

Table des matières

1	Définitions, notations et rappels	2
1.1	Bibliographie	2
1.2	Espaces métriques complets	2
1.3	Applications	4
1.4	Espaces vectoriels normés et préhilbertiens.	6
1.5	Espaces de Banach	7
1.5.1	Exemples d'espaces de Banach	8
1.6	Critère de Cauchy pour les séries.	11
2	Différentielle d'une application	13
2.1	Quelques définitions	13
2.2	Exemples	16
2.2.1	Application affine	16
2.2.2	Un exemple en dimension finie	16
2.2.3	Un exemple en dimension infinie	16
2.3	Différentielles partielles	17
2.3.1	Différentiabilité dans une « direction »	17
2.3.2	Dérivées partielles et matrice jacobienne	18
2.4	Différentiabilité et opérations classiques	19
2.4.1	Combinaison linéaire et composition	19
2.4.2	Applications bilinéaires et produits	20
2.4.3	Inverse	22
2.4.4	Applications multilinéaires	24
3	Théorème des accroissements finis	25
3.1	Théorème des accroissements finis	25
3.2	Applications de classe \mathcal{C}^1	27
3.3	Sur la différentiabilité au sens de Gâteaux	29
3.4	Propriétés des fonctions \mathcal{C}^1	30
4	Théorème d'inversion locale	33
4.1	Théorème d'inversion locale	33
4.2	Le théorème des fonctions implicites	37
5	Différentielles d'ordre supérieur	40
5.1	Différentielle seconde	40
5.1.1	Définition	40

Chapitre 1

Définitions, notations et rappels

1.1 Bibliographie

- *Calcul différentiel et équations différentielles*, Sylvie Benzoni-Gavage, éditions Dunod.
- *Cours et exercices de calcul différentiel pour la licence*, Bazanfaré Mahaman, éditions Ellipses.
- *Équations différentielles ordinaires avec applications*, Basem Attili et Rima Cheaytoui, éditions Ellipses.
- Pour aller plus loin dans la théorie : *Calcul différentiel*, Henri Cartan, éditions Hermann
- Pour « visualiser » les équations différentielles et leurs applications au chaos : www.chaos-math.org/fr/les-chapitres

1.2 Espaces métriques complets

Dans toute la section, (X, d) est un espace métrique.

Définition 1.1 (suites de Cauchy). Soit $(u_k) \subset X$ une suite d'éléments de X . On dit que (u_k) est une suite de Cauchy si

$$\forall \varepsilon > 0 \exists k_0 \geq 0 \text{ tel que pour } j, k \geq k_0, \text{ on a } d(u_j, u_k) < \varepsilon.$$

Proposition 1.2. Toute suite convergente d'éléments de X est de Cauchy.

Démonstration. Soit $(u_k) \subset X$ une suite convergente et soit $\varepsilon > 0$. Il existe $k_0 \geq 0$ tel que

$$k \geq k_0 \implies d(u_k, x) < \varepsilon/2.$$

Par inégalité triangulaire, pour $j, k \geq k_0$, on a

$$d(u_j, u_k) \leq d(u_j, x) + d(x, u_k) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

Proposition 1.3. Soit (u_k) une suite de Cauchy.

- a) La suite (u_k) est bornée.
- b) Si (u_k) admet une valeur d'adhérence, alors elle converge vers cette valeur d'adhérence. En particulier, (u_k) a au plus une valeur d'adhérence.

Démonstration.

(a) Il suffit d'appliquer la définition avec $\varepsilon = 1$, il existe $k_0 \geq 0$ tel que

$$d(u_k, u_{k_0}) \leq 1, \text{ pour } k \geq k_0.$$

On pose alors

$$r = 1 + \max \{d(u_j, u_{k_0}) : 0 \leq j < k_0\}.$$

On a $(u_k) \subset BF(u_{k_0}, r)$. Ainsi (u_k) est bornée.

(b) Soit x une valeur d'adhérence de (u_k) . Soit u_φ une sous-suite convergent vers x . Montrons que u converge vers x .

Soit $\varepsilon > 0$. Il existe j_0 tel que pour $j \geq j_0$, on a $u_{\varphi(j)} \in B(x, \varepsilon/2)$.

Comme u est de Cauchy, il existe k_0 tel que pour $k, l \geq k_0$, on a $d(u_l, u_k) < \varepsilon/2$.

Soit $j \geq 0$ tel que $j \geq j_0$ et $\varphi(j) \geq k_0$. Pour $k \geq \varphi(j)$, on a

$$d(u_k, x) \leq d(u_k, u_{\varphi(j)}) + d(u_{\varphi(j)}, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

On a bien $u \rightarrow x$. □

Définition 1.4 (Espaces métriques complets). *Si toute suite de Cauchy de X est convergente, on dit que (X, d) est un espace métrique **complet**.*

Exemple 1.5.

\mathbb{Q} muni de la distance usuelle n'est pas complet. Par exemple la suite

$$u_k := \sum_{j=0}^k \frac{1}{j!} \in \mathbb{Q}$$

est bien une suite de rationnels qui converge dans \mathbb{R} vers $e = \exp(1)$. Elle est donc de Cauchy. Par contre $e \notin \mathbb{Q}$.

Théorème 1.6. *L'ensemble \mathbb{R} des nombres réels est complet.*

Démonstration. Si $(x_k)_{k \geq 0} \subset \mathbb{R}$ est une suite de Cauchy pour la distance usuelle, alors en particulier, elle est bornée. Par le théorème de Bolzano-Weierstrass, (x_k) admet une sous-suite convergente. La suite (x_k) admet donc une valeur d'adhérence et comme elle est de Cauchy, elle converge. □

Théorème 1.7. \mathbb{R}^n est complet.

Démonstration. Si $(x^{(k)}) \subset \mathbb{R}^n$ est une suite de Cauchy, alors pour $1 \leq j \leq n$, les suites $((x_j^{(k)}))_k$ sont aussi des suites de Cauchy d'éléments de \mathbb{R} . Comme \mathbb{R} est complet, elles convergent. Ceci implique la convergence de $(x^{(k)})$ dans \mathbb{R}^n . □

Théorème 1.8. *Les espaces métriques compacts sont complets.*

Démonstration. Soit (K, d) un espace métrique compact et soit $(x_k) \subset K$ une suite de Cauchy. Par compacité de K , cette suite admet une valeur d'adhérence et comme elle est de Cauchy, elle converge. □

Théorème 1.9. *Soit $Y \subset X$.*

a) *Si (Y, d) est complet alors Y est fermé.*

b) *Si (X, d) est complet et Y fermé alors (Y, d) est complet.*

En particulier si (X, d) est un espace métrique complet on a l'équivalence

$$Y \text{ fermé} \iff (Y, d) \text{ complet.}$$

Théorème 1.10. Si (X, d) et (Y, δ) sont des espaces métriques et que (Y, δ) est **complet** alors $C_b(X, Y)$ muni de la distance

$$\delta_\infty(f, g) := \sup \{ \delta(f(x), g(x)) : x \in X \}$$

est complet.

Démonstration. Soit $(f_k) \subset C_b(X, Y)$ une suite de Cauchy. On voit facilement que pour tout $x \in X$, la suite $(f_k(x))$ est de Cauchy dans (Y, δ) . Comme (Y, δ) est complet, il existe $y \in Y$ tel que $f(x_k) \rightarrow y$ dans Y . Notons $y =: f(x)$.

Montrons que $f_k \xrightarrow{u} f$. Soit $\varepsilon > 0$. Il existe $k_0 \geq 0$ tel que pour tous $k, l \geq k_0$ on ait

$$\delta(f_k(x), f_l(x)) < \varepsilon \quad \forall x \in X.$$

En faisant $l \rightarrow \infty$, on obtient : pour $k \geq k_0$,

$$\delta(f_k(x), f(x)) \leq \varepsilon \quad \forall x \in X.$$

On conclut que $f_k \xrightarrow{u} f$. Comme une limite uniforme de fonctions continues est continue on a $f \in C_b(X, Y)$. Ceci achève la démonstration. \square

1.3 Applications

Le théorème suivant permet de s'assurer de l'existence d'une limite sans connaître sa valeur explicitement.

Théorème 1.11 (critère de Cauchy). Soit (X, d) et (Y, δ) deux espaces métriques. On suppose que (Y, δ) est **complet**. Soit $A \subset X$, $f : A \rightarrow Y$ et $x \in \bar{A}$. Alors f admet une limite au point x si et seulement si

$$\forall \varepsilon > 0 \exists r > 0 \text{ tel que } [x' \in B(x, r) \text{ et } x'' \in B(x, r)] \implies \delta(f(x'), f(x'')) < \varepsilon.$$

Corollaire 1.12 (prolongement des fonctions lipschitziennes). Soit (X, d) et (Y, δ) deux espaces métriques. On suppose que (Y, δ) est **complet**. Soit $A \subset X$ et $f : A \rightarrow Y$ une fonction lipschitzienne. Alors :

- f admet une limite en tout point de \bar{A} .
- le prolongement par continuité de f sur \bar{A} est une fonction lipschitzienne (avec la même constante de Lipschitz).

Le résultat suivant, très important, est à la base de nombreux théorèmes d'existence. Par exemple, il va être utilisé pour démontrer l'existence d'une solution unique au problème suivant. Étant donné un intervalle $I \subset \mathbb{R}$, $x_0 \in I$, $y_0 \in \mathbb{R}$ et $f \in \text{Lip}(\mathbb{R}, \mathbb{R})$, on cherche une solution $y \in C^1(I, \mathbb{R})$ de

$$\begin{cases} y'(x) &= f(y(x)) & \forall x \in I, \\ y(x_0) &= y_0. \end{cases}$$

Ce type de problème formé d'une équation différentielle ($y' = f(y)$) et d'une donnée initiale ($y(x_0) = y_0$) est appelé problème de Cauchy.

Théorème 1.13 (Théorème du point fixe de Banach-méthode des approximations successives de Picard). Soit (X, d) un espace métrique **complet** et soit $f : X \rightarrow X$ une application contractante, i.e. qu'il existe une constante $0 < K < 1$ telle que f soit K -Lipschitzienne :

$$d(f(x), f(y)) \leq Kd(x, y), \quad \forall x, y \in X.$$

Alors f admet un unique point fixe dans X :

$$\exists! x \in X \text{ tel que } f(x) = x.$$

De plus toute suite d'éléments de X construite par la relation de récurrence $x_{k+1} := f(x_k)$ converge vers ce point fixe.

Démonstration.

1. Unicité. Soit $x, y \in X$ deux points fixes de f , i.e. $x = f(x)$ et $y = f(y)$. Comme f est K -Lipschitzienne, on a

$$d(x, y) = d(f(x), f(y)) \leq Kd(x, y) \implies (1 - K)d(x, y) \leq 0.$$

Comme $1 - K > 0$, on en déduit $d(x, y) = 0$ et donc $x = y$.

2. Existence. Soit $x_0 \in X$. Nous définissons de manière récursive pour $k \geq 0$,

$$x_{k+1} := f(x_k). \tag{1.1}$$

On a alors, pour $k \geq 0$,

$$d(x_{k+2}, x_{k+1}) = d(f(x_{k+1}), f(x_k)) \leq Kd(x_{k+1}, x_k).$$

On obtient par récurrence l'estimation :

$$d(x_{k+1}, x_k) \leq K^k d(x_1, x_0) \text{ pour } k \geq 0. \tag{1.2}$$

Maintenant, par inégalité triangulaire, on a pour $l \geq k$,

$$d(x_l, x_k) \leq d(x_k, x_{k+1}) + d(x_{k+1}, x_{k+2}) + \dots + d(x_{l-1}, x_l).$$

En utilisant (1.2) pour estimer chacun des termes, il vient

$$\begin{aligned} d(x_l, x_k) &\leq (K^k + K^{k+1} + \dots + K^{l-1})d(x_1, x_0) \\ &\leq \left(\sum_{j \geq k} K^j \right) d(x_1, x_0) = \frac{K^k}{1 - K} d(x_1, x_0) \end{aligned}$$

Comme $0 \leq K < 1$, le membre de droite converge vers 0 quand $k \rightarrow \infty$. On en déduit que la suite (x_k) est de Cauchy. Par hypothèse, (X, d) est complet et donc il existe $x^* \in X$ tel que $x_k \xrightarrow{k \rightarrow \infty} x^*$. En passant à la limite dans l'identité $x_{k+1} = f(x_k)$, on conclut par continuité de f que $x^* = f(x^*)$. Nous venons de montrer que f admet au moins x^* comme point fixe. De plus toute suite récurrente de la forme (1.1) converge vers un point fixe de f . Nous avons vu au point 1 que ce point fixe était unique. On en déduit que toute suite de la forme (1.1) converge vers x^* , l'unique point fixe de f . \square

Exemple 1.14. L'équation $x = \cos x$ a une seule solution sur \mathbb{R} . En effet, comme $\cos x \in [-1, 1]$ pour $x \in \mathbb{R}$, on peut se restreindre à cet intervalle. On cherche donc les points fixes de $f : [-1, 1] \rightarrow [-1, 1]$, $x \mapsto \cos x$. Sur cet intervalle, on a

$$\forall x \in [-1, 1], |f'(x)| \leq \sin 1 < 1.$$

Donc f est K -Lipschitzienne, avec $K = \sin 1 < 1$. Elle admet donc un unique point fixe. On conclut que l'équation $\cos x = x$ admet une solution et une seule sur \mathbb{R} .

1.4 Espaces vectoriels normés et préhilbertiens.

Définition 1.15 (norme). Soit E un espace vectoriel sur \mathbb{R} . Une norme sur E est une application $N : E \rightarrow \mathbb{R}_+$ qui satisfait les trois axiomes suivants :

- i. Pour tout $x \in E$, on a $N(x) = 0 \iff x = 0$.
- ii. Pour tout $x \in E$ et $\lambda \in \mathbb{R}$, on a $N(\lambda x) = |\lambda|N(x)$.
- iii. Pour tous $x, y \in E$ on a $N(x + y) \leq N(x) + N(y)$.

On dit que E muni de N est un espace vectoriel normé.

Remarque 1.16. La plupart du temps, une norme est notée $\|x\|$ ou $|x|$ plutôt que $N(x)$.

Exemple 1.17. Soit $E = \mathbb{R}^n$. Pour $x \in E$, posons

- $N_\infty(x) = |x|_\infty := \max_{1 \leq k \leq n} |x_k|$,
- $N_1(x) = |x|_1 := \sum_{k=1}^n |x_k|$.

Il est facile de vérifier que $|\cdot|_\infty$ et $|\cdot|_1$ sont des normes sur E .

Exemple 1.18. Soit $l^\infty(\mathbb{R})$ l'ensemble des suites bornées d'éléments de \mathbb{R} . Pour $u = (u_k)_{k \geq 0}$, posons

$$N_\infty(u) = |u|_\infty = \sup_{k \geq 0} |u_k|.$$

$|\cdot|_\infty$ est une norme sur $l^\infty(\mathbb{R})$.

Exemple 1.19. Soit $l^1(\mathbb{R})$ l'ensemble des suites sommables d'éléments de \mathbb{R} . Pour $u = (u_k)_{k \geq 0}$, posons

$$N_1(u) = |u|_1 = \sum_{k=0}^{\infty} |u_k|.$$

N_1 est une norme sur $l^1(\mathbb{R})$.

Exemple 1.20. Soit I un intervalle de \mathbb{R} et $C_b(I)$ l'ensemble des fonctions continues et bornées sur I à valeurs dans \mathbb{R} . Pour $f \in C_b(I)$, on pose

$$N_\infty(f) = \|f\|_\infty = \sup_{x \in I} |f(x)|.$$

On vérifie facilement que $\|\cdot\|_\infty$ définit une norme sur $C_b(I)$.

Soit E un espace vectoriel et considérons une application de $b : E \times E \rightarrow \mathbb{R}$ notée $b(x, y) = \langle x, y \rangle$ pour $x, y \in E$.

Définition 1.21 (produit scalaire). On dit que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur E si

- i. $\langle x, y \rangle = \langle y, x \rangle$ pour tous $x, y \in E$,
- ii. $\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle x, z \rangle$ pour tous $x, y, z \in E$ et $\lambda \in \mathbb{R}$,
- iii. $\langle x, x \rangle \geq 0$ pour tout $x \in E$,
- iv. $\langle x, x \rangle = 0 \implies x = 0$ pour tout $x \in E$.

On dit que E muni de $\langle \cdot, \cdot \rangle$ est un espace préhilbertien. La norme associée est

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Proposition 1.22. Si $(E, \langle \cdot, \cdot \rangle)$ est un espace préhilbertien de norme associée $|\cdot|$ on a les identités remarquables

$$\begin{aligned} \langle x + y, x - y \rangle &= |x|^2 - |y|^2 \\ |x + y|^2 &= |x|^2 + 2\langle x, y \rangle + |y|^2 \\ |x - y|^2 &= |x|^2 - 2\langle x, y \rangle + |y|^2 \end{aligned}$$

Proposition 1.23 (formule de polarisation, cas réel). *On déduit des identités ci-dessus que pour $x, y \in E$, on a*

$$\langle x, y \rangle = (|x + y|^2 - |x - y|^2)/4.$$

Ainsi il suffit de connaître la norme associée $|\cdot|$ pour déterminer le produit scalaire.

Exemple 1.24. Si $E = \mathbb{R}$, le produit $\langle t, s \rangle := ts$ définit un produit scalaire. Plus généralement, si $a > 0$, $a \neq 1$ est fixé, $\langle t, s \rangle := ats$ définit un autre produit scalaire.

Exemple 1.25. Si $E = \mathbb{R}^n$, le produit scalaire euclidien usuel

$$\langle x, y \rangle := \sum_{k=1}^n x_k y_k$$

est bien un produit scalaire.

Exemple 1.26. Si $E = l^2(\mathbb{R})$ est l'ensemble des suites à valeurs réelles et de carré sommable. L'application définie pour $u = (u_k)_{k \geq 0}$ et $v = (v_k)_{k \geq 0}$ par

$$\langle u, v \rangle := \sum_{k \geq 0} u_k v_k$$

définit un produit scalaire sur $l^2(\mathbb{R})$. \square

Exemple 1.27. Si $E = \mathcal{M}_n(\mathbb{R})$ est l'ensemble des matrices réelles de taille $n \times n$, on peut définir le produit scalaire suivant sur E :

$$\langle M, N \rangle := \sum_{k=1}^n \sum_{l=1}^n M_{k,l} N_{k,l}.$$

Exemple 1.28. Si $E = C([0, 1])$ est l'ensemble des fonctions continues sur $[0, 1]$, on définit un produit scalaire sur E par

$$\langle f, g \rangle := \int_0^1 f(t)g(t) dt.$$

Exemple 1.29. Soit $n \geq 0$ et soit $E = \mathbb{R}_n[X]$ l'espace des polynômes à coefficients réels de degré plus petit que n . Alors

$$\langle P, Q \rangle := \int_0^1 P(t)Q(t) dt.$$

définit un produit scalaire sur $\mathbb{R}_n[X]$.

Définition 1.30. *Un espace préhilbertien de dimension finie est appelé espace Euclidien.*

Théorème 1.31. *Dans un espace de dimension finie, toutes les normes sont équivalentes.*

1.5 Espaces de Banach

Définition 1.32. a) *Un espace vectoriel normé complet est appelé un espace de Banach.*

b) *Un espace préhilbertien complet est appelé un espace de Hilbert.*

Les propriétés établies pour les espaces métriques complets sont bien sûr valables pour les espaces de Banach.

Proposition 1.33.

