Topologie et calcul différentiel

Xavier BLANC,
Laboratoire Jacques-Louis Lions,
Université Paris Diderot

18 décembre 2013

Table des matières

Première partie

Topologie

Chapitre 1

Espaces vectoriels normés

Dans ce chapitre comme dans tout le cours, les espaces vectoriels sont des espaces sur le corps $\mathbb{K}$, qui désigne soit $\mathbb{R}$ soit $\mathbb{C}$.

1.1 Normes

Définition 1.1 *Soit E un espace vectoriel sur le corps $\mathbb{K}$. On appelle norme sur E une application $N : E \longrightarrow \mathbb{R}^+$ telle que*

(i) $N(x) = 0 \iff x = 0$;

(ii) $\forall x \in E$, $\forall \lambda \in \mathbb{K}$, $N(\lambda x) = |\lambda| N(x)$;

(iii) $\forall x \in E$, $\forall y \in E$, $N(x + y) \leq N(x) + N(y)$.

Pour désigner la propriété (ii), on dit que N est positivement homogène de degré 1. La propriété (iii) s'appelle l'inégalité triangulaire.

Définition 1.2 *Si E est un $\mathbb{K}$-espace vectoriel, et que N est une norme sur E, le couple (E, N) est appelé espace vectoriel normé.*

Usuellement, on note par défaut $\|\cdot\|$ la norme : $N(x) = \|x\|$. Ainsi, l'affirmation "soit E un espace vectoriel normé" est un abus de langage signifiant "soit $(E, \|\cdot\|)$ un espace vectoriel normé".

Définition 1.3 *Soit E un $\mathbb{K}$-espace vectoriel. On appelle semi-norme sur E une application $N : E \longrightarrow \mathbb{R}^+$ les propriétés (ii) et (iii) de la définition 1.1. Le couple (E, N) est alors appelé espace vectoriel semi-normé.*

Notons que, comme une semi-norme reste positivement homogène, on a toujours que $N(0) = 0$. Ainsi, seule l'implication "vers la gauche" de (i) reste vraie pour une semi-norme.

Remarque 1.4 (importante) *Si N est une semi-norme sur un espace vectoriel E, alors on a l'inégalité suivante :*

$$\forall x \in E, \quad \forall y \in E, \quad N(x - y) \geq |N(x) - N(y)|. \tag{1.1}$$

Démonstration. On a, d'après l'inégalité triangulaire, $N(x) = N(x-y+y) \leq N(x-y) + N(y)$. Ainsi,

$$N(x-y) \geq N(x) - N(y).$$

De même, en échangeant les rôles de x et y, on a $N(y) = N(y-x+x) \leq N(y-x) + N(x)$, donc

$$N(y-x) \geq N(y) - N(x).$$

Ainsi,

$$N(x-y) \geq \max(N(x) - N(y), N(y) - N(x)) = |N(x) - N(y)|.$$

car, pour tout réel a, $|a| = \max(a, -a)$. □

Citons quelques exemples classiques d'espaces vectoriels normés :

Exemple 1.5 *Soit n un entier naturel non nul. Sur $E = \mathbb{K}^n$, les applications suivantes sont des normes :*

1. $\|x\|_1 = \sum_{j=1}^{n} |x_j|$.
2. $\|x\|_2 = \left(\sum_{j=1}^{n} |x_j|^2\right)^{1/2}$. *Il s'agit de la norme euclidienne.*
3. $\|x\|_\infty = \sup_{1\leq i\leq n} |x_j|$.

Démonstration. Nous ne donnons la preuve que pour la norme $\|\cdot\|_2$. L'adaptation aux autres cas est facile.

Commençons pas démontrer la propriété (i), et supposons donc que $\|x\|_2 = 0$. Ceci implique que

$$\sum_{j=1}^{n} |x_j|^2 = 0.$$

Il s'agit d'une somme de nombres positifs. Ils sont donc tous nuls. Donc $x = 0$.

Démontrons ensuite l'homogénéité, c'est-à-dire la propriété (ii). Soit donc $x \in E$ et $\lambda \in \mathbb{K}$. On a alors :

$$\|\lambda x\|_2 = \left(\sum_{j=1}^{n} |\lambda x_j|^2\right)^{1/2} = \left(|\lambda|^2 \sum_{j=1}^{n} |x_j|^2\right)^{1/2} = |\lambda| \left(\sum_{j=1}^{n} |x_j|^2\right)^{1/2} = |\lambda| \|x\|_2.$$

Enfin, démontrons l'inégalité triangulaire. On commence pour cela par démontrer l'inégalité de Cauchy-Schwarz :

$$\forall x \in E, \quad \forall y \in E, \quad \left|\sum_{j=1}^{n} x_j \overline{y_j}\right| \leq \|x\|_2 \|y\|_2. \tag{1.2}$$

Pour cela, on remarque que, pour tout $t \in \mathbb{R}$, on a l'égalité suivante :

$$\|x+ty\|_2^2 = \sum_{j=1}^{n} \left[|x_j|^2 + 2t\mathrm{Re}\,(x_j\overline{y_j}) + t^2|y_j|^2\right] = \|x\|_2^2 + 2t\sum_{j=1}^{n}\mathrm{Re}\,(x_j\overline{y_j}) + t^2\|y\|_2^2.$$

Il est clair que le membre de gauche de l'inégalité est toujours positif. Le membre de droite, lui, est un polynôme de degré 2 en t. Son discriminant doit donc être positif. Ainsi,

$$4\left[\mathrm{Re}\left(\sum_{j=1}^{n} x_j\overline{y_j}\right)\right]^2 - 4\|x\|_2^2\|y\|_2^2 \leq 0,$$

D'où

$$\left|\mathrm{Re}\left(\sum_{j=1}^{n} x_j\overline{y_j}\right)\right| \leq \|x\|_2\|y\|_2.$$

Ceci vaut pour tout $x \in E$ et tout $y \in E$. On l'applique donc en changeant y en

$$y' = e^{i\theta}y, \quad \theta = \arg\left(\sum_{j=1}^{n} x_j\overline{y_j}\right), \quad \text{c'est-à-dire} \quad \sum_{j=1}^{n} x_j\overline{y_j} = \left|\sum_{j=1}^{n} x_j\overline{y_j}\right| e^{i\theta}.$$

On obtient donc

$$\left|\mathrm{Re}\left(\sum_{j=1}^{n} x_j\overline{y'_j}\right)\right| \leq \|x\|_2\|y'\|_2 = \|x\|_2\|y\|_2,$$

et d'autre part,

$$\sum_{j=1}^{n} x_j\overline{y'_j} = \sum_{j=1}^{n} x_j\overline{y_j}e^{-i\theta} = \left|\sum_{j=1}^{n} x_j\overline{y_j}\right|.$$

On obtient donc bien

$$\left|\sum_{j=1}^{n} x_j\overline{y_j}\right| \leq \|x\|_2\|y\|_2,$$

c'est-à-dire (1.2). Muni de cette inégalité, nous pouvons maintenant prouver l'inégalité triangulaire :

$$\begin{aligned}
\|x+y\|_2^2 = \sum_{j=1}^{n} |x_j+y_j|^2 &= \sum_{j=1}^{n} |x_j|^2 + 2\mathrm{Re}(x_j\overline{y_j}) + |y_j|^2 \\
&= \|x\|_2^2 + 2\mathrm{Re}\left(\sum_{j=1}^{n} x_j\overline{y_j}\right) + \|y\|_2^2 \\
&\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2,
\end{aligned}$$

où, pour passer de la deuxième à la troisième ligne, on a utilisé l'inégalité de Cauchy-Schwarz. □

Exemple 1.6 *Sur $E = \mathbb{K}[X]$ espace des polynômes à coefficients dans $\mathbb{K}$, la quantité*

$$\|P\|_1 = \int_0^1 |P(t)|dt,$$

est une norme.

Démonstration. On rappelle que si une fonction continue positive est d'intégrale nulle sur $[0,1]$ alors cette fonction est nulle sur $[0,1]$. Ainsi, un polynôme tel que $\|P\|_1 = 0$ est nul sur l'intervalle $[0,1]$, car il s'agit d'une fonction continue. Ceci implique donc que tous les réels de $[0,1]$ sont racines de P. Ce polynôme a donc une infinité de racines, et ne peut être que le polynôme nul. Ceci prouve le point (i). Pour l'homogénéité, on suppose que $P \in \mathbb{K}[X]$ et que $\lambda \in \mathbb{K}$. On a alors

$$\|\lambda P\|_1 = \int_0^1 |\lambda||P(t)|dt = |\lambda| \int_0^1 |P(t)|dt = |\lambda|\|P\|_1.$$

Ceci prouve (ii). Enfin, pour l'inégalité triangulaire, on commence par constater que, si P et Q sont dans E, alors, pour tout $t \in [0,1]$, on a $|P(t) + Q(t)| \leq |P(t)| + |Q(t)|$. On intègre cette inégalité sur $[0,1]$, et on obtient alors

$$\begin{aligned}\|P+Q\|_1 &= \int_0^1 |P(t)+Q(t)|dt \\ &\leq \int_0^1 (|P(t)| + |Q(t)|)\, dt = \int_0^1 |P(t)|dt + \int_0^1 |Q(t)|dt = \|P\|_1 + \|Q\|_1.\end{aligned}$$

□

Exemple 1.7 *On suppose que $a < b$ sont deux réels distincts, et on pose $E = C^0([a,b],\mathbb{K})$ l'espace des fonctions continues bornées de $[a,b]$ dans $\mathbb{K}$. Sur cet espace vectoriel, la quantité*

$$\|f\|_\infty = \sup_{a \leq x \leq b} |f(x)|,$$

est une norme.

La preuve de ce fait est élémentaire, et laissée en exercice.

Exemple 1.8 (exemples de semi-normes) *Donnons quelques exemples de semi-normes qui ne sont pas des normes (les preuves sont laissées en exercice) :*

1. *Sur $E = \mathbb{K}^n$, $N(x) = \sum_{i=1}^{n-1} |x_i|$.*
2. *Sur $C^0([a,b],\mathbb{K})$ (cf exemple 1.7), $N(x) = \sup_{a \leq x,y \leq b} |f(x) - f(y)|$.*
3. *Sur $C^1([a,b],\mathbb{K})$, l'ensemble des fonctions dérivables sur $[a,b]$, à dérivées continues, $N(f) = \int_a^b |f'(t)|dt$.*

Proposition 1.9 *Soit E un espace vectoriel et soient $N_1, N_2, \dots, N_p$ des normes sur E. Alors*

$$N(x) = \sup_{1 \le i \le p} N_i(x),$$

et

$$\tilde{N}(x) = \sum_{i=1}^{p} N_i(x),$$

sont des normes sur E.

Démonstration. On commence par le point (ii) de la défintion 1.1. Supposons donc que $x \in E$ et $\lambda \in \mathbb{K}$. Alors on calcule :

$$N(\lambda x) = \sup_{1 \le i \le p} N_i(\lambda x) = \sup_{1 \le i \le p} |\lambda| N_i(x) = |\lambda| \sup_{1 \le i \le p} N_i(x) = |\lambda| N(x).$$

De même,

$$\tilde{N}(\lambda x) = \sum_{i=1}^{p} N_i(\lambda x) = \sum_{i=1}^{p} |\lambda| N_i(x) = |\lambda| \sum_{i=1}^{p} N_i(\lambda x) = |\lambda| \tilde{N}(x).$$

En ce qui concerne le point (iii), on suppose que $x \in E$ et $y \in E$, et on a $N_i(x+y) \le N_i(x) + N_i(y)$, pour tout i entre 1 et p. Donc, d'une part,

$$\begin{aligned} N(x+y) = \sup_{1 \le i \le p} N_i(x+y) &\le \sup_{1 \le i \le p} (N_i(x) + N_i(y)) \\ &\le \sup_{1 \le i \le p} N_i(x) + \sup_{1 \le i \le p} N_i(y) = N(x) + N(y), \end{aligned}$$

et d'autre part,

$$\begin{aligned} \tilde{N}(x+y) &= \sum_{i=1}^{p} N_i(x+y) \\ &\le \sum_{i=1}^{p} (N_i(x) + N_i(y)) = \sum_{i=1}^{p} N_i(x) + \sum_{i=1}^{p} N_i(y) = \tilde{N}(x) + \tilde{N}(y). \end{aligned}$$

Enfin, si $N(x) = 0$, on a que pour tout i, $N_i(x) = 0$, donc $x = 0$. De même pour $\tilde{N}$. □

La proposition 1.9 reste également vraie pour des semi-normes :

Proposition 1.10 *Soit E un espace vectoriel et soient $N_1, N_2, \dots, N_p$ des semi-normes sur E. Alors*

$$N(x) = \sup_{1 \le i \le p} N_i(x),$$

et

$$N(x) = \sum_{i=1}^{p} N_i(x),$$

sont des semi-normes sur E.

La preuve de ce résultat est laissée en exercice : il suffit de reproduire celle de la proposition 1.9.

Exercice 1.1 *Soit E un espace vectoriel et soient $N_1, N_2, \ldots, N_p$ des semi-normes sur E. Démontrer que la quantité*

$$N(x) = \left(\sum_{i=1}^{p} N_i(x)^2 \right)^{1/2}$$

est une semi-norme sur E. Même question dans le cas de normes.

1.2 Boules, ensembles ouverts et fermés

Définition 1.11 *Soit E un espace vectoriel normé, soit $x \in E$ et $r \in \mathbb{R}^{+*}$. On appelle :*

- *boule ouverte de centre x et de rayon r l'ensemble*

$$B_r(x) = \{y \in E, \quad \|x - y\| < r\},$$

- *boule fermée de centre x et de rayon r l'ensemble*

$$\overline{B_r}(x) = \{y \in E, \quad \|x - y\| \leq r\},$$

- *sphère de centre x et de rayon r l'ensemble*

$$S_r(x) = \{y \in E, \quad \|x - y\| = r\},$$

Exemple 1.12 *Dans $\mathbb{R}$, muni de la norme $N(x) = |x|$, la boule ouverte de centre x et de rayon r est l'intervalle $B_r(x) =]x - r, x + r[$. La boule fermée ce centre x et de rayon r est l'intervalle $\overline{B_r}(x) = [x - r, x + r]$.*

Définition 1.13 *Soit E un espace vectoriel normé, et soit $A \subset E$. On dit que*

- *A est ouvert si $\forall x \in A, \quad \exists r > 0, \quad B_r(x) \subset A$,*
- *A est fermé si son complémentaire $A^C = E \setminus A$ est ouvert.*

En particulier, $\emptyset$ est un ouvert.

Rappelons la définition du complémentaire, utilisée ci-dessus :

$$A^C = E \setminus A = \{x \in E \text{ tels que } x \notin A\}.$$

Notons que "A est fermé" ne signifie pas "A n'est pas ouvert". D'ailleurs, un ensemble peut très bien être ni ouvert ni fermé.

Exemple 1.14 *Soient $a < b$ deux réels. Alors, dans $\mathbb{R}$ muni de la norme valeur absolue,*

1. *l'intervalle $]a, b[$ est ouvert ;*

2. *l'intervalle* $[a,b]$ *est fermé.*

Démonstration. Soit $x \in]a,b[$. On pose alors $r = \min(|x-a|, |x-b|)$. Ce réel est strictement positif car $a < x < b$, par définition. Prouvons que $B_r(x) \subset]a,b[$: soit $y \in B_r(x)$. Alors $|x-y| < r \leq |x-a|$ d'une part, et $|x-y| < r \leq |x-b|$ d'autre part. Ainsi,

$$y = y - x + x \leq |y-x| + x < |x-b| + x = b - x + x = b,$$

car $x < b$. De même,

$$y = y - x + x \geq -|y-x| + x > -|x-a| + x = -(x-a) + x = a,$$

car $x > a$. Donc $y \in]a,b[$. On a donc prouvé que, pour tout $y \in]a,b[$, il existe $r > 0$ tel que $B_r(x) \subset]a,b[$. L'intervalle $]a,b[$ est donc bien un ouvert.

Démontrons maintenant que $[a,b]$ est un fermé. On veut donc prouver que A^C est un ouvert. Notons que

$$A^C =]-\infty, a[\cup]b, \infty[.$$

Soit dont $x \in A^C$. Alors $x < a$ ou bien $x > b$. Dans le premier cas, on pose $r = a - x > 0$, et on a

$$B_r(x) =]x-r, x+r[=]2x-a, a[\subset]-\infty, a[\subset A^C.$$

Si maintenant $x > b$, on pose $r = x - b > 0$, et on a

$$B_r(x) =]x-r, x+r[=]b, 2x-b[\subset]b, +\infty[\subset A^C.$$

Ainsi, A^C est ouvert, donc A est fermé. □

Remarque 1.15 *Dans la démonstration ci-dessus, on a en fait prouvé que les intervalles* $]-\infty, a[$ *et* $]b, +\infty[$ *sont des ouverts.*

Proposition 1.16 *Soit* E *un espace vectoriel normé. Alors toute boule ouverte est un ouvert.*

Démonstration. Soit $B_r(x)$ une boule ouverte. Soit $y \in B_r(x)$: $\|x-y\| < r$. On pose $r' = r - \|x-y\| > 0$, et on va montrer que $B_{r'}(y) \subset B_r(x)$, ce qui prouvera que $B_r(x)$ est un ouvert. Soit donc $z \in B_{r'}(y)$. On a alors $\|z-y\| < r'$. Donc

$$\|z-x\| = \|z-y+y-x\| \leq \|z-y\| + \|y-x\| < r' + \|y-x\| = r,$$

par définition de r'. Ainsi, $z \in B_r(x)$, pour tout $z \in B_{r'}(y)$, donc $B_{r'}(y) \subset B_r(x)$. □

Proposition 1.17 *Soit* E *un espace vectoriel normé. Alors toute boule fermée est un fermé.*

Démonstration. Soit $\overline{B_r}(x)$ une boule fermée. Pour montrer qu'il s'agit d'un fermé, nous devons prouver que son complémentaire A est ouvert, avec

$$A = \{y \in E, \quad \|x - y\| > r\}.$$

Soit donc $y \in A$. On pose alors $r' = \|x - y\| - r > 0$, et on va prouver que $B_{r'}(y) \subset A$, ce qui prouvera que A est ouvert, et terminera la démonstration. Soit donc $z \in B_{r'}(y)$. Alors $\|z - y\| < r'$. Donc

$$\|z-x\| = \|z-y+y-x\| \geq \|y-x\|-\|z-y\| > \|x-y\|-r' = \|x-y\|-\|x-y\|+r = r.$$

Ainsi, $z \in A$, pour tout $z \in B_{r'}(y)$, donc $B_{r'}(y) \subset A$. □

Définition 1.18 *Soit E un espace vectoriel normé, $x \in E$ et $A \subset E$. On dit que A est un voisinage de x s'il contient un ouvert contenant x.*

Donnons quelques exemples de voisinages de 0 dans $\mathbb{R}$ muni de la norme valeur absolue :
- $]-1, 1[$ est un voisinage de 0, car il est lui-même ouvert.
- $]-1, 1]$ est un voisinage de 0, car il contient $]-1, 1/2[$.
- $[-1, 1[$ est un voisinage de 0, car il contient $]-1/2, 1[$.
- $[0, 1[$ n'est pas un voisinage de 0. En effet, s'il était voisinage de 0, il contiendrait un ouvert contenant 0, donc une boule ouverte contenant 0. Il existerait donc $r > 0$ tel que $]-r, r[\subset [0, 1[$. Ce qui est contradictoire.

Plus généralement, pour tout $\delta > 0$ et tout $\varepsilon > 0$, l'intervalle $[-\delta, \varepsilon]$ est un voisinage de 0.

Il est clair, d'après la définition 1.18 et la définition d'un ouvert, qu'on a

Proposition 1.19 *Soit E un espace vectoriel normé, $x \in E$ et $A \subset E$. A est un voisinage de x si et seulement s'il existe $r > 0$ tel que $B_r(x) \subset A$.*

Démonstration. Si A est voisinage de x, alors A contient un ouvert B qui contient x. Comme $x \in B$ ouvert, il existe $r > 0$ tel que $B_r(x) \subset B \subset A$. Réciproquement, si A contient $B_r(x)$ pour un certain $r > 0$, comme $B_r(x)$ est ouvert, A est voisinage de x. □

Proposition 1.20 *Soit E un espace vectoriel normé. Dans E, on a alors les propriétés suivantes :*
(i) Une réunion d'ouverts est un ouvert.
(ii) Une intersection finie d'ouverts est un ouvert.
(iii) $\emptyset$ et E sont ouverts et fermés.

Démonstration. *(i)* Soient des ouverts $(O_i)_{i \in I}$, et soit $O = \bigcup_{i \in I} O_i$. Soit $x \in O$. Alors, par définition de l'union, il existe $i_0 \in I$ tel que $x \in O_{i_0}$. Comme ce dernier ensemble est ouvert, on sait qu'il existe $r > 0$ tel que $B_r(x) \subset O_{i_0}$. Ainsi,

$$B_r(x) \subset O_{i_0} \subset O,$$

ce qui prouve que O est un ouvert.
(ii) Soient des ouverts $(O_i)_{1\leq i\leq p}$, et soit $O = \bigcap_{i=1}^{p} O_i$. Pour tout $x \in O$, on sait que $x \in O_i$, donc il existe $r_i > 0$ tel que $B_{r_i}(x) \subset O_i$. On pose donc $r = \min_{1\leq i\leq p} r_i$, et on a alors

$$\forall i, \quad B_r(x) \subset B_{r_i}(x) \subset O_i.$$

Donc $B_r(x) \subset O$. Ainsi, O est bien un ouvert.
(iii) $\emptyset$ est ouvert par définition. Donc E est fermé. De plus, pour tout $x \in E$, on a $B_1(x) \subset E$ par définition. Donc E est ouvert. Donc $\emptyset$ est fermé. □

Remarque 1.21 *Une intersection infinie d'ouvert peut ne pas être un ouvert. Considérons par exemple, dans $\mathbb{R}$, les ouverts*

$$U_n = \left]-\frac{1}{n}, \frac{1}{n}\right[, \quad n \in \mathbb{N}^*.$$

Alors il est clair que

$$U = \bigcap_{n\in\mathbb{N}^*} U_n = \{0\},$$

qui n'est pas un ouvert.

Démonstration. Comme, pour tout $n \geq 1$, $0 \in U_n$, on a bien $0 \in U$. De plus, si $x \in U$, alors, pour tout $n \in \mathbb{N}^*$, on $-\frac{1}{n} < x < \frac{1}{n}$. En prenant la limite $n \to \infty$ de cette inégalité, on obtient $0 \leq x \leq 0$, donc $x = 0$. Ainsi, $U = \{0\}$. Si maintenant cet ensemble était ouvert, il existerait $r > 0$ tel que $]-r, r[\subset O = \{0\}$, ce qui n'est pas le cas. □

Une conséquence directe de la proposition 1.20 est le résultat suivant :

Corollaire 1.22 *Soit E un espace vectoriel normé. Dans E, on a alors les propriétés suivantes :*
(i) Une réunion finie de fermés est un fermé.
(ii) Une intersection de fermés est un fermé.

Proposition 1.23 *Soit E un espace vectoriel normé. Soit $A \subset E$. Alors A est ouvert si et seulement s'il est voisinage de chacun de ses points.*

Démonstration. Si A est ouvert, il est clair qu'il est voisinage de chacun de ses points. Réciproquement, si A est voisinage de chaque $x \in A$, alors il existe un ouvert U tel que $x \in U \subset A$. Comme U est ouvert, il existe $r > 0$ tel que $B_r(x) \subset U$, donc $B_r(x) \subset A$. Ceci valant pour tout $x \in A$, A est ouvert. □

1.3 Continuité

Définition 1.24 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés, $f : E \longrightarrow F$, et $x_0 \in E$. On dit que f est continue en x_0 si*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 \quad \textit{tel que} \quad \|x_0 - x\|_E < \delta \Rightarrow \|f(x_0) - f(x)\|_F < \varepsilon.$$

Définition 1.25 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés, et $f : E \longrightarrow F$. On dit que f est continue sur E (ou tout simplement continue) si elle est continue en tout point de E.*

Avant de donner la proposition suivante, rappelons la définition de l'image réciproque d'un ensemble par une application : pour toute application f d'un ensemble E dans un ensemble F, et pour tout $A \subset F$, on note

$$f^{-1}(A) = \{x \in E, \quad \text{tel que} \quad f(x) \in A\}.$$

L'ensemble $f^{-1}(A)$ est appelé l'image réciproque de A par f. Contrairement à ce que la notation suggère, cette définition ne nécessite pas que l'application réciproque de f existe.

Proposition 1.26 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés, et $f : E \longrightarrow F$. Alors on a équivalence entre :*

(i) f est continue sur E.

(ii) $\forall A$ ouvert de F, $f^{-1}(A)$ est un ouvert de E.

(iii) $\forall A$ fermé de F, $f^{-1}(A)$ est un fermé de E.

(iv) $\forall x \in E$, $\forall V$ voisinage de $f(x)$, $f^{-1}(V)$ est voisinage de x.

Démonstration. Nous allons démontrer que $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$, ce qui prouvera l'équivalence.
$(i) \Rightarrow (ii)$: soit $A \subset F$, A ouvert. On pose $B = f^{-1}(A)$. Pour montrer que B est ouvert, on prend $x_0 \in B$, et on veut démontrer qu'une boule ouverte centrée en x_0 est incluse dans B. Comme A est ouvert et que $f(x_0) \in A$, on sait, par définition, qu'il existe $\varepsilon > 0$ tel que $B_\varepsilon(f(x_0)) \subset A$. De plus, par définition de la continuité, on sait qu'il existe $\delta > 0$ tel que $\|x_0 - x\|_E < \delta \Rightarrow \|f(x_0) - f(x)\|_F < \varepsilon$. Ainsi, si $x \in B_\delta(x_0)$, on a $f(x) \in B_\varepsilon(f(x_0)) \subset A$. Donc $B_\delta(x_0) \subset B$, et B est donc bien un ouvert.
$(ii) \Rightarrow (iii)$: supposons que B est fermé. On a alors

$$\begin{aligned}\left(f^{-1}(B)\right)^C = \{x \in E, \quad / \quad f(x) \in B\}^C &= \{x \in E, \quad / \quad f(x) \notin B\} \\ &= \left\{x \in E, \quad / \quad f(x) \in B^C\right\} = f^{-1}(B^C). \quad (1.3)\end{aligned}$$

Par hypothèse, B^C est ouvert, donc $\left(f^{-1}(B)\right)^C$ aussi. Donc $f^{-1}(B)$ est fermé.
$(iii) \Rightarrow (iv)$: soit $x \in E$, et soit V un voisinage de $f(x)$ dans F. Alors, par définition d'un voisinage, il existe $\varepsilon > 0$ tel que $B_\varepsilon(f(x)) \subset V$. De plus, $B = B_\varepsilon(f(x))^C$ est fermé, donc, par hypothèse, $f^{-1}(B)$ l'est aussi, d'où $\left(f^{-1}(B)\right)^C$ est un ouvert. Si on reproduit le calcul (1.3) ci-dessus, on obtient aussi que $f^{-1}\left(B^C\right) = \left(f^{-1}(B)\right)^C$ est un ouvert, et donc que

$$f^{-1}\left(B_\varepsilon(f(x))\right) \text{ est ouvert.}$$

Comme de plus $x \in f^{-1}\left(B_\varepsilon(f(x))\right) \subset f^{-1}(V)$, ceci prouve bien que $f^{-1}(V)$ est un voisinage de x.

$(iv) \Rightarrow (i)$: soit $x_0 \in E$, et soit $\varepsilon > 0$. Par définition, $B_\varepsilon(f(x_0))$ est un voisinage de $f(x_0)$. Donc $f^{-1}(B_\varepsilon(f(x_0)))$ est un voisinage de x_0. Donc, toujours par définition d'un voisinage, il existe $\delta > 0$ tel que $B_\delta(x_0) \subset f^{-1}(B_\varepsilon(f(x_0)))$. Autrement dit,

$$\|x - x_0\|_E < \delta \Rightarrow \|f(x) - f(x_0)\|_F < \varepsilon.$$

Ceci termine la démonstration. □

Finissons cette section par deux notions liées à la continuité :

Définition 1.27 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés. Soit $f : E \longrightarrow F$. On dit que f est uniformément continue sur E si*

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \textit{tel que} \quad \|x - y\|_E < \delta \Rightarrow \|f(x) - f(y)\|_F < \varepsilon.$$

Pour $A \subset E$, on dira que f est uniformément continue sur A si cette propriété est vraie pour $x \in A$, $y \in A$.

Il est clair qu'une application uniformément continue sur E est continue. Mais remarquons bien la différence avec la notion de continuité : δ ici ne dépend pas du point x où on examine la continuité. En particulier, une application peut être continue sans être uniformément continue. Ainsi, la fonction $f(x) = x^2$ de $\mathbb{R}$ dans $\mathbb{R}$ est continue, mais n'est pas uniformément. En effet, si elle l'était, on aurait,

$$\forall \varepsilon > 0, \ \exists \delta > 0, \quad |x - y| < \delta \Rightarrow \left|x^2 - y^2\right| < \varepsilon.$$

En particulier, en choisissant $y = x + \delta/2$, on a $x^2 - y^2 = \delta x + \delta^2/4$, ce qui contredit l'inégalité ci-dessus dès que $x \geq \varepsilon/\delta - \delta/4$.

Définition 1.28 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés, Soit $f : E \longrightarrow F$. On dit que f est lipschitzienne s'il existe $K \geq 0$ tel que*

$$\forall x \in E, \quad \forall y \in E, \quad \|f(x) - f(y)\|_F \leq K\|x - y\|_E.$$

Pour $A \subset E$, on dira que f est lipschitzienne sur A si cette propriété est vraie pour $x \in A$, $y \in A$.

Notons qu'une application lipschitzienne est toujours uniformément continue (le démontrer en exercice). Il existe en revanche des applications qui sont uniformément continues mais pas lipschitziennes. Par exemple, l'application $f(x) = \sqrt{|x|}$ de $\mathbb{R}$ dans $\mathbb{R}$. En effet, on a

$$f(x) - f(y) = \sqrt{|x|} - \sqrt{|y|} = \frac{|x| - |y|}{\sqrt{|x|} + \sqrt{|y|}}.$$

Ainsi,

$$|f(x) - f(y)| = \frac{\big||x| - |y|\big|}{\sqrt{|x|} + \sqrt{|y|}} \leq \sqrt{\big||x| - |y|\big|}\,\frac{\sqrt{\big||x| - |y|\big|}}{\sqrt{|x|} + \sqrt{|y|}} \leq \sqrt{\big||x| - |y|\big|}.$$

Ainsi, $\delta = \varepsilon^2$ convient dans la définition de la continuité uniforme. Si maintenant f était lipschitzienne, on aurait, en prenant $y = 0$ dans la définition,

$$\forall x \in \mathbb{R}, \quad \sqrt{|x|} \leq K|x|,$$

d'où $1 \leq K\sqrt{|x|}$, pour tout $x \neq 0$. Ceci est contradictoire (prendre par exemple $x = 1/(4K^2)$.)

1.4 Continuité : le cas des applications linéaires

Dans le cas où l'application f est linéaire, la propriété de continuité se caractérise de façon beaucoup plus simple, comme nous allons le voir.

Rappelons tout d'abord la définition d'une application linéaire :

Définition 1.29 *Soient E et F deux espaces vectoriels. Soit f une application de E dans F. On dit que f est linéaire si*

$$\forall x \in E,\ \forall y \in E,\ \forall \lambda \in \mathbb{K},\ \forall \mu \in \mathbb{K}, \quad f(\lambda x + \mu y) = \lambda f(x) + \mu f(y).$$

Notons, que, dans la définition ci-dessus, on peut se restreindre à $\lambda = 1$ (le démontrer en exercice). Bien sûr, la notion d'application linéaire est uniquement liée à la structure d'espace vectoriel, et il n'y a donc aucun besoin que les espaces vectoriels soient normés pour la définir.

Définition 1.30 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires continues de E dans F.*

Proposition 1.31 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés. Soit f une application linéaire de E dans F. Alors les propriétés suivantes sont équivalentes :*

(i) f est continue en 0.
(ii) f est continue.
(iii) f est uniformément continue sur E.
(iv) f est lipschitzienne sur E.
(v) f est bornée sur la boule unité fermée.
(vi) f est bornée sur la sphère unité.
(vii) $\exists M > 0$ tel que $\forall x \in E$, $\|f(x)\|_F \leq M\|x\|_E$.

Démonstration. Il est évident que $(iv) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i)$. On démontre maintenant les implications manquantes :
$(i) \Rightarrow (vii)$: comme f est continue en 0, on applique la définition de la continuité pour $\varepsilon = 1$. On sait donc qu'il existe $\delta > 0$ tel que

$$\|x\|_E < \delta \ \Rightarrow\ \|f(x)\|_F < 1.$$

Soit maintenant $y \in E$ quelconque. Si $y = 0$, alors $f(y) = 0$, et on a évidemment $f(y)_F \leq M\|y\|_E$ pour tout $M > 0$. Si maintenant $y \neq 0$, posons $x = \delta\frac{y}{2\|y\|_E}$. Alors $\|x\|_E = \delta/2 < \delta$, donc $\|f(x)\|_F < 1$. Par linéarité de f, on a donc

$$\frac{\delta}{2\|y\|_E}\|f(y)\|_F < 1, \text{ donc } \|f(y)\|_F < \frac{2}{\delta}\|y\|_E.$$

Ceci valant pour tout $y \in E$, on a donc prouvé (vii) avec $M = 2/\delta$.
$(vii) \Rightarrow (vi)$ est évident.
$(vi) \Rightarrow (v)$: supposons que f est bornée par une constante C sur la sphère unité. Soit $x \in B_1(0)$ la boule unité. Si $x = 0$, alors $f(x) = 0$ par linéarité de f. Si $x \neq 0$, alors $\frac{x}{\|x\|_E}$ est sur la sphère unité. Donc

$$\left\| f\left(\frac{x}{\|x\|_E}\right)\right\|_F \leq C,$$

d'où, par linéarité de f, $\|f(x)\|_F \leq C\|x\|_E \leq C$, puisque x est dans la boule unité.
$(v) \Rightarrow (iv)$: soient deux vecteurs x et y de E. Si $x = y$, on a évidemment $\|f(x) - f(y)\|_F = 0 \leq M\|x - y\|_E$, pour toute valeur M réelle positive. Si $x \neq y$, alors on a, par linéarité de f,

$$\begin{aligned}\|f(x) - f(y)\|_F = \|f(x-y)\|_F &= \left\| \|x-y\|_E f\left(\frac{x-y}{\|x-y\|_E}\right)\right\|_F \\ &= \|x-y\|_E \left\| f\left(\frac{x-y}{\|x-y\|_E}\right)\right\|_F \leq M\|x-y\|_E,\end{aligned}$$

où M est un majorant de f sur la sphère unité. □

Corollaire 1.32 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés. On suppose que F n'est pas réduit à $\{0\}$. Soit f une application linéaire de E dans F. Alors*

$$f \text{ continue } \iff \sup_{x\neq 0}\frac{\|f(x)\|_F}{\|x\|_E} < +\infty \iff \sup_{\|x\|_E=1}\|f(x)\|_F < +\infty.$$

Définition 1.33 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés. On suppose que E n'est pas réduit à $\{0\}$. Soit f une application linéaire continue de E dans F. On appelle norme de f la quantité*

$$\|f\|_{\mathcal{L}(E,F)} = \sup_{x\neq 0}\frac{\|f(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E=1}\|f(x)\|_F.$$

Démonstration. Prouvons d'abord l'égalité. Comme f est continue, on sait, d'après le corollaire 1.32 que ces deux bornes supérieures sont finies. De plus, on a

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

D'autre part, si $x \neq 0$, on a

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

puisque $x/\|x\|_E$ est de norme 1. Ceci valant pour tout $x \neq 0$, on en déduit

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

ce qui prouve l'égalité.

Démontrons maintenant que $\|\cdot\|_{\mathcal{L}(E,F)}$ est bien une norme :

Tout d'abord, si $\|f\|_{\mathcal{L}(E,F)} = 0$, alors pour tout $x \neq 0$, $f(x) = 0$. Comme de plus $f(0) = 0$, on en déduit que f est l'application nulle.

Démontrons maintenant l'homogénéité : soit $f \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{R}$. Alors

$$\begin{aligned}\|\lambda f\|_{\mathcal{L}(E,F)} = \sup_{x\neq 0} \frac{\|\lambda f(x)\|_F}{\|x\|_E} &= \sup_{x\neq 0} |\lambda| \frac{\|f(x)\|_F}{\|x\|_E} \\ &= |\lambda| \sup_{x\neq 0} \frac{\|f(x)\|_F}{\|x\|_E} = |\lambda| \|f\|_{\mathcal{L}(E,F)}.\end{aligned}$$

Enfin, démontrons l'inégalité triangulaire : pour tout f et tout g dans $\mathcal{L}(E,F)$,

$$\begin{aligned}\|f+g\|_{\mathcal{L}(E,F)} &= \sup_{\|x\|_E=1} \|f(x)+g(x)\|_F \leq \sup_{\|x\|_E=1} (\|f(x)\|_F + \|g(x)\|_F) \\ &\leq \sup_{\|x\|_E=1} \|f(x)\|_F + \sup_{\|x\|_E=1} \|g(x)\|_F = \|f\|_{\mathcal{L}(E,F)} + \|g\|_{\mathcal{L}(E,F)}\end{aligned}$$

Ainsi, $\|\cdot\|_{\mathcal{L}(E,F)}$ est bien une norme. □

Proposition 1.34 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels normés. Soit f une application linéaire continue de E dans F. Alors*

$$\forall x \in E, \quad \|f(x)\|_F \leq \|f\|_{\mathcal{L}(E,F)} \|x\|_E.$$

La preuve (très facile) de ce résultat est laissée en exercice. On peut également démontrer que, parmi tous les M vérifiant

$$\forall x \in E, \quad \|f(x)\|_F \leq M \|x\|_E,$$

$\|f\|_{\mathcal{L}(E,F)}$ est le plus petit possible.

Proposition 1.35 *Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ des espaces vectoriels normés. Supposons que $f \in \mathcal{L}(E,F)$ et $g \in \mathcal{L}(F,G)$. Alors $g \circ f$ est continue, et*

$$\|g \circ f\|_{\mathcal{L}(E,G)} \leq \|f\|_{\mathcal{L}(E,F)} \|g\|_{\mathcal{L}(F,G)}.$$

Démonstration. On a :

$$\|(g \circ f)(x)\|_G = \|g(f(x))\|_G \leq \|g\|_{\mathcal{L}(F,G)}\|f(x)\|_F \leq \|g\|_{\mathcal{L}(F,G)}\|f\|_{\mathcal{L}(E,F)}\|x\|_E.$$

Ce qui prouve le résultat. □

Une caractérisation similaire est possible pour les applications bilinéaires. Commençons par donner leur définition.

Définition 1.36 *Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ des espaces vectoriels. Soit f une application de $E \times F$ dans G. On dit que f est bilinéaire si*

$$\forall u \in E,\ \forall v \in E,\ \forall \lambda \in \mathbb{K},\ \forall w \in F, \quad f(u + \lambda v, w) = f(u, w) + \lambda f(v, w),$$

et

$$\forall u \in E,\ \forall w \in F,\ \forall \lambda \in \mathbb{K},\ \forall x \in F, \quad f(u, w + \lambda x) = f(u, w) + \lambda f(u, x),$$

Cette définition équivaut à dire que f est linéaire par rapport à sa première variable quand la deuxième est fixée, et linéaire par rapport à sa deuxième variable quand sa première est fixée. C'est (très) différent de dire que f est une application linéaire de $E \times F$ dans G.

Exemple 1.37 *Le plus simple exemple d'application bilinéaire est l'application f de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x, y) = axy$, où a est un réel fixé. On peut d'ailleurs démontrer (le faire en exercice) que toute application bilinéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ est de cette forme.*

Définition 1.38 *Soient $(E, \|\cdot\|_E)$ et $(F, \|\cdot\|_F)$ des espaces vectoriels. On appelle forme bilinéaire une application bilinéaire de $E \times F$ dans $\mathbb{K}$.*

Exemple 1.39 *Un exemple de forme bilinéaire dans le cas $E = F = \mathbb{R}^n$ est le produit scalaire : l'application $B : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par*

$$f(x, y) = \sum_{i=1}^{n} x_i y_i$$

est une forme bilinéaire.

Comme dans le cas des applications linéaires, il existe une caractérisation simples des applications bilinéaires continues.

Proposition 1.40 *Soient* $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ *et* $(G, \|\cdot\|_G)$ *des espaces vectoriels normés. On munit* $E \times F$ *de la norme*

$$\forall (u, v) \in E \times F, \quad \|(u, v)\|_{E\times F} = \|u\|_E + \|v\|_F.$$

Supposons que $f : E \times F \longrightarrow G$ *est une application bilinéaire. Alors on a équivalence entre*

(i) f *est continue en* $(0, 0)$.

(ii) f *est continue.*

(iii) f *est bornée sur le produit des boules unités* $B_E \times B_F$ *de* E *et* F.

(iv) f *est bornée sur le produit des sphères unités* $S_E \times S_F$ *de* E *et* F.

(v) $\exists M > 0$, $\forall u \in E$, $\forall v \in F$, $\|f(u, v)\|_G \leq M\|u\|_E\|v\|_F$.

Démonstration. On calque la preuve sur celle de la proposition 1.31. Il est clair, ici encore, que $(ii) \Rightarrow (i)$. On prouve les implications manquantes :
$(i) \Rightarrow (v)$: comme f est continue en $(0, 0)$, on a, en application la définition de la continuité avec $\varepsilon = 1$, l'existence de $\delta > 0$ tel que

$$\|u\|_E + \|v\|_F < \delta \;\Rightarrow\; \|f(u, v)\|_G < 1.$$

Soit maintenant $(x, y) \in E \times F$. Si $x = 0$ ou $y = 0$, alors on a bien entendu $f(x, y) = 0$, et donc $\|f(x, y)\|_G \leq M\,(\|x\|_E\|y\|_F)$ pour tout $M > 0$. Si maintenant $x \neq 0$ et $y \neq 0$, alors $\|x\|_E \neq 0$ et $\|y\|_F \neq 0$, et on pose $(u, v) = \left(\delta\frac{x}{4\|x\|_E}, \delta\frac{y}{4\|y\|_F}\right) \in E \times F$. Il est clair que $\|u\|_E < \delta/2$, et que $\|v\|_F < \delta/2$. En particulier, $\|u\|_E + \|v\|_F < \delta$. Donc, en appliquant l'inégalité ci-dessus, on obtient $\|f(u, v)\|_G < 1$. Donc, par bilinéarité de f, on obtient

$$\|f(x, y)\|_G < \frac{16}{\delta^2}\|x\|_E\|y\|_F.$$

$(v) \Rightarrow (iv)$ est évident.
$(iv) \Rightarrow (iii)$: supposons que f est bornée par une constante C sur le produit $S_E \times S_F$ des sphères unité. Soit $(x, y) \in B_E \times B_F$, le produit des sphères unité. Si $\|x\|_E = 0$ ou $\|y\|_F = 0$ alors $x = 0$ ou $y = 0$, donc $f(x, y) = 0$. Sinon, on pose

$$(u, v) = \left(\frac{x}{\|x\|_E}, \frac{y}{\|x\|_E}\right) \in S_E \times S_F.$$

Donc $\|f(u, v)\|_G \leq C$, d'où par bilinéarité de f,

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

$(iii) \Rightarrow (ii)$: soit M le majorant de f sur $B_E \times B_F$. Alors, pour tout $(x, y) \in E \times F$ tel que $x \neq 0$ et $y \neq 0$, on a

$$\begin{aligned}\|f(x, y)\|_G &= \left\|f\left(2\|x\|_E\frac{x}{2\|x\|_E}, 2\|y\|_E\frac{y}{2\|y\|_E}\right)\right\|_G \\ &= 4\|x\|_E\|y\|_F\left\|f\left(\frac{x}{2\|x\|_E}, \frac{y}{2\|y\|_E}\right)\right\|_G \leq 4M\|x\|_E\|y\|_F.\end{aligned}$$

Cette inégalité reste évidemment vraie si $x = 0$ ou $y = 0$.

Soient maintenant deux vecteurs (u, v) et (x, y) de $E \times F$. Supposons que $\|u - x\|_E \leq 1$ et que $\|y - v\|_F \leq 1$. Alors on a

$$\begin{aligned}
\|f(x,y) - f(u,v)\|_G &= \|f(x,y) - f(u,y) + f(u,y) - f(u,v)\|_G \\
&\leq \|f(x,y) - f(u,y)\|_G + \|f(u,y) - f(u,v)\|_G \\
&= \|f(x-u,y)\|_G + \|f(u,y-v)\|_G \\
&\leq 4M\left(\|x-u\|_E\|y\|_F + \|u\|_E\|y-v\|_F\right) \\
&\leq 4M\|y\|_F\|x-u\|_E + 4M(\|x\|_E + 1)\|y-v\|_F.
\end{aligned}$$

Soit $\varepsilon > 0$. On pose

$$\delta = \min\left(1, \frac{\varepsilon}{8M\|y\|_F}, \frac{\varepsilon}{8M(\|x\|_E + 1)}\right),$$

de sorte que si $(u, v) \in B_\delta^{E\times F}(x, y)$, alors $\|x - u\|_E < 1$ et $\|y - v\|_F < 1$, et l'inégalité précédente est vraie. De plus, elle implique

$$\|f(x,y) - f(u,v)\|_G < 4M\|y\|_F\delta + 4M(\|x\|_E + 1)\delta \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Ceci prouve bien la continuité. □

De même que dans le cas linéaire, on a

Corollaire 1.41 *Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ des espaces vectoriels normés. On suppose que E et F ne sont pas réduits à $\{0\}$. Soit f une application bilinéaire de $E \times F$ dans G. Alors*

$$\begin{aligned}
f \text{ continue} \quad &\iff \sup_{(x,y)\neq(0,0)} \frac{\|f(x,y)\|_F}{\|x\|_E\|y\|_F} < +\infty \\
&\iff \sup_{\|x\|_E = \|y\|_F = 1} \|f(x,y)\|_F < +\infty.
\end{aligned}$$

Et de même que dans le cas linéaire, on peut définir une norme sur l'espace vectoriel des applications bilinéaires de $E \times F$ dans G :

Définition 1.42 *Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ des espaces vectoriels normés. On appelle $\mathcal{B}(E\times F, G)$ l'ensemble des applications bilinéaires continues de $E \times F$ dans G.*

Définition 1.43 *Soient $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ et $(G, \|\cdot\|_G)$ des espaces vectoriels normés. On suppose que E et F ne sont pas réduits à $\{0\}$. Soit $f \in \mathcal{B}(E\times F, G)$. On appelle norme de f la quantité*

$$\|f\|_{\mathcal{B}(E\times F,G)} = \sup_{(x,y)\neq(0,0)} \frac{\|f(x,y)\|_F}{\|x\|_E\|y\|_F} = \sup_{\|x\|_E = \|y\|_F = 1} \|f(x,y)\|_F.$$

C'est une norme sur l'espace $\mathcal{B}(E \times F, G)$.

La preuve de ce résultat reproduit celle de la définition 1.33, nous la laissons en exercice.

On a en particulier le résultat suivant :

Corollaire 1.44 *Soient* $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ *et* $(G, \|\cdot\|_G)$ *des espaces vectoriels normés. Soit* $f \in \mathcal{B}(E \times F, G)$. *Alors*

$$\forall x \in E, \ \forall y \in F, \quad \|f(x,y)\|_G \leq \|f\|_{\mathcal{B}(E\times F,G)} \|x\|_E \|y\|_F.$$

Remarque 1.45 *Tout ce qui précède est également valable si on munit* $E \times F$ *de la norme* $N(x,y) = \sup(\|x\|_E, \|y\|_F)$.

Finissons cette section par le cas général des applications multilinéaires :

Définition 1.46 *Soient* $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$, $\ldots$ $(E_p, \|\cdot\|_{E_p})$ *des espaces vectoriels normés. On munit* $E = E_1 \times \cdots \times E_p$ *de la norme*

$$\|(u_1, u_2, \ldots, u_p)\|_E = \sum_{i=1}^{p} \|u_i\|_{E_i}.$$

Soit $(G, \|\cdot\|_G)$ *un espace vectoriel normé. On dit que* $f : E_1 \times \ldots E_p \longrightarrow G$ *est p-linéaire si elle est linéaire par rapport à chaque* x_i.

Proposition 1.47 *Soient* $(E_1, \|\cdot\|_{E_1})$, $(E_2, \|\cdot\|_{E_2})$, $\ldots$ $(E_p, \|\cdot\|_{E_p})$ *des espaces vectoriels normés. On munit* $E = E_1 \times \cdots \times E_p$ *de la norme*

$$\|(u_1, u_2, \ldots, u_p)\|_E = \sum_{i=1}^{p} \|u_i\|_{E_i}.$$

Soit $(G, \|\cdot\|_G)$ *un espace vectoriel normé. Soit* $f : E_1 \times \ldots E_p \longrightarrow G$ *p-linéaire.*

On alors équivalence entre

(i) f *est continue.*

(ii) f *est continue en* $(0, \ldots, 0)$.

(iii) f *est bornée sur* $B_{E_1} \times B_{E_2} \times \cdots \times B_{E_p}$.

(iv) f *est bornée sur* $S_{E_1} \times S_{E_2} \times \cdots \times S_{E_p}$.

