COURS DE L3 : CALCUL DIFFÉRENTIEL

Laurent BRUNEAU
Université de Cergy-Pontoise

Table des matières

Chapitre 1

Espaces vectoriels normés et espaces métriques

1.1 Notion d'espace vectoriel normé

Dans tout ce cours les espaces vectoriels considérés seront des $\mathbb{R}$-espaces vectoriels. Cependant, la plupart des résultats que nous verrons sont également vrais si on considère des $\mathbb{C}$-espaces vectoriels.

Définition 1.1. *Soit E un espace vectoriel. Une application $N : E \to \mathbb{R}^+$ est une norme si*

1. *$N(x) = 0$ si et seulement si $x = 0$,*
2. *$N(\lambda x) = |\lambda| N(x)$ pour tout $\lambda \in \mathbb{R}$ et pour tout $x \in E$,*
3. *$N(x + y) \leq N(x) + N(y)$ pour tous $x, y \in E$.*

Un espace vectoriel E muni d'une norme N, noté (E, N) sera appelé un espace vectoriel normé, en abrégé evn.

Notation : Les normes seront souvent notées $\| \;\|$.

Exemple 1.1. *Soit $E = \mathbb{R}^n$. On note $x = (x_1, \ldots, x_n)$ un élément de E. Les applications ci-dessous sont des normes sur E :*

- $x \mapsto \|x\|_1 = \displaystyle\sum_{i=1}^{n} |x_i|$,
- $x \mapsto \|x\|_2 = \sqrt{\displaystyle\sum_{i=1}^{n} |x_i|^2}$,
- $x \mapsto \|x\|_\infty = \displaystyle\max_{i=1,\ldots,n} |x_i|$.

Exemple 1.2. *Soit $E = C^0([a, b])$ l'espace des fonctions continues de $[a, b]$ dans $\mathbb{R}$. Les applications ci-dessous sont des normes sur E :*

- $f \mapsto \|f\|_1 = \displaystyle\int_a^b |f(t)| \mathrm{d}t$,
- $f \mapsto \|f\|_2 = \left(\displaystyle\int_a^b |f(t)|^2 \mathrm{d}t \right)^{1/2}$,

- $f \mapsto \|f\|_\infty = \sup_{t\in[a,b]} |f(t)|$.

Exemple 1.3. *Sur* $E = \ell^1(\mathbb{N}) := \{x = (x_n)_n \in \mathbb{R}^{\mathbb{N}} \mid \sum_n |x_n| < \infty\}$, *l'application* $x \mapsto \|x\|_1 = \sum_{n\in\mathbb{N}} |x_n|$ *est une norme.*

Exercice 1.1. Soit N une norme sur un espace vectoriel E. Montrer que pour tous $x, y \in E$ on a $|N(x) - N(y)| \leq N(x - y)$ (parfois appelé inégalité triangulaire inversée).

Définition 1.2. *Soit E un espace vectoriel et N_1, N_2 deux normes sur E. On dit que N_1 et N_2 sont équivalentes s'il existe $C, C' \in \mathbb{R}_+^*$ tels que*

$$\forall x \in E, \quad CN_1(x) \leq N_2(x) \leq C'N_1(x).$$

Exemple 1.4. *Les normes $\| \ \|_1$ et $\| \ \|_\infty$ dans l'Exemple 1.1 ci-dessus sont équivalentes.*

Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Pour tout i on a $|x_i| \leq \|x\|_\infty$ et donc

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \leq n\|x\|_\infty.$$

Par ailleurs, il existe i_0 tel que $\max_{i=1,\dots,n} |x_i| = |x_{i_0}|$ et donc

$$\|x\|_\infty = |x_{i_0}| \leq \sum_{i=1}^{n} |x_i| = \|x\|_1.$$

Pour tout $x \in \mathbb{R}^n$ on a $\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$, ces deux normes sont donc équivalentes.

Exercice 1.2. Montrer que les normes $\| \ \|_2$ et $\| \ \|_\infty$ sont équivalentes. Que peut-on en déduire sur les normes $\| \ \|_1$ et $\| \ \|_2$?

Le fait que les normes $\| \ \|_1$, $\| \ \|_2$ et $\| \ \|_\infty$ sur $\mathbb{R}^n$ soient équivalentes est en fait un cas particulier du théorème important suivant :

Théorème 1.3. *Si E est de dimension finie alors toutes les normes sur E sont équivalentes.*

On démontrera ce Théorème dans le Chapitre 2 (fin de la Section 2.1).

Exemple 1.5. *Dans l'Exemple 1.2, les normes $\| \ \|_1$ et $\| \ \|_\infty$ ne sont pas équivalentes.*

Si $f \in C^0([a, b])$, pour tout $t \in [a, b]$ on a $|f(t)| \leq \|f\|_\infty$ et donc

$$\|f\|_1 = \int_a^b |f(t)|\mathrm{d}t \leq \int_a^b \|f\|_\infty \mathrm{d}t = (b - a)\|f\|_\infty.$$

On montre que par contre on ne peut pas trouver $C > 0$ tel que pour tout f on ait $\|f\|_\infty \leq C\|f\|_1$. Pour cela il suffit de construire une suite $(f_n)_n$ telle que $\|f_n\|_\infty = 1$ pour tout n mais $\|f_n\|_1 \to 0$. Soit f_n définie par

$$f_n(t) = \begin{cases} na + 1 - nt & \text{si } a \leq t < a + \frac{1}{n}, \\ 0 & \text{sinon.} \end{cases}$$

On a bien $\|f_n\|_\infty = 1$ tandis que $\int_a^b |f_n(t)|\mathrm{d}t = \frac{1}{2n}$ (représenter graphiquement la fonction f_n).

Définition 1.4. *Soient E et F deux espaces vectoriels, on appelle produit de E et F, noté $E \times F$, l'ensemble $\{(x, y) \,|\, x \in E$ et $y \in F\}$.*

Proposition 1.5. *Sur $E \times F$ on définit les lois*

- *Pour tous (x, y) et (x', y') dans $E \times F$, $(x, y) + (x', y') := (x + x', y + y')$,*
- *Pour tous (x, y) dans $E \times F$ et λ dans $\mathbb{R}$, $\lambda(x, y) := (\lambda x, \lambda y)$.*

$E \times F$ muni de ces lois est un espace vectoriel.

Exercice 1.3. Démontrer la proposition ci-dessus.

Proposition 1.6. *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ deux evn. Les applications*

$$E \times F \ni (x, y) \mapsto \|x\|_E + \|y\|_F \quad \text{et} \quad E \times F \ni (x, y) \mapsto \max(\|x\|_E, \|y\|_F)$$

définissent des normes sur l'espace vectoriel $E \times F$ et elles sont équivalentes.

Exercice 1.4. Démontrer la proposition ci-dessus.

Remarque 1.1. *De la même façon si $(E_1, \| \ \|_1), \ldots, (E_n, \| \ \|_n)$ sont des evn alors les applications $E_1 \times \cdots \times E_n \ni (x_1, \ldots, x_n) \mapsto \max\{\|x_1\|_{E_1}, \ldots, \|x_n\|_{E_n}\}$ et $E_1 \times \cdots \times E_n \ni (x_1, \ldots, x_n) \mapsto \|x_1\|_{E_1} + \ldots + \|x_n\|_{E_n}$ définissent des normes sur l'espace vectoriel $E_1 \times \cdots \times E_n$ et elles sont équivalentes.*

L'importance de la notion de normes équivalentes apparaitra par la suite quand on abordera les notions de convergence, continuité, etc. dans les evn.

1.2 Notions d'espaces métriques

Définition 1.7. *Soit E un ensemble. Une application $d : E \times E \to \mathbb{R}_+$ est appelée distance si elle vérifie :*

1. *$d(x, y) = d(y, x)$ pour tous x, y dans E,*
2. *$d(x, y) = 0$ si et seulement si $x = y$,*
3. *$d(x, z) \leq d(x, y) + d(y, z)$ pour tous x, y, z dans E.*

Un ensemble E muni d'une distance d, noté (E, d), est appelé espace métrique, en abrégé e.m.

Exemple 1.6. *Si $(E, \| \ \|)$ est un evn, alors $d(x, y) := \|x - y\|$ définit une distance sur E. On dira que d est la distance associée à la norme $\| \ \|$.*

Si $A \subset E$ (pas nécessairement un sous-espace) est non vide alors (A, d) est un espace métrique.

L'exemple ci-dessus fournit une classe importante d'espaces métriques et dans la suite de ce cours tous les espaces métriques que nous rencontrerons seront de cette forme. Ce ne sont cependant pas les seuls.

Exemple 1.7. *Soit E un ensemble. L'application d définie sur $E \times E$ par $d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{sinon} \end{cases}$ est une distance.*

Exercice 1.5. 1) Sur $\mathbb{R}$ on définit $d(x,y) = |\arctan(x) - \arctan(y)|$. Montrer que d est une distance sur $\mathbb{R}$.

2) Sur $\mathbb{R}_+$ soit $d(x,y) = |x^2 - y^2|$. Montrer que d définit une distance sur $\mathbb{R}_+$. Est-ce que d est une distance sur $\mathbb{R}$?

Définition 1.8. *Soit E un ensemble et d_1, d_2 deux distances sur E. On dit que d_1 et d_2 sont équivalentes s'il existe $C, C' \in \mathbb{R}_+^*$ tels que*

$$\forall x, y \in E, \quad Cd_1(x,y) \leq d_2(x,y) \leq C'd_1(x,y).$$

Exemple 1.8. *Si E est un espace vectoriel et N_1, N_2 deux normes sur E, alors les distances d_1 et d_2 associées respectivement à N_1 et N_2 sont équivalentes si et seulement si les normes N_1 et N_2 le sont.*

Exercice 1.6. Sur $\mathbb{R}$ on considère les distances $d_1(x,y) = |x-y|$ et $d_2(x,y) = |\arctan(x) - \arctan(y)|$. Montrer que pour tous x, y on a $d_2(x,y) \leq d_1(x,y)$ mais que ces deux distances ne sont pas équivalentes.

1.3 Topologie élémentaire des e.m.

Dans ce cours on s'intéressera à l'étude de fonctions $f : E \to F$ où E et F sont deux espaces vectoriels, ou éventuellement définies uniquement sur une partie de E. Dans cette section on introduit les notions (boules, ouverts, etc) préalables nécessaires à cette étude. Ces notions sont cependant identiques dans le cas des espaces métriques. Dans toute cette section (E,d) notera donc un espace métrique mais il faut avoir en tête que E est un evn ou une partie d'un evn et que d est la distance associée à la norme.

Définition 1.9. *Soit $x \in E$ et $r \geq 0$. On appelle boule ouverte, resp. fermée, de centre x et de rayon r l'ensemble $B(x,r) = \{y \in E \mid d(x,y) < r\}$, resp $\bar{B}(x,r) = \{y \in E \mid d(x,y) \leq r\}$.*

Exemple 1.9. *Si $\mathbb{R}$ est muni de $d(x,y) = |x-y|$ on a $B(x,r) =]x-r, x+r[$ et $\bar{B}(x,r) = [x-r, x+r]$.*

Remarque 1.2. *La notion de boule dépend du choix de la distance. Si d_1 et d_2 sont deux distances sur E alors en général $B_{d_1}(x,r) \neq B_{d_2}(x,r)$. De même la notion de boule dépend de l'espace lui-même. Par exemple, dans $(\mathbb{R}, |\ |)$ on a $B(1,2) =]-1,3[$ alors que dans $(\mathbb{R}_+, |\ |)$ on a $B(1,2) = [0,3[$.*

Exercice 1.7. Sur $\mathbb{R}^2$, déterminer et représenter graphiquement la boule $B(0,1)$ pour la distance issue de chacune des normes de l'Exemple 1.1.

Définition 1.10. *Soit $x \in E$. Un ensemble $V \subset E$ est appelé un voisinage de x s'il existe $r > 0$ tel que $B(x,r) \subset V$.*

Exemple 1.10. *Si $\mathbb{R}$ est muni de $d(x,y) = |x-y|$, l'ensemble $V =]-1,1]$ est un voisinage de 0 mais ce n'est pas un voisinage de 1 bien que $1 \in V$.*

Définition 1.11. *Un ensemble $O \subset E$ est dit ouvert si pour tout $x \in O$ il existe $r > 0$ tel que $B(x,r) \subset O$, i.e. si O est voisinage de chacun de ses points.*

Remarque 1.3. *L'ensemble vide est un ensemble ouvert ainsi que E lui même.*

Définition 1.12. *Un ensemble $F \subset E$ est dit fermé si son complémentaire $F^c = E \backslash F$ est ouvert.*

Exemple 1.11. *Pour tout $x \in E$ et $r > 0$ la boule ouverte $B(x, r)$ est un ouvert et la boule fermée $\bar{B}(x, r)$ est un fermé.*

Soit $y \in B(x, r)$ alors $d = d(x, y) < r$. Soit $r' = r - d > 0$, si $z \in B(y, r')$ alors $d(x, z) \leq d(x, y) + d(y, z) < d + r' = r$ donc $z \in B(x, r)$. Autrement dit $B(y, r') \subset B(x, r)$ et donc $B(x, r)$ est ouvert.

Montrons maintenant que $\bar{B}(x, r)$ est fermé c'est-à-dire que son complémentaire $\{y \in E \,|\, d(x, y) > r\}$ est ouvert. Soit y tel que $d = d(x, y) > r$ et $r' = d - r > 0$. Si $z \in B(y, r')$ alors $d(x, y) \leq d(x, z) + d(z, y)$, i.e. $d < d(x, z) + r' \;\Leftrightarrow\; d(x, z) > d - r' = r$ et donc $z \notin \bar{B}(x, r)$. D'où $E \backslash \bar{B}(x, r)$ est ouvert et $\bar{B}(x, r)$ est un fermé.

Proposition 1.13. *Toute réunion d'ensembles ouverts est un ouvert et toute intersection finie d'ensembles ouverts est un ouvert. Toute intersection de fermés est un fermé et toute réunion finie de fermés est un fermé.*

Démonstration. On montre le résultats sur les ouverts, celui sur les fermés en découle directement (le complémentaire d'une intersection est la réunion des complémentaires et vice versa).

Soit donc $(O_i)_{i \in I}$ une famille d'ouverts, $O = \cup_{i \in I} O_i$, et soit $x \in O$. Par définition de l'union il existe $i \in I$ tel que $x \in O_i$. Comme O_i est ouvert il existe $r > 0$ tel que $B(x, r) \subset O_i \subset O$ ce qui prouve que O est ouvert.

Soit maintenant $O_1, \ldots, O_n$ des ouverts, $O = \cap_{i=1}^{n} O_i$ et soit $x \in O$ (si O est vide il n'y a rien à montrer). Par définition, pour tout $i = 1, \ldots, n$, $x \in O_i$ qui est ouvert donc il existe $r_i > 0$ tel que $B(x, r_i) \subset O_i$. Soit $r = \min\{r_1, \ldots, r_n\} > 0$. Pour tout i on a $B(x, r) \subset B(x, r_i) \subset O_i$ et donc $B(x, r) \subset O$. □

Exemple 1.12. *On se place dans $(\mathbb{R}, |\ |)$. Pour $n \geq 1$ soit $O_n =]-\frac{1}{n}, \frac{1}{n}[$. Pour tout n l'ensemble O_n est ouvert dans $\mathbb{R}$. Par contre $O = \cap_{n \geq 1} O_n = \{0\}$ ne l'est pas.*

À la vue de cet exemple, quelle étape de la démonstration ci-dessus n'est plus vraie si on a une intersection infinie d'oueverts ?

Définition 1.14. *Soit $A \subset E$.*

- *On appelle intérieur de A l'ensemble $\overset{\circ}{A} := \{x \in A \,|\, \exists r > 0,\ B(x, r) \subset A\}$.*
- *On appelle adhérence de A l'ensemble $\bar{A} := \{x \in E \,|\, \forall r > 0,\ B(x, r) \cap A \neq \emptyset\}$.*

Remarque 1.4. *Si A est ouvert alors $\overset{\circ}{A} = A$ et si A est fermé alors $\bar{A} = A$.*

Exemple 1.13. *On se place dans $(\mathbb{R}, |\ |)$. Soit $A = [0, 1[$ alors $\overset{\circ}{A} =]0, 1[$ et $\bar{A} = [0, 1]$.*

Toujours dans $(\mathbb{R}, |\ |)$, on a $\overset{\circ}{\mathbb{Q}} = \emptyset$ et $\bar{\mathbb{Q}} = \mathbb{R}$. En effet pour tout $x \in \mathbb{R}$ et $r > 0$ il existe $q \in \mathbb{Q}$ tel que $|x - q| < r$ (voir cours de L1) ce qui montre que $B(x, r) \cap \mathbb{Q} \neq \emptyset$ et donc $\bar{\mathbb{Q}} = \mathbb{R}$. De même, pour tout $x \in \mathbb{R}$ et $r > 0$ il existe $y \in \mathbb{R} \backslash \mathbb{Q}$ tel que $|x - y| < r$ (voir cours de L1) ce qui montre que $B(x, r)$ n'est pas inclus dans $\mathbb{Q}$ et donc $\overset{\circ}{\mathbb{Q}} = \emptyset$.

Proposition 1.15. *Soit $A \subset E$. Alors $\overset{\circ}{A}$ est le plus grand ouvert inclus dans A et $\bar{A}$ est le plus petit fermé contenant A.*

Dans la plupart des livres c'est cette caractérisation qui est prise comme définition de l'inérieur et de l'adhérence. Cependant, dans la pratique on utilise plutôt directement la Définition 1.14.

Démonstration. On va montrer que $\overset{\circ}{A}$ est le plus grand ouvert inclus dans A. Il faut donc montrer que $\overset{\circ}{A}$ est ouvert et que si $O \subset A$ est un ouvert alors $O \subset \overset{\circ}{A}$. Soit donc $O \subset A$ un ouvert. Si $x \in O$ alors $x \in A$ et il existe $r > 0$ tel que $B(x, r) \subset O \subset A$, d'où $x \in \overset{\circ}{A}$. Montrons maintenant que $\overset{\circ}{A}$ est ouvert. Soit donc $x \in \overset{\circ}{A}$, on cherche $r > 0$ tel que $B(x, r) \subset \overset{\circ}{A}$. Par définition de $\overset{\circ}{A}$ il existe $r > 0$ tel que $B(x, r) \subset A$. On va montrer qu'on a en fait $B(x, r) \subset \overset{\circ}{A}$. Soit $y \in B(x, r)$ et $d = d(x, y) < r$. Si $r' = r - d$ on a, cf Exemple 1.11, $B(y, r') \subset B(x, r) \subset A$ et donc $y \in \overset{\circ}{A}$. □

Exercice 1.8. Montrer que $\bar{A}$ est le plus petit fermé contenant A.

Corollaire 1.16. *Un ensemble A est ouvert si et seulement $\overset{\circ}{A} = A$ et il est fermé si et seulement si $\bar{A} = A$.*

Tout comme pour la notion de boule, les notions de voisinage, ouvert, fermé, intérieur et adhérence dépendent a priori du choix de la distance (ainsi que de l'espace E lui-même). Cependant, on a

Proposition 1.17. *Si d_1 et d_2 sont deux distances équivalentes alors les notions de voisinage, ouvert, fermé, intérieur et adhérence coïncident, i.e. tout ouvert pour d_1 est un ouvert pour d_2 et réciproquement, et de même pour les fermés, les voisinages etc.*

Démonstration. On montre le résultat pour les ouverts. La démonstration est similaire pour les autres notions et laissée à titre d'exercice.

Par définition il existe $C, C' > 0$ tels que pour tous x, y dans E on a

$$Cd_1(x, y) \leq d_2(x, y) \leq C'd_1(x, y). \tag{1.1}$$

Soit O un ouvert pour d_1 on montre que O est ouvert pour d_2. Soit donc $x \in O$, on cherche $r > 0$ tel que $d_2(x, y) < r \ \Rightarrow y \in O$. On sait que O ouvert pour d_1 donc il existe $r_1 > 0$ tel que $d_1(x, y) < r_1 \ \Rightarrow y \in O$. En utilisant la première inégalité dans (1.1), si y vérifie $d_2(x, y) < r_1 C$ alors $d_1(x, y) \leq \dfrac{d_2(x, y)}{C} < r_1$ et donc $y \in O$. Ainsi $r = r_1 C > 0$ convient. □

Exercice 1.9. Démontrer la Proposition 1.17.

1.4 Convergence dans les espaces métriques

La définition suivante est la généralisation naturelle de la notion de limite de suite dans le cadre des espaces métriques.

Définition 1.18. *Soit $u = (u_n)_n \in E^{\mathbb{N}}$ une suite d'éléments de E. On dit que la suite u converge s'il existe $\ell \in E$ tel que*

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geq N \quad d(u_n, \ell) < \epsilon,$$

autrement dit si la suite de nombres réels $(d(u_n, \ell))_n$ tend vers 0 (dans $\mathbb{R}$).

Tout comme pour les suites de nombres réels, on a

Proposition 1.19. *Soit $u = (u_n)_n$ une suite. Si elle converge alors sa limite est unique.*

Exercice 1.10. Montrer la proposition 1.19. Indication : c'est la même preuve que dans $\mathbb{R}$.

Exercice 1.11. Soit $(u_n)_n$ telle que $u_n \to \ell$ et $x \in E$. Montrer que $d(u_n, x) \to d(\ell, x)$. (Indication : utiliser l'inégalité triangulaire.)

Attention ! Tout comme dans la section précédente, la notion de limite dépend du choix de la distance. Quand on étudiera la convergence d'une suite, on fera toujours bien attention au choix de la distance et s'il y a ambiguité on précisera pour quelle distance cette convergence a lieu.

Exemple 1.14. *Soit $E = C^0([0,1])$, on considère les (distances issues des) normes $\|f\|_1 = \displaystyle\int_0^1 |f(t)|\mathrm{d}t$ et $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$. Soit $(f_n)_n$ la suite de fonctions définies par $f_n(t) = t^n$. On va montrer que cette suite converge pour $\|\ \|_1$ (vers la fonction nulle) mais qu'elle ne converge pas pour $\|\ \|_\infty$.*

Pour tout $n \in \mathbb{N}$ on a $\|f_n - 0\|_1 = \int_0^1 |t^n - 0|\mathrm{d}t = \int_0^1 t^n \mathrm{d}t = \frac{1}{n+1}$ qui tend vers 0, donc la suite $(f_n)_n$ converge vers la fonction nulle si E est muni de la norme $\|\ \|_1$.

On munit maintenant E de la norme $\|\ \|_\infty$, on va montrer que la suite $(f_n)_n$ ne converge pas. On raisonne par contradiction. Soit $f \in E$ telle que $\|f_n - f\|_\infty \to 0$. En particulier, pour tout $t \in [0,1]$ on a $|f_n(t) - f(t)| \to 0$ (convergence uniforme implique convergence simple). Hors $f_n(t) = t^n$ tend vers 0 si $t \in [0,1[$ et vers 1 si $t = 1$. On en déduit que nécessairement $f(t) = 0$ si $t \in [0,1[$ et que $f(1) = 1$ ce qui contredit la continuité de f.

Remarque 1.5. *On peut rencontrer la situation où une suite $(u_n)_n$ converge pour deux distances d_1 et d_2 mais vers des limites différentes !*

Ce qui fait que dans l'exemple ci-dessus on ait convergence pour une norme mais pas pour l'autre vient du fait que celles-ci ne sont pas équivalentes. La proposition ci-dessous montre toute l'importance de la notion de distances/normes équivalentes.

Proposition 1.20. *Soient d_1 et d_2 des distances équivalentes et $(u_n)_n$ une suite. La suite $(u_n)_n$ converge pour d_1 si et seulement si elle converge pour d_2. Dans ce cas la limite est la même.*

Démonstration. Soient C, C' tels que $Cd_1(x,y) \leq d_2(x,y) \leq C'd_1(x,y)$ pour tous x, y dans E. Si $\ell \in E$ on a donc, pour tout n,

$$Cd_1(u_n, \ell) \leq d_2(u_n, \ell) \leq C'd_1(u_n, \ell),$$

ce qui prouve que $d_1(u_n, \ell) \to 0$ si et seulement si $d_2(u_n, \ell) \to 0$ par le Théorème des gendarmes (noter qu'une distance est toujours positive). $\square$

La proposition suivante est souvent très utile pour montrer qu'un ensemble est fermé (ou ouvert en passant au complémentaire).

Proposition 1.21. *Soit (E, d) un espace métrique et $F \subset E$. L'ensemble F est fermé si et seulement pour toute suite $(u_n)_n \in F^{\mathbb{N}}$, si $(u_n)_n$ converge alors sa limite est dans F.*

Démonstration. On suppose d'abord que F est un fermé. Soit $(u_n)_n$ une suite de F et $\ell \in E$ tels que $u_n \to \ell$. On a donc,

$$\forall \epsilon > 0, \exists N \in \mathbb{N} \,/\, \forall n \geq N, d(u_n, \ell) < \epsilon.$$

On veut montrer que $\ell \in F$. On raisonne par l'absurde. On suppose donc que $\ell \notin F$. Comme F est fermé, F^c est un ouvert et donc il existe $r > 0$ tel que $B(\ell, r) \subset F^c$, i.e. $B(\ell, r) \cap F = \emptyset$. Avec $\epsilon = r$ on a donc pour tout $n \geq N$, $u_n \in B(\ell, r)$. Or $u_n \in F$, ce qui contredit $B(\ell, r) \cap F = \emptyset$.

Réciproquement, on raisonne par contraposition. Si F n'est pas fermé, F^c n'est pas ouvert. Il existe donc $\ell \in F^c$ tel que pour tout $r > 0$, $B(\ell, r) \not\subset F^c$, i.e. $B(\ell, r) \cap F \neq \emptyset$. On prend $r = 1/n$. Il existe donc $u_n \in B(\ell, 1/n) \cap F$. On obtient donc une suite $(u_n)_n$ d'éléments de F qui vérifie pour tout n, $d(u_n, \ell) < 1/n$ et donc qui converge vers ℓ. On a donc construit une suite d'éléments de F qui converge mais dont la limite n'est pas dans F. $\square$

On rappelle qu'une sous-suite, ou suite extraite, d'une suite $(u_n)_n$ est une suite $(u_{\varphi(n)})_n$ où $\varphi : \mathbb{N} \to \mathbb{N}$ est strictement croissante et que si la suite $(u_n)_n$ converge vers ℓ alors toute sous-suite de $(u_n)_n$ converge aussi vers ℓ (vous l'avez vu dans $\mathbb{R}$, c'est identique dans les espaces métriques).

Exercice 1.12. On dit que ℓ est une valeur d'adhérence de la suite $(u_n)_n$ s'il existe une sous-suite de $(u_n)_n$ qui converge vers ℓ et on note A l'ensemble des valeurs d'adhérence de $(u_n)_n$. On note aussi $U_n := \{u_k,\, k \geq n\}$. Montrer que $A = \cap_{n \in \mathbb{N}} \bar{U}_n$.

Définition 1.22. *Un ensemble $K \subset E$ est dit compact si toute suite $(u_n)_n$ d'éléments de K admet une sous-suite qui converge dans K (i.e. la limite est aussi dans K). Autrement dit, il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante et $\ell \in K$ tels que $\lim_{n \to \infty} u_{\varphi(n)} = \ell$.*

Comme pour les notions d'ouvert, fermé, etc. la notion d'ensemble compact dépend de la distance (ou norme).

Exercice 1.13. Si d_1 et d_2 sont équivalentes, montrer que K est compact pour d_1 si et seulement si il est compact pour d_2.

On rappelle qu'un ensemble A est borné s'il est inclus dans une boule (dans un evn celà revient à dire qu'il existe $M > 0$ tel que pour tout $x \in A$ on a $\|x\| \leq M$). Dans $\mathbb{R}^n$, et plus généralement dans tout espace de dimension finie (pour n'importe quelle norme puisque celles-ci sont toutes équivalentes en dimension finie), les ensembles compacts sont les ensembles fermés et bornés (cf cours de L2, analyse dans $\mathbb{R}^n$). Ce n'est pas forcément le cas en général, voir Exemple 1.15, mais on a cependant

Proposition 1.23. *Si $K \subset E$ est compact alors K est fermé et borné.*

Démonstration. La preuve est la même que dans $\mathbb{R}^n$. On montre d'abord que K est borné. En effet sinon, étant donné $x \in E$, on pourrait construire une suite $(u_n)_n$ telle que $d(u_n, x) \geq n$ pour tout n. En particulier $\lim_{n\to\infty} d(u_n, x) = +\infty$. Par ailleurs, puisque K est compact, il existe φ strictement croissante et $\ell \in K$ tels que $\lim_{n\to\infty} u_{\varphi(n)} = \ell$. On a alors, voir Exercice 1.11, $\lim_{n\to\infty} d(u_{\phi(n)}, x) = d(\ell, x)$ ce qui contredit $\lim_{n\to\infty} d(u_n, x) = +\infty$.