- a) \mathbb{R}^n et plus généralement tout espace vectoriel normé de dimension finie est un espace de Banach.
- b) Tout espace euclidien (i.e. un espace préhilbertien de dimension finie) est un espace de Hilbert.
- c) Si $(E_1, \|\cdot\|_1), \dots, (E_N, \|\cdot\|_N)$ sont des espaces de Banach, alors l'espace produit muni de la norme $\|(x_1, \dots, x_N)\|_{\max} := \max \{\|x_j\|_j : j = 1, \dots, N\}$ est un espace de Banach.
- d) Si $(E, \|\cdot\|)$ est un espace de Banach et F est un sous-espace vectoriel **fermé** de E alors $(F, \|\cdot\|)$ est un espace de Banach.

1.5.1 Exemples d'espaces de Banach

Proposition 1.34. Soit $(E, \|\cdot\|)$ un espace de Banach, soit I un ensemble et soit $l^\infty(I, E)$ l'espace des fonctions bornées de I à valeurs dans E . Pour $f \in l^\infty(I, E)$ on pose

$$\|f\|_\infty := \sup \{|f(x)| : x \in I\}.$$

Alors, $(l^\infty(I, E), \|\cdot\|_\infty)$ est un espace de Banach.

Dans le cas particulier $I = \mathbf{N}$ et $\mathbb{R} = E$ on retrouve l'espace des suites bornées d'éléments de \mathbb{R}

$$l^\infty(\mathbf{N}, \mathbb{R}) = l^\infty(\mathbb{R}), \quad \|\cdot\|_\infty = |\cdot|_\infty.$$

On a donc :

Corollaire 1.35. L'espace $(l^\infty(\mathbb{R}), |\cdot|_\infty)$ des suites bornées de \mathbb{R} est un espace de Banach.

Démonstration.

Notons $G = l^\infty(I, E)$. Montrons que l'e.v.n. $(G, \|\cdot\|_\infty)$ est complet. Soit $(f_j)_j$ une suite de Cauchy dans G , i.e.

$$\forall \varepsilon > 0 \exists l_0 \geq 0 \quad \text{tel que} \quad j, l \geq l_0 \implies \|f_j - f_l\|_\infty < \varepsilon.$$

Comme pour tout $x \in I$, $|f_j(x) - f_l(x)| \leq \|f_j - f_l\|_\infty$, chacune des suites $(f_j(x))_{j \geq 0} \subset E$ est de Cauchy dans E et est donc convergente. Notons $f(x) := \lim_{j \rightarrow \infty} f_j(x)$.

Pour conclure, nous allons montrer que $f \in G$ et que $\|f_j - f\|_\infty \rightarrow 0$.

1. Montrons que $f \in E$.

La suite (f_j) étant de Cauchy, elle est bornée dans G . Soit $C \geq 0$ telle que

$$\forall j \geq 0 \quad \|f_j\|_\infty \leq C.$$

Pour tout $x \in I$, on a $|f(x)| = \lim_{j \rightarrow \infty} \underbrace{|f_j(x)|}_{\leq \|f_j\|_\infty \leq C} \leq C$. Donc $\|f\|_\infty \leq C$ et on a bien

$f \in G$.

2. Montrons que $\|f_j - f\|_\infty \rightarrow 0$.

Soit $\varepsilon > 0$ et soit $l_0 \geq 0$ tel que

$$j, l \geq l_0 \implies \|f_j - f_l\|_\infty \leq \varepsilon \implies \forall x \in I, |f_j(x) - f_l(x)| \leq \varepsilon.$$

En faisant $l \rightarrow \infty$ à $j \geq l_0$ fixé, on obtient qu'il existe $l_0 \geq 0$ tel que

$$j \geq l_0 \implies \forall x \in I, |f_j(x) - f(x)| \leq \varepsilon \implies \|f_j - f\|_\infty \leq \varepsilon.$$

Comme $\varepsilon > 0$ est arbitraire, on conclut que $f_j \rightarrow f$ pour la norme $\|\cdot\|_\infty$. □

Proposition 1.36. Soit $(E, |\cdot|)$ un espace de Banach et (X, d) est un espace métrique alors $C_b(X, E)$ est un sous-espace fermé de $(l^\infty(X, E), \|\cdot\|_\infty)$.

Donc $(C_b(X, E), \|\cdot\|_\infty)$ est un espace de Banach.

Proposition 1.37. Soit $(E, |\cdot|)$ un espace de Banach et (X, d) est un espace métrique. On munit $\text{Lip}(X, E)$ de la norme

$$\|f\|_{\text{Lip}} := \|f\|_\infty + \sup \left\{ \frac{|f(x) - f(y)|}{\delta(x, y)} : x, y \in X, x \neq y \right\}.$$

L'espace $(\text{Lip}(X, E), \|\cdot\|_{\text{Lip}})$ est un espace de Banach.

Proposition 1.38. Soit $(E, |\cdot|_E)$ et $(F, |\cdot|_F)$ des espaces vectoriels normés. On munit l'espace $L(E, F)$ des applications linéaires continues de E dans F de la norme

$$\|\ell\| := \sup \{ \|\ell(x)\|_F : x \in E, \|x\|_E \leq 1 \} = \sup_{x \in E \setminus \{0\}} \frac{\|\ell(x)\|_F}{\|x\|_E}.$$

Si F est un Banach alors $L(E, F), \|\cdot\|_{E, F}$ est un Banach.

En particulier, on peut faire $E = \mathbb{R}$ dans les propositions précédentes : les espaces $(C_b(X, \mathbb{R}), \|\cdot\|_\infty)$, $(\text{Lip}(X, \mathbb{R}), \|\cdot\|_{\text{Lip}})$ et $(F', \|\cdot\|_{F'})$ sont des espaces de Banach.

Revenons aux espaces de suites.

Théorème 1.39. a) L'espace de suites $(l^1(\mathbb{R}), |\cdot|_1)$ est un Banach.

b) L'espace de suites $(l^2(\mathbb{R}), |\cdot|_2)$ est un Hilbert.

c) Plus généralement, pour $p \geq 1$, l'espace $(l^p(\mathbb{R}), |\cdot|_p)$ est un espace de Banach.

Nous rappelons que pour $u = (u_k)_{k \geq 0}$ une suite d'éléments de \mathbb{R} , on définit

$$|u|_p := \left(\sum_{k \geq 0} |u_k|^p \right)^{\frac{1}{p}} \in [0, \infty],$$

et

$$l^p(\mathbb{R}) = \{(u_k)_{k \geq 0} \subset \mathbb{R} : |u|_p < \infty\}.$$

Démonstration. Il suffit de montrer le (c). Soit $1 \leq p < \infty$ et soit $(u^{(j)}) \subset l^p(\mathbb{R})$ une suite de Cauchy dans $(l^p(\mathbb{R}), |\cdot|_p)$.

Commençons par remarquer que les suites sommables sont bornées et donc $l^p(\mathbb{R}) \subset l^\infty(\mathbb{R})$. De plus on a l'estimation,

$$\forall x \in l^p(\mathbb{R}) \quad |x|_\infty \leq |x|_p.$$

On voit donc que la suite $(u^{(j)})$ est de Cauchy dans $(l^\infty(\mathbb{R}), |\cdot|_\infty)$. Nous venons de voir que $(l^\infty(\mathbb{R}), |\cdot|_\infty)$ est un espace de Banach, nous en déduisons qu'il existe $x \in l^\infty(\mathbb{R})$ tel que

$$u^{(j)} \xrightarrow{j \rightarrow \infty} x \text{ dans } (l^\infty(\mathbb{R}), |\cdot|).$$

Nous allons montrer maintenant que $x \in l^p(\mathbb{R})$ et que $u^{(j)} \rightarrow x$ dans $l^p(\mathbb{R})$.

1. $x \in l^p(\mathbb{R})$ Comme $(u^{(j)})$ est de Cauchy dans $(l^p(\mathbb{R}), |\cdot|_p)$ elle est bornée dans cet espace : il existe $C \geq 0$ telle que

$$|u^{(j)}|_p \leq C, \quad \forall j \geq 0.$$

D'autre part, la convergence de $u^{(j)}$ vers x dans $l^\infty(\mathbb{R})$ entraîne la convergence de chaque composante : pour $k \geq 0$ fixé, on a $u_k^{(j)} \xrightarrow{j \rightarrow \infty} x_k$ dans \mathbb{R} . En particulier, pour $K \geq 0$, on a

$$\sum_{k=0}^K |x_k|^p = \lim_{j \rightarrow \infty} \left(\underbrace{\sum_{k=0}^K |u_k^{(j)}|^p}_{\leq C^p} \right) \leq C^p.$$

La série de terme général $(|x_k|^p)$ est donc sommable. On en déduit que $x \in l^p(\mathbb{R})$.

2. $u^{(j)} \rightarrow x$ dans $l^p(\mathbb{R})$ Fixons $\varepsilon > 0$. Il existe $l_0 \geq 0$ tel que

$$j, l \geq l_0 \implies |u^{(j)} - u^{(l)}|_p < \varepsilon. \quad (1.3)$$

Comme $u^{(l_0)} - x \in l^p(\mathbb{R})$, il existe k_0 tel que

$$\left(\sum_{k \geq k_0+1} |u_k^{(l_0)} - x_k|^p \right)^{\frac{1}{p}} < \varepsilon. \quad (1.4)$$

Soit maintenant, $j \geq l_0$, on a

$$(|u^{(j)} - x|_p)^p = \sum_{k=0}^{k_0} |u_k^{(j)} - x_k|^p + \sum_{k \geq k_0+1} |u_k^{(j)} - x_k|^p \quad (1.5)$$

Or, par inégalité triangulaire dans $l^p(\mathbb{R})$ (inégalité de Minkowski) on a

$$\begin{aligned} \left(\sum_{k \geq k_0+1} |u_k^{(j)} - x_k|^p \right)^{\frac{1}{p}} &\leq \underbrace{\left(\sum_{k \geq k_0+1} |u_k^{(j)} - u_k^{(l_0)}|^p \right)^{\frac{1}{p}}}_{\leq |u^{(j)} - u^{(l_0)}|_p < \varepsilon \text{ par (1.3)}} + \underbrace{\left(\sum_{k \geq k_0+1} |u_k^{(l_0)} - x_k|^p \right)^{\frac{1}{p}}}_{< \varepsilon \text{ par (1.4)}} \\ &< 2\varepsilon. \end{aligned}$$

En injectant cette estimation dans (1.5), on obtient

$$(|u^{(j)} - x|_p)^p = \sum_{k=0}^{k_0} |u_k^{(j)} - x_k|^p + (2\varepsilon)^p.$$

Finalement, comme à k fixé $u_k^{(j)} - x_k \xrightarrow{j \rightarrow \infty} 0$, on a $\sum_{k=0}^{k_0} |u_k^{(j)} - x_k|^p \xrightarrow{j \rightarrow \infty} 0$. On en déduit qu'il existe j_0 tel que pour $j \geq j_0$,

$$|u^{(j)} - x|_p \leq 3\varepsilon.$$

Comme $\varepsilon > 0$ est arbitraire, on en conclut que $u^{(j)} \rightarrow x$ dans $(l^p(\mathbb{R}), |\cdot|_p)$. \square

Exemple 1.40. Soit Ω une partie de \mathbb{R}^n . Les espaces de Lebesgue $L^p(\Omega)$ sont des espaces de Banach pour tout $p \in [1, +\infty]$. L'espace de Lebesgue $L^2(\Omega)$ muni du produit scalaire

$$\langle f, g \rangle = \int_{\Omega} f(x)g(x) dx$$

est un espace de Hilbert. Cependant, l'espace $C(\Omega)$ des fonctions continues n'est pas un espace de Hilbert muni de ce produit scalaire. Par exemple, si on considère $\Omega = [0, 1]$ et la suite de fonctions

$$f_n(x) = \begin{cases} 1, & x \leq 1/2, \\ 1 - n(x - 1/2), & 1/2 < x < 1/2 + 1/n, \\ 0, & x \geq 1/2 + 1/n, \end{cases}$$

on obtient que (f_n) est une suite de Cauchy pour la norme

$$\|f\|_2 = \left(\int_0^1 f^2(x) dx \right)^{1/2},$$

mais la suite n'a pas de limite dans $C(\Omega)$.

Définition 1.41. On dit que $u \in L(E, F)$ est un isomorphisme de E sur F si u appartient à l'ensemble :

$$\text{Isom}(E, F) = \{u \in L(E, F); \exists v \in L(F, E), v \circ u = Id_E, u \circ v = Id_F\},$$

noté aussi $\text{Inv}(E, F)$.

Si E et F sont des espaces de Banach, on a le résultat suivant (voir H. Brezis, Analyse Fonctionnelle p.18) :

Théorème 1.42 (Théorème de Banach). Si E et F sont des espaces de Banach, la réciproque d'une application linéaire continue et bijective de E sur F , est continue. Autrement dit, si $\ell \in L(E, F)$ est bijective, alors $\ell \in \text{Isom}(E, F)$.

En dimension finie toutes les applications linéaires sont continues. Le théorème n'a donc d'intérêt qu'en dimension infinie.

Définition 1.43. On dit que f est un homéomorphisme de U vers V si f est continue et s'il existe g continue de V vers U telle que

$$g(f(u)) = u, \forall u \in U \quad \text{et} \quad f(g(v)) = v, \forall v \in V.$$

Autrement dit un homéomorphisme est une bijection f telle que f et la bijection réciproque soient toutes les deux continues. On note alors $f \in \text{Hom}(U, V)$.

On dit que E et F sont homéomorphes s'il existe un homéomorphisme de E vers F .

1.6 Critère de Cauchy pour les séries.

Soit $(E, \|\cdot\|)$ un espace vectoriel normé.

Définition 1.44. Soit $(u_n) \subset E$ une suite.

a) On dit que la série de terme général (u_n) est convergente, si la suite

$$S_n := \sum_{k=0}^n u_k$$

est convergente. On note $\sum_{k \geq 0} u_k$ sa limite.

b) On dit que la série est **absolument** convergente si

$$\sum_{k \geq 0} \|u_k\|$$

est convergente.

Autrement dit si il existe $C \geq 0$ telle que $\forall n \in \mathbf{N}$, $\sum_{k=0}^n \|u_k\| \leq C$.

Proposition 1.45 (critère de Cauchy pour les séries). *Dans un espace de Banach, toute série absolument convergente est convergente.*

Démonstration. Soit $(E, \|\cdot\|)$ un espace de Banach et $(u_n) \subset E$ le terme général d'une série absolument convergente. Pour $n \geq 0$, on pose $S_n := \sum_{k=0}^n u_k$. Nous devons montrer que (S_n) est convergente. Comme E est complet il nous suffit d'établir que (S_n) est de Cauchy.

Soit $j \geq l \geq 0$, on calcule

$$\|S_j - S_l\| = \left\| \sum_{k=l+1}^j u_k \right\| \leq \sum_{k=l+1}^j \|u_k\| \leq \sum_{k \geq l+1} \|u_k\| \xrightarrow{l \rightarrow \infty} 0.$$

(Pour le passage à la limite $j \rightarrow \infty$, on a utilisé le fait que la série $\sum \|u_k\|$ était convergente.) On en déduit que (S_n) est bien une suite de Cauchy et est donc convergente. \square

Définition 1.46. Soit $n \in \mathbf{N}$ et $f_n, f : U \rightarrow F$. On dit que la suite de fonctions (f_n) converge uniformément vers f sur U ssi

$$\forall \varepsilon > 0, \exists n_0 \in \mathbf{N}, \forall n \geq n_0, \sup_{x \in U} \|f_n(x) - f(x)\| \leq \varepsilon.$$

Définition 1.47. Soit $f_n : U \rightarrow F$ une suite de fonctions pour $n \in \mathbf{N}$. On dit que la série de fonctions $\sum_{n=0}^{+\infty} f_n$ est normalement convergente sur U si la série numérique $\sum_{n=0}^{+\infty} a_n$ converge, avec $a_n = \sup\{\|f_n(x)\| ; x \in U\}$.

En utilisant la proposition 1.45, on en déduit la proposition suivante.

Proposition 1.48. Soit $f_n : U \rightarrow F$ une suite de fonctions. Si la série $\sum_{n=0}^{+\infty} f_n$ est normalement convergente sur U , alors pour tout $x \in U$, la série $\sum_{n=0}^{+\infty} f_n(x)$ est convergente. De plus, la série converge uniformément sur U , c'est à dire $\sum_{n=0}^m f_n$ converge vers S uniformément sur U lorsque $m \rightarrow +\infty$, où $S(x) = \sum_{n=0}^{+\infty} f_n(x)$.

Chapitre 2

Différentielle d'une application

Dans tout ce cours les espaces considérés sont des *espaces vectoriels normés* sur \mathbb{R} qu'on note en général E et F .

2.1 Quelques définitions

Commençons par le

Lemme 2.1. *Si $u : E \rightarrow F$ est linéaire continue, et vérifie $\lim_{h \rightarrow 0} \frac{\|u \cdot h\|}{\|h\|} = 0$, alors $u = 0$.*

Preuve : L'hypothèse signifie que pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que $\|h\| < \alpha$ implique $\|u \cdot h\| \leq \varepsilon \|h\|$. Soit $x \in E$. Si $x \neq 0$, le vecteur $h = \frac{\alpha x}{2\|x\|}$ a pour norme $\frac{\alpha}{2} < \alpha$, donc $\|u(\frac{\alpha x}{2\|x\|})\| \leq \varepsilon \left\| \frac{\alpha x}{2\|x\|} \right\|$, ce qui s'écrit, en utilisant la linéarité, $\|u \cdot x\| \leq \varepsilon \|x\|$. On en déduit $\|u\| \leq \varepsilon$ pour tout $\varepsilon > 0$, d'où $\|u\| = 0$ et donc $u = 0$. ■

Définition 2.2. *Soit U un ouvert de E et $f : U \rightarrow F$ une application. L'application f est dite différentiable (au sens de Fréchet) en $a \in U$ s'il existe une application linéaire continue $u : E \rightarrow F$ telle que*

$$\lim_{x \rightarrow a} \frac{f(x) - f(a) - u(x - a)}{\|x - a\|} = 0.$$

L'application f est dite différentiable dans U si elle l'est en chaque point de U .

Il est utile de remarquer que, contrairement à f qui n'est supposée définie que sur l'ouvert U , l'application u , comme toute application linéaire, est définie sur l'espace vectoriel E tout entier. On rappelle d'ailleurs que dans un espace normé le sous-espace vectoriel engendré par un ouvert non vide est toujours l'espace entier.

Le lemme 2.1 justifie de donner un nom à l'application linéaire continue u .

Définition 2.3. *Si f est différentiable en a , l'application linéaire continue u est appelée différentielle de f en a . On la note $df(a)$.*

On peut alors expliciter la définition sous la forme : si f est différentiable en a , pour tout $\varepsilon > 0$, il existe $r > 0$ tel que $\|h\|_E < r$ implique $a + h \in U$ et

$$\|f(a + h) - f(a) - df(a) \cdot h\|_F < \varepsilon \|h\|_E.$$

Remarque 2.4. On notera $\|\cdot\|$ la norme de E et la norme de F , s'il n'y a pas de confusion possible. L'image par une application linéaire $u : E \rightarrow F$ d'un vecteur $h \in E$ sera souvent notée $u \cdot h$, parfois $u(h)$ ou encore $u[h]$.

Proposition 2.5. Si $f : U \rightarrow F$ est différentiable en $a \in U$, alors pour tout $h \in E$,

$$\lim_{t \rightarrow 0} \frac{f(a + th) - f(a)}{t} = df(a) \cdot h$$

Preuve : On écrit

$$\frac{f(a + th) - f(a)}{t} = \frac{f(a + th) - f(a) - df(a)(th)}{t} + df(a) \cdot h$$

et on utilise la définition en remarquant que si $t \rightarrow 0$ alors $th \rightarrow 0$. ■

Il est essentiel de signaler tout de suite que la réciproque n'est pas vraie. L'existence d'une limite pour chaque $h \in E$ permet de deviner $df(a)$ mais ne dispense de montrer ni la linéarité ni la continuité.