(v) $\exists M > 0, \ \forall (x_1, \ldots, x_p) \in E_1 \times \cdots \times E_p, \ \|f(x_1, \ldots x_p)\| \leq M \prod_{i=1}^{p} \|x_i\|_{E_i}$

La preuve de ce résultat est laissée en exercice. Il s'agit d'une généralisation simple de celle de la proposition 1.40.

1.5 Équivalence de normes

Définition 1.48 *Soit E un espace vectoriel muni de deux normes N_1 et N_2. On dit qu'elles sont équivalentes si*

$$\exists M \geq m > 0, \quad \textit{tel que} \quad \forall x \in E, \quad mN_2(x) \leq N_1(x) \leq MN_2(x).$$

On note $N_1 \sim N_2$.

Remarquons que $N_1 \sim N_2$ si et seulement si $N_2 \sim N_1$. Il est d'ailleurs facile de montrer que la relation $\sim$ est une relation d'équivalence. De plus, deux normes équivalentes définissent les mêmes topologies :

Proposition 1.49 *Soit E un espace vectoriel muni de deux normes N_1 et N_2. Elles sont équivalentes si et seulement si elles définissent les mêmes ouverts.*

Démonstration. On note $B_r^{N_1}(x)$ la boule de centre x et de rayon $r > 0$ pour la norme N_2, et $B_r^{N_2}(x)$ celle pour la norme N_2.

Supposons $N_1 \sim N_2$. Soit U_1 un ouvert pour N_1. Soit $x \in U_1$. Alors, par définition d'un ouvert, on sait qu'il existe $\varepsilon > 0$ tel que $B_\varepsilon^{N_1}(x) \subset U_1$. On sait que $N_1(x-y) \leq MN_2(x-y)$, donc

$$B_{\varepsilon/M}^{N_2}(x) \subset B_\varepsilon^{N_1}(x) \subset U_1,$$

donc U_1 est bien un ouvert pour N_2. Il est clair que la même preuve s'applique pour U_2 ouvert pour la norme N_2, et permet de prouver que U_2 est ouvert pour N_1.

Réciproquement, supposons que N_1 et N_2 définissent les mêmes ouverts. On pose alors

$$\begin{array}{rcl} f \ : \ (E, N_1) & \longrightarrow & (E, N_2), \\ x & \longmapsto & x. \end{array}$$

Il est clair que f est une application linéaire. De plus, pour tout $x \in E$, si V_2 est un ouvert pour N_2, alors $f^{-1}(V_2) = V_2$ est un ouvert pour N_1. Donc f est continue. Donc, d'après la proposition 1.31, on sait qu'il existe $M > 0$ tel que $\forall x \in E, \quad N_2(f(x)) \leq MN_1(x)$, c'est-à-dire,

$$\forall x \in E, \quad N_2(x) \leq MN_1(x).$$

On reproduit ensuite la preuve ci-dessus en échangeant les rôles de N_1 et N_2, et on obtient l'inégalité inverse. Ce qui prouve que N_1 et N_2 sont bien équivalentes. □

Corollaire 1.50 *Soit E un espace vectoriel muni de deux normes N_1 et N_2. Elles sont équivalentes si et seulement si elles définissent les mêmes voisinages.*

Corollaire 1.51 *Soit F un espace vectoriel normé. Soit E un espace vectoriel muni de deux normes N_1 et N_2. Si elles sont équivalentes, toute application de E dans F continue pour l'une est continue pour l'autre.*

Démonstration. Si N_1 et N_2 sont équivalentes, on sait d'après la proposition 1.49 qu'un voisinage de x_0 pour N_1 est un voisinage de x_0 pour N_2, et inversement. De plus, une application est continue si et seulement si pour tout $x_0 \in E$ et pour tout voisinage V de $f(x_0)$, $f^{-1}(V)$ est un voisinage de x_0. Ainsi, si f est continue pour N_1, elle l'est pour N_2, et inversement. □

1.6 Exemples d'espaces vectoriels normés

1.6.1 Espaces de suites

Définition 1.52 *On note, pour tout réel $p \geq 1$,*

$$\ell^p(\mathbb{K}) = \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}},\ \sum_{n\geq 0} |u_n|^p < +\infty \right\}.$$

$$\ell^\infty(\mathbb{K}) = \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}},\ \sup_{n\geq 0} |u_n| < +\infty \right\}.$$

Nous allons démontrer que chacun d'entre eux est un espace vectoriel, et construire des normes sur ces espaces vectoriels.

Définition 1.53 *Pour tout réel $p \geq 1$, et pour $u \in \ell^p(\mathbb{K})$, on note*

$$\|u\|_p = \left(\sum_{n\geq 0} |u_n|^p \right)^{1/p},$$

et

$$\|u\|_\infty = \sup_{n\geq 0} |u_n|.$$

Proposition 1.54 *L'ensemble $\ell^\infty(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel, et la quantité $\|\cdot\|_\infty$ définie ci-dessus est une norme sur cet espace.*

Démonstration. Il est clair qu'une combinaison linéaire de deux suites bornées est une suite bornée, donc $\ell^\infty(\mathbb{K})$ est bien un espace vectoriel.

On démontre successivement les trois axiomes de définition des normes. Tout d'abord, on a évidemment $\|0\|_\infty = 0$. Ensuite, si $\|u\|_\infty = 0$, alors pour tout $n \in \mathbb{N}$, on a $|u_n| \leq 0$, et u est donc la suite nulle.

Démontrons maintenant l'homogénéité : soit $u \in \ell^\infty(\mathbb{K})$ et $\lambda \in \mathbb{K}$. Alors λu est évidemment une suite bornée, et on a

$$\|\lambda u\|_\infty = \sup_{n\geq 0} |\lambda u_n| = \sup_{n\geq 0} |\lambda||u_n| = |\lambda| \sup_{n\geq 0} |u_n| = |\lambda|\|u\|_\infty,$$

car λ ne dépend pas de n.

Enfin, l'inégalité triangulaire se démontre facilement : si $u \in \ell^\infty(\mathbb{K})$ et $v \in \ell^\infty(\mathbb{K})$, alors

$$\|u+v\|_\infty = \sup_{n\geq 0}(|u_n + v_n|) \leq \sup_{n\geq 0}|u_n|+|v_n| \leq \sup_{n\geq 0}|u_n|+\sup_{n\geq 0}|v_n| = \|u\|_\infty+\|v\|_\infty.$$

Ceci termine la démonstration. □

Pour démontrer que la quantité $\|\cdot\|_p$ est bien une norme sur $\ell^p(\mathbb{K})$ pour $p \geq 1$, on commence par démontrer l'inégalité de Young :

Proposition 1.55 *Soient $p > 1$, et $q > 1$ tels que $\frac{1}{p}+\frac{1}{q}=1$. Alors, pour tout couple de réels positifs a et b, on a*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}. \tag{1.4}$$

Démonstration. Si $a = 0$ ou $b = 0$, l'inégalité est évidemment vraie. On peut donc supposer que $a > 0$ et $b > 0$. La fonction logarithme est concave, donc

$$\forall x > 0, \quad \forall y > 0, \quad \forall \lambda \in [0,1], \quad \ln(\lambda x + (1-\lambda)y) \geq \lambda \ln(x) + (1-\lambda)\ln(y).$$

On applique cette inégalité avec $\lambda = \frac{1}{p}$, $x = a^p$ et $y = b^q$. On obtient alors

$$\ln\left(\frac{a^p}{p}+\frac{b^q}{q}\right) \geq \frac{1}{p}\ln(a^p) + \frac{1}{q}\ln(b^q) = \ln(a)+\ln(b) = \ln(ab).$$

La fonction exponentielle étant croissante, on peut l'appliquer à cette dernière inégalité, ce qui donne bien (1.4). □

Remarque 1.56 *L'inégalité de Young est souvent donnée avec l'hypothèse $p \geq 1$, en utilisant les conventions suivantes : si $p = 1$, alors $q = \infty$, et dans ce cas $\frac{b^q}{q} = +\infty$ si $b > 1$, et $\frac{b^q}{q} = 0$ si $b \leq 1$. L'inégalité devient donc triviale pour $p = 1$.*

Proposition 1.57 (inégalité de Hölder) *Soit $p \geq 1$, et $q \geq 1$ tels que $\frac{1}{p}+\frac{1}{q}=1$. Supposons que $u \in \ell^p$ et $v \in \ell^q$. Alors $uv \in \ell^1$, et*

$$\|uv\|_1 \leq \|u\|_p\|v\|_q.$$

Démonstration. Si $p = 1$, on a $u \in \ell^1$ et $v \in \ell^\infty$, d'où $uv \in \ell^1$, et l'inégalité

$$\|uv\|_1 = \sum_{n\geq 0}|u_n||v_n| \leq \sum_{n\geq 0}|u_n|\|v\|_\infty = \|u\|_1\|v\|_\infty.$$

On suppose maintenant $p > 1$. On commence par traiter le cas particulier où $\|u\|_p = \|v\|_q = 1$. Fixons $N \in \mathbb{N}$, et appliquons l'inégalité de Young :

$$\sum_{n=0}^N |u_n v_n| \leq \sum_{n=0}^N \left(\frac{|u_n|^p}{p} + \frac{|u_n|^q}{q}\right) \leq \frac{\|u\|_p^p}{p} + \frac{\|v\|_q^q}{q} = \frac{1}{p}+\frac{1}{q} = 1.$$

Donc, en faisant tendre N vers l'infini, on obtient que $uv \in \ell^1(\mathbb{K})$, et que $\|uv\|_1 \leq 1$.

Si maintenant u ou v ne sont pas de norme 1, on commence par remarquer que si $u = 0$ ou $v = 0$, alors $uv = 0 \in \ell^1$ et l'inégalité est vraie. On suppose donc que $u \neq 0$ et $v \neq 0$. On pose alors

$$\tilde{u} = \frac{u}{\|u\|_p}, \quad \tilde{v} = \frac{v}{\|v\|_q}.$$

D'après ce qu'on vient de démontrer, $\tilde{u}\tilde{v} \in \ell^1$, et $\|\tilde{u}\tilde{v}\|_1 \leq 1$. Donc, par homogénéité de la norme,

$$\frac{\|uv\|_1}{\|u\|_p\|v\|_q} = \|\tilde{u}\tilde{v}\|_1 \leq 1,$$

ce qui prouve l'inégalité voulue. □

Nous allons maintenant démontrer l'inégalité de Minkovski :

Proposition 1.58 (inégalité de Minkovski) *Soit $p \geq 1$. Si $u \in \ell^p(\mathbb{K})$ et $v \in \ell^p(\mathbb{K})$, alors $u + v \in \ell^p(\mathbb{K})$, et on a*

$$\|u + v\|_p \leq \|u\|_p + \|v\|_p. \tag{1.5}$$

Démonstration. Commençons par le cas $p = 1$. Alors, pour $N \in \mathbb{N}$ fixé, on a

$$\sum_{n=0}^{N} |u_n + v_n| \leq \sum_{n=0}^{N} (|u_n| + |v_n|) = \sum_{n=0}^{N} |u_n| + \sum_{n=0}^{N} |v_n|.$$

Donc, en faisant tendre N vers l'infini, on obtient d'une part que $u + v \in \ell^1$, et d'autre part l'inégalité (1.5).

Supposons maintenant $p > 1$. Fixons à nouveau $N \in \mathbb{N}$. On a, en supposant que $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned}
\sum_{n=0}^{N} |u_n + v_n|^p &= \sum_{n=0}^{N} |u_n + v_n| \; |u_n + v_n|^{p-1} \\
&\leq \sum_{n=0}^{N} |u_n| \; |u_n + v_n|^{p-1} + \sum_{n=0}^{N} |v_n| \; |u_n + v_n|^{p-1} \\
&\leq \left(\sum_{n=0}^{N} |u_n|^p \right)^{1/p} \left(\sum_{n=0}^{N} |u_n + v_n|^{q(p-1)} \right)^{1/q} \\
&\qquad + \left(\sum_{n=0}^{N} |v_n|^p \right)^{1/p} \left(\sum_{n=0}^{N} |u_n + v_n|^{q(p-1)} \right)^{1/q},
\end{aligned}$$

grâce à l'inégalité de Hölder. De plus, $q(p-1) = pq\left(1 - \frac{1}{p}\right) = \frac{pq}{q} = p$, et donc

$$\sum_{n=0}^{N} |u_n + v_n|^p \leq (\|u\|_p + \|v\|_p) \left(\sum_{n=0}^{N} |u_n + v_n|^p \right)^{1-1/p}.$$

On suppose N assez grand pour que le membre de gauche de cette inégalité ne soit pas nul. On peut donc diviser par le deuxième facteur du membre de droite, et on obtient

$$\left(\sum_{n=0}^{N}|u_n+v_n|^p\right)^{1/p} \leq \|u\|_p + \|v\|_p.$$

Ainsi, $u+v \in \ell^p$, et, en faisant tendre N vers l'infini, on obtient l'inégalité de Minkovski. □

Proposition 1.59 *Soit $p \geq 1$. Alors l'ensemble $\ell^p(\mathbb{K})$ est un $\mathbb{K}$-espace vectoriel, et la quantité $\|\cdot\|_p$ de la définition 1.53 est une norme sur cet espace.*

Démonstration. L'inégalité de Minkovski implique si $u \in \ell^p(\mathbb{K})$ et $v \in \ell^p(\mathbb{K})$, alors $u+v \in \ell^p(\mathbb{K})$. De plus, on a

$$\forall u \in \ell^p(\mathbb{K}), \quad \forall \lambda \in \mathbb{K}, \quad \forall N \in \mathbb{N}, \quad \sum_{n=1}^{N}|\lambda u_n| = |\lambda|\sum_{n=0}^{N}|u_n|,$$

Ce qui prouve tout d'abord que $\lambda u \in \ell^p(\mathbb{K})$, et de plus, en faisant tendre N vers l'infini, que $\|\lambda u\|_p = |\lambda|\|u\|_p$. Donc $\ell^p(\mathbb{K})$ est bien un espace vectoriel, et l'application $u \to \|u\|_p$ est homogène. Il est clair que $\|u\|_p = 0 \iff u = 0$, et l'inégalité triangulaire est exactement équivalente à l'inégalité de Minkovski. □

Proposition 1.60 *Soient deux réels p_1 et p_2 tels que $1 \leq p_1 \leq p_2$. Alors*

$$\ell^{p_1}(\mathbb{K}) \subset \ell^{p_2}(\mathbb{K}) \subset \ell^{\infty}(\mathbb{K}).$$

Démonstration. Tout d'abord, il est clair que si $u \in \ell^{p_2}(\mathbb{K})$, alors u est une suite bornée. Donc $\ell^{p_2}(\mathbb{K}) \subset \ell^{\infty}(\mathbb{K})$. Supposons maintenant que $u \in \ell^{p_1}(\mathbb{K})$. Alors, d'après ce qu'on vient de voir, $u \in \ell^{\infty}(\mathbb{K})$. De plus, pour tout entier $N \geq 0$, on a

$$\sum_{n=0}^{N}|u_n|^{p_2} = \sum_{n=0}^{N}|u_n|^{p_1}|u_n|^{p_2-p_1} \leq \|u\|_{p_1}^{p_1}\|u\|_{\infty}^{p_2-p_1}.$$

Ceci prouve que $u \in \ell^{p_2}(\mathbb{K})$, d'où l'inclusion $\ell^{p_1}(\mathbb{K}) \subset \ell^{p_2}(\mathbb{K})$. □

Corollaire 1.61 *Soient deux réels p_1 et p_2 tels que $p_1 \leq p_2$. Alors l'application*

$$\begin{array}{rccc} I \ : & \ell^{p_1}(\mathbb{K}) & \longrightarrow & \ell^{p_2}(\mathbb{K}), \\ & u & \longmapsto & u. \end{array}$$

est linéaire continue, et de norme 1.

Démonstration. Il est clair que I est linéaire. De plus, l'inégalité ci-dessus implique que

$$\|u\|_{p_2}^{p_2} \leq \|u\|_{p_1}^{p_1}\|u\|_{\infty}^{p_2-p_1}.$$

Par ailleurs, il est clair que, par définition de $\|\cdot\|_{p_1}$, on a, pour tout $n \in \mathbb{N}$, $|u_n| \leq \|u\|_{p_1}$. Ceci implique $\|u\|_\infty \leq \|u\|_{p_1}$. En insérant cette inégalité dans l'inégalité ci-dessus, on en déduit

$$\|u\|_{p_2}^{p_2} \leq \|u\|_{p_1}^{p_2}, \quad \text{donc} \quad \|u\|_{p_2} \leq \|u\|_{p_1}.$$

Ceci prouve que I est continue, et que $\|I\|_{\mathcal{L}(\ell^{p_1}(\mathbb{K}),\ell^{p_2}(\mathbb{K}))} \leq 1$. Pour montrer que cette inégalité est en fait une égalité, nous exhibons une suite u telle que $\|u\|_{p_1} = \|u\|_{p_2}$. On choisit $u_n = 1$ si $n = 0$, $u_n = 0$ sinon. Cette suite est évidemment dans $\ell^{p_1}(\mathbb{K})$, et $\|u\|_{p_1} = 1 = \|u\|_{p_2}$. □

1.6.2 Espaces de fonctions continues

Définition 1.62 *Soit $I = [a,b]$ un intervalle borné de $\mathbb{R}$. Soit $E = C^0(I,\mathbb{K})$. On définit les quantités suivantes :*

$$\forall f \in E, \quad \forall p \geq 1, \quad \|f\|_p = \int_a^b |f(t)|^p dt,$$
$$\forall f \in E, \qquad \|f\|_\infty = \sup_{x\in I} |f(x)|.$$

Nous allons démontrer qu'il s'agit de normes sur E. Pour cela, on commence par démontrer l'inégalité de Hölder dans ce cas-là :

Proposition 1.63 (inégalité de Hölder) *Soient deux réels p et q tels que $p \geq 1$, $q \geq 1$ et $\frac{1}{p} + \frac{1}{q} = 1$ (par convention, $q = \infty$ si $p = 1$ et inversement). Alors*

$$\forall f \in E, \quad \forall g \in E, \quad \left|\int_a^b fg\right| \leq \|f\|_p \|g\|_q.$$

Démonstration. Commençons par le cas où $p = 1$. Alors l'inégalité se démontre simplement :

$$\left|\int_a^b fg\right| \leq \int_a^b |fg| \leq \int_a^b |f|\|g\|_\infty = \|f\|_1\|g\|_\infty.$$

Pour le cas $p > 1$, ici encore, on commence par le cas $\|f\|_p = \|g\|_q = 1$, et on utilise l'inégalité de Young :

$$\left|\int_a^b fg\right| \leq \int_a^b |fg| \leq \int_a^b \frac{|f|^p}{p} + \int_a^b \frac{|g|^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Si maintenant $f = 0$ ou $g = 0$, l'inégalité est triviale. Supposons donc que $f \neq 0$ et que $g \neq 0$, de sorte que $\|f\|_p \neq 0$ et $\|g\|_q \neq 0$. On applique alors l'inégalité ci-dessus à $\tilde{f} = \frac{f}{\|f\|_p}$ et $\tilde{g} = \frac{g}{\|g\|_q}$:

$$\frac{1}{\|f\|_p\|g\|_q}\left|\int_a^b fg\right| = \left|\int_a^b \tilde{f}\tilde{g}\right| \leq 1.$$

Ainsi,
$$\left|\int_a^b fg\right| \leq \|f\|_p \|g\|_q.$$
□

Proposition 1.64 (inégalité de Minkowski) *Soit $p \geq 1$. Alors pour tout $f \in E$ et tout $g \in E$, on a*
$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

Démonstration. Le cas $p = 1$ est très facile. On suppose que $p > 1$. Alors on a

$$\begin{aligned}\int_a^b |f+g|^p = \int_a^b |f+g||f+g|^{p-1} &\leq \int_a^b |f||f+g|^{p-1} + \int_a^b |g||f+g|^{p-1} \\ &\leq \|f\|_p \left(\int_a^b |f+g|^{q(p-1)}\right)^{1/q} + \|g\|_p \left(\int_a^b |f+g|^{q(p-1)}\right)^{1/q},\end{aligned}$$

en appliquant l'inégalité de Hölder. De plus, $q(p-1) = qp\left(1-\frac{1}{p}\right) = qp\frac{1}{q} = p$, et donc
$$\|f+g\|_p^p \leq (\|f\|_p + \|g\|_q)\, \|f+g\|_p^{p/q},$$
d'où
$$\|f+g\|_p^{p-p/q} \leq \|f\|_p + \|g\|_q.$$
Comme $p - p/q = 1$, on obtient bien l'inégalité triangulaire. □

Proposition 1.65 *Les quantités $\|\cdot\|_p$ sont des normes sur E pour tout $p \geq 1$. La quantité $\|\cdot\|_\infty$ est une norme sur E.*

Démonstration. Commençons par le cas $p = \infty$. Il est clair que $\|f\|_\infty = 0 \iff f = 0$. De plus, pour tout réel λ et toute fonction $f \in E$,

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

Enfin, pour tout f et g dans E, on a

$$\begin{aligned}\|f+g\|_\infty = \sup_{x\in I} |f(x)+g(x)| &\leq \sup_{x\in I} (|f(x)| + |g(x)|) \\ &\leq \sup_{x\in I} |f(x)| + \sup_{x\in I} |g(x)| \leq \|f\|_\infty + \|f\|_\infty.\end{aligned}$$

Nous traitons maintenant le cas p réel (avec $p \geq 1$.) Il est clair que $\|f\|_p = 0 \iff f = 0$, et que l'homogénéité est vérifiée. Pour l'inégalité triangulaire, on applique l'inégalité de Minkowski. □

Chapitre 2

Espaces métriques

2.1 Distances

Définition 2.1 *Soit E un ensemble, et $d : E \times E \to \mathbb{R}^+$ une application. On dit que d est une distance sur E si*

(i) $\forall (x,y) \in E \times E, \quad d(x,y) = 0 \iff x = y.$

(ii) $\forall (x,y) \in E \times E, \quad d(x,y) = d(y,x).$

(iii) $\forall (x,y,z) \in E \times E \times E, \quad d(x,z) \leq d(x,y) + d(y,z).$

Le couple (E,d) est alors appelé espace métrique.

La propriété (i) est évidemment très voisine de la propriété (i) de définition des normes (définition 1.1). La propriété (ii) est en général appelée symétrie de l'application distance, et la propriété (iii) est l'inégalité triangulaire. Il s'agit d'une généralisation de la propriété (iii) de la définition 1.1.

On a le lien suivant entre espaces vectoriels normés et espaces métriques :

Proposition 2.2 *Soit E un espace vectoriel normé. Soit $A \subset E$ non vide. Alors A muni de la distance*

$$\begin{array}{rccl} d \; : & E \times E & \longrightarrow & \mathbb{R}^+, \\ & (x,y) & \longmapsto & \|x-y\|, \end{array} \tag{2.1}$$

est un espace métrique.

Démonstration. Commençons par démontrer (i) : comme $\|\cdot\|$ est une norme, on a

$$d(x,y) = 0 \iff \|x-y\| = 0 \iff x-y = 0 \iff x = y.$$

Pour démontrer (ii), on utilise l'homogénéité de la norme :

$$d(x,y) = \|x-y\| = \|y-x\| = d(y,x).$$

Enfin, (iii) découle directement de l'inégalité triangulaire pour $\|\cdot\|$:

$$d(x,z) = \|x-z\| = \|x-y+y-z\| \leq \|x-y\| + \|y-z\| = d(x,y) + d(y,z).$$

L'application d est donc bien une distance. $\square$

Citons maintenant quelques exemples de distances qui ne sont pas de la forme (2.1).

$$\begin{array}{rrcl} d\ : & \mathbb{R}\times\mathbb{R} & \longrightarrow & \mathbb{R}^+, \\ & (x,y) & \longmapsto & \left|x^3-y^3\right|, \end{array}$$

est une distance. Par ailleurs, si E est un espace vectoriel normé, les applications suivantes sont des distances :

$$\begin{array}{rrcl} d\ : & E\times E & \longrightarrow & \mathbb{R}^+, \\ & (x,y) & \longmapsto & \min\left(1,\|x-y\|\right), \end{array}$$

et

$$\begin{array}{rrcl} d\ : & E\times E & \longrightarrow & \mathbb{R}^+, \\ & (x,y) & \longmapsto & \frac{\|x-y\|}{1+\|x-y\|}, \end{array} \tag{2.2}$$

sont des distances. Pour les deux premiers cas, nous laissons la preuve en exercice. Démontrons que (2.2) définit bien une distance. Tout d'abord, il est clair que (i) est vérifiée, car $\|\cdot\|$ est une norme. De plus, il est aussi clair que d définie par (2.2) est symétrique, par homogénéité de la norme (propriété (ii) de la définition 1.1). Enfin, démontrons l'inégalité triangulaire. On sait, tout d'abord, puisque $\|\cdot\|$ est une norme, que

$$\forall (x,y,z)\in E\times E\times E,\quad \|x-z\|\leq\|x-y\|+\|y-z\|.$$

Par ailleurs, l'application $f:\mathbb{R}^+\to\mathbb{R}^+$ définie par

$$f(t)=\frac{t}{1+t},$$

est continue et strictement croissante. En effet, elle est dérivable, et sa dérivée est égale à

$$f'(t)=\frac{1}{1+t}-\frac{t}{(1+t)^2}=\frac{1+t-t}{(1+t)^2}=\frac{1}{(1+t)^2}\geq 0.$$

On a donc

$$\begin{aligned} d(x,z)=f(\|x-z\|)\leq f\left(\|x-y\|+\|y-z\|\right) &= \frac{\|x-y\|+\|y-z\|}{1+\|x-y\|+\|y-z\|} \\ &= \frac{\|x-y\|}{1+\|x-y\|+\|y-z\|}+\frac{\|y-z\|}{1+\|x-y\|+\|y-z\|} \\ &\leq \frac{\|x-y\|}{1+\|x-y\|}+\frac{\|y-z\|}{1+\|y-z\|}=d(x,y)+d(y,z), \end{aligned}$$

ce qui prouve l'inégalité triangulaire.

Remarque 2.3 (importante) *De même que pour le cas des normes et semi-normes, on a que si d est une distance sur un ensemble E, alors*

$$\forall (x,y,z)\in E\times E\times E,\quad d(x,z)\geq\left|d(x,y)-d(y,z)\right|. \tag{2.3}$$

Démonstration. On a, d'après l'inégalité triangulaire, $d(x,y) \leq d(x,z) + d(z,y)$. Ainsi,

$$d(x,z) \geq d(x,y) - d(y,z).$$

De même, en échangeant les rôles de x et z, on a $d(z,y) \leq d(z,x) + d(x,y)$, donc

$$d(x,z) \geq d(y,z) - d(x,y).$$

Donc

$$d(x,z) \geq \max(d(x,y) - d(y,z), d(y,z) - d(x,y)) = |d(x,y) - d(y,z)|,$$

car, pour tout réel a, $|a| = \max(a, -a)$. □

Proposition 2.4 *Soit* $(E_1, d_1), \ldots, (E_p, d_p)$ *des espaces métriques. Soit*

$$E = E_1 \times E_2 \times \cdots \times E_p.$$

Alors l'application

$$\begin{array}{rccl} d : & E \times E & \longrightarrow & \mathbb{R}^+, \\ & (x,y) & \longmapsto & \max\limits_{1 \leq i \leq p} d_i(x_i, y_i), \end{array}$$

est une distance sur E.

Démonstration. Commençons par démontrer la propriété (i) de la définition 2.1. Si $d(x,y) = 0$, alors on a, pour tout entier i entre 1 et p, $d_i(x_i, y_i) = 0$. Donc $x_i = y_i$. Ainsi, $x = y$. La réciproque est immédiate. La preuve du point (ii) est triviale. Enfin, pour démontrer l'inégalité triangulaire, on suppose donnés trois vecteurs x, y, z de E, c'est-à-dire

$$x = (x_1, \ldots x_p), \quad y = (y_1, \ldots y_p), \quad z = (z_1, \ldots z_p).$$

et on écrit :

$$\forall i, \quad d_i(x_i, z_i) \leq d_i(x_i, y_i) + d_i(y_i, z_i).$$

Ceci implique que

$$\max_{1 \leq i \leq p} d_i(x_i, z_i) \leq \max_{1 \leq i \leq p} d_i(x_i, y_i) + \max_{1 \leq i \leq p} d_i(y_i, z_i),$$

donc $d(x,z) \leq d(x,y) + d(y,z)$, et termine la démonstration. □

2.2 Boules, ensembles ouverts et fermés

Nous reprenons dans cette section les notions développées dans la section 1.2 du chapitre sur les espaces vectoriels normés. Comme nous allons le voir, les notions en question ne sont pas liés à la notion de norme, mais bien à celle de distance.

Définition 2.5 *Soit (E, d) un espace métrique, soit $x \in E$ et $r \in \mathbb{R}^{+*}$. On appelle :*

- *boule ouverte de centre x et de rayon r l'ensemble*

$$B_r(x) = \{y \in E, \quad d(x, y) < r\},$$

- *boule fermée de centre x et de rayon r l'ensemble*

$$\overline{B_r}(x) = \{y \in E, \quad d(x, y) \leq r\},$$

- *sphère de centre x et de rayon r l'ensemble*

$$S_r(x) = \{y \in E, \quad d(x, y) = r\},$$

Exemple 2.6 *Dans $\mathbb{R}$, muni de la distance $d(x, y) = |x - y|$, la boule ouverte de centre x et de rayon r est l'intervalle $B_r(x) =]x - r, x + r[$. La boule fermée ce centre x et de rayon r est l'intervalle $\overline{B_r}(x) = [x - r, x + r]$.*

Donnons maintenant un exemple moins intuitif :

Exemple 2.7 *Munissons $\mathbb{K}$ de la distance suivante, dite "discrète" :*

$$\forall (x, y) \in \mathbb{K}^2, \quad d(x, y) = \begin{cases} 0 & \text{si } x = y, \\ 1 & \text{si } x \neq y. \end{cases}$$

Alors, on a, pour tout $x \in \mathbb{K}$ et tout $r > 0$,

$$B_r(x) = \begin{cases} \{x\} & \text{si } r \leq 1, \\ \mathbb{K} & \text{si } r > 1, \end{cases} \qquad \overline{B_r}(x) = \begin{cases} \{x\} & \text{si } r < 1, \\ \mathbb{K} & \text{si } r \geq 1. \end{cases}$$

Démonstration. En effet, il est clair que $x \in B_r(x) \subset \overline{B_r}(x)$ pour tout $r > 0$. De plus, si $r \leq 1$, et que $y \in B_r(x)$, alors $d(x, y) < 1$, ce qui implique, d'après la définition de d, que $d(x, y) = 0$, donc que $x = y$. Ainsi, $B_r(x) = \{x\}$. Supposons maintenant $r > 1$. Alors, pour tout $y \in \mathbb{K}$, $d(x, y) \leq 1 < r$, et donc $B_r(x) = \mathbb{K}$.

Démontrons maintenant l'égalité concernant $\overline{B_r}(x)$. Si $r < 1$, alors $y \in \overline{B_r}(x)$ implique $d(x, y) \leq r < 1$, ce qui, toujours d'après la définition de d, implique $d(x, y) = 0$, donc $x = y$. D'où $\overline{B_r}(x) = \{x\}$. Et si $r \geq 1$, alors pour tout $y \in \mathbb{K}$, on a $d(x, y) \leq 1 \leq r$, ce qui prouve que $\overline{B_r}(x) = \mathbb{K}$. □

Définition 2.8 *Soit (E, d) un espace métrique, et soit $A \subset E$. On dit que*

- *A est ouvert si $\forall x \in A, \quad \exists r > 0, \quad B_r(x) \subset A$,*
- *A est fermé si son complémentaire $A^C = E \setminus A$ est ouvert.*

En particulier, $\emptyset$ est un ouvert.

Exemple 2.9 *Soient $a < b$ deux réels. Alors, dans $\mathbb{R}$ muni de la distance $d(x, y) = |x - y|$,*

1. *l'intervalle $]a, b[$ est ouvert ;*

2. *l'intervalle $[a, b]$ est fermé.*

La preuve est bien entendu exactement la même que dans le cas des espaces vectoriels normés (exemple 1.14).

Proposition 2.10 *Soit (E, d) un espace métrique. Alors toute boule ouverte de E est un ouvert.*

Démonstration. On reprend exactement la preuve du cas des espaces vectoriels normés (proposition 1.16). Soit donc $x \in E$, et $r > 0$. Supposons que $y \in B_r(x)$. Alors $d(x, y) < r$, et on pose

$$r' = r - d(x, y) > 0.$$

Montrons alors que $B_{r'}(y) \subset B_r(x)$, ce qui prouvera que $B_r(x)$ est bien un ouvert. Soit $z \in B_{r'}(y)$. On a alors, en utilisant l'inégalité triangulaire,

$$d(z, x) \leq d(z, y) + d(y, x) < r' + d(x, y) = r - d(x, y) + d(x, y) = r,$$

ce qui prouve que $z \in B_r(x)$, et donc, puisque ceci vaut pour tout $z \in B_{r'}(y)$, que $B_{r'}(y) \subset B_r(x)$. □

On a également :

Proposition 2.11 *Soit (E, d) un espace métrique. Alors toute boule fermée de E est un fermé.*

Démonstration. Ici encore, on reprend la preuve du cas normé (proposition 1.17). Soit donc $x \in E$, $r > 0$, et A le complémentaire de $\overline{B_r}(x)$. Autrement dit,

$$A = \{y \in E, \quad d(x, y) > r\}.$$

Soit $y \in A$. On pose alors $r' = d(x, y) - r > 0$, et on va prouver que $B_{r'}(y) \subset A$, ce qui prouvera que A est ouvert, et terminera la démonstration. Soit donc $z \in B_{r'}(y)$. Alors $d(z, y) < r'$. Donc, en appliquant (2.3),

$$d(z, x) \geq d(x, y) - d(y, z) > d(x, y) - r' = d(x, y) - d(x, y) + r = r.$$

Ainsi, $z \in A$, pour tout $z \in B_{r'}(y)$, donc $B_{r'}(y) \subset A$. □

Définition 2.12 *Soit (E, d) un espace métrique, $x \in E$ et $A \subset E$. On dit que A est un voisinage de x s'il contient un ouvert contenant x.*

Proposition 2.13 *Soit (E, d) un espace métrique. Dans E, on a alors les propriétés suivantes :*

(i) Une réunion d'ouverts est un ouvert.

(ii) Une intersection finie d'ouverts est un ouvert.

(iii) $\emptyset$ et E sont ouverts et fermés.

Démonstration. La preuve du cas des espaces vectoriels normés se reproduit telle quelle, car seule la notion de boule ouverte y intervient :
(i) Soient des ouverts $(O_i)_{i\in I}$, et soit $O = \bigcup_{i\in I} O_i$. Soit $x \in O$. Alors, par définition de l'union, il existe $i_0 \in I$ tel que $x \in O_{i_0}$. Comme ce dernier ensemble est ouvert, on sait qu'il existe $r > 0$ tel que $B_r(x) \subset O_{i_0}$. Ainsi,

$$B_r(x) \subset O_{i_0} \subset O,$$

ce qui prouve que O est un ouvert.
(ii) Soient des ouverts $(O_i)_{1\leq i\leq p}$, et soit $O = \bigcap_{i=1}^{p} O_i$. Pour tout $x \in O$, on sait que $x \in O_i$, donc il existe $r_i > 0$ tel que $B_{r_i}(x) \subset O_i$. On pose donc $r = \min_{1\leq i\leq p} r_i$, et on a alors

$$\forall i, \quad B_r(x) \subset B_{r_i}(x) \subset O_i.$$

Donc $B_r(x) \subset O$. Ainsi, O est bien un ouvert.
(iii) $\emptyset$ est ouvert par définition. Donc E est fermé. De plus, pour tout $x \in E$, on a $B_1(x) \subset E$ par définition. Donc E est ouvert. Donc $\emptyset$ est fermé. □

Remarque 2.14 *Ici encore, une intersection infinie d'ouvert peut ne pas être un ouvert. Le contre-exemple est bien sûr le même que dans le cas des espaces vectoriels normés : dans $\mathbb{R}$ muni de la distance $d(x, y) = |x - y|$, soient les ouverts*

$$U_n = \left]-\frac{1}{n}, \frac{1}{n}\right[, \quad n \in \mathbb{N}^*.$$

Alors il est clair que

$$U = \bigcap_{n\in\mathbb{N}^*} U_n = \{0\},$$

qui n'est pas un ouvert.

Il convient cependant de noter que ceci dépend fortement de la distance. En effet, si on utilise celle de l'exemple 2.7, alors tout sous-ensemble de $\mathbb{K}$ est un ouvert, puisque les boules ouvertes sont soit des singletons, soit $\mathbb{K}$ tout entier. Donc, dans ce cas très particulier, une intersection infinie d'ouverts est un ouvert.

Corollaire 2.15 *Soit E un espace métrique. Dans E, on a alors les propriétés suivantes :*
(i) Une réunion finie de fermés est un fermé.
(ii) Une intersection de fermés est un fermé.

Proposition 2.16 *Soit E un espace métrique. Soit $A \subset E$. Alors A est ouvert si et seulement s'il est voisinage de chacun de ses points.*

Démonstration. Si A est ouvert, il est clair qu'il est voisinage de chacun de ses points. Réciproquement, si A est voisinage de chaque $x \in A$, alors il existe un ouvert U tel que $x \in U \subset A$. Comme U est ouvert, il existe $r > 0$ tel que $B_r(x) \subset U$, donc $B_r(x) \subset A$. Ceci valant pour tout $x \in A$, A est ouvert. □

2.3 Continuité

Définition 2.17 *Soient (E_1, d_1) et (E_2, d_2) des espaces métriques, $f : E_1 \longrightarrow E_2$, et $x_0 \in E_1$. On dit que f est continue en x_0 si*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 \quad \text{tel que} \quad d_1(x_0, x) < \delta \Rightarrow d_2\left(f(x_0), f(x)\right) < \varepsilon.$$

On dit que f est continue sur E_1 si elle est continue en tout point de E_1.

Proposition 2.18 *Soient (E_1, d_1) et (E_2, d_2) des espaces métriques, et $f : E_1 \longrightarrow E_2$. Alors on a équivalence entre :*

(i) f est continue sur E_1.

(ii) $\forall A$ ouvert de E_2, $f^{-1}(A)$ est un ouvert de E_1.

(iii) $\forall A$ fermé de E_2, $f^{-1}(A)$ est un fermé de E_1.

(iv) $\forall x \in E_1$, $\forall V$ voisinage de $f(x)$, $f^{-1}(V)$ est voisinage de x.

Démonstration. Ici encore, la preuve est une adaptation facile de celle du cas des espaces vectoriels normés. Nous allons démontrer que $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$, ce qui prouvera l'équivalence.
$(i) \Rightarrow (ii)$: soit $A \subset E_2$, A ouvert. On pose $B = f^{-1}(A)$. Pour montrer que B est ouvert, on prend $x_0 \in B$, et on veut démontrer qu'une boule ouverte centrée en x_0 est incluse dans B. Comme A est ouvert et que $f(x_0) \in A$, on sait, par définition, qu'il existe $\varepsilon > 0$ tel que $B_\varepsilon(f(x_0)) \subset A$. De plus, par définition de la continuité, on sait qu'il existe $\delta > 0$ tel que $d_1(x_0, x) < \delta \Rightarrow d_2\left(f(x_0), f(x)\right) < \varepsilon$. Ainsi, si $x \in B_\delta(x_0)$, on a $f(x) \in B_\varepsilon(f(x_0)) \subset A$. Donc $B_\delta(x_0) \subset B$, et B est donc bien un ouvert.
$(ii) \Rightarrow (iii)$: supposons que B est fermé. On a alors

$$\begin{aligned}\left(f^{-1}(B)\right)^C = \{x \in E_1, \quad / \quad f(x) \in B\}^C &= \{x \in E_1, \quad / \quad f(x) \notin B\} \\ &= \left\{x \in E_1, \quad / \quad f(x) \in B^C\right\} = f^{-1}(B^C). \quad (2.4)\end{aligned}$$

Par hypothèse, B^C est ouvert, donc $\left(f^{-1}(B)\right)^C$ aussi. Donc $f^{-1}(B)$ est fermé.
$(iii) \Rightarrow (iv)$: soit $x \in E_1$, et soit V un voisinage de $f(x)$ dans E_2. Alors, par définition d'un voisinage, il existe $\varepsilon > 0$ tel que $B_\varepsilon(f(x)) \subset V$. De plus, $B = B_\varepsilon(f(x))^C$ est fermé, donc, par hypothèse, $f^{-1}(B)$ l'est aussi, d'où $\left(f^{-1}(B)\right)^C$ est un ouvert. Si on reproduit le calcul (2.4) ci-dessus, on obtient aussi que $f^{-1}\left(B^C\right) = \left(f^{-1}(B)\right)^C$ est un ouvert, et donc que

$$f^{-1}\left(B_\varepsilon(f(x))\right) \text{ est ouvert.}$$

Comme de plus $x \in f^{-1}\left(B_\varepsilon(f(x))\right) \subset f^{-1}(V)$, ceci prouve bien que $f^{-1}(V)$ est un voisinage de x.
$(iv) \Rightarrow (i)$: soit $x_0 \in E_1$, et soit $\varepsilon > 0$. Par définition, $B_\varepsilon\left(f(x_0)\right)$ est un voisinage de $f(x_0)$. Donc $f^{-1}\left(B_\varepsilon\left(f(x_0)\right)\right)$ est un voisinage de x_0. Donc, toujours

par définition d'un voisinage, il existe $\delta > 0$ tel que $B_\delta(x_0) \subset f^{-1}(B_\varepsilon(f(x_0)))$. Autrement dit,

$$d_1(x, x_0) < \delta \Rightarrow d_2(f(x), f(x_0)) < \varepsilon.$$

Ceci termine la démonstration. □

De même que dans le cas des espaces vectoriels normés, on peut définir la notion d'uniforme continuité et d'applications lipschitziennes :

Définition 2.19 *Soient (E_1, d_1) et (E_2, d_2) des espaces métriques, $f : E_1 \longrightarrow E_2$. On dit que f est uniformément continue sur E_1 si*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 \quad \textit{tel que} \quad d_1(x, y) < \delta \Rightarrow d_2(f(x), f(y)) < \varepsilon.$$

Définition 2.20 *Soient (E_1, d_1) et (E_2, d_2) des espaces métriques, $f : E_1 \longrightarrow E_2$. On dit que f est lipschitzienne sur E_1 si*

$$\exists M > 0 \quad \textit{tel que} \quad \forall x \in E_1, \quad \forall y \in E_1, \quad d_2(f(x), f(y)) \leq M d_1(x, y).$$

Ici encore, f lipschitzienne implique f uniformément continue qui implique f continue, mais les réciproques sont fausses (voir les contre-exemples dans le chapitre sur les espaces vectoriels normés).

2.4 Distances équivalentes

Contrairement au cas des normes, il y a plusieurs notions d'équivalence de distance.

Définition 2.21 *Soit E un ensemble muni de deux distances d_1 et d_2. On dit que ces distances sont :*

(i) topologiquement équivalentes si elles définissent les mêmes ouverts ;

(ii) uniformément équivalentes si l'application

$$\begin{array}{rccc} I : & (E, d_1) & \longrightarrow & (E, d_2), \\ & x & \longmapsto & x, \end{array}$$

est uniformément continue et que sa réciproque est uniformément continue ;

(iii) Lipschitz-équivalentes s'il existe deux constantes $M \geq m > 0$ telles que

$$\forall x \in E, \quad \forall y \in E, \quad m d_1(x, y) \leq d_2(x, y) \leq M d_1(x, y).$$

Remarque 2.22 *Dans la défintion 2.21, la propriété (i) équivaut exactement à dire que l'application I est continue de (E, d_1) dans (E, d_2) d'inverse continue. Il suffit pour le prouver d'appliquer la proposition 2.18. De plus, la propriété (iii) est exactement équivalente à dire que I et sa réciproque sont lipschitziennes.*

Proposition 2.23 *Dans la définition 2.21, (iii) ⇒ (ii) ⇒ (i).*

Démonstration. Supposons que d_1 et d_2 sont Lipschitz-équivalentes. Alors l'application I vérifie $d_2(I(x), I(y)) = d_2(x, y) \leq M d_1(x, y)$. Donc I est lipschitzienne, donc uniformément continue. Notons maintenant J l'application réciproque de I. On a alors $d_1(J(x), J(y)) = d_1(x, y) \leq \frac{1}{m} d_2(x, y)$. Donc J est lipschitzienne, donc uniformément continue.

Supposons maintenant que I et sa réciproque sont uniformément continues. Alors en particulier I est continue. Si U est un ouvert pour d_2, alors $U = I^{-1}(U)$ est aussi un ouvert pour d_1, d'après la proposition 2.18. Réciproquement, toujours en notant J l'application réciproque de I, qui est continue, si V et un ouvert pour d_1, alors $V = J^{-1}(V)$ est un ouvert pour d_2. $\square$

Notons bien que les réciproques de la proposition 2.23 sont fausse. En effet, dans $\mathbb{R}$, on pose

$$d_1(x, y) = |x - y|, \quad d_2(x, y) = \left|x^3 - y^3\right|, \quad d_3(x, y) = \frac{|x - y|}{1 + |x - y|}.$$

Nous avons déjà vu que d_1, d_2 et d_3 sont des distances. Nous allons montrer que

Lemme 2.24 *Dans $\mathbb{R}$, les distances définies ci-dessus vérifient :*

1. *d_1 et d_2 sont topologiquement équivalentes, mais pas uniformément équivalentes ;*
2. *d_1 et d_3 sont uniformément équivalentes, mais pas lipschitz-équivalentes.*

Démonstration. Commençons pas démontrer le premier point. Pour cela, nous établissons tout d'abord l'encadrement :

$$\forall x \in \mathbb{R}, \quad \forall y \in \mathbb{R}, \quad \frac{1}{2}(x^2 + y^2)|x - y| \leq |x^3 - y^3| \leq \frac{3}{2}(x^2 + y^2)|x - y|. \quad (2.5)$$

On écrit $|x^3 - y^3| = |(x - y)(x^2 + xy + y^2)| = |x - y| \, |x^2 + xy + y^2|$, donc

$$\left(x^2 - |xy| + y^2\right)|x - y| \leq \left|x^3 - y^3\right| \leq \left(x^2 + |xy| + y^2\right)|x - y|.$$

Comme on sait que $|xy| \leq \frac{1}{2}(x^2 + y^2)$, on obtient bien le résultat. Supposons maintenant que U est un ouvert pour d_1. Alors, pour tout $x \in U$, il existe $r > 0$ tel que $B_r^{d_1}(x) =]x - r, x + r[\subset U$. Si $x = 0$, alors la boule de rayon δ de centre x pour d_2 est exactement $B_\delta^{d_2}(x) = \left]-\delta^{1/3}, \delta^{1/3}\right[$. Donc, en prenant $\delta = r^3$, on obtient que $B_\delta^{d_2}(x) \subset U$. Si maintenant $x \neq 0$, alors (2.5) implique que $|x - y| \leq 2|x^3 - y^3|/x^2$, donc que $B_{rx^2/2}^{d_2}(x) \subset B_r^{d_1}(x) \subset U$. Ainsi, U est également un ouvert pour d_2.

Réciproquement, si U est un ouvert pour d_2, alors, pour tout $x \in U$, il existe $r > 0$ tel que $B_r^{d_2}(x) \subset U$. Si $x = 0$, alors $B_r^{d_2}(x) = \left]-r^{1/3}, r^{1/3}\right[= B_{r^{1/3}}^{d_1}(x)$, donc $B_{r^{1/3}}^{d_1}(x) \subset U$. Si $x \neq 0$, alors l'inégalité (2.5) implique, pour tout $\delta > 0$ et tout y tel que $|x - y| < \delta$,

$$|x^3 - y^3| \leq \frac{3}{2}\left(x^2 + (|x| + \delta)^2\right)|x - y| < \frac{3}{2}\left(x^2 + (|x| + \delta)^2\right)\delta.$$

Pour $\delta > 0$ assez petit, ce majorant est plus petit que r. Donc $B_\delta^{d_1}(x) \subset B_r^{d_2}(x) \subset U$. Ainsi, U est un ouvert pour d_1. On a donc prouvé que d_1 et d_2 définissent les mêmes ouverts.

Supposons maintenant que d_1 et d_2 sont uniformément équivalentes. Ceci implique que

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad |x-y| < \delta \Rightarrow \left|x^3 - y^3\right| < \varepsilon.$$

Fixons donc $y = x + \delta/2$, on alors

$$\forall x \in \mathbb{R}, \quad \left(x + \frac{\delta}{2}\right)^3 - x^3 < \varepsilon.$$

Or

$$\left(x + \frac{\delta}{2}\right)^3 - x^3 = \frac{3}{2}x^2\delta + \frac{3}{4}x\delta^2 + \frac{1}{8}\delta^3 \geq \frac{3}{2}x^2\delta.$$

Ainsi, en prenant $x > 0$, on obtient

$$\forall x > 0, \quad 3x^2\delta < 2\varepsilon.$$

Ceci est contradictoire : il suffit pour le voir de choisir $x = \sqrt{2\varepsilon/(3\delta)}$.

