overline{B(0,R)} \subset U$. Par l'inégalité des accroissements finis appliquée à $f - \mathrm{Id}$, pour tout $x \in \overline{B(0,R)}$, $\|f(x) - x\| \leqslant \frac{1}{2}\|x\| \leqslant \frac{R}{2}$ et donc $\|T_y(x)\| \leqslant \|y\| + \frac{R}{2}$. Si on suppose que $\|y\| < \frac{R}{2}$, alors T_y envoie $\overline{B(0,R)}$ dans $B(0,R)$ et y est $\frac{1}{2}$-contractante, de nouveau par l'inégalité des accroissements finis.
La boule $\overline{B(0,R)}$ est fermée dans le Banach E, donc complète. Par le théorème de Picard, T_y admet sur cette boule un unique point fixe $g(y) = T_y(g(y)) \in B(0,R)$. En d'autres termes, $g(y)$ est l'unique antécédent de y dans $B(0,R)$. Notons $V = B(0, \frac{R}{2})$ et $U' = B(0,R) \cap f^{-1}(V)$. Comme f est continue, $f^{-1}(V)$ est un voisinage ouvert de 0, et donc U' aussi. L'application $f_{|U'} : U' \to V$ est alors bien définie et bijective, d'inverse g.
Continuité de l'inverse. Soient y_1 et $y_2 \in V$, on note $x_i = g(y_i) \in B(0,R)$. On a alors

$$\|x_1 - x_2\| = \|T_{y_1}(x_1) - T_{y_2}(x_2)\| \leqslant \|T_{y_1}(x_1) - T_{y_1}(x_2)\| + \|T_{y_1}(x_2) - T_{y_2}(x_2)\| \leqslant \frac{1}{2}\|x_1 - x_2\| + \|y_1 - y_2\|$$

puisque T_{y_1} est $\frac{1}{2}$-contractante. Donc $\frac{1}{2}\|x_1 - x_2\| \leqslant \|y_1 - y_2\|$ et $\|g(y_1) - g(y_2)\| \leqslant 2\|y_1 - y_2\|$. □

Corollaire 5.9 (Inversion globale). *Si $f : U \to F$ est $\mathcal{C}^k$ et $D_x f$ est inversible pour tout $x \in U$ alors f est une application ouverte : l'image d'un ouvert est ouverte.*
Si de plus f est injective, alors f est un $\mathcal{C}^k$-difféomorphisme de U vers $f(U)$.

Démonstration. En exercice. □

5.3 Le théorème des fonctions implicites

Dans cette section, on considère E_1, E_2 et F trois espaces de Banach. Soient $U_i \subset E_i$ des ouverts et $f : U_1 \times U_2 \to F$ différentiable en $(x,y) \in U_1 \times U_2$. On note $\partial_i f(x,y) \in \mathcal{L}(E_i, F)$ les applications linéaires définies par $\partial_1 f(x,y) = D_{(x,y)} f(\cdot, 0)$ et $\partial_2 f(x,y) = D_{(x,y)} f(0, \cdot)$. Lorsque $E_1 = \mathbb{R} = E_2$ ces notations sont cohérentes avec celles introduites pour les dérivées partielles. En général, ces différentielles partielles déterminent la différentielle $D_{(x,y)} f : (u,v) \mapsto \partial_1 f(x,y) \cdot u + \partial_2 f(x,y) \cdot v$.

Théorème 5.10 (Fonctions implicites). *Soit $f : U_1 \times U_2 \to F$ de classe $\mathcal{C}^k$ et $(a,b) \in U_1 \times U_2$. On suppose que $f(a,b) = 0$ et $\partial_2 f(a,b) \in \mathrm{Iso}(E_2, F)$. Alors il existe $V_i \subset U_i$, des voisinages ouverts de a et b respectivement, et $\varphi : V_1 \to V_2$ de classe $\mathcal{C}^k$ telle que $\varphi(a) = b$ et $\forall (x,y) \in V_1 \times V_2$, $f(x,y) = 0 \iff y = \varphi(x)$. De plus, $D_a \varphi = -\partial_2 f(a,b)^{-1} \circ \partial_1 f(a,b)$.*

Démonstration. On définit $g : U_1 \times U_2 \to U_1 \times F$ par $g : (x,y) \mapsto (x, f(x,y))$. Alors g est $\mathcal{C}^k$ et $D_{(a,b)} g : E_1 \times E_2 \to E_1 \times F$ est l'application $(u,v) \mapsto (u, \partial_1 f(a,b) \cdot u + \partial_2 f(a,b) \cdot v)$. Comme $\partial_2 f(a,b)$ est un isomorphisme, l'application linéaire suivante est l'inverse de $D_{(a,b)} g$:

$$(u,h) \longmapsto \big(u, \partial_2 f(a,b)^{-1} \cdot \big(h - \partial_1 f(a,b) \cdot u\big)\big).$$