Remarquons cependant que cette proposition permet de s'assurer que dans le cas particulier où $E = F = \mathbb{R}$, f est différentiable en a si et seulement si elle est dérivable en a . Si c'est le cas l'application linéaire $df(a) : \mathbb{R} \rightarrow \mathbb{R}$ est $df(a) \cdot t = f'(a)t$.

Proposition 2.6. Si f est différentiable en a , elle est continue en a .

Rappel. (Inégalités triangulaires) Pour tout $x, y \in E$,

- a) $\|x + y\| \leq \|x\| + \|y\|$;
- b) $|\|x\| - \|y\|| \leq \|x - y\|$;
- c) $\|x\| - \|y\| \leq \|x - y\|$ et $\|y\| - \|x\| \leq \|x - y\|$.

Preuve En choisissant $\varepsilon = 1$, on voit qu'il existe $\alpha > 0$ tel que si $\|x - a\| < \alpha$, alors $x \in U$ et $\|f(x) - f(a) - df(a)(x - a)\| < \|x - a\|$. En utilisant l'inégalité triangulaire, on obtient que

$$\|f(x) - f(a)\| < \|df(a)(x - a)\| + \|x - a\| \leq (1 + \|df(a)\|)\|x - a\|,$$

ce qui montre que f est « localement lipschitzienne » en a , donc continue. ■

La démonstration de la proposition suivante est laissée en exercice.

Proposition 2.7. Si on remplace les normes choisies sur E et F par des normes équivalentes, la différentiabilité d'une application en un point et la valeur de la différentielle sont inchangées.

Rappelons que sur un espace vectoriel de dimension finie toutes les normes sont équivalentes et que cette propriété est caractéristique des espaces de dimension finie.

La proposition suivante décrit le cas où l'espace F est un produit d'espaces normés, $F = F_1 \times F_2 \times \cdots \times F_n$, muni de la norme produit :

$$\|(y_1, \dots, y_n)\| = \max_{j=1 \dots n} \|y_j\|$$

ou de toute norme équivalente. Définir une application f de U dans F c'est se donner n applications $f_j : U \rightarrow F_j$, de sorte que pour tout $x \in U$,

$$f(x) = (f_1(x), \dots, f_n(x)).$$

Les applications f_j sont appelées *composantes* de f . On vérifie aisément que lorsque $U = E$, f est linéaire si et seulement si chaque composante f_j l'est. On sait que la continuité de f est équivalente à celle de chacune des composantes. La proposition suivante montre que la même propriété est vraie en ce qui concerne la différentiabilité.

Proposition 2.8. *L'application f est différentiable en $a \in U$ si et seulement si chaque composante f_j l'est. Dans ce cas*

$$df(a) = (df_1(a), \dots, df_n(a))$$

où $df_j(a) \in L(E, F_j)$ est la différentielle de f_j en a .

Preuve Si f est différentiable en a , $df(a) : E \rightarrow F_1 \times \dots \times F_n$ a n composantes : $df(a) = (u_1, \dots, u_n)$ où $u_j : E \rightarrow F_j$ est linéaire continue. De plus, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, si $\|h\| < \alpha$, $a + h \in U$ et

$$\|f(a+h) - f(a) - df(a) \cdot h\| < \varepsilon \|h\|.$$

En explicitant la norme de F et en utilisant les composantes de f et de $df(a)$ ceci s'écrit :

$$\max_{j=1 \dots n} \|f_j(a+h) - f_j(a) - u_j \cdot h\| < \varepsilon \|h\|.$$

Cela prouve que, pour chaque j , f_j est différentiable en a et que $df_j(a) = u_j$.

Réciproquement, supposons chaque f_j différentiable en a . L'application $u = (df_1(a), \dots, df_n(a))$ est linéaire et continue d'après ce qui est rappelé. De plus en écrivant simultanément pour les n composantes la différentiabilité en a , on voit que pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, si $\|h\| < \alpha$, $a + h \in U$ et pour $j = 1 \dots n$, $\|f_j(a+h) - f_j(a) - df_j(a) \cdot h\| < \varepsilon \|h\|$. Alors

$$\|f(a+h) - f(a) - u \cdot h\| = \max_{j=1 \dots n} \|f_j(a+h) - f_j(a) - df_j(a) \cdot h\| < \varepsilon \|h\|$$

ce qui montre que f est différentiable en a et que $df(a) = u$. ■

Ceci s'applique en particulier au cas où F est de *dimension finie*. En appelant n cette dimension, le choix d'une base $\mathcal{W} = (w_1, \dots, w_n)$ de F permet d'identifier F à \mathbb{R}^n . Les composantes de f sont alors appelées *fonctions coordonnées*. Leurs différentielles (lorsqu'elles existent) sont des *formes* linéaires continues sur E .

L'utilisation de la base \mathcal{W} conduit aux écritures suivantes. Pour tout $x \in U$,

$$f(x) = \sum_{j=1}^n f_j(x)w_j$$

et pour tout $h \in E$,

$$df(a) \cdot h = \sum_{j=1}^n (df_j(a) \cdot h)w_j$$

où $df_j(a) \cdot h$ est un scalaire.

Si de plus $E = \mathbb{R}$, $f : U \subset \mathbb{R} \rightarrow \mathbb{R}^n$ est donnée par $f(x) = (f_1(x), \dots, f_n(x))$. Elle est différentiable en a si et seulement si chaque fonction coordonnée est dérivable en a . Dans ce cas $df(a) \cdot t = t(f'_1(a), \dots, f'_n(a))$ est le produit du scalaire t par le vecteur de \mathbb{R}^n dont les composantes sont les dérivées en a des composantes de f . On note $f'(a)$ ce vecteur que l'on appelle *vecteur dérivée* de f en a . On a donc $df(a) \cdot t = tf'(a)$ et en particulier $f'(a) = df(a) \cdot 1$.

Remarquons que, plus généralement, lorsque $E = \mathbb{R}$, si F est un espace normé quelconque, $f : U \subset \mathbb{R} \rightarrow F$ est différentiable en a si et seulement si $\lim_{t \rightarrow 0} \frac{f(a+t) - f(a)}{t}$ existe. C'est alors un vecteur de F que l'on note $f'(a)$ et qu'on appelle vecteur dérivée de f en a . La différentielle de f en a est alors l'application linéaire (automatiquement continue) $df(a) : \mathbb{R} \rightarrow F$ définie par $df(a) \cdot t = tf'(a)$. En particulier $df(a) \cdot 1 = f'(a)$. Autrement dit lorsque $E = \mathbb{R}$ (et dans ce cas seulement) la proposition 2.5 a une réciproque.

2.2 Exemples

2.2.1 Application affine

Une application affine $A : E \rightarrow F$ est de la forme $A(x) = b + u(x)$ où $b \in F$ et u est une application linéaire de E dans F . Pour $a, x \in E$,

$$(1) \quad A(x) - A(a) = u(x) - u(a) = u(x - a)$$

par linéarité. Donc A est continue en a si et seulement si u est continue en 0, donc, puisque u est linéaire, en tout point. Dans ce cas la formule (1) montre que A est différentiable en tout point de E et que, pour tout $a \in E$,

$$dA(a) = u.$$

La différentielle d'une application affine continue est donc *indépendante* de a . En particulier si u est linéaire continue, pour tout $x \in E$, $du(x) = u$ et donc $du(x)[h] = u[h]$.

2.2.2 Un exemple en dimension finie

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'application définie par $f(x, y) = (x + y, xy)$. Soit $(a, b) \in \mathbb{R}^2$. Pour $(h, k) \in \mathbb{R}^2$, on a

$$f(a + h, b + k) - f(a, b) = (h + k, hb + ka + hk) = u(h, k) + \varphi(h, k)$$

où $u(h, k) = (h + k, hb + ka)$ et $\varphi(h, k) = (0, hk)$.

L'application u est linéaire continue (on est en dimension finie). Sa matrice dans les bases canoniques est $Jf(a, b) = \begin{pmatrix} 1 & 1 \\ b & a \end{pmatrix}$.

Choisissons sur \mathbb{R}^2 la norme $\|(h, k)\| = \max(|h|, |k|)$. La fonction $\varphi(h, k)$ vérifie $\|\varphi(h, k)\| = |hk| \leq \|(h, k)\|^2$. Donc $\lim_{(h,k) \rightarrow 0} \frac{\varphi(h, k)}{\|(h, k)\|} = 0$, ce qui montre que l'application f est différentiable en (a, b) , sa différentielle en (a, b) étant l'application linéaire de \mathbb{R}^2 dans \mathbb{R}^2 de matrice $Jf(a, b)$.

2.2.3 Un exemple en dimension infinie

Appelons E l'espace vectoriel des fonctions continûment dérivables sur $[0, 1]$ à valeurs dans \mathbb{R} , muni de la norme $\|f\| = \|f\|_\infty + \|f'\|_\infty$. Soit F l'espace vectoriel des fonctions continues sur $[0, 1]$ à valeurs dans \mathbb{R} , muni de la norme $\|\cdot\|_\infty$. On considère l'application $H : E \rightarrow F$ donnée par $H(f) = f' + f^2$. Fixons $f \in E$. On a pour $h \in E$,

$$H(f + h) - H(f) = h' + 2fh + h^2$$

L'application définie sur E par $u(h) = h' + 2fh$ est visiblement linéaire, mais cette fois il faut montrer sa continuité. On a

$$\|u(h)\|_\infty = \|h' + 2fh\|_\infty \leq \|h'\|_\infty + 2\|f\|_\infty \|h\|_\infty \leq (1 + 2\|f\|_\infty) \|h\|_\infty$$

ce qui établit le résultat.

D'autre part $H(f + h) - H(f) - u(h) = h^2$ et $\|h^2\|_\infty \leq \|h\|_\infty^2 \leq \|h\|^2$, de sorte que

$$\lim_{h \rightarrow 0} \frac{H(f + h) - H(f) - u(h)}{\|h\|} = 0.$$

On a donc prouvé la différentiabilité de H en f , la différentielle étant l'application linéaire

continue de E dans F définie par $dH(f) \cdot h = h' + 2fh$.

Remarquons que si on choisit sur E la norme $\|\cdot\|_\infty$ au lieu de $\|\cdot\|$, l'application u n'est pas continue et H n'est pas différentiable. Il est en effet impossible de déterminer une constante $C > 0$ telle que $\|h' + 2fh\|_\infty \leq C\|h\|_\infty$ pour tout $h \in E$, puisqu'en choisissant $h_n = x^n$, on devrait avoir pour tout n , $C \geq \frac{\|h'_n + 2fh_n\|_\infty}{\|h_n\|_\infty} = \sup_{x \in [0,1]} |nx^{n-1} + 2x^n f(x)| \geq |n + 2f(1)|$ et que cette quantité tend vers l'infini avec n .

2.3 Différentielles partielles

2.3.1 Différentiabilité dans une « direction »

Soit $f : U \rightarrow F$ et $a \in U$. Soit $v \in E$, $v \neq 0$, la fonction d'une variable $\varphi_v : t \mapsto f(a+tv)$ est définie au voisinage de 0 (U est ouvert). Elle est différentiable en 0 si et seulement si

$$\varphi'_v(0) = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}$$

existe. On dit alors que f admet une *dérivée partielle* en a dans la direction v (appelé aussi dérivée directionnelle), égale à $\varphi'_v(0)$. C'est un vecteur de F . Lorsque la limite existe, on note

$$\partial_v f(a) = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}.$$

Définition 2.9. On dit que $f : U \rightarrow F$ est *Gâteaux différentiable* en $a \in U$ (ou *G-différentiable* en a), s'il existe une application linéaire continue $u : E \rightarrow F$ telle que $\forall v \in E$ la dérivée directionnelle $\partial_v f(a)$ existe et est égale à $u(v)$, i.e.

$$\partial_v f(a) = \lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t} = u(v), \quad \forall v \in E.$$

On vérifie alors que, si une telle application u existe elle est unique. On note

$$u = d_G f(a)$$

D'après la Proposition 2.5, c'est le cas pour tout $v \neq 0$ si f est différentiable en a et $\varphi'_v(0) = d_G f(a) \cdot v = df(a) \cdot v$.

Il n'y a pas de réciproque à cette propriété sauf si $E = \mathbb{R}$, comme le montre l'exemple suivant en dimension 2.

Prenons $U = \mathbb{R}^2$, $F = \mathbb{R}$ et f définie par $f(x, y) = \frac{x^2 y}{x^2 + y^2}$ si $(x, y) \neq (0, 0)$ et $f(0, 0) = 0$. Étudions f au point $(0, 0)$.

En remarquant que pour tout $t \in \mathbb{R}$ et tout $v = (x, y) \in \mathbb{R}^2$, on a $f(tv) = tf(v)$, on en déduit, pour $t \neq 0$, $\frac{f(tv) - f(0)}{t} = f(v)$ et donc f admet $f(v)$ pour dérivée partielle en $(0, 0)$ dans la direction v , pour tout $v \neq 0$. Pourtant f n'est pas différentiable en $(0, 0)$. Si elle l'était, le calcul précédent montre (Proposition 2.5) que sa différentielle serait f . Mais f n'est pas linéaire et ne saurait donc être une différentielle.

Plus généralement, si $E_1 \neq 0$ est un sous-espace vectoriel de E , il existe un voisinage ouvert U_1 de 0 dans E_1 tel que si $h_1 \in U_1$, $a + h_1 \in U$: ceci parce que U est ouvert et l'application $E_1 \rightarrow E$ qui à h_1 associe $a + h_1$ continue. On peut alors définir l'application partielle $\varphi_1 : U_1 \rightarrow F$ donnée par $\varphi_1(h_1) = f(a + h_1)$ et se poser le problème de sa différentiabilité en 0.

Définition 2.10. Si φ_1 est différentiable en 0, on dit que f admet une différentielle partielle en a dans la direction E_1 , égale à $d\varphi_1(0)$.

Bien sûr $d\varphi_1(0) \in L(E_1, F)$.

Proposition 2.11. Si $f : U \rightarrow F$ est différentiable en $a \in U$, elle admet en a dans toute direction $E_1 \neq 0$ de E une différentielle partielle égale à la restriction à E_1 de la différentielle de f en a .

Preuve On sait que quel que soit $\varepsilon > 0$, il existe $\delta > 0$ tel que si $h \in E$ vérifie $\|h\| < \delta$, alors $\|f(a+h) - f(a) - df(a) \cdot h\| < \varepsilon \|h\|$. C'est vrai a fortiori si $h_1 \in E_1$ vérifie $\|h_1\| < \delta$ et cela s'écrit alors $\|\varphi_1(h_1) - \varphi_1(0) - df(a)|_{E_1} \cdot h_1\| < \varepsilon \|h_1\|$ et le résultat suit en remarquant que $df(a)|_{E_1} \in L(E_1, F)$. ■

Bien sûr il n'y a pas de réciproque.

Si $E = E_1 \times E_2$ muni de la norme produit, E_1 et E_2 peuvent être vus comme des sous-espaces vectoriels de E en identifiant E_1 à $E_1 \times 0 = \{(h_1, 0) \mid h_1 \in E_1\}$ (resp. E_2 à $0 \times E_2$) et en remarquant que la norme dans E de $(h_1, 0)$ est égale à la norme dans E_1 de h_1 . Si $f : U \rightarrow F$ admet en $a \in U$ une différentielle partielle dans la direction E_1 (resp. E_2), on la note $d_1f(a)$ (resp. $d_2f(a)$). Si f est différentiable en $a \in U$, la proposition 2.11 permet d'écrire l'utile formule suivante. Si $h = (h_1, h_2) \in E = E_1 \times E_2$, puisque $h = (h_1, 0) + (0, h_2)$,

$$df(a) \cdot h = df(a) \cdot (h_1, 0) + df(a) \cdot (0, h_2) = d_1f(a) \cdot h_1 + d_2f(a) \cdot h_2.$$

2.3.2 Dérivées partielles et matrice jacobienne

Lorsque E est de dimension finie, on l'identifie à \mathbb{R}^m en y choisissant une base $\mathcal{V} = (v_1, \dots, v_m)$. Alors $f : U \rightarrow F$ s'identifie à une fonction de m variables réelles (x_1, \dots, x_m) : les coordonnées dans la base \mathcal{V} . Soit $a = (a_1, \dots, a_m) \in U$, la dérivée partielle de f en a dans la direction v_j (si elle existe) est le vecteur de F qui est la dérivée en a_j de la fonction d'une variable

$$x_j \mapsto f(a_1, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_m).$$

On le note $\frac{\partial f}{\partial x_j}(a)$. On a donc si $h = \sum_{j=1}^m h_j v_j$,

$$df(a) \cdot h = \sum_{j=1}^m h_j \frac{\partial f}{\partial x_j}(a)$$

où les h_j sont des scalaires et les $\frac{\partial f}{\partial x_j}(a)$ des vecteurs de F . On a aussi la formule

$$\frac{\partial f}{\partial x_j}(a) = df(a) \cdot v_j$$

Si de plus F est de dimension finie et identifié à \mathbb{R}^n par le choix d'une base $\mathcal{W} = (w_1, \dots, w_n)$, la fonction f se représente par les n fonctions coordonnées (f_1, \dots, f_n) où $f_i : U \rightarrow \mathbb{R}$ est une fonction de m variables à valeurs réelles. Les f_i sont différentiables en a si et seulement si f l'est et dans ce cas pour tout $h \in E$,

$$df_i(a) \cdot h = \sum_{i=1}^n (df_i(a) \cdot h) w_i$$

où, si $h = \sum_{j=1}^m h_j v_j$, alors $df_i(a) \cdot h = \sum_{j=1}^m h_j \frac{\partial f_i}{\partial x_j}(a)$.

Cette fois $\frac{\partial f_i}{\partial x_j}(a)$ est un réel et on peut aussi écrire $df_i(a) \cdot h = \sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(a) h_j$.

L'application linéaire $df(a)$ est donc celle qui associe au vecteur h de E le vecteur $df(a) \cdot h = \sum_{i=1}^n \left(\sum_{j=1}^m \frac{\partial f_i}{\partial x_j}(a) h_j \right) w_i$ de F . En d'autres termes la matrice de l'application linéaire $df(a)$ dans les bases \mathcal{V} et \mathcal{W} est la matrice à n lignes et m colonnes notée $Jf(a)$ donnée par

$$Jf(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{1 \leq i \leq n, 1 \leq j \leq m}$$

où i est l'indice ligne et j l'indice colonne.

Définition 2.12. La matrice $Jf(a)$ est appelée matrice jacobienne de f en a dans les bases \mathcal{V} et \mathcal{W} .

Une ligne de la matrice jacobienne est constituée des dérivées partielles d'une même composante de f et une colonne par les dérivées partielles par rapport à une même variable. Remarquons que, si $E = E_1 \times E_2$, ce qui correspond, via l'identification déjà faite, à supposer que (v_1, \dots, v_r) est une base de E_1 et (v_{r+1}, \dots, v_m) une base de E_2 , la matrice $Jf(a)$ est de la forme

$$Jf(a) = (J_1 f(a), J_2 f(a))$$

où $J_i f(a)$ est la matrice jacobienne de $d_i f(a)$, ($i = 1, 2$).

Définition 2.13. Si $n = m$, le déterminant de la matrice (carrée) $Jf(a)$ est appelé jacobien de f en a et noté $jf(a)$.