Démontrons maintenant le deuxième point. On souhaite donc prouver que I est uniformément continue, donc que

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \text{tel que} \quad |x-y| < \delta \Rightarrow \frac{|x-y|}{1+|x-y|} < \varepsilon.$$

Il suffit pour cela de choisir $\delta = \varepsilon$. En effet, on a alors, si $|x-y| < \delta = \varepsilon$,

$$\frac{|x-y|}{1+|x-y|} < \frac{\varepsilon}{1+|x-y|} \leq \varepsilon.$$

Pour démontrer que la réciproque de I est uniformément continue, on doit prouver que

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \text{tel que} \quad \frac{|x-y|}{1+|x-y|} < \delta \Rightarrow |x-y| < \varepsilon.$$

Pour cela, on choisit $\delta = \min\left(1, \frac{\varepsilon}{2}\right)$. On a alors, si $d_3(x,y) < \delta$,

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

Supposons maintenant que d_1 et d_3 sont lipschitz-équivalentes. Alors il existerait $M > 0$ tel que

$$\forall x \in \mathbb{R}, \quad \forall y \in \mathbb{R}, \quad |x-y| \leq M\frac{|x-y|}{1+|x-y|}.$$

On fixe alors $y = 0$, et on obtient alors que $|x|(1+|x|) \leq M|x|$, donc

$$\forall x > 0, \quad 1 + x \leq M,$$

ce qui est contradictoire (on peut par exemple prendre $x = M$). Donc d_1 et d_3 ne sont pas lipschitz-équivalentes. $\square$

Chapitre 3

Espaces topologiques

3.1 Ouverts, fermés, voisinages

Définition 3.1 *Soit E un ensemble. On appelle topologie de E un ensemble $\mathcal{O}$ de parties de E tel que :*

(i) E et $\emptyset$ appartiennent à $\mathcal{O}$;

(ii) Toute intersection finie d'éléments de $\mathcal{O}$ appartient à $\mathcal{O}$;

(iii) Toute union d'éléments de $\mathcal{O}$ est dans $\mathcal{O}$.

Le couple $(E, \mathcal{O})$ est alors appelé espace topologique. Les éléments de $\mathcal{O}$ sont appelés les ouverts de E.

Définition 3.2 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $A \subset E$. On dit que A est fermé si $A^C \in \mathcal{O}$.*

Par passage au complémentaire dans la définition 3.1, on obtient facilement la propriété suivante :

Proposition 3.3 *Soit $(E, \mathcal{O})$ un espace topologique. Alors*

(i) Toute union finie de fermés est un fermé ;

(ii) Toute intersection de fermés est un fermé.

Exemple 3.4 *Citons des exemples de topologies issus des chapitres précédents :*

- *si E est un espace vectoriel normé, les ouverts associés à sa norme sont une topologie (voir la proposition 1.20).*
- *Si E est un espace métrique, les ouverts définis par sa distance sont une topologie (voir la proposition 2.13).*

Par ailleurs, on aussi les exemples suivants.

- *L'ensemble $\mathcal{P}(E)$ de toutes les parties de E est une topologie. On l'appelle la topologie discrète.*
- *L'ensemble $\{E, \emptyset\}$ est une topologie. On l'appelle la topologie triviale.*

Exercice 3.1 *Soit E un ensemble quelconque. On le munit de la distance*

$$\forall (x,y) \in E^2, \quad d(x,y) = \begin{cases} 0 & si\ x = y, \\ 1 & si\ x \neq y. \end{cases}$$

Démontrer que la topologie associée est la topologie discrète.

Nous définissons ensuite la notion de voisinage de la même façon que pour les espaces métriques et les espaces vectoriels normés :

Définition 3.5 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $x \in E$. Un sous-ensemble V de E est appelé voisinage de x si*

$$\exists U \in \mathcal{O} \quad \text{tel que} \quad x \in U \subset V.$$

Autrement dit, V est voisinage de x si et seulement s'il contient un ouvert qui contient x.

Définition 3.6 *Soit $(E, \mathcal{O})$ un espace topologique. On dit qu'il est séparé si pour tout $x \neq y$ dans E, il existe V voisinage de x et W voisinage de y tels que $V \cap W = \emptyset$.*

Proposition 3.7 *Tout espace métrique est séparé.*

Démonstration. Soit (E, d) un espace métrique. Soient $x \neq y$ des éléments de E. On a alors $d(x, y) = r > 0$, par définition d'une distance. On pose alors

$$V = B_{r/2}(x), \quad W = B_{r/2}(x).$$

Les ensembles V et W sont des boules ouvertes, donc des ouverts. Comme $x \in V$ et que $y \in W$, V est voisinage de x et W voisinage de y. De plus, $V \cap W = \emptyset$. En effet, s'il existait $z \in V \cap W$, on aurait $d(x, z) < r/2$ et $d(y, z) < r/2$. Donc, d'après l'inégalité triangulaire,

$$d(x,y) \leq d(x,z) + d(y,z) < \frac{r}{2} + \frac{r}{2} = r = d(x,y),$$

ce qui est contradictoire. □

Exemple 3.8 *Soit E un ensemble contenant au moins deux éléments. On le munit de la topologie triviale. Alors E n'est pas séparé.*

Démonstration. En effet, pour tout $x \in E$, le seul ouvert qui contient x est E lui-même. Donc si V est voisinage de x, alors $E \subset V$, donc $V = E$. Ainsi, si $x \neq y$, et si V et W sont des voisinages de x et y respectivement, alors $V = W = E$, et on ne peut donc avoir $V \cap W = \emptyset$. □

Corollaire 3.9 *La topologie triviale n'est pas métrisable, c'est-à-dire qu'il n'existe aucune distance dont elle est issue.*

Démonstration. S'il existait une distance d dont la topologie triviale est issue, alors, en appliquant la proposition 3.7, l'espace en question serait séparé. $\square$

Proposition 3.10 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $A \subset E$. Alors A est ouvert si et seulement s'il est voisinage de chacun de ses points.*

Démonstration. Supposons que A est ouvert. Alors pour tout $x \in A$, A contient un ouvert (A lui-même) qui contient x.

Réciproquement, si A est voisinage de chacun de ses points, alors, pour tout $x \in A$, on sait qu'il existe un ouvert U_x tel que $x \in U_x \subset A$. Posons alors

$$U = \bigcup_{x \in A} U_x.$$

Cet ensemble est une union d'ouverts, donc un ouvert. De plus, $U \subset A$ car pour tout $x \in A$, $U_x \subset A$. Enfin, pour tout $x \in A$, $x \in U_x \subset U$. Donc $A \subset U$. D'où $A = U$ est un ouvert. $\square$

3.2 Intérieur et adhérence

Définition 3.11 *Soit $(E, \mathcal{O})$ un espace topologique et soit $A \subset E$. Soit $x \in E$. On dit que x est adhérent à A si tout voisinage V de x vérifie $A \cap V \neq \emptyset$.*

Notons tout de suite que tout point de A est adhérent à A.

Exemple 3.12 *Soit $(E, \|\cdot\|)$ un espace vectoriel normé, $x \in E$ et $r > 0$. On pose $A = B_r(x)$. Alors tout point $y \in E$ tel que $\|x - y\| = r$ est adhérent à A.*

Démonstration. Si V est un voisinage de y, alors V contient un ouvert qui contient y, donc une boule ouverte $B_\delta(y)$, avec $\delta > 0$. Quitte à diminuer δ, on peut supposer $\delta < r$. Posons alors

$$z = \frac{\delta}{2r} x + \left(1 - \frac{\delta}{2r}\right) y,$$

qui vérifie donc

$$\|z - x\| = \left\| \underbrace{\left(1 - \frac{\delta}{2r}\right)}_{>0} (y - x) \right\| = \left(1 - \frac{\delta}{2r}\right) \|x - y\| = \left(1 - \frac{\delta}{2r}\right) r = r - \frac{\delta}{2} < r,$$

et

$$\|z - y\| = \left\| \frac{\delta}{2r}(x - y) \right\| = \frac{\delta}{2r} \|x - y\| = \frac{\delta}{2} < \delta.$$

$\square$

Exercice 3.2 *Dans $\mathbb{R}$, on considère la distance d définie par*

$$d(x, y) = \min(1, |x - y|).$$

Soit $x \in \mathbb{R}$, et $A = B_1(x)$. Démontrer qu'il existe $y \in \mathbb{R}$ tel que $d(x, y) = 1$ mais y n'est pas adhérent à A.

Exemple 3.13 *Dans $\mathbb{R}$ muni de la topologie associée à la norme valeur absolue, on considère $A = [0, 1[$. Alors 1 est adhérent à A.*

Définition 3.14 *Soit $(E, \mathcal{O})$ un espace topologique, et soit $A \subset E$. Soit $x \in E$. On dit que x est intérieur à A si A est voisinage de x.*

Tout point intérieur à A est bien entendu un élément de A.

Exemple 3.15 *Si A est ouvert, alors $\forall x \in A$, x est intérieur à A.*

Exemple 3.16 *Dans un espace métrique, si $A = \overline{B_r}(x)$, avec $r > 0$, alors tout point $y \in B_r(x)$ est intérieur à A.*

Démonstration. Si $y \in B_r(x)$, alors $y \in B_r(x) \subset \overline{B_r}(x) = A$, et $B_r(x)$ est un ouvert, d'après la proposition 2.10. A est donc bien un voisinage de x. □

Exemple 3.17 *Dans $\mathbb{R}$ muni de la topologie associée à la norme valeur absolue, on considère $A = [0, 1[$. Alors tout point de $]0, 1[$ est intérieur à A. Mais 0 n'est pas intérieur à A.*

Définition 3.18 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $A \subset E$.*
- *On appelle adhérence de A, notée $\overline{A}$, l'ensemble des points adhérents à A.*
- *On appelle intérieur de A, notée $\mathring{A}$, l'ensemble des points intérieurs à A.*

Il est donc clair que

$$\mathring{A} \subset A \subset \overline{A}.$$

De plus, $\overline{A}$ est le plus petit fermé contenant A, et $\mathring{A}$ est le plus grand ouvert contenu dans A, autrement dit,

Proposition 3.19 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $A \subset E$.*

(i) $\overline{A}$ est fermé, et si F est un fermé tel que $A \subset F$, alors $\overline{A} \subset F$.

(ii) $\mathring{A}$ est un ouvert, et si U est un ouvert tel que $U \subset A$, alors $U \subset \mathring{A}$.

Démonstration. (i) : supposons F fermé tel que $A \subset F$. Si $x \in \overline{A}$ mais que $x \notin F$, alors $x \in F^C$ qui est un ouvert, donc un voisinage de x. Donc $A \cap F^C \neq \emptyset$. Ce qui est contradictoire. Donc $x \in F$, pour tout $x \in \overline{A}$. D'où $\overline{A} \subset F$.

Il reste à prouver que $\overline{A}$ est fermé. Pour cela, on démontre que $\left(\overline{A}\right)^C$ est ouvert. Soit donc $x \in \left(\overline{A}\right)^C$. Alors, par définition de $\overline{A}$, il existe V voisinage de x tel que $V \cap A = \emptyset$. Donc il existe U ouvert tel que $U \subset V \subset A^C$. Donc

$A \subset U^C$, qui est fermé. Donc, d'après ce qu'on vient de voir, $\overline{A} \subset U^C$. D'où $U \subset \left(\overline{A}\right)^C$, ce qui prouve bien que $\left(\overline{A}\right)^C$ est un ouvert.
(ii) : soit U un ouvert tel que $U \subset A$. Soit $x \in U$. Alors, par définition, A est voisinage de x. Donc $x \in \mathring{A}$. Ainsi, $U \subset \mathring{A}$.

Démontrons à présent que $\mathring{A}$ est un ouvert. Soit donc $x \in \mathring{A}$, alors A est voisinage de x, donc il existe U ouvert tel que $x \in U \subset A$. Donc, d'après ce qui précède, $U \subset \mathring{A}$. Donc $\mathring{A}$ est voisinage de chacun de ses points : c'est un ouvert. □

Remarque 3.20 *Ceci prouve donc que :*
- $\overline{A}$ *est l'intersection de tous les fermés contenant* A*.*
- $\mathring{A}$ *est l'union de tous les ouverts contenus dans* A*.*

En effet, si on note F l'intersection de tous les fermés contenant A, alors, la proposition 3.3 implique que F est fermé, donc $\overline{A} \subset F$. Par ailleurs, $\overline{A}$ est un fermé contenant A, donc $F \subset \overline{A}$.

De même si on note U l'union de tous les ouverts contenus dans A, alors la définition 3.1 implique que U est un ouvert, donc que $U \subset \mathring{A}$. Comme de plus $\mathring{A}$ est lui-même un ouvert, $\mathring{A} \subset U$.

Exemple 3.21 *Donnons quelques exemples simples dans le cas de* $\mathbb{R}$ *muni de la topologie usuelle :*
- *si* $a < b$ *sont des réels, alors* $\overline{[a, b[} = [a, b]$*.*
- *Si* $a < b$ *sont des réels, alors* $\overset{\circ}{\widehat{[a, b]}} =]a, b[$*.*
- *Si* $A = \{x\}$*, alors* $\overline{A} = A$ *et* $\mathring{A} = \emptyset$*.*
- *Si* $A =]0, +\infty[$*, alors* $\overline{A} = [0, +\infty[$ *et* $\mathring{A} = A$*.*
- *Si* $A = [0, +\infty[$*, alors* $\overline{A} = A$ *et* $\mathring{A} =]0, +\infty[$*.*

Corollaire 3.22 *Soit* $(E, \mathcal{O})$ *un espace topologique, et soit* $A \subset E$*. Alors :*
- A *ouvert* $\iff A = \mathring{A}$*,*
- A *fermé* $\iff A = \overline{A}$*.*

Démonstration. Si $A = \mathring{A}$, il est clair que A est ouvert car $\mathring{A}$ l'est. Réciproquement, si A est ouvert, alors A est un ouvert contenu dans A, donc, en appliquant la proposition 3.19, $A \subset \mathring{A}$. Comme par ailleurs, par définition de l'intérieur, $\mathring{A} \subset A$, on a bien que $A = \mathring{A}$.

Si $A = \overline{A}$, alors A est fermé car $\overline{A}$ l'est. Réciproquement, si A est fermé, alors A est un fermé contenant A, donc, en appliquant la proposition 3.19, $\overline{A} \subset A$. Par ailleurs, par définition de l'adhérence, $A \subset \overline{A}$, donc $A = \overline{A}$. □

Définition 3.23 *Soit* $(E, \mathcal{O})$ *un espace topologique. Soit* $A \subset E$*. On dit que* A *est dense dans* E *si tout point de* E *est adhérent à* A *(autrement dit si* $\overline{A} = E$*.)*

Exemple 3.24
- $\mathbb{Q}$ *est dense dans* $\mathbb{R}$ *(muni de la topologie usuelle).*
- E *est dense dans* E*.*

- *Dans l'ensemble des matrices carrées de taille n (muni de la topologie associée à une de ses normes), l'ensemble des matrices inversibles est dense.*
- *On pose $E = C^0([a, b])$ l'ensemble des fonctions continues sur l'intervalle borné $[a, b]$. On munit E de la norme*

$$\|f\|_1 = \int_a^b |f(t)|dt.$$

Alors l'ensemble des fonctions affines par morceaux est dense dans E.

Définition 3.25 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $A \subset E$. On appelle frontière de A l'ensemble* $\mathrm{Fr}(A) = \overline{A} \setminus \mathring{A}$.

3.3 Topologie induite et topologie produit

Proposition 3.26 *Soit $(E, \mathcal{O})$ un espace topologique. Soit $F \subset E$. Alors $\mathcal{O}' = \{U \cap F, \quad U \in \mathcal{O}\}$ est une topologie sur F*

Notation 3.27 *Sous ces hypothèses, $(F, \mathcal{O}')$ est appelé sous-espace topologique de $(E, \mathcal{O})$. La topologie $\mathcal{O}'$ est appelée topologie induite de E sur F.*

Démonstration. Vérifions les axiomes de définition d'une toplogie (définition 3.1) :
(i) : $\emptyset = F \cap \emptyset$, donc $\emptyset \in \mathcal{O}'$. Et $F = F \cap E$, donc $F \in \mathcal{O}'$.
(ii) : soit $(U_i)_{1 \leq i \leq p}$ une famille finie d'éléments de $\mathcal{O}'$. Alors, pour tout i, il existe $V_i \in \mathcal{O}$ tel que $U_i = F \cap V_i$. Donc

$$\bigcap_{i=1}^{p} U_i = \bigcap_{i=1}^{p} (V_i \cap F) = \left(\bigcap_{i=1}^{p} V_i \right) \cap F = V \cap F,$$

avec $V \in \mathcal{O}$. Donc

$$\bigcap_{i=1}^{p} U_i \in \mathcal{O}'.$$

(iii) : soit $(U_i)_{i \in I}$ une famille d'éléments de $\mathcal{O}'$. Alors, pour tout $i \in I$, il existe $V_i \in \mathcal{O}$ tel que $U_i = F \cap V_i$. Donc

$$\bigcup_{i \in I} U_i = \bigcup_{i \in I} (V_i \cap F) = \left(\bigcup_{i \in I} V_i \right) \cap F = V \cap F,$$

avec $V \in \mathcal{O}$. Donc

$$\bigcup_{i \in I} U_i \in \mathcal{O}'.$$

□

Proposition 3.28 *Soit (E, d) est un espace métrique. On note $\mathcal{O}$ la topologie associée à d. Soit $F \subset E$. Alors la topologie induite sur F est celle associée à la distance $d_{|F \times F}$ restreinte à F.*

Démonstration. Soit $U \in \mathcal{O}'$. On souhaite montrer que U est un ouvert de F pour la distance d, donc que pour tout $x \in U$, U est voisinage de x dans (F, d). Autrement dit qu'il existe $r > 0$ tel que $B_r^F(x) \subset U$, avec

$$B_r^F(x) = \{y \in F, \quad d(x,y) < r\} = B_r^E(x) \cap F,$$

et

$$B_r^E(x) = \{y \in E, \quad d(x,y) < r\}.$$

Soit donc $x \in U$. On sait que $U = V \cap F$, avec V ouvert de (E, d). Comme $x \in V$, il existe donc $r > 0$ tel que $B_r^E(x) \subset V$. D'où l'on déduit que

$$B_r^F(x) = B_r^E(x) \cap F \subset V \cap F = U,$$

donc U est bien un ouvert de (F, d).

Réciproquement, si U est un ouvert pour (F, d), alors on sait que, pour tout $x \in U$, il existe $r_x > 0$ tel que $B_{r_x}^F(x) \subset U$, autrement dit que

$$B_{r_x}^E(x) \cap F \subset U.$$

On pose alors

$$V = \bigcup_{x \in U} B_{r_x}^E(x),$$

qui est une union d'ouverts de (E, d), donc un ouvert. On a alors

$$U = V \cap F.$$

En effet, si $x \in U$, alors $x \in B_{r_x}^F(x) = B_{r_x}^E(x) \cap F \subset V \cap F$. Réciproquement, si $x \in V \cap F$, il existe $y \in U$ tel que $x \in B_{r_y}^E(y) \cap F = B_{r_y}^F(y) \subset U$. Comme V est un ouvert de E, on a bien que $U \in \mathcal{O}'$. □

Proposition 3.29 *Soient $(E_i, \mathcal{O}_i)_{1 \le i \le n}$ des espaces topologiques. On pose $E = E_1 \times E_2 \times \cdots \times E_n$, et on définit*

$$\mathcal{O} = \{U \subset E, \ / \ \forall x \in U, \quad \exists (U_1, U_2, \ldots U_n) \in \mathcal{O}_1 \times \mathcal{O}_2 \times \cdots \mathcal{O}_n, \\ \text{avec} \quad x \in U_1 \times U_2 \times \cdots U_n \subset U\}.$$

Alors $\mathcal{O}$ est une topologie sur E.

Démonstration. On prouve successivement les propriétés (i), (ii) et (iii) de la définition 3.1 :
(i) : $\emptyset \in \mathcal{O}$ car $\emptyset = \emptyset \times \emptyset \times \cdots \times \emptyset$. De plus, $E \in \mathcal{O}$ car $E_1 \times E_2 \times \cdots \times E_n \subset E$.
(ii) Si $(V^i)_{i \in I}$ est une famille d'éléments de $\mathcal{O}$, on pose

$$V = \bigcup_{i \in I} V^i.$$

Alors, si $x \in V$, on sait qu'il existe $i_0 \in I$ tel que $x \in V^{i_0}$. De plus, par définition de $\mathcal{O}$, il existe $(U_1^{i_0}, U_2^{i_0}, \dots U_n^{i_0}) \in \mathcal{O}_1 \times \mathcal{O}_2 \times \cdots \mathcal{O}_n$ tel que

$$x \in U_1^{i_0} \times U_2^{i_0} \times \cdots \times U_n^{i_0} \subset V^{i_0} \subset V.$$

Donc $V \in \mathcal{O}$.
(iii) Si $(V^i)_{1\leq i\leq p}$ est une famille finie d'éléments de $\mathcal{O}$. On pose

$$V = \bigcap_{1\leq i\leq p} V^i.$$

Si $x \in V$, alors on sait que, pour tout $1 \leq i \leq p$,

$$\exists \left(U_1^i, U_2^i, \dots U_n^i\right) \in \mathcal{O}_1 \times \mathcal{O}_2 \times \cdots \mathcal{O}_n, \quad \text{avec} \quad x \in U_1^i \times U_2^i \times \cdots U_n^i \subset V^i.$$

Ainsi,

$$x \in \left[\left(\bigcap_{i=1}^p U_1^i\right) \times \left(\bigcap_{i=1}^p U_2^i\right) \times \cdots \times \left(\bigcap_{i=1}^p U_n^i\right)\right]$$
$$= \bigcap_{i=1}^p \left(U_1^i \times U_2^i \times \cdots U_n^i\right) \subset \bigcap_{i=1}^p V^i = V,$$

donc $V \in \mathcal{O}$. □

3.4 Continuité

Définition 3.30 *Soient $(E, \mathcal{O})$ et $(F, \mathcal{O}')$ des espaces topologiques. On considère $f : E \longrightarrow F$, et $x \in E$. On dit que f est continue en x si pour tout V voisinage de $f(x)$, $f^{-1}(V)$ est voisinage de x.*

Définition 3.31 *Soient $(E, \mathcal{O})$ et $(F, \mathcal{O}')$ des espaces topologiques. On considère $f : E \longrightarrow F$. On dit que f est continue sur E si pour tout ouvert U de F, $f^{-1}(U)$ est un ouvert de E.*

Proposition 3.32 *Soient $(E, \mathcal{O})$ et $(F, \mathcal{O}')$ des espaces topologiques. On considère $f : E \longrightarrow F$. Alors on a équivalence entre les propriétés suivantes :*

(i) f est continue sur E.

(ii) $\forall x \in E$, f est continue en x.

(iii) Pour tout fermé A de F, $f^{-1}(A)$ est un fermé de E.

Démonstration.
$(i) \Rightarrow (ii)$: supposons que f est continue sur E. Soit $x \in E$. Soit V un voisinage de $f(x)$. Alors il existe U ouvert de F tel que $f(x) \in U \subset V$. Donc

$$x \in f^{-1}(U) \subset f^{-1}(V).$$

Or, par hypothèse, $f^{-1}(U)$ est un ouvert. Donc $f^{-1}(V)$ est bien un voisinage de x, et f est donc continue en x.
$(ii) \Rightarrow (i)$: si f est continue en tout point x de E, et si U est un ouvert de F, supposons que $x \in f^{-1}(U)$. Alors U est voisinage de $f(x)$, puisque $f(x) \in U$ et que U est ouvert. Donc $f^{-1}(U)$ est voisinage de x. Ceci vaut pour tout $x \in f^{-1}(U)$, donc $f^{-1}(U)$ est bien un ouvert de E.
$(i) \Rightarrow (iii)$ si A est un fermé, alors A^C et un ouvert. Donc $f^{-1}(A^C)$ est un ouvert. Or $f^{-1}(A^C) = f^{-1}(A)^C$, car

$$f^{-1}(A^C) = \left\{x \in E, \quad f(x) \in A^C\right\} = \{x \in E, \quad f(x) \notin A\}$$
$$= \{x \in E, \quad f(x) \in A\}^C .$$

Donc $f^{-1}(A)^C$ est un ouvert, donc $f^{-1}(A)$ est un fermé.
$(iii) \Rightarrow (i)$: si A est ouvert, on applique le même raisonnement que précédemment, et on a $f^{-1}(A)^C = f^{-1}(A^C)$ qui est fermé car A^C est fermé, donc $f^{-1}(A)$ est un ouvert. □

Définition 3.33 *Soient $(E, \mathcal{O})$ et $(F, \mathcal{O}')$ des espaces topologiques. On considère $f : E \longrightarrow F$. On dit que f est un homéomorphisme si f est bijective, continue, et que f^{-1} est continue.*

Notation 3.34 *Dans ce cas, on dit que les espaces topologiques $(E, \mathcal{O})$ et $(F, \mathcal{O}')$ sont homéomorphes.*

Exemple 3.35 *On considère $E = \mathbb{R}$ et $F = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, munis de la topologie associée à la distance usuelle $d(x, y) = |x - y|$. L'application*

$$\begin{array}{rccl} f : & \mathbb{R} & \longrightarrow & \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[, \\ & x & \longmapsto & \operatorname{Arctan}(x), \end{array}$$

est un homéomorphisme. Plus généralement, tout intervalle ouvert de $\mathbb{R}$ muni de la distance usuelle est homéomorphe à $\mathbb{R}$ muni de la distance ususelle.

Démonstration. En effet, il est clair que f est continue. De plus, f est dérivable et sa dérivée est égale à

$$f'(x) = \frac{1}{1 + x^2} > 0, \quad \forall x \in \mathbb{R}.$$

Donc f est strictement croissante, donc injective. Elle est surjective car si $y \in F$, alors $\tan(y) \in \mathbb{R}$ vérifie $f(x) = y$. Enfin, son application réciproque est

$$\begin{array}{rccl} f : & \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[& \longrightarrow & \mathbb{R}, \\ & x & \longmapsto & \tan(x), \end{array}$$

qui est bien continue. □

Ainsi, $\mathbb{R}$ muni de sa distance usuelle est homéomorphe à $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ muni de cette même distance.

Enfin, démontrons que la composée de deux applications continues est continue :

Proposition 3.36 *Soient $(E, \mathcal{O})$, $(F, \mathcal{O}')$, $(G, \mathcal{O}'')$ des espaces topologiques. Soient deux applications $f : E \longrightarrow F$ et $g : F \longrightarrow G$ qui sont continues. Alors l'application*

$$g \circ f : E \longrightarrow G$$

est continue.

Démonstration. Soit U un ouvert de G. On a alors

$$\begin{aligned}(g \circ f)^{-1}(U) &= \{x \in E, \quad g(f(x)) \in U\}, \\ &= \left\{x \in E, \quad f(x) \in g^{-1}(U)\right\} = f^{-1}\left(g^{-1}(U)\right).\end{aligned}$$

Or $f^{-1}(U)$ est un ouvert car f est continue. Donc $g^{-1}\left(f^{-1}(U)\right)$ est un ouvert car g est continue. Donc $g \circ f$ est bien continue. □

Remarque 3.37 *La proposition ci-dessus admet également un version locale : si $x \in E$, et si f est continue en x et g continue en $f(x)$, alors $g \circ f$ est continue en x. Nous laissons la preuve en exercice.*

3.5 Convergence de suites

Définition 3.38 *Soit $(E, \mathcal{O})$ un espace topologique, et soit $(x_n)_{n \in \mathbb{N}}$ une suite à valeurs dans E. Soit $x \in E$. On dit que $(x_n)_{n \in \mathbb{N}}$ converge vers x si*

$$\forall V \text{ voisinage de } x, \quad \exists n_0 \in \mathbb{N} \text{ tel que } \forall n \geq n_0, \quad x_n \in V.$$

Notation 3.39 *On écrit alors $x_n \underset{n \to +\infty}{\longrightarrow} x$.*

Dans un espace métrique (donc a fortiori dans un espace vectoriel normé) cette notion correspond à la notion usuelle de convergence :

Proposition 3.40 *Soit (E, d) un espace métrique, et soit $(x_n)_{n \in \mathbb{N}}$ une suite de E. Alors*

$$x_n \underset{n \to +\infty}{\longrightarrow} x \iff \lim_{n \to +\infty} d(x, x_n) = 0.$$

Démonstration. Supposons tout d'abord que $x_n \underset{n \to +\infty}{\longrightarrow} x$. Soit alors $\varepsilon > 0$. Alors $B_\varepsilon(x)$ est un voisinage de x. Donc, en appliquant la définition 3.38, on sait qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad x_n \in B_\varepsilon(x).$$

Autrement dit,

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N}, \quad \text{tel que} \quad \forall n \geq n_0, \quad x_n \in d(x, x_n) < \varepsilon.$$

C'est-à-dire que $\lim\limits_{n \to +\infty} d(x, x_n) = 0$.

Réciproquement, si $\lim_{n\to+\infty} d(x, x_n) = 0$, et si V est un voisinage de x, alors il existe $\varepsilon > 0$ tel que $B_\varepsilon(x) \subset V$. On sait qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad d(x, x_n) < \varepsilon,$$

c'est-à-dire

$$\forall n \geq n_0, \quad x_n \in B_\varepsilon(x) \subset V.$$

Donc $x_n \underset{n\to+\infty}{\longrightarrow} x$. $\square$

Proposition 3.41 *Soit $(E, \mathcal{O})$ un espace topologique séparé. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. Si elle converge, sa limite est unique. On la note $\lim_{n\to+\infty} x_n$.*

Démonstration. Procédons par l'absurde : supposons que la suite $(x_n)_{n\in\mathbb{N}}$ a deux limites $x \neq y$. Alors, comme E est séparé, il existe V voisinage de x et W voisinage de y tels que $V \cap W = \emptyset$. Par ailleurs, puisque la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x, on sait que

$$\exists n_0 \in \mathbb{N} \text{ tel que } \forall n \geq n_0, \quad x_n \in V.$$

De même,

$$\exists n_1 \in \mathbb{N} \text{ tel que } \forall n \geq n_1, \quad x_n \in W.$$

Donc, si on choisit $n \geq \max(n_0, n_1)$ on obtient $x_n \in V \cap W$, ce qui est contradictoire puisque cet ensemble est vide. $\square$

Attention, si l'espace n'est pas séparé, la limite n'est en général pas unique, comme le montre l'exercice suivant :

Exercice 3.3 *On munit $\mathbb{R}$ de la topologie triviale. Soit $(x_n)_{n\in\mathbb{N}}$ une suite réelle. Montrer que, pour cette topologie, on a*

$$\forall x \in \mathbb{R}, \quad x_n \underset{n\to+\infty}{\longrightarrow} x.$$

Proposition 3.42 *Soient $(E, \mathcal{O})$ et $(E, \mathcal{O}')$ des espaces topologiques, $x \in E$ et soit $f : E \to F$ continue en x. Si $(x_n)_{n\in\mathbb{N}}$ est une suite de E, alors*

$$x_n \underset{n\to+\infty}{\longrightarrow} x \qquad \Rightarrow \qquad f(x_n) \underset{n\to+\infty}{\longrightarrow} f(x).$$

Démonstration. Supposons donc que $(x_n)_{n\in\mathbb{N}}$ soit une suite de E qui converge vers x. Soit V un voisinage de $f(x)$ dans F. Par continuité de f en x, on sait que $W = f^{-1}(V)$ est un voisinage de x. Donc, puisque la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x, on sait que

$$\exists n_0 \in \mathbb{N} \text{ tel que } \forall n \geq n_0, \quad x_n \in W.$$

Ceci équivaut très exactement, étant donné la définition de $W = f^{-1}(V)$, à

$$\exists n_0 \in \mathbb{N} \text{ tel que } \forall n \geq n_0, \quad f(x_n) \in V.$$

C'est-à-dire que $f(x_n) \underset{n\to+\infty}{\longrightarrow} f(x)$. □

Ce résultat autorise bien entendu un des deux espaces E ou F à être non séparé. Cependant, comme le montre l'exercice suivant, si F est séparé (et que f est injective), alors E aussi.

Exercice 3.4 *Soient $(E, \mathcal{O}$ et $(F, \mathcal{O}')$ des espaces topologiques. On suppose que chacun contient au moins deux points, et qu'il existe $f : E \to F$ continue injective.*

1. *Démontrer que si F est séparé, alors E est séparé.*
2. *Donner un exemple où E est séparé, mais pas F.*

Dans le cas d'espaces métriques, on a un résultat plus précis :

Proposition 3.43 *Soient E et F des espaces métriques. Soit f une application de E dans F et $x \in E$. Alors on a équivalence entre les deux propriétés :*

(i) f est continue en x.

(ii) f transforme toute suite convergente vers x en une suite convergente vers $f(x)$.

Démonstration. La preuve de $(i) \Rightarrow (ii)$ est une application simple de la proposition 3.42. Démontrons donc sa réciproque par l'absurde : on suppose que (ii) est vraie, mais que f n'est pas continue en x. Alors

$$\exists \varepsilon > 0, \text{ tel que } \forall \delta > 0, \exists y \in B_\delta(x), \text{ tel que } f(y) \notin B_\varepsilon(f(x)).$$

Soit donc un tel $\varepsilon > 0$. En choisissant $\delta = \frac{1}{n+1}$, on construit une suite $(y_n)_{n\in\mathbb{N}}$ de E telle que

$$\forall n \in \mathbb{N}, \quad y_n \in B_{\frac{1}{n+1}}(x), \quad f(y_n) \notin B_\varepsilon(f(x)).$$

Autrement dit,

$$\forall n \in \mathbb{N}, \quad d_E(y_n, x) < \frac{1}{n+1}, \quad d_F(f(y_n), f(x)) > \varepsilon.$$

Ainsi, $y_n \underset{n\to+\infty}{\longrightarrow} x$, mais $(f(y_n))_{n\in\mathbb{N}}$ ne converge pas vers $f(x)$, et on obtient bien une contradiction. □

Exercice 3.5 *Soient E et F des espaces métriques. Soit f une application de E dans F. Démontrer qu'on a alors équivalence entre*

(i) f est continue sur E.

(ii) f transforme toute suite convergente en une suite convergente.

Définition 3.44 *Soit E un ensemble. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. On appelle suite extraite de $(x_n)_{n\in\mathbb{N}}$ une suite de la forme*

$$y_n = x_{\varphi(n)},$$

où $\varphi : \mathbb{N} \to \mathbb{N}$ est une application strictement croissante.

Exemple 3.45 *La suite* $(y_n)_{n\in\mathbb{N}}$ *définie par* $y_n = x_n$ *est une suite extraite de* $(x_n)_{n\in\mathbb{N}}$. *De même pour les suites définies par :*

$$y_n = x_{2n}, \quad y_n = x_{2n+1}, \quad \text{ou} \quad y_n = x_{2^n}.$$

En revanche, la suite définie par

$$y_n = \begin{cases} x_n & \text{si } n \text{ est pair,} \\ x_{n-1} & \text{sinon,} \end{cases}$$

n'est pas une suite extraite de $(x_n)_{n\in\mathbb{N}}$, *car l'application* φ *qui correspond n'est pas strictement croissante.*

Proposition 3.46 *Soit* $(E, \mathcal{O})$ *un espace topologique, et soit* $(x_n)_{n\in\mathbb{N}}$ *une suite à valeurs dans* E. *Si elle converge vers* $x \in E$, *alors toute suite extraite de* $(x_n)_{n\in\mathbb{N}}$ *converge vers* x.

Démonstration. Remarquons tout d'abord que, si $\varphi : \mathbb{N} \to \mathbb{N}$ est strictement croissante, alors $\forall n \in \mathbb{N}, \quad \varphi(n) \geq n$. Ceci peut se démontrer par récurrence : on a évidemment $\varphi(0) \geq 0$, et si $\varphi(n) \geq n$, alors $\varphi(n+1) > \varphi(n) \geq n$, donc, puisque $\varphi(n+1)$ est un entier, $\varphi(n+1) \geq n+1$.

Supposons maintenant que $(x_n)_{n\in\mathbb{N}}$ converge vers x, et que $(y_n)_{n\in\mathbb{N}}$ est une suite extraite de $(x_n)_{n\in\mathbb{N}}$. Alors il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que $\forall n \in \mathbb{N}, \quad y_n = x_{\varphi(n)}$. Soit alors V un voisinage de x. On sait donc que

$$\exists n_0 \in \mathbb{N}, \text{ tel que } \forall n \geq n_0, \quad x_n \in V.$$

Donc, en utilisant la propriété ci-dessus, on sait que si $n \geq n_0$, alors $\varphi(n) \geq n_0$, et donc $y_n = x_{\varphi(n)} \in V$. Ceci prouve donc que

$$\forall n \geq n_0, \quad y_n \in V.$$

Donc la suite $(y_n)_{n\in\mathbb{N}}$ converge bien vers x. □

Définition 3.47 *Soit* $(E, \mathcal{O})$ *un espace topologique. Soit* $(x_n)_{n\in\mathbb{N}}$ *une suite à valeurs dans* E. *On appelle* <u>*valeur d'adhérence*</u> *de* $(x_n)_{n\in\mathbb{N}}$ *tout* $x \in E$ *tel qu'il existe une suite extraite de* $(x_n)_{n\in\mathbb{N}}$ *qui converge vers* x.

Remarque 3.48 *Un point de l'adhérence de l'ensemble* $\{x_n, \ n \in \mathbb{N}\}$ *n'est en général pas une valeur d'adhérence de la suite* $(x_n)_{n\in\mathbb{N}}$. *Par exemple, si* $(x_n)_{n\in\mathbb{N}}$ *converge vers une limite* x *unique, il s'agit de sa seule valeur d'adhérence. Alors que l'ensemble*

$$\overline{\{x_n, \ n \in \mathbb{N}\}}$$

contient chaque terme x_n *de la suite.*

Proposition 3.49 *Soit* E *un espace métrique. Soit* $(x_n)_{n\in\mathbb{N}}$ *une suite à valeurs dans* E. *Soit* $x \in E$. *Alors* x *est valeur d'adhérence de* $(x_n)_{n\in\mathbb{N}}$ *si et seulement si*

$$\forall \varepsilon > 0, \quad \forall n \in \mathbb{N}, \quad \exists p \geq n \text{ tel que } d(x, x_n) < \varepsilon.$$

Démonstration. Supposons que x soit valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$. Alors on sait qu'il existe $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que la suite $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ converge vers x. Soit alors $\varepsilon > 0$, et $n \in \mathbb{N}$. On sait qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall q \geq n_0, \quad d\left(x, x_{\varphi(q)}\right) < \varepsilon.$$

Donc, en prenant $p = \varphi(\max(n, n_0))$, on a bien $p \geq \varphi(n) \geq n$, et $p \geq \varphi(n_0) \geq n_0$, donc $d(x, x_p) < \varepsilon$.

Réciproquement, supposons que

$$\forall \varepsilon > 0, \quad \forall n \in \mathbb{N}, \quad \exists p \geq n \text{ tel que } d(x, x_p) < \varepsilon.$$

On construit par récurrence une application $\varphi : \mathbb{N} \to \mathbb{N}$ strictement croissante telle que

$$d(x, x_{\varphi(n)}) < \frac{1}{n+1}.$$

Pour $n = 0$, il suffit d'appliquer la propriété ci-dessus avec $\varepsilon = 1$ et $n = 0$. On sait donc qu'il existe $p_0 \in \mathbb{N}$ tel que $d(x, x_{p_0}) < 1$. On pose $\varphi(0) = p_0$. Si maintenant on suppose que $\varphi(n)$ a été construit, on applique la propriété ci-dessus avec $\varepsilon = 1/(n+2)$ et $\varphi(n)+1$ à la place de n. Ainsi, il existe $p_{n+1} \geq \varphi(n)+1$ tel que $d(x, x_{p_{n+1}}) < 1/(n+2)$. On pose $\varphi(n+1) = p_{n+1}$. L'application φ ainsi construite est donc bien strictement croissante, et la suite extraite correspondante converge bien vers x. $\square$

Proposition 3.50 *Soit E un espace métrique, et $A \subset E$. Alors on a équivalence entre*

(i) A est fermé.

(ii) Pour toute suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans A et convergente, $\lim\limits_{n\to+\infty} x_n \in A$.

Démonstration. Montrons d'abord que $(i) \Rightarrow (ii)$. Supposons donc que A est fermé, et soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans A qui converge vers $x \in E$. Si $x \notin A$, alors $x \in A^C$ qui est ouvert, donc voisinage de x. Donc il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, $x_n \in A^C$. Ce qui est contradictoire. Donc $x \in A$.

Montrons maintenant que $(ii) \Rightarrow (i)$. Si A n'est pas fermé, alors $A \subsetneq \overline{A}$, d'après le corollaire 3.22. Donc il existe $x \in \overline{A}$ tel que $x \notin A$. En particulier, par définition de l'adhérence,

$$\forall \varepsilon > 0, \quad B_\varepsilon(x) \cap A \neq \emptyset.$$

On choisit donc $\varepsilon = 1/n$, pour $n \geq 1$. On sait alors que

$$\forall n \in \mathbb{N}^*, \quad \exists x_n \in A, \quad x_n \in B_{1/n}(x).$$

En particulier, $x_n \in A$ et $d(x, x_n) \leq 1/n$. Donc $(x_n)_{n\in\mathbb{N}^*}$ est une suite à valeurs dans A qui converge vers $x \notin A$. Ceci est contradictoire, donc A est bien un fermé. $\square$

Remarque 3.51 *Dans un espace topologique,* $(i) \Rightarrow (ii)$ *reste vrai, mais la réciproque peut être fausse, comme le montre l'exercice suivant.*

Exercice 3.6 *Dans* $\mathbb{R}$*, on considère* $\mathcal{O} \subset \mathcal{P}(E)$ *défini par*

$$A \in \mathcal{O} \iff \big(A^C = \mathbb{R} \text{ ou } A^C \text{ est fini ou dénombrable.}\big)$$

1. *Démontrer que* $\mathcal{O}$ *est une topologie sur* $\mathbb{R}$.
2. *On suppose que* $A = [0,1]$*. Démontrer que, pour la topologie* $\mathcal{O}$*,* $\overline{A} = \mathbb{R}$.
3. *Démontrer que, pour la topologie* $\mathcal{O}$*,*

$$\Big(x_n \underset{n\to+\infty}{\longrightarrow} x\Big) \Rightarrow \Big(\{n \in \mathbb{N} \quad \text{tel que} \quad x_n = x\} \text{ est infini.}\Big)$$

4. *En déduire que si* $(x_n)_{n\in\mathbb{N}}$ *est une suite convergente à valeurs dans* $A = [0,1]$*, alors sa limite est dans* A.

Chapitre 4

Espaces compacts

4.1 Espaces topologiques compacts

Définition 4.1 *Soit E un ensemble, et soit $\mathcal{R} = (O_i)_{i\in I}$ une famille de parties de E. On dit que $\mathcal{R}$ est un recouvrement de E si*

$$E \subset \bigcup_{i\in I} O_i.$$

Définition 4.2 *Soit E un ensemble et $\mathcal{R} = (O_i)_{i\in I}$ un recouvrement de E. On appelle sous-recouvrement de $\mathcal{R}$ un ensemble de la forme $\mathcal{S} = (O_i)_{i\in J}$, avec $J \subset I$, tel que $\mathcal{S}$ soit un recouvrement de E :*

$$E \subset \bigcup_{i\in J} O_i.$$

Définition 4.3 *Soit E un espace topologique. On dit que E est compact s'il est séparé et si de tout recouvrement d'ouverts de E, on peut extraire un sous-recouvrement fini.*

Proposition 4.4 *Soit E un espace topologique séparé. Alors les deux propriétés suivantes sont équivalents :*

(i) E est compact.

(ii) Pour toute famille $(F_i)_{i\in I}$ de fermés non vides stable par intersection finie, $\bigcap_{i\in I} F_i \neq \emptyset$.

Démonstration. $(i) \Rightarrow (ii)$: soit $(F_i)_{i\in I}$ une famille de fermés non vides, stable par intersection finie. Supposons que $\bigcap_{i\in I} F_i = \emptyset$. Alors

$$E = \left(\bigcap_{i\in I} F_i\right)^C = \bigcup_{i\in I} F_i^C.$$

Comme E est compact, on sait donc qu'il existe un sous-recouvrement fini de E. Donc il existe $J \subset I$, J fini, tel que

$$E \subset \bigcup_{i\in J} F_i^C, \quad \text{donc} \quad F = \bigcap_{i\in J} F_i = \emptyset.$$

Or ceci est contradictoire car F est dans la famille $(F_i)_{i\in I}$, dont aucun élément n'est l'ensemble vide.
$(ii) \Rightarrow (i)$: soit $\mathcal{R} = (O_i)_{i\in I}$ un recouvrement d'ouverts de E. Supposons qu'on ne puisse pas en extraire un sous-recouvrement fini. Alors pour tout $J \subset I$, J fini, on a

$$E \neq \bigcup_{i\in J} O_i, \quad \text{d'où} \quad \bigcap_{i\in J} O_i^C \neq \emptyset.$$

On pose donc, pour tout J fini inclus dans I,

$$F_J = \bigcap_{i\in J} O_i^C.$$

La famille $(F_J)_{\substack{J\subset I,\\ J \text{ fini}}}$ est donc une famille de fermés non vides, stable par intersection finie. Donc

$$\bigcap_{\substack{J\subset I,\\ J \text{ fini}}} F_J \neq \emptyset.$$

Or cet ensemble est égal à

$$\bigcap_{\substack{J\subset I,\\ J \text{ fini}}} F_J = \bigcap_{\substack{J\subset I,\\ J \text{ fini}}} \left(\bigcap_{i\in J} O_i^C\right) = \bigcap_{i\in I} O_i^C = \left(\bigcup_{i\in I} O_i\right)^C = E^C = \emptyset,$$

ce qui est contradictoire. □

Proposition 4.5 *Soit E un espace topologique compact, et soit F un fermé de E. Alors F est compact (pour la topologie induite sur F).*

Démonstration. Démontrons d'abord que F est séparé. Si x et y sont deux points distincts de F, comme E est séparé par hypothèse, on sait qu'il existe V voisinage de x dans E et W voisinage de y dans E tels que $V \cap W = \emptyset$. Posons alors

$$V' = V \cap F, \quad W' = W \cap F.$$

Alors V' et W' sont des voisinages de x et y, respectivement, dans F. De plus,

$$V' \cap W' \subset V \cap W = \emptyset.$$

Donc F est bien séparé.

Soit maintenant un recouvrement $\mathcal{R} = (O_i)_{i \in I}$ d'ouverts de F. Alors

$$\forall i \in I, \quad O_i = F \cap U_i, \quad U_i \text{ ouvert de } E.$$

On considère donc la famille d'ouverts $F^C \cup (U_i)_{i \in I}$. Il s'agit d'un recouvrement de E, car

$$E = F^C \cup F \subset F^C \cup \left(\bigcup_{i \in I} O_i \right) \subset F^C \cup \left(\bigcup_{i \in I} U_i \right).$$

De plus, chaque élément de cette famille est un ouvert de E : les U_i par définition, et F^C car F est fermé. Comme E est compact, on peut donc en extraire un sous-recouvrement fini : il existe $J \subset I$, J fini, tel que

$$E = F^C \cup \left(\bigcup_{i \in J} U_i \right).$$

Donc

$$F \subset \bigcup_{i \in J} U_i, \quad \text{d'où} \quad F \subset \bigcup_{i \in J} O_i.$$

Donc F est bien compact. □

Remarque 4.6 *Notons que dans l'énoncé ci-dessus, on a pris soin d'utiliser la topologie induite pour caractériser la compacité de F. Cependant, cela revient au même si on utilise la topologie de E. En effet, on a équivalence entre*

(a) De tout recouvrement de F par une famille d'ouverts de F, on peut extraire un recouvrement fini.

(b) De tout recouvrement de F par une famille d'ouverts de E, on peut extraire un recouvrement fini.

Démonstration. Montrons d'abord que $(a) \Rightarrow (b)$: soit $(O_i)_{i \in I}$ une famille d'ouverts de E qui recouvre F. Alors, en posant $U_i = O_i \cap F$, on obtient une famille d'ouverts de F, telle que

$$\bigcup_{i \in I} U_i = F.$$

Comme F vérifie (a), on en extrait un sous-recouvrement fini. Donc il existe $J \subset I$, J fini, tel que

$$F = \bigcup_{i \in J} U_i = \bigcup_{i \in J} (F \cap O_i), \quad \text{c'est-à-dire} \quad F \subset \bigcup_{i \in J} O_i.$$

Donc F vérifie bien (b).

