U.S.P.N - Sup Galilée
Formation Ingénieurs MACS

Année scolaire 2024-2025

Analyse

Le 12 septembre 2024

Bibliographie

[1] J. Faraut, *Calcul intégral*, 2006, EDP Sciences.

[2] J.P. Marco, *Analyse pour la licence*, Dunod.

[3] J.P. Marco et autres, *Mathématiques L3, Analyse* , Pearson Education France.

[4] F. Rouvière, *Petit guide de calcul différentiel à l'usage de la licence et de l'Agrégation*, Cassini.

Table des matières

Chapitre 1

Topologie

1.1 Un mot de topologie générale

On commence par se donner un ensemble E.

Définition 1.1.1. *On appelle topologie sur E toute famille de parties de E, soit $\mathcal{T}$, vérifiant les propriétés suivantes :*

(i) $\varnothing \in \mathcal{T}$ *et* $E \in \mathcal{T}$

(ii) $\mathcal{T}$ *est stable par réunion.*

(iii) $\mathcal{T}$ *est stable par intersection finie.*

Définition 1.1.2. *Le couple* $(E, \mathcal{T})$ *est appelé espace topologique ; les éléments de* $\mathcal{T}$ *sont alors appelés ouverts de cet espace topologique.*

Traduisons les assertions *(ii)* et *(iii)* :

(ii) Si $(\Omega_i)_{i \in I}$ est une famille d'ouverts de E : $\Omega = \bigcup_{i \in I} \Omega_i$ est un ouvert de E.

(iii) Si $\Omega_1, \ldots, \Omega_n$ est une famille finie d'ouverts de E : $\Omega = \bigcap_{i=1}^{n} \Omega_i$ est un ouvert de E.

Exemple 1 : Soit E un ensemble. Alors $\mathcal{O} = \{\varnothing, E\}$ est une topologie sur E, appelée *topologie grossière*.

Exemple 2 : Soit E un ensemble. Alors $\mathcal{O} = \mathcal{P}(E)$ est une topologie sur E, appelée *topologie discrète*.

Exemple 3 : Donnons à présent un exemple plus courant : la topologie usuelle de $\mathbb{R}$.
Une partie Ω de $\mathbb{R}$ est dite ouverte lorsque, pour tout x dans Ω, il existe un réel α strictement positif tel que : $]x-\alpha, x+\alpha[\subset \Omega$.

Démonstration : *(i)* $\mathbb{R}$ et $\varnothing$ sont ouverts.

(ii) Soit $(\Omega_i)_{i \in I}$ une famille d'ouverts de $\mathbb{R}$. Soit $\Omega = \bigcup_{i \in I} \Omega_i$. Par définition d'une réunion : $\exists i \in I, x \in \Omega_i$. Comme Ω_i est ouvert : $(\exists \alpha > 0)(]x-\alpha, x+\alpha[\subset \Omega_i)$. De là : $]x-\alpha, x+\alpha[\subset \Omega_i \subset \Omega$.

(iii) Soient $\Omega_1, \ldots, \Omega_n$, n ouverts de $\mathbb{R}$ et $\Omega = \bigcap_{i=1}^{n} \Omega_i$. Soit $x \in \Omega$. Par définition :

$$(\forall i \in [\![1, n]\!])(\exists \alpha_i > 0)(]x-\alpha_i, x+\alpha_i[\subset \Omega_i)$$

Posons $\alpha = \min_{1 \leqslant i \leqslant n} \alpha_i$. Alors $]x-\alpha, x+\alpha[\subset \Omega_i$ pour tout i dans $[\![1, n]\!]$. Donc : $]x-\alpha, x+\alpha[\subset \Omega$.

□

Cette caractérisation des ouverts de $\mathbb{R}$ nous permet d'exhiber un premier exemple d'une intersection infinie d'ouverts qui n'est pas ouverte. Pour $n \geqslant 1$ on considère les ouverts :

$$\Omega_n = \left]-\frac{1}{n}, \frac{1}{n}\right[.$$

Il vient : $\bigcap_{n\geqslant 1}\Omega_n = \{0\}$ et $\{0\}$ n'est pas ouvert. En effet : $\forall\alpha > 0,]-\alpha,\alpha[\not\subset \{0\}$.

Définition 1.1.3. *Une partie F de E est dite fermée lorsque son complémentaire dans E est un ouvert de E.*

On obtient alors les propriétés de réunion et d'intersection duales de celles pour les ouverts.

Proposition 1.1.4. *On a :*
(i) $\varnothing$ et E sont fermés.
(ii) Toute intersection de fermés est fermée.
(iii) Toute réunion finie de fermés est fermée.

Démonstration : *(i)* Les complémentaires de $\varnothing$ et de E sont respectivement E et $\varnothing$ qui sont tous deux ouverts.
(ii) Soit $(F_i)_{i\in I}$ une famille de fermés de E. Pour tout i dans I on note : $O_i = E\backslash F_i$. Alors, O_i est ouvert dans E et donc $\bigcup_{i\in I} O_i$ est ouverte dans E. De là :

$$E\backslash\bigcup_{i\in I} O_i = \bigcap_{i\in I} E\backslash O_i = \bigcap_{i\in I} F_i \quad \text{est fermée.}$$

(iii) Soit $F_1,\ldots,F_n$ une famille finie de fermés de E. Pour tout i dans $[\![1,n]\!]$ on note : $O_i = E\backslash F_i$. Alors, O_i est ouvert dans E et donc $\bigcap_{1\leqslant i\leqslant n} O_i$ est ouverte dans E. De là :

$$E\backslash\bigcap_{i=1}^{n} O_i = \bigcup_{i=1}^{n} E\backslash O_i = \bigcup_{i=1}^{n} F_i \quad \text{est fermée.}$$

□

Exemples :
— Un singleton est fermé. Le plus simple pour le montrer est d'utiliser la caractérisation séquentielle des fermés que nous verrons plus tard.
— $\mathbb{N}$ et $\mathbb{Z}$ sont fermés dans $\mathbb{R}$. En effet, leurs complémentaires sont des réunions d'intervalles ouverts de $\mathbb{R}$ donc des ouverts.

Remarque : Il faut bien prendre garde au fait qu'être fermé n'est pas le contraire d'être ouvert. En effet il existe des parties ouvertes et fermées, à commencer par E et $\varnothing$. Dans la topologie discrète, toute partie est ouverte et fermée. D'autres part, il peut aussi y avoir des parties ni ouverte, ni fermée. Par exemple : $[0,1[$.

1.2 Espaces métriques

1.2.1 Distances et normes

1.2.1.1 Distances

Définition 1.2.1. *Soit E un ensemble non vide. On appelle distance sur E toute application $d : E \times E \longrightarrow \mathbb{R}_+$ vérifiant :*
(i) $\forall (x,y) \in E^2, d(x,y) = 0 \Longleftrightarrow x = y$ (axiome de séparation)
(ii) $\forall (x,y) \in E^2, d(y,x) = d(x,y)$ (symétrie)
(iii) $\forall (x,y,z) \in E^3, d(x,z) \leqslant d(x,y) + d(y,z)$ (inégalité triangulaire)
Le couple (E,d) est alors appelé espace métrique.

Proposition 1.2.2. *On a :*
1) $\forall p \geqslant 2, \forall (x_1, \ldots, x_p) \in E^p, d(x_1, x_p) \leqslant \sum_{i=1}^{p-1} d(x_i, x_{i+1})$
2) $\forall (x,y,z) \in E^3, |d(x,y) - d(x,z)| \leqslant d(y,z)$

Démonstration : 1) Par récurrence sur p en utilisant le (iii) de la définition.
2) Soit $\forall (x,y,z) \in E^3 : d(x,y) \leqslant d(x,z) + d(z,y)$, d'où : $d(x,y) - d(x,z) \leqslant d(y,z)$. De même, $d(x,z) - d(x,y) \leqslant d(y,z)$ et on obtient bien l'inégalité cherchée.

□

Exemple 1 : On se place sur $E = \mathbb{R}^n$. Soient $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ deux vecteurs de E. On pose :

$$d_1(x,y) = \sum_{i=1}^{n} |x_i - y_i| \quad d_2(x,y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{\frac{1}{2}} \quad d_\infty(x,y) = \sup_{1 \leqslant i \leqslant n} |x_i - y_i|$$

d_1, d_2 et d_∞ sont des distances sur E.

Démonstration : On vérifie pour d_1, d_2 et d_∞ les axiomes (i), (ii) et (iii). Pour d_1 :

$$\begin{aligned} d_1(x,y) = 0 &\iff \sum_{i=1}^{n} |x_i - y_i| = 0 \iff \forall i \in [\![1,n]\!], |x_i - y_i| = 0 \\ &\iff \forall i \in [\![1,n]\!],\ x_i = y_i \iff x = y. \end{aligned}$$

Puis : $d_1(y,x) = \sum_{i=1}^{n} |y_i - x_i| = \sum_{i=1}^{n} |x_i - y_i| = d_1(x,y)$. Enfin, soit $z = (z_1, \ldots, z_n)$: $\forall i \in [\![1,n]\!], |x_i - y_i| \leqslant |x_i - z_i| + |z_i - y_i|$ d'où $d_1(x,y) \leqslant \sum_{i=1}^{n} |x_i - z_i| + \sum_{i=1}^{n} |z_i - y_i|$. Donc $d_1(x,y) \leqslant d_1(x,z) + d_1(z,y)$.

Pour d_2 : Les axiomes (i) et (ii) se vérifient comme pour d_1. (iii) vient de l'inégalité de Minkowski appliquée à $|| \ ||_2$. Pour d_∞ :

$$\begin{aligned} d_\infty(x,y) = 0 &\iff \sup_{1 \leqslant i \leqslant n} |x_i - y_i| = 0 \iff \forall i \in [\![1,n]\!], |x_i - y_i| = 0 \\ &\iff \forall i \in [\![1,n]\!], x_i = y_i \iff x = y. \end{aligned}$$

De plus : $d_\infty(y,x) = \sup_{1 \leqslant i \leqslant n} |y_i - x_i| = \sup_{1 \leqslant i \leqslant n} |x_i - y_i| = d_\infty(x,y)$.
Enfin : $\forall i \in [\![1,n]\!], |x_i - y_i| \leqslant |x_i - z_i| + |z_i - y_i|$. D'où par passage à la borne supérieure : $d_\infty(x,y) \leqslant d_\infty(x,z) + d_\infty(z,y)$.

□

Exemple 2 : Soit E un ensemble non vide. Pour x, y dans E on pose $d(x,y) = 0$ si $x = y$ et $d(x,y) = 1$ si $x \neq y$. d est alors une distance sur E appelée distance discrète. Celle-ci vérifie l'inégalité suivante :

$$\forall (x,y,z) \in E^3,\ d(x,z) \leqslant \max[d(x,y), d(y,z)]$$

appelée inégalité ultramétrique. On peut ainsi définir ce que l'on appelle une distance ultramétrique qui au lieu de vérifier l'inégalité triangulaire vérifie l'inégalité ultramétrique.

Exemple 3 : Soit $E = B(X,\mathbb{R})$ l'espace vectoriel des applications bornées de X vers $\mathbb{R}$, où X est un ensemble non vide quelconque. Alors pour $(f,g) \in E^2$: $d(f,g) = \sup_{x\in X} |f(x) - g(x)|$ définit une distance sur E.

Démonstration : Soit $(f,g) \in E$.

$$d(f,g) = 0 \iff \sup_{x\in X} |f(x) - g(x)| = 0 \iff \forall x \in X, |f(x) - g(x)| = 0 \iff f = g.$$

On a bien aussi $d(g,f) = d(f,g)$. Enfin si $(f,g,h) \in E^3$, on a : $\forall x \in X, |f(x) - g(x)| \leqslant |f(x) - h(x)| + |h(x) - g(x)|$, ce qui entraine par passage à la borne supérieure :

$$d(f,g) \leqslant d(f,h) + d(h,g).$$

□

Pour $p \geqslant 1$, on se donne $(E_1, d_1), \dots, (E_p, d_p)$ des espaces métriques. On a alors :

Théorème 1.2.3. *L'espace produit $E = E_1 \times \ldots \times E_p$ muni de l'application :*

$$d : \begin{pmatrix} E^2 & \to & \mathbb{R}^+ \\ (x,y) & \mapsto & \max_{1\leqslant i\leqslant p} d_i(x_i, y_i) \end{pmatrix}$$

est un espace métrique appelé espace produit de $(E_1, d_1), \dots, (E_p, d_p)$.

Démonstration : Comme max d'éléments positifs, d est bien positive.

Si $x = y$, on a : $\forall i \in [\![1,p]\!],\ x_i = y_i$. D'où : $\forall i \in [\![1,p]\!], d_i(x_i,y_i) = 0$. Ainsi, $d(x,y)$ est le max d'une famille de nombres nuls donc : $d(x,y) = 0$. Réciproquement, si $d(x,y) = 0$, alors : $\forall i \in [\![1,p]\!],\ d_i(x_i,y_i)$ et $\forall i \in [\![1,p]\!],\ x_i = y_i$. Donc : $x = y$. La symétrie est claire : $d(x,y) = d(y,x)$.

Enfin : $\forall i \in [\![1,p]\!], d_i(x_i,z_i) \leqslant d_i(x_i,y_i) + d_i(y_i,z_i) \leqslant d(x,y) + d(y,z)$. De là, par passage au max on a : $d(x,z) \leqslant d(x,y) + d(y,z)$. D'où l'inégalité triangulaire.

□

1.2.1.2 Normes

Soit $\mathbb{K}$ un corps valué. En pratique $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Soit E un $\mathbb{K}$-espace vectoriel.

Définition 1.2.4. *On appelle norme sur E toute application $N : E \longrightarrow \mathbb{R}_+$ vérifiant :*

(i) $\forall x \in E, N(x) = 0 \iff x = 0$ *(axiome de séparation)*

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

(iii) $\forall (x,y) \in E^2, N(x+y) \leqslant N(x) + N(y)$ *(inégalité triangulaire)*

Le couple (E,N) est alors appelé espace vectoriel normé.

Notation : Nous noterons indifféremment N(x) ou $||x||$ pour désigner la norme du vecteur x.

Proposition 1.2.5. *On a :*
1) $\forall p \geqslant 1, \forall (x_1, \ldots, x_p) \in E^p, N(x_1 + \ldots + x_p) \leqslant \sum_{i=1}^{p} N(x_i)$
2) $\forall x \in E,\ N(-x) = N(x)$
3) $\forall x \in E\backslash\{0\}, N(\frac{x}{N(x)}) = 1$
4) $\forall (x, y) \in E^2, |N(x) - N(y)| \leqslant N(x - y)$

Démonstration : 1) Par récurrence sur p en utilisant le (iii) de la définition.
2) Immédiat par (ii) de la définition.
3) Immédiat par (ii) de la définition.
4) On a : $N(x) = N(x - y + y) \leqslant N(x - y) + N(y)$. D'où : $N(x) - N(y) \leqslant N(x - y)$. De même : $N(y) - N(x) \leqslant N(y - x) = N(x - y)$ par 2). D'où le résultat.

□

Exemple 1 : On se place sur $E = \mathbb{C}^n$. Soit $x = (x_1, \ldots, x_n)$ un vecteur de E. On pose :

$$N_1(x) = \sum_{i=1}^{n} |x_i| \quad N_p(x) = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \quad N_\infty(x) = \sup_{1 \leqslant i \leqslant n} |x_i|$$

N_1, N_2 et N_∞ sont des normes sur E. On les note aussi souvent $|| \ ||_1$, $|| \ ||_p$ et $|| \ ||_\infty$.

Démonstration : On vérifie pour N_1, N_p et N_∞ les axiomes (i), (ii) et (iii).

Pour N_1 :

$$N_1(x) = 0 \Leftrightarrow \sum_{i=1}^{n} |x_i| = 0 \Leftrightarrow \forall i \in [\![1, n]\!], |x_i| = 0 \iff \forall i \in [\![1, n]\!],\ x_i = 0 \iff x = 0.$$

Puis : $N_1(\lambda x) = \sum_{i=1}^{n} |\lambda x_i| = |\lambda| \sum_{i=1}^{n} |x_i| = |\lambda| N_1(x)$. Enfin : $\forall i \in [\![1, n]\!],\ |x_i + y_i| \leqslant |x_i| + |y_i|$. D'où $N_1(x + y) \leqslant \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i|$ et $N_1(x + y) \leqslant N_1(x) + N_1(y)$.
Pour N_p : Les axiomes (i) et (ii) se vérifient comme pour N_1. (iii) vient de l'inégalité de Minkowski.
Pour N_∞ :

$$N_\infty(x, y) = 0 \Leftrightarrow \sup_{1 \leqslant i \leqslant n} |x_i| = 0 \Leftrightarrow \forall i \in [\![1, n]\!], |x_i| = 0 \Leftrightarrow \forall i \in [\![1, n]\!], x_i = 0 \Leftrightarrow x = 0.$$

De plus :

$$N_\infty(\lambda x) = \sup_{1 \leqslant i \leqslant n} |\lambda x_i| = |\lambda| \sup_{1 \leqslant i \leqslant n} |x_i| = |\lambda| N_\infty(x).$$

Enfin : $\forall i \in [\![1, n]\!], |x_i + y_i| \leqslant |x_i| + |y_i|$. Par passage à la borne supérieure : $N_\infty(x + y) \leqslant N_\infty(x) + N_\infty(y)$.

□

Exemple 2 : Soit $E = B(X, \mathbb{C})$ l'espace vectoriel des applications bornées de X vers $\mathbb{C}$, où X est un ensemble non vide quelconque. Alors pour $f \in E$:

$$N(f) = \sup_{x \in X} |f(x)|$$

définit une norme sur E.

Démonstration : Soit $f \in E$,

$$N(f) = 0 \Leftrightarrow \sup_{x \in X} |f(x)| = 0 \Leftrightarrow \forall x \in X, |f(x)| = 0 \Leftrightarrow f = 0.$$

On a bien aussi $N(\lambda f) = |\lambda| N(f)$. Enfin si $(f, g) \in E^2$, on a :

$$\forall x \in X, |f(x) + g(x)| \leqslant |f(x)| + |g(x)|,$$

ce qui entraine, par passage à la borne supérieure, $N(f + g) \leqslant N(f) + N(g)$.

□

Exemple 3 : Soit $E = C^o([a, b], \mathbb{C})$ l'espace vectoriel des applications continues de $[a, b]$ dans $\mathbb{C}$. Alors, pour $f \in E$ et $p \geqslant 1$,

$$||f||_p = \left(\int_a^b |f|^p \right)^{\frac{1}{p}}$$

est une norme sur E.

Démonstration : *(i)* $||f||_p = 0 \Leftrightarrow \int_a^b |f|^p = 0 \Leftrightarrow |f|^p = 0$ car $|f|^p$ est positive et continue.

(ii)$||\lambda f||_p = \left(\int_a^b |\lambda f|^p \right)^{\frac{1}{p}} = \left(|\lambda|^p \int_a^b |f|^p \right)^{\frac{1}{p}} = \lambda ||f||_p$

(iii) C'est l'inégalité de Minkowski pour les fonctions.

□

Proposition 1.2.6. *On reprend les notations de l'exemple 1. Soit $x \in \mathbb{K}^n$ fixé. On a :*

$$\lim_{p \to +\infty} ||x||_p = ||x||_\infty$$

Démonstration : On a :

$$\forall p \geqslant 1, \ ||x||_\infty \leqslant ||x||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}} \leqslant n^{\frac{1}{p}} ||x||_\infty$$

D'où le résultat en faisant tendre p vers $+\infty$.

□

Exemple 4 : Lorsque l'on considère un produit d'espaces vectoriels normés, la norme correspondante à la distance produit est :

$$|| \ x \ || = \max_{1 \leqslant i \leqslant p} || \ x_i \ ||_i$$

1.2.1.3 Boules et sphères

Nous définissons alors les notions de boules et de sphère.

Définition 1.2.7. (E, d) *est un espace métrique.*

On appelle boule ouverte de centre a et de rayon $r > 0$ l'ensemble :

$$B(a, r) = \{x \in E, d(a, x) < r\}$$

On appelle boule fermée de centre a et de rayon $r > 0$ l'ensemble :

$$B_f(a,r) = \{x \in E, d(a,x) \leqslant r\}$$

(avec $B_f(a,0) = \{a\}$).
On appelle sphère de centre a et de rayon $r > 0$ l'ensemble :

$$S(a,r) = \{x \in E, d(a,x) = r\}$$

On donne ici une caractérisation des boules ouvertes dans un espace produit.

Proposition 1.2.8. *Soit $a = (a_1, \ldots, a_p)$. Soit $r > 0$. On a :*

$$B_d(a,r) = \prod_{i \in [1,p]} B_{d_i}(a_i, r)$$

Démonstration : On a :

$$\begin{aligned} (x_1, \ldots, x_n) \in B_d(a,r) &\Leftrightarrow \max[(d_1(x_1,a_1), \ldots, d_p(x_p,a_p)] < r \\ &\Leftrightarrow \forall i \in [\![1,p]\!],\ d_i(x_i,a_i) < r. \end{aligned}$$

□

Définition 1.2.9. *$(E, || \ ||)$ est un espace vectoriel normé.*
On appelle boule ouverte de centre a et de rayon $r > 0$ l'ensemble :

$$B(a,r) = \{x \in E, ||a - x|| < r\}$$

On appelle boule fermée de centre a et de rayon $r > 0$ l'ensemble :

$$B_f(a,r) = \{x \in E, ||a - x|| \leqslant r\}$$

(avec $B_f(a,0) = \{a\}$).
On appelle sphère de centre a et de rayon $r > 0$ l'ensemble :

$$S(a,r) = \{x \in E, ||a - x|| = r\}$$

Proposition 1.2.10. *On a pour $a \in E$ et $r > 0$:*
(i) $B(a,r) = a + B(0,r)$, $B_f(a,r) = a + B_f(0,r)$ *et* $S(a,r) = a + S(0,r)$.
(ii) $B(0,r) = rB(0,1)$, $B_f(0,r) = rB_f(0,1)$ *et* $S(0,r) = rS(0,1)$.

Démonstration : *(i)* $x \in B(a,r) \Leftrightarrow ||x - a|| < r \Leftrightarrow x - a \in B(0,r) \Leftrightarrow x \in a + B(0,r)$.
De même pour les boules fermées et les sphères.
(ii) $x \in B(0,r) \Leftrightarrow ||x|| < r \Leftrightarrow \frac{1}{r}||x|| < 1 \Leftrightarrow ||\frac{1}{r}x|| < 1 \Leftrightarrow \frac{1}{r}x \in B(0,1) \Leftrightarrow x \in rB(0,1)$.
De même pour les boules fermées et les sphères.

□

Proposition 1.2.11. *Soit E un $\mathbb{K}$-espace vectoriel, N_1 et N_2 deux normes sur E. Soient B_1 la boule unité pour N_1 et B_2 la boule unité pour N_2. Alors on a équivalence entre :*
(i) $N_1 = N_2$.
(ii) $B_1 = B_2$.

Démonstration : $(i) \Rightarrow (ii)$: Clair !

$(ii) \Rightarrow (i)$: Pour $x = 0$, on a bien $N_1(0) = N_2(0) = 0$. Soit donc $x \in E \backslash \{0\}$ et $\varepsilon > 0$ Alors, $\frac{x}{N_1(x)+\varepsilon} \in B_1$ et $\frac{x}{N_1(x)+\varepsilon} \in B_2$. Donc :

$$N_2\left(\frac{x}{N_1(x)+\varepsilon}\right) < 1 \Rightarrow N_2(x) < N_1(x) + \varepsilon.$$

On fait tendre ε vers 0 et il vient : $N_2(x) \leqslant N_1(x)$. Par symétrie, $N_1(x) \leqslant N_2(x)$ et $N_1 = N_2$.

□

Le résultat reste vrai avec les boules unités fermées si ce n'est que la démonstration est plus simple, il est inutile d'introduire un ε.

1.2.1.4 Distances et normes équivalentes

On introduit la notion de distances équivalentes.

Définition 1.2.12. *On dit que deux distances d et d' sont équivalentes s'il existe deux réels a et b strictement positifs tels que :*

$$\forall (x, y) \in E^2, ad(x, y) \leqslant d'(x, y) \leqslant bd(x, y).$$

De même, on introduit la notion de normes équivalentes.

Définition 1.2.13. *On dit que deux normes N_1 et N_2 sont équivalentes s'il existe deux réels C_1 et C_2 strictement positifs tels que :*

$$\forall x \in E, C_1 N_1(x) \leqslant N_2(x) \leqslant C_2 N_1(x).$$

1.2.2 Topologie des espaces métriques

1.2.2.1 Ouverts, fermés

On va maintenant donner la définition d'un ouvert et d'un fermé pour les espaces métriques.

Définition 1.2.14. *Soit (E, d) un espace métrique. On dit que la partie Ω de E est ouverte lorsque :*

$$\forall x \in \Omega, \exists r_x > 0, B(x, r_x) \subset \Omega$$

Proposition 1.2.15. *Les ouverts ainsi définis constituent une topologie sur l'ensemble E appelée topologie de l'espace métrique (E, d).*

Démonstration : (i) E et $\varnothing$ sont ouverts.

(ii) Soit $(\Omega_i)_{i \in I}$ une famille d'ouverts de (E, d). On pose $\Omega = \bigcup_{i \in I} \Omega_i$. Si $x \in \Omega$, alors $\forall i \in I, x \in \Omega_i$. Soit $i_0 \in I$, on sait que Ω_{i_0} est ouvert, d'où l'existence de $r > 0$ tel que $B(x, r) \subset \Omega_{i_0} \subset \Omega$. On en déduit que Ω est un ouvert de E.

(iii) Soient $\Omega_1, \ldots, \Omega_n$ des ouverts de E et $x \in \Omega = \bigcap_{i \in [\![1,n]\!]} \Omega_i$.

$$\forall i \in [\![1, n]\!], \exists r_i > 0, B(x, r_i) \subset \Omega_i$$

On pose $R = \min(r_1, \ldots, r_n) > 0$ (fini). Il vient : $\forall i \in [\![1, n]\!], B(x, R) \subset \Omega_i$ et $B(x, R) \subset \Omega$. Donc Ω est ouvert.

□

Définition 1.2.16. *Une partie F de (E, d) est dite fermée si son complémentaire est ouvert.*

Proposition 1.2.17. *Soit (E, d) un espace métrique.*
1) Toute boule ouverte de (E, d) est un ouvert de E.
2) Toute boule fermée de (E, d) est un fermé de E.

Démonstration : 1) Soit $B(a, r)$ $(r > 0)$ une boule ouverte de E. Soient $x \in B(a, r)$ et $r_x = r - d(a, x)$. Si $y \in B(x, r_x)$, il vient : $d(a, y) \leqslant d(a, x) + d(x, y)$. Ainsi : $d(a, y) < d(a, x) + r_x = r$ d'où $y \in B(a, r)$, c'est à dire que : $B(x, r_x) \subset B(a, r)$.
2) Soit $\Omega = E \backslash B_f(a, r)$. Soit $x \in \Omega$. On sait par définition de Ω que $d(a, x) > r$. Soit $r_x = d(a, x) - r$. Si $y \in B(x, r_x), d(a, x) \leqslant d(a, y) + d(y, x)$. D'où : $d(a, x) - r_x < d(a, y)$ et $d(a, y) > r$ c'est à dire $y \in \Omega$. Donc $B(x, r_x) \subset \Omega$ ce qui montre que Ω est ouvert.

□

Corollaire 1.2.18. *Les ouverts de E sont les réunions de boules ouvertes.*

Démonstration : Une boule ouverte est ouverte et une réunion d'ouverts est ouverte donc une réunion de boules ouvertes est ouverte.
Réciproquement, avec les notations de la définition des ouverts, si Ω est un ouvert de E :

$$\Omega = \bigcup_{x \in \Omega} B(x, r_x)$$

et donc tout ouvert de E est réunion de boules ouvertes.

□

1.2.2.2 Voisinages

Définition 1.2.19. *On appelle voisinage de $a \in E$ toute partie V contenant un ouvert contenant a.*

Traduction dans un espace métrique : $V \in V(a) \iff \exists r > 0, B(a, r) \subset V$.

Proposition 1.2.20 (Caractère séparé de la topologie d'un espace métrique)**.** *Si a et b sont deux points distincts de E, il existe des voisinages U et V de a et b respectivements, tels que $U \cap V = \varnothing$*

Démonstration : Soit $r = d(a, b)$. On pose $\rho = \frac{1}{2} d(a, b).$, $U = B(a, \rho) \in V(a)$ et $V = B(b, \rho) \in V(b)$. Si $x \in U \cap V$, il vient : $d(a, b) \leqslant d(a, x) + d(x, b) < \rho + \rho = r$. Contradiction, donc $U \cap V = \varnothing$.

□

1.2.2.3 Intérieur, adhérence, densité

On se donne dans cette partie $(E, \mathcal{T})$ un espace topologique.

1.2.2.3.1 Intérieur

Définition 1.2.21. *Soit A une partie de E. Le point a de A est dit intérieur à A lorsqu'il existe un ouvert ω tel que :*

$$a \in \omega \ \text{ et } \ \omega \subset A.$$

Définition 1.2.22. *L'intérieur de A est l'ensemble des points intérieurs à A. On le note Int(A) ou $\mathring{A}$.*

Proposition 1.2.23. *Soit A une partie de E.*
(i) L'intérieur de A est un ouvert.
(ii) $\overset{\circ}{A}$ est le plus grand ouvert contenu dans A.

Démonstration : (i) Montrons que $\overset{\circ}{A}$ est un voisinage de chacun de ses points. Soit $a \in \overset{\circ}{A}$. Il existe par définition ω ouvert de E tel que $a \in \omega$ *et* $\omega \subset A$. Si $b \in \omega$, b est aussi un point intérieur. Ainsi, $\omega \subset \overset{\circ}{A}$ *et* $\overset{\circ}{A} \in V(a)$.
(ii) Si Ω est un ouvert contenu dans A alors pour tout $a \in \Omega$, l'inclusion $\Omega \subset A$ amène $a \in \overset{\circ}{A}$, et donc $\Omega \subset A$.

□

Corollaire 1.2.24. *$\overset{\circ}{A}$ est la réunion de tous les ouverts contenus dans A.*

Démonstration : La réunion de tous les ouverts contenus dans A est un ouvert contenu dans A et c'est le plus grand au sens de l'inclusion.

□

Corollaire 1.2.25. *A est ouvert dans E si et seulement si $A = \overset{\circ}{A}$.*

Démonstration : On suppose que $A = \overset{\circ}{A}$. Comme $\overset{\circ}{A}$ est un ouvert de E, A est ouvert dans E. Réciproquement, si A ouvert dans E, A est le plus grand ouvert contenu dans A donc $\overset{\circ}{A} = A$.

□

Proposition 1.2.26. *Soit A et B deux parties de E.*
(i) Si $A \subset B$, alors $\overset{\circ}{A} \subset \overset{\circ}{B}$.
(ii) $\overset{\circ}{\overline{A \cap B}} = \overset{\circ}{A} \cap \overset{\circ}{B}$.
(iii) $(\overset{\circ}{A} \cup \overset{\circ}{B}) \subset \overset{\circ}{\overline{A \cup B}}$.
(iv) $\overset{\circ}{\overset{\circ}{A}} = \overset{\circ}{A}$

Démonstration : (i) On suppose $A \subset B$ et soit $x \in \overset{\circ}{A}$. Il existe alors par définition un ouvert ω tel que $x \in \omega$ et $\omega \subset A$. Comme $A \subset B$, $\omega \subset B$ et $x \in \overset{\circ}{B}$, ce qui montre que $\overset{\circ}{A} \subset \overset{\circ}{B}$.
(ii) On a : $(A \cap B) \subset A$ d'où $\overset{\circ}{\overline{A \cap B}} \subset \overset{\circ}{A}$. De même : $(A \cap B) \subset B$ d'où $\overset{\circ}{\overline{A \cap B}} \subset \overset{\circ}{B}$. Donc : $\overset{\circ}{\overline{A \cap B}} \subset (\overset{\circ}{A} \cap \overset{\circ}{B})$. $\overset{\circ}{A} \cap \overset{\circ}{B}$ est un ouvert contenu dans $A \cap B$, donc $(\overset{\circ}{A} \cap \overset{\circ}{B}) \subset \overset{\circ}{\overline{A \cap B}}$. Ainsi $\overset{\circ}{\overline{A \cap B}} = \overset{\circ}{A} \cap \overset{\circ}{B}$.
(iii) $A \subset (A \cup B)$ d'où $\overset{\circ}{A} \subset \overset{\circ}{\overline{A \cup B}}$. De même : $B \subset (A \cup B)$ d'où $\overset{\circ}{B} \subset \overset{\circ}{\overline{A \cup B}}$. Donc : $(\overset{\circ}{A} \cup \overset{\circ}{B}) \subset \overset{\circ}{\overline{A \cup B}}$
(iv) $\overset{\circ}{A}$ est ouvert donc par le corollaire précédent $\overset{\circ}{\overset{\circ}{A}} = \overset{\circ}{A}$.

□

Exemple : Si $A = \mathbb{Q}$ et $B = \mathbb{R} \backslash \mathbb{Q}$, alors la réunion de A et B vaut $\mathbb{R}$ d'intérieur égal à $\mathbb{R}$. Or A et B sont d'intérieurs vides donc la réunion de leurs intérieurs est vide.

1.2.2.3.2 Adhérence On commence par donner la définition d'un point adhérent à une partie A de E.

Définition 1.2.27. *On dit que le point x de E est adhérent à A lorsque :*

$$\forall V \in V(x), V \cap A \neq \emptyset$$

Remarque : Il convient de toujours garder à l'esprit qu'un point adhérent à A n'est pas forcément dans A.

Définition 1.2.28. *L'ensemble des points adhérents à A est appellé l'adhérence de A. On le note $\bar{A}$.*

Remarque 1 : On a bien sûr $A \subset \bar{A}$.
Remarque 2 : On parle encore parfois de la fermeture de A pour l'adhérence de A.
On donne alors une première caractérisation de $\bar{A}$.

Proposition 1.2.29. *$\bar{A}$ est le plus petit fermé de E contenant A.*

Démonstration : On montre tout d'abord que $\bar{A}$ est fermé en montrant que son complémentaire est ouvert. Soit x dans $E\backslash\bar{A}$. Alors il existe un voisinage V de x tel que : $V \cap A = \emptyset$. Cela signifie que $V \subset E\backslash A$. On prend ω un ouvert contenu dans V et contenant x. Alors, si $y \in \omega$, ω est un voisinage de y tel que $\omega \cap A = \emptyset$ et donc y est dans $E\backslash\bar{A}$. De là, $\omega \subset E\backslash\bar{A}$ et donc $E\backslash\bar{A}$ est ouvert. Donc $\bar{A}$ est fermé.

Soit F un fermé de E contenant A. Alors, pour tout x n'appartenant pas à F, $\omega = E\backslash F$ est un voisinage de x tel que $\omega \cap F = \emptyset$. A fortiori : $\omega \cap A = \emptyset$ car $A \subset F$. Ainsi, x n'appartient pas à $\bar{A}$. Donc si x n'appartient pas à F, x n'appartient pas à $\bar{A}$. Donc $\bar{A} \subset F$ et $\bar{A}$ est bien le plus petit fermé contenant A.

□

Corollaire 1.2.30. *$\bar{A}$ est l'intersection de tous les fermés contenant A.*

Démonstration : En effet, une telle intersection est fermée et contient A. De plus c'est le plus petit au sens de l'inclusion.

□

Corollaire 1.2.31. *A est fermée dans E si et seulement si $A = \bar{A}$.*

Démonstration : On suppose que $A = \bar{A}$. Comme $\bar{A}$ est fermée, A est fermée.
Réciproquement, si A est fermée, A est un fermé contenant A et c'est le plus petit. Donc $A = \bar{A}$.

□

On a alors des propriétés de réunion et d'intersection analogues à celles pour l'intérieur.

Proposition 1.2.32. *Soient A et B deux parties de E.*
(i) Si $A \subset B$ alors $\bar{A} \subset \bar{B}$.
(ii) $\overline{A \cup B} = \bar{A} \cup \bar{B}$.
(iii) $\overline{A \cap B} \subset \bar{A} \cap \bar{A}$.
(iv) $\bar{\bar{A}} = \bar{A}$.

Démonstration : *(i)* Soit $x \in \bar{A}$. Tout voisinage de x rencontre B donc $x \in \bar{B}$. Ou bien comme autre preuve : $\bar{B}$ est un fermé contenant A donc comme $\bar{A}$ est le plus petit fermé contenant A, $\bar{A} \subset \bar{B}$.

(ii) On a : $A \subset (A \cup B) \subset (\overline{A \cup B})$. Donc $\overline{A \cup B}$ est un fermé contenant A donc $\bar{A} \subset \overline{A \cup B}$ (on aurait pu appliquer directement *(i)* à $A \subset (A \cup B)$ mais autant voir une deuxième fois l'idée de la preuve). De même, $\bar{B} \subset \overline{A \cup B}$. Donc $(\bar{A} \cup \bar{B}) \subset (\overline{A \cup B})$. En sens inverse, $\bar{A} \cup \bar{B}$ est un fermé contenant $A \cup B$ donc $\overline{A \cup B} \subset (\bar{A} \cup \bar{B})$.

(iii) On a : $(A \cap B) \subset A$. D'où par *(i)*, $\overline{A \cap B} \subset \bar{A}$. De même, $\overline{A \cap B} \subset \bar{A}$. Donc : $\overline{A \cap B} \subset (\bar{A} \cap \bar{B})$.

(iv) $\bar{A}$ est fermée donc par le corollaire précédent : $\bar{\bar{A}} = \bar{A}$.

□

On peut maintenant donner un résultat reliant l'intérieur et l'adhérence.

Proposition 1.2.33. *On a :* $\overline{E\backslash A} = E\backslash \overset{\circ}{A}$ *et* $E\backslash\bar{A} = \mathrm{Int}(E\backslash A)$.

Démonstration : Soit $x \in E$. On a :

$$\begin{aligned} x \in \overline{E\backslash A} &\iff \forall V \in V(x), V \cap (E\backslash A) \neq \varnothing \iff \forall V \in V(x), V \not\subset A \\ &\iff \neg(\exists V \in V(x), V \subset A) \iff \neg(x \in \overset{\circ}{A}) \iff x \in E\backslash\overset{\circ}{A}. \end{aligned}$$

D'où la première égalité. La deuxième égalité s'obtient en appliquant la première à $E\backslash A$.

□

1.2.2.3.3 Frontière

Définition 1.2.34. *On appelle frontière de A l'ensemble* $\bar{A} \cap \overline{E\backslash A}$. *On le note* $Fr(A)$ *ou* ∂A.

Remarque : Comme intersection de deux fermés, la frontière de A est un fermé de E.

Proposition 1.2.35. *On a :* $\partial A = \bar{A}\backslash\overset{\circ}{A}$.

Démonstration : On a : $E\backslash(\overline{E\backslash A}) = \mathrm{Int}(E\backslash(E\backslash A)) = \mathrm{Int}(A)$. Donc si un point de E est dans l'adhérence du complémentaire de A, il n'est pas dans l'intérieur de A. Donc si un point de E est dans ∂A, il est dans $\bar{A}$ et n'est pas dans $\mathrm{Int}(A)$.

□

On peut aussi introduire la notion d'extérieur de A qui est l'intérieur du complémentaire de A ou encore le complémentaire de $\bar{A}$. On le note habituellement $\mathrm{Ext}(A)$. On a alors que $\overset{\circ}{A}$, ∂A et $\mathrm{Ext}(A)$ forment une partition de E (ce qui évident au vu de la proposition ci-dessus). On peut alors donner quelques exemples.

Exemples : On munit $\mathbb{R}$ de sa topologie naturelle. Alors, on a :

— $\mathbb{Q}$ est d'intérieur vide et d'adhérence égale à $\mathbb{R}$ donc $\partial\mathbb{Q} = \mathbb{R}$ et $\mathrm{Ext}(\mathbb{Q}) = \varnothing$.

— On prend $A = [0,1]$. Alors, $\overset{\circ}{A} =]0,1[$, $\partial A = \{0,1\}$ et $\mathrm{Ext}(A) =]-\infty, 0[\cup]1, \infty[$.

1.2.2.3.4 Densité

Définition 1.2.36. *Soient A et B deux parties de E. On dit que A est dense dans B lorsque $\bar{A} \supset B$.*

Définition 1.2.37. *Une partie A de E qui est dense dans E est dite partout dense.*

Proposition 1.2.38. *Une partie A de E est partout dense si et seulement si elle rencontre tout ouvert non vide ω de E. C'est à dire :*

$$(\forall \omega \in \mathcal{T}, \omega \neq \emptyset) \Longrightarrow (\omega \cap A \neq \emptyset)$$

Démonstration : Si $\bar{A} = E$ et si ω est un ouvert non vide de E, on choisit $x \in \omega$; ω est alors un voisinage de x d'où $x \in \bar{A}$, et $\omega \cap A \neq \emptyset$.

En sens inverse : si $x \in E$ et si $V \in V(x)$ il existe (par définition de V) un ouvert ω tel que $x \in \omega$ et que$\omega \subset V$. Par hypothèse, $\omega \cap A \neq \emptyset$ d'où $V \cap A \neq \emptyset$, et $x \in \bar{A}$.

□

Donnons maintenant deux exemples :

Exemple 1 : $\mathbb{Q}$ et $\mathbb{R}\backslash\mathbb{Q}$ sont denses dans $\mathbb{R}$. En effet, si $x \in \mathbb{R}$ et si $V \in V(x)$ il existe $\alpha > 0$ tel que $]x-\alpha; x+\alpha[\subset V$. Alors $]x-\alpha; x+\alpha[\cap\mathbb{Q} \neq \emptyset$ et $]x-\alpha; x+\alpha[\cap(\mathbb{R}\backslash\mathbb{Q}) \neq \emptyset$.

Exemple 2 : L'ensemble des rationnels dyadiques est dense dans $\mathbb{R}$. En effet : On rappelle qu'un rationnel dyadique s'écrit sous la forme $r = \frac{p}{2^q}$ avec $(p, q) \in \mathbb{Z}^2$.

Soient a et b deux réels avec $a < b$. Comme $\mathbb{R}$ est archimédien, il existe un entier q tel que : $\frac{1}{2^q} < (b - a)$. Posons $p = E(2^q a) + 1$. On a :

$$a2^q < p < a2^q + 1 \text{ et } a < r < a + \frac{1}{2^q} < b.$$

On a donc trouvé entre deux réels quelconques un rationnel dyadique, donc l'ensemble de ces rationnels rencontre tout ouvert non vide de $\mathbb{R}$. Donc, l'ensemble des rationnels dyadiques est dense dans $\mathbb{R}$.

1.2.2.4 Points d'accumulation

Soit A une partie de (E, d).

Définition 1.2.39. *On dit que que le point a de E est un point d'accumulation de A si tout voisinage de a rencontre A en au moins un point différent de A, soit :*

$$\forall V \in V(a), V \cap (A\backslash\{a\}) \neq \emptyset.$$

On note A^{acc} l'ensemble des points d'accumulation de A.

Proposition 1.2.40. *On a : $A^{acc} \subset \bar{A}$ et plus précisement, $\bar{A} = A \cup A^{acc}$.*

Démonstration : Soit $a \in A^{acc}$. Alors $\forall V \in V(a), V \cap (A\backslash\{a\}) \neq \emptyset$. En particulier $\forall V \in V(a), V \cap A \neq \emptyset$ et par définition, $a \in \bar{A}$.

Comme $A^{acc} \subset \bar{A}$ et $A \subset \bar{A}$, on a : $(A \cup A^{acc}) \subset \bar{A}$. Réciproquement, soit $a \in \bar{A}$, a n'appartenant pas à A. Si V est un voisinage de a, il rencontre A et forcément en un point différent de a car a n'appartient pas à A. Donc $a \in A^{acc}$. D'où : $\bar{A} \subset (A \cup A^{acc})$.