La condition $jf(a) \neq 0$ exprime que $df(a)$ est un isomorphisme de E sur F , ou que $df(a) \in \text{Isom}(E, F)$.

2.4 Différentiabilité et opérations classiques

2.4.1 Combinaison linéaire et composition

Le premier résultat est immédiat.

Proposition 2.14. Soit f_1 et f_2 deux applications définies sur un ouvert U d'un espace normé E à valeurs dans un espace normé F , différentiables en $a \in U$. Alors pour tous scalaires λ et μ , l'application $f = \lambda f_1 + \mu f_2$ est différentiable en a et $df(a) = \lambda df_1(a) + \mu df_2(a)$.

La proposition suivante montre que la notion de différentielle se comporte vis-à-vis de la composition de la façon la plus simple possible.

Proposition 2.15. Soit $f : U \rightarrow F$, $a \in U$, V un ouvert de F contenant $f(a) = b$ et $g : V \rightarrow G$. Si f est différentiable en a et g en b , alors $g \circ f$ est différentiable en a et

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

Preuve On remarque d'abord que la composée de deux applications linéaires continues étant linéaire continue, puisque $df(a) \in L(E, F)$ et $dg(b) \in L(F, G)$, $dg(b) \circ df(a) \in L(E, G)$.

Puisque f est continue en a , il existe $r > 0$ tel que pour $x \in Bo(a, r)$, $f(x) \in V$ donc $g \circ f$ est définie sur cette boule. De plus si r est assez petit et $\|h\| < r$, la preuve de la proposition 2.6 montre que

$$\|f(a+h) - f(a)\| \leq (1 + \|df(a)\|)\|h\|.$$

D'autre part, pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que, pour tout $k \in F$, si $\|k\| < \alpha$,

$$\|g(b+k) - g(b) - dg(b) \cdot k\| \leq \varepsilon\|k\|.$$

Or $g \circ f(a+h) = g(b+k)$ où $k = f(a+h) - b = f(a+h) - f(a)$.

Si $\|h\| < \alpha_1 = \min\left(r, \frac{\alpha}{1 + \|df(a)\|}\right)$, on a donc $\|k\| < \alpha$ et par suite :

$$\|g \circ f(a+h) - g \circ f(a) - dg(b) \circ df(a)(h)\| = \|g(b+k) - g(b) - dg(b) \cdot k + dg(b) \cdot [k - df(a) \cdot h]\| \leq \varepsilon\|f(a+h) - f(a)\| + \|dg(b)\| \|f(a+h) - f(a) - df(a) \cdot h\|.$$

Mais il existe $\eta > 0$ tel que $\|h\| < \eta$ implique $\|f(a+h) - f(a) - df(a) \cdot h\| < \varepsilon\|h\|$. Si donc $\|h\| < \min(\alpha_1, \eta)$, on a

$$\|g \circ f(a+h) - g \circ f(a) - dg(b) \circ df(a)(h)\| < \varepsilon(1 + \|df(a)\| + \|dg(b)\|)\|h\|.$$

C'est le résultat annoncé. ■

Dans le cas où les espaces E et F sont tous deux de dimension finie, cette proposition s'explique en termes de matrices jacobiniennes :

$$J(g \circ f)(a) = Jg(b) Jf(a)$$

Il s'agit bien sûr du produit des deux matrices.

2.4.2 Applications bilinéaires et produits

Proposition 2.16. *Soit E, F et G des espaces normés et $B : E \times F \rightarrow G$ une application bilinéaire. Alors B est continue si et seulement s'il existe $C > 0$ tel que pour tout $x \in E$ et tout $y \in F$, $\|B(x, y)\| \leq C\|x\| \|y\|$. Dans ce cas B est différentiable dans $E \times F$ et pour tout $(x, y) \in E \times F$, $dB(x, y)$ est l'application linéaire continue de $E \times F$ dans G donnée par*

$$(h, k) \mapsto dB(x, y) \cdot (h, k) = B(x, k) + B(h, y).$$

Preuve L'affirmation concernant la continuité se montre comme son analogue pour les applications linéaires. Si B est continue, elle l'est en $(0, 0)$ et il existe $\alpha > 0$ tel que si $\|(x, y)\| = \max(\|x\|, \|y\|) < \alpha$, alors $\|B(x, y)\| \leq 1$. Si $x \in E$, $x \neq 0$ et $y \in F$, $y \neq 0$, $(u, v) = \left(\frac{\alpha x}{2\|x\|}, \frac{\alpha y}{2\|y\|}\right)$ vérifie $\|(u, v)\| = \frac{\alpha}{2} < \alpha$. D'où $\|B(u, v)\| = \left\|B\left(\frac{\alpha x}{2\|x\|}, \frac{\alpha y}{2\|y\|}\right)\right\| \leq 1$ ou encore, en utilisant la bilinéarité et la propriété convenable d'une norme,

$$\|B(x, y)\| \leq \frac{4}{\alpha^2} \|x\| \|y\|$$

ce qui permet de choisir $C = \frac{4}{\alpha^2}$.

Montrons maintenant que cette propriété implique la différentiabilité de B . Cela établira a fortiori sa continuité. (Une preuve directe est un excellent exercice...) Tout d'abord

une vérification facile montre que l'application $u : (h, k) \mapsto B(x, k) + B(h, y)$ est linéaire. La majoration

$$\|u \cdot (h, k)\| \leq C(\|x\| \|k\| + \|h\| \|y\|) \leq C(\|x\| + \|y\|)\|(h, k)\|$$

permet d'affirmer la continuité de u (et aussi $\|u\| \leq C(\|x\| + \|y\|)$).

Ensuite si $(h, k) \in E \times F$, la bilinéarité de B permet d'écrire

$$B(x + h, y + k) - B(x, y) - u \cdot (h, k) = B(h, k)$$

et la majoration $\|B(h, k)\| \leq C\|h\| \|k\| \leq C\|(h, k)\|^2$ montre que $\frac{B(h, k)}{\|(h, k)\|} \rightarrow 0$ quand $\|(h, k)\| \rightarrow 0$. ■

Exemples

1. Soit E un espace normé et $L(E)$ l'espace vectoriel normé des applications linéaires continues de E dans E . L'application $B : L(E) \times L(E) \rightarrow L(E)$ définie par $(u, v) \mapsto u \circ v$ est bilinéaire continue car $\|u \circ v\| \leq \|u\| \|v\|$. Elle est donc différentiable dans $L(E) \times L(E)$ et sa différentielle en (u, v) est donnée par

$$(h, k) \mapsto dB(u, v) \cdot (h, k) = u \circ k + h \circ v.$$

2. Si E est muni d'un produit scalaire $E \times E \rightarrow \mathbb{R}$ noté $\langle x, y \rangle$ et si la norme sur E est la norme associée : $\|x\| = \sqrt{\langle x, x \rangle}$, la forme bilinéaire $B : E \times E \rightarrow \mathbb{R}$, $B(x, y) = \langle x, y \rangle$ est continue donc différentiable en tout point. Lorsque $E = \mathbb{R}^m$, la norme $\|\cdot\|_2$ est obtenue de cette façon à partir du produit scalaire usuel.

En combinant les propositions 2.16 et 2.15 on obtient nombre de résultats classiques. En voici quelques uns.

1. Produit de deux fonctions numériques

Proposition 2.17. *Si f et $g : U \rightarrow \mathbb{R}$ sont différentiables en $a \in U$, leur produit $\varphi = fg$, défini par $\varphi(x) = f(x)g(x)$, est différentiable en a . La différentielle est la forme linéaire continue donnée, pour $h \in E$, par*

$$d\varphi(a) \cdot h = g(a) df(a) \cdot h + f(a) dg(a) \cdot h$$

Preuve On a $\varphi = B \circ \Phi$ où $\Phi : U \rightarrow \mathbb{R}^2$ a pour composantes f et g et B est la forme bilinéaire continue $\mathbb{R}^2 \rightarrow \mathbb{R}$ donnée par $B(s, t) = st$. La proposition 2.8 montre que Φ est différentiable en a et que $D\Phi(a) = (df(a), dg(a))$. D'autre part la proposition 2.15 donne pour tout $h \in E$, $d\varphi(a) \cdot h = dB(\Phi(a)) \circ D\Phi(a)(h) = dB(f(a), g(a)) \cdot (df(a) \cdot h, dg(a) \cdot h) = B(f(a), dg(a) \cdot h) + B(df(a) \cdot h, g(a)) = f(a) dg(a) \cdot h + g(a) df(a) \cdot h$. ■

2. Produit d'une application par une fonction numérique

Proposition 2.18. *Si $f : U \rightarrow F$ et $\varphi : U \rightarrow \mathbb{R}$ sont différentiables en a , alors l'application $g = \varphi f$, c'est-à-dire $g(x) = \varphi(x)f(x)$ est différentiable en a et pour tout $h \in E$,*

$$dg(a) \cdot h = \varphi(a) df(a) \cdot h + (d\varphi(a) \cdot h) f(a).$$

(Dans cette formule, $\varphi(a)$ et $d\varphi(a) \cdot h$ sont des scalaires tandis que $f(a)$ et $df(a) \cdot h$ sont des vecteurs de F .)

Preuve La preuve est identique en écrivant $g = B \circ \Phi$ où $\Phi : U \rightarrow \mathbb{R} \times F$ a pour composantes (φ, f) et B est l'application bilinéaire de $\mathbb{R} \times F \rightarrow F$ donnée par $B(\lambda, y) = \lambda y$. La continuité de B est conséquence de $\|\lambda y\| = |\lambda| \|y\|$. ■

3. Produit scalaire

Proposition 2.19. *Supposons l'espace F muni d'une norme associée à un produit scalaire. Soit $f, g : U \rightarrow F$ des fonctions différentiables en $a \in U$. L'application $\varphi : U \rightarrow \mathbb{R}$ définie par $\varphi(x) = \langle f(x), g(x) \rangle$ est différentiable en a et sa différentielle est la forme linéaire continue sur E donnée par*

$$d\varphi(a) \cdot h = \langle df(a) \cdot h, g(a) \rangle + \langle f(a), dg(a) \cdot h \rangle$$

Preuve On écrit $\varphi = B \circ \Phi$ où $\Phi = (f, g) : U \rightarrow F \times F$ et B est le produit scalaire. ■

Remarque Si E est de dimension finie et $F = \mathbb{R}$, la formule

$$df(a) \cdot h = \sum_{j=1}^m \frac{\partial f}{\partial x_j}(a) h_j$$

peut se lire

$$df(a) \cdot h = \langle \nabla f(a), h \rangle$$

où $\nabla f(a)$ (appelé *gradient* de f en a) est le vecteur de E

$$\nabla f(a) = \sum_{j=1}^m \frac{\partial f}{\partial x_j}(a) v_j$$

si $\mathcal{V} = (v_1, \dots, v_m)$ est la base choisie dans E .

2.4.3 Inverse

Ce mot peut désigner

- la fonction $\frac{1}{f(x)}$ lorsque la fonction $f : U \rightarrow \mathbb{R}$ ne s'annule pas,
- l'application f^{-1} lorsque f est une bijection.

Il ne faut pas confondre ces deux situations... et il y a un théorème pour chacune d'elles.

Proposition 2.20. *Soit $f : U \rightarrow \mathbb{R}$ une fonction différentiable en $a \in U$ telle que $f(a) \neq 0$. Alors la fonction $g(x) = \frac{1}{f(x)}$ est définie au voisinage de a , différentiable en a et, pour tout $h \in E$,*

$$dg(a) \cdot h = -\frac{1}{f(a)^2} df(a) \cdot h$$

Preuve Puisque $f(a) \neq 0$ et que f est continue en a (puisque différentiable), il existe un voisinage V de a tel que, pour tout $x \in V$, $f(x) \neq 0$. L'application g est définie sur V et s'écrit $g = \iota \circ f$ où $\iota : \mathbb{R}^* \rightarrow \mathbb{R}^*$ est donnée par $\iota(y) = \frac{1}{y}$. C'est une fonction dérivable en

tout point et $\iota'(y) = -\frac{1}{y^2}$. Autrement dit la différentielle de ι en y est l'application linéaire

$d\iota(y) : \mathbb{R} \rightarrow \mathbb{R}$ définie par $d\iota(y) \cdot t = -\frac{t}{y^2}$ et tout découle de la proposition 2.15. ■

Proposition 2.21. *Soit U un ouvert de E , V un ouvert de F et $f : U \rightarrow V$ un homéomorphisme. On suppose que f est différentiable en a et que $df(a) \in \text{Isom}(E, F)$, c'est-à-dire qu'elle est linéaire, continue, bijective et que son inverse est continue. Alors $f^{-1} : V \rightarrow E$ est différentiable en $b = f(a)$ et*

$$df^{-1}(b) = df(a)^{-1}.$$

Remarquons que la condition $df(a) \in \text{Isom}(E, F)$ suppose déjà qu'il y a des isomorphismes entre E et F et ceci est exigeant : en dimension finie cela suppose que F et E ont la même dimension, en dimension infinie que F (resp. E) soit un Banach dès que E (resp. F) l'est.

Preuve Si f^{-1} est différentiable en b , la proposition 2.15 donne pour sa différentielle la formule indiquée. En effet la proposition 2.15 permet de déduire de l'égalité $f \circ f^{-1} = id$ (sur V) et de $f^{-1}(b) = a$ l'égalité $df(a) \circ df^{-1}(b) = id$ (sur F). De même les égalités $f^{-1} \circ f = id$ (sur U) et $f(a) = b$ donnent $df^{-1}(b) \circ df(a) = id$ (sur E). Il faut donc établir qu'en posant, pour $k \in F$ de norme assez petite,

$$\varphi(k) = f^{-1}(b+k) - f^{-1}(b) - df(a)^{-1} \cdot k$$

on a $\lim_{k \rightarrow 0} \frac{\varphi(k)}{\|k\|} = 0$.

Comme f est différentiable en a , si $\varepsilon > 0$ est fixé, il existe $\alpha > 0$ tel que si $h \in E$ vérifie $\|h\| < \alpha$, alors $a+h \in U$ et

$$(*) \quad \|f(a+h) - f(a) - df(a) \cdot h\| < \varepsilon \|h\|$$

Comme f^{-1} est continue en b , il existe $\eta > 0$ tel que $\|k\| < \eta$ implique $b+k \in V$ et $\|f^{-1}(b+k) - f^{-1}(b)\| < \alpha$ et donc, en posant $h = f^{-1}(b+k) - f^{-1}(b)$ de sorte que $a+h = f^{-1}(b+k)$, $(*)$ s'écrit $\|b+k - b - df(a) \cdot h\| < \varepsilon \|h\|$. D'où

$$\|df(a) \cdot h\| \leq \|k\| + \varepsilon \|h\|.$$

Mais $h = df(a)^{-1} \circ df(a) \cdot h$ et donc

$$\|h\| \leq \|df(a)^{-1}\| \|df(a) \cdot h\| \leq \|df(a)^{-1}\| (\|k\| + \varepsilon \|h\|)$$

ou encore

$$(1 - \varepsilon \|df(a)^{-1}\|) \|h\| \leq \|df(a)^{-1}\| \|k\|.$$

Si on prend $\varepsilon < \frac{1}{2\|df(a)^{-1}\|}$, on a donc

$$(**) \quad \|h\| \leq 2\|df(a)^{-1}\| \|k\|.$$

Mé زالور

$$\begin{aligned} \varphi(k) &= h - df(a)^{-1} \cdot k = h - df(a)^{-1} \cdot [(f(a+h) - f(a))] \\ &= -df(a)^{-1} \cdot [f(a+h) - f(a) - df(a) \cdot h] \end{aligned}$$

et donc en utilisant $(*)$ et $(**)$:

$$\|\varphi(k)\| \leq \|df(a)^{-1}\| \varepsilon \|h\| \leq 2\|df(a)^{-1}\|^2 \varepsilon \|k\| \quad \blacksquare$$

2.4.4 Applications multilinéaires

On peut généraliser la proposition 2.16 aux applications multilinéaires de la manière suivante.

Proposition 2.22. *Soit E_1, E_2, \dots, E_n, F des espaces normés et $\Phi : E_1 \times E_2 \times \dots \times E_n \rightarrow F$ une application multilinéaire (où n -linéaire). Alors Φ est continue si et seulement s'il existe $C > 0$ tel que pour tout $x = (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ on a que*

$$\|\Phi(x)\| \leq C \|x_1\| \cdots \|x_n\|.$$

Dans ce cas Φ est différentiable et pour tout $x = (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$ sa différentielle est donnée par :

$$d\Phi(x)(h) = \sum_{j=1}^n \Phi(x_1, \dots, x_{j-1}, h_j, x_{j+1}, \dots, x_n), \quad \text{pour tout } h = (h_1, \dots, h_n) \in E_1 \times \dots \times E_n.$$

Chapitre 3

Théorème des accroissements finis et applications de classe \mathcal{C}^1

3.1 Théorème des accroissements finis

C'est un résultat qui concerne les fonctions d'une seule variable, même s'il a des conséquences générales. Pour les fonctions d'une variable on emploie de façon équivalente les mots dérivable et différentiable. Rappelons que si I est un ouvert de \mathbb{R} et $f : I \rightarrow F$ une fonction dérivable en $t \in I$, on note $f'(t)$ le vecteur de F qui est la dérivée de f en t .

Définition 3.1. Soit $[\alpha, \beta]$ un segment de \mathbb{R} et $f : [\alpha, \beta] \rightarrow F$. On dit que f est dérivable sur $[\alpha, \beta]$ si elle l'est dans $] \alpha, \beta[$ et si $\lim_{t \rightarrow \alpha^+} \frac{f(t) - f(\alpha)}{t - \alpha}$ et $\lim_{t \rightarrow \beta^-} \frac{f(t) - f(\beta)}{t - \beta}$ existent. On note (abusivement) $f'(\alpha)$ et $f'(\beta)$ les vecteurs de F ainsi définis.