Montrons maintenant la réciproque : soit $(U_i)_{i\in I}$ une famille d'ouverts de F qui recouvre F. Alors $U_i = F \cap O_i$, avec O_i ouvert de E. Et on a

$$F = \bigcup_{i\in I} U_i = \bigcup_{i\in I} (F \cap O_i), \quad \text{donc} \quad F \subset \bigcup_{i\in I} O_i.$$

Donc, en appliquant (b), on en déduit qu'il existe un sous-recouvrement de F par un sous-ensemble fini de $(O_i)_{i\in I}$. Donc il existe $J \subset I$, J fini, tel que

$$F \subset \bigcup_{i\in J} O_i, \quad \text{d'où} \quad F = \bigcup_{i\in J} (O_i \cap F) = \bigcup_{i\in J} U_i,$$

d'où l'on déduit que F vérifie (a). □

Définition 4.7 *Soit E un espace topologique, et soit $A \subset E$. On dit que que A est relativement compact si son adhérence $\overline{A}$ est compacte.*

Proposition 4.8 *Soit E un espace topologique séparé, et soit $A \subset E$ un compact. Alors A est fermé.*

Démonstration. Nous allons prouver que A^C est un ouvert : soit $x \in A^C$. Pour tout $a \in A$, $x \neq a$ par définition, donc il existe V_a ouvert de E contenant a et W_a ouvert de E contenant x tels que

$$\forall a \in A, \quad V_a \cap W_a = \emptyset.$$

De plus,

$$A \subset \bigcup_{a\in A} V_a,$$

donc $(A \cap V_a)_{a\in A}$ est un recouvrement de A. De plus, chaque élément de cette famille est un ouvert de A, donc on peut en extraire un sous-recouvrement fini : il existe $a_1, a_2, \ldots, a_p$ des points de A tels que

$$A \subset \bigcup_{i=1}^{p} (A \cap V_{a_i}), \quad \text{donc} \quad A \subset \bigcup_{i=1}^{p} V_{a_i}.$$

On pose donc

$$W = \bigcap_{i=1}^{p} W_{a_i}.$$

On a alors $x \in W$ et

$$W \cap A \subset W \cap \left(\bigcup_{i=1}^{p} V_{a_i} \right) = \left(\bigcap_{i=1}^{p} W_{a_i} \right) \cap \left(\bigcup_{i=1}^{p} V_{a_i} \right) \subset \bigcup_{i=1}^{p} (W_{a_i} \cap V_{a_i}) = \emptyset.$$

Donc $x \in W \subset A^C$. Comme de plus W est un ouvert, A^C est donc voisinage de x. Donc A^C est un ouvert. □

Corollaire 4.9 *Soit E un espace topologique compact. Soit $F \subset E$. Alors*

$$F \text{ compact} \iff F \text{ fermé.}$$

Démonstration. On vient de voir que si F est compact, alors F est fermé (même si E n'est pas compact). Par ailleurs, la proposition 4.5 donne la réciproque. □

Proposition 4.10 *Soit E un espace topologique. Soient K un compact de E et F un fermé de E. Alors $K \cap F$ est un compact.*

Démonstration. Comme F est fermé, $K \cap F$ est un fermé de K. Comme K est compact, $K \cap F$ est donc compact. □

Proposition 4.11 *Dans un espace topologique séparé, toute union finie de compacts est compact.*

Démonstration. Soit E un espace topologique séparé, et soient $K_1, K_2, \dots, K_p$ des compacts de E. Posons

$$K = \bigcup_{i=1}^{p} K_i.$$

Considérons maintenant un recouvrement de K par des ouverts $(O_i)_{i\in I}$. Alors, pour tout entier j entre 1 et p, la famille $(O_i)_{i\in I}$ est un recouvrement d'ouverts de K_j. On peut donc en extraire un recouvrement fini :

$$\forall j \in \{1, 2, \dots, p\}, \quad K_j \subset \bigcup_{i\in I_j} O_i,$$

où $I_j \subset I$ est fini. On pose alors $M = \bigcup_{j=1}^{p} I_j$, qui est fini, est on a donc :

$$\forall j \in \{1, 2, \dots, p\}, \quad K_j \subset \bigcup_{i\in I_j} O_i \subset \bigcup_{i\in M} O_i.$$

Donc $K \subset \bigcup_{i\in M} O_i$, et on a ainsi un sous-recouvrement fini de K. □

La notion de compacité est préservée par les homéomorphismes :

Proposition 4.12 *Soient E et F des espaces topologiques homéomorphes. Alors*

$$E \text{ compact} \iff F \text{ compact.}$$

Démonstration. Supposons que E est compact. Soit $(O_i)_{i\in I}$ un recouvrement d'ouverts de F. En notant f un homéomorphisme de E dans F, on sait que $f^{-1}(O_i)$ est ouvert, pour tout $i \in I$. De plus,

$$\bigcup_{i\in I} f^{-1}(O_i) = f^{-1}\left(\bigcup_{i\in I} O_i\right) = f^{-1}(F) = E.$$

Donc la famille $\left(f^{-1}(O_i)\right)_{i\in I}$ est un recouvrement d'ouverts de E. On en extrait un sous-recouvrement fini : il existe $J \subset I$, J fini, tel que

$$E = \bigcup_{i\in J} f^{-1}(O_i).$$

D'où l'on déduit, puisque f est bijective,

$$F = f(E) = f\left(\bigcup_{i\in J} f^{-1}(O_i)\right) = \bigcup_{i\in J} f\left(f^{-1}(O_i)\right) = \bigcup_{i\in J} O_i.$$

Donc F est bien un compact. On prouve la réciproque de la même façon, en échangeant le rôle de E et F, et en utilisant f^{-1} à la place de f. □

Théorème 4.13 *Soit E et F des espaces topologiques tels que F est séparé. Soit $f : E \to F$ une application continue. Soit $K \subset E$ un compact. Alors $f(K)$ est compact.*

Démonstration. Soit $(O_i)_{i\in I}$ un recouvrement d'ouverts de $f(K)$. Alors $f^{-1}(O_i)$ est ouvert pour tout $i \in I$, et

$$K \subset f^{-1}(f(K)) \subset f^{-1}\left(\bigcup_{i\in I} O_i\right) = \bigcup_{i\in I} f^{-1}(O_i).$$

Ainsi, $\left(f^{-1}(O_i)\right)_{i\in I}$ est un recouvrement d'ouverts de K. On en extrait un sous-recouvrement fini :

$$K \subset \bigcup_{i\in J} f^{-1}(O_i), \quad J \subset I, \quad J \text{ fini}.$$

D'où l'on déduit

$$f(K) \subset f\left(\bigcup_{i\in J} f^{-1}(O_i)\right) = \bigcup_{i\in J} f\left(f^{-1}(O_i)\right) \subset \bigcup_{i\in J} O_i.$$

Donc $f(K)$ est bien compact. □

Corollaire 4.14 *Soit E un espace topologique compact, et F un espace topologique séparé. Soit $f : E \to F$ une application bijective continue. Alors f est un homéomorphisme.*

Démonstration. Il suffit de démontrer que l'application réciproque $g = f^{-1}$ de f est continue. Soit donc K un fermé de E. Comme E est compact, K est compact. Donc, en appliquant le théorème 4.13, $f(K) = g^{-1}(K)$ est compact. Donc $g^{-1}(K)$ est fermé. Donc g est bien continue. □

4.2 Espaces métriques compacts

Théorème 4.15 (de Bolzano-Weierstrass.) *Soit E un espace métrique. Alors on a équivalence entre les propriétés suivantes :*

(i) E est compact.

(ii) Toute suite de E admet une valeur d'adhérence.

(iii) De toute suite de E on peut extraire une suite convergente.

Démonstration. $(i) \Rightarrow (ii)$: soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. Supposons qu'elle n'a pas de valeur d'adhérence. Alors

$$\forall x \in E, \quad \exists O_x \quad \text{ouvert tel que} \quad \{n \in \mathbb{N}, \ / \ x_n \in O_x\} \ \text{ est fini.}$$

On a $E \subset \displaystyle\bigcup_{x\in E} O_x$. Comme E est compact, on en extrait un sous-recouvrement fini :

$$E \subset \bigcup_{i=1}^{p} O_{y_i}.$$

Comme chacun des O_{y_i} ne contient qu'un nombre fini de termes de la suite, E lui-même n'en contient qu'un nombre fini. Ce qui est contradictoire.
$(ii) \Rightarrow (iii)$ est évident.
$(iii) \Rightarrow (i)$: soit $(O_i)_{i\in I}$ un recouvrement d'ouverts de E. Nous allons d'abord démontrer que

$$\exists r > 0 \quad / \quad \forall x \in E, \quad \exists i \in I, \quad B_r(x) \subset O_i. \tag{4.1}$$

On procède par l'absurde, et on suppose que

$$\forall r > 0, \quad \exists x \in E \quad / \quad \forall i \in I, \quad B_r(x) \not\subset O_i.$$

On choisit donc $r = 1/(n+1)$, et on construit donc une suite $(x_n)_{n\in\mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}, \quad \forall i \in I, \quad B_{\frac{1}{n+1}}(x_n) \not\subset O_i.$$

Comme E vérifie (iii), on peut donc extraire de $(x_n)_{n\in\mathbb{N}}$ une suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ convergente. On note x sa limite, et, par définition de la famille $(O_i)_{i\in I}$, on sait donc qu'il existe $i_0 \in I$ tel que $x \in O_{i_0}$. Mais, par définition d'un ouvert, on sait aussi que

$$\exists \rho > 0 \quad / \quad B_\rho(x) \subset U_{i_0}.$$

D'autre part, on sait qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad d(x_{\varphi(n)}, x) < \frac{\rho}{2}.$$

On choisit donc $n \geq n_0$ tel que $\frac{1}{\varphi(n)+1} < \frac{\rho}{2}$. Ainsi,

$$B_{\frac{1}{\varphi(n)+1}}(x_{\varphi(n)}) \subset B_\rho(x) \subset O_{i_0},$$

ce qui est contradictoire. Donc (4.1) est bien vérifiée.
Nous allons maintenant démontrer que

$$\forall r > 0, \quad \exists n \in \mathbb{N}, \quad \exists (x_1, x_2, \ldots, x_n) \in E^n \quad / \quad E \subset \bigcup_{i=1}^{n} B_r(x_i). \tag{4.2}$$

Là encore, on procède par l'absurde. On suppose donc que

$$\exists r > 0, \quad \forall n \in \mathbb{N}, \quad \forall (x_1, x_2, \ldots, x_n) \in E^n, \quad E \not\subset \bigcup_{i=1}^{n} B_r(x_i).$$

Soit alors $x_0 \in E$. On sait donc que $E \not\subset B_r(x_0)$. Soit donc $x_1 \in E \setminus B_r(x_0)$. On construit ainsi par récurrence une suite $(x_n)_{n\in\mathbb{N}}$: si $x_1, x_2, \ldots, x_n$ ont été construits, on sait que $E \not\subset \bigcup_{i=1}^{n} B_r(x_i)$, donc on choisit $x_{n+1} \in E \setminus \bigcup_{i=1}^{n} B_r(x_i)$. Donc cette suite vérifie

$$\forall n \in \mathbb{N}, \quad \forall p \in \mathbb{N}^*, d(x_n, x_{n+p}) \geq r.$$

Or on peut en extraire une suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ qui converge. En particulier, cette suite vérifie donc

$$d(x_{\varphi(n)}, x_{\varphi(n+1)}) \underset{n\to+\infty}{\longrightarrow} 0.$$

Ceci est contradctoire avec l'inégalité précédente, car cette dernière implique $d(x_{\varphi(n)}, x_{\varphi(n+1)}) \geq r$.

Nous pouvons maintenant utiliser (4.1) et (4.2) pour démontrer (i). En effet, on sait donc que

$$\forall x \in E, \quad \exists i_x \in I \quad / \quad B_r(x) \subset O_{i_x}.$$

De plus, il existe $(x_1, x_2, \ldots, x_n) \in E^n$ tel que

$$E \subset \bigcup_{j=1}^{n} B_r(x_j) \subset \bigcup_{j=1}^{n} O_{i_{x_j}},$$

et on a donc bien construit un sous-recouvrement fini de E. □

Proposition 4.16 *Soient $E_1, E_2, \ldots, E_p$ des espaces métriques compacts. Alors le produit $E = \prod_{i=1}^{p} E_i$ est compact.*

Démonstration. Commençons par prouver le résultat pour $p = 2$: soit $(x_n)_{n\in\mathbb{N}}$ une suite de $E = E_1 \times E_2$. Alors

$$\forall n \in \mathbb{N}, \quad x_n = (x_n^1, x_n^2), \quad x_n^1 \in E_1, \quad x_n^2 \in E_2.$$

Donc, puisque E_1 est compact, on peut extraire une suite $\left(x_{\varphi(n)}^1\right)_{n\in\mathbb{N}}$ de $\left(x_n^1\right)_{n\in\mathbb{N}}$ qui converge :

$$x_{\varphi(n)}^1 \underset{n\to+\infty}{\longrightarrow} x^1 \in E_1.$$

De plus, E_2 est compact également, donc un peut extraire de la suite $\left(x^2_{\varphi(n)}\right)_{n\in\mathbb{N}}$ une suite $\left(x^2_{\varphi(\psi(n))}\right)_{n\in\mathbb{N}}$ qui converge vers une limite $x^2 \in E_2$. La suite extraite $\left(x^1_{\varphi(\psi(n))}\right)_{n\in\mathbb{N}}$ converge elle aussi (vers x^1), et donc, en posant $x = (x^1, x^2)$,

$$d\left(x_{\varphi(\psi(n))}, x\right) = \max\left[d_1\left(x^1_{\varphi(\psi(n))}, x^1\right), d_2\left(x^2_{\varphi(\psi(n))}, x^2\right)\right] \underset{n\to+\infty}{\longrightarrow} 0.$$

On a donc extrait une suite convergente de la suite $(x_n)_{n\in\mathbb{N}}$. Donc, d'après le théorème de Bolzano-Weierstrass (théorème 4.15), E est compact.

On prouve ensuite le cas général par récurrence sur p : si $p = 1$, il n'y a rien à prouver. Si le résultat est vrai pour p donné, alors, pour prouver le résultat au rang $p + 1$, on écrit

$$E = E_1 \times \cdots \times E_{p+1} = \underbrace{(E_1 \times \ldots E_p)}_{F} \times E_{p+1},$$

et F est compact grâce à l'hypothèse de récurrence. Donc E est bien compact, d'après ce qu'on vient de prouver. Ceci achève la récurrence. □

Théorème 4.17 (de Heine.) *Soient E et F deux espaces métriques. Soit $f : E \to F$ une application continue. Si E est compact, f est uniformément continue.*

Démonstration. On procède par l'absurde. Si f n'est pas uniformément, continue, alors

$$\exists \varepsilon > 0 \quad / \quad \forall \delta > 0, \quad \exists (x, y) \in E^2 \quad / \quad d_E(x, y) < \delta \quad \text{et} \quad d_F(f(x), f(y)) > \varepsilon.$$

Ce réel ε étant fixé, on choisi alors $\delta = 1/(n + 1)$, pour $n \in \mathbb{N}$. On a donc

$$\forall n \in \mathbb{N}, \quad \exists (x_n, y_n) \in E^2 \quad / \quad d_E(x_n, y_n) < \frac{1}{n+1} \quad \text{et} \quad d_F(f(x), f(y)) > \varepsilon.$$

Comme E est compact, $E \times E$ également, donc on peut extraire une suite de $[(x_n, y_n)]_{n\in\mathbb{N}}$, que l'on note $\left[(x_{\varphi(n)}, y_{\varphi(n)}\right]_{n\in\mathbb{N}}$, et qui converge vers $(x, y) \in E^2$. On a

$$\forall n \in \mathbb{N}, \quad d_E\left(x_{\varphi(n)}, y_{\varphi(n)}\right) \leq \frac{1}{\varphi(n) + 1} \underset{n\to+\infty}{\longrightarrow} 0,$$

donc $x = y$. Par ailleurs, on a

$$\forall n \in \mathbb{N}, \quad d_F(f(x_{\varphi(n)}), f(y_{\varphi(n)})) > \varepsilon,$$

d'où $d(f(x), f(y)) \geq \varepsilon > 0$. Ceci est contradictoire avec l'égalité $x = y$. □

4.3 Parties compactes de $\mathbb{R}$

Dans toute cette section, on utilise sur $\mathbb{R}$ la topologie usuelle, c'est-à-dire la topologie définie par la distance $d(x, y) = |x - y|$. On rappelle alors les faits suivants, qui seront utiles dans la suite :

– Tout ensemble majoré de $\mathbb{R}$ admet une borne supérieure.
– Tout ensemble minoré de $\mathbb{R}$ admet une borne inférieure.
– Toute suite croissante majorée converge, et sa limite est sa borne supérieure.
– Toute suite décroissante minorée converge, et sa limite est sa borne inférieure.

En particulier, une conséquence simple des deux propriétés précédentes, que nous laissons en exercice, est la convergence des suites adjacentes :

Lemme 4.18 *Soient deux suites réelles* $(u_n)_{n\in\mathbb{N}}$ *et* $(v_n)_{n\in\mathbb{N}}$ *telles que*

(a) $(u_n)_{n\in\mathbb{N}}$ *est décroissante.*
(b) $(v_n)_{n\in\mathbb{N}}$ *est croissante.*
(c) $\forall n \in \mathbb{N}, \quad v_n \leq u_n$.
(d) $\lim_{n\to+\infty}(u_n - v_n) = 0$.

Alors ces deux suites convergent est ont la même limite.

Nous laissons la preuve de ce résultat en exercice.

Définition 4.19 *Soit* (E, d) *un espace métrique, et soit* $A \subset E$*. On dit que* A *est borné s'il existe* $x_0 \in E$ *et* $M > 0$ *tels que*

$$\forall x \in A, \quad d(x, x_0) \leq M.$$

Les deux cas particuliers suivants sont importants :

$$\forall A \subset \mathbb{R}, \quad A \text{ borné} \iff \exists M > 0 \quad / \quad A \subset [-M, M].$$

Et pour tout entier $n \neq 0$,

$$\forall A \subset \mathbb{R}^n, \quad A \text{ borné} \iff \exists M > 0 \quad / \quad A \subset [-M, M]^n.$$

Théorème 4.20 (de Borel-Lebesgue.) *L'intervalle* $[0, 1]$ *muni de la distance* $d(x, y) = |x - y|$ *est compact.*

Démonstration. Nous allons démontrer que de toute suite à valeurs dans $[0, 1]$, on peut extraire une suite qui converge vers un élément de $[0, 1]$. Soit donc une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans $[0, 1]$. On procède par dichotomie : soit

$$a_0 = 0, \quad b_0 = 1, \quad c_0 = \frac{1}{2}(a_0 + b_0),$$

et

$$N_0 = \{n \in \mathbb{N}, \quad x_n \in [a_0, c_0[\}, \quad P_0 = \{n \in \mathbb{N}, \quad x_n \in [c_0, b_0]\}.$$

Alors $P_0 \cup N_0 = \mathbb{N}$, donc l'un au moins de ces ensembles est infini. On définit donc a_1 et b_1 de la façon suivante :

$$\begin{aligned} \text{si } N_0 \text{ est infini,} \quad & a_1 = a_0, \quad b_1 = c_0, \\ \text{sinon,} \quad & a_1 = c_0, \quad b_1 = b_0. \end{aligned}$$

Par récurrence, on construit ainsi les suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ telles que l'ensemble $\{k \in \mathbb{N}, \quad x_k \in [a_n, b_n]\}$ est infini :

$$c_n = \frac{1}{2}(a_n + b_n),$$

$$N_n = \{k \in \mathbb{N}, \quad x_k \in [a_n, c_n[\}, \quad P_n = \{k \in \mathbb{N}, \quad x_k \in [c_n, b_n]\}.$$

Pour tout $n \in \mathbb{N}$, l'ensemble $N_n \cup P_n$ est infini, donc N_n ou P_n est infini, et on pose

$$\begin{aligned} \text{si } N_n \text{ est infini,} \quad & a_{n+1} = a_n, \quad b_{n+1} = c_n, \\ \text{sinon,} \quad & a_{n+1} = c_n, \quad b_{n+1} = b_n. \end{aligned}$$

On a de plus que $(a_n)_{n\in\mathbb{N}}$ est une suite croissante, et que $(b_n)_{n\in\mathbb{N}}$ est décroissante. De plus,

$$b_{n+1} - a_{n+1} = \frac{1}{2}(b_n - a_n), \quad \text{et} \quad b_0 - a_0 = 1.$$

donc, par récurrence, $b_n - a_n = 2^{-n}$. Les suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont donc adjacentes : elles convergent donc vers une limite commune $x \in [0, 1]$:

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} b_n = x.$$

Maintenant, on construit une suite extraite de $(x_n)_{n\in\mathbb{N}}$ qui converge vers x :

$$\varphi(0) = 0, \quad \text{et} \quad \forall k \geq 1, \quad \varphi(k) = \inf\{n \geq \varphi(k-1) + 1, \quad x_n \in [a_k, b_k]\}.$$

Ainsi, la suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$ vérifie

$$\forall k \in \mathbb{N}, \quad a_k \leq x_{\varphi(k)} \leq b_k.$$

Donc cette suite converge vers x. □

Proposition 4.21 *Les parties compactes de $\mathbb{R}$ sont ses parties fermées bornées.*

Démonstration. Nous avons donc deux implications à montrer : d'une part que si $A \subset \mathbb{R}$ est compact, alors il est fermé borné, et d'autre part sa réciproque.

Commençons par l'implication directe. Si A est compact, alors il est fermé d'après la propriété 4.8. De plus, supposons que A n'est pas borné. Alors

$$\forall M \in \mathbb{R}, \quad \exists x \in A \quad / \quad |x| > M.$$

On construit alors une suite de A de la façon suivante :
- $x_0 \in A$ tel que $|x_0| \geq 1$.
- $\forall n \in \mathbb{N}$, $x_{n+1} \in A$ tel que $|x_{n+1}| \geq 2|x_n|$.

Pour chacune de ces étapes, on sait qu'un tel réel existe d'après ce qu'on vient de voir. Comme A est compact, on sait, d'après le théorème de Bolzano-Weierstrass (théorème 4.15), qu'on peut en extraire une sous-suite convergente. Soit une telle sous-suite : il existe une application $\varphi : \mathbb{N} \to \mathbb{N}$ croissante strictement telle que

$$x_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} x \in A.$$

En particulier,

$$x_{\varphi(n+1)} - x_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} 0.$$

Or, d'après ce qui précède, on a

$$\begin{aligned}\left|x_{\varphi(n+1)} - x_{\varphi(n)}\right| &\geq \left|x_{\varphi(n+1)}\right| - \left|x_{\varphi(n)}\right| \geq 2^{\varphi(n+1)-\varphi(n)}\left|x_{\varphi(n)}\right| - \left|x_{\varphi(n)}\right| \\ &= \left(2^{\varphi(n+1)-\varphi(n)} - 1\right)\left|x_{\varphi(n)}\right| \geq 2^{\varphi(n+1)-\varphi(n)} - 1 \geq 1,\end{aligned}$$

ce qui est contradictoire. Donc A est borné.

Pour démontrer la réciproque, on commence par remarquer que pour tout réel $M > 0$, l'intervalle fermé $[-M, M]$ est compact car il est homéomorphe à $[0, 1]$, qui est compact d'après le théorème de Borel-Lebesgue. En effet, l'application

$$\begin{array}{rccl} f : & [0,1] & \longrightarrow & [-M, M], \\ & x & \longmapsto & -M + 2Mx, \end{array}$$

est une application bijective et continue. Comme $[0, 1]$ est compact, il s'agit d'un homéomorphisme d'après le corollaire 4.14.

De plus, si A est fermé borné, par définition, on sait qu'il existe $M > 0$ tel que $A \subset [-M, M]$. A est donc un fermé de $[-M, M]$ qui est compact, donc un compact. □

On généralise sans problème cette preuve au cas de $\mathbb{R}^n$:

Proposition 4.22 *Soit $n \in \mathbb{N}^*$, et soit $\mathbb{R}^n$ muni de la norme infinie. Ses parties compactes sont ses fermés bornés.*

Démonstration. Rappelons que la norme infinie est définie par

$$\forall x \in \mathbb{R}^n, \quad \|x\|_\infty = \sup_{1\leq j\leq n} |x_j|.$$

Soit $A \subset \mathbb{R}^n$ un compact. Alors A est fermé, et si A n'est pas borné, alors

$$\forall M > 0, \quad \exists x \in A \quad / \quad \|x\|_\infty > M.$$

On construit alors une suite de A de la façon suivante :
- $x_0 \in A$ tel que $\|x_0\|_\infty \geq 1$.
- $\forall n \in \mathbb{N}$, $x_{n+1} \in A$ tel que $\|x_{n+1}\|_\infty \geq 2\|x_n\|_\infty$.

Pour chacune de ces étapes, on sait qu'un tel vecteur existe d'après ce qu'on vient de voir. Comme A est compact, on sait, d'après le théorème de Bolzano-Weierstrass (théorème 4.15), qu'on peut en extraire une sous-suite convergente. Soit une telle sous-suite : il existe une application $\varphi : \mathbb{N} \to \mathbb{N}$ croissante strictement telle que

$$x_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} x \in A.$$

En particulier,

$$x_{\varphi(n+1)} - x_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} 0.$$

Or, d'après ce qui précède, on a

$$\begin{aligned} \left\|x_{\varphi(n+1)} - x_{\varphi(n)}\right\|_\infty &\geq \left\|x_{\varphi(n+1)}\right\|_\infty - \left\|x_{\varphi(n)}\right\|_\infty \\ &\geq 2^{\varphi(n+1)-\varphi(n)} \left\|x_{\varphi(n)}\right\|_\infty - \left\|x_{\varphi(n)}\right\|_\infty = \left(2^{\varphi(n+1)-\varphi(n)} - 1\right) \left\|x_{\varphi(n)}\right\|_\infty \\ &\geq 2^{\varphi(n+1)-\varphi(n)} - 1 \geq 1, \end{aligned}$$

ce qui est contradictoire.

Pour démontrer la réciproque, on commence par remarquer que pour tout réel $M > 0$, l'ensemble $[-M, M]^n$ est un produit de compacts, donc il est compact pour la topologie produit. Comme de plus cette topologie est issue de la norme $\|\cdot\|_\infty$, cet ensemble est un compact pour cette norme.

De plus, si A est fermé borné, par définition, on sait qu'il existe $M > 0$ tel que $\forall x \in A$, $\|x\|_\infty \leq M$. En particulier, pour tout $i \in \{1, 2, \cdots, n\}$, $|x_i| \leq M$. Donc $A \subset [-M, M]^n$. A est donc un fermé de $[-M, M]^n$ qui est compact, donc un compact. □

Proposition 4.23 *Soit (E, d) un espace topologique compact, et soit $f : E \to \mathbb{R}$ une application continue. Alors f est bornée et atteint ses bornes.*

Démonstration. Par application du théorème 4.13, $f(E)$ est compact, donc fermé borné. Donc f est bornée. Soit

$$M = \sup\{f(x), \quad x \in E\}.$$

Par définition d'une borne supérieure,

$$\forall \varepsilon > 0, \quad \exists y \in f(E) \quad / \quad M - \varepsilon \leq y \leq M.$$

En particulier, on sait donc que

$$\forall n \in \mathbb{N}, \quad \exists x_n \in E \quad / \quad M - \frac{1}{n+1} \leq y_n \leq M.$$

Par compacité de $f(E)$, on peut extraire une sous-suite de $(y_n)_{n\in\mathbb{N}}$ qui converge : il existe $\varphi : \mathbb{N} \to \mathbb{N}$ croissante strictement telle que

$$y_{\varphi(n)} \underset{n\to+\infty}{\longrightarrow} y \in f(E).$$

D'autre part, d'après ce qui précède,

$$y_n \underset{n\to+\infty}{\longrightarrow} M.$$

Autrement dit, $M \in f(E)$. Donc, par définition de $f(E)$, il existe $x \in E$ tel que $f(x) = M$. Donc la borne supérieure de f est atteinte en x.

Pour démontrer que la borne inférieure est atteinte, on change f en $-f$, car alors

$$\inf \{f(x), \quad x \in E\} = -\sup \{-f(x), \quad x \in E\},$$

qui est atteint, d'après ce qui précède, puisque $-f$ est continue. □

Proposition 4.24 *Dans l'espace vectoriel $\mathbb{R}^n$, $n \in \mathbb{N}^*$, toutes les normes sont équivalentes.*

Démonstration. On définit

$$\|x\|_\infty = \sup_{1\leq i\leq n} |x_i|, \quad \text{où} \quad x = (x_1, x_2, \ldots x_n).$$

On a vu dans l'exemple 1.5 que $\|\cdot\|_\infty$ est une norme sur $\mathbb{R}^n$. Soit N une norme quelconque sur $\mathbb{R}^n$. Nous allons démontrer qu'elle est équivalente à $\|\cdot\|_\infty$. Pour cela, on pose

$$S = \{x \in E, \quad \|x\|_\infty = 1\}.$$

Il s'agit d'un fermé de $(\mathbb{R}^n, \|\cdot\|_\infty)$, car c'est l'image réciproque par la norme de $\{1\}$, qui est un fermé de $\mathbb{R}$, et c'est un borné. Donc S est compact. On appelle $(e_1, e_2, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$, et on a alors

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

d'où

$$\forall x \in \mathbb{R}^n, \quad \forall y \in \mathbb{R}^n, \quad |N(x) - N(y)| \leq N(x-y) \leq M\|x-y\|_\infty,$$

où $M = \displaystyle\sum_{i=1}^n N(e_i)$. Ainsi, $N : S \to \mathbb{R}^+$ est continue. Elle y atteint donc ses bornes, et en particulier sa borne inférieure, que l'on note $m \geq 0$. Il existe donc $x \in S$ tel que $N(x) = m$. Donc $m \neq 0$, puisque $x \neq 0$. Soit maintenant $z \in \mathbb{R}^n$ quelconque. Si $z = 0$, alors $N(z) \geq m\|z\|_\infty$. Et sinon, on pose

$$x = \frac{z}{\|z\|_\infty}, \quad \text{d'où} \quad N(x) = \frac{N(z)}{\|z\|_\infty} \geq m.$$

Ainsi, $N(z) \geq m\|z\|_\infty$. Comme on a déjà vu plus haut que $N(z) \leq M\|z\|_\infty$, ceci prouve bien que N et $\|\cdot\|_\infty$ sont équivalentes. Donc toutes les normes sur $\mathbb{R}^n$ sont équivalentes à $\|\cdot\|_\infty$. Elles sont donc équivalentes entre elles. □

Proposition 4.25 *Soit $(E, \|\cdot\|)$ un espace vectoriel normé de dimension finie. Alors toutes les normes sur E sont équivalentes.*

Démonstration. Soit $(e_1, e_2, \dots e_n)$ une base de E. Considérons l'application

$$\begin{array}{rccl} f : & \mathbb{R}^n & \longrightarrow & E, \\ & x = (x_1, x_2, \dots x_n) & \longmapsto & \displaystyle\sum_{i=1}^{n} x_i e_i. \end{array}$$

Alors l'application N définie par $N(x) = \|f(x)\|$ est une norme sur $\mathbb{R}^n$, puisqu'elle vérifie les trois axiomes de la définition 1.1 :

(i) $\forall x \in \mathbb{R}^n$, $N(x) = 0 \iff f(x) = 0 \iff x = 0$ car $(e_1, e_2, \dots, e_n)$ est une base de E.

(ii) $\forall x \in \mathbb{R}^n$, $\forall \lambda \in \mathbb{R}$, $N(\lambda x) = \|f(\lambda x)\| = \|\lambda f(x)\| = |\lambda|\|f(x)\| = |\lambda| N(x)$, car f est linéaire.

(iii) $\forall x \in \mathbb{R}^n$, $\forall y \in \mathbb{R}^n$, $N(x+y) = \|f(x+y)\| = \|f(x) + f(y)\| \leq \|f(x)\| + \|f(y)\| = N(x) + N(y)$.

De même, si N' est une autre norme sur E, l'application $N'(f(x))$ est une norme sur $\mathbb{R}^n$. Or d'après la proposition 4.24, elle sont équivalentes. Donc il existe $M \geq m > 0$ tels que

$$\forall x \in \mathbb{R}^n, \quad mN'(f(x)) \leq \|f(x)\| \leq MN'(f(x)).$$

De plus, f est bijective car $(e_1, e_2, \dots, e_n)$ et une base de E. Donc l'encadrement ci-dessus implique

$$\forall y \in \mathbb{R}^n, \quad mN'(y) \leq \|y\| \leq MN'(y).$$

Les normes $\|\cdot\|$ et N' sont donc bien équivalentes. □

Notons que, dans la preuve ci-dessus, l'application f est une application linéaire bijective qui vérifie

$$\forall x \in \mathbb{R}^n, \quad m\|x\|_\infty \leq \|f(x)\| \leq M\|x\|_\infty,$$

puisque N est une norme sur $\mathbb{R}^n$, donc qu'elle est équivalente à $\|\cdot\|_\infty$. Ceci implique en particulier que f est continue, et que f^{-1} l'est aussi, car elle vérifie :

$$\forall y \in E, \quad \|f^{-1}(y)\|_\infty \leq \frac{1}{m}\|y\|.$$

On a donc démontré également :

Corollaire 4.26 *Soit E un espace vectoriel normé de dimension $n \in \mathbb{N}$. Alors E est homéomorphe à $\mathbb{R}^n$.*

Remarque 4.27 *Citons un exemple d'application linéaire non continue : on pose $E = \mathbb{R}[X]$ espace vectoriel des polynômes à coefficients réels. On munit cet espace de la norme*

$$\|P\|_1 = \int_0^1 |P(t)|dt.$$

(On a vu dans l'exemple 1.6 qu'il s'agit bien d'une norme.) Soit l'application

$$\begin{array}{rccc} f: & E & \longrightarrow & \mathbb{R}, \\ & P & \longmapsto & P(2). \end{array}$$

Cette application est linéaire, et n'est pas continue.

Démonstration. Montrons d'abord que f est linéaire : soient P et Q dans E et $\lambda \in \mathbb{R}$. Alors on a

$$f(P + \lambda Q) = (P + \lambda Q)(2) = P(2) + \lambda Q(2) = f(P) + \lambda f(Q).$$

Montrons maintenant qu'elle n'est pas continue. Pour cela, on considère la suite d'éléments de E définie par

$$\forall n \in \mathbb{N}, \quad P_n(x) = x^n.$$

On a alors

$$\|P_n\|_1 = \int_0^1 t^n dt = \frac{1}{n+1}, \quad f(P_n) = 2^n.$$

Ainsi,

$$\frac{|f(P_n)|}{\|P_n\|_1} = \frac{2^n}{n+1} \underset{n\to+\infty}{\longrightarrow} +\infty.$$

Or, si f était continue, il existerait $M \geq 0$ tel que

$$\forall n \in \mathbb{N}, \quad |f(P_n)| \leq M\|P_n\|_1,$$

ce qui est contradictoire. □

Théorème 4.28 *Soient E et F des espaces vectoriels normés tels que E est de dimension finie. Soit f une application linéaire de E dans F. Alors f est continue.*

Démonstration. Soit $(e_1, e_2, \ldots, e_n)$ une base de E. Alors on sait que l'application $N : E \to \mathbb{R}^+$ définie par

$$\forall x \in E, \quad N(x) = \sum_{i=1}^n |x_i|, \quad \text{où} \quad x = \sum_{i=1}^n x_i e_i,$$

est une norme sur E (voir la démonstration de la proposition 4.25). Nous allons démontrer que f est continue pour cette norme, ce qui prouvera qu'elle est continue pour toute les normes sur E, d'après la proposition 4.25. On a alors :

$$\forall x \in E, \quad \|f(x)\|_F = \left\| \sum_{i=1}^n x_i f(e_i) \right\|_F \leq \sum_{i=1}^n |x_i| \|f(e_i)\|_F \leq M N(x),$$

avec

$$M = \sup_{1 \leq i \leq n} \|f(e_i)\|_F.$$

Donc, d'après le point (vii) de la proposition 1.31, f est continue. □

Chapitre 5

Espaces connexes

La notion mathématique de connexité correspond intuitivement au fait qu'un ensemble est "en un seul morceau". Par exemple, le sous-ensemble $[0,1]$ de $\mathbb{R}$ est connexe, mais pas le sous-ensemble $[0,1] \cup [2,3]$. Nous allons voir comment utiliser les outils de topologie développés jusqu'à présent pour définir mathématiquement cette notion.

5.1 Connexité

Définition 5.1 *Soit E un espace topologique. On dit que E est connexe si ses seules parties à la fois ouvertes et fermées sont E et $\emptyset$.*

Définition 5.2 *Soit E un espace topologique. Une partie A de E est dite connexe si elle l'est pour la topologie induite par celle de E.*

Citons, avant d'aller plus loin, des exemples d'ensembles connexes (ou non) :

Exemple 5.3 *$\mathbb{R}$ muni de sa topologie usuelle est connexe.*

Dans cette preuve, nous utiliserons le fait que $\mathbb{R}$ est un ensemble totalement ordonné, et que toute partie majorée de $\mathbb{R}$ admet une borne supérieure.
Démonstration. Supposons que $A \subset \mathbb{R}$ est à la fois ouvert et fermé, et que A n'est ni $\mathbb{R}$ ni le vide. Alors $\exists x \in A$ et $\exists y \in A^C$. Quitte à échanger le rôle de A et A^C, on peut supposer que $x < y$. Comme $y \notin A$,

$$A' = A \cap]-\infty, y[= A \cap]-\infty, y] .$$

Cet ensemble est donc à la fois un ouvert et un fermé de A. De plus, il est majoré. Donc il admet une borne supérieure : soit $z = \sup A'$. Par définition, on a :

$$\forall \delta > 0, \quad \exists z' \in A' \quad / \quad z - \delta \leq z' \leq z.$$

En appliquant cela pour $\delta = \frac{1}{n}$, on obtient donc :

$$\forall n \in \mathbb{N}^*, \quad \exists z_n \in A', \quad z - \frac{1}{n} \leq z_n \leq z.$$

Donc $\lim_{n\to+\infty} z_n = z$, d'où $z \in A'$ puisque A' est fermé. Mais comme A' est ouvert, on sait qu'il existe $\delta > 0$ tel que $]z-\delta, z+\delta[\subset A'$. En particulier, $z + \delta/2 \in A'$, ce qui est contradictoire avec la définition de z. □

Exemple 5.4 *Dans un espace métrique (E, d), si x et y vérifient $d(x, y) \geq 1$, alors $B_{1/2}(x) \cup B_{1/2}(y)$ (muni de la topologie induite) n'est pas connexe. En particulier, dans $\mathbb{R}$, $]0, 1[\cup]2, 3[$ n'est pas connexe.*

Démonstration. Soit $A = B_{1/2}(x) \cup B_{1/2}(y)$. Posons $r = d(x, y) > 1$. On a alors

$$A \cap B_{1/2}(x) = B_{1/2}(x),$$

donc $B_{1/2}(x)$ est un ouvert de A. De plus,

$$B_{1/2}(x) = A \cap B_{1/2}(x) \subset A \cap \overline{B}_{1/2}(x),$$

et, si $z \in A \cap \overline{B}_{1/2}(x)$, alors $d(x, z) \leq 1/2$ et $z \in A = B_{1/2}(x) \cup B_{1/2}(y)$. Mais z ne peut être dans $B_{1/2}(y)$, car on aurait alors

$$d(x, y) \leq d(x, z) + d(z, y) < \frac{1}{2} + \frac{1}{2} = 1,$$

ce qui est impossible, par hypothèse. Donc $z \in B_{1/2}(x)$. On a donc montré que $A \cap \overline{B}_{1/2}(x) \subset B_{1/2}(x)$, donc que

$$B_{1/2}(x) = A \cap \overline{B}_{1/2}(x),$$

donc $B_{1/2}(x)$ est un fermé de A. □

Exemple 5.5 *L'ensemble $\mathbb{Q}$ muni de la topologie associé à la métrique usuelle $d(x, y) = |x - y|$ n'est pas connexe.*

Démonstration. La topologie utilisée ici correspond à la topologie induite par la topologie usuelle de $\mathbb{R}$. On rappelle que $\sqrt{2} \notin \mathbb{Q}$. Donc

$$I = \left]-\infty, \sqrt{2}\right] \cap \mathbb{Q} = \left]-\infty, \sqrt{2}\right[\cap \mathbb{Q}$$

est à la fois ouvert et fermé. Il n'est pas vide car il contient 0. Si $\mathbb{Q}$ était connexe, on aurait donc $I = \mathbb{Q}$. Or $2 \in \mathbb{Q}$ mais $2 \notin I$. □

Proposition 5.6 *Soit E un espace topologique. Les trois propriétés suivantes sont équivalentes :*

(i) E est connexe.

(ii) Il n'existe pas de partition non triviale de E en deux ouverts.

(iii) Il n'existe pas de partition non triviale de E en deux fermés.

(iv) Si $f : E \longrightarrow \{0, 1\}$ est continue (on munit $\{0, 1\}$ de la topologie discrète), alors f est constante.

Démonstration. $(i) \Rightarrow (ii)$: supposons que $E = O_1 \cup O_2$, $O_1 \cap O_2 = \emptyset$, avec O_1 et O_2 ouverts. Alors $O_1 = O_2^C$ est donc fermé. Donc, puisque E est connexe, $O_1 = E$ ou $O_1 = \emptyset$. Cette partition est donc triviale.
$(ii) \Rightarrow (iii)$: supposons que $E = F_1 \cup F_2$, $F_1 \cap F_2 = \emptyset$, avec F_1 et F_2 ouverts. Alors $F_2 = F_1^C$ et $F_1 = F_2^C$ sont des ouverts. Donc cette partition est triviale.
$(iii) \Rightarrow (iv)$: soit f une application continue de E dans $\{0,1\}$ muni de la topologie discrète. Alors $\{0\}$ et $\{1\}$ sont des fermés de $\{0,1\}$, donc $f^{-1}(\{0\})$ et $f^{-1}(\{1\})$ sont des fermés de E. De plus,

$$f^{-1}(\{0\}) \cap f^{-1}(\{1\}) = \emptyset,$$

et

$$f^{-1}(\{0\}) \cup f^{-1}(\{1\}) = f^{-1}(\{0\} \cup \{1\}) = f^{-1}(\{0,1\}) = E.$$

Cette partition est donc trivale, d'où $f^{-1}(\{0\}) = \emptyset$ ou $f^{-1}(\{1\}) = \emptyset$. Donc f est constante.
$(iv) \Rightarrow (i)$: supposons que $A \subset E$ est à la fois ouvert et fermé. On construit alors la fonction $f : E \longrightarrow \{0,1\}$ en posant

$$f(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \in A^C. \end{cases}$$

Vérifions que l'image réciproque par f de toute partie de $\{0,1\}$ est bien un ouvert :

$$f^{-1}(\emptyset) = \emptyset, \quad f^{-1}(\{0\}) = A^C, \quad f^{-1}(\{1\}) = A, \quad f^{-1}(\{0,1\}) = E,$$

qui sont tous des ouverts. Donc f est continue. Elle est donc constante. Si elle vaut 1, $A = E$, et si elle vaut 0, $A = \emptyset$. □

Proposition 5.7 *Soit E et F deux espaces topologiques. On suppose que E est connexe. Soit $f : E \longrightarrow F$ une application continue. Alors $f(E)$ est connexe.*

Remarque 5.8 *Dans le résultat ci-dessus, en cohérence avec la définition 5.2, la topologie utilisée sur $f(E)$ est celle induite par F.*

Démonstration. Soit $A \subset f(E)$ à la fois ouvert et fermé. Alors $A = f(E) \cap G = f(E) \cap O$ avec G fermé de F et O ouvert de F. On a alors

$$f^{-1}(A) = f^{-1}(f(E) \cap G) = f^{-1}(G), \quad \text{et} \quad f^{-1}(A) = f^{-1}(f(E) \cap O) = f^{-1}(O).$$

Comme f est continue, cet ensemble est donc à la fois ouvert et fermé dans E. Donc $f^{-1}(A) = E$ ou $f^{-1}(A) = \emptyset$. Dans le premier cas, $A = f(E)$. Et dans le deuxième, on a

$$f^{-1}(A) = \emptyset \quad \text{et} \quad A \subset f(E),$$

donc $A = \emptyset$. Ainsi, $f(E)$ est bien connexe. □

Théorème 5.9 *Soit E un espace topologique. Soit $(A_i)_{i\in I}$ une famille de parties connexes de E telles que $\bigcap_{i\in I} A_i \neq \emptyset$. Alors $A = \bigcup_{i\in I} A_i$ est connexe.*

Démonstration. Soit $a \in \bigcap_{i\in I} A_i$. Soit $f : A \longrightarrow \{0,1\}$ une application continue (pour la topologie discrète sur $\{0,1\}$. Notons $f_i : A_i \longrightarrow \{0,1\}$ la restriction de f à A_i. Elle est continue, et comme A_i est connexe, elle doit être constante. En particulier,

$$\forall x \in A_i, \quad f(x) = f(a).$$

Mais ceci vaut pour tout $i \in I$. Donc

$$\forall x \in A, \quad f(x) = f(a).$$

Ceci prouve que f est une fonction constante sur A. □

Proposition 5.10 *Soient E et F deux espaces topologiques connexes. Alors $E \times F$ est connexe.*

Démonstration. Si E ou F est vide, alors $E \times F = \emptyset$ est connexe. Supposons donc que E et F sont non vides.

Soit $f : E \times F \longrightarrow \{0,1\}$ une application continue. Soient $(x,y) \in E \times F$ et $(x',y') \in E \times F$. Alors l'application $f_{|\{x\}\times F}$ est continue, et l'ensemble $\{x\} \times F$ est homéomorphe à F, qui est connexe. Il est donc connexe. Ainsi, $f_{|\{x\}\times F}$ est constante. De même, $f_{|E\times\{y'\}}$ est constante. On en déduit donc que

$$f(x,y) = f(x,y') = f(x',y').$$

La première égalité est vrai car $f_{|\{x\}\times F}$ est constante, et la deuxième car $f_{|E\times\{y'\}}$ est constante. □

Bien entendu, une récurrence immédiate permet de démontrer qu'un produit fini d'espaces connexes est connexe :

Corollaire 5.11 *Soient $E_1, E_2, \dots E_n$ des espaces topologiques connexes. Alors $E = \prod_{i=1}^{n} E_i$ est connexe.*

Remarque 5.12 *Dans la preuve de la proposition 5.10, on a utilisé que $\{x\}\times F$ était homéomorphe à F. Ceci se démontre en considérant l'application $g : F \longrightarrow \{x\} \times F$ définie par $g(y) = (x,y)$, qui est bien un homéomorphisme.*

Démonstration. Il est clair que g est bijective. Soit maintenant un ouvert U de $\{x\} \times F$. Alors $U = (\{x\} \times F) \cap V$, où V est un ouvert de $E \times F$. On a alors

$$g^{-1}(U) = \{y \in F, \quad (x,y) \in V\}.$$

Soit donc $y \in g^{-1}(U)$: $(x,y) \in V$, qui est ouvert pour la topologie produit. Donc, d'après la proposition 3.29 il existe O_1 ouvert de E et O_2 ouvert de F

tels que $(x,y) \in O_1 \times O_2 \subset V$. En particulier, $x \in O_1 \subset f^{-1}(V)$. donc $f^{-1}(V)$ est bien un ouvert.

Montrons maintenant que g^{-1} est continue. Soit donc U un ouvert de F, et considérons $g(U) = \{(x,y), \quad y \in U\}$. On a alors $g(U) = (E \times U) \cap (\{x\} \times F)$, qui est bien un ouvert de $\{x\} \times F$. $\square$

Proposition 5.13 *Soit E un espace topologique, $A \subset E$ connexe. Si $A \subset B \subset \overline{A}$, alors B est connexe.*

Démonstration. Soit $f : B \longrightarrow \{0,1\}$ continue pour la topologie discrète. En particulier, f est continue sur A. Comme A est connexe, f est constante sur A. Supposons par exemple que $f = 0$ sur A. L'ensemble $f^{-1}(\{0\})$ est un fermé de B, car f est continue. Donc il existe F fermé de E tel que $f^{-1}(\{0\}) = B \cap F$. $A \subset f^{-1}(\{0\}) \subset B \cap F \subset F$, donc $\overline{A} \subset F$ puisque F est fermé. Donc $B = B \cap \overline{A} \subset B \cap F = f^{-1}(\{0\})$. Donc f est nulle sur B. $\square$

5.2 Parties connexes de $\mathbb{R}$

Théorème 5.14 *Soit $A \subset \mathbb{R}$. Alors A connexe $\Longleftrightarrow$ A est un intervalle.*

Nous allons utiliser la caractérisation suivante des intervalles de $\mathbb{R}$:

$$A \subset \mathbb{R} \text{ est un intervalle} \iff \big(\forall (x,y) \in A^2 \quad / \quad x \leq y, \quad [x,y] \subset A\big).$$

De plus, nous allons utiliser, comme nous l'avons fait jusqu'à présent, que toute partie majorée de $\mathbb{R}$ admet une borne supérieure.

Démonstration. $\Rightarrow$: si A est connexe, et que A est le vide ou un singleton, alors A est évidemment un intervalle. Supposons donc que A a au moins deux éléments $x < y$. Si $[x,y] \not\subset A$, alors

$$\exists z \in [x,y] \quad / \quad z \notin A.$$

En particulier, $z \neq x$ et $z \neq y$. On pose alors

$$B =]-\infty, z] \cap A =]-\infty, z[\, \cap A,$$

qui est un ouvert et un fermé de A. Donc $B = \emptyset$ ou $B = A$. Mais $B \neq \emptyset$ car $x \in B$, et $B \neq A$ car $y \notin B$. On aboutit à une contradiction : $[x,y] \subset A$, et A est donc un intervalle.