□

Corollaire 1.2.41. *A est fermée si et seulement si elle contient tous ses points d'accumulation.*

Démonstration : A est fermée si et seulement si $A = \bar{A}$. Or, $A^{acc} \subset \bar{A}$, donc si A est fermée, elle contient tous ses points d'accumulation.
Réciproquement, si A contient tous ses points d'accumulation, $A \cup A^{acc} = A$ et $A = \bar{A}$.

□

Proposition 1.2.42. *L'ensemble des points d'accumulation de A est un fermé de E.*

Démonstration : Soit (x_n) une suite d'éléments de A^{acc} qui converge vers x sans stationner à x. Soit V un voisinage de x. Comme (x_n) converge vers x sans stationner à x : $\exists n \in \mathbb{N}, x_n \in V \backslash \{x\}$. Or, si on considère un voisinage W de x_n inclus dans $V \backslash \{x\}$, alors $W \backslash \{x_n\}$ intersecte A car x_n est un point d'accumulation de A. Donc $V \backslash \{x\} \cap A \neq \varnothing$ et $x \in A^{acc}$.

□

On utilise ici la caractérisation séquentielle de la fermeture que nous allons voir un peu plus loin dans ce chapitre.

Proposition 1.2.43. *Si a est un point d'accumulation de A, pour tout voisinage V de A, $V \cap A$ est infini.*

Démonstration : Soit $V \in V(a)$. Supposons que $V \cap A$ soit fini. Quitte à réduire V, on pose $V = B(a, r)$. Soit $x_0 \in V \cap A$ tel que

$$d(a, x_0) = \min_{t \in (V \backslash \{a\}) \cap A} d(a, t).$$

Posons $d = \frac{d(a, x_0)}{2}$, alors $V' = B(a, d) \in V(a)$ d'où $(V' \backslash \{a\}) \cap A \neq \varnothing$. Soit $x \in (V' \backslash \{a\}) \cap A$, $x \in B(a, d) \Rightarrow d(a, x) < d$. Mais $V' \subset V$. D'où $x \in V$. Absurde ! Donc $V \cap A$ est infini.

□

1.2.2.5 Points isolés

Définition 1.2.44. *On dit que $a \in A$ est un point isolé de A s'il existe un voisinage a qui ne rencontre A qu'en a, soit :*

$$\exists V \in V(a), V \cap A = \{a\}$$

On note A^i l'ensemble des points isolés de A.

Définition 1.2.45. *Une partie A de E est dite discrète si tous ses points sont isolés.*

Proposition 1.2.46. *On a : $\bar{A} = A^i \cup A^{acc}$.*

Démonstration : Comme $A \subset \bar{A}$, on se ramène à montrer que $\bar{A} \subset (A^i \cup A^{acc})$. Soit $a \in \bar{A}$. Alors tout voisinage de a rencontre A.
On suppose que a n'est pas un point d'accumulation de A. Alors : $\exists V \in V(a), V \cap (A \backslash \{a\}) = \varnothing$. Or $V \cap A \neq \varnothing$. Donc : $V \cap A = \{a\}$ et $a \in A^i$. Donc $a \in (A^i \cup A^{acc})$.
Réciproquement, $A^{acc} \subset \bar{A}$ et $A^i \subset A$, donc $(A^i \cup A^{acc}) \subset \bar{A}$.

□

On en déduit qu'un fermé se partage en l'ensemble de ses points d'accumulation et l'ensemble de ses points isolés. D'autre part, il faut préciser que cette réunion est disjointe.

Exemple 1 : $\mathbb{Z}$ est discret, et plus généralement, les sous-groupes additifs $a\mathbb{Z}$ sont discrets. En effet si a est un point de $\mathbb{Z}$ et que l'on considère comme voisinage de a, $V = [a - \frac{1}{2}, a + \frac{1}{2}]$, alors $V \cap \mathbb{Z} = \{a\}$.
$\mathbb{Z}$ n'a pas de point d'accumulation. En effet, soit $a \in \mathbb{R}$. Si $a \in \mathbb{Z}$ on reprend V comme ci-dessus et on a : $V \cap (\mathbb{Z}\backslash\{a\}) = \emptyset$. Si a n'est pas dans $\mathbb{Z}$ on considère $b = \inf_{z\in\mathbb{Z}} |a - z|$. Soit alors $V = [a - \frac{b}{2}, a + \frac{b}{2}]$. On a alors : $V \cap (\mathbb{Z}\backslash\{a\}) = \emptyset$. Donc, dans les deux cas, a n'est pas un point d'accumulation. Donc $\mathbb{Z}$ n'a pas de point d'accumulation.

Exemple 2 : Soit $A = \{\frac{1}{n}, n \geqslant 1\}$. Alors A est discrète. En effet, soit $a = \frac{1}{n}$ dans A et $V = [a - \frac{1}{2^n}, a + \frac{1}{2^n}]$. Alors, $V \cap A = \{a\}$ et donc a est un point isolé.
Le seul point d'accumulation de A est 0. Cela provient d'une caractérisation séquentielle des points d'accumulation valable dans les espaces métriques appliquée au fait que $(\frac{1}{n})_{n\in\mathbb{N}}$ tend vers 0.
On remarque en particulier qu'un point d'accumulation peut ne pas appartenir à la partie A, c'est pour cela que dans la définition on prend à priori a, point d'accumulation, dans E.

Exemple 3 : Soit $A = [a, b]$, a et b étant deux réels avec $a < b$. Alors : $A^{acc} = [a, b]$. En effet, si $x \in [a, b]$, tout voisinage de x coupe $A\backslash\{x\}$ et x est dans A^{acc}. Réciproquement, comme A est fermée, elle contient tous ses points d'accumulation et : $A = A^{acc}$. De là, comme $A = \bar{A} = A^i \cup A^{acc}$ et que cette réunion est disjointe, $A^i = \emptyset$.

1.2.2.6 Diamètre, parties bornées

Soit (E, d) un espace métrique.

Définition 1.2.47. *Soit A une partie non vide de E. Le diamètre de A est l'élément de* $[0, +\infty]$ *défini par :*

$$\delta(A) = \sup_{(x,y)\in A\times A} d(x, y) \quad \text{et} \quad \delta(\emptyset) = 0.$$

Exemples : Dans $(\mathbb{R}, |\ |)$: $\delta(]0, 1[) = 1$ et $\delta(]0, +\infty[) = +\infty$.
Dans $(\mathbb{R}^2, ||\ ||_2)$: le diamètre d'un carré de côté 1 est $\sqrt{2}$. Le diamètre d'un triange équilatéral de côté 1 est 1. Le diamètre d'un cercle est son diamètre usuel (le double de son rayon. . .).

Proposition 1.2.48. *Soient A et B deux parties non vides de E. Alors :*

$$\delta(A \cup B) \leqslant \delta(A) + \delta(B) + d(A, B).$$

Démonstration : Soit $(x, y) \in (A \cup B)^2$. Soient $z \in A$ et $t \in B$. Alors : $d(x, y) \leqslant d(x, z) + d(z, t) + d(t, y)$. On obtient alors l'inégalité recherchée par passage au sup pour $(x, y, z, t) \in (A \cup B)^4$.

□

Définition 1.2.49. *On dit que la partie A de E est bornée lorsque* $\delta(A) < +\infty$.

Théorème 1.2.50. *Soit A une partie de E,* $E \neq \emptyset$. *Il y a équivalence entre :*
(i) A est bornée.
(ii) Il existe une boule de E contenant A.
(iii) Pour tout $a \in E$, *il existe* $R > 0$ *tel que :* $A \subset B(a, R)$.

Démonstration : $(iii) \Rightarrow (ii)$: E est non vide.

$(ii) \Rightarrow (i)$: Si A est incluse dans $B(a,R)$ et si x et y sont dans A, alors :

$$d(x,y) \leqslant d(x,a) + d(a,y) \leqslant 2R.$$

Par passage au sup il vient : $\delta(A) \leqslant 2R$ et A est bornée.

$(i) \Rightarrow (iii)$: Soit $a \in E$. On fixe $x \in A$ (si A est vide c'est fini. . .). Soit alors $y \in A$. On a :

$$d(a,y) \leqslant d(a,x) + d(x,y) \leqslant d(a,x) + \delta(A).$$

Posons $R = d(a,x) + \delta(A)$. Alors, $y \in B(a,R)$ et $A \subset B(a,R)$.

□

Remarque : On retrouve la caractérisation usuelle que l'on avait pour les parties bornées de $\mathbb{R}$ à savoir, A est bornée si et seulement si il existe un réel M tel que : $\forall x \in A, |x| \leqslant M$.

1.2.2.7 Distance induite

On se donne (E,d) un espace métrique et A une partie de E.

Définition 1.2.51. *La distance induite par d sur A est la restriction δ de d à $A \times A$.*

Lemme 1.2.52. *On a : $B_\delta(a,r) = B_d(a,r) \cap A$.*

Démonstration : Si $a \in A$ alors : $B_\delta(a,r) = \{x \in A | \delta(a,x) < r\} = \{x \in A | d(a,x) < r\}$.

□

1.2.3 Suites dans un espace métrique

Soit (E,d) un espace métrique.

1.2.3.1 Convergence d'une suite

Définition 1.2.53. *Soit $(u_n)_{n\in\mathbb{N}}$ une suite d'éléments de E. On dit que (u_n) converge vers a lorsque :*

$$(\forall \varepsilon > 0)(\exists n_\varepsilon \in \mathbb{N})(\forall n \geqslant n_\varepsilon)(d(u_n,a) \leqslant \varepsilon)$$

Théorème 1.2.54 (Unicité de la limite)**.** *Si (u_n) converge vers a et vers b, alors $a = b$.*

Démonstration : Si $a \neq b$, $\exists r > 0,\ B(a,r) \cap B(b,r) = \varnothing$ car dans un espace métrique, la topologie induite par la distance est séparée. Or, comme (u_n) converge vers a et b on a :

$$\exists n_1 \in \mathbb{N},\ \forall n \geqslant n_1,\ d(u_n,a) < r$$

$$\exists n_2 \in \mathbb{N},\ \forall n \geqslant n_2,\ d(u_n,b) < r$$

Pour $n \geqslant \max(n_1,n_2)$, $u_n \in B(a,r) \cap B(b,r)$ ce qui est exclu.

□

On se donne $(E_1,d_1),\ldots,(E_p,d_p)$ des espaces métriques.

Théorème 1.2.55. *Soit $(x_n) \in E^{\mathbb{N}}$. Alors, (x_n) converge si et seulement si pour tout i dans $[\![1,p]\!]$, $(x_n^{(i)})$ converge. Dans ce cas :*

$$l = \lim_{n\to+\infty} x_n \Leftrightarrow \forall i \in [\![1,p]\!],\ l_i = \lim_{n\to+\infty} x_n^{(i)}.$$

Démonstration : Dans le sens direct : si (x_n) converge vers l, il vient :

$$\exists \lim_{n\to+\infty} \max[d_1(x_n^{(1)}, l_n^{(1)}), \ldots, d_p(x_n^{(p)}, l_n^{(p)})] = 0.$$

A fortiori, chaque suite $(x_n^{(i)})$ converge vers $l_n^{(i)}$ dans (E_i, d_i). La réciproque se montre exactement de la même façon.

□

1.2.3.2 Valeurs d'adhérence

Théorème 1.2.56. *Soit $(u_n) \in E^{\mathbb{N}}$. Soit $a \in E$. Les trois propositions suivantes sont équivalentes :*
(i) $(\forall \varepsilon > 0)(\forall n_\varepsilon \in \mathbb{N})(\exists n \geqslant n_\varepsilon)(d(a,u_n) \leqslant \varepsilon)$
(ii) $(\forall \varepsilon > 0)(\mathcal{A}_\varepsilon = \{n \in \mathbb{N} | d(a,u_n) \leqslant \varepsilon\}$ *est infini).*
(iii) Il existe une suite extraite de (u_n) qui converge vers a.

Définition 1.2.57. *Lorsqu'un point a de E vérifie l'une des trois assertions équivalentes ci-dessus, on dit que a est une valeur d'adhérence de la suite (u_n).*

Démonstration : $(i) \iff (ii)$: Si $\mathcal{A}_\varepsilon$ est une partie de $\mathbb{N}$, dire que $\mathcal{A}_\varepsilon$ est infini équivaut à dire que $\mathcal{A}_\varepsilon$ n'est pas majorée.
$(iii) \Rightarrow (i)$: Soit $\varphi : \mathbb{N} \to \mathbb{N}$, strictement croissante et telle que $(u_{\varphi(n)})$ converge vers a. Soit $\varepsilon > 0$, il existe $n_\varepsilon \in \mathbb{N}, \forall n \geqslant n_\varepsilon, d(u_{\varphi(n)}, a) \leqslant \varepsilon$. Si $N \in \mathbb{N}$ on choisit m tel que $m \geqslant n_\varepsilon$ et $\varphi(m) \geqslant N$. Pour $n = \varphi(m)\ :\ d(u_n, a) \leqslant \varepsilon$.
$(ii) \Rightarrow (iii)$: On construit $(u_{\varphi(n)})$ par récurrence. Pour $\varepsilon = 1 : d(u_{\varphi(0)}, a) \leqslant 1$, pour $\varepsilon = \frac{1}{2}$, $N = \varphi(0)+1$, $\varphi(1) > N$ et $d(u_{\varphi(1)}, a) < \frac{1}{2}, \ldots$, pour $\varepsilon = \frac{1}{n+1} : N = \max(\varphi(0), \ldots, \varphi(n-1)) + 1$, $\varphi(n) \geqslant N$ tel que $d(a, u_{\varphi(n)}) \leqslant \frac{1}{n+1}$.
On a alors $d(a, u_{\varphi(n)}) \longrightarrow 0$, c'est à dire que $(u_{\varphi(n)})$ converge vers a.

□

Proposition 1.2.58. *Soit (u_n) une suite d'éléments de E. L'ensemble de ses valeurs d'adhérence est :*

$$X = \bigcap_{m\in\mathbb{N}} \overline{\{u_n | n \geqslant m\}}.$$

Démonstration : On note : $U_m = \{u_n | n \geqslant m\}$.

$$\begin{aligned} a \in \bigcap_{m\in\mathbb{N}} \overline{U_m} &\iff \forall m \in \mathbb{N}, a \in \overline{U_m} \\ &\iff \forall m \in \mathbb{N}, \forall \varepsilon > 0, B(a,\varepsilon) \cap U_m \neq \varnothing \\ &\iff \forall m \in \mathbb{N}, \forall \varepsilon > 0, \exists n \geqslant m, d(a,u_n) < \varepsilon \\ &\iff a \text{ est une valeur d'adhérence} \end{aligned}$$

□

Corollaire 1.2.59. *Soit (u_n) une suite d'éléments de E. Alors l'ensemble de ses valeurs d'adhérence est fermé dans E.*

Démonstration : $\forall m \in \mathbb{N}, \overline{U_m}$ est fermé et une intersection de fermés est fermée.

□

Proposition 1.2.60. *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de points de E et soit $A = \{u_n | n \in \mathbb{N}\}$. Alors, tout point d'accumulation de A est valeur d'adhérence de la suite (u_n).*

Démonstration : Soit a un point d'accumulation de A. Alors tout voisinage V de a rencontre A en une infinité de points. En particulier, si V est un voisinage de a, $\{n \in \mathbb{N} | u_n \in V\}$ est infini. Donc, pour tout $N \in \mathbb{N}$ il existe $n \geqslant N$ tel que $u_n \in V$. Donc a est une valeur d'adhérence de (u_n).

□

1.2.3.3 Suites de Cauchy

On se donne (E, d) un espace métrique.

Définition 1.2.61. *On dit que la suite $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ est de Cauchy lorsque :*

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N},\ \forall m \geqslant N_\varepsilon,\ \forall n \geqslant N_\varepsilon,\ d(x_n, x_m) \leqslant \varepsilon.$$

Comme on peut le constater, le fait d'être de Cauchy pour une suite est une notion qui ne peut pas s'énoncer en terme de voisinages. Ainsi ce n'est pas une notion topologique mais une notion métrique. On peut donner une autre écriture parfois plus utile du fait d'être de Cauchy.

Définition 1.2.62. *On dit que la suite $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$ est de Cauchy lorsque :*

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N},\ \forall n \geqslant N_\varepsilon,\ \forall p \geqslant 1,\ d(x_{n+p}, x_n) \leqslant \varepsilon.$$

En particulier, si (x_n) est de Cauchy, $(x_{n+1} - x_n)$ tend vers 0. Mais la réciproque n'est pas vraie en général. On peut alors donner quelques propriétés des suites de Cauchy.

Proposition 1.2.63. *Une suite de Cauchy est bornée.*

Démonstration : On prend $\varepsilon = 1$. Il existe $N \in \mathbb{N}$ tel que : $\forall m, n \geqslant N,\ d(x_m, x_n) \leqslant 1$. Alors, en particulier, on a : $\forall n \geqslant N,\ d(x_n, x_N) \leqslant 1$. De là, soit $R = \max(1, d(x_o, x_N), \ldots, d(x_{N-1}, x_N))$. On a alors : $\forall n \in \mathbb{N},\ x_n \in B_f(x_N, R)$.

□

On va maintenant donner deux résultats sur le lien entre être de Cauchy et être convergente.

Proposition 1.2.64. *Une suite convergente est de Cauchy.*

Démonstration : Soit l la limite de notre suite convergente (x_n). Soit $\varepsilon > 0$. Alors, il existe $N \in \mathbb{N}$ tel que : $\forall n \geqslant \mathbb{N},\ d(x_n, l) \leqslant \varepsilon$. Soient alors n et m des entiers supérieurs à N. On a :

$$d(x_n, x_m) \leqslant d(x_n, l) + d(x_m, l) \leqslant 2\varepsilon$$

Donc (x_n) est de Cauchy.

□

Attention, la réciproque est fausse en général. Par exemple, si on prend comme espace métrique $\mathbb{Q}$ muni de la distance usuelle induite par la valeur absolue, on peut considérer la suite d'éléments de $\mathbb{Q}$: $x_n = \frac{1}{0!} + \ldots + \frac{1}{n!}$. Alors, on a : pour $m > n$:

$$\begin{aligned} 0 < x_m - x_n &\leqslant \frac{1}{(n+1)!} + \ldots + \frac{1}{m!} \leqslant \frac{1}{n!}\left(\frac{1}{(n+1)} + \frac{1}{(n+1)^2} + \ldots\right) \\ &= \frac{1}{n!}\left(\frac{1}{n+1}\left(\frac{1}{1-\frac{1}{n+1}}\right)\right) \leqslant \frac{1}{n\,n!} \end{aligned}$$

Donc la suite (x_n) est de Cauchy, mais il est connu qu'elle converge vers e qui est irrationnel. Donc ce n'est pas une suite de $\mathbb{Q}$ convergente.

Par contre, comme réciproque, on a :

Proposition 1.2.65. *Soit (x_n) une suite de Cauchy. Si (x_n) possède une valeur d'adhérence a, (x_n) converge vers a.*

Démonstration : Soit $\varepsilon > 0$. Comme (x_n) est de Cauchy, on a :

$$\exists N_\varepsilon \in \mathbb{N},\ \forall m \geqslant N_\varepsilon,\ \forall n \geqslant N_\varepsilon,\ d(x_m, x_n) \leqslant \varepsilon$$

Comme a est valeur d'adhérence de (x_n), on a :

$$\forall N \in \mathbb{N},\ \exists m \geqslant N,\ d(x_m, a) \leqslant \varepsilon$$

De là, avec $N = N_\varepsilon$, il vient : $\exists m \geqslant N_\varepsilon,\ d(x_m, a) \leqslant \varepsilon$. Alors, pour tout $n \geqslant N_\varepsilon$, on a : $d(x_m, x_n) \leqslant \varepsilon$. D'où :

$$\forall n \geqslant N_\varepsilon,\ d(x_n, a) \leqslant d(x_n, x_m) + d(x_m, a) \leqslant 2\varepsilon.$$

□

1.2.3.4 Caractérisations séquentielles

Théorème 1.2.66 (Caractérisation séquentielle de l'adhérence). *Soient A une partie de E et a un point de E. On a équivalence entre :*

(i) $a \in \bar{A}$.

(ii) Il existe une suite de points de A qui converge vers a.

Démonstration : $(i) \Rightarrow (ii)$ On a :

$$\begin{aligned} a \in \bar{A} \Rightarrow &\quad \forall U \in V(a), U \cap A \neq \varnothing \Rightarrow \quad \forall \varepsilon > 0, B(a,\varepsilon) \cap A \neq \varnothing \\ \Rightarrow &\quad \forall n \in \mathbb{N}^*, B(a, \frac{1}{n}) \cap A \neq \varnothing \Rightarrow \quad \forall n \in \mathbb{N}^*, \exists u_n \in A, d(a, u_n) < \frac{1}{n}. \end{aligned}$$

Alors $(u_n) \in A^{\mathbb{N}}$ et (u_n) converge vers a.

$(i) \Rightarrow (ii)$ Soit $U \in V(a)$. Puisque (u_n) converge vers a, tous les u_n sont dans U à partir d'un certain rang. A fortiori : $U \cap A \neq \varnothing$.

□

Proposition 1.2.67 (Caractérisation séquentielle de la densité). *Soit A une partie de A. On a équivalence entre :*

(i) A est dense dans E.

(ii) Tout point de E est limite d'une suite de points de A.

Démonstration : On rappelle que A est dense dans E signifie que $\bar{A} = E$.

$(i) \Rightarrow (ii)$ Par le théorème de caractérisation séquentielle de l'adhérence, tout point de $\bar{A}$ est limite d'une suite de points de A. Donc tout point de E est limite d'une suite de points de A.

$(ii) \Rightarrow (i)$ Comme $A \subset E$, $\bar{A} \subset E$ (car E est l'espace ambiant). Réciproquement, soit $a \in E$. Alors, a est limite d'une suite de points de A. Donc $a \in \bar{A}$. Donc A est dense dans E.

□

Proposition 1.2.68 (Caractérisation séquentielle des fermés). *Soit F une partie de E. On a équivalence entre :*

(i) F est fermée.

(ii) Toute suite convergente de points de F a sa limite dans F.

Démonstration : $(i) \Rightarrow (ii)$ Soit (u_n) une suite de points de F qui converge mettons vers $a \in E$. Alors, par le théorème de caractérisation séquentielle de l'adhérence, $a \in \bar{F}$. Or, comme F est fermé, $F = \bar{F}$ et a est dans F.

$(ii) \Rightarrow (i)$ Soit $a \in \bar{F}$. Il existe une suite de points de F qui converge vers a. Donc, a est dans F. Donc : $\bar{F} \subset F$. Donc F est fermée.

□

Proposition 1.2.69 (Caractérisation séquentielle des points d'accumulation). *Soit A une partie de E et a un point de E. Alors, a est un point d'accumulation de A si et seulement si il existe une suite de points deux à deux distincts de A convergeant vers a.*

Démonstration : Soit (a_n) une suite de points de A deux à deux distincts qui converge vers a. Soit $V \in \mathcal{V}(a)$. Il existe $N \in \mathbb{N}$ tel que pour tout $n \geqslant N, a_n \in V$. Comme les $a_n, n \geqslant N$ sont deux à deux distincts et en nombre infini, il y en a au moins un qui est distinct de a et donc a est un point d'accumulation de A.

Réciproquement, soit a un point d'accumulation de A. On construit notre suite par récurrence. La boule $V_0 = B(a, 1)$ rencontre A en un point a_0 distinct de a. Soit alors, $V_1 = B(a, \varepsilon_1)$ avec $\varepsilon_1 = \frac{1}{2}d(a, a_0) > 0$. On choisit a_1, nécessairement distinct de a_0, dans $V_1 \cap (A \backslash \{a\})$. On suppose ainsi $a_0, a_1, \ldots, a_{n-1}$ construits. On pose $\varepsilon_n = \frac{1}{2}d(a, a_{n-1})$ et on choisit a_n dans $B(a, \varepsilon_n) \cap (A \backslash \{a\})$. La suite ainsi construite par récurrence est une suite de points deux à deux distincts et converge vers a car $d(a, a_n) \leqslant \frac{d(a,a_0)}{2^n}$.

□

Ces caractérisations sont fondamentales car en pratique ce sont elles qui sont constamment utilisées. Ainsi, sauf cas exceptionnel, pour montrer qu'un espace est fermé, on essaiera toujours de le faire par caractérisation séquentielle. Ce sont des critères pratiques.

Proposition 1.2.70. *On a : $\overline{B(a, r)} \subset B_f(a, r)$.*

Démonstration : Soit $x \in \overline{B(a, r)}$. Il existe une suite de points de $B(a, r)$, mettons (x_n) qui converge vers x. On a alors : $\forall n \in \mathbb{N},\ d(a, x_n) < r$ et par passage à la limite, $d(a, x) \leqslant r$ et x est dans $B_f(a, r)$.

□

1.3 Limites et continuité des fonctions dans les espaces métriques

Historiquement, la topologie s'est développée pour donner une définition rigoureuse à la notion de limite et de continuité.

1.3.1 Limites

1.3.1.1 Notion de limite

Nous allons tout d'abord donner une définition très générale de la notion de limite.

Définition 1.3.1. *Soient $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques. Soit A une partie de E, f une application de A vers F et $a \in \overline{A}$. On dit que f admet pour limite l lorsque x tend vers a selon A lorsque :*

$$\forall V \in V(l),\ \exists U \in V(a),\ f(U \cap A) \subset V$$

Nous noterons, en cas d'existence d'une limite l de f en a selon A :

$$l = \lim_{x \to a, x \in A} f(x)$$

On peut tout de suite préciser que le fait d'avoir une limite en a est une propriété locale ce qui signifie que si U_0 est un voisinage de a et si f admet une limite en a selon $U_0 \cap A$, elle admet une limite en a selon A.

En effet, si $V \in V(l),\ \exists U \in V(a),\ f(U \cap (U_0 \cap A)) \subset V$, soit en posant $U' = U \cap U_0$, on a : $f(U' \cap A) \subset V$.

Définition 1.3.2. *Soient (E, d) et (F, δ) deux espaces métriques. Soient A une partie de E, $a \in \overline{A}$ et f une application de A dans F. On dit que f admet l pour limite lorsque x tend vers a selon A lorsque :*

$$\forall \varepsilon > 0,\ \exists \eta \in \mathbb{R},\ \forall x \in A,\ d(x, a) < \eta \Rightarrow \delta(f(x), l) < \varepsilon$$

Définition 1.3.3. *Soient $(E, \|\ \|_E)$ et $(F, \|\ \|_F)$ deux espaces vectoriels normés. Soient A une partie de E, $a \in \overline{A}$ et f une application de A dans F. On dit que f admet l pour limite lorsque x tend vers a selon A lorsque :*

$$\forall \varepsilon > 0,\ \exists \eta \in \mathbb{R},\ \forall x \in A,\ \|x - a\|_E < \eta \Rightarrow \|f(x) - l\|_F < \varepsilon$$

Proposition 1.3.4. *Si f admet l comme limite en a selon A, $l \in \overline{f(A)}$.*

Démonstration : Soit $V \in V(l)$. Il existe $U \in V(a),\ f(U \cap A) \subset V$. Alors, on a :

$$f(U \cap A) \subset f(U \cap A) \cap f(A) \subset V \cap f(A)$$

Or, comme $a \in \overline{A}$, $U \cap A \neq \varnothing$ et $f(U \cap A) \neq \varnothing$ et donc : $V \cap f(A) \neq \varnothing$. Comme cela est valable pour tout voisinage de l, on a bien : $l \in \overline{f(A)}$.

□

1.3.1.2 Caractérisation séquentielle

On se donne (E, d) et (F, δ) deux espaces métriques. Soient $A \subset E$, $a \in \overline{A}$ et $f : A \to F$. On a alors :

Théorème 1.3.5. *L'application f admet une limite en a selon A si et seulement si pour toute suite $(u_n)_{n \in \mathbb{N}}$ d'éléments de A convergeant vers a, $(f(u_n))_{n \in \mathbb{N}}$ converge.*

Lemme 1.3.6. *On suppose que pour toute suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de A convergeant vers a, $(f(u_n))_{n\in\mathbb{N}}$ converge. Si (u_n) et (v_n) sont deux suites d'éléments de A qui convergent vers a, alors $(f(u_n))$ et $(f(v_n))$ ont la même limite.*

Démonstration : Soit (w_n) la suite définit par : $w_{2n} = u_n$ et $w_{2n+1} = v_n$. Alors, (w_n) converge vers a. On suppose que $(f(u_n))$ converge vers l, $(f(v_n))$ converge vers l' et $(f(w_n))$ converge vers l''. Alors, $(f(w_n))$ admet comme extraction $(f(u_n))$ et $(f(v_n))$ donc compte tenu du fait que toute suite extraite d'une suite convergente et convergente et à la même limite, on a $l = l' = l''$.

□

On peut alors passer à la démonstration du théorème.

Démonstration : On commence par le sens direct. On suppose que f admet une limite l en a selon A. Soit $(u_n) \in A^{\mathbb{N}}$ convergeant vers a. Soit $\varepsilon > 0$. On a :

$$\exists\alpha > 0,\ \forall x \in A,\ d(a,x) < \alpha \Rightarrow \delta(f(x),l) < \varepsilon.$$

Or, la suite (u_n) converge vers a donc :

$$\exists N \in \mathbb{N},\ \forall n \geqslant N,\ d(a,u_n) < \alpha.$$

Alors, pour $n \geqslant N$, $u_n \in A \cap B(a,\alpha)$. Donc, en prenant $x = u_n$, on a : $\delta(f(u_n),l) < \varepsilon$. Donc, $(f(u_n))$ converge vers l.

Réciproquement, compte tenu du lemme, on note l la limite commune de toutes les suites $(f(u_n))$ où (u_n) est une suite d'éléments de A qui converge vers a. Montrons que f possède alors la limite l en a selon A. On suppose par l'absurde que ce n'est pas le cas :

$$\exists\varepsilon > 0,\ \forall\alpha > 0,\ f(B(a,\alpha) \cap A) \not\subseteq B(l,\varepsilon)$$

On prend $\alpha = \frac{1}{n}$, $n \in \mathbb{N}^*$. Alors, il existe un élément u_n de A tel que :

$$d(a,u_n) < \frac{1}{n},\ \text{et}\ \delta(f(u_n),l) \geqslant \varepsilon$$

De là, $(u_n)_{n\in\mathbb{N}}$ ainsi définie est une suite d'éléments de A qui converge vers a mais $(f(u_n))$ ne converge pas vers l. Ceci contredit l'hypothèse que l'on fait initialement, d'où le résultat.

□

La démonstration du théorème nous permet de donner une version précisée du théorème de caractérisation séquentielle parfois plus pratique que la forme générale que l'on a énoncé :

Théorème 1.3.7. *L'application f admet une limite l en a selon A si et seulement si pour toute suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de A convergeant vers a, $(f(u_n))_{n\in\mathbb{N}}$ converge vers l.*

Remarque : La démonstration du théorème nous permet de généraliser le sens direct du théorème, à savoir que si f admet une limite l en a selon A, alors pour toute suite (u_n) d'éléments de A convergeant vers a, $(f(u_n))$ converge vers l, au cadre des espaces topologiques et sans se restreindre à celui des espaces métriques. La réciproque du théorème quant à elle nécessite tout de même le cadre métrique.

1.3.1.3 Composition des limites

On se donne $(E, \mathcal{T})$, $(F, \mathcal{P})$ et $(G, \mathcal{D})$ trois espaces topologiques. On se donne aussi $f : E \to F$ et $g : F \to G$ deux fonctions. On a alors :

Théorème 1.3.8. *Soient A une partie de E, B une partie de F, $a \in \overline{A}$, $b \in \overline{B}$. On suppose que f admet la limite b en a selon A et que g admet la limite l en b selon B. Alors, si $f(A) \subset B$, $g \circ f$ admet la limite l en a selon A.*

Démonstration : Soit $W \in V(l)$. Alors : $\exists V \in V(b),\ g(V \cap B) \subset W$ et $\exists U \in V(a), f(U \cap A) \subset V$.
On a aussi : $f(A) \subset B$ et donc : $f(U \cap A) \subset B$. De là : $f(U \cap A) \subset B \cap V$. Appliquons g :

$$(g \circ f)(U \cap A) \subset g(B \cap V) \subset W$$

□

1.3.1.4 Limites dans $\mathbb{R}$

On se place tout d'abord dans le cas où seul le but est réel. On considère donc $(E, \mathcal{T})$ un espace topologique et f une application de $A \subset E$ dans $\mathbb{R}$. On a alors les deux résultats intéressants suivants.

Proposition 1.3.9. *Si f admet $l \in \mathbb{R}^*$ comme limite en a selon A, alors il existe un voisinage V_o de a et un réel η strictement positif tel que :*

$$\forall x \in V_o \cap A,\ |f(x)| \geqslant \eta > 0$$

Démonstration : Prenons dans la définition métrique de l'existence d'une limite, $\varepsilon = \frac{|l|}{2}$. Alors, il existe un voisinage V_o de a tel que :

$$\forall x \in V_o \cap A,\ |l| - |f(x)| \leqslant |f(x) - l| \leqslant \frac{|l|}{2}$$

Posons : $\eta = \frac{|l|}{2}$. Alors : $\eta > 0$ et on a bien :

$$\forall x \in V_o \cap A,\ |f(x)| \geqslant \eta$$

□

Proposition 1.3.10. *Si f admet une limite l en $a \in \overline{\mathbb{R}} \cap \overline{A}$ alors f est bornée au voisinage de a.*

Démonstration : Dans la définition métrique de l'existence d'une limite, on prend $\varepsilon = 1$. Alors :

$$\exists V \in V(a),\ \forall x \in V \cap A,\ |f(x) - l| \leqslant 1$$

Alors : $\forall x \in V \cap A,\ |f(x)| \leqslant 1 + |l|$.

□

Bien évidemment, ce dernier résultat s'étend immédiatement au cas où le but est un espace vectoriel normé.

1.3.2 Continuité

1.3.2.1 Continuité locale

Définition 1.3.11. *Soient $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques. Soit A une partie de E, f une application de A vers F et $a \in \overline{A}$. On dit que f est continue en a selon A lorsque :*

$$\forall V \in V(f(a)),\ \exists U \in V(a),\ f(U \cap A) \subset V.$$

Cela revient bien sûr à dire que f admet la limite $f(a)$ en a selon A. Ainsi, la plupart des résultats de cette partie découleront directement de ceux énoncés pour les limites. Autre implication de cela, la continuité en un point est une notion locale d'où le titre de cette section.

On peut bien sûr donner immédiatement l'écriture de l'assertion "*être continue en a*" en termes métriques et dans un espace vectoriel normé.

Définition 1.3.12. *Soient (E, d) et (F, δ) deux espaces métriques. Soient A une partie de E, $a \in \overline{A}$ et f une application de A dans F. On dit que f est continue en a selon A lorsque :*

$$\forall \varepsilon > 0,\ \exists \eta \in \mathbb{R},\ \forall x \in A,\ d(x,a) < \eta \Rightarrow \delta(f(x), f(a)) < \varepsilon$$

Définition 1.3.13. *Soient $(E, \|\ \|_E)$ et $(F, \|\ \|_F)$ deux espaces vectoriels normés. Soient A une partie de E, $a \in \overline{A}$ et f une application de A dans F. On dit que f est continue en a selon A lorsque :*

$$\forall \varepsilon > 0,\ \exists \eta \in \mathbb{R},\ \forall x \in A,\ \|x - a\|_E < \eta \Rightarrow \|f(x) - f(a)\|_F < \varepsilon$$

On peut alors donner une caractérisation de la continuité en terme de voisinages.

Proposition 1.3.14. *Avec les mêmes notations que dans la définition de la continuité, on a équivalence entre :*

(i) f est continue en a.
(ii) Pour tout voisinage V de $f(a)$, $f^{-1}(V)$ est un voisinage de a.
(iii) Pout tout ouvert ω contenant $f(a)$, $f^{-1}(\omega)$ est un voisinage de a.

Démonstration : $(i) \Rightarrow (ii)$: Si V est un voisinage de $f(a)$, puisque f est continue en a :

$$\exists U \in V(a),\ f(U) \subset V.$$

De là, avec U ouvert :

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

et donc $f^{-1}(V)$ est un voisinage de a.
$(ii) \Rightarrow (i)$: Soit $V \in V(a)$. On pose : $U = f^{-1}(V)$. Alors, par hypothèse, U est un voisinage de a et on a : $f(U) \subset V$. Donc f est continue en a.
$(ii) \Rightarrow (iii)$: ω est un voisinage de $f(a)$.
$(iii) \Rightarrow (ii)$: Si V est un voisinage de $f(a)$, il existe un ouvert ω tel que : $a \in \omega$ et $\omega \subset V$. Alors : $f^{-1}(\omega) \subset f^{-1}(V)$. Or, par hypothèse, $f^{-1}(\omega)$ est un voisinage de a, donc $f^{-1}(V)$ est un voisinage de a.

□

Proposition 1.3.15. *Soit f une application de $(E, \mathcal{T})$ espace topologique dans $(F, \|\ \|_F)$ espace vectoriel normé. Si f est continue en a, f est bornée au voisinage de a.*

Démonstration : C'est bien évidemment le même résultat que pour les limites de fonctions réelles.

□

1.3.2.2 Caractérisation séquentielle

On se donne (E, d) et (F, δ) deux espaces métriques. Soient $A \subset E$, $a \in \overline{A}$ et $f : A \to F$. On a alors :

Théorème 1.3.16. *L'application f est continue en a selon A si et seulement si pour toute suite $(u_n)_{n\in\mathbb{N}}$ d'éléments de A convergeant vers a, $(f(u_n))_{n\in\mathbb{N}}$ converge vers $f(a)$.*

Démonstration : Cela résulte immédiatement de la version précisée du théorème de caractérisation séquentielle de l'existence d'une limite en a pour f avec $l = f(a)$.

□

Remarque : Compte tenu de la remarque déjà faite pour la caractérisation séquentielle des limites, le sens direct de ce théorème se généralise sans peine au cadre des espaces topologiques.

Ce résultat est absolument fondamental car en pratique on étudie souvent des suites de la forme $v_n = f(u_n)$ (avec f continue) où (u_n) converge vers une limite connue et où l'on peut en déduire la convergence et la limite de (v_n) à partir de la continuité de f. On l'utilise aussi constamment lorsque l'on cherche à appliquer la caractérisation séquentielle des fermés.

1.3.2.3 Composition

On se donne $(E, \mathcal{T})$, $(F, \mathcal{P})$ et $(G, \mathcal{D})$ trois espaces topologiques. On se donne aussi $f : E \to F$ et $g : F \to G$ deux fonctions. On a alors :

Théorème 1.3.17. *Soient A une partie de E, B une partie de F, $a \in \overline{A}$. On suppose que f est continue en a selon A et que g est continue en $f(a)$ selon B. Alors, si $f(A) \subset B$, $g \circ f$ est continue en a selon A.*

Démonstration : Cela provient immédiatement du théorème de composition des limites.

□

1.3.2.4 Continuité globale

On commence par une définition.

Définition 1.3.18. *Soient $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques. Une application f de E dans F est dite continue sur E lorsqu'elle est continue en tout point de E.*

Remarque : On peut ainsi généraliser le résultat de composition des fonctions continues en affirmant que la composée de deux fonctions continues est une fonction continue.

On passe maintenant au théorème fondamental de caractérisation de la continuité globale :

Théorème 1.3.19. *Soient $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques, f une application de E dans F. On a équivalence entre :*
(i) f est continue sur E.
(ii) Pour tout ouvert Ω de F, $f^{-1}(\Omega)$ est un ouvert de E.
(iii) Pour tout fermé Y de F, $f^{-1}(Y)$ est fermé de E.

Démonstration : On montre tout d'abord l'équivalence entre (ii) et (iii) par complémentarité. On a les relations :

$$f^{-1}(F \backslash \Omega) = E \backslash f^{-1}(\Omega)$$

et

$$f^{-1}(F \backslash Y) = E \backslash f^{-1}(Y)$$

D'où l'équivalence de (ii) et de (iii) car tout fermé est complémentaire d'un ouvert et tout ouvert est complémentaire d'un fermé.

On va maintenant montrer que (i) implique (ii). Soit Ω un ouvert de F. Soit $a \in f^{-1}(\Omega)$. Alors $f(a) \in \Omega$ ouvert, donc Ω est un voisnage de $f(a)$. De là, comme f est continue en a, par caractérisation de la continuité locale par les voisinages, $f^{-1}(\Omega)$ est un voisinage de a. Ainsi, $f^{-1}(\Omega)$ est voisinage de chacun de ses points donc est ouvert.

Il nous reste à montrer la réciproque, à savoir (ii) implique (i). Soit $a \in E$ et $b = f(a)$. Soit $V \in V(b)$ et soit $U = f^{-1}(V)$. Il existe par définition d'un voisinage, un ouvert Ω de F tel que : $b \in \Omega$ et $\Omega \subset V$. Alors, comme Ω est un ouvert de F, $\omega = f^{-1}(\Omega)$ est un ouvert de E. De plus, comme $b \in \Omega$, $a \in \omega$. Comme de plus, $\omega \subset U$, U est un voisinage de a. Alors, par caractérisation de la continuité locale par les voisinages, l'image réciproque de tout voisinage de $f(a)$ par f étant un voisinage de a, f est continue en a. Cela est valable pour tout a dans E, donc f est continue sur E.

$\square$

Ce théorème s'utilise en pratique aussi bien pour montrer qu'une application est globalement continue que pour montrer qu'un ensemble donné est ouvert ou fermé en le voyant comme image réciproque respectivement d'un ouvert ou d'un fermé par une application continue. C'est d'ailleurs dans cette seconde utilisation qu'il se révèle le plus utile.

Nous verrons divers utilisations de ce théorème dans les prochains chapitres, en particulier au chapitre des opérateurs. On peut tout de même déjà donner quelques exemples d'ensembles ouverts ou fermés, ces caractères fermés et ouverts se déduisant immédiatement de ce théorème.

Exemples : On se donne f et g deux applications de E espace topologique dans $\mathbb{R}$, continues. On se donne aussi λ et μ deux réels en supposant de plus $\lambda < \mu$.

Alors les ensembles suivants sont ouverts :

$$\{x \in E,\ f(x) > \lambda\} = f^{-1}(]\lambda, +\infty[)$$
$$\{x \in E,\ f(x) < \lambda\} = f^{-1}(]-\infty, \lambda[)$$
$$\{x \in E,\ \lambda < f(x) < \mu\} = f^{-1}(]\lambda, \mu[)$$

Puis, les ensembles suivants sont fermés :

$$\{x \in E,\ f(x) \geqslant \lambda\} = f^{-1}([\lambda, +\infty[)$$
$$\{x \in E,\ f(x) \leqslant \lambda\} = f^{-1}(]-\infty, \lambda])$$
$$\{x \in E,\ \lambda \leqslant f(x) \leqslant \mu\} = f^{-1}([\lambda, \mu])$$
$$\{x \in E,\ f(x) = \lambda\} = f^{-1}(\{\lambda\})$$

Enfin, l'ensemble suivant est ouvert :

$$\{x \in E,\ f(x) > g(x)\} = (f-g)^{-1}(]0, +\infty[)$$

Et les ensembles suivants sont fermés :

$$\{x \in E,\ f(x) \geqslant g(x)\} = (f-g)^{-1}([0, +\infty[)$$
$$\{x \in E,\ f(x) = g(x)\} = (f-g)^{-1}(\{0\})$$

Voici donc un moyen rapide de montrer que des ensembles que l'on rencontre régulièrement en analyse réelle sout ouverts ou fermés.

1.3.2.5 Projections et plongements

On se donne $(E_1, d_1), \dots, (E_p, d_p)$ des espaces métriques.

