Chapitre I

Calcul différentiel dans les espaces de Banach

Le calcul différentiel dans $\mathbb{R}$ et dans $\mathbb{R}^n$ était un des sujets traités au cours 'Analyse I' (voir le livre [HW95]). Le premier chapitre du cours 'Analyse II' a comme but d'étendre ce calcul aux espaces plus généraux. Ceci nous permet non seulement d'obtenir des résultats plus généraux avec des applications intéressantes, mais nous obtenons en même temps une meilleure compréhension du calcul différentiel dans $\mathbb{R}^n$.

Après un rappel sur la différentiabilité dans $\mathbb{R}^n$, nous donnons la définition d'un espace de Banach et nous discutons les différences essentielles entre $\mathbb{R}^n$ et les espaces de Banach de dimension infinie. Nous étendons ensuite la notion d'application différentiable aux espaces de Banach, nous donnons plusieurs exemples, et nous abordons les sujets suivants: le théorème des accroissements finis, le théorème du point fixe de Banach, le théorème d'inversion locale, ainsi que le théorème des fonctions implicites.

I.1 Rappel sur la différentiabilité dans $\mathbb{R}^n$

Pour une fonction à une variable $f : (a,b) \to \mathbb{R}$, la dérivée au point x_0 est définie par

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{1.1}$$

De toute évidence, cette définition n'a pas de sens pour des fonctions à plusieurs variables.

Pour une fonction $f : U \to \mathbb{R}^m$ (U étant un ouvert de $\mathbb{R}^n$), on dit que $f(x)$ est *différentiable* en $x_0 \in U$, s'il existe une application linéaire $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$, telle que

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)\|x - x_0\|, \tag{1.2}$$

où la fonction $r : U \to \mathbb{R}^m$ (qui depend du paramètre x_0) satisfait $r(x) \to 0$ pour $x \to x_0$. On a vue dans [HW95, IV.3] que cette définition ne dépend pas de la norme choisie et que l'application linéaire est donnée par la *matrice jacobienne*

$$f'(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}, \tag{1.3}$$

où $x = (x_1, \dots, x_n)^T$ et $f(x) = (f_1(x), \dots, f_m(x))^T$. L'exemple suivant montre qu'il n'est pas toujours avantageux de représenter l'application linéaire $f'(x)$ à l'aide de la matrice jacobienne.

Exemple 1.1 Identifions l'espace $\mathbb{R}^{n\cdot n}$ avec l'ensemble des matrices carrées de dimension n, et considérons l'application $f : \mathbb{R}^{n\cdot n} \to \mathbb{R}^{n\cdot n}$ définie par $f(X) = X^2$. La propriété $(X_0 + H)^2 = X_0^2 + X_0H + HX_0 + H^2$ suggère la définition

$$f'(X_0)H = X_0H + HX_0, \qquad r(X) = (X - X_0)^2/\|X - X_0\|. \tag{1.4}$$

Pour démontrer que l'application linéaire $f'(X_0)$ de (1.4) est vraiment la dérivée de $f(X)$, il faut voir que $r(X) \to 0$ si $X \to X_0$. Ceci découle du fait que, pour la norme euclidienne dans $\mathbb{R}^{n\cdot n}$, on a $\|(X - X_0)^2\| \le \|X - X_0\|^2$.

I.2 Espaces de Banach et de Hilbert

Essayons d'étendre la définition (1.2) de la différentiabilité à une fonction $f : E \to F$. Dans les espaces E et F il faut savoir additionner, soustraire, multiplier avec un nombre réel et il faut avoir à disposition une norme. Rappelons qu'une *norme* sur un espace vectoriel E est une application $\|\cdot\| : E \to \mathbb{R}$ qui vérifie les trois propriétés:

(N1) $\|x\| \ge 0$ et $\|x\| = 0 \Leftrightarrow x = 0$,

(N2) $\|\lambda x\| = |\lambda| \cdot \|x\|$,

(N3) $\|x + y\| \le \|x\| + \|y\|$ (inégalité du triangle).

Un cas particulier important d'une norme est

$$\|x\| := \sqrt{\langle x, x\rangle}, \tag{2.1}$$

où $\langle\cdot,\cdot\rangle : E \times E \to \mathbb{R}$ est un *produit scalaire*, c.-à-d.,

(PS1) $\langle x, x\rangle \ge 0$ et $\langle x, x\rangle = 0 \Leftrightarrow x = 0$ (définie positive),

(PS2) $\langle x, y\rangle = \langle y, x\rangle$ (symétrique),

(PS3) $\langle \lambda x_1 + \mu x_2, y\rangle = \lambda\langle x_1, y\rangle + \mu\langle x_2, y\rangle$ (lináire).

Comme pour la norme euclidienne dans $\mathbb{R}^n$ on vérifie que $\|\cdot\|$, donné par (2.1), satisfait les propriétés (N1), (N2), (N3) d'une norme (Exercice 3).

Dès qu'on a donné une norme sur E, on peut définir la convergence de suites (voir [HW95, IV.1]). On dit qu'une suite $\{x_n\}$ à valeurs dans E *converge vers* $a \in E$, si

$$\forall\varepsilon > 0 \quad \exists N \ge 1 \quad \forall n \ge N \quad \|x_n - a\| < \varepsilon. \tag{2.2}$$

Elle est une *suite de Cauchy*, si

$$\forall\varepsilon > 0 \quad \exists N \ge 1 \quad \forall n, m \ge N \quad \|x_n - x_m\| < \varepsilon. \tag{2.3}$$

Définition 2.1 Un espace vectoriel E muni d'une norme $\|\cdot\|$ s'appelle *espace vectoriel normé*. Il est *complet* si chaque suite de Cauchy dans E est convergente. Un espace vectoriel normé et complet s'appelle un *espace de Banach*.

Un espace de Banach, dont la norme est donnée par un produit scalaire, s'appelle un *espace de Hilbert*.

Le premier exemple d'un espace de Hilbert est l'espace $I\!R^n$ avec comme norme la norme euclidienne. La complétude est une conséquence du Théorème IV.1.8 de [HW95]. Avec la norme $\|x\|_1 = \sum_{i=1}^n |x_i|$ (ou $\|x\|_\infty = \max_i |x_i|$) l'espace $I\!R^n$ n'est pas un espace de Hilbert (Exercice 4), mais il est un espace de Banach, car toutes les normes sont équivalentes dans $I\!R^n$ ([HW95, Théorème IV.2.4]).

Proposition 2.2 *Soit* $\mathcal{C}([0,1]) := \{f : [0,1] \to I\!R\ ;\ f$ *est continue*$\}$. *Avec la norme*

$$\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|, \tag{2.4}$$

$\mathcal{C}([0,1])$ *est un espace de Banach. Par contre, avec une des normes*

$$\|f\|_1 = \int_0^1 |f(t)|\,dt \qquad ou \qquad \|f\|_2 = \sqrt{\int_0^1 |f(t)|^2\,dt}, \tag{2.5}$$

l'espace $\mathcal{C}([0,1])$ *n'est pas complet.*

Remarque. La norme $\|f\|_2$ est définie à partir du produit scalaire $\langle f,g\rangle = \int_0^1 f(t)g(t)\,dt$, mais $\mathcal{C}([0,1])$ avec cette norme n'est pas un espace de Hilbert.

Démonstration. Comme dans $I\!R^n$, en remplaçant la somme finie par une intégrale, on vérifie que $\|f\|_1$, $\|f\|_2$ et $\|f\|_\infty$ sont des normes sur $\mathcal{C}([0,1])$.

a) Soit $\{f_n\}_{n\ge1}$ une suite de Cauchy pour la norme $\|f\|_\infty$. Alors, la propriété

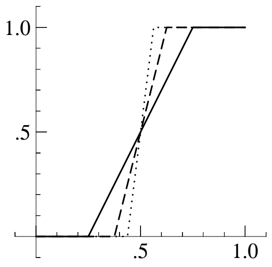
$$|f_n(t) - f_m(t)| \le \|f_n - f_m\|_\infty < \varepsilon \qquad \text{pour} \quad n,m \ge N \tag{2.6}$$

implique que $\{f_n(t)\}_{n\ge1}$ est une suite de Cauchy dans $I\!R$. Comme $I\!R$ est complet ([HW95, Théorème III.1.8]), la suite $\{f_n(t)\}_{n\ge1}$ converge vers un élément de $I\!R$, qu'on dénote par $f(t)$. Il reste donc à démontrer que la fonction $f(t)$, définie de cette manière, est continue. En passant à la limite $m \to \infty$ dans (2.6), nous obtenons

$$|f_n(t) - f(t)| \le \varepsilon \qquad \text{pour} \quad n \ge N,$$

où N dépend de $\varepsilon > 0$ mais pas de $t \in [0,1]$. Alors, la suite $\{f_n\}$ converge uniformément vers f, et la continuité de f est une conséquence de [HW95, Théorème III.4.2].

b) Considérons la suite $\{f_n\}$ dans $\mathcal{C}([0,1])$, où $f_n(t)$ est linéaire par morceaux, 0 sur $[0, 1/2 - 1/n]$ et 1 sur $[1/2 + 1/n, 1]$ (voir le dessin, $n = 4, 8, 16$). Pour $m \ge n$ on a que $\|f_n - f_m\|_1 < 1/n$ et $\|f_n - f_m\|_2 < 1/\sqrt{n}$. La suite $\{f_n\}$ est alors une suite de Cauchy. Comme la fonction limite $f(t)$ n'est pas continue sur $[0,1]$, cette suite ne converge pas dans $\mathcal{C}([0,1])$. □



Proposition 2.3 *Pour un ensemble arbitraire A considérons l'espace*

$$\mathcal{B}(A) := \{f : A \to I\!R\ ;\ f \text{ est bornée}\} \qquad avec \qquad \|f\|_\infty = \sup_{t\in A} |f(t)|. \tag{2.7}$$

$\mathcal{B}(A)$ *est un espace de Banach.*

La *démonstration* de cette proposition est presque la même que pour la partie (a) de la Proposition 2.2. Nous omettons les détails. □

Toutes les notions et définitions des paragraphes IV.1 et IV.2 de [HW95] peuvent être étendues aux espaces de Banach:

- Deux normes $\|\cdot\|_p$ et $\|\cdot\|_q$ sont *équivalentes*, s'il existe des constantes positives C_1 et C_2 pour lesquelles $C_1\|x\|_p \le \|x\|_q \le C_2\|x\|_p$ pour tout $x \in E$.
- Une *boule* de centre a et de rayon r est l'ensemble $B_r(a) = \{x \in E\ ;\ \|x-a\| < r\}$.
- Un ensemble $V \subset E$ est un *voisinage* de $a \in E$, s'il existe un $\varepsilon > 0$ tel que $B_\varepsilon(a) \subset V$.
- Un ensemble $U \subset E$ est *ouvert*, si U est voisinage de chacun de ses éléments, c.-à-d., $\forall x \in U\ \ \exists\varepsilon > 0\ \ B_\varepsilon(x) \subset U$.
- Un ensemble $V \subset E$ est *fermé*, si la limite de chaque suite convergente $\{x_n\}$ avec $x_n \in V$, est dans V.
- Un ensemble $K \subset E$ est *compact*, si chaque suite $\{x_n\}$ avec $x_n \in K$ possède une sous-suite qui converge vers un élément de K.
- Soient E et F deux espaces normés et $U \subset E$. Une fonction $f : U \to F$ est *continue* en $x_0 \in U$ si $\forall\varepsilon > 0\ \ \exists\delta > 0\ \ \forall x \in U\ :\ \|x - x_0\| < \delta \quad \|f(x) - f(x_0)\| < \varepsilon$.
- Soient E un espace normé. L'application $x \mapsto \|x\|$ est continue (car $\big|\|x\| - \|x_0\|\big| \le \|x - x_0\|$).

La majorité des résultats de [HW95, IV.1 et IV.2] reste vraie si l'on remplace $\mathbb{R}^n$ et $\mathbb{R}^m$ par des espaces de Banach. Neanmoins, il faut faire attention, car certaines propriétés sont perdues si la dimension de l'espace de Banach est infinie (comme c'est le cas pour les exemples des Propositions 2.2 et 2.3). A l'aide de contre-exemples, nous démontrons que:

- La boule fermée $\{x \in E\ ;\ \|x\| \le 1\}$ n'est pas forcément compacte.
- Deux normes dans le même espace ne sont pas toujours équivalentes.
- Le théorème de Bolzano-Weierstrass "*chaque suite bornée possède une sous-suite convergente*" n'est plus vrai.
- La caractérisation "*K compact $\Leftrightarrow$ K fermé et borné*" n'est plus valable.

Considérons l'espace $\mathcal{B}(\mathbb{R})$ de (2.7) et définissons $f_n \in \mathcal{B}(\mathbb{R})$ par $f_n(t) = 1$ si $t \in [n, n+1)$ et par $f_n(t) = 0$ si $t \notin [n, n+1)$. La suite $\{f_n\}$ est bornée (on a $\|f_n\|_\infty \le 1$) et elle satisfait $\|f_n - f_m\|_\infty = 1$ pour $n \ne m$. Par conséquent, cette suite ne peut pas avoir une sous-suite convergente. Ce contre-exemple montre en même temps que la boule fermée n'est pas compacte, que le théorème de Bolzano-Weierstrass n'est pas vrai sur $\mathcal{B}(\mathbb{R})$, et que la caractérisation ci-dessus de la compacticité n'est pas valable.

La Proposition 2.2 montre que dans l'espace $\mathcal{C}([0,1])$ les normes $\|f\|_\infty$ et $\|f\|_1$ ne sont pas équivalentes. Au cas contraire, la suite $\{f_n\}$ de la partie (b) de la démonstration serait aussi une suite de Cauchy pour la norme $\|\cdot\|_\infty$, ce qui impliquerait la convergence de $\{f_n\}$ vers une fonction continue $f \in \mathcal{C}([0,1])$.

I.3 Applications linéaires

Les applications linéaires jouent un rôle important dans la définition (1.2) de la différentiabilité. Dans $\mathbb{R}^n$, elles sont toujours continues, même uniformement continues ([HW95, Théorème IV.2.6]). Nous verrons dans ce paragraphe que ceci n'est pas toujours le cas si la dimension de l'espace vectoriel est infinie.

Proposition 3.1 *Pour une application linéaire $A : E \to F$ (E et F sont des espaces vectoriels normés), les conditions suivantes sont équivalentes:*

(a) *A est continue en tout point de E;*

(b) *A est continue à l'origine $0 \in E$;*

(c) *$\|A(x)\|$ est bornée sur la boule-unité $\{x \in E \; ; \; \|x\| \leq 1\}$.*

Démonstration. Il est clair que (a) $\Rightarrow$ (b). Montrons que (b) $\Rightarrow$ (c): la continuité de $A(x)$ à l'origine implique que pour $\varepsilon = 1$ il existe un $\delta > 0$ tel que $\|A(y)\| \leq 1$ pour $\|y\| \leq \delta$. En utilisant la linéarité de A, on obtient alors

$$\|A(x)\| = \Big\|A\Big(\frac{1}{\delta}\,\delta\,x\Big)\Big\| = \frac{1}{\delta}\,\|A(\delta x)\| \leq \frac{1}{\delta} \qquad \text{pour} \quad \|x\| \leq 1.$$

On vient donc de prouver que $\|A(x)\|$ est bornée sur la boule-unité.