Proposition 3.2 (Théorème des accroissements finis). Soit $[\alpha, \beta]$ un segment de \mathbb{R} et $f : [\alpha, \beta] \rightarrow F$ une application dérivable sur ce segment. Soit $\varphi : [\alpha, \beta] \rightarrow \mathbb{R}$ une fonction dérivable sur ce même segment. On suppose que, pour tout $t \in [\alpha, \beta]$, $\|f'(t)\| < \varphi'(t)$. Alors

$$\|f(\beta) - f(\alpha)\| \leq \varphi(\beta) - \varphi(\alpha).$$

Preuve Commençons par remarquer que la fonction φ est strictement croissante sur l'intervalle $[\alpha, \beta]$ puisque sa dérivée est strictement positive, ceci permet d'être rassuré malgré l'absence de valeur absolue dans le membre de droite. Supposons $\|f(\beta) - f(\alpha)\| > \varphi(\beta) - \varphi(\alpha)$. Soit $m = \frac{\alpha + \beta}{2}$ le milieu de $[\alpha, \beta]$. L'une des deux inégalités $\|f(\beta) - f(m)\| > \varphi(\beta) - \varphi(m)$ ou $\|f(m) - f(\alpha)\| > \varphi(m) - \varphi(\alpha)$ a lieu (utiliser l'inégalité triangulaire). Soit $[\alpha_1, \beta_1]$ l'intervalle $[\alpha, m]$ ou $[m, \beta]$ sur lequel l'inégalité est vérifiée. On a donc $\|f(\beta_1) - f(\alpha_1)\| > \varphi(\beta_1) - \varphi(\alpha_1)$, $[\alpha_1, \beta_1] \subset [\alpha, \beta]$ et $\beta_1 - \alpha_1 = \frac{1}{2}(\beta - \alpha)$. On recommence avec le segment $[\alpha_1, \beta_1]$ au lieu de $[\alpha, \beta]$. Par récurrence on construit une suite de segments emboîtés $[\alpha_n, \beta_n]$ tels que $\beta_n - \alpha_n = \frac{1}{2^n}(\beta - \alpha)$ et $\|f(\beta_n) - f(\alpha_n)\| > \varphi(\beta_n) - \varphi(\alpha_n)$. Les suites (α_n) et (β_n) sont adjacentes et convergent vers une même limite $\gamma \in [\alpha, \beta]$. Par hypothèse $\|f'(\gamma)\| < \varphi'(\gamma)$ et il est possible de choisir $\varepsilon > 0$ tel que $\|f'(\gamma)\| + 2\varepsilon < \varphi'(\gamma)$. En exprimant la dérivabilité de f et de φ en γ on peut trouver $\eta > 0$ tel que, si $t \in [\alpha, \beta]$ vérifie $|t - \gamma| < \eta$, alors

$$\|f(t) - f(\gamma) - (t - \gamma)f'(\gamma)\| < \varepsilon |t - \gamma|$$

et

$$|\varphi(t) - \varphi(\gamma) - (t - \gamma)\varphi'(\gamma)| < \varepsilon |t - \gamma|.$$

Puisque les suites (α_n) et (β_n) convergent vers γ , pour n convenable on a $|\alpha_n - \gamma| < \eta$ et $|\beta_n - \gamma| < \eta$. Alors :

$$\|f(\alpha_n) - f(\gamma) - (\alpha_n - \gamma)f'(\gamma)\| < \varepsilon(\gamma - \alpha_n)$$

$$\|f(\beta_n) - f(\gamma) - (\beta_n - \gamma)f'(\gamma)\| < \varepsilon(\beta_n - \gamma)$$

$$|\varphi(\alpha_n) - \varphi(\gamma) - (\alpha_n - \gamma)\varphi'(\gamma)| < \varepsilon(\gamma - \alpha_n)$$

$$|\varphi(\beta_n) - \varphi(\gamma) - (\beta_n - \gamma)\varphi'(\gamma)| < \varepsilon(\beta_n - \gamma).$$

On en déduit (inégalité triangulaire) :

$$\|f(\beta_n) - f(\alpha_n)\| \leq \varepsilon(\beta_n - \alpha_n) + \|(\beta_n - \alpha_n)f'(\gamma)\| = (\beta_n - \alpha_n)(\|f'(\gamma)\| + \varepsilon)$$

Vu le choix de ε , on a aussi

$$\|f(\beta_n) - f(\alpha_n)\| < (\beta_n - \alpha_n)(\varphi'(\gamma) - \varepsilon).$$

D'autre part

$$(\gamma - \alpha_n)\varphi'(\gamma) < \varphi(\gamma) - \varphi(\alpha_n) + \varepsilon(\gamma - \alpha_n)$$

et

$$(\beta_n - \gamma)\varphi'(\gamma) < \varphi(\beta_n) - \varphi(\gamma) + \varepsilon(\beta_n - \gamma)$$

d'où

$$(\beta_n - \alpha_n)\varphi'(\gamma) < \varphi(\beta_n) - \varphi(\alpha_n) + \varepsilon(\beta_n - \alpha_n)$$

En combinant avec l'inégalité précédente on obtient

$$\|f(\beta_n) - f(\alpha_n)\| < \varphi(\beta_n) - \varphi(\alpha_n)$$

ce qui contredit la définition du segment $[\alpha_n, \beta_n]$. ■

Si E est un espace vectoriel et $a, b \in E$, on définit le *segment* $[a, b]$ comme l'image de l'application affine $\gamma : [0, 1] \rightarrow E$ donnée par $\gamma(t) = (1 - t)a + tb = a + t(b - a)$.

Corollaire 3.3. *Soit U un ouvert de E et $f : U \rightarrow F$ une application différentiable. Si $a, b \in U$ sont tels que $[a, b] \subset U$ et s'il existe $M > 0$ tel que, pour tout $t \in [0, 1]$, $\|df(\gamma(t))\| \leq M$, alors $\|f(b) - f(a)\| \leq M \|b - a\|$.*

Preuve Soit $g = f \circ \gamma : [0, 1] \rightarrow F$, c'est-à-dire $g(t) = f(a + t(b - a))$. Pour tout $t \in [0, 1]$ on a

$g'(t) = d(f \circ \gamma)(t) \cdot 1 = df(\gamma(t)) \circ D\gamma(t)(1) = df(\gamma(t)) \cdot \gamma'(t) = df(\gamma(t)) \cdot (b - a)$ donc $\|g'(t)\| \leq M \|b - a\|$ et on applique la proposition 3.2 à g sur $[0, 1]$ avec $\varphi(t) = M' \|b - a\| t$ où $M' > M$. On obtient

$$\|g(1) - g(0)\| \leq \varphi(1) - \varphi(0)$$

c'est-à-dire

$$\|f(b) - f(a)\| \leq M' \|b - a\|$$

pour tout $M' > M$ donc aussi $\|f(b) - f(a)\| \leq M \|b - a\|$. ■

Rappelons qu'une partie A d'un espace vectoriel E est *convexe* si pour tous $a, b \in A$ le segment $[a, b] \subset A$. Les boules d'un espace normé sont convexes.

Corollaire 3.4 (Théorème des accroissements finis). *Soit U un ouvert convexe de E et $f : U \rightarrow F$ une application différentiable. Alors, pour tout $x, y \in U$,*

$$\|f(x) - f(y)\| \leq \sup_{z \in U} \|df(z)\| \|x - y\|$$

Preuve Ce résultat n'a d'intérêt que lorsque $\sup_{z \in U} \|df(z)\|$ est fini. Dans ce cas, pour tout choix de deux points de U la convexité assure qu'on peut appliquer le corollaire précédent au segment défini par ces deux points. ■

Le théorème des accroissements finis a la très importante conséquence suivante.

Proposition 3.5. *Soit U un ouvert connexe de E et $f : U \rightarrow F$ une application différentiable telle que $\forall x \in U, df(x)$ est l'application nulle, alors f est constante.*

Preuve Fixons $a \in U$ et définissons $\Omega = \{x \in U \mid f(x) = f(a)\}$. Comme f est continue, Ω est un fermé de U . D'autre part si $x \in \Omega \subset U$, il existe $\rho > 0$ tel que $Bo(x, \rho) \subset U$. Une boule étant convexe, si $y \in Bo(x, \rho)$, on peut appliquer le corollaire 3.4 au segment $[x, y]$. Comme $\sup_{z \in Bo(x, \rho)} \|df(z)\| = 0$, on en déduit $f(y) = f(x) = f(a)$ donc $y \in \Omega$. Ceci montre que Ω est aussi un ouvert. Comme U est connexe et $\Omega \neq \emptyset$ (il contient a), c'est que $\Omega = U$ et donc que f est constante sur U . ■

3.2 Applications de classe \mathcal{C}^1

Soit U un ouvert de E et $f : U \rightarrow F$ une application différentiable dans U . On peut définir l'application $df : U \rightarrow L(E, F)$ qui associe à x l'application linéaire continue $df(x)$. Puisque $L(E, F)$ est aussi un espace normé, on se trouve dans la même situation qu'au point de départ et on peut parler de la continuité ou de la différentiabilité de df en un point.

Définition 3.6. *Soit U un ouvert de E . On dit que $f : U \rightarrow F$ est de classe \mathcal{C}^1 (ou continûment différentiable) sur U si*

- *f est différentiable dans U ,*
- *l'application df est continue sur U .*

La deuxième propriété s'explique en : pour tout $x \in U$ et tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour tout $y \in U$ si $\|y - x\| < \alpha$, alors $\|df(x) - df(y)\| \leq \varepsilon$. Cette inégalité équivaut à la propriété : pour tout $h \in E$, $\|df(x) \cdot h - df(y) \cdot h\| \leq \varepsilon \|h\|$.

Remarque

Si E et F sont tous deux de dimension finie la continuité (en x) de df équivaut à celle de l'application $x \mapsto Jf(x)$ ou encore à la continuité comme fonction de $x = (x_1, \dots, x_m)$ de chacune des dérivées partielles $\frac{\partial f_i}{\partial x_j}(x)$.

Proposition 3.7. *Toute application de classe \mathcal{C}^1 sur un ouvert U est localement lipschitzienne, c'est-à-dire que, pour tout $a \in U$, il existe $\rho > 0$ et $k > 0$ tels que $\forall x, y \in Bo(a, \rho)$, on a $\|f(x) - f(y)\| \leq k \|x - y\|$.*

Preuve Noter que k et ρ dépendent de a . Soit $k > \|df(a)\|$. Par continuité de df , il existe une boule centrée en a sur laquelle $\|df(x)\| \leq k$. On peut alors appliquer à cette boule, qui est convexe, le corollaire 3.6. ■

Le plat de résistance de ce chapitre c'est le théorème suivant qui permet, en présence de continuité, de « remonter » de la différentiabilité partielle à la différentiabilité.

Théorème 3.8. Soit U un ouvert de $E_1 \times E_2$ et $f : U \rightarrow F$. Alors f est \mathcal{C}^1 sur U si et seulement si f admet en tout point $x \in U$ une différentielle partielle dans la direction E_1 et une différentielle partielle dans la direction E_2 et, de plus, $d_1f : U \rightarrow L(E_1, F)$ et $d_2f : U \rightarrow L(E_2, F)$ sont continues.

Preuve Dans le sens direct l'existence des différentielles partielles est connue. Montrons la continuité de d_1f sur U . Pour $x, a \in U$, on a

$$\begin{aligned} \|d_1f(x) - d_1f(a)\| &= \sup_{h_1 \in E_1, \|h_1\|=1} \|d_1f(x) \cdot h_1 - d_1f(a) \cdot h_1\| \\ &= \sup_{h_1 \in E_1, \|h_1\|=1} \|df(x) \cdot (h_1, 0) - df(a) \cdot (h_1, 0)\| \\ &\leq \sup_{h \in E_1 \times E_2, \|h\|=1} \|df(x) \cdot h - df(a) \cdot h\| \\ &= \|df(x) - df(a)\| \end{aligned}$$

et cette dernière quantité peut être rendue aussi petite que l'on veut en prenant $\|x - a\|$ assez petit.

Pour établir la réciproque, le plus difficile est de prouver la différentiabilité. Soit $a \in U$, Le « candidat » est connu : c'est l'application $u : E_1 \times E_2 \rightarrow F$ donnée pour $h = (h_1, h_2) \in E$ par

$$u \cdot h = d_1f(a) \cdot h_1 + d_2f(a) \cdot h_2$$

Clairement u est linéaire. Sa continuité est conséquence de la majoration

$$\begin{aligned} \|u \cdot h\| &\leq \|d_1f(a) \cdot h_1\| + \|d_2f(a) \cdot h_2\| \\ &\leq \|d_1f(a)\| \|h_1\| + \|d_2f(a)\| \|h_2\| \\ &\leq (\|d_1f(a)\| + \|d_2f(a)\|) \|h\| \end{aligned}$$

Il faut maintenant montrer que la fonction de $h = (h_1, h_2)$ définie au voisinage de 0 par

$$\varphi(h) = f(a + h) - f(a) - d_1f(a) \cdot h_1 - d_2f(a) \cdot h_2$$

vérifie $\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0$. Supposons $a = (a_1, a_2)$ et posons $b = (a_1 + h_1, a_2)$. On peut écrire

$$\varphi(h) = \varphi_1(h) + \varphi_2(h)$$

où

$$\varphi_1(h) = f(b) - f(a) - d_1f(a) \cdot h_1$$

et

$$\varphi_2(h) = f(a + h) - f(b) - d_2f(a) \cdot h_2.$$

Par définition de la différentielle partielle en a dans la direction E_1 , on a $\lim_{h_1 \rightarrow 0} \frac{\varphi_1(h)}{\|h_1\|} = 0$.

En fait φ_1 ne dépend de h que par l'intermédiaire de h_1 et $\|h_1\| \leq \|h\|$ donc, a fortiori,

$$\lim_{h \rightarrow 0} \frac{\varphi_1(h)}{\|h\|} = 0.$$

C'est pour étudier $\varphi_2(h)$ qu'on utilise le théorème des accroissements finis. On introduit l'application $x_2 \mapsto g(x_2)$ avec

$$g(x_2) = f(a_1 + h_1, x_2) - d_2f(a) \cdot x_2.$$

Elle est définie et différentiable dans un voisinage U_2 de a_2 dans E_2 . De plus $dg(x_2) = d_2f(a_1 + h_1, x_2) - d_2f(a)$. La continuité de d_2f en a permet, pour tout $\varepsilon > 0$, de déterminer $\eta > 0$ tel que $\|x - a\| < \eta$ implique

$$\|d_2f(x) - d_2f(a)\| < \varepsilon.$$

Donc, si $x = (a_1 + h_1, x_2)$ avec $\|h_1\| < \eta$ et $\|x_2 - a_2\| < \eta$, on a $\|dg(x_2)\| < \varepsilon$. En particulier, si $\|h_2\| < \eta$, cette majoration est valable pour tout $x_2 \in [a_2, a_2 + h_2]$ et le théorème des accroissements finis appliqué à g sur ce segment donne

$$\|g(a_2 + h_2) - g(a_2)\| \leq \varepsilon \|h_2\|.$$

Or $g(a_2 + h_2) = f(a + h) - d_2f(a) \cdot (a_2 + h_2)$ et $g(a_2) = f(b) - d_2f(a) \cdot a_2$ de sorte que, utilisant la linéarité de $d_2f(a)$, ceci s'écrit

$$\|f(a + h) - f(b) - d_2f(a) \cdot h_2\| = \|\varphi_2(h)\| \leq \varepsilon \|h_2\| \leq \varepsilon \|h\|.$$

On a donc établi que f est différentiable dans U et que pour tout $x \in U$ et tout $h = (h_1, h_2) \in E$,

$$df(x) \cdot h = d_1f(x) \cdot h_1 + d_2f(x) \cdot h_2.$$

Il reste à prouver la continuité de df sur U . Si x et $a \in U$,

$$\begin{aligned} \|df(x) - df(a)\| &= \sup_{\|h\| \leq 1} \|df(x) \cdot h - df(a) \cdot h\| \\ &= \sup_{\|h_1\| \leq 1, \|h_2\| \leq 1} \|(d_1f(x) - d_1f(a)) \cdot h_1 + (d_2f(x) - d_2f(a)) \cdot h_2\| \\ &\leq \sup_{\|h_1\| \leq 1} \|(d_1f(x) - d_1f(a)) \cdot h_1\| + \sup_{\|h_2\| \leq 1} \|(d_2f(x) - d_2f(a)) \cdot h_2\| \\ &\leq \|d_1f(x) - d_1f(a)\| + \|d_2f(x) - d_2f(a)\| \end{aligned}$$

et cette dernière quantité tend vers 0 lorsque x tend vers a puisque les différentielles partielles sont continues en a . ■

Voici un très utile corollaire.

Corollaire 3.9. *Soit E et F deux espaces vectoriels de dimension finie. Soit U un ouvert de E . L'application $f : U \rightarrow F$ est de classe \mathcal{C}^1 si et seulement si pour tout $i \in \{1, \dots, \dim F\}$ et tout $j \in \{1, \dots, \dim E\}$, $\frac{\partial f_i}{\partial x_j}$ est définie et continue dans U .*

3.3 Sur la différentiabilité au sens de Gâteaux

Soit

$$\begin{cases} f(x, y) = \left(\frac{x^2 y}{x^4 + y^2}\right)^2, \text{ si } y \neq 0, \\ f(x, 0) = 0. \end{cases}$$

On vérifie que f est G -différentiable en $(0, 0)$, mais f n'est pas continue en zéro. Ainsi, la différentiabilité (au sens de Fréchet) implique la G -différentiabilité, mais la réciproque est fautive en général car la G -différentiabilité n'implique pas la continuité. En revanche, si on sait que f est différentiable au sens de Gâteaux et si l'application

$$\begin{aligned} U &\mapsto L(E, F) \\ x &\mapsto d_G f(x) \end{aligned}$$

est continue, alors f est différentiable (au sens de Fréchet) sur U et les deux notions coïncident. En effet, on a le résultat suivant.

Théorème 3.10. Soit $f : U \rightarrow F$ un application différentiable au sens de Gâteaux. On suppose que l'application $x \mapsto d_G f(x)$ est continue sur U , alors f est de classe \mathcal{C}^1 sur U et

$$df(x) = d_G f(x), \quad \forall x \in U.$$

Preuve Soit $a \in U$. On pose $R(h) = f(a+h) - f(a) - d_G f(a) \cdot h$. Cette application est G-différentiable et

$$d_G R(h) : k \rightarrow d_G f(a+h) \cdot k - d_G f(a) \cdot k.$$

En regardant la preuve du théorème des accroissements finis, on peut voir qu'il reste vrai si f est G-différentiable. En utilisant donc ce théorème avec le segment $[0, h]$ et le fait que $R(0) = 0$, on obtient :

$$\|R(h)\| \leq \sup_{0 \leq t \leq 1} \|d_G R(th)\| \|h\|.$$

Or

$$\|d_G R(th)\| = \|d_G f(a+th) - d_G f(a)\|,$$

on a

$$\|R(h)\| \leq \sup_{0 \leq t \leq 1} \|d_G f(a+th) - d_G f(a)\| \|h\|.$$

Le fait que l'application $x \mapsto d_G f(x)$ est continue implique que

$$\sup_{0 \leq t \leq 1} \|d_G f(a+th) - d_G f(a)\| \rightarrow 0,$$

lorsque $\|h\| \rightarrow 0$, ce qui montre que f est différentiable et $df(a) = d_G f(a)$. La continuité de df est une conséquence immédiate de la continuité de $d_G f$. ■

Remarque 3.11. En pratique, il est parfois plus facile de vérifier la G-différentiabilité. Si on veut prouver que f est \mathcal{C}^1 il suffit donc de prouver qu'elle est différentiable au sens de Gâteaux, puis de vérifier que la différentielle $d_G f$ est continue.

3.4 Propriétés des fonctions \mathcal{C}^1

Les points non démontrés de la proposition qui suit sont sans piège.

Proposition 3.12. 1. Toute combinaison linéaire d'applications de classe \mathcal{C}^1 sur un ouvert U est de classe \mathcal{C}^1 sur U .

2. La composée de deux applications de classe \mathcal{C}^1 est de classe \mathcal{C}^1 .

3. Toute application affine continue, toute application n -linéaire continue est de classe \mathcal{C}^1 .

4. Si f et $g : U \rightarrow \mathbb{R}$ sont de classe \mathcal{C}^1 sur l'ouvert U , leur produit est aussi de classe \mathcal{C}^1 sur U et il en est de même pour les autres produits étudiés au chapitre 1.

5. Si $f : U \rightarrow \mathbb{R}$ est \mathcal{C}^1 et si, pour tout $x \in U$, $f(x) \neq 0$, alors la fonction $\frac{1}{f}$ est de classe \mathcal{C}^1 sur U .

Preuve

- On sait que si $f : U \rightarrow F$ est différentiable et si $g : V \rightarrow G$ est différentiable où V est un ouvert de F tel que $f(U) \subset V$, alors $\varphi = g \circ f$ est différentiable et, pour tout $x \in U$, $d\varphi(x) = dg(f(x)) \circ df(x)$. Cette formule peut se lire $d\varphi = B \circ \Phi$ où

$$\begin{aligned} \Phi : U &\longrightarrow L(F, G) \times L(E, F) \\ x &\longmapsto (dg(f(x)), df(x)) \end{aligned}$$

autrement dit $\Phi = (dg \circ f, df)$. Et B est l'application bilinéaire continue

$$\begin{aligned} L(F, G) \times L(E, F) &\longrightarrow L(E, G) \\ (u, v) &\longmapsto u \circ v \end{aligned}$$

Comme f est \mathcal{C}^1 , df est continue sur U . Comme g est \mathcal{C}^1 , dg est continue sur V et, puisque f est continue sur U et que $f(U) \subset V$, la composée $dg \circ f$ est continue sur U . L'application Φ est continue sur U puisque ses deux composantes le sont. Enfin $d\varphi$ est continue comme composée de deux fonctions continues.