$\Leftarrow$: soit I un intervalle de $\mathbb{R}$. Ici encore, si I est vide ou réduit à un singleton, il est évidemment connexe. Soient donc $a < b$ dans I. Soit de plus $f : I \longrightarrow \{0,1\}$ une application continue. Alors $f^{-1}(f(a))$ est ouvert et fermé dans I. Donc

$$B = f^{-1}(f(a)) \cap [a,b]$$

est un fermé. Il est borné, et admet donc une borne supérieure $c = \sup B$. On a vu dans la preuve de l'exemple 5.3 ci-dessus que nécessairement, $c \in B$. Comme B est aussi ouvert de $[a,b]$, on sait que

$$\exists \varepsilon > 0 \quad / \quad]c-\varepsilon, c+\varepsilon[\, \cap [a,b] \subset B.$$

Si $c < b$, alors $\exists x \in]c, b[$ tel que $x \in B$, ce qui est impossible. Donc $c \geq b$. Comme par ailleurs on sait que $c \leq b$, on obtient $c = b$. En particulier, $f(b) = f(a)$. Ceci valant pour tout $a < b$ dans I, f est constante sur I. Donc I est connexe. □

Corollaire 5.15 *Pour tout entier $n \geq 1$, l'espace vectoriel normé $\mathbb{R}^n$ est connexe.*

Démonstration. La preuve se fait par récurrence : le cas $n = 1$ est traité dans l'exemple 5.3. Pour passer de $\mathbb{R}^n$ à $\mathbb{R}^{n+1}$, on constate que $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$, et on applique la proposition 5.10. □

Corollaire 5.16 *Pour tout entier $n \geq 1$, l'espace vectoriel normé $\mathbb{C}^n$ est connexe.*

Démonstration. Ici encore, on procède par récurrence. Pour $n = 1$, on pose

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{C}, \\ (x, y) & \longmapsto & x + iy. \end{array}$$

Cette application est continue car

$$\begin{aligned} |f(x, y) - f(x', y')| = |x - x' + i(y - y')| &= \sqrt{(x - x')^2 + (y - y')^2} \\ &= \|(x, y) - (x', y')\|. \end{aligned}$$

De plus, elle est surjective, donc en appliquant la proposition 5.7, $\mathbb{C} = f(\mathbb{R}^2)$ est connexe, puisque $\mathbb{R}^2$ est connexe. Il suffit ensuite d'appliquer la proposition 5.10 pour passer du rang n au rang $n + 1$ de la récurrence. □

Théorème 5.17 (des valeurs intermédiaires) *Soit E un espace topologique connexe, et $f : E \longrightarrow \mathbb{R}$ une application continue. Soient des réels $a < b$. S'il existe $x \in E$ et $y \in E$ tels que $f(x) = a$ et $f(y) = b$, alors, pour tout $c \in [a, b]$, il existe $z \in E$ tel que $f(z) = c$.*

Démonstration. On sait, d'après la proposition 5.7, que $f(E)$ est une partie connexe de $\mathbb{R}$. Donc $f(E)$ est un intervalle, d'après le théorème 5.14. Donc si $a < b$ sont dans $f(E)$, alors $[a, b] \subset f(E)$. □

5.3 Connexité par arcs

Définition 5.18 *Soit E un espace topologique, et soient x et y deux points de E. On appelle chemin de x à y une application $\gamma : [0, 1] \longrightarrow E$ continue telle que $\gamma(0) = x$ et $\gamma(1) = y$.*

Définition 5.19 *Soit E un espace topologique. E est dit connexe par arcs si pour tout x dans E et tout y dans E, il existe un chemin de x à y.*

Exemple 5.20 *Si E est un espace vectoriel normé, alors il est connexe par arcs.*

Démonstration. Soient x et y deux points de E. Alors l'application

$$\begin{array}{rccc} \gamma : & [0,1] & \longrightarrow & E \\ & t & \longmapsto & x + t(y-x), \end{array}$$

vérifie $\gamma(0) = x$ et $\gamma(1) = y$. De plus elle est lipschitzienne, car

$$\forall t_1 \in [0,1], \quad \forall t_2 \in [0,1], \quad \|\gamma(t_1) - \gamma(t_2)\| = \|(t_1 - t_2)(y-x)\| \leq |t_1 - t_2| \|x - y\|.$$

En particulier, elle est continue. □

Exercice 5.1 *Généraliser la preuve ci-dessus pour démontrer que toute partie convexe d'un espace vectoriel normé est connexe par arcs.*

Théorème 5.21 *Soit E un espace topologique connexe par arcs. Alors il est connexe.*

Démonstration. Si E n'est pas connexe, alors $E = O_1 \cup O_2$, où O_1 et O_2 sont des ouverts non vides, et $O_1 \cap O_2 = \emptyset$. Soit donc $x \in O_1$ et $y \in O_2$. Alors, puisque E est connexe par arcs, il existe un chemin γ de x à y. En particulier, γ est continue, donc $\gamma([0,1])$ est connexe. Mais

$$\gamma([0,1]) = [O_1 \cap \gamma([0,1])] \cup [O_2 \cap \gamma([0,1])],$$

où le premier ensemble est un ouvert non vide de $\gamma([0,1])$ (il contient x), et le deuxième un ouvert non vide de $\gamma([0,1])$ (il contient y). On aboutit ainsi à une contradiction. □

Proposition 5.22 *Soit E un espace vectoriel normé, et $A \subset E$. Si A est ouvert et connexe, alors il est connexe par arcs.*

Démonstration. Si A est vide, il est évidemment connexe. Sinon, soit $x \in A$. On pose

$$U = \{y \in A \quad / \quad \exists \gamma : [0,1] \longrightarrow A \quad \text{continue} \quad / \quad \gamma(0) = x, \quad \gamma(1) = y\},$$

l'ensemble des points de A tels qu'il existe un chemin de x à y.

Montrons que U est un ouvert de A : si $y \in U$, alors il existe un chemin γ de x à y. De plus, $y \in U \subset A$ qui est ouvert, donc $\exists \varepsilon > 0$ tel que $B_\varepsilon(y) \subset A$. Soit $y' \in B_\varepsilon(y)$. On définit alors le chemin suivant :

$$\begin{array}{rccc} \tilde{\gamma} : & [0,1] & \longrightarrow & A \\ & t & \longmapsto & \begin{cases} \gamma(2t) & \text{si } t \leq \frac{1}{2}, \\ y + 2\left(t - \frac{1}{2}\right)(y' - y) & \text{si } t > \frac{1}{2}. \end{cases} \end{array}$$

Cette application est continue sur $[0, 1/2[$ car γ l'est, et sur $]1/2, 1]$ car

$$\forall (t, t') \in \left]\frac{1}{2}, 1\right]^2, \quad \|\tilde{\gamma}(t) - \tilde{\gamma}(t')\| = 2|t - t'| \, \|y - y'\|.$$

Enfin, elle est continue en $1/2$ car $\gamma(1) = y$. De plus $\tilde{\gamma}(1) = y'$: il s'agit donc d'un chemin de x à y'. Donc $y' \in U$. Donc $B_\varepsilon(x) \subset U$, ce qui prouve que U est bien un ouvert de A.

Montrons que U est un fermé de A : si $(y_n)_{n\in\mathbb{N}}$ est une suite de U qui converge vers $y \in A$, alors, comme A est ouvert, on sait que $\exists\varepsilon > 0$ tel que $B_\varepsilon(y) \subset A$. De plus, on sait aussi qu'il existe $n \in \mathbb{N}$ tel que $y_n \in B_\varepsilon(y)$. Soit γ_n un chemin de x à y_n. On pose alors

$$\begin{array}{rccl} \tilde{\gamma}: & [0,1] & \longrightarrow & A \\ & t & \longmapsto & \begin{cases} \gamma_n(2t) & \text{si } t \leq \frac{1}{2}, \\ y_n + 2\left(t - \frac{1}{2}\right)(y' - y_n) & \text{si } t > \frac{1}{2}. \end{cases} \end{array}$$

Pour les mêmes raisons que précédemment, cette application est continue, et vérifie $\tilde{\gamma}(0) = x$ et $\tilde{\gamma}(1) = y$. Donc $y \in U$.

Nous avons donc montré que U était un ouvert et un fermé de A. De plus il est non vide car $x \in U$. Donc, puisque A est connexe, $U = A$. Donc A est connexe par arcs. □

Remarque 5.23 *L'hypothèse sur E dans la proposition 5.22 n'est pas optimale : il suffit en fait que E soit un espace topologique localement connexe par arcs, c'est-à-dire vérifiant la propriété :*

$$\forall x \in E, \quad \forall V \quad \textit{voisinage de} \quad x,$$
$$\exists U \subset V \quad \textit{ouvert connexe par arcs} \quad / \quad x \in U.$$

Remarque 5.24 *La proposition 5.22 est fausse lorsque E n'est pas ouvert. Par exemple, le sous-ensemble E de $\mathbb{R}^2$ (muni de la topologie usuelle) défini par*

$$E = \Big(\{0\} \times [-1,1]\Big) \cup \left\{\left(x, \sin\frac{1}{x}\right), \quad x > 0\right\},$$

est connexe mais pas connexe par arcs.

Démonstration. Soit l'application F définie par

$$\begin{array}{rccl} F: & \mathbb{R}^{+*} & \longrightarrow & \mathbb{R}^2, \\ & x & \longmapsto & \left(x, \sin\frac{1}{x}\right). \end{array}$$

F est continue pour les topologies usuelles de $\mathbb{R}$ et $\mathbb{R}^2$. Comme $\mathbb{R}^{+*}$ est connexe, $F(\mathbb{R}^{+*})$ est connexe, d'après la proposition 5.7. Ainsi,

$$F(\mathbb{R}^{+*}) = \left\{\left(x, \sin\frac{1}{x}\right), \quad x > 0\right\}$$

est connexe. D'après la proposition 5.13, son adhérence l'est aussi.

Montrons que $\overline{F(\mathbb{R}^{+*})} = E$: tout d'abord, si $(x,y) \in E$, alors soit $x = 0$, soit $y = \sin(1/x)$. Dans le deuxième cas, $(x,y) \in F(\mathbb{R}^{+*}) \subset \overline{F(\mathbb{R}^{+*})}$. Dans le premier cas, $y \in [-1,1]$, et on construit la suite suivante :

$$\forall n \in \mathbb{N}^*, \quad x_n = \frac{1}{\arcsin(y) + 2n\pi}.$$

Ce quotient est bien défini car $\arcsin(y) \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, donc $\arcsin(y) + 2n\pi > 0$. Alors

$$F(x_n) = \left(\frac{1}{\arcsin(y) + 2n\pi}, y\right) \underset{n\to+\infty}{\longrightarrow} (0, y),$$

donc $(0, y) \in \overline{F(\mathbb{R}^{+*})}$. Réciproquement, si $(x, y) \in \overline{F(\mathbb{R}^{+*})}$, alors il existe $(x_n)_{n\in\mathbb{N}}$ une suite de $\mathbb{R}^{+*}$ telle que $F(x_n) \underset{n\to+\infty}{\longrightarrow} (x, y)$. En particulier, $x_n \underset{n\to+\infty}{\longrightarrow} x$. Donc $x \geq 0$. Si $x > 0$, comme la fonction $x \longmapsto \sin(1/x)$ est continue sur $\mathbb{R}^{+*}$, $y = \sin(1/x)$ et donc $(x, y) \in E$. Si $x = 0$, alors $y_n = \sin(1/x_n)$ est une suite de $[-1, 1]$ qui converge. Sa limite est dans $[-1, 1]$, donc $(x, y) \in \{0\} \times [-1, 1] \subset E$.

On a donc prouvé que $E = \overline{F(\mathbb{R}^{+*})}$, qui est connexe.

Supposons maintenant que E est connexe par arcs. En particulier, il existe un chemin continu de $(0, 0)$ à $(\pi, 0)$. On note $t \longmapsto \gamma(t)$ ce chemin. Sa première composante, $x(t)$ est continue, car l'application $(x, y) \longmapsto x$ est continue de $\mathbb{R}^2$ dans $\mathbb{R}$. De plus,

$$\forall t \in [0, 1], \quad \gamma(t) = (x(t), y(t)),$$

et $x(0) = 0$ et $x(1) = \pi$. De plus,

$$\forall t \in [0, 1] \quad / \quad x(t) > 0, \quad y(t) = \sin\left(\frac{1}{x(t)}\right).$$

On pose

$$t_0 = \sup\{t \geq 0 \quad / \quad x(t) = 0\}.$$

Alors $t_0 < 1$, et $\forall t > t_0, \quad x(t) > 0$, et $x(t_0) = 0$. On pose alors

$$m = \sup_{t\in[t_0,1]} x(t).$$

Cette borne supérieure existe car l'application x est continue et $[t_0, 1]$ est compacte, donc x est bornée sur $[t_0, 1]$. On définit par récurrence deux suites $(t_n)_{n\in\mathbb{N}^*}$ et $(t'_n)_{n\in\mathbb{N}^*}$ telles que

$$\forall n \in \mathbb{N}^*, \quad x(t_n) = \min\left(m, \frac{1}{n\pi}\right), \quad x(t'_n) = \min\left(m, \frac{1}{2n\pi + \frac{\pi}{2}}\right), \qquad (5.1)$$

et

$$\lim_{n\to+\infty} t_n = \lim_{n\to+\infty} t'_n = t_0. \qquad (5.2)$$

Pour $n = 1$, l'existence de t_1 et t'_1 est évidente. Et si t_n et t'_n existent, alors il existe toujours $t_{n+1} \in [t_0, t_n]$ et $t'_{n+1} \in [t_0, t'_n]$ vérifiant (5.1) au rang $n + 1$. Ceci permet de construire les deux suites vérifiant (5.1). Ces suites sont donc décroissantes minorées. Donc elle convergent. Si $t = \lim t_n > t_0$, alors $x(t) > 0$, ce qui est contradictoire. Donc on a bien la première convergence de (5.2). La deuxième se prouve de la même façon. On a donc, pour $n \geq 1/(m\pi)$:

$$\gamma(t_n) = \left(\frac{1}{n\pi}, 0\right) \underset{n\to+\infty}{\longrightarrow} (0, 0), \quad \gamma(t'_n) = \left(\frac{1}{2n\pi + \frac{\pi}{2}}, 1\right) \underset{n\to+\infty}{\longrightarrow} (0, 1).$$

Donc, par continuité de γ et (5.2), $\gamma(t_0) = (0, 0) = (0, 1)$. Ceci est contradictoire. E n'est donc pas connexe par arcs. □

Chapitre 6

Espaces complets

6.1 Suites de Cauchy

Définition 6.1 *Soit (E, d) un espace métrique. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. On dit que $(x_n)_{n\in\mathbb{N}}$ est de Cauchy si*

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \quad \forall p \in \mathbb{N}, \quad d\left(x_{n+p}, x_n\right) \leq \varepsilon.$$

Citons des formes équivalentes de la définition 6.1 :

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \forall q \geq n_0, \quad d(x_n, x_q) \leq \varepsilon.$$

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \quad \sup_{q \geq n_0} d(x_n, x_q) \leq \varepsilon.$$

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \sup_{n \geq n_0,\ q \geq n_0} d(x_n, x_q) \leq \varepsilon.$$

$$\lim_{n\to+\infty} \left(\sup_{q \geq n} d(x_n, x_q) \right) = 0.$$

Une autre façon de dire cela est que la suite $d(x_{n+p}, x_n)$ converge vers 0 quand $n \to +\infty$, uniformément en $p \in \mathbb{N}$.

Exercice 6.1 *Démontrer que les quatre propriétés ci-dessus sont bien équivalentes à la définition 6.1.*

Exemple 6.2 *Soient les suites réelles $(x_n)_{n\in\mathbb{N}^*}$, $(y_n)_{n\in\mathbb{N}^*}$, $(z_n)_{n\in\mathbb{N}^*}$, définies par*

$$\forall n \in \mathbb{N}, \quad x_n = \frac{1}{n}, \quad y_n = \sqrt{n}, \quad z_n = \sum_{k=1}^{n} \frac{1}{k}.$$

Alors $(x_n)_{n\in\mathbb{N}^*}$ est de Cauchy, mais pas $(y_n)_{n\in\mathbb{N}^*}$ ni $(z_n)_{n\in\mathbb{N}^*}$.

Démonstration. Commençons pas prouver que $(x_n)_{n\in\mathbb{N}}$ est de Cauchy : on a

$$\forall n\in\mathbb{N},\quad \forall p\in\mathbb{N},\quad x_{n+p}-x_n=\frac{1}{n+p}-\frac{1}{n}=\frac{p}{n(n+p)}.$$

En particulier,

$$\forall n\in\mathbb{N},\quad \forall p\in\mathbb{N},\quad 0\le x_{n+p}-x_n\le\frac{1}{n}.$$

Cette borne supérieure ne dépend pas de p, et tend vers 0 quand $n\to+\infty$. Donc la suite $(x_n)_{n\in\mathbb{N}}$ est bien de Cauchy.

Etudions maintenant la suite $(y_n)_{n\in\mathbb{N}}$:

$$\forall n\in\mathbb{N},\quad \forall p\in\mathbb{N},\quad y_{n+p}-y_n=\sqrt{n+p}-\sqrt{n}=\frac{p}{\sqrt{n}+\sqrt{n+p}}.$$

En particulier, on a bien $\lim_{n\to+\infty}(y_{n+p}-y_n)=0$, mais cette limite n'est pas uniforme en p. En effet, on a, en choisissant $p=n$,

$$y_{2n}-y_n=\frac{n}{\sqrt{n}+\sqrt{n+n}}=\frac{\sqrt{n}}{1+\sqrt{2}}\underset{n\to+\infty}{\longrightarrow}+\infty.$$

Pour ce qui concerne $(z_n)_{n\in\mathbb{N}}$, on procède de la même façon, en étudiant la différence $z_{2n}-z_n$:

$$z_{2n}-z_n=\sum_{k=n+1}^{2n}\frac{1}{k}\ge\sum_{k=n+1}^{2n}\frac{1}{2n}=\frac{n}{2n}=\frac{1}{2}.$$

Donc la suite $(z_n)_{n\in\mathbb{N}}$ n'est pas de Cauchy. □

Proposition 6.3 *Soit (E,d) un espace métrique, et soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans E. Alors toute suite extraite de $(x_n)_{n\in\mathbb{N}}$ est de Cauchy.*

Démonstration. Soit une suite extraite de $(x_n)_{n\in\mathbb{N}}$. Elle s'écrit donc $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$, avec $\varphi:\mathbb{N}\to\mathbb{N}$ strictement croissante. On sait qu'alors

$$\forall n\in\mathbb{N},\quad \varphi(n)\ge n.$$

(Voir pour cela la preuve de la proposition 3.46.) Soit $\varepsilon>0$. On sait qu'il existe $n_0\in\mathbb{N}$ tel que

$$\forall n\ge n_0,\quad \forall p\ge n_0,\quad d(x_n,x_p)\le\varepsilon.$$

Si $n\ge n_0$ et $p\ge n_0$, alors $\varphi(n)\ge n\ge n_0$, et $\varphi(p)\ge p\ge n_0$, donc

$$d\left(x_{\varphi(n)},x_{\varphi(p)}\right)\le\varepsilon.$$

La suite extraite considérée est donc bien de Cauchy. □

Proposition 6.4 *Soit (E,d) un espace métrique, et soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. Si elle converge dans E, elle est de Cauchy.*

Démonstration. Si la suite $(x_n)_{n\in\mathbb{N}}$ converge, on a alors, en notant x sa limite,

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \quad d(x, x_n) \leq \frac{\varepsilon}{2}.$$

Donc

$$\forall n \geq n_0, \quad \forall p \in \mathbb{N}, \quad d(x, x_n) \leq \frac{\varepsilon}{2} \quad \text{et} \quad d(x, x_{n+p}) \leq \frac{\varepsilon}{2}.$$

Donc, en appliquant l'inégalité triangulaire, on a donc :

$$d(x_n, x_{n+p}) \leq d(x_n, x) + d(x, x_{n+p}) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Ce qui prouve que $(x_n)_{n\in\mathbb{N}}$ est de Cauchy. □

Remarque 6.5 *La réciproque est fausse. Par exemple, si $E =]0,1[$ est muni de la distance usuelle, la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = \frac{1}{n+1}$ est une suite de Cauchy d'après l'exemple 6.2, mais ne converge pas dans E.*

Démonstration. Supposons que la suite $(x_n)_{n\in\mathbb{N}}$ converge dans E. Alors il existerait $x \in E$ tel que

$$\lim_{n\to+\infty} \left| \frac{1}{n+1} - x \right| = 0.$$

D'où $x = 0$, ce qui est contradictoire avec l'hypothèse $x \in E =]0,1[$. □

En revanche, une suite de Cauchy qui admet une valeur d'adhérence est convergente :

Proposition 6.6 *Soit (E, d) un espace métrique, et soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans E. Alors*

$$(x_n)_{n\in\mathbb{N}} \text{ converge} \iff (x_n)_{n\in\mathbb{N}} \text{ a une valeur d'adhérence.}$$

Démonstration. Si $x_n \underset{n\to+\infty}{\longrightarrow} a$, alors a est valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$.

Réciproquement, si a est une valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$, alors,

$$\forall \varepsilon > 0, \quad \forall n_0 \in \mathbb{N}, \quad \exists q \geq n_0 \quad / \quad d(x_q, a) \leq \frac{\varepsilon}{2}.$$

De plus, puisque $(x_n)_{n\in\mathbb{N}}$ est de Cauchy,

$$\forall \varepsilon > 0, \quad \exists n_1 \in \mathbb{N}, \quad / \quad \forall n \geq n_1, \quad \forall q \geq n_1, \quad d(x_n, x_q) \leq \frac{\varepsilon}{2}.$$

On choisit donc $q \geq \max(n_0, n_1)$, qui vérifie $d(x_q, a) \leq \varepsilon/2$, et on a donc :

$$\forall n \geq n_1, \quad d(x_n, a) \leq d(x_n, x_q) + d(x_q, a) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Donc la suite $(x_n)_{n\in\mathbb{N}}$ converge vers a. □

Proposition 6.7 *Soient E et F deux espaces métriques, et soit $f : E \to F$ une application uniformément continue. Soit $(x_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E. Si elle est de Cauchy dans E, alors la suite $\left(f(x_n)\right)_{n\in\mathbb{N}}$ est de Cauchy dans F.*

Démonstration. L'uniforme continuité de f équivaut à :

$$\forall\varepsilon > 0, \quad \exists\delta > 0 \quad / \quad d_E(x, y) \leq \delta \quad \Rightarrow \quad d_F(f(x), f(y)) \leq \varepsilon.$$

Par ailleurs, on sait que $(x_n)_{n\in\mathbb{N}}$ est de Cauchy, donc

$$\exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \quad \forall p \in \mathbb{N}, \quad d_E(x_n, x_{n+p}) \leq \delta.$$

Donc, pour cette valeur de n_0, on a

$$\forall n \geq n_0, \quad \forall p \in \mathbb{N}, \quad d_F\left(f(x_n), f(x_{n+p})\right) \leq \varepsilon.$$

Ceci prouve bien que la suite $\left(f(x_n)\right)_{n\in\mathbb{N}}$ est de Cauchy dans F. □

6.2 Complétude

Définition 6.8 *Soit (E, d) un espace métrique. On dit que (E, d) est complet si toute suite de Cauchy de E converge dans E.*

Remarque 6.9 *Il ne s'agit pas d'une notion topologique, au sens où deux espaces métriques peuvent être homéomorphes alors qu'un est complet et l'autre non. En effet, $]0, 1[$ muni de la distance usuelle n'est pas complet, alors qu'il est homéomorphe à $\mathbb{R}$ muni de la distance usuelle (voir l'exemple 3.35), qui, lui, est complet (voir le théorème 6.14 ci-dessous).*

Définition 6.10 *Soit E un espace vectoriel normé. On dit que E est un espace de Banach s'il est complet pour la distance associée à sa norme.*

Proposition 6.11 *Tout espace métrique compact est complet.*

Démonstration. Soit (E, d) un espace métrique compact. Soit une suite de Cauchy dans (E, d). D'après le théorème de Bolzano-Weierstrass, elle a une valeur d'adhérence. Donc, d'après la proposition 6.6, elle converge. □

Proposition 6.12 *Soit (E, d) un espace métrique complet. Soit $F \subset E$. Alors (F, d) est complet si et seulement si F est fermé dans E.*

Démonstration. Supposons que (F, d) est complet. Soit une suite $(x_n)_{n\in\mathbb{N}}$ à valeurs dans F, qui converge dans E vers $x \in E$. Alors, elle est de Cauchy pour la distance d, donc de Cauchy dans (F, d). Comme cet espace est complet, elle converge dans F. Sa limite pour d est x. Donc $x \in F$. Donc F est fermé.

Réciproquement, supposons que F est un fermé de E. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans (F, d). Elle est de Cauchy pour la distance d, donc dans (E, d). Donc elle converge dans (E, d) puisque cet espace est complet. Soit x sa limite. Comme F est fermé, $x \in F$. Donc (F, d) est bien complet. □

Proposition 6.13 *Soient* $(E_1, d_1), (E_2, d_2), \dots (E_p, d_p)$ *des espaces métriques complets. Alors le produit* $E = \prod_{i=1}^{n} E_i$ *muni de la distance produit*

$$\forall x \in E, \quad \forall y \in E, \quad d(x, y) = \sup_{1 \le i \le p} d_i(x_i, y_i)$$

est complet.

Démonstration. Soit $(X_n)_{n\in\mathbb{N}}$ une suite à valeurs dans E, que l'on suppose être de Cauchy. Alors pour tout $n \in \mathbb{N}$, $X_n = \left(x_1^n, x_2^n, \dots x_p^n\right)$, avec $x_i^n \in E_i$. De plus,

$$\forall i \in \{1, 2, \dots, p\}, \quad \forall n \in \mathbb{N}, \quad \forall q \in \mathbb{N}, \quad d_i(x_i^n, x_i^{n+q}) \le d(X_n, X_{n+q}).$$

Donc pour chaque i, la suite $(x_i^n)_{n\in\mathbb{N}}$ est de Cauchy. Donc elle converge dans (E_i, d_i). On appelle x_i sa limite, et $X = (x_1, x_2, \dots x_p) \in E$. On a alors

$$d(X_n, X) = \max_{1 \le i \le p} d_i(x_i^n, x_i) \underset{n\to+\infty}{\longrightarrow} 0.$$

Donc $(X_n)_{n\in\mathbb{N}}$ converge vers X dans (E, d). Donc (E, d) est complet. □

Théorème 6.14 $\mathbb{R}$ *muni de la distance usuelle est complet.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $\mathbb{R}$. Alors cette suite est bornée. En effet, en appliquant la définition d'une suite de Cauchy pour $\varepsilon = 1$, on a

$$\exists n_0 \in \mathbb{N} \quad / \quad \forall n \ge n_0, \quad |x_{n_0} - x_n| \le 1.$$

On pose alors

$$M = 1 + \max_{1 \le q \le n_0} |x_q|,$$

et on a alors

$$\forall n \in \mathbb{N}, \quad |x_n| \le M.$$

L'ensemble $[-M, M]$ est une partie compacte de $\mathbb{R}$, d'après la proposition 4.21. Donc la suite $(x_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence, d'après le théorème de Bolzano-Weierstrass. On applique alors la proposition 6.6, et on déduit que la suite $(x_n)_{n\in\mathbb{N}}$ converge. □

Corollaire 6.15 *L'espace vectoriel normé* $\mathbb{R}^n$ *est complet.*

Démonstration. On rappelle que toutes les normes sont équivalentes sur $\mathbb{R}^n$, donc il suffit de prouver le résultat pour la norme $\|\cdot\|_\infty$. On applique la proposition 6.13, et on obtient le résultat. □

Corollaire 6.16 *L'ensemble* $\mathbb{C}$ *des nombres complexes, muni de la distance* $d(z, z') = |z - z'|$, *où* $|z|$ *désigne le module de* z, *est complet.*

Démonstration. Remarquons qu'on ne peut pas invoquer le fait que $\mathbb{C}$ est homéomorphe à $\mathbb{R}^2$, car cela n'implique pas que $\mathbb{C}$ est complet, même si $\mathbb{R}^2$ l'est. Soit donc une suite de Cauchy $(z_n)_{n\in\mathbb{N}}$ dans $\mathbb{C}$. On note $(x_n)_{n\in\mathbb{N}}$ la suite de ses parties réelles, et $(y_n)_{n\in\mathbb{N}}$ la suite de ses parties imaginaires. Chacune d'entre elle est de Cauchy dans $\mathbb{R}$, car

$$\forall n \in \mathbb{N}, \quad \forall p \in \mathbb{N}, \quad |x_n - x_{n+p}| \leq |z_n - z_{n+p}| \quad \text{et} \quad |y_n - y_{n+p}| \leq |z_n - z_{n+p}|.$$

Donc les deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ convergent dans $\mathbb{R}$. Remarquons maintenant que l'application

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{C}, \\ (x,y) & \longmapsto & x+iy, \end{array}$$

est continue. Donc, puisque $z_n = f(x_n, y_n)$, la suite $(z_n)_{n\in\mathbb{N}}$ est convergente. □

Proposition 6.17 *Soit (K, d_K) un espace métrique compact et (F, d_F) un espace métrique complet. Alors l'espace métrique $C(K, F)$ des applications continues de K dans F muni de la distance*

$$d(f,g) = \sup_{x\in K} d_F(f(x), g(x)),$$

est complet.

Démonstration. Vérifions tout d'abord que d est bien une distance. On vérifie donc les trois propriétés de la définition 2.1.
$(i) : \forall (f,g) \in C(K,F)^2, \quad d(f,g) = 0 \iff (\forall x \in K, \quad d_F(f(x), g(x)) = 0) \iff (\forall x \in K, \quad f(x) = g(x)) \iff f = g.$
$(ii) : \forall (f,g) \in C(K,F)^2,$

$$d(f,g) = \sup_{x\in K} d_F(f(x), g(x)) = \sup_{x\in K} d_F(g(x), f(x)) = d(g,f).$$

$(iii) : \forall (f,g,h) \in C(K,F)^3$, on a

$$\forall x \in K, \quad d_F(f(x), h(x)) \leq d_F(f(x), g(x)) + d_F(g(x), h(x)) \leq d(f,g) + d(g,h).$$

Donc, en prenant la borne supérieure par rapport à x dans cet inégalité, on obtient

$$d(f,h) \leq d(f,g) + d(g,h).$$

Démontrons maintenant que cette distance rend l'espace $C(K,F)$ complet. Soit donc $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans cet espace. Alors, on a, pour tout $x \in K$,

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \quad \forall p \in \mathbb{N}, \quad \forall x \in K,$$
$$d_F(f_n(x), f_{n+p}(x)) \leq d(f_n, f_{n+p}) \leq \frac{\varepsilon}{3}.$$

Ainsi, la suite $(f_n(x))_{n\in\mathbb{N}}$ est de Cauchy dans F. Donc elle converge. On note $f(x)$ sa limite. On a ainsi construit une application $f : K \to F$ telle que

$$\forall x \in K, \quad f_n(x) \underset{n\to+\infty}{\longrightarrow} f(x),$$

cette convergence ayant lieu pour la distance d_F. On remarque ensuite que la majoration ci-dessus est uniforme en x et en p. En faisant tendre p vers l'infini dans cette inégalité, on obtient

$$\forall \varepsilon > 0, \quad \exists n_0 \in \mathbb{N} \quad / \quad \forall n \geq n_0, \quad \forall x \in K, \quad d_F(f_n(x), f(x)) \leq \frac{\varepsilon}{3}. \tag{6.1}$$

Soient maintenant x, y dans K. L'inégalité triangulaire implique

$$\begin{aligned} d_F(f(x), f(y)) &\leq d_F(f(x), f_{n_0}(x)) + d_F(f_{n_0}(x), f_{n_0}(y)) + d_F(f_{n_0}(y), f(y)) \\ &\leq \frac{2\varepsilon}{3} + d_F(f_{n_0}(x), f_{n_0}(y)). \end{aligned}$$

La fonction f_{n_0} est continue, donc uniformément continue d'après le théorème de Heine. Donc il existe $\delta > 0$ tel que

$$\forall (x,y) \in K^2, \quad d_K(x,y) \leq \delta \Rightarrow d_F(f_{n_0}(x), f_{n_0}(y)) \leq \frac{\varepsilon}{3}.$$

Ainsi,

$$\forall (x,y) \in K^2, \quad d_K(x,y) \leq \delta \Rightarrow d_F(f(x), f(y)) \leq \varepsilon.$$

Donc f est uniformément continue sur K, donc continue. Enfin, en prenant la borne supérieure en x dans l'inégalité (6.1), on obtient que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f pour la distance d. □

En appliquant ce résultat au cas où $K = [a,b]$ est un intervalle borné, on a donc :

Corollaire 6.18 *Soient $a < b$ deux réels. Soit $E = C([a,b], \mathbb{K})$ l'ensemble des fonctions continues. Muni de la norme*

$$\forall f \in E, \quad \|f\|_\infty = \sup_{a\leq x\leq b} |f(x)|,$$

cet espace est complet.

Rappelons la définition des espaces de suites $\ell^p(\mathbb{K})$ (définition 1.52) : pour tout réel $p \geq 1$,

$$\ell^p(\mathbb{K}) = \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}, \ \sum_{n\geq 0} |u_n|^p < +\infty \right\},$$

et

$$\ell^\infty(\mathbb{K}) = \left\{ (u_n)_{n\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}}, \ \sup_{n\geq 0} |u_n| < +\infty \right\}.$$

Sur ces espaces vectoriels, on rappelle la définition des normes correspondantes $\|\cdot\|_p$ (définition 1.53 et proposition 1.59) : pour tout réel $p \geq 1$, et pour $u \in \ell^p(\mathbb{K})$,

$$\|u\|_p = \left(\sum_{n\geq 0} |u_n|^p\right)^{1/p},$$

et

$$\|u\|_\infty = \sup_{n\geq 0} |u_n|.$$

On a alors

Théorème 6.19 (de Riesz-Fischer) *L'espace $\ell^p(\mathbb{K})$ muni de la norme $\|\cdot\|_p$ est un espace de Banach.*

Démonstration. La proposition 1.59 assure que $(\ell^p(\mathbb{K}), \|\cdot\|_p)$ est un espace vectoriel normé. Il reste donc à démontrer qu'il est complet. Soit $(u^k)_{k\in\mathbb{N}}$ une suite de Cauchy dans $\ell^p(\mathbb{K})$. Alors pour chaque $k \in \mathbb{N}$, $u^k = \left(u_n^k\right)_{n\in\mathbb{N}}$ est une suite de $\ell^p(\mathbb{K})$, c'est-à-dire que

$$\sum_{n\geq 0} \left|u_n^k\right|^p < +\infty \quad \text{si} \quad p \geq 1,$$

$$\sup_{n\geq 0} \left|u_n^k\right| < +\infty \quad \text{si} \quad p = +\infty.$$

De plus, comme cette suite est de Cauchy, on a

$$\forall \varepsilon > 0, \quad \exists K \in \mathbb{N} \quad / \quad \forall k, j \geq K, \quad \left\|u^k - u^j\right\|_p \leq \varepsilon.$$

En particulier, pour tout $n \in \mathbb{N}$, on a $\left|u_n^k - u_n^j\right| \leq \left\|u^k - u^j\right\|_p$, donc

$$\forall \varepsilon > 0, \quad \exists K \in \mathbb{N} \quad / \quad \forall k, j \geq K, \quad \left|u_n^k - u_n^j\right| \leq \varepsilon.$$

Donc la suite de nombres réels $\left(u_n^k\right)_{k\in\mathbb{N}}$ est de Cauchy dans $\mathbb{K}$. D'après le théorème 6.14 (dans le cas réel) et le corollaire 6.16 (dans le cas complexe), cette suite converge. On note $u_n \in \mathbb{K}$ sa limite :

$$\forall n \in \mathbb{N}, \quad \lim_{k\to+\infty} u_n^k = u_n.$$

A partir de maitenant, on suppose p fini, on traitera le cas $p = +\infty$ ensuite. On a, pour $k, j \geq K$:

$$\left\|u^k - u^j\right\|_p \leq \varepsilon, \tag{6.2}$$

donc $\left\|u^j\right\|_p \leq \varepsilon + \left\|u^k\right\|_p$, d'où

$$\left(\sum_{n=0}^{N} |u_n^j|^p\right)^{1/p} \leq \left\|u^j\right\|_p \leq \varepsilon + \|u^k\|_p.$$

En faisant tendre j vers l'infini dans cette inégalité, on obtient donc

$$\left(\sum_{n=0}^{N}|u_n|^p\right)^{1/p}\leq \varepsilon+\|u^k\|_p.$$

Comme le majorant ne dépend pas de N, on obtient donc que $u\in\ell^p(\mathbb{K})$ en faisant tendre N vers l'infini. Enfin, en reprenant (6.2), on a, pour $k,j\geq K$,

$$\forall N\in\mathbb{N},\quad \left(\sum_{n=0}^{N}\left|u_n^k-u_n^j\right|^p\right)^{1/p}\leq\left\|u^k-u^j\right\|_p\leq\varepsilon.$$

En faisant tendre j vers l'infini, on en déduit

$$\forall k\geq K,\quad\forall N\in\mathbb{N},\quad\left(\sum_{n=0}^{N}\left|u_n^k-u_n\right|^p\right)^{1/p}\leq\varepsilon.$$

On fait maintenant tendre N vers l'infini dans cette inégalité, ce qui donne

$$\forall k\geq K,\quad\left\|u-u^k\right\|_p\leq\varepsilon.$$

Donc la suite $\left(u^k\right)_{k\in\mathbb{N}}$ converge bien vers u dans $\ell^p(\mathbb{K})$.

Traitons maintenant le cas $p=+\infty$: l'inégalité (6.2) devient alors, pour tout $k,j\geq K$, $\left\|u^k-u^j\right\|_\infty\leq\varepsilon$, donc

$$\forall N\in\mathbb{N},\quad\sup_{0\leq n\leq N}\left|u_n^k-u_n^j\right|\leq\varepsilon. \tag{6.3}$$

On en déduit immédiatement que

$$\sup_{0\leq n\leq N}\left|u_n^j\right|\leq\varepsilon+\sup_{0\leq n\leq N}\left|u_n^k\right|\leq\varepsilon+\left\|u^k\right\|_\infty.$$

Donc, en prenant la limite $j\to+\infty$, on obtient que $u\in\ell^\infty(\mathbb{K})$. De plus, en faisant $j\to+\infty$ directement dans (6.3), on déduit

$$\sup_{0\leq n\leq N}\left|u_n-u_n^k\right|\leq\varepsilon.$$

En faisant là aussi tendre N vers l'infini,

$$\forall k\geq K,\quad\left\|u-u^k\right\|_\infty\leq\varepsilon.$$

Donc la suite $\left(u^k\right)_{k\in\mathbb{N}}$ converge bien vers u dans $\ell^\infty(\mathbb{K})$. □

Nous avons vu (corollaire 6.18) que l'ensemble des fonctions continues est complet pour la norme uniforme. En revanche, il ne l'est pas pour la norme p, définie par (voir la définition 1.62)

$$\forall f\in C([a,b],\mathbb{K}),\quad\|f\|_p=\left(\int_a^b|f(t)|^p dt\right)^{1/p}.$$

Pour prouver cela, on chosit $a = 0$ et $b = 2$ (ce qui suit se généralise sans problème à tous les couples $a < b$), et on construit une suite de Cauchy pour cette norme, mais qui ne converge pas dans $C([0,2],\mathbb{K})$:

$$\forall n \in \mathbb{N}, \quad f_n(x) = \begin{cases} x^n & \text{si } x \leq 1, \\ 1 & \text{si } x > 1. \end{cases}$$

Pour tout $n \in \mathbb{N}$, cette fonction est continue. De plus, on a, pour tous les entiers n et q,

$$\begin{aligned} \|f_n - f_{n+q}\|_p^p = \int_0^1 \left|x^n - x^{n+q}\right|^p dx &= \int_0^1 \left(x^n - x^{n+q}\right)^p dx \\ &\leq \int_0^1 x^{np} dx = \frac{1}{1+np}. \end{aligned}$$

On en déduit donc que

$$\forall n \in \mathbb{N}, \quad \forall q \in \mathbb{N}, \quad \|f_n - f_{n+q}\|_p \leq \frac{1}{(1+np)^{1/p}}.$$

Cette borne supérieure ne dépend pas de $q \in \mathbb{N}$, et elle tend vers 0 quand $n \to +\infty$. Donc la suite $(f_n)_{n\in\mathbb{N}}$ est de Cauchy. Supposons maintenant que cette suite converge vers $f \in C([0,2],\mathbb{K})$. Alors on a

$$\lim_{n\to+\infty} \left(\int_0^2 |f(x) - f_n(x)|^p \, dx \right) = 0.$$

Donc, pour tout $\delta \in]0,1[$,

$$0 \leq \int_0^\delta |f(x) - x^n|^p \, dx \leq \left(\int_0^2 |f(x) - x^n|^p \, dx \right) \underset{n\to+\infty}{\longrightarrow} 0.$$

Par ailleurs, la suite de fonctions $x \mapsto |f(x) - x^n|^p$ converge uniformément vers $|f(x)|^p$ sur $[0,\delta]$. Donc

$$\int_0^\delta |f(x)|^p dx = 0, \quad \forall \delta \in]0,1[.$$

Comme f est continue, elle est donc nulle sur $[0,\delta]$. Ceci valant pour tout $\delta < 1$, on en déduit

$$\forall x \in [0,1], \quad f(x) = 0.$$

De même, on a

$$\lim_{n\to+\infty} \int_1^2 |f(x) - 1| dx = 0,$$

donc $f = 1$ sur $[1,2]$. Donc f ne peut pas être continue.

6.3 Théorème du point fixe

Définition 6.20 *Soient E, F des espaces métriques. Soit $f : E \to F$ une application. Elle est dite contractante si*

$$\exists k \in [0,1[\quad / \quad \forall x \in E, \quad \forall y \in E, \quad d_F(f(x)f(y)) \leq k d_E(x,y).$$

Remarque 6.21 *Une application contractante et donc en particulier Lipschitzienne, donc continue.*

Théorème 6.22 (du point fixe de Picard) *Soit E un espace métrique complet non vide. Soit $f : E \to E$ une application contractante. Alors*

$$\exists! a \in E \quad / \quad f(a) = a.$$

Démonstration. Commençons par prouver l'existence de a : comme $E \neq \emptyset$, on sait qu'il existe $x_0 \in E$. On construit par récurrence la suite des itérés de f à partir de x_0 :

$$\forall n \in \mathbb{N}, \quad x_{n+1} = f(x_n).$$

On a donc, puisque f est contractante,

$$\forall n \in \mathbb{N}^*, \quad d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq k d(x_n, x_{n-1}).$$

On en déduit par récurrence que

$$\forall n \in \mathbb{N}, \quad d(x_{n+1}, x_n) \leq k^n d(x_1, x_0).$$

En effet, cette relation est vraie pour $n = 0$. Si de plus elle est vraie au rang $n-1$, on en déduit

$$d(x_{n+1}, x_n) \leq k d(x_n, x_{n-1}) \leq k k^{n-1} d(x_1, x_0) = k^n d(x_1, x_0).$$

Démontrons maintenant que $(x_n)_{n \in \mathbb{N}}$ est une suite de Cauchy. On calcule :

$$\forall n \in \mathbb{N}, \quad \forall p \in \mathbb{N}, \quad d(x_n, x_{n+p}) \leq \sum_{q=0}^{p-1} d(x_{n+q}, x_{n+q+1}) \leq \sum_{q=0}^{p-1} k^{n+q} d(x_1, x_0).$$

On calcule cette somme géométrique : puisque $k \in [0,1[$,

$$\sum_{q=0}^{p-1} k^{n+q} = k^n \sum_{q=0}^{p-1} k^q = k^n \frac{1-k^p}{1-k},$$

Donc

$$\forall n \in \mathbb{N}, \quad \forall p \in \mathbb{N}, \quad d(x_n, x_{n+p}) \leq k^n \frac{1-k^p}{1-k} d(x_1, x_0) \leq k^n \frac{1}{1-k} d(x_1, x_0).$$

Ce majorant est indépendant de p, et tend vers 0 quand $n \to +\infty$. Donc $(x_n)_{n \in \mathbb{N}}$ est de Cauchy. Donc elle converge, puisque E est complet. Soit $a \in E$ sa limite.

En passant à la limite dans $x_{n+1} = f(x_n)$, on obtient, puisque f est continue, $a = f(a)$.

Démontrons maitenant l'unicité du point fixe a : s'il existe un autre point fixe $b \in E$, on a $f(b) = b$, donc

$$d(a, b) = d(f(a), f(b)) \leq kd(a, b).$$

Comme $k < 1$, cette inégalité ne peut être satisfaite que si $d(a, b) = 0$, autrement dit $a = b$. □

Corollaire 6.23 *Sous les hypothèses du théorème 6.22, l'unique point fixe a est limite de toute suite de la forme $x_{n+1} = f(x_n)$.*

En anticipant un peu sur la partie calcul différentiel de ce cours, on a la remarque suivante :

Remarque 6.24 *Dans le cas $E = \mathbb{R}$ (ou bien E est un intervalle de $\mathbb{R}$), et si l'application f est dérivable, un critère simple pour prouver que f est contractante, est de démontrer que $\forall x \in E, \quad |f'(x)| \leq k < 1$.*

Démonstration. En effet, on a, pour tous $x < y$ dans E,

$$f(y) - f(x) = \int_x^y f'(t)dt,$$

donc

$$|f(y) - f(x)| \leq \int_x^y |f'(t)|dt \leq \int_x^y kdt = k(y - x).$$

Donc f est bien contractante. □

Citons quelques exemples d'application de ce théorème, très important en analyse :

- le théorème de Cauchy-Lipschitz, qui permet de démontrer l'existence et l'unicité d'une solution à une équation différentielle ;
- le théorème d'inversion locale, ainsi que son "cousin" le théorème des fonctions implicites ;
- la méthode de Newton pour résoudre une équation de la forme $g(x) = 0$. Il s'agit de construire une fonction f dérivable, telle que si $g(x) = 0$, alors $f(x) = x$ et $f'(x) = 0$. Ceci assurera qu'au voisinage du point fixe a de f, f est contractante car on aura $|f'(x)| \leq k < 1$ pour x proche de a (voir la remarque 6.24). Une façon d'obtenir cette propriété est d'utiliser la fonction
$$f(x) = x - \frac{g(x)}{g'(x)}.$$

Exercice 6.2 *Démontrer que, si g est dérivable sur $E \subset \mathbb{R}$ et que sa dérivée ne s'annule pas, alors la fonction f définie ci-dessus vérifie les propriétés voulues, à savoir :*

$$\forall x \in E, \quad g(x) = 0 \Rightarrow (f(x) = x \quad et \quad f'(x) = 0).$$

6.4 Prolongement par continuité

Rappelons la définition 3.23 : pour tout espace topologique E, $A \subset E$ est dit dense dans E si $\overline{A} = E$.

Théorème 6.25 *Soient E et F deux espaces métriques tels que F est complet. Soit $A \subset E$ une partie dense de E. Soit $f : A \to F$ une application uniformément continue. Alors $\exists! \tilde{f} : E \to F$ continue telle que $\tilde{f}_{|A} = f$.*

Démonstration. On commence par démontrer l'existence. Pour cela, on construit d'abord $\tilde{f}$. Soit $x \in E$. Puisque A est dense dans E, il existe $(x_n)_{n\in\mathbb{N}}$ une suite de A qui converge vers x. Cette suite est donc de Cauchy dans E, donc, d'après la proposition 6.7, la suite $(f(x_n))_{n\in\mathbb{N}}$ est de Cauchy dans F. Comme F est complet, elle converge. On appelle donc $\tilde{f}(x)$ sa limite :

$$\tilde{f}(x) = \lim_{n\to+\infty} f(x_n).$$

Pour démontrer que $\tilde{f}$ est bien une application, il nous faut vérifier que cette limite ne dépend pas de la suite $(x_n)_{n\in\mathbb{N}}$ utilisée pour la définir. Soit donc une autre suite $(y_n)_{n\in\mathbb{N}}$ qui converge elle aussi vers x. Alors, pour tout $\varepsilon > 0$, on sait qu'il existe $\delta > 0$ tel que

$$\forall (x,y) \in A^2, \quad d_E(x,y) \leq \delta \quad \Rightarrow \quad d_F(f(x), f(y)) \leq \varepsilon.$$

Par ailleurs, les convergences des deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ impliquent qu'il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad d_E(x_n, x) \leq \frac{\delta}{2}, \quad \text{et} \quad d_E(y_n, x) \leq \frac{\delta}{2}.$$

En appliquant l'inégalité triangulaire, on en déduit que $d(x_n, y_n) \leq d(x_n, x) + d(x, y_n) \leq \delta$, donc que

$$\forall n \geq n_0, \quad d_F(f(x_n), f(y_n)) \leq \varepsilon.$$

Ainsi,

$$d_F\left(\lim_{n\to+\infty} f(x_n), \lim_{n\to+\infty} f(y_n)\right) = \lim_{n\to+\infty} d_F(f(x_n), f(y_n)) = 0.$$

Ceci prouve que $\tilde{f}$ est bien une application. Démontrons maintenant qu'elle est continue : soit $\varepsilon > 0$. On sait alors qu'il existe $\delta > 0$ tel que

$$\forall (x,y) \in A^2, \quad d_E(x,y) \leq \delta \quad \Rightarrow \quad d_F(f(x), f(y)) \leq \varepsilon.$$

Soit alors un couple $(x,y) \in E^2$ tel que $d_E(x,y) \leq \delta/3$. On sait qu'il existe deux suite $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ de A qui convergent vers x et y, respectivement. Donc il existe $n_0 \in \mathbb{N}$ tel que

$$\forall n \geq n_0, \quad d_E(x_n, x) \leq \frac{\delta}{3}, \quad d_E(y_n, y) \leq \frac{\delta}{3}.$$

Ainsi,

$$\forall n \geq n_0, \quad d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n) \leq \frac{\delta}{3} + \frac{\delta}{3} + \frac{\delta}{3} = \delta.$$

Donc $d_F(f(x_n), f(y_n)) \leq \varepsilon$. De plus,

$$\tilde{f}(x) = \lim_{n \to +\infty} f(x_n), \quad \tilde{f}(y) = \lim_{n \to +\infty} f(y_n),$$

et donc en faisant tendre n vers l'infini, on obtient

$$d_F(\tilde{f}(x), \tilde{f}(y)) \leq \varepsilon.$$

Ceci prouve que $\tilde{f}$ est uniformément continue, donc continue.