Pour démontrer (c) $\Rightarrow$ (a), nous fixons arbitrairement un $x_0 \in E$ et nous supposons que $\|A(x)\|$ est bornée par M sur la boule-unité. Alors, on a pour $x \neq x_0$

$$\|A(x) - A(x_0)\| = \|A(x - x_0)\| = \|x - x_0\| \cdot \Big\|A\Big(\frac{x - x_0}{\|x - x_0\|}\Big)\Big\| \leq M\,\|x - x_0\|,$$

ce qui implique la continuité de A en x_0. □

Surtout dans les espaces de dimension infinie, nous appelons une application linéaire aussi *opérateur linéaire*. Nous écrivons souvent Ax au lieu de $A(x)$ et nous disons aussi *application bornée* pour une application continue (à cause de la propriété (c) dans la proposition précédente). Une application linéaire A de $\mathbb{R}^n$ dans $\mathbb{R}^m$ est représentée par une matrice, qu'on dénote par la même lettre A. L'expression Ax peut donc être interprétée comme la valeur de l'application A au point x, ou bien comme le produit de la matrice A avec le vecteur x.

Exemple 3.2 Il existe des opérateurs linéaires non-continus. Considérons l'espace $\mathcal{C}([0,1])$ avec la norme $\|f\|_1$ de (2.5) et les fonctions $f_n \in \mathcal{C}([0,1])$ données par $f_n(t) = n - n^2t/2$ si $t \in [0, 2/n]$ et $f_n(t) = 0$ si $t \geq 2/n$. L'application $A : \mathcal{C}([0,1]) \to \mathbb{R}$ définie par $A(f) = f(0)$ est linéaire mais non-bornée, car $|A(f_n)| = n$ et $\|f_n\|_1 = 1$ pour tout n.

Définition 3.3 Soient E et F des espaces vectoriels normés. On dénote par $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F. Pour un élément $A \in \mathcal{L}(E,F)$, on définit

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}. \tag{3.1}$$

Pour $E = F$, on écrit aussi $\mathcal{L}(E)$ à la place de $\mathcal{L}(E,E)$.

Cette définition signifie que $\|A\|$ est le plus petit nombre réel tel que

$$\|Ax\| \leq \|A\| \cdot \|x\| \qquad \text{pour tout} \;\; x \in E. \tag{3.2}$$

L'inégalité (3.2) est fondamentale pour tous les calculs avec des applications linéaires.

Proposition 3.4 *L'espace $\mathcal{L}(E,F)$ muni de (3.1) est un espace vectoriel normé. Si F est complet, alors $\mathcal{L}(E,F)$ est aussi complet.*

Démonstration. Les propriétés (N1) et (N2) d'une norme sont faciles à vérifier. Démontrons (N3): pour $A, B \in \mathcal{L}(E,F)$ nous avons

$$\|(A+B)x\| \leq \|Ax\| + \|Bx\| \leq (\|A\| + \|B\|)\|x\|.$$

En divisant cette relation par $\|x\|$ et en prenant le supremum, nous obtenons l'inégalité du triangle $\|A+B\| \leq \|A\| + \|B\|$.

La démonstration de la complétude et similaire à celle de la partie (a) de la Proposition 2.2. Soit $\{A_n\}$ une suite de Cauchy dans $\mathcal{L}(E,F)$. Pour $\|x\| \leq 1$ on obtient

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \cdot \|x\| \leq \varepsilon \cdot \|x\| \leq \varepsilon \qquad \text{pour} \;\; n, m \geq N, \tag{3.3}$$

ce qui implique que $\{A_n x\}$ est une suite de Cauchy dans F. Cet espace étant complet, la suite $\{A_n x\}$ possède une limite qu'on denote par Ax. De cette manière, on obtient une application linéaire $A : E \to F$. En passant à la limite $m \to \infty$ dans (3.3) et en divisant par $\|x\|$, on voit que $A_n - A$ (et alors aussi A) est bornée et que A est la limite de $\{A_n\}$. □

Proposition 3.5 *Soit $I \in \mathcal{L}(E)$ l'identité (c.-à-d., $Ix = x$) et considérons des applications linéaires $A \in \mathcal{L}(E,F)$ et $B \in \mathcal{L}(G,E)$. Alors, on a*

$$\|I\| = 1, \qquad \|AB\| \leq \|A\| \cdot \|B\|. \tag{3.4}$$

Démonstration. La propriété $\|I\| = 1$ est évidente. Pour démontrer l'estimation de $\|AB\|$, nous appliquons deux fois l'inégalité fondamentale (3.2),

$$\|(AB)x\| \leq \|A\| \cdot \|Bx\| \leq \|A\| \cdot \|B\| \cdot \|x\|.$$

Ensuite, nous divisons cette relation par $\|x\|$ et nous prenons le supremum sur $x \neq 0$. □

Exemple 3.6 Soit A une matrice $m \times n$, c.-à-d., $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, et notons par

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} \tag{3.5}$$

l'expression (3.1) si l'on utilise la même norme dans les deux espaces $\mathbb{R}^n$ et $\mathbb{R}^m$. Alors, on a les formules explicites

$$\begin{aligned} \|A\|_1 &= \max_{j=1,\ldots,n} \Big(\sum_{i=1}^m |a_{ij}| \Big), \qquad \|A\|_\infty = \max_{i=1,\ldots,m} \Big(\sum_{j=1}^n |a_{ij}| \Big), \\ \|A\|_2 &= \sqrt{\text{plus grande valeur propre de } A^T A}. \end{aligned}$$

Démonstration. Pour la norme $\|x\|_1$, on a

$$\|Ax\|_1 = \sum_{i=1}^m \Big| \sum_{j=1}^n a_{ij} x_j \Big| \leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| \cdot |x_j| = \sum_{j=1}^n \Big(\sum_{i=1}^m |a_{ij}| \Big) |x_j| \leq \max_{j=1,\ldots,n} \Big(\sum_{i=1}^m |a_{ij}| \Big) \cdot \|x\|_1.$$

On en déduit que $\|A\|_1 \leq \max_j (\sum_i |a_{ij}|)$. Pour montrer l'égalité, on choisit un j_0 avec $\max_j (\sum_i |a_{ij}|) = \sum_i |a_{ij_0}|$ et on pose $x = (0, \ldots, 0, 1, 0, \ldots, 0)^T$, où 1 est à la position j_0. Avec ce choix de x, on a égalité dans l'estimation ci-dessus, ce qui démontre que $\|A\|_1$ ne peut pas être plus petit que $\max_j(\sum_i |a_{ij}|)$. La formule pour la norme $\|x\|_\infty$ se démontre de la même manière.

La matrice A^TA étant symétrique et semi-définie positive ($x^TA^TAx = \|Ax\|_2 \geq 0$), il existe une matrice orthogonale U ($U^TU = I$) telle que $U^TA^TAU = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, où $\lambda_i \geq 0$ sont les valeurs propres de A^TA. On obtient alors avec la transformation $x = Uy$ et en utilisant $\|x\|_2 = \|y\|_2$ que

$$\|Ax\|_2^2 = x^TA^TAx = y^TU^TA^TAUy = \sum_{i=1}^n \lambda_i |y_i|^2 \leq \lambda_{\max} \|y\|_2^2 = \lambda_{\max} \|x\|_2^2.$$

Ceci implique $\|A\|_2 \leq \sqrt{\lambda_{\max}}$. Pour montrer l'égalité, on pose x égal au vecteur propre de A^TA qui correspond à la valeur propre $\lambda_{\max}$. □

Exemple 3.7 Pour la matrice

$$A = \begin{pmatrix} 4 & 3 \\ -2 & 5 \\ 4 & 1 \end{pmatrix} \qquad \text{on a} \qquad \begin{array}{l} \|A\|_1 = \max(10, 9) = 10 \\ \|A\|_2 = \ldots = \sqrt{(71 + \sqrt{145})/2} \approx 6.4437 \\ \|A\|_\infty = \max(7, 7, 5) = 7\,. \end{array}$$

Exemple 3.8 Soit $k : [0,1] \times [0,1] \to I\!R$ une fonction continue à deux variables. Nous considérons l'opérateur linéaire $A : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$, défini par

$$(Af)(t) = \int_0^1 k(t,s) f(s)\, ds, \tag{3.6}$$

et les normes de la Proposition 2.2. Avec la notation (3.5) nous avons

$$\begin{aligned} \|A\|_1 &= \max_{s\in[0,1]} \int_0^1 |k(t,s)|\, dt, \qquad \|A\|_\infty = \max_{t\in[0,1]} \int_0^1 |k(t,s)|\, ds, \\ \|A\|_2 &\leq \sqrt{\int_0^1 \int_0^1 |k(t,s)|^2\, ds\, dt}. \end{aligned}$$

Les formules pour $\|A\|_1$ et $\|A\|_\infty$ sont obtenues comme dans l'Exemple 3.6, l'estimation pour $\|A\|_2$ comme dans [HW95, Théorème IV.2.6].

Proposition 3.9 *Soit E un espace de Banach et supposons que l'opérateur $A \in \mathcal{L}(E)$ satisfasse $\|A\| < 1$. Alors, $I - A$ est inversible, $(I-A)^{-1}$ est continue, et on a*

$$(I - A)^{-1} = I + A + A^2 + A^3 + \ldots \tag{3.7}$$

(série géométrique ou "série de Neumann").

Démonstration. Montrons d'abord que $A_n := I + A + A^2 + \ldots + A^n$ est une suite de Cauchy dans $\mathcal{L}(E)$. En utilisant $\|A^n\| \leq \|A\|^n$, on obtient pour $m > n$ que

$$\|A_n - A_m\| = \|A^{n+1} + \ldots + A^m\| \leq \|A^{n+1}\| + \ldots + \|A^m\| \leq \|A\|^{n+1} + \ldots + \|A\|^m \leq \frac{\|A\|^{n+1}}{1 - \|A\|}$$

(observons que $\|A\|^{n+1} \to 0$ si $n \to \infty$). Comme $\mathcal{L}(E)$ est complet (voir la Proposition 3.4), la suite $\{A_n\}$ converge vers un $B \in \mathcal{L}(E)$. En passant à la limite $n \to \infty$ dans l'identité $A_n(I - A) = (I - A)A_n = I - A^{n+1}$, nous obtenons $B(I - A) = (I - A)B = I$, ce qui démontre l'inversibilité de $I - A$ et la formule (3.7). □

I.4 Différentiabilité dans les espaces normés

Nous donnons maintenant une généralisation directe de la définition (1.2).

Définition 4.1 Soient E, F deux espaces vectoriels normés et $U \subset E$ un ouvert. On dit que $f : U \to F$ est *différentiable* en $a \in U$ s'il existe une application linéaire continue $f'(a) \in \mathcal{L}(E, F)$ telle que

$$f(x) = f(a) + f'(a)(x-a) + r(x)\|x-a\|, \tag{4.1}$$

où la fonction $r : U \to F$ satisfait $r(x) \to 0$ pour $x \to a$.[1]

Par rapport à la définition usuelle dans $I\!R^n$, on demande en plus que l'application linéaire $f'(a)$ soit continue. Ceci est important, car sinon une fonction différentiable pourrait ne pas être continue (voir les exemples ci-dessous).

Si $f : E \to F$ est linéaire et continue, on a $f'(a) = f$ quelque soit $a \in E$. Une application linéaire non-bornée n'est pas différentiable (e.g., l'opérateur de l'Exemple 3.2).

Exemple 4.2 Soient $E = I\!R^3$, $F = I\!R^2$ et $f : E \to F$ donnée par

$$f(x) = \begin{pmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 \\ x_1 x_2 x_3 \end{pmatrix}.$$

Sa dérivée au point $a = (a_1, a_2, a_3)^T$ est l'application linéaire $f'(a) \in \mathcal{L}(I\!R^3, I\!R^2)$ définie par la matrice

$$f'(a) = \begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ a_2 a_3 & a_1 a_3 & a_1 a_2 \end{pmatrix}.$$

On vérifie aisement que l'application $r(x)$ de (4.1) satisfait $r(x) \to 0$ si $x \to a$. Comme toutes les normes sont équivalentes sur $I\!R^n$, il ne faut pas préciser la norme avec laquelle on travaille. De plus, une application linéaire de $I\!R^n$ dans $I\!R^m$ est automatiquement continue.

Exemple 4.3 Considérons l'espace $\mathcal{C}([0,1])$ avec la norme $\|\cdot\|_\infty$, notons ses éléments par $x(t)$ ou $a(t)$, et étudions l'application $f : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$,

$$f(x)(t) = \int_0^1 k(t,s) g(x(s))\, ds, \tag{4.2}$$

où $k : [0,1] \times [0,1] \to I\!R$ est continue et $g : I\!R \to I\!R$ est 2 fois continûment différentiable. Pour une fonction $a \in \mathcal{C}([0,1])$ donnée, cherchons la dérivée $f'(a)$. Avec $h \in \mathcal{C}([0,1])$ on a

$$\begin{aligned} \Big(f(a+h) - f(a)\Big)(t) &= \int_0^1 k(t,s)\Big(g(a(s)+h(s)) - g(a(s))\Big)\, ds \\ &= \int_0^1 k(t,s)\Big(g'(a(s)) + \frac{1}{2} g''(\alpha(s)) h(s)\Big) h(s)\, ds, \end{aligned}$$

où $\alpha(s)$ est une valeur entre $a(s)$ et $a(s) + h(s)$. Ce calcul montre que l'application linéaire $f'(a) : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$, définie par

$$\Big(f'(a)h\Big)(t) := \int_0^1 k(t,s) g'(a(s)) h(s)\, ds,$$

est un bon candidat pour la dérivée de f. Cette application est continue (voir Exemple 3.8). De plus, le reste $\|f(a+h) - f(a) - f'(a)h\|_\infty$ peut être estimé par $Const\|h\|_\infty^2$, car $g''(x)$ est bornée dans l'intervalle contenant $a(s)$ et $a(s) + h(s)$ pour tout $s \in [0,1]$ et pour toute h avec $\|h\|_\infty \le 1$. Ceci démontre que la fonction non-linéaire (4.2) est différentiable en a.

1. Il est possible d'étendre la caractérisation de Carathéodory [HW95, Lemma IV.3.5], mais elle nécessite la connaissance du Théorème de Hahn-Banach de l'analyse fonctionelle.

Définition 4.4 Une application $A : E \to F$ (où E et F sont des espaces vectoriels normés) est un *isomorphisme*, si A est linéaire, bijective, continue et si l'inverse A^{-1} est continu.[2] On dénote

$$\mathrm{GL}(E,F) := \{A \in \mathcal{L}(E,F)\ ;\ A \text{ est un isomorphisme }\}. \tag{4.3}$$

Pour $E = F$, on écrit aussi $\mathrm{GL}(E)$ à la place de $\mathrm{GL}(E,E)$.[3]

Lemme 4.5 *Si E et F sont des espaces de Banach, l'ensemble* $\mathrm{GL}(E,F)$ *est ouvert dans* $\mathcal{L}(E,F)$.