- On sait que si $f : U \rightarrow \mathbb{R}$ est différentiable et ne s'annule pas, la fonction $g = \frac{1}{f}$ est définie et différentiable en tout point de U , sa différentielle en $x \in U$ étant la forme linéaire continue sur E : $dg(x) = -\frac{1}{f(x)^2} df(x)$. Cette formule signifie que dg est le produit de la fonction continue $-\frac{1}{f^2}$ et de l'application continue (puisque f est \mathcal{C}^1) df , d'où la continuité de dg sur U . ■

Dans la suite de ce chapitre on suppose que les espaces sont des Banach.

Proposition 3.13. *Si E et F sont deux espaces de Banach, l'ensemble $\text{Isom}(E, F)$ est un ouvert de $L(E, F)$.*

Preuve Le cas $\text{Isom}(E, F) = \emptyset$ n'est pas exclu, mais c'est bien un ouvert. Sinon, soit $u \in \text{Isom}(E, F)$ et $h \in L(E, F)$, on peut écrire, puisque u est inversible

$$u + h = u \circ (id_E + u^{-1} \circ h) = u \circ v$$

où $v = id_E + u^{-1} \circ h$ appartient à $L(E)$ (abréviation de $L(E, E)$) qui est un Banach puisque E en est un. Si $\|u^{-1} \circ h\| < 1$, ce qui a sûrement lieu si $\|h\| < r = \frac{1}{\|u^{-1}\|}$, la série de terme général $(-1)^n (u^{-1} \circ h)^n$ est normalement convergente donc convergente et sa somme est l'inverse de v qui appartient donc à $\text{Isom}(E)$. Mais alors $u \circ v = u + h$ appartient à $\text{Isom}(E, F)$. On vient de montrer que la boule ouverte de $L(E, F)$ de centre u de rayon r est incluse dans $\text{Isom}(E, F)$. ■

Remarque 3.14. En particulier, lorsque $E = F$ est un Banach, la preuve précédente montre que la boule ouverte de centre id_E et de rayon 1 du Banach $L(E)$ est contenue dans $\text{Isom}(E)$. Autrement dit, si $v \in L(E)$ et

$$\|id_E - v\| < 1$$

alors $v \in \text{Isom}(E)$.

Proposition 3.15. *Soit E et F deux espaces de Banach tels que $\text{Isom}(E, F) \neq \emptyset$. L'application*

$$\begin{aligned} \mathcal{J} : \text{Isom}(E, F) &\longrightarrow \text{Isom}(F, E) \\ u &\longmapsto u^{-1} \end{aligned}$$

est différentiable et $D\mathcal{J}(u)$ est l'application (linéaire continue) de $L(E, F)$ dans $L(F, E)$ donnée, pour $h \in L(E, F)$, par

$$D\mathcal{J}(u) \cdot h = -u^{-1} \circ h \circ u^{-1}.$$

Preuve On reprend les notations précédentes. Pour $u \in \text{Isom}(E, F)$ et $h \in L(E, F)$ avec $\|h\| < r$, on a

$$\begin{aligned} \mathcal{J}(u+h) - \mathcal{J}(u) &= (u+h)^{-1} - u^{-1} = (u \circ v)^{-1} - u^{-1} = \\ v^{-1} \circ u^{-1} - u^{-1} &= (id_E + u^{-1} \circ h)^{-1} \circ u^{-1} - u^{-1} \\ &= \sum_{n=0}^{\infty} (-1)^n (u^{-1} \circ h)^n \circ u^{-1} - u^{-1} \\ &= \sum_{n=1}^{\infty} (-1)^n (u^{-1} \circ h)^n \circ u^{-1} \\ &= U(h) + \varphi(h) \end{aligned}$$

où $U(h) = -u^{-1} \circ h \circ u^{-1}$ et $\varphi(h) = \sum_{n=2}^{\infty} (-1)^n (u^{-1} \circ h)^n \circ u^{-1}$, toutes ces égalités étant valables sous réserve de convergence des séries écrites.

L'application $U : L(E, F) \rightarrow L(F, E)$ ainsi définie est clairement linéaire et la majoration

$$\|U(h)\| \leq \|u^{-1}\|^2 \|h\|$$

montre qu'elle est continue (et de norme inférieure ou égale à $\|u^{-1}\|^2$).

Montrons que la série qui définit φ est absolument convergente. En effet, par récurrence sur n , on voit que $\|(u^{-1} \circ h)^n \circ u^{-1}\| \leq \|u^{-1}\|^n \|h\|^n$, qui est le terme général d'une série (géométrique) convergente. Comme $L(E)$ est un Banach, la série définissant $\varphi(h)$ converge et les calculs précédents sont justifiés. De plus, en passant à la limite dans les sommes partielles, on a la majoration

$$\|\varphi(h)\| \leq \sum_{n=2}^{\infty} \|u^{-1}\|^{n+1} \|h\|^n = \frac{\|h\|^2 \|u^{-1}\|^3}{1 - \|h\| \|u^{-1}\|}.$$

Cette dernière égalité montre que $\lim_{h \rightarrow 0} \frac{\varphi(h)}{\|h\|} = 0$. ■

Corollaire 3.16. *Sous les mêmes hypothèses, l'application \mathcal{J} est continue.*

Proposition 3.17. *Soit U un ouvert d'un espace de Banach E , V un ouvert d'un espace de Banach F et $f : U \rightarrow V$ un homéomorphisme de classe \mathcal{C}^1 tel que pour tout $x \in U$, $df(x) \in \text{Isom}(E, F)$, alors f^{-1} est \mathcal{C}^1 sur V .*

Preuve On sait déjà que f est différentiable en tout point de V et que, pour tout $y \in V$, $df^{-1}(y) = df(f^{-1}(y))^{-1}$. Cette formule peut se lire $df^{-1} = \mathcal{J} \circ df \circ f^{-1}$ et fait apparaître df^{-1} comme composée de trois applications continues, d'où le résultat. ■

Chapitre 4

Théorème d'inversion locale et théorème des fonctions implicites

Dans tout ce chapitre les espaces considérés sont des Banach.

4.1 Théorème d'inversion locale

La définition suivante est naturelle.

Définition 4.1. Soit U un ouvert de E , V un ouvert de F . Une application $f : U \rightarrow V$ est un *difféomorphisme* (resp. \mathcal{C}^1 -difféomorphisme) si c'est une bijection telle que f et f^{-1} sont différentiables (resp. de classe \mathcal{C}^1).

Dans tout ce chapitre on considère U un ouvert de E non vide.

Remarque 4.2. Clairement tout difféomorphisme est un homéomorphisme.

S'il existe un difféomorphisme f d'un ouvert U de E sur un ouvert V de F , alors $\text{Isom}(E, F)$ doit être non vide. En effet, puisque $f^{-1} \circ f = \text{id}_U$, en différentiant on a, si $x \in U$ et $y = f(x)$, $\text{d}f^{-1}(y) \circ \text{d}f(x) = \text{id}_E$. De même, puisque $f \circ f^{-1} = \text{id}_V$, on a $\text{d}f(x) \circ \text{d}f^{-1}(y) = \text{id}_F$. Ceci montre que $\text{d}f(x) \in \text{Isom}(E, F)$. Si E est de dimension finie cela impose que F est aussi de dimension finie et que E et F sont de même dimension.

Le dernier résultat du chapitre précédent permet d'énoncer la proposition suivante.

Proposition 4.3. Si $f : U \rightarrow V$ est un homéomorphisme différentiable (resp. de classe \mathcal{C}^1) sur U et tel que pour tout $x \in U$, $\text{d}f(x) \in \text{Isom}(E, F)$ alors f est un difféomorphisme (resp. \mathcal{C}^1 -difféomorphisme) de U sur V .

Il est en général plus facile d'étudier l'inversibilité de l'application linéaire $\text{d}f(x)$ que celle de f : en dimension finie il suffit de calculer un déterminant (le jacobien de f en x).

En dimension 1 la condition $\text{d}f(x) \in \text{Isom}(E, F)$ s'exprime par $f'(x) \neq 0$. C'est une condition « ouverte » : si f' est continue et si la condition est réalisée en un point, elle l'est au voisinage. L'analogie lorsque E et F sont des Banach, en particulier en dimension finie, est le fait que $\text{Isom}(E, F)$ est un ouvert de $L(E, F)$. Toujours en dimension 1, on peut même déduire de $f'(a) \neq 0$ et de la continuité de f' qu'au voisinage de a la fonction f est strictement monotone et réalise donc une *bijection* d'un voisinage de a sur son image. Autrement dit on peut *localement* « remonter » du fait que la différentielle est bijective au fait que la fonction l'est. Le théorème d'inversion locale permet d'obtenir cette propriété pour les Banach.

Théorème 4.4 (Théorème d'inversion locale). *Soit E et F deux espaces de Banach et U un ouvert de E . Soit $f : U \rightarrow F$ une application de classe \mathcal{C}^1 . Soit $a \in U$ tel que $df(a) \in \text{Isom}(E, F)$, alors il existe un voisinage ouvert U_1 de a et un voisinage ouvert V_1 de $f(a)$ tels que $U_1 \subset U$ et $f : U_1 \rightarrow V_1$ est un \mathcal{C}^1 -difféomorphisme.*

Preuve Commençons par montrer qu'il suffit d'établir le théorème lorsque $E = F$, $a = f(a) = 0$ et $df(0) = id_E$. En effet, posons

$$\tilde{f}(x) = df(a)^{-1} \cdot [f(a+x) - f(a)].$$

On vérifie que

- l'application \tilde{f} est définie sur $\tilde{U} = \{x \in E \mid a+x \in U\}$ qui est un voisinage ouvert de 0 et $\tilde{f} : \tilde{U} \rightarrow E$ est de classe \mathcal{C}^1 ,
- en 0, $\tilde{f}(0) = df(a)^{-1} \cdot 0 = 0$,
- $D\tilde{f}(x) = df(a)^{-1} \circ df(a+x)$. En particulier $D\tilde{f}(0) = df(a)^{-1} \circ df(a) = id_E$.

Si on a établi le théorème pour \tilde{f} , on « revient » à f en remarquant que $f(a+x) = f(a) + df(a) \cdot \tilde{f}(x)$ ou encore $f(x) = f(a) + df(a) \circ \tilde{f}(x-a) = A \circ \tilde{f}(x-a)$ où $A : E \rightarrow F$ est la *bijection affine continue* $h \mapsto f(a) + df(a) \cdot h$. Si donc \tilde{f} est un \mathcal{C}^1 -difféomorphisme de \tilde{U}_1 (voisinage ouvert de 0) sur \tilde{V}_1 (voisinage ouvert de 0), alors f est un \mathcal{C}^1 -difféomorphisme de $U_1 = a + \tilde{U}_1$ (voisinage ouvert de a) sur $V_1 = A(\tilde{V}_1)$ qui est un voisinage ouvert de $A(0) = f(a)$ dans F .

On se place dans le cas particulier indiqué. Montrer que f est bijective de U_1 sur V_1 c'est montrer qu'on peut résoudre, pour y fixé dans V_1 , l'équation $f(x) = y$ de façon unique si l'on impose que la solution trouvée x appartient à U_1 . On le fait en montrant qu'on peut appliquer le théorème du point fixe à l'application $g_y(x) = x - f(x) + y$ dans une *boule fermée* convenable, puisqu'une boule fermée d'un Banach est un espace métrique complet.

Pour tout $y \in E$, l'application g_y est de classe \mathcal{C}^1 sur U et $dg_y(x) = id_E - df(x)$ est *indépendante de y* et nulle en $x = 0$. Donc par continuité, il existe $r > 0$ tel que, si x vérifie $\|x\| \leq r$, alors $x \in U$ et pour tout $y \in E$, $\|dg_y(x)\| \leq \frac{1}{2}$. Le théorème des accroissements finis permet d'en déduire que pour tous $x, x' \in Bf(0, r)$ et tout $y \in E$,

$$(*) \quad \|g_y(x) - g_y(x')\| \leq \frac{1}{2}\|x - x'\|.$$

Comme $g_y(0) = y$, on a en particulier $\|g_y(x) - y\| \leq \frac{1}{2}\|x\|$ pour tout $x \in Bf(0, r)$. Supposons que y vérifie $\|y\| < \frac{1}{2}r$. Alors, si $x \in Bf(0, r)$,

$$\|g_y(x)\| = \|g_y(x) - y\| + \|y\| \leq \frac{1}{2}\|x\| + \|y\| < r.$$

Donc g_y envoie la boule fermée $Bf(0, r)$ sur la boule ouverte $Bo(0, r)$ et, a fortiori, sur la boule fermée elle-même. L'inégalité (*) montre que g_y est contractante et a donc un unique point fixe. Ceci permet de *définir* une application $\varphi : V_1 \rightarrow E$ où $V_1 = Bo(0, \frac{1}{2}r)$ en associant à $y \in V_1$ l'unique point fixe $\varphi(y)$ de g_y dans $Bf(0, r)$. Comme ce point fixe est image par g_y d'un point de la boule fermée $Bf(0, r)$, il appartient en fait à la boule ouverte $Bo(0, r)$. On va montrer que $U_1 = \varphi(V_1)$ vérifie $U_1 = f^{-1}(V_1) \cap Bo(0, r)$ et est donc un voisinage ouvert de 0 et que les applications $f : U_1 \rightarrow V_1$ et $\varphi : V_1 \rightarrow U_1$ sont inverses l'une de l'autre. En effet :

- D'une part, si $y \in V_1$, $\varphi(y) \in Bo(0, r)$ et vérifie $g_y(\varphi(y)) = \varphi(y) = \varphi(y) - f(\varphi(y)) + y$. D'où (*) $f(\varphi(y)) = y$. Si donc $x \in U_1$, de sorte que $x = \varphi(y)$ où $y \in V_1$, on a $x \in Bo(0, r)$ et $f(x) = y$, ce qui prouve l'inclusion $U_1 \subset f^{-1}(V_1) \cap Bo(0, r)$ et permet d'interpréter (*) en $f \circ \varphi = id_{V_1}$.

- D'autre part, si $x \in f^{-1}(V_1) \cap Bo(0, r)$, alors $\|x\| < r$ et $f(x) \in V_1$. Mais $g_{f(x)}(x) = x - f(x) + f(x) = x$, ce qui montre que x est point fixe dans $Bf(0, r)$ de $g_{f(x)}$. Donc par unicité on a $(**) x = \varphi(f(x))$. Ceci montre que $x \in U_1$, ce qui prouve l'inclusion $f^{-1}(V_1) \cap Bo(0, r) \subset U_1$ et permet d'interpréter $(**)$ en $\varphi \circ f = id_{U_1}$.

Pour prouver que φ est de classe \mathcal{C}^1 sur V_1 , il suffit de montrer qu'elle est *continue*. En effet, pour $x \in U_1$, $\|id_E - df(x)\| \leq \frac{1}{2} < 1$ et la remarque 3.14 montre que $df(x) \in \text{Isom}(E)$. On peut alors appliquer à l'homéomorphisme $f : U_1 \rightarrow V_1$ la proposition 4.3.

Pour établir la continuité de φ , nous allons montrer qu'elle est lipschitzienne de rapport 2 sur V_1 . Si $y, y' \in V_1$, alors $x = \varphi(y)$ et $x' = \varphi(y')$ appartiennent à $Bo(0, r)$ et, pour $y_0 \in E$ fixé (quelconque), en utilisant (\star) , on a :

$$\|x - x' - f(x) + f(x')\| = \|g_{y_0}(x) - g_{y_0}(x')\| \leq \frac{1}{2}\|x - x'\|$$

d'où $\|x - x'\| \leq \|f(x) - f(x')\| + \frac{1}{2}\|x - x'\|$ et $\|x - x'\| \leq 2\|f(x) - f(x')\|$ ou encore

$$\|\varphi(y) - \varphi(y')\| \leq 2\|y - y'\|. \quad \blacksquare$$

On adopte parfois la terminologie suivante, utile lorsqu'on veut privilégier l'espace source.

Définition 4.5. Soit $f : U \rightarrow F$ une application de classe \mathcal{C}^1 sur un ouvert d'un Banach E . On dit que f est un \mathcal{C}^1 -difféomorphisme local au voisinage de $a \in U$ s'il existe un voisinage ouvert U_a de a et un ouvert V de F tel que la restriction de f à U_a soit un \mathcal{C}^1 -difféomorphisme de U_a sur V .

Le théorème d'inversion locale peut alors s'énoncer comme suit :

Théorème 4.6. Soit E et F deux espaces de Banach et U un ouvert de E . Une application $f : U \rightarrow F$ de classe \mathcal{C}^1 est un \mathcal{C}^1 -difféomorphisme local au voisinage de $a \in U$ si et seulement si $df(a) \in \text{Isom}(E, F)$.

Lorsque $E = F$ est de dimension finie, un \mathcal{C}^1 -difféomorphisme local s'appelle aussi un *changement de coordonnées*.

Le corollaire suivant explicite *en dimension finie* une généralisation locale \mathcal{C}^1 du théorème de la base incomplète de l'algèbre linéaire.

Corollaire 4.7. Soit U un ouvert de \mathbb{R}^n , p un entier tel que $1 \leq p \leq n$ et f_1, \dots, f_p des fonctions numériques de classe \mathcal{C}^1 . Soit $a \in U$, les deux propriétés suivantes sont équivalentes :

- les p formes linéaires $df_1(a), \dots, df_p(a)$ sont linéairement indépendantes,
- il existe $n - p$ fonctions numériques f_{p+1}, \dots, f_n de classe \mathcal{C}^1 sur U telles que $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ soit un \mathcal{C}^1 -difféomorphisme local au voisinage de a .

Preuve Si $p = n$, c'est une simple reformulation du théorème d'inversion locale si on se souvient qu'une matrice est inversible si et seulement si les n formes linéaires définies par ses lignes sont indépendantes. Si $p < n$ et si les p vecteurs $df_1(a), \dots, df_p(a)$ du dual de \mathbb{R}^n sont indépendants, le théorème de la base incomplète montre qu'il existe $n - p$ formes linéaires sur \mathbb{R}^n , u_{p+1}, \dots, u_n telles que $(df_1(a), \dots, df_p(a), u_{p+1}, \dots, u_n)$ soit une base du dual de \mathbb{R}^n . Définissons $f : U \rightarrow \mathbb{R}^n$ par $f = (f_1, \dots, f_n)$ en posant, pour $j \geq p + 1$, $f_j = u_{j|_U}$. La fonction numérique f_j ($p + 1 \leq j \leq n$) est de classe \mathcal{C}^1 et vérifie, pour tout $x \in U$, $df_j(x) = u_j$. On est ainsi ramené au cas $n = p$.