Donc $D_{(a,b)}g$ est un isomorphisme et on peut appliquer le théorème d'inversion locale. Comme on a $g(a,b) = (a,0)$, il existe donc un voisinage $W \subset U_1 \times U_2$ de (a,b) et $R > 0$ tels que g induise un $\mathcal{C}^k$-difféomorphisme de W vers $B(a,R) \times B(0,R) \subset E_1 \times F$.
Vue la forme de g, il existe $\Phi : B(a,R) \times B(0,R) \to E_2$ de classe $\mathcal{C}^k$ tel que $g^{-1} : (x,z) \mapsto (x, \Phi(x,z))$ et on a

$$(x,y) \in W \text{ et } f(x,y) = 0 \iff x \in B(a,R) \text{ et } (x,y) = g^{-1}(x,0) \iff x \in B(a,R) \text{ et } y = \Phi(x,0).$$

Notons $\varphi = \Phi(\cdot, 0) : B(a,R) \to E_2$, les zéros de f dans W sont les points du graphe de φ. Comme W est ouvert, il existe $V_1 \subset B(a,R)$ (resp. V_2) voisinage de a (resp. b) tel que $(a,b) \in V_1 \times V_2 \subset W$. Comme $\varphi(a) = b$ et φ est continu en a, quitte à remplacer V_1 par $V_1 \cap \varphi^{-1}(V_2)$, on peut supposer que $\varphi(V_1) \subset V_2$. Alors $\varphi_{|V_1} : V_1 \to V_2$ est l'application recherchée.
L'application $x \mapsto f(x, \varphi(x))$ est nulle sur V_1 par définition, en la différentiant, on obtient que $\partial_1 f(x, \varphi(x)) + \partial_2 f(x, \varphi(x)) \circ D_x\varphi = 0$ pour tout $x \in V_1$. Si $\partial_2 f(x, \varphi(x))$ est inversible, alors on obtient $D_x\varphi = -\partial_2 f(x, \varphi(x))^{-1} \circ \partial_1 f(x, \varphi(x))$. Pour $x = a$ c'est le résultat recherché. □

Exercice 5.11. Soient $P \in \mathbb{R}_d[X]$ et $x \in \mathbb{R}$ tels que $P(x) = 0$ et $P'(x) \neq 0$. Montrer qu'il existe un voisinage U de P dans $\mathbb{R}_d[X]$ et $\varphi : U \to \mathbb{R}$ de classe $\mathcal{C}^\infty$ telle que pour tout $Q \in U$, $Q(\varphi(Q)) = 0$.

Indication. Considérer l'application $\mathrm{ev} : \mathbb{R}_d[X] \times \mathbb{R}$ définie par $(Q,y) \mapsto Q(y)$.

5.4 Le théorème de la submersion

Dans cette section on va travailler dans $\mathbb{R}^n$ pour se faciliter la vie. Soit $U \subset \mathbb{R}^n$ un ouvert.

Définition 5.12. Soit $f : U \to \mathbb{R}^m$ une application différentiable, on dit que f est une *submersion* (resp. une *immersion*) si pour tout $x \in U$, $D_x f$ est surjective (resp. injective).

Remarque 5.13. Si f est une submersion (resp. immersion) alors $n \geqslant m$ (resp. $n \leqslant m$). Si f est $\mathcal{C}^1$ et $D_a f$ est surjective (resp. injective), alors il existe U voisinage de a sur lequel f est une submersion (resp. immersion), par continuité des déterminants extraits de taille maximale de la jacobienne.

Théorème 5.14 (Submersion)**.** *Soient $f : U \to \mathbb{R}^m$ une application $\mathcal{C}^k$ et $a \in U$ tel que $D_a f$ soit surjective, alors il existe des voisinages U' de a et V de 0 dans $\mathbb{R}^n$ et un $\mathcal{C}^k$-difféomorphisme $\varphi : V \to U'$ tel que $f \circ \varphi : (x_1, \dots, x_n) \mapsto f(a) + (x_1, \dots, x_m)$.*

Démonstration. Quitte à considérer $x \mapsto f(a+x) - f(a)$, on peut ramener au cas $a = 0$ et $f(a) = 0$. Ensuite, quitte à changer de base à la source (ce qui revient à précomposer par un élément de $GL_n(\mathbb{R})$), on peut supposer que $D_0 f = \begin{pmatrix} I_m & 0 \end{pmatrix}$.
On considère l'application $g : (x_1, \dots, x_n) \mapsto (f(x_1, \dots, x_n), x_{m+1}, \dots, x_n)$ de U dans $\mathbb{R}^m \times \mathbb{R}^{n-m}$. Cette application est $\mathcal{C}^k$ et $D_0 g = \begin{pmatrix} I_m & 0 \\ 0 & I_{n-m} \end{pmatrix}$. D'après le théorème d'inversion locale, il existe $U' \subset U$ et V des voisinages de 0 dans $\mathbb{R}^n$ tels que $g : U' \to V$ soit un $\mathcal{C}^k$-difféomorphisme. Notons $g^{-1} = \varphi = (\varphi_1, \dots, \varphi_n)$. Pour tout $x = (x_1, \dots, x_n) \in V$, on a

$$(x_1, \dots, x_n) = g \circ \varphi(x) = (f \circ \varphi(x), \varphi_{m+1}(x), \dots, \varphi_n(x)).$$