Démonstration. Pour $A \in \mathrm{GL}(E,F)$ l'application $A+H$ est un isomorphisme pour tout H avec $\|H\| \le 1/\|A^{-1}\|$. Ceci découle de $A+H = A(I+A^{-1}H)$, de $\|A^{-1}H\| \le \|A^{-1}\|\,\|H\| < 1$, et de la Proposition 3.9. □

Exemple 4.6 Considérons l'application $f(X) = X^{-1}$ de $\mathrm{GL}(E)$ dans $\mathcal{L}(E)$ (où E est un espace de Banach). Avec la série de Neumann (Proposition 3.9) nous obtenons

$$(A+H)^{-1} = (I+A^{-1}H)^{-1}A^{-1} = A^{-1} - A^{-1}HA^{-1} + \sum_{j\ge 2}(-A^{-1}H)^j A^{-1},$$

dont la partie linéaire est

$$f'(A)H = -A^{-1}HA^{-1}. \tag{4.4}$$

La continuité de $f'(A)$ est une conséquence de $\|f'(A)H\| \le \|A^{-1}\|^2\|H\|$ (en utilisant la Proposition 3.1). Le reste $\sum_{j\ge 2}(-A^{-1}H)^j A^{-1}$ peut être estimé par $\|H\|^2\|A^{-1}\|^3/(1-\|A^{-1}H\|)$. Divisé par $\|H\|$, il tend vers 0 si $\|H\| \to 0$. Ceci démontre la différentiabilité de $f(X) = X^{-1}$ ainsi que la formule (4.4).

Pour deux espaces vectoriels E_1, E_2 de normes respectives $\|\cdot\|_1, \|\cdot\|_2$, considérons le *produit cartésien* $E_1 \times E_2$. Si on le munit de la norme

$$\|(x_1,x_2)\| := \|x_1\|_1 + \|x_2\|_2, \tag{4.5}$$

on obtient un espace vectoriel normé. C'est un espace de Banach si les espaces E_1 et E_2 sont des espaces de Banach.

Exemple 4.7 Soit $B : E_1 \times E_2 \to F$ une *application bilinéaire bornée*, c.-à-d., elle satisfait

$$\|B(x_1,x_2)\| \le M\,\|x_1\|_1\|x_2\|_2 \qquad \text{pour tout} \quad (x_1,x_2) \in E_1 \times E_2. \tag{4.6}$$

Une telle application est différentiable et sa dérivée est donnée par

$$B'(a_1,a_2)(h_1,h_2) = B(a_1,h_2) + B(h_1,a_2). \tag{4.7}$$

La continuité de $B'(a_1,a_2)$ résulte de $\|B'(a_1,a_2)(h_1,h_2)\| \le M(\|a_1\|_1\|h_2\|_2 + \|h_1\|_1\|a_2\|_2) \le M\max(\|a_1\|_1, \|a_2\|_2)\|(h_1,h_2)\|$. La différentiabilité de B est alors une conséquence de

$$B(a_1+h_1, a_2+h_2) - B(a_1,a_2) - B'(a_1,a_2)(h_1,h_2) = B(h_1,h_2),$$

car $\|B(h_1,h_2)\| \le M\|h_1\|_1\|h_2\|_2 \le M(\|h_1\|_1 + \|h_2\|_2)^2/4 = M\|(h_1,h_2)\|^2/4$.

---

2. La continuité de A^{-1} est une conséquence des hypothèses sur A. La démonstration de ce résultat est difficile et utilise le "théorème sur les applications ouvertes" (analyse fonctionnelle).

3. GL est une abréviation pour "general linear group".

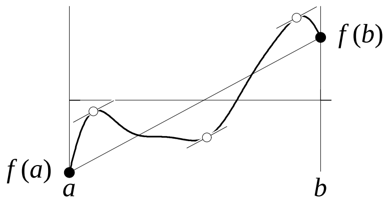
Résumons quelques règles du calcul différentiel:

- *Linéarité de la dérivée.* Soient f, g deux applications $U \to F$ (U un ouvert de E) différentiables en $a \in U$, et soit $\lambda \in I\!R$. Alors, on a
$$(f+g)'(a) = f'(a) + g'(a), \qquad (\lambda f)'(a) = \lambda f'(a). \tag{4.8}$$
- *Dérivée d'une fonction composée.* Soient E, F, G trois espaces vectoriels normés, soit U un ouvert de E, et soit V un ouvert de F. On considère deux applications, $f : U \to F$ différentiable en $a \in U$ et $g : V \to G$ différentiable en $b = f(a) \in V$ (on suppose $f(U) \subset V$). Alors, $g \circ f$ est différentiable et on a
$$(g \circ f)'(a) = g'(f(a)) \circ f'(a). \tag{4.9}$$
- *Formule de Leibniz.* Soient E, F_1, F_2, G des espaces vectoriels normés, U un ouvert de E, $f : U \to F_1$ et $g : U \to F_2$ des applications différentiables en $a \in U$, et $B : F_1 \times F_2 \to G$ une application bilinéaire bornée (c.-à-d., elle satisfait (4.6)). Alors, la fonction $p(x) := B(f(x), g(x))$ est différentiable en a et on a
$$p'(a)h = B(f(a), g'(a)h) + B(f'(a)h, g(a)). \tag{4.10}$$

La démonstration de (4.8) est triviale et celle de (4.9) est la même que pour des fonctions dans $I\!R^n$. La formule (4.10) est une conséquence de (4.7) et de (4.9), car $p = B \circ d$ où $d : U \to F_1 \times F_2$ est donnée par
$$d(x) = (f(x), g(x)) \qquad \text{avec pour dérivée} \qquad d'(x)h = (f'(x)h, g'(x)h).$$

I.5 Théorème des accroissements finis

> La terminologie des «accroissements finis» s'explique par des raisons historiques: la notion d'accroissements «finis» s'oppose à celle d'accroissements «infinitésimaux» ... (H. Cartan 1967)



Pour une fonction $f : [a, b] \to I\!R$, le théorème des accroissements finis (ou théorème de Lagrange) affirme qu'il existe un $\xi \in (a, b)$ tel que
$$f(b) - f(a) = f'(\xi)(b - a),$$
si f est continue sur $[a, b]$ et différentiable sur (a, b).

Pour une fonction $f : [a, b] \to I\!R^m$, l'affirmation sous cette forme n'est plus vraie. Un contre-exemple (déjà donné dans [HW95, IV.3]) est la fonction $f(x) = (f_1(x), f_2(x))^T$ où $f_1(x) = \cos x$, $f_2(x) = \sin x$, et $[a, b] = [0, 2\pi]$. On a que $f(a) = f(b)$, mais il n'y existe pas de ξ avec $f'(\xi) = 0$. Par contre, on voit que l'inégalité $\|f(b) - f(a)\| \le \|f'(\xi)\| \cdot \|b - a\|$ reste vraie dans cet exemple.

Dans un premier théorème, nous considérons des fonctions $f : [a, b] \to F$ où F est un espace vectoriel normé. La démonstration va être différente de celle de [HW95, Théorème IV.3.7], car nous n'avons pas de produit scalaire à disposition dans F. Ensuite, nous étendons le résultat à des fonctions $f : U \to F$, où $U \subset E$, et E est un autre espace vectoriel normé.

Lemme 5.1 *Soient $a < b$ deux réels, F un espace vectoriel normé, $f : [a, b] \to F$ et $g : [a, b] \to I\!R$ deux applications continues sur $[a, b]$ et différentiables sur (a, b). Supposons que $\|f'(t)\| \le g'(t)$ pour $a < t < b$. Alors,*
$$\|f(b) - f(a)\| \le g(b) - g(a).$$

Démonstration. Pour un $\varepsilon > 0$ donné, considérons l'inégalité

$$\|f(t) - f(a)\| \le g(t) - g(a) + \varepsilon(t-a) + \varepsilon. \tag{5.1}$$

Elle est satisfaite pour $t = a$ et aussi, par continuité de f et g, dans un voisinage de a. Donc, le suprémum

$$c := \sup \{ t \in [a,b] \ ; \ (5.1) \text{ est satisfaite} \}$$

existe et on a que $c > a$.

Montrons que l'inégalité (5.1) est satisfaite pour c. Par définition du suprémum, il existe une suite $\{t_n\}$, $t_n \to c$, telle que $\|f(t_n) - f(a)\| \le g(t_n) - g(a) + \varepsilon(t_n - a) + \varepsilon$. En passant à la limite $n \to \infty$, la continuité de f et g montre l'inégalité (5.1) pour $t = c$.

Supposons que $c < b$. La différentiabilité de f et g au point c implique qu'il existe un $\eta > 0$ tel que

$$\begin{aligned} \|f(t) - f(c)\| &\le \|f'(c)\|(t-c) + \frac{\varepsilon}{2}(t-c) \\ g(t) - g(c) &\ge g'(c)(t-c) - \frac{\varepsilon}{2}(t-c) \end{aligned}$$

pour tout $t \in [c, c+\eta]$. L'hypothèse $\|f'(c)\| \le g'(c)$ implique alors que

$$\|f(t) - f(c)\| \le g'(c)(t-c) + \frac{\varepsilon}{2}(t-c) \le g(t) - g(c) + \varepsilon(t-c).$$

Ceci, en plus de l'inégalité (5.1) pour $t = c$, implique que

$$\begin{aligned} \|f(t) - f(a)\| &\le \|f(t) - f(c)\| + \|f(c) - f(a)\| \\ &\le g(t) - g(c) + \varepsilon(t-c) + g(c) - g(a) + \varepsilon(c-a) + \varepsilon \\ &= g(t) - g(a) + \varepsilon(t-a) + \varepsilon \end{aligned}$$

pour tout $t \in [c, c+\eta]$, ce qui contredit la définition de c. Alors, on a $c = b$. Si l'on laisse tendre ε vers 0 dans (5.1) avec $t = b$, on obtient l'affirmation du théorème. □

Considérons maintenant la situation où f est définie sur un ouvert U d'un espace vectoriel normé E, qui n'est plus nécessairement $\mathbb{R}$. Pour $a, b \in E$, on appelle *segment* d'extrémités a et b l'ensemble des points $x \in E$ de la forme $x = a + t(b-a)$ avec $0 \le t \le 1$.

Théorème 5.2 (Théorème des accroissements finis) *Soient E, F des espaces vectoriels normés et $U \subset E$ un ouvert. Si $f : U \to F$ est différentiable dans U, et si le segment d'extrémités a et b est contenu dans U, on a*

$$\|f(b) - f(a)\| \le \sup_{0<t<1} \|f'(a + t(b-a))\| \cdot \|b-a\|. \tag{5.2}$$

Démonstration. L'application $h(t) := f(a + t(b-a))$ est une application différentiable de $[0,1]$ dans F et on a $h'(t) = f'(a + t(b-a))(b-a)$, donc

$$\|h'(t)\| \le \|f'(a + t(b-a))\| \, \|(b-a)\|.$$

Il suffit alors d'appliquer le Lemme 5.1, en remplaçant a par 0, b par 1, f par h, et $g(t)$ par Mt où $M = \sup_{0<s<1} \|f'(a + s(b-a))\| \, \|b-a\|$. □

On dit qu'un sous-ensemble D d'un espace vectoriel est *convexe* si, quels que soient $a, b \in D$, le segment $\{a + t(b-a)\,;\, t \in [0,1]\}$ est dans D.

Corollaire 5.3 *Soient E, F des espaces vectoriels normés, $U \subset E$ un ouvert, $f : U \to F$ différentiable dans U, et $D \subset U$ un ensemble convexe. Alors,*

$$\|f(x) - f(y)\| \le \sup_{z \in D} \|f'(z)\| \cdot \|x - y\| \qquad \textit{pour} \quad x, y \in D.$$

Démonstration. Conséquence immédiate du Théorème 5.2. □

Un ouvert U d'un espace vectoriel normé est dit *connexe*, si deux points quelconques de U peuvent être joints par une ligne brisée dans U. Rappelons qu'une *ligne brisée* est l'union fini de segments $\cup_{i=1}^{m}\{a_{i-1} + t(a_i - a_{i-1})\,;\, t \in [0,1]\}$.

Corollaire 5.4 *Soient E, F des espaces vectoriels normés, U un ouvert connexe, et $f : U \to F$ une application différentiable dans U. Si $f'(x) = 0$ pour tout $x \in U$, alors f est constante.*

Démonstration. Fixons $a \in U$ et prenons $x \in U$ arbitraire. Comme U est connexe, il existe une suite finie $a = a_0, a_1, \ldots, a_m = x$ telle que les segments $\{a_{i-1} + t(a_i - a_{i-1})\,;\, t \in [0,1]\}$ sont dans U. Le Théorème 5.2 implique alors que

$$\|f(a_i) - f(a_{i-1})\| \le \sup_{0<t<1} \|f'(a_{i-1} + t(a_i - a_{i-1}))\| \cdot \|a_i - a_{i-1}\| = 0.$$

Donc, $f(a_i) = f(a_{i-1})$ et par conséquent aussi $f(x) = f(a)$. □

I.6 Théorème du point fixe de Banach

Considérons le problème de résoudre une équation nonlinéaire dans un espace de Banach. Pour des fonctions f et g de E dans E, le problème peut être formulé de la manière suivante:

- $g(x) = 0$, on cherche un zéro de g, ou
- $f(x) = x$, on cherche un point fixe de f.

En posant $f(x) = x + g(x)$ ou $f(x) = x + Ag(x)$ avec $A \in \mathrm{GL}(E)$, on peut passer d'une formulation à l'autre. Le théorème suivant donne une condition suffisante pour l'existence et l'unicité d'une solution à ce problème.

Théorème 6.1 (Théorème du point fixe de Banach, 1922) *Soient E un espace de Banach, $D \subset E$ fermé, et $f : D \to E$ une application satisfaisant*

(a) *$f(D) \subset D$,*

(b) *f est une contraction sur D, c.-à-d., il existe un $\alpha < 1$ tel que*

$$\|f(x) - f(y)\| \le \alpha\,\|x - y\| \qquad \textit{pour} \quad x, y \in D. \tag{6.1}$$

Alors, f possède un unique point fixe dans D.

Remarques. La condition (6.1) implique que f est uniformément continue sur D. Inversément, si f est différentiable dans un voisinage de D, si D est convexe, et si $\sup_{x\in D} \|f'(x)\| < 1$, alors f est une contraction. Ceci est une conséquence du Corollaire 5.3.

Démonstration. Unicité. Soient x et y deux points fixes, c.-à-d., $f(x) = x$ et $f(y) = y$. La contractivité implique que

$$\|x - y\| = \|f(x) - f(y)\| \leq \alpha \|x - y\|$$

avec $\alpha < 1$, ce qui est possible seulement si $x = y$.

Existence. Prenons $x_0 \in D$ arbitraire et considérons l'itération $x_{n+1} = f(x_n)$. L'hypothèse $f(D) \subset D$ implique que $x_n \in D$ pour tout $n \geq 0$. Montrons que $\{x_n\}$ est une suite de Cauchy. On a que $\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq \alpha\|x_n - x_{n-1}\|$ et, en appliquant cette inégalité itérativement, on obtient que

$$\|x_{n+1} - x_n\| \leq \alpha^n \|x_1 - x_0\|.$$

Pour $m \geq n$ on en déduit que

$$\begin{aligned} \|x_m - x_n\| &\leq \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \ldots + \|x_{n+1} - x_n\| \\ &\leq \Big(\alpha^{m-1} + \alpha^{m-2} + \ldots + \alpha^n\Big)\|x_1 - x_0\| \leq \frac{\alpha^n}{1-\alpha} \|x_1 - x_0\|. \end{aligned}$$

Alors, $\{x_n\}$ est une suite de Cauchy (observons que $\alpha^n \to 0$). Comme l'espace E est complet, cette suite converge vers un $x \in E$. La limite x est dans D, car D est fermé. En prenant la limite $n \to \infty$ dans $x_{n+1} = f(x_n)$ et en utilisant la continuité de f, on obtient $x = f(x)$, c.-à-d., x est un point fixe de f. □

La démonstration du théorème du point fixe de Banach est constructive et nous conduit à l'algorithme suivant.