La réciproque est claire. \blacksquare

Remarques

1. Dans le théorème d'inversion locale, l'hypothèse de **continuité** de df est indispensable. On pourrait croire qu'en demandant seulement à f d'être différentiable, on aurait le même résultat, sans la continuité de l'application inverse. L'exemple suivant prouve qu'il n'en est rien, même en dimension finie. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = \frac{1}{2}x + x^2 \sin \frac{1}{x}$ lorsque $x \neq 0$ et $f(0) = 0$. Un dessin permet de se convaincre que f n'est injective sur aucun intervalle contenant 0. Pourtant f est dérivable en tout point et $f'(0) = \lim_{x \rightarrow 0} \frac{f(x)}{x} = \frac{1}{2} \neq 0$. Mais pour $x \neq 0$, $f'(x) = \frac{1}{2} + 2x \sin \frac{1}{x} - \cos \frac{1}{x}$, qui n'a pas de limite quand $x \rightarrow 0$.
2. On ne saurait trop insister sur le fait qu'il s'agit d'un résultat **local**. Même si f , de classe \mathcal{C}^1 , vérifie $df(x) \in \text{Isom}(E, F)$ pour tout $x \in U$, on ne peut pas en déduire que f est une bijection de U sur son image. Voici un contre-exemple. L'application $f : U =]0, +\infty[\times \mathbb{R} \rightarrow \mathbb{R}^2$ définie par $f(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta)$ est de classe \mathcal{C}^1 et en tout point de U le jacobien vaut $\det Jf(\rho, \theta) = \rho \neq 0$. Mais la périodicité en θ montre que f n'est pas injective sur U . Le théorème affirme que si (ρ_0, θ_0) est fixé dans U , il existe $U_1 \subset U$, $(U_1 \ni (\rho_0, \theta_0))$, tel que $f|_{U_1}$ soit injective. On peut en effet prendre $U_1 =]0, +\infty[\times]\theta_0 - \alpha, \theta_0 - \alpha + 2\pi[$, avec $0 < \alpha < 2\pi$. Un examen scrupuleux de la preuve montre que l'ouvert qui y est construit est plus petit que celui-ci.

Comme le suggère l'exemple qui vient d'être traité, c'est l'**injectivité** de f qui pose problème. Les deux résultats qui suivent sont d'une grande importance pratique. Le deuxième explicite comment passer du résultat local à un résultat global.

Théorème 4.8 (Théorème de l'image ouverte). *Si $f : U \rightarrow F$ de classe \mathcal{C}^1 est telle que pour tout $x \in U$, $df(x) \in \text{Isom}(E, F)$, alors $f(U)$ est un ouvert de F et, plus généralement, l'image par f de tout ouvert de E inclus dans U est un ouvert de F .*

Preuve Soit $y_0 \in f(U)$. Il existe $x_0 \in U$ tel que $y_0 = f(x_0)$. Le théorème appliqué au point x_0 donne un ouvert $U_1 \subset U$, $U_1 \ni x_0$, et un ouvert $V_1 \ni y_0$ tels que f de U_1 sur V_1 soit bijective. En particulier, pour tout $y \in V_1$, il existe $x \in U_1 \subset U$ tel que $y = f(x)$. Autrement dit $V_1 \subset f(U)$. Ceci montre que $f(U)$ est un ouvert de F . Bien sûr l'argument s'applique à tout ouvert de E inclus dans U . ■

Théorème 4.9 (Théorème d'inversion globale sur l'image). *Soit $f : U \rightarrow f(U)$ une application de classe \mathcal{C}^1 . Si*

1. *f est injective et*
2. *pour tout $x \in U$, $df(x) \in \text{Isom}(E, F)$,*

alors f est un \mathcal{C}^1 -difféomorphisme (sur son image $f(U)$)

Preuve Notons que le théorème 4.8 montre que l'image $V = f(U)$ est un ouvert de F et l'hypothèse 1. permet d'affirmer que f est une bijection de U sur son image V . Il reste à montrer que l'application $f^{-1} : V \rightarrow U$ ainsi définie est continue. Les autres affirmations seront en effet cette fois encore conséquence de la proposition 4.3. Or la continuité de f^{-1} équivaut au fait que f est ouverte et le théorème 4.8 assure cette propriété. ■

Proposition 4.10. *Soit U connexe et $f : U \rightarrow F$. Si*

1. *pour tout $x \in U$, $df(x) \in \text{Isom}(E, F)$, et*
2. *$f(U)$ est fermé,*

alors f est surjective, i.e. $f(U) = F$.

Démonstration. Le théorème 4.8 montre que l'image $V = f(U)$ est un ouvert de F , V est donc un ouvert et fermé de F . Or f est continue, on a donc que V est connexe, ce qui implique que $V = F$. □

Un conséquence immédiate du théorème 4.9 et de la proposition 4.10 est le théorème d'inversion global suivant.

Théorème 4.11 (Théorème d'inversion globale). *Soit $f : U \rightarrow F$ telle que*

1. U est connexe,
2. f est injective,
3. pour tout $x \in U$, $df(x) \in \text{Isom}(E, F)$, et
4. $f(U)$ est fermé,

alors f est un \mathcal{C}^1 -difféomorphisme globale.

4.2 Le théorème des fonctions implicites

Le théorème d'inversion locale peut s'interpréter comme la possibilité de résoudre *localement* pour chaque valeur de y de façon unique l'équation en $x : f(x) - y = 0$ et d'affirmer la régularité de la *fonction* $x(y)$ ainsi définie. La condition à satisfaire, en plus d'une « régularité » convenable de f , est l'inversibilité de l'application linéaire que constitue la différentielle, ce que l'on peut interpréter comme la possibilité de résoudre *le même problème* dans le cas *linéaire*. Plus généralement le théorème des fonctions implicites va donner des conditions permettant de résoudre $f(x, y) = 0$ par une fonction $y(x)$ convenable (on a échangé les lettres x et y pour se conformer à l'habitude). Cette fois aussi l'idée directrice est que ce qu'on sait faire dans le cadre linéaire pourra être fait *localement* dans le cadre \mathcal{C}^1 .

On sait que, si dans un système *linéaire* de n équations à $n+m$ inconnues un déterminant d'ordre n extrait de la matrice du système est non nul, on peut choisir convenablement n des inconnues et les exprimer comme *fonctions linéaires* des m autres inconnues. Si la matrice de ce système est la matrice jacobienne en un point d'une application d'un ouvert de \mathbb{R}^{n+m} ($n+m$ est le nombre de variables) dans \mathbb{R}^n (n est le nombre d'équations), la condition précédente s'interprète comme la possibilité, pour un ordre convenable des variables, de décomposer \mathbb{R}^{n+m} en $\mathbb{R}^m \times \mathbb{R}^n$ de sorte que la matrice jacobienne de la *différentielle partielle* dans la direction du facteur \mathbb{R}^n (deuxième facteur) soit inversible. Le théorème d'inversion locale peut être traduit dans ce langage. Reprenons en effet les hypothèses du théorème 4.4 et introduisons l'application $\varphi : F \times U \rightarrow F$ définie par $\varphi(y, z) = f(z) - y$. On a, pour tout $y \in F$, $d_2\varphi(y, a) = df(a)$. L'hypothèse peut donc se lire $d_2\varphi(b, a) \in \text{Isom}(E, F)$ avec $(b, a) = (f(a), a)$, point où la relation $\varphi(b, a) = 0$ est réalisée. La conclusion du théorème peut se lire : il existe un voisinage ouvert U_1 de a , un voisinage ouvert V de b et une application de classe \mathcal{C}^1 , $g : V \rightarrow U_1$ telle que pour tout $y \in V$, $z = g(y)$ est l'unique solution *dans* U_1 de $\varphi(y, z) = 0$. Autrement dit on peut résoudre *localement* l'équation $\varphi(y, z) = 0$ par une application de classe \mathcal{C}^1 , $z = g(y)$. Le théorème suivant montre que cette situation est générale.

Théorème 4.12. *Soit E, F, G trois espaces de Banach, Ω un ouvert de $E \times F$ et $f : \Omega \rightarrow G$ une application de classe \mathcal{C}^1 . Soit $(a, b) \in \Omega$ tel que $d_2f(a, b) \in \text{Isom}(F, G)$. Alors il existe un ouvert U de E contenant a , un ouvert V de F contenant b et une application $g : U \rightarrow V$ tels que*

1. $U \times V \subset \Omega$
2. pour tout $x \in U$, l'équation $f(x, y) = f(a, b)$ admet $y = g(x)$ pour unique solution dans V .

De plus l'application g est de classe \mathcal{C}^1 dans U et, si U est assez petit, pour tout $x \in U$,

$$dg(x) = -d_2f(x, g(x))^{-1} \circ d_1f(x, g(x)).$$

Remarques

1. L'hypothèse implique que F et G ont même dimension (finie ou non).
2. Il est important de bien comprendre l'affirmation 2. qui *ne dit pas* que l'équation $f(x, y) = f(a, b)$ n'a qu'une solution : elle n'en a qu'une dans V (c'est-à-dire proche de b) mais peut en avoir ailleurs.
3. La propriété 2. peut aussi s'exprimer par l'égalité :

$$\{(x, y) \in U \times V \mid f(x, y) = f(a, b)\} = \{(x, g(x)) \mid x \in U\} = \Gamma_g$$

en désignant par Γ_g le *graphe* de la fonction g .

Preuve Le préambule du théorème laisse deviner que sa preuve n'est pas très différente de celle du théorème d'inversion locale. En effet on va déduire ce résultat du précédent. On aurait pu aussi prouver d'abord ce théorème et en déduire, comme cas particulier, le théorème 4.4.

Définissons l'application $\Phi : \Omega \rightarrow E \times G$ par $\Phi(x, y) = (x, f(x, y))$ et montrons qu'on peut lui appliquer le théorème d'inversion locale au point (a, b) . Elle est clairement de classe \mathcal{C}^1 sur Ω et sa différentielle au point (a, b) est donnée, pour $(h, k) \in E \times F$, par

$$D\Phi(a, b) \cdot (h, k) = (h, df(a, b) \cdot (h, k)) = (h, d_1f(a, b) \cdot h + d_2f(a, b) \cdot k).$$

On résout « à la main » l'équation $D\Phi(a, b) \cdot (h, k) = (h', k')$. On trouve $h = h'$ puis $d_2f(a, b) \cdot k = k' - d_1f(a, b) \cdot h'$ et donc

$$k = d_2f(a, b)^{-1} \cdot [k' - d_1f(a, b) \cdot h'].$$

Autrement dit l'application $D\Phi(a, b)$ est inversible, son inverse étant

$$D\Phi(a, b)^{-1} \cdot (h', k') = (h', d_2f(a, b)^{-1} \cdot [k' - d_1f(a, b) \cdot h']).$$

Il faut aussi montrer la continuité de l'application linéaire $D\Phi(a, b)^{-1}$. Examinons la deuxième composante. On a

$$d_2f(a, b)^{-1} \cdot [k' - d_1f(a, b) \cdot h'] = d_2f(a, b)^{-1} \cdot k' - d_2f(a, b)^{-1} \circ d_1f(a, b)(h')$$

d'où

$$\begin{aligned} \|d_2f(a, b)^{-1} \cdot [k' - d_1f(a, b) \cdot h']\| &\leq \|d_2f(a, b)^{-1}\|(\|k'\| + \|d_1f(a, b)\| \|h'\|) \\ &\leq \underbrace{\|d_2f(a, b)^{-1}\|(1 + \|d_1f(a, b)\|)}_M \|h', k'\| \end{aligned}$$

et donc $\|D\Phi(a, b)^{-1} \cdot (h', k')\| \leq \max(1, M)\|(h', k')\|$. On a ainsi montré que $D\Phi(a, b)$ appartient à $\text{Isom}(E \times F, E \times G)$.

On a $\Phi(a, b) = (a, f(a, b))$ et le théorème d'inversion locale appliqué à Φ au point (a, b) donne un voisinage ouvert Ω_1 de (a, b) , $\Omega_1 \subset \Omega$ et un voisinage ouvert W de $(a, f(a, b))$ tels que $\Phi : \Omega_1 \rightarrow W$ soit un \mathcal{C}^1 -difféomorphisme. Quitte à diminuer de manière cohérente Ω_1 et W , on peut supposer que $\Omega_1 = U_1 \times V$ où U_1 est un voisinage ouvert de a dans E et V un voisinage ouvert de b dans F . L'application $\Phi^{-1} : W \rightarrow U_1 \times V$ a alors deux composantes : $\Phi^{-1} = (g_1, g_2)$ où $g_1 : W \rightarrow U_1$ et $g_2 : W \rightarrow V$ sont de classe \mathcal{C}^1 . On a alors la chaîne d'équivalences :

$$\begin{aligned} (x, y) \in U_1 \times V \quad \text{et} \quad f(x, y) = f(a, b) \\ \iff \\ (x, y) \in U_1 \times V \quad \text{et} \quad \Phi(x, y) = (x, f(a, b)) \\ \iff \\ (x, y) = (g_1(x, f(a, b)), g_2(x, f(a, b))) \quad \text{et} \quad (x, f(a, b)) \in W \end{aligned}$$

Considérons $U = \{x \in U_1 \mid (x, f(a, b)) \in W\}$. C'est un *ouvert* de U_1 (puisque W est ouvert), donc de E , contenant a et tel que $U \times V \subset U_1 \times V \subset \Omega$. De plus la chaîne d'équivalence indiquée montre que l'application $g : U \rightarrow V$ donnée par $g(x) = g_2(x, f(a, b))$ vérifie

$$\begin{aligned} (x, y) \in U \times V \quad \text{et} \quad f(x, y) = f(a, b) \\ \Updownarrow \\ x \in U \quad \text{et} \quad y = g(x). \end{aligned}$$

La définition de g montre qu'elle est de classe \mathcal{C}^1 dans U .

Enfin de la relation $f(x, g(x)) = f(a, b)$ valable pour tout $x \in U$, on tire par différentiation que, pour tout $h \in E$,

$$d_1 f(x, g(x)) \cdot h + d_2 f(x, g(x)) \circ dg(x)(h) = 0.$$

Si U est assez petit pour qu'en tout point $d_1 f(x, g(x))$ (qui dépend continûment de x) appartienne à l'ouvert $\text{Isom}(F, G)$, on en déduit

$$dg(x) = -d_2 f(x, g(x))^{-1} \circ d_1 f(x, g(x)). \quad \blacksquare$$

Insistons sur l'importance de se familiariser avec les diverses façons d'utiliser ce théorème. Il est souvent nécessaire de découvrir la bonne décomposition de E en produit de deux facteurs. Cette décomposition peut dépendre du point au voisinage duquel on souhaite faire l'étude. Par exemple prenons $U = E = \mathbb{R}^2$, $F = \mathbb{R}$ et $f(x, y) = x^2 + y^2 - R^2$, où $R > 0$, de sorte que $f(x, y) = 0$ est l'équation cartésienne du cercle (C) de centre 0 et de rayon R . Si $A = (a, b)$ est un point de ce cercle alors :

- si A n'est sur aucun des deux axes de coordonnées ($ab \neq 0$), on peut décrire les points de (C) voisins de A , soit en exprimant y comme fonction de x , soit en exprimant x comme fonction de y . L'écriture de cette fonction dépend de A . On a $y(x) = \sqrt{R^2 - x^2}$ si $b > 0$, $y(x) = -\sqrt{R^2 - x^2}$ si $b < 0$. On a aussi $x(y) = \sqrt{R^2 - y^2}$ si $a > 0$ et $x(y) = -\sqrt{R^2 - y^2}$ si $a < 0$. Dans tous les cas la fonction est de classe \mathcal{C}^1 sur l'ouvert $] -R, +R[$.

On remarque que $\frac{\partial f}{\partial y}(a, b) = 2b \neq 0$ et que $\frac{\partial f}{\partial x}(a, b) = 2a \neq 0$ et qu'on peut voir l'un ou l'autre résultat comme une application du théorème 4.12.

- Si $A = (0, \pm R)$, seule l'écriture de y comme fonction de x est possible au voisinage de A , en conformité avec le fait que $\frac{\partial f}{\partial y}(A) \neq 0$ mais que $\frac{\partial f}{\partial x}(A) = 0$. Les conclusions sont analogues (en échangeant x et y) aux points $(\pm R, 0)$.

Remarquons que *tous* les points du cercle sont susceptibles d'être étudiés de l'une ou l'autre façon.

Chapitre 5

Différentielles d'ordre supérieur

5.1 Différentielle seconde

5.1.1 Définition

Soit E et F des espaces normés, U un ouvert de E et $f : U \rightarrow F$ une application de classe \mathcal{C}^1 sur U . L'application $df : U \rightarrow L(E, F)$ définie par $x \mapsto df(x)$ est continue. Si elle est différentiable en $a \in U$, on dit que f est *deux fois différentiable* en a . Dans ce cas $d(df)(a)$ est une application linéaire continue de E dans $L(E, F)$, c'est-à-dire un élément de $L(E, L(E, F))$.

Lemme 5.1. *L'espace normé $L(E, L(E, F))$ s'identifie à l'espace normé des applications bilinéaires continues de $E \times E$ dans F , noté $L^2(E, F)$.*

Rappelons qu'une application bilinéaire $B : E \times E \rightarrow F$ est continue si et seulement s'il existe $C > 0$ tel que pour tous $h, k \in E$, $\|B(h, k)\| \leq C\|h\| \|k\|$. La *norme* de B est, par définition, le « meilleur » C ou encore :

$$\|B\| = \sup_{\|h\|=\|k\|=1} \|B(h, k)\|.$$

En dimension finie, il n'y a pas lieu de se préoccuper des normes et l'identification est la suivante. On définit l'application linéaire (à vérifier) :

$$\Phi : L(E, L(E, F)) \rightarrow L^2(E, F)$$

en associant à T l'application bilinéaire $B_T : E \times E \rightarrow F$ définie par

$$B_T(h, k) = T(h) \cdot k$$

On définit l'application linéaire (à vérifier)

$$\Psi : L^2(E, F) \rightarrow L(E, L(E, F))$$

en associant à B l'application $T_B : E \rightarrow L(E, F)$ qui envoie $h \in E$ sur l'application linéaire $T_B(h) : E \rightarrow F$ donnée par $T_B(h) \cdot k = B(h, k)$ (pour tout $k \in E$). On vérifie sans (trop de) peine que Φ et Ψ sont inverses l'une de l'autre. Si E ou F n'est plus de dimension finie il faut ajouter la vérification, fastidieuse mais sans piège, de la continuité de ces applications linéaires et la conservation de la norme (Φ et Ψ sont des isométries).

On fera systématiquement cette identification et on note $d^2f(a)$ l'application bilinéaire continue à laquelle s'identifie $d(df)(a)$. On a donc pour $h, k \in E$,

$$d^2f(a)(h, k) = \underbrace{\overbrace{d(df)(a)}^{\in L(E, L(E, F))}}_{\in L(E, F)}(h) \cdot k$$

Exemples

1. Soit $u : E \rightarrow F$ une application linéaire continue. Alors, pour tout $x \in E$, $du(x) = u$ donc du est une application constante et sa différentielle est nulle en tout point. Autrement dit u est deux fois différentiable en tout point de E et $d^2u(x) = 0$ pour tout $x \in E$.
2. Soit $B : E \times E \rightarrow F$ une application bilinéaire continue. On sait que B est différentiable en tout point $(x, y) \in E \times E$ et que, pour tout $(h, k) \in E \times E$,

$$dB(x, y) \cdot (h, k) = B(x, k) + B(h, y).$$

Donc l'application $dB : (x, y) \mapsto dB(x, y)$ est linéaire continue donc différentiable et, pour tout $(x, y) \in E \times E$,

$$d(dB)(x, y) = dB.$$

Cela s'écrit, pour (h, k) et (h', k') appartenant à $E \times E$,

$$\begin{aligned} d^2B(x, y)((h, k), (h', k')) &= d(dB)(x, y)(h, k) \cdot (h', k') \\ &= dB(h, k) \cdot (h', k') \\ &= B(h, k') + B(h', k) \end{aligned}$$

et l'application $d^2B : E \times E \rightarrow L^2(E \times E, F)$ obtenue est constante.