En particulier $f \circ \varphi(x_1, \dots, x_n) = (x_1, \dots, x_m)$. □

Ce théorème affirme que, à difféomorphisme à la source près, une submersion $f : U \to \mathbb{R}^m$ ressemble localement à la projection sur les m premières coordonnées. En particulier, son lieu d'annulation

ressemble au sous-espace vectoriel $\{0\}\times\mathbb{R}^{n-m}$. C'est le point de départ de la théorie des sous-variétés différentielles.
Il existe un théorème de l'immersion, dual du précédent. Il affirme qu'une immersion ressemble localement à une injection linéaire, à difféomorphisme près.

Théorème 5.15 (Immersion)**.** *Soient $f : U \to \mathbb{R}^m$ une application $\mathcal{C}^k$ et $a \in U$ tel que D_af soit injective, alors il existe des voisinages V de $f(a)$ et W de 0 dans $\mathbb{R}^m$ et un $\mathcal{C}^k$-difféomorphisme $\psi : V \to W$ tel que $\psi \circ f : (x_1, \dots, x_n) \mapsto (x_1 - a_1, \dots, x_n - a_n, 0, \dots, 0)$.*

Démonstration. En exercice, en utilisant le même genre d'idée que celle du thm. 5.14. □

Pour finir, énonçons le théorème du rang constant dont les deux précédents théorèmes sont des cas particuliers. Il affirme que si D_xf est de rang constant au voisinage de a, alors f est conjuguée à sa différentielle D_af modulo difféomorphismes (i.e. changements de coordonnées) à la source et au but. C'est une façon de formaliser l'idée que f ressemble localement à sa différentielle.

Théorème 5.16 (Rang constant)**.** *Soient $f : U \to \mathbb{R}^m$ une application $\mathcal{C}^k$ et $a \in U$. On suppose que, pour tout $x \in U$, D_xf est de rang exactement $r \leqslant \min(m, n)$. Alors il existe deux $\mathcal{C}^k$-difféomorphismes, φ d'un voisinage U' de 0 dans $\mathbb{R}^n$ vers un voisinage de a et ψ d'un voisinage de $f(a)$ vers un voisinage de 0 dans $\mathbb{R}^m$ tels que, pour tout $x \in U'$, $\psi \circ f \circ \varphi(x) = D_af \cdot x$.*

Démonstration. Voir par exemple [Rou09, exo. 74]. □

Conclusion et ouverture

D'autres sujets auraient eu leur place dans ces notes si on disposait de suffisamment de temps pour les aborder. Deux thèmes que l'on n'a pas abordé mais qu'il faut maitriser pour l'Agrégation sont :

- le cas particulier des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$ (théorème de Rolle, égalité des accroissements finis, lien entre sens de variation et signe de la dérivée, ...) ;
- la recherche et l'étude des extrema et points critiques d'une fonction à valeurs dans $\mathbb{R}$.

Pour conclure voici quelques thèmes dans le prolongement de ce cours qui peuvent être explorés par les étudiant·e·s qui se sentent à l'aise avec le calcul différentiel.

- Les sous-variétés de $\mathbb{R}^n$, et plus loin les variétés différentielles. Voir [Lau11, chap. III].
- Le lemme de Morse [Lau11, II.5.5] est un classique de l'Agrégation, même si son intérêt mathématique n'est pas évident à justifier à ce niveau.
- La théorie des équations différentielles contient de belles applications du calcul différentiel, y compris parfois en dimension infinie. On en trouvera plusieurs dans [Lau11, chap. IV], dont le théorème de redressement du champ [Lau11, IV.2.5] (ou théorème de la boite de flot) et le caractère $\mathcal{C}^k$ du flot [Lau11, IV.1.2].
- Le calcul des variations, qui est une importante motivation pour faire du calcul différentiel en dimension infinie. On peut consulter [SR08, XIV.4] ou [Rou09, exo 47 et 133] par exemple, pour une première approche de ces questions.
- Un exemple de calcul des variations accessible est de chercher à établir l'équation des géodésiques dans $\mathbb{R}^n$. Le problème de minimiser la longueur d'un chemin dont on a fixé les extrémités (problème global), se traduit via le calcul des variations par une équation différentielle que doit satisfaire le paramétrage (problème local). Le résultat n'est cependant pas bien surprenant, le plus court chemin de a à b est le segment $[a, b]$ ...