Méthode des approximations successives Pour résoudre un problème $x = f(x)$ dans un espace de Banach, cette méthode est définie par:

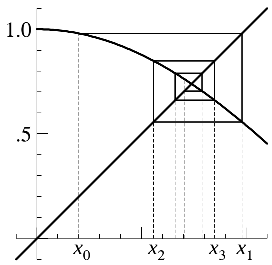
- on choisit x_0 arbitrairement,
- on applique l'itération $x_{n+1} = f(x_n)$.

Sous les hypothèses du théorème, cet algorithme converge vers la solution unique du problème. Souvent, les hypothèses sont difficiles à vérifier, mais on peut quand même appliquer cet algorithme. Si on a convergence, on est sûr d'avoir trouvé une solution (si f est continue). Elle n'est pas nécessairement unique.

Exemple 6.2 Prenons la fonction $f(x) = \cos x$ sur $D = [0, 1]$. Cette fonction est une contraction, car $|f'(x)| = |\sin x| \leq \sin 1 < 1$ pour $x \in D$. Un autre exemple est la fonction $f(x) = e^x/4$ sur $D = [0, 1.1]$. Pour cette fonction on a $|f'(x)| = e^x/4 \leq e^{1.1}/4 < 1$ sur D. Les itérations sont illustrées dans la Fig. 6.1.

Exemple 6.3 Considérons le système de deux équations nonlinéaires à deux inconnues

$$x = 2 + (x^2 + y^2)/20, \qquad y = 1 - xy^3/10.$$



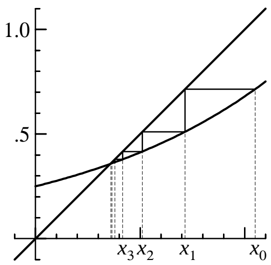


FIG. 6.1 – *Méthode des approximations successives pour* $f(x) = \cos x$ *et* $f(x) = e^x/4$

Sans vérifier les hypothèses du théorème du point fixe de Banach, appliquons la méthode itérative avec comme valeurs initiales $x_0 = 2, y_0 = 1$. On obtient alors

$$\begin{aligned} x_1 &= 2.2500000, & y_1 &= 0.8000000, \\ x_2 &= 2.2851250, & y_2 &= 0.8848000, \\ x_3 &= 2.3002333, & y_3 &= 0.8417129, \\ x_4 &= 2.2999777, & y_4 &= 0.8628284, \end{aligned}$$

et on observe convergence vers la solution $x = 2.3014505$, $y = 0.8557662$ du système.

Exemple 6.4 Considérons l'espace de Banach $\mathcal{C}([0,1])$ avec la norme $\|\cdot\|_\infty$, et le problème à point fixe

$$y(t) = f(t) + \lambda \int_0^1 k(t,s)y(s)\,ds, \tag{6.2}$$

où f et k sont données et λ est tel que $|\lambda| \cdot \max_{0\le t\le 1} \int_0^1 |k(t,s)|\,ds < 1$ (voir l'Exemple 3.8). La méthode des approximations successives s'écrit comme suit: on prend une approximation initiale, par exemple $y_0(t) = f(t)$, et on itère selon

$$y_{n+1}(t) = f(t) + \lambda \int_0^1 k(t,s)y_n(s)\,ds.$$

La suite $\{y_n(t)\}$ converge vers la solution unique de (6.2).

Considérons maintenant le problème de résoudre $f(x) = y$ pour un vecteur donné y, et cherchons des résultats sur l'existence et l'unicité (locale) d'une solution, c.-à-d., nous cherchons des résultats sur la bijectivité de la fonction f. De plus, on aimerait savoir si la solution dépend continûment du paramètre y. Ceci est résumé dans la définition suivante.

Définition 6.5 Soient E, F des espaces vectoriels normés, $U \subset E$ et $V \subset F$. Une application $f : U \to V$ est un *homéomorphisme* de U sur V, si f est bijective, continue, et si l'inverse $f^{-1} : V \to U$ est continue.

Méthode de Newton simplifiée Supposons $f(a) = b$ et considérons le problème $f(x) = y$ pour un y donné proche de b. Soit x_0 une approximation de la solution cherchée (par exemple $x_0 = a$). L'idée est de linéariser le problème autour de x_0 et de résoudre le problème linéarisé pour obtenir une meilleure approximation x_1:

$$f(x_0) + f'(x_0)(x_1 - x_0) = y \qquad \text{ou} \qquad x_1 = x_0 - f'(x_0)^{-1}(f(x_0) - y).$$

En itérant cette procédure, on obtient la *méthode de Newton*

$$x_{n+1} = x_n - f'(x_n)^{-1}(f(x_n) - y)$$

(voir Fig. 6.2). Elle peut être intérpretée comme la méthode des approximations successives appliquée à $x = g(x)$ où $g(x) = x - f'(x)^{-1}(f(x) - y)$.

En pratique, on remplace souvent la dérivée $f'(x_n)$ par une approximation A qui ne dépend pas de x_n. Ceci simplifie et le calcul numérique et la théorie (étude de convergence, voir le théorème suivant). Nous considérons alors la *méthode de Newton simplifiée*, qui s'écrit comme $x_{n+1} = g(x_n)$ où

$$g(x) := x - A^{-1}(f(x) - y). \tag{6.3}$$

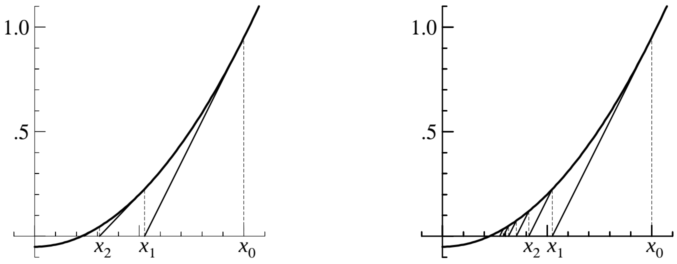


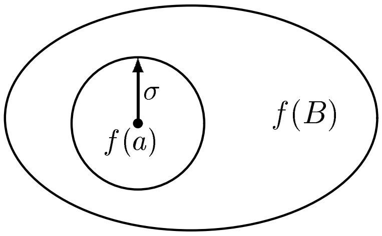
FIG. 6.2 – *Méthode de Newton (gauche) et méthode de Newton simplifiée (droite)*

Proposition 6.6 *Soient E et F des espaces de Banach, $B := \{x \in E\,;\, \|x - a\| \le \rho\}$, $A : E \to F$ un isomorphisme, et supposons que $f : B \to F$ satisfasse*

$$\|x - z - A^{-1}(f(x) - f(z))\| \le \alpha\, \|x - z\| \qquad \textit{pour} \quad x, z \in B \tag{6.4}$$

avec $\alpha < 1$. Alors, on a

- *f est un homéomorphisme de B sur $f(B)$,*
- *avec $\sigma := \rho(1-\alpha)/\|A^{-1}\|$ on a $\{y \in F\,;\, \|y - f(a)\| \le \sigma\} \subset f(B)$,*
- *pour $y \in F$ satisfaisant $\|y - f(a)\| \le \sigma$, l'itération $x_0 = a$, $x_{n+1} = g(x_n)$ (méthode de Newton simplifiée) converge vers la solution de $f(x) = y$.*



Démonstration. a) L'application $g(x)$ de (6.3) est une contraction sur B par notre hypothèse. La continuité de $f(x)$ est donc une conséquence de

$$\|f(x) - f(z)\| = \|A(x - z) - A(g(x) - g(z))\| \le \|A\|(1 + \alpha)\|x - z\|.$$

De la définition de $g(x)$, nous déduisons $x - z = g(x) - g(z) + A^{-1}(f(x) - f(z))$ et nous obtenons l'estimation $\|x - z\| \le \alpha\, \|x - z\| + \|A^{-1}\|\, \|f(x) - f(z)\|$. On a donc

$$\|x - z\| \le \frac{\|A^{-1}\|}{1 - \alpha}\, \|f(x) - f(z)\| \qquad \text{pour} \quad x, z \in B. \tag{6.5}$$

La fonction $f : B \to f(B)$ est surjective par définition. La propriété (6.5) démontre qu'elle est injective et donc aussi bijective. En posant $u = f(x)$ et $v = f(z)$ dans (6.5), on obtient

$$\|f^{-1}(u) - f^{-1}(v)\| \le \frac{\|A^{-1}\|}{1 - \alpha}\, \|u - v\|, \tag{6.6}$$

ce qui démontre la continuité de f^{-1} sur $f(B)$.

b) Considérons maintenant un $y \in F$ avec $\|y - f(a)\| \le \sigma$. Montrons que $g(B) \subset B$. Ceci découle de $g(x) - a = g(x) - g(a) + g(a) - a = g(x) - g(a) - A^{-1}(f(a) - y)$, car

$$\|g(x) - a\| \le \alpha\, \|x - a\| + \|A^{-1}\| \cdot \|f(a) - y\| \le \alpha\rho + \|A^{-1}\|\sigma = \rho$$

pour $x \in B$. Comme B est fermé, on peut appliquer le théorème du point fixe de Banach. Pour un tel y il existe donc un $x \in B$ satisfaisant $g(x) = x$, c.-à-d., $f(x) = y$. $\square$

Si $f(x)$ est différentiable en a (le centre de la boule B), il est naturel de choisir $A = f'(a)$ pourvu que cette application soit un isomorphisme. Mais la différentiabilité au point a ne garantit pas que l'estimation (6.4) soit vérifiée avec un $\alpha < 1$. Nous considérons alors une condition plus forte que la différentiabilité.

Définition 6.7 Soient E, F des espaces vectoriels normés, $U \subset E$ un ouvert et $a \in U$. On dit que $f : U \to F$ est *strictement différentiable* en a s'il existe une application linéaire continue $A : E \to F$ telle que

$$f(x) - f(z) = A(x - z) + r(x, z)\|x - z\|, \tag{6.7}$$

où la fonction $r : U \times U \to F$ satisfait $r(x, z) \to 0$ si $(x, z) \to (a, a)$.

En posant $z = a$ dans (6.7) on retrouve (4.1) avec $A = f'(a)$. Ainsi "strictement différentiable en a" entraîne "différentiable en a", ce qui est heureux pour la terminologie choisie. On voit aussi de (6.7) qu'une fonction strictement différentiable en a est continue dans tout un voisinage de a.

Exemple 6.8 Il existe des fonctions qui sont différentiables en a, mais pas strictement différentiables en a. On peut prendre une fonction qui est différentiable en a, mais qui n'est continue dans aucun voisinage de a (exemple 5 de [HW95, III.6]). Un autre exemple est la fonction $f(x) = x^2 \cos(1/x)$ en $a = 0$. Les suites $x_n = (2n\pi)^{-1}$ et $z_n = ((2n+1)\pi)^{-1}$ convergent vers 0, mais $(f(x_n) - f(z_n))/(x_n - z_n) = (x_n^2 + z_n^2)/(x_n - z_n)$ ne converge pas vers $f'(0) = 0$ si $n \to \infty$.

Proposition 6.9 *Soient E, F des espaces vectoriels normés, $U \subset E$ un ouvert et $a \in U$. Si $f : U \to F$ est différentiable dans U et si l'application $f' : U \to \mathcal{L}(E, F)$ est continue au point a, alors f est strictement différentiable au point a.*

Démonstration. Appliquons le théorème des accroissements finis à $g(x) := f(x) - f'(a)x$. Comme la dérivée de $g(x)$ est $g'(x) = f'(x) - f'(a)$, on a

$$\|f(x) - f(z) - f'(a)(x - z)\| \le \sup_{0<t<1} \|f'(z + t(x - z)) - f'(a)\| \cdot \|x - z\|.$$

La continuité de $f'(x)$ au point a implique que l'expression $\sup_{0<t<1} \|f'(z+t(x-z)) - f'(a)\|$ tend vers 0 si $(x, z) \to (a, a)$. □

Exemple 6.10 Reprenons la fonction $f(X) = X^{-1}$ de l'Exemple 4.6, pour laquelle la dérivée est $f'(X)H = -X^{-1}HX^{-1}$. Montrons que cette fonction est strictement différentiable dans $\mathrm{GL}(E)$. Comme $f(X)$ est continue en $A \in \mathrm{GL}(E)$, on a que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que $\|X^{-1} - A^{-1}\| < \varepsilon$ si $\|X - A\| < \delta$. En écrivant

$$f'(X)H - f'(A)H = -X^{-1}HX^{-1} + A^{-1}HA^{-1} = -X^{-1}H(X^{-1} - A^{-1}) - (X^{-1} - A^{-1})HA^{-1},$$

on obtient $\|(f'(X) - f'(A))H\| \le \varepsilon(\|X^{-1}\| + \|A^{-1}\|)\|H\|$. On en déduit $\|f'(X) - f'(A)\| \le \varepsilon(\|X^{-1}\| + \|A^{-1}\|)$ et donc la continuité de $f'(X)$ au point A. Ceci nous permet d'appliquer la Proposition 6.9.

Pour toutes les fonctions des exemples du paragraphe I.4, on peut démontrer sans difficulté que $f'(x)$, considérée comme fonction de x, est continue. Alors, la Proposition 6.9 implique qu'elles sont strictement différentiables.

I.7 Théorème d'inversion locale

Nous poursuivons l'étude de la résolution d'équations nonlinéaires $f(x) = 0$ (ou $f(x) = y$ pour un y donné) dans un espace de Banach.

Exemple 7.1 Considérons le système nonlinéaire

$$x_1^2 + x_2^3 - 3x_1 = y_1, \qquad x_1^4 - x_1^2 x_2 + 2 = y_2. \tag{7.1}$$

Pour $(x_1, x_2) = (1, 2)$ on obtient $(y_1, y_2) = (6, 1)$. La question à laquelle on aimerait trouver une réponse est la suivante: *pour* (y_1, y_2) *proche de* $(6, 1)$*, existe-t-il une solution* (x_1, x_2) *de (7.1) qui est proche de* $(1, 2)$*? Est-elle localement unique?*

Nous disons qu'une application $f : U \to F$ (où E, F sont des espaces vectoriels normés et $U \subset E$ est un ouvert) est un *homéomorphisme local près de* $a \in U$, s'il existe un voisinage ouvert $U' \subset U$ de a et un voisinage ouvert V' de $f(a)$ tels que la restriction $f|_{U'}$ est un homéomorphisme de U' sur V' (voir la Défintion 6.5).

Lemme 7.2 *Soient E, F des espaces de Banach, $U \subset E$ un ouvert et $a \in U$. Si $f : U \to F$ est strictement différentiable en a et si $f'(a)$ est un isomorphisme de E sur F, alors f est un homéomorphisme local près de a. De plus, l'application inverse f^{-1} est strictement différentiable en $b = f(a)$ et on a*

$$(f^{-1})'(b) = f'(a)^{-1}. \tag{7.2}$$

Démonstration. L'idée est d'appliquer la Proposition 6.6 avec $A = f'(a)$. La fonction f est strictement différentiable en a. Ceci signifie que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$\|f(x) - f(z) - f'(a)(x - z)\| \le \varepsilon \,\|x - z\| \qquad \text{pour} \quad x, z \in B, \tag{7.3}$$

où $B = \{x \in E\,;\, \|x - a\| \le \delta\}$. Alors, on a aussi

$$\|x - z - f'(a)^{-1}(f(x) - f(z))\| \le \varepsilon \,\|f'(a)^{-1}\| \cdot \|x - z\| \qquad \text{pour} \quad x, z \in B. \tag{7.4}$$

On fixe $\varepsilon > 0$ tel que $\varepsilon\|f'(a)^{-1}\| \le 1/2$. Pour le δ correspondant, f est un homéomorphisme de B sur $f(B)$ (Proposition 6.6) et $f(B)$ contient l'ouvert $V' = \{y \in F\,;\, \|y - b\| < \sigma\}$ où $\sigma := \delta/(2\|f'(a)^{-1}\|)$. Donc, f est aussi un homéomorphisme de l'ouvert $U' := f^{-1}(V')$ sur V'.

Pour montrer que f^{-1} est strictement différentiable en $b = f(a)$, nous posons $x = f^{-1}(u)$ et $z = f^{-1}(v)$ dans (7.4). Ceci donne pour $u, v \in V'$ que

$$\|f^{-1}(u) - f^{-1}(v) - f'(a)^{-1}(u - v)\| \le \varepsilon\|f'(a)^{-1}\| \cdot \|f^{-1}(u) - f^{-1}(v)\| \le 2\varepsilon\|f'(a)^{-1}\|^2\|u - v\|$$

(ici on a utilisé l'estimation (6.6) avec $A = f'(a)$ et $\alpha = 1/2$). Donc f^{-1} est strictement différentiable en b et la dérivée est donnée par (7.2). □

Exemple 7.3 Pour le problème de l'Exemple 7.1 nous avons

$$f'(a) = \begin{pmatrix} 2a_1 - 3 & 3a_2^2 \\ 4a_1^3 - 2a_1a_2 & -a_1^2 \end{pmatrix} = \begin{pmatrix} -1 & 12 \\ 0 & -1 \end{pmatrix}.$$

Pour la norme euclidienne, on calcule $\|f'(a)^{-1}\| \approx 12.1$. Si l'on pose $\varepsilon = 0.04$, on peut trouver un $\delta > 0$ tel que (7.3) est vérifié. Pour tout $y \in \mathbb{R}^2$ avec $\|y - b\| \le 2\delta/12.1$ le système (7.1) possède alors une solution x qui satisfait $\|x - a\| \le \delta$. Dans cette boule de rayon δ, il n'y a pas d'autre solution. De plus, la démonstration du Lemme 7.2 montre que la méthode de Newton simplifiée converge vers cette solution.

Le but suivant est d'étudier si la solution de $f(x) = y$ dépend de manière différentiable de y. Plus précisement, nous étudions les conditions sous lesquelles la fonction inverse est différentiable dans tout un voisinage de $b = f(a)$.

Définition 7.4 Soient E, F des espaces vectoriels normés, $U \subset E$ et $V \subset F$ des ouverts.

- $f : U \to F$ est *continûment différentiable* (ou *de classe* $\mathcal{C}^1$) sur U, si f est différentiable en tout point de U et si $f' : U \to \mathcal{L}(E, F)$ est continue.
- $f : U \to V$ est un *difféomorphisme* (de classe $\mathcal{C}^1$) de U sur V, si f est bijective, continûment différentiable, et si l'inverse $f^{-1} : V \to U$ est continûment différentiable.
- $f : U \to F$ est un *difféomorphisme local près de* $a \in U$, s'il existe un voisinage ouvert $U' \subset U$ de a et un voisinage ouvert V' de $f(a)$ tels que f est un difféomorphisme de U' sur V'.

Remarquons qu'un homéomorphisme (voir la Définition 6.5) qui est continûment différentiable n'est pas toujours un difféomorphisme. La fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3$ nous sert comme contre-exemple, car l'inverse $f^{-1}(y) = \sqrt[3]{y}$ est continu mais pas différentiable à l'origine.

Une fonction qui est continûment différentiable sur U est strictement différentiable sur U. Ceci est une conséquence de la Proposition 6.9.

Théorème 7.5 (Théorème d'inversion locale) *Soient E, F des espaces de Banach, $U \subset E$ un ouvert et $a \in U$. Une application $f : U \to F$ de classe $\mathcal{C}^1$ est un difféomorphisme local près de a, si et seulement si $f'(a)$ est un isomorphisme de E sur F. De plus, on a*

$$(f^{-1})'(y) = f'(x)^{-1} \qquad \textit{pour } y = f(x) \textit{ dans un voisinage de } b = f(a).$$

Démonstration. "$\Rightarrow$": Si $f : U \to F$ est un difféomorphisme local près de a, on peut dériver l'identité $f^{-1}(f(x)) = x$. Ceci donne

$$(f^{-1})'(y)\, f'(x) = I \qquad \text{avec} \quad y = f(x) \tag{7.5}$$

dans un voisinage de a. Par conséquent, $f'(a)$ est inversible. L'inverse $f'(a)^{-1} = (f^{-1})'(b)$ est une application bornée car f^{-1} est supposée différentiable en b.

"$\Leftarrow$": Le Lemme 7.2 implique que f est un homéomorphisme local près de a. Il reste à démontrer que f^{-1} est continûment différentiable dans un voisinage de $b = f(a)$. Comme $f'(x)$ est proche de $f'(a)$ (continuité de $f' : U \to \mathcal{L}(E, F)$) et $\mathrm{GL}(E, F)$ est ouvert (voir le Lemme 4.5), $f'(x)$ est un isomorphisme pour tout x dans un voisinage U' de a. On peut donc appliquer le Lemme 7.2 à chaque point $x \in U'$, ce qui implique la différentiabilité de f^{-1} dans $V' := f(U')$. En dérivant l'identité $f^{-1}(f(x)) = x$, on obtient (7.5) et donc aussi $(f^{-1})'(y) = f'(f^{-1}(y))^{-1}$ pour $y \in V'$. La fonction $(f^{-1})' : V' \to \mathcal{L}(F, E)$, étant la composition des applications continues f^{-1}, f' et $(\cdot)^{-1}$, est par conséquent continue. $\square$

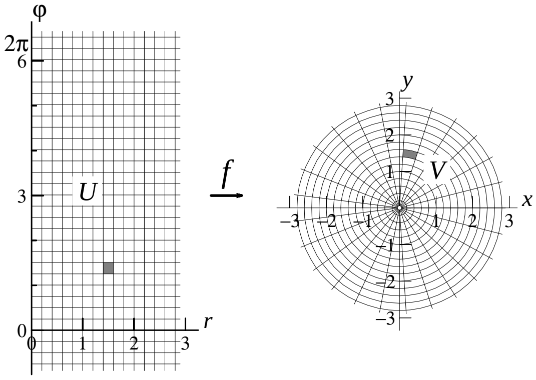
Un résultat global est le suivant.

Corollaire 7.6 *Soient E, F des espaces de Banach, $U \subset E$ un ouvert et $f : U \to F$ continûment différentiable sur U. Alors, f est un difféomorphisme de U sur $f(U)$, si et seulement si (i) f est injective sur U et (ii) $f'(x)$ est un isomorphisme pour tout $x \in U$.*

Démonstration. Ce résultat est une conséquence du Théorème 7.5, car la différentiabilité est une propriété locale. □

Exemple 7.7 (Coordonnées polaires) Soient $U = \{(r,\varphi) \in \mathbb{R}^2\,;\, r > 0\}$, $V = \mathbb{R}^2 \setminus \{(0,0)\}$ et $f : U \to V$ donnée par

$$(r,\varphi) \mapsto (r\cos\varphi, r\sin\varphi).$$



Cette application est un difféomorphisme local près de chaque point de U, car

$$f'(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}$$

et $\det(f'(r,\varphi)) = r > 0$. Elle n'est pas injective (on a $f(r,\varphi+2\pi) = f(r,\varphi)$). Donc, elle n'est pas un difféomorphisme de U sur V. Si l'on restreint les ensembles U et V à

$$U_0 = \{(r,\varphi)\,;\, r > 0, -\pi < \varphi < \pi\}, \qquad V_0 = \mathbb{R}^2 \setminus \{(x,0)\,;\, x \le 0\},$$

f devient un difféomorphisme de U_0 sur V_0.

Exemple 7.8 (Coordonnées sphèriques) L'application

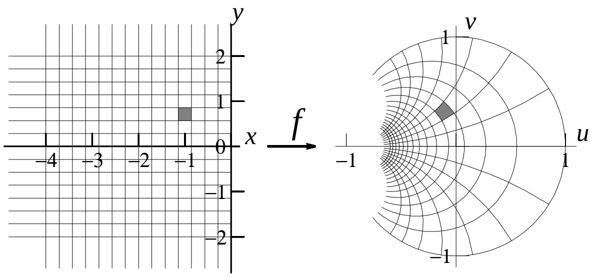
$$(r,\varphi,\theta) \mapsto (r\cos\varphi\sin\theta, r\sin\varphi\sin\theta, r\cos\theta)$$

est un difféomorphisme de $U = \{(r,\varphi,\theta)\,;\, r > 0, -\pi < \varphi < \pi, 0 < \theta < \pi\}$ sur l'ensemble $V = \mathbb{R}^3 \setminus \{(x,0,z)\,;\, x \le 0, z \in \mathbb{R}\}$, car le déterminant de la matrice Jacobienne est $-r^2\sin\theta \ne 0$.

Exemple 7.9 (Transformation de Cayley) L'application

$$(x,y) \mapsto \Big(\frac{1-x^2-y^2}{(1-x)^2+y^2}\,,\, \frac{2y}{(1-x)^2+y^2}\Big) \tag{7.6}$$

est un difféomorphisme du demi-plan gauche $\{(x,y)\,;\, x < 0\}$ sur le disque ouvert $\{(u,v)\,;\, u^2+v^2 < 1\}$. Pour démontrer ceci, nous identifions $\mathbb{R}^2$ avec le plan complexe (en posant $z = x+iy$ et $w = u+iv$) et nous remarquons que l'application (7.6) est équivalente à



$$w = \frac{1+z}{1-z} \qquad \text{avec inverse} \qquad z = \frac{w-1}{w+1}.$$

Donc, l'application (7.6) est bijective et elle est continûment différentiable (comme fonction rationnelle), ainsi que son inverse.

I.8 Théorème des fonctions implicites

Dans le paragraphe précédent nous avons considéré le problème de résoudre $f(y) = x$ et nous avons trouvé des conditions suffisantes permettant d'écrire la solution sous la forme $y = g(x)$ (attention: nous avons inversé les rôles de x et y). Le but de ce paragraphe est d'étendre ce résultat au problème

$$f(x, y) = 0 \tag{8.1}$$

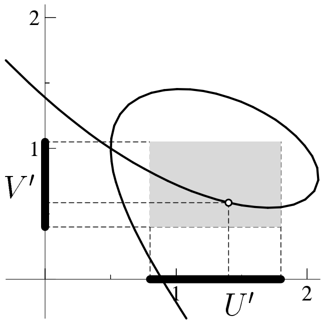
où x, y et $f(x, y)$ sont dans des espaces de Banach. On cherche à savoir si l'équation (8.1) peut être résolue pour obtenir $y = g(x)$ tel que (du moins localement)

$$f(x, y) = 0 \qquad \Longleftrightarrow \qquad y = g(x). \tag{8.2}$$

Exemple 8.1 Le cas où x, y et $f(x, y)$ sont dans $I\!R$ a été traité au cours d'Analyse I (voir [HW95, Théorème IV.3.8]). Considérons par exemple la fonction

$$f(x, y) = 16x^3 - 84x^2 + 162x - 89 + 27y^3 + 54xy^2 - 108y^2 + 36x^2y - 180xy + 162y.$$

Pour un point (a, b) satisfaisant $f(a, b) = 0$ et $\frac{\partial f}{\partial y}(a, b) \neq 0$, il existe des voisinages U' de a, V' de b et une fonction différentiable $g : U' \to V'$ tels que (8.2) est vraie pour $(x, y) \in U' \times V'$. Les points sur la courbe ayant une tangente verticale ou horizontale peuvent être trouvés par la condition $\frac{\partial f}{\partial y}(a, b) = 0$ et $\frac{\partial f}{\partial x}(a, b) = 0$ respectivement. Au point du croisement on a nécessairement $f(a, b) = 0$, $\frac{\partial f}{\partial y}(a, b) = 0$ et $\frac{\partial f}{\partial x}(a, b) = 0$ (3 conditions pour 2 inconnues).



Exemple 8.2 Une fonction $f(x_1, x_2, y) = 0$ représente une surface dans $I\!R^3$. Quand peut-on écrire cette équation sous la forme $y = g(x_1, x_2)$?

Le système de deux fonctions

$$f_1(x, y_1, y_2) = 0, \qquad f_2(x, y_1, y_2) = 0$$

représente l'intersection de deux surfaces dans $I\!R^3$. Sous quelle condition peut-on écrire cette intersection sous la forme $y_1 = g_1(x), y_2 = g_2(x)$, ce qui représenterait une courbe dans $I\!R^3$?

Pour une fonction $f : U \to G$, $U \subset E \times F$ (où E, F, G sont des espaces de Banach) on définit la *dérivée partielle* $\frac{\partial f}{\partial y}(a, b)$ comme la dérivée de l'application $h(y) = f(a, y)$, où a est considéré comme un paramètre fixé, c.-à-d., $\frac{\partial f}{\partial y}(a, b) := h'(b)$. Si f est de classe $\mathcal{C}^1$, h est aussi de classe $\mathcal{C}^1$, car $h = f \circ \lambda$ où l'injection $\lambda : F \to E \times F$, définie par $\lambda(y) = (a, y)$, est continûment différentiable.

Dans le cas où $x \in I\!R^m$, $y \in I\!R^n$ et $f : I\!R^m \times I\!R^n \to I\!R^n$ est donnée par

$$f(x, y) = \begin{pmatrix} f_1(x_1, \ldots, x_m, y_1, \ldots, y_n) \\ \vdots \\ f_n(x_1, \ldots, x_m, y_1, \ldots, y_n) \end{pmatrix} \quad \text{on a} \quad \frac{\partial f}{\partial y}(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \cdots & \frac{\partial f_1}{\partial y_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1}(x, y) & \cdots & \frac{\partial f_n}{\partial y_n}(x, y) \end{pmatrix}.$$

Théorème 8.3 (Théorème des fonctions implicites) *Soient E, F et G des espaces de Banach, $U \subset E$ et $V \subset F$ des ouverts et $f : U \times V \to G$ une application de classe $\mathcal{C}^1$. Supposons qu'en $(a, b) \in U \times V$*

$$f(a,b) = 0 \qquad \textit{et} \qquad \frac{\partial f}{\partial y}(a,b) \ \textit{ est un isomorphisme de } F \textit{ sur } G.$$

(i) Il existe alors un voisinage U' de a, un voisinage V' de b et une application unique $g : U' \to V'$ tels que $f(x, g(x)) = 0$ pour $x \in U'$.

(ii) L'application $g : U' \to V'$ est de classe $\mathcal{C}^1$ et on a

$$g'(x) = -\Big(\frac{\partial f}{\partial y}\big(x, g(x)\big)\Big)^{-1}\frac{\partial f}{\partial x}\big(x, g(x)\big). \tag{8.3}$$

Démonstration. L'idée est de considérer l'application

$$F : U \times V \to E \times G \qquad \text{définie par} \qquad F(x,y) = \big(x, f(x,y)\big)$$

et d'appliquer le théorème d'inversion locale. Cette application est de classe $\mathcal{C}^1$ et a pour dérivée

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

De plus, elle satisfait $F(a,b) = (a,0)$. Puisque $\frac{\partial f}{\partial y}(a,b)$ est un isomorphisme, $F'(a,b)$ est inversible avec pour inverse

$$F'(a,b)^{-1}(\hat{h},\hat{k}) = \Big(\hat{h}, \frac{\partial f}{\partial y}(a,b)^{-1}\Big(\hat{k} - \frac{\partial f}{\partial x}(a,b)\hat{h}\Big)\Big).$$

Cet inverse est continu car $\frac{\partial f}{\partial x}(a,b)$ et $\frac{\partial f}{\partial y}(a,b)^{-1}$ le sont. D'après le théorème d'inversion locale (Théorème 7.5), F est un difféomorphisme de classe $\mathcal{C}^1$ d'un voisinage de (a,b) sur un voisinage de $(a,0)$. On peut supposer qu'ils contiennent $U' \times V'$ et $U' \times W'$, respectivement, où U', V' et W' sont des voisinages de a, b et $0 \in G$. Quitte à réduire U', on peut aussi supposer que $F^{-1}(U' \times \{0\}) \subset U' \times V'$. Le difféomorphisme inverse F^{-1} est de la forme $F^{-1}(x,z) = (x, \hat{g}(x,z))$ et on a donc $f(x, \hat{g}(x,z)) = z$. L'application $g(x) := \hat{g}(x,0)$ est l'application cherchée.

Comme $F^{-1}(x,z)$ est de classe $\mathcal{C}^1$, les applications $\hat{g}(x,z)$ et $g(x) = \hat{g}(x,0)$ sont aussi de classe $\mathcal{C}^1$. On obtient finalement la formule (8.3) en dérivant l'identité $f(x, g(x)) = 0$ par rapport à x. □

Exemple 8.4 Soit $p_a(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$ un polynôme à coefficients réels $a = (a_0, a_1, \ldots, a_n)$. *Est-ce que la racine x_0 de $p_a(x) = 0$ est une fonction différentiable de $a_0, \ldots, a_n$?* Pour répondre à cette question, nous considérons la fonction

$$f(x, a_0, \ldots, a_n) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = p_a(x)$$

à $n+2$ variables pour laquelle

$$\frac{\partial f}{\partial x}(x, a_0, \ldots, a_n) = a_1 + 2a_2 x + \ldots + n a_n x^{n-1} = p_a'(x).$$

Si pour des coefficients $a^* = (a_0^*, \ldots, a_n^*)$ on a $p_{a^*}(x_0^*) = 0$ et $p'_{a^*}(x_0^*) \neq 0$, l'équation $p_a(x) = 0$ possède, pour a proche de a^*, une solution $x_0(a)$ qui est proche de x_0^* et qui est une fonction différentiable de $a_0, \ldots, a_n$.

En particulier, le polynôme $p_\varepsilon(x) = x^6 - x^5 + x^3 - 1 + \varepsilon$ possède un zéro proche de $x_0^* = 1$ qui dépend différentiablement de ε (on a $x_0(\varepsilon) = 1 - \varepsilon/4 + \mathcal{O}(\varepsilon^2)$). Par contre, les zéros du polynôme $p_\varepsilon(x) = x^2 - \varepsilon$ satisfont $x_0(\varepsilon) = \pm\sqrt{\varepsilon}$ pour $\varepsilon > 0$ (une fonction non-différentiable à l'origine) et pour $\varepsilon < 0$ le polynôme n'a même pas de zéro réel.

I.9 Applications bilinéaires et multilinéaires

Les dérivées d'ordre supérieur sont des applications bilinéaires (pour la 2ème dérivée), trilinéaires (pour la 3ème dérivée) et multilinéaires (pour la $n$ème dérivée). C'est la raison pour laquelle nous allons étudier ces applications un peu plus en détail.

Rappelons qu'une application

$$A : E_1 \times \ldots \times E_n \to F$$

(où $E_1, \ldots, E_n, F$ sont des espaces vectoriels normés) est dite *multilinéaire* (bilinéaire si $n = 2$) si, pour chaque $i \in \{1, \ldots, n\}$ et pour chaque $a_i \in E_i$ l'application partielle $x_i \mapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$ est linéaire. L'espace $E_1 \times \ldots \times E_n$, muni de la norme

$$\|(x_1, \ldots, x_n)\| := \max\{\|x_1\|_1, \ldots, \|x_n\|_n\}, \tag{9.1}$$

est un espace vectoriel normé. C'est un espace de Banach si les espaces $E_1, \ldots, E_n$ sont des espaces de Banach. Cela a donc un sens d'étudier la continuité d'une application multilinéaire (voir l'exemple 4.7).

Proposition 9.1 *Pour une application multilinéaire $A : E_1 \times \ldots \times E_n \to F$ ($E_1, \ldots, E_n$ et F sont des espaces vectoriels normés), les conditions suivantes sont équivalentes:*

(a) *A est continue en tout point de $E_1, \ldots, E_n$;*

(b) *A est continue à l'origine $(0, \ldots, 0) \in E_1 \times \ldots \times E_n$;*

(c) *$\|A(x_1, \ldots, x_n)\|$ est bornée sur la boule-unité de $E_1 \times \ldots \times E_n$.*

Démonstration. La démonstration procède comme celle de la Proposition 3.1. Pour démontrer (c) $\Rightarrow$ (a), nous écrivons

$$\begin{aligned} A(x_1, \ldots, x_n) - A(a_1, \ldots, a_n) &= A(x_1 - a_1, x_2, \ldots, x_n) \\ &+ A(a_1, x_2 - a_2, x_3, \ldots, x_n) + \ldots + A(a_1, \ldots a_{n-1}, x_n - a_n) \end{aligned}$$

et estimons chaque terme séparement. □

Définition 9.2 Soient $E_1, \ldots, E_n$ et F des espaces vectoriels normés. On dénote par $\mathcal{L}(E_1, \ldots, E_n; F)$ l'ensemble des applications multilinéaires continues de $E_1 \times \ldots \times E_n$ dans F. Pour un élément $A \in \mathcal{L}(E_1, \ldots, E_n; F)$, on définit

$$\|A\| = \sup_{\|(x_1,\ldots,x_n)\| \leq 1} \|A(x_1, \ldots, x_n)\| = \sup_{x_1 \neq 0, \ldots, x_n \neq 0} \frac{\|A(x_1, \ldots, x_n)\|}{\|x_1\|_1 \cdot \ldots \cdot \|x_n\|_n}. \tag{9.2}$$

Si $E_1 = \ldots = E_n = E$, on écrit aussi $\mathcal{L}^n(E; F)$ au lieu de $\mathcal{L}(E, \ldots, E; F)$.

L'espace $\mathcal{L}(E_1, \ldots, E_n; F)$ muni de (9.2) est un espace vectoriel normé. Si F est complet, alors $\mathcal{L}(E_1, \ldots, E_n; F)$ est aussi complet. Cette affirmation est démontrée exactement comme pour la Proposition 3.4. Le résultat de la proposition suivante est la base pour l'interprétation de la deuxième dérivée d'une fonction comme application bilinéaire.

Proposition 9.3 (Isométrie naturelle) *Soient E, F, G des espaces vectoriels normés. Alors, l'application*

$$\psi : \mathcal{L}(E, \mathcal{L}(F, G)) \to \mathcal{L}(E, F; G),$$

définie par $\psi(A) = B$ où $B(x, y) = (Ax)y$, est un isomorphisme qui est également une isométrie, c.-à-d., on a $\|\psi(A)\| = \|A\|$ pour tout $A \in \mathcal{L}(E, \mathcal{L}(F, G))$.

Démonstration. L'application ψ est linéaire et bijective. Son inverse est donné par $\psi^{-1}(B) = A$ où, pour $x \in E$, $Ax \in \mathcal{L}(F, G)$ est l'application $y \mapsto B(x, y)$. Pour démontrer la continuité de ψ et de ψ^{-1}, il suffit de voir que $\|\psi(A)\| = \|A\|$.

En utilisant l'estimation (3.2) une fois pour Ax et une deuxième fois pour A, nous obtenons

$$\|B(x, y)\| = \|(Ax)y\| \leq \|Ax\| \cdot \|y\| \leq \|A\| \cdot \|x\| \cdot \|y\|.$$

Ceci impliqe que $B = \psi(A)$ satisfait $\|B\| \leq \|A\|$.

D'autre part, la Définition 9.2 montre que

$$\|(Ax)y\| = \|B(x, y)\| \leq \|B\| \cdot \|x\| \cdot \|y\|.$$

La norme de l'application linéaire Ax satisfait donc $\|Ax\| \leq \|B\| \cdot \|x\|$ et, par conséquent, $\|A\| \leq \|B\|$. Ces deux inégalités montrent que l'application ψ est une isométrie. □

On peut donc identifier les applications linéaires continues de E dans $\mathcal{L}(F, G)$ avec les applications bilinéaires de $E \times F$ dans G.

Exemple 9.4 Une matrice C (avec n colonnes et m lignes) peut être identifiée avec l'application bilinéaire $B : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ définie par $B(x, y) = x^T C y$. Elle peut également être identifiée avec un élément de $\mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ de la manière suivante: pour $x \in \mathbb{R}^m$, le vecteur $x^T C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ définit l'application $x^T C : y \mapsto x^T C y$.

L'affirmation de la proposition précédente peut être généralisée à des applications multilinéaires. Si $E_1, \ldots, E_n$ et F sont des espaces vectoriels normés, l'application

$$\psi : \mathcal{L}(E_1, \mathcal{L}(E_2, \ldots, E_n; F)) \to \mathcal{L}(E_1, \ldots, E_n; F), \tag{9.3}$$

définie par $\psi(A) = B$ où $B(x_1, \ldots, x_n) = (Ax_1)(x_2, \ldots, x_n)$, est un isomorphisme et une isométrie. La démonstration de ce fait est identique à celle de la Proposition 9.3.

I.10 Dérivées d'ordre supérieur

La dérivée d'une fonction $f : U \to F$ (avec $U \subset E$) est une application $f' : U \to \mathcal{L}(E, F)$. Comme $\mathcal{L}(E, F)$ est un espace vectoriel normé, rien ne nous empèche de considérer la différentiabilité de f'.

Définition 10.1 Soient E, F deux espaces vectoriels normés, $U \subset E$ un ouvert et $f : U \to F$ différentiable dans un voisinage de $a \in U$. On dit que f est *deux fois différentiable* en a si

$f' : U \to \mathcal{L}(E, F)$ est différentiable en a. La dérivée de f' satisfait $(f')'(a) \in \mathcal{L}(E, \mathcal{L}(E, F))$. En utilisant l'identification de la Proposition 9.3, on définit la *deuxième dérivée* de f en a comme l'application bilinéaire

$$f''(a)(h,k) := \big((f')'(a)h\big)k.$$

Dans la situation de la définition précédente, considérons l'application $g_k : U \to F$ définie par $g_k(x) := f'(x)k$ (avec un $k \in E$ fixé). Alors, par la formule (4.10) de Leibniz (on a $g_k(x) = B(f'(x), k)$ où $B : \mathcal{L}(E,F) \times E \to F$ est l'application bilinéaire $B(A,v) = Av$) on obtient

$$g_k'(a)h = \big((f')'(a)h\big)k = f''(a)(h,k).$$

Exemple 10.2 Considérons une fonction $f : I\!R^n \to I\!R^m$ et notons $x = (x_1, \ldots, x_n)^T$. Pour $k \in I\!R^n$ fixé, la dérivée de la fonction $g_k(x) := f'(x)k = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x)\, k_j$ est donnée par $g_k'(x)h = \sum_{i=1}^n (\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\, k_j) h_i$. La deuxième dérivée de f est donc

$$f''(x)(h,k) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x)\, h_i k_j.$$

Pour la fonction $f : I\!R^3 \to I\!R^2$ de l'exemple 4.2, la deuxième dérivée est

$$f''(a)(h,k) = \begin{pmatrix} 2h_1k_1 + 2h_2k_2 + 2h_3k_3 \\ a_1(h_2k_3 + h_3k_2) + a_2(h_1k_3 + h_3k_1) + a_3(h_1k_2 + h_2k_1) \end{pmatrix}.$$

Exemple 10.3 Pour la fonction $f(X) = X^{-1}$ de $\mathrm{GL}(E)$ dans $\mathcal{L}(E)$ l'Exemple 4.6 montre que $g_K(X) := f'(X)K = -X^{-1}KX^{-1}$ pour $K \in \mathcal{L}(E)$. La formule de Leibniz nous donne alors

$$f''(A)(H,K) = A^{-1}HA^{-1}KA^{-1} + A^{-1}KA^{-1}HA^{-1}.$$

Théorème 10.4 *Soient E, F deux espaces vectoriels normés, $U \subset E$ un ouvert, et supposons que $f : U \to F$ soit deux fois différentiable en $a \in U$. Alors, l'application bilinéaire $f''(a) : E \times E \to F$ est symétrique, c.-à-d.,*

$$f''(a)(h,k) = f''(a)(k,h) \qquad \textit{pour} \quad h,k \in E.$$

Démonstration. Nous allons montrer que

$$f''(a)(h,k) \;=\; \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2}\Big(f(a+\varepsilon h+\varepsilon k) - f(a+\varepsilon h) - f(a+\varepsilon k) + f(a)\Big). \qquad (10.1)$$

Puisque l'expression de droite est symétrique en h et k, il est de même de $f''(a)(h,k)$.

Pour démontrer (10.1), nous considérons l'application

$$g_u(v) := f(a+u+v) - f(a+u) - f(a+v) + f(a) - f''(a)(u,v)$$

où u et v sont suffisamment petits en norme. Comme $g_u(0) = 0$, le théorème des accroissements finis (Théroème 5.2) implique que

$$\|g_u(v)\| \le \sup_{0<t<1} \|g_u'(tv)\| \cdot \|v\|. \qquad (10.2)$$

Il reste donc à estimer la dérivée

$$g_u'(v) = f'(a+u+v) - f'(a+v) - f''(a)(u,\cdot)$$

(la notation $f''(a)(u,\cdot)$ est utilisée pour l'application $h \mapsto f''(a)(u,h)$). Le fait que f est deux fois différentiable en a (c.-à-d., $x \mapsto f'(x)$ est différentiable en a) implique que

$$\begin{aligned} f'(a+u+v) &= f'(a) + f''(a)(u+v,\cdot) + r(u+v)\|u+v\| \\ f'(a+v) &= f'(a) + f''(a)(v,\cdot) + r(v)\|v\| \end{aligned}$$

où $r(v) \to 0$ si $v \to 0$. La soustraction de ces deux équations donne pour $g_u'(v)$ la formule $g_u'(v) = r(u+v)\|u+v\| - r(v)\|v\|$ et on obtient pour $0 < t < 1$

$$\|g_u'(tv)\| \le \Big(\|r(u+tv)\| + \|r(tv)\|\Big)\Big(\|u\| + \|v\|\Big). \tag{10.3}$$

On déduit des estimations (10.2) et (10.3) que $\|g_u(v)\|/(\|u\|+\|v\|)^2 \to 0$ si $\|u\|+\|v\| \to 0$. En posant $u = \varepsilon h$ et $v = \varepsilon k$ on obtient donc l'affirmation (10.1). □

La définition de la troisième dérivée est analogue à Définition 10.1. Supposons qu'une fonction $f : U \to F$ (E, F des espaces vectoriels normés et $U \subset E$ un ouvert) soit 2 fois différentiable dans un voisinage de $a \in U$. On dit que f est 3 fois différentiable en a si $f'' : U \to \mathcal{L}^2(E;F)$ est différentiable en a. La dérivée satisfait $(f'')'(a) \in \mathcal{L}(E, \mathcal{L}^2(E;F))$ et, en utilisant l'isométrie entre $\mathcal{L}(E, \mathcal{L}^2(E;F))$ et $\mathcal{L}^3(E;F)$ on définit la *3ème dérivée de f* en a comme l'application trilinéaire

$$f'''(a)(h,k,l) := \Big((f'')'(a)h\Big)(k,l).$$

De manière évidente on définit par récurrence la $p$ème dérivée $f^{(p)}(a)$ comme application multilinéaire $f^{(p)}(a) \in \mathcal{L}^p(E;F)$.

Pour un calcul pratique de la 3ème dérivée, on peut utiliser la formule

$$g_{kl}'(a)h = \Big((f'')'(a)h\Big)(k,l) = f'''(a)(h,k,l) \tag{10.4}$$

où $g_{kl} : U \to F$ est définie par $g_{kl}(x) := f''(x)(k,l)$ et $k, l \in E$ sont des vecteurs fixés. La 3ème dérivée de f peut également être interprétée comme la deuxième dérivée de $g_l(x) := f'(x)l$, c.-à-d.,

$$g_l''(a)(h,k) = \Big((f')''(a)(h,k)\Big)l = f'''(a)(h,k,l). \tag{10.5}$$

Les deux formules (10.4) et (10.5) montrent qu'on peut échanger k et l ainsi que h et k sans changer la valeur de $f'''(a)(h,k,l)$. On a donc le résultat suivant.

Corollaire 10.5 *Si, sous les hypothèses du Théorème 10.4, la fonction $f : U \to F$ est p fois différentiable en $a \in U$, l'application multilinéaire $f^{(p)}(a)$ est symétrique, c.-à-d.,*

$$f^{(p)}(a)(h_1,\ldots,h_p) = f^{(p)}(a)(h_{\sigma(1)},\ldots,h_{\sigma(p)})$$

où σ est une permutation quelconque de $\{1,\ldots,p\}$. □

Exemple 10.6 La 3ème dérivée d'une application $f : \mathbb{R}^n \to \mathbb{R}^m$ est donnée par

$$f'''(x)(h,k,l) = \sum_{i=1}^n \sum_{j=1}^n \sum_{m=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_m}(x)\, h_i k_j l_m.$$

Définition 10.7 Soient E, F des espaces vectoriels normés, $U \subset E$ et $V \subset F$ des ouverts.

- $f : U \to F$ est *p fois continûment différentiable* (ou *de classe $\mathcal{C}^p$*) sur U, si f est p fois différentiable en tout point de U et si $f^{(p)} : U \to \mathcal{L}^p(E; F)$ est continue.
- $f : U \to F$ est *indéfiniment différentiable* (ou *de classe $\mathcal{C}^\infty$*) sur U, si f est de classe $\mathcal{C}^p$ pour tout p.
- $f : U \to V$ est un *difféomorphisme* (de classe $\mathcal{C}^p$) de U sur V, si f est bijective, p fois continûment différentiable, et si l'inverse $f^{-1} : V \to U$ est p fois continûment différentiable.

Exemple 10.8 La fonction $f(X) = X^{-1}$ de $\mathrm{GL}(E)$ dans $\mathcal{L}(E)$ est indéfiniment différentiable. En effet, en dérivant la formule de l'Exemple 10.3 on obtient

$$\begin{aligned} f'''(A)(H, K, L) &= -A^{-1}HA^{-1}KA^{-1}LA^{-1} - A^{-1}HA^{-1}LA^{-1}KA^{-1} \\ &\quad -A^{-1}LA^{-1}HA^{-1}KA^{-1} - A^{-1}KA^{-1}HA^{-1}LA^{-1} \\ &\quad -A^{-1}KA^{-1}LA^{-1}HA^{-1} - A^{-1}LA^{-1}KA^{-1}HA^{-1}. \end{aligned}$$

On voit par récurrence que chaque dérivée est une combinaison linéaire d'expressions qui sont un produit alterné de A^{-1} avec des applications linéaires constantes.

Proposition 10.9 *Soient E, F des espaces vectoriels normés, $U \subset E$, $V \subset F$ des ouverts et $f : U \to V$ un difféomorphisme de classe $\mathcal{C}^1$. Si f est de classe $\mathcal{C}^p$, alors la fonction inverse $f^{-1} : V \to U$ est aussi de classe $\mathcal{C}^p$.*

Démonstration. Par le Théorème 7.5 on sait que

$$(f^{-1})'(y) = f'(f^{-1}(y))^{-1},$$

c.-à-d., $(f^{-1})'$ est la composition de trois applications, notamment $y \mapsto f^{-1}(y)$, $x \mapsto f'(x)$ et $A \mapsto A^{-1}$. Toutes les trois applications sont continûment différentiables. Alors, f^{-1} est de classe $\mathcal{C}^2$. La démonstration pour $p \geq 3$ est analogue. □

Une conséquence de cette proposition est la suivante: si dans le théorème d'inversion locale (Théorème 7.5) la fonction f est de classe $\mathcal{C}^p$, alors son inverse est automatiquement de classe $\mathcal{C}^p$. Similairement, si dans le théorème des fonctions implicites (Théorème 8.3) la fonction f est de classe $\mathcal{C}^p$, alors l'application g est aussi de classe $\mathcal{C}^p$.

I.11 Exercices

1. Soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ donnée par
$$f(x) = \begin{pmatrix} x_1x_2^2 - 3x_1x_2x_3 \\ x_1^2 + x_2^2 + x_3^2 \end{pmatrix},$$
et $g : \mathbb{R}^n \to \mathbb{R}$ donnée par $g(x) = x^T Ax + b^T x$, où A est une matrice $n \times n$ et $b \in \mathbb{R}^n$. Calculer les deux premières dérivées de f et g.
2. Montrer que l'application $N : \mathbb{R}^2 \to \mathbb{R}$ donnée par $N(x) = \sqrt{x_1^2 - x_1x_2 + 4x_2^2}$, définit une norme sur $\mathbb{R}^2$ et montrer (explicitement) que cette norme est équivalente à la norme $\|\cdot\|_\infty$.
3. Démontrer que (2.1) est une norme.

4. Si la norme découle d'un produit scalaire (2.1), alors on a
$$\|x+y\|+\|x-y\| = 2(\|x\|+\|y\|) \qquad \text{(identité du parallélogramme)}$$
Montrer à l'aide de contre-exemples que $\mathbb{R}^n$ munit de la norme $\|\cdot\|_1$ ou $\|\cdot\|_\infty$ n'est pas un espace de Hilbert.

5. On considère les ensembles suivants de $\mathbb{R}^2$:
$$\begin{aligned} A &= \Big\{(x,y)\in\mathbb{R}^2,\ 0 < x+y \le 1\Big\}\ , \\ B &= \Big\{(x,y)\in\mathbb{R}^2,\ \max(|x|,|y|) < 3\Big\}\ , \\ C &= \Big\{(x,y)\in\mathbb{R}^2,\ (x^2+y^2)^2-2x^2+2y^2=0\Big\}\ , \\ D &= \Big\{(x,y)\in\mathbb{R}^2,\ y=0,\ x\in\{1/n, n=1,2,\ldots\}\Big\}\ . \end{aligned}$$
Quels ensembles sont-ils ouverts, fermés, bornés, compacts? Justifier.

6. Pour $\alpha \ge 0$ considérons la fonction $f:\mathbb{R}^2\to\mathbb{R}$ donnée par
$$f(x,y) = \begin{cases} \frac{xy}{(x^2+y^2)^\alpha} & \text{si } (x,y)\ne(0,0) \\ 0 & \text{si } (x,y)=(0,0)\ . \end{cases}$$
   (a) Montrer que $\partial f/\partial x$ et $\partial f/\partial y$ existent partout et sont telles que les applications $x\mapsto \frac{\partial f}{\partial x}(x,b)$, $y\mapsto \frac{\partial f}{\partial y}(a,y)$ sont continues pour tout $a,b\in\mathbb{R}$.
   (b) Si $\alpha \ge 1/2$, les applications $\partial f/\partial x$ et $\partial f/\partial y$ sont discontinues en (0,0).
   (c) $f(x,y)$ est continue à l'origine si et seulement si $\alpha < 1$.
   (d) $f(x,y)$ est dérivable à l'origine si et seulement si $\alpha < 1/2$.
   (e) Pour $\alpha = 1$, trouver un ouvert $U\subset\mathbb{R}$ et un fermé $F\subset\mathbb{R}$ tels que $f^{-1}(U)$ ne soit pas ouvert et $f^{-1}(F)$ ne soit pas fermé.

7. Soit E un espace de Hilbert et soit $A\in\mathcal{L}(E)$. Montrer que
$$\|A\| = \sup_{u\ne0,v\ne0}\frac{\langle u,Av\rangle}{\|u\|\cdot\|v\|}$$
où $\langle\cdot,\cdot\rangle$ est le produit scalaire et $\|\cdot\|$ est la norme induite par le produit scalaire.

8. Démontrer que $\|f\|_2$ est une norme sur $\mathcal{C}([0,1])$.

9. Démontrer que, pour un ensemble arbitraire A,
$$\mathcal{B}(A) = \{f:A\to\mathbb{R}\ ;\ f \text{ est bornée}\} \quad \text{avec} \quad \|f\|_\infty = \sup_{t\in A}|f(t)|\ ,$$
est un espace de Banach (Proposition 2.3).

10. Le rayon spectral $\rho(A)$ d'une matrice $n\times n$, A, est défini comme suit:
$$\rho(A) = \max\{|\lambda|\ ;\ \lambda \text{ valeur propre de } A\}\ .$$
Soit $E=\mathbb{R}^n$ muni d'une norme et soit $A\in\mathcal{L}(E)$. Montrer que pour toute norme sur $\mathbb{R}^n$ on a l'inégalité suivante $\|A\|\ge\rho(A)$, et que si A est symmétrique on l'égalité $\|A\|_2=\rho(A)$.

11. Soit
$$A = \begin{pmatrix} 0.999 & 1000 \\ 0 & 0.999 \end{pmatrix}\ .$$
   (a) Calculer le rayon spectral $\rho(A)$ ainsi que $\|A\|_1$, $\|A\|_2$ et $\|A\|_\infty$.
   (b) Trouver une norme $\|\cdot\|$ sur $\mathbb{R}^2$ telle que, pour cette norme, on ait $\|A\|\le1$.

12. Soit l'ensemble des matrices $n\times n$ que l'on identifie avec $\mathbb{R}^{n\cdot n}$ et que l'on muni de la norme $\|A\| = \sup_{\|x\|\le1}\|Ax\|$. Considérons l'application linéaire $l:\mathbb{R}^{n\cdot n}\to\mathbb{R}^{n\cdot n}$ donnée par $l(H)=X_0H+HX_0$ où X_0 est une matrice fixée (voir Exemple 1.1). Calculer la norme de l'application l.

13. Dans la situation de l'Exemple 3.8 démontrer que
$$\|A\|_\infty \;=\; \max_{t\in[0,1]} \int_0^1 |k(t,s)|\, ds.$$
Indication. Considérer en premier que la fonction $k(t,s)$ possède qu'un nombre fini de zéros pour tout t.
14. On considère l'équation intégrale suivante, dite équation intégrale de Fredholm (2ème espèce):
$$y(t) \;=\; f(t) + \int_0^1 k(t,s)y(s)ds \;, \tag{11.1}$$
où $f \in \mathcal{C}([0,1])$ et $k \in \mathcal{C}([0,1]\times[0,1])$ sont données et y est l'inconnue. On suppose que
$$\max_{t\in[0,1]} \int_0^1 |k(t,s)|ds \;<\; 1$$
   (a) Démontrer que (11.1) possède une unique solution dans $\mathcal{C}([0,1])$.
   (b) En utilisant la série de Neumann, écrire la solution sous la forme
$$y(t) = f(t) + \int_0^1 r(t,s)f(s)ds$$
15. Résoudre
$$y(t) = f(t) + \frac{1}{2}\int_0^1 e^{t-s}y(s)ds \;. \tag{11.2}$$
   (a) On pose $a \equiv \int_0^1 e^{-s}y(s)ds$. En multipliant (11.2) par e^{-t} et en intégrant, on trouve a et donc aussi la solution $y(t) = f(t) + ae^t/2$.
   (b) En utilisant la formule de l'exercice précédent.
16. Soit E un espace de Banach, on considère l'application $f : \mathcal{L}(E) \to \mathcal{L}(E)$ donnée par $f(X) = X^3$. Calculer la dérivée de f. Montrer par des exemples, qu'en général, $f'(X) \neq 3X^2$.
17. Soit E un espace de Banach, on considère l'application $\exp : \mathcal{L}(E) \to \mathcal{L}(E)$ donnée par la série suivante:
$$\exp(X) \;=\; \sum_{n=0}^{\infty} \frac{1}{n!}X^n \;.$$
   (a) Montrer que l'application exp est bien définie.
   (b) Calculer la derivée, $\exp'$ de l'application exp.
   (c) Montrer par des exemples que, en général, $\exp'(X) \neq \exp(X)$.

   Indication. Si cela vous aide, supposez que $E = \mathbb{R}^n$.
18. Calculer la dérivée de la fonction $f : \mathrm{GL}(\mathbb{R}^n) \to \mathcal{L}(\mathbb{R}^n)$ donnée par $f(X) = (X^TX)^{-1}$.
19. Soit $A : \mathbb{R} \to \mathbb{R}^{n\cdot n}$ une fonction différentiable à valeur dans les matrices $n \times n$. Montrer que
$$\frac{d}{dx}\Big(A(x)\Big)^{-1} \;=\; -A(x)^{-1}A'(x)A(x)^{-1}$$
Indication. Utiliser l'Exemple 4.6 ainsi que la formule pour la dérivée de fonctions composées.
20. (a) Soit $M : E_1 \times \ldots \times E_n \to F$ une application multilinéaire où $E_1, \ldots, E_n, F$ sont des espaces vectoriels normés. Calculer la dérivée de M.
   (b) Soit $A : \mathbb{R} \to \mathbb{R}^{3\cdot 3}$ une fonction matricielle différentiable dont les colonnes sont $A_1(x)$, $A_2(x)$ et $A_3(x)$. Montrer que
$$\frac{d}{dx}\det\Big(A(x)\Big) \;=\; \det\left(A_1', A_2, A_3\right) + \det\left(A_1, A_2', A_3\right) + \det\left(A_1, A_2, A_3'\right)$$
21. Soit $f(x,y) = \ln(x^2+y^2)$ pour $x^2+y^2 > 0$ et $g(x,y) = (xy, \sqrt{x}/y)^T$ pour $x > 0, y > 0$. Calculer la dérivée de $h(x,y) = (f\circ g)(x,y) = \ln(x^2y^2 + x/y^2)$ directement et ensuite avec la formule de différentiation des fonctions composées.
22. Soit $f : [a,b] \to \mathbb{R}$ continue, $x_1, x_2, \ldots, x_n \in (a,b)$ et f différentiable sur l'ensemble $D = (a,b)\backslash\{x_1, x_2, \ldots, x_n\}$. Démontrer ou donner des contre-exemples aux affirmations suivantes:
   (a) Il existe un $\xi \in D$ tel que $f(b) - f(a) \;=\; f'(\xi)(b-a)$.

   (b) On a l'estimation $|f(b) - f(a)| \leq \sup_{t\in D} |f'(t)|(b-a)$.

23. Considérons la fonction $f : I\!R \to I\!R$ définie par
$$f(x) \;=\; \begin{cases} x + e^{-x/2} & \text{si } x \geq 0 \\ e^{x/2} & \text{si } x \leq 0 \end{cases} \;.$$
   (a) Montrer que $|f(x) - f(y)| < |x - y|$ pour $x \neq y$.
   (b) Montrer que la fonction $f(x)$ ne possède pas de point fixe.

   Est-ce une contradiction au théorème du point fixe?

24. On considère $\mathcal{C}([0,1])$ avec la norme $\|\cdot\|_\infty$ et l'application $f : \mathcal{C}([0,1]) \to \mathcal{C}([0,1])$ donnée par
$$f(x)(t) \;=\; \lambda \int_0^1 k(t,s) g(x(s)) ds$$
   où $k \in \mathcal{C}([0,1] \times [0,1])$ et $g : I\!R \to I\!R$ est suffisamment différentiable. Trouver une condition sur λ sous laquelle f est une contraction sur le fermé
$$D \;=\; \{x \in \mathcal{C}([0,1]) \;;\; \|x\|_\infty \leq 1\} \;.$$

25. Montrer que l'hypothèse "D fermé" ne peut en général pas être omise dans le théorème du point fixe de Banach: trouver un D non fermé et une application $f : D \to E$ tels que $f(D) \subset D$ et f est une contraction sur D, mais f ne possède pas de point fixe dans D.

26. Soit U un ouvert connexe de $I\!R$ et E un espace vectoriel normé. Soit $f : U \to E$ une application qui satisfait
$$\|f(x) - f(y)\| \;\leq\; C\|x - y\|^2 \quad \text{pour tout } x, y \in U \;,$$
   où $C > 0$. Montrer que f est constante sur U.

27. Soit la fonction $f : I\!R^2 \to I\!R^3$ donnée par
$$f(x_1, x_2) \;=\; \begin{pmatrix} x_1^2 + x_2^2 + 5 \\ 3x_1x_2 + 2 \\ x_1^3 + x_2^3 \end{pmatrix} \;,$$
   et soit $K = \{(x_1, x_2) \in I\!R^2 \mid |x_1| \leq 1, |x_2| \leq 2\}$. Trouver une constante L telle que
$$\|f(x) - f(y)\|_\infty \;\leq\; L\|x - y\|_\infty \quad \text{pour } x, y \in K \;.$$
   Remarque. L s'appelle *constante de Lipschitz.*

28. Soit $A \in \mathcal{L}(E)$ avec $\|A\| < 1$ et considérons le problème
$$x \;=\; Ax + b \;.$$
   Motrer que la k-ième approximation, x_k, obtenue par la méthode des approximations successives (avec $x_0 = b$) est égale a l'approximation obtenue en tronquant la série de Neumann pour $(I - A)^{-1}$ après k termes. Trouver ainsi une connection entre l'Exemple 6.4 et l'Exercice 14.

29. Estimation de l'erreur pour la méthode des approximations successives. Soit x^* un point fixe donné par cette méthode. Démontrer
$$\|x^* - x_k\| \leq \frac{\alpha}{1-\alpha}\|x_k - x_{k-1}\| \;. \tag{11.3}$$
   Pour les fonctions de l'Exemple 6.2, faire quelques itérations et utiliser (11.3) pour obtenir une estimation rigoureuse de l'erreur.

30. Soit E, F des espaces vectoriels normés et $U \subset E$ un fermé. On se pose la question suivante: soit $f : U \to F$ injective et continue, est-ce que f est un homéomorphisme de U sur $f(U)$?
   (a) Démontrer que ceci est vrai si $E = I\!R^n$ et U borné (utiliser la démonstration de [HW95, Théorème III.3.9]).
   (b) Montrer par un contre-exemple que ce résultat n'est plus vrai dans les espaces vectoriels de dimension infinie (considérer par exemple l'identité dans un espace muni de 2 normes non-équivalentes).

31. Considérons la fonction $f : [-1,1] \to \mathbb{R}$ donnée par
$$f(x) \;=\; \begin{cases} \frac{1}{n+1}\left(2|x| - \frac{1}{n}\right) & \text{si} \quad \frac{1}{n} \le |x| \le \frac{1}{n-1} \\ 0 & \text{si} \quad x = 0 \end{cases} .$$
Montrer que f est strictement différentiable à l'origine, mais qu'elle n'est différentiable dans aucun voisinage de 0.
Remarque. $f(x)$ satisfait $f(1/n) = 1/n(n+1)$ et elle est linéaire sur $[1/n, 1/(n+1)]$.

32. Soit $A \in \mathcal{L}(E)$ un isomorphisme. Pour calculer son inverse, on applique la méthode de Newton à l'équation $X^{-1} - A = 0$. Montrer que l'on obtient ainsi l'itération
$$X_{k+1} \;=\; X_k + X_k(I - AX_k)$$
et vérifier que la suite $\{X_k\}$ converge vers A^{-1} si $\|I - AX_0\| < 1$.
Indication. $E_k := I - AX_k$ satisfait $E_{k+1} = E_k E_k$.

33. Soit E, F, G des espaces vectoriels normés. Soit $U \subset E$ un ouvert et $a \in U$ et soit $f : U \to G$ strictement différentiable en a. Soit $V \subset F$ un ouvert avec $b = f(a) \in V$ et soit $g : V \to G$ strictement différentiable en $b = f(a)$. Démontrer que $g \circ f$ est strictement différentiable en a et que $(g \circ f)'(a) \;=\; g'(f(a))f'(a)$.

34. Considérons la fonction $f : \mathbb{R}^2 \to \mathbb{R}^2$ donnée par
$$f_1(x_1, x_2) \;=\; x_1, \qquad f_2(x_1, x_2) \;=\; \begin{cases} x_2 - x_1^2 & \text{si} \quad x_1^2 \le x_2 \\ (x_2^2 - x_1^2 x_2)/x_1^2 & \text{si} \quad 0 \le x_2 \le x_1^2 \\ -f_2(x_1, -x_2) & \text{si} \quad x_2 \le 0 \end{cases}$$
Montrer que
   (a) f est partout différentiable.
   (b) $f'(0)$ est inversible (calculer $f'(0)$).
   (c) f n'est pas strictement différentiable à l'origine.
   (d) il n'existe pas de voisinage de l'origine où f est injective (i.e. f n'est pas un homéomorphisme près de l'origine).

35. Donner un difféomorphisme (a) de l'intervalle $(0,1)$ avec $\mathbb{R}$, (b) de l'intervalle $(0,1)$ avec la demi-droite $\mathbb{R}_+^*$.

36. Soient le "cube" $C = \{x \in \mathbb{R}^n \,;\, 0 < x_i < 1, i = 1, \ldots, n\}$ et le "simplexe" $S = \{y \in \mathbb{R}^n \,;\, y_i > 0\,,\ \sum_{i=1}^n y_i < 1\}$ et on considère l'application $\varphi : C \to \mathbb{R}^n$ donnée par $\varphi(x) = y$ avec $y_k = (1 - x_1) \cdots (1 - x_{k-1}) x_k$ pour $k = 1, \ldots, n$.
   (a) Pour $n = 2$ dessiner les images par φ des droites parallèles aux axes.
   (b) Vérifier que $\sum_{i=1}^n y_i = 1 - \prod_{i=1}^n (1 - x_j)$.
   (c) Montrer que φ est un difféomorphisme de C sur S et trouver son inverse.

37. Les coordonnées sphériques de $\mathbb{R}^n$ sont données par $\psi(r, \theta_1, \ldots \theta_{n-1}) = x$ où
$$\begin{aligned} x_1 &= r\cos\theta_1 \\ x_2 &= r\sin\theta_1\cos\theta_2 \\ x_3 &= r\sin\theta_1\sin\theta_2\cos\theta_3 \\ &\;\;\vdots \\ x_{n-1} &= r\sin\theta_1\sin\theta_2\ldots\sin\theta_{n-2}\cos\theta_{n-1} \\ x_n &= r\sin\theta_1\sin\theta_2\ldots\sin\theta_{n-2}\sin\theta_{n-1}\ . \end{aligned}$$
Montrer que
$$\det\psi' = r^{n-1}\sin^{n-2}\theta_1\sin^{n-3}\theta_2\cdots\sin\theta_{n-2}\ ,$$
et montrer que la restriction de l'application ψ à $(0,\infty) \times (0,\pi)^{n-2} \times (-\pi,\pi)$ est un difféomorphisme.

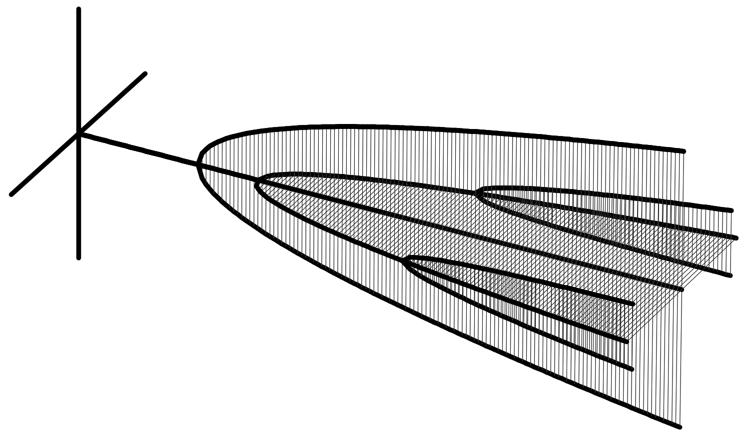
38. Soit α_n le volume de la sphère $\{x \in \mathbb{R}^n\,;\, \|x\|_2 \le 1\}$ dans $\mathbb{R}^n$. En utilisant les résultats de l'exercice 37, le théorème sur les changements de coordonnées ([HW95, Théorème IV.5.7]) et la définition de la fonction Gamma $\Gamma(x)$ ([HW95, Définition III.8.10]) montrer que
$$\alpha_n = \frac{\pi^{n/2}}{\Gamma(n/2+1)}\ .$$
Montrer également que $\alpha_n \to 0$ si $n \to \infty$.
Indication. Pour $I_k \equiv \int_0^\pi \sin^k t\,dt$ on a $(k+1)I_{k+1} = kI_{k-1}$.

39. Dans le plan avec coordonnées cartésiennes (x, y) on considère les deux courbes
$$\begin{aligned} C_1 &= \{(x,y)\,;\, x^2 - y = 0\} \\ C_2 &= \{(x,y)\,;\, x - y^2 = 0\}\ . \end{aligned}$$
Trouver un difféomorphisme local près de $(0, 0)$ tel que
$$\begin{aligned} f(C_1) &= \{(u,v)\,;\, v = 0\} \\ f(C_2) &= \{(u,v)\,;\, u = 0\}\ . \end{aligned}$$

40. Pour tout $a \in \mathbb{R}$ on considère la fonction $f_a : \mathbb{R} \to \mathbb{R}$ donnée par $f_a(x) = ax - \sin x$. Pour quelles valeurs de a la fonction f_a est-elle un homéomorphisme? un difféomorphisme? un difféomorphisme local près de l'origine?

41. Soit $f : \mathbb{R}^4 \to \mathbb{R}^2$ l'application définie par
$$\begin{aligned} f_1(x,y,u,v) &= u^3 + vx + y \\ f_2(x,y,u,v) &= uy + v^3 - x \end{aligned}$$
et soit $P_0 = (x_0, y_0, u_0, v_0)$ un point de $f^{-1}(\{0\})$.
    (a) Ecrire une condition sur P_0 qui permet d'obtenir localement près de ce point l'ensemble $f^{-1}(\{0\})$ comme le graphe d'une fonction g de la forme $u = g_1(x, y)$, $v = g_2(x, y)$.
    (b) Calculer $\frac{\partial g_1}{\partial x}(x_0, y_0)$ si $P_0 = (0, -1, 1, 1)$.

42. On considère le système
$$\begin{aligned} \lambda c_1 &= 3c_1 + \frac{9}{4}c_1^3 + \frac{9}{2}c_1 c_2^2 \\ \lambda c_2 &= 2c_2 + 3c_1^2 c_2 + \frac{3}{2}c_2^2. \end{aligned}$$
Trouver les points de bifurcations des courbes $c_1(\lambda)$ et $c_2(\lambda)$. Une condition nécessaire pour un point de bifurcation est que le théorème de fonctions implicites ne peut pas être appliqué.



43. Démontrer que le théorème d'inversion locale (Théorème 7.5) peut être déduit du théorème des fonctions implicites.

44. Soit $S = \{(x,y,z)\,;\, xz + \sin(xy) + \cos(xz) = 1\}$. Déterminer si dans un voisinage de $(0, 1, 1)$ l'ensemble S peut être écrit sous la forme
$$z = f(x,y) \quad \text{ou} \quad y = g(x,z) \quad \text{ou} \quad x = h(y,z)\ .$$

45. Calculer la norme du déterminant vu comme application multilinéaire
$$\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R} \qquad (n \text{ facteurs})\,,$$
où $\mathbb{R}^n$ est munie de la norme euclidienne. En déduire l'inégalité $|\det A| \le n^{n/2}$ (de Hadamard) pour une matrice $A = (a_{ij})_{1 \le i,j \le n}$ satisfaisant $|a_{ij}| \le 1$.

46. Soient $\mathbb{R}^n$ et $\mathbb{R}^m$ munis de la norme euclidienne. Soit $b : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ l'application bilinéaire donnée par $b(x,y) = x^T By$, où B est une matrice arbitraire. Montrer que la norme de b satisfait $\|b\| = \|B\|_2$.
47. Soient $f : E \to F$ et $g : f \to G$ (où E, F et G sont des espaces vectoriels normés) deux applications deux fois différentiables. Calculer $(g \circ f)''(x)(h,k)$.
48. Soit $b : E \times F \to G$ (où E, F et G sont des espaces vectoriels normés) une application bilinéaire continue. Calculer sa deuxième dérivée.
49. Calculer la deuxième dérivée de $f : \mathbb{R}^2 \to \mathbb{R}^3$ donnée par
$$f(x) = \left(x_1^3 + x_1 x_2^3, \mathrm{e}^{(x_1 - x_2)}, \cos(x_1 x_2)\right)^T.$$
50. Soit $X \in \mathrm{GL}(\mathbb{R}^n)$, calculer toutes les dérivées de la fonction $f(X) = X^3$.