Démontrons maintenant l'unicité : supposons qu'il existe $g : E \to F$ et $h : E \to F$ continues telles que $g_{|A} = h_{|A} = f$. Alors, pour tout $x \in E$, il existe une suite $(x_n)_{n \in \mathbb{N}}$ de a qui converge vers x. Donc

$$\forall n \in \mathbb{N}, \quad g(x_n) = h(x_n), \quad \text{et} \quad \begin{cases} \lim\limits_{n \to +\infty} h(x_n) = h(x) \\ \lim\limits_{n \to +\infty} g(x_n) = g(x). \end{cases}$$

Donc, puisque E est un espace métrique, donc est séparé, $h(x) = g(x)$. □

Remarque 6.26 *Comme on vient de le voir dans la preuve, $\tilde{f}$ n'est pas seulement continue, mais uniformément continue.*

Exemple 6.27 *L'application*

$$\begin{array}{cccc} f : & \mathbb{R}^* & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & e^{-\frac{1}{x^2}}, \end{array}$$

est uniformémement continue (le démontrer en exercice), donc admet un unique prolongement $\tilde{f}$ sur $\mathbb{R}$. De plus, $\lim\limits_{x \to 0} f(x) = 0$, donc ce prolongement est défini par $\tilde{f}(0) = 0$, c'est-à-dire

$$\forall x \in \mathbb{R}, \quad \tilde{f}(x) = \begin{cases} e^{-\frac{1}{x^2}} & si\ x \neq 0, \\ 0 & si\ x = 0. \end{cases}$$

Remarque 6.28 *Il est courant, par abus de notation, d'appeler f le prolongement de f par continuité.*

Définition 6.29 *Soit E un espace vectoriel normé. Soit $(x_n)_{n \in \mathbb{N}}$ une suite de E. On dit que la série de terme général x_n est absolument convergente si $\sum\limits_{n \geq 0} \|x_n\| < +\infty$.*

Proposition 6.30 *Soit E un espace vectoriel normé. Il est complet si et seulement si toute série absolument convergente est convergente.*

Démonstration. Supposons que E est complet. Si la série de terme général x_n converge absolument, alors la suite des sommes partielles $(S_n)_{n\in\mathbb{N}}$ définie par

$$\forall n \in \mathbb{N}, \quad S_n = \sum_{k=0}^{n} x_k$$

vérifie

$$\forall n \in \mathbb{N}, \quad \forall p \in \mathbb{N}, \quad \|S_n - S_{n+p}\| \leq \sum_{k=n+1}^{n+p} \|x_k\| \leq \sum_{k\geq n+1} \|x_k\|.$$

Ce majorant est indépendant de p, et converge vers 0 quand $n \to +\infty$. Donc la suite $(S_n)_{n\in\mathbb{N}}$ est de Cauchy dans F, donc converge.

Réciproquement, supposons que toute série absolument convergente de E converge. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans E. Alors, pour tout $k \in \mathbb{N}$, il existe $n_k \in \mathbb{N}$ tel que

$$\forall n \geq n_k, \quad \forall p \in \mathbb{N}, \quad \|x_{n+p} - x_n\| \leq \frac{1}{2^k}.$$

On définit alors $\varphi : \mathbb{N} \to \mathbb{N}$ de la façon suivante :

$$\varphi(0) = n_0, \quad \forall k \geq 1, \quad \varphi(k+1) = \max(\varphi(k), n_{k+1}).$$

Ainsi, $\varphi : \mathbb{N} \to \mathbb{N}$ est une application strictement croissante, qui vérifie

$$\forall k \in \mathbb{N}, \quad \forall p \in \mathbb{N}, \quad \left\|x_{\varphi(k)+p} - x_{\varphi(k)}\right\| \leq \frac{1}{2^k}.$$

En particulier,

$$\forall k \in \mathbb{N}, \quad \forall p \in \mathbb{N}, \quad \left\|x_{\varphi(k+1)} - x_{\varphi(k)}\right\| \leq \frac{1}{2^k}.$$

Donc la série de terme général $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ est absolument convergente, donc convergente. La suite $(x_n)_{n\in\mathbb{N}}$ a donc une valeur d'adhérence. Comme elle est de Cauchy, elle converge d'après la proposition 6.6. $\square$

Deuxième partie

Calcul différentiel

Chapitre 7

Fonctions d'une variable réelle

7.1 Continuité et dérivabilité

Définition 7.1 *Soit I un intervalle de $\mathbb{R}$, soit F un espace vectoriel normé. Soit $f : I \to F$ une application. Soit $x_0 \in I$. On dit que f est dérivable en x_0 si* $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} \dfrac{f(x) - f(x_0)}{x - x_0}$ *existe. On l'appelle alors la dérivée de f au point x_0, et on la note*

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

Exemple 7.2 *Si f est une application polynômiale, elle est dérivable en tout point de $\mathbb{R}$.*

Exemple 7.3 *Soit $f : \mathbb{R}^+ \to \mathbb{R}$ définie par $f(x) = \sqrt{x}$ n'est pas dérivable en 0.*

Démonstration. En effet, on a :

$$\forall x > 0, \quad \frac{f(x) - f(0)}{x - 0} = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}} \underset{x \to 0}{\longrightarrow} +\infty.$$

$\square$

Proposition 7.4 *Soit I un intervalle de $\mathbb{R}$, soit F un espace vectoriel normé. Soit $f : I \to F$ une application. Soit $x_0 \in I$. Si f est dérivable en x_0, alors elle est continue en x_0.*

Démonstration. Notons $f'(x_0)$ la dérivée de f en x_0. Alors on sait qu'il existe $\delta > 0$ tel que

$$\forall x \in I \cap \,]x_0 - \delta, c_0 + \delta[, \quad x \neq x_0, \quad \left\| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right\|_F \leq 1.$$

Donc, d'après la deuxième forme de l'inégalité triangulaire (remarque 1.4),

$$\forall x \in I\cap\,]x_0-\delta, c_0+\delta[, \quad x \neq x_0, \quad \left\| \frac{f(x)-f(x_0)}{x-x_0} \right\|_F \leq \|f'(x_0)\|_F + 1.$$

On en déduit donc que

$$\forall x \in I\cap\,]x_0-\delta, c_0+\delta[, \quad \|f(x)-f(x_0)\|_F \leq (\|f'(x_0)\|_F + 1)\,|x-x_0|.$$

Soit maintenant $\varepsilon > 0$. On choisit alors

$$\eta = \min\left(\delta, \frac{\varepsilon}{1+\|f'(x_0)\|_F}\right) > 0,$$

et on a bien

$$\forall x \in I\cap\,]x_0-\eta, x_0+\eta[, \quad \|f(x)-f(x_0)\|_F \leq \varepsilon.$$

Donc f est bien continue en 0. □

Définition 7.5 *Soit I un intervalle de $\mathbb{R}$, soit F un espace vectoriel normé. Soit $f : I \to F$ une application. On dit que f est dérivable sur I si elle est dérivable en tout point de I.*

Définition 7.6 *Soit I un intervalle de $\mathbb{R}$, soit F un espace vectoriel normé. On note :*

- *$C^0(I,F)$ l'ensemble des applications continues de I dans F.*
- *$C^1(I,F)$ l'ensemble des applications de I dans F qui sont dérivables et dont la dérivée est continue :*

$$C^1(I,F) = \left\{ f : I \to F, \quad f \text{ dérivable sur } I, \quad f' \in C^0(I,F) \right\}.$$

Théorème 7.7 (de Rolle) *Soient $x < y$ deux réels, et soit $f : [x,y] \to \mathbb{R}$ une application continue telle que $f \in C^1(]x,y[,\mathbb{R})$. On suppose de plus que $f(x) = f(y)$. Alors*

$$\exists z \in]x,y[\quad / \quad f'(z) = 0.$$

Démonstration. Si f est une application constante, alors $f' = 0$ par définition de la dérivée, donc la conclusion du théorème est vraie. Supposons donc f non constante. Il existe donc $t \in]x,y[$ tel que $f(t) \neq f(x) = f(y)$. Supposons par exemple que $f(t) > f(x) = f(y)$. L'ensemble $[x,y]$ est un fermé borné de $\mathbb{R}$, donc un compact, donc f est bornée et atteint ses bornes sur $[x,y]$: il existe $z \in [x,y]$ tel que

$$f(z) = \max_{t\in[x,y]} f(t).$$

Comme de plus $f(x) = f(y) < \max_{t\in]x,y[} f(t)$, on a donc $z \in]x,y[$, et

$$f(z) = \max_{t\in]x,y[} f(t).$$

D'où
$$\forall z' \in]x,y[, \quad f(z) \geq f(z'),$$
donc
$$\forall z' \in]x,z[, \quad \frac{f(z)-f(z')}{z-z'} \geq 0,$$
$$\forall z' \in]z,y[, \quad \frac{f(z)-f(z')}{z-z'} \leq 0.$$
En faisant tendre z' vers z dans chacune de ces inégalités, on déduit de la première que $f'(z) \geq 0$, et de la deuxième que $f'(z) \leq 0$. Donc $f'(z) = 0$.

Si maintenant $f(t) < f(x) = f(y)$, on appplique le même raisonnement en utilisant le min à la place du max. □

Remarque 7.8 *Le théorème de Rolle est faux si $f : I \to F$ avec F espace vectoriel normé tel que $F \neq \mathbb{R}$.*

Démonstration. Citons comme contre-exemple
$$\begin{array}{rccl} f: & \mathbb{R} & \longrightarrow & \mathbb{R}^2, \\ & t & \longmapsto & \begin{pmatrix} x \\ y \end{pmatrix}. \end{array}$$
Cette application vérifie $f(0) = f(2\pi)$. Elle est bien dérivable sur tout $\mathbb{R}$, et sa dérivée vaut
$$f'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix},$$
qui vérifie $\|f'(t)\| = \sqrt{\cos^2 t + \sin^2 t} = 1$. Donc f' ne s'annule pas sur $]0, 2\pi[$. □

Proposition 7.9 (formule des accroissements finis) *Soit I un intervalle de $\mathbb{R}$, et $f \in C^1(I, \mathbb{R})$. Alors pour tout $x < y$ dans I,*
$$\exists z \in]x,y[\quad / \quad \frac{f(x)-f(y)}{x-y} = f'(z).$$

Démonstration. Soit $\phi : I \to \mathbb{R}$ définie par
$$\forall z \in I, \quad \phi(z) = f(z) - f(x) - \frac{f(x)-f(y)}{x-y}(z-x).$$
Alors ϕ est continue sur $[x,y]$, C^1 sur $]x,y[$, et vérifie
$$\phi(x) = \phi(y) = 0.$$
On applique donc le théorème de Rolle à ϕ : il existe $z \in]x,y[$ tel que $\phi'(z) = 0$. Or
$$\phi'(z) = f'(z) - \frac{f(x)-f(y)}{x-y},$$
ce qui donne la conclusion. □

Corollaire 7.10 *Soit I un intervalle de $\mathbb{R}$, et $f \in C^1(I,\mathbb{R})$. Alors pour tout $x < y$ dans I,*

$$|f(x) - f(y)| \leq \left(\sup_{x \leq z \leq y} |f'(z)| \right) |x - y|.$$

Démonstration. Comme f' est continue sur $[x, y]$, elle est bornée et atteint ses bornes. Donc $M = \sup\limits_{x \leq z \leq y} |f'(z)|$ existe. On applique donc la formule des accroissements finis, et on obtient

$$|f(x) - f(y)| \leq |f'(z)|\ |x - y| \leq M|x - y|.$$

□

Proposition 7.11 (inégalité des accroissements finis) *Soit I un intervalle de $\mathbb{R}$, F un espace vectoriel normé, et $f \in C^1(I, F)$. On suppose de plus qu'il existe $M \geq 0$ tel que*

$$\forall x \in I, \quad \|f'(x)\| \leq M.$$

Alors

$$\forall x \in I, \quad \forall y \in I, \quad \|f(x) - f(y)\| \leq M\|x - y\|.$$

Démonstration. On commence par le cas où $x = 0$ et $y = 1$. Soit $\varepsilon > 0$. On pose

$$S_\varepsilon = \{t \in [0,1] \quad / \quad \forall s \in [0,t], \quad \|f(s) - f(0)\| \leq (M + \varepsilon)s\}.$$

L'ensemble S_ε contient 0, donc n'est pas vide. De plus, $S_\varepsilon \subset [0,1]$. On définit alors

$$t_0 = \sup S_\varepsilon.$$

Alors

$$\forall t \in [0, t_0[, \quad t \in S_\varepsilon, \quad \text{donc} \quad \|f(t) - f(0)\| \leq (M + \varepsilon)t.$$

Supposons que $t_0 < 1$. Alors, pour tout $n \in \mathbb{N}^*$, il existe $t_n \in \left[t_0, t_0 + \frac{1}{n}\right]$ tel que $t_n \notin S_\varepsilon$. Donc

$$\exists s_n \in [t_0, t_n] \subset \left[t_0, t_0 + \frac{1}{n}\right] \quad / \quad \|f(s_n) - f(0)\| > (M + \varepsilon)s_n.$$

Donc, en appliquant l'inégalité triangulaire (remarque 1.4), on a

$$\begin{aligned}\|f(s_n) - f(t_0)\| &\geq \|f(s_n) - f(0)\| - \|f(0) - f(t_0)\| \\ &\geq (M + \varepsilon)s_n - (M + \varepsilon)t_0 = (M + \varepsilon)(s_n - t_0).\end{aligned}$$

Remarquons enfin que $s_n \neq t_0$, car, par définition de t_0, $\|f(t_0) - f(0)\| \leq (M + \varepsilon)t_0$. Donc, en divisant l'inégalité ci-dessus par $s_n - t_0$, on obtient

$$\left\| \frac{f(s_n) - f(t_0)}{s_n - t_0} \right\| \geq M + \varepsilon,$$

donc, puisque $s_n \to t_0$ quand $n \to +\infty$, et que f est dérivable en t_0,

$$\|f'(t_0)\| \geq M + \varepsilon.$$

Ceci est contradictoire avec l'hypothèse. Donc $t_0 = 1$. Donc $S_\varepsilon = [0, 1]$, ce qui termine la preuve.

Traitons maintenant le cas où x et y sont quelconques. On pose alors

$$\begin{array}{llll} g: & [0,1] & \longrightarrow & F \\ & t & \longmapsto & f(x + t(y - x)). \end{array}$$

L'application g est bien dérivable, et pour tout $t \in [0, 1]$, $g'(t) = f'(x + t(y - x))(y - x)$. En particulier, on a $|g'(t)| \leq M|x - y|$. Donc, en appliquant ce qui précède à g, on obtient le résultat. □

7.2 Lien avec l'intégrale

Lemme 7.12 *Soit I un intervalle de $\mathbb{R}$, et soient $a < b$ dans I. On suppose que $f \in C^0(I, F)$. Alors*

$$\left\| \int_a^b f(t)dt \right\| \leq \int_a^b \|f(t)\|dt.$$

Démonstration. Comme f est continue, on peut définir son intégrale par somme de Riemann :

$$\int_a^b f(t)dt = \lim_{N \to +\infty} \left(\frac{1}{N+1} \sum_{i=0}^{N} f\left(a + \frac{i}{N}(b - a)\right) \right).$$

Soit $\varepsilon > 0$. Il existe donc $N_0 \in \mathbb{N}$ tel que

$$\forall N \geq N_0, \quad \left\| \int_a^b f(t)dt - \frac{1}{N+1} \sum_{i=0}^{N} f\left(a + \frac{i}{N}(b - a)\right) \right\| \leq \varepsilon.$$

Donc, en appliquant l'inégalité triangulaire,

$$\left\| \int_a^b f(t)dt \right\| \leq \frac{1}{N+1} \sum_{i=0}^{N} \left\| f\left(a + \frac{i}{N}(b - a)\right) \right\| + \varepsilon.$$

L'application $t \mapsto \|f(t)\|$ est continue, donc le membre de droite de cette inégalité converge, quand $N \to +\infty$, vers $\displaystyle\int_a^b \|f(t)\|dt$. Donc

$$\left\| \int_a^b f(t)dt \right\| \leq \int_a^b \|f(t)\|dt + \varepsilon.$$

Ceci valant pour tout $\varepsilon > 0$, on obtient le résultat. □

Proposition 7.13 *Soit I un intervalle de $\mathbb{R}$, F un espace vectoriel normé. Soit $f \in C^1(I,F)$. Soit $a \in I$. Alors, l'application $g : I \to F$ définie par*

$$\forall x \in I, \quad g(x) = \int_a^x f(t)dt,$$

est dans $C^1(I,F)$ et vérifie

$$\forall x \in I, \quad g'(x) = f(x).$$

Démonstration. Démontrons tout d'abord que g est continue : pour tout $(x,y) \in I^2$, tel que $x < y$, on a

$$\|g(x) - g(y)\| = \left\|\int_x^y f(t)dt\right\| \leq \int_x^y \|f(t)\|dt \leq \sup_{z\in[x,y]} \|f(z)\|(y-x) \underset{x\to y}{\longrightarrow} 0.$$

De plus, si $x_0 \in I$, on a, pour tout $x \neq x_0$,

$$\begin{aligned}\left\|\frac{g(x)-g(x_0)}{x-x_0} - f(x_0)\right\| &= \left\|\frac{1}{x-x_0}\int_{x_0}^x (f(t)-f(x_0))\,dt\right\| \\ &\leq \frac{1}{|x-x_0|}\int_{x_0}^x \|f(t)-f(x_0)\|dt.\end{aligned}$$

Soit $\varepsilon > 0$. Il existe alors $\delta > 0$ tel que

$$\forall t \in]x_0-\delta, x_0+\delta[\cap I, \quad \|f(t)-f(x_0)\| \leq \varepsilon.$$

Donc

$$\forall x \in]x_0-\delta, x_0+\delta[\cap I, \quad \left\|\frac{g(x)-g(x_0)}{x-x_0} - f(x_0)\right\| \leq \varepsilon.$$

Ceci prouve que g est dérivable en x_0, et que $g'(x_0) = f(x_0)$. □

7.3 Formules de Taylor

Définition 7.14 *Soit I un intervalle de $\mathbb{R}$, F un espace vectoriel normé, f une application de I dans F. On définit par récurrence les dérivées successives de f :*

- *$f^{(0)} = f$,*
- *$\forall k \in \mathbb{N}$, si $f^{(k)}$ est définie et est dérivable sur I, on pose $f^{(k+1)} = \left(f^{(k)}\right)'$.*

Si les $f^{(k)}$ existent pour tout $k \leq n$, on dit que f est n fois dérivable. Si de plus $f^{(n)}$ est continue, on dit qu'elle est n fois continûment détivable si de plus $f^{(n)}$ est continue. On note $C^n(I,F)$ l'ensemble des fonctions n fois continûment dérivables sur I.

Proposition 7.15 (formule de Taylor-Young) *Soit I un intervalle de $\mathbb{R}$, F un espace vectoriel normé, $f \in C^n(I,F)$, et $a \in I$. Alors il existe $\varepsilon : I \to F$ continue telle que $\lim_{x\to a} \varepsilon(x) = 0$ et*

$$\forall x \in I, \quad f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \varepsilon(x)(x-a)^n.$$

Démonstration. La preuve se fait par récurrence sur n :

- si $n = 0$, c'est la continuité de f.
- si la propriété ci-dessus est vraie au rang n, et que $f \in C^{n+1}(I,F)$, on applique l'hypothèse de récurrence à f' :

$$f'(x) = \sum_{k=0}^{n} \frac{f^{(k+1)}(a)}{k!}(x-a)^k + \varepsilon(x)(x-a)^n.$$

On intègre ensuite cette égalité de a à x :

$$\int_a^x f'(t)dt = \sum_{k=0}^{n} \frac{f^{(k+1)}(a)}{k!} \int_a^x (y-a)^k dy + \int_a^x \varepsilon(y)(y-a)^n dy.$$

En appliquant la proposition 7.13, on obtient donc

$$f(x) - f(a) = \sum_{k=0}^{n} \frac{f^{(k+1)}(a)}{k!} \int_a^x (y-a)^k dy + \int_a^x \varepsilon(y)(y-a)^n dy.$$

D'où

$$f(x) = \sum_{k=0}^{n+1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \varepsilon(y)(y-a)^n dy.$$

Posons

$$\overline{\varepsilon}(x) = \frac{1}{(x-a)^{n+1}} \int_a^x \varepsilon(y)(y-a)^n dy.$$

Soit $\delta > 0$. Alors il existe $h > 0$ tel que

$$\forall y \in]a-h, a+h[\cap I, \quad \|\varepsilon(y)\| \le \delta.$$

D'où, si $x \in]a-h, a+h[\cap I$,

$$\|\overline{\varepsilon}(x)\| \le \frac{1}{|x-a|^{n+1}} \int_a^x \delta |y-a|^n dy = \frac{\delta}{n+1} \le \delta.$$

Donc $\lim_{x\to a} \overline{\varepsilon}(x) = 0$. De même, $\overline{\varepsilon}$ est continue sur I.

□

Citons également le résultat suivant, valable uniquement si $F = \mathbb{R}$:

Proposition 7.16 (formule de Taylor-Lagrange) *Soit I un intervalle de $\mathbb{R}$, $f : I \to \mathbb{R}$ une application $(n+1)$ fois dérivable, et $a \in I$. Alors, pour tout $x \in I$, il existe ξ entre a et ξ tel que*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Remarque 7.17 *Dans le résultat ci-dessus, il est indispensable que la fonction soit à valeurs réelles. Le résultat est faux si l'espace d'arrivée est un espace vectoriel normé quelconque.*

Enfin, on a également

Proposition 7.18 (formule de Taylor-Laplace) *Soit I un intervalle de $\mathbb{R}$, F un espace vectoriel normé, $f \in C^{n+1}(I, F)$, et $a \in I$. Alors*

$$\forall x \in I, \quad f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n dt.$$

La preuve de ces deux derniers résultats se fait également par récurrence.

Chapitre 8

Fonctions de plusieurs variables

8.1 Dérivées directionnelles et dérivées partielles

Définition 8.1 *Soit U un ouvert de $\mathbb{R}^n$, soit F un espace vectoriel normé, et soit $f : U \to F$ une application. Pour $a \in U$ et $i \in \{1, \ldots, n\}$, on pose*

$$U_i^a = \{t \in \mathbb{R} \quad / \quad (a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n) \in U\}.$$

Soit

$$\begin{array}{rccl} \tau_i^a : & U_i^a & \longrightarrow & \mathbb{R}^n \\ & t & \longmapsto & (a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n). \end{array}$$

Alors l'application $f \circ \tau_i^a$ est appelée i-ème application partielle de f en a.

En d'autres termes, l'application $f_i^a = f \circ \tau_i^a$ est définie par

$$\begin{array}{rccl} f_i^a : & U_i^a & \longrightarrow & F \\ & t & \longmapsto & f(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_n). \end{array}$$

Proposition 8.2 *Soit U un ouvert de $\mathbb{R}^n$, soit F un espace vectoriel normé, et soit $f : U \to F$ une application continue. Alors toutes ses applications partielles sont continues.*

Démonstration. Soit $a \in U$ et $i \in \{1, \ldots, n\}$. Comme $f_i^a = f \circ \tau_i^a$, il suffit de prouver que τ_i^a est continue. Or $\tau_i^a(t) = te_i + a - a_ie_i$, où e_i est le i-ème vecteur de la base canonique. Cette application est donc la somme d'une application linéaire, donc continue, d'après le théorème 4.28, et d'une constante. Elle est donc continue. □

Remarque 8.3 *La réciproque de cette propriété est fausse.*

Démonstration. Considérons pour le prouver l'application

$$\begin{array}{rccll} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} & \\ & (x,y) & \longmapsto & \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0), \\ & (0,0) & \longmapsto & 0. & \end{array}$$

Alors, pour $a = (0,0)$, on calcule facilement :

$$f_1^a(t) = f(t,0) = 0, \quad f_2^a(t) = f(0,t) = 0.$$

Chacune d'entre elle est évidemment continue. Cependant, en posant

$$x_n = \frac{\cos\theta}{n}, \quad y_n = \frac{\sin\theta}{n},$$

on a

$$f(x_n, y_n) = \cos\theta \sin\theta,$$

qui ne converge pas, quand $n \to +\infty$, vers $f(0,0) = 0$, si $\theta \notin \frac{\pi}{2}\mathbb{Z}$. Donc f n'est pas continue en 0. □

Définition 8.4 *Soient U un ouvert de $\mathbb{R}^n$, F un espace vectoriel normé, $f : U \longrightarrow F$ et $a \in U$. Pour tout $v \in \mathbb{R}^n$, on dit que f admet une dérivée en a suivant le vecteur v si $\lim\limits_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv)-f(a)}{t}$ existe. On la note :*

$$\frac{\partial f}{\partial v}(a) = \lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv)-f(a)}{t}.$$

Définition 8.5 *Soient U un ouvert de $\mathbb{R}^n$, F un espace vectoriel normé, $f : U \longrightarrow F$ et $a \in U$. Si f admet une dérivée partielle suivant le i-ème vecteur e_i de la base canonique, on dit que f admet une dérivée partielle par rapport à la variable x_i en a. On la note*

$$\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i}(a).$$

Remarque 8.6 *Il existe des fonctions dont toutes les dérivées partielles existent, mais qui n'ont pas de dérivée directionnelle selon certains vecteurs.*

Démonstration. Un exemple possible est le suivant :

$$\begin{array}{rccll} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} & \\ & (x,y) & \longmapsto & \frac{xy}{|x|} & \text{si } x \neq 0, \\ & (0,y) & \longmapsto & 0. & \end{array}$$

Alors $f(t,0) = 0 = f(0,t)$, donc

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0.$$

Mais $f(t,t) = |t|$. Donc
$$\frac{f(t,t) - f(0,0)}{t} = \frac{|t|}{t},$$
qui n'a pas de limite quand $t \to 0$. □

Si f admet des dérivées partielles en a, on définit sa matrice jacobienne :

Définition 8.7 *Soient U un ouvert de $\mathbb{R}^n$, $f : U \longrightarrow \mathbb{R}^p$ et $a \in U$. On pose $f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_p(x) \end{pmatrix}$. Si toutes le dérivées partielles de f existent en a, alors il en est de même pour chaque f_j, $1 \leq j \leq p$. La matrice*
$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \cdots & \cdots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}$$
est appelée matrice jacobienne de f en a. Si $p = n$, on appelle jacobien de f en a le réel
$$\mathrm{Jac}_f(a) = \det(Df(a)).$$

Démonstration. Démontrons que si f admet des dérivées partielles, alors f_j aussi : on suppose donc que
$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(a + te_i) - f(a)}{t} = \frac{\partial f}{\partial x_i}(a).$$

Cette limite a lieu dans $\mathbb{R}^p$. Comme l'application
$$\begin{array}{cccc} \chi_j : & \mathbb{R}^p & \longrightarrow & \mathbb{R} \\ & \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} & \longmapsto & x_j \end{array}$$
est linéaire, donc continue d'après le théorème 4.28, on a donc
$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f_j(a + te_i) - f(a)}{t} = \frac{\partial f_j}{\partial x_i}(a).$$

Donc f_j admet bien des dérivées partielles en a, qui sont les composantes de celles de f. □

Avant d'aller plus loin, examinons quelques cas particuliers :

- si $n = p = 1$, f est une application de $\mathbb{R}$ dans $\mathbb{R}$, et dire que f admet des dérivées partielles (en réalité, une dérivée partielle) équivaut à dire que f est dérivable. De plus, $Df(a) = f'(a)$.

– Si $p = 1$ et n est quelconque, alors f est une application de $\mathbb{R}^n$ dans $\mathbb{R}$. Dans ce cas, la matrice $Df(a)$ est un vecteur ligne :

$$Df(a) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \frac{\partial f}{\partial x_2}(a) & \cdots & \frac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

Il est d'usage de noter $\nabla f(a) = (Df(a))^T$, le vecteur colonne des dérivées partielles de f au points a.

Donnons de plus un exemple de calcul explicite des dérivées partielles d'une fonction : soit

$$f : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & \begin{pmatrix} \sin\left(x^2+y^2\right) \\ xy \end{pmatrix} \end{array}.$$

L'application f a des dérivées partielles, car pour tout x fixé, l'application $y \mapsto f\begin{pmatrix} x \\ y \end{pmatrix}$ est dérivable, et pour tout y fixé, l'application $x \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$ l'est également. De plus, on a

$$f_1\begin{pmatrix} x \\ y \end{pmatrix} = \sin(x^2+y^2), \quad f_2\begin{pmatrix} x \\ y \end{pmatrix} = xy,$$

donc

$$\frac{\partial f_1}{\partial x}\begin{pmatrix} x \\ y \end{pmatrix} = 2x\cos(x^2+y^2), \quad \frac{\partial f_1}{\partial y}\begin{pmatrix} x \\ y \end{pmatrix} = 2y\cos(x^2+y^2),$$

$$\frac{\partial f_2}{\partial x}\begin{pmatrix} x \\ y \end{pmatrix} = y, \quad \frac{\partial f_2}{\partial y}\begin{pmatrix} x \\ y \end{pmatrix} = x.$$

Donc la jacobienne de f au point $\begin{pmatrix} x \\ y \end{pmatrix}$ vaut

$$Df\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x\cos(x^2+y^2) & 2y\cos(x^2+y^2) \\ y & x \end{pmatrix}.$$

Et le jacobien en ce même point vaut

$$\mathrm{Jac}_f\begin{pmatrix} x \\ y \end{pmatrix} = 2\left(x^2-y^2\right)\cos\left(x^2+y^2\right).$$

On constatera que, pour calculer la dérivée partielle par rapport à une variable x, on fixe toutes les autres et on calcule la dérivée de f par rapport à x.

8.2 Différentiabilité

Rappelons que, pour deux espaces vectoriels normés, on note $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F.

Définition 8.8 *Soient E et F des espaces vectoriels normés, et soit U un ouvert de E. Soient $f : U \to F$ et $a \in U$. On dit que f est différentiable en a s'il existe $L \in \mathcal{L}(E, F)$ telle que*

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - f(a) - L(x-a)\|_F}{\|x-a\|_E} = 0.$$

L'application L est appelée différentielle de f en a. On la note

$$L = df(a).$$

Remarque 8.9 *Il est courant de noter $df(a)(h) = df(a)h$, pour ne pas alourdir les notations.*

La proposition suivante justifie la notation $L = df(a)$ ci-dessus :

Proposition 8.10 *Soient E et F des espaces vectoriels normés, et soit U un ouvert de E. Soient $f : U \to F$ et $a \in U$. Si f est différentiable en a, sa différentielle est unique.*

Démonstration. Soient L et L' deux éléments de $\mathcal{L}(E, F)$ tels que

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - f(a) - L(x-a)\|_F}{\|x-a\|_E} = \lim_{\substack{x \to a \\ x \neq a}} \frac{\|f(x) - f(a) - L'(x-a)\|_F}{\|x-a\|_E} = 0.$$

En appliquant l'inégalité triangulaire, on a

$$\|L(x-a) - L'(x-a)\|_F \leq \|f(x) - f(a) - L(x-a)\|_F + \|f(x) - f(a) - L'(x-a)\|_F.$$

Donc $\displaystyle\lim_{\substack{x \to a \\ x \neq a}} \frac{\|L(x-a) - L'(x-a)\|_F}{\|x-a\|_E} = 0$. Autrement dit

$$\lim_{\substack{y \to 0 \\ y \neq 0}} \frac{\|(L-L')(y)\|_F}{\|y\|_E} = 0.$$

Fixons $z \in E$, $z \neq 0$. On a donc

$$\forall t \neq 0, \quad \frac{\|(L-L')(z)\|_F}{\|z\|_E} = \frac{\|(L-L')(tz)\|_F}{\|tz\|_E} \underset{t \to 0}{\longrightarrow} 0.$$

Donc $(L - L')z = 0$, pour tout $z \in E \setminus \{0\}$. Comme cette application est linéaire, elle est aussi nulle en 0. Donc $L = L'$. □

Remarque 8.11 *Si $E = F = \mathbb{R}$, alors f est dérivable en a si et seulement si f est différentiable en a, et sa différentielle est l'application définie par*

$$\forall t \in \mathbb{R}, \quad df(a)(t) = f'(a)t.$$

Démonstration. Supposons que f est dérivable en a. Par définition de la dérivée,

$$f'(a) = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}.$$

Donc

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a) - f'(a)(x - a)}{x - a} = 0.$$

Autrement dit,

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{|f(x) - f(a) - f'(a)(x - a)|}{|x - a|} = 0,$$

ce qui prouve bien que f est différentiable en a, et que $df(a)$ est l'application de multiplication par $f'(a)$.

Supposons maintenant que f est différentiable en a : on a donc

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{|f(x) - f(a) - df(a)(x - a)|}{|x - a|} = 0.$$

L'application $df(a)$ est linéaire de $\mathbb{R}$ dans $\mathbb{R}$, donc elle est de la forme $df(a)(x) = \lambda x$, pour un certain $\lambda \in \mathbb{R}$ (le démontrer en exercice). On en déduit donc

$$\lim_{\substack{x \to a \\ x \neq a}} \left| \frac{f(x) - f(a)}{x - a} - \lambda \right| = 0.$$

Donc f est bien dérivable, et $\lambda = f'(a)$, c'est-à-dire $df(a)(x) = f'(a)x$. □

Remarque 8.12 *Si f est linéaire continue, alors elle est différentiable en tout $a \in E$, et $df(a) = f$.*

Démonstration. Pour tout $a \in E$ et tout $x \in E$, on a $f(x) - f(a) = f(x - a)$, donc $f(x) - f(a) - f(x - a) = 0$. Ainsi,

$$\forall a \in E, \quad \forall x \in E \setminus \{a\}, \quad \frac{\|f(x) - f(a) - f(x - a)\|_F}{\|x - a\|_E} = 0.$$

En particulier, la limite de ce quotient est nulle, ce qui prouve le résultat voulu. □

Exercice 8.1 *Si f est une application affine continue, c'est-à-dire $f(x) = g(x) + u$, où $g \in \mathcal{L}(E, F)$ et $u \in F$ est fixé, démontrer que f est différentiable et que $\forall a \in E, \quad df(a) = g$.*

Remarque 8.13 *Si $f : E \times E \to F$ est une application bilinéaire continue, alors elle est différentiable en tout point $(a, b) \in E \times E$, et*

$$\forall (h, k) \in E^2, \quad df(a, b)(h, k) = f(h, b) + f(a, k).$$

Démonstration. Pour tout $(x,y) \in E^2$, on a

$$f(x,y) = f(x-a+a, y-b+b) = f(a,b)+f(x-a,b)+f(a,y-b)+f(x-a,y-b).$$

Donc

$$f(x,y) - f(a,b) - f(x-a,b) - f(x,y-b) = f(x-a,y-b).$$

Donc, puisque f est bilinéaire continue, d'après la proposition 1.40, il existe une constante $C > 0$ telle que

$$\begin{aligned}\forall (x,y) \in E^2, \quad & \|f(x,y) - f(a,b) - f(x-a,b) - f(x,y-b)\|_F \\ & = \|f(x-a,y-b)\|_F \le C\|x-a\|_E \,\|y-b\|_E \\ & \le C\left(\|x-a\|_E + \|y-b\|_E\right)^2 = C\|(x,y)-(a,b)\|^2_{E\times E}.\end{aligned}$$

On obtient donc

$$\forall (x,y) \in E^2, \quad \frac{\|f(x,y) - f(a,b) - f(x-a,b) - f(x,y-b)\|_F}{\|(x,y)-(a,b)\|_{E\times E}} \le C\|(x,y)-(a,b)\|_{E\times E}.$$

Ce quotient tend donc vers 0 quand $(x,y) \to (a,b)$, ce qui prouve bien que f est différentiable en (a,b) et que $df(a,b)(h,k) = f(h,b) + f(a,k)$. □

Proposition 8.14 *Soient E et F des espaces vectoriels normés, et soit U un ouvert de E. Soit $f : U \to F$ une application différentiable en $a \in U$. Alors f est continue en a.*

Démonstration. Par application de la définition de la différentielle, on sait qu'il existe $\delta > 0$ tel que

$$\forall x \in B_\delta(a) \cap U, \quad \|f(x) - f(a) - df(a)(x-a)\|_F \le \|x-a\|_E.$$

Donc, par application de l'inégalité triangulaire,

$$\forall x \in B_\delta(a) \cap U, \quad \|f(x) - f(a)\|_F \le \|df(a)(x-a)\|_F + \|x-a\|_E.$$

Comme l'application linéaire $df(a)$ est continue, la proposition 1.40 implique qu'il existe une constante $C > 0$ telle que

$$\forall x \in B_\delta(a)\cap U, \quad \|f(x)-f(a)\|_F \le C\|x-a\|_E+\|x-a\|_E = (C+1)\|x-a\|_E \underset{x\to a}{\longrightarrow} 0.$$

Donc f est bien continue en a. □

Définition 8.15 *Soient E et F des espaces vectoriels normés. Soit U un ouvert de E, et $f : U \to F$.*

- *On dit que f est différentiable sur U si elle l'est en tout point de U.*
- *On dit que f est continûment différentiable sur U si elle est différentiable sur U et si l'application*

$$\begin{array}{rccc} df : & U & \longrightarrow & \mathcal{L}(E,F) \\ & a & \longmapsto & df(a) \end{array}$$

est continue.

Remarque 8.16 *Si $E = F = \mathbb{R}$, dire que f est continûment différentiable équivaut à dire que f est de classe C^1.*

Proposition 8.17 *Soient E et F des espaces vectoriels normés, et soit U un ouvert de E. Soit $f : U \to F$ une application différentiable en $a \in U$. Alors pour tout $v \in E$, f admet une dérivée directionnelle selon v en a. De plus,*

$$\forall v \in E, \quad \frac{\partial f}{\partial v}(a) = df(a)(v).$$

Démonstration. Appliquons la définition de la différentielle à $x = a + h$:

$$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{\|f(a+h) - f(a) - df(a)(h)\|_F}{\|h\|_E} = 0.$$

Donc, pour $v \neq 0$ fixé, on a

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{\|f(a+tv) - f(a) - df(a)(tv)\|_F}{\|tv\|_E} = 0.$$

C'est-à-dire, puisque $df(a)$ est linéaire,

$$\lim_{\substack{t\to 0\\ t\neq 0}} \left(\frac{1}{\|v\|_E} \left\| \frac{f(a+tv) - f(a)}{t} - df(a)(v) \right\|_F \right) = 0.$$

Donc

$$\lim_{\substack{t\to 0\\ t\neq 0}} \frac{f(a+tv) - f(a)}{t} = df(a)(v).$$

Donc la dérivée directionnelle suivant v existe, et vaut $df(a)(v)$. □

Ce résultat permet de faire le lien entre matrice jacobienne et différentielle :

Proposition 8.18 *Soient n, p des entiers naturels non nuls. Soit U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^p$, et soit $a \in U$. On suppose que f est différentiable en a. Alors elle admet des dérivées partielles, et sa matrice jacobienne $Df(a)$ est la matrice de l'application linéaire $df(a)$ dans la base canonique. Autrement dit,*

$$\forall h \in \mathbb{R}^n, \quad df(a)(h) = Df(a)h.$$

Démonstration. On applique la proposition précédente avec $v = e_i$, où $(e_i)_{1\leq i\leq n}$ est la base canonique de $\mathbb{R}^n$. La fonction f admet donc des dérivées partielles en a. De plus,

$$\forall i \in \{1, 2, \ldots, n\}, \quad \frac{\partial f}{\partial x_i}(a) = df(a)(e_i).$$

Soit maintenant $h \in \mathbb{R}^n$:

$$h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_{i=1}^{n} h_i e_i.$$

Donc

$$df(a)(h) = df(a)\left(\sum_{i=1}^{n} h_i e_i\right) = \sum_{i=1}^{n} h_i df(a)(e_i) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(e_i)h_i.$$

En particulier,

$$\forall j \in \{1, 2, \dots, p\}, \quad (df(a)(h))_j = \sum_{i=1}^{n} \frac{\partial f_j}{\partial x_i}(a)h_i = \sum_{i=1}^{n} [Df(a)]_{ji}\, h_i,$$

c'est-à-dire $df(a)(h) = Df(a)h$. □

Remarque 8.19 *Il est possible que toutes les dérivées partielles de f existent, mais que f ne soit pas différentiable.*

Démonstration. Considérons la fonction

$$\begin{array}{rccll} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R} & \\ & (x,y) & \longmapsto & \frac{x^3+y^3}{x^2+y^2} & \text{si} \quad (x,y) \neq (0,0), \\ & (0,0) & \longmapsto & 0. & \end{array}$$

On a

$$\forall x \in \mathbb{R}, \quad f(x,0) = x, \quad \text{donc} \quad \frac{\partial f}{\partial x}(0,0) = 1.$$

$$\forall y \in \mathbb{R}, \quad f(0,y) = y, \quad \text{donc} \quad \frac{\partial f}{\partial y}(0,0) = 1.$$

Supposons maintenant que f est différentiable en $(0,0)$. Alors

$$df(0,0)(h) = Df(0,0)h = (1,1)\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = h_1 + h_2,$$

d'après la proposition précédente. Soit maintenant $t \in \mathbb{R}$. On calcule :

$$\frac{f(t,t) - f(0,0) - df(0,0)(t,t)}{\|(t,t)\|} = \frac{t - 2t}{\sqrt{2}|t|} = -\frac{t}{|t|\sqrt{2}},$$

qui ne tend pas vers 0 quand $t \to 0$. Ce qui est contradictoire. Donc f n'est pas différentiable en $(0,0)$ □

Remarquons qu'il est également possible de reprendre la remarque 8.6. En effet, il s'agit d'une fonction qui a des dérivées partielles en $(0,0)$. Si elle était différentiable en ce point, elle aurait aussi une dérivée directionnelle selon le vecteur $(1,1)$. Or ce n'est pas le cas, comme on l'a vu à la remarque 8.6.

Proposition 8.20 *Soient E et F des espaces vectoriels normés. Soit U un ouvert de E. Soit $f, g : U \to F$ deux applications différentiables en $a \in U$. Alors, pour tout $(\lambda, \mu) \in \mathbb{R}^2$, l'application $\lambda f + \mu g$ est différentiable en a, et*

$$d(\lambda f + \mu g)(a) = \lambda df(a) + \mu dg(a).$$

Démonstration. Posons $F = \lambda f + \mu g$. Alors on a

$$\begin{aligned}\forall x \in U, \quad & \|F(x) - F(a) - [\lambda df(a) + \mu dg(a)]\,(x-a)\|_F \\ &= \|\lambda f(x) - \lambda f(a) - \lambda df(a)(x-a) + \mu g(x) - \mu g(a) - \mu dg(a)(x-a)\|_F \\ &\leq |\lambda|\,\|f(x) - f(a) - df(a)(x-a)\|_F + |\mu|\|g(x) - g(a) - dg(a)(x-a)\|_F.\end{aligned}$$

Donc, puisque g et f sont différentiables au point a, on en déduit que

$$\lim_{\substack{x \to a \\ x \neq a}} \left(\frac{\|F(x) - F(a) - [\lambda df(a) + \mu dg(a)]\,(x-a)\|_F}{\|x-a\|_E} \right) = 0.$$

Donc F est bien différentiable en a, et $dF(a) = \lambda df(a) + \mu dg(a)$. □

Proposition 8.21 *Soient E, F, G des espaces vectoriels normées, U un ouvert de E, V un ouvert de F. On suppose que $f : U \to V$, $g : V \to G$ et $a \in U$. Si f est différentiable en a et g différentiable en $f(a)$, alors $g \circ f$ est différentiable en a, et*

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

Démonstration. Par définition de la différentiabilité,

$$\forall x \in U, \quad f(x) = f(a) + df(a)(x-a) + \varepsilon\left(\|x-a\|_E\right)\|x-a\|_E,$$

avec $\lim_{t \to 0} \varepsilon(t) = 0$. Donc

$$g(f(x)) = g\left(f(a) + \underbrace{df(a)(x-a) + \varepsilon\left(\|x-a\|_E\right)\|x-a\|_E}_{=y-f(a)} \right).$$

Comme g est différentiable en $f(a)$, on en déduit

$$g(f(x)) = g(f(a)) + dg(f(a))(y - f(a)) + \tilde{\varepsilon}\left(\|y - f(a)\|_F\right)\|y - f(a)\|_F,$$

avec $\lim_{t \to 0} \tilde{\varepsilon}(t) = 0$. Par linéarité de $dg(f(a))$, on en déduit que

$$\begin{aligned}g(f(x)) = g(f(a)) + dg(f(a))(df(a)(x-a)) + \|x-a\|_E dg(f(a))\left[\varepsilon\left(\|x-a\|_E\right)\right] \\ + \tilde{\varepsilon}\left(\|y - f(a)\|_F\right)\|y - f(a)\|_F.\end{aligned}$$

Remarquons que, puisque $dg(f(a))$ et $df(a)$ sont des applications linéaires continues, il existe $C > 0$ tel que

$$\forall z \in F, \quad \|dg(a)(z)\|_G \leq C\|z\|_F, \quad \text{et} \quad \forall x \in E, \quad \|df(a)(x)\|_F \leq C\|x\|_E$$

Par ailleurs, il existe $\delta > 0$ tel que

$$\forall t \in [0, \delta[, \quad \|\varepsilon(t)\| \leq 1 \quad \text{et} \quad \|\tilde{\varepsilon}(t)\|_G \leq 1.$$

Donc

$$\begin{aligned}\forall x \in B_\delta(a), \quad \|y - f(a)\|_F &\leq \|df(a)(x-a)\|_F + \|x-a\|_E \,\|\varepsilon\left(\|x-a\|_E\right)\|_F \\ &\leq C\|x-a\|_E + \|x-a\|_E = (C+1)\|x-a\|_E,\end{aligned}$$

Et

$$\forall x \in B_\delta(a), \quad \left\|dg(f(a))\left[\varepsilon\left(\|x-a\|_E\right)\right]\right\|_G \leq C\,\|\varepsilon\left(\|x-a\|_E\right)\|_F \leq C\|x-a\|_E.$$

Ainsi,

$$\begin{aligned}\forall x \in B_\delta(a) \cap U, \quad &\|g(f(x)) - g(f(a)) - dg(f(a))(df(a)(x-a))\|_G \\ &\leq \|x-a\|_E^2 + \tilde{\varepsilon}\left(\|y - f(a)\|_F\right)(C+1)\|x-a\|_E.\end{aligned}$$

On vient de voir que $\|y - f(a)\|_F \to 0$ quand $x \to a$, donc

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{\|g(f(x)) - g(f(a)) - dg(f(a))(df(a)(x-a))\|_G}{\|x-a\|_E} = 0.$$

Ceci prouve que $g \circ f$ est différentiable en a, et que sa différentielle vérifie $d(g \circ f)(a) = dg(f(a)) \circ df(a)$. □

Exemple 8.22 *Soit U un ouvert de $\mathbb{R}^n$, et $f : U \to \mathbb{R}^p$ différentiable en $a \in U$. Soit $\gamma :]-h, h[\to U$, avec $h > 0$, telle que $\gamma(0) = a$. On suppose de plus que γ est dérivable en 0, et que $\gamma'(0) = v \in \mathbb{R}^n$. Alors, en appliquant ce qui précède, $f \circ \gamma$ est différentiable en 0. D'autre part,*

$$\forall t \in \mathbb{R}, \quad d(f \circ \gamma)(0)(t) = df(a)\left(\gamma'(0)t\right) = t df(a)(\gamma'(0)) = t df(a)(v).$$

Donc

$$\left(f \circ \gamma\right)'(0) = df(a)\left(\gamma'(0)\right) = df(a)(v).$$

En choisissant γ $\gamma(t) = a + tv$, on retrouve la définition de $\frac{\partial f}{\partial v}(a)$.