Montrons maintenant que K est fermé. On utilise la Proposition 1.21. Soit $(u_n)_n$ une suite de K qui converge vers $\ell \in E$. On veut montrer que $\ell \in K$. Comme K est compact, il existe une sous-suite $(u_{\varphi(n)})_n$ qui converge et dont la limite est dans K. Par ailleurs, puisque $(u_n)_n$ converge vers ℓ, la sous-suite $(u_{\varphi(n)})_n$ converge également vers ℓ. Ce qui prouve que $\ell \in K$. □

Exemple 1.15. *Soit $E = \mathbb{R}[X]$, si $P(X) = \sum_{k=0}^{n} a_k X^k$ on définit $N(P) = \max\{|a_0|, \dots, |a_n|\}$. N définit bien une norme sur E (montrez le).*

Soit $K = \bar{B}(0, 1)$ la boule fermée de centre 0 et de rayon 1. K est fermé et borné. On montre cependant que K n'est pas compact. On considère la suite $(P_n)_n$ définie par $P_n(X) = X^n$. Pour tout n on a $P_n \in K$, cependant cette suite n'admet pas de sous-suite convergente. On raisonne par l'absurde, on suppose que $P_{\varphi(n)} \to Q$ où φ est strictement croissante et $Q \in K$. Si $Q = \sum_{k=0}^{q} a_k X^k$ où q est le degré de Q, alors pour tout $n > q$ on a $\varphi(n) > q$, donc $P_{\varphi(n)} - Q = X^{\varphi(n)} - \sum_{k=0}^{q} a_k X^k$ et ainsi $N(P_{\varphi(n)} - Q) \geq 1$ ce qui contredit $P_{\varphi(n)} \to Q$.

Proposition 1.24. *Soit K un compact et $F \subset K$ un fermé, alors F est compact.*

Démonstration. Soit $(u_n)_n$ une suite d'éléments de F. On montre qu'elle possède une sous-suite convergeant dans F. Comme $F \subset K$, $(u_n)_n$ est une suite d'éléments de K qui est compact donc elle possède une sous-suite $(u_{\varphi(n)})_n$ qui converge vers $\ell \in K$. La suite $(u_{\varphi(n)})_n$ est une suite de F qui converge, et comme F est fermé sa limite ℓ est dans F. La suite $(u_n)_n$ possède donc bien une sous-suite qui converge dans F. □

Finalement on rappelle la notion de suite de Cauchy.

Définition 1.25. *Une suite $(u_n)_n \in E^{\mathbb{N}}$ est dite de Cauchy si*

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N},\ \forall n, m \geq N \quad d(u_n, u_m) < \epsilon.$$

On sait que dans $\mathbb{R}^n$ (et donc dans tout espace de dimension finie) une suite est de Cauchy si et seulement si elle converge. En général on a toujours

Proposition 1.26. *Si une suite converge alors elle est de Cauchy.*

Exercice 1.14. Démontrer la Proposition.

Cependant la réciproque n'est plus forcément vraie.

Exemple 1.16. *Soit $E = C^0([0,1])$ muni de la norme $\|f\|_1 = \int_0^1 |f(x)|\mathrm{d}x$. On considère la suite $(f_n)_{n\geq 2}$ définie par*

$$f_n(x) = \begin{cases} 0 & \text{si } x < \frac{1}{2}, \\ n(x-\frac{1}{2}) & \text{si } \frac{1}{2} \leq x \leq \frac{1}{2} + \frac{1}{n} \\ 1 & \text{si } x > \frac{1}{2} + \frac{1}{n} \end{cases}$$

Cette suite est de Cauchy mais ne converge pas.

Pour tous $n \geq m$ on calcule facilement $\|f_n - f_m\|_1 = \dfrac{1}{2}\left(\dfrac{1}{n} - \dfrac{1}{m}\right)$ (faites un dessin !) et donc la suite est de Cauchy.

Supposons maintenant qu'elle converge et notons f sa limite. Pour tout n on a

$$\int_0^{1/2} |f(x)|\mathrm{d}x = \int_0^{1/2} |f(x) - f_n(x)|\mathrm{d}x \leq \|f - f_n\|_1 \to 0,$$

ce qui prouve que $\int_0^{1/2} |f(x)|\mathrm{d}x = 0$. Comme la fonction $|f|$ est positive <u>et continue</u> sur $[0, \frac{1}{2}]$ on en déduit que f est nulle sur $[0, \frac{1}{2}]$.

Soit $\frac{1}{2} < a \leq 1$, pour n assez grand on a $\frac{1}{2} + \frac{1}{n} < a$ et donc

$$\int_a^1 |f(x) - 1|\mathrm{d}x = \int_a^1 |f(x) - f_n(x)|\mathrm{d}x \leq \|f - f_n\|_1 \to 0.$$

On en déduit que $f(x) = 1$ pour tout $x \in [a, 1]$. Ceci étant vrai pour tout $a > \frac{1}{2}$ on a ainsi $f(x) = 1$ si $x > \frac{1}{2}$.

Si la suite $(f_n)_n$ convergeait sa limite f vaudrait 0 sur $[0, \frac{1}{2}]$ et 1 sur $]\frac{1}{2}, 1]$ ce qui contredit sa continuité.

Définition 1.27. *Un espace métrique (E, d) est dit complet si toute suite de Cauchy converge. Si $(E, \| \ \|)$ est un evn et qu'il est complet pour la distance associée à la norme on dira que c'est un espace de Banach.*

Les espaces de Banach joueront un rôle important par la suite.

Exemple 1.17. *Tout evn de dimension finie est un espace de Banach.*

Exemple 1.18. *L'espace $E = C^0([0,1])$ muni de la norme $\|f\|_\infty = \sup\limits_{x\in[0,1]} |f(x)|$ est complet.*

Soit $(f_n)_n$ une suite de Cauchy dans E, on montre qu'elle converge. Soit $x \in [0,1]$, pour tous n, m on a $|f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty$ et on en déduit que la suite de nombres réels $(f_n(x))_n$ est de Cauchy (dans $\mathbb{R}$!). Comme $\mathbb{R}$ est complet cette suite converge, on note $f(x)$ sa limite. On va montrer que la fonction f ainsi définie est dans E et que $\|f_n - f\|_\infty \to 0$.

Soit $\epsilon > 0$. Comme la suite $(f_n)_n$ est de Cauchy dans E il existe $N \in \mathbb{N}$ tel que pour tous $n, m \geq N$ on ait $\|f_n - f_m\|_\infty < \epsilon$. On a donc

$$\forall n, m \geq N,\ \forall x \in [0,1],\ |f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty < \epsilon.$$

En faisant tendre m vers l'infini on en déduit que

$$\forall n \geq N,\ \forall x \in [0,1],\ |f_n(x) - f(x)| \leq \epsilon,$$

et donc $\|f_n - f\|_\infty \leq \epsilon$ *ce qui prouve que* $\|f_n - f\|_\infty \to 0$.

Il reste à montrer que $f \in E$. *Les fonctions* f_n *sont continues et* $\|f_n - f\|_\infty \to 0$ *signifie que la suite de fonctions* $(f_n)_n$ *converge uniformément vers* f. *Cette dernière est donc continue (voir cours de L2 sur les suites de fonctions).*

Comme le montrent les exemples ci-dessus, le fait pour un espace d'être complet ou non dépend aussi du choix de la norme. Cependant on a

Proposition 1.28. *Soient* d_1 *et* d_2 *deux métriques équivalentes, alors* (E, d_1) *est complet si et seulement si* (E, d_2) *est complet.*

Démonstration. On montre que si (E, d_1) est complet alors (E, d_2) aussi. La réciproque s'obtient en inversant les rôles de d_1 et d_2. Soit donc $(u_n)_n$ une suite de Cauchy pour d_2, i.e.

$$\forall \epsilon > 0, \ \exists N \in \mathbb{N}, \ \forall n, m \geq N \quad d_2(u_n, u_m) < \epsilon.$$

Puisque d_1 et d_2 sont équivalentes il existe $C, C' > 0$ tels que pour tous x, y on ait $Cd_1(x, y) \leq d_2(x, y) \leq C'd_1(x, y)$. En particulier pour $n, m \geq N$ on a $d_1(u_n, u_m) \leq C^{-1}d_2(u_n, u_m) < C^{-1}\epsilon$. La suite $(u_n)_n$ est donc de Cauchy aussi pour d_1. Comme (E, d_1) est complet il existe $\ell \in E$ tel que $u_n \to \ell$ pour d_1. Les distances d_1 et d_2 sont équivalentes donc, Proposition 1.20, $u_n \to \ell$ aussi pour d_2. □

Chapitre 2

Continuité dans les evn

Dans tout ce chapitre $(E, \| \ \|)$ notera un espace vectoriel normé.

2.1 Fonctions continues

Définition 2.1. *Soient $(E, \| \ \|_E)$, $(F, \| \ \|_F)$ deux evn, $A \subset E$ et $f : A \to F$ une application. On dit que f est continue en $x_0 \in A$ si*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall x \in A,\ \|x - x_0\|_E < \delta \implies \|f(x) - f(x_0)\|_F < \epsilon.$$

La fonction f est dite continue sur A si f est continue en tout point $x_0 \in A$.

Exercice 2.1. Montrer que f est continue en x_0 si et seulement si pour toute suite $(x_n)_n \in A^{\mathbb{N}}$ qui converge vers x_0 (dans A pour la norme $\| \ \|_E$) la suite $(f(x_n))_n$ converge vers $f(x_0)$ (dans F pour la norme $\| \ \|_F$).

Exercice 2.2. Soit $(E, \| \ \|)$ un evn. Montrer que l'application $x \mapsto \|x\|$ est continue de E dans $\mathbb{R}$.

Tout comme la notion de convergence des suites, la notion de continuité dépend des normes choisies (dans E et dans F).

Exemple 2.1. *Soit $E = F = \ell^1(\mathbb{N}) = \{u = (u_n)_n \mid \sum_n |u_n| < \infty\}$. On note $\|u\|_1 = \sum |u_n|$ et $\|u\|_\infty = \sup\limits_{n \in \mathbb{N}} |u_n|$ (vérifier que ce sont des normes sur $\ell^1(\mathbb{N})$). On considère $\Phi : E \to F$ définie par $\Phi(u) = u^2$, i.e. l'image par Φ de la suite $(u_n)_n$ est la suite $(u_n^2)_n$.*

L'application Φ est bien définie, i.e. si $u \in E$ alors $\Phi(u) \in F$. En effet, puisque la série $\sum |u_n|$ converge la suite $(u_n)_n$ tend vers 0 et est donc bornée : il existe $M > 0$ tel que $|u_n| \leq M$ pour tout n. On a alors $|u_n^2| \leq M|u_n|$ pour tout n et par comparaison des séries à termes positifs la série $\sum |u_n^2|$ converge.

On montre que Φ est continue en 0 de $(E, \| \ \|_1)$ dans $(F, \| \ \|_1)$ ainsi que de $(E, \| \ \|_\infty)$ dans $(F, \| \ \|_\infty)$, mais pas de $(E, \| \ \|_\infty)$ dans $(F, \| \ \|_1)$. On remarque que $\Phi(0) = 0$.

- *De $(E, \| \ \|_1)$ dans $(F, \| \ \|_1)$. On remarque que pour tout n on a $|u_n| \leq \|u\|_1$ et donc $\|\Phi(u)\|_1 = \sum |u_n|^2 \leq \sum \|u\|_1 \times |u_n| = \|u\|_1^2$. Soit $\epsilon > 0$, on pose $\delta = \sqrt{\epsilon}$. Si $\|u\|_1 < \delta$ alors on a bien $\|\Phi(u)\|_1 < \epsilon$.*
- *De $(E, \| \ \|_\infty)$ dans $(F, \| \ \|_\infty)$, c'est le même raisonnement que ci-dessus.*

– *De* $(E, \| \|_\infty)$ *dans* $(F, \| \|_1)$. *Pour* $N \in \mathbb{N}$, *soit* $u^{(N)}$ *la suite définie par* $u_n^{(N)} = \frac{1}{\sqrt{N+1}}$ *si* $n \leq N$ *et* $u_n^{(N)} = 0$ *sinon. On a* $\|u^{(N)}\|_\infty = \frac{1}{\sqrt{N+1}}$ *et donc* $u^{(N)} \to 0$ *dans* E. *Par contre* $\|\Phi(u^{(N)})\|_1 = \sum |u_n^{(N)}|^2 = 1$ *et donc* $\Phi(u^{(N)})$ *ne tend pas vers* $\Phi(0)$ *dans* $(F, \| \|_1)$. *La fonction* Φ *n'est donc pas continue en* 0.

On a les même propriétés pour les fonctions continues sur des evn quelconques que dans $\mathbb{R}$ ou $\mathbb{R}^n$.

Proposition 2.2. *Soient* $A \subset E$ *et* $x \in A$. *Si* $f, g : A \to F$ *sont continues en* x *alors* $f + g$ *est continue en* x.

Proposition 2.3. *Soient* $A \subset E$, $x \in A$, $f : A \to F$ *continue en* x *et* $g : F \to G$ *continue en* $f(x)$. *Alors la fonction* $g \circ f : A \to G$ *est continue en* x.

Exercice 2.3. Démontrer ces deux propositions.

Proposition 2.4. *Soit* $A \subset E$ *et* $f : A \to F$. *Les propriétés suivantes sont équivalentes*

1. f *est continue sur* A,
2. *Pour tout ouvert* U *de* F, $f^{-1}(U)$ *est un ouvert de* A,
3. *Pour tout fermé* V *de* F, $f^{-1}(V)$ *est un fermé de* A.

Démonstration. L'équivalence entre 2. et 3. est immédiate par passage au complémentaire. Montrons que 1. et 2. sont équivalentes.

Supposons que f est continue. Soit U un ouvert de F et $x_0 \in f^{-1}(U)$. On a donc $f(x_0) \in U$ qui est ouvert donc il existe $\epsilon > 0$ tel que $B(f(x_0), \epsilon) \subset U$. Soit $\delta > 0$ tel $\|x - x_0\| < \delta \;\Rightarrow \|f(x) - f(x_0)\| < \epsilon$. Si $x \in B(x_0, \delta)$ on a donc $f(x) \in B(f(x_0), \epsilon) \subset U$ et donc $x \in f^{-1}(U)$, i.e. $B(x_0, \delta) \subset f^{-1}(U)$, ce dernier est donc ouvert.

Réciproquement, supposons que 2. est vraie. Soit $x_0 \in A$ et $\epsilon > 0$. La boule $B(f(x_0), \epsilon)$ est un ouvert de F donc $f^{-1}(B(f(x_0), \epsilon))$ est un ouvert de A. Or $x_0 \in f^{-1}(B(f(x_0), \epsilon))$ donc il existe $\delta > 0$ tel que $B(x_0, \delta) \subset f^{-1}(B(f(x_0), \epsilon))$, i.e. si $\|x - x_0\| < \delta$ on a $\|f(x) - f(x_0)\| < \epsilon$. C'est précisément la définition de la continuité. □

Théorème 2.5. *Soit* $K \subset E$ *un compact et* $f : K \to F$ *une fonction continue. Alors* $f(K)$ *est un compact de* F.

Démonstration. Soit $(y_n)_n$ une suite de $f(K)$. Par définition, pour tout n il existe $x_n \in K$ tel que $f(x_n) = y_n$. La suite $(x_n)_n$ est dans K qui est compact donc elle admet une sous-suite $(x_{\varphi(n)})_n$ convergeant vers $x \in K$. Comme f est continue, la suite $f(x_{\varphi(n)}) = y_{\varphi(n)}$ converge vers $f(x) \in f(K)$. La suite $(y_n)_n$ possède donc une sous-suite convergente dans $f(K)$, ce qui prouve que $f(K)$ est compact. □

Le théorème important qui suit, et que vous avez déjà rencontré dans $\mathbb{R}^n$, découle du précédent.

Théorème 2.6. *Soit* $K \subset E$ *un compact et* $f : K \to \mathbb{R}$ *une fonction continue. Alors* f *est bornée et atteint ses bornes, i.e. il existe* $x^m, x^M \in K$ *tels que pour tout* $x \in K$ *on ait* $f(x^m) \leq f(x) \leq f(x^M)$.

Démonstration. D'après le théorème précédent l'ensemble $f(K)$ est un compact de $\mathbb{R}$ et donc en particulier est borné, i.e. la fonction f est bornée. On montre maintenant qu'elle atteint ses bornes. On montre le résultat pour la borne supérieure, i.e. qu'il existe $x^M \in K$ tel que $M := \sup_{x\in K} f(x) = f(x^M)$.

Soit donc M la borne supérieure de $f(K)$. Il existe $(y_n)_n$ dans $f(K)$ telle que $y_n \to M$. Comme $f(K)$ est compact il est fermé et donc, Proposition 1.21, $M \in f(K)$, i.e. il existe $x_M \in K$ tel que $f(x_M) = M$. □

On rappelle également la notion de fonction uniformément continue

Définition 2.7. *Soit $A \subset E$ et $f : A \to F$. On dit que f est uniformément continue sur A si*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall x, y \in A,\ \|x - y\|_E < \delta \implies \|f(x) - f(y)\|_F < \epsilon.$$

Dire que f est continue sur A s'écrit

$$\forall y \in A,\ \forall \epsilon > 0,\ \exists \delta > 0,\ \forall x \in A,\ \|x - y\|_E < \delta \implies \|f(x) - f(y)\|_F < \epsilon.$$

La notion de continuité uniforme est plus forte que celle de continuité : on a intervertit les quantificateurs $\forall y \in A$ et $\exists \delta > 0$. Pour une fonction uniformément continue, le δ à partir duquel $\|f(x) - f(y)\|_F < \epsilon$ est satisfait ne dépend pas du point y considéré.

Exemple 2.2. *La fonction $f\colon]0, 1] \to \mathbb{R}$ définie par $f(x) = \dfrac{1}{x}$ est continue mais n'est pas uniformément continue.*

Théorème 2.8. *Toute fonction définie et continue sur un compact K est uniformément continue sur K.*

Démonstration. On va raisonner par l'absurde. On suppose que f n'est pas uniformément continue. On a donc

$$\exists \epsilon > 0,\ \forall \delta > 0, \exists x, y \in K,\ \|x - y\|_E < \delta \text{ et } \|f(y) - f(x)\|_F \geq \epsilon.$$

Pour tout entier $n \geq 1$, on applique ceka avec $\delta = \frac{1}{n}$. On construit ainsi deux suites $(x_n)_n$ et $(y_n)_n$ d'éléments de K tels que $\|x_n - y_n\|_E < \frac{1}{n}$ et $\|f(x_n) - f(y_n)\|_F \geq \epsilon$ pour tout k.

Puisque K est compact, il existe φ strictement croissante et $x \in K$ tels que la suite $(x_{\varphi(n)})_n$ converge vers x. De plus, pour tout n,

$$\|x_{\varphi(n)} - y_{\varphi(n)}\|_E < \frac{1}{\varphi(n)} \leq \frac{1}{n} \quad \text{et} \quad \|f(x_{\varphi(n)}) - f(y_{\varphi(n)})\|_F \geq \epsilon.$$

On considère maintenant la suite $(y_{\varphi(n)})_n$ qui est une suite d'éléments de K. Il existe donc ψ strictement croissante et $y \in K$ tels que la suite $y_{\varphi(\psi(n))}$ converge vers y. Si on pose $\phi = \varphi \circ \psi$, on a donc $x_{\phi(n)}$ qui est une sous-suite de $x_{\varphi(n)}$ et donc converge vers x, et $y_{\phi(n)}$ qui converge vers y. De plus, pour tout n,

$$\|x_{\phi(n)} - y_{\phi(n)}\|_E < \frac{1}{\phi(n)} \leq \frac{1}{n} \quad \text{et} \quad \|f(x_{\phi(n)}) - f(y_{\phi(n)})\|_F \geq \epsilon.$$

La première inégalité entraine que $x = y$, tandis que la seconde, puisque f est continue, entraine $\|f(x) - f(y)\|_F \geq \epsilon$, ce qui est absurde. □

On termine cette section en démontrant le Théorème 1.3.

Démonstration. Soit E un espace de dimension finie. On note n sa dimension et on se donne une base $(e_1, \ldots, e_n)$ de E. Pour $x = \sum_{i=1}^{n} x_i e_i$ on note $\|x\| = \max_i |x_i|$. On montre facilement que l'application $x \mapsto \|x\|$ est une norme sur E (vérifiez le). On va montrer que n'importe quelle norme N sur E est équivalente à $\| \ \|$ (elles seront donc toutes équivalentes deux à deux). Soit donc N une norme, on cherche $C, C' > 0$ tels que pour tout $x \in E$ on ait

$$C\|x\| \leq N(x) \leq C'\|x\|$$

Soit $x \in E$, en utilisant l'inégalité triangulaire on a

$$N(x) = N\left(\sum_{i=1}^{n} x_i e_i\right) \leq \sum_{i=1}^{n} N(x_i e_i) = \sum_{i=1}^{n} |x_i| N(e_i) \leq \|x\| \times \sum_{i=1}^{n} N(e_i).$$

En posant $C' = \sum_{i=1}^{n} N(e_i) > 0$ on a donc montré que pour tout $x \in E$ on avait bien $N(x) \leq C'\|x\|$.

On considère maintenant N comme application de $(E, \| \ \|)$ dans $(\mathbb{R}, | \ |)$. On montre que N est continue. Soit $x_0 \in E$, pour tout $x \in E$ on a

$$|N(x) - N(x_0)| \leq N(x - x_0) \leq C'\|x - x_0\|.$$

Étant donné $\epsilon > 0$ si on pose $\delta = \dfrac{\epsilon}{C'}$ on a bien $\|x - x_0\| < \delta \Rightarrow |N(x) - N(x_0)| < \epsilon$ et donc N est continue.

On pose $K := \{x \in E \mid \|x\| = 1\}$. L'ensemble K est un fermé borné dans $(E, \| \ \|)$ donc il est compact (dimension finie). La fonction N admet donc une borne inférieure C sur l'ensemble K et celle-ci est atteinte, i.e. il existe $x_0 \in K$ tel que pour tout $x \in K$ on ait $C = N(x_0) \leq N(x)$. Comme $x_0 \neq 0$ on a donc $C > 0$. On va montrer que pour tout $x \in E$ on a $C\|x\| \leq N(x)$. C'est vrai pour $x = 0$. Soit donc $x \neq 0$, on a $\dfrac{x}{\|x\|} \in K$ donc

$$C \leq N\left(\frac{x}{\|x\|}\right) = \frac{1}{\|x\|} N(x) \iff C\|x\| \leq N(x).$$

□

2.2 Applications linéaires continues

Parmi les application d'un espace vectoriel E dans un espace vectoriel F il y en a qui jouent un rôle important, ce sont les applications linéaires. Dans cette section on s'intéresse aux applications linéaires continues d'un evn dans un autre. Si E et F sont deux espaces vectoriels, on notera $\mathcal{L}(E,F)$ l'ensemble des applications linéaires de E dans F (ou plus simplement $\mathcal{L}(E)$ si $E=F$). C'est en fait un espace vectoriel.

La proposition suivante est souvent très utile pour savoir si une application linéaire est continue.

Proposition 2.9. *Soient E et F deux evn et $f \in \mathcal{L}(E,F)$. Les assertions suivantes sont équivalentes*

1. *f est continue sur E,*
2. *f est continue en 0,*
3. *f est bornée sur la boule unité fermée de E, i.e. $\{f(x) \mid x \in E, \|x\|_E \leq 1\}$ est borné,*
4. *f est bornée sur la sphère unité de E, i.e. $\{f(x) \mid x \in E, \|x\|_E = 1\}$ est borné,*
5. *Il existe $C \geq 0$ tel que pour tout $x \in E$,*

$$\|f(x)\|_F \leq C\|x\|_E.$$

Démonstration. 1. $\Rightarrow$ 2. est évident. On suppose maintenant que f est continue en 0. Soit $x \in E$ et $(x_n)_n$ telle que $x_n \to x$. On a $x_n - x \to 0$ et comme f est continue en 0 on a $f(x_n - x) \to f(0) = 0$ (f est linéaire). Comme $f(x_n - x) = f(x_n) - f(x)$ on en déduit que $f(x_n) \to f(x)$ et donc f est continue en x. On a donc aussi 2. $\Rightarrow$ 1.

On montre maintenant que 2. $\Rightarrow$ 3. $\Rightarrow$ 4. $\Rightarrow$ 5. $\Rightarrow$ 2.

2. $\Rightarrow$ 3. On applique la définition de la continuité en 0 avec $\epsilon = 1$. Il existe $\delta > 0$ tel que $\|x\|_E < \delta \;\Rightarrow\; \|f(x)\|_F < 1$. Soit $x \in E$ tel que $\|x\|_E \leq 1$. On a alors $y = \frac{\delta}{2}x$ vérifie $\|y\|_E \leq \frac{\delta}{2} < \delta$ donc $\|f(y)\|_F < 1$. Comme f est linéaire on a $f(y) = \frac{\delta}{2}f(x)$ et donc $\|f(y)\|_F = \frac{\delta}{2}\|f(x)\|_F < 1$ d'où $\|f(x)\|_F < \frac{2}{\delta}$ pour tout x dans la boule unité.

3. $\Rightarrow$ 4. Comme la sphère unité est incluse dans la boule unité le résultat est immédiat.

4. $\Rightarrow$ 5. f est bornée sur la sphère unité, soit donc C tel que $\|f(x)\|_F \leq C$ pour tout $x \in E$ tel que $\|x\|_E = 1$. Soit $x \in E$. Si $x = 0$ alors $\|f(x)\|_F \leq C\|x\|_E$ est évident. Si $x \neq 0$, $\dfrac{x}{\|x\|_E}$ est de norme 1 donc

$$\left\| f\left(\frac{x}{\|x\|_E}\right) \right\|_F \leq C \quad\Longleftrightarrow\quad \frac{1}{\|x\|_E}\|f(x)\|_F \leq C \quad\Longleftrightarrow\quad \|f(x)\|_F \leq C\|x\|_E.$$

5. $\Rightarrow$ 2. est immédiat. □

L'ensemble des applications linéaires continues forme un sous-espace vectoriel de $\mathcal{L}(E,F)$ que l'on notera $\mathcal{L}_c(E,F)$ ou $L(E,F)$. Lorsque $E=F$ on notera simplement $\mathcal{L}_c(E)$ ou $L(E)$. Bien que la notation ne l'indique pas, on fera bien attention au fait que $L(E,F)$ ne dépend pas que de E et F mais aussi des normes $\|\ \|_E$ et $\|\ \|_F$ choisies sur ces espaces.

Remarque 2.1. *Les caractérisations 3. et 4. font qu'on parle aussi d'application linéaire bornée au lieu d'application linéaire continue. Cela signifie la même chose.*

Exemple 2.3. *On considère $E = F = \mathbb{R}[X]$ munis de la norme suivante : si $P(X) = \sum_{k=0}^{n} a_k X^k$, $\|P\| = \max\{|a_0|, \dots, |a_n|\}$.*

Soit $f : \mathbb{R}[X] \to \mathbb{R}[X]$ définie par $f(P)(X) = XP(X)$. f est bien une application linéaire. Si $P(X) = \sum_{k=0}^{n} a_k X^k$ alors $f(P)(X) = \sum_{k=0}^{n} a_k X^{k+1} = \sum_{k=1}^{n+1} a_{k-1} X^k$. On a donc $\|f(P)\| = \|P\|$ et la propriété 5. ci-dessus est vraie. Donc f est continue (pour la norme choisie).

Soit $g : \mathbb{R}[X] \to \mathbb{R}[X]$ définie par $g(P) = P'$. g est bien une application linéaire. Pour tout entier n, soit $P_n(X) = X^n$. On a $\|P_n\| = 1$ tandis que $g(P_n)(X) = nX^{n-1}$ et donc $\|g(P_n)\| = n$. L'application g n'est pas bornée sur la sphère unité, elle n'est donc pas continue (pour la norme choisie).

Exercice 2.4. Sur $\mathbb{R}[X]$, si $P(X) = \sum_{k=0}^{n} a_k X^k$, on définit $\|P\| = \sum_{k=0}^{n} |a_k| \times k!$. Montrer que cela définit bien une norme et que $g : P \mapsto P'$ est continue de $(\mathbb{R}[X], \| \, \|)$ dans $(\mathbb{R}[X], \| \, \|)$.

Attention ! Cette caractérisation de la continuité ne s'applique qu'aux applications linéaires.

Exercice 2.5. Soit $E = \mathbb{R}$ muni de la valeur absolue et $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sqrt{|x|}$. Vérifier que f est continue mais que 5. n'est pas vraie.

Exercice 2.6. Soit $E = \mathbb{R}$ muni de la valeur absolue et $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x$ si $x \in [-1, 1]$ et $f(x) = 0$ sinon. Vérifiez que 3., 4. et 5. sont vraies mais que f n'est pas continue.

Lorsqu'on est en dimension finie, la situation est plus simple : la continuité des applications linéaires est en fait automatique.

Proposition 2.10. *Soient $(E, \| \, \|_E)$ et $(F, \| \, \|_F)$ deux evn. Si E est de dimension finie alors toute application linéaire de E dans F est continue.*

Démonstration. Soit $f : E \to F$ une application linéaire et $(e_1, \ldots, e_n)$ une base de E. Si $x = \sum_{i=1}^{n} x_i e_i$ alors $\|x\| = \max |x_i|$ définit une norme sur E, et puisque E est de dimension finie elle est équivalente à $\| \, \|_E$. Soit $c > 0$ tel que $\|x\| \leq c\|x\|_E$ pour tout $x \in E$. On a alors

$$\|f(x)\|_F \leq \sum_{i=1}^{n} |x_i| \|f(e_i)\|_F \leq \|x\| \sum_{i=1}^{n} \|f(e_i)\|_F \leq C\|x\|_E$$

avec $C := c \sum_{i=1}^{n} \|f(e_i)\|_F$, ce qui prouve que f est continue. □

Proposition 2.11. *Soit $f \in \mathcal{L}(E, F)$, alors on a*

$$\sup_{x \in E, \, \|x\|_E = 1} \|f(x)\|_F = \sup_{x \in E, \, \|x\|_E \leq 1} \|f(x)\|_F = \sup_{x \in E, \, x \neq 0} \frac{\|f(x)\|_F}{\|x\|_E}.$$

Ces quantités sont finies si et seulement si f est continue et on note alors $|||f|||_{L(E,F)}$, ou simplement $|||f|||$ s'il n'y a pas d'ambiguité, leur valeur commune.

Attention ! Bien qu'on ne l'indique pas explicitement pour ne pas trop alourdir la notation, on fera bien attention au fait que la quantité $|||f|||_{L(E,F)}$ dépend des normes choisies sur les espaces E et F (voir l'Exemple 2.4 et l'Exercice 2.9).

Démonstration. Le fait que ces quantités soient finies si et seulement si f est continue est exactement le contenu de la Proposition 2.9.

On a $\{x \in E, \|x\|_E = 1\} \subset \{x \in E, \|x\|_E \leq 1\}$ et donc $\sup\limits_{x\in E,\, \|x\|_E=1} \|f(x)\|_F \leq \sup\limits_{x\in E,\, \|x\|_E\leq 1} \|f(x)\|_F$.

Si $x = 0$ on a $f(x) = 0$ et donc $\|f(x)\|_F \leq \sup\limits_{x\in E,\, x\neq 0} \dfrac{\|f(x)\|_F}{\|x\|_E}$. Si $\|x\|_E \leq 1$ et $x \neq 0$ alors $1 \leq \dfrac{1}{\|x\|_E}$ et donc $\|f(x)\|_F \leq \dfrac{\|f(x)\|_F}{\|x\|_E}$. D'où

$$\sup_{x\in E,\, \|x\|_E\leq 1} \|f(x)\|_F \leq \sup_{x\in E,\, \|x\|_E\leq 1} \frac{\|f(x)\|_F}{\|x\|_E} \leq \sup_{x\in E,\, x\neq 0} \frac{\|f(x)\|_F}{\|x\|_E}.$$

Finalement si $x \in E$ est non nul on a

$$\frac{\|f(x)\|_F}{\|x\|_E} = \left\| f\left(\frac{x}{\|x\|_E}\right)\right\|_F \leq \sup_{x\in E,\, \|x\|_E=1} \|f(x)\|_F,$$

et donc $\sup\limits_{x\in E,\, x\neq 0} \dfrac{\|f(x)\|_F}{\|x\|_E} \leq \sup\limits_{x\in E,\, \|x\|_E=1} \|f(x)\|_F$.

On a montré

$$\sup_{x\in E,\, \|x\|_E=1} \|f(x)\|_F \leq \sup_{x\in E,\, \|x\|_E\leq 1} \|f(x)\|_F \leq \sup_{x\in E,\, x\neq 0} \frac{\|f(x)\|_F}{\|x\|_E} \leq \sup_{x\in E,\, \|x\|_E=1} \|f(x)\|_F,$$

et donc ces trois quantités sont égales. □

On notera que pour tout $x \in E$ on a l'inégalité

$$\|f(x)\|_F \leq |||f|||_{L(E,F)} \times \|x\|_E. \tag{2.1}$$

Exercice 2.7. Soit $f \in L(E, F)$, montrer que $|||f|||$ est la plus petite constante C vérifiant $\|f(x)\|_F \leq C\|x\|_E$ pour tout $x \in E$.

Proposition 2.12. *L'application $f \mapsto |||f|||$ définit une norme sur $L(E, F)$.*

Exercice 2.8. Démontrer la proposition.

Exemple 2.4. *Soit $E = F = \mathbb{R}^n$ munis de la norme $\|x\|_1 = \sum\limits_{i=1}^{n} |x_i|$. Si $f \in \mathcal{L}(\mathbb{R}^n)$ on note $A = (a_{ij})_{i,j}$ la matrice associée dans la base canonique. On sait que f est continue (on est en dimension finie), et on cherche à calculer sa norme. Soit $x \in E$, alors $(f(x))_i = \sum\limits_{j=1}^{n} a_{ij}x_j$ et donc*

$$\|f(x)\|_1 = \sum_{i=1}^{n} \left|\sum_{j=1}^{n} a_{ij}x_j\right| \leq \sum_{i,j=1}^{n} |a_{ij}||x_j| = \sum_{j=1}^{n} \left(|x_j| \times \sum_{i=1}^{n} |a_{ij}|\right) \leq \max_j \left(\sum_{i=1}^{n} |a_{ij}|\right) \times \|x\|_1.$$

On en déduit que $|||f||| \leq \max_j \sum_{i=1}^n |a_{ij}|$.

Pour montrer que $|||f||| = \max_j \sum_{i=1}^n |a_{ij}|$, *il suffit de trouver un vecteur* x *tel que toutes les inégalités ci-dessus soient des égalités. Soit* j_0 *tel que* $\sum_{i=1}^n |a_{ij_0}| = \max_j \sum_{i=1}^n |a_{ij}|$, *on prend* x *tel que* $x_i = \delta_{ij_0}$, *i.e.* $x_i = 1$ *si* $i = j_0$ *et* 0 *sinon. On a alors* $\|x\|_1 = 1$ *et* $f(x)_i = a_{ij_0}$ *pour tout* i *d'où* $\|f(x)\|_1 = \sum_{i=1}^n |a_{ij_0}| = \max_j \sum_{i=1}^n |a_{ij}|$ *et donc* $|||f||| = \max_j \sum_{i=1}^n |a_{ij}|$.

Exercice 2.9. Même situation que dans l'exemple ci-dessus mais on munit cette fois $\mathbb{R}^n$ de la norme $\|x\|_\infty = \max\limits_{i=1,\ldots,n} |x_i|$. Montrer que $|||f||| = \max_i \sum_{j=1}^n |a_{ij}|$.

On termine cette section avec une caractérisation des applications multi-linéaires continues. On rappelle

Définition 2.13. *Soient* $E_1, \ldots, E_n, F$ *des espaces vectoriels. Une application* $f : E_1 \times \cdots \times E_n \to F$ *est dite* $n-$*linéaire si pour tout* $i = 1, \ldots, n$ *et tout* $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in E_1 \times \cdots \times E_{i-1} \times E_{i+1} \times \cdots \times E_n$ *l'application* $f_i : E_i \to F$ *définie par* $f_i(x_i) = f(x_1, \ldots, x_n)$ *est linéaire.*

Si $n = 2$ on parle d'application bilinéaire. Par exemple, un produit scalaire sur un espace vectoriel E est une application bilinéaire (symétrique, définie, positive) de $E \times E$ dans $\mathbb{R}$.

Si $E_1, \ldots, E_n$ sont des evn on rappelle (voir Section 1.1) que l'application

$$E = E_1 \times \cdots \times E_n \ni (x_1, \ldots, x_n) \mapsto \|(x_1, \ldots, x_n)\|_E := \max\{\|x_1\|_{E_1}, \ldots, \|x_n\|_{E_n}\}$$

définit une norme sur $E_1 \times \cdots \times E_n$. On a alors la caractérisation suivante des applications multilinéaires qui est le pendant de la Proposition 2.9 pour les applications linéaires.

Proposition 2.14. *Soient* $E_1, \ldots, E_n$ *et* F *des evn et* f *une application* $n-$*linéaire de* $E = E_1 \times \cdots \times E_n$ *dans* F. *Les assertions suivantes sont équivalentes*

1. f *est continue sur* E,
2. f *est continue en* 0,
3. f *est bornée sur la boule unité fermée de* E, *i.e.* $\{f(x) \mid x \in E, \|x\|_E \leq 1\}$ *est borné,*
4. f *est bornée sur la sphère unité de* E, *i.e.* $\{f(x) \mid x \in E, \|x\|_E = 1\}$ *est borné,*
5. *Il existe* $C \geq 0$ *tel que pour tout* $x = (x_1, \ldots, x_n) \in E$,

$$\|f(x)\|_F \leq C\|x_1\|_{E_1} \times \cdots \times \|x_n\|_{E_n}.$$

Démonstration. On montre que 1. $\Rightarrow$ 2. $\Rightarrow$ 3. $\Rightarrow$ 4. $\Rightarrow$ 5. $\Rightarrow$ 1. Les implications 1. $\Rightarrow$ 2. et 3. $\Rightarrow$ 4. sont immédiates.

2. $\Rightarrow$ 3. On procède comme pour la Proposition 2.9. On applique la définition de la continuité en 0 avec $\epsilon = 1$. Il existe $\delta > 0$ tel que $\|x\|_E < \delta \;\Rightarrow\; \|f(x)\|_F < 1$. Soit $x \in E$ tel que $\|x\|_E \leq 1$. On a alors $y = \frac{\delta}{2}x$ vérifie $\|y\|_E \leq \frac{\delta}{2} < \delta$ donc $\|f(y)\|_F < 1$. Comme f est n-linéaire on a $f(y) = \left(\frac{\delta}{2}\right)^n f(x)$ et donc $\|f(y)\|_F = \left(\frac{\delta}{2}\right)^n \|f(x)\|_F < 1$ d'où $\|f(x)\|_F < \left(\frac{2}{\delta}\right)^n$ pour tout x dans la boule unité.

4. $\Rightarrow$ 5. f est bornée sur la sphère unité, soit donc C tel que $\|f(x)\|_F \leq C$ pour tout $x \in E$ tel que $\|x\|_E = 1$. Soit $x = (x_1, \ldots, x_n) \in E$. S'il existe i tel que $x_i = 0$ on a $f(x) = 0$

et donc $\|f(x)\|_F \leq C\|x_1\|_{E_1} \times \cdots \times \|x_n\|_{E_n}$ est vrai. Sinon, soit $y := \left(\frac{x_1}{\|x_1\|_{E_1}}, \ldots, \frac{x_n}{\|x_n\|_{E_n}}\right)$. Par construction $\|y\|_E = 1$ donc $\|f(y)\|_F \leq C$. Or

$$f(y) = \frac{1}{\|x_1\|_{E_1} \times \cdots \times \|x_n\|_{E_n}} f(x),$$

et donc $\|f(x)\|_F \leq C\|x_1\|_{E_1} \times \cdots \times \|x_n\|_{E_n}$.

5. $\Rightarrow$ 1. Soit $x = (x_1, \ldots, x_n) \in E$ et $(x^{(k)})_k \in E^{\mathbb{N}}$ telle que $x^{(k)} \to x$, on montre que $f(x^{(k)}) \to f(x)$. On notera $x^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$. On écrit

$$f(x^{(k)}) - f(x) = \sum_{j=1}^{n} f(x_1^{(k)}, \ldots, x_{j-1}^{k)}, x_j^{(k)} - x_j, x_{j+1}, \ldots, x_n).$$

En utilisant l'inégalité triangulaire puis 5. on en déduit que

$$\begin{aligned} \|f(x^{(k)}) - f(x)\|_F &\leq \sum_{j=1}^{n} \|f(x_1^{(k)}, \ldots, x_{j-1}^{k)}, x_j^{(k)} - x_j, x_{j+1}, \ldots, x_n)\|_F \\ &\leq C \sum_{j=1}^{n} \|x_1^{(k)}\|_{E_1} \cdots \|x_{j-1}^{k)}\|_{E_{j-1}} \|x_j^{(k)} - x_j\|_{E_j} \|x_{j+1}\|_{E_{j+1}} \cdots \|x_n\|_{E_n}. \end{aligned}$$

Dans chaque E_j la norme est continue (voir Exercice 2.2) donc dans le membre de droite les termes de la forme $\|x_j^{(k)}\|_{E_j}$ tendent (lorsque $k \to \infty$) vers $\|x_j\|_{E_j}$ tandis que $\|x_j^{(k)} - x_j\|_{E_j} \to 0$, et donc $\|f(x^{(k)}) - f(x)\|_F \to 0$. $\square$

Remarque 2.2. *Dans le cas des applications bilinéaires de $E_1 \times E_2$ dans F la condition 5. s'écrit $\|f(x_1, x_2)\|_F \leq C\|x_1\|_{E_1}\|x_2\|_{E_2}$.*

Proposition 2.15. *Soit f une application n-linéaire continue de $E_1 \times \cdots \times E_n$ dans F. Alors*

$$\sup_{x \in E,\, \|x\|_E = 1} \|f(x)\|_F = \sup_{x \in E,\, \|x\|_E \leq 1} \|f(x)\|_F = \sup_{\substack{x=(x_1,\ldots,x_n)\in E,\\ x_1\neq 0,\ldots,x_n\neq 0}} \frac{\|f(x)\|_F}{\|x_1\|_{E_1} \times \cdots \times \|x_n\|_{E_n}}.$$

On note $|||f|||$ cette valeur. L'application $f \mapsto |||f|||$ définit une norme sur l'espace vectoriel des applications n-linéaires continues de $E_1 \times \cdots \times E_n$ dans F.

Exercice 2.10. Démontrer la proposition ci-dessus.

2.3 Séries à valeurs evn

Une fois qu'on a la notion de suite, resp. suite convergente, la notion de série, resp. série convergente, est identique à celle des séries de nombres réels ou complexes.

Définition 2.16. *Soit $(E, \| \ \|)$ un evn et $(u_n)_n \in E^{\mathbb{N}}$. On dit que la série $\sum u_n$ converge (dans E et pour la norme $\| \ \|$) si la suite des sommes partielles $\left(\sum_{n=0}^{N} u_n\right)_N$ converge (dans E et pour la norme $\| \ \|$), sinon on dit qu'elle diverge. Lorsqu'elle converge on note $\sum_{n\in\mathbb{N}} u_n$ la valeur de cette série (c'est la limite des sommes partielles).*

Définition 2.17. *Soit $(E, \| \ \|)$ un evn et $(u_n)_n \in E^{\mathbb{N}}$. On dit que la série $\sum u_n$ est normalement convergente si la série $\sum \|u_n\|$ converge (dans $\mathbb{R}$!).*

On a alors le lien suivant entre ces deux notions

Théorème 2.18. *Soit $(E, \| \ \|)$ est un espace de Banach et $(u_n)_n \in E^{\mathbb{N}}$. Si la série $\sum u_n$ est normalement convergente alors elle converge. De plus on a $\left\|\sum_{n\in\mathbb{N}} u_n\right\| \leq \sum_{n\in\mathbb{N}} \|u_n\|$.*

Démonstration. Comme E est un Banach, pour montrer que la série $\sum u_n$ converge, i.e. que la suite $\left(\sum_{n=0}^{N} u_n\right)_N$ converge, il suffit de montrer qu'elle vérifie le critère de Cauchy :

$$\forall \epsilon > 0,\ \exists N \in \mathbb{N},\ \forall q \geq p \geq N,\ \left\|\sum_{n=p}^{q} u_n\right\| < \epsilon.$$

Soit donc $\epsilon > 0$. La série $\sum \|u_n\|$ converge donc elle vérifie le critère de Cauchy (dans $\mathbb{R}$), i.e. il existe $N \in \mathbb{N}$ tel que pour tous $q \geq p \geq N$ on a

$$\sum_{n=p}^{q} \|u_n\| < \epsilon.$$

L'inégalité triangulaire montre alors que si $q \geq p \geq N$ on a

$$\left\|\sum_{n=p}^{q} u_n\right\| \leq \sum_{n=p}^{q} \|u_n\| < \epsilon.$$

La série $\sum u_n$ vérifie le critère de Cauchy et est donc convergente.

D'après l'inégalité triangulaire, pour tout N on a $\left\|\sum_{n=0}^{N} u_n\right\| \leq \sum_{n=0}^{N} \|u_n\|$. L'inégalité voulue s'obtient alors en passant à la limite $N \to \infty$ et en utilisant la continuité de la norme. □

Vous avez déjà rencontré la notion de série normalement convergente dans le cadre des séries de fonctions : si I est un intervalle et $(f_n)_n$ une suite de fonctions de I dans $\mathbb{R}$ alors on dit que la série de fonctions $\sum f_n$ converge normalement sur I s'il existe une suite $(a_n)_n$ de nombres positifs tels $\sum a_n$ converge et telle que pour tout $x \in I$ et tout n on ait $|f_n(x)| \leq a_n$. Par ailleurs, si une série de fonctions converge normalement sur I alors elle converge uniformément (et donc simplement) sur I, i.e. $\left\|\sum_{n=0}^{N} f_n - f\right\|_\infty \to 0$ lorsque $N \to \infty$ où $\|g\|_\infty = \sup_{x\in I} |g(x)|$.

Prenons $I = [a, b]$ et notons $E = C^0([a, b])$ l'espace des fonctions continues sur $[a, b]$. Muni de la norme $\|f\|_\infty = \sup\limits_{x\in[a,b]} |f(x)|$ c'est un espace de Banach (voir Exemple 1.18). Si $(f_n)_n \in E^{\mathbb{N}}$, dire que la série $\sum f_n$ (c'est une série de fonctions) est normalement convergente au sens de la Définition 2.17 signifie que la série $\sum \|f_n\|_\infty$ converge. Au sens que vous avez vu en cours de séries c'est dire qu'il existe une suite $(a_n)_n$ de nombres positifs tels $\sum a_n$ converge et telle que $\|f_n\|_\infty \leq a_n$ pour tout n. C'est exactement la même chose : si $\sum a_n$ converge, par comparaison de séries à termes positifs $\sum \|f_n\|_\infty$ aussi, et si on suppose que $\sum \|f_n\|_\infty$ converge il suffit de prendre $a_n = \|f_n\|_\infty$ (la définition vue en cours de séries est "il existe une suite $(a_n)_n$").

Dans la section précédente on a vu la notion d'application linéaire continue. On rencontrera par la suite, en particulier dans les Théorèmes d'inversion locale et des fonctions implicites, la notion importante d'homéomorphisme (application continue, inversible et dont l'inverse est continu). Étant donnée une application linéaire, on a vu comment étudier sa continuité mais il est important d'avoir un critère pour montrer qu'elle est inversible et aussi que son inverse est continue : l'inverse d'une application n'a pas forcément une expression aussi simple que l'application elle même, et si on peut éviter d'avoir à la calculer... (penser au calcul de l'inverse d'une matrice $n \times n$). On va utiliser la notion de série normalement convergente pour montrer le critère suivant.

Théorème 2.19. *Soit $(E, \|\ \|)$ un espace de Banach et $u \in L(E)$ tel que $|||u||| < 1$. Alors l'application $I - u$ est inversible, son inverse $(I - u)^{-1}$ est continue et on a $|||(I - u)^{-1}||| \leq \frac{1}{1-|||u|||}$.*

Remarque 2.3. *Si $u \in L(E)$ est inversible, son inverse n'est pas nécessairement continue. Par exemple, soit $E = \mathbb{R}[X]$ muni de la norme $\|P\| = \max\{|a_0|, \ldots, |a_n|\}$ si $P(X) = \sum\limits_{k=0}^{n} a_k X^k$, et $u : E \to E$ définie par $(u(P))(X) = \sum_{k=0}^{n} \frac{a_k}{k+1} X^k$. Alors u est continue (montrez-le) et bijective. Son inverse est $(u^{-1}(P))(X) = \sum_{k=0}^{n} (k + 1)a_k X^k$ qui n'est pas continue.*

La fin de cette section a pour but de montrer le Théorème 2.19. L'objet clé sera la série $\sum u^n$, où $u^n := \underbrace{u \circ \cdots \circ u}_{\text{n fois}}$. Puisque $u \in L(E)$, par composition $u^n \in L(E)$ pour tout n et on considèrera ainsi une série dans l'evn $(L(E), |||\ |||)$. Il est donc important (cf Théorème 2.18) de savoir quand ce dernier est un Banach.

Proposition 2.20. *Soient $(E, \|\ \|_E)$ et $(F, \|\ \|_F)$ deux evn. Si F est un Banach alors $(L(E, F), |||\ |||)$ est un Banach.*

Corollaire 2.21. *Si $(E, \|\ \|)$ est un Banach alors $L(E)$ est un Banach.*

Exercice 2.11. Le but est de démontrer la proposition ci-dessus. Soit $(f_n)_n$ une suite de Cauchy dans $(L(E,F), ||| \ |||)$.

a) Montrer que pour tout $x \in E$ la suite $(f_n(x))_n$ est de Cauchy dans $(F, \| \ \|_F)$ (on pourra utiliser (2.1)), puis qu'elle converge. On note $f(x)$ sa limite.

b) On note $f : x \mapsto f(x)$. Montrer que $f \in \mathcal{L}(E,F)$.

c) Montrer que $f \in L(E,F)$ et que $|||f_n - f||| \to 0$ (on pourra s'inspirer du raisonnement de l'Exemple 1.18).

d) Conclure.

Exercice 2.12. Soit $(E, \| \ \|)$ un evn. Montrer que l'espace $L(E,\mathbb{R})$ des formes linéaires continues sur E est un Banach.

Proposition 2.22. *Soient $(E, \| \ \|_E)$, $(F, \| \ \|_F)$ et $(G, \| \ \|_G)$ trois evn, $u \in L(E,F)$ et $v \in L(F,G)$, alors $v \circ u \in L(E,G)$ et $|||v \circ u|||_{L(E,G)} \leq |||u|||_{L(E,F)} \times |||v|||_{L(F,G)}$.*

Démonstration. $v \circ u$ est la composée d'applications linéaires continues donc est elle même linéaire continue. Il reste juste à montrer l'inégalité entre les normes. Soit $x \in E$, on a $\|u(x)\|_F \leq |||u|||_{L(E,F)}\|x\|_E$. Par ailleurs, pour tout $y \in F$ on a $\|v(y)\|_G \leq |||v|||_{L(F,G)}\|y\|_F$, donc en particulier

$$\|v \circ u(x)\|_G \leq |||v|||_{L(F,G)}\|u(x)\|_F \leq |||v|||_{L(F,G)}|||u|||_{L(E,F)}\|x\|_E.$$

Comme $|||v \circ u|||_{L(E,G)}$ est la plus petite constante C vérifiant $\|v \circ u(x)\|_G \leq C\|x\|_E$ pour tout $x \in E$ on a bien $|||v \circ u|||_{L(E,G)} \leq |||u|||_{L(E,F)} \times |||v|||_{L(F,G)}$. □

Corollaire 2.23. *Si $(E, \| \ \|)$ est un evn alors pour tout $u \in L(E)$ et pour tout $n \in \mathbb{N}$ on a*

$$|||\underbrace{u \circ \cdots \circ u}_{n \text{ fois}}||| \leq |||u|||^n.$$

Remarque 2.4. *Si $(E, \| \ \|)$ est un evn, $L(E)$ muni de $||| \ |||$ est un evn mais on a aussi une autre loi, la composition $\circ$. Une norme N sur $L(E)$ qui vérifie $N(v \circ u) \leq N(u)N(v)$ pour tous $u, v \in L(E)$ est appelée norme d'algèbre. La proposition ci-dessus montre que la norme $||| \ |||$ est une telle norme. Il peut parfois être utile d'avoir cette propriété supplémentaire. En particulier, si l'espace sur lequel on travaille est $M_n(\mathbb{R})$, on a de telles normes (a priori une pour chacun des choix de normes sur $\mathbb{R}^n$, voir Exemple 2.4 et Exercice 2.9) et comme par ailleurs on est dimension finie toutes les normes sont équivalentes, donc on ne perd rien à choisir une telle norme.*

Le Théorème 2.19 découle du théorème suivant.

Théorème 2.24. *Soit $(E, \| \ \|)$ un espace de Banach et $u \in L(E)$ tel que $|||u||| < 1$. Alors la série $\sum u^n$, où $u^n := \underbrace{u \circ \cdots \circ u}_{n \text{ fois}}$, est normalement convergente (dans $L(E)$) et donc convergente. De plus, sa somme vérifie*

$$(I - u) \circ \sum u^n = \sum u^n \circ (I - u) = I.$$

Démonstration du Théorème 2.19. On a $\sum u^n \in L(E)$ et vérifie $(I-u)\circ\sum u^n = \sum u^n \circ (I-u) = I$. Ca signifie exactement que $I-u$ est inversible et que son inverse est $\sum u^n$ qui est donc continue. De plus, d'après le Théorème 2.18, on a

$$\left|\left|\left|\sum_{n\in\mathbb{N}} u^n\right|\right|\right| \leq \sum_{n\in\mathbb{N}} |||u^n||| \leq \sum_{n\in\mathbb{N}} |||u|||^n = \frac{1}{1-|||u|||}.$$

□

Démonstration du Théorème 2.24. D'après le Corollaire 2.23, pour tout n on a $|||u^n||| \leq |||u|||^n$. Comme $|||u||| < 1$, la série $\sum |||u|||^n$ converge et donc $\sum |||u^n|||$ aussi. La série $\sum u^n$ est donc normalement convergente. Comme E est un Banach, $(L(E), |||\ |||)$ aussi et la série $\sum u^n$ converge. Soit $N \in \mathbb{N}$, on a

$$(I-u)\circ\sum_{n=0}^{N} u^n = \sum_{n=0}^{N} u^n \circ (I-u) = I - u^{N+1}.$$

Puisque $|||u^{N+1}||| \leq |||u|||^{N+1} \to 0$ on a donc

$$\lim_{N\to\infty} (I-u)\circ\sum_{n=0}^{N} u^n = \lim_{N\to\infty} \sum_{n=0}^{N} u^n \circ (I-u) = I.$$

Il reste à montrer que $\lim\limits_{N\to\infty} \sum\limits_{n=0}^{N} u^n \circ (I-u) = \sum\limits_{n\in\mathbb{N}} u^n \circ (I-u)$ et $\lim\limits_{N\to\infty} (I-u)\circ\sum\limits_{n=0}^{N} u^n = (I-u)\circ\sum\limits_{n\in\mathbb{N}} u^n$. On montre le premier, le second est similaire. On écrit

$$\begin{aligned}
\left|\left|\left|\sum_{n=0}^{N} u^n \circ (I-u) - \sum_{n\in\mathbb{N}} u^n \circ (I-u)\right|\right|\right| &= \left|\left|\left|\left(\sum_{n=0}^{N} u^n - \sum_{n\in\mathbb{N}} u^n\right)\circ (I-u)\right|\right|\right| \\
&\leq \underbrace{\left|\left|\left|\sum_{n=0}^{N} u^n - \sum_{n\in\mathbb{N}} u^n\right|\right|\right|}_{\overset{N\to\infty}{\longrightarrow} 0} \times |||I-u|||,
\end{aligned}$$

ce qui prouve que $\sum\limits_{n\in\mathbb{N}} u^n \circ (I-u) = \lim\limits_{N\to\infty} \sum\limits_{n=0}^{N} u^n \circ (I-u)$. □

Chapitre 3

Différentiabilité

Dans tout le chapitre f sera une fonction définie sur un <u>ouvert</u> U d'un evn $(E, \| \|_E)$ et à valeurs dans un evn $(F, \| \|_F)$.

3.1 Fonctions différentiables - Différentielle

Définition 3.1. *Soit $f : U \to F$ et $x_0 \in U$. On dit que f est différentiable en x_0 s'il existe $L \in L(E, F)$ telle que*

$$\forall \epsilon > 0,\ \exists \delta > 0,\ \forall h \in E,\ \|h\|_E < \delta \quad \Rightarrow \quad \|f(x_0 + h) - f(x_0) - L(h)\|_F < \epsilon \|h\|_E.$$

De façon équivalente, f est différentiable en x_0 s'il existe $L \in L(E, F)$

$$\lim_{h \to 0_E} \frac{1}{\|h\|_E}(f(x_0 + h) - f(x_0) - L(h)) = 0. \tag{3.1}$$

On écrira encore $f(x_0 + h) = f(x_0) + L(h) + o(h)$.

Remarque 3.1. *Puisque $x_0 \in U$ qui est ouvert, la fonction $h \mapsto f(x_0 + h)$ est bien définie sur un voisinage de 0, i.e. si $\|h\|_E$ est assez petite.*

Proposition 3.2. *Soit $f : U \to F$ et $x_0 \in U$. Si f est différentiable en x_0 alors l'application L est unique. Elle est appelée différentielle de f en x_0 et est notée $Df(x_0)$, ou encore $\mathrm{d}f(x_0)$, $\mathrm{d}f_{x_0}$ ou Df_{x_0}.*

Démonstration. Soient L_1, L_2 dans $L(E, F)$ telles que

$$\lim_{h \to 0_E} \frac{1}{\|h\|_E}(f(x_0 + h) - f(x_0) - L_1(h)) = \lim_{h \to 0_E} \frac{1}{\|h\|_E}(f(x_0 + h) - f(x_0) - L_2(h)) = 0.$$

Par différence on a donc $\lim\limits_{h \to 0_E} \dfrac{1}{\|h\|_E}(L_1(h) - L_2(h)) = 0$. Soit $x \in E$ non nul, on a $x_n := \frac{1}{n}x \to 0$ et donc

$$\frac{1}{\|x\|_E}(L_1(x) - L_2(x)) = \frac{1}{\|x_n\|_E}\left(L_1\left(x_n\right) - L_2\left(x_n\right)\right) \xrightarrow{n \to \infty} 0,$$

i.e. $L_1(x) = L_2(x)$. □

Attention ! La notion de différentiabilité, et de différentielle, fait intervenir la notion de limite. Tout comme pour la continuité elle dépend donc du choix des normes sur E et sur F ! Lorsqu'il peut y avoir ambiguïté sur le choix de celle(s)-ci on précisera bien pour quelle norme la fonction est différentiable.

Définition 3.3. *Soit $f : U \to \mathbb{R}$. On dit que f est différentiable sur U si f est différentiable en tout $x_0 \in U$. On appelle alors différentielle de f l'application* $Df : \begin{array}{ccc} U & \to & L(E,F) \\ x & \mapsto & Df(x) \end{array}$.

Lorsque f est différentiable, en *chaque* point $x_0 \in U$ la différentielle $Df(x_0)$ de f en x_0 est une application linéaire de E dans F (f est elle définie sur un ouvert $U \subset E$ à valeurs dans F), et la différentielle de f est l'application définie sur U qui à $x_0 \in U$ associe $Df(x_0) \in L(E,F)$. On fera également bien attention au fait que dans la définition de différentiable on impose à l'application linéaire L d'être continue ! ! La raison est simplement qu'on souhaite avoir

Proposition 3.4. *Soit $f : U \to F$ et $x_0 \in U$. Si f est différentiable en x_0 alors f est continue en x_0.*

Démonstration. Si $x_n \to x_0$, en écrivant $f(x_n) = f(x_0) + Df(x_0)(x_n - x_0) + o(x_n - x_0)$, on a immédiatement $f(x_n) \to f(x_0)$ et donc la continuité de f en x_0. □

Si on avait $f(x_0 + h) = f(x_0) + L(h) + o(h)$ avec L qui n'est pas continue, en particulier elle ne serait pas continue en 0 et donc f ne serait pas continue en x_0. Si E est de dimension finie on a vu, Proposition 2.10, que toutes les applications linéaires de E dans F étaient continues, et cela simplifie l'étude de la différentiabilité.

Lorsque $E = \mathbb{R}$, resp. $\mathbb{R}^n$, et $F = \mathbb{R}$ on retrouve évidemment les notions de fonctions dérivables, resp. différentiables, telles que vue en L1-L2. (Dans ces deux cas on est en dimension finie donc toutes les normes sont équivalentes et toutes les applications linéaires sont continues.)

- Si $E = F = \mathbb{R}$, tout élément $L \in L(E,F)$ est de la forme $x \mapsto \ell x$ où $\ell \in \mathbb{R}$. La fonction f est donc différentiable en x_0 s'il existe $\ell \in \mathbb{R}$ tel que

$$\lim_{h \to 0} \frac{1}{|h|}(f(x_0 + h) - f(x_0) - \ell h) = 0 \quad \Longleftrightarrow \quad \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \ell,$$

  ce qui est précisément la définition de dérivable en x_0. On fera juste attention au fait que $\ell = f'(x_0)$ est le nombre dérivé de f en x_0 tandis que $Df(x_0) \in L(\mathbb{R}, \mathbb{R})$ est l'application $h \mapsto f'(x_0)h$.
- $E = \mathbb{R}^n$ et $F = \mathbb{R}$. Vous avez vu en L2 la définition suivante de fonction différentiable :

  f est différentiable en $a = (a_1, \ldots, a_n)$ si toutes les dérivées partielles de f au point a existent et si

$$\lim_{h \to 0} \frac{1}{\|h\|} \left(f(a+h) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i \right) = 0.$$

Si f vérifie la propriété ci-dessus, alors elle est différentiable au sens de la Définition 3.1 avec $L(h) := \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i$. Inversement, si f est différentiable au sens de la Définition 3.1, en prenant $h = (0, \ldots, 0, h_i, 0, \ldots, 0)$ on en déduit que f admet une dérivée partielle par rapport à la $i-$ème variable en a, avec $\frac{\partial f}{\partial x_i}(a) = Df(a)(h_i)$, puis que

$$\frac{1}{\|h\|}\left(f(a+h) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i\right) = \frac{1}{\|h\|}\left(f(a+h) - f(a) - Df(a)(h)\right) \overset{h\to 0}{\longrightarrow} 0.$$

En particulier $Df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i$.

Dans $\mathbb{R}^n$ le calcul des dérivées partielles nous donne le candidat pour la différentielle si f est différentiable. Pour un espace E général la Définition 3.1 semble difficile à appliquer puisqu'il faut trouver une application linéaire continue L qui satisfasse (3.1). La notion suivante généralise celle de dérivée partielle et sera utile dans la pratique pour trouver l'application linéaire L candidate à être la différentielle (voir Exemple 3.1).

Définition 3.5. *Soit $f : U \to F$, $x_0 \in U$ et $v \in E$. On dit que f est dérivable en x_0 dans la direction v si la fonction d'une variable réelle $f_v : t \mapsto f(x_0 + tv)$, définie au voisinage de 0, est dérivable en 0.*

Remarque 3.2. *Si $E = \mathbb{R}^n$, $F = \mathbb{R}$ et $(e_1, \ldots, e_n)$ désigne la base canonique de $\mathbb{R}^n$, dire que f est dérivable en x_0 dans la direction e_i signifie que f admet une dérivée partielle par rapport à la $i-$ème variable en x_0.*

Proposition 3.6. *Si $f : U \to F$, $x_0 \in U$ et f est différentiable en x_0, alors f est dérivable en x_0 dans n'importe quelle direction v et on a $f_v'(0) = Df(x_0)(v)$.*

Démonstration. Puisque f est différentiable en x_0, pour t assez petit on a

$$f(x_0 + tv) = f(x_0) + Df(x_0)(tv) + o(tv) = f(x_0) + tDf(x_0)(v) + o(t)$$

et donc $\frac{f_v(t) - f_v(0)}{t} = Df(x_0)(v) + o(1) \overset{t\to 0}{\longrightarrow} 0.$ □

Attention ! La réciproque est fausse. Une fonction peut être dérivable en un point x_0 dans toutes les directions sans être différentiable.

Exercice 3.1. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \frac{y^2}{x}$ si $x \neq 0$ et $f(0, y) = 0$. Montrer que f est dérivable en $(0, 0)$ dans toutes les directions mais qu'elle n'y est pas différentiable.

Remarque 3.3. *Si $E = \mathbb{R}$ il n'y a essentiellement qu'une seule direction possible (à multiplication par un scalaire près) et par abus de langage on parlera parfois de fonction dérivable au lieu de différentiable. Par ailleurs, on a alors $Df(x_0)(h) = f'(x_0)h$ où $f'(x_0) := \lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$.*

Dans la pratique, pour voir si une fonction est différentiable en x_0 et calculer $Df(x_0)$ on pourra suivre le schéma suivant :

1. Étant donné $v \in E$, montrer que f est dérivable dans la direction v en x_0, i.e. $f_v'(0)$ existe. (Si f est différentiable elle doit être dérivable dans toutes les directions.)
2. Montrer que l'application $L : v \mapsto f_v'(0)$ est linéaire continue. (Si f est différentiable on doit avoir $f_v'(0) = Df(x_0)(v)$.)
3. Montrer que $f(x_0 + h) - f(x_0) - L(h) = o(h)$.

Exemple 3.1. *Soit $f : M_n(\mathbb{R}) \ni M \mapsto M^2 \in M_n(\mathbb{R})$. On veut montrer que f est différentiable sur $M_n(\mathbb{R})$ et calculer $Df(M)$. On est en dimension finie donc toutes les normes sont équivalentes et la différentiabilité de f ne dépendra de la norme choisie. Afin d'avoir le maximum de propriétés on munit $M_n(\mathbb{R})$ d'une norme d'application linéaire (on parlera aussi de norme matricielle), par exemple $\|M\| = \max_j \sum_{i=1}^n |m_{ij}|$ (c'est la norme application linéaire associée à la norme $\| \ \|_1$ sur $\mathbb{R}^n$, cf Exemple 2.4). L'avantage de choisir une norme de ce type est qu'en plus de l'inégalité triangulaire on a aussi la propriété de norme d'algèbre : si $M, N \in M_n(\mathbb{R})$ on a $\|MN\| \leq \|M\| \times \|N\|$. Soit $M \in M_n(\mathbb{R})$, pour étudier la différentiabilité de f en M on va suivre le schéma ci-dessus.*

1. *Soit $H \in M_n(\mathbb{R})$, on étudie la dérivabilité en $t = 0$ de $f_H(t) = (M + tH)^2$. On a*

$$\frac{1}{t}(f_H(t)-f_H(0)) = \frac{1}{t}(M^2+tMH+tHM+t^2H^2-M^2) = MH+HM+tH^2 \xrightarrow{t\to 0} MH+HM.$$

2. *Soit $L : M_n(\mathbb{R}) \to M_n(\mathbb{R})$ définie par $L(H) = MH + HM$. On vérifie facilement que L est une application linéaire. Par ailleurs, $M_n(\mathbb{R})$ est de dimension finie donc L est continue.*
3. *On regarde si $f(M + H) - f(M) - L(H) = o(H)$. On a*

$$\frac{1}{\|H\|}\|(M+H)^2 - M^2 - L(H)\| = \frac{\|H^2\|}{\|H\|} \leq \frac{\|H\|^2}{\|H\|} = \|H\| \xrightarrow{H\to 0} 0.$$

On voit sur ce dernier calcul l'avantage d'avoir utilisé une norme d'application linéaire.

En conclusion, f est différentiable en M et $Df(M) : H \mapsto MH + HM$.

On a bien sûr les propriétés usuelles de somme et de composition pour les fonctions différentiables sur des evn quelconques.

Proposition 3.7. *Soient $f, g : U \to F$ et $x_0 \in U$. Si f et g sont différentiables en x_0 alors $f + g$ aussi et on a $D(f + g)(x_0) = Df(x_0) + Dg(x_0)$.*

Exercice 3.2. Démontrer la proposition.

Proposition 3.8. *Soient $U \subset E$, $V \subset F$ des ouverts, $f : U \to F$ telle que $f(U) \subset V$, $g : V \to G$ et $x_0 \in U$. Si f est différentiable en x_0 et g est différentiable en $f(x_0)$. Alors la fonction $g \circ f : U \to G$ est différentiable en x_0 et on a $D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0)$.*

Démonstration. Les applications linéaires $Df(x_0) : E \to F$ et $Dg(f(x_0)) : F \to G$ sont continues donc l'application linéaire $Dg(f(x_0)) \circ Df(x_0) : E \to G$ aussi. Il suffit donc de montrer que

$$(g \circ f)(x_0 + h) - (g \circ f)(x_0) - (Dg(f(x_0)) \circ Df(x_0))(h) = o(h).$$

Puisque f est différentiable en x_0 on a $h' := f(x_0 + h) - f(x_0) = Df(x_0)(h) + \|h\|\epsilon_1(h)$ avec $\epsilon_1(h) \to 0$ lorsque $h \to 0$. On notera que

$$\|h'\| \leq \|h\|(|||Df(x_0)||| + \|\epsilon_1(h)\|), \tag{3.2}$$

et en particulier $h' \to 0$ lorsque $h \to 0$.

Par ailleurs g est différentiable en $f(x_0)$ donc

$$\begin{aligned} (g \circ f)(x_0 + h) &= g(f(x_0) + h') \\ &= (g \circ f)(x_0) + Dg(f(x_0))(h') + \|h'\|\epsilon_2(h') \\ &= (g \circ f)(x_0) + (Dg(f(x_0)) \circ Df(x_0))(h) + \|h\|Dg(f(x_0))(\epsilon_1(h)) + \|h'\|\epsilon_2(h'). \end{aligned}$$

Il reste donc à montrer que $\|h\|Dg(f(x_0))(\epsilon_1(h)) + \|h'\|\epsilon_2(h') = o(h)$. On a, en utilisant (3.2),

$$\frac{\big\| \|h\|Dg(f(x_0))(\epsilon_1(h)) + \|h'\|\epsilon_2(h')\big\|}{\|h\|} \leq |||Dg(f(x_0))||| \cdot \|\epsilon_1(h)\| + (|||Df(x_0)||| + \|\epsilon_1(h)\|)\|\epsilon_2(h')\|$$

Puisque ϵ_1 et ϵ_2 tendent vers 0 en 0 et que $h' \to 0$ lorsque $h \to 0$, le membre de droite tend vers 0 ce qui prouve le résultat. □

On termine cette section avec deux cas particuliers de différentiabilité, celle des applications linéaires et celle de l'application $A \mapsto A^{-1}$ définie sur l'ensemble des matrices carrées inversibles $GL_n(\mathbb{R})$.

Proposition 3.9. *Soit $f \in L(E, F)$, alors f est différentiable sur E et pour tout $x_0 \in E$ on a $Df(x_0) = f$.*

Démonstration. Puisque $f \in L(E, F)$ il suffit de montrer que pour tout x_0 on a $f(x_0 + h) - f(x_0) - f(h) = o(h)$ lorsque $h \to 0$. C'est immédiat puisque f est linéaire et donc $f(x_0 + h) - f(x_0) - f(h) = 0$. □

Exercice 3.3. Soit $f : M_n(\mathbb{R}) \ni M \mapsto \mathrm{Tr}(M^2) \in \mathbb{R}$. Montrer que f est différentiable sur $M_n(\mathbb{R})$ et que pour tout M on a $Df(M) : H \mapsto 2\mathrm{Tr}(MH)$.

Proposition 3.10. *L'application $f : GL_n(\mathbb{R}) \ni A \mapsto A^{-1} \in M_n(\mathbb{R})$ est différentiable et pour tout $A \in GL_n(\mathbb{R})$ on a $Df(A)(H) = -A^{-1}HA^{-1}$.*

Démonstration. On identifiera $M_n(\mathbb{R})$ et $L(\mathbb{R}^n)$ (une matrice A est associée à l'application linéaire dont elle est la matrice représentative dans la base canonique) et on prendra sur $M_n(\mathbb{R})$ une norme matricielle.

On commence par vérifier que $GL_n(\mathbb{R})$ est bien un ouvert de $M_n(\mathbb{R})$. Soit $A \in GL_n(\mathbb{R})$, pour tout $H \in M_n(\mathbb{R})$ on a $A + H = A(I + A^{-1}H)$. Puisque A est inversible, $A + H$ est aussi inversible si et seulement si $I + A^{-1}H$ l'est. D'après le Théorème 2.19 c'est le cas dès que $\|A^{-1}H\| < 1$. Comme $\|A^{-1}H\| \leq \|A^{-1}\|\|H\|$, on en déduit que si $\|H\| < \|A^{-1}\|^{-1}$ alors $A+H$ est inversible. Autrement dit, si $A \in GL_n(\mathbb{R})$ alors $B(A, \|A^{-1}\|^{-1}) \subset GL_n(\mathbb{R})$ et donc ce dernier est un ouvert.

Soit $A \in GL_n(\mathbb{R})$, l'application $H \mapsto -A^{-1}HA^{-1}$ est bien linéaire et donc continue puisqu'on est en dimension finie. Enfin pour tout $H \in B(A, \|A^{-1}\|^{-1})$ (pour que $A+H$ soit inversible), on a

$$\begin{aligned} f(A+H) - f(A) + A^{-1}HA^{-1} &= (A+H)^{-1} - A^{-1} + A^{-1}HA^{-1} \\ &= (I + A^{-1}H)^{-1}A^{-1} - A^{-1} + A^{-1}HA^{-1} \\ &= \left(\sum_{k\geq 0}(-A^{-1}H)^k\right) A^{-1} - A^{-1} + A^{-1}HA^{-1} \\ &= \sum_{k\geq 2}(-A^{-1}H)^k A^{-1}, \end{aligned}$$

où on a utilisé le Théorème 2.24 pour passer de la deuxième à la troisième ligne. Ainsi

$$\|f(A+H) - f(A) + A^{-1}HA^{-1}\| \leq \sum_{k\geq 2} \|A^{-1}\|^{k+1}\|H\|^k = \frac{\|A^{-1}\|^3\|H\|^2}{1 - \|A^{-1}\|\|H\|},$$

et donc $f(A+H) - f(A) + A^{-1}HA^{-1} = o(H)$. □

Remarque 3.4. *Ici la différentielle de f est donnée dans l'énoncé de la proposition et on a donc directement vérifié que (3.1) était satisfaite. On aurait pu aussi appliquer le schéma indiqué précédemment pour retrouver l'expression de $Df(A)$.*

3.2 Les accroissements finis

On rappelle le Théorème des accroissements finis pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

Théorème 3.11. *Soit $f : [a, b] \to \mathbb{R}$ continue sur $[a, b]$, dérivable sur $]a, b[$. Alors il existe $c \in]a, b[$ tel que $f(b) - f(a) = f'(c)(b - a)$.*

L'inégalité des accroissements finis en découle alors directement

Théorème 3.12. *Soit I un intervalle et $f : I \to \mathbb{R}$ dérivable. Alors pour tous $a, b \in I$ on a $|f(b) - f(a)| \leq \sup_{x\in I} |f'(x)| \times |b - a|$.*

Pour rappel, ces théorèmes sont fondamentaux dans l'étude des fonctions d'une variable. Par exemple, le fait qu'une fonction dont la dérivée est nulle, resp. positive/négative, sur un intervalle soit constante, resp. croissante/décroissante, en est une conséquence. Il est donc naturel de se demander ce que deviennent ces théorèmes pour des fonctions d'un evn dans un evn.

Théorème 3.13. *Soit $f : [a, b] \to E$ une fonction différentiable, alors $\|f(b) - f(a)\| \leq \sup_{x\in I} |||Df(x)||| \times (b - a)$.*

Démonstration. Soit $M > \sup_{x\in I} \|Df(x)\|$, on va montrer que $\|f(b) - f(a)\| \leq M(b-a)$. Il suffira alors de faire tendre M vers $\sup_{x\in I} |||Df(x)|||$.

On pose $I = \{x \in [a,b] \,|\, \forall y \leq x,\ \|f(y) - f(a)\| \leq M(y-a)\}$. On a évidemment $a \in I$ et, par construction, I est un intervalle. Soit $c = \sup I$. On a ainsi $I = [a, c[$ ou $I = [a, c]$. La fonction f est différentiable donc continue, et on en déduit que $c \in I$ (on prend $(c_n)_n \in I^{\mathbb{N}}$ telle que $c_n \to c$ et on passe à la limite dans l'inégalité) d'où $I = [a, c]$.

Il reste à montrer que $c = b$. On raisonne par l'absurde. Si $c < b$, pour tout $h > 0$ tel que $c + h < b$ on a

$$\|f(c+h) - f(c)\| = \|Df(c)(h) + o(h)\| \leq |||Df(c)|||h + o(h).$$

Puisque $M > \sup_{x\in I} |||Df(x)||| \geq |||Df(c)|||$, il existe h_0 tel que pour $h \leq h_0$ on ait

$$|||Df(c)|||h + o(h) \leq Mh.$$

Finalement on obtient

$$\|f(c+h) - f(a)\| \leq \|f(c+h) - f(c)\| + \|f(c) - f(a)\| \leq Mh + M(c-a) = M(c+h-a)$$

pour $h \leq h_0$ et donc $c + h_0 \in I$ ce qui contredit la définition de c. □

Corollaire 3.14. *(Inégalité des accroissements finis) Soit U un ouvert de E et $f : U \to F$ différentiable sur U, alors pour tous $a, b \in U$ tels que $[a, b] \subset U$ on a*

$$\|f(b) - f(a)\|_F \leq \sup_{x\in[a,b]} |||Df(x)||| \times \|b-a\|_E.$$

Remarque 3.5. *La notation $[a, b] \subset U$ signifie que pour tout $t \in [0, 1]$ on a $(1-t)a + tb \in U$.*

Démonstration. Soit $g : [0, 1] \to F$ définie par $g(t) = f((1-t)a + tb)$. La fonction g est différentiable (dérivable) et $Dg(t)(h) = Df((1-t)a + tb)(b-a)h$. D'où

$$|||Dg(t)||| = \|Df((1-t)a+tb)(b-a)\| \leq |||Df((1-t)a+tb)|||\times\|b-a\|_E \leq \sup_{x\in[a,b]} |||Df(x)|||\times\|b-a\|_E.$$

D'après le Théorème 3.13 on a donc

$$\|f(b) - f(a)\|_F = \|g(1) - g(0)\|_F \leq \sup_{t\in[0,1]} |||Dg(t)||| \leq \sup_{x\in[a,b]} |||Df(x)||| \|b-a\|_E.$$

□

En appliquant le corollaire à la fonction $g(x) := f(x) - Df(a)(x-a)$, dont la différentielle est $Dg(x) = Df(x) - Df(a)$ on obtient

Corollaire 3.15. *Soit U un ouvert de E et $f : U \to F$ différentiable sur U, alors pour tous $a, b \in U$ tels que $[a, b] \subset U$ on a*

$$\|f(b) - f(a) - Df(a)(b-a)\|_F \leq \sup_{x\in[a,b]} |||Df(x) - Df(a)||| \times \|b-a\|_E.$$

Si la fonction f est à valeurs dans un evn quelconque on voit donc que l'inégalité des accroissements finis est toujours vraie. Par contre le Théorème 3.11 ne l'est plus, dans le sens où il n'existe pas forcément $c \in [a, b]$ tel que $f(b) - f(a) = Df(c)(b - a)$. Il reste vrai si f est définie sur un evn et à valeurs *réelles*.

Proposition 3.16. *Soit U un ouvert de E et $f : U \to \mathbb{R}$ différentiable sur U, alors pour tous $a, b \in U$ tels que $[a, b] \subset U$ il existe $c \in]a, b[$ tel que $f(b) - f(a) = Df(c)(b - a)$.*

Démonstration. Comme dans le corollaire précédent on se ramène en fait aux fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Soit $g : [0, 1] \to \mathbb{R}$ définie par $g(t) = f((1 - t)a + tb)$. La fonction g est dérivable et $g'(t) = Df((1 - t)a + tb)(b - a)$. D'après le Théorème 3.11 il existe $s \in]0, 1[$ tel que $g(1) - g(0) = g'(s)$, i.e. $f(b) - f(a) = Df(c)(b - a)$ avec $c = (1 - s)a + sb$. □

Exemple 3.2. *Soit $f : [0, 2\pi] \to \mathbb{C}$ définie par $f(t) = \mathrm{e}^{it}$. La fonction f est différentiable et on a $Df(t)(h) = i\mathrm{e}^{it}h$ d'où $|||Df(t)||| = 1$ pour tout t. On peut vérifier qu'on a bien $|f(t) - f(s)| \leq t - s$ pour tous $0 \leq s \leq t \leq 2\pi$.*

Par contre le théorème 3.11 n'est plus vrai ! En effet, $f(2\pi) - f(0) = 0$ mais pour tout $t \in [0, 2\pi]$ on a $Df(t)(2\pi) = 2i\pi\mathrm{e}^{it} \neq 0$.

Une conséquence importante de l'inégalité des accroissements finis est la suivante. On rappelle qu'un ensemble U est convexe si pour tous $a, b \in U$ on a $[a, b] \subset U$.

Proposition 3.17. *Soit $U \subset E$ un ouvert convexe et $f : U \to F$ différentiable. Alors $Df \equiv 0$ si et seulement si f est constante.*

La notation $Df \equiv 0$ signifie que pour tout $x \in U$ on a $Df(x) = 0$, i.e. $Df(x)$ est l'application linéaire de E dans F constante égale à 0.

Démonstration. Si f est constante on montre facilement que $Df \equiv 0$. Réciproquement, soient $a, b \in U$, on a $[a, b] \subset U$ et donc

$$\|f(b) - f(a)\|_F \leq \sup_{x \in [a,b]} |||Df(x)||| \times \|b - a\|_E = 0.$$

La fonction f prend la même valeur en deux points quelconques de U, elle est constante. □

On rappelle la définition de maximum/minimum local d'une fonction à valeurs réelles.

Définition 3.18. *Soit $f : U \to \mathbb{R}$. On dit que f possède un minimum, resp. maximum, local en x_0 s'il existe un voisinage V de x_0 tel que pour tout $x \in V$ on ait $f(x) \geq f(x_0)$, resp. $f(x) \leq f(x_0)$. De façon équivalente, il existe $\epsilon > 0$ tel que si $\|x - x_0\| < \epsilon$ alors $f(x) \geq f(x_0)$, resp. $f(x) \leq f(x_0)$.*

Proposition 3.19. *Soit $f : U \to \mathbb{R}$. Si f possède un extremum local en $x_0 \in U$ alors $Df(x_0) = 0$, i.e. c'est l'application nulle de E dans $\mathbb{R}$ (on dit que x_0 est un point critique de f).*

Démonstration. Soit $h \in E$. La fonction $g(t) := f(x_0 + th)$ définie au voisinage de $t = 0$ admet un extremum local en 0 et y est dérivable donc (voir cours de L1) $g'(0) = 0$. Or $g'(0) = Df(x_0)(h)$. On a donc $Df(x_0)(h) = 0$ pour tout $h \in E$, i.e. $Df(x_0) = 0$. □

Remarque 3.6. *Si $E = \mathbb{R}^n$ alors $Df(x_0) = 0$ si et seulement si toutes les dérivées partielles de f s'annulent en x_0.*

3.3 Fonctions de classe C^1 et différentielles partielles

Si f est une fonction différentiable sur un ouvert U, sa différentielle Df est une application définie sur $U \subset E$ et à valeurs dans $L(E, F)$. Muni de $|||\ |||_{L(E,F)}$ ce dernier est un evn, on peut donc s'intéresser à la continuité (ou à la différentiabilité de Df, voir Section 3.4).

Définition 3.20. *Soit $f : U \to F$ une application différentiable. On dira que f est de classe C^1 si $Df : U \to L(E, F)$ est continue.*

Tout comme pour la continuité et la différentiabilité, la somme et la composée de fonctions de classe C^1 est aussi de classe C^1.

Exemple 3.3. *On continue l'Exemple 3.1. La fonction $f(M) = M^2$ est différentiable sur $M_n(\mathbb{R})$ est sa différentielle au point M est $Df(M) : H \mapsto MH + HM$. Soit $M \in M_n(\mathbb{R})$ et $(M_k)_k$ telle que $M_k \to M$ dans $M_n(\mathbb{R})$. On a*

$$\begin{aligned} |||Df(M_k) - Df(M)||| &= \sup_{H \in M_n(\mathbb{R}),\ \|H\|=1} \|(Df(M_k) - Df(M))(H)\| \\ &= \sup_{H \in M_n(\mathbb{R}),\ \|H\|=1} \|M_k H + HM_k - MH - HM\| \\ &\leq \sup_{H \in M_n(\mathbb{R}),\ \|H\|=1} (\|(M_k - M)H\| + \|H(M_k - M)\| \\ &\leq \sup_{H \in M_n(\mathbb{R}),\ \|H\|=1} 2\|M_k - M\|\|H\| \\ &= 2\|M_k - M\|, \end{aligned}$$

et donc $Df(M_k) \to Df(M)$ dans $L(M_n(\mathbb{R}))$ ce qui prouve que Df est continue en M. La fonction f est bien de classe C^1.

Lorsque $E_1, \ldots, E_n$ sont des evn, $U \subset E := E_1 \times \cdots \times E_n$ et $f : U \to F$ on a une notion de différentielle partielle comme on a des dérivées partielles pour les fonctions définies sur $\mathbb{R}^n$. Celle-ci jouera un rôle important au Chapitre 4 dans le Théorème des fonctions implicites.

Définition 3.21. *Soient $U \subset E = E_1 \times \cdots \times E_n$, $f : U \to F$ et $x = (x_1, \ldots, x_n) \in U$. On dit que f admet une différentielle partielle en x par rapport à la $i-$ème variable si l'application $f_i(y) := f(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n)$ définie au voisinage de $x_i \in E_i$ est différentiable au point x_i. On note $D_i f(x) = Df_i(x_i) \in L(E_i, F)$ sa différentielle et elle est appelée différentielle partielle de f en x par rapport à la $i-$ème variable.*

Proposition 3.22. *Soient $U \subset E = E_1 \times \cdots \times E_n$, $f : U \to F$ et $x = (x_1, \ldots, x_n) \in U$. Si f est différentiable en x alors f admet une différentielle partielle en x par rapport à chacune des variables et on a $D_i f(x)(h_i) = Df(x)(0, \ldots, 0, h_i, 0, \ldots, 0)$.*

Démonstration. L'application $h_i \mapsto Df(x)(0, \ldots, 0, h_i, 0, \ldots, 0)$ est clairement linéaire continue. Par ailleurs, en notant $h = (0, \ldots, 0, h_i, 0, \ldots, 0)$ on a

$$f_i(x_i+h_i)-f_i(x_i)-Df(x)(0, \ldots, 0, h_i, 0, \ldots, 0) = f(x+h)-f(x)-Df(x)(h) = o(h) = o(h_i).$$

□

Tout comme l'existence de dérivées partielles n'entraine pas la différentiabilité dans $\mathbb{R}^n$, l'existence de différentielle partielle par rapport à chacune des variables n'entraine pas la différentiabilité ($\mathbb{R}^n$ est le cas particulier où $E_1 = \cdots = E_n = \mathbb{R}$), i.e. la réciproque de la proposition ci-dessus est fausse. Cependant, si on suppose qu'en plus les différentielles partielles sont continues on a alors

Proposition 3.23. *Si f admet une différentielle partielle par rapport à chaque variable et si pour tout i l'application $U \ni x \mapsto D_i f(x) \in L(E_i, F)$ est continue, alors f est de classe C^1.*

Remarque 3.7. *Dans le cas des fonctions de $\mathbb{R}^n \to \mathbb{R}$, c'est en fait ça que vous avez pris comme définition de fonction de classe C^1 : si toutes les dérivées partielles existent et sont continues.*

Démonstration. On fait la démonstration dans le cas de $f : E_1 \times E_2 \to F$. On commence par montrer que f est différentiable. Soit $x = (x_1, x_2)$ et $h = (h_1, h_2)$, on a

$$\begin{aligned} &\|f(x+h) - f(x) - D_1 f(x)(h_1) - D_2 f(x)(h_2)\| \\ &\quad \leq \|f(x+h) - f(x_1, x_2 + h_2) - D_1 f(x)(h_1)\| + \|f(x_1, x_2 + h_2) - f(x) - D_2 f(x)(h_2)\| \end{aligned}$$

On applique ensuite l'inégalité des accroissements finis aux fonctions $f_1(y_1) := f(y_1, x_2 + h_2) - D_1 f(x)(y_1)$ entre x_1 et $x_1 + h_1$ et $f_2(y_2) := f(x_1, y_2) - D_2 f(x)(y_2)$ entre x_2 et $x_2 + h_2$ (ces fonctions sont bien différentiables et même de classe C^1). On a

$$Df_1(y_1)(k_1) = D_1 f(y_1, x_2 + h_2)(k_1) - D_1 f(x)(k_1) \text{ et } Df_2(y_2)(k_2) = D_2 f(x_1, y_2)(k_2) - D_2 f(x)(k_2),$$

d'où,

$$\begin{aligned} &\|f(x+h) - f(x) - D_1 f(x)(h_1) - D_2 f(x)(h_2)\| \\ &\quad \leq \sup_{y_1 \in [x_1, x_1 + h_1]} |||D_1 f(y_1, x_2 + h_2) - D_1 f(x_1, x_2)||| \cdot \|h_1\| \\ &\qquad + \sup_{y_2 \in [x_2, x_2 + h_2]} |||D_2 f(x_1, y_2) - D_2 f(x_1, x_2)||| \cdot \|h_2\| \end{aligned}$$

Puisque $D_1 f$ et $D_2 f$ sont continues en x on en déduit que $f(x+h) - f(x) - D_1 f(x)(h_1) - D_2 f(x)(h_2) = o(h)$. Comme l'application

$$h = (h_1, h_2) \mapsto D_1 f(x)(h_1) + D_2 f(x)(h_2)$$

est linéaire continue, cela prouve que f est différentiable en x avec $Df(x)(h) = D_1 f(x)(h_1) + D_2 f(x)(h_2)$. La continuité de $D_1 f$ et $D_2 f$ entraine finalement celle de Df. □

Proposition 3.24. *Soit $f : E_1 \times \cdots \times E_n \to F$ une application $n-$linéaire continue. Alors f est différentiable sur $E_1 \times \cdots \times E_n$ et*

$$Df(x_1, \ldots, x_n)(h_1, \ldots, h_n) = \sum_{i=1}^{n} f(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_n).$$

En particulier, f admet des différentielles partielles par rapport à chaque variable et

$$D_i f(x)(h_i) = f(x_1, \ldots, x_{i-1}, h_i, x_{i+1}, \ldots, x_n).$$

Exercice 3.4. Démontrer la proposition.

3.4 Différentielles d'ordre supérieur. Étude d'extremas

Définition 3.25. *Soit $f : U \to F$ une application différentiable. On dira que f est deux fois différentiable si $Df : U \to L(E, F)$ est différentiable. La différentielle de Df au point $x \in U$ est appelée différentielle seconde de f au point x, c'est un élément de $L(E, L(E, F))$ et elle est notée $D^2 f(x)$.*

Comme précédemment la somme et la composée de fonctions deux fois différentiables est aussi deux fois différentiable. Il ne semble pas forcément évident de manipuler un objet tel que la différentielle seconde : c'est une application linéaire (continue) définie sur E mais à valeurs dans les applications linéaires (continues) de E dans F. La proposition suivante permet en fait d'identifier cet espace avec celui $L_2(E, F)$ des applications bilinéaires (continues) de E dans F.

Proposition 3.26. *L'application $\varphi : L(E, L(E, F)) \to L_2(E, F)$ définie par $(\varphi(f))(x, y) := (f(x))(y)$ est un isomorphisme et une isométrie, i.e. $|||\varphi(f)|||_{L_2(E,F)} = |||f|||_{L(E,L(E,F))}$. Son inverse est l'application $\psi : L_2(E, F) \to L(E, L(E, F))$ définie par $(\psi(g)(x))(y) := g(x, y)$.*

Démonstration. On montre facilement que si $f \in L(E, L(E, F))$ alors $\varphi(f)$ est une application bilinéaire et que $f \mapsto \varphi(f)$ est linéaire. Montrons que $\varphi(f)$ est continue. Soient $x, y \in E$ on a

$$\|(\varphi(f))(x, y)\|_F = \|(f(x))(y)\|_F \leq |||f(x)|||_{L(E,F)}\|y\|_E \leq |||f|||_{L(E,L(E,F))}\|x\|_E\|y\|_E.$$

En appliquant la Proposition 2.14 on en déduit que $\varphi(f)$ est continue. De plus on a, d'après la Proposition 2.15, $|||\varphi(f)|||_{L_2(E,F)} \leq |||f|||_{L(E,L(E,F))}$.

De même on montre facilement que si $g \in L_2(E, F)$ alors $\psi(g) \in \mathcal{L}(E, \mathcal{L}(E, F))$. Pour tous x, y on a

$$\|(\psi(g)(x))(y)\|_F \leq |||g|||_{L_2(E,F)}\|x\|_E\|y\|_E$$

donc $\psi(g)(x) \in L(E, F)$ et $|||\psi(g)(x)|||_{L(E,F)} \leq |||g|||_{L_2(E,F)}\|x\|_E$ ainsi $\psi(g) \in L(E, L(E, F))$ et $|||\psi(g)|||_{L(E,L(E,F))} \leq |||g|||_{L_2(E,F)}$. Autrement dit $\psi : L_2(E, F) \to L(E, L(E, F))$.

On vérifie facilement que $\varphi \circ \psi = id$ et que $\psi \circ \varphi = id$ ce qui prouve que φ est bien un isomorphisme de $L(E, L(E, F))$ dans $L_2(E, F)$. On a déjà $|||\varphi(f)|||_{L_2(E,F)} \leq |||f|||_{L(E,L(E,F))}$. Pour montrer l'inégalité inverse on écrit

$$|||f|||_{L(E,L(E,F))} = |||\psi \circ \varphi(f)|||_{L(E,L(E,F))} \leq |||\varphi(f)|||_{L_2(E,F)}.$$

$\square$

Par récurrence, on définit les fonctions n-fois différentiables ainsi que la différentielle d'ordre n que l'on identifie, de la même façon que ci-dessus, à une application n-linéaire continue de E dans F.

Exemple 3.4. *Soit $f : M_n(\mathbb{R}) \to \mathbb{R}$ définie par $f(A) = \mathrm{Tr}(A^3)$. On montre alors que pour tout A, $Df(A) : K \mapsto 3\mathrm{Tr}(A^2K)$. Pour calculer la différentielle seconde de f en A on écrit*

$$\begin{aligned} Df(A + tH) - Df(A) : K \quad &\mapsto \quad 3\mathrm{Tr}((A + tH)^2K) - 3\mathrm{Tr}(A^2K) \\ &\qquad = 3t\mathrm{Tr}(AHK) + 3t\mathrm{Tr}(HAK) + 3t^2\mathrm{Tr}(H^2K), \end{aligned}$$

et donc $\lim_{t\to 0}\dfrac{Df(A+tH)-Df(A)}{t} = L(H)$ *où* $L(H) : K \mapsto 3\mathrm{Tr}(AHK+HAK)$. *On vérifie que* L *est une application linéaire de* $M_n(\mathbb{R})$ *dans* $L(M_n(\mathbb{R}),\mathbb{R})$. *Elle est continue (dimension finie). Finalement,* $Df(A+H)-Df(A)-L(H) : K \mapsto 3\mathrm{Tr}(H^2K)$. *On a*

$$|\mathrm{Tr}(H^2K)| = |\sum_{i=1}^n (H^2K)_{ii}| \leq n\sup_i |(H^2K)|_i i \leq n\|H^2K\| \leq n\|H\|^2\cdot\|K\|,$$

d'où $|||Df(A+H)-Df(A)-L(H)||| \leq 3n\|H\|^2 = o(H)$.

Conclusion : f *est deux fois différentiable et* $D^2f(A)(H,K) = 3\mathrm{Tr}(AHK+HAK)$.

Cas des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$

Si $f : U \subset \mathbb{R}^n \to \mathbb{R}$ est différentiable, sa différentielle est

$$Df(x) : (h_1,\ldots,h_n) \mapsto \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)h_i.$$

On montre alors que si f est deux fois différentiable alors f admet des dérivées partielles secondes et que pour $h=(h_1,\ldots,h_n)$ et $k=(k_1,\ldots,k_n)$, via l'isomorphisme précédent, on a

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

La matrice $H_f(x) := \left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x)\right)_{i,j=1,\ldots,n}$ est la matrice associée à la forme bilinéaire $D^2f(x)$ dans la base canonique et est appelée matrice hessienne de f au point x, i.e. pour tous $h,k\in\mathbb{R}^n$ on a $D^2f(x)(h,k) = {}^t hH_f(x)k$.

Vous avez vu que si f est deux fois différentiable en x alors

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(x) = \frac{\partial^2 f}{\partial x_j\partial x_i}(x),$$

c'est le Théorème de Schwarz. Cela se traduit par le fait que la différentielle seconde de f au point x est une forme bilinéaire symétrique, ou encore que la matrice $H_f(x)$ est symétrique. C'est en fait vrai dans un cadre plus général.

Théorème 3.27. *(Théorème de Schwarz). Si* $f : U \to F$ *est deux fois différentiable en* x *alors* $D^2f(x)$*, vue comme application bilinéaire, est symétrique, i.e. pour tous* $h,k\in E$ *on a* $D^2f(x)(h,k) = D^2f(x)(k,h)$.

Démonstration. Pour $h,k\in E$ assez petits, on définit

$$\Delta(h,k) = f(x+h+k) - f(x+h) - f(x+k) + f(x).$$

On donne d'abord l'idée de la preuve. On pose, pour $y\in[0,h]$, $g(y) := f(x+y+k)-f(x+y)$ de sorte que $\Delta(h,k) = g(h)-g(0) = Dg(0)(h)+o(h)$. Par ailleurs, par définition de g, on a $Dg(0)(h) = Df(x+k)(h)-Df(x)(h) = D^2f(x)(k,h)+$reste. D'où $\Delta(h,k) = D^2f(x)(k,h)+$

reste. En intervertissant les rôles de h et k (Δ est symétrique en h, k) on montre de même que $\Delta(h,k) = D^2f(x)(h,k) +$ reste. Toute la difficulté est alors de montrer que les restes se "compensent".

Soit donc g comme ci-dessus. On applique l'inégalité des accroissements finis, Corollaire 3.15. On a alors

$$\|\Delta(h,k) - Dg(0)(h)\| = \|g(h) - g(0) - Dg(0)(h)\| \leq \sup_{y\in[0,h]} |||Dg(y) - Dg(0)||| \cdot \|h\|. \quad (3.3)$$

Pour tout $y \in [0,h]$ on écrit

$$\begin{aligned} Dg(y) - D^2f(x)(k) &= Df(x+y+k) - Df(x+y) - D^2f(x)(k) \\ &= \big(Df(x+y+k) - Df(x) - D^2f(x)(y+k)\big) \\ &\quad - \big(Df(x+y) - Df(x) - D^2f(x)(y)\big)\,. \end{aligned}$$

Puisque f est deux fois différentiable en x, pour tout $\epsilon > 0$, si h et k sont assez petits on a donc pour tout $y \in [0,h]$

$$|||Dg(y) - D^2f(x)(k)||| \leq \epsilon(\|y+k\| + \|y\|) \leq \epsilon(2\|y\| + \|k\|) \leq \epsilon(2\|h\| + \|k\|). \quad (3.4)$$

En particulier

$$|||Dg(0) - D^2f(x)(k)||| \leq \epsilon(2\|h\| + \|k\|). \quad (3.5)$$

En combinant (3.3)-(3.4)-(3.5) on obtient

$$\begin{aligned} \|\Delta(h,k) - D^2f(x)(k,h)\| &\leq \|\Delta(h,k) - Dg(0)(h)\| + \|Dg(0)(h) - D^2f(x)(k,h)\| \\ &\leq \sup_{y\in[0,h]} |||Dg(y) - Dg(0)||| \cdot \|h\| + |||Dg(0) - D^2f(x)(k)||| \cdot \|h\| \\ &\leq (2\epsilon(2\|h\| + \|k\|))\|h\| + \epsilon(2\|h\| + \|k\|)\|h\| \\ &\leq \epsilon(6\|h\|^2 + 3\|h\|\|k\|) \\ &\leq \epsilon 9\|(h,k)\|^2_{E\times E}. \end{aligned}$$

En intervertissant les rôles de h et k on montre de même que pour h, k assez petits on a $\|\Delta(h,k) - D^2f(x)(h,k)\| \leq \epsilon 9\|(h,k)\|^2_{E\times E}$ et donc

$$\|D^2(f)(x)(h,k) - D^2f(x)(k,h)\| \leq \epsilon 18\|(h,k)\|^2_{E\times E}.$$

Puisque $D^2f(x)$ est bilinéaire l'inégalité ci-dessus est vraie pour tous h, k dans E. Si on note $B(h,k) := D^2(f)(x)(h,k) - D^2f(x)(k,h)$ on obtient ainsi que

$$|||B|||_{L_2(E,F)} \leq 18\epsilon.$$

En faisant tendre ϵ vers 0 on conclut que $B \equiv 0$ et donc que $D^2(f)(x)(h,k) = D^2f(x)(k,h)$ □

Exemple 3.5. *On reprend la fonction $f(A) = \mathrm{Tr}(A^3)$ définie sur $M_n(\mathbb{R})$. On a montré qu'elle était deux fois différentiable et que $D^2f(A)(H,K) = 3\mathrm{Tr}(AHK + HAK)$. En utilisant la propriété $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ on a $D^2f(A)(H,K) = 3\mathrm{Tr}(KAH + HAK)$ et on vérifie ainsi que $D^2f(A)$ est bien symétrique.*

Tout comme pour les fonctions d'une variable, on a une formule de Taylor.

Théorème 3.28. *Si $f : U \to F$ est deux fois différentiable alors*

$$f(x+h) = f(x) + Df(x)(h) + \frac{1}{2}D^2f(x)(h,h) + o(\|h\|^2).$$

Démonstration. Soit $R(h) := f(x+h) - f(x) - Df(x)(h) - \frac{1}{2}D^2f(x)(h,h)$. Cette fonction est bien définie au voisinage de 0, elle est de classe C^1 et $R(0) = 0$. Il faut montrer que $R(h) = o(\|h\|^2)$, c'est-à-dire que pour tout $\epsilon > 0$, il existe $\delta > 0$ tel que si $\|h\| < \delta$ alors $\|R(h)\| < \epsilon\|h\|^2$.

Soit donc $\epsilon > 0$. On a

$$\begin{aligned} DR(h)(k) &= Df(x+h)(k) - Df(x)(k) - \frac{1}{2}D^2f(x)(h,k) - \frac{1}{2}D^2f(x)(k,h) \\ &= Df(x+h)(k) - Df(x)(k) - D^2f(x)(h,k), \end{aligned}$$

où on a utilisé le Théorème de Schwarz, autrement dit

$$DR(h) = Df(x+h) - Df(x) - D^2f(x)(h).$$

Puisque f est deux fois différentiable, par définition, $DR(h) = o(\|h\|)$ c'est-à-dire qu'il existe $\delta > 0$ tel que si $\|h\| < \delta$ alors $|||DR(h)||| < \epsilon\|h\|$. D'après l'inégalité des accroissements finis on a donc, pour $\|h\| < \delta$,

$$\|R(h)\| = \|R(h) - R(0)\| \leq \sup_{y\in[0,h]} |||DR(y)||| \cdot \|h\| \leq \sup_{y\in[0,h]} \epsilon\|y\| \cdot \|h\| \leq \epsilon\|h\|^2.$$

□

La différentielle seconde est utile pour étudier les extremas d'une fonction à valeurs réelles. On rappelle que pour une fonction $f : I \to \mathbb{R}$ où I est un intervalle ouvert, si f possède un minimum, resp. maximum, local en x_0 alors $f'(x_0) = 0$ et $f''(x_0) \geq 0$, resp. $f''(x_0) \leq 0$. Réciproquement si x_0 est tel que $f'(x_0) = 0$ et $f''(x_0) > 0$, resp. $f''(x_0) < 0$, alors f possède un minimum, resp. maximum, local strict en x_0. Ces résultats découlent de la formule de Taylor à l'ordre 2 : $f(x_0 + h) = f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + o(h^2)$. En effet, si x_0 est un minimum local on sait que $f'(x_0) = 0$, et on a donc pour h assez petit

$$0 \leq f(x_0+h) - f(x_0) = \frac{h^2}{2}f''(x_0) + o(h^2).$$

En divisant par h^2 et en faisant tendre h vers 0 on obtient bien $f''(x_0) \geq 0$. Réciproquement, si $f'(x_0) = 0$ et $f''(x_0) > 0$ on a

$$f(x_0+h) - f(x_0) = h^2\left(\frac{f''(x_0)}{2} + o(1)\right),$$

et il existe $h_0 > 0$ tel que pour $|h| < h_0$ le terme dans la parenthèse est strictement positif ce qui prouve que $f(x_0+h) > f(x_0)$ pour $h \in]-h_0, h_0[$ et non nul.

Pour une fonction $f : U \to \mathbb{R}$ on a vu, voir Proposition 3.19, que si f possède un extremum local en x_0 alors $Df(x_0) = 0$. On rappelle également la notion de forme bilinéaire (définie) positive, resp. négative.

Définition 3.29. *Soit $L : E \times E \to \mathbb{R}$ une forme bilinéaire. On dit que L est positive, resp. négative, si pour tout $h \in E$ on a $L(h,h) \geq 0$, resp. $L(h,h) \leq 0$. Si $L(h,h) > 0$, resp. $L(h,h) < 0$, pour tout $h \in E\backslash\{0\}$ on dit que L est définie positive, resp. définie négative.*

Remarque 3.8. *Si $E = \mathbb{R}$ on a $D^2f(x)(h,k) = f''(x)hk$ donc $D^2f(x_0)$ est (définie) positive, resp. négative, si et seulement si $f''(x_0)$ est (strictement) positif, resp. négatif.*

On a alors le résultat suivant dans $\mathbb{R}^n$ (et donc dans n'importe quel espace E de dimension finie).

Proposition 3.30. *Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}$ une fonction deux fois différentiable.*

- *Si x_0 est un minimum, resp. maximum, local de f alors $Df(x_0) = 0$ et $D^2f(x_0)$ est positive, resp. négative.*
- *Si x_0 est tel que $Df(x_0) = 0$ et $D^2f(x_0)$ est définie positive, resp. définie négative, alors x_0 est un minimum, resp. maximum, local strict de f.*

Attention, la différentielle seconde (hessienne) en un point x_0 peut n'être ni positive ni négative. Par exemple la matrice $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ vérifie ${}^t\begin{pmatrix} 1 \\ 0 \end{pmatrix} A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 1 > 0$ tandis que ${}^t\begin{pmatrix} 0 \\ 1 \end{pmatrix} A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1 < 0$. Un point critique x_0 dont la hessienne n'est ni positive ni négative est appelé point selle ou point col.

Lemme 3.31. *Soit $A \in M_n(\mathbb{R})$ une matrice symétrique. A est définie positive si et seulement si il existe $m > 0$ tel que pour tout $h \in \mathbb{R}^n$ on ait ${}^thAh \geq m\,{}^thh = m\|h\|_2^2$.*

Ce lemme affirme que si A est définie positive non seulement $h \mapsto {}^thAh > 0$ mais qu'il est minoré par une fonction quadratique (on dit que A est coercive).

Démonstration. On note $S = \{h \in \mathbb{R}^n \mid \|h\|_2 = 1\}$ la sphère unité de $\mathbb{R}^n$ pour la norme $\| \|_2$. En particulier S est compact. Pour tout $h \neq 0$ on a $\frac{{}^thAh}{{}^thh} = {}^t\left(\frac{h}{\|h\|_2}\right) A \left(\frac{h}{\|h\|_2}\right)$ et donc

$$m := \inf_{h \neq 0} \frac{{}^thAh}{{}^thh} = \inf_{h \neq 0} {}^t\left(\frac{h}{\|h\|_2}\right) A \left(\frac{h}{\|h\|_2}\right) = \inf_{h \in S} {}^thAh = \min_{h \in S} {}^thAh,$$

où, à la dernière égalité, on a utilisé la compacité de S et la continuité de $h \mapsto {}^thAh$. Comme A est définie positive on a donc $m > 0$. $\square$

Démonstration de la Proposition. Étant donné $h \in E$, pour t dans un voisinage de 0 on définit $g_h(t) = f(x_0 + th)$ (g_h est une fonction d'une variable). On suppose que x_0 est un minimum local de f. On en déduit que 0 est un minimum local de g_h donc $g_h''(0) \geq 0$. Or $g_h'(t) = Df(x_0 + th)(h)$ donc $g_h''(0) = D^2f(x_0)(h,h)$. Pour tout h on a $D^2f(x_0)(h,h) \geq 0$ et donc $D^2f(x_0)$ est positive.

Réciproquement, si $Df(x_0) = 0$ et $D^2f(x_0)$ est définie positive. Soit $m > 0$ tel que $D^2f(x_0)(h,h) \geq m\|h\|_2^2$. D'après la formule de Taylor à l'ordre 2, pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que si $\|h\|_2 < \delta$ (on est en dimension finie donc toutes les normes sont équivalentes) alors

$$|\underbrace{f(x_0 + h) - f(x_0) - Df(x_0)(h) - \frac{1}{2}D^2f(x_0)(h,h)}_{=:R(h)}| \leq \epsilon\|h\|_2^2.$$

En prenant $\epsilon = \frac{m}{4}$ on a, pour $\|h\| < \delta$,

$$\begin{aligned} f(x_0+h) &= f(x_0) + Df(x_0)(h) + \frac{1}{2}D^2f(x_0)(h,h) + R(h) \\ &\geq f(x_0) + \frac{1}{2}m\|h\|_2^2 - \frac{m}{4}\|h\|_2^2 = f(x_0) + \frac{m}{4}\|h\|_2^2, \end{aligned}$$

et donc x_0 est bien un minimum local strict. □

Pour chercher les extremas d'une fonction de n variables il faut donc

1. chercher ses points critiques,
2. en chaque point critique, étudier si sa hessienne (différentielle seconde) est définie positive ou négative.

Pour étudier si la hessienne est positive ou négative on pourra utiliser le résultat suivant

Proposition 3.32. *Soit A une matrice symétrique. Elle est diagonalisable dans $\mathbb{R}$. De plus, A est (définie) positive, resp. négative, si toutes ses valeurs propres sont (strictement) positives, resp. négatives.*

En particulier en dimension 2, puisque la somme des valeurs propres est la trace de A et leur produit le déterminant de A, une matrice symétrique A est définie positive, resp. négative, si et seulement si $\det(A) > 0$ et $\mathrm{Tr}(A) > 0$, resp. $\mathrm{Tr}(A) < 0$.

Exemple 3.6. *Soit $f(x,y) = 2(x-y)^2 - x^4 - y^4$ définie sur $\mathbb{R}^2$. La fonction f est clairement deux fois différentiable. On a*

$$\frac{\partial f}{\partial x}(x,y) = 4(x-y) - 4x^3, \quad \frac{\partial f}{\partial y}(x,y) = 4(y-x) - 4y^3.$$

Si (x,y) est un point critique, en additionant les deux dérivées partielles on en déduit que $x^3 + y^3 = 0$ et donc $x = -y$. Ainsi (x,y) est point critique si et seulement si $x = -y$ et $8x - 4x^3 = 0$. On trouve ainsi 3 points critiques : $(0,0)$, $(\sqrt{2}, -\sqrt{2})$ et $(-\sqrt{2}, \sqrt{2})$.

On calcule ensuite $H_f(x,y) = \begin{pmatrix} 4-12x^2 & -4 \\ -4 & 4-12y^2 \end{pmatrix}$ et on étudie cette dernière en chacun des points critiques :

- *En $(\sqrt{2}, -\sqrt{2})$ on a $H_f(\sqrt{2}, -\sqrt{2}) = \begin{pmatrix} -20 & -4 \\ -4 & -20 \end{pmatrix}$, qui a pour déterminant $384 > 0$ et pour trace $-40 < 0$. Elle est donc définie négative et $(\sqrt{2}, -\sqrt{2})$ est un maximum local strict.*
- *En $(-\sqrt{2}, \sqrt{2})$ on a $H_f(-\sqrt{2}, +\sqrt{2}) = H_f(\sqrt{2}, -\sqrt{2})$ et donc $(-\sqrt{2}, \sqrt{2})$ est aussi un maximum local strict.*
- *En $(0,0)$ on a $H_f(0,0) = \begin{pmatrix} 4 & -4 \\ -4 & 4 \end{pmatrix}$ qui a pour déterminant 0 et pour trace 8. Elle a donc une valeur propre positive et une valeur propre nulle. Si $(0,0)$ est un extremum c'est un minimum local. Or $f(0,0) = 0$ et pour tout x on a $f(x,x) = -2x^4 < 0$ donc $(0,0)$ ne peut pas être un minimum, ce n'est donc pas un extremum de f.*

Attention ! On fera bien attention que la Proposition 3.30 n'est a priori valable qu'en dimension finie. En dimension quelconque seule la première partie est vraie (voir la preuve). Pour la réciproque il faut des hypothèses un peu plus fortes.

Exercice 3.5. Sur $E = \ell^1(\mathbb{N}^*)$ muni de la norme $\|u\| = \sum_{n\geq 1} |u_n|$ on définit

$$f(u) = \sum_{n\geq 1} \frac{u_n^2}{n^2} - u_n^4.$$

a) Montrer que f est bien définie.

b) Montrer que f est différentiable et que $Df(u)(h) = \sum_{n\geq 1} \left(\frac{2u_n}{n^2} - 4u_n^3\right) h_n$. Vérifier que la suite nulle est point critique de f.

c) Montrer que f est deux fois différentiable et que $D^2f(u)(h,h) = \sum_{n\geq 1} \left(\frac{2}{n^2} - 12u_n^2\right) h_n^2$. En déduire que $D^2f(0)$ est définie positive.

d) Montrer que 0 n'est pas un minimum local de f. Indication : étant donné $\epsilon > 0$, pour $N \in \mathbb{N}$ tel que $\frac{1}{N} < \epsilon$ considérer la suite $u = (u_n)_n$ telle que $u_n = 0$ si $n \neq N$ et $u_N = \epsilon$.

Dans cet exemple, le problème vient du fait que dans chaque direction fixée 0 est un minimum local mais la taille du voisinage sur lequel on a $f(h) \geq f(0)$ dépend de cette direction et peut être arbitrairement petit.

Chapitre 4

Théorèmes d'inversion locale et des fonctions implicites

Dans ce chapitre on va s'intéresser à deux théorèmes importants : le Théorème d'inversion locale et le Théorème des fonctions implicites. On peut résumer grossièrement ces deux théorèmes de la façon suivante : étant donnés une fonction $f : U \to F$ et $b \in F$ peut-on résoudre l'équation $f(x) = b$? Pour comprendre la philosophie de ce que l'on va faire prenons le cas de la dimension finie et supposons que $f : \mathbb{R}^n \to \mathbb{R}^m$ est linéaire, autrement dit on cherche à résoudre un système linéaire de m équations à n inconnues (si A est la matrice de f dans une base choisie l'équation s'écrit $Ax = b$). On peut distinguer 3 cas :

- si $m = n$, on a autant d'équations que d'inconnues et pourvu que A (et donc f) soit inversible on aura alors une unique solution pour tout b donnée par $x = f^{-1}(b)$ (ou encore $x = A^{-1}b$).
- si $m < n$, on a moins d'équations que d'inconnues et on s'attend typiquement à avoir une infinité de solutions. Si on écrit $x = (x_1, x_2) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ et $A = (A_1\ A_2)$ où $A_1 \in M_{m,n-m}(\mathbb{R})$ et $A_2 \in M_m(\mathbb{R})$, pourvu que A_2 soit inversible on pourra exprimer x_2 en fonction de x_1 : $x_2 = A_2^{-1}(b - A_1 x_1)$, i.e. on exprime m variables en fonction des $n - m$ autres, pour chaque choix des $n - m$ premières variables (ici de $x_1 \in \mathbb{R}^{n-m}$) on trouve les valeurs des m autres (ici de $x_2 \in \mathbb{R}^m$).
- si $m > n$, on a plus d'équations que d'inconnues et dans ce cas on résout en général n d'entre elles (on se ramène au 1er cas) et on regarde si la solution trouvée est compatible avec les équations restantes.

On voudrait savoir ce qui reste de cela si $f : \mathbb{R}^n \to \mathbb{R}^m$ n'est plus linéaire mais seulement différentiable. L'idée est alors que pour x proche de x_0 on écrira

$$f(x) = f(x_0) + Df(x_0)(x - x_0) + o(x - x_0) \simeq f(x_0) + Df(x_0)(x - x_0),$$

c'est-à-dire de se ramener à un système linéaire (et de contrôler les termes de reste, c'est là toute la difficulté !). On ne pourra bien sur pas résoudre l'équation sur tout U mais seulement au voisinage d'un point x_0 et pourvu que l'on ait une information sur $Df(x_0)$:

- si $n = m$ et $Df(x_0)$ est inversible on pourra trouver une fonction réciproque f^{-1} au voisinage de x_0, c'est le Théorème d'inversion locale.
- si $m < n$, on décompose comme ci-dessus $x \in \mathbb{R}^n$ en $(x_1, x_2) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ et on écrit $Df(x)(h_1, h_2) = D_1f(x)(h_1) + D_2f(x)(h_2)$. Si $D_2f(x_0)$ est inversible on pourra exprimer x_2 en fonction de x_1 au voisinage de x_0, c'est le Théorème des fonctions

implicites. Par exemple, dans le cas de $f : \mathbb{R}^2 \to \mathbb{R}$ on s'intéresse à des équations du type $f(x, y) = 0$. Typiquement l'ensemble des solutions forme une courbe dans le plan (voir l'Exemple 4.4) et la question est de savoir si on peut décrire celle-ci comme le graphe d'une fonction $y = \varphi(x)$ ou alors $x = \psi(y)$ (au moins localement). L'équation $f(x, y) = 0$ définit y en fonction de x de façon implicite alors que $y = \varphi(x)$ est explicite.

4.1 Le Théorème du point fixe

Chercher la réciproque d'une fonction f revient à résoudre une équation de la forme $f(x) = y$ où y est donné. On souhaite alors montrer qu'il y a une unique solution (f est bijective) et éventuellement à trouver celle-ci (trouver f^{-1}). Le théorème suivant sera l'outil clé de la preuve du théorème d'inversion locale.

Théorème 4.1. *(Théorème du point fixe de Picard) Soit E un espace de* Banach *et $F \subset E$ un fermé (non-vide). Soit $\varphi : F \to E$ telle que $\varphi(F) \subset F$ et φ soit contractante, i.e. il existe $k \in [0, 1[$ tel que pour tous $x, y \in F$ on ait*

$$\|\varphi(y) - \varphi(x)\| \leq k\|y - x\|. \tag{4.1}$$

Alors φ admet un unique point fixe dans F, i.e. il existe un unique $x \in F$ tel que $\varphi(x) = x$.

Démonstration. Soit $x_0 \in F$. On considère la suite $(x_n)_n$ définie par la relation de récurrence $x_{n+1} = \varphi(x_n)$. L'hypothèse $\varphi(F) \subset F$ assure que cette suite est bien définie pour tout n. On va montrer qu'elle converge, sa limite x sera la solution cherchée.

D'après (4.1), pour tout $n \in \mathbb{N}$ on a

$$\|x_{n+2} - x_{n+1}\| = \|\varphi(x_{n+1}) - \varphi(x_n)\| \leq k\|x_{n+1} - x_n\|.$$

On montre alors par récurrence que pour tout $n \in \mathbb{N}$ on a

$$\|x_{n+1} - x_n\| \leq k^n\|x_1 - x_0\|. \tag{4.2}$$

– C'est immédiat pour $n = 0$.
– Soit $n \in \mathbb{N}$, on suppose que $\|x_{n+1} - x_n\| \leq k^n\|x_1 - x_0\|$. On a alors

$$\|x_{n+2} - x_{n+1}\| \leq k\|x_{n+1} - x_n\| \leq k^{n+1}\|x_1 - x_0\|.$$

Comme $k \in [0, 1[$, on déduit de (4.2) par comparaion des séries à termes positifs que la série $\sum_n \|x_{n+1} - x_n\|$ est convergente. La série $\sum_n (x_{n+1} - x_n)$ est ainsi normalement convergente et donc convergente puisque E est un Banach. Cela signifie que la suite de terme général $\sum_{n=0}^{N}(x_{n+1} - x_n) = x_{N+1} - x_0$ converge, autrement dit la suite $(x_n)_n$ converge. On note x sa limite. Comme F est un fermé on a bien $x \in F$. Par ailleurs, la propriété (4.1) implique que φ est continue et en passant à la limite dans l'égalité $x_{n+1} = \varphi(x_n)$ on obtient donc $x = \varphi(x)$.

Finalement, si x et y sont deux points fixes on a

$$\|x - y\| = \|\varphi(x) - \varphi(y)\| \leq k\|x - y\|.$$

Comme $k < 1$ on en déduit que nécessairement $x = y$ ce qui prouve l'unicité. □

Remarque 4.1. *Pour montrer que la suite $(x_n)_n$ converge on est passé par la notion de série normalement convergente. On peut aussi montrer directement à partir de (4.2) que la suite $(x_n)_n$ est de Cauchy et donc converge puisque E est un Banach.*

Exercice 4.1. Montrer la remarque ci-dessus.

Remarque 4.2. *Dans la preuve on construit le point fixe comme limite d'une suite définie par récurrence dont le point de départ est arbitraire. L'unicité du point fixe montre que quelque soit le choix de x_0 la suite $(x_n)_n$ convergera vers celui-ci. Cela donne un algorithme pour trouver des valeurs approchées de ce dernier. Par ailleurs on a*

$$\|x - x_N\| = \left\| \sum_{n \geq N} (x_{n+1} - x_n) \right\| \leq \sum_{n \geq N} \|x_{n+1} - x_n\| \leq \sum_{n \geq N} k^n \|x_1 - x_0\| = k^N \frac{\|x_1 - x_0\|}{1 - k},$$

ce qui permet d'estimer l'erreur entre la valeur approchée x_N et x.

Remarque 4.3. *Un moyen pratique pour montrer que φ est contractante est de calculer sa différentielle (dérivée) et d'utiliser l'inégalité des accroissements finis.*

Exemple 4.1. *Soit $E = C^0([0,1])$ muni de la norme $\| \ \|_\infty$. On considère $\varphi : E \to E$ définie par*

$$(\varphi(f))(x) = \frac{1}{2} \int_0^x f(t) \mathrm{d}t + \frac{1}{4} \cos(f(x)).$$

On veut montrer que φ possède un unique point fixe. Soient $f, g \in E$, pour tout $x \in [0,1]$ on a

$$\begin{aligned} |(\varphi(f))(x) - (\varphi(g))(x)| &\leq \frac{1}{2} \int_0^x |f(t) - g(t)| \mathrm{d}t + \frac{1}{4} |\cos(f(x)) - \cos(g(x))| \\ &\leq \frac{1}{2} \int_0^x \|f - g\|_\infty \mathrm{d}t + \frac{1}{4} |f(x) - g(x)| \\ &\leq \frac{3}{4} \|f - g\|_\infty. \end{aligned}$$

(Pour passer de la première à la deuxième ligne on a utilisé l'inégalité des accroissements finis pour la fonction $\cos : \mathbb{R} \to \mathbb{R}$.) La fonction φ est donc contractante avec $k = \frac{3}{4}$. Comme E muni de la norme $\| \ \|_\infty$ est un Banach, on peut appliquer le théorème du point fixe, ce qui prouve que φ possède un unique point fixe.

4.2 Le Théorème d'inversion locale

Si $f : I \to \mathbb{R}$ est une fonction continue d'une variable, on sait que f est bijective de I sur $f(I)$ si et seulement si elle est strictement monotone (voir cours de L1). Si de plus f est dérivable il suffit donc de s'assurer que $f' > 0$ ou bien $f' < 0$. Dans ce cas f^{-1} est dérivable et on a $(f^{-1})'(y) = \dfrac{1}{f'(f^{-1}(y))}$.

Supposons maintenant f non seulement dérivable mais de classe C^1. Si on a $f'(x) \neq 0$, alors comme f' est continue il existe $\epsilon > 0$ tel que f' soit de signe constant sur l'intervalle

$]x-\epsilon, x+\epsilon[$ (théorème des valeurs intermédiaires). Ainsi f sera bijective de $]x-\epsilon, x+\epsilon[$ dans $f(]x-\epsilon, x+\epsilon[)$, on dira que f est localement inversible au voisinage de x. Par ailleurs la fonction $f^{-1} : f(]x-\epsilon, x+\epsilon[) \to]x-\epsilon, x+\epsilon[$ sera aussi de classe C^1. C'est ce genre de résultat que l'on va généraliser à des fonctions d'un evn dans un evn.

Définition 4.2. *Une application $f : E \to F$ est un homéomorphisme si elle est continue, bijective et si son inverse est continue.*

Théorème 4.3. *(Inversion locale) Soient E, F des Banach, $U \subset E$ un ouvert et $f : U \to F$ de classe C^1. Soit $x_0 \in U$ tel que $Df(x_0)$ soit un homéomorphisme. Alors il existe un voisinage $V \subset U$ de x_0, un voisinage W de $f(x_0)$ et $g : W \to V$ de classe C^1 telle que*

$$f \circ g = id_W \text{ et } g \circ f = id_V.$$

Autrement dit, la fonction $f : V \to W$ est bijective et son inverse est de classe C^1 (on dit que f est un difféomorphisme de V dans W). Par ailleurs, pour tout $y \in W$ on a $Dg(y) = (Df(g(y)))^{-1}$.

Attention ! L'hypothèse que f soit de classe C^1 et pas seulement différentiable est nécessaire. C'est déjà vrai pour les fonctions d'une variable. La fonction f définie par $f(x) = \frac{x}{2} + x^2 \sin(\frac{1}{x})$ si $x \neq 0$ et $f(0) = 0$ est dérivable sur $\mathbb{R}$ (vérifiez-le) et on a $f'(0) = \frac{1}{2} \neq 0$. Cependant f n'est bijective sur aucun voisinage de 0 : sa dérivée $f'(x) = \frac{1}{2} + 2x\sin(\frac{1}{x}) - \cos(\frac{1}{x})$ n'est de signe constant sur aucun voisinage de 0 (pour tout entier n on a $f'(\frac{1}{2n\pi}) = -\frac{1}{2} < 0$ et $f'(\frac{1}{(2n+1)\pi}) = \frac{3}{2} > 0$).

Avant de démontrer ce théorème, regardons ce qu'il devient en dimension finie, i.e. pour $f : \mathbb{R}^n \to \mathbb{R}^d$. On a bien sur $\mathbb{R}^n$ et $\mathbb{R}^d$ qui sont des Banach. Par ailleurs, si f est différentiable, pour que $Df(x) : \mathbb{R}^n \to \mathbb{R}^d$ soit inversible il faut que $n = d$ (théorème du rang) et, si on note $J_f(x)$ sa matrice dans la base canonique (matrice jacobienne), i.e. si $f(x) = (f_1(x), \ldots, f_n(x))$ avec $x = (x_1, \ldots, x_n)$ on a $J_f(x) = \left(\frac{\partial f_i}{\partial x_j}(x)\right)_{1 \le i,j \le n}$, alors $Df(x)$ est inversible si et seulement $\det(J_f(x)) \neq 0$ (c'est le jacobien de f au point x). Dans ce cas son inverse est automatiquement continue puisqu'on est en dimension finie. Autrement dit on a la version suivante du théorème d'inversion locale

Théorème 4.4. *Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}^n$ de classe C^1. Soit $x_0 \in U$ tel que $J_f(x_0)$ soit inversible. Alors il existe un voisinage $V \subset U$ de x_0, un voisinage W de $f(x_0)$ et $g : W \to V$ de classe C^1 telle que*

$$f \circ g = id_W \text{ et } g \circ f = id_V.$$

Autrement dit, la fonction $f : V \to W$ est bijective et son inverse est de classe C^1 (on dit que f est un difféomorphisme de V dans W). Par ailleurs, pour tout $y \in W$ on a $Dg(y) = (Df(g(y)))^{-1}$.

Exemple 4.2. *Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $f(x,y) = \big(\sin(\frac{y}{2}) + x, \sin(\frac{x}{2}) + y\big)$. La fonction f est de classe C^1 et pour tout $(x,y) \in \mathbb{R}^2$ on a*

$$J_f(x,y) = \begin{pmatrix} 1 & \frac{1}{2}\cos(\frac{y}{2}) \\ \frac{1}{2}\cos(\frac{x}{2}) & 1 \end{pmatrix},$$

et donc $\det(J_f(x,y)) = 1 - \frac{1}{4}\cos(\frac{x}{2})\cos(\frac{y}{2}) \neq 0$. *La fonction* f *est donc localement inversible en tout point de* $\mathbb{R}^2$.

Attention ! Contrairement au cas des fonctions de une variable, le fait que $Df(x)$ soit un homéomorphisme pour tout $x \in U$ n'implique pas que f est bijective de U sur $f(U)$. La fonction $f : \mathbb{R}^2 \ni (x,y) \mapsto (\mathrm{e}^x\cos(y), \mathrm{e}^x\sin(y)) \in \mathbb{R}^2$ est de classe C^1. Pour tout $(x,y) \in \mathbb{R}^2$ on a $J_f(x,y) = \begin{pmatrix} \mathrm{e}^x\cos(y) & -\mathrm{e}^x\sin(y) \\ \mathrm{e}^x\sin(y) & \mathrm{e}^x\cos(y) \end{pmatrix}$, et donc $J_f(x,y)$ est inversible puisque $\det(J_f(x,y)) = \mathrm{e}^{2x} \neq 0$. Cependant f n'est pas injective (elle est 2π-périodique dans la variable y).

Démonstration du Théorème 4.3. On note $y_0 = f(x_0)$. On souhaite montrer que si y est proche de y_0 (y est dans un voisinage W de y_0) alors il admet un unique antécédent proche de x_0 (unique dans un voisinage V de x_0). Étant donné $x \in U$, l'application

$$\varphi : z \mapsto f(x) + Df(x_0)(z - x)$$

est une approximation de f, au moins si x est proche de x_0 (on écrit $f(z) \simeq f(x) + Df(x)(z - x) \simeq f(x) + Df(x_0)(z - x)$ où on a utilisé la différentiabilité de f en x puis la continuité de Df en x_0). Pour trouver un antécédent de y par f on va d'abord en chercher un par φ. Puisque $Df(x_0)$ est inversible, pour tout $y \in F$ il existe un unique $z \in E$ tel que $\varphi(z) = y$. On note $G_y(x)$ cet antécédent. On a en fait

$$G_y(x) = x + (Df(x_0))^{-1}(y - f(x)). \tag{4.3}$$

On remarque alors que $y = f(x)$ si et seulement si $G_y(x) = x$. Autrement dit, trouver un antécédent de y par f revient en fait à trouver un point fixe de G_y. On va montrer que si y est suffisament proche de $f(x_0)$ l'application G_y sera contractante sur un voisinage fermé bien choisi de x_0 et aura donc un unique point fixe <u>dans ce fermé</u>, y aura un unique antécédent par f au voisinage de x_0.

Pour montrer que G_y est contractante il suffit de montrer qu'elle est différentiable et de majorer la norme de sa différentielle par $k \in [0,1[$, cf Remarque 4.3. La fonction G_y est bien différentiable pour tout y comme composée de fonctions différentiables, et on a

$$DG_y(x) = I - (Df(x_0))^{-1} \circ Df(x) = (Df(x_0))^{-1} \circ (Df(x_0) - Df(x)). \tag{4.4}$$

Comme Df est continue, il existe $\epsilon > 0$ tel que pour tout $x \in \bar{B}(x_0,\epsilon)$ on ait $|||Df(x) - Df(x_0)||| < \frac{1}{2|||(Df(x_0))^{-1}|||}$ et donc,

$$\forall x \in \bar{B}(x_0,\epsilon),\ \forall y \in F,\ |||DG_y(x)||| < \frac{1}{2}.$$

Pour affirmer que G_y est contractante sur $\bar{B}(x_0,\epsilon)$ et pouvoir utiliser le théorème du point fixe il faut encore s'assurer que

$$G_y(\bar{B}(x_0,\epsilon)) \subset \bar{B}(x_0,\epsilon).$$

D'après l'inégalité des accroissements finis, si $x \in \bar{B}(x_0,\epsilon)$ on a $\|G_y(x) - G_y(x_0)\| < \frac{1}{2}\|x - x_0\| \leq \frac{\epsilon}{2}$. Pour s'assurer que $G_y(x) \in \bar{B}(x_0,\epsilon)$ il suffit (inégalité triangulaire) que

$$\|G_y(x_0) - x_0\| < \frac{\epsilon}{2},$$

c'est-à-dire que $\|(Df(x_0))^{-1}(y-y_0)\| < \frac{\epsilon}{2}$, cf (4.3). On définit donc

$$W := \left\{ y \in F \,|\, y - y_0 \in Df(x_0)\left((B(0, \frac{\epsilon}{2})\right)\right\} \text{ et } V := \{x \in B(x_0,\epsilon) \,|\, f(x) \in W\}.$$

D'après la Proposition 2.4, puisque f et $Df(x_0)$ sont continues V et W sont des ouverts contenant respectivement x_0 et y_0, et pour tout $y \in W$ l'application G_y est une contraction de $\bar{B}(x_0,\epsilon)$ dans $B(x_0,\epsilon) \subset \bar{B}(x_0,\epsilon)$. D'après le théorème du point fixe, il existe un unique $x \in \bar{B}(x_0,\epsilon)$ tel que $G_y(x) = x \iff y = f(x)$. Comme G_y est à valeurs dans $B(x_0,\epsilon)$ on a nécessairement $x \in B(x_0,\epsilon)$ et donc $x \in V$. En résumé, pour tout $y \in W$ il existe un unique $x = g(y) \in V$ tel que $f(x) = y$ ce qui montre que f est bijective de V dans W de réciproque $g : W \to V$.

Il faut maintenant montrer que g est de classe C^1 et que $Dg(y) = (Df(g(y)))^{-1}$. On montre d'abord que g est continue. Soient $y, y' \in W$, on note $x = g(y)$ et $x' = g(y')$. On a

$$\begin{aligned}
\|x - x'\| &= \|G_y(x) - G_{y'}(x')\| \\
&\leq \|G_y(x) - G_y(x')\| + \|G_y(x') - G_{y'}(x')\| \\
&\leq \frac{1}{2}\|x - x'\| + \|(Df(x_0))^{-1}(y - y')\| \\
&\leq \frac{1}{2}\|x - x'\| + |||(Df(x_0))^{-1}||| \times \|(y - y')\|.
\end{aligned}$$

On a donc

$$\|g(y) - g(y')\| = \|x - x'\| \leq 2|||(Df(x_0))^{-1}||| \times \|(y - y')\|, \tag{4.5}$$

ce qui prouve que g est continue.

Pour montrer que g est différentiable sur W, on va montrer que $Df(x)$ est inversible d'inverse continue pour tout $x \in V$ et que $L = (Df(f^{-1}(y))^{-1}$ vérifie (3.1) pour la fonction g. Soit $x \in V$ et $y = f(x)$, d'après (4.4), on a

$$Df(x) = Df(x_0) \circ (I - DG_y).$$

$Df(x_0)$ est inversible d'inverse continue par hyptohèse et $I - DG_y$ aussi d'après le Théorème 2.19 puisque $|||DG_y||| < \frac{1}{2} < 1$ $(y \in W)$. Ainsi $Df(x)$ est inversible et son inverse est continue. Soit maintenant $y \in W$ et $k \in F$ tel que $y + k \in W$. On note $x = g(y)$ et h tel que $x + h = g(y + k)$, i.e. $f(x + h) = y + k$. On a $g(y + k) - g(y) = h$ et par ailleurs, puisque f est différentiable

$$k = f(x + h) - f(x) = Df(x)(h) + \|h\|\epsilon(h)$$

avec $\lim\limits_{h\to 0} \epsilon(h) = 0$. Donc $h = (Df(x))^{-1}(k) - \|h\|(Df(x))^{-1}(\epsilon(h))$ et ainsi

$$g(y + k) - g(y) - (Df(x))^{-1}(k) = -\|h\|(Df(x))^{-1}(\epsilon(h)).$$

On en déduit que

$$\begin{aligned}
&\|g(y + k) - g(y) - (Df(x))^{-1}(k)\| \\
&\leq \|h\| \times \|(Df(x))^{-1}(\epsilon(h))\| \\
&\leq \left(2|||(Df(x_0))^{-1}||| \times \|k\|\right) \times \left(|||(Df(x))^{-1}||| \times \|\epsilon(g(y + k) - g(y))\|\right),
\end{aligned}$$

où on a utilisé (4.5) avec $y' = y + k$ pour majorer $\|h|\|$ en terme de $\|k\|$ (autrement dit pour montrer qu'un $o(h)$ est aussi un $o(k)$). Comme g est continue on $\lim_{k\to 0} \epsilon(g(y+k) - g(y)) = 0$ et donc $g(y+k) - g(y) - (Df(x))^{-1}(k) = o(k)$ ce qui prouve que g est différentiable en y avec $Dg(y) = (Df(x))^{-1} = (Df(f^{-1}(y)))^{-1}$.

Finalement, comme $f^{-1} = g$, Df et $L \mapsto L^{-1}$ sont continues cela prouve que Dg est continue et donc g est de classe C^1. □

Remarque 4.4. *On a montré dans la Proposition 3.10 que l'application $A \mapsto A^{-1}$ était différentiable, et donc continue, en dimension finie uniquement. C'est en fait vrai sur n'importe quel espace de Banach et la preuve est la même (l'application $L \mapsto L^{-1}$ est alors définie sur l'ensemble des homéomorphismes de E dans F).*

Exemple 4.3. *On va utiliser le théorème d'inversion locale pour résoudre l'équation différentielle $f'(x) + f^2(x) = g(x)$ sur $[0, 1]$ avec condition initiale $f(0) = 0$ où g est une fonction continue donnée suffisament petite dans un sens à préciser.*

On note $E = C^1_0([0, 1])$ l'ensemble des fonctions f de classe C^1 (si f est solution elle est dérivable et donc continue, on a alors $f' = g - f^2$ qui est continue donc f est C^1) telles que $f(0) = 0$. On le munit de la norme $\|f\|_E = \sup_{x\in[0,1]} |f'(x)|$ (vérifier que c'est bien une norme). Et soit $F = C^0([0, 1])$ muni de la norme $\| \|_\infty$. On définit $\varphi : E \to F$ par $\varphi(f) = f' + f^2$. Résoudre l'équation revient à trouver $f \in E$ tel que $\varphi(f) = g$. On peut noter que $\varphi(0) = 0$, donc on va chercher à appliquer le théorème d'inversion locale à la fonction φ au point $0 \in E$ (g petit signifie alors g est dans un voisinage de $0 \in F$ pour $\| \|_\infty$).

On sait que $(F, \| \|_\infty)$ est un Banach. On peut montrer que $(E, \| \|_E)$ aussi (voir Exercice 4.2). Pour pouvoir appliquer le Théorème d'inversion locale il faut montrer que φ est de classe C^1 et que $D\varphi(0)$ est un homéomorphisme. On montre que φ est différentiable et que pour tout $f \in E$ on a $D\varphi(f)(h) = h' + 2fh$ (pour montrer la continuité de $D\varphi(f)$ on pourra remarquer que si $f \in E$ on a $|f(x)| \le \int_0^x |f'(t)|\mathrm{d}t \le \|f\|_E$ et donc $\|f\|_\infty \le \|f\|_E$). On a alors

$$|||D\varphi(f) - D\varphi(g)||| = \sup_{\|h\|_E\le 1} \|D\varphi(f)(h) - D\varphi(g)(h)\| \le \sup_{\|h\|_E\le 1} 2\|f - g\|_\infty \|h\|_\infty \le 2\|f - g\|_E,$$

ce qui prouve que $D\varphi$ est continue et donc φ est C^1.

Finalement on a $D\varphi(0) : E \ni h \mapsto h' \in F$. Tout $k \in F$ possède exactement un antécédent par $D\varphi(0)$, la fonction $x \mapsto \int_0^x k(t)\mathrm{d}t$, donc $D\varphi(0)$ est inversible. Par ailleurs on a pour tout $k \in F$

$$\left\|(D\varphi(0))^{-1}(k)\right\|_E = \|((D\varphi(0))^{-1}(k))'\|_\infty = \|k\|_\infty$$

ce qui prouve que $(D\varphi(0))^{-1}$ est continue (de norme 1) et donc $D\varphi(0)$ est bien un homéomorphisme.

D'après le Théorème d'inversion locale, il existe un voisinage V de 0 dans E et un voisinage W de 0 dans F tels que $\varphi : V \to W$ soit bijective, i.e. pour tout $g \in W$ il existe un unique $f \in V$ tel que $f' + f^2 = g$.

Remarque : cela montre l'existence d'une unique solution f seulement au voisinage de 0 (on peut en fait montrer qu'il n'y en a pas d'autre du tout).

Exercice 4.2. On veut montrer que l'espace $E = C_0^1([0,1])$ muni de $\|f\|_E = \sup_{x\in[0,1]} |f'(x)|$ est un Banach. Soit $(f_n)_n$ une suite de Cauchy dans E.

a) Montrer que $(f_n')_n$ converge dans $(C^0([0,1]), \|\ \|_\infty)$. On note g sa limite.

b) Soit f définie sur $[0,1]$ par $f(x) = \int_0^x g(t)\mathrm{d}t$. Vérifier que $f \in E$ et montrer que $(f_n)_n$ converge vers f.

c) Conclure.

4.3 Le Théorème des fonctions implicites

Considérons une fonction $f : U \to \mathbb{R}$ où $U \subset \mathbb{R}^2$. Étant donné $\lambda \in \mathbb{R}$, on considère l'ensemble $E_\lambda = \{(x,y) \in U \,|\, f(x,y) = \lambda\}$. De façon générique, l'ensemble E_λ est une “courbe”, appelée courbe de niveau de la fonction f (pour le niveau λ).

Exemple 4.4. *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^2 + y^2$. Si $\lambda < 0$ alors $E_\lambda = \emptyset$, $E_0 = \{(0,0)\}$, et si $\lambda > 0$ alors E_λ est le cercle de centre $(0,0)$ et de rayon $\sqrt{\lambda}$.*

L'objectif ici est de “savoir” si on peut trouver une fonction de une variable dont cette courbe (ou au moins une partie de cette courbe) serait le graphe, i.e. existe-t-il une fonction φ_λ telle $(x,y) \in E_\lambda$ si et seulement si $y = \varphi_\lambda(x)$? Dans le cas de l'exemple précédent, on sait que l'on ne peut pas décrire toute la courbe à l'aide d'une fonction. Si par exemple $\lambda = 1$, les deux points $(0,1)$ et $(0,-1)$ sont sur le cercle, mais on ne peut pas trouver de fonction $\varphi(x)$ telle que $\varphi(0)$ soit égal à la fois à 1 et à -1. On pourra par contre décrire (par exemple) le demi-cercle $C^+ = \{(x,y) \in \mathbb{R} \times \mathbb{R}^+ \,|\, x^2 + y^2 = 1\}$ comme étant le graphe de la fonction $\varphi(x) = \sqrt{1-x^2}$ définie sur l'intervalle $[-1,1]$.

De façon plus précise, la question que l'on se pose est la suivante :

> Etant donnés une fonction $f(x,y)$, un nombre réel λ et un point (x_0,y_0) tel que $f(x_0,y_0) = \lambda$, peut-on trouver un intervalle (ouvert) I contenant x_0, un intervalle (ouvert) J contenant y_0 et une fonction $\varphi_\lambda(x)$ définie sur I tels que pour tout $(x,y) \in I \times J$ on ait $f(x,y) = \lambda$ si et seulement si $y = \varphi_\lambda(x)$?

On reprend l'exemple précédent, et on fixe $\lambda = R^2 > 0$ (il n'y a rien à faire sinon). On se donne un point (x_0,y_0) sur le cercle de centre $(0,0)$ et de rayon R, en particulier $x_0 \in [-R,R]$. On peut alors distinguer 3 cas :

1. $x_0 \in]-R,R[$ et $y_0 > 0$. Le point (x_0,y_0) est sur le demi-cercle “supérieur”. On peut alors choisir $I =]-R,R[$, $J =]0,+\infty[$ et $\varphi_\lambda(x) = \sqrt{R^2-x^2}$.
2. $x_0 \in]-R,R[$ et $y_0 < 0$. Le point (x_0,y_0) est sur le demi-cercle “inférieur”. On peut alors choisir $I =]-R,R[$, $J =]-\infty,0[$ et $\varphi_\lambda(x) = -\sqrt{R^2-x^2}$.
3. $x_0 = \pm R$ et alors $y_0 = 0$. On considère par exemple $x_0 = R$. Si I est un intervalle ouvert qui contient x_0, il existe $\epsilon > 0$ tel que $[R-\epsilon, R] \subset I$, et de même il existe $\eta > 0$ tel que $[-\eta,\eta] \subset J$. Quitte à diminuer un peu ϵ, les points $(R-\epsilon, \sqrt{R^2-(R-\epsilon)^2})$ et $(R-\epsilon, -\sqrt{R^2-(R-\epsilon)^2})$ sont sur le cercle et les nombres $\sqrt{R^2-(R-\epsilon)^2}$ et $-\sqrt{R^2-(R-\epsilon)^2}$ sont dans J. On devrait alors avoir en même temps $\varphi_\lambda(R-\epsilon) = \sqrt{R^2-(R-\epsilon)^2}$ et $\varphi_\lambda(R-\epsilon) = -\sqrt{R^2-(R-\epsilon)^2}$, ce qui est impossible.

Quelle différence y a-t-il entre les points tels que $x_0 = \pm R$ et les autres ? Graphiquement, celà se voit très bien : ce sont les points où la courbe possède des tangentes verticales. En termes de la fonction f celà se traduit par $\dfrac{\partial f}{\partial y}(x,y) = 0$.

Théorème 4.5. *Soient E, F, G des Banach, $U \subset E \times F$ un ouvert et $f : U \to G$ de classe C^1. Soit $(x_0, y_0) \in U$ tel que $D_2 f(x_0, y_0) \in L(F, G)$ soit un homéomorphisme. Alors il existe des voisinages V de x_0, W de y_0 et une unique application $\varphi : V \to W$ de classe C^1 tels que*

$$\forall (x,y) \in V \times W, \quad f(x,y) = f(x_0,y_0) \iff y = \varphi(x).$$

De plus, pour tout $x \in V$ on a

$$D\varphi(x) = -(D_2 f(x, \varphi(x)))^{-1} \circ D_1 f(x, \varphi(x)). \tag{4.6}$$

En particulier en x_0 on obtient $D\varphi(x_0) = -(D_2 f(x_0, y_0))^{-1} \circ D_1 f(x_0, y_0)$.

Démonstration. L'idée de la démonstration est d'appliquer le théorème d'inversion locale. Pour ça on définit $F : U \to E \times G$ par $F(x,y) = (x, f(x,y))$. La fonction F est de classe C^1 sur U et on a

$$DF(x,y)(h,k) = (h, Df(x,y)(h,k)) = (h, D_1 f(x,y)(h) + D_2 f(x,y)(k)). \tag{4.7}$$

Montrons que $DF(x_0, y_0)$ est un homéomorphisme. Si $(h', k') \in E \times G$ on a

$$\begin{aligned} DF(x_0,y_0)(h,k) = (h',k') &\iff \left\{ \begin{array}{rcl} h &=& h' \\ D_1 f(x_0,y_0)(h) + D_2 f(x_0,y_0)(k) &=& k' \end{array} \right. \\ &\iff \left\{ \begin{array}{rcl} h &=& h' \\ k &=& (D_2 f(x_0,y_0))^{-1}(k' - D_1 f(x_0,y_0)(h')) \end{array} \right. . \end{aligned}$$

$DF(x_0, y_0)$ est donc inversible et

$$(DF(x_0,y_0))^{-1}(h',k') = (h', (D_2 f(x_0,y_0))^{-1}(k' - D_1 f(x_0,y_0)(h')))$$

qui est continue par composition d'applications linéaires continues. $DF(x_0, y_0)$ est donc bien un homéomorphisme.

D'après le théorème d'inversion locale il existe des voisinages U_1 de (x_0, y_0) et U_2 de $F(x_0, y_0) = (x_0, f(x_0, y_0))$ et $G : U_2 \to V_2$ de classe C^1 inverse de F. Quitte à diminuer U_2 on peut de plus supposer que U_1 est de la forme $U_1 = V \times W$ (on remplace U_2 par $F(U \times V) = G^{-1}(U \times V)$ qui est ouvert puisque G est continue). Si $(a,b) \in U_2$ on a

$$(x,y) = G(a,b) \quad \iff \quad (a,b) = F(x,y) = (x, f(x,y)),$$

autrement dit nécessairement $x = a$ et donc G est de la forme $G(a,b) = (a, g(a,b))$. Quitte à restreindre V, pour tout $x \in V$ on a $(x, f(x_0, y_0)) \in U_2$ donc il existe un unique $y = g(x, f(x_0, y_0)) \in W$ tel que $F(x,y) = (x, f(x_0, y_0))$ c'est-à-dire $f(x,y) = f(x_0, y_0)$. On définit alors $\varphi : V \to W$ par $\varphi(x) = g(x, f(x_0, y_0))$ qui est bien de classe C^1.

Il reste à montrer (4.6). Pour tout $x \in V$, $h(x) = f(x, \varphi(x)) = f(x_0, y_0)$, i.e. h est constante, sa différentielle est donc nulle. Or, par différentiation de fonctions composées on a

$$D_1 f(x, \varphi(x)) + D_2 f(x, \varphi(x)) \circ D\varphi(x) = Dh(x) = 0.$$

D'autre part, d'après le Théorème d'inversion locale on sait que $DF(x, y)$ est inversible sur $V \times W$ et donc, cf (4.7), $D_2 f(x, \varphi(x))$ est inversible. On en déduit (4.6). □

Comme pour l'inversion locale, lorsqu'on est en dimension finie on peut voir ce que devient ce théorème. On considère donc $f : U \to \mathbb{R}^m$ avec $U \subset \mathbb{R}^n \simeq \mathbb{R}^{n-m} \times \mathbb{R}^m$ (pour que $D_2 f(x_0, y_0)$ soit inversible il faut que F et G aient même dimension). On notera $p = n - m$ et $x = (x_1, \ldots, x_p) \in \mathbb{R}^p$ et $y = (y_1, \ldots, y_m) \in \mathbb{R}^m$.

Théorème 4.6. *Soient $U \subset \mathbb{R}^p \times \mathbb{R}^m$ un ouvert et $f := (f_1, \ldots, f_m) : U \to \mathbb{R}^m$ de classe C^1. Soit $(a, b) \in U$ tel que*

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \cdots & \frac{\partial f_1}{\partial y_m}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(a,b) & \cdots & \frac{\partial f_m}{\partial y_m}(a,b) \end{pmatrix}$$

soit inversible. Alors il existe des voisinages V de a et W de b et une unique application $\varphi : V \to W$ de classe C^1 tels que

$$\forall (x, y) \in V \times W, \quad f(x, y) = f(a, b) \iff y = \varphi(x).$$

De plus on a

$$J_\varphi(a) = -\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & \cdots & \frac{\partial f_1}{\partial y_m}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(a,b) & \cdots & \frac{\partial f_m}{\partial y_m}(a,b) \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a,b) & \cdots & \frac{\partial f_1}{\partial x_p}(a,b) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a,b) & \cdots & \frac{\partial f_m}{\partial x_p}(a,b) \end{pmatrix}.$$

Remarque 4.5. *Dans le cas $p = m = 1$ on retrouve bien l'hypothèse $\frac{\partial f}{\partial y}(a, b) \neq 0$, et on a alors*

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))}. \tag{4.8}$$

En particulier,

$$\varphi'(a) = -\frac{\frac{\partial f}{\partial x}(a, b)}{\frac{\partial f}{\partial y}(a, b)}.$$

Dans le cas où $\frac{\partial f}{\partial y}(a, b) = 0$ mais $\frac{\partial f}{\partial x}(a, b) \neq 0$ on peut intervertir les rôles de x et y et exprimer x en fonction de y au lieu de y en fonction de x.

Exemple 4.5. *On considère la courbe du plan définie par l'équation $f(x, y) := y^3 - xy - 1 = 0$. Le point $(0, 1)$ est sur la courbe. On montre qu'au voisinage de ce point, la courbe est le graphe d'une fonction $y = \varphi(x)$.*

La fonction f est bien de classe C^1. Par ailleurs $\frac{\partial f}{\partial y}(x, y) = 3y^2 - x$ d'où $\frac{\partial f}{\partial y}(0, 1) = 3 \neq 0$ et on peut donc appliquer le Théorème des fonctions implicites et exprimer y en fonction de x au voisinage de $(0, 1)$.

Le théorème ne donne par contre aucune expression de φ. En utilisant (4.8) la fonction φ vérifie

$$\varphi'(x) = \frac{\varphi(x)}{3\varphi(x)^2 - x} \quad \Longleftrightarrow \quad \varphi'(x) \times (3\varphi(x)^2 - x) - \varphi(x) = 0, \tag{4.9}$$

qui est une équation différentielle pas plus facile à résoudre que $f(x, y) = 0$. On peut cependant utiliser cette dernière pour obtenir un DL de φ en 0 à tout ordre. On sait que φ est C^1 donc, en utilisant (4.9), φ' aussi donc φ est C^2. Par récurrence on en déduit que φ est C^∞ et admet donc un DL à tout ordre.

Par ailleurs on a $\varphi(0) = 1$ par définition de φ et donc $\varphi'(0) = \frac{1}{3}$ d'après (4.9), d'où on obtient $\varphi(x) = 1 + \frac{x}{3} + o(x)$. En dérivant la deuxième expression de (4.9) on a

$$\varphi''(x) \times (3\varphi(x)^2 - x) + \varphi'(x) \times (6\varphi(x)\varphi'(x) - 1) - \varphi'(x) = 0.$$

On en déduit que $\varphi''(0) = 0$ et donc $\varphi(x) = 1 + \frac{x}{3} + o(x^2)$. En continuant ainsi on peut calculer les dérivées successives de φ en 0 et ainsi obtenir un DL à tout ordre.

Exemple 4.6. *On considère l'équation $f(x, y, z) = y + (x+y+z)^2 + x^4 + y^4 + z^4 - 2 = 0$. C'est l'équation d'une surface S dans $\mathbb{R}^3$ (si $f(x, y, z) = ax + by + cz$ est linéaire $f(x, y, z) = 0$ est l'équation d'un plan, pas d'une droite). Le point $P_0 = (x_0, y_0, z_0) = (0, 0, 1)$ est sur S. On a 1 équation et 3 variables, on va chercher à exprimer une variable en fonction des deux autres au voisinage de ce point. La fonction f est bien de classe C^1, on a*

$$\frac{\partial f}{\partial z}(x, y, z) = 2(x + y + z) + 4z^3$$

et donc $\dfrac{\partial f}{\partial z}(0, 0, 1) = 6 \neq 0$. Il existe donc des voisinages $V \subset \mathbb{R}^2$ de $(0, 0)$, $W \subset \mathbb{R}$ de 1 et $\varphi : V \to W$ tels que pour tout $(x, y, z) \in V \times W$ on ait $f(x, y, z) = 0$ si et seulement si $z = \varphi(x, y)$. On a exprimé, au voisinage de P_0, la surface S comme le graphe de la fonction φ.

On a $\varphi(0, 0) = 1$ et par ailleurs $D\varphi(0, 0) = -\left(\frac{\partial f}{\partial z}(0, 0, 1)\right)^{-1} D_1 f(0, 0, 1)$, c'est-à-dire

$$\frac{\partial \varphi}{\partial x}(0, 0) = -\frac{\frac{\partial f}{\partial x}(0, 0, 1)}{\frac{\partial f}{\partial z}(0, 0, 1)} = -\frac{1}{3} \quad \text{et} \quad \frac{\partial \varphi}{\partial y}(0, 0) = -\frac{\frac{\partial f}{\partial y}(0, 0, 1)}{\frac{\partial f}{\partial z}(0, 0, 1)} = -\frac{1}{2}.$$

D'où $\varphi(x, y) = 1 - \dfrac{x}{3} - \dfrac{y}{2} + o(|x| + |y|)$ au voisinage de $(0, 0)$.

Chapitre 5

Extrema liés

Dans ce chapitre on cherche à étudier les extrema d'une fonction $f : U \subset \mathbb{R}^n \to \mathbb{R}$, non pas sur U mais lorsque les variables $x_1, \dots, x_n$ sont liées par une (ou plusieurs) contrainte du type $g(x_1, \dots, x_n) = 0$. Pour voir comment ce genre de situations peut apparaitre, commençons par un exemple simple.

Soit $f(x, y) = 3x + 4y$, on cherche les extrema de f sur le disque unité $D = \{(x, y) \in \mathbb{R} \,|\, x^2 + y^2 \leq 1\}$. Ce dernier est un compact de $\mathbb{R}^2$ et comme f est continue on sait qu'elle admet un minimum et un maximum. Pour les déterminer on calcule la différentielle de f, on a $\dfrac{\partial f}{\partial x}(x, y) = 3$ et $\dfrac{\partial f}{\partial y}(x, y) = 4$. On remarque que la différentielle ne s'annule jamais. Cela n'est pas en contradiction avec l'existence d'un minimum et un maximum ! La Proposition 3.19 ne s'applique que pour une fonction définie sur un *ouvert* et ici D ne l'est pas. Ce qu'on peut en déduire c'est que sur n'importe quel ouvert inclus dans D il n'y a pas de point critique et donc pas d'extremum. Le plus grand ouvert dans D est $\overset{\circ}{D} = \{(x, y) \in \mathbb{R} \,|\, x^2 + y^2 < 1\}$. Le minimum et le maximum de f se trouvent donc nécessairement sur $D \setminus \overset{\circ}{D} = \{(x, y) \in \mathbb{R} \,|\, x^2 + y^2 = 1\}$. Pour les déterminer on cherche donc le minimum et le maximum de $f(x, y) = 3x + 4y$ sous la contrainte que $g(x, y) := x^2 + y^2 - 1 = 0$.

La première idée est d'exprimer y en fonction de x (ou l'inverse) pour que $g(x, y) = 0$, c'est-à-dire que sur la contrainte on a $y = \varphi(x)$ puis d'étudier la fonction de une variable $h(x) = f(x, \varphi(x))$. Ici on peut facilement le faire, $y = \varphi(x) = \pm\sqrt{1 - x^2}$. En général, pourvu que soit $\dfrac{\partial g}{\partial x}$ soit $\dfrac{\partial g}{\partial y}$ ne s'annule pas, on pourra appliquer le Théorème des fonctions implicites qui assure l'existence, au moins locale, d'une telle fonction $y = \varphi(x)$ ou bien $x = \psi(y)$. Comme on l'a vu au chapitre précédent on n'a en général pas d'expression pour cette dernière, cependant on va voir que ce n'est pas nécessaire.

Si x est un extremum de h on aura

$$0 = h'(x) = \frac{\partial f}{\partial x}(x, \varphi(x)) + \varphi'(x) \frac{\partial f}{\partial y}(x, \varphi(x)).$$

Mais on sait, cf (4.8), que $\varphi'(x) = -\dfrac{\frac{\partial g}{\partial x}(x, \varphi(x))}{\frac{\partial g}{\partial y}(x, \varphi(x))}$, et donc

$$\frac{\partial f}{\partial x}(x, \varphi(x)) = \frac{\frac{\partial f}{\partial y}(x, \varphi(x))}{\frac{\partial g}{\partial y}(x, \varphi(x))} \times \frac{\partial g}{\partial x}(x, \varphi(x)).$$

Par ailleurs

$$\frac{\partial f}{\partial y}(x,\varphi(x)) = \frac{\frac{\partial f}{\partial y}(x,\varphi(x))}{\frac{\partial g}{\partial y}(x,\varphi(x))} \times \frac{\partial g}{\partial y}(x,\varphi(x))$$

est évidente. Autrement dit on a

$$Df(x,\varphi(x)) = \lambda Dg(x,\varphi(x)).$$

avec $\lambda = \dfrac{\frac{\partial f}{\partial y}(x,\varphi(x))}{\frac{\partial g}{\partial y}(x,\varphi(x))}$. Là où on a un extremum la différentielle de f n'est pas forcément nulle mais est proportionnelle à la différentielle de g. Finalement, plutôt que d'écrire l'identité ci-dessus en terme de φ (que l'on ne connait toujours pas), on écrira plutôt le système

$$\begin{cases} Df(x,y) &= \lambda Dg(x,y), \\ g(x,y) &= 0. \end{cases}$$

Voyons ce que celà donne sur l'exemple. On a

$$\frac{\partial f}{\partial x}(x,y) = 3,\ \frac{\partial f}{\partial y}(x,y) = 4, \frac{\partial g}{\partial x}(x,y) = 2x,\ \frac{\partial g}{\partial y}(x,y) = 2y.$$

On constate que $\dfrac{\partial g}{\partial x}(x,y)$ et $\dfrac{\partial g}{\partial y}(x,y)$ sont tous les deux nuls uniquement en $(0,0)$ qui ne satisfait pas la contrainte $g(x,y) = 0$. En chaque point (x,y) de la contrainte on pourra donc appliquer la stratégie ci-dessus (exprimer y en fonction de x ou l'inverse via le Théorème des fonctions implicites). On cherche donc x, y et λ tels que

$$\begin{cases} 3 &= 2\lambda x & \left(\frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x}\right) \\ 4 &= 2\lambda y & \left(\frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}\right) \\ x^2+y^2-1 &= 0 & (g(x,y)=0) \end{cases}$$

La première équation impose que $\lambda \neq 0$ et donc on a $x = \frac{3}{2\lambda}$ tandis que $y = \frac{2}{\lambda}$. En remplaçant dans la dernière équation on obtient $\lambda^2 = \frac{25}{4}$ et donc $\lambda = \pm\frac{5}{2}$. On trouve alors les couples (x,y) correspondants : $\left(\frac{3}{5},\frac{4}{5}\right)$ et $\left(-\frac{3}{5},-\frac{4}{5}\right)$. On sait qu'il doit y avoir un maximum et un minimum, l'un de ces points correspond au maximum et l'autre au minimum. Pour décider il suffit de calculer la valeur de f en chacun de ces deux points. On trouve $f\left(\frac{3}{5},\frac{4}{5}\right) = 5$ et $f\left(-\frac{3}{5},-\frac{4}{5}\right) = -5$. Le maximum de f sur D est donc 5 et son minimum -5.

On va voir que la même stratégie s'applique pour des fonctions de n variables et lorsqu'il peut y avoir plusieurs contraintes. Commençons par le cas d'une seule contrainte. On a vu que l'important était de pouvoir appliquer le Théorème des fonctions implicites en n'importe quel point de la contrainte.

Définition 5.1. *Soit $U \subset \mathbb{R}^n$ un ouvert et $g : U \to \mathbb{R}$. On dit que $\Gamma := \{x \in U \mid g(x) = 0\}$ est une contrainte régulière si pour tout $x \in U$ il existe $i \in \{1,\ldots,n\}$ tel que $\dfrac{\partial g}{\partial x_i}(x) \neq 0$. En d'autres termes, si pour tout $x \in U$ on a $Dg(x) \neq 0$, ou encore (puisque $Dg(x) : \mathbb{R}^n \to \mathbb{R}$) si $Dg(x)$ est surjective.*

Théorème 5.2. *Soit U un* ouvert *de $\mathbb{R}^n$, $f, g : U \to \mathbb{R}$ de classe C^1 telles que $\Gamma := \{x \in U \,|\, g(x) = 0\}$ soit une contrainte régulière. Si f possède un extremum local relatif à la contrainte Γ en $a \in \Gamma$ alors il existe (un unique) $\lambda \in \mathbb{R}$ tel que $Df(a) = \lambda Dg(a)$. Le nombre λ est appelé un multiplicateur de Lagrange.*

Attention ! Il est important que f et g soient définies sur un ouvert ! On a vu dans le cas $n = 2$ que l'idée était de chercher un extremum de $h(x) = f(x, \varphi(x))$, φ étant donnée par le Théorème des fonctions implicites, et pour cela d'avoir $h'(x) = 0$. À nouveau, ceci n'est vrai que si x est dans un ouvert ! Il ne faut pas confondre l'ensemble U sur lequel sont définies f et g avec la contrainte Γ sur laquelle on cherche les extrema de f. Dans l'exemple précédent, $U = \mathbb{R}^2$ est bien ouvert tandis que $\Gamma = \{(x, y) \in \mathbb{R}^2 \,|\, x^2 + y^2 - 1 = 0\}$ est lui un compact.

Démonstration. Supposons que f possède un extremum local relatif à la contrainte Γ en a. Puisque la contrainte est régulière, il existe i tel que $\dfrac{\partial g}{\partial x_i}(a) \neq 0$. Quitte à intervertir l'ordre des variables on peut supposer que $i = n$. On écrit $a = (a', a_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$.

D'après le Théorème des fonctions implicites, il existe des voisinages (ouverts) V de a' et W de a_n et une fonction $\varphi : V \to W$ de classe C^1 tels que pour tout $x = (x', x_n) \in V \times W$ on ait $g(x) = 0$ si et seulement si $x_n = \varphi(x')$. Autrement dit, $x \in (V \times W) \cap \Gamma$ si et seulement si $x_n = \varphi(x')$. La fonction $h(x') = f(x', \varphi(x'))$ définie sur l'ouvert V possède donc un extremum local en a' et on a ainsi $Dh(a') = 0$. Si on note $D_1 f(x)$ la différentielle partielle de f par rapport à $x' \in \mathbb{R}^{n-1}$ on a, par différentiation de fonctions composées,

$$0 = Dh(a') = D_1 f(a) + \frac{\partial f}{\partial x_n}(a) \times D\varphi(a'). \tag{5.1}$$

Par ailleurs on a, cf (4.6),

$$D\varphi(a') = -(D_2 g)(a))^{-1} \circ D_1 g(a) = -\left(\frac{\partial g}{\partial x_n}(a)\right)^{-1} \times D_1 g(a).$$

(Ici la “deuxième variable” est $x_n \in \mathbb{R}$ donc $D_2 g(a) : \mathbb{R} \ni k \to \dfrac{\partial g}{\partial x_n}(a)k \in \mathbb{R}$.) En remplaçant $D\varphi(a')$ dans (5.1) on a alors

$$D_1 f(a) = \frac{\partial f}{\partial x_n}(a) \times \left(\frac{\partial g}{\partial x_n}(a)\right)^{-1} \times D_1 g(a).$$

Comme par ailleurs

$$\frac{\partial f}{\partial x_n}(a) = \frac{\partial f}{\partial x_n}(a) \times \left(\frac{\partial g}{\partial x_n}(a)\right)^{-1} \times \frac{\partial g}{\partial x_n}(a)$$

est évidente, on obtient

$$Df(a) = \lambda Dg(a)$$

avec $\lambda = \dfrac{\partial f}{\partial x_n}(a) \times \left(\dfrac{\partial g}{\partial x_n}(a)\right)^{-1}$.

L'unicité de λ est évidente puisque $Dg(a) \neq 0$. $\square$

Passons maintenant au cas où on a plusieurs contraintes $g_1, \dots, g_m$. On prendra toujours $m < n$ (si on a $m \geq n$ on a au moins autant de contraintes que la dimension de l'espace de départ et l'ensemble sur lequel on aura à chercher les extrema de f sera typiquement vide ou un nombre fini de points). On notera $g = (g_1, \dots, g_m) : U \subset \mathbb{R}^n \to \mathbb{R}^m$.

Définition 5.3. *Soit $U \subset \mathbb{R}^n$ un ouvert et $g : U \to \mathbb{R}^m$. On dit que $\Gamma := \{x \in U \mid g(x) = 0\}$ est une contrainte régulière si pour tout $x \in U$ la différentielle $Dg(x)$ est surjective, i.e. $Dg(x)$ est de rang m.*

Cette définition est la généralisation de la Définition 5.1, l'idée étant toujours de pouvoir appliquer le Théorème des fonctions implicites en n'importe quel point de la contrainte Γ.

Théorème 5.4. *Soit U un* ouvert *de $\mathbb{R}^n$, $f : U \to \mathbb{R}$ et $g : u \to \mathbb{R}^m$ de classe C^1 telles que $\Gamma := \{x \in U \mid g(x) = 0\}$ soit une contrainte régulière. Si f possède un extremum local relatif à la contrainte Γ en $a \in \Gamma$ alors il existe (un unique) $\Lambda \in L(\mathbb{R}^m, \mathbb{R})$ tel que $Df(a) = \Lambda \circ Dg(a)$. De façon équivalente, il existe $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ tels que*

$$Df(a) = \lambda_1 Dg_1(a) + \cdots + \lambda_m Dg_m(a).$$

Les λ_j sont appelés des multiplicateurs de Lagrange.

Démonstration. Supposons que f possède un extremum local relatif à la contrainte Γ en a. Puisque la contrainte est régulière, il existe m colonnes de $J_g(a)$ qui sont indépendantes ($J_g(a)$ est de rang m). Quitte à intervertir l'ordre des variables, on peut supposer que ce sont les m dernières. Si on note $x = (y, z) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$, et $a = (b, c)$, cela signifie que $D_2 g(a)$ est inversible et on peut alors appliquer le Théorème des fonctions implicites au point a.

Il existe des voisinages (ouverts) V de b et W de c et une fonction $\varphi : V \to W$ de classe C^1 tels que pour tout $x = (y, z) \in V \times W$ on ait $g(x) = 0$ si et seulement si $z = \varphi(y)$. Autrement dit, $x \in (V \times W) \cap \Gamma$ si et seulement si $z = \varphi(y)$. La fonction $h(y) = f(y, \varphi(y))$ définie sur l'ouvert V possède donc un extremum local en b et on a $Dh(b) = 0$. Si on note $D_1 f(x)$ la différentielle partielle de f par rapport à $y \in \mathbb{R}^{n-m}$ et $D_2 f(x)$ celle par rapport à $z \in \mathbb{R}^m$ on a, par différentiation de fonctions composées,

$$0 = Dh(b) = D_1 f(a) + D_2 f(a) \circ D\varphi(b).$$

Par ailleurs on a, cf (4.6),

$$D\varphi(b) = -(D_2 g)(a))^{-1} \circ D_1 g(a),$$

et donc

$$D_1 f(a) = D_2 f(a) \circ (D_2 g)(a))^{-1} \circ D_1 g(a).$$

Comme par ailleurs

$$D_2 f(a) = D_2 f(a) \circ (D_2 g)(a))^{-1} \circ D_2 g(a)$$

est évidente, on obtient

$$Df(a) = \Lambda Dg(a)$$

avec $\Lambda = D_2 f(a) \circ (D_2 g)(a))^{-1}$. On a bien $\Lambda \in L(\mathbb{R}^m, \mathbb{R})$ puisque $D_2 g(a) \in L(\mathbb{R}^m, \mathbb{R}^m)$, et donc $(D_2 g)(a))^{-1} \in L(\mathbb{R}^m, \mathbb{R}^m)$, et $D_2 f(a) \in L(\mathbb{R}^m, \mathbb{R})$.

L'unicité de Λ découle du fait que $Dg(a)$ est de rang m. En effet, si $\Lambda : (h_1, \dots, h_m) \mapsto \lambda_1 h_1 + \cdots \lambda_m h_m$ on a

$$Df(a) = \lambda_1 Dg_1(a) + \cdots + \lambda_m Dg_m(a),$$

et $Dg_1(a), \dots, Dg_m(a)$ sont linéairement indépendantes puisque $Dg(a)$ est de rang m, ce qui prouve l'unicité des λ_j et donc de Λ. □

Dans la pratique, pour chercher les extrema d'une fontion de n variables lorsqu'on a m contraintes on aura à résoudre un système de $n+m$ équations (n équations exprimant que $Df(a) = \lambda_1 Dg_1(a) + \cdots + Dg_m(a)$, soit $\displaystyle\frac{\partial f}{\partial x_i}(a) = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i}(a)$ pour tout $i = 1, \dots, n$, et m contraintes) à $n+m$ inconnues ($x_1, \dots, x_n$ et $\lambda_1, \dots, \lambda_m$).

Exemple 5.1. *On veut trouver les extrema de la fonction $f(x,y,z) = x^2 + y^2 + z^2$ sous la contrainte*

$$\left\{ \begin{array}{rcl} 3x + y &=& 1, \\ x^2 + 2z^2 &=& 4. \end{array} \right.$$

Si on note $g = (g_1, g_2)$ avec $g_1(x,y,z) = 3x + y - 1$ et $g_2(x,y,z) = x^2 + 2z^2 - 4$, on cherche donc les extrema de f sur l'ensemble $\Gamma := \{(x,y,z) \in \mathbb{R}^3 \,|\, g(x,y,z) = 0\}$. L'ensemble Γ est fermé dans $\mathbb{R}^3$ et il est borné (en effet si $(x,y,z) \in \Gamma$ on doit avoir $x^2 \leq 4$, duquel on déduit $y^2 = (1-3x)^2 \leq 49$ et $z^2 \leq 2$). Donc Γ est compact et comme f est continue elle possède un minimum et un maximum global.

On vérifie ensuite que Γ est une contrainte régulière. On a

$$J_g(x,y,z) = \begin{pmatrix} 3 & 1 & 0 \\ 2x & 0 & 4z \end{pmatrix}.$$

On voit que $J_g(x,y,z)$ est toujours au moins de rang 1 (la première ligne n'est jamais nulle) et qu'elle est de rang 2 sauf si $\begin{pmatrix} 2x \\ 0 \\ 4z \end{pmatrix} = \lambda \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$, ce qui n'est le cas que si $x = z = 0$. Comme $g_2(0,y,0) = -4 \neq 0$, de tels points ne sont pas sur la contrainte et donc en chaque point de la contrainte J_g est de rang 2. On peut donc appliquer le Théorème 5.4, si (x,y,z) est un extremum de f sur Γ il existe λ_1, λ_2 tels que $Df(x,y,z) = \lambda_1 Dg_1(x,y,z) + \lambda_2 Dg_2(x,y,z)$. On doit donc résoudre le système suivant

$$\left\{ \begin{array}{rcll} 2x &=& 3\lambda_1 + 2x\lambda_2 & \left(\frac{\partial f}{\partial x} = \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2 \frac{\partial g_2}{\partial x}\right) \\ 2y &=& \lambda_1 & \left(\frac{\partial f}{\partial y} = \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2 \frac{\partial g_2}{\partial y}\right) \\ 2z &=& 4z\lambda_2 & \left(\frac{\partial f}{\partial z} = \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2 \frac{\partial g_2}{\partial z}\right) \\ 3x + y - 1 &=& 0 & (g_1(x,y,z) = 0) \\ x^2 + 2z^2 - 4 &=& 0 & (g_2(x,y,z) = 0) \end{array} \right.$$

La 3è équation donne $z = 0$ ou $\lambda_2 = \frac{1}{2}$. Si $z = 0$ on déduit de la dernière équation que $x = \pm 2$ puis on trouve y à partir de la 4è, λ_1 de la 2è et enfin λ_2 de la 1è. On obtient ainsi les points $(2, -5, 0)$ avec $\lambda_1 = -10$ et $\lambda_2 = \frac{17}{2}$ et $(-2, 7, 0)$ avec $\lambda_1 = 14$ et $\lambda_2 = \frac{23}{2}$. Si $\lambda_2 = \frac{1}{2}$

on a alors le système

$$\left\{\begin{array}{rcl} x & = & 3\lambda_1 \\ 2y & = & \lambda_1 \\ 3x+y-1 & = & 0 \\ x^2+2z^2-4 & = & 0 \end{array}\right.$$

et on trouve que $\lambda_1 = \frac{2}{19}$, $x = \frac{6}{19}$, $y = \frac{1}{19}$ *et* $z^2 = \frac{704}{361}$*. On a ainsi les deux points* $\left(\frac{6}{19}, \frac{1}{19}, \pm\frac{8\sqrt{11}}{19}\right)$ *avec* $\lambda_1 = \frac{2}{19}$ *et* $\lambda_2 = \frac{1}{2}$.

Pour finir on calcule la valeur de f *en chacun de ces points :*

$$f(2,-5,0) = 29,\ \ f(-2,7,0) = 53,\ \ f\left(\frac{6}{19}, \frac{1}{19}, \frac{8\sqrt{11}}{19}\right) = f\left(\frac{6}{19}, \frac{1}{19}, \pm\frac{8\sqrt{11}}{19}\right) = \frac{741}{361}.$$

On sait que au moins l'un de ces points correspond au minimum et au moins l'un au maximum. On a finalement $\min f = \frac{741}{361}$ *(atteint en deux points) et* $\max f = 53$.

Chapitre 6

Équation d'Euler-Lagrange

L'objet de ce chapitre est de trouver le minimum de quantités du type

$$F(y) = \int_a^b f(x, y(x), y'(x))\mathrm{d}x, \tag{6.1}$$

où $y \in C^1([a,b])$ et pour lesquelles on fixe en général des contraintes du type $y(a) = y_a$ et $y(b) = y_b$, i.e. les valeurs de la fonction y aux bords de l'intervalle $[a,b]$.

Commençons par un exemple simple. On souhaite trouver parmi toutes les courbes décrites comme le graphe d'une fonction et passant par deux points donnés celle dont la longueur est minimum (la réponse est bien entendu "un segment de droite"). La longueur de la courbe d'équation $y = g(x)$ entre les points de coordonnées $(a, y(a))$ et $(b, y(b))$ (on supposera $a < b$) est donnée par $L = \int_a^b \sqrt{1 + g'(x)^2}\mathrm{d}x$. On cherche donc à trouver le minimum de la "fonctionnelle" F définie par

$$F(y) = \int_a^b \sqrt{1 + y'(x)^2}\mathrm{d}x$$

sur l'ensemble $\mathcal{E} := \{y \in C^1([a,b]) \,|\, y(a) = y_a, y(b) = y_b\}$.

Étant donnée $f : U \ni (x, y, p) \mapsto f(x, y, p) \in \mathbb{R}$ une fonction de classe C^1 définie sur un ouvert $U \subset \mathbb{R}^3$, on s'intéressera donc à la minimisation de quantités du type (6.1) sur des ensembles de la forme

$$\mathcal{A} := \{y \in C^1([a,b]) \,|\, y(a) = y_a,\ y(b) = y_b,\ (x, y(x), y'(x) \in U\ \forall x \in [a,b]\}, \tag{6.2}$$

avec y_a et y_b fixés. On cherche donc le minimum d'une "fonction" définie sur un sous-ensemble de $C^1([a,b])$. On munira ce dernier de la norme

$$\|y\|_{C^1} := \|y\|_\infty + \|y'\|_\infty.$$

Muni de cette norme c'est un espace de Banach. Pour trouver un minimum de F l'idée est de calculer sa différentielle et de chercher quand cette dernière s'annule. Cependant on cherche le minimum sur l'ensemble $\mathcal{A}$ qui n'est pas un ouvert de C^1. Ce n'est pas non plus un sous-espace vectoriel de ce dernier (on aurait alors directement étudié F sur $\mathcal{A}$ muni de la norme $\|\ \|_{C^1}$).

Soit $\mathcal{E} := \{y \in C^1([a,b]) \,|\, y(a) = y_a,\ y(b) = y_b\}$. L'ensemble $\mathcal{E}$ est un sous-espace dit *affine* de $C^1([a,b])$: l'ensemble $E_0 = \{y \in C^1([a,b]) \,|\, y(a) = y(b) = 0\}$ est un sous-espace vectoriel de C^1 et si y_0 est un élément fixé de $\mathcal{E}$ alors $y \in \mathcal{E}$ si et seulement si $y - y_0 \in E_0$. On dit que E_0 est la direction de $\mathcal{E}$. Étant donné un evn E_0 et un espace affine de direction E_0, on munit ce dernier de la distance induite par la norme sur E_0 : pour tous $x, y \in \mathcal{E}$, $d(x,y) = \|x-y\|_{E_0}$. On a alors les adaptations directes suivantes de la Définition 3.1 et de la Proposition 3.19

Définition 6.1. *Soit $f : \mathcal{A} \to F$ où $\mathcal{A}$ est un ouvert d'un espace affine $\mathcal{E}$ de direction E_0. On dit que f est différentiable en x_0 s'il existe $L \in L(E_0, F)$ telle que pour tout $h \in E_0$ tel que $x_0 + h \in \mathcal{A}$ on ait $f(x_0+h) = f(x_0) + L(h) + o(h)$. Lorsqu'elle existe L est unique, elle est appelée différentielle de f en x_0 et notée $Df(x_0)$.*

Remarque 6.1. *Dire que f est différentiable en x_0 est en fait équivalent à dire que la fonction $g : h \mapsto f(x_0+h)$, qui est définie sur l'ouvert $U := \mathcal{A} - \{x_0\} = \{h \in E_0 \,|\, x_0 + h \in \mathcal{A}\}$ de l'evn E_0, est différentiable en 0. Par ailleurs on a alors $Df(x_0) = Dg(0)$.*

On admettra que que l'ensemble $\mathcal{A}$ défini par (6.2) est un ouvert de l'espace affine $\mathcal{E}$ (si f est définie sur tout $\mathbb{R}^3$ on a simplement $\mathcal{A} = \mathcal{E}$).

Proposition 6.2. *Soit $f : \mathcal{A} \to \mathbb{R}$ une fonction différentiable sur un ouvert $\mathcal{A}$ d'un espace affine $\mathcal{E}$. Si f possède un extremum local en x_0 alors $Df(x_0) = 0$.*

Démonstration. La fonction $g : E_0 \ni h \mapsto f(x_0+h)$ est définie au voisinage de 0 et admet un extremum local en 0. Comme elle est différentiable en 0 on a $Df(x_0) = Dg(0) = 0$. □

Proposition 6.3. *Soit U un ouvert de $\mathbb{R}^3$ et $f : U \to \mathbb{R}$ de classe C^1. Alors la fonction $F : \mathcal{A} \to \mathbb{R}$ définie par (6.1) est différentiable sur $\mathcal{E}$ et*

$$DF(y) : h \mapsto \int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x.$$

Démonstration. L'application

$$E_0 \ni h \mapsto \int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x$$

est clairement linéaire et on vérifie qu'elle est continue. En effet, on a

$$\begin{aligned} &\left| \int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x \right| \\ &\qquad \leq \int_a^b \left(\left| \frac{\partial f}{\partial y}(x, y(x), y'(x)) \right| + \left| \frac{\partial f}{\partial p}(x, y(x), y'(x)) \right| \right) \mathrm{d}x \times \|h\|_{C^1}. \end{aligned}$$

Soit $y \in \mathcal{A}$ fixé et $h \in E_0$ assez petit, i.e. tel que $y + h \in \mathcal{A}$. On écrit

$$\begin{aligned} &F(y+h) - F(y) - \int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x \\ &= \int_a^b \Bigg(f(x, y(x)+h(x), y'(x)+h'(x)) - f(x, y(x), y'(x)) - \frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) \\ &\qquad\qquad - \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \Bigg) \mathrm{d}x \\ &= \int_a^b \Big(f(x, y(x)+h(x), y'(x)+h'(x)) - f(x, y(x), y'(x)) - Df(x, y(x), y'(x))(0, h(x), h'(x)) \Big). \end{aligned}$$

En notant $u(x) = (x, y(x), y'(x))$ et $v(x) = (x, y(x) + h(x), y'(x) + h'(x))$, d'après l'inégalité des accroissements finis (Corollaire 3.15) on a donc

$$\begin{aligned}
&\left| F(y+h) - F(y) - \int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x \right| \\
&\qquad \leq \int_a^b \left(\sup_{k \in [u(x), v(x)]} |||Df(k) - Df(u(x))||| \cdot \underbrace{\|v(x) - u(x)\|_{\mathbb{R}^3}}_{=|h(x)|+|h'(x)|} \right) \mathrm{d}x \\
&\qquad \leq \left(\int_a^b \sup_{k \in [u(x), v(x)]} |||Df(k) - Df(u(x))||| \mathrm{d}x \right) \|h\|_{C^1}.
\end{aligned}$$

Comme y est de classe C^1 l'ensemble $\Gamma := \{(x, y(x), y'(x)) \mid x \in [a, b]\}$ est un compact de $\mathbb{R}^3$ (muni par exemple de la norme $\|(x, y, p)\|_{\mathbb{R}^3} = |x| + |y| + |p|$). La fonction f est C^1 donc Df est continue. Elle est donc uniformément continue sur Γ (Théorème 2.8). Soit $\epsilon > 0$, il existe $\delta > 0$ tel que pour tous $u, v \in \Gamma$ on ait

$$\|u - v\|_{\mathbb{R}^3} < \delta \implies |||Df(u) - Df(v)||| < \epsilon.$$

Si $\|h\|_{C^1} < \delta$, pour tout $x \in [a, b]$ on a $\|(0, h(x), h'(x))\|_{\mathbb{R}^3} = |h(x)| + |h'(x)| \leq \|h\|_{C^1} < \delta$ et donc pour tout $k \in [u(x), v(x)]$ on a

$$\|k - u(x)\|_{\mathbb{R}^3} \leq \|(0, h(x), h'(x))\|_{\mathbb{R}^3} < \delta.$$

On en déduit que pour tout $x \in [a, b]$ on a

$$\sup_{k \in [u(x), v(x)]} |||Df(k) - Df(u(x))||| < \epsilon,$$

et donc

$$\left| F(y+h) - F(y) - \int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x \right| < \epsilon(b-a)\|h\|_{C^1}.$$

On a bien montré que $F(y + h) - F(y) - DF(y)(h) = o(h)$. □

Pour trouver le minimum de F on va donc commencer par chercher les points critiques de F c'est-à-dire que l'on doit trouver les fonctions $y \in \mathcal{A}$ telles que pour tout $h \in C^1([a, b])$ avec $h(a) = h(b) = 0$ on ait

$$\int_a^b \left(\frac{\partial f}{\partial y}(x, y(x), y'(x))h(x) + \frac{\partial f}{\partial p}(x, y(x), y'(x))h'(x) \right) \mathrm{d}x = 0.$$

Théorème 6.4. *Soit $y \in \mathcal{A}$ un point critique de F, alors y vérifie l'équation dite de Euler-Lagrange : pour tout $x \in [a, b]$ on a*

$$\frac{\partial f}{\partial y}(x, y(x), y'(x)) = \frac{\mathrm{d}}{\mathrm{d}x} \frac{\partial f}{\partial p}(x, y(x), y'(x)).$$

Attention ! La fonction $f(x,y,p)$ est une fonction de 3 variables et $\frac{\partial f}{\partial y}(x,y(x),y'(x))$ et $\frac{\partial f}{\partial p}(x,y(x),y'(x))$ représentent les dérivées partielles de f par rapport aux variables y et p respectivement, ces dernières étant calculées au point $(x,y(x),y'(x))$.

La notation $\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial p}(x,y(x),y'(x))$ signifie alors la dérivée de la fonction d'une variable $x \mapsto \frac{\partial f}{\partial p}(x,y(x),y'(x))$.

Par exemple, si $f(x,y,p) = y^2 + xp + p^2$ on a $\frac{\partial f}{\partial y}(x,y,p) = 2y$ donc $\frac{\partial f}{\partial y}(x,y(x),y'(x)) = 2y(x)$, $\frac{\partial f}{\partial p}(x,y,p) = x + 2p$ donc $\frac{\partial f}{\partial p}(x,y(x),y'(x)) = x + 2y'(x)$ et ainsi $\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial p}(x,y(x),y'(x)) = 1 + 2y''(x)$. L'équation de Euler-Lagrange correspondante est alors

$$\frac{\partial f}{\partial y}(x,y(x),y'(x)) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial p}(x,y(x),y'(x)) \quad \Longleftrightarrow \quad 2y(x) = 1 + 2y''(x).$$

Voyons ce que cela donne sur l'exemple ci-dessus. On a $f(x,y,p) = \sqrt{1+p^2}$ qui est une fonction de classe C^1 sur $\mathbb{R}^3$. On a $\frac{\partial f}{\partial y}(x,y,p) = 0$ et $\frac{\partial f}{\partial p}(x,y,p) = \frac{p}{\sqrt{1+p^2}}$. L'équation de Euler-Lagrange s'écrit donc

$$\frac{\mathrm{d}}{\mathrm{d}x}\frac{y'(x)}{\sqrt{1+(y'(x))^2}} = 0.$$

Autrement dit $\frac{y'(x)}{\sqrt{1+(y'(x))^2}}$ est constante. On note α cette valeur. On en déduit que $(y'(x))^2 = \frac{\alpha^2}{1-\alpha^2}$ est aussi constante et donc y' aussi (c'est une conséquence du théorème des valeurs intermédiaires). Autrement dit y est une fonction affine et son graphe sur $[a,b]$ est bien un segment de droite.

Lemme 6.5. *Soit $g \in C^0([a,b])$ telle que pour tout $h \in E_0$ on ait $\int_a^b g(x)h(x)\mathrm{d}x = 0$, alors $g = 0$.*

On commence par démontrer le Théorème à partir du lemme, on démontrera le lemme ensuite.

Démonstration. Pour simplifier la preuve on supposera que $\frac{\partial f}{\partial p}$ n'est pas seulement continue mais de classe C^1. Étant donnée $h \in E_0$ on a donc

$$\int_a^b \frac{\partial f}{\partial y}(x,y(x),y'(x))h(x)\mathrm{d}x + \int_a^b \frac{\partial f}{\partial p}(x,y(x),y'(x))h'(x)\mathrm{d}x = 0.$$

On effectue une intégration par parties dans la seconde intégrale. On a alors

$$\begin{aligned}&\int_a^b \frac{\partial f}{\partial p}(x,y(x),y'(x))h'(x)\mathrm{d}x \\ &= \frac{\partial f}{\partial p}(b,y(b),y'(b))h(b) - \frac{\partial f}{\partial p}(a,y(a),y'(a))h(a) - \int_a^b \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial p}(x,y(x),y'(x)) \times h(x)\mathrm{d}x.\end{aligned}$$

Puisque $h \in E_0$ on a $h(a) = h(b) = 0$ et donc

$$\begin{aligned}0 &= \int_a^b \frac{\partial f}{\partial y}(x,y(x),y'(x))h(x)\mathrm{d}x + \int_a^b \frac{\partial f}{\partial p}(x,y(x),y'(x))h'(x)\mathrm{d}x \\ &= \int_a^b \left(\frac{\partial f}{\partial y}(x,y(x),y'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial p}(x,y(x),y'(x))\right) h(x)\mathrm{d}x.\end{aligned}$$

C'est vrai pour tout $h \in E_0$ et le résultat découle donc du lemme. $\square$

Démonstration du lemme. On montre que g est nulle sur $]a,b[$. La continuité de g assure alors qu'on aura également $g(a) = g(b) = 0$. On raisonne par l'absurde, supposons donc qu'il existe $x_0 \in]a,b[$ tel que $g(x_0) \neq 0$. Quitte à changer g en $-g$ on peut supposer que $g(x_0) > 0$. Comme g est continue, il existe $\alpha > 0$ tel que $g(x) \geq 0$ pour tout $x \in [x_0 - \alpha, x_0 + \alpha]$. L'idée est alors de choisir $h \in E_0$ non nulle telle que $h(x) \geq 0$ sur ce voisinage et h soit nulle en dehors. On prend par exemple la fonction h définie par

$$h(x) = \begin{cases} 1 + \sin\left(\dfrac{\pi}{\alpha}(x - x_0) + \dfrac{\pi}{2}\right) & \text{si} \qquad x \in [x_0 - \alpha, x_0 + \alpha], \\ 0 & \text{sinon.} \end{cases}$$

On a alors

$$0 = \int_a^b g(x)h(x)\mathrm{d}x = \int_{x_0-\alpha}^{x_0+\alpha} g(x)h(x)\mathrm{d}x.$$

Sur l'intervalle $[x_0 - \alpha, x_0 + \alpha]$ la fonction gh est positive et continue, elle est donc nulle. Or $g(x_0)h(x_0) = 2g(x_0)$ ce qui contredit $g(x_0) > 0$. $\square$

Exemple 6.1. *Soient $a < b$ et $y_a, y_b > 0$. Trouver les points critiques de la fonctionnelle $F(y) := 2\pi \int_a^b y(x)\sqrt{1 + (y'(x))^2}\mathrm{d}x$ sur $\{y \in C^1([a,b]) \,|\, y(a) = y_a,\ y(b) = y_b\}$. (Cela correspond à chercher la courbe donnée par le graphe d'une fonction $y = y(x)$ et passant par les points $A(a, y_a)$ et $B(b, y_b)$ telle que la surface de révolution engendrée par cette dernière autour de l'axe des abscisses ait une surface minimale.)*

La fonction $f(x,y,p) = 2\pi y\sqrt{1+p^2}$ est de classe C^1 sur $\mathbb{R}^3$ donc les points critiques de F vérifie l'équation d'Euler-Lagrange. On a $\frac{\partial f}{\partial y}(x,y,p) = 2\pi\sqrt{1+p^2}$ et $\frac{\partial f}{\partial p}(x,y,p) = 2\pi y \frac{p}{\sqrt{1+p^2}}$. Donc y vérifie

$$\begin{aligned} & \sqrt{1+(y')^2} = \left(\frac{yy'}{\sqrt{1+(y')^2}}\right)' \\ \iff \quad & \sqrt{1+(y')^2} = \frac{(yy'' + (y')^2)\sqrt{1+(y')^2} - yy'\frac{y'}{\sqrt{1+(y')^2}}}{1+(y')^2} \\ \iff \quad & (1+(y')^2)^2 = (yy'' + (y')^2)(1+(y')^2) - y(y')^2. \end{aligned}$$

Comme dans l'exemple ci-dessus, l'équation d'Euler-Lagrange est typiquement une équation différentielle du second ordre (on y dérive $\dfrac{\partial f}{\partial p}(x, y(x), y'(x))$) très difficile à résoudre. Il y a cependant certains cas un peu plus simples :

1. La fonction f ne dépend pas de p, i.e. $f = f(x,y)$. Dans ce cas l'équation d'Euler-Lagrange s'écrit $\dfrac{\partial f}{\partial y}(x, y(x)) = 0$ et il "suffit" d'exprimer y en fonction de x. Par exemple, on voudrait minimiser $\displaystyle\int_1^2 xy(x) - \mathrm{e}^{y(x)}\,\mathrm{d}x$ avec $y(1) = 0$ et $y(2) = 2$. On a $f(x,y,p) = xy - \mathrm{e}^y$ et l'équation d'Euler-Lagrange s'écrit $x - \mathrm{e}^{y(x)} = 0$ soit $y(x) = \ln(x)$. Cette dernière ne vérifie pas les conditions au bord ($y(2) \neq 2$) et il n'y a donc pas de point critique.

2. La fonction f ne dépend pas de y, i.e. $f = f(x,p)$. Dans ce cas l'équation d'Euler-Lagrange s'écrit $\dfrac{\mathrm{d}}{\mathrm{d}x}\dfrac{\partial f}{\partial p}(x,y'(x)) = 0$, donc $\dfrac{\partial f}{\partial p}(x,y'(x))$ est constante. On essaie alors d'exprimer $y'(x)$ en fonction de x puis d'en calculer une primitive. Par exemple, on voudrait minimiser $\displaystyle\int_1^2 x^2y'(x)-(y'(x))^2\,\mathrm{d}x$ avec $y(0)=0$ et $y(1)=1$. On a $f(x,y,p) = x^2p-p^2$ et l'équation d'Euler-Lagrange s'écrit $\frac{\mathrm{d}}{\mathrm{d}x}(x^2-2y'(x)) = 0$ soit $x^2-2y'(x) = C$ où $C\in\mathbb{R}$ est une constante. On en déduit que $y'(x) = \frac{x^2-C}{2}$ et donc $y(x) = \frac{x^3}{6}-\frac{Cx}{2}+D$ où $D\in\mathbb{R}$. On cherche ensuite C et D pour satisfaire $y(0)=0$ et $y(1)=1$. On trouve que $D=0$ puis $C=-\frac{5}{3}$, il y a donc un unique point critique qui est la fonction $y(x) = \frac{x^3+5x}{6}$.

Dans le cas où f ne dépend pas de x, i.e. $f = f(y,p)$, l'équation d'Euler-Lagrange reste une équation différentielle du second ordre. Cependant dans ce cas on peut se ramener à une équation du 1er ordre.

Théorème 6.6. *Étant donnée $f(x,y,p)$ de classe C^1, on définit la fonction*

$$H(x,y,p) = p\frac{\partial f}{\partial p}(x,y,p) - f(x,y,p),$$

appelée fonction de Hamilton. Soit $y\in\mathcal{A}$ un point critique de F, alors pour tout $x\in[a,b]$ on a

$$\frac{\mathrm{d}}{\mathrm{d}x}H(x,y(x),y'(x)) = -\frac{\partial f}{\partial x}(x,y(x),y'(x)).$$

Si en particulier f ne dépend pas de x, i.e. $f = f(y,p)$ alors la fonction $H(x,y(x),y'(x))$ est constante.

Démonstration. Pour simplifier on suppose que toutes les fonctions sont de classe C^2. On calcule

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}x}f(x,y(x),y'(x))\\
&= \frac{\partial f}{\partial x}(x,y(x),y'(x)) + y'(x)\frac{\partial f}{\partial y}(x,y(x),y'(x)) + y''(x)\frac{\partial f}{\partial p}(x,y(x),y'(x))\\
&= \frac{\partial f}{\partial x}(x,y(x),y'(x)) + y'(x)\times\frac{\mathrm{d}}{\mathrm{d}x}\frac{\partial f}{\partial p}(x,y(x),y'(x)) + \frac{\mathrm{d}}{\mathrm{d}x}y'(x)\times\frac{\partial f}{\partial p}(x,y(x),y'(x))\\
&= \frac{\partial f}{\partial x}(x,y(x),y'(x)) + \frac{\mathrm{d}}{\mathrm{d}x}\left(y'(x)\frac{\partial f}{\partial p}(x,y(x),y'(x))\right),
\end{aligned}$$

d'où

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\underbrace{y'(x)\frac{\partial f}{\partial p}(x,y(x),y'(x)) - f(x,y(x),y'(x))}_{=H(x,y(x),y'(x))}\right) = -\frac{\partial f}{\partial x}(x,y(x),y'(x))$$

□

Exemple 6.2. *On reprend l'Exemple 6.1. On a* $f(x, y, p) = 2\pi y\sqrt{1 + p^2}$ *d'où* $H(x, y, p) = -2\pi\frac{y}{\sqrt{1+p^2}}$. *Comme* f *ne dépend pas de* x *on a*

$$\begin{aligned} H(x, y(x), y'(x)) = H_0 \text{ (constante)} \quad &\Longleftrightarrow \quad \frac{y(x)}{\sqrt{1 + y'(x)^2}} = \omega \\ &\Longleftrightarrow \quad y(x)^2 = \omega^2(1 + y'(x)^2), \end{aligned}$$

où ω *est une constante. On s'est ramené à une équation différentielle du 1er ordre. En la réécrivant sous la forme*

$$\frac{y'(x)}{\sqrt{\frac{y(x)^2}{\omega^2} - 1}} = 1,$$

on peut montrer que les solutions sont de la forme $y(x) = \omega \cosh\left(\frac{x - x_0}{\omega}\right)$. *Il faut ensuite trouver* x_0 *et* ω *pour satisfaire les conditions au bord* $y(a) = y_a$ *et* $y(b) = y_b$.