5.1.2 Lemme de Schwarz

Dans l'identification précédente les vecteurs h et k ne jouent pas a priori le même rôle. En fait il n'en est rien.

Proposition 5.2 (Lemme de Schwarz). *Si $f : U \rightarrow F$ est deux fois différentiable en a , l'application bilinéaire $d^2f(a)$ est symétrique.*

Preuve Soit $r > 0$ assez petit pour que $Bo(a, 2r) \subset U$ et que df y soit définie. Posons pour $\|h\| < r$ et $\|k\| < r$,

$$\varphi(h, k) = f(a + h + k) - f(a + h) - f(a + k) + f(a) - d^2f(a)(h, k)$$

et montrons que $\frac{\varphi(h, k)}{\|(h, k)\|^2} \rightarrow 0$ quand $\|(h, k)\| \rightarrow 0$. En effet la différentiabilité de df en a s'exprime par le fait que, quel que soit $\varepsilon > 0$, il existe $\alpha > 0$ (on peut imposer $\alpha < r$) tel que si $\|h\| < 2\alpha$, alors

$$\|df(a + h) - df(a) - d(df)(a) \cdot h\| < \varepsilon \|h\|$$

(il s'agit au premier membre de la norme d'un élément de $L(E, F)$).

Pour h fixé vérifiant $\|h\| < \alpha$, l'application $\varphi_h : k \mapsto \varphi(h, k)$ est définie dans la boule

$Bo(0, \alpha)$, à valeurs dans F et différentiable en tout point de cette boule. De plus (dans $L(E, F)$) on a

$$\begin{aligned} d\varphi_h(k) &= df(a+h+k) - df(a+k) - d(df)(a) \cdot h \\ &= df(a+h+k) - df(a) - d(df)(a) \cdot (h+k) - (df(a+k) - df(a) - d(df)(a) \cdot k). \end{aligned}$$

Et donc, puisque $\|h+k\| < 2\alpha$ et $\|k\| < 2\alpha$,

$$\|d\varphi_h(k)\| < \varepsilon(\|h+k\| + \|k\|) < 3\varepsilon\|(h, k)\|$$

La boule $Bo(0, \alpha)$ étant convexe, on peut appliquer le théorème des accroissements finis à φ_h sur le segment $[0, k]$. On obtient

$$\|\varphi_h(k) - \varphi_h(0)\| = \|\varphi(h, k)\| \leq 3\varepsilon\|(h, k)\| \|k\| \leq 3\varepsilon\|(h, k)\|^2$$

et le résultat annoncé suit.

Mais cela implique que l'application bilinéaire définie par

$$B(h, k) = d^2f(a)(h, k) - d^2f(a)(k, h)$$

vérifie $\frac{B(h, k)}{\|(h, k)\|^2} \rightarrow 0$ quand $\|(h, k)\| \rightarrow 0$, puisque $B(h, k) = \varphi(h, k) - \varphi(k, h)$. Or, la bilinéarité permet d'en déduire que $B = 0$. On raisonne comme dans le lemme 2.1 : si $(h_0, k_0) \neq (0, 0)$, pour tout réel $t \neq 0$,

$$\frac{B(th_0, tk_0)}{\|(th_0, tk_0)\|^2} = \frac{B(h_0, k_0)}{\|(h_0, k_0)\|^2}.$$

Mais, par hypothèse le terme de gauche tend vers 0 quand $t \rightarrow 0$. Cela oblige $B(h_0, k_0) = 0$. Or, dire que $B = 0$, c'est dire que $d^2f(a)$ est symétrique. ■

5.1.3 Dérivées partielles secondes

Lorsque E et F sont de dimension finie, le choix de bases $\mathcal{V} = (v_1, \dots, v_m)$ dans E et $\mathcal{W} = (w_1, \dots, w_n)$ dans F permet de se donner l'application bilinéaire $d^2f(a)$ en se donnant les composantes dans \mathcal{W} des m^2 vecteurs $d^2f(a)(v_i, v_j)$ ($i \leq j$). Si f est donnée par ses composantes $f = (f_1, \dots, f_n)$, cela revient à se donner pour chaque $\ell = 1, \dots, n$, les m^2 scalaires

$$d^2f_\ell(a)(v_i, v_j) = d(df_\ell)(a)(v_i) \cdot v_j$$

Mais pour tout $x \in U$, on a par définition

$$df_\ell(x)(v_i) = \frac{\partial f_\ell}{\partial x_i}(x)$$

et c'est donc la dérivée partielle par rapport à j de cette fonction de x que l'on calcule pour obtenir le scalaire

$$d^2f_\ell(a)(v_i, v_j) = \frac{\partial}{\partial x_j} \left(\frac{\partial f_\ell}{\partial x_i} \right) (a)$$

que l'on note $\boxed{\frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(a)}$ et que l'on appelle dérivée partielle seconde de f_ℓ en a par rapport à x_j et x_i .

Le théorème de Schwarz s'énonce maintenant : pour tout $1 \leq i, j \leq m$ et tout $1 \leq \ell \leq n$,

$$\boxed{\frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(a) = \frac{\partial^2 f_\ell}{\partial x_i \partial x_j}(a)}$$

Lorsque $n = 1$, c'est-à-dire si $f : U \rightarrow \mathbb{R}$, $d^2 f(a)$ est une *forme* bilinéaire symétrique dont la matrice est appelée matrice *hessienne* de f en a . C'est la matrice carrée symétrique constituée par les dérivées partielles secondes de f en a :

$$Hf(a) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(a) \right)_{1 \leq i, j \leq m}.$$

Lorsque F est de dimension n , ceci s'applique à chacune des composantes f_ℓ de f .

5.1.4 Fonctions de classe \mathcal{C}^2

Définition 5.3. On dit que f est de classe \mathcal{C}^2 dans U si f admet en tout point de U une différentielle seconde et si $d^2 f : U \rightarrow L^2(E, F)$ est continue.

Noter que l'existence de $d^2 f(x)$ pour tout $x \in U$ impose à f d'être de classe \mathcal{C}^1 dans U .

Le résultat suivant s'obtient sans peine.

Proposition 5.4. Si E et F sont de dimension finie, f est de classe \mathcal{C}^2 sur U si et seulement si chaque composante f_ℓ de f admet en tout point de U des dérivées partielles secondes et chacune des fonctions $x \mapsto \frac{\partial^2 f_\ell}{\partial x_j \partial x_i}(x)$ est continue dans U .

5.2 Fonctions de classe \mathcal{C}^p

Rien n'empêche de continuer et d'adopter la définition inductive suivante.

Définition 5.5. Soit E et F des espaces normés et U un ouvert de E .

1. On dit que $f : U \rightarrow F$ est de classe \mathcal{C}^0 si elle est continue.
2. On dit que $f : U \rightarrow F$ est de classe \mathcal{C}^p ($p \geq 1$) si elle est différentiable dans U et si $df : U \rightarrow L(E, F)$ est de classe \mathcal{C}^{p-1} .
3. On dit que $f : U \rightarrow F$ est de classe \mathcal{C}^∞ si elle est de classe \mathcal{C}^p pour tout $p \in \mathbb{N}$.

Pour concrétiser les différentielles successives, il faut généraliser l'isomorphisme entre $L(E, L(E, F))$ et $L^2(E, F)$ en définissant

$$L^p(E, F) = \{ \text{applications } p\text{-linéaires continues de } E^p = \underbrace{E \times \cdots \times E}_p \rightarrow F \}.$$

La continuité d'une application p -linéaire φ équivaut à l'existence de $C > 0$ tel que pour $(h_1, \dots, h_p) \in E^p$, $\|\varphi(h_1, \dots, h_p)\| \leq C \|h_1\| \|h_2\| \cdots \|h_p\|$. Ceci permet aussi de définir une norme sur $L^p(E, F)$. Le résultat clé est le fait que $L(E, L^{p-1}(E, F))$ s'identifie (avec sa norme) à $L^p(E, F)$.

On peut alors identifier la différentielle d'ordre p de f en a à une application p -linéaire continue de $E^p \rightarrow F$ notée $d^p f(a)$. L'analogie du lemme de Schwarz, est la

Proposition 5.6. Si f est p fois différentiable en a , alors $d^p f(a)$ est symétrique, c'est-à-dire que pour toute permutation σ de $\{1, \dots, p\}$, et tout $(h_1, \dots, h_p) \in E^p$,

$$d^p f(a)(h_{\sigma(1)}, h_{\sigma(2)}, \dots, h_{\sigma(p)}) = d^p f(a)(h_1, \dots, h_p).$$

A titre d'exemple, on peut retenir que toute application bilinéaire continue B est de classe \mathcal{C}^∞ et que $d^p B = 0$ pour tout $p \geq 3$. Comment ce résultat se généralise-t-il ?

Proposition 5.7. *Soient E, F et G des espaces normés, U un ouvert de E et V un ouvert de G . Si $f : U \rightarrow F$ et $g : V \rightarrow G$ sont de classe \mathcal{C}^p et si $f(U) \subset V$, alors $g \circ f$ est de classe \mathcal{C}^p .*

Preuve Elle se fait par récurrence sur $p \geq 1$. Si $p = 1$, c'est un résultat connu et, pour tout $x \in U$, $d(g \circ f)(x) = dg(f(x)) \circ df(x)$. Cette formule s'interprète aussi en

$$d(g \circ f) = B \circ \Phi$$

où $\Phi = (dg \circ f, df)$ et où $B : L(F, G) \times L(E, F) \rightarrow L(E, G)$ est donnée par $B(u, v) = u \circ v$. Comme B est bilinéaire continue, elle est de classe \mathcal{C}^∞ . Supposons alors le résultat établi jusqu'à l'ordre $p - 1$. Puisque dg est de classe \mathcal{C}^{p-1} et que f est de classe \mathcal{C}^p , a fortiori \mathcal{C}^{p-1} , l'hypothèse de récurrence montre déjà que les deux composantes de Φ , donc Φ , sont de classe \mathcal{C}^{p-1} . Un deuxième usage de l'hypothèse de récurrence montre alors que $B \circ \Phi$ est de classe \mathcal{C}^{p-1} . Mais cela veut dire que $d(g \circ f)$ est de classe \mathcal{C}^{p-1} et donc que $g \circ f$ est de classe \mathcal{C}^p . ■

La formule de Faa di Bruno permet d'exprimer les différentielles successives d'une fonction composée à l'aide de celles des facteurs. C'est vite très compliqué, mais on peut vérifier à titre d'exercice que

$$d^2(g \circ f)(a)(h, k) = d^2g(f(a))(df(a) \cdot h, df(a) \cdot k) + dg(f(a)) \cdot d^2f(a)(h, k)$$

formule qui généralise « visiblement » la formule

$$(g \circ f)''(x) = g''(f(x))f'(x)^2 + g'(f(x))f''(x).$$

5.3 Formules de Taylor

On connaît la formule de Taylor à une variable. Soit I un ouvert de \mathbb{R} contenant le segment $[0, 1]$ et $\varphi : I \rightarrow \mathbb{R}^n$ une fonction de classe \mathcal{C}^p ($p \geq 1$). Alors

$$\varphi(1) = \sum_{k=0}^{p-1} \frac{1}{k!} \varphi^{(k)}(0) + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} \varphi^{(p)}(t) dt$$

Théorème 5.8 (Taylor avec reste intégrale). *Soit U un ouvert d'un espace normé E et $f : U \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^p ($p \geq 1$). Si a et b sont deux points de U tels que $[a, b] \subset U$, on pose $h = b - a$, $h^k = \underbrace{(h, \dots, h)}_k$ (pour $k \in \mathbb{N}^*$) et, pour $t \in [0, 1]$,*

$\gamma(t) = a + th$. Alors

$$f(b) = f(a) + df(a) \cdot h + \frac{1}{2} d^2f(a)(h^2) + \dots + \frac{1}{(p-1)!} d^{p-1}f(a)(h^{p-1}) + \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} d^p f(\gamma(t))(h^p) dt.$$

Preuve La fonction $\varphi(t) = f(a+th) = f \circ \gamma(t)$ est de classe \mathcal{C}^p sur un ouvert de \mathbb{R} contenant $[0, 1]$. On a $\gamma'(t) = h$ et pour tout $k \geq 2$, $\gamma^{(k)}(t) = 0$. On en déduit $\varphi'(t) = df(\gamma(t)) \cdot h$, $\varphi''(t) = d^2f(\gamma(t))(h^2)$ et, plus généralement, $\varphi^{(k)}(t) = d^k f(\gamma(t))(h^k)$. La formule est alors une simple retranscription de la formule rappelée en remarquant que $\varphi(0) = f(a)$ et $\varphi(1) = f(b)$. ■

Dans le cas $n = 1$, le théorème de la moyenne permet de simplifier le reste intégral.

Théorème 5.9 (théorème de la moyenne). Soit $a, b \in \mathbb{R}$, $a < b$, $f, g : [a, b] \rightarrow \mathbb{R}$ deux fonctions continues. On suppose que g ne change pas de signe. Alors $\exists c \in]a, b[$ tel que

$$\int_a^b f(x)g(x)dx = f(c) \int_a^b g(x)dx. \quad (5.1)$$

En particulier, si $g \equiv 1$, $\exists c \in]a, b[$ tel que

$$f(c) = \frac{1}{b-a} \int_a^b f(x)dx.$$

Démonstration. Soit $I = \int_a^b g(x)dx$. Sans perte de généralité on suppose $g \geq 0$ et $I > 0$. On pose

$$m = \min_{[a,b]} f \quad \text{et} \quad M = \max_{[a,b]} f.$$

Ainsi $m \leq \frac{1}{I} \int_a^b f(x)g(x)dx \leq M$. Le théorème des valeurs intermédiaires implique l'existence de $c \in [a, b]$ tel que

$$f(c) = \frac{1}{I} \int_a^b f(x)g(x)dx,$$

ce qui équivaut à (5.1). Pour conclure, il faut montrer qu'on peut prendre $c \neq a$ et $c \neq b$, ce qu'on laisse en exercice au lecteur. \square

On obtient comme conséquence directe :

Corollaire 5.10 (Taylor–Lagrange pour $n = 1$). Soit $f : U \rightarrow \mathbb{R}$ une application de classe \mathcal{C}^p . Si a et b sont deux points de U tels que $[a, b] \subset U$, on pose $h = b - a$. Alors il $t_0 \in]0, 1[$ tel que

$$f(b) = f(a) + df(a) \cdot h + \frac{1}{2} d^2 f(a)(h^2) + \cdots + \frac{1}{(p-1)!} d^{p-1} f(a)(h^{p-1}) + \frac{1}{p!} d^p f(a + t_0 h)(h^p)$$

Démonstration. En utilisant le théorème de la moyenne, on obtient qu'il existe $t_0 \in]0, 1[$ tel que

$$\int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} d^p f(\gamma(t))(h^p) dt = d^p f(\gamma(t_0))(h^p) \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} dt = \frac{1}{p!} d^p f(\gamma(t_0))(h^p).$$

\square

Corollaire 5.11 (Taylor–Young). Soit $f : U \rightarrow \mathbb{R}^n$ une application de classe \mathcal{C}^p et $a \in U$. Alors

$$f(b) = f(a) + df(a) \cdot h + \cdots + \frac{1}{p!} d^p f(a)(h^p) + o(\|h\|^p).$$

Démonstration. Soit $r > 0$ tel que $B_0(a, r) \subset U$. Si $\|h\| < r$, on peut appliquer le théorème précédent au segment $[a, a + h]$ et il suffit de montrer que

$$\int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} d^p f(\gamma(t))(h^p) dt = \frac{1}{p!} d^p f(a)(h^p) + R(h)$$

avec $\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|^p} = 0$. Or, puisque $\int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} dt = \frac{1}{p!}$, on a

$$R(h) = \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} [d^p f(\gamma(t))(h^p) - d^p f(a)(h^p)] dt.$$

Mais puisque $x \mapsto d^p f(x)$ est continue en a , pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que si $\|x - a\| < \alpha$, $\|d^p f(x) - d^p f(a)\| < \varepsilon$. Au premier membre de cette inégalité, il s'agit de la norme de l'application p -linéaire continue $d^p f(x) - d^p f(a)$. Si elle est plus petite que ε , alors, pour tout $v \in E$, on a $\|d^p f(x)(v^p) - d^p f(a)(v^p)\| \leq \varepsilon \|v\|^p$. Si on suppose que $\|h\| < \alpha$, alors $x = \gamma(t) = a + th$ vérifie $\|x - a\| = t\|h\| < \alpha$. On en déduit

$$\|R(h)\| \leq \varepsilon \|h\|^p \int_0^1 \frac{(1-t)^{p-1}}{(p-1)!} dt = \frac{\varepsilon}{p!} \|h\|^p$$

ce qui est le résultat cherché. \square

Remarque 5.12 (Cas $U \subset \mathbb{R}^d$). Si $f : U \subset \mathbb{R}^d \rightarrow F$ est une application p -fois différentiable, alors pour tout $j_1, \dots, j_p \in \{1, \dots, d\}$

$$d^p f(x)(e_{j_1}, \dots, e_{j_p}) = \frac{\partial^p f}{\partial x_{j_1} \dots \partial x_{j_p}}(x),$$

où $\{e_1, \dots, e_d\}$ est la base canonique de \mathbb{R}^d . L'ordre de dérivation est indifférent puisque $d^p f(x)$ est symétrique. De plus, par multilinéarité, on obtient :

$$d^p f(x)(h^p) = \sum_{j_1=1}^d \sum_{j_2=1}^d \dots \sum_{j_p=1}^d h_{j_1} h_{j_2} \dots h_{j_p} \frac{\partial^p f}{\partial x_{j_1} \dots \partial x_{j_p}}(x),$$

où h_l est la composante l -ième du vecteur h .

Exemple 5.13. Soit $f(x, y) = x^2 y + x \cos y$. Alors

$$\begin{aligned} f_x &= 2xy + \cos y, & f_y &= x^2 - x \sin y, \\ f_{xy} &= f_{yx} = 2x - \sin y, & f_{xx} &= 2y, & f_{yy} &= -x \cos y, \\ f_{xxy} &= 0, & f_{xyy} &= -\cos y, & f_{xxx} &= 0, & f_{yyy} &= x \sin y. \end{aligned}$$

Ainsi la formule de Taylor–Lagrange avec $p = 3$, nous donne le développement de Taylor suivant en $(1, 0)$:

$$f(1+h, k) = 1 + h + k + 2hk - \frac{1}{2}k^2 - \frac{1}{2} \cos(ck)hk^2 + \frac{1}{6}(1+ch) \sin(ck)k^3,$$

pour certain $c \in]0, 1[$ qui dépend de h et k . Autrement dit, en prenant $x = 1 + h$ et $y = k$, le polynôme de Taylor P_2 d'ordre 2 associé à f en $(1, 0)$ est

$$P_2(x, y) = x + y + 2(x-1)y.$$

Essayons de trouver $r > 0$ tel que pour tout $(x, y) \in B((1, 0), r)$:

$$|f(x, y) - P_2(x, y)| \leq 10^{-2}.$$

On a que pour tout $(x, y) \in B((1, 0), r)$

$$\begin{aligned} |f(x, y) - P_2(x, y)| &= \left| -\frac{1}{2} \cos(cy)(x-1)y^2 + \frac{1}{6}(1+c(x-1)) \sin(cy)y^3 \right| \\ &\leq \frac{1}{2}r^3 + \frac{1}{6}(1+r)r^3 = \frac{2r^3}{3} + \frac{r^4}{6}. \end{aligned}$$

Si $r < 1$, alors $r^4 < r^3$ et il suffit donc d'imposer que $5/6r^3 < 10^{-2}$, ce qui donne

$$r < \left(\frac{3}{250} \right)^{1/3} \sim 0.229.$$