Corollaire 8.23 *Soient n, p, q des entiers naturels non nuls. Soient U un ouvert de $\mathbb{R}^n$ et V un ouvert de $\mathbb{R}^p$. Soient $f : U \to V$, $g : V \to \mathbb{R}^q$ et $a \in U$. On suppose que f est différentiable en a et que g est différentiable en $f(a)$. Alors la jacobienne de $g \circ f$ en a vérifie :*

$$D\left(g \circ f\right)(a) = Dg(f(a))Df(a).$$

Autrement dit, on a l'égalité :

$$\forall i \in \{1, \ldots, n\}, \quad \forall j \in \{1, \ldots, q\} \quad \frac{\partial (g \circ f)_j}{\partial x_i}(a) = \sum_{k=1}^{p} \frac{\partial g_i}{\partial x_k}(f(a)) \frac{\partial f_k}{\partial x_i}(a).$$

Démonstration. D'après la proposition 8.21, on a

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

Or la matrice de la composée de deux applications linéaires est le produit des matrices. Comme, d'après la proposition 8.18, $Df(a)$ est la matrice de $df(a)$ et $Dg(f(a))$ celle de $dg(f(a))$, on en déduit que $D(g \circ f)(a) = Dg(f(a))Df(a)$. □

Exemple 8.24 *Si f est une application affine, c'est-à-dire de la forme $f(x) = Ax + y_0$, avec $y_0 \in \mathbb{R}^p$ et $A \in \mathbb{R}^{p\times n}$, alors la jacobienne de f en a est A (voir par exemple l'exercice 8.1). Donc*

$$D\,(g \circ f)\,(a) = Dg(Ax + y_0)A.$$

8.3 Inégalité des accroissements finis

Définition 8.25 *Soit E un espace vectoriel normé. Soit $(a, b) \in E^2$. On appelle segment fermé d'extrémités a et b l'ensemble*

$$[a, b] = \{ta + (1 - t)b,\ t \in [0, 1]\}\,.$$

On appelle segment ouvert d'extrémités a et b l'ensemble

$$]a, b[= \{ta + (1 - t)b,\ t \in]0, 1[\}\,.$$

Théorème 8.26 (Inégalité des accroissements finis) *Soient E et F des espaces vectoriels normés, soit U un ouvert de E. Soient $a \in U$, $b \in U$, et $f : U \to F$ une application continue. On suppose de plus que f est différentiable en tout point de $]a, b[$, et que*

$$M = \sup_{0<t<1} \|df(ta + (1 - t)b)\|_{\mathcal{L}(E,F)} < +\infty.$$

Alors

$$\|f(b) - f(a)\|_F \leq M\|b - a\|_E.$$

Démonstration. On définit l'application

$$\begin{array}{rccl} \varphi : & [0, 1] & \longrightarrow & F \\ & t & \longmapsto & f(ta + (1 - t)b). \end{array}$$

Cette application est continue car f l'est et que $t \to ta + (1 - t)b$ est une application continue de $[0, 1]$ dans E. De plus, elle est différentiable sur $]0, 1[$ par application de la proposition 8.21, et sa différentielle vérifie

$$\forall s \in \mathbb{R}, \quad \forall t \in]0, 1[, \quad d\varphi(t)(s) = df(at + (1 - t)b)(a - b)s.$$

Autrement dit, d'après la remarque 8.11, cette application φ est dérivable, et

$$\forall t \in]0, 1[, \quad \varphi'(t) = df(at + (1 - t)b)(a - b).$$

L'hypothèse faite sur f implique que

$$\forall t \in]0, 1[, \quad \|\varphi'(t)\|_F \leq \|df(at + (1 - t)b)\|_{\mathcal{L}(E,F)}\|a - b\|_E \leq M\|a - b\|_E.$$

On applique maintenant la proposition 7.11 à l'application φ :

$$\|\varphi(1) - \varphi(0)\|_F \leq M\|a - b\|_E,$$

ce qui donne le résultat, car $\varphi(1) = f(b)$ et $\varphi(0) = f(a)$. □

On a déjà vu qu'une application peut admettre des dérivées partielles sans être différentiable. Le résultat suivant permet d'établir une condition sur les dérivées partielles pour qu'une fonction soit différentiable :

Théorème 8.27 *Soit F un espace vectoriel normé. Soient U un ouvert de $\mathbb{R}^n$ et $a \in U$. Soit $f : U \to F$. On suppose que f admet des dérivées partielles sur un voisinage de a, et que ces dérivées partielles sont continues. Alors f est différentiable en a, et*

$$\forall h \in \mathbb{R}^n, \quad df(a)(h) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i.$$

Démonstration. Pour tout $h \in \mathbb{R}^n$, on pose

$$\phi(h) = f(a+h) - f(a) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i.$$

On souhaite prouver que $\dfrac{\|\phi(h)\|_F}{\|h\|_{\mathbb{R}^n}} \underset{\substack{h\to 0\\ h\neq 0}}{\longrightarrow} 0$. Pour cela, on pose, pour tout indice $i \in \{1, 2, \ldots, n\}$,

$$\forall t \in [0,1], \quad g_i(t) = f(a_1, a_2, \ldots, a_i + th_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a) t h_i.$$

Alors

$$\begin{aligned} g_i(1) - g_i(0) = {} & f(a_1, a_2, \ldots, a_i + h_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) \\ & - f(a_1, a_2, \ldots, a_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a) h_i. \end{aligned}$$

Donc, en sommant toutes ces inégalités pour i de 1 à n, on obtient

$$\forall h \in \mathbb{R}^n, \quad \phi(h) = \sum_{i=1}^{n} \left(g_i(1) - g_i(0)\right).$$

Il nous suffit donc de prouver que chaque terme de cette somme vérifie l'asymptotique voulue, c'est-à-dire que

$$\forall i \in \{1, 2, \ldots, n\}, \quad \frac{\|g_i(1) - g_i(0)\|_F}{\|h\|_{\mathbb{R}^n}} \underset{\substack{h\to 0\\ h\neq 0}}{\longrightarrow} 0.$$

Pour cela, nous allons majorer la dérivée de la fonction g_i et appliquer l'inégalité des accroissements finis : g_i est dérivable car f l'est, et

$$\forall t \in [0,1], \quad g_i'(t) = \frac{\partial f}{\partial x_i}(a_1, \ldots, a_i + th_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) h_i - \frac{\partial f}{\partial x_i}(a) h_i.$$

Appliquons la continuité de l'application $\frac{\partial f}{\partial x_i}$: pour tout $\varepsilon > 0$, il existe $\delta_i > 0$ tel que

$$\begin{aligned}\|h\|_{\mathbb{R}^n} \leq \delta_i \quad &\Rightarrow \quad \forall t \in [0,1], \\ &\left\| \frac{\partial f}{\partial x_i}(a_1, \ldots, a_i + th_i, a_{i+1} + h_{i+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_i}(a) \right\|_F \leq \varepsilon.\end{aligned}$$

On pose donc $\delta = \min_{1 \leq i \leq n} \delta_i$, et on a

$$\|h\|_{\mathbb{R}^n} \leq \delta \quad \Rightarrow \quad \|g_i'(t)\|_F \leq \varepsilon.$$

Donc, en appliquant l'inégalité des accroissements finis,

$$\|h\|_{\mathbb{R}^n} \leq \delta \quad \Rightarrow \quad \|g_i(1) - g_i(0)\|_F \leq \varepsilon |h_i|.$$

En sommant sur i, on obtient donc, pour tout $h \in \mathbb{R}^n$ tel que $\|h\|_{\mathbb{R}^n} \leq \delta$,

$$\|\phi(h)\|_F = \left\| \sum_{i=1}^n g_i(1) - g_i(0) \right\|_F \leq \sum_{i=1}^n \|g_i(1) - g_i(0)\|_F \leq \varepsilon \sum_{i=1}^n |h_i|.$$

Ainsi,

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{\|\phi(h)\|_F}{\|h\|_{\mathbb{R}^n}} = 0.$$

Ceci prouve que f est différentiable en a, et que sa différentielle est bien donnée par la formule voulue. □

Corollaire 8.28 *Soit F un espace vectoriel normé, et soit U un ouvert de $\mathbb{R}^n$. Soit $f : U \to F$. Alors $f \in C^1(U, F)$ si et seulement si f admet des dérivées partielles en tout point de U, et que ces dérivées partielles sont des fonctions continues.*

Démonstration. Supposons que $f \in C^1(U, F)$. Alors, d'après la définition 8.15, elle est différentiable en tout point a de U, et l'application $a \to df(a)$ est continue de U dans $\mathcal{L}(\mathbb{R}^n, F)$. En particulier, d'après la proposition 8.17, f admet des dérivées partielles en tout point $a \in U$, et

$$\forall a \in U, \quad \frac{\partial f}{\partial x_i}(a) = df(a)(e_i).$$

Si maintenant a et b sont des points de U, on en déduit

$$\begin{aligned}\left\| \frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right\|_F &= \|df(a)(e_i) - df(b)(e_i)\|_F = \|[df(a) - df(b)]\,(e_i)\|_F \\ &\leq \|df(a) - df(b)\|_{\mathcal{L}(\mathbb{R}^n, F)} \|e_i\|_{\mathbb{R}^n} \underset{a \to b}{\longrightarrow} 0.\end{aligned}$$

Donc les dérivées partielles sont bien continues.

Réciproquement, si f admet des dérivées partielles, et qu'elles sont continues, la proposition précédente prouve que f est différentiable. De plus, pour tous points a et b de U, on a

$$\forall h \in \mathbb{R}^n \quad \|[df(a) - df(b)]\,(h)\|_F = \left\| \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right) h_i \right\|_F .$$

Donc, en appliquant l'inégalité triangulaire, on obtient

$$\begin{aligned} \forall h \in \mathbb{R}^n \quad \|[df(a) - df(b)]\,(h)\|_F &\le \sum_{i=1}^{n} \left\| \left(\frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right) h_i \right\|_F \\ &\le \sum_{i=1}^{n} \left\| \frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right\|_F |h_i| \le \|h\|_{\mathbb{R}^n} \sum_{i=1}^{n} \left\| \frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right\|_F . \end{aligned}$$

Donc

$$\sup_{h \in \mathbb{R}^n \setminus \{0\}} \frac{\|[df(a) - df(b)]\,(h)\|_F}{\|h\|_{\mathbb{R}^n}} \le \sum_{i=1}^{n} \left\| \frac{\partial f}{\partial x_i}(a) - \frac{\partial f}{\partial x_i}(b) \right\|_F \underset{a \to b}{\longrightarrow} 0.$$

Ce qui prouve bien que f est de classe C^1. □

Proposition 8.29 *Soient E et F des espaces vectoriels normés, et soit U un ouvert connexe de E. Soit $f : U \to F$. Alors*

$$f \text{ est constante sur } U \quad \Longleftrightarrow \quad \Big(\forall a \in U, \quad df(a) = 0 \Big).$$

Démonstration. Supposons que f est constante sur U. Alors

$$\forall x \in U, \quad \forall a \in U, \quad f(x) - f(a) = 0,$$

ce qui prouve, d'après la définition 8.8, que f est différentiable en a et que $df(a) = 0$.

Réciproquement, supposons que df est nulle. On fixe $a \in U$, et on pose

$$V = \{x \in U \quad / \quad f(x) = f(a)\} .$$

En d'autres termes, $V = f^{-1}(\{f(a)\})$, donc, puisque f est continue, V est un fermé de U. De plus, si $x \in V$, on sait qu'il existe $r > 0$ tel que $B_r(x) \subset U$. Donc

$$\forall y \in B_r(x), \quad [x, y] \subset B_r(x) \subset U.$$

Donc, en appliquant l'inégalité des accroissements finis,

$$\|f(x) - f(y)\|_F \le \sup_{z \in [x,y]} \|df(z)\|_{\mathcal{L}(E,F)} \|x - y\|_E = 0,$$

puisque $\forall z \in U$, $df(z) = 0$. Donc $B_r(x) \subset V$. Ceci prouve que V est un ouvert. Comme U est connexe, V est donc soit vide, soit égal à U. On sait qu'il n'est pas vide car $a \in V$. Donc $V = U$. Donc f est constante sur U. □

8.4 Différentielles d'ordre supérieur

Définition 8.30 *Soient E et F des espaces vectoriels normés, et soit U un ouvert de E. Soit $a \in U$, et soit $f : U \to F$.*
- *On dit que f est deux fois différentiable en a si elle est différentiable sur un voisinage de a, et si l'application $x \to df(x)$ est une application différentiable en a.*
- *On dit que f est deux fois différentiable sur U si elle est deux fois différentiable en tout point de U.*

Remarquons que la différentielle de df en a, que l'on note $d(df)(a)$, est une application linéaire continue de E dans $\mathcal{L}(E, F)$. Autrement dit,

$$d(df)(a) \in \mathcal{L}(E, \mathcal{L}(E, F)).$$

D'où

$$\forall h \in E, \quad d(df)(a)(h) \in \mathcal{L}(E, F).$$

Donc l'application

$$\begin{array}{ccc} E \times E & \longrightarrow & F \\ (h, k) & \longmapsto & [d(df)(a)(h)]\,(k), \end{array}$$

est linéaire en k pour h fixé, et linéaire en h pour k fixé. C'est donc une application bilinéaire continue. On la note $d^2f(a)$, et on l'identifie à $d(df)(a)$:

$$d^2f(a) : \begin{array}{ccc} E \times E & \longrightarrow & F \\ (h, k) & \longmapsto & [d(df)(a)(h)]\,(k). \end{array}$$

On a donc $d^2f(a) \in \mathcal{B}(E \times E, F)$, au sens de la définition 1.42.

Définition 8.31 *Sous les hypothèses de la définition précédente, on dit que f est deux fois continûment différentiable sur U si elle est deux fois différentiable en tout point de U et que l'application $a \to d^2f(a)$ est continue de U dans $\mathcal{B}(E \times E, F)$.*

Précisons la norme que l'on utilise sur l'espace $\mathcal{L}(E, \mathcal{L}(E, F))$: si $\phi \in \mathcal{L}(E, \mathcal{L}(E, F))$, alors

$$\|\phi\|_{\mathcal{L}(E,\mathcal{L}(E,F))} = \sup_{h \in E\setminus\{0\}} \frac{\|\phi(h)\|_{\mathcal{L}(E,F)}}{\|h\|_E} = \sup_{h \in E\setminus\{0\}} \sup_{k \in E\setminus\{0\}} \frac{\|\phi(h)(k)\|_F}{\|h\|_E \|k\|_E}.$$

Autrement dit, lorsqu'on identifie ϕ à l'application bilinéaire $(h, k) \to \phi(h)(k)$, on a

$$\|\phi\|_{\mathcal{L}(E,\mathcal{L}(E,F))} = \sup_{(h,k) \in E^2,\ h \neq 0,\ k \neq 0} \frac{\|\phi(h, k)\|_F}{\|h\|_E \|k\|_E}.$$

La définition 1.43 assure qu'il s'agit d'une norme sur $\mathcal{B}(E \times E, F)$.

Théorème 8.32 *Soient E et F des espaces vectoriels normés. Soit U un ouvert de E et $f : U \to F$ une application deux fois différentiable en $a \in U$. Alors $d^2f(a)$ est une application bilinéaire symétrique continue.*

Démonstration. On a déjà vu que $d^2f(a)$ est bilinéaire continue. Il reste à démontrer qu'elle est symétrique. Posons

$$\forall (h,k) \in E^2, \quad D(h,k) = f(a+h+k) - f(a+h) - f(a+k) - f(a).$$

On sait que U est un ouvert, donc il existe $r > 0$ tel que $B_r(a) \subset U$. Donc $D(h,k)$ est bien définie pour tout $(h,k) \in B_{r/2}(a)^2$. De plus, elle est symétrique :

$$\forall (h,k) \in B_{r/2}(a)^2, \quad D(h,k) = D(k,h).$$

Posons

$$\forall t \in [0,1], \quad \psi(t) = f(a+h+tk) - f(a+h) - f(a+tk) + f(a) - d^2f(a)(h,tk).$$

Alors $\psi(0) = 0$ et $\psi(1) = D(h,k) - d^2f(a)(h,k)$. De plus, ψ est dérivable, et

$$\begin{aligned}
\psi'(t) &= df(a+h+tk)(k) - df(a+tk)(k) - d^2f(a)(h,k) \\
&= df(a+h+tk)(k) - df(a)(k) - d^2f(a)(h+tk,k) \\
&\quad + df(a)(k) - df(a+tk)(k) - d^2f(a)(h,k) + d^2f(a)(h+tk,k) \\
&= \left[df(a+h+tk) - df(a) - d(df(a))(h+tk)\right](k) \\
&\quad + \left[df(a) - df(a+tk) + d(df)(a)(tk)\right](k).
\end{aligned}$$

Donc, en appliquant l'inégalité triangulaire, on a

$$\begin{aligned}
\|\psi'(t)\|_F \leq{}& \|df(a+h+tk) - df(a) - d(df)(a)(h+tk)\|_{\mathcal{L}(E,F)}\|k\|_E \\
&+ \|df(a) - df(a+tk) - d(df)(a)(tk)\|_{\mathcal{L}(E,F)}\|k\|_E.
\end{aligned}$$

Donc

$$\begin{aligned}
\frac{\|\psi'(t)\|_F}{\|h\|_E^2 + \|k\|_E^2} \leq{}& \frac{\|df(a+h+tk) - df(a) - d(df)(a)(h+tk)\|_{\mathcal{L}(E,F)}}{\|h+tk\|_E} \frac{\|k\|_E\|h+tk\|_E}{\|h\|_E^2 + \|k\|_E^2} \\
&+ \frac{\|df(a) - df(a+tk) - d(df)(a)(tk)\|_{\mathcal{L}(E,F)}}{\|tk\|_E} \frac{\|k\|_E\|tk\|_E}{\|h\|_E^2 + \|k\|_E^2}.
\end{aligned}$$

On prouve facilement, en utilisant l'inégalité triangulaire, que $\|h+tk\|_E \leq 2\|h\|_E^2 + 2\|k\|_E^2$. Par ailleurs, les taux d'accroissements ci-dessus tendent vers 0 quand h et k tendent vers 0 car f est deux fois différentiable en a. Donc

$$\lim_{(h,k)\to(0,0)} \frac{\|\psi'(t)\|_F}{\|h\|_E^2 + \|k\|_E^2} = 0.$$

Donc, en appliquant l'inégalité des accroissements finis,

$$\frac{\|D(h,k) - d^2f(a)(h,k)\|_F}{\|h\|_E^2 + \|k\|_E^2} = \frac{\|\psi(1) - \psi(0)\|_F}{\|h\|_E^2 + \|k\|_E^2} \underset{(h,k)\to(0,0)}{\longrightarrow} 0.$$

Fixons maintenant $(h,k) \in B_{r/2}(a)^2$. On a, pour tout $\lambda \in [0,1]$,

$$\lim_{\lambda \to 0} \frac{d^2 f(a)(\lambda h, \lambda k) - d^2 f(a)(\lambda k, \lambda h)}{\lambda^2} = 0.$$

Donc, puisque $d^2 f(a)$ est bilinéaire,

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

Il s'agit donc bien d'une application bilinéaire symétrique. □

Proposition 8.33 *Soit U un ouvert de $\mathbb{R}^n$, soit F un espace vectoriel normé, et soit $f : U \to F$. On suppose que $a \in U$ et que f est deux fois différentiable en a. Alors les dérivées partielles $\dfrac{\partial f}{\partial x_i}$ sont différentiables en a, et on note leurs dérivées partielles en a :*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(a).$$

On les appelles dérivées partielles secondes.

Démonstration. On sait que l'application $x \mapsto df(x)$ est différentiable sur un voisinage V de a, par définition. Donc

$$\forall x \in V, \quad \frac{\|df(x+h) - df(x) - d^2 f(x)(h)\|_{\mathcal{L}(\mathbb{R}^n,F)}}{\|h\|_{\mathbb{R}^n}} \underset{\substack{h\to 0\\ h\neq 0}}{\longrightarrow} 0.$$

Donc pour tout $i \in \{1,2,\ldots,n\}$,

$$\forall x \in V, \quad \frac{\|df(x+h)(e_i) - df(x)(e_i) - d^2 f(x)(h,e_i)\|_F}{\|h\|_{\mathbb{R}^n}} \underset{\substack{h\to 0\\ h\neq 0}}{\longrightarrow} 0.$$

Ceci prouve que l'application

$$\begin{array}{rccl} \frac{\partial f}{\partial x_i} : & V & \longrightarrow & F \\ & x & \longmapsto & df(x)(e_i) = \frac{\partial f}{\partial x_i}(x), \end{array}$$

est différentiable, et que

$$\forall h \in \mathbb{R}^n, \quad d\left(\frac{\partial f}{\partial x_i}\right)(x)(h) = d^2 f(x)(h,e_i).$$

Donc

$$\forall j \in \{1,2,\ldots,n\}, \quad \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x) = d\left(\frac{\partial f}{\partial x_i}\right)(x)(e_j) = d^2 f(x)(e_j,e_i).$$

□

Proposition 8.34 *Soit U un ouvert de $\mathbb{R}^n$, soit F un espace vectoriel normé, et soit $f : U \to F$. On suppose que $a \in U$ et que f est deux fois différentiable ne a. Alors*

$$\forall (h,k) \in \mathbb{R}^n \times \mathbb{R}^n, \quad d^2 f(a)(h,k) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i k_j.$$

Démonstration. On vient de voir que

$$\forall j \in \{1,2,\ldots,n\}, \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(a) = df(a)(e_i, e_j).$$

Donc, par bilinéarité de $d^2 f(a)$, on en déduit

$$\begin{aligned}\forall (h,k) \in \mathbb{R}^n \times \mathbb{R}^n, \quad d^2 f(a)(h,k) &= \sum_{i=1}^{n} \sum_{j=1}^{n} h_i k_j d^2 f(a)(e_i, e_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i k_j.\end{aligned}$$

□

Théorème 8.35 (de Schwarz) *Soit U un ouvert de $\mathbb{R}^n$, soit F un espace vectoriel normé, et soit $f : U \to F$. On suppose que $a \in U$ et que f est deux fois différentiable en a. Alors*

$$\forall i \in \{1,2,\ldots,n\}, \quad \forall j \in \{1,2,\ldots,n\}, \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Démonstration. Ceci est une conséquence immédiate de l'égalité

$$\forall j \in \{1,2,\ldots,n\}, \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(a) = df(a)(e_i, e_j),$$

et du fait que $d^2 f(a)$ est symétrique. □

Définition 8.36 *Soit U un ouvert de $\mathbb{R}^n$, et soit $f : U \to \mathbb{R}$. On suppose que $a \in U$ et que f est deux fois différentiable en a. Alors on appelle matrice hessienne de f en a la matrice*

$$D^2 f(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)_{\substack{1 \le i \le n \\ 1 \le j \le n}}.$$

Notons que $D^2 f(a)$ est la matrice de la forme bilinéaire $d^2 f(a)$ dans la base canonique, comme le montre la proposition 8.34. De plus, la matrice hessienne est symétrique, d'après le théorème de Schwarz.

On peut généraliser sans problème la notion de différentielle à n'importe quel ordre. On note $\mathcal{L}^k(E,F)$ les applications k-linéaires de E^k dans F : une application est alors dite k-fois différentiable en a si elle est $k-1$-différentiable sur un voisinage V de a, et si l'application

$$\begin{array}{rccl} d^{k-1}f: & V & \longrightarrow & \mathcal{L}^{k-1}(E,F) \\ & x & \longmapsto & d^{k-1}f(x) \end{array}$$

est différentiable en a. Ici, la norme utilisé sur $\mathcal{L}^{k-1}(E,F)$ est définie par

$$\forall L \in \mathcal{L}^k(E,F), \quad \|L\|_{\mathcal{L}^k(E,F)} = \sup_{\substack{(X_1,\dots,X_k)\in E^k \\ (X_1,\dots,X_k)\neq(0,\dots,0)}} \frac{\|f(X_1,X_2,\dots,X_k)\|_F}{\|X_1\|_E\|X_2\|_E\dots\|X_k\|_E}.$$

8.5 Formules de Taylor

Comme dans le cas de la dimension un on a les formules de Taylor :

Proposition 8.37 (formule de Taylor-Young) *Soient E et F des espaces vectoriels normés. Soit U un ouvert de E, et soit $f : U \to F$ une application n fois différentiable en a. Alors il existe un voisinage V de a et une application $\varepsilon : V \to F$ telle que $\lim_{x\to a} \varepsilon(x) = 0$, et*

$$\forall x \in V, \quad f(x) = f(a) + \sum_{k=1}^{n} \frac{d^k f(a)(x-a,\dots,x-a)}{k!} + \varepsilon(x-a)\|x-a\|_F^n.$$

Démonstration. Soit $\varphi : [0,1] \to F$ l'application définie par

$$\forall t \in [0,1], \quad \varphi(t) = f(a+th).$$

Comme composée d'applications n fois différentiables, φ est n fois différentiable, donc n fois dérivable, et

$$\forall k \in \{1,\dots,n\}, \quad \varphi^{(k)}(t) = d^k f(a+th)(h,h,\dots,h).$$

On applique ensuite la formule de Taylor-Young à φ (proposition 7.15), et on obtient le résultat. $\square$

De même, en appliquant la formule de Taylor-Lagrange à la fonction φ ci-dessus, on a

Proposition 8.38 (formule de Taylor-Lagrange) *Soit E un espace vectoriel normé. Soit U un ouvert de E, et soit $f : U \to \mathbb{R}$ une application $n+1$ fois différentiable sur U. Soit $x \in U$ tel que $[a,x] \subset U$. Alors il existe $\xi \in]0,1[$ tel que*

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{d^k f(a)(x-a,\dots,x-a)}{k!} + \frac{d^{n+1} f(a+\xi h)(x-a,\dots,x-a)}{(n+1)!}.$$

Remarque 8.39 *Comme dans le cas de la proposition 7.16, ce résultat est faux si F n'est pas l'espace $\mathbb{R}$,*

Enfin, on a de la même manière :

Proposition 8.40 (formule de Taylor avec reste intégral) *Soient E et F des espaces vectoriels normés. Soit U un ouvert de E, et soit $f : U \to F$ une application de classe C^{n+1} sur U. Soit x tel que $[a, x] \subset U$. Alors*

$$\begin{aligned} f(x) = f(a) + \sum_{k=1}^{n} & \frac{d^k f(a)(x-a, \dots, x-a)}{k!} \\ & + \int_0^1 \frac{(1-t)^n}{n!} d^{n+1} f(a + t(x-a))(x-a, \dots, x-a) dt. \end{aligned}$$

Citons le cas particulier des formules de Taylor à l'ordre 2 :

– Taylor-Young :

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + \varepsilon(h) \|h\|_E^2.$$

– Taylor-Lagrange (avec f à valeurs réelles) :

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \xi h) h_i h_j, \quad \text{avec} \quad \xi \in]0, 1[.$$

– Taylor avec reste intégral :

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i + \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \int_0^1 (1-t) \frac{\partial^2 f}{\partial x_i \partial x_j}(a + th) dt.$$

Chapitre 9

Problèmes d'extrema

9.1 Définitions

Définition 9.1 *Soit U un ouvert de $\mathbb{R}^n$, et soit $f : U \to \mathbb{R}$. On suppose que f est différentiable en $a \in U$. On dit que a est un point critique de f si $df(a) = 0$.*

Remarque 9.2 *La propriété "a est point critique de f" signifie exactement*

$$\forall i \in \{1, 2 \ldots, n\}, \quad \frac{\partial f}{\partial x_i}(a) = 0.$$

Définition 9.3 *Soit U un ouvert de $\mathbb{R}^n$, et soit $f : U \to \mathbb{R}$. On suppose que $a \in U$. On dit que*
- *f admet un minimum global en a si $\forall x \in U, \quad f(x) \geq f(a)$.*
- *f admet un maximum global en a si $\forall x \in U, \quad f(x) \leq f(a)$.*
- *f admet un minimum local en a s'il existe V voisinage de a tel que $\forall x \in V, \quad f(x) \geq f(a)$.*
- *f admet un maximum local en a s'il existe V voisinage de a tel que $\forall x \in V, \quad f(x) \leq f(a)$.*
- *f admet un extremum global en a si elle admet un maximum global ou un minimum global en a.*
- *f admet un extremum local en a si elle admet un maximum local ou un minimum local en a.*

Dans toutes les définitions ci-dessus, on parlera de minimum/maximum/extremum local ou global strict si les inégalités sont strictes pour $x \neq a$.

Proposition 9.4 *Soit U un ouvert de $\mathbb{R}^n$, et soit $f : U \to \mathbb{R}$. On suppose que $a \in U$. Si f admet un extremum local en a, alors $df(a) = 0$.*

Démonstration. Soit $h \in \mathbb{R}^n$. Par application de la formule de Taylor-Young à l'ordre 1, il existe une application ε tendant vers 0 en 0 telle que

$$f(a + th) = f(a) + df(a)(th) + \varepsilon(th)\|th\|_{\mathbb{R}^n}.$$

Supposons que a est un minimum local. Alors, pour t assez petit, on a $f(a+th) \geq f(a)$, donc

$$tdf(a)(h) + \varepsilon(th)|t|\|h\|_{\mathbb{R}^n} \geq 0.$$

Si $t > 0$, ceci implique

$$df(a)(h) \geq -\varepsilon(th)\|h\|_{\mathbb{R}^n} \underset{t\to 0}{\longrightarrow} 0,$$

donc $df(a)(h) \geq 0$. De même, pour $t < 0$, on obtient $df(a)(h) \leq 0$. Donc $df(a) = 0$. Le cas d'un maximum local se traite de la même façon. □

Remarque 9.5 *La réciproque est fausse : il existe des applications dont la différentielle s'annule en a sans que f n'admette un extremum local en a.*

Démonstration. Un exemple simple est $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3$. $f'(0) = 0$, mais 0 n'est pas extremum local de f, car

$$\forall x > 0, \quad f(-x) < f(0) = 0 < f(x).$$

□

Remarque 9.6 *La proposition 9.4 est fausse si U n'est pas un ouvert.*

Démonstration. On peut prendre $U = [0,1]$ et $f(x) = x$. Alors 0 est minimum local de f (c'est même un minimum global), mais $f'(0) = 1 \neq 0$. □

9.2 Extrema est compacité

Rappelons que si K est compact et $f : K \to \mathbb{R}$ continue, alors elle est bornée et atteint ses bornes. Ceci signifie qu'elle admet un minimum global et un maximum global.

Proposition 9.7 *Soit E un espace vectoriel normé. Soit $f : E \to \mathbb{R}$ telle que $\lim\limits_{\|x\|\to+\infty} f(x) = +\infty$. Alors f admet un minimum global.*

Démonstration. Par définition, on sait que

$$\forall M \in \mathbb{R}, \quad \exists R > 0 \quad / \quad \|x\| \geq R \Rightarrow f(x) \geq M.$$

On choisit $M = f(0)$. On a donc

$$\forall x \in \mathbb{R}^n, \quad \|x\| \geq R \Rightarrow f(x) \geq f(0).$$

Soit $K = \overline{B_R}(0)$: c'est un fermé borné de $\mathbb{R}^n$, donc un compact. Donc f admet un minimum global sur K. Notons a le point de K où elle l'atteint. On a donc :

$$\forall x \in K, \quad f(x) \geq f(a),$$

et

$$\forall x \in K^C, \quad f(x) \geq M \geq f(0) \geq f(a).$$

Donc $\forall x \in \mathbb{R}^n, \quad f(x) \geq f(a)$. □

9.3 Utilisation de la différentielle seconde

Théorème 9.8 *Soit U un ouvert de $\mathbb{R}^n$, et soit $f : U \to \mathbb{R}$ une application différentiable sur U et deux fois différentiable en $a \in U$. On suppose que a est un point critique de f. Alors*

(i) Si $D^2 f(a)$ est définie positive, f admet un minimum local en a.

(ii) Si $D^2 f(a)$ est définie négative, f admet un maximum local en a.

(iii) Si $D^2 f(a) \neq 0$ n'est ni positive ni négative, alors a n'est ni un point de minimum local de f ni un point de maximum local de f. On dit alors que a est un point selle de f.

Démonstration. (i) : on applique la formule de Taylor-Young au voisinage de a : il existe $r > 0$ tel que $B_r(a) \subset U$, donc

$$\begin{aligned}\forall h \in B_r(0), \quad f(a+h) &= f(a) + df(a)(h) + d^2 f(a)(h,h) + \varepsilon(h)\|h\|^2 \\ &= f(a) + h^T D^2 f(a) h + \varepsilon(h)\|h\|^2,\end{aligned}$$

avec $\lim\limits_{h\to 0} \varepsilon(h) = 0$. Comme $D^2 f(a)$ est définie positive, il existe $\alpha > 0$ tel que

$$\forall h \in \mathbb{R}^n, \quad h^T D^2 f(a) h \geq \alpha \|h\|^2.$$

Par ailleurs, il existe $r' \leq r$ tel que

$$\forall h \in B_{r'}(0), \quad \varepsilon(h) \leq \alpha.$$

On en déduit donc

$$\forall h \in B_{r'}(0), \quad f(a+h) - f(a) \geq \alpha\|h\|^2 - \alpha\|h\|^2 = 0.$$

Donc f admet un minimum local en a.
(ii) : on applique le point (i) à $g = -f$. Comme $D^2 g(a) = -D^2 f(a)$, elle est définie positive, et g admet donc un minimum local en a. Par définition, ceci signifie que f admet un maximum local en a.
(iii) : si $D^2 f(a) \neq 0$ n'est ni positive ni négative, elle a au moins une valeur propre non nulle. Si elle est positive, au moins une autre est négative, sinon $D^2 f(a)$ serait positive. De même, si elle est négative, au moins une autre est positive. Donc il existe $\lambda_+ > 0$ et $\lambda_- < 0$ tels que

$$\exists h_+ \in \mathbb{R}^n, \quad \exists h_- \in \mathbb{R}^n \quad / \quad D^2 f(a) h_+ = \lambda_+ h_+, \quad D^2 f(a) h_- = \lambda_- h_-.$$

Donc, pour t assez petit, on peut appliquer la formule de Taylor-Young :

$$f(a + th_+) = f(a) + t^2 h_+^T D^2 f(a) h_+ + t^2 \|h_+\|^2 \varepsilon(th_+),$$

avec $\lim\limits_{h\to 0} \varepsilon(h) = 0$. Pour $t > 0$ assez petit, $\varepsilon(th_+) \geq -\frac{1}{2}\lambda_+$, donc

$$f(a+th_+) \geq f(a) + t^2 \lambda_+ h_+^T h_+ - \frac{1}{2}\lambda_+ t^2 \|h_+\|^2 = f(a) + \frac{1}{2} t^2 \lambda_+ \|h_+\|^2 > f(a).$$

De même,

$$f(a+th_-) = f(a) + t^2 h_-^T D^2 f(a) h_- + t^2 \|h_-\|^2 \varepsilon(th_+),$$

avec $\lim_{h\to 0} \varepsilon(h) = 0$. Pour $t > 0$ assez petit, $\varepsilon(th_+) \leq -\frac{1}{2}\lambda_-$, donc

$$f(a-th_-) \leq f(a) + t^2\lambda_- h_-^T h_- - \frac{1}{2}\lambda_- t^2 \|h_-\|^2 = f(a) + \frac{1}{2} t^2 \lambda_- \|h_-\|^2 < f(a).$$

Pour tout voisinage V de a, il existe $t > 0$ tel que $a + th_+ \in V$ et $a + th_- \in V$. Comme on vient de voir, toujours pour t assez petit, que $f(a - th_-) < f(a) < f(a + th_+)$, a ne peut pas être un point d'extremum de f. □

Remarque 9.9 *Si a est un point critique de f tel que $D^2 f(a)$ est positive mais pas définie, alors on ne peut pas conclure : a peut ne pas être un point de minimum de f.*

Démonstration. Un peut prendre les deux exemples suivants :

$$\begin{array}{rccl} f: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}, \\ & (x_1, x_2) & \longmapsto & x_1^2 + x_2^3, \end{array}$$

et

$$\begin{array}{rccl} g: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}, \\ & (x_1, x_2) & \longmapsto & x_1^2 + x_2^4. \end{array}$$

Chacune de ces fonctions est deux fois différentiable. On calcule facilement leurs différentielles respectives :

$$\frac{\partial f}{\partial x_1}(x_1,x_2) = 2x_1, \quad \frac{\partial f}{\partial x_2}(x_1,x_2) = 3x_2^2, \quad \frac{\partial g}{\partial x_1}(x_1,x_2) = 2x_1, \quad \frac{\partial g}{\partial x_2}(x_1,x_2) = 4x_2^3.$$

Donc

$$df(x_1,x_2)(h_1,h_2) = 2x_1h_1 + 2x_2^2h_2, \quad dg(x_1,x_2)(h_1,h_2) = 2x_1h_2 + 4x_2^3h_2.$$

En particulier, $df(0,0) = dg(0,0) = 0$. Le calcul des matrices hessiennes donne

$$D^2 f(x_1,x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 6x_2 \end{pmatrix}, \quad D^2 g(x_1,x_2) = \begin{pmatrix} 2 & 0 \\ 0 & 12x_2^2 \end{pmatrix}.$$

En particulier,

$$D^2 f(0,0) = D^2 g(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

qui est une matrice positive, mais pas définie positive. Or la fonction g admet bien un minimum local (global en fait) en $(0,0)$, car

$$\forall (x_1,x_2) \in \mathbb{R}^2, \quad g(x_1,x_2) = x_1^2 + x_2^4 \geq 0 = g(0,0).$$

En revanche, f n'admet pas d'extremum en $(0,0)$, car

$$\forall t > 0, \quad f(0,t) = t^3 > 0 = f(0,0), \quad \text{et} \quad \forall t < 0, \quad f(0,t) = t^3 < 0 = f(0,0).$$

□

9.4 Le cas convexe

Dans toute cette partie, on suppose que l'ensemble U est un ouvert convexe de $\mathbb{R}^n$, c'est-à-dire que

$$\forall x \in U, \quad \forall y \in U, \quad [x, y] \subset U.$$

Autrement dit, pour tout couple $(x, y) \in U^2$, et pour tout $\lambda \in [0, 1]$, $\lambda x + (1 - \lambda)y \in U$.

Définition 9.10 *Soit $f : U \to \mathbb{R}$. On dit que f est convexe si*

$$\forall (x, y) \in U^2, \quad \forall \lambda \in [0, 1], \quad f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

On dit que f est concave si $-f$ est convexe.

Définition 9.11 *Soit $f : U \to \mathbb{R}$. On dit que f est strictement convexe si*

$$\forall (x, y) \in U^2, \quad x \neq y, \quad \forall \lambda \in]0, 1[, \quad f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y).$$

On dit que f est strictement concave si $-f$ est strictement convexe.

Proposition 9.12 *Soit $f : U \to \mathbb{R}$. Elle est convexe si et seulement si elle vérifie, pour tout $n \in \mathbb{N}^*$*

$$\forall (\alpha_1, \ldots, \alpha_n) \in [0, 1]^n \quad / \quad \sum_{i=1}^{n} \alpha_i = 1, \quad \forall (x_1, \ldots, x_n) \in U^n,$$
$$f\left(\sum_{i=1}^{n} \alpha_i x_i\right) \leq \sum_{i=1}^{n} \alpha_i f(x_i).$$

Démonstration. Si f vérifie la propriété ci-dessus, il suffit de l'appliquer pour $n = 2$ pour obtenir que f est convexe.

Supposons que f est convexe. Alors la propriété ci-dessus est vraie pour $n = 2$ (elle est toujours vraie pour $n = 1$, même si f n'est pas convexe). Démontrons qu'elle est vraie pour tout $n \in \mathbb{N}^*$ par récurrence. Si elle est vraie au rang n, on suppose que

$$(\alpha_1, \ldots, \alpha_{n+1}) \in [0, 1]^{n+1}, \quad \sum_{i=1}^{n+1} \alpha_i = 1, \quad (x_1, \ldots, x_{n+1}) \in U^{n+1}.$$

Si $\alpha_{n+1} = 1$, alors l'inégalité est trivialement vraie. Sinon, on écrit

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \alpha_i x_i\right) &= f\left((1 - \alpha_{n+1})\frac{1}{1 - \alpha_{n+1}} \sum_{i=1}^{n} \alpha_i x_i + \alpha_{n+1} x_{n+1}\right) \\ &\leq (1 - \alpha_{n+1}) f\left(\sum_{i=1}^{n} \frac{\alpha_i}{1 - \alpha_{n+1}} x_i\right) + \alpha_{n+1} f(x_{n+1}), \end{aligned}$$

car f est convexe. Par ailleurs, $\frac{\alpha_i}{1-\alpha_{n+1}} \geq 0$ pour tout i, et

$$\sum_{i=1}^{n} \frac{\alpha_i}{1-\alpha_{n+1}} = \frac{\left(\sum_{i=1}^{n+1} \alpha_i\right) - \alpha_{n+1}}{1-\alpha_{n+1}} = \frac{1-\alpha_{n+1}}{1-\alpha_{n+1}} = 1.$$

On peut donc appliquer l'hypothèse de récurrence, et on obtient

$$\begin{aligned} f\left(\sum_{i=1}^{n+1} \alpha_i x_i\right) &\leq (1-\alpha_{n+1}) f\left(\frac{1}{1-\alpha_{n+1}} \sum_{i=1}^{n} \alpha_i x_i\right) + \alpha_{n+1} f(x_{n+1}) \\ &\leq \frac{1-\alpha_{n+1}}{1-\alpha_{n+1}} \sum_{i=1}^{n} \alpha_i f(x_i) + \alpha_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} \alpha_i f(x_i). \end{aligned}$$

Ceci prouve bien la propriété au rang $n+1$, et termine la récurrence. □

Proposition 9.13 *Soit U un ouvert convexe de $\mathbb{R}^n$. Soit $f : U \to \mathbb{R}$ une application convexe. Alors f est continue.*

Démonstration. On choisit comme norme la norme 1 :

$$\|x\| = \sum_{i=1}^{n} |x_i|.$$

Soit $a \in U$. Alors il existe $r > 0$ tel que $\overline{B_r}(a) \subset U$. On désigne par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{R}^n$. On a donc, pour tout $i \in \{1, 2, \ldots, n\}$,

$$a + re_i \in U, \quad a - re_i \in U.$$

Soit $x \in \overline{B_r}(a)$: alors $x = a + ry$, avec $\|y\| \leq 1$.

$$x = (1 - \|y\|)a + \|y\|a + r\sum_{i=1}^{n} y_i e_i = (1 - \|y\|)a + \sum_{i=1}^{n} |y_i| a + \sum_{i=1}^{n} r|y_i|\sigma_i e_i,$$

où σ_i désigne le signe de y_i : $\sigma_i = 1$ si $y_i > 0$, $\sigma_i = -1$ si $y_i < 0$. Ainsi,

$$x = \left(1 - \sum_{i=1}^{n} |y_i|\right) a + \sum_{i=1}^{n} |y_i| \left(a + r\sigma_i e_i\right).$$

On applique la proposition 9.12, et on a donc

$$f(x) \leq (1 - \|y\|) f(a) + \sum_{i=1}^{n} |y_i| f(a + r\sigma_i e_i).$$

On pose alors $M = \sup_{1 \leq i \leq n} \left(\sup(f(a + re_i), f(a - re_i))\right)$, et on a donc

$$f(x) \leq (1 - \|y\|) f(a) + M\|y\|.$$

Autrement dit,

$$\forall x \in B_r(a), \quad f(x) \le \left(1 - \frac{\|x-a\|}{r}\right) f(a) + \frac{\|x-a\|}{r} M.$$

De plus,

$$\begin{aligned} f(a) = f\left(\frac{1}{2}(a+ry) + \frac{1}{2}(a-ry)\right) &\le \frac{1}{2} f(a+ry) + \frac{1}{2} f(a-ry) \\ &\le \frac{1}{2} f(x) + \frac{1}{2}\left(1 - \frac{1}{r}\|a-x\|\right) f(a) + \frac{M}{2r}\|a-x\|. \end{aligned}$$

Donc

$$f(x) \ge \left(1 + \frac{\|x-a\|}{r}\right) f(a) - \frac{\|x-a\|}{r} M.$$

En regroupant les deux inégalités, on obtient donc

$$\left(1 + \frac{\|x-a\|}{r}\right) f(a) - \frac{\|x-a\|}{r} M \le f(x) \le \left(1 - \frac{\|x-a\|}{r}\right) f(x) + \frac{\|x-a\|}{r} M.$$

Donc $\lim_{x \to a} f(x) = f(a)$. Ceci valant pour tout $a \in U$, l'application f est bien continue sur U. □

Remarque 9.14 *On généralise sans problème le résultat ci-dessus à n'importe quel espace vectoriel normé de dimension finie. En revanche, il est faux en dimension infinie. En effet, en dimension infinie, il existe des applications linéaires non continues. Or une application linéaire est trivialement convexe.*

Remarque 9.15 *Ce résultat est faux si U n'est pas ouvert.*

Démonstration. L'exemple suivant est une fonction convexe non continue sur $[0,1]$:

$$\forall x \in [0,1], \quad f(x) = \begin{cases} 1 & \text{si } x = 0, \\ 0 & \text{si } 0 < x < 1, \\ 2 & \text{si } x = 1. \end{cases}$$

En effet, il est clair qu'elle n'est pas continue en 0 et 1. De plus, si $x \in [0,1]$ et $y \in [0,1]$, et si $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)y) = 0 \quad \text{si} \quad \lambda \in]0,1[,$$

car alors $\lambda x + (1-\lambda)y \in]0,1[$. En particulier,

$$f(\lambda x + (1-\lambda)y) = 0 \le \lambda f(x) + (1-\lambda) f(y),$$

puisque $f(x) \ge 0$, $f(y) \ge 0$ et $\lambda \in [0,1]$. Examinons maintenant les cas $\lambda = 0$ et $\lambda = 1$: si $\lambda = 0$, alors

$$f(\lambda x + (1-\lambda)y) = f(y) = \lambda f(x) + (1-\lambda) f(y),$$

et si $\lambda = 1$, alors

$$f(\lambda x + (1-\lambda)y) = f(x) = \lambda f(x) + (1-\lambda)f(y).$$

Au final, f est bien convexe. $\square$

Proposition 9.16 *Soit U un ouvert convexe de E, espace vectoriel normé. Soit $f : U \to \mathbb{R}$ différentiable. Alors f est convexe si et seulement si*

$$\forall a \in U, \quad \forall x \in U, \quad f(x) \geq f(a) + df(a)(x-a).$$

Démonstration. Supposons que f est convexe. Comme U est convexe, $[a, x] \subset U$. Soit

$$\begin{array}{rccl} \varphi : & [0,1] & \longrightarrow & \mathbb{R}, \\ & t & \longmapsto & f(a + t(x-a)). \end{array}$$

Cette application est convexe, car pour t et s dans $[0, 1]$, et $\lambda \in [0, 1]$, on a

$$\begin{aligned} \varphi(\lambda t + (1-\lambda)s) &= f(a + \lambda t(x-a) + (1-\lambda)s(x-a)) \\ &= f(\lambda a + \lambda t(x-a) + (1-\lambda)a + (1-\lambda)s(x-a)) \\ &= f(\lambda(a + t(x-a)) + (1-\lambda)(a + s(x-a))) \\ &\leq \lambda f(a + t(x-a)) + (1-\lambda)f(a + s(x-a)) \\ &= \lambda\varphi(t) + (1-\lambda)\varphi(s). \end{aligned}$$

Donc, pour tout $h \in [0, 1]$ on a

$$\varphi(h) = \varphi(h1 + (1-h)0) \leq h\varphi(1) + (1-h)\varphi(0).$$

D'où

$$\frac{\varphi(h) - \varphi(0)}{h} \leq \varphi(1) - \varphi(0).$$

L'application φ est une composée d'applications différentiables, donc elle est différentiable, donc dérivable, en tout point de $[0, 1]$. Donc, en faisant tendre h vers 0 dans l'inégalité ci-dessus, on obtient

$$\varphi'(0) \leq \varphi(1) - \varphi(0) = f(x) - f(a).$$

De plus, $\forall t \in [0, 1], \quad \varphi'(t) = df(a + t(x-a))(x-a)$, et donc

$$df(a)(x-a) \leq f(x) - f(a).$$

Réciproquement, supposons que f vérifie cette inégalité. Soient $x \in U$, $y \in U$ et $\lambda \in [0, 1]$. On pose $a = \lambda x + (1-\lambda)y$, et on applique l'inégalité :

$$\begin{aligned} f(x) = f(\lambda x + (1-\lambda)y + (1-\lambda)x - (1-\lambda)y) &= f(a + (1-\lambda)(x-y)) \\ &\geq f(a) + (1-\lambda)df(a)(x-y). \end{aligned}$$

De même,

$$f(y) = f(\lambda x + (1-\lambda)y - \lambda x + \lambda y) = f(a + \lambda(y-x))$$
$$\geq f(a) + \lambda df(a)(y-x).$$

On mutliplie la première inégalité par λ et la deuxième pas $1-\lambda$, et on somme :

$$\lambda f(x) + (1-\lambda)f(y)$$
$$\geq \lambda f(a) + (1-\lambda)f(a) + \lambda(1-\lambda)df(a)(x-y) + \lambda(1-\lambda)df(a)(y-x) = f(a).$$

Donc f est bien convexe. □

La même preuve permet également de démontrer :

Proposition 9.17 *Soit U un ouvert convexe de E, espace vectoriel normé. Soit $f : U \to \mathbb{R}$ différentiable. Alors f est strictement convexe si et seulement si*

$$\forall a \in U, \quad \forall x \in U \setminus \{a\}, \quad f(x) > f(a) + df(a)(x-a).$$

Il est possible d'utiliser la différentielle seconde pour caractériser la convexité :

Théorème 9.18 *Soit U un ouvert convexe de $\mathbb{R}^n$. Soit $f : U \to \mathbb{R}$ une application deux fois dérivable sur U. On a alors :*

(i) f est convexe $\iff$ $D^2 f(x)$ est positive pour tout $x \in U$.

(ii) Si $D^2 f(x)$ est définie positive pour tout $x \in U$, alors f est strictement convexe.