Définition 1.3.20. *Les applications $p_i : E \to E_i,\ x \mapsto x_i$ sont appelées projections canoniques.*

Remarque : Lorsque les E_i sont des espaces vectoriels normés, les p_i sont linéaires.

Proposition 1.3.21. *Les p_i sont 1-lipschitziennes et ouvertes.*

Démonstration : Soient $x = (x_1, \dots, x_p)$ et $y = (y_1, \dots, y_p)$ dans E. On a :

$$d_i(p_i(x), p_i(y)) = d_i(x_i, y_i) \leqslant d(x, y)$$

En particulier, les p_i sont continues.
Soit Ω un ouvert de E. On veut $p_i(\Omega)$ ouvert. Soit $a_i \in p_i(\Omega)$. Il existe alors des $a_{j \neq i}$ tels que : $a = (a_1, \dots, a_p) \in \Omega$. Or, comme Ω est ouvert, il existe $r > 0$ tel que : $B_d(a, r) \subset \Omega$. Or : $B_d(a, r) = \prod_{1 \leqslant i \leqslant p} B_{d_i}(a_i, r)$. Donc : $p_j(\prod_{1 \leqslant i \leqslant p} B_{d_i}(a_i, r)) \subset p_j(\Omega)$. Or : $p_j(\prod_{1 \leqslant i \leqslant p} B_{d_i}(a_i, r)) = B_{d_j}(a_j, r)$. Donc : $B_{d_i}(a_i, r) \subset p_i(\Omega)$ qui est donc ouvert.

□

Par contre, les projections ne sont pas fermées en général. Comme contre-exemple, il suffit de prendre le graphe de $x \mapsto \frac{1}{x}$ qui est fermé et sa projection sur l'axe des abscisse qui est égale à $\mathbb{R}^*$ qui est ouvert.

Définition 1.3.22. *Les applications $j_i : E_i \to E,\ x \mapsto (a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p)$ sont appelées plongements canoniques.*

Proposition 1.3.23. *Les j_i sont des isométries. En particulier elles sont continues.*

Démonstration : On a : $d(j_i(x), j_i(x')) = d_i(x, x')$.

□

1.3.2.6 Limites dans les espaces produits

On se donne $(E_1, d_1), \dots, (E_p, d_p)$ des espaces métriques.

Théorème 1.3.24. *Soit (F, δ) un espace métrique, A une partie de F et $a \in \overline{A}$. Soit $f : A \to E$. Alors, f possède une limite en a si et seulement si les composantes $f_1, \dots, f_p$ de f ont une limite en a selon A et dans ce cas :*

$$l = \lim_{x \to a, x \in A} f(x) \Leftrightarrow \forall i \in [\![1, p]\!],\ l_i = \lim_{x \to a, x \in A} f_i(x).$$

Démonstration : Soit $x \in F$. On a :

$$f(x) = (f_1(x), \dots, f_p(x)) = ((p_1 \circ f)(x), \dots, (p_p \circ f)(x))$$

De là, si f a une limite l, par composition des limites, chaque $p_i \circ f$ a une limite : $p_i(l) = l_i$.
Réciproquement, on suppose que chaque f_i admet une limite en a, soit l_i. On se donne alors $\varepsilon > 0$. Il existe $\eta > 0$ tel que :

$$\forall x \in A,\ \delta(a, x) < \eta \Rightarrow d_i(l_i, f_i(x)) < \varepsilon$$

et ce pour tout i dans $[1, p]$. Par passage au max, il vient alors :

$$\forall x \in A,\ \delta(a, x) < \eta \Rightarrow d(l, f(x)) < \varepsilon$$

□

1.3.2.7 Continuité et espaces produits

On se donne $(E_1, d_1), \dots, (E_p, d_p)$ des espaces métriques.

Théorème 1.3.25. *Soit f une application de (F, δ) dans (E, d) et $a \in F$. On a alors équivalence entre :*
(i) f est continue en a (respectivement sur F).
(ii) Chaque f_i est continue en a (respectivement sur F).

Démonstration : C'est une simple reformulation du résultat sur les limites en prenant : $l = f(a)$ et $l_i = f_i(a)$.

□

Proposition 1.3.26. *Soit :*

$$f : \left(\begin{array}{ccc} E & \to & F \\ (x_1, \dots, x_p) & \mapsto & f(x_1, \dots, x_p) \end{array} \right)$$

Si f est continue en $a = (a_1, \dots, a_p)$, les applications partielles :

$$f_{i-partielle} : \left(\begin{array}{ccc} E_i & \to & E \\ x & \mapsto & (a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_p) \end{array} \right)$$

sont continues en a_i.

Démonstration : Les plongements canoniques sont continues donc par composition, $f \circ j_i$ est continue.

□

Il faut ici prendre garde à la réciproque, ce n'est pas parce que toutes les applications partielles sont continues que f est continue, comme on peut le voir en considérant l'exemple de la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par : $f(0,0) = 0$ et $f(x,y) = \frac{xy}{x^2+y^2}$. Alors f n'est pas continue en 0 car : $\forall x \in \mathbb{R}^*,\ f(x,x) = \frac{1}{2} \neq 0$.

1.3.2.8 Principe de prolongement des identités

On se donne $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques, f et g deux applications de E dans F continues sur E. On a alors le résultat :

Théorème 1.3.27 (Principe de prolongement des identités). *S'il existe une partie A de E dense dans E telle que : $\forall x \in A, f(x) = g(x)$, alors $f = g$.*

Démonstration : Soit $x \in E$. Comme A est dense dans E, il existe une suite $(x_n)_{n \in \mathbb{N}}$ d'éléments de A qui converge vers x. Or, par hypothèse : $\forall n \in \mathbb{N},\ f(x_n) = g(x_n)$. Alors par passage à la limite, licite par continuité de f et g :

$$f(x) = \lim_{n \to +\infty} f(x_n) = \lim_{n \to +\infty} g(x_n) = g(x).$$

D'où : $\forall x \in E,\ f(x) = g(x)$.

□

On peut rapidement donner une autre preuve de ce résultat, un peu plus courte :

Démonstration : L'ensemble $\{x \in E \mid f(x) = g(x)\}$ est un fermé par continuité de f et de g, il est donc égal à son adhérence. Or, il est dense dans E donc son adhérence est E et donc cet ensemble est égal à E, ce qui prouve le résultat.

□

Un exemple typique d'utilisation de ce résultat est l'étude des morphismes de $\mathbb{R}$ que nous verrons au chapitre sur la topologie de $\mathbb{R}$.

1.3.2.9 Prolongement par continuité

On termine cette section par un résultat de prolongement par continuité bien connu dans le cadre réel et que étendons ici au cadre des espaces topologiques séparés.

On se donne $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques, F étant de plus séparé, A une partie non vide de E, f une application de A dans F et $a \in \overline{A} \backslash A$. On a alors :

Théorème 1.3.28. *On suppose que f admette la limite l en a selon A. Alors, il existe un unique prolongement g de f à $A \cap \{a\}$ qui soit continue en a. g est défini par : $\forall x \in A,\ g(x) = f(x)$ et $g(a) = l$. On dit que g est le prolongement par continuité de f en a.*

Démonstration : L'existence de g est précisée par l'énoncé car g ainsi définie est évidemment continue. L'unicité de g vient de l'unicité de la limite dans un espace séparé.

□

1.3.2.10 Homéomorphismes

On arrive à présent à l'un des concepts les plus importants en topologie, celui d'homéomorphie.

Définition 1.3.29. *Soient $(E, \mathcal{T})$ et $(F, \mathcal{P})$ deux espaces topologiques. On dit que $f : E \to F$ est un homéomorphisme de E dans F si f est bijective, continue et si sa réciproque f^{-1} est aussi continue.*

Définition 1.3.30. *Deux espaces topologiques $(E, \mathcal{T})$ et $(F, \mathcal{P})$ sont dit homéomorphes s'il existe un homéomorphisme de $(E, \mathcal{T})$ dans $(F, \mathcal{P})$.*

L'idée est que deux espaces homéomorphes ont exactement les mêmes propriétés topologiques, mais pas nécéssairement les mêmes propriétés métriques.

Proposition 1.3.31. *Un homéomorphisme est une application ouverte et fermée.*

Cela signifie que l'image directe d'un ouvert par un homéomorphisme est un ouvert et que l'image directe d'un fermé est un fermé.

Démonstration : Soit U un ouvert de E. Alors : $f(U) = (f^{-1})^{-1}(U)$. D'où, comme f^{-1} est continue, $f(U)$ est ouvert. Idem pour les fermés.

□

Exemples : Dans $\mathbb{R}$, tout segment $[a, b]$ est homéomorphe à $[0, 1]$. En effet, soit f définie par : $f(t) = (1 - t)a + tb$. f est bijective de $[0, 1]$ dans $[a, b]$, continue sur $[0, 1]$ et de réciproque : $f^{-1}(t) = \frac{1}{b-a}t - \frac{a}{b-a}$ elle aussi continue.

On a ensuite que la fonction *tan* est un homéomorphisme de $]-\frac{\pi}{2}, \frac{\pi}{2}[$ dans $\mathbb{R}$. Ainsi, $\mathbb{R}$ est homéomorphe à tout ses segments (par transitivité de l'homéomorphie).

On termine par un exemple de bijection continue dont la réciproque n'est pas continue. Soient $A = [0, 1[\cup \{2\}$ et $B = [0, 1[$. Soit f de A dans B définie par : $\forall x \in [0, 1[,\ f(x) = x$ et $f(2) = 1$. Alors f est bijective et continue sur A. Mais f^{-1} est discontinue en 1.

1.3.3 Uniforme continuité

On se donne (E, d) et (F, δ) deux espaces métriques.

Définition 1.3.32. *Une application $f : E \to F$ est dite uniformément continue lorsque :*

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall (x, y) \in E^2,\ d(x, y) \leqslant \eta \Rightarrow \delta(f(x), f(y)) \leqslant \varepsilon.$$

Définition 1.3.33. *Le réel η défini dans la définition de l'uniforme continuité est appelé module d'uniforme continuité de f pour ε sur E.*

Remarque : L'uniforme continuité est une notion métrique et non une notion topologique, on ne peut l'exprimer en terme de voisinages.

Proposition 1.3.34. *Si $f : E \to F$ est uniformément continue sur E, alors elle est continue sur E.*

Démonstration : Soit $\varepsilon > 0$. Soit $a \in E$. Alors :

$$\exists \eta > 0,\ \forall x \in E, d(a, x) \leqslant \eta \Rightarrow \delta(f(x), f(a)) \leqslant \varepsilon$$

Donc :

$$\forall x \in E, \exists \eta > 0,\ d(a, x) \leqslant \eta \Rightarrow \delta(f(x), f(a)) \leqslant \varepsilon$$

et f est continue en a.

□

On donne enfin un résultat de recollement.

Proposition 1.3.35. *Soient I et J deux parties de E non disjointes ($I \cap J \neq \varnothing$). Alors $f : E \to F$ est uniformément continue sur $I \cup J$ si et seulement si elle est uniformément continue sur I et sur J.*

Démonstration : Dans le sens direct : si η est un module d'uniforme continuité de f pour ε sur $I \cup J$ c'est aussi un module d'uniforme continuité de f pour ε sur I et sur J.

Réciproquement, soit η_1 un module d'uniforme continuité de f pour ε sur I et soit η_2 un module d'uniforme continuité de f pour ε sur J. Soit alors $\eta = \min(\eta_1, \eta_2)$. Alors, η est un module d'uniforme continuité de f pour ε sur $I \cup J$.

□

1.3.3.1 Caractérisation séquentielle

On se donne (E, d) et (F, δ) deux espaces métriques. On a le résultat suivant :

Théorème 1.3.36. *Soit $f : E \to F$. Alors, on a équivalence entre :*
(i) f est uniformément continue sur E.
(ii) Pour tout choix de suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ d'éléments de E, on a :

$$d(x_n, y_n) \to 0 \Rightarrow \delta(f(x_n), f(y_n)) \to 0$$

Démonstration : $(i) \Rightarrow (ii)$: Soit $\varepsilon > 0$. Soit η un module d'uniforme continuité de f pour ε. Soient $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ deux suites d'éléments de E telles que : $d(x_n, y_n) \to 0$. Alors, à partir d'un certain rang, $d(x_n, y_n) \leqslant \eta$. Alors : $\delta(f(x_n), f(y_n)) \leqslant \varepsilon$. Ainsi :

$$\exists N \in \mathbb{N},\ \forall n \geqslant N,\ \delta(f(x_n), f(y_n)) \leqslant \varepsilon$$

Donc : $\delta(f(x_n), f(y_n)) \to 0$.
$(ii) \Rightarrow (i)$: On raisonne par contraposée. On a :

$$\exists \varepsilon > 0,\ \forall \eta > 0, \exists ((x_n),(y_n)) \in (E^{\mathbb{N}})^2,\ d(x_n,y_n) < \eta \text{ et } \delta(f(x_n),f(y_n)) \geqslant \varepsilon$$

On prend : $\eta = \frac{1}{n}$. Alors on a l'existence de deux suites (x_n) et (y_n) telles que : $\forall n \in \mathbb{N}$, $d(x_n,y_n) < \frac{1}{n}$ et $\delta(f(x_n),f(y_n)) \geqslant \varepsilon$. Alors ces deux suites conviennent pour la négation de (ii).

□

En pratique, c'est à l'aide de ce théorème que l'on montre qu'une application n'est pas uniformément continue. On peut aussi s'en servir pour montrer qu'une application est uniformément continue mais c'est moins fréquent, on utilise d'autres outils que nous verrons dans la suite comme le théorème de Heine ou bien le caractère lipschitzien.

1.3.3.2 Composition

On se donne (E,d) (F,δ) et (G,μ) trois espaces métriques.

Théorème 1.3.37. *Si f est une application de E dans F uniformément continue sur E et si g est une application de F dans G uniformément continue sur F, alors l'application $g \circ f$ est uniformément continue sur E.*

Démonstration : Soit $\varepsilon > 0$. Soit η_1 un module d'uniforme continuité de g pour ε. Soit η_2 un module d'uniforme continuité de f pour η_1. Alors, η_2 est un module d'uniforme continuité de $g \circ f$ pour ε. En effet, soient x et y dans E. Si $d(x,y) \leqslant \eta_2$ alors $\delta(f(x),f(y)) \leqslant \eta_1$ et $\mu(g(f(x)),g(f(y))) \leqslant \varepsilon$.

□

1.3.3.3 Fonctions lipschitziennes

Nous allons maintenant passer à une classe de fonctions uniforméments continues très importante tant elle intervient dans nombre de grands théorèmes de l'analyse tels le théorème du point fixe de Picard ou encore le théorème de Cauchy-Lipschitz, la classe des fonctions lipschitziennes.

On se donne (E,d) et (F,δ) deux espaces métriques.

Définition 1.3.38. *Une application $f : E \to F$ est dite k-lipschitzienne s'il existe un réel $k > 0$ tel que :*

$$\forall (x,y) \in E^2,\ \delta(f(x),f(y)) \leqslant kd(x,y)$$

On va tout de suite donner une interprétation graphique du caractère lipschitzien dans le cas de fonctions de $\mathbb{R}$ dans $\mathbb{R}$. En fait, une fonction de $\mathbb{R}$ dans $\mathbb{R}$ est lipschitzienne si et seulement si en tous ses points, les pentes de ses tangentes sont plus petites que la constante de Lipschitz k. Cette interprétation graphique nous amène à la proposition :

Proposition 1.3.39. *Soit I un intervalle de $\mathbb{R}$. Soit $f : I \to \mathbb{R}$, dérivable. On a équivalence entre :*
(i) f' est bornée sur I par $k \in \mathbb{R}^+$.
(ii) f est k-lipschitzienne sur I.

Démonstration : $(i) \Rightarrow (ii)$: On applique l'inégalité des accroissements finis :

$$\forall (x,y) \in I^2,\ |f(x) - f(y)| \leqslant k|x-y|$$

$(ii) \Rightarrow (i)$: On a :

$$\forall (x,y) \in I^2,\ x \neq y \Rightarrow \left|\frac{f(y)-f(x)}{y-x}\right| \leqslant k$$

De là, en faisant tendre y vers x, il vient : $|f'(x)| \leqslant k$.

□

On termine en donnant quelques exemples d'applications lipschitziennes.

Exemples : Soit un espace vectoriel normé $(E, \|\ \|)$. Alors, l'application $x \mapsto ||x||$ est 1-lipschitzienne. En effet, par l'inégalité triangulaire, on a :

$$|\ \|x\| - \|y\|\ | \leqslant \|x-y\|$$

Toutes les fonctions constantes, toutes les fonctions affines de $\mathbb{R}$ dans $\mathbb{R}$ sont lipschitziennes car dérivables et de dérivée bornée.

Enfin, toujours parmi les fonctions de $\mathbb{R}$ dans $\mathbb{R}$, la fonction Arctan admet pour dérivée $x \mapsto \frac{1}{1+x^2}$ bornée par 1 sur $\mathbb{R}$. Donc la fonction Arctan est 1-lipschitzienne sur $\mathbb{R}$.

1.3.3.4 Exemples de fonctions uniformément continues

On va donner ici quelques exemples de fonctions uniforméments continues.

Le premier exemple est absolument fondamental car pour montrer qu'une application est uniformément continue, on montre souvent qu'elle est en fait k-lipschtzienne. En effet, on a le résultat :

Proposition 1.3.40. *Une application f k-lipschitzienne est uniformément continue.*

Démonstration : Soit $\varepsilon > 0$. Alors, $\eta = \frac{\varepsilon}{k+1}$ est un module d'uniforme continuité de f pour ε. En effet :

$$\forall (x,y) \in E^2,\ d(x,y) \leqslant \eta \Rightarrow d(f(x), f(y)) \leqslant kd(x,y) \leqslant \frac{k\varepsilon}{k+1} \leqslant \varepsilon$$

□

De manière un peu plus générale, on définie les applications höldérienne.

Définition 1.3.41. *On dit que f est höldérienne d'ordre $\alpha > 0$ lorsque :*

$$\exists k \in \mathbb{R},\ \forall (x,y) \in \mathbb{R}^2, d(f(x), f(y)) \leqslant kd(x,y)^\alpha$$

On a alors le résultat :

Proposition 1.3.42. *Soit f une application höldérienne d'ordre α. Alors f est uniformément continue.*

Démonstration : Soit $\varepsilon > 0$. Alors, $\eta = \left(\frac{\varepsilon}{k+1}\right)^{\frac{1}{\alpha}}$ est un module d'uniforme continuité de f pour ε.

□

L'exemple fondamental d'application höldérienne est : $t \mapsto t^\alpha$ pour $0 < \alpha \leqslant 1$. En effet, si $0 \leqslant x < y$, on a :

$$y^\alpha - x^\alpha = x^\alpha \left(\left(\frac{y}{x} \right)^\alpha - 1 \right)$$

Or : $u^\alpha - 1 \leqslant (u-1)^\alpha$ pour $u > 1$ (cela se montre à la main en écrivant $u = 1 + h$, puis en considérant les variations de la différence des deux membres). On a alors :

$$y^\alpha - x^\alpha \leqslant (y - x)^\alpha$$

Puis, par symétrie des rôles de x et de y :

$$|y^\alpha - x^\alpha| \leqslant |y - x|^\alpha$$

Cet exemple nous fournit un exemple de fonction uniformément continue et non lipschitzienne. En effet, considérons $f \ : \ x \mapsto \sqrt{x}$ sur $[0, 1]$. Alors f est $\frac{1}{2}$-höldérienne donc uniformément continue mais n'est pas k-lipschitzienne. Si c'était le cas, on aurait : $\forall x \in]0, 1]$, $|f(x) - f(0)| \leqslant kx$, d'où, $\forall x \ ]0, 1]$, $|\sqrt{x}| \leqslant kx$ et donc : $\forall x \ ]0, 1]$, $\frac{1}{\sqrt{x}} \leqslant k$ ce qui est absurde lorsque l'on fait tendre x vers 0.

Enfin, on donne un dernier exemple d'application non uniformément continue. Justifions que la fonction *tan* n'est pas uniformément continue sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Pour cela on considère les suites : $x_n = \frac{\pi}{2} - \frac{1}{n^2}$ et $y_n = \frac{\pi}{2} - \frac{1}{n}$. Alors, on a bien que $|x_n - y_n|$ tend vers 0. De plus :

$$\tan x_n = -\frac{\cos -\frac{1}{n^2}}{\sin -\frac{1}{n^2}} = \frac{\cos \frac{1}{n^2}}{\sin \frac{1}{n^2}} \sim n^2$$

$$\tan y_n = -\frac{\cos -\frac{1}{n}}{\sin -\frac{1}{n}} = \frac{\cos \frac{1}{n}}{\sin \frac{1}{n}} \sim n$$

Donc : $|\tan y_n - \tan x_n| \sim |n^2 - n|$ tend vers $+\infty$.

Ce dernier exemple nous montre l'utilité de la caractérisation séquentielle pour nier des uniformes continuités.

1.4 Espaces vectoriels normés

1.4.1 Topologie des espaces vectoriels normés

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

Définition 1.4.1. *L'application $d : E \times E \to \mathbb{R}^+$ définie par $d(x, y) = ||x - y||$ est une distance sur E appelée distance associée à la norme $|| \ ||$. La topologie associée à cette distance est appelée topologie définie par la norme $|| \ ||$.*

Proposition 1.4.2 (Invariance par translation)**.** *La distance d associée à la norme $|| \ ||$ est invariante par translation.*

Démonstration : Si $a \in E$ et $(x, y) \in E^2$ alors :

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

$\tau_a(x)$ désignant l'image du vecteur x par la translation de vecteur a, $\tau_a : x \mapsto x + a$.

□

Le fait que l'on puisse associer à $|| \ ||$ une distance muni canoniquement E d'une structure d'espace métrique. On va donc retrouver pour les espaces vectoriels normés toutes les propriétes topologiques des espaces métriques.

Nous allons tout d'abord donner un ensemble de résultats concernant l'intérieur et l'adhérence des boules ouvertes et fermées.

Proposition 1.4.3. *On a pour $a \in E$ et $r > 0$:*
(i) $\overline{B(a, r)} = B_f(a, r)$
(ii) $\mathrm{Int}(B_f(a, r)) = B(a, r)$
(iii) $\mathrm{Fr}(B(a, r)) = \mathrm{Fr}(B_f(a, r)) = S(a, r)$

Démonstration : *(i)* On a : $B(a, r) \subset B_f(a, r)$ donc $\overline{B(a, r)} \subset \overline{B_f(a, r)}$. Or $B_f(a, r)$ est fermée donc $\overline{B(a, r)} \subset B_f(a, r)$.

En sens inverse, soit $x \in B_f(a, r)$. Si $x \in B(a, r)$, alors $x \in \overline{B(a, r)}$ et c'est fini. On suppose donc que x n'est pas dans $B(a, r)$. Alors : $x \in S(a, r)$ et $||x - a|| = r$. A présent, montrons que :

$$\forall \delta > 0, \delta < \frac{||x - a||}{2}, B(x, \delta) \cap B(a, r) \neq \varnothing.$$

Soit $\delta > 0$ et $y = x + \frac{\delta}{2} \frac{a-x}{||a-x||}$. On a : $||x - y|| = ||\frac{\delta}{2} \frac{a-x}{||a-x||}|| = \frac{\delta}{2} < \delta$. Donc : $y \in B(x, \delta)$. De plus :

$$\begin{aligned} ||y - a|| &= ||\left(1 - \frac{\delta}{2||x - a||}\right)(x - a)|| = \left(1 - \frac{\delta}{2||x - a||}\right)||x - a|| \\ &= ||x - a|| - \frac{\delta}{2} = r - \frac{\delta}{2} < r. \end{aligned}$$

Donc : $y \in B(a, r)$. On a donc bien :

$$\forall \delta > 0, \delta < \frac{||x - a||}{2}, B(x, \delta) \cap B(a, r) \neq \varnothing$$

et $x \in \overline{B(a, r)}$. D'où le point (i).

Autre preuve : Soit $x \in B_f(a,r)$. On reprend $x \in S(a,r)$. On a :

$$(y \in [a,x]) \Leftrightarrow (\exists \lambda \in [0,1])(y = \lambda x + (1-\lambda)a) \Leftrightarrow (\exists \lambda \in [0,1])(y = a + \lambda(x-a)).$$

On introduit donc pour $\lambda = 1 - \frac{1}{n}$ la suite (x_n) : $x_n = a + (1-\frac{1}{n})(x-a)$. On a : $\forall n \in \mathbb{N}, x_n \in B(a,r)$. En effet : $||x_n - a|| = ||\left(1-\frac{1}{n}\right)(x-a)|| = \left(1-\frac{1}{n}\right)||x-a|| = \left(1-\frac{1}{n}\right)r < r$. De plus : $x_n \to x$. En effet : $||x_n - x|| = ||\frac{1}{n}(x-a)|| = \frac{1}{n}r \to 0$. On a donc bien que : $x \in \overline{B(a,r)}$.

(ii) On a : $B(a,r) \subset B_f(a,r)$ d'où : $\mathrm{Int}(B(a,r)) \subset \mathrm{Int}(B_f(a,r))$. Or $B(a,r)$ est ouverte donc : $B(a,r) \subset Int(B_f(a,r))$. Inversement, soit $x \in Int(B_f(a,r))$. Alors : $\exists \rho > 0, B(x,\rho) \subset B_f(a,r)$. Soit alors : $y = x + \frac{\rho}{2}\frac{x-a}{||x-a||}$. On a : $y \in B(x,\rho)$, donc $y \in B_f(a,r)$. Or :

$$\begin{aligned} y \in B_f(a,r) &\iff ||y-a|| \leqslant r \iff ||\left(1+\frac{\rho}{2||x-a||}\right)(x-a)|| \leqslant r \\ &\iff \left(1+\frac{\rho}{2||x-a||}\right)||x-a|| \leqslant r \iff ||x-a|| \leqslant r - \frac{\rho}{2} < r. \end{aligned}$$

Donc $x \in B(a,r)$ et : $\mathrm{Int}(B_f(a,r)) \subset B(a,r)$.

(iii) On a : $\mathrm{Fr}(B_f(a,r)) = \overline{B_f(a,r)} \backslash \mathrm{Int}(B_f(a,r))$. Or, d'après *(ii)*, $\mathrm{Int}(B_f(a,r)) = B(a,r)$. De là : $\mathrm{Fr}(B_f(a,r)) = B_f(a,r) \backslash B(a,r) = S(a,r)$. De même : $\mathrm{Fr}(B(a,r)) = S(a,r)$.

□

Proposition 1.4.4. *Pour $a \in E$ et $r > 0$, on a : $\delta(B(a,r)) = \delta(B_f(a,r)) = 2r$.*

Démonstration : Comme $B(a,r) \subset B_f(a,r)$, $\delta(B(a,r)) \leqslant \delta(B_f(a,r))$. Soit $(x,y) \in B_f(a,r)^2$. On a : $||x-y|| \leqslant ||x-a|| + ||y-a|| \leqslant 2r$. Donc : $\delta(B_f(a,r)) \leqslant 2r$. Il reste à montrer que $\delta(B(a,r)) \geqslant 2r$. Soit $u \in E$ tel que : $||u|| = 1$. Soit ε tel que : $0 < \varepsilon < r$. On introduit $x_\varepsilon = a + (r-\varepsilon)u$ et $y_\varepsilon = 2a - x_\varepsilon = a - (r-\varepsilon)u$. Alors,$(x_\varepsilon, y_\varepsilon) \in B(a,r)^2$ et $||x_\varepsilon - y_\varepsilon|| = 2r - 2\varepsilon$. D'où : $\delta(B(a,r)) \geqslant 2r - 2\varepsilon$. Cela est valable pour tout $\varepsilon > 0$ donc : $\delta(B(a,r)) \geqslant 2r$. D'où le résultat voulu.

□

A présent on se donne quelques résultats classiques.

Proposition 1.4.5. *Soit F un sous-espace vectoriel de E. Alors, $\overline{F}$ est un sous-espace vectoriel de E.*

Démonstration : Soit $(x,y) \in \overline{F}^2$. Il existe une suite de points de F, soit (x_n) qui converge vers x et une suite de points de F, soit (y_n) qui converge vers y. Alors, $(x_n + y_n)$ est une suite de points de F qui converge vers $x+y$. Donc : $x+y \in \overline{F}$. De même avec λx. Donc $\overline{F}$ est bien un sev de E.

□

Proposition 1.4.6. *Un hyperplan d'un espace vectoriel normé est soit fermé soit dense.*

Démonstration : On a : $H \subset \overline{H} \subset E$. Or $\overline{H}$ est un sous-espace vectoriel de E. Donc : $\overline{H} = E$ ou $\overline{H} = H$.

□

Proposition 1.4.7. *Soit A une partie non vide de E. Si $\overset{\circ}{A} \neq \emptyset$, alors $Vect(A) = E$.*

Démonstration : Quitte à faire une translation on peut supposer que $0 \in \overset{\circ}{A}$ (comme $\overset{\circ}{A} \neq \emptyset$, on peut prendre a dans $\overset{\circ}{A}$). Soit alors $r > 0$ tel que : $B(0, r) \subset A$. Soit $X \in E$. On pose : $x = \frac{r}{2}\frac{X}{||X||}$. Alors, $||x|| = \frac{r}{2}$. De là : $X = \frac{2}{r||x||}x$ et $X \in Vect(B(0, r))$. Donc : $E \subset Vect(B(0, r))$. Or, comme $B(0, r) \subset A$, alors $Vect(B(0, r)) \subset Vect(A)$ et $E \subset Vect(A)$. On a donc bien le résultat voulu.

□

1.4.2 Applications linéaires continues

Nous commençons par définir l'espace des opérateurs bornés entre espaces vectoriels normés.

Proposition 1.4.8. *Soient $(E, \| \|_E)$ et $(F, \| \|_F)$ deux espaces vectoriels normés. Soit $T \in \mathcal{L}(E, F)$. Alors T est continue si et seulement si l'une des conditions suivantes est vérifiée :*

1. *T est continue en zéro.*
2. *T est uniformément continue.*
3. *T est bornée sur $B_f(0, 1)$.*
4. *T est bornée sur $S(0, 1)$.*
5. *T est lipschitzienne.*

Démonstration : Clairement T continue implique la propriété 1. Montrons que $1 \Rightarrow 2$. Si on suppose 1, alors

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ ||z||_E < \eta \Rightarrow ||T(z)||_F < \varepsilon.$$

On prend alors $z = x - y$ et comme $T(x - y) = T(x) - T(y)$ par linéarité de T, il vient :

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ ||x - y||_E < \eta \Rightarrow ||T(x) - T(y)||_F < \varepsilon.$$

D'où 2. Montrons que $2 \Rightarrow 3$. Si on a 2, en particulier T est continue en 0 et pour $\varepsilon = 1$,

$$\exists \eta > 0,\ ||z||_E < \eta \Rightarrow ||T(z)||_F < 1.$$

Donc, si $||x|| \leqslant 1$, on a $||\frac{\eta}{2}x||_E < \eta$ et $||T(\frac{\eta}{2}x)||_F < 1$, ce qui donne

$$\forall x \in E,\ ||x|| \leqslant 1 \ \Rightarrow\ ||T(x)||_F < \frac{2}{\eta}.$$

On a donc bien 3. Bien entendu 3 implique 4. Enfin si on suppose 4 on a :

$$\exists M > 0,\ \forall z \in E,\ ||z|| \leqslant 1 \ \Rightarrow\ ||T(z)||_F < M.$$

On applique cela à $z = \frac{y-x}{||y-x||_E}$ pour $y \neq x$ deux vecteurs E pour obtenir

$$\forall x, y \in E,\ x \neq y,\ ||T(z)||_F = \left\| T(\frac{y - x}{||y - x||_E}) \right\|_F \leqslant M$$

soit encore par linéarité de T, $||T(y - x)||_F \leqslant M||y - x||_E$. Cette dernière inégalité reste vraie pour $x = y$. D'où le point 5. Enfin, le caractère lipschitzien implique la continuité.

□

Définition 1.4.9. *Si $(E, \| \|_E)$ et $(F, \| \|_F)$ sont des espaces vectoriels normés, un opérateur borné de E dans F est une application linéaire continue $T : E \to F$, donc, telle que*

$$\exists C > 0,\ \forall u \in E,\ \|Tu\|_F \leqslant C \|u\|_E .$$

On désigne par $\mathcal{L}_c(E, F)$ l'ensemble des opérateurs bornés de E dans F. Lorsque $E = F$, on notera $\mathcal{L}_c(E) = \mathcal{L}_c(E, E)$. $\mathcal{L}_c(E, F)$ est un espace vectoriel sur lequel on introduit la norme,

$$\|T\|_{\mathcal{L}_c(E,F)} = \sup_{u \in E \setminus \{0\}} \frac{\|Tu\|_F}{\|u\|_E} = \sup_{\|u\|_E = 1} \|Tu\|_F .$$

Remarquons que

$$\|T\|_{\mathcal{L}_c(E,F)} = \inf\{k \geqslant 0 \text{ tels que } T \text{ est } k\text{-liptschitzienne }\}.$$

La topologie induite par cette norme sur $\mathcal{L}_c(E, F)$ est appelée *topologie uniforme des opérateurs*. Si $(F, \| \|_F)$ est un espace de Banach, $(\mathcal{L}_c(E, F), \| \|_{\mathcal{L}_c(E,F)})$ est aussi un espace de Banach. De plus, la norme $\| \|_{\mathcal{L}_c(E,E)}$ est une norme d'algèbre sur $(\mathcal{L}_c(E), +, ., \circ)$ et, plus généralement, si $(E, \| \|_E)$, $(F, \| \|_F)$ et $(G, \| \|_G)$ sont des espaces vectoriels normés et si $T_1 \in \mathcal{L}_c(E, F)$ et $T_2 \in \mathcal{L}_c(F, G)$, alors $T_2 \circ T_1 \in \mathcal{L}_c(E, G)$ et

$$\|T_2 \circ T_1\|_{\mathcal{L}_c(E,G)} \leqslant \|T_2\|_{\mathcal{L}_c(F,G)} \ \|T_1\|_{\mathcal{L}_c(E,F)} .$$

Nous noterons T_2T_1 la composée $T_2 \circ T_1$ de deux opérateurs $T_1 \in \mathcal{L}_c(E, F)$ et $T_2 \in \mathcal{L}_c(F, G)$.

S'il n'y a aucune ambiguité sur l'espace E sur lequel est définie T, nous noterons également $||| \ ||| = || \ ||_{\mathcal{L}_c(E)}$, appelée norme triple.

Les propriétés spécifiques des applications linéaires continues se généralisent au cas des applications bilinéaires continues.

Proposition 1.4.10. *Soient E_1, E_2, F trois espaces vectoriels normés, et $E = E_1 \times E_2$. Soit b une application bilinéaire de $E \to F$. Alors b est continue si et seulement si l'une des conditions suivantes est vérifiée :*

1. *b est continue en zéro.*
2. *il existe k tel que $\|b\,(x_1, x_2)\| \leqslant k \|x_1\| \|x_2\|$.*
3. *b est bornée sur le produit des boules unité fermées.*
4. *b est bornée sur le produit des sphères unité.*

Démonstration : La continuité implique 1. 2 $\Longrightarrow$ 1 est trivial. Montrons que 1 $\Longrightarrow$ 2 : on sait que

$$\exists \eta > 0,\ \forall x \in E, \|x\| < \eta \Longrightarrow \|b(x)\| < 1$$

On l'applique à $x = \left(\eta y_1/\left(2\|y_1\|\right), \eta y_2/\left(2\|y_2\|\right)\right)$, et on obtient

$$\|b\,(y_1, y_2)\| \leqslant \frac{2}{\eta} \|y_1\| \|y_2\| .$$

2 $\Longrightarrow$ 3 et 3 $\Longrightarrow$ 4 sont triviaux. Montrons que 4 $\Longrightarrow$ 2 : il suffit de considérer le vecteur de la sphère unité $x = (y_1/\|y_1\|, y_2/\|y_2\|)$ et d'écrire $\|b(x)\| \leqslant M$. Enfin, montrons que 2 implique la continuité de b : on écrit

$$\begin{aligned} b(x+h) - b(x) &= b\,(x_1 + h_1, x_2 + h_2) - b\,(x_1, x_2 + h_2) + b\,(x_1, x_2 + h_2) - b(x) \\ &= b\,(h_1, x_2 + h_2) + b\,(x_1, h_2) \end{aligned}$$

d'où

$$\|b(x+h) - b(x)\| \leqslant k \left(\|h_1\| \|x_2 + h_2\| + \|x_1 + h_1\| \|h_2\|\right) .$$

□

1.4.3 Dualité

Définition 1.4.11. *Soit $(E, \|.\|)$ un espace vectoriel normé. On appelle dual topologique de E, noté E', l'ensemble des formes linéaires continues sur E.*

Définition 1.4.12. *Soit $(E, \|.\|)$ un espace vectoriel normé. Soit $u \in E'$. La norme de u est par définition le réel positif défini par :*

$$||u||_{E'} = \sup_{x \neq 0} \frac{|u(x)|}{||x||} = \sup_{||x||=1} |u(x)|.$$

Les formes linéiares continues caractérisent les hyperplans fermés.

Théorème 1.4.13. *Une forme linéaire $\phi \in E^*$ est continue si et seulement si son noyau* $\ker \phi$ *est un hyperplan fermé* (i.e. *si et seulement si il n'est pas dense).*

Démonstration : Condition nécessaire : si ϕ est continu, alors $\ker \phi = \phi^{-1}(0)$ est fermé comme image réciproque d'un fermé par une application conrinue.
Condition suffisante : par l'absurde, supposons que $\ker \phi$ est fermé et que ϕ n'est pas continue en zéro. Alors

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x \in E,\ \|x\| < \eta \text{ et } |\phi(x)| > \varepsilon.$$

On considère $\eta = \frac{1}{n}$ et on construit une suite x_n de points de E tels que $\|x_n\| < \frac{1}{n}$ et $|\phi(x_n)| > \varepsilon$. Considérons alors $x \notin \ker \phi$, on a la décomposition en somme directe $E = \ker \phi \oplus \mathbb{K}a$ où $\mathbb{K}$ est le corps sous-jacent ($\mathbb{R}$ ou $\mathbb{C}$). On peut donc écrire $x_n = x'_n + \lambda_n a$, et $\phi(x_n) = \lambda_n \phi(a)$, donc pour tout n, $|\lambda_n| > \varepsilon / |\phi(a)|$. Puisque $|\lambda_n|$ est bornée inférieurement, la suite $(1/|\lambda_n|)$ est bornée, et

$$\lambda_n^{-1} x'_n = \lambda_n^{-1} x_n - a \to -a$$

puisque x_n tend vers zéro. Il y a donc contradiction avec le fait que $\ker \phi$ est fermé.

□

Les formes linéaires continues permettent également de calculer la distance à l'hyperplan fermé qu'elles définissent.

Théorème 1.4.14. *Soit H un hyperplan fermé d'un espace vectoriel normé $(E, || \, ||_E)$, et $\phi \in E'$ telle que* $\ker \phi = H$*. Alors, pour tout $a \in E$,*

$$d(a, H) = \frac{|\phi(a)|}{\|\|\phi\|\|}$$

Notons que si l'on se donne H, alors il est facile de construire une application linéaire ϕ telle que $\ker \phi = H$, en considérant une décomposition en somme directe $E = H \oplus \mathbb{K}b$, où $b \notin H$, et si $x = x' + \mu_x b, \phi(x) = \mu_x$. Lorsque H est fermé, un tel ϕ est continu d'après le théorème précédent.

Démonstration : Considérons $x \in H$. On a :

$$|\phi(x - a)| = |\phi(a)| \leqslant \|\|\phi\|\| \|x - a\|$$

donc $\|x - a\| \geqslant |\phi(a)| / \|\|\phi\|\|$ pour tout $x \in H$, ainsi $d(a, H) \geqslant |\phi(a)| / \|\|\phi\|\|$. Réciproquement, par définition de la norme triple,

$$\forall \varepsilon > 0, \exists y \in E \backslash \{0\},\ |\phi(y)| \geqslant (\|\|\phi\|\| - \varepsilon)\|y\|$$

Si $a \notin H$, alors $E = H \oplus \mathbb{K}a$, donc $y = x + \lambda a$. Comme $\phi(y) \neq 0, \lambda \neq 0$, donc $|\phi(y)| = |\lambda| \cdot |\phi(\frac{x}{\lambda} + a)| = |\lambda| \cdot |\phi(a)| \geqslant (|||\phi||| - \varepsilon)\|y\| = (|||\phi||| - \varepsilon)|\lambda| \cdot \|\frac{x}{\lambda} + a\|$, soit

$$\left\|a - \left(-\tfrac{x}{\lambda}\right)\right\| \leqslant \frac{|\phi(a)|}{|||\phi||| - \varepsilon}, \text{ c'est-à-dire que } d(a, H) \leqslant \frac{|\phi(a)|}{|||\phi||| - \varepsilon}$$

puisque $-\frac{x}{\lambda} \in H$, ceci pour tout $\varepsilon > 0$, ce qui achève la démonstration.

□

1.5 Espaces compacts

1.5.1 Généralités

Nous commençons par donner une définition très générale de la notion de compacité dans un espace topologique quelconque.

Définition 1.5.1 (Borel-Lebesgue)**.** *Un espace topologique $(E, \mathcal{T})$ est dit compact si de tout recouvrement ouvert de E on peut extraire un recouvrement fini (propriété de Borel-Lebesgue),* i.e *:*

$$\text{si } \forall i \in I, \Omega_i \text{ est ouvert et } E = \bigcup_{i \in I} \Omega_i, \quad \text{alors } \exists J \subset I, \text{fini tel que } E = \bigcup_{i \in J} \Omega_i.$$

Ou encore, de toute famille de fermés dont l'intersection est vide, on peut extraire une famille finie dont l'intersection est vide.

Si l'on se restreint au cadre des espaces métrique, on retrouve la propriété de Bolzano-Weierstraß.

Théorème 1.5.2. *Un espace métrique (E, d) est compact si et seulement si toute suite admet une valeur d'adhérence dans E. C'est ce que l'on appelle la propriété de Bolzano-Weierstraß.*

Remarquons que cette propriété est équivalente au fait que toute partie infinie de E admet un point d'accumulation.

Démonstration : Commençons par la condition suffisante, i.e. BL $\Longrightarrow$ BW. Considérons (u_n) une suite, et $A_p = \{u_n \text{ t.q. } n \geqslant p\}$. On a que l'ensemble des valeurs d'adhérence de la suite (u_n) est l'intersection $\bigcap_p \bar{A}_p$. Par l'absurde, si cette intersection est vide (i.e., pas de valeur d'adhérence), alors il existe un nombre fini de p, $p_1 < \ldots < p_n$, tels que $\bar{A}_{p_1} \cap \cdots \cap \bar{A}_{p_n} = \varnothing$. Comme $A_{i+1} \subset A_i$, cela signifie que $\bar{A}_{p_1} = \varnothing$ ce qui est une contradiction.

Condition nécessaire, BW $\Longrightarrow$ BL : la démonstration se fait en trois temps.

On montre que si E est compact, alors E est précompact, c'est-à-dire que pour tout $\varepsilon > 0$, il existe un recouvrement fini de E par des boules ouvertes de rayon ε. Par l'absurde : on construit une suite (x_n) par récurrence, telle que $x_n \notin \bigcup_{1 \leqslant i \leqslant n-1} B_o(x_i, \varepsilon)$. (x_n) admet une valeur d'adhérence, a, limite de la suite extraite $\left(x_{\sigma(n)}\right)$. Or par définition de la suite, on aura $d\left(x_{\sigma(n)}, x_{\sigma(m)}\right) \geqslant \varepsilon$, donc la suite n'est pas de Cauchy : contradiction !

— Soit $(\Omega_i)_{i \in I}$ un recouvrement ouvert de E. Alors il existe un $r \in \mathbb{R}_+^*$ tel que pour tout $x \in E$, $B(x, r)$ est incluse dans au moins l'un des Ω_i. On le démontre par l'absurde : supposons que

$$\forall r > 0, \ \exists x \in E, \ \forall i \in I, \ B(x, r) \nsubseteq \Omega_i.$$

On prend $r = \frac{1}{n}$ et on construit une suite $(x_n)_{n \in \mathbb{N}}$ telle que

$$\forall n \in \mathbb{N}, \forall i \in I, \ B(x_n, \frac{1}{n}) \nsubseteq \Omega_i.$$

La suite $(x_n)_{n\in\mathbb{N}}$ admet une valeur d'adhérence a, limite de la sous-suite $(x_{\varphi(n)})_{n\in\mathbb{N}}$. Par ailleurs, il existe un $j \in I$ tel que $a \in \Omega_j$ et une boule $B(a, r') \subset \Omega_j$. Pour n assez grand, $x_{\varphi(n)} \in B(a, \frac{r'}{2})$, donc en imposant en plus que $\varphi(n) < \frac{2}{r'}$, on $B(x_{\varphi(n)}, \frac{1}{\varphi(n)}) \subset B(a, r') \subset \Omega_j$, d'où une contradiction !

— Pour finir, on empile les deux propriétés précédentes : soit $(\Omega_i)_{i\in I}$ un recouvrement ouvert de E. Il existe un $r > 0$ vérifiant la propriété précédente. On applique alors la première propriété avec $\varepsilon = r$, et on en déduit l'existence d'un recouvrement fini par des boules ouvertes, chacune étant incluse dans un Ω_i.

□

On dira qu'une partie A de E est compacte si (A, d_A) est un espace métrique compact (i.e. toute suite de A a une valeur d'adhérence dans A).

Proposition 1.5.3. *Toute partie compacte est fermée et bornée.*

Démonstration : Une partie compacte est fermée : considérons une suite $(x_n) \in A^{\mathbb{N}}$, convergeant vers $x \in \bar{A}$. Elle admet une valeur d'adhérence dans A. Ce ne peut être que x donc $x \in \bar{A}$. Montrons que A est bornée : par l'absurde, considérons $a \in A$. Alors il existe une suite (x_n) d'éléments de A tels que pour tout $n \in \mathbb{N}$, $d(x_n, a) > n$. Mais par ailleurs il existe une sous-suite convergente, donc borné : contradiction !

□

Proposition 1.5.4. *Soit E un espace métrique compact. Alors A est une partie compacte si et seulement si A est fermée dans E.*

Démonstration : La condition nécessaire résulte du théorème précédent. Pour la condition suffisante, si A est fermée, une suite de A est une suite de E donc convergente, mais alors sa limite est dans $\bar{A} = A$.

□

Proposition 1.5.5. *Toute réunion finie de compacts est un compact, et toute intersection de compacts est un compact.*

Un espace E est dit localement compact si tout point x de E admet un voisinage compact.

Théorème 1.5.6. *Soit E_1, E_2 des espaces compacts, alors $E = E_1 \times E_2$ est compact.*

Démonstration : Considérons une suite de terme principal $x_n = (y_n, z_n)$: par compacité la suite (y_n) admet une valeur d'adhérence, il existe donc une sous-suite $\left(y_{\sigma_1(n)}\right)$ convergente. Considérons la suite $z_{\sigma_1(n)}$: par compacité, il existe une sous-suite convergente, $\left(z_{\sigma_2(\sigma_1(n))}\right)$. La suite $x_{\phi(n)} = \left(y_{\sigma_2(\sigma_1(n))}, z_{\sigma_2(\sigma_1(n))}\right)$ est convergente. D'où le résultat.

□

Plus généralement, un théorème dû à Tychonov assure qu'un produit quelconque d'espaces compact est compact, mais sa démonstration est beaucoup plus difficile et il nécessite également de définir la topologie produit en général.

1.5.2 Fonctions continues sur les espaces compacts

Proposition 1.5.7. *L'image d'une partie compacte par une fonction continue est une partie compacte.*

Démonstration : Soit A une partie compacte et soit f une fonction continue. Soit $(y_n)_{n \in \mathbb{N}}$ une suite d'éléments de $f(A)$. Alors pour chaque $n \in \mathbb{N}$, il existe $x_n \in A$, $y_n = f(x_n)$. Comme A est compacte on peut extraire de (x_n) une sous-suite $(x_{\varphi(n)})$ qui converge vers $x \in A$. Par continuité de f, $y_{\varphi(n)} = f(x_{\varphi(n)})$ tend vers $f(x) \in f(A)$ lorsque n tend vers l'infini. Ainsi on a pu extraire de (y_n) une sous-suite qui converge dans $f(A)$, donc $f(A)$ est compacte.

□

On se place à présent dans le cadre des fonctions numériques. Rappelons que les parties compactes de $\mathbb{R}$ sont exactement les fermés bornés (nous allons redémontrer ce résultat en dimension finie quelconque plus loin).

Théorème 1.5.8. *Soit E un espace compact non vide. Toute fonction numérique continue, i.e. une fonction $f \in \mathcal{C}(E, \mathbb{R})$ est bornée sur E et atteint ses bornes.*

Démonstration : D'après le théorème précédent, l'image de E par f continue est un compact, donc c'est une fermé borné de $\mathbb{R}$, et les bornes ont une image réciproque.

□

Théorème 1.5.9 (Heine). *Soient (E, d) et (E', d') deux espaces métriques. Toute fonction continue sur un compact A de E est uniformément continue sur ce compact.*

Démonstration : Par l'absurde, si une telle fonction f n'est pas uniformément continue sur A,

$$\exists \varepsilon > 0,\ \forall \eta > 0,\ \exists x, y \in A,\ d(x, y) < \eta \text{ et } d'(f(x), f(y)) > \varepsilon.$$

Prenons $\eta = \frac{1}{n}$: il existe donc deux suites d'éléments de A telles que pour tout $n \geqslant 1$,

$$d\left(x_n, y_n\right) < \frac{1}{n} \text{ et } d'\left(f\left(x_n\right), f\left(y_n\right)\right) > \varepsilon$$

Par compacité de A, il existe une suite extraite de x_n qui converge vers une valeur a, notons cette sous-suite $\left(x_{\sigma(n)}\right)$. On a donc aussi $\left(y_{\sigma(n)}\right)$ convergente vers a. Par continuité de f, les suites image convergent, vers la même valeur $f(a)$, ce qui est une contradiction !

□

Proposition 1.5.10. *Si f est une bijection continue d'un espace métrique compact E dans un espace métrique compact E', alors l'application réciproque f^{-1} est continue, c'est-à-dire que f est un homéomorphisme.*

Démonstration : Toute partie A fermée de E est compacte, donc son image $f(A)$ l'est aussi, en particulier $f(A)$ est fermée. Mais alors, comme $f(A) = \left(f^{-1}\right)^{-1}(A)$, l'image réciproque par f^{-1} d'un fermé est un fermé.

□

Proposition 1.5.11. *Soit E un espace métrique, et A une partie compacte. Alors pour tout $x \in E$ il existe $a \in A$ tel que $d(x, A) = d(x, a)$, c'est-à-dire que la distance à un compact est atteinte.*

Démonstration : La fonction de $A \to \mathbb{R}$ qui à a associe $d(x,a)$ est continue sur A compact donc atteint ses bornes.

$\square$

Proposition 1.5.12. *Si A est une partie compacte de E et F une partie fermée telles que $d(A,F) = 0$, alors $A \cap F \neq \varnothing$.*

Démonstration : Si la distance est nulle, il existe deux suites $(x_n) \in A^{\mathbb{N}}$ et $(y_n) \in F^{\mathbb{N}}$ telles que $d(x_n, y_n) < \frac{1}{n}$. Il existe une suite extraite de (x_n) qui converge vers $x \in A$, donc la suite extraite associée de (y_n) converge aussi vers x. Donc $x \in \bar{F} = F$.

$\square$

De même, on montre que la distance entre deux compacts est atteinte.

1.5.3 Compacité et dimension finie

L'intérêt de travailler avec des espaces vectoriels normés plutôt qu'avec des espaces métriques est que l'espace vectoriel sous-jacent possède en plus une structure algébrique. Dans le cas où cet espace vectoriel est de dimension finie, on obtient un certains nombre de résultats de topologie intéressants dont la démonstration de plusieurs d'entre eux reposent sur la notion de compacité.

Proposition 1.5.13. *Les parties compactes de $\mathbb{K}^n$ (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) sont exactement les parties fermées bornées.*

Démonstration : On a déjà fait la condition nécessaire. Pour la condition suffisante, si $A \in \mathbb{R}^n$ est bornée, on a $A \in [\alpha_1, \beta_1] \times \cdots [\alpha_n, \beta_n]$. Chaque intervalle est compact, leur produit l'est aussi, donc A est un fermé dans un compact. Pour $\mathbb{C}^n$, il suffit de l'identifier à $\mathbb{R}^{2n}$.

$\square$

Théorème 1.5.14 (Equivalence des normes)**.** *Toutes les normes sur un espace vectoriel de dimension finie sont équivalentes.*

Démonstration : Tout d'abord, toute norme sur $\mathbb{K}^n$ est majorée par la norme $\|\cdot\|_\infty$, c'est à dire que si N est une norme quelconque et si $(e_1, \dots e_n)$ désigne la base canonique de $\mathbb{K}^n$,

$$\forall x \in \mathbb{K}^n,\ N(x) = N\left(\sum_{i=1}^{n} x_i e_i\right) \leqslant \sum_{i=1}^{n} |x_i| N(e_i) \leqslant \left(\sum_{i=1}^{n} N(e_i)\right) \|x\|_\infty.$$

Pour la minoration, considérons la sphère $\mathcal{S}_\infty = \{x \mid \|x\|_\infty = 1\}$. L'application qui à x associe $N(x)$ est lipschitzienne de $(E, \|\cdot\|_\infty)$ dans $\mathbb{R}$. La sphère $\mathcal{S}_\infty$ est un compact de $(E, \|\cdot\|_\infty)$, donc N atteint ses bornes sur $\mathcal{S}_\infty$. La borne inférieure m est nécessairement non nulle (sinon il existerait $x \in \mathcal{S}_\infty$ tel que $\|x\|_\infty = 0$), et on a : $\forall x \in \mathcal{S}_\infty,\ m \leqslant N(x)$. Donc en prenant $x = y/\|y\|_\infty$, il vient : $m\|y\|_\infty \leqslant N(y)$.

Pour un espace de dimension finie n quelconque, on utilise l'isométrie entre cet espace et $\mathbb{K}^n$ en prenant une base $(e_1, \dots, e_n)$ et en définissant ϕ comme l'application de $\mathbb{K}^n \to E$ qui à $(x_1, \dots, x_n)$ associe $x = \sum_i x_i e_i$.

$\square$

Proposition 1.5.15. *Les parties compactes d'un espace E de dimension finie sont exactement les fermés bornés.*

Démonstration : Conséquence des résultats précédents, en utilisant l'isométrie.

□

Proposition 1.5.16. *Soit E un espace vectoriel normé de dimension finie, F une partie fermée, et $x \in E$. Alors la distance $d(x, F)$ est atteinte.*

Démonstration : Soit $x_0 \in F$, et $A = F \cap B_f(x_0, \|x - x_0\|)$. A est compacte et $d(x, F) = d(x, A)$.

□

Théorème 1.5.17 (Riesz). *Soit $(E, || \, ||)$ un espace vectoriel normé. E est de dimension finie si et seulement si la sphère unité pour $|| \, ||$ est compacte.*

Démonstration : La condition nécessaire est facile, $S(0, 1)$ est fermée bornée. On va montrer la condition suffisante par l'absurde. Supposons que E soit de dimension infinie. Il existe une famille dénombrable de vecteurs de norme 1, (e_n) qui est une famille libre. Considérons $E_n = \operatorname{Vect}(e_1, \ldots, e_n)$. Il existe $h_n \in E_n$ tel que $d(e_{n+1}, E_n) = d(e_{n+1}, h_n) \neq 0$. Considérons $f_n = (e_{n+1} - h_n) / \|e_{n+1} - h_n\|$, c'est un vecteur de la sphère unité, tel que $f_n \in E_{n+1} \backslash E_n$. Comme $d(0, f_n) = 1$, on a $d(f_n, E_n) \leqslant 1$. Par ailleurs,

$$\forall x \in E_n, d(f_n, x) = \frac{e_{n+1} - (h_n - \|e_{n+1} - h_n\| \, x)}{\|e_{n+1} - h_n\|} \geqslant \frac{d(e_{n+1}, E_n)}{d(e_{n+1}, h_n)} = 1$$

Donc la distance $d(f_n, E_n) = 1$, ce qui signifie en particulier que

$$\forall n, p, \ n > p \Longrightarrow \|f_n - f_p\| \geqslant 1$$

donc aucune suite extraite ne peut être de Cauchy, d'où une contradiction.

□

Théorème 1.5.18. *En dimension finie, $\mathcal{L}(E, F) = \mathcal{L}_c(E, F)$. Idem pour les applications bilinéaires.*

Démonstration : En prenant une base, on voit immédiatement que toute application linéaire est lipschitzienne.

□

1.6 Espaces connexes et convexes

1.6.1 Définition et composantes connexes

Définition 1.6.1. *Un espace topologique $(E, \mathcal{T})$ est dit connexe lorsque le seul recouvrement ouvert de E en deux ouverts disjoints est $\{\varnothing, E\}$.*

Autrement dit, E ne peut pas s'écrire comme une réunion disjointe de deux ouverts strictement inclus dans E. De manière équivalente, si $(E, \mathcal{T})$ est connexe, E ne peut pas s'écrire comme une réunion disjointe de deux fermés strictement inclus dans E.

Exemple 1.6.2. *Les parties connexes de $\mathbb{R}$ sont les intervalles.*

Proposition 1.6.3. *Un espace topologique $(E, \mathcal{T})$ est connexe si et seulement si les seules parties de E à la fois ouvertes et fermées sont $\emptyset$ et E.*

Démonstration : Supposons E connexe et soit $\mathcal{O}$ une partie de E à la fois ouverte et fermée. Alors $\mathcal{O} \cup (E \backslash \mathcal{O})$ est un recouvrement de E par deux ouverts disjoints donc $\mathcal{O} = \emptyset$ ou $\mathcal{O} = E$. Réciproquement, si on se donne ouverts disjoints $\mathcal{O}_1$ et $\mathcal{O}_2$ tels que $E = \mathcal{O}_1 \cup \mathcal{O}_2$, alors $\mathcal{O}_1 = E \backslash \mathcal{O}_2$ et $\mathcal{O}_1$ est à la fois ouverte et fermée. D'où $\mathcal{O}_1 = \emptyset$ ou $\mathcal{O}_1 = E$ et E est connexe.

□

Remarquons qu'un espace topologique homéomorphe à un espace topologique connexe est lui aussi connexe.

Proposition 1.6.4. *Soit $(C_i)_{i \in I}$ une famille de parties connexes d'un espace topologique $(E, \mathcal{T})$.*

1. *Si $\cap_{i \in I} C_i \neq \emptyset$, alors $C = \cup_{i \in I} C_i$ est connexe.*
2. *S'il existe $i_0 \in I$ tel que, pour tout $i \in I, C_i \cap C_{i_0} \neq \emptyset$, alors $C = \cup_{i \in I} C_i$ est connexe.*
3. *Si I est au plus dénombrable et si pour tout $i \in I$ (I est assimilé à $\mathbb{N}$), $C_{i+1} \cap C_i \neq \emptyset$, alors $C = \cup_{i \in I} C_i$ est connexe.*

Démonstration : 1. Soient U et V deux ouverts disjoints de C, de réunion C. Pour $i \in I$, on note $U_i = C_i \cap U$ et $V_i = C_i \cap V$. Pour chaque i, U_i et V_i forment un recouvrement de C_i en deux ouverts disjoints et la connexité de C_i entraîne donc que l'un des deux est vide. Supposons par exemple que $(\cap_{i \in I} C_i) \cap U \neq \emptyset$. Alors pour tout $i \in I$, $U_i \neq \emptyset$ donc $V_i = \emptyset$. Donc $C \subset U$ et $V = \emptyset$. Il n'existe donc pas de recouvrement de C en deux ouverts disjoints non vides, donc C est connexe.

2. Par le point 1., pour tout $i \in I$, $C_i \cup C_{i_0}$ est connexe. Or $C = \cup_{i \in I} C_i = \cup_{i \in I}(C_i \cup C_{i_0})$ et comme l'intersection $\cap_{i \in I}(C_i \cup C_{i_0})$ n'est pas vide car elle contient C_{i_0}, par le point 1. encore, $\cup_{i \in I}(C_i \cup C_{i_0})$ est connexe, donc C l'est.

3. Par réccurence, pour tout $k \in \mathbb{N}$, $D_k = \cup_{i=0}^{k} C_i$ est connexe par le point 2. On considère alors la famille $(D_k)_{k \in \mathbb{N}}$. Chaque D_k est connexe et $\cap_{k \in \mathbb{N}} D_k \neq \emptyset$ car C_0 est inclus dans cette intersection. La réunion des D_k est donc connexe par le point 1, d'où le résultat.

□

Nous définissons à présent la notion de composantes connexes d'un espace topologique.

Définition 1.6.5. *Soit $(E, \mathcal{T})$ un espace topologique. Deux points x et y de E sont dit connectés (et on note alors $x \mathcal{C} y$) lorsque x et y appartiennent à une même partie connexe de E. La relation $\mathcal{C}$ ainsi définie est une relation d'équivalence et les classes d'équivalence pour cette relation sont appelées les* composantes connexes *de E.*

Remarquons que la composante connexe d'un point x de E est la plus grande partie connexe de E qui contient x.

1.6.2 Connexité et continuité

Nous voyons à présent les liens entre fonctions continues et connexité.

Théorème 1.6.6. *Soient $(E, \mathcal{T}_E)$ et $(F, \mathcal{T}_F)$ deux espaces topologiques avec $(E, \mathcal{T}_E)$ connexe. Si $f : E \to F$ est continue, alors $f(E)$ est connexe.*

Démonstration : Soient U et V deux ouverts disjoints de $f(E)$ qui recouvrent $f(E)$. Alors, puisque f est continue, $A = f^{<-1>}(U)$ et $B = f^{<-1>}(V)$ sont deux ouverts disjoints qui recouvrent E. Or E est connexe donc A ou B est vide. Supposons que ce soit A. Alors $U = f(A)$ est vide, d'où la connexité de $f(E)$.

□

Appliqué aux connexes de $\mathbb{R}$ qui sont les intervalles, on retrouve le théorème des valeurs intermédiaires.

On peut également utiliser ce résultat pour montrer qu'un espace topologique $(E, \mathcal{T})$ est connexe si et seulement si les seules applications continues de E dans $\{0, 1\}$ sont les constantes.

En utilisant la continuité des projections, on montre également qu'un produit d'espaces topologiques est connexe pour la topologie produit si et seulement si chacun de ses facteurs est un espace connexe.

1.6.3 Connexité par arcs

Nous intorduisons à présent une notion de connexité plus faible mais très utile en pratique.

Définition 1.6.7. *Soit $(E, \mathcal{T})$ un espace topologique. On appelle arc à valeurs dans E toute application continue de $[0, 1]$ dans E.*

Attention à bien prendre garde à la différence entre un arc (une application) et son image. Sinon, on n epeut distinguer l'arc dont l'image est le cercle unité et qui le parcourt une seule fois de l'arc de même image qui parcourt deux fois le cercle...

Définition 1.6.8. *Un espace topologique $(E, \mathcal{T})$ est connexe par arcs si pour tout couple de points $(x, y) \in E^2$, il existe un arc $\gamma : [0, 1] \to E$ tel que $\gamma(0) = x$ et $\gamma(1) = y$. On dit alors que γ joint les points x et y.*

Proposition 1.6.9. *Tout espace topologique connexe par arcs est connexe.*

Démonstration : Soient f une fonction continue de E dans $\{0, 1\}$ et soit $y \in E$. Pour $x \in E$ donné, considérons un arc γ qui joint x et y. Alors la composée $f \circ \gamma$ est continue de $[0, 1]$ dans $\{0, 1\}$. Comme $[0, 1]$ est connexe, $f \circ \gamma$ est alors constante puisque continue, donc $f(x) = f(y)$. Mais comme x est arbitraire, f est constante et E est donc connexe.

□

Il faut prendre garde au fait que la réciproque est fausse : il existe des connexes non connexes par arcs (penser à l'adhérence du graphe de $x \mapsto \sin(\frac{1}{x})$ sur $]0, 1]$).

Corollaire 1.6.10. *L'espace $\mathbb{R}$ n'est pas homémomorphe à l'espace $\mathbb{R}^n$ pour $n \geqslant 2$ (et pour les topologies usuelles).*

Démonstration : Supposons le contraire et considérons un homéomorphisme $\phi : \mathbb{R} \to \mathbb{R}^n$. Soit $a \in \mathbb{R}$. Alors la restriction à $\mathbb{R} \backslash \{a\}$ de ϕ définit un homéomorphisme de $\mathbb{R} \backslash \{a\}$ sur $\mathbb{R}^n \backslash \{\phi(a)\}$. Mais, pour $n \geqslant 2$, $\mathbb{R}^n \backslash \{\phi(a)\}$ est connexe par arcs donc connexe, alors que $\mathbb{R} \backslash \{a\}$ ne l'est pas.

□

On peut étendre le résultat à $\mathbb{R}^n$ et $\mathbb{R}^m$ pour $n \neq m$, mais c'est plus difficile. Il s'agit d'un théorème dû à Brouwer.

Montrons enfin un résultat d'existence qui généralise le théorème des valeurs intermédiaires.

Proposition 1.6.11 (Passage des douanes)**.** *Soient $(E, \mathcal{T})$ un espace topologique, A une partie de E et B une partie connexe de E qui rencontre $\overset{\circ}{A}$ et $E \backslash \overline{A}$. Alors B rencontre la frontière de A.*

Démonstration : Les trois parties $\overset{\circ}{A}$, $E \backslash \overline{A}$ et $\mathrm{Fr}(A)$ sont disjointes et recouvrent E. Laes parties $\overset{\circ}{A}$ et $E \backslash \overline{A}$ sont ouvertes dans E. Si B ne rencontre pas $\mathrm{Fr}(A)$, les parties $\overset{\circ}{A} \cap B$, et $(E \backslash \overline{A}) \cap B$ forment un recouvrmeent de B en deux ouverts disjoints et non vides. Cela contredit la connexité de B. Donc $B \cap \mathrm{Fr}(A) \neq \emptyset$.

□

1.6.4 Convexité

Soient E un espace vectoriel et x et y deux vecteurs de E. On appelle segment $[x, y]$ l'ensemble

$$[x, y] = \{tx + (1 - t)y \mid t \in [0, 1]\}.$$

Définition 1.6.12. *On dit qu'une partie A d'un espace vectoriel E est convexe si pour tous $x, y \in A$, le segment $[x, y]$ est inclus dans A.*

On peut alors montrer les propriétés suivantes :

1. Une partie $A \subset E$ est convexe si et seulement si l'ensemble des barycentres positifs de points de A est inclus dans A.
2. Si l'on définit l'enveloppe convexe d'une partie F comme le plus petit convexe contenant F, alors c'est l'ensemble des barycentres positifs de points de F. C'est également l'intersection de tous les convexes contenant F.

1.7 Espaces complets

1.7.1 Généralités

1.7.1.1 Définition

Toutes les suites de Cauchy ne convergent pas. Par contre, dans certains espaces, toutes les suites de Cauchy convergent, ces espaces sont ceux que nous allons à présent étudier.

Définition 1.7.1. *On dit que l'espace métrique (E, d) est complet lorsque toute suite de Cauchy de cet espace est convergente dans cet espace.*

Nous allons ici travailler naturellement dans le cadre métrique. Comme on l'a déjà remarqué dans la première partie, tenter de travailler avec des suites de Cauchy dans un espace topologique général n'aurait aucun sens. Ainsi, parler d'espace complet n'a de sens que dans le cadre métrique.

Remarque : Il faut faire très attention dans la définition au fait que la limite de notre suite de Cauchy doit être dans l'espace de départ. Cela peut nous aider à montrer qu'un espace n'est pas complet comme on l'a déjà fait pour $\mathbb{Q}$, mais nous assure aussi l'appartenance d'une limite de suite à un espace complet donné.

On peut donner comme exemples connus d'espaces complets, $(\mathbb{R}, |\ |)$ et $(\mathbb{C}, |\ |)$. Ce sont des résultats classiques de première année qui sont en fait une conséquence d'un résultat très général qui nous dit que tout espace vectoriel normé de dimension finie est complet.

On peut enfin donner une première propriété :

Proposition 1.7.2. *Si l'espace métrique (E, d) est complet et si δ est une distance équivalente à d, (E, δ) est complet.*

Démonstration : Soient $a, b > 0$ tels que : $ad \leqslant \delta \leqslant bd$. Soit (x_n) une suite de Cauchy de (E, δ). On a :

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N},\ \forall m \geqslant N_\varepsilon,\ \forall n \geqslant N_\varepsilon,\ \delta(x_n, x_m) \leqslant \varepsilon$$

Mais alors :

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N},\ \forall m \geqslant N_\varepsilon,\ \forall n \geqslant N_\varepsilon,\ ad(x_n, x_m) \leqslant \varepsilon$$

Donc, (x_n) est une suite de Cauchy de (E, d) complet, elle converge donc dans E. Donc toute suite de Cauchy de (E, δ) converge dans E, (E, δ) est donc complet.

□

Remarque : Il faut prendre garde qu'un même ensemble E peut être complet pour une certaine métrique et ne plus l'être pour une autre.

Définition 1.7.3. *Un espace vectoriel normé complet est appelé espace de Banach.*

1.7.1.2 Sous-espaces

On se donne F une partie de l'espace métrique (E, d). On a alors :

Proposition 1.7.4. *Si F est complète pour la métrique induite, alors F est fermée.*

Démonstration : Soit $a \in \overline{F}$. Soit alors, par caractérisation séquentielle de l'adhérence, une suite (x_n) de points de F qui converge vers a. Alors, (x_n) est de Cauchy et donc puisque F est complet, elle converge vers $a' \in F$. On a alors :

$$d(a, a') \leqslant d(x_n, a) + d(x_n, a') \to 0$$

Donc : $d(a, a') = 0 \Rightarrow a = a'$. Donc : $a \in F$. Donc : $\overline{F} \subset F$ et F est fermée.

□

Proposition 1.7.5. *Si (E, d) est de plus complet, si F est fermée dans E, F est complète pour la métrique induite.*

Démonstration : Soit $(x_n) \in F^{\mathbb{N}}$ de Cauchy pour d_F. (x_n) est alors de Cauchy pour d donc converge vers $a \in E$. Or, on sait par caractérisation séquentielle de l'adhérence que $a \in \overline{F}$. Or, F est fermée, donc $a \in F$. Ainsi, (F, d_F) est complet.

□

On peut résumer les deux propriétés que l'on vient de montrer :

Théorème 1.7.6. *Une partie A d'un espace métrique complet est complète pour la métrique induite si et seulement si elle est fermée.*

Notons que les espaces compacts sont toujours complets.

Proposition 1.7.7. *Tout espace métrique (E, d) compact est complet.*

Démonstration : Une suite de Cauchy qui admet une valeur d'adhérence est convergente.

□

1.7.1.3 Produit fini d'espaces complets

On donne ici un premier résultat relativement évident sur la complétude en espace produit.

Théorème 1.7.8. *Soient $(E_1, d_1), \dots, (E_p, d_p)$ p espaces métriques complets. Alors le produit de ces p espaces métriques, soit (E, d), est un espace métrique complet.*

Démonstration : Soit $((x_{1n}, \dots, x_{pn}))_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. On suppose que $((x_{1n}, \dots, x_{pn}))_{n\in\mathbb{N}}$ est de Cauchy. Alors :

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \in \mathbb{N},\ \forall m \geqslant n \geqslant n_\varepsilon,\ \max(d_1(x_{1n}, x_{1m}), \dots, d_p(x_{pn}, x_{pm})) \leqslant \varepsilon$$

Ainsi : $\forall i \in [1, p],\ d_i(x_{in}, x_{im}) \leqslant \varepsilon$. Alors, les $(x_{in})_{n\in\mathbb{N}}$ sont de Cauchy dans (E_i, d_i) donc convergent dans E_i, mettons vers a_i. Alors, $(x_{1n}, \dots, x_{pn})_{n\in\mathbb{N}}$ converge vers $(a_1, \dots, a_p) \in E$. Donc $(x_{1n}, \dots, x_{pn})_{n\in\mathbb{N}}$ est de Cauchy dans (E, d).

□

1.7.1.4 Théorème des fermés emboîtés

On commence par un lemme de caractérisation des suites de Cauchy.

Lemme 1.7.9. *Soit (u_n) une suite à valeurs dans (E, d) espace métrique. On note alors :*

$$C_n(u) = \overline{\{u_p \mid p \geqslant n\}}$$

Alors, (u_n) est de Cauchy si et seulement si la suite des diamètres $\delta(C_n(u))$ tend vers 0.

Démonstration : On suppose que (u_n) est de Cauchy. Soit $\varepsilon > 0$. Soit $N \in \mathbb{N}$ tel que :

$$\forall n, m \geqslant N,\ d(u_n, u_m) \leqslant \varepsilon$$

Soit alors u_n et u_m dans $C_N(u)$. Alors : $d(u_n, u_m) \leqslant \varepsilon$. Donc : $\delta(C_N(u)) \leqslant \varepsilon$. Or : $\forall n \geqslant N,\ C_n(u) \subset C_N(u)$. Donc :

$$\forall n \geqslant N,\ \delta(C_n(u)) \leqslant \delta(C_N(u)) \leqslant \varepsilon$$

Ainsi, $(\delta(C_n(u)))$ tend vers 0.
Réciproquement, si $(\delta(C_n(u)))$ tend vers 0 :

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N,\ \delta(C_n(u)) \leqslant \varepsilon$$

De là, soit $n, m \geqslant N$. On a : $u_n \in C_N(u)$ et $u_m \in C_N(u)$. Ainsi :

$$d(u_n, u_m) \leqslant \delta(C_N(u)) \leqslant \varepsilon$$

Donc, (u_n) est de Cauchy.

□

Remarque : Avec ces notations, on retrouve le fait que l'ensemble des valeurs d'adhérences de (u_n) est l'intersection des $C_n(u)$ ce que l'on avait vu au chapitre sur les espaces métriques, au paragraphe sur les valeurs d'adhérences.

On peut alors énoncer le théorème de Cantor-Baire, encore connu sous le nom de théorème des fermés emboîtés lorsque l'on ne considère pas sa réciproque.

Théorème 1.7.10 (Cantor-Baire). *On a équivalence entre :*

(i) L'espace métrique (E, d) est complet.

(ii) Toute suite de fermés non vides de E, décroissante et dont le diamètre tend vers 0 a une intersection non vide.

Démonstration : $(i) \Rightarrow (ii)$ On se donne (F_n) une suite décroissante de fermés non vides dont le diamètre tend vers 0. Comme chaque F_n est non vide, on peut définir une suite (u_n) d'éléments de E telle que pour tout n, $u_n \in F_n$. Alors, pour tout n, on a : $C_n(u) \subset F_n$. De ce fait, par domination, la suite $\delta(C_n(u))$ tend vers 0 et par le lemme, (u_n) est de Cauchy. Comme E est complet, la suite (u_n) converge, mettons vers l. En particulier, l est une valeur d'adhérence de (u_n) donc compte tenu de la remarque ci-dessus, l est dans l'intersection des $C_n(u)$ qui est donc non vide. Or, comme pour tout n, $C_n(u) \subset F_n$, l'intersection des $C_n(u)$ est incluse dans celle des F_n qui est donc non vide.

$(ii) \Rightarrow (i)$ Soit (u_n) une suite de Cauchy de E. Alors la suite $(C_n(u))$ est une suite décroissante de fermés dont le diamètre tend vers 0 par le lemme. Alors, l'intersection des $C_n(u)$ est non vide. Soit a un élément de cette intersection. Alors a est une valeur d'adhérence de (u_n). Ainsi, (u_n) est une suite de Cauchy qui possède une valeur d'adhérence, elle est donc convergente. Donc, (E, d) est complet.

□

On va alors clairement énoncer le théorème des fermés emboîtés qui est le sens utile en pratique.

Théorème 1.7.11 (Fermés emboîtés). *Soit (F_n) une suite de parties de l'espace complet (E, d) vérifiant :*

(i) Chaque F_n est fermé et non vide.

(ii) La suite (F_n) décroît.

(iii) La suite des diamètres $(\delta(F_n))$ tend vers 0.

Alors, on a :

$$\bigcap_{n \in \mathbb{N}} F_n \neq \emptyset$$

Remarque : Il faut prendre garde à toutes les hypothèses. Par exemple la suite $([n, +\infty[)$ est une suite décroissante de fermés non vide dont l'intersection est vide. En effet, le diamètre de chacun de ces fermés est $+\infty$ et donc la suite des diamètres ne tend pas vers 0.

Ce résultat nous servira plus loin dans la démonstration du théorème de Baire. En attendant, il peut être le moyen de montrer qu'un espace métrique n'est pas complet en exhibant une suite décroissante de fermés dont le diamètre tend vers 0 et dont l'intersection est vide.

1.7.2 Complétion topologique

Soit (E, d) un espace métrique. On cherche à plonger (E, d) dans un espace complet dont la distance prolonge celle de E.

On note $\mathcal{C}$ l'ensemble des suites de Cauchy de E.

Lemme 1.7.12. *Soient $U = (u_n)$ et $V = (v_n)$ dans $\mathcal{C}$. Alors, la suite $(d(u_n, v_n))_{n \in \mathbb{N}}$ converge. On notera $\delta(U, V)$ sa limite.*

Démonstration : Soit $\varepsilon > 0$. Alors, comme U et V sont de Cauchy, on a :

$$\exists N_1 \in \mathbb{N},\ \forall n, m \geqslant N_1,\ d(u_n, u_m) \leqslant \varepsilon$$

et
$$\exists N_2 \in \mathbb{N},\ \forall n, m \geqslant N_2,\ d(v_n, v_m) \leqslant \varepsilon$$

Soit alors $N = \max(N_1, N_2)$. Alors, pour $n, m \geqslant N$, on a :

$$|d(u_n, v_n) - d(u_m, v_m)| \leqslant d(u_n, u_m) + d(v_n, v_m) \leqslant 2\varepsilon$$

En effet :
$$d(u_n, v_n) \leqslant d(u_n, u_m) + d(u_m, v_m) + d(v_m, v_n)$$

et :
$$d(u_n, v_n) - d(u_m, v_m) \leqslant d(u_n, u_m) + d(v_n, v_m)$$

De même dans l'autre sens pour avoir l'inégalité avec la valeur absolue.
On a alors montré que $(d(u_n, v_n))_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{R}$ complet donc est convergente.

□

Lemme 1.7.13. *L'application δ ainsi définie est une semi-distance sur $\mathcal{C}$.*

Démonstration : (i) Si $U = V$, on a : $\forall n \in \mathbb{N},\ d(u_n, v_n) = 0$ et donc par passage à la limite : $\delta(U, V) = 0$.
(ii) On a : $\forall n \in \mathbb{N},\ d(u_n, v_n) = d(v_n, u_n)$ et donc par passage à la limite : $\delta(U, V) = \delta(V, U)$.
(iii) On a : $\forall n \in \mathbb{N},\ d(u_n, v_n) \leqslant d(u_n, w_n) + d(w_n, v_n)$ et donc par passage à la limite :

$$\delta(U, V) \leqslant \delta(U, W) + \delta(W, V)$$

□

On introduit la relation d'équivalence sur $\mathcal{C}$ associée à la semi-distance δ :

$$(U \sim V) \Leftrightarrow (\delta(U, V) = 0)$$

On a alors le résultat d'existence suivant :

Théorème 1.7.14. *L'espace $(\mathcal{C}/\sim, \delta)$ est un espace métrique complet. De plus, il existe une injection naturelle $i\ :\ E \to \mathcal{C}/\sim$, isométrique et telle que $i(E)$ est dense dans $\mathcal{C}/\sim$.*

Démonstration : Tout d'abord, comme δ est une semi-distance sur $\mathcal{C}$, $\mathcal{C}/\sim$ est un espace métrique d'après le résultat que l'on a déjà vu au chapitre sur les espaces métriques à propos des semi-distances.

Pour tout $\alpha \in E$, on note $(\alpha) \in \mathcal{C}$ la suite constante égale à α. Soit i l'application de E dans $\mathcal{C}/\sim$ qui à $\alpha \in E$ associe la classe de (α). Alors :

$$\delta(i(\alpha), i(\beta)) = \delta((\alpha), (\beta)) = d(\alpha, \beta)$$

Donc i est isométrique donc est injective. En effet, toute isométrie est injective. Soit f une isométrie et soient x et y tels que : $f(x) = f(y)$. Alors : $\delta(f(x), f(y)) = 0 = d(x, y)$ et : $x = y$.

Montrons alors que $i(E)$ est dense dans $\mathcal{C}/\sim$. Soit $\overline{U} \in \mathcal{C}/\sim$ et $U = (u_n) \in \mathcal{C}$ un représentant de $\overline{U}$. On va montrer que $\overline{U}$ est limite de la suite $(i(u_n))_{n\in\mathbb{N}}$. Soit $\varepsilon > 0$.

Comme (u_n) est de Cauchy, il existe $N \in \mathbb{N}$ tel que pour tous $n, m \geqslant N$, on ait $d(u_n, u_m) \leqslant \varepsilon$. On fixe alors $m \geqslant N$. On a alors :

$$\delta(\overline{U}, i(u_m)) = \delta(U, (u_m)) = \lim_{n \to +\infty} d(u_n, u_m) \leqslant \varepsilon$$

Ceci étant vrai pour tout $p \geqslant N$, on en déduit que :

$$\lim_{m \to +\infty} \delta(\overline{U}, i(u_m)) = 0$$

et que :

$$\lim_{m \to +\infty} i(u_m) = \overline{U}$$

Ainsi, on a montré que tout élément de $\mathcal{C}/\sim$ est limite de points de $i(E)$ donc par caractérisation séquentielle de la densité, $i(E)$ est dense dans $\mathcal{C}/\sim$.

Il nous reste à montrer le plus important, à savoir le fait que $(\mathcal{C}/\sim, \delta)$ est complet. Soit $(\alpha_n)_{n \in \mathbb{N}}$ une suite de Cauchy de $(\mathcal{C}/\sim, \delta)$. Comme $i(E)$ est dense dans $\mathcal{C}/\sim$, pour tout entier n il existe un point x_n de E tel que : $\delta(\alpha_n, i(x_n)) < \frac{1}{n}$. Alors, l'inégalité :

$$\begin{aligned} d(x_n, x_p) &= \delta(i(x_n), i(x_p)) \\ &\leqslant \delta(i(x_n), \alpha_n) + \delta(\alpha_n, \alpha_p) + \delta(\alpha_p, i(x_p)) \\ &\leqslant \delta(\alpha_n, \alpha_p) + \frac{1}{n} + \frac{1}{p} \end{aligned}$$

montre que la suite (x_n) est de Cauchy dans E (preuve en 3ε). Notons alors α la suite $(\overline{x}_n)_{n \in \mathbb{N}}$ élément de $\mathcal{C}/\sim$.
Montrons que la suite (α_n) converge vers α. On a :

$$\delta(\alpha_n, \alpha) \leqslant \delta(\alpha_n, i(x_n)) + \delta(i(x_n), \alpha) \leqslant \frac{1}{n} + \delta(i(x_n), \alpha)$$

Donc, il nous suffit de montrer que $(\delta(i(x_n), \alpha))$ tend vers 0. Soit $\varepsilon > 0$. La suite (x_n) étant de Cauchy dans E :

$$\exists N \in \mathbb{N},\ \forall p \geqslant N,\ \forall q \geqslant N,\ d(x_p, x_q) \leqslant \varepsilon$$

Donc, si on fixe $n \geqslant N$, il vient :

$$\delta(i(x_n), \alpha) = \lim_{p \to +\infty} d(x_n, x_p) \leqslant \varepsilon$$

et ceci pour tout $n \geqslant N$. Alors, on passe à la limite pour n en $+\infty$ et on obtient :

$$\lim_{n \to +\infty} \delta(i(x_n), \alpha) \leqslant \varepsilon$$

et donc cette limite est nulle. On a alors bien que (α_n) tend vers α par majoration. Ainsi, toute suite de Cauchy de $(\mathcal{C}/\sim, \delta)$ est convergente, donc $(\mathcal{C}/\sim, \delta)$ est complet.

□

Définition 1.7.15. *L'espace $(\mathcal{C}/\sim, \delta)$ est appelé complété topologique de (E, d).*

Ainsi, E est plongé dans un espace complet dont la métrique prolonge celle de E.
On termine par un résultat d'unicité du complété topologique à isométrie bijective près.

Théorème 1.7.16. *Soient (E_1, d_1) et (E_2, d_2) deux espaces métriques complets tels qu'il existe une isométrie i_1 de E dans E_1 avec $i_1(E)$ dense dans E_1 et une isométrie i_2 de E dans E_2 avec $i_2(E)$ dense dans E_2. Alors, il existe une unique isométrie φ de E_1 dans E_2, bijective et vérifiant, pour tout x dans E : $\varphi(i_1(x)) = i_2(x)$.*

Démonstration : Pour cette démonstration, on admet temporairement le théorème de prolongement uniformément continu que nous verrons plus loin.

On définit φ sur $i_1(E)$ en posant : $\varphi(i_1(x)) = i_2(x)$ pour tout x dans E. L'application φ restreinte à $i_1(E)$ est isométrique car :

$$\begin{aligned} \forall (x,y) \in E^2,\ d_2(\varphi(i_1(x)), \varphi(i_1(y))) &= d_2(i_2(x), i_2(y)) \\ &= d(x,y) \\ &= d_1(i_1(x), i_1(y)) \end{aligned}$$

Ainsi φ est uniformément continue sur $i_1(E)$ (évident, par caractérisation séquentielle, toute isométrie est uniformément continue). Comme $i_1(E)$ est dense dans E_1 et que E_2 est complet, il existe un prolongement de φ sur E_1 (que l'on note encore φ) qui est uniformément continu sur E_1. De plus φ est isométrique sur $i_1(E)$ et comme $i_1(E)$ est dense dans E_1 et que φ est continue, φ est isométrique sur E_1 tout entier (écrire que tout élément de E est limite d'une suite d'éléments de $i_1(E)$). En particulier, φ est injective.

Il nous reste à montrer que φ est surjective. Soit $\beta \in E_2$. Comme $i_2(E)$ est dense dans E_2, il existe une suite $(\beta_n) = (i_2(x_n))$ de $i_2(E)$ qui converge vers β. De plus, pour tous $p,\ q \in \mathbb{N}$:

$$d_1(i_1(x_p), i_1(x_q)) = d(x_p, x_q) = d_2(i_2(x_p), i_2(x_q)) = d_2(\beta_p, \beta_q)$$

La suite $(i_1(x_n))$ est donc de Cauchy dans E_1. Comme E_1 est complet, elle converge. Soit α sa limite. Comme φ est continue :

$$\varphi(\alpha) = \lim_{n\to+\infty} \varphi(i_1(x_n)) = \lim_{n\to+\infty} i_2(x_n) = \lim_{n\to+\infty} \beta_n = \beta$$

D'où la surjectivité de φ. Ainsi, φ est bijective. Enfin, l'unicité de φ vient de l'unicité du prolongement uniformément continue sur E_1.

□

On peut alors parler "du" complété de E.

Remarque : C'est un peu de cette manière que l'on construit $\mathbb{R}$ à partir de $\mathbb{Q}$ à la différence près que $\mathbb{R}$ "n'existant" pas encore, le réel $\delta(U, V)$ n'est pas défini. On définit alors la relation d'équivalence par : $U \sim V \Leftrightarrow \lim_{n\to+\infty} u_n - v_n = 0$, la notion de limite étant définie dans $\mathbb{Q}$.

1.7.3 Espaces complets de référence

1.7.3.1 Espace vectoriels normés de dimension finie

Théorème 1.7.17. *Tout espace vectoriel normé de dimension finie est complet.*

Démonstration : Toute suite de Cauchy est bornée et est donc dans une boule fermée qui est compacte (dimension finie). Ainsi elle admet une valeur dadhérence dans cette boule fermée et converge donc.

□

En particulier, tout sous-espace de dimension finie d'un espace vectoriel normé est complet et est donc fermé.

Notons que l'on peut aussi déduire le théorème précédent du fait que tout $\mathbb{K}$-espace vectoriel E de dimension finie est isomorphe à $\mathbb{K}^n$ où $n = \dim E$.

1.7.3.2 L'espace $(\mathcal{B}(X,E), \| \ \|_\infty)$ des applications bornées

Soit X un ensemble et $(E, \| \ \|)$ un espace de Banach. On introduit $\mathcal{B}(X,E)$ l'espace des applications bornées de X vers E. Pour $f \in \mathcal{B}(X,E)$, on pose :

$$\|f\|_\infty = \sup_{x \in X} \|f\|$$

On a alors :

Théorème 1.7.18. *$(\mathcal{B}(X,E), \| \ \|_\infty)$ est un espace complet.*

Démonstration : On commence par se donner (f_n) une suite de fonctions bornées de X dans E, de Cauchy pour $\|\|_\infty$. Alors :

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \in \mathbb{N},\ \forall m \geqslant n \geqslant n_\varepsilon,\ \sup_{x \in X} \|f_m(x) - f_n(x)\| \leqslant \varepsilon$$

D'où :

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \in \mathbb{N},\ \forall m \geqslant n \geqslant n_\varepsilon,\ \forall x \in X,\ \|f_m(x) - f_n(x)\| \leqslant \varepsilon$$

Ainsi, pour tout x dans X, $(f_n(x))$ est une suite de Cauchy dans E complet, elle converge donc, mettons vers f.

(i) Montrons que f est bornée. $\varepsilon = 1$ donne en particulier n_1 tel que :

$$\forall n \geqslant n_1,\ \forall x \in X,\ \|f_n(x) - f_{n_1}(x)\| \leqslant 1$$

On fixe $x \in X$ et on a par passage à la limite : $\|f(x) - f_{n_1}(x)\| \leqslant 1$. Soit :

$$\forall x \in X,\ \|f(x)\| \leqslant 1 + \|f_{n_1}(x)\| \leqslant 1 + \|f_{n_1}\|_\infty$$

Donc f est bornée.

(ii) Montrons que (f_n) tend vers f pour $\| \ \|_\infty$. Soit $\varepsilon > 0$.

$$\exists n_\varepsilon \in \mathbb{N},\ \forall m \geqslant n \geqslant n_\varepsilon,\ \forall x \in X,\ \|f_m(x) - f_n(x)\| \leqslant \varepsilon$$

On fixe $x \in X$ et $n \geqslant n_\varepsilon$. Alors :

$$\forall m \geqslant n,\ \|f_m(x) - f_n(x)\| \leqslant \varepsilon$$

On fait tendre m vers $+\infty$ et il vient : $\|f(x) - f_n(x)\| \leqslant \varepsilon$. De là :

$$\forall n \geqslant n_\varepsilon,\ \forall x \in X,\ \|f(x) - f_n(x)\| \leqslant \varepsilon$$

Donc : $\forall n \leqslant n_\varepsilon,\ \|f - f_n\|_\infty \leqslant \varepsilon$.

Conclusion : Toute suite de Cauchy de $(\mathcal{B}(X,E), \| \ \|_\infty)$ converge dans $\mathcal{B}(X,E)$, donc l'espace $(\mathcal{B}(X,E), \| \ \|_\infty)$ est complet.

□

1.7.3.3 L'espace $(\mathcal{C}_b(X,E), \| \|_\infty)$ des applications continues et bornées

Soit X un espace topologique, $(E, \| \|)$ un espace de Banach. Si $\mathcal{C}(X,E)$ désigne l'ensemble des fonctions continues de X dans E, alors on définit l'ensemble des fonctions continues et bornées de X dans E par :

$$\mathcal{C}_b(X,E) = \mathcal{C}(X,E) \cap \mathcal{B}(X,E)$$

On a alors :

Théorème 1.7.19. $(\mathcal{C}_b(X,E), \| \|_\infty)$ *est complet.*

Démonstration : Comme $(\mathcal{C}_b(X,E), \| \|_\infty)$ est un sous-espace de $(\mathcal{B}(X,E), \| \|_\infty)$ qui est complet, il nous suffit de montrer que $(\mathcal{C}_b(X,E), \| \|_\infty)$ est fermé.

Soit (f_n) un suite d'éléments de $\mathcal{C}_b(X,E)$ convergeant pour $\| \|_\infty$ vers $f \in \mathcal{B}(X,E)$ (car $(\mathcal{B}(X,E), \| \|_\infty)$ est complet donc fermé). On a alors :

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \in \mathbb{N},\ \forall n \geqslant n_\varepsilon,\ \sup_{x \in X} \|f_n(x) - f(x)\| \leqslant \varepsilon$$

Donc, (f_n) converge uniformément vers f sur X donc f est continue. Ainsi, f est continue et bornée sur X donc est dans $\mathcal{C}_b(X,E)$ qui est donc fermé.

□

1.7.3.4 L'espace $(\ell^1(\mathbb{N},\mathbb{C}), \| \|_1)$ des séries absolument convergentes

On considère l'espace :

$$\ell^1(\mathbb{N},\mathbb{C}) = \{(a_n) \in \mathbb{C}^{\mathbb{N}} \mid \sum |a_n| converge\}$$

Evidemment, on pourrait prendre à la place de $\mathbb{C}$ n'importe quel espace de Banach, mais dans un souci de simplification des notations, on se place dans le cas complexe.

Lorsque $(a_n) \in \ell^1(\mathbb{N},\mathbb{C})$, on pose :

$$||(a_n)||_1 = \sum_{n=0}^{+\infty} |a_n|$$

Alors, clairement, $|| \ ||_1$ est une norme sur $\ell^1(\mathbb{N},\mathbb{C})$ (en passant par les sommes partielles) et on a le résultat :

Théorème 1.7.20. *L'espace* $(\ell^1(\mathbb{N},\mathbb{C}),\ || \ ||_1)$ *est un espace de Banach.*

Démonstration : Soit $s_p = ((a_n)_p) = (a_{n,p})$ une suite d'éléments de $\ell^1(\mathbb{N},\mathbb{C})$ qui est de Cauchy pour $|| \ ||_1$.

(i) Montrons que pour chaque n (on fixe n), $p \mapsto a_{n,p}$ converge. On a :

$$\forall (p,p') \in \mathbb{N},\ |a_{n,p} - a_{n,p'}| \leqslant \sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p'}| = ||s_p - s_{p'}||_1$$

Comme (s_p) est de Cauchy, $p \mapsto a_{n,p}$ l'est aussi ; comme $\mathbb{C}$ est complet, elle converge. On note alors a_n la limite de $p \mapsto a_{n,p}$.

(ii) Montrons que (a_n) est dans $\ell^1(\mathbb{N},\mathbb{C})$. Pour cela, on passe par les sommes finies. Soit $N \in \mathbb{N}$. Avec $\varepsilon = 1$, le fait que (s_p) est de Cauchy se traduit par :

$$\exists p_1 \in \mathbb{N},\ \forall p' \geqslant p \geqslant p_1,\ \sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p'}| \leqslant 1$$

En particulier :

$$\forall p \geqslant p_1,\ \sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p_1}| \leqslant 1$$

Soit :

$$\forall p \geqslant p_1,\ \sum_{n=0}^{+\infty} |a_{n,p}| \leqslant 1 + \sum_{n=0}^{+\infty} |a_{n,p_1}|$$

On note :

$$M = 1 + \sum_{n=0}^{+\infty} |a_{n,p_1}|$$

Il vient alors :

$$\forall p \geqslant p_1,\ \sum_{n=0}^{N} |a_{n,p}| \leqslant M$$

On peut alors, puisque l'on considère une somme finie, faire tendre p vers $+\infty$ pour obtenir enfin :

$$\sum_{n=0}^{N} |a_n| \leqslant M$$

Ainsi, $\sum |a_n|$ est une série à termes positifs dont les sommes partielles sont bornées, elle est donc convergente et : $(a_n) \in \ell^1(\mathbb{N},\mathbb{C})$.

(iii) On note : $s = (a_n)$. Montrons alors que (s_p) converge vers s pour $|| \ ||_1$. On va de nouveau passer par les sommes finies, soit donc $N \in \mathbb{N}$.
Soit $\varepsilon > 0$. Il existe alors $p_\varepsilon \in \mathbb{N}$ tel que :

$$\forall p' \geqslant p \geqslant p_\varepsilon,\ \sum_{n=0}^{+\infty} |a_{n,p'} - a_{n,p}| \leqslant \varepsilon$$

D'où :

$$\forall p' \geqslant p \geqslant p_\varepsilon,\ \sum_{n=0}^{N} |a_{n,p'} - a_{n,p}| \leqslant \varepsilon$$

De là, en faisant tendre p' vers $+\infty$:

$$\forall p \geqslant p_\varepsilon,\ \sum_{n=0}^{N} |a_{n,p} - a_n| \leqslant \varepsilon$$

On peut alors faire tendre N vers $+\infty$:

$$\forall p \geqslant p_\varepsilon,\ \sum_{n=0}^{+\infty} |a_{n,p} - a_n| \leqslant \varepsilon$$

On a donc : $\forall p \geqslant p_\varepsilon,\ ||s_p - s||_1 \leqslant \varepsilon$. Donc (s_p) converge vers s pour $|| \ ||_1$.

Conclusion : Toute suite de Cauchy d'éléments de $(\ell^1(\mathbb{N},\mathbb{C}), || \ ||_1)$ converge dans $\ell^1(\mathbb{N},\mathbb{C})$ donc $\ell^1(\mathbb{N},\mathbb{C})$ est complet.

□

1.7.3.5 L'espace $(\ell^q(\mathbb{N},\mathbb{C}), \| \|_q)$

On considère l'espace :

$$\ell^q(\mathbb{N},\mathbb{C}) = \{(a_n) \in \mathbb{C}^{\mathbb{N}} \mid \sum |a_n|^q converge\}$$

Evidemment, on pourrait prendre à la place de $\mathbb{C}$ n'importe quel espace de Banach, mais dans un souci de simplification des notations, encore une fois on se place dans le cas complexe.

Lorsque $(a_n) \in \ell^q(\mathbb{N},\mathbb{C})$, on pose :

$$||(a_n)||_q = \left(\sum_{n=0}^{+\infty} |a_n|^q\right)^{\frac{1}{q}}$$

Alors, clairement, $|| \; ||_q$ est une norme sur $\ell^q(\mathbb{N},\mathbb{C})$ (en passant par les sommes partielles) et on a le résultat :

Théorème 1.7.21. *L'espace $(\ell^q(\mathbb{N},\mathbb{C}),\ || \; ||_q)$ est un espace de Banach.*

Démonstration : Soit $s_p = ((a_n)_p) = (a_{n,p})$ une suite d'éléments de $\ell^q(\mathbb{N},\mathbb{C})$ qui est de Cauchy pour $|| \; ||_q$.

(i) Montrons que pour chaque n (on fixe n), $p \mapsto a_{n,p}$ converge. Tout d'abord, on a , pour n fixé et $p, p' \in \mathbb{N}$:

$$|a_{n,p} - a_{n,p'}|^q \leqslant \sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p'}|^q$$

Il vient :

$$\forall (p,p') \in \mathbb{N},\ |a_{n,p} - a_{n,p'}| \leqslant \left(\sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p'}|^q\right)^{\frac{1}{q}} = ||s_p - s_{p'}||_q$$

Comme (s_p) est de Cauchy, $p \mapsto a_{n,p}$ l'est aussi ; comme $\mathbb{C}$ est complet, elle converge. On note alors a_n la limite de $p \mapsto a_{n,p}$.

(ii) Montrons que (a_n) est dans $\ell^q(\mathbb{N},\mathbb{C})$. Pour cela, on passe par les sommes finies. Soit $N \in \mathbb{N}$. Avec $\varepsilon = 1$, le fait que (s_p) est de Cauchy se traduit par :

$$\exists p_1 \in \mathbb{N},\ \forall p' \geqslant p \geqslant p_1,\ \left(\sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p'}|^q\right)^{\frac{1}{q}} \leqslant 1$$

En particulier :

$$\forall p \geqslant p_1,\ \left(\sum_{n=0}^{+\infty} |a_{n,p} - a_{n,p_1}|^q\right)^{\frac{1}{q}} \leqslant 1$$

Soit :

$$\forall p \geqslant p_1,\ \left(\sum_{n=0}^{+\infty} |a_{n,p}|^q\right)^{\frac{1}{q}} \leqslant 1 + \left(\sum_{n=0}^{+\infty} |a_{n,p_1}|^q\right)^{\frac{1}{q}}.$$

On note alors :

$$M = 1 + \left(\sum_{n=0}^{+\infty} |a_{n,p_1}|^q\right)^{\frac{1}{q}}$$

Il vient alors :

$$\forall p \geqslant p_1, \ \left(\sum_{n=0}^{N} |a_{n,p}|^q\right)^{\frac{1}{q}} \leqslant M$$

On peut alors, puisque l'on considère une somme finie, faire tendre p vers $+\infty$ pour obtenir enfin :

$$\sum_{n=0}^{N} |a_n|^q \leqslant M^q$$

Ainsi, $\sum |a_n|^q$ est une série à termes positifs dont les sommes partielles sont bornées, elle est donc convergente et : $(a_n) \in \ell^q(\mathbb{N}, \mathbb{C})$.

(iii) On note : $s = (a_n)$. Montrons alors que (s_p) converge vers s pour $|| \ ||_q$. On va de nouveau passer par les sommes finies, soit donc $N \in \mathbb{N}$.
Soit $\varepsilon > 0$. Il existe alors $p_\varepsilon \in \mathbb{N}$ tel que :

$$\forall p' \geqslant p \geqslant p_\varepsilon, \ \left(\sum_{n=0}^{+\infty} |a_{n,p'} - a_{n,p}|^q\right)^{\frac{1}{q}} \leqslant \varepsilon$$

D'où :

$$\forall p' \geqslant p \geqslant p_\varepsilon, \ \left(\sum_{n=0}^{N} |a_{n,p'} - a_{n,p}|^q\right)^{\frac{1}{q}} \leqslant \varepsilon$$

De là, en faisant tendre p' vers $+\infty$:

$$\forall p \geqslant p_\varepsilon, \ \left(\sum_{n=0}^{N} |a_{n,p} - a_n|^q\right)^{\frac{1}{q}} \leqslant \varepsilon$$

On peut alors faire tendre N vers $+\infty$:

$$\forall p \geqslant p_\varepsilon, \ \left(\sum_{n=0}^{+\infty} |a_{n,p} - a_n|^q\right)^{\frac{1}{q}} \leqslant \varepsilon$$

On a donc : $\forall p \geqslant p_\varepsilon, \ ||s_p - s||_q \leqslant \varepsilon$. Donc (s_p) converge vers s pour $|| \ ||_q$.

Conclusion : Toute suite de Cauchy d'éléments de $(\ell^q(\mathbb{N}, \mathbb{C}), || \ ||_q)$ converge dans $\ell^q(\mathbb{N}, \mathbb{C})$ donc $\ell^q(\mathbb{N}, \mathbb{C})$ est complet.

□

Remarque : Evidemment, ce théorème inclus le paragraphe précédent, mais il est bon de voir la démonstration dans le cas particulier $q = 1$ pour mieux comprendre le cas général.

1.7.4 Applications de la complétude

1.7.4.1 Caractérisation de la complétude par les séries

Dans toute la suite on considère $(E, || \ ||_E)$ un espace vectoriel normé et $(u_n) \in E^{\mathbb{N}}$.
On commence par un petit rappel :

Définition 1.7.22. *La série $\sum u_n$ est dite absolument convergente lorsque $\sum ||u_n||_E$ converge.*

On a alors le théorème :

Théorème 1.7.23. *$(E, || \, ||_E)$ est complet si et seulement si toute série à coefficients dans E et absolument convergente, converge dans E.*

Démonstration : On part de $(E, || \, ||_E)$ complet. Si $\sum u_n$ est absolument convergente, alors $\sum ||u_n||$ vérifie le critère de Cauchy :

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \in \mathbb{N},\ \forall m > n \geqslant n_\varepsilon,\ \sum_{k=n+1}^{m} ||u_k|| \leqslant \varepsilon$$

A fortiori :

$$\forall \varepsilon > 0,\ \exists n_\varepsilon \in \mathbb{N},\ \forall m > n \geqslant n_\varepsilon,\ ||\sum_{k=n+1}^{m} u_k|| \leqslant \varepsilon$$

Donc, $n \mapsto \sum_{k=0}^{n} u_k$ est de Cauchy. Elle converge donc dans E complet.

Réciproquement, soit (x_n) une suite de Cauchy de E. On prend $\varepsilon = \frac{1}{2^n}$. Il existe alors $\varphi(n) \in \mathbb{N}$ tel que :

$$\forall (m,p) \in \mathbb{N}^2,\ (m \geqslant \varphi(n) \text{ et } p \geqslant \varphi(n)) \Rightarrow (||x_m - x_p|| \leqslant \frac{1}{2^n})$$

Quitte à extraire, on peut toujours faire en sorte que φ soit croissante. Soit alors : $u_n = x_{\varphi(n+1)} - x_{\varphi(n)}$. Il vient :

$$\forall n \in \mathbb{N},\ ||u_n|| = ||x_{\varphi(n+1)} - x_{\varphi(n)}|| \leqslant \frac{1}{2^n}$$

De là, par domination, $\sum u_n$ est absolument convergente et donc $\sum u_n$ est convergente. Mais alors, $(x_{\varphi(n)})$ est aussi convergente (c'est le terme général de la suite des sommes partielles de $\sum u_n$) et donc (x_n) est convergente car elle est de Cauchy et possède une valeur d'adhérence. Ainsi, toute suite de Cauchy de E est convergente, E est donc complet.

□

1.7.4.2 Critère de Cauchy pour les fonctions

On cherche à donner pour les fonctions l'équivalent du critère de convergence pour les suites qu'est le fait, dans un espace complet, d'être de Cauchy. On commence par définir le critère de Cauchy pour une fonction en un point a d'une partie d'un espace métrique.

Définition 1.7.24. *Soient (E,d) et (F,δ) deux espaces métriques, A une partie de E, $a \in \overline{A}$ et $f : A \to F$. On dit que f vérifie le critère de Cauchy en a si :*

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall (x,y) \in A^2,\ (d(a,x) \leqslant \alpha \text{ et } d(a,y) \leqslant \alpha) \Rightarrow \delta(f(x), f(y)) \leqslant \varepsilon$$

On a alors le résultat fondamental suivant :

Théorème 1.7.25 (Critère de Cauchy)**.** *Si f possède une limite en a, alors elle vérifie le critère de Cauchy en a. Réciproquement, si on suppose de plus que (F,δ) est complet, alors si f vérifie le critère de Cauchy en a, f possède une limite en a.*

Démonstration : Soit $\varepsilon > 0$. Puisque f possède une limite l en a, il existe $\alpha > 0$ tel que :

$$\forall (x,y) \in (B(a,\alpha) \cap A)^2,\ \delta(l, f(x)) \leqslant \varepsilon \text{ et } \delta(l, f(y)) \leqslant \varepsilon$$

De là :

$$\forall (x,y) \in (B(a,\alpha) \cap A)^2,\ \delta(f(x), f(y)) \leqslant 2\varepsilon$$

Et donc f vérifie le critère de Cauchy en a.

Réciproquement, soient $\varepsilon > 0$, un réel $\alpha > 0$ adapté à ε pour le critère de Cauchy en a et (u_n) une suite de points de A convergeant vers a. On a alors :

$$\exists n_\alpha \in \mathbb{N},\ \forall m, n \geqslant n_\alpha,\ d(a, u_m) \leqslant \alpha \text{ et } d(a, u_n) \leqslant \alpha$$

On a alors, comme f vérifie le critère de Cauchy en a :

$$\delta(f(u_n), f(u_m)) \leqslant \varepsilon$$

Ainsi, $(f(u_n))$ est de Cauchy dans F complet, elle converge donc. Alors, par caractérisation séquentielle de l'existence d'une limite, f possède une limite en a selon A.

□

Remarque : Comme pour les suites, l'énorme intérêt du critère de Cauchy est que l'on peut déterminer l'existence d'une limite sans pour autant la connaître.

On énonce le critère de Cauchy à l'infini dans le cas particulier où $E = \overline{\mathbb{R}}$ et $+\infty \in \overline{A}$. On garde les notations précédentes.

Corollaire 1.7.26. *Si (F, δ) est complet, alors on a équivalence entre :*
(i) f possède une limite en $+\infty$.
(ii) $\forall \varepsilon > 0,\ \exists M > 0,\ \forall (x,y) \in A^2,\ (x \geqslant M \text{ et } y \geqslant M) \Rightarrow \delta(f(x), f(y)) \leqslant \varepsilon$

Démonstration : C'est une application du Critère de Cauchy.

□

Enfin, on s'intéresse au cas des intégrales :

Corollaire 1.7.27. *Pour u mesurable,*

$$\int_a^{+\infty} u \text{ converge} \Leftrightarrow \forall \varepsilon > 0,\ \exists M > 0,\ \forall y \geqslant x \geqslant M,\ \left| \int_x^y u(t)dt \right| \leqslant \varepsilon$$

Démonstration : On applique le Critère de Cauchy en $+\infty$ à :

$$f(x) = \int_a^x u(t)dt$$

□

1.7.4.3 Théorème du point fixe de Picard

Définition 1.7.28. *On dit que f est strictement contractante ou est une contraction stricte lorsqu'il existe au moins un réel $k \in [0,1[$ tel que f soit k-lipschitzienne.*

On a alors :

Théorème 1.7.29 (Point fixe de Picard). *Soit (E,d) un espace métrique, soit A une partie non vide de E et soit f une application de A dans E. Si on a :*
(i) $f(A) \subset A$.
(ii) f est une contraction stricte.
(iii) (A, d_A) est complète.
Alors f possède un point fixe unique $l \in A$ et pour tout $u_o \in A$, la suite définie par la donnée de u_o et la relation de récurrence $u_{n+1} = f(u_n)$ converge vers l.

Démonstration : Soit $k \in [0,1[$ un rapport de Lipschitz de f. On a :

$$\forall n \in \mathbb{N},\ d(u_{n+1}, u_n) = d(f(u_n), f(u_{n-1})) \leqslant kd(u_n, u_{n-1}) \leqslant \ldots \leqslant k^n d(u_1, u_o)$$

Soit $n \in \mathbb{N}$ et $p \geqslant 2$:

$$\begin{array}{rcl} d(u_{n+p}, u_n) & \leqslant & d(u_{n+p}, u_{n+p-1}) + \ldots + d(u_{n+1}, u_n) \\ & \leqslant & (k^{n+p-1} + \ldots + k^n) d(u_1, u_o) \\ & = & k^n \left(\dfrac{1-k^p}{1-k} \right) d(u_1, u_o) \\ & \leqslant & \dfrac{k^n}{1-k} d(u_1, u_o) \end{array}$$

Comme $k \in [0,1[$, (k^n) tend vers 0, donc (u_n) est de Cauchy. De là, comme (A, d_A) est complet, (u_n) converge, mettons vers $l \in A$. Puis, par continuité de f, on a : $f(l) = l$.

Montrons l'unicité de l. Si l' est un autre point fixe :

$$0 \leqslant d(l, l') = d(f(l), f(l')) \leqslant kd(l, l')$$

avec $k < 1$. Donc : $d(l, l') = 0$ et donc : $l = l'$.

□

Ce théorème est un résultat fondamental en analyse, il intervient entre autre dans la démonstration du théorème de Cauchy-Lipschitz dans la théorie des équations différentielles, et surtout il intervient dans celle du théorème d'inversion locale et de son corollaire, le théorème des fonctions implicites, base de la géométrie différentielle. C'est en partie du fait de ce théorème que l'analyse se fait naturellement dans des espaces complets.

1.7.4.4 Théorème de Baire

Théorème 1.7.30 (de Baire). *Soit (E,d) un espace métrique complet et soit (Ω_n) une suite d'ouverts denses de E. Alors, $A = \bigcap_{n \in \mathbb{N}} \Omega_n$ est une partie dense de E.*

Démonstration. Ceci revient à démontrer que si on prend $B(x,r)$ une boule ouverte ($r > 0$) quelconque, alors $B(x,r) \cap A \neq \varnothing$.

La preuve est constructive : remarquons que comme Ω_0 est dense, elle rencontre cette boule : $\overline{\Omega_0} = E \Rightarrow \Omega_0 \cap B(x,r) \neq \varnothing$.

On choisit donc $x_0 \in \Omega_0 \cap B(x,r)$ et $r_0 > 0$ tels que $B_f(x_0,r_0) \subset \Omega_0 \cap B(x,r)$. On impose de plus $r_0 < 1$. $\overline{\Omega_1} = E \Rightarrow \Omega_1 \cap B(x_0,r_0)$ est un ouvert (intersection de deux ouverts) non vide (même argument qu'au début). On choisit donc x_1 dans l'intersection et $r_1 > 0$ en imposant $r_1 < 1/2$ de telle façon que $B_f(x_1,r_1) \subset \Omega_1 \cap B(x_0,r_0)$. On suppose construits $B(x_0,r_0),\ldots,B(x_n,r_n)$ avec $r_i < 1/2^i$ et $B_f(x_i,r_i) \subset \Omega_{i-1} \cap B(x_{i-1},r_{i-1}) \qquad i \geqslant 1$. $\overline{\Omega_{n+1}} = E \Rightarrow \Omega_{n+1} \cap B(x_n,r_n)$ est un ouvert non vide. On choisit x_{n+1}, $r_{n+1} > 0$ avec $r_{n+1} < 1/2^{n+1}$ tels que $B_f(x_{n+1},r_{n+1}) \subset \Omega_n \cap B(x_n,r_n)$ et on a enfin construits correctement par récurrence la suite des $B_f(x_n,r_n) = F_n$. Ces F_n forment une suite décroissante pour l'inclusion de fermés non vides. De plus, leur diamètre tend vers 0. On applique donc le lemme des fermés emboîtés et on en déduit $\bigcap_{n\in\mathbb{N}} F_n = \{a\}$. Pour tout n, $a \in B_f(x_n,r_n) \subset \Omega_n$: $a \in \bigcap_{n\in\mathbb{N}} \Omega_n = A$. Pour tout n, $a \in B_f(x_n,r_n) \subset B(x,r)$: $a \in B(x,r)$. C'est donc que $A \cap B(x,r) \neq \varnothing$, ce qu'il fallait démontrer. □

Il y a une formulation de ce théorème en terme de fermés que l'on obtient par passage au complémentaire.

Théorème 1.7.31. *Soit (E,d) un espace métrique complet et soit (F_n) une suite de fermés d'intérieurs vides de E. Alors, $F = \bigcup_{n\in\mathbb{N}} F_n$ est d'intérieur vide dans E.*

1.7.4.5 Théorème de prolongement uniformément continu et théorème de Tietze

On commence cette section sur les résultats de prolongements en donnant un premier résultat de prolongement continu sur une partie dense.

Théorème 1.7.32. *Soient (E,d) et (F,δ) deux espaces métriques, A une partie dense de E. Soit $f : A \to F$. On suppose que f est continue et que :*

$$\forall x \in E\backslash A,\ \lim_{y\to x,\ y\in A} f(y) \text{ existe.}$$

Alors, il existe une unique fonction $g : E \to F$ continue et telle que sa restriction à A soit égale à f.

Démonstration : On définit g de la manière suivante : $\forall x \in A,\ g(x) = f(x)$ et :

$$\forall x \in E\backslash A,\ g(x) = \lim_{y\to x,\ y\in A} f(y)$$

Montrons que g ainsi définie est continue sur E. Soit $x \in E$ et (x_n) une suite de points de E qui converge vers x. On a :

$$\forall n \in \mathbb{N},\ \lim_{y\to x_n,\ y\in A} f(y) = g(x_n)$$

Dans la définition métrique de la limite, on prend $\varepsilon = \frac{1}{n}$. D'où l'existence d'un réel $\eta_n > 0$ tel que :

$$\forall y \in A,\ d(x_n,y) \leqslant \eta_n \Rightarrow \delta(g(x_n),f(y)) \leqslant \frac{1}{n}$$

Alors, par densité de A dans E : $B(x_n,\frac{1}{n}) \cap A \neq \varnothing$. Soit alors $y_n \in B(x_n,\frac{1}{n}) \cap B(x_n,\eta_n) \cap A$. On a donc :

$$\forall n \in \mathbb{N}^*,\ \exists y_n \in A,\ d(x_n,y_n) \leqslant \frac{1}{n} \text{ et } \delta(g(x_n),f(y_n)) < \frac{1}{n}$$

De là, la relation $d(x, y_n) \leqslant d(x, x_n) + d(x_n, y_n) \leqslant \frac{1}{n} + d(x, x_n)$ montre que (y_n) tend vers x et donc que $(f(y_n))$ converge vers $g(x)$. Or, on a de plus les inégalités :

$$\delta(g(x_n), g(x)) \leqslant \delta(g(x_n), f(y_n)) + \delta(f(y_n), g(x)) \leqslant \frac{1}{n} + \delta(f(y_n), g(x))$$

Il vient donc que $(g(x_n))$ converge vers $g(x)$. Ceci étant vrai pour toute suite (x_n) de E tendant vers x, par caractérisation séquentielle de la continuité, g est continue en x. Ceci étant vrai pour tout x dans E, g est continue sur E.

Unicité : Si g et h sont deux fonctions continues sur E qui coïncident avec f sur A, elles coïncident sur A qui est dense dans E, donc par le principe de prolongement des identités, $g = h$.

□

On va maintenant étendre ce résultat au cas des fonctions uniformément continues.

Théorème 1.7.33. *Soient (E, d) et (F, δ) deux espaces métriques, A une partie dense de E et f une application de A vers F. Si F est complet, il existe une application uniformément continue g de E vers F qui prolonge f sur E.*

Démonstration : Unicité : Si h est un autre prolongement continue de f à E, h et g sont deux applications continues qui coïncident sur la partie dense A, donc $h = g$.

Existence : Soit $a \in E$. Alors, $a \in \overline{A} = E$ et d'autre part, si l'on se donne $\varepsilon > 0$, il existe par uniforme continuité de f, un réel $\alpha > 0$ tel que :

$$\forall (x, y) \in A^2,\ d(x, y) \leqslant \alpha \Rightarrow \delta(f(x), f(y)) \leqslant \varepsilon$$

A fortiori :

$$\forall (x, y) \in (A \cap B(a, \frac{\alpha}{2}))^2,\ \delta(f(x), f(y)) \leqslant \varepsilon$$

D'où, par le critère de Cauchy en a, f possède une limite en a, soit $g(a)$. Montrons alors que g ainsi définie est uniformément continue.
Soit $(a, b) \in E^2$ avec $d(a, b) < \frac{\alpha}{3}$, puis (x_n) et (y_n) deux suites d'éléments de A qui convergent respectivement vers a et b et avec : $d(x_n, a) < \frac{\alpha}{3}$ et $d(y_n, b) < \frac{\alpha}{3}$. On a alors par l'inégalité triangulaire : $d(x_n, y_n) < \alpha$. On a alors puisque α est un module d'uniforme continuité de f pour ε : $\delta(f(x_n), f(y_n)) \leqslant \varepsilon$. Or, par existence d'une limite en a et en b pour f, on a que $(f(x_n))$ converge vers $g(a)$ et que $(f(y_n))$ converge vers $g(b)$. D'où, par passage à la limite :

$$\delta(g(a), g(b)) \leqslant \varepsilon$$

Donc, g est uniformément continue sur E et prolonge bien f.

□

On va maintenant donner deux exemples d'utilisation de ce résultat.

Exemple 1 : Soit (E, d) un espace métrique complet et $f : [a, b[\to E$ une application uniformément continue. On suppose de plus que $b < +\infty$. Alors f possède une limite en b et un prolongement uniformément continue en b.

Exemple 2 : Soit $\mathcal{E} = Esc([a,b],E)$ l'espace vectoriel des fonctions en escalier de $[a,b]$ dans E espace métrique complet. Soit alors $\mathcal{C}$ l'espace des fonctions continues par morceaux de $[a,b]$ dans E. Soit alors :

$$\varphi : \begin{pmatrix} \mathcal{E} & \to & E \\ u & \mapsto & \int_a^b u(t)dt \end{pmatrix}$$

Alors, on a :

$$||\varphi(u)||_E \leqslant (b-a) \sup_{t\in[a,b]} ||u(t)|| = (b-a)||u||_\infty$$

donc φ est une forme linéaire continue donc uniformément continue sur $\mathcal{E}$. Alors φ possède un prolongement uniformément continue à $\overline{\mathcal{E}}^{|| \, ||_\infty}$ dans $\mathcal{C}$. Or : $\overline{\mathcal{E}}^{|| \, ||_\infty} = \mathcal{C}$. Donc φ possède un prolongement uniformément continue à $\mathcal{C}$. Et voilà, on vient de reconstruire l'intégrale de Riemann.

On va maintenant énoncer un théorème de prolongement par continuité plus général que les précédents, le théorème de Tietze, énoncé ici sous sa forme précisée. On indiquera ensuite la forme non précisée qui évidemment se déduit de la précisée.

Théorème 1.7.34 (Tietze-Urysohn)**.** *Soient (E,d) un espace métrique, A un fermé de E et f une application continue et bornée de A dans $\mathbb{R}$. Alors, il existe une fonction continue $g : E \to \mathbb{R}$ qui prolonge f et telle que :*

$$\sup_{x\in E} g(x) = \sup_{x\in A} f(x) \ \ et \ \ \inf_{x\in E} g(x) = \inf_{x\in A} f(x)$$

Démonstration : Si f est constante, le résultat est évident, sinon, quitte à remplacer f par $\alpha f + \beta$ avec $\alpha \neq 0$ et β des réels bien choisis, on peut supposer $\inf_{y\in A} f(y) = 1$ et $\sup_{y\in A} f(y) = 2$. On considère alors la fonction $g : E \to \mathbb{R}$ définie par : $g(x) = f(x)$ si $x \in A$ et :

$$\forall x \in E\backslash A, \ g(x) = \frac{1}{d(x,A)} \inf_{y\in A}(f(y)d(x,A))$$

Cette dernière expression est bien définie car A étant fermé, on a $d(x,A) = 0$ si et seulement si $x \in A$. Les inégalités $1 \leqslant f(x) \leqslant 2$ pour tout x dans A montrent que $1 \leqslant g(x) \leqslant 2$ pour tout x dans E. Comme la restriction de g à A est f, on a même : $\inf_{x\in E} g(x) = 1$ et $\sup_{x\in E} g(x) = 2$. Il nous reste à montrer la continuité de g.

Soit x un point de E. Si $x \in \overset{\circ}{A}$, comme f est continue sur A, g est continue en x. Supposons que $x \in E\backslash A$. Sur l'ouvert $E\backslash A$, on peut écrire :

$$g(z) = \frac{h(z)}{d(z,A)} \text{ où } h(z) = \inf_{y\in A}(f(y)d(z,y))$$

L'application $z \mapsto d(z,A)$ étant continue (voir annexe sur la distance à une partie), il nous suffit de montrer la continuité de h sur $E\backslash A$. Soit $r > 0$ tel que $B(x,r) \subset E\backslash A$. Pour tout $z \in B(x,r)$ et pour tout $y \in A$, on a :

$$h(x) \leqslant f(y)d(x,y) \leqslant f(y)(d(x,z) + d(z,y)) \leqslant 2d(x,z) + f(y)d(z,y)$$

et ceci étant vrai pour tout $y \in A$, on en déduit, en ne considérant que les expressions des extrémités et en prenant la borne inférieure sur les $y \in A$ que : $h(x) \leqslant h(z) + 2d(x,z)$. De même, $h(z) \leqslant h(x) + 2d(z,x)$, et donc finalement :

$$|h(x) - h(z)| \leqslant 2d(x,z)$$

Ceci prouve la continuité de h en $x \in E \backslash A$.
Il nous reste à étudier le cas où $x \in Fr(A)$. Soit $\varepsilon > 0$. Comme f est continue en $x \in A$:

$$\exists r > 0,\ \forall y \in A \cap B(x, r),\ |f(y) - f(x)| \leqslant \varepsilon$$

Si $z \in E \backslash A$ et $d(x, z) \leqslant \frac{r}{4}$, on a, en notant $C = A \cap B(x, r)$:

$$\forall y \in A \backslash C,\ d(z, y) \geqslant d(x, y) - d(x, z) \geqslant \frac{3r}{4}$$

et :

$$\inf_{y \in A \backslash C} f(y) d(z, y) \geqslant \frac{3r}{4}$$

D'autre part, $f(x)d(x, z) \leqslant 2d(x, z) \leqslant \frac{r}{2}$ et donc :

$$\inf_{y \in A} f(y) d(z, y) = \inf_{y \in C} f(y) d(z, y)$$

Et comme $f(x) - \varepsilon \leqslant f(y) \leqslant f(x) + \varepsilon$ pour tout $y \in C$, on en déduit :

$$(f(x) - \varepsilon) d(z, A) \leqslant \inf_{y \in A} (f(y) d(z, y)) \leqslant (f(x) + \varepsilon) d(z, A)$$

Donc, $g(x) - \varepsilon \leqslant g(z) \leqslant g(x) + \varepsilon$ pour tout $z \in E \backslash A$ tel que $d(x, z) \leqslant \frac{r}{4}$. Ceci reste vrai si $z \in A$ et $d(x, z) \leqslant \frac{r}{4}$ car c'est vrai pour f. Ainsi, g est continue en $x \in Fr(A)$.
Donc g en continue en tout x dans E, d'où le résultat voulu.

$\square$

Remarque : On appelle ce théorème forme précisée car le résultat inclus un résultat sur la coïncidence des inf et des sup.

On a la forme générale de ce résultat que nous ne démontrerons pas :

Théorème 1.7.35 (Tietze)**.** *Soient (E, d) un espace métrique, A un fermé de E et $f : A \to \mathbb{R}$ continue. Alors, f admet un prolongement continu $g : E \to \mathbb{R}$.*

Chapitre 2

Espaces de Hilbert

La notion d'espace de Hilbert constitue une généralisation naturelle de la notion d'espace *euclidien* (espace vectoriel réel de dimension finie muni d'une forme bilinéaire symétrique définie positive). Expliquons cela à l'aide d'un exemple. Le produit scalaire usuel sur $\mathbb{R}^2$ est défini par $\big((x_1, x_2) \,|\, (y_1, y_2)\big) = x_1 y_1 + x_2 y_2$. Si l'on voulait définir un produit scalaire sur l'espace vectoriel $\mathbb{R}^{\mathbb{N}} = \{(x_k)_{k\in\mathbb{N}} \;:\; x_k \in \mathbb{R}\}$ des suites quelconques de nombres réels, on devrait donc naturellement poser

$$\big((x_k)_{k\in\mathbb{N}} \,|\, (y_k)_{k\in\mathbb{N}}\big) = \sum_{k=0}^{+\infty} x_k y_k. \tag{2.1}$$

Mais on constate que cela ne fonctionne pas car la série de terme général $x_n y_n$ n'est en général pas convergente. Une première idée pour remédier à cela est de se limiter au sous-espace $\mathbb{R}^{(\mathbb{N})} = \{(x_k)_{k\in\mathbb{N}} \,|\, \exists\, k_0 \in \mathbb{N} \text{ tel que } x_k = 0 \text{ pour } k > k_0\} \subset \mathbb{R}^{\mathbb{N}}$ des suites presque nulles. La somme $\sum_{k=0}^{+\infty} x_k y_k$ est alors toujours une somme finie et l'expression (2.1) définit une forme bilinéaire symétrique définie positive sur $\mathbb{R}^{(\mathbb{N})}$. La norme associée est

$$\|\mathbf{x}\| = \sqrt{\sum_{k=0}^{+\infty} x_k^2} \quad \text{(somme finie)}. \tag{2.2}$$

Mais il se produit alors le phénomène suivant : la suite $(\mathbf{x}^{(n)})_{n\in\mathbb{N}}$ d'éléments de $\mathbb{R}^{(\mathbb{N})}$ définie par $\mathbf{x}^{(n)} = (1, \frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^n}, 0, \ldots)$ ne converge pas dans $\mathbb{R}^{(\mathbb{N})}$, sinon sa limite serait $y = (1, \frac{1}{2}, \ldots, \frac{1}{2^n}, \frac{1}{2^{n+1}}, \ldots)$ et un tel y n'est pas élément de $\mathbb{R}^{(\mathbb{N})}$. La suite $(\mathbf{x}^{(n)})_{n\in\mathbb{N}}$ est pourtant de Cauchy car, pour tous $n, p \in \mathbb{N}, p \geqslant 1$, on a $\|\mathbf{x}^{(n)} - \mathbf{x}^{(n+p)}\| = \sqrt{\sum_{k=n+1}^{n+p} \left(\frac{1}{2^k}\right)^2} \leqslant \sqrt{\sum_{k=n+1}^{+\infty} \frac{1}{4^k}} = \frac{1}{\sqrt{3.2^n}}$ qui tend bien vers 0 quand n tend vers l'infini. Autrement dit, *l'espace* $\mathbb{R}^{(\mathbb{N})}$ *n'est pas complet pour la norme associée au produit scalaire (2.1) dont nous l'avons muni.* C'est problématique car, comme nous le verrons plus loin, si l'on a un espace muni d'un produit scalaire, il est bon que cet espace soit complet si l'on veut pouvoir disposer d'une notion de projection orthogonale sur l'un de ses sous-espace vectoriel fermé (ceci constituant une autre généralisation souhaitable lors du passage de la dimension finie à la dimension infinie). Il nous faut donc *compléter* l'espace $(\mathbb{R}^{(\mathbb{N})}, \|.\|)$ construit ci-dessus. Pour cela, on peut par exemple remarquer que le plus grand sous-espace vectoriel de $\mathbb{R}^{\mathbb{N}}$ sur lequel l'expression (2.2) définit une norme est le sous-espace $\ell^2(\mathbb{N}, \mathbb{R}) = \{(x_k)_{k\in\mathbb{N}} \,|\, \sum_{k=0}^{+\infty} x_k^2 < +\infty\}$ des suites de carré sommable. Nous verrons plus loin que cet espace vectoriel normé est complet.

Ainsi, *l'espace* $\ell^2(\mathbb{N}, \mathbb{R})$, muni du produit scalaire $\big((x_k)_{k\in\mathbb{N}} \,|\, (y_k)_{k\in\mathbb{N}}\big) = \sum_{k=0}^{+\infty} x_k y_k$ et de la norme associée, *constitue la généralisation naturelle des espaces euclidiens* $\mathbb{R}^n$ *munis de leur produit scalaire* *usuel*. Cet exemple conduit à la notion générale d'espace de Hilbert (un espace vectoriel réel ou complexe muni d'un produit scalaire tel que la norme associée en fasse un espace complet), que nous présentons dans ce chapitre.

Dans toute la suite, on désigne par $\mathbb{K}$ un corps qui est soit le corps $\mathbb{R}$ des nombres réels soit le corps $\mathbb{C}$ des nombres complexes. Pour $\alpha \in \mathbb{K}$, on pose $\overline{\alpha} = \alpha$ si $\mathbb{K} = \mathbb{R}$ et $\overline{\alpha}$ = le complexe conjugué de α si $\mathbb{K} = \mathbb{C}$. On désigne par $\mathcal{H}$ un espace vectoriel sur $\mathbb{K}$. Si $\alpha \in \mathbb{K}$, la notation $\alpha \geqslant 0$ signifie que α est positif si $\mathbb{K} = \mathbb{R}$ et que α est un réel positif si $\mathbb{K} = \mathbb{C}$.

2.1 Espaces préhilbertiens et espaces de Hilbert

2.1.1 Produit scalaire et produit scalaire hermitien

Définition 2.1.1. *Un produit scalaire (hermitien) est une application de $\mathcal{H} \times \mathcal{H} \to \mathbb{C}$, $(x, y) \mapsto (x|y)$ vérifiant :*

(i) $\forall x, y, z \in \mathcal{H}$, $\forall \lambda, \mu \in \mathbb{K}$, $(\lambda x + \mu y|z) = \lambda(x|z) + \mu(y|z)$.

(ii) $\forall x, y \in \mathcal{H}$, $(x|y) = \overline{(y|x)}$.

(iii) $\forall x \in \mathcal{H}$, $x \neq 0 \Rightarrow (x|x) > 0$.

Vocabulaire. Une application de $\mathcal{H} \times \mathcal{H} \to \mathbb{C}$, $(x, y) \mapsto (x|y)$ vérifiant (i) et (ii) est aussi appelée *forme hermitienne*. Si elle vérifie (iii) on parle aussi de forme hermitienne définie positive.
Lorsque $\mathbb{K} = \mathbb{R}$ on parlera tout simplement de produit scalaire, lorsque $\mathbb{K} = \mathbb{C}$ on parlera de produit scalaire hermitien. Par abus de langage il nous arrivera souvent de parler de produit scalaire tout simplement, même lorsque l'on est sur $\mathbb{C}$.

Remarque. En combinant (i) et (ii), on a aussi la propriété suivante :

(i') $\forall x, y, z \in \mathcal{H}$, $\forall \lambda, \mu \in \mathbb{K}$, $(x|\lambda y + \mu z) = \overline{\lambda}(x|y) + \overline{\mu}(x|z)$.

Par ailleurs, sans conjugués, on ne pourrait avoir (iii) dans le cas où $\mathbb{K} = \mathbb{C}$.

Notation. Nous noterons $\|.\|$ l'application de $\mathcal{H}$ dans $\mathbb{K}$ définie par $\|x\| = \sqrt{(x \mid x)} \in [0, +\infty[$ et nous appellerons $\|x\|$ la *norme* du vecteur du vecteur x.

Exemples.

1. $\mathbb{K}^n$ muni de la forme hermitienne définie positive $(x \mid y) = \sum_{k=1}^{n} x_k \overline{y_k}$.
2. $\mathbb{K}^{(\mathbb{N})} = \{(x_k)_{k \in \mathbb{N}} \mid \exists\, k_0 \in \mathbb{N}$ tel que $x_k = 0$ pour $k > k_0\}$ (l'espace vectoriel des suites presque nulles d'éléments de $\mathbb{K}$) muni de la forme hermitienne définie positive
$$\big((x_k)_{k \in \mathbb{N}} \mid (y_k)_{k \in \mathbb{N}}\big) = \sum_{k=0}^{+\infty} x_k \overline{y_k} \text{ (somme finie).}$$
3. L'application
$$(f, g) \mapsto (f \mid g) = \int_X f(x)\overline{g}(x)\mathrm{d}x$$
est une forme hermitienne définie positive sur l'espace vectoriel
$$L^2(X, \mathbb{K}) = \left\{ f : X \to \mathbb{C} \;\middle|\; \int_X |f|^2 \mathrm{d}x < +\infty \right\}$$
des fonctions de carré intégrable définies sur X. En effet, l'inégalité $|f\overline{g}| \leqslant \frac{1}{2}(|f|^2 + |g|^2)$ garantit l'intégrabilité de la fonction $f\overline{g}$, puisque $|f|^2$ et $|g|^2$ sont intégrables. De plus, le caractère hermitien positif de $(. \mid .)$ se vérifie immédiatement. Enfin, si $\int_X |f|^2 \mathrm{d}x = 0$ alors, $|f|^2$étant positive, elle est nulle presque partout donc $f = 0$ presque partout. Autrement dit, $f = 0$ dans $L^2(X, \mathbb{K})$ et $(. \mid .)$ est donc définie positive.

4. Soit I un ensemble quelconque, non vide, et soit
$$\ell^2(I,\mathbb{K}) = \left\{ \mathbf{x} = (x_i)_{i\in I} \mid x_i \in \mathbb{K} \text{ pour tout i et } \sum_{i\in I} |x_i|^2 < +\infty \right\}$$
l'espace vectoriel des familles de carré sommable d'éléments de $\mathbb{K}$. Alors, l'application
$$(x,y) \mapsto (\mathbf{x} \mid \mathbf{y}) = \sum_{i\in I} x_i\overline{y_i} \in \mathbb{K}$$
est une forme hermitienne définie positive sur $\ell^2(I,\mathbb{K})$. En effet, puisque $|x_i\overline{y_i}| \leqslant \frac{1}{2}(|x_i|^2 + |y_i|^2)$, la famille à termes positifs $(|x_iy_i|)_{i\in I}$ est sommable car les familles $(|x_i|^2)_{i\in I}$ et $(|y_i|^2)_{i\in I}$ sont sommables. Par conséquent, la famille $(x_i\overline{y_i})_{i\in I}$ est absolument sommable donc sommable et la quantité $(\mathbf{x} \mid \mathbf{y})$ est donc bien définie. Le caractère hermitien positif se vérifie immédiatement. Enfin, la famille sommable $(|x_i|^2)_{i\in I}$étant à termes positifs, sa somme est nulle si et seulement si tous ses termes sont nuls et $(.\,|\,.)$ est donc définie positive.
5. L'application $(A,B) \mapsto \mathrm{Tr}(A^t\overline{B})$ est un produit scalaire hermitien sur $\mathcal{M}_\mathrm{n}(\mathbb{C})$.

Définition 2.1.2. *On appelle espace préhilbertien (réel ou complexe) tout couple $(\mathcal{H},(.\,|\,.))$ où $\mathcal{H}$ est une espace vectoriel et $(.\,|\,.)$ est un produit scalaire sur $\mathcal{H}$.*

2.1.2 Inégalité de Cauchy-Schwarz et norme

Nous allons maintenant démontrer l'une des inégalités les plus importantes en analyse, l'inégalité de Cauchy-Schwarz.

Proposition 2.1.3 (Inégalité de Cauchy-Schwarz). *Si $(.\,|\,.)$ est une forme hermitienne positive sur $\mathcal{H}$, on a, pour tous $x,y \in \mathcal{H}$*
$$|(x \mid y)|^2 \leqslant \|x\|^2\,\|y\|^2\,.$$
L'inégalité ci-dessus, dite inégalité de Cauchy-Schwarz, est uneégalité si et seulement s'il existe $\alpha,\beta \in \mathbb{K}$ non tous deux nuls tels que $\|\alpha x + \beta y\|^2 = 0$.

Démonstration : Fixons $x,y \in \mathcal{H}$. La forme hermitienne $(.\,|\,.)$étant positive, on a, pour tout $\lambda \in \mathbb{K}$:
$$0 \leqslant (x-\lambda y \mid x-\lambda y) = (x \mid x) - \lambda(y \mid x) - \overline{\lambda}(x \mid y) + |\lambda|^2(y \mid y).$$
Écrivons $(y \mid x) = be^{i\theta}$ avec $b \geqslant 0$ et $\theta \in \mathbb{R}$ et posons $\lambda = te^{-i\theta}$ avec $t \in \mathbb{R}$ (si $\mathbb{K} = \mathbb{R}$, on a $\theta = 0$ donc un tel λ est bien dans $\mathbb{R}$).

L'inégalité ci-dessus devient
$$\forall t \in \mathbb{R}, (y \mid y)t^2 - 2bt + (x \mid x) \geqslant 0.$$
Notons $P(t) = (y \mid y)t^2 - 2bt + (x,x)$. Puisque $(y \mid y) \in \mathbb{R}$ et $(x \mid x) \in \mathbb{R}$, P est un polynôme à coefficients réels de degré $\leqslant 2$ et l'inégalité ci-dessus montre que P est positif. P est donc soit constant, soit de degré 2.

Premier cas : *deg* $P = 0$. Dans ce cas, on a $(y \mid y) = 0$ et $b = 0$. Donc $(x \mid y) = be^{-i\theta} = 0$ et l'inégalité de Cauchy-Schwarz est simplement l'égalité $0 = 0$ et est donc vraie. De plus, on a $\|0 \times x - y\|^2 = \|y\|^2 = (y \mid y) = 0$.

Second cas : *deg* $P = 2$. Alors le discriminant $\Delta = 4b^2 - 4(x \mid x)(y \mid y)$ de P est $\leqslant 0$, ce qui entraîne l'inégalité de Cauchy-Schwarz puisque $b^2 = |(x \mid y)|^2$. En outre, cette inégalité est uneégalité si et seulement si $\Delta = 0$, ce qui signifie qu'il existe $t_0 \in \mathbb{R}$ tel que $P(t_0) = 0$ soit $\|x - t_0 e^{-i\theta} y\|^2 = 0$.

Réciproquement, supposons $\|\alpha x + \beta y\|^2 = 0$ avec par exemple $\alpha \neq 0$. D'après l'inégalité de Cauchy-Schwarz, on a alors $(\alpha x + \beta y \,|\, y) = 0$ et $(\alpha x + \beta y \,|\, x) = 0$, soit $\alpha(x \,|\, y) + \beta(y \,|\, y) = 0$ et $\alpha(x \,|\, x) + \beta(y \,|\, x) = 0$.

Premier cas : $\beta = 0$. Alors $(x \,|\, y) = 0$ et $(x \,|\, x) = 0$, donc l'inégalité de Cauchy-Schwarz est une égalité.

Second cas : $\beta \neq 0$. Alors $(x \,|\, y) = -\frac{\beta}{\alpha}(y|y)$ et $\overline{(x \,|\, y)} = (y \,|\, x) = -\frac{\alpha}{\beta}(x \,|\, x)$, donc $|(x \,|\, y)|^2 = (x \,|\, x)(y \,|\, y) = \|x\|^2\|y\|^2$ et l'inégalité de Cauchy-Schwarz est encore une égalité.

□

On en déduit la proposition importante suivante.

Proposition 2.1.4. *Si $(.\,|\,.)$ est un produit scalaire sur $\mathcal{H}$, l'application*

$$\|.\| : \begin{array}{ccc} \mathcal{H} & \longrightarrow & [0, +\infty[\\ x & \longmapsto & \sqrt{(x \,|\, x)} \end{array}$$

est une norme sur $\mathcal{H}$.

Démonstration : Montrons l'inégalité triangulaire (le reste est une simple vérification). On a $\|x + y\|^2 = (x + y \,|\, x + y) = \|x\|^2 + 2\mathrm{Re}\,(x \,|\, y) + \|y\|^2$. D'après Cauchy-Schwarz, $\mathrm{Re}\,(x \,|\, y) \leqslant |(x \,|\, y)| \leqslant \|x\|\|y\|$, donc $\|x + y\|^2 \leqslant (\|x\| + \|y\|)^2$.

□

Cette proposition est très importante car munissant $\mathcal{H}$ d'une norme naturellement associée au produit scalaire, nous définissons ainsi une topologie particulière sur $(\mathcal{H}, (.\,|\,.))$, celle associée à son produit scalaire. Elle conduira plus tard à la définition d'espace de Hilbert.

2.1.3 Identités de polarisation

On fixe $\big(\mathcal{H}, (.\,|\,.)\big)$ un espace préhilbertien sur $\mathbb{K}$.

Proposition 2.1.5 (Identités de polarisation)**.** *Soient $x, y \in \mathcal{H}$. Si $\mathbb{K} = \mathbb{R}$, alors*

$$(x \,|\, y) = \frac{\|x + y\|^2 + \|x - y\|^2}{2}$$

et si $\mathbb{K} = \mathbb{C}$, alors

$$(x \,|\, y) = \frac{\|x + y\|^2 - \|x - y\|^2}{4} + i\,\frac{\|x + iy\|^2 - \|x - iy\|^2}{4}.$$

Démonstration : Pour $\mathbb{K} = \mathbb{R}$, on a $\|x + y\|^2 = \|x\|^2 + 2(x \,|\, y) + \|y\|^2$ et $\|x - y\|^2 = \|x\|^2 - 2(x \,|\, y) + \|y\|^2$, d'où l'identité de polarisation en sommant ces deux égalités.

Pour $\mathbb{K} = \mathbb{C}$, on a de même $\|x + y\|^2 - \|x - y\|^2 = 4\mathrm{Re}\,(x \,|\, y)$, donc $\|x + iy\|^2 - \|x - iy\|^2 = 4\mathrm{Re}\,(x \,|\, iy) = 4\mathrm{Re}\,(-i(x \,|\, y)) = 4\mathrm{Im}\,(x \,|\, y)$. On en déduit l'identité de polarisation dans ce cas à l'aide de la relation $(x \,|\, y) = \mathrm{Re}\,(x \,|\, y) + i\,\mathrm{Im}\,(x \,|\, y)$.

□

Proposition 2.1.6 (Identité du parallélogramme). *Pour tous $x, y \in \mathcal{H}$, on a*

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\,\|x\|^2 + \|y\|^2\,\big).$$

Démonstration : On a

$$\|x+y\|^2 = \|x\|^2 + 2\mathrm{Re}\,(x\,|\,y) + \|y\|^2$$

et

$$\|x-y\|^2 = \|x\|^2 - 2\mathrm{Re}\,(x\,|\,y) + \|y\|^2\,,$$

donc

$$\|x+y\|^2 + \|x-y\|^2 = 2\big(\,\|x\|^2 + \|y\|^2\,\big).$$

□

L'identité du parallélogramme exprime que la somme des carrés des longueurs des diagonales d'un parallélogramme est le double de la somme des carrés des longueurs des côtés adjacents. On peut montrer que, réciproquement, si $\|.\|$ est une norme vérifiant cette identité, alors les quantités

$$(x\,|\,y) = \frac{1}{2}\big(\,\|x+y\|^2 - \|x-y\|^2\,\big)$$

si $\mathbb{K} = \mathbb{R}$ et

$$(x\,|\,y) = \frac{1}{4}\big(\,\|x+y\|^2 - \|x-y\|^2 + i\,\|x+iy\|^2 - i\,\|x-iy\|^2\,\big)$$

si $\mathbb{K} = \mathbb{C}$, définissent des produits scalaires sur le $\mathbb{K}$-espace vectoriel $\mathcal{H}$, respectivement lorsque $\mathbb{K} = \mathbb{R}$ et lorsque $\mathbb{K} = \mathbb{C}$ (ce résultat est parfois appelé le théorème de Jordan-von Neumann).

Proposition 2.1.7 (Identité de la médiane). *Pour tous $x, y, z \in \mathcal{H}$, on a*

$$\|x-y\|^2 + \|x-z\|^2 = \frac{1}{2}\,\|z-y\|^2 + 2\left\|x - \frac{y+z}{2}\right\|^2.$$

Démonstration : On a $\|x-y\|^2 + \|x-z\|^2 = 2\|x\|^2 + \|y\|^2 + \|z\|^2 - 2\mathrm{Re}\,(x\,|\,y) - 2\mathrm{Re}\,(x\,|\,z)$ et $\frac{1}{2}\|z-y\|^2 + 2\|x - \frac{y+z}{2}\|^2 = \|z\|^2 + \|y\|^2 + 2\|x\|^2 - \mathrm{Re}\,(y\,|\,z) + \mathrm{Re}\,(y\,|\,z) - 2\mathrm{Re}\,(x\,|\,y+z)$, d'où l'égalité. L'identité de la médiane exprime la longueur de la médiane $[xm]$ où $m = \frac{y+z}{2}$ en fonction des longueurs des trois côtés du triangle (xyz) : $\left\|x - \frac{y+z}{2}\right\| = \sqrt{\frac{\|x-y\|^2 + \|x-z\|^2 - \frac{1}{2}\|y-z\|^2}{2}}$.

□

2.1.4 Orthogonalité et théorème de Pythagore

Définition 2.1.8 (Orthogonalité). *Deux éléments x et y d'un espace de Hilbert $\big(\mathcal{H}, (.\,|\,.)\big)$ sont dits orthogonaux lorsque leur produit scalaire est nul : $(x\,|\,y) = 0$. Nous noterons cela $x \perp y$. Deux parties A et B de $\mathcal{H}$ sont dites orthogonales si $(x\,|\,y) = 0$ pour tout $x \in A$ et tout $y \in B$. Nous noterons cela $A \perp B$.*

Lorsque $\mathcal{H} = \mathbb{R}^2$ est muni du produit scalaire usuel, deux vecteurs sont bien orthogonaux lorsque l'angle non orienté qu'ils définissent est de mesure $\frac{\pi}{2}$.

Proposition 2.1.9. *Pour toute partie $A \subset \mathcal{H}$, on note*

$$A^{\perp} = \{x \in \mathcal{H} \,|\, \forall y \in A, (x\,|\,y) = 0\}.$$

$A^{\perp}$ est un sous-espace vectoriel fermé de $\mathcal{H}$, appelé l'orthogonal de A dans $\mathcal{H}$, qui vérifie

1. *si $B \subset A$, alors $A^\perp \subset B^\perp$;*
2. $A \subset (A^\perp)^\perp$ *;*
3. $\overline{A}^\perp = A^\perp$ *;*
4. $A^\perp = \mathrm{vect}(A)^\perp = \overline{\mathrm{vect}(A)}^\perp$ *;*
5. $\overline{\mathrm{vect}(A)} \cap A^\perp = \{0\}$.

Démonstration : $A^\perp$ est l'intersection, pour $y \in A$, des noyaux des formes linéaires continues $\Psi_y : x \mapsto (x \,|\, y)$ qui sont des sous-espaces vectoriels fermés de $\mathcal{H}$ et est donc un sous-espace vectoriel fermé de $\mathcal{H}$.

Si $B \subset A$ et $x \in A^\perp$ alors on a en particulier, pour tout $y \in B$, $(x \,|\, y) = 0$, donc $x \in B^\perp$.

1. Si $y \in A$ alors, pour tout $x \in A^\perp$, $(y \,|\, x) = \overline{(x \,|\, y)} = 0$, donc $y \in (A^\perp)^\perp$, donc $A \subset (A^\perp)^\perp$.
2. $A \subset \overline{A}$ donc $\overline{A}^\perp \subset A^\perp$ d'après le point 1. Réciproquement, si $x \in A^\perp$ et $y = \lim_{n\to+\infty} y_n$ avec $y_n \in A$ alors, par continuité de $(.\,|\,.)$, on a $(x \,|\, y) = \lim_{n\to+\infty}(x \,|\, y_n) = 0$, donc $x \in \overline{A}^\perp$, donc $A^\perp \subset \overline{A}^\perp$.
3. $A \subset \mathrm{vect}(A)$ donc $(\mathrm{vect}(A))^\perp \subset A^\perp$ d'après le point 1. Réciproquement, si $x \in A^\perp$ et $y = \sum_{i=1}^n y_i \in \mathrm{vect}(A)$ avec $y_i \in A$ pour tout i, alors $(x \,|\, y) = \sum_{i=1}^n (x \,|\, y_i) = 0$, donc $x \in (\mathrm{vect}(A))^\perp$, de sorte que $A^\perp \subset (\mathrm{vect}(A))^\perp$, d'où $A^\perp = (\mathrm{vect}(A))^\perp$ et ceci est égal à $\overline{\mathrm{vect}(A)}^\perp$ d'après le point 3.
4. Soit $x \in \overline{\mathrm{vect}(A)} \cap A^\perp$. Alors $x \in A^\perp = \overline{\mathrm{vect}(A)}^\perp$ d'après le point 4, donc $(x \,|\, x) = 0$, donc $x = 0$.

□

Une conséquence immédiate de la notion d'orthogonalité est la relation suivante.

Théorème 2.1.10 (Pythagore). *Si $x_1, ..., x_n$ sont deux à deux orthogonaux dans $\mathcal{H}$, alors*

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2 .$$

Démonstration : Le cas $n = 2$ se démontre à partir de l'identité $\|x_1 + x_2\|^2 = \|x_1\|^2 + 2\mathrm{Re}\,(x_1 \,|\, x_2) + \|x_2\|^2$ et du fait que $|\mathrm{Re}\,(x_1 \,|\, x_2)| = 0$. On en déduit la relation de Pythagore par récurrence sur n.

□

2.1.5 Espaces de Hilbert

2.1.5.1 Définition

Définition 2.1.11 (Espace de Hilbert.). *Un espace de Hilbert $(\mathcal{H}, (.\,|\,.))$ sur $\mathbb{K}$ est un $\mathbb{K}$-espace vectoriel $\mathcal{H}$ muni d'un produit scalaire $(.\,|\,.)$ tel que l'espace vectoriel normé $\left(\mathcal{H}, \|.\| = \sqrt{(.\,|\,.)}\right)$ soit complet.*

2.1.5.2 Premiers exemples

Reprenons les exemples rencontrés plus haut.

1. $\mathbb{K}^n$ muni du produit scalaire $(x|y) = \sum_{k=1}^n x_k \overline{y_k}$ est complet pour la norme associée car c'est un espace vectoriel normé de dimension finie.

2. Comme nous l'avons vu en introduction, $\mathbb{K}^{(\mathbb{N})}$ n'est *pas* complet pour la norme (2.2) associée au produit scalaire (2.1).
3. La norme sur $L^2(X, \mathbb{C})$ associée au produit scalaire dont nous l'avons muni est $\|f\| = \sqrt{\int_X |f|^2 d\mu}$. Cet espace vectoriel normé est complet d'après le théorème de Riesz-Fischer.
4. Le sous-espace vectoriel de $L^2(X, \mathbb{C})$ composé des fonctions continues à support compact n'est pas complet car il n'est pas fermé dans $L^2(X, \mathbb{C})$. En effet, ce sous-espace est dense dans $L^2(X, \mathbb{C})$.
5. La norme sur $\ell^2(I, \mathbb{K})$ associée au produit scalaire dont nous l'avons muni est $\|\mathbf{x}\| = \sqrt{\sum_{i \in I} |x_i|^2}$. Pour montrer la complétude de cet espace vectoriel normé, on peut soit recourir à l'exemple précédent soit effectuer une démonstration directe.
6. $\mathcal{M}_\mathrm{n}(\mathbb{C})$ muni de son produit scalaire défini à partir de la trace est complet car de dimension finie.

2.2 Théorème de la projection orthogonale

Le théorème que nous présentons dans ce chapitre est l'un des principaux résultats de la théorie des espaces de Hilbert au sens où il permet la construction de nombreux objets, aussi bien en probabilité et statistiques qu'en analyse fonctionnelle.

2.2.1 Projection sur un convexe fermé

Théorème 2.2.1 (Projection orthogonale sur un convexe complet)**.** *Soit $\big(\mathcal{H}, (.\,|\,.)\big)$ un $\mathbb{K}$-espace vectoriel muni d'un produit scalaire et soit C une partie convexe complète de $\mathcal{H}$. Alors, pour tout $x \in \mathcal{H}$, il existe un unique $\in C$ tel que $\|x - \Pi_C(x)\| = \mathrm{dist}(x, C) = \inf_{y \in C} \|x - y\|$. Le point $\Pi_C(x) \in C$ est appelé le projeté orthogonal de x sur C.*

Démonstration : Notons $\delta = \mathrm{dist}(x, C)$. Nous allons séparer les rôles de la convexité et de la complétude.

Utilisation de la convexité. Nous allons montrer que la convexité de C entraîne que, pour tout $\varepsilon > 0$, l'intersection C_ε de C avec la boule fermée $B(x, \delta + \varepsilon]$, est bornée, et que diam C_ε tend vers 0 lorsque ε tend vers 0. Soient u et v dans C_ε. Appliquons l'identité de la médiane (voir Proposition 2.1.7) aux points x, u, v et notons $m = \frac{u+v}{2}$ le milieu de u et v :

$$\|x - u\|^2 + \|x - v\|^2 = 2\,\|x - m\|^2 + \frac{1}{2}\,\|u - v\|^2 .$$

Comme C est convexe, $m \in C$, donc $\|x - m\|^2 \geqslant \delta^2$. On en déduit

$$\|u - v\|^2 = 2(\|x - u\|^2 + \|x - v\|^2) - 4\,\|x - m\|^2 \leqslant 4((\delta + \varepsilon)^2 - \delta^2) = 8\,\delta\,\varepsilon + 4\,\varepsilon^2.$$

Comme u et v sont des points quelconques de C_ε, on en déduit que

$$\mathrm{diam}\, C_\varepsilon \leqslant 8\,\delta\,\varepsilon + 4\,\varepsilon^2,$$

donc $\lim_{\varepsilon \to 0} \mathrm{diam}\, C_\varepsilon = 0$.

2. *Utilisation de la complétude*. Considérons maintenant la suite $(C_{1/n})_{n \in \mathbb{N}^*}$. C'est une suite décroissante de parties fermées non vides de C, dont les diamètres tendent vers 0. Comme C est complet, l'intersection de cette suite est non vide et se réduit à un point, que l'on notera $\Pi_C(x)$. Comme $\Pi_C(x)$ est dans chacune des boules $B(x, \delta + 1/n]$, $\|x - \Pi_C(x)\| \leqslant \delta$, et, comme $\Pi_C(x) \in C$, $\|x - \Pi_C(x)\| \geqslant \delta$ donc $\|x - \Pi_C(x)\| = \delta$ et la distance

est atteinte au point $\Pi_C(x)$. Réciproquement, si cette distance est atteinte au point $z \in C$, z est dans chacune des boules $B(x, \delta + 1/n]$, donc dans l'intersection $\{\Pi_C(x)\}$ de ces boules donc $z = \Pi_C(x)$.

□

Remarquons que le théorème 2.2.1 est vrai sans supposer que la norme associée au produit scalaire $(.\,|\,.)$ fait de $\mathcal{H}$ un espace complet. Cela permet par exemple de projeter orthogonalement sur le sous-espace vectoriel $\mathbb{K}^{n+1} = \{(x_0, ..., x_n, 0, ...) \ : \ x_i \in \mathbb{K}\}$ de $\mathbb{K}^{(\mathbb{N})}$, ce sous-espace étant bien complet car de dimension finie. Le théorème 2.2.1 est cependant le plus souvent utilisé dans le cas où $\big(\mathcal{H}, (.\,|\,.)\big)$ est un espace de Hilbert et où $C = F$ est un sous-espace vectoriel *fermé* (donc complet) de $\mathcal{H}$. Dans ce cas, on a la *caractérisation* suivante du projeté orthogonal d'un point x de $\mathcal{H}$.

Théorème 2.2.2 (Projection sur un sous-espace fermé dans un espace de Hilbert)**.** *Soit F un sous-espace vectoriel fermé d'un espace de Hilbert $\big(\mathcal{H}, (.\,|\,.)\big)$ et soit $x \in \mathcal{H}$. On note $\Pi_F(x)$ le projeté orthogonal de x sur F. Alors, pour $y \in \mathcal{H}$, les conditions suivantes sont équivalentes :*

1. *$y = \Pi_F(x)$;*
2. *$y \in F$ et $(x - y) \in F^\perp$.*

Démonstration : Supposons que $y = \Pi_F(x)$. Alors $y \in F$ par définition. De plus, pour tout $z \in F$, $y + z \in F$ car F est un sous-espace vectoriel de $\mathcal{H}$. Donc, par définition de $y = \Pi_F(x)$,

$$\|x - y\|^2 \leqslant \|x - (y + z)\|^2 = \|(x - y) - z\|^2 = \|x - y\|^2 - 2\mathrm{Re}\,(x - y\,|\,z) + \|z\|^2.$$

Donc $2\mathrm{Re}\,(x - y\,|\,z) \leqslant \|z\|^2$ pour tout $z \in F$. En remplaçant z par $-z$, on obtient $-2\mathrm{Re}\,(x - y\,|\,z) \leqslant \|z\|^2$, de sorte que $|2\mathrm{Re}\,(x - y\,|\,z)| \leqslant \|z\|^2$.
Écrivons $(x - y\,|\,z) = re^{i\theta}$ avec $r \geqslant 0$. Le même raisonnement que ci-dessus appliqué à $te^{i\theta}z$ avec $t \in \mathbb{R}$ quelconque entraîne : $2|t||\mathrm{Re}\,\big(e^{-i\theta}(x - y\,|\,z)\big)| \leqslant \|te^{i\theta}z\|^2$, soit $2|t|r \leqslant t^2\|z\|^2$, soit, pour tout $t \neq 0$, $r \leqslant \frac{1}{2}|t|\|z\|^2$. En faisant tendre t vers 0, on voit que l'on a nécessairement $r = 0$, donc $(x - y\,|\,z) = 0$, et ceci pour tout $z \in F$, c'est-à-dire que $(x - y) \in F^\perp$.

Réciproquement, supposons que $y \in F$ et $(x - y) \in F^\perp$. Alors, pour tout $z \in F$, $(x - y) \perp (y - z)$, et par Pythagore, $\|x - z\|^2 = \|x - y + y - z\|^2 = \|x - y\|^2 + \|y - z\|^2 \geqslant \|x - y\|^2$. Donc $\|x - y\| = \inf_{z \in F} \|x - z\| = dist(x, F)$, c'est-à-dire que $y = \Pi_F(x)$ d'après le théorème 2.2.1.

□

2.2.2 Projecteurs orthogonaux

Nous allons maintenant vérifier que si F est un sous-espace vectoriel fermé dans un espace de Hilbert, l'application $x \mapsto \Pi_F(x)$ est un projecteur orthogonal au sens que l'on connaît déjà en dimension finie.

Proposition 2.2.3. *Soit F un sous-espace vectoriel fermé d'un espace de Hilbert $\big(\mathcal{H}, (.\,|\,.)\big)$ et soit Π_F la projection orthogonale sur F. Alors, l'application Π_F est une application linéaire continue qui vérifie*

1. *$\Pi_F \circ \Pi_F = \Pi_F$;*
2. *$\mathrm{Im}\,\Pi_F = F$ et $\mathrm{Ker}\,\Pi_F = F^\perp$;*
3. *$Id_{\mathcal{H}} - \Pi_F$ est la projection orthogonale sur $F^\perp$;*
4. *$\mathrm{Ker}\,\Pi_F \oplus \mathrm{Im}\,\Pi_F = \mathcal{H}$.*

Démonstration : On commence par montrer le fait que l'application Π_F est une application linéaire continue. Soient $x, y \in \mathcal{H}$ et $\lambda, \mu \in \mathbb{K}$. Alors, pour tout $z \in F$,

$$\big((\lambda x + \mu y) - (\lambda \Pi_F(x) + \mu \Pi_F(y)) \,|\, z\big) = \lambda(x - \Pi_F(x) \,|\, z) + \mu(y - \Pi_F(y) \,|\, z) = 0.$$

Puisque $(\lambda \Pi_F + \mu \Pi_F(y)) \in F$, on a donc, d'après le théorème 2.2.2, $\lambda \Pi_F(x) + \mu \Pi_F(y) = \Pi_F(\lambda x + \mu y)$ et Π_F est donc bien linéaire.

Pour tout $x \in \mathcal{H}$, on a $x = (x - \Pi_F(x)) + \Pi_F(x)$, donc, d'après la relation de Pythagore, $\|x\|^2 = \|x - \Pi_F(x)\|^2 + \|\Pi_F(x)\|^2 \geqslant \|\Pi_F(x)\|^2$, donc $\|\Pi_F(x)\| \leqslant \|x\|$, ce qui montre que Π_F est continue de norme plus petite que 1.

On a $\operatorname{Im} \Pi_F \subset F$ et, si $y \in F$, $\Pi_F(y) = y$ par définition de Π_F. Donc $\Pi_F \circ \Pi_F = \Pi_F$.

2. Ce qui précède montreégalement que $\operatorname{Im} \Pi_F = F$. De plus, si $\Pi_F(x) = 0$, alors $x = x - \Pi_F(x) \in F^\perp$. Réciproquement, si $x \in F^\perp$, alors $(x - 0) \in F^\perp$ avec $0 \in F$, donc $\Pi_F(x) = 0$ d'après le théorème 2.2.2. Donc $\operatorname{Ker} \Pi_F = F^\perp$.
3. On a $\Pi_F^2 = \Pi_F$ d'après le point 1. Donc, pour tout $x \in \mathcal{H}$, $(x - \Pi_F(x)) \in \operatorname{Ker} \Pi_F = F^\perp$ d'après le point 2. De plus, $x - (x - \Pi_F(x)) = \Pi_F(x) \in F \subset (F^\perp)^\perp$ d'après la proposition 2.1.9. Donc, d'après le théorème 2.2.2, $(Id_{\mathcal{H}} - \Pi_F)(x) = x - \Pi_F(x)$ est le projeté orthogonal de x sur $F^\perp$.
4. Soit $x \in \mathcal{H}$. Alors $x = (x - \Pi_F(x)) + \Pi_F(x) \in \operatorname{Ker} \Pi_F + \operatorname{Im} \Pi_F$ de sorte que $\mathcal{H} = \operatorname{Ker} \Pi_F + \operatorname{Im} \Pi_F$ et cette somme est directe d'après le point 2 puisque $F^\perp \cap F = \{0\}$ d'après la proposition 2.1.9.

□

2.2.3 Supplémentaire orthogonal

On montre dans cette section qu'un sous-espace vectoriel fermé d'un espace de Hilbert possède toujours un supplémentaire orthogonal.

Théorème 2.2.4 (Supplémentaire orthogonal)**.** *Soit F un sous-espace vectoriel fermé d'un espace de Hilbert $\big(\mathcal{H}, (.\,|\,.)\big)$. Alors,*

$$F \overset{\perp}{\oplus} F^\perp = \mathcal{H}.$$

Démonstration : L'égalité $F \oplus F^\perp = \mathcal{H}$ se déduit immédiatement des point 2 et 4 de la proposition 2.2.3. Il nous reste à vérifier que F et $F^\perp$ sont bien l'orthogonal l'un de l'autre. Toujours d'après la proposition 2.2.3, $(Id_{\mathcal{H}} - \Pi_F)$ est la projection orthogonale sur $F^\perp$, donc $(F^\perp)^\perp = \operatorname{Ker}(Id_{\mathcal{H}} - \Pi_F) = \operatorname{Im} \Pi_F = F$ (où l'égalité $\operatorname{Ker}(Id_{\mathcal{H}} - \Pi_F) = \operatorname{Im} \Pi_F$ vient du fait que $\Pi_F^2 = \Pi_F$). Comme $(F^\perp)^\perp = F$, on a bien $F \overset{\perp}{\oplus} F^\perp = \mathcal{H}$.

□

On a plus généralement, lorsque l'on ne suppose plus que F est un sous-espace vectoriel fermé de $\mathcal{H}$, le résultat suivant.

Proposition 2.2.5. *Soit $\big(\mathcal{H}, (.\,|\,.)\big)$ un espace de Hilbert. Si $A \subset \mathcal{H}$ est une partie quelconque de $\mathcal{H}$, alors*

$$\overline{\operatorname{vect}(A)} \oplus A^\perp = \mathcal{H} \quad \text{et} \quad (A^\perp)^\perp = \overline{\operatorname{vect}(A)}.$$

En particulier, si $A = F$ est un sous-espace vectoriel de $\mathcal{H}$, alors

$$\overline{F} \oplus F^\perp = \mathcal{H} \quad \text{et} \quad (F^\perp)^\perp = \overline{F}.$$

Démonstration : Posons $F = \overline{\text{vect}(A)}$. D'après la proposition 2.1.9, on a $A^\perp = F^\perp$. Or F est un sous-espace vectoriel fermé de $\mathcal{H}$, donc $F \oplus F^\perp = \mathcal{H}$ et $(F^\perp)^\perp = F$ d'après le théorème 2.2.4, c'est-à-dire que $\overline{\text{vect}(A)} \oplus A^\perp = \mathcal{H}$ et $(A^\perp)^\perp = \overline{\text{vect}(A)}$.

L'énoncé pour $A = F$, un sous-espace vectoriel non nécessairement fermé de $\mathcal{H}$, s'en déduit alors immédiatement en remarquant que dans ce cas $\text{vect}(F) = F$.

□

Le résultat précédent, important en pratique, a un corollaire sur la caractérisation des sous-espaces vectoriels denses dans $\mathcal{H}$. Ce corollaire est très utile en pratique car il donne une caractérisation géométrique de la densité qui est une propriété souvent difficile à vérifier dans un espace vectoriel normé quelconque qui n'est pas un espace de Hilbert.

Corollaire 2.2.6. *Si F est un sous-espace vectoriel de $\mathcal{H}$, alors F est dense dans $\mathcal{H}$ si et seulement si $F^\perp = \{0\}$.*

Démonstration : Le sous-espace vectoriel F est dense dans $\mathcal{H}$ si et seulement si $\overline{F} = \mathcal{H}$, soit encore d'après la proposition 2.2.5, $F^\perp)^\perp = \mathcal{H}$. Or, $F^\perp)^\perp = \mathcal{H}$ si et seulement si $F^\perp = \{0\}$.

□

Nous finissons ce chapitre par une application de ce dernier corollaire à la bijectivité des applications coercives.

Proposition 2.2.7. *Soit $T : \mathcal{H} \to \mathcal{H}$ une application linéaire continue. On suppose de plus qu'il existe $C > 0$ tel que*

$$\forall u \in \mathcal{H},\ |(Tu|u)| \geqslant C\|u\|^2. \tag{2.3}$$

Alors T est bijective.

Vocabulaire. Lorsque T vérifie (2.3), on dit que T est *coercive*.

Démonstration : Pour montrer que T est bijective, il suffit de montrer qu'elle est injective et surjective. Pour la surjectivité, nous montrerons que l'image de T est fermée et dense dans $\mathcal{H}$, donc qu'elle vaut $\mathcal{H}$ tout entier.

Injectivité. Soit $u \in \mathcal{H}$. Supposons que $Tu = 0$. Alors, $0 = |(Tu|u)| \geqslant C\|u\|^2 \geqslant 0$, donc $\|u\|^2 = 0$ et $u = 0$.

Image de T fermée. Soit $(v_n)_{n\in\mathbb{N}}$ une suite de points de $\operatorname{Im} T$ qui converge vers $v \in \mathcal{H}$. Montrons que $v \in \operatorname{Im} T$. Pour tout $n \in \mathbb{N}$, soit $u_n \in \mathcal{H}$ tel que $v_n = Tu_n$. Comme $(v_n)_{n\in\mathbb{N}}$ converge dans $\mathcal{H}$ elle est en particulier de Cauchy :

$$\forall \varepsilon > 0,\ \exists N \geqslant 1,\ \forall n, p \geqslant N,\ \|Tu_n - Tu_p\| \leqslant \varepsilon.$$

En utilisant la coercivité de T et l'inégalité de Cauchy-Schwarz, on a

$$\forall n, p,\ C\|u_n - u_p\|^2 \leqslant |(T(u_n - u_p)|u_n - u_p)| \leqslant \|T(u_n - u_p)\| \cdot \|u_n - u_p\|.$$

En supposant $\|u_n - u_p\| \neq 0$, il vient

$$\forall n, p \geqslant N,\ C\|u_n - u_p\| \leqslant \|Tu_n - Tu_p\| \leqslant \varepsilon.$$

D'où,

$$\forall n, p \geqslant N,\ \|u_n - u_p\| \leqslant \frac{1}{C}\varepsilon.$$

Donc la suite $(u_n)_{n\in\mathbb{N}}$ est de Cauchy dans $\mathcal{H}$, donc converge vers $u \in \mathcal{H}$. Or, T est continue, donc $(v_n) = (Tu_n)$ converge vers Tu. Par unicité de la limité, $v = Tu \in \operatorname{Im} T$. Donc $\operatorname{Im} T$ est fermée.

Image de T dense. Si $u \in (\operatorname{Im} T)^\perp$, alors $Tu \perp u$, donc $0 = |(Tu|u)| \geqslant C\|u\|^2 \geqslant 0$ et $u = 0$. Donc $(\operatorname{Im} T)^\perp \subset \{0\}$ et ainsi $(\operatorname{Im} T)^\perp = \{0\}$. Par le corollaire 2.2.6, $\operatorname{Im} T$ est dense dans $\mathcal{H}$.

□

Nous utiliserons ce résultat pour démontrer le théorème de Lax-Milgram.

2.3 Dualité dans les espaces de Hilbert

2.3.1 Théorème de Riesz-Fréchet

Le théorème de représentation de Riesz-Fréchet caractérise le dual topologique d'un espace de Hilbert : il montre comment on obtient *toutes* les formes linéaires continues sur un espace de Hilbert.

Théorème 2.3.1 (Riesz-Fréchet)**.** *Soit $\big(\mathcal{H}, (.\,|\,.)\big)$ un espace de Hilbert. Pour toute forme linéaire continue φ sur $\mathcal{H}$, il existe un unique vecteur $v \in \mathcal{H}$ tel que :*

$$\forall u \in \mathcal{H},\ \varphi(u) = (u|v).$$

De plus,

$$|||\varphi||| = \|v\|.$$

Démonstration : Si $\varphi = 0$, alors $v = 0$ convient. On peut donc supposer que $\varphi \neq 0$.

Soit $u_0 \in \mathcal{H}$ tel que $\varphi(u_0) \neq 1$. En particulier, comme $\varphi \neq 0$, $\operatorname{Ker}\varphi \neq \mathcal{H}$, et puisque $(\operatorname{Ker}\varphi) \oplus (\operatorname{Ker}\varphi)^\perp = \mathcal{H}$, on peut trouver unélément non nul v_0 de $(\operatorname{Ker}\varphi)^\perp$.
Par ailleurs, comme $\varphi(u_0) \neq 1$, pour tout $u \in \mathcal{H}$, $u - \varphi(u)u_0 \in \operatorname{Ker}\varphi$. En particulier, $u - \varphi(u)u_0 \perp v_0$. Soit encore,

$$0 = (u - \varphi(u)u_0|v_0) = (u|v_0) - \varphi(u)(u_0|v_0).$$

De là, si $(u_0|v_0) = 0$, on en déduit que pour tout $u \in \mathcal{H}$, $(u|v_0) = 0$ ce qui entraînerait que $v_0 = 0$. Cela est absurde, donc $(u_0|v_0) \neq 0$. Alors, en réutilisant la relation ci-dessus, on obtient,

$$\forall u \in \mathcal{H},\ \varphi(u) = \frac{(u|v_0)}{(u_0|v_0)} = (u|v)$$

où l'on a posé $v = \frac{v_0}{(u_0|v_0)}$. Cela prouve l'existence de v.

Pour prouver l'unicité de v, supposons qu'il existe deux vecteurs v_1 et v_2 vérifiant pour tout $u \in \mathcal{H}$, $\varphi(u) = (u|v_1) = (u|v_2)$. Alors, pour tout $u \in \mathcal{H}$, $(u|v_1 - v_2) = 0$ ce qui entraîne que $v_1 - v_2 = 0$ soit encore $v_1 = v_2$.

Enfin, on a

$$|||\varphi||| = \sup_{||u||=1} |\varphi(u)| = \sup_{||u||=1} |(u|v)| = ||v||.$$

En effet, par Cauchy-Schwarz, $|(u|v)| \leqslant ||u|| \cdot ||v||$ d'où $\sup_{||u||=1} |(u|v)| \leqslant ||v||$. De plus, pour $u = \frac{v}{||v||}$, le sup est atteint, d'où l'égalité.

□

2.3.2 Existence et propriétés de l'adjoint d'un opérateur borné

Nous allons maintenant définir l'adjoint d'un opérateur borné qui généralise à la dimension quelconque la transposée d'une matrice réelle ou la transconjuguée d'une matrice complexe.

Proposition 2.3.2. *Soient $\mathcal{H}$ un espace de Hilbert et soit $T \in \mathcal{L}(\mathcal{H})$. Il existe un unique opérateur $T^* \in \mathcal{L}(H)$ tel que*

$$\forall u \in H,\ \forall v \in H,\ (Tu|v) = (u|T^*v). \tag{2.4}$$

Démonstration : Soit $v \in \mathcal{H}$. Alors, $\ell_v : u \mapsto (Tu|v)$ est une forme linéaire continue sur $\mathcal{H}$. En effet, on applique l'inégalité de Cauchy-Schwarz et on utilise le caractère borné de T. Par le théorème de Riesz de représentation des formes linéaires continues, il existe un unique vecteur $w \in \mathcal{H}$ tel que, pour tout $u \in \mathcal{H}$, $\ell_v(u) = (u|w)$. Posons $T^* : \mathcal{H} \to \mathcal{H}$, $T^*v = w$.

T^* est linéaire. En effet, si $v_1, v_2 \in \mathcal{H}$ et $\lambda_1, \lambda_2 \in \mathbb{C}$, alors soit $v = \lambda_1 v_1 + \lambda_2 v_2$ et $w_1 = T^*(v_1)$, $w_2 = T^*(v_2)$, $T^*(v) = w$. Alors,

$$\begin{aligned}\forall u \in \mathcal{H},\ (u|w) &= (Tu|v) = (Tu|\lambda_1 v_1 + \lambda_2 v_2) = \overline{\lambda}_1 (Tu|v_1) + \overline{\lambda}_2 (Tu|v_2)\\ &= \overline{\lambda}_1 (u|w_1) + \overline{\lambda}_2 (u|w_2) = (u|\lambda_1 w_1 + \lambda_2 w_2).\end{aligned}$$

Donc, $w - \lambda_1 w_1 - \lambda_2 w_2 \in \mathcal{H}^\perp = \{0\}$ et $w = \lambda_1 w_1 + \lambda_2 w_2$ et T^* est linéaire.

T^* est borné. En effet, soient $u, v \in \mathcal{H}$, $\|u\| \leqslant 1$ et $\|v\| \leqslant 1$. Alors,

$$|(u|T^*v)| = |(Tu|v)| \leqslant \|Tu\|\ \|v\| \leqslant |||T|||.$$

Donc, en prenant $u = \frac{T^*v}{\|T^*v\|}$, pour tout $v \in \mathcal{H}$, $\|v\| \leqslant 1$ et $T^*v \neq 0$, $\|T^*v\| \leqslant |||T|||$. Si $v \in \mathcal{H}$ est tel que $T^*v = 0$, l'inégalité est encore vérifiée. On obtient donc $|||T^*||| \leqslant |||T|||$ et T^* est borné.

Enfin, pour l'unicité, si T_1^* et T_2^* vérifient (2.4), alors, pour tous $u, v \in \mathcal{H}$, $(u|(T_1^* - T_2^*)v) = 0$, donc $T_1^* - T_2^* = 0$.

□

Définition 2.3.3 (Adjoint)**.** *L'opérateur borné $T^* \in \mathcal{L}(\mathcal{H})$ est appelé adjoint de l'opérateur T.*

Proposition 2.3.4 (Propriétés algébriques de l'adjoint)**.** *Soient $T, T_1, T_2 \in \mathcal{L}(\mathcal{H})$ et $\lambda \in \mathbb{C}$. Alors,*

1. $(T_1 + T_2)^* = T_1^* + T_2^*$ *;*
2. $(\lambda T)^* = \overline{\lambda} T^*$ *;*
3. $(T_1 T_2)^* = T_2^* T_1^*$ *;*
4. $(T^*)^* = T$ *;*
5. *si T a un inverse borné T^{-1}, T^* a aussi un inverse borné et $(T^*)^{-1} = (T^{-1})^*$.*

Démonstration : Les deux premiers points proviennent de la semi-linéarité à droite du produit scalaire. Pour le troisième point, onécrit, pour tous $u, v \in \mathcal{H}$, $(T_1 T_2 u|v) = (T_2 u|T_1^* v) = (u|T_2^* T_1^* v)$. Le quatrième point s'obtient en remarquant que, dans (2.4), les vecteurs u et v jouent le même rôle et $(Tu|v) = (u|T^*v)$ pour tous $u, v \in \mathcal{H}$ si et seulement si $(T^*u|v) = (u|Tv)$ pour tous $u, v \in H$ en passant aux conjugués. Enfin, pour le dernier point, de $TT^{-1} = I = T^{-1}T$ on déduit par passage à l'adjoint que $T^*(T^{-1})^* = I^* = I = I^* = (T^{-1})^* T^*$.

□

Proposition 2.3.5 (Propriétés métriques de l'adjoint)**.** *Soit $T \in \mathcal{L}(\mathcal{H})$. Alors,*

1. $|||T^*||| = |||T|||$*;*
2. $|||T^*T||| = |||T|||^2$.

Démonstration : D'après la démonstration de la proposition 2.3.2, $|||T^*||| \leqslant |||T|||$. Puis, en appliquant cette inégalité à l'opérateur borné T^* et en utilisant le fait que $(T^*)^* = T$, on obtient aussi $|||T||| \leqslant |||T^*|||$, ce qui prouve le premier point. Pour le second point, on a tout d'abord, $|||T^*T||| \leqslant |||T^*||| \cdot |||T||| = |||T|||^2$. Réciproquement, si $u \in \mathcal{H}$, $\|u\| = 1$,

$$\|Tu\|^2 = (Tu|Tu) = (T^*Tu|u) \leqslant |||T^*T|||,$$

donc $|||T|||^2 \leqslant |||T^*T|||$.

□

Proposition 2.3.6 (Propriétés géométriques de l'adjoint)**.** *Soit $T \in \mathcal{L}(\mathcal{H})$. Alors,*

1. $\mathrm{Ker}\, T^* = (\mathrm{Im} T)^\perp$ *et* $(\mathrm{Ker}\, T^*)^\perp = \overline{\mathrm{Im} T}$*;*
2. *si $F \subset \mathcal{H}$ est un sous-espace vectoriel stable par T, alors $F^\perp$ est stable par T^*.*

Démonstration : u appartient à $(\mathrm{Im}T)^\perp$ si et seulement si, pour tout $v \in \mathcal{H}$, $(u|Tv) = 0$, ce quiéquivaut à ce que, pour tout $v \in \mathcal{H}$, $(T^*u|v) = 0$. C'estéquivalent à $T^*u = 0$, soit encore $u \in \mathrm{Ker}\, T^*$. La seconde propriété provient de la proposition 2.2.5.

Pour le second point, soit $v \in F^\perp$ et $u \in F$. Alors $Tu \in F$, donc $(T^*v|u) = (v|Tu) = 0$. Donc $T^*v \in F^\perp$.

□

Définition 2.3.7. *Un opérateur $T \in \mathcal{L}(\mathcal{H})$ est dit auto-adjoint lorsque $T = T^*$.*

Les opérateurs auto-adjoints sont la généralisation des matrices symétriques à la dimension infinie. Ils jouent un rôle majeur en analyse fonctionnelle et en physique mathématique (en particulier en mécanique quantique). On peut démontrer un théorème de structure sur ces opérateurs qui affirme que tout opérateur auto-adjoint est diagonalisable, en un sens à préciser en dimension infinie. Un premier exemple d'opérateur auto-adjoint est celui de projecteur orthogonal.

Proposition 2.3.8. *Soit $T \in \mathcal{L}(\mathcal{H})$ un opérateur auto-adjoint. Alors, pour tout $u \in \mathcal{H}$, $(Tu|u) \in \mathbb{R}$.*

Démonstration : Comme $T = T^*$, $(Tu|u) = (u|Tu) = \overline{(Tu|u)}$ et $(Tu|u) \in \mathbb{R}$.

□

2.3.3 Théorème de Lax-Milgram

Nous terminons par une autre application du théorème de Riesz-Fréchet.

Théorème 2.3.9 (Lax-Milgram)**.** *Soit $a : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ une application linéaire à gauche et semi-linéaire à droite. On suppose que a est continue au sens où*

$$\exists C_1 > 0,\ \forall u, v \in \mathcal{H},\ |a(u,v)| \leqslant C||u|| \cdot ||v||.$$

On suppose également que a est coercive :

$$\exists C_2 > 0,\ \forall u \in \mathcal{H},\ |a(u,u)| \geqslant C_2||u||^2.$$

Alors, pour toute forme linéaire continue φ sur $\mathcal{H}$, il existe $v_\varphi \in \mathcal{H}$ unique tel que :

$$\forall u \in \mathcal{H},\ a(u, v_\varphi) = \varphi(u).$$

Démonstration : Pour tout $v \in \mathcal{H}$, soit $\phi_v : \mathcal{H} \to \mathbb{C}$ la forme linéaire définie par $\phi_v(u) = a(u,v)$. Alors, ϕ_v est une forme linéaire continue et d'après le théorème de Riesz-Fréchet, il existe un unique vecteur w_{ϕ_v} dans $\mathcal{H}$ tel que, pour tout $u \in \mathcal{H}$, $\phi_v(u) = (u|w_{\phi_v})$.
Soit maintenant l'opérateur $T : \mathcal{H} \to \mathcal{H}$ défini par, pour tout $v \in \mathcal{H}$, $Tv = w_{\phi_v}$. Par définition,

$$\forall u \in \mathcal{H},\ \forall v \in \mathcal{H},\ a(u,v) = (u|Tv).$$

De plus, montrons que T est une application linéaire continue et coercive.

Linéarité. Soient $u, v_1, v_2 \in \mathcal{H}$. On a

$$(u|T(v_1+v_2)) = a(u,v_1+v_2) = a(u,v_1) + a(u,v_2) = (u|Tv_1) + (u|Tv_2) = (u|Tv_1+Tv_2).$$

Comme cela est vrai pour tout $u \in \mathcal{H}$, on $T(v_1+v_2) = Tv_1 + Tv_2$. De même, si $\lambda \in \mathbb{C}$ et si $u, v \in \mathcal{H}$, $(u|T(\lambda v)) = a(u,\lambda v) = \overline{\lambda} a(u,v) = \overline{\lambda}(u|Tv) = (u|\lambda Tv)$. Là encore on en déduit que $T(\lambda v) = \lambda Tv$.

Continuité. On a (voir preuve du théorème de Riesz et utilisation de Cauchy-Schwarz)

$$\forall v \in \mathcal{H},\ ||Tv|| = \sup_{||u||=1} |(u|Tv)| = \sup_{||u||=1} |a(u,v)| \leqslant C||v|| \cdot ||u|| = C||v||.$$

D'où la continuité de T.

Coercivité. On a : $\forall u \in \mathcal{H},\ |(Tu|u)| = |(u|Tu)| = |a(u,u)| \geqslant C||u||^2$.
Comme T est linéaire continue et coercive, elle est bijective. Soit donc φ une forme linéaire continue sur $\mathcal{H}$. Par le théorème de Riesz-Fréchet, il existe un unique vecteur $w_\varphi \in \mathcal{H}$ tel que, pour tout $u \in \mathcal{H}$, $\varphi(u) = (u|w_\varphi)$. Soit v_φ l'unique antécédent de w_φ par T : $Tv_\varphi = w_\varphi$. Alors,

$$\forall u \in \mathcal{H},\ \varphi(u) = (u|w_\varphi) = (u|Tv_\varphi) = a(u,v_\varphi).$$

Ainsi, v_φ est le vecteur recherché.

□

2.4 Bases hilbertiennes

Il s'agit ici de généraliser la notion de *base orthonormée* d'un espace euclidien ou hermitien, qui sont de dimension finie par définition. Le passage à la dimension infinie entraîne l'apparition de phénomènes nouveaux : les *bases hilbertiennes* que nous allons définir ne sont *jamais* des bases algébriques (familles libres et génératrices) de l'espace vectoriel sous-jacent à un espace de Hilbert *lorsque celui-ci n'est pas de dimension finie*.

2.4.1 Définition et inégalité de Bessel

Définition 2.4.1 (Base hilbertienne)**.** *Soit $\big(\mathcal{H}, (.\,|\,.)\big)$ un espace de Hilbert. Une famille $(e_i)_{i\in I}$ de vecteurs de $\mathcal{H}$ est dite*

1. *orthogonale si $(e_i\,|\,e_j) = 0$ pour tous $i, j \in I$ tels que $i \neq j$;*
2. *orthonormée si $(e_i\,|\,e_j) = 0$ pour tous $i, j \in I$ tels que $i \neq j$ et $(e_i\,|\,e_i) = 1$ pour tout $i \in I$;*
3. *totale si vect$\big((e_i)_{i\in I}\big)$ est dense dans $\mathcal{H}$.*

On appelle base hilbertienne de $\mathcal{H}$ une famille orthonormée totale de vecteurs de $\mathcal{H}$.

Proposition 2.4.2. *1. Toute famille orthonormée est libre.*

2. $(e_i)_{i\in I}$ est une famille totale si et seulement si la condition $\big((x\,|\,e_i) = 0$ pour tout $i \in I\big)$ implique $x = 0$.

Démonstration : 1. Rappelons qu'une famille est dite libre lorsque toute combinaison linéaire *finie* nulle est à coefficients tous nuls. Soit $J \subset I$ une partie finie et soit $\sum_{j\in J} \lambda_j e_j = 0$ une combinaison linéaire nulle. Alors, pour tout $j_0 \in J$, on a $0 = (\sum_{j\in J} \lambda_j e_j\,|\,e_{j_0}) = \sum_{j\in J} \lambda_j (e_j\,|\,e_{j_0}) = \lambda_{j_0}$.

2. Notons $F = vect\big((e_i)_{i\in I}\big)$. Rappelons que F est dense dans $\mathcal{H}$ si et seulement si $F^\perp = \{0\}$. Soit $x \in \mathcal{H}$ tel que, pour tout $i \in I$, $(x\,|\,e_i) = 0$. Alors $x \in F^\perp$, donc $x = 0$. Réciproquement, soit $x \in F^\perp$. Alors en particulier $(x\,|\,e_i) = 0$ pour tout $i \in I$, donc $x = 0$ par hypothèse. Donc $F^\perp = \{0\}$.

□

Théorème 2.4.3 (Orthonormalisation de Gram-Schmidt)**.** *Considérons $(f_n)_n$ une suite totale d'un espace de Hilbert séparable H. Construisons une famille $(e_n)_n$ par le procédé suivant :*

— *$e_0 = f_0/\|f_0\|$ (on peut toujours se ramener à f_0 non nul).*

— *Si $e_0, \ldots, e_n$ ont été définis, on définit e_{n+1} en considérant f_m le premier vecteur de la famille $(f_n)_n$ qui est indépendant de $(e_0, \ldots, e_n)$, et on pose*

$$g_{n+1} = f_m - \sum_{0\leqslant i\leqslant n} (f_m|e_i)e_i, \quad e_{n+1} = \frac{g_{n+1}}{\|g_{n+1}\|}.$$

La famille $(e_n)_n$ ainsi obtenue est une base hilbertienne de H.

Démonstration : Quitte à enlever les éléments de la famille qui sont des combinaisons linéaires (finies !) des éléments précédents, on peut considérer que la famille $(f_n)_n$ est linéairement indépendante. Le procédé d'orthonormalisation nous fournit alors simplement une base orthonormée de $G_N = \mathrm{Vect}(f_0, \ldots, f_N)$ et les espaces G_N et $\mathrm{Vect}(e_0, \ldots, e_N)$ coïncident.

□

On admet le résultat suivant qui relève du théorème de Zorn.

Théorème 2.4.4. *Tout espace de Hilbert non réduit à $\{0\}$ possède une base hilbertienne.*

Remarque 2.4.5. *Comme on a besoin du lemme de Zorn (axiome de choix), le procédé n'est plus constructif, ce qui n'a pas un grand intérêt pratique. Par ailleurs, en utilisant ce même lemme de Zorn, on peut construire une base algébrique de H vu comme un simple espace vectoriel. Les deux n'ont aucune raison de coïncider : on peut par exemple montrer que toute base algébrique de ℓ^2 a un cardinal qui est la puissance du continu (le cardinal de $\mathbb{R}$). Notons qu'il n'est pas très difficile de montrer qu'une telle base ne peut avoir un cardinal dénombrable : en effet, si c'est le cas, l'espace ℓ^2 est la réunion des espaces vectoriels de dimension finie engendrés par les N premiers vecteurs de la base. Ces espaces sont d'intérieur vide, donc leur réunion est aussi d'intérieur vide (conséquence du théorème de Baire). Mais ℓ^2 n'est certainement pas d'intérieur vide...*

Proposition 2.4.6. *Tout espace de Hilbert dénombrable est isométrique à $\ell^2(\mathbb{N})$.*

Démonstration : Il suffit de mettre en relation les bases hilbertiennes pour obtenir une isométrie.

□

On termine cette partie par un résultat qui peut être très pratique pour la suite.

Proposition 2.4.7 (Inégalité de Bessel). *Si $(e_j)_{j\in J}$ est une famille orthonormée d'éléments de $\mathcal{H}$, alors, pour tout $x \in \mathcal{H}$, la famille $\big(|(x\,|\,e_j)|^2\big)_{j\in J}$ est sommable dans $\mathbb{R}$ et on a*

$$\sum_{j\in J} |(x\,|\,e_j)|^2 \leqslant \|x\|^2 .$$

Démonstration : Pour toute partie finie $F \subset J$, posons $y = x - \sum_{k\in F}(x\,|\,e_k)e_k$. Alors $(y\,|\,e_l) = 0$ pour tout $l \in F$, donc

$$\|x\|^2 = \|y\|^2 + \Big\|\sum_{k\in F}(x\,|\,e_k)e_k\Big\|^2 = \|y\|^2 + \sum_{k\in F} |(x\,|\,e_k)|^2 \geqslant \sum_{k\in F} |(x\,|\,e_k)|^2.$$

L'ensemble des sommes finies d'eléments de la famille de nombres réels positifs $\big(|(x\,|\,e_j)|^2\big)_{j\in J}$ est donc majoré. Cette famille est donc sommable dans $\mathbb{R}$ et on a bien $\sum_{j\in J} |(x\,|\,e_j)|^2 \leqslant \|x\|^2$.

□

2.4.2 Projection sur une base hilbertienne

On généralise ici la notion de coordonnées dans une base algébrique en dimension finie au cas des coordonnées dans une bases hilbertienne en dimension infinie.

Théorème 2.4.8. *Soit $(e_j)_{j\in J}$ une famille orthonormée d'éléments de $\mathcal{H}$. On note $F_J = \overline{vect\big((e_j)_{j\in J}\big)}$ et Π_{F_J} la projection orthogonale sur le sous-espace vectoriel fermé F_J. Alors, pour tout $x \in \mathcal{H}$, la famille $\big((x|e_j)e_j\big)_{j\in J}$ est sommable dans $\mathcal{H}$ et on a*

$$\sum_{j\in J}(x\,|\,e_j)e_j = \Pi_{F_J}(x).$$

Démonstration :
1. *Sommabilité de $\big((x\,|\,e_j)e_j\big)_{j\in J}$.* On va montrer que la famille $\big((x|e_j)e_j\big)_{j\in J}$ est sommable dans $\mathcal{H}$, qui est complet, en montrant qu'elle vérifie la condition de Cauchy. Soit $\varepsilon > 0$. D'après l'inégalité de Bessel, la famille $\big(|(x\,|\,e_j)|^2\big)_{j\in J}$ est sommable dans $\mathbb{R}$. Elle vérifie donc la condition de Cauchy : il existe une partie finie J_0 de J telle que, pour toute partie finie $J_1 \subset J$ ne rencontrant pas J_0, $\sum_{j\in J_1} |(x\,|\,e_j)|^2 < \varepsilon^2$. Puisque la famille $(e_j)_{j\in J}$ est orthonormée, on a

$$\|\sum_{j\in J_1}(x\,|\,e_j)e_j\|^2 = \sum_{j\in J_1}\big(|(x|e_j)|^2\|e_j\|^2\big) = \sum_{j\in J_1}|(x\,|\,e_j)|^2$$

et on en déduit que

$$\Big\|\sum_{j\in J_1}(x\,|\,e_j)e_j\Big\| = \sqrt{\sum_{j\in J_1}|(x\,|\,e_j)|^2} < \varepsilon.$$

Donc $\big((x\,|\,e_j)e_j\big)_{j\in J}$ vérifie la condition de Cauchy et est donc sommable.

2. *Détermination de la somme.* La famille $\big((x\,|\,e_j)e_j\big)_{j\in J}$étant sommable, une quantité au plus dénombrable des $(x\,|\,e_j)e_j$, donc des $(x\,|\,e_j)$, sont non nuls. Par définition de $F_J = \overline{vect\big((e_j)_{j\in J}\big)}$, on a donc bien $\sum_{j\in J}(x\,|\,e_j)e_j \in F_J$. De plus, si $y \in F_J$, alors $y = \sum_{l\in J} y_l e_l$ avec une quantité au plus dénombrable des y_l qui soient non nuls et on a donc, par

continuité de $(.\,|\,.)$, pour tout $j \in J$, $(e_j\,|\,y) = \sum_{l\in J} \overline{y_l}(e_j\,|\,e_l) = \overline{y_j}$. Par suite, toujours par continuité de $(.\,|\,.)$, on a

$$\begin{aligned}\Big(x - \sum_{j\in J}(x\,|\,e_j)e_j\,\Big|\,y\Big) = (x\,|\,y) - \sum_{j\in J}(x\,|\,e_j)(e_j\,|\,y) &= (x\,|\,y) - \sum_{j\in J}(x\,|\,e_j)\overline{y_j}\\ &= (x\,|\,y) - \sum_{j\in J}(x\,|\,y_je_j)\\ &= (x\,|\,y) - (x\,|\,\sum_{j\in J}y_je_j)\\ &= (x\,|\,y) - (x\,|\,y) = 0,\end{aligned}$$

et ceci pour tout $y \in F_J$. Donc $x - \sum_{j\in J}(x\,|\,e_j)e_j \in F_J^{\perp}$. D'après le théorème 2.2.2, on a donc bien $\sum_{j\in J}(x\,|\,e_j)e_j = \Pi_{F_J}(x)$.

□

Lorsque l'on applique ce théorème au cas d'une base hilbertienne on obtient le résultat suivant qui généralise directement la notion de coordonnées en dimension finie.

Corollaire 2.4.9. *Soit $(e_i)_{i\in I}$ une base hilbertienne de $\mathcal{H}$. Pour tout $x \in \mathcal{H}$, la famille $\big((x\,|\,e_i)e_i\big)_{i\in I}$ est sommable dans $\mathcal{H}$ et on a*

$$\sum_{i\in I}(x\,|\,e_i)e_i = x.$$

On dit que les $x_i := (x|e_i)$ sont les coordonnées *de x dans la base hilbertienne $(e_i)_{i\in I}$.*

Démonstration : C'est le théorème 2.4.8 appliqué dans le cas où la famille $(e_j)_{j\in J}$ est en outre totale, de sorte qu'ici $F_J = \mathcal{H}$ et $\Pi_{F_J} = Id_{\mathcal{H}}$.

□

Ainsi, étant donnée une base hilbertienne $(e_i)_{i\in I}$ de $\mathcal{H}$, tout vecteur x de $\mathcal{H}$ est *limite* d'une suite bien déterminée d'éléments de $vect\big((e_i)_{i\in I}\big)$. Plus généralement, si F est un sous-espace vectoriel fermé de $\mathcal{H}$ et $\big((e_j)_{j\in J}\big)$ une base hilbertienne de F, le théorème 2.4.8 donne une formule explicite pour le projeté orthogonal sur F d'un vecteur x de $\mathcal{H}$. Ces formules généralisent des formules déjà connues en dimension finie : les sommes finies y sont remplacées par des limites.

2.4.3 Théorème de Parseval

Corollaire 2.4.10. *Soit $(e_i)_{i\in I}$ une base hilbertienne de $\mathcal{H}$.*

1. *Pour tous $x, y \in \mathcal{H}$, la famille $\big((x\,|\,e_i)(e_i\,|\,y)\big)_{i\in I}$ est sommable dans $\mathbb{K}$ et on a*

$$\sum_{i\in I}(x\,|\,e_i)(e_i\,|\,y) = (x\,|\,y).$$

   En coordonnées dans la base hilbertienne $(e_i)_{i\in I}$, cela s'écrit

$$(x|y) = \sum_{i\in I} x_i\overline{y_i}.$$

2. *Pour tout $x \in \mathcal{H}$, la famille $\big(|(x\,|\,e_i)|^2\big)_{i\in I}$ est sommable dans $\mathbb{R}$ et on a*

$$\sum_{i\in I}|(x\,|\,e_i)|^2 = \|x\|^2 \quad \textit{(égalité de Parseval).}$$

Démonstration : 1. On a, pour tout $i \in I$, $|(x \mid e_i)(e_i \mid y)| \leqslant \frac{1}{2}\left(|(x \mid e_i)|^2 + |(e_i \mid y)|^2\right)$. Or les familles $\left(|(x \mid e_i)|^2\right)_{i\in I}$ et $\left(|(e_i \mid y)|^2\right)_{i\in I}$ sont sommables d'après l'inégalité de Bessel (voir Proposition 2.4.7). Donc $\left((x \mid e_i)(e_i \mid y)\right)_{i\in I}$ est absolument sommable dans $\mathbb{K}$, donc sommable dans $\mathbb{K}$. Puisque, d'après le corollaire 2.4.9, $x = \sum_{i\in I}(x \mid e_i)e_i$ et $y = \sum_{j\in I}(y \mid e_j)e_j$ (avec une quantité au plus dénombrable de termes non nuls dans chacune de ces sommes), on a, par continuité de $(. \mid .)$,

$$(x \mid y) = \sum_{i\in I}\sum_{j\in I}(x \mid e_i)\overline{(y \mid e_j)}(e_i \mid e_j) = \sum_{i\in I}(x \mid e_i)(e_j \mid y_j).$$

2. C'est le point 1 lorsque $x = y$.

□

Ce résultat n'est autre que le théorème de Parseval que vous connaissez pour les série de Fourier énoncé dans le cadre général des bases hilbertienne. Il sera rappelé dans le prochain chapitre.

Chapitre 3

Séries de Fourier

3.1 Série de Fourier d'une fonction périodique

L'objectif de ce chapitre est de montrer que l'on peut, sous certaines hypothèses, décomposer toute fonction périodique comme une somme (infinie) de fonctions périodiques élémentaires, les $x \mapsto \mathrm{e}^{inx}$ pour $n \in \mathbb{Z}$ ou les $x \mapsto \cos(nx)$ et $x \mapsto \sin(nx)$ pour $n \in \mathbb{N}$.

Ainsi, on souhaite pouvoir écrire, pour $f : \mathbb{R} \to \mathbb{C}$, 2π-périodique,

$$\forall x \in \mathbb{R},\ f(x) = \sum_{n \in \mathbb{Z}} c_n \mathrm{e}^{\mathrm{i}nx}$$

ou

$$\forall x \in \mathbb{R},\ f(x) = a_0 + \sum_{n \geqslant 1} a_n \cos(nx) + b_n \sin(nx).$$

L'idée de cette décomposition est de Joseph Fourier (1768-1830) qui l'utilisa dans son étude de l'équation de la chaleur.

Si on considère un polynôme trigonométrique

$$P = \sum_{n=0}^{N} a_n \cos(n\cdot) + b_n \sin(n\cdot)$$

ou une série trigonométrique

$$\sum_{n \geqslant 0} a_n \cos(n\cdot) + b_n \sin(n\cdot)$$

qui converge uniformément sur une période (par exemple $[-\pi, \pi]$) vers f, on peut essayer de trouver une expression des coefficients a_n et b_n en fonction de P ou de f.

Nous allons pour cela utiliser la propriété d'orthogonalité de la famille $\{(\cos(n\cdot), \sin(n\cdot)\}_{n \geqslant 0}$.

Soit $m \neq 0$. Par linéarité de la somme et convergence uniforme (qui permet d'intégrer la série terme à terme) pour calculer

$$\int_{-\pi}^{\pi} f(x) \cos(mx) \mathrm{d}x$$

il suffit de calculer pour tout n,

$$\int_{-\pi}^{\pi} (a_n \cos(nx)) \cos(mx) \mathrm{d}x = a_n \int_{-\pi}^{\pi} \frac{1}{2} (\cos(m+n)x + \cos(m-n)x) \mathrm{d}x$$

Si $m \neq n$, on trouve 0. Si $m = n$, on trouve

$$\int_{-\pi}^{\pi} (a_n \cos(nx)) \cos(nx) \mathrm{d}t = a_n \int_{-\pi}^{\pi} \frac{1}{2}(\cos(2nx) + 1) \mathrm{d}t = \pi a_n.$$

Un calcul analogue montre que les termes de la forme

$$\int_{-\pi}^{\pi} (b_n \sin(nx)) \cos(mx) \mathrm{d}x$$

sont tous nuls.

Si $m = 0$, on obtient directement que $\int_{-\pi}^{\pi} f(x) \mathrm{d}x = 2\pi a_0$.

On a donc

$$\int_{-\pi}^{\pi} f(x) \cos(mx) \mathrm{d}x = \begin{cases} \pi a_m & \text{si} \quad m \neq 0 \\ 2\pi a_0 & \text{si} \quad m = 0 \end{cases}$$

On peut faire de même pour trouver les coefficients b_n en multipliant par $\sin(mx)$ dans l'intégrale.

On peut effectuer le même calcul pour obtenir les coefficients complexe c_n. Cela conduit à poser :

Définition 3.1.1. *Soit f une fonction continue par morceaux et périodique de période 2π. Les coefficients de Fourier complexes de f sont :*

$$\forall n \in \mathbb{Z}, \; c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) \mathrm{e}^{-\mathrm{i}nt} \mathrm{d}t.$$

Les coefficients de Fourier réels de f sont :

$$a_0(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \mathrm{d}t, \; \forall n \geqslant 1, \; a_n(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) \mathrm{d}t$$

$$\text{et} \;\; \forall n \geqslant 1, \; b_n(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) \mathrm{d}t.$$

Les coefficients de Fourier d'une fonction périodique ont un comportement asymptotique qui dépend de la régularité de f.

Lemme 3.1.2 (Riemann-Lebesgue). *Soit f une fonction intégrable sur un intervalle $[a, b]$ (ou sur $\mathbb{R}$). Alors*

$$\int_a^b f(x) \mathrm{e}^{\mathrm{i}nx} \mathrm{d}x \xrightarrow[n \to +\infty]{} 0.$$

Appliqué à une fonction f continue par morceaux et 2π-périodique, on obtient que :

$$c_n(f) \xrightarrow[|n| \to +\infty]{} 0$$

et

$$a_n(f) \xrightarrow[n \to +\infty]{} 0, \quad \text{et} \quad b_n(f) \xrightarrow[n \to +\infty]{} 0.$$

On peut aller plus loin, on considérant des fonctions de classe C^p.

Proposition 3.1.3. *Soit f une fonction 2π-périodique et de classe C^p sur $\mathbb{R}$. Alors, à l'infini,*

$$c_n(f) = o\left(\frac{1}{|n|^p}\right),$$

$$a_n(f) = o\left(\frac{1}{n^p}\right) \quad et \quad b_n(f) = o\left(\frac{1}{n^p}\right).$$

Ce qu'il faut retenir, c'est que le comportement asymptotique des coefficients de Fourier est directement lié à la régularité de f.

Définition 3.1.4. *Soit f une une fonction continue par morceaux et périodique de période 2π.*

1. *La série de Fourier complexe de f est la série indexée par $\mathbb{Z}$,*

$$\sum_{n\in\mathbb{Z}} c_n(f)\mathrm{e}^{-\mathrm{i}n\cdot}.$$

2. *La série de Fourier réelle de f est la série*

$$a_0(f) + \sum_{n\geqslant 1} a_n(f)\cos(n\cdot) + b_n(f)\sin(n\cdot).$$

3. *Pour tout $n \in \mathbb{N}$ et tout $x \in \mathbb{R}$, on pose*

$$S_n(f)(x) = \sum_{k=-n}^{n} c_k(f)\mathrm{e}^{\mathrm{i}kx}$$

dans le cadre complexe et

$$S_n(f)(x) = \sum_{k=0}^{n} a_k(f)\cos(kx) + b_k(f)\sin(kx)$$

dans le cadre réel. Ce sont les sommes partielles des séries de Fourier complexes et réelles respectivement.

3.2 Convergence ponctuelle des séries de Fourier

A l'aide des formules des coefficients de Fourier, on peut associer à chaque fonction périodique une suite de coefficients. Pour que cette notion soit utile, il faut pouvoir "reconstruire" la fonction périodique à partir uniquement de la donnée de ses coefficients de Fourier.

On se pose donc la question de la convergence de la série de Fourier (réelle ou complexe) d'une fonction périodique f vers la fonction f.

Cette question admet plusieurs réponses, suivant la régularité de f et suivant le type de convergence voulue. Nous commençons par donner une réponse en terme de convergence simple et uniforme, nous verrons ensuite ce qu'il se passe pour la convergence en moyenne quadratique.

Théorème 3.2.1 (Dirichlet 1). *Soit f une fonction 2π-périodique et C^1 par morceaux. Alors sa série de Fourier converge simplement vers $\tilde{f}$ la régularisée de f définie par :*

$$\forall x \in \mathbb{R},\ \tilde{f}(x) = \frac{1}{2}(f(x^+) + f(x^-)).$$

Démonstration : On considère la somme partielle

$$S_N(f)(x) = \sum_{|n|\leqslant N} \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{in(x-t)}dt = \frac{1}{2\pi}\int_0^{2\pi} f(t)\left(\sum_{|n|\leqslant N} e^{in(x-t)}\right)dt,$$

et en réexprimant la somme trigonométrique, on obtient

$$S_N(f)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\frac{\sin(n+\frac{1}{2})(x-t)}{\sin(\frac{x-t}{2})}dt.$$

Par changement de variable,

$$S_N(f)(x) = \frac{1}{2\pi}\int_x^{x+2\pi} f(x-u)\frac{\sin(n+\frac{1}{2})u}{\sin(\frac{u}{2})}du,$$

et par périodicité,

$$S_N(f)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(x-u)\frac{\sin(n+\frac{1}{2})u}{\sin(\frac{u}{2})}du.$$

Notons qu'on peut aussi intégrer sur $]-\pi,\pi[$ et symétriser, pour obtenir

$$S_N(f)(x) = \frac{1}{2\pi}\int_0^{\pi}(f(x+u)+f(x-u))\frac{\sin(n+\frac{1}{2})u}{\sin(\frac{u}{2})}du.$$

C'est la formule de Dirichlet. Montrons donc la convergence simple : si on prend $f = 1$ dans la formule ci-dessus, il vient

$$S_N(1) = 1 = \frac{1}{2\pi}\int_0^{\pi} 2\frac{\sin(n+\frac{1}{2})u}{\sin(\frac{u}{2})}du.$$

Donc

$$\begin{aligned} R_N(f) =& S_N(f) - \frac{1}{2}(f(x^+)+f(x^-)) \\ =& \frac{1}{2\pi}\int_0^{\pi}(f(x+u)+f(x-u)-f(x^+)-f(x^-))\frac{\sin(n+\frac{1}{2})u}{\sin(\frac{u}{2})}du. \end{aligned}$$

Appelons $g(u) = (f(x+u)+f(x-u)-f(x^+)-f(x^-))\,/\sin(\frac{u}{2})$. Cette fonction est continue par hypothèse (écrire un DL de g en $u = 0$), donc on peut appliquer le lemme de Riemann-Lebesgue et conclure à la convergence simple.

□

Ainsi, sous les hypotèses du théorème de Dirichlet, en tout point de continuité de f, sa série de Fourier converge simplement vers f. En tout point de discontinuité de f, sa série de Fourier converge vers la moyenne de ses limites à gauche et à droite (celles-ci existent car la fonction est C^1 par morceaux donc continue par morceaux).

Proposition 3.2.2. *Si $(c_n(f))_{n\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$, alors $S_N(f) = \sum_{|n|\leqslant N} c_n(f)e_n$ converge uniformément vers f sur $\mathbb{R}$.*

Démonstration : Il s'agit d'une application directe du critère de Cauchy uniforme.

□

Remarque 3.2.3. *Cette condition est facilement vérifiée si par exemple $f \in C^2$, puisqu'alors $c_n(f) = \mathcal{O}(\frac{1}{n^2})$.*

Théorème 3.2.4 (Dirichlet 2)**.** *Soit f une fonction 2π-périodique, continue et C^1 par morceaux. Alors sa série de Fourier converge uniformément vers f sur tout compact.*

Démonstration : Pour simplifier, on fera juste C^1, sinon il faut découper en intervalles sur lesquels la fonction est C^1. On intègre par parties le coefficient,

$$c_n(f) = \frac{1}{2\pi in}\int_0^{2\pi} f'(t)e^{-int}dt = \frac{1}{in}c_n(f').$$

On utilise Parseval puisque f' est continue, donc dans $L^2([0,2\pi])$. Alors $\sum_n |c_n(f')|^2$ converge, donc en utilisant Cauchy-Schwarz, on en déduit la convergence dans ℓ^1 des $c_n(f)$. On applique enfin la proposition précédente.

□

Le simple fait de rajouter la continuité de f fait que la convergence n'est pas seulement simple vers f, mais elle devient uniforme.

Ce théorème donne un excellent cadre à la reconstruction des fonctions car la convergence uniforme permet aussi, par exemple, de faire des intégrations terme à terme de la série de Fourier.

Théorème 3.2.5 (Fejer)**.** *Soient $n \geqslant 1$, $x \in \mathbb{R}$. On pose*

$$\sigma_n(f)(x) = \frac{S_1(f)(x) + S_2(f)(x) + ... + S_n(f)(x)}{n}$$

que l'on appelle la moyenne de Césaro de la suite $(S_n(f))_{n\geqslant 0}$.

1. *Si f est continue, $(\sigma_n(f))_{n\geqslant 1}$ converge uniformément vers f.*

2. *De plus,*

$$\int_{-\pi}^{\pi} |\sigma_n(f)(x) - f(x)|^2 \mathrm{d}x \xrightarrow[n\to+\infty]{} 0.$$

Démonstration : En utilisant la formule de Dirichlet, il vient

$$\begin{aligned}\sigma_n(f) &= \frac{1}{2\pi}\int_0^{\pi}(f(x+u)+f(x-u))\sum_0^{n-1}\frac{\sin(k+\frac{1}{2})u}{\sin(\frac{u}{2})}du\\ &= \frac{1}{2\pi n}\int_0^{\pi}(f(x+u)+f(x-u))\frac{\sin^2(\frac{nu}{2})}{\sin^2(\frac{u}{2})}du.\end{aligned}$$

On a gagné le caractère positif de la fonction sous le signe somme. Par continuité uniforme,

$$\forall\varepsilon > 0, \forall x \in \mathbb{R}, \exists\delta,\ \forall u \in \mathbb{R},\ |f(x+u) + f(x-u) - 2f(x)| < \varepsilon.$$

Donc, en posant $h(x,u) = f(x+u) + f(x-u) - 2f(x)$

$$\left|\frac{1}{2\pi n}\int_0^\delta h(x,u)\frac{\sin^2(\frac{nu}{2})}{\sin^2(\frac{u}{2})}du\right| \leqslant \frac{\varepsilon}{2\pi n}\int_0^\delta \frac{\sin^2(\frac{nu}{2})}{\sin^2(\frac{u}{2})}du,$$

donc

$$\left|\frac{1}{2\pi n}\int_0^\delta h(x,u)\frac{\sin^2(\frac{nu}{2})}{\sin^2(\frac{u}{2})}du\right| \leqslant \frac{\varepsilon}{2\pi n}\int_0^\pi \frac{\sin^2(\frac{nu}{2})}{\sin^2(\frac{u}{2})}du = \frac{\varepsilon}{2},$$

en se souvenant que $\sigma_n(1) = 1$. Pour l'autre partie, on remarque que $h(x,u)$ est une fonction bornée (disons par M), et on minore le $\sin^2(\frac{nu}{2})$ par $\sin^2(\frac{\delta}{2})$:

$$\left|\frac{1}{2\pi n}\int_\delta^{2\pi} h(x,u)\frac{\sin^2(\frac{nu}{2})}{\sin^2(\frac{u}{2})}du\right| \leqslant \frac{1}{2\pi n\sin^2(\delta/2)}\int_\delta^\pi |h(x,u)|du \leqslant \frac{C}{n},$$

d'où l'on en déduit la convergence uniforme et le second point.

□

Le théorème de Féjer nous dit donc que si f n'est pas C^1 par morceaux mais tout de même continue, on peut conserver une convergence uniforme de la série de Fourier vers sa somme, mais avec une convergence "au sens de Césaro".

Une conséquence du théorème de Féjer est que toute fonction continue et périodique sur $\mathbb{R}$ est limite uniforme de polynômes trigonométriques. Par restriction, toute fonction continue sur un intervalle compact est encore limite uniforme de polynômes trigonométriques. Il s'agit du théorème de Weierstraß trigonométrique.

3.3 Théorie L^2 des séries de Fourier

Nous allons voir dans ce chapitre comment les résultat obtenus au chapitre précédent sur les bases hilbertiennes s'appliquent directement au cas de l'espace de Hilbert $L^2_{\mathbb{C}}([0,2\pi])$ pour donner les résultats classiques sur la théorie L^2 des séries de Fourier.

3.3.1 Base hilbertienne de $L^2_{\mathbb{C}}([0,2\pi])$

On rappelle que le $\mathbb{C}$-espace vectoriel $L^2_{\mathbb{C}}([0,2\pi])$ des fonctions de carré intégrable sur $[0,2\pi]$ et à valeurs dans $\mathbb{C}$ muni du produit scalaire

$$\forall f,g \in L^2_{\mathbb{C}}([0,2\pi]),\ (f|g) = \int_0^{2\pi} f(x)\overline{g(x)}\mathrm{d}x,$$

est un espace de Hilbert. On note $|| \ ||_2$ la norme associée à son produit scalaire.

Définition 3.3.1. *Un polynôme trigonométrique est une fonction de la forme*

$$\forall x \in \mathbb{R},\ f(x) = \sum_{n=-N}^{N} a_n \mathrm{e}^{\mathrm{i}nx},\ \text{avec } a_n \in \mathbb{C} \text{ et } N \in \mathbb{N}.$$

Théorème 3.3.2. *Soit $\mathcal{P}$ le $\mathbb{C}$-espace vectoriel des polynômes trigonométrique sur $\mathbb{R}$. Alors $\mathcal{P}$ est dense dans $L^2_{\mathbb{C}}([0,2\pi])$ pour la norme induite par le produit scalaire.*

Démonstration : $\mathcal{P}$ est dense dans les fonctions continues sur $[0, 2\pi]$ et on peut montrer que cet espace de fonctions continues est dense dans $L^2_{\mathbb{C}}([0, 2\pi])$.

□

On déduit de ce théorème la construction d'une base hilbertienne explicite pour $L^2_{\mathbb{C}}([0, 2\pi])$.

Proposition 3.3.3. *On définit, pour tout $n \in \mathbb{Z}$, la fonction e_n par*

$$\forall x \in [0, 2\pi],\ e_n(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{inx}.$$

Alors $(e_n)_{n\in\mathbb{Z}}$ est une base hilbertienne de $\big(L^2_{\mathbb{C}}([0, 2\pi]), (.\,|\,.)\big)$.

Démonstration : Tout d'abord, il est clair que $||e_n||_2 = 1$ puisque $|e_n(x)| = 1$ pour tout x. Puis, si $m \neq n$, on a

$$(e_m|e_n) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{imx} \mathrm{e}^{-inx} \mathrm{d}x = 0.$$

Donc $(e_n)_{n\in\mathbb{Z}}$ est une famille orthonormée. Elle est de plus totale d'après le théorème 3.3.2.

□

3.3.2 Convergence L^2 et théorème de Parseval

On va maintenant appliquer les résultats du chapitre précédent à la base hilbertienne $(e_n)_{n\in\mathbb{Z}}$. On remarque que l'on peut interpréter les coefficients de Fourier en terme de produit scalaire. En effet, il est immédiat que :

$$\forall n \in \mathbb{Z},\ c_n(f) = \frac{1}{\sqrt{2\pi}} (f|e_n).$$

On remarque également que

$$\forall N \in \mathbb{N},\ S_N(f) = \sum_{n=-N}^{N} (f|e_n) e_n.$$

On peut donc interpréter $S_N f$ comme étant la projection orthogonale sur le sous-espace vectoriel de dimension finie vect$(e_n,\ -N \leqslant n \leqslant N)$ (qui est donc fermé).

On peut montrer en utilisant directement les résultats du chapitre précédent que la suite des sommes partielles de Fourier de f converge vers f dans $L^2_{\mathbb{C}}([0, 2\pi])$ et que l'on a une relation de Parseval entre la norme L^2 de f et la somme des modules au carré des coefficients de Fourier de f.

Théorème 3.3.4. *Soient f et g dans $L^2_{\mathbb{C}}([0, 2\pi])$. Alors,*

1.
$$\lim_{N\to\infty} \|f - S_N f\|_2^2 = 0,$$

2.
$$(f|g) = 2\pi \sum_{n\in\mathbb{Z}} c_n(f)\overline{c_n(g)},$$

3.

$$\sum_{n\in\mathbb{Z}} |c_n(f)|^2 = \frac{1}{2\pi}||f||_2^2.$$

Démonstration : Le point 1 n'est autre que le corollaire 2.4.9 appliqué à la base hilbertienne $(e_n)_{n\in\mathbb{Z}}$. Les points 2 et 3 sont les points 1 et 2 du corollaire 2.4.10 eux aussi appliqués à la base hilbertienne $(e_n)_{n\in\mathbb{Z}}$.

□

3.4 Espaces de Sobolev

De nombreux problèmes issus de la physique mathématique se modélisent par des équations différentielles ou plus généralement des équations aux dérivées partielles. Le cadre fonctionnel des espaces de Hilbert est très pratique pour aborder ce type de questions, mais les espaces de fonctions comme C^1, C^n ne sont pas hilbertiens, et L^2 n'est pas suffisant. Ceci motive l'introduction d'espaces hilbertiens qui se comportent "bien" par rapport aux dérivations.

Définition 3.4.1. *Soit $I =]a, b[$ un intervalle (non nécessairement borné). On appelle espace de Sobolev $H^1(I)$ l'ensemble des fonctions continues et de carré intégrable sur I, telles que $f \in H^1(I)$ puisse s'écrire*

$$f(x) = C + \int_c^x g(t)dt,$$

où $C \in \mathbb{C}$, $c \in I$, et $g \in L^2(I)$.

Proposition 3.4.2. *La fonction g associée à f dans la définition précédente est unique dans L^2.*

Démonstration : Si $f(x) = C' + \int_{c'}^x h(t)dt$, alors, pour $c \leqslant c'$,

$$\int_{c'}^x (g-h)dt = C' - C + \int_{c'}^c g(t)dt = 0,$$

en prenant $x = c'$. Donc $g - h$ est orthogonale à toutes les indicatrices $\mathbf{1}_{[c',x]}$ donc par combinaison linéaire à toutes les fonctions en escalier de I. Par densité de ces dernières, on en déduit que $g - h = 0$.

□

Cela permet de dire que g est la dérivée faible (ou généralisée, ou au sens L^2) de f, on la notera par abus de notation, f'.

Proposition 3.4.3. *Pour toute fonction $f \in H^1(I)$, on a*

$$|f(x) - f(y)| \leqslant \sqrt{|x-y|}\|f'\|_2 \quad \text{et } \|f\|_\infty \leqslant (\|f\|_2 + \|f'\|_2).$$

Démonstration : La première inégalité est une conséquence de l' inégalité de Schwarz appliquée à

$$f(x) - f(y) = \int_y^x f'(t)dt.$$

On remarque donc que f est non seulement continue, mais hölderienne d'exposant $\frac{1}{2}$. Ensuite, recouvrons I par des intervalles finis J. On a, pour $x, y \in J$

$$|f(x)| \leqslant |f(y)| + \int_y^x |f'(t)|dt \leqslant |f(y)| + |J|^{\frac{1}{2}}\|f'\|_2,$$

donc en intégrant sur J par rapport à y, on obtient

$$|J||f(x)| \leqslant \int_J |f(y)|dy + |J|^{\frac{3}{2}}\|f'\|_2 \leqslant |J|^{\frac{1}{2}}\|f\|_2 + |J|^{\frac{3}{2}}\|f'\|_2$$

donc,

$$\sup_{x\in J} |f(x)| \leqslant |J|^{-\frac{1}{2}}\|f\|_2 + |J|^{\frac{1}{2}}\|f'\|_2.$$

En optimisant sur $|J|$ on obtient, par l'inégalité arithmético-géométrique,

$$\|f\|_\infty \leqslant 2\sqrt{\|f\|_2\|f'\|_2} \leqslant \|f\|_2 + \|f'\|_2.$$

Notons aussi que la première propriété assure que si a (resp. b) est fini, f a une limite finie quand $x \to a$ (resp. $x \to b$) (en utilisant le critère de Cauchy).

□

Théorème 3.4.4. *L'espace $H^1(I)$ est un espace de Hilbert pour le produit scalaire*

$$(f|g)_{H^1} = (f|g)_{L^2} + (f'|g')_{L^2}.$$

Démonstration : On vérifie aisément qu'il s'agit bien d'un produit scalaire et qu'ainsi $H^1(I)$ est préhilbertien. Reste à montrer qu'il est complet. Considérons une suite de Cauchy (f_n) dans $H^1(I)$: compte-tenu de la définition de la norme associée au produit scalaire, (f_n) et (f_n') sont de Cauchy dans $L^2(I)$, donc elles convergent dans $L^2(I)$ vers des limites f et g. Pour montrer la convergence dans H^1, il faut montrer que f est continue et que g est sa dérivée faible. En utilisant la proposition précédente, on voit que (f_n) est une suite de Cauchy pour la norme uniforme, donc elle converge uniformément vers une fonction continue (qui ne peut être que f par unicité de la limite). Par ailleurs,

$$\left|\int_y^x f_n' - \int_y^x g\right| \leqslant |x-y|^{\frac{1}{2}}\|f_n' - g\|_2$$

donc on peut passer à la limite dans

$$f_n(x) = f_n(y) + \int_y^x f_n',$$

ce qui achève la preuve.

□

Remarque 3.4.5. *En utilisant la proposition précédente, on voit que l'injection de $H^1(I)$ dans $C_b(I)$ (les fonctions continues bornées sur I) est continue, puisque*

$$\|f\|_\infty \leqslant \sqrt{\|f\|_2^2 + \|f'\|_2^2} = \|f\|_{H^1}.$$

Proposition 3.4.6 (Intégration par parties)**.** *Considérons $f, g \in H^1(I)$, $\alpha < \beta \in \bar{I}$. Alors*

$$\int_\alpha^\beta f'g + g'f = [fg]_\alpha^\beta,$$

et $fg \in H^1$, avec

$$(fg)' = f'g + g'f.$$

Démonstration : Supposons que α et β sont finis : on a

$$f(x) = f(\beta) - \int_x^\beta f'(t)dt, \text{ et } g(x) = g(\alpha) + \int_\alpha^x g'(t)dt.$$

On calcule l'intégrale de la fonction de deux variables $f'(x)g'(y)$ sur le domaine $D = \{\alpha < y < x < \beta\}$ (la fonction est bien dans $L^1(D)$). Par Fubini, on peut intégrer d'abord par rapport à x ou par rapport à y, et les deux sont égales, ce qui donne la formule demandée. La formule de dérivation du produit en est alors une conséquence. Si l'un des α, β est infini, on veut passer à la limite : pour cela, montrons que lorsque a ou b est infini, alors nécessairement $f \in H^1(I)$ a une limite nulle quand $x \to a$ (ou b) : prenons $c \in I$, on vient de montrer que

$$f(x)^2 - f(c)^2 = 2\int_c^x f(t)f'(t)dt,$$

et l'intégrale de droite est absolument convergente lorsque $x \to a$ (par Cauchy-Schwarz), donc f a une limite quand $x \to a$. Mais cette limite est forcément nulle, ou f ne serait pas de carré intégrable.

□

Définition 3.4.7. *On appelle $H_0^1(I)$ le sous-espace de $H^1(I)$ constitué des fonctions qui s'annulent aux extrémités finies de H^1.*

Remarquons que compte-tenu de la continuité de l'injection de $H^1(I)$ dans $C_b(I)$, $H_0^1(I)$ est fermé, c'est donc un sous-espace hilbertien.

Proposition 3.4.8. *Soient f et g deux fonctions de $L^2(I)$ telles que*

$$\forall h \in H_0^1(I), \int_I f(x)w'(x)dx = -\int_I g(x)w(x)dx,$$

alors f est égale presque partout à une fonction de $H^1(I)$ (noté aussi f par abus de notation), et $f' = g$.

Démonstration : On la fera uniquement pour I fini. Supposons que $I = [a,b]$, notons $G(x) = \int_a^x g(t)dt$. On intègre par parties dans le membre de droite, et $\int_I (f(x) - G(x))w'(x)dx = 0$. Notons $C = (b-a)^{-1}\int_a^b (f-G)(x)dx$. Alors pour tout $c \in \mathbb{C}$ et toute $w \in H_o^1$, on a

$$\int_I ((f-G)(x) - C)(c + w'(x))dx = 0.$$

Il suffit alors de voir que toute fonction $h \in L^2(I)$ peut s'écrire $c + w'$ avec $w \in H_0^1(I)$. Une telle fonction w est donnée par :

$$w(x) = \int_a^x h(t)dt - \frac{x-a}{b-a}\int_a^b h(t)dt.$$

Finalement, $f(x) - G(x) - C$ est orthogonale à toutes les fonctions de L^2, donc $f = G + C$ presque partout. Dans le cas où I n'est pas fini, on peut refaire le raisonnement sur tout sous-intervalle fini $[\alpha, \beta]$ et vérifier que la constante K est indépendante de β.

□

Définition 3.4.9. *On appelle $H_p^1(0, 2\pi)$ le sous-espace de $H^1(0, 2\pi)$ constitué des fonctions f qui vérifient $f(0) = f(2\pi)$.*

C'est, comme $H_0^1(I)$, un sous-espace hilbertien. On peut alors faire le lien entre cet espace et les séries de Fourier.

Théorème 3.4.10. *Une fonction f 2π-périodique est dans $H_p^1(0,2\pi)$ si et seulement si ses coefficients de Fourier vérifient*

$$\sum_n n^2 |c_n(f)|^2 < \infty.$$

Démonstration : Remarquons que la suite $(e_n^1)_n$ des fonctions $e_n^1 = (1+n^2)^{-\frac{1}{2}} e_n$ est une famille orthogonale de $H_p^1(0,2\pi)$ (remarquer que $e'_n(x) = ine_n(x)$). Considérons $f \in H_p^1$, par définition sa dérivée est dans L^2 et on peut calculer ses coefficients de Fourier en utilisant le théorème d'intégration par parties, pour $n \neq 0$:

$$c_n(f') = \frac{1}{2\pi} \int_0^{2\pi} f'(t) e^{-int} dt = inc_n(f),$$

donc comme $c_n(f')$ est dans $\ell^2(\mathbb{Z})$, la suite $(c_n(f))$ vérifie bien la condition du théorème. Réciproquement, si la condition est vérifiée, on calcule la norme H^1 d'un paquet de Cauchy,

$$\|S_q(f) - S_p(f)\|_{H^1}^2 = \sum_{n=p+1}^{q} (1+n^2)|c_n(f)|^2,$$

donc les sommes partielles convergent vers $\tilde{f}$ dans H^1. La convergence dans H^1 implique celle dans L^2, donc $f = \tilde{f}$ presque partout (avec égalité si l' on suppose au préalable que f est continue). Pour finir, on remarque que la somme partielle $S_N(f)$ est

$$S_N(f) = \sum_{|n| \leqslant N} (f|e_n^1)_{H^1} e_n^1,$$

ce qui montre que la famille $(e_n^1)_n$ est totale, donc est une base hilbertienne de $H_p^1(0,2\pi)$.

□

On peut définir des espaces de Sobolev d'ordre plus élevé, $H^2, H^3, \ldots$ de la même manière. Nous en verrons une généralisation l'an prochain dans le cours de Distributions.

Chapitre 4

Calcul différentiel

Commençons par quelques rappels :

- Si (E, d) est un espace métrique et F un espace vectoriel normé, l'ensemble $C_b^0(E, F)$ des fonctions continues bornées de X dans F est un espace vectoriel normé si on le munit de la norme de la convergence uniforme,

$$\|f\|_{C^0} = \|f\|_{C^0(X,F)} = \|f\|_\infty = \sup_{x \in X} \|f(x)\|_F.$$

- Un espace vectoriel normé $(E, \|\cdot\|)$ est appelé espace de Banach s'il est complet pour la distance $d(v, w) = \|v - w\|_E$, c' est-à-dire si toute suite de Cauchy (u_n) de E vérifiant par définition,

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n, m \geqslant N,\ \|u_n - u_m\|_E < \varepsilon,$$

converge vers un $u \in E$.

- Une propriété facile mais très utile des espaces de Banach est la suivante : si (u_n) est une suite de E telle que,

$$\sum_{n=1}^{\infty} \|u_n\|_E < \infty,$$

alors la série $\sum u_n$ converge dans E, c'est-à-dire que la suite $\sum_{n=1}^{N} u_n$ converge dans E quand N tend vers l'infini, vers un élément que l'on convient de noter $\sum_{n=1}^{\infty} u_n$

- Si (E, d) est un espace métrique et F un espace de Banach l'ensemble $C_b^0(X, F)$ des fonctions continues et bornées de X dans F est un espace de Banach pour la norme $\|\cdot\|_\infty$ que nous avons définie auparavant.

- De même si E est un espace vectoriel normé et F un espace de Banach, l'ensemble $\mathcal{L}_c(E, F)$ des applications linéaires continues de E dans F est un espace de Banach si on le munit de la norme d'opérateurs $|||\cdot||| = \|\cdot\|_{\mathcal{L}_c(E,F)}$.

- Une conséquence de ce dernier fait est la suivante : l'ensemble des applications linéaires continues inversibles et dont l'inverse est continu d'un espace de Banach E dans lui-même est un ouvert.

- En fait si u est une application linéaire continue inversible d'un espace de Banach E dans lui-même, son inverse est automatiquement continue (mais c'est un théorème sensiblement plus difficile à démontrer. La démonstration nécessite le théorème de Baire).

Montrons l'avant-dernière propriété.

Lemme 4.0.1. *Supposons que $u \in \mathcal{L}_c(E, E)$ soit inversible (et donc d'inverse continu) et notons h un élément de $\mathcal{L}_c(E, E)$ tel que, $|||h||| < |||u^{-1}|||^{-1}$. Alors, $(u + h)^{-1}$ existe et est dans $\mathcal{L}_c(E)$.*

Démonstration : D' après l'hypothèse faite sur h, $|||u^{-1} \circ h||| \leqslant |||u^{-1}||| \cdot |||h||| \leqslant |||u^{-1}||| \cdot |||u^{-1}|||^{-1} < 1$. Donc,

$$\sum_{k=0}^{\infty} |||u^{-1} \circ h|||^k < \infty.$$

Par conséquent, $\sum_{k=0}^{N}(-1)^k(u^{-1} \circ h)^k$ a une limite quand N tend vers l'infini. Mais un calcul simple montre que, (en notant Id l'application identité),

$$(\sum_{k=0}^{N}(-1)^k(u^{-1} \circ h)^k) \circ (\mathrm{Id} + (u^{-1} \circ h)) = \mathrm{Id} + (-1)^{N+1}(u^{-1} \circ h)^{N+1},$$

et comme $|||(u^{-1} \circ h)^{N+1}||| \leqslant |||u^{-1} \circ h|||^{N+1}$ tend vers 0 on a bien,

$$(\sum_{k=0}^{\infty}(-1)^k(u^{-1} \circ h)^k) \circ (\mathrm{Id} + (u^{-1} \circ h)) = \mathrm{Id}.$$

Par conséquent,

$$[(\sum_{k=0}^{\infty}(-1)^k(u^{-1} \circ h)^k) \circ u^{-1}] \circ (u + h) = \mathrm{Id},$$

et on a donc calculé un inverse à gauche pour $u + h$ c'est-à-dire un inverse pour $u + h$ (remarquer que si $w \circ v$ =Id avec v inversible, alors $v \circ (w \circ v) \circ v^{-1} = Id = v \circ w$).

□

4.1 Différentielle

Dans tout ce qui suit nous noterons E, F deux espaces de Banach. Le plus souvent, il s'agira en fait d'espaces de dimension finie, $\mathbb{R}^n, \mathbb{R}^p$, mais il est important de noter que l'on ne se restreint pas à ces seuls cas.

4.1.1 Définition

Définition 4.1.1. *Soient $f : E \to F$ une application et $x_0 \in E$. Nous dirons que f est différentiable en x_0 s'il existe une application linéaire continue $A_{x_0} : E \to F$ telle que la limite suivante est nulle :*

$$\lim_{||v||_E \to 0} \frac{\|f(x_0 + v) - (f(x_0) + A_{x_0}v)\|_F)}{\|v\|_E} = 0.$$

Si elle existe, une telle application linéaire est unique. On la note $A_{x_0} = D_{x_0}f$.

En d'autres termes, $f(x_0) + D_{x_0}f(v)$ est, uniformément en v au voisinage de $v = 0$, une bonne approximation affine de $f(x_0 + v)$ à l'ordre $o(\|v\|_E)$.

Démonstration : Montrons qu'effectivement A_{x_0} est unique : s'il existe B_{x_0} vérifiant la même propriété, alors : $A_{x_0}v - B_{x_0}v = \|v\|\varepsilon(v)$. Prenons $v = \lambda w$ où w est un vecteur fixé de E : $(A_{x_0} - B_{x_0})w = \|w\|\varepsilon(\lambda w)$, et donc en faisant tendre λ vers zéro, on a bien $A_{x_0} - B_{x_0} = 0$. L'application linéaire $D_{x_0}f$ étant continue (i.e vérifiant $\|D_{x_0}f(v)\|_F \leqslant C.\|v\|_E$) il est clair que f est alors continue en x_0.

□

Remarquons que l'hypothèse de continuité de la différentielle est automatique en dimension finie mais ne l'est pas en dimension infinie et doit donc être vérifiée dans ce dernier cas.

Remarque 4.1.2. *La notation de la différentielle n'est pas universelle : $D_{x_0}f$ peut être notée $Df(x_0)$, $df(x_0)$, ou df_{x_0} voire $f'(x_0)$ (qui est clairement un abus de notation). On notera Df ou df l'application de E dans $\mathcal{L}_c(E,F)$ qui à x_0 associe $D_{x_0}f$.*

Définition 4.1.3. *Lorsque $D_{x_0}f = 0$ (application linéaire nulle), on dit que x_0 est un point critique de la fonction f.*

Il est commode parfois de disposer d'une notion plus faible de différentiabilité qui a le mérite d'avoir un sens quand les espaces E et F ne sont plus nécessairement des Banach ou même normés.

Définition 4.1.4. *Une application $f : E \to F$ entre deux espaces vectoriels normés est dite Gateaux-différentiable en $x_0 \in E$, s'il existe une application linéaire $D_{x_0}f : E \to F$ telle que pour tout vecteur $v \in E$ la limite suivante existe et est nulle :*

$$\lim_{t\to 0}\left\|\frac{f(x_0+tv)-(f(x_0)+D_{x_0}f(tv))}{t}\right\|_F = 0.$$

Dans ce cas on note

$$\lim_{t\to 0}\frac{f(x_0+tv)-f(x_0)}{t} = D_{x_0}f(v).$$

Nous appellerons cette valeur la dérivée de f dans la direction v et on peut également la noter $\frac{\partial f}{\partial v}(x_0)$. Plus généralement, on peut donc définir la dérivée de f dans la direction v comme la limite précédente.

Si f est Gateaux-différentiable en x_0, la dérivée de f existe donc dans toutes les directions. On a alors le résultat immédiat :

Proposition 4.1.5. *Une application $f : E \to F$ différentiable en x_0 est Gateaux-différentiable en ce point.*

Remarque 4.1.6. *La réciproque est bien sûr fausse, comme le montre l'exemple suivant : considérons la fonction de deux variables $f(x,y) = xy^2(x^2+y^4)^{-1}$ pour $(x,y) \neq 0$ et $f(0,0) = 0$. La dérivée de f en $(0,0)$ dans la direction (α,β) est $\beta\alpha^{-1}$ (ou 0 si $\alpha = 0$). Cependant, si $x = y^2$, alors $f(x,y) = \frac{1}{2}$ donc la fonction n'est pas différentiable en $(0,0)$, puisqu'elle n'y est même pas continue !*

4.1.2 Composition

Proposition 4.1.7. *Soient E, F, G trois espaces de Banach, U et V des ouverts respectivement de E et F et $f : U \to F$ et $g : V \to G$ deux applications continues sur U et V respectivement avec $f(U) \subset V$. Supposons que f et g sont différentiables respectivement en $x_0 \in U$ et en $y_0 = f(x_0) \in V$. Alors $g \circ f : E \to G$ est différentiables en x_0 et son application linéaire tangente $D_{x_0}(g\circ f) \in \mathcal{L}_c(E,G)$ en x_0 est la composition des applications linéaires $D_{f(x_0)}g$ et $D_{x_0}f$:*

$$D_{x_0}(g\circ f) = D_{f(x_0)}g \circ D_{x_0}f.$$

Démonstration : Notons $h = g\circ f$. On a alors pour tout v assez petit,

$$\begin{aligned} h(x_0+v)-h(x_0) &= g(f(x_0+v))-g(f(x_0)) \\ &= g(f(x_0)+(f(x_0+v)-f(x_0)))-g(f(x_0)) \\ &= D_{f(x_0)}g(w)+\|w\|\varepsilon(w), \end{aligned}$$

où on a noté $w = f(x_0 + v) - f(x_0)$, et où $\lim_{z\to 0} \varepsilon(z) = 0$. On a alors,

$$w = D_{x_0} f(v) + \|v\|\eta(v),$$

avec $\lim_{v\to 0} \eta(v) = 0$, ce qui entraîne en particulier, si v est suffisamment petit, que, $\|w\| \leqslant (\|D_{x_0} f\| + 1).\|v\|$. On a donc,

$$h(x_0 + v) - h(x_0) = D_{f(x_0)} g.[D_{x_0} f(v) + \|v\|\eta(v)] + \|w\|\varepsilon(w),$$

ce qui montre, vue la majoration sur $\|w\|$ en fonction de $\|v\|$,

$$h(x_0 + v) - h(x_0) = (D_{f(x_0)} g \circ D_{x_0} f)(v) + o(\|v\|).$$

La proposition est donc démontrée.

□

4.1.3 Exemples

4.1.3.1 Les applications linéaires et bilinéaires

Si $f : E \to F$ est une application linéaire continue, sa dérivée $D_a f$ en tout point $a \in E$ est égale à f.

Rappelons que $B : E \times F \to G$ est une application bilinéaire continue si elle est bilinéaire et si elle vérifie pour une constante $C > 0$,

$$\forall (u, v) \in E \times F \ \|B(u, v)\|_G \leqslant C\|u\|_E \|v\|_F.$$

Une telle application est différentiable en tout point $(u_0, v_0) \in E \times F$ et son application linéaire tangente est donnée par,

$$\forall (h, k) \in E \times F,\ D_{(u_0,v_0)} B(h, k) = B(u_0, k) + B(h, v_0).$$

4.1.3.2 Différentielle de l'inverse

On va étudier la différentiabilité et faire le calcul de l'application linéaire tangente de l'application

$$\begin{array}{rccc} \Phi : & GL(E) & \to & GL(E) \\ & u & \mapsto & u^{-1}. \end{array}$$

Soit $u \in GL(E)$ et soit $h \in \mathcal{L}_c(E)$ tel que $|||h||| < |||u^{-1}|||^{-1}$. Alors,

$$u + h = u(\mathrm{Id} + u^{-1}h),$$

et la série,

$$\mathrm{Id} - (u^{-1}h) + (u^{-1}h)^2 + \cdots + (-1)^n (u^{-1}h)^n + \cdots$$

est convergente et converge vers $(\mathrm{Id} + u^{-1}h)^{-1}$. On a donc pour h suffisamment petit,

$$\begin{aligned} (u + h)^{-1} &= (\mathrm{Id} - u^{-1}h)^{-1} u^{-1} \\ &= u^{-1} - u^{-1}hu^{-1} + o(\|u\|). \end{aligned}$$

La dérivée de $\Phi : u \to u^{-1}$ en tout point u inversible est donc l'application,

$$D_u \Phi(h) = -u^{-1}hu^{-1}.$$

En effet, pour tout $h \in \mathcal{L}_c(E)$, $|||-u^{-1}hu^{-1}||| \leqslant |||u^{-1}|||^2 |||h|||^2$ et cette application linéaire est bien continue.

4.1.3.3 Lien avec la notion usuelle de dérivée

Soit $f : (a, b) \to \mathbb{R}$ une fonction définie sur un intervalle $(a, b) \subset \mathbb{R}$ et à valeurs réelles. Supposons que f admette une application linéaire tangente en $t_0 \in (a, b)$:

$$f(t_0 + h) = f(t_0) + D_{t_0}f(h) + o(h),$$

où $D_{t_0}f$ est une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$. Celle-ci est nécessairement de la forme $D_{t_0}f(h) = \lambda h$ où λ est un réel. On a donc,

$$\lim_{h\to 0} \frac{f(t_0 + h) - f(t_0)}{h} = \lambda.$$

Le réel λ s'identifie donc à la dérivée usuelle de f en t_0 que l'on note $f'(t_0)$. La différentielle de f en t_0 est donc, pour tout $h \in \mathbb{R}$, $D_{t_0}f(h) = f'(t_0)h$.

4.1.4 Dérivées partielles

Soient E, F, G des espaces vectoriels normés, W un voisinage d'un point $z_0 = (x_0, y_0) \in E \times F$ et $f : W \to G$ une application différentiable en $z_0 = (x_0, y_0)$ d'application linéaire tangente en z_0, $D_{z_0}f : E \times F \to G$. Nous conviendrons de noter $D_{z_0}f(v, w)$ son action sur $(v, w) \in E \times F$. On peut toujours supposer que W contient un ouvert de la forme $U \times V$ où U est un ouvert de E contenant x_0 et V un ouvert de F contenant y_0. Introduisons alors les applications,

$$g_{z_0} : \begin{array}{ccc} U & \to & G \\ x & \mapsto & f(x, y_0), \end{array}$$

et,

$$h_{z_0} : \begin{array}{ccc} V & \to & G \\ y & \mapsto & f(x_0, y). \end{array}$$

Proposition 4.1.8. *L'application $g_{z_0} : U \to G$, (resp.h_{z_0}) est différentiable en x_0 et son application linéaire tangente $D_{z_0}g(x_0) \in \mathcal{L}_c(E, G)$ (resp.$D_{z_0}h(y_0) :\in \mathcal{L}_c(F, G)$) vérifie pour tout $v \in E$ (resp. tout $w \in F$),*

$$D_{z_0}g(x_0)(v) = D_{(x_0,y_0)}f(v, 0), \text{ (resp. } D_{z_0}h(y_0)(w) = D_{(x_0,y_0)}f(0, w)).$$

L'application $D_{z_0}g(x_0)$ (resp. $D_{z_0}h(y_0)$) s'appelle la dérivée partielle de f en x_0 (resp. y_0) suivant E (resp. F) et se note $D_x f(x_0, y_0)$ (resp. $D_y f(x_0, y_0)$)) ou encore $\partial_x f(x_0, y_0)$ (resp. $\partial_y f(x_0, y_0)$).

Remarque 4.1.9. *On peut montrer, avec les notations précédentes, que si H est un espace vectoriel normé, $\phi : H \to E, \psi : H \to F$ sont continues sur un voisinage W de $a_0 \in H$, différentiables en $a_0 \in H$ et si $x_0 = \phi(a_0), y_0 = \psi(a_0)$ alors l'application $\theta(\cdot) : W \to G$ définie par $\theta(a) = f(\phi(a), \psi(a))$ est différentiable en a_0 et on a,*

$$D_{a_0}\theta = D_x f(x_0, y_0) \circ D_{a_0}\phi + D_y f(x_0, y_0) \circ D_{a_0}\psi$$

4.2 Le cas de la dimension finie

Définition 4.2.1. *Considérons $E = \mathbb{R}^p$, $F = \mathbb{R}$, et $f \in \mathcal{F}(\Omega, F)$ où Ω est un ouvert de E. Soit $x_0 \in E$. Si f est différentiable en x_0, il existe un unique vecteur de E, noté $\nabla f(x_0)$, tel que pour tout $h \in \mathbb{R}^p$, $D_{x_0}f(h) = (\nabla f(x_0)|h)$, où $(\cdot|\cdot)$ est le produit scalaire euclidien sur $\mathbb{R}^p$.*

C'est une simple conséquence du théorème de Riesz (et on pourrait donc généraliser la définition au cas où E est un espace de Hilbert).

Considérons maintenant le cas où $E = \mathbb{R}^p$, $F = \mathbb{R}^n$, et $f = (f_1, \ldots, f_n)$ où les f_i sont les coordonnées de f. Il est clair que f est différentiable si et seulement si tous les f_i le sont. Si p_i est l'application qui à $x \in E$ associe sa i-ème coordonnée x_i, p_i est une application linéaire, donc sa différentielle est elle-même, et l'on convient de la noter dx_i. dx_i est donc l'application qui à $h \in E$ associe $dx_i(h) = h_i$. Alors, à j fixé,

$$D_{x_0} f_j = \sum_{i=1}^{p} \frac{\partial f_j}{\partial x_i}(x_0) dx_i,$$

et par extension, on écrira

$$D_{x_0} f = (D_{x_0} f_1, \ldots, D_{x_0} f_n) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(x_0) dx_i.$$

On peut donc écrire

$$D_{x_0} f(h) = \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant p}} \begin{pmatrix} h_1 \\ \vdots \\ h_p \end{pmatrix}.$$

La matrice de la formule précédente est appelée la matrice jacobienne de l'application différentiable f et son déterminant (quand il existe, c.à.d. lorsque $E = F$) le jacobien. Compte-tenu de la formule de composition des différentielles, on voit que la matrice jacobienne de la composée $g \circ f$ n'est autre que le produit des matrices jacobiennes de g et f.

4.3 Inégalité des accroissements finis

Théorème 4.3.1. *Soit $f : U \to F$ différentiable sur un ouvert U convexe de E. Alors pour tout $a, b \in U$,*

$$\|f(b) - f(a)\|_F \leqslant \left(\sup_{x \in [a,b]} |||D_x f||| \right) . \|b - a\|_E,$$

où $[a, b] = \{ta + (1 - t)b, \, t \in [0, 1]\}$ (comme U est convexe $[a, b] \subset U$).

Démonstration : Notons $M = \sup_{x \in [a,b]} |||D_x f|||$ et supposons $M < \infty$ (sinon la conclusion du théorème est triviale). Soit en outre $\varepsilon > 0$ et posons,

$$A_\varepsilon = \{t \in [0, 1], \, \|f(a + t(b - a)) - f(a)\|_F \leqslant (M + \varepsilon)t\|b - a\|_E\}.$$

L'ensemble A_ε est fermé (puisque f est continue), contient 0 et étant borné, il est compact. Notons t_0 son plus grand élément qui, comme A_ε est compact, appartient à A_ε. Supposons par l'absurde que $t_0 < 1$. On a alors,

$$\|f(a + t_0(b - a)) - f(a)\|_F \leqslant (M + \varepsilon)t_0\|b - a\|_E.$$

Par ailleurs, d'après la définition de t_0 il existe une suite (t_n) qui converge vers t_0 telle que pour tout n, $t_n > t_0$ et qui est donc composée d'éléments n'appartenant pas à A_ε, c'est-à-dire,

$$\forall n, \; \|f(a + t_n(b - a)) - f(a)\|_F > (M + \varepsilon)t_n\|b - a\|_E.$$

En soustrayant ces deux inégalités il vient,

$$\begin{aligned}(M+\varepsilon)(t_n-t_0)\|b-a\| &< \|f(a+t_n(b-a))-f(a)\|-\|f(a+t_0(b-a))-f(a)\| \\ &\leqslant \|f(a+t_n(b-a))-f(a+t_0(b-a))\|,\end{aligned}$$

soit,

$$M+\varepsilon < \frac{\|f(a+t_n(b-a))-f(a+t_0(b-a))\|_F}{(t_n-t_0)}.$$

En faisant tendre n vers l'infini,

$$M+\varepsilon \leqslant \|D_{a+t_0(b-a)}f(b-a)\|_F,$$

ce qui contredit la définition de M. Par conséquent pour tout $\varepsilon > 0$, $A_\varepsilon = [0,1]$. Si on fait à présent tendre ε vers 0, on obtient la conclusion du théorème.

□

1. Le même théorème est vrai si on suppose seulement U connexe (il faut alors prendre le sup sur U tout entier). En effet, dans un espace vectoriel normé la connexité d'un ouvert est équivalente à sa connexité par arcs et on peut choisir l'arc joignant deux points quelconques affine par morceaux.
2. Nous attirons l'attention sur le fait qu'en général (à la différence des applications dérivables d'un intervalle de $\mathbb{R}$ dans $\mathbb{R}$) il n'existe pas toujours $c \in (a,b)$ pour lequel $f(b)-f(a) = (b-a)f'(c)$.
3. On peut donner des versions plus générales de ce théorème. Signalons par exemple celle-ci : supposons qu'il existe $\phi : [0,1] \to \mathbb{R}$ dérivable et telle que pour tout $t \in [0,1]$ on ait,
$$\|Df(a+t(b-a)).(b-a)\|_F \leqslant \phi'(t),$$
alors,
$$\|f(b)-f(a)\|_F \leqslant \phi(b)-\phi(a).$$
La démonstration de ce fait s'effectue de la même manière que précédemment.

Nous donnons à présent une condition suffisante de différentiabilité qui est très utile, surtout en dimension finie. Nous allons en fait donner une condition sur la notion plus forte de fonction de classe C^1.

Définition 4.3.2. *Nous dirons qu'une application f d'un ouvert U de E dans F est C^1 en x_0 si elle est différentiable sur un voisinage V de x_0 et si l'application*

$$D.f \ : \ \begin{array}{ccc} V & \to & \mathcal{L}_c(E,F) \\ x & \mapsto & D_x f \end{array}$$

est continue en x_0.

On a alors la caractérisation suivante des fonctions C^1.

Théorème 4.3.3. *Soient U un ouvert d'un espace produit $E_1 \times \cdots \times E_n$ (en particulier d' un espace de dimension finie) et soit $f : U \to F$. On a alors équivalence entre les deux assertions suivantes :*

1. *f est C^1 en $a \in U$.*
2. *pour tout $i \in [\![1,n]\!]$, les dérivées partielles $D_{x_i}f = \frac{\partial f}{\partial x_i}$ existent et sont continues en a.*

Démonstration : Le fait que 1. implique 2. est clair. Montrons l'implication inverse dans le cas où $n = 2$ (le cas général s'effectuant de la même manière). On a,

$$\begin{aligned}\|f(a_1+h_1,a_2+h_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1 - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2\| \\ \leqslant \|f(a_1+h_1,a_2+h_2) - f(a_1+h_1,a_2) - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2\| \\ + \|f(a_1+h_1,a_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1\|.\end{aligned}$$

Par définition de la différentiabilité de f en (a_1, a_2) on a déjà,

$$\|f(a_1+h_1,a_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1\| = o(\|h_1\|).$$

Par ailleurs l'application g définie dans un voisinage de 0 par,

$$g(h_2) = f(a_1+h_1,a_2+h_2) - f(a_1+h_1,a_2) - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2,$$

est dérivable sur un voisinage de 0 et vérifie,

$$D(\tilde{h})g = \frac{\partial f}{\partial x_2}(a_1+h_1,a_2+\tilde{h}) - \frac{\partial f}{\partial x_2}(a_1,a_2) = \varepsilon(\|h_1\| + \|\tilde{h}\|),$$

où, puisque $\frac{\partial f}{\partial x_2}$ est continue, $\varepsilon(s)$ tend vers 0 avec s (noter l'abus de notation: ε dépend de $h_1, \tilde{h}$ et pas seulement de leurs normes). On a donc d'après le théorème des accroissements finis,

$$\begin{aligned}\|f(a_1+h_1,a_2+h_2) - f(a_1+h_1,a_2) - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2\| &= \|g(h_2) - g(0)\| \\ &\leqslant \|h_2\|.\varepsilon(\|h_1\| + \|h_2\|).\end{aligned}$$

Au total,

$$\begin{aligned}&\|f(a_1+h_1,a_2+h_2) - f(a_1,a_2) - \frac{\partial f}{\partial x_1}(a_1,a_2).h_1 - \frac{\partial f}{\partial x_2}(a_1,a_2).h_2\| \\ &\leqslant \|h_2\|\varepsilon(\|h_1\| + \|h_2\|) + o(\|h_1\|) = o(\|h_1\| + \|h_2\|),\end{aligned}$$

ce qui démontre le théorème.

□

4.4 Différentielles d'ordres supérieurs

4.4.1 Définition

Supposons que $f : E \to F$ soit différentiable sur un ouvert U de E. En tout point $x \in U$ on peut donc considérer l'application linéaire tangente $D_x f \in \mathcal{L}_c(E, F)$ et ce dernier espace (des opérateurs continus de E dans F) est également un espace de Banach quand on le munit de la norme d'opérateurs.

Définition 4.4.1. *Nous dirons que f est deux fois différentiable en x_0 si f est différentiable sur un voisinage U de x_0 et si l'application $Df : U \to \mathcal{L}_c(E,F)$ qui à x associe $D_x f$ est différentiable en x_0. On notera $D^2_{x_0} f$ l'application linéaire tangente correspondante qui est un élément de $\mathcal{L}_c(E, \mathcal{L}_c(E,F))$ et peut être vu comme un élément de $\mathcal{L}_c(E \times E, F)$.*

De façon générale, f est dite k fois dérivable en x_0 si elle est $k-1$ fois dérivable sur un voisinage U de x_0 et si sa dérivée $(k-1)$-ième $D^{(k-1)} f : U \to \mathcal{L}_c(E \times \cdots \times E, F)$ (application $(k-1)$-linéaire continue) est dérivable en x_0. Alors, $D^k_x f$ est, pour tout $x \in U$, une application k-linéaire de $E \times \cdots \times E$ dans F.

Enfin, si $D^k f$ est continue sur U, on dit que f est de classe C^k sur U.

Nous noterons $D^k_{x_0} f(h_1, \ldots, h_k)$ l'action de $D^k_{x_0} f$ sur un vecteur $(h_1, \ldots, h_k) \in E^k$.

Considérons le cas particulier où $k = 2$. Dire que f est deux fois différentiable en x_0, c'est dire que f est différentiable dans un voisinage de $U \subset E$ de x_0 et que l'application $D : x \mapsto D_x f$, de $U \to \mathcal{L}_c(E,F)$ est différentiable en $x = x_0$. On note alors $D^2_{x_0} f = D_{x_0}(Df)$ qui est donc une application linéaire de E dans $\mathcal{L}_c(E,F)$, et on peut l'identifier à un élément de $\mathcal{L}_c(E \times E, F)$. Fixons $h \in E$. L'application $x \to D_x f(h)$ de $E \to F$ est différentiable, de différentielle $D^2_x f(h)$, qui est donc une application linéaire de E dans F, i.e.

$$D^2_x f(h) : \ k \in E \mapsto D^2_x f(h)(k) = D^2_x f(h,k) \in F.$$

On peut donc voir $D^2_x f$ comme une application bilinéaire, et on va voir qu'elle est en fait symétrique.

4.4.2 Théorème de Schwarz

Théorème 4.4.2. *Si f est k fois différentiable en $x_0 \in U$, alors $D^k_{x_0} f$ est une application k- linéaire symétrique. Cela siginife que pour toute permutation σ des indices $\{1, \ldots, k\}$ et tout élément $(h_1, \ldots, h_k) \in E^k$,*

$$D^k_{x_0} f(h_{\sigma(1)}, \ldots, h_{\sigma(k)}) = D^k_{x_0} f(h_1, \ldots, h_k).$$

Démonstration : Nous prouvons le théorème dans le cas $k = 2$, la démonstration dans le cas général étant la même.

On montre tout d'abord que pour h_1, h_2 suffisamment petits,

$$\|(f(x_0 + h_1 + h_2) - f(x_0 + h_1) - f(x_0 + h_2) + f(x_0)) - D^2 f(x_0).h_2.h_1\| \tag{4.1}$$
$$= (\|h_1\|^2 + \|h_2\|^2)\varepsilon(\|h_1\| + \|h_2\|) \tag{4.2}$$

où $\varepsilon(\cdot)$ tend vers 0 en 0. Notons,

$$S(h_1, h_2) = f(x_0 + h_1 + h_2) - f(x_0 + h_1) - f(x_0 + h_2) + f(x_0),$$

et pour $t \in [0,1]$,

$$T(t) = S(th_1, h_2) - D^2_{x_0} f(h_2, th_1).$$

On a alors,

$$\partial_t T(t) = D_{x_0 + th_1 + h_2} f(h_1) - D_{x_0 + th_1} f(h_1) - D^2_{x_0} f(h_2, h_1),$$

et puisque Df est dérivable en x_0,

$$\begin{aligned} D_{x_0+th_1+h_2} f(h_1) - D_{x_0+th_1}(h_1) &= D_{x_0} f(h_1) + D^2_{x_0} f(th_1 + h_2, h_1) + o(\|th_1 + h_2\|\|h_1\|) \\ &\quad - (D_{x_0} f(h_1) + D^2_{x_0} f(th_1, h_1) + o(\|h_2\|\|h_1\|) \\ &= D^2_{x_0} f(h_2, h_1) + o((t\|h_1\| + \|h_2\|)\|h_1\|) + o(\|h_2\|\|h_1\|), \end{aligned}$$

et finalement, comme $t \in [0,1]$, $\partial_t T(t) = o(\|h_1\|^2 + 2\|h_2\| \cdot \|h_1\|)$.
L'inégalité des accroissements finis montre que,

$$S(h_1, h_2) = \|T(1) - T(0)\| \leqslant (\|h_1\|^2 + \|h_2\|^2)\varepsilon(\|h_1\| + \|h_2\|),$$

où $\varepsilon(\cdot)$ tend vers 0 en 0, et l'inégalité (4.2) est démontrée.
Puisque $S(h_1, h_2)$ est une expression symétrique en h_1, h_2 on a,

$$\|D^2_{x_0} f(h_1, h_2) - D^2_{x_0} f(h_2, h_1)\| \leqslant 2(\|h_1\|^2 + \|h_2\|^2)\varepsilon(\|h_1\| + \|h_2\|),$$

et si on remplace h_1, h_2 par th_1, th_2 et que l'on divise l'expression alors obtenue par t^2, on obtient la conclusion du théorème en laissant t tendre vers zéro.

□

4.4.3 Formules de Taylor

Nous donnons les trois versions classiques, formule de Taylor-Young, inégalité de Taylor-Lagrange et formule de Taylor avec reste intégral.

Théorème 4.4.3. *Supposons que $f : U \to F$ est une application n fois dérivable en $x_0 \in U$. Alors,*

$$\|f(x_0 + h) - f(x_0) - D_{x_0} f(h) - \cdots - \frac{1}{n!} D^n_{