Démonstration. Commençons par prouver (i). Si f est convexe, alors on a, d'après la proposition 9.16,

$$\forall x \in U, \quad \forall a \in U, \quad f(x) \geq f(a) + df(a)(x-a).$$

Pour $a \in U$ fixé, on sait, puisque U est ouvert, qu'il existe $r > 0$ tel que $B_r(a) \subset U$. Donc, pour tout $h \in \mathbb{R}^n$, on peut appliquer cette inégalité à $x = a + th$, pour $t > 0$ assez petit :

$$f(a+th) \geq f(a) + tdf(a)(h).$$

D'autre part, par application de la formule de Taylor-Young à l'ordre 2, on obtient,

$$f(a+th) = f(a) + tdf(a)(h) + \frac{1}{2}t^2 d^2 f(a)(h,h) + t^2 \varepsilon(th)\|h\|^2_{\mathbb{R}^n},$$

avec $\lim\limits_{h \to 0} \varepsilon(h) = 0$. Donc

$$h^T D^2 f(a) h = d^2 f(a)(h,h) \geq -2\varepsilon(th)\|h\|^2_{\mathbb{R}^n} \underset{t\to 0}{\longrightarrow} 0.$$

Ceci valant pour tout $h \in \mathbb{R}^n$, la matrice $D^2 f(a)$ est donc bien positive.

Réciproquement, si $D^2f(a)$ est positive, on applique Taylor-Lagrange,

$$\forall x \in U, \quad \forall a \in U, \quad f(x)$$
$$= f(a) + df(a)(x-a) + \overbrace{d^2f(a+\xi(x-a))(x-a,x-a)}^{=(x-a)^T D^2 f(a+\xi(x-a))(x-a)\geq 0}$$
$$\geq f(a) + df(a)(x-a).$$

Donc, d'après la proposition 9.16, f est convexe.

Pour prouver (i), il suffit de reprendre la fin de la preuve ci-dessus : comme $D^2f(a)$ est définie positive, l'inégalité est stricte sauf si $x = a$. Toujours d'après la proposition 9.16, on en déduit que f est strictement convexe. □

Remarque 9.19 *Dans le cas $n = 1$, on retrouve que lorsque f est deux fois dérivable, elle est convexe si et seulement si $f'' \geq 0$.*

La preuve du résultat suivant est très facile est laissée en exercice :

Proposition 9.20
- *Une somme de fonctions convexes est convexe.*
- *Si f est convexe et $\alpha \geq 0$, alors αf est convexe.*
- *Toute forme quadratique positive est convexe.*
- *Toute application affine est convexe.*

Théorème 9.21 *Soit E un espace vectoriel normé. Soit A une partie convexe de E, et soit $f : A \to \mathbb{R}$ une application convexe. Si $a \in \mathring{A}$ est un point critique de f, alors c'est un minimum global. Lorsque f est strictement convexe, ce minimum est strict.*

Démonstration. Soit $a \in \mathring{A}$. Alors il existe $r > 0$ tel que $B_r(a) \subset A$. Alors, pour $y \in B_r(a)$, on a

$$f(y) \geq f(a) + df(a)(y-a) = f(a).$$

Soit maintenant $x \in U$. Alors, pour t assez petit, on a $y = a + t(x-a) \in B_r(a)$. Donc

$$f(a + t(x-a)) \geq f(a).$$

Quitte à diminuer t un peu plus, on peut toujours supposer que $t \in]0,1]$. Par convexité,

$$f(a + t(x-a)) = f((1-t)a + tx) \leq (1-t)f(a) + tf(x).$$

Donc

$$f(a) \leq (1-t)f(a) + tf(x), \text{ d'où } f(a) \leq f(x).$$

Donc a est bien un point de minimum global.

Si de plus la convexité est stricte, alors le point de minimum de f est unique. En effet, s'il en existait deux distincts x et y, on aurait

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) < \frac{1}{2}f(x) + \frac{1}{2}f(y) = \min_{x\in A} f(x),$$

ce qui est contradictoire. $\square$

En changeant f en $-f$, on a facilement le résultat suivant :

Théorème 9.22 *Soit E un espace vectoriel normé. Soit A une partie convexe de E, et soit $f : A \to \mathbb{R}$ une application concave. Si $a \in \mathring{A}$ est un point critique de f, alors c'est un maximum global. Lorsque f est strictement concave, ce maximum est strict.*

9.5 Etude d'extrema - exemple

Pour recenser les extrema d'une application $f : A \to \mathbb{R}$, on procèdera de la façon suivante (on décrit la méthode pour chercher les minima, c'est la même pour chercher les maxima) :

1. Utilisation de la compacité : A est-il compact ? Sinon, dans le cas où A est non borné, a-t-on $\lim\limits_{\|x\|\to+\infty} f(x) = +\infty$? Ces arguments permettent de démontrer l'existence d'un point de minimum (voir la proposition 4.23 et la proposition 9.7).
2. Utilisation de la convexité : si f est convexe, et si $a \in \mathring{A}$ est point critique, alors a est un point de minimum de f, d'après le théorème 9.21.
3. Utilisation de la jacobienne (si f est différentiable). On calcule $Df(a)$ pour tout $a \in \mathring{A}$. Les minima sont à chercher parmi les points critiques de f, d'après la proposition 9.4.
4. Utilisation de la hessienne. Si f est deux fois différentiable, si $Df(a) = 0$ et $D^2f(a)$ est définie positive, alors a est un minimum local.
5. Points de la frontière : Si A n'est pas ouvert, l'étude doit se faire au cas par cas pour les $a \in \partial A$.

Exemple Soit la fonction

$$\begin{array}{rcl} f : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \longmapsto & x^4 + y^4 - 2(x-y)^2. \end{array}$$

Cette application est continue car il s'agit d'une application polynômiale.

Appliquons les points qui précèdent successivement :

1. Démontrons que $\lim\limits_{\|X\|\to+\infty} f(X) = +\infty$: on a

$$x^4 + y^4 \geq \frac{1}{2}\left\|\begin{pmatrix} x \\ y \end{pmatrix}\right\|^4, \tag{9.1}$$

car

$$x^4+y^4 \geq 2x^2y^2, \quad \text{d'où} \quad x^4+y^4 \geq \frac{1}{2}\left(x^4+y^4+2x^2y^2\right) = \frac{1}{2}\left(x^2+y^2\right)^2.$$

Par ailleurs, on a également

$$(x-y)^2 \leq 2\left\|\begin{pmatrix}x\\y\end{pmatrix}\right\|^2. \tag{9.2}$$

En effet, $x^2+y^2+2xy = (x+y)^2 \geq 0$, d'où

$$x^2+y^2-2xy = 2x^2+2y^2-\left(x^2+y^2-2xy\right) \leq 2x^2+2y^2.$$

On déduit donc de (9.1) et (9.2) que

$$\forall \begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2, \quad f\begin{pmatrix}x\\y\end{pmatrix} \geq \frac{1}{2}\left\|\begin{pmatrix}x\\y\end{pmatrix}\right\|^4 - 2\left\|\begin{pmatrix}x\\y\end{pmatrix}\right\|^2.$$

Ceci prouve bien que $\lim\limits_{\|X\|\to+\infty} f(X) = +\infty$, donc, d'après la proposition 9.7, f admet au moins un minimum global.

2. La fonction f n'est pas convexe : si elle l'était, on aurait

$$f\begin{pmatrix}0\\0\end{pmatrix} \leq \frac{1}{2}f\begin{pmatrix}1\\0\end{pmatrix} + \frac{1}{2}f\begin{pmatrix}-1\\0\end{pmatrix},$$

or $f\begin{pmatrix}0\\0\end{pmatrix} = 0$, $f\begin{pmatrix}1\\0\end{pmatrix} = -1$, et $f\begin{pmatrix}-1\\0\end{pmatrix} = -1$, ce qui est contradictoire.

3. Avant de rechercher les points critiques de f, il convient tout d'abord de noter que cette application est bien continue et différentiable sur $\mathbb{R}^2$. En effet, en tant que fonction de x, à y fixé, c'est une application polynômiale, donc de classe C^1. De même, en tant que fonction de y, à x fixée, c'est une application polynômiale, donc de classe C^1. De plus, on calcule facilement les dérivées partielles :

$$\frac{\partial f}{\partial x}\begin{pmatrix}x\\y\end{pmatrix} = 4x^3-4x+4y, \quad \frac{\partial f}{\partial y}\begin{pmatrix}x\\y\end{pmatrix} = 4y^3-4y+4x.$$

Ces applications sont des applictions continues de $\begin{pmatrix}x\\y\end{pmatrix}$. Donc, d'après le théorème 8.27, f est différentiable.
Ainsi, les points critiques de f sont les solutions du système

$$\begin{cases} x^3-x+y &= 0, \\ y^3+x-y &= 0. \end{cases}$$

En sommant ces deux équations, on obtient $x^3+y^3=0$, donc $x=-y$. Si on reporte la valeur de y, les deux équations deviennent équivalentes, et s'écrivent $x^3 - 2x = 0$, c'est-à-dire $x(x^2 - 2) = 0$. Les solutions de cette équations sont $x = 0$, $x = \sqrt{2}$ et $x = -\sqrt{2}$, donc les points critiques de f sont

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \end{pmatrix}.$$

Les points de minimum de f sont nécessairement parmi ces trois points, d'après la proposition 9.4. On calcule la valeur de f en ces points :

$$f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0, \quad f\begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \end{pmatrix} = f\begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \end{pmatrix} = -8.$$

Donc f est minimale en les points $\begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \end{pmatrix}$ et $\begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \end{pmatrix}$, et la valeur correspondante est -8.

4. Ici encore, on remarque tout d'abord que f est bien de classe C^2. En effet, ses dérivées partielles sont de classe C^1 pour les mêmes raisons que précédemment. On calcule facilement sa matrice hessienne :

$$\frac{\partial^2 f}{\partial x^2}\begin{pmatrix} x \\ y \end{pmatrix} = 12x^2 - 4, \quad \frac{\partial^2 f}{\partial y^2}\begin{pmatrix} x \\ y \end{pmatrix} = 12y^2 - 4, \quad \frac{\partial^2 f}{\partial x \partial y}\begin{pmatrix} x \\ y \end{pmatrix} = 4.$$

Bien sûr, d'après le théorème de Schwarz, $\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial^2 f}{\partial y \partial x}$. Donc

$$D^2 f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 12x^2 - 4 & 4 \\ 4 & 12y^2 - 4 \end{pmatrix}.$$

Pour savoir si le point $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ est un extremum local, on calcule la hessienne en ce point :

$$D^2 f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix}.$$

Les deux colonnes de cette matrice sont proportionnelles, donc elle est singulière : 0 est donc valeur propre. On est donc dans un cas où le théorème 9.8 ne s'applique pas. Il faut donc étudier localement la fonction f au voisinage de l'origine :

$$\forall x \in \mathbb{R}^{+*}, \quad f\begin{pmatrix} x \\ x \end{pmatrix} = 2x^4 > 0, \quad f\begin{pmatrix} x \\ -x \end{pmatrix} = 2x^4 - 4x^2 = 2x^2(x^2 - 1).$$

Cette dernière valeur est strictement négative si $x < 1$. Donc l'origine n'est ni un point de maximum local, ni un point de minimum local.

En conclusion, la fonction f a deux minima globaux :

$$f\begin{pmatrix} \sqrt{2} \\ -\sqrt{2} \end{pmatrix} = f\begin{pmatrix} -\sqrt{2} \\ \sqrt{2} \end{pmatrix} = -8.$$

L'origine est son seul troisième point critique, et n'est ni un maximum local, ni un minimum local.

Chapitre 10

Théorème des fonctions implicites

10.1 Théorème d'inversion locale

Définition 10.1 *Soient E et F des espaces vectoriels normés, soit U un ouvert de E, V un ouvert de F. Soit $f : U \to V$.*

- *On dit que f est un homéomorphisme si elle est bijective, continue, et si f^{-1} est continue.*
- *On dit que f est un difféomorphisme si elle est bijective, de classe C^1, et si f^{-1} est de classe C^1.*
- *Soit $k \in \mathbb{N}^*$. On dit que f est un C^k-difféomorphisme si elle est bijective, de classe C^k, et si f^{-1} est de classe C^k.*

Notons que la définition d'un homéomorphisme a déjà été donnée dans le chapitre 3 (définition 3.33).

Remarque 10.2 *Si f est un difféomorphisme, alors f est un homéomorphisme. La réciproque est fausse, même si f est de classe C^1.*

En effet, l'application $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^3$ est bien un homéomorphisme de classe C^1. Son inverse est la fonction $x \to \sqrt[3]{x}$ et n'est pas dérivable en 0, donc f n'est pas un difféomorphisme.

Proposition 10.3 *Soient E et F des espaces de Banach, soit U un ouvert de E et V un ouvert de F. Soit $f : U \to V$ un homéomorphisme de classe C^1. Alors f est un difféomorphisme si et seulement si $\forall x \in U$, $df(x)$ est un isomorphisme de E dans F. Dans ce cas,*

$$\forall x \in U, \quad d\left(f^{-1}\right)(f(x)) = (df(x))^{-1}.$$

Démonstration. Supposons tout d'abord que f est un difféomorphisme. Alors f^{-1} est de classe C^1, et vérifie

$$\forall x \in U, \quad f^{-1}(f(x)) = x.$$

Donc la proposition 8.21 implique que

$$\forall x \in U, \quad d\left(f^{-1}\right)(f(x)) \circ df(x) = I_E,$$

où I_E désigne l'application identité de E dans E. De même, puisque $\forall y \in V, \quad f(f^{-1}(y)) = y$, on en déduit que $df(f^{-1}(y)) \circ d\left(f^{-1}\right)(y) = I_F$, où I_F désigne l'application identité de F dans F. Donc, en posant $y = f(x)$,

$$\forall x \in U, \quad df(x) \circ d\left(f^{-1}\right)(f(x)) = I_F.$$

Ces deux égalités prouvent que $df(x)$ est bien un isomorphisme d'espaces vectoriels, et que son application réciproque est donnée par $d\left(f^{-1}\right)(f(x))$.

Réciproquement, supposons que f est un homéomorphisme et que $\forall x \in U$, $df(x)$ est un isomorphisme. Soit $y \in V$. Alors, puisque f est bijective, il existe $x \in U$ tel que $f(x) = y$. Pour $h \in F$, on pose

$$\Delta(h) = f^{-1}(y+h) - f^{-1}(y) - (df(x))^{-1}(h).$$

Comme f est différentiable en x, on sait que

$$\forall k \in E, \quad f(x+k) = f(x) + df(x)(k) + \varepsilon(k)\|k\|_E,$$

où $\varepsilon : E \to F$ est une application qui tend vers 0 en 0. On applique cela à $k = f^{-1}(y+h) - f^{-1}(y)$:

$$f\left(x + f^{-1}(y+h) - f^{-1}(y)\right) = f(x) + df(x)\left(f^{-1}(y+h) - f^{-1}(y)\right) + \varepsilon(k)\|k\|_E.$$

Le membre de gauche de cette égalité s'écrit $f\left(x + f^{-1}(y+h) - f^{-1}(y)\right) = f\left(x + f^{-1}(y+h) - x\right) = f\left(f^{-1}(y+h)\right) = y + h$. Donc

$$y + h = y + df(x)\left(f^{-1}(y+h) - f^{-1}(y)\right) + \varepsilon(k)\|k\|_E.$$

On soustrait y à cette égalité, puis on lui applique $df(x)^{-1}$, ce qui donne

$$df(x)^{-1}(h) = f^{-1}(y+h) - f^{-1}(y) + df(x)^{-1}(\varepsilon(k))\|k\|_E.$$

Autrement dit,

$$\Delta(h) = -df(x)^{-1}(\varepsilon(k))\|k\|_E.$$

Comme on sait, par hypothèse, que $df(x)^{-1}$ est linéaire continue, il existe donc $M > 0$ tel que

$$\|\Delta(h)\|_F \leq M\|\varepsilon(k)\|_F\|k\|_F.$$

Nous aurons donc prouvé que f^{-1} est différentiable en y si on démontre que $\|k\|_E \leq C\|h\|_F$, pour une constante $C > 0$ indépendante de h et k. Pour cela, on écrit $k + x = f^{-1}(y+h)$, donc

$$f(x+k) = y + h.$$

En utilisant à nouveau la définition de la différentielle, on déduit

$$df(x)(k) + \varepsilon(k)\|k\|_E = h.$$

Ainsi, par application de l'inégalité triangulaire,

$$\|h\|_F \geq \|df(x)(k)\|_F - \|\varepsilon(k)\|_F \|k\|_E.$$

Comme l'application $df(x)$ est inversible d'inverse continue, il existe $\alpha > 0$ tel que $\|df(x)(k)\|_F \geq \alpha \|k\|_E$, pour tout $k \in E$. Donc

$$\|h\|_F \geq (\alpha - \|\varepsilon(k)\|_F)\, \|k\|_E.$$

Comme, puisque f^{-1} est continue, k tend vers 0 quand h tend vers 0, on a bien $\varepsilon(k) \to 0$ quand $h \to 0$. Donc, pour h assez petit, on a k suffisamment petit, pour avoir $\|\varepsilon(k)\|_F \leq \alpha/2$. D'où

$$\|h\|_F \geq \frac{\alpha}{2} \|k\|_E.$$

On obtient donc

$$\|\Delta(h)\|_E \leq M \|\varepsilon(k)\|_F \frac{2}{\alpha} \|h\|_F.$$

On remarque à nouveau que $\varepsilon(k) \to 0$ quand $h \to 0$. Ceci prouve donc que f^{-1} est différentiable, et que

$$\forall x \in U, \quad d\left(f^{-1}\right)(f(x)) = (df(x))^{-1}.$$

Il reste à prouver que l'application $x \mapsto df(x)^{-1}$ est continue de U dans $\mathcal{L}(F, E)$. Tout d'abord, on sait que $x \mapsto df(x)$ est continue de U dans $\mathcal{L}(E, F)$. De plus, si on note $\mathcal{GL}(E, F)$ l'ensemble des applications linéaires continues inversibles d'inverses continues de E dans F (i.e l'ensemble des isomorphismes de E dans F), alors

$$\begin{array}{rccc} \Phi: & \mathcal{GL}(E,F) & \longrightarrow & \mathcal{GL}(F,E) \\ & L & \longmapsto & L^{-1} \end{array}$$

est continue, comme le démontre le lemme ci-dessous. □

Lemme 10.4 *Soient E et F des espaces de Banach. Soit $\mathcal{GL}(E, F)$ l'ensemble des isomorphismes de E dans F. Alors l'application*

$$\begin{array}{rccc} \Phi: & \mathcal{GL}(E,F) & \longrightarrow & \mathcal{GL}(F,E) \\ & L & \longmapsto & L^{-1} \end{array}$$

est continue.

Démonstration. Rappelons tout d'abord que la norme sur $\mathcal{L}(E, F)$ est définie par la définition 1.33, que nous rappelons ici :

$$\|L\|_{\mathcal{L}(E,F)} = \sup_{x \in E \setminus \{0\}} \frac{\|L(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E = 1} \|L(x)\|_F.$$

De plus, on a, si G est un espace vectoriel normé :

$$\forall A \in \mathcal{L}(E, F), \quad \forall B \in \mathcal{L}(F, G), \quad \|B \circ A\|_{\mathcal{L}(E,G)} \leq \|A\|_{\mathcal{L}(E,F)} \|B\|_{\mathcal{L}(F,G)}.$$

Dans la suite, on notera $BA = B \circ A$.

Soit $L \in \mathcal{GL}(E,F)$. Soit $H \in \mathcal{L}(E,F)$ tel que

$$\|H\|_{\mathcal{L}(E,F)} < \frac{1}{\|L^{-1}\|_{\mathcal{L}(E,F)}}.$$

Nous allons tout d'abord montrer que $L+H \in \mathcal{GL}(E,F)$. Pour cela, on écrit

$$L + H = L\left(I_E + L^{-1}H\right),$$

où I_E est l'application identité de E dans E. Il suffit donc de montrer que $L^{-1}H \in \mathcal{GL}(E,E)$. Pour cela, on pose

$$K_n = \sum_{k=0}^{n} (-1)^k \left(L^{-1}H\right)^k = I_E - L^{-1}H + \left(L^{-1}H\right)^2 + \cdots + (-1)^n \left(L^{-1}H\right)^n.$$

On utilise la convention que $\left(L^{-1}H\right)^0 = I_E$. Un calcul simple montre que

$$\begin{aligned}
\left(I_E + L^{-1}H\right) K_n &= \sum_{k=0}^{n} (-1)^k \left(L^{-1}H\right)^k + \sum_{k=0}^{n} (-1)^k \left(L^{-1}H\right)^{k+1} \\
&= I_E + \sum_{k=1}^{n} (-1)^k \left(L^{-1}H\right)^k + \sum_{k=0}^{n-1} (-1)^k \left(L^{-1}H\right)^{k+1} + (-1)^n \left(L^{-1}H\right)^{n+1} \\
&= I_E + \left(\sum_{k=1}^{n} \left[(-1)^k + (-1)^{k-1}\right] \left(L^{-1}H\right)^k\right) + (-1)^n \left(L^{-1}H\right)^{n+1} \\
&= I_E + (-1)^n \left(L^{-1}H\right)^{n+1}.
\end{aligned}$$

De plus, on a, pour tout entier k,

$$\left\|\left(L^{-1}H\right)^k\right\|_{\mathcal{L}(E,E)} \leq \left(\left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}\right)^k \underset{k\to+\infty}{\longrightarrow} 0,$$

puisque $\left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)} < 1$. D'autre part, toujours à cause de cette estimation, la série de terme général $(-1)^k \left(L^{-1}H\right)^k$ est normalement convergente, car la norme de son terme général est majorée par celui d'une série géométrique convergente. Donc, puisque $\mathcal{L}(E,E)$ est complet, la proposition 6.30 implique que la suite $(K_n)_{n\in\mathbb{N}}$ converge dans $\mathcal{L}(E,E)$. Soit K sa limite. On a alors, d'après ce qui précède, $\left(I_E + L^{-1}H\right) K = I_E$, donc

$$(L+H)KL^{-1} = L^{-1}\left(I_E + L^{-1}H\right) KL^{-1} = LL^{-1} = I_F.$$

On peut refaire exactement la même preuve pour obtenir

$$KL^{-1}(L+H) = K\left(I_E + L^{-1}H\right) = I_E.$$

Donc $L+H$ est bien inversible, et son inverse KL^{-1} est dans $\mathcal{L}(F,E)$. Donc $L+H \in \mathcal{GL}(E,F)$.

Il reste à prouver que, quand $H \to 0$, $(L+H)^{-1} \to L^{-1}$. Pour cela, on remarque que

$$\begin{aligned}
\|K_n - I_E\|_{\mathcal{L}(E,E)} &\le \sum_{k=1}^{n} \left(\left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}\right)^k \\
&= \left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)} \frac{1-\left(\left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}\right)^n}{1-\|L^{-1}\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}} \\
&\le \left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)} \frac{1}{1-\|L^{-1}\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}}.
\end{aligned}$$

Donc, puisque la suite $(K_n)_{n\in\mathbb{N}}$ converge vers K pour la norme $\|\cdot\|_{\mathcal{L}(E,E)}$ et que le majorant est indépendant de n, on peut passer à la limite et obtenir

$$\|K - I_E\|_{\mathcal{L}(E,E)} \le \left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)} \frac{1}{1-\|L^{-1}\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}}.$$

Ceci implique que

$$\begin{aligned}
\left\|(L+H)^{-1} - L^{-1}\right\|_{\mathcal{L}(F,E)} &= \left\|KL^{-1} - L^{-1}\right\|_{\mathcal{L}(F,E)} \\
&\le \|K - I_E\|_{\mathcal{L}(E,E)} \left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \\
&\le \left\|L^{-1}\right\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)} \frac{\left\|L^{-1}\right\|_{\mathcal{L}(F,E)}}{1-\|L^{-1}\|_{\mathcal{L}(F,E)} \|H\|_{\mathcal{L}(E,F)}}.
\end{aligned}$$

Donc

$$\lim_{H\to 0} \left\|(L+H)^{-1} - L^{-1}\right\|_{\mathcal{L}(F,E)} = 0,$$

ce qui prouve bien que l'application Φ est continue. □

Remarque 10.5 *Dans le cas où E et F sont $\mathbb{R}^n$ et $\mathbb{R}^p$, respectivement, le lemme ci-dessus devient trivial, car l'application $A \to A^{-1}$ qui à une matrice inversible associe son inverse est une fonction continue des coefficients de A. En effet,*

$$A^{-1} = \frac{1}{\det(A)} \operatorname{com}(A)^T,$$

où com(A) est la comatrice de A. Cette matrice est une fonction polynômiale des coefficients de A, donc continue.

Théorème 10.6 (d'inversion locale) *Soient E et F des espaces de Banach. Soit U un ouvert de E, et soit $f : U \to F$ une application de classe C^1. Soit $a \in U$ tel que $df(a)$ est un isomorphisme de E dans F. Alors il existe un voisinage U_1 de a dans U et un voisinage V_1 de $f(a)$ dans F tels que $f : U_1 \to V_1$ est un difféomorphisme.*

Démonstration. On commence par se ramener au cas où $E = F$, $a = f(a) = 0$ et $df(a)$ est l'application identité. Pour cela, on pose

$$\begin{array}{rccl} h: & E & \longrightarrow & E \\ & x & \longmapsto & df(a)^{-1}(f(x+a) - f(a)), \end{array}$$

qui est une application de $U - a$ dans E. On a bien

$$dh(x) = df(a)^{-1} \circ df(x+a),$$

donc $dh(0) = I_E$, l'application identité de E dans E. De plus, $h(0) = 0$. Enfin, le fait que $f : U_1 \to V_1$ soit un difféomorphisme équivaut au fait que $h : U_1 - a \to W_1$ soit un difféomorphisme, avec $W_1 = df(a)^{-1}(V_1 - f(a)) \subset E$.

Démontrons maintenant que pour tout $y \in F$ proche de 0 l'équation $h(x) = y$ admet une unique solution. On a :

$$h(x) = y \iff x - h(x) + y = x.$$

On pose donc $\varphi : U - a \to E$ définie par $\varphi(x) = x - h(x) + y$. Cette application est de classe C^1, et $d\varphi(x) = I_E - dh(x)$. En particulier, $d\varphi(0) = 0$. Donc

$$\exists \varepsilon > 0 \quad / \quad \forall x \in B_\varepsilon(0), \quad \|d\varphi(x)\|_{\mathcal{L}(E,F)} \leq \frac{1}{2}.$$

Quitte à diminuer ε, on peut suppose $B_\varepsilon(0) \subset U - a$, et donc, par application de l'inégalité des accroissements finis,

$$\forall x \in B_\varepsilon(0), \quad \|\varphi(x) - \varphi(0)\|_E \leq \frac{1}{2}\|x - 0\|_E \leq \frac{\varepsilon}{2}.$$

Donc $\varphi(B_\varepsilon(0)) \subset B_{\varepsilon/2}(\varphi(0)) = B_{\varepsilon/2}(y)$. En particulier, pour $\|y\| \leq \frac{\varepsilon}{2}$, on a donc

$$\varphi(B_\varepsilon(0)) \subset B_\varepsilon(0).$$

De plus, toujours par application de l'inégalité des accroissements finis,

$$\forall x \in B_\varepsilon(0), \quad \forall x' \in B_\varepsilon(0), \quad \|\varphi(x) - \varphi(x')\|_E \leq \frac{1}{2}\|x - x'\|_E.$$

Donc l'application φ est contractante. Par application du théorème du point fixe (théorème 6.22), il existe un unique $x \in B_\varepsilon(0)$ tel que $\varphi(x) = x$. Donc

$$\forall y \in B_{\varepsilon/2}, \quad \exists! x \in B_\varepsilon(0) \quad / \quad y = h(x).$$

Démontrons maintenant que h^{-1} est continue de $B_{\varepsilon/2}(0)$ sur $U_1 = B_\varepsilon(0) \cap h^{-1}(B_{\varepsilon/2}(0))$: soient y, y' dans $B_{\varepsilon/2}(0)$. On sait qu'il existe x et x' dans $B_\varepsilon(0)$ tels que $y = h(x)$, $y' = h(x')$. On a donc, en utilisant l'inégalité triangulaire,

$$\begin{aligned} &\|y - y'\|_E = \|h(x) - h(x')\|_E = \|h(x) - x - h(x') + x' + x - x'\|_E \\ &\geq \|x - x'\|_E - \|h(x) - x - (h(x') - x')\|_E = \|x - x'\|_E - \|\varphi(x) - \varphi(x')\|_E \\ &\qquad\qquad\qquad\qquad \geq \|x - x'\|_E - \frac{1}{2}\|x - x'\|_E = \frac{1}{2}\|x - x'\|_E. \end{aligned}$$

Autrement dit, $\|h(y) - h(y')\|_E \leq 2\|x - x'\|_E$. Donc h est lipschitzienne, donc continue.

Il reste maintenant à appliquer la proposition 10.3 pour conclure que h est un difféomorphisme. □

Exemple 10.7 *En dimension 1, ce théorème prouve que si $f : \mathbb{R} \to \mathbb{R}$ est une application C^1 tel que $f'(x_0) \neq 0$, alors il existe un voisinage U_1 de x_0 et un voisinage V_1 de $f(x_0)$ tels que $f : U_1 \to V_1$ est un difféomorphisme. En particulier, $f^{-1} : V_1 \to U_1$ est de classe C^1.*

Remarque 10.8 *Il convient de remarquer que $df(a) \neq 0$ n'est pas suffisant pour appliquer le théorème d'inversion locale, sauf en dimension 1 (comme on vient de le voir) : il faut que l'application linéaire $df(a)$ soit un isomorphisme.*

Remarque 10.9 *En prolongement de la remarque précédente, l'hypothèse du théorème implique que $df(a)$ est un isomorphisme de E dans F, donc que E et F sont isomorphes. Dans le cas d'espaces de dimension finie, ils doivent donc avoir la même dimension.*

Remarque 10.10 *Dans le cas où $E = F = \mathbb{R}^n$, le fait que $df(a)$ soit un isomorphisme équivaut au fait que sa matrice dans les bases canoniques soit inversibles, c'est-à-dire*

$$\det(Df(a)) \neq 0.$$

Exemple 10.11 *Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ l'application définie par $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y^4 \\ y + x^3 y \end{pmatrix}$. Cette application est bien de classe C^1, et sa jacobienne s'écrit*

$$\forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \quad Df\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 4y^3 \\ 3x^2 y & 1 + x^3 \end{pmatrix}.$$

En particulier,

$$Df\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Donc f est un difféomorphisme sur un voisinage U de l'origine. Autrement dit, pour (a, b) suffisamment proche de $(0, 0)$, le système

$$\begin{cases} x + y^4 &= a \\ y + x^3 y &= b \end{cases}$$

admet une unique solution (x, y) qui dépend de façon C^1 de (a, b). Si on cherchait à résoudre directement le système ci-dessus, on pourrait remplacer $x = a - y^4$ par sa valeur dans la deuxième équation. Ceci donnerait une équation de degré 13 à résoudre en y.

Théorème 10.12 (d'inversion globale) *Soit E un espace de Banach, et soit U un ouvert de E. Soit $f : U \to E$ une application de classe C^1 telle que f est bijective de U dans $f(U)$ et que $\forall x \in U$, $df(x)$ est un isomorphisme de E dans E. Alors $f(U)$ est un ouvert, et f est un difféomorphisme de U dans $f(U)$.*

Démonstration. Démontrons tout d'abord que $f(U)$ est un ouvert. Soit donc $b \in f(U)$. Par définition, il existe $a \in U$ tel que $b = f(a)$. Par hypothèse, $df(a)$ est un isomorphisme. Le théorème d'inversion locale s'applique : il existe U' voisinage de a et V' voisinage de $f(a)$ tels que f est un difféomorphisme de U' dans V'. Donc $V' = f(U') \subset f(U)$. Donc $f(U)$ est bien un ouvert. On sait déjà que f est bijective de U dans $f(U)$, et qu'elle est C^1. La proposition 10.3 permet de prouver qu'il s'agit d'un difféomorphisme. □

Exemple 10.13 *L'application ϕ définie par*

$$\begin{array}{cccc} \phi : & \mathbb{R}^{+*} \times]-\pi, \pi[& \longrightarrow & \mathbb{R}^2 \setminus (\mathbb{R}^{-*} \times \{0\}) \\ & \begin{pmatrix} r \\ \theta \end{pmatrix} & \longmapsto & \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}, \end{array}$$

est un difféomorphisme.

Démonstration. Il est clair que ϕ est bijective. De plus, elle est de classe C^1, et ses dérivées partielles s'écrivent

$$\frac{\partial\phi}{\partial r}\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}, \quad \frac{\partial\phi}{\partial\theta} = \begin{pmatrix} -r\sin\theta \\ r\cos\theta \end{pmatrix}.$$

Donc la jacobienne s'écrit $D\phi\begin{pmatrix} r \\ \theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$. Donc

$$\det\left(D\phi\begin{pmatrix} r \\ \theta \end{pmatrix}\right) = r\left(\cos^2\theta + \sin^2\theta\right) = r \neq 0.$$

Ceci prouve que la différentielle est inversible en tout point, donc, d'après le théorème d'inversion globale, ϕ est un difféomorphisme. □

On peut donc maintenant utiliser ce changement de variables polaires pour calculer le gradient d'une fonction dans ces coordonnées : soit $g : \mathbb{R}^2 \to \mathbb{R}$ une application de classe C^1, et posons

$$\forall (r, \theta) \in \mathbb{R}^{+*} \times]-\pi, \pi[, \quad h(r, \theta) = g(r\cos\theta, r\sin\theta).$$

On a donc $h = g \circ \phi$. Cette application est donc différentiable, et donc, grâce au corollaire 8.23, la jacobienne de h est le produit des jacobiennes de ϕ et de g.

$$Dh\begin{pmatrix} r \\ \theta \end{pmatrix} = Dg\left(\phi\begin{pmatrix} r \\ \theta \end{pmatrix}\right) D\phi\begin{pmatrix} r \\ \theta \end{pmatrix}.$$

C'est-à-dire

$$\begin{pmatrix} \frac{\partial h}{\partial r} & \frac{\partial h}{\partial\theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{pmatrix} \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Ou encore

$$\begin{cases} \dfrac{\partial h}{\partial r}\begin{pmatrix} r \\ \theta \end{pmatrix} = \cos\theta \dfrac{\partial g}{\partial x}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} + \sin\theta \dfrac{\partial g}{\partial y}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}, \\ \dfrac{\partial h}{\partial\theta}\begin{pmatrix} r \\ \theta \end{pmatrix} = -r\sin\theta \dfrac{\partial g}{\partial x}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} + r\cos\theta \dfrac{\partial g}{\partial y}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}. \end{cases}$$

On peut inverser cette formule, et on obtient

$$\left\{\begin{array}{l} \dfrac{\partial g}{\partial x}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \cos\theta \dfrac{\partial h}{\partial r}\begin{pmatrix} r \\ \theta \end{pmatrix} - \dfrac{\sin\theta}{r}\dfrac{\partial h}{\partial \theta}\begin{pmatrix} r \\ \theta \end{pmatrix}, \\ \dfrac{\partial g}{\partial y}\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \sin\theta \dfrac{\partial h}{\partial r}\begin{pmatrix} r \\ \theta \end{pmatrix} + \dfrac{\cos\theta}{r}\dfrac{\partial h}{\partial \theta}\begin{pmatrix} r \\ \theta \end{pmatrix}. \end{array}\right.$$

En particulier, en posant

$$\vec{e}_r = (\cos\theta, \sin\theta), \quad \vec{e}_\theta = (-\sin\theta, \cos\theta),$$

on obtient donc

$$\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right)\begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} = \frac{\partial h}{\partial r}\begin{pmatrix} r \\ \theta \end{pmatrix}\vec{e}_r + \frac{1}{r}\frac{\partial h}{\partial \theta}\begin{pmatrix} r \\ \theta \end{pmatrix}\vec{e}_\theta.$$

Cette formule correspond à la formule classique utilisée en mécanique.

10.2 Théorème des fonctions implicites

Théorème 10.14 *Soient E, F, G des espaces de Banach. Soit Ω un ouvert de $E \times F$, et soit $f : \Omega \to G$ une application de classe C^1. Soit $(x_0, y_0) \in \Omega$ tel que $f(x_0, y_0) = 0$. On suppose de plus que*

$$\forall (h,k) \in E \times F, \quad df(x_0, y_0)(h,k) = A(h) + B(k),$$

avec $A \in \mathcal{L}(E,G)$ et $B \in \mathcal{L}(F,G)$.

Si B est un isomorphisme de F dans G, alors il existe U voisinage de x_0 et V voisinage de y_0 tels que

$$\exists! \varphi \in C^1(U,V) \quad / \quad \forall (x,y) \in U \times V, \quad f(x,y) = 0 \iff y = \varphi(x).$$

De plus,

$$d\varphi(x) = -B(x, \varphi(x))^{-1} A(x, \varphi(x)),$$

où $df(x,y)(h,k) = A(x,y)(h) + B(x,y)(k)$.

Remarque 10.15 *La différentielle $df(x,y)$ de f en (x,y) peut toujours s'écrire $df(x,y)(h,k) = A(x,y)(h) + B(x,y)(k)$, car*

$$df(x,y)(h,k) = df(x,y)(h,0) + df(x,y)(0,k).$$

Démonstration. Soit

$$\begin{array}{rccl} g : & \Omega & \longrightarrow & E \times G \\ & (x,y) & \longmapsto & (x, f(x,y)). \end{array}$$

Alors g est de classe C^1 (le démontrer en exercice) et

$$\forall (h,k) \in E \times F, \quad dg(x,y)(h,k) = (h, df(x,y)(h,k)).$$

Donc

$$\forall (h,k) \in E \times F, \quad dg(x_0,y_0)(h,k) = (h, A(x_0,y_0)h + B(x_0,y_0)k)\,.$$

Cette application linéaire est continue. De plus, elle est inversible. En effet,

$$dg(x_0,y_0)(h,k) = (\alpha,\beta) \iff \begin{cases} h = \alpha, \\ A(x_0,y_0)h + B(x_0,y_0)k = \beta \end{cases}$$
$$\iff \begin{cases} h = \alpha, \\ k = B(x_0,y_0)^{-1}(\beta - A(x_0,y_0)\alpha). \end{cases}$$

Enfin, l'inverse de $dg(x_0,y_0)$ est continue, car $\|h\|_E = \|\alpha\|_E$, et

$$\|k\|_F \leq C\|\beta\|F + C'\|\alpha\|_E,$$

puisque $B(x_0,y_0)^{-1}$ et $A(x_0,y_0)$ sont continues.

On peut donc appliquer le théorème d'inversion locale : il existe U' un voisinage de (x_0,y_0) dans $E \times F$ et V' un voisinage de $(x_0,0)$ dans $E \times G$ tels que $g : U' \to V'$ est un difféomorphisme. En particulier, $g^{-1} : V' \to U'$ est de classe C^1. On pose alors

$$U = \{x \in E \quad / \quad \exists y \in G, \quad (x,y) \in V'\}\,,$$

et

$$V = \{y \in F \quad / \quad \exists x \in E, \quad (x,y) \in U'\}\,.$$

L'ensemble U est donc bien un voisinage de x_0, et V un voisinage de y_0. Soit

$$\begin{array}{rccl} \varphi : & U & \longrightarrow & V \\ & x & \longmapsto & \left(g^{-1}(x,0)\right)_2, \end{array}$$

c'est-à-dire que $\varphi(x)$ est l'unique $y \in V$ tel que $g(x,y) = (x,0)$. On a donc bien

$$\varphi(x) = y \iff g(x,y) = (x,0) \iff (x, f(x,y)) = (x,0) \iff f(x,y) = 0.$$

Donc $f(x,\varphi(x)) = 0$. Enfin, par dérivation des fonctions composées, on a, en posant $df(x,y)(h,k) = A(x,y)(h) + B(x,y)(k)$,

$$B(x,\varphi(x))\,[d\varphi(x)(h)] + A(x,\varphi(x))(h) = 0,$$

ce qui termine la preuve. $\square$

Exemple 10.16 *Si $E = F = G = \mathbb{R}$, alors*

$$\begin{array}{rccl} f : & \mathbb{R} \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & f(x,y), \end{array}$$

et l'hypothèse que B est un isomorphisme équivaut à ce que l'application

$$\begin{array}{rccl} B : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & h & \longmapsto & \frac{\partial f}{\partial y}(x_0,y_0)h \end{array}$$

est inversible. Ceci équivaut donc à $\frac{\partial f}{\partial y}(x_0,y_0) \neq 0$. Autrement dit, la courbe d'équation $f(x,y) = 0$ ne doit pas avoir de tangente verticale en (x_0,y_0).

Comme exemple concret, on peut considérer

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

Alors $\frac{\partial f}{\partial y}(x_0, y_0) = 2y_0$. Donc tant que $y_0 \neq 0$, on peut définir la courbe d'équation $f(x, y) = 0$ par une application $y = \varphi(x)$ au voisinage de (x_0, y_0). Un calcul explicite donne

$$f(x, y) = 0 \iff y^2 = 1 - x^2 \iff y = \mathrm{sgn}(y_0)\sqrt{1 - x^2}.$$

Ici, sgn désigne la fonction signe. On a donc

$$\varphi(x) = \mathrm{sgn}(y_0)\sqrt{1 - x^2},$$

qui est bien définie au voisinage de x_0. Pour $y_0 = 0$, cette fonction n'est pas bien définie.

Remarque 10.17 *Il se peut que φ existe sans que B soit inversible.*

Démonstration. On peut par exemple prendre $E = F = G = \mathbb{R}$, et la fonction $f(x, y) = x - y^3$. Alors l'équation $f(x, y) = 0$ se résout explicitement : $y = \sqrt[3]{x}$. Donc $\varphi(x) = \sqrt[3]{x}$ est bien définie pour tout $x \in \mathbb{R}$. Cependant, $\frac{\partial f}{\partial y}(x, y) = 3y^2$, qui est nulle en $(x_0, y_0) = (0, 0)$. Donc $B = 0$ en ce point, et n'est donc pas inversible. □

10.3 Extrema sous contraintes

10.3.1 Cas d'une seule contrainte

Théorème 10.18 *Soit U un ouvert de $\mathbb{R}^n$, et soient $f, g : U \to \mathbb{R}$ des applications de classe C^1. On pose*

$$\Gamma = \{x \in U, \quad g(x) = 0\} = g^{-1}(\{0\}).$$

On suppose que $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$, et que $dg(a) \neq 0$. Alors

$$\exists \lambda \in \mathbb{R} \quad / \quad df(a) + \lambda dg(a) = 0.$$

Le réel λ est appelé multiplicateur de Lagrange associé à la contrainte $g(x) = 0$.

L'interprétation géométrique de ce résultat est qu'au point a, puisqu'il est un point d'extremum local pour f sur Γ, le vecteur $Df(a)$ doit être orthogonal à Γ, donc colinéaire à $Dg(a)$.

Démonstration. Comme $dg(a) \neq 0$, la jacobienne $Dg(a)$ est non nulle. Quitte à renuméroter les variables, on peut toujours supposer que $\frac{\partial g}{\partial x_n}(a) \neq 0$. On pose alors

$$a' = (a_1, \dots a_{n-1}).$$

On applique le théorème des fonctions implicites à l'application g, vue comme application de $\mathbb{R}^{n-1}\times\mathbb{R}$ dans $\mathbb{R}$. On a bien $\frac{\partial g}{\partial x_n}(a)\neq 0$, donc il existe V voisinage de a et V' voisinage de a' et $\varphi : V'\to\mathbb{R}$ tel que

$$\forall x\in V,\quad x\in\Gamma \iff x_n=\varphi(x').$$

On définit donc

$$\begin{array}{rccl} F: & V' & \longrightarrow & \mathbb{R} \\ & x' & \longmapsto & f(x',\varphi(x')). \end{array}$$

Comme $f_{|\Gamma}$ admet un extremum local en a, F admet un extremum local en a'. Donc, d'après la proposition 9.4, $dF(a')=0$. Or

$$\forall i\in\{1,2,\ldots,n-1\},\quad \frac{\partial F}{\partial x_i}(a')=\frac{\partial f}{\partial x_i}(a)+\frac{\partial f}{\partial x_n}(a)\frac{\partial\varphi}{\partial x_i}(a').$$

Par ailleurs, le théorème des fonctions implicites implique également que

$$0=\frac{\partial g}{\partial x_i}(a)+\frac{\partial g}{\partial x_n}(a)\frac{\partial\varphi}{\partial x_i}(a').$$

Donc, en reportant cette équation dans l'équation $\frac{\partial F}{\partial x_i}(a')=0$, on a

$$0=\frac{\partial f}{\partial x_i}(a)-\frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial g}{\partial x_n}(a)}\frac{\partial g}{\partial x_i}(a).$$

En posant $\lambda=-\frac{\frac{\partial f}{\partial x_n}(a)}{\frac{\partial g}{\partial x_n}(a)}$, on obtient donc le résultat. □

Exemple 10.19 *Calculons les extrema de $f(x,y)=x+y$ sur la courbe*

$$\Gamma=\left\{(x,y)\in\mathbb{R}^2\quad/\quad x^4+y^4=1\right\}.$$

On calcule facilement la jacobienne de $g(x,y)=x^4+y^4-1$:

$$Dg(x,y)=\begin{pmatrix}4x^3 & 4y^3\end{pmatrix},$$

qui n'est jamais nulle sur Γ. De plus, Γ est un fermé borné, donc un compact de $\mathbb{R}^2$. Donc f, qui est continue, admet un minimum et un maximum sur Γ. En un point d'extremum (a,b), on doit avoir

$$Df(a,b)+\lambda Dg(a,b)=0.$$

Comme $Df(a,b)=(1,1)$, on obtient

$$\left\{\begin{array}{l}\lambda 4a^3=1,\\ \lambda 4b^3=1.\end{array}\right.$$

Donc $a=b$. Ainsi, $2a^4=1$ puisque $(a,b)\in\Gamma$. Donc

$$(a,b)=\left(2^{-1/4},2^{-1/4}\right)\quad\text{ou}\quad(a,b)=-\left(2^{-1/4},2^{-1/4}\right).$$

Dans le premier cas, f vaut $2^{3/4}$, dans le deuxième cas elle vaut $-2^{3/4}$. Le premier point est donc le point de maximum, le second le point de minimum de f sur Γ.

10.3.2 Cas de plusieurs contraintes

Théorème 10.20 *Soit U un ouvert de $\mathbb{R}^n$, et soient $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}^p$ des applications de classe C^1. On pose $\Gamma = g^{-1}(\{0\})$. Si $f_{|\Gamma}$ admet un extremum local en $a \in \Gamma$, et si $dg(a)$ est surjective, alors*

$$\exists(\lambda_1, \lambda_2, \dots, \lambda_p) \in \mathbb{R}^p \quad / \quad df(a) = \sum_{i=1}^{p} \lambda_i dg_i(a).$$

Les réels λ_i, pour $1 \leq i \leq p$, sont appelés multiplicateurs de Lagranges associés aux contraintes g_i.

Dans l'énoncé ci-dessus, on a noté g_i la composante i de l'application g.
Démonstration. L'application linéaire $dg(a)$ est surjective, donc son rang vaut p. Donc $\dim(\ker(dg(a))) = n - p$. Soit F un supplémentaire de $\ker(dg(a))$. Alors $\dim F = p$, et

$$\mathbb{R}^n = F \oplus \ker(dg(a)).$$

Quitte à changer de variables, on peut supposer que

$$\ker(dg(a)) = \mathbb{R}^{n-p} \times \{0\}, \quad F = \{0\} \times \mathbb{R}^p.$$

Dans la première égalité, 0 désigne le vecteur nul de $\mathbb{R}^p$, et dans la deuxième, celui de $\mathbb{R}^{n-p}$. Pour tout $x \in \mathbb{R}^n$, on pose donc $x = (x_K, x_F)$, avec $(x_K, 0, \dots, 0) \in \ker(dg(a))$ et $(0, \dots, 0, x_F) \in F$. On a alors $g(a) = 0$, et la différentielle de g par rapport à x_F, notée $d_{x_F} g(a)$, est inversible. On peut donc appliquer le théorème des fonctions implicites : il existe U voisinage de a_F, V voisinage de a_K et $\varphi : U \to V$ difféomorphisme tels que

$$\forall (y, z) \in U \times V, \quad (y, z) \in \Gamma \iff z = \varphi(y).$$

On pose alors

$$\begin{array}{rccl} F: & U & \longrightarrow & \mathbb{R} \\ & y & \longmapsto & f(y, \varphi(y)). \end{array}$$

Cette application admet un extremum en a_F. Donc

$$d_y f(a_F, \varphi(a_F)) + d_z f(a) \circ d\varphi(a_F) = 0,$$

où on note $d_y f$ la différentielle par rapport à $y \in \ker(dg(a))$ et $d_z f$ celle par rapport à $z \in F$. On a également

$$d_y g(a) + d_z g(a) \circ d\varphi(a_F) = 0, \text{ donc } d\varphi(a_F) = (d_z g(a))^{-1} \circ d_y g(a).$$

L'écriture matricielle de ces égalités donne

$$D_y f(a) + D_z f(a) D\varphi(a_F) = 0,$$

c'est-à-dire

$$D_y f(a) = D_z f(a) \left(D_z g(a)\right)^{-1} D_y g(a).$$

Ainsi, en posant $\lambda = (\lambda_1, \ldots, \lambda_p) = D_z g(a)^{-1} D_z f(a) \in \mathbb{R}^p$, on déduit

$$\frac{\partial f}{\partial y_j}(a) = \sum_{i=1}^{p} \lambda_i \frac{\partial g_i}{\partial y_j}(a),$$

ce qui est bien le résultat voulu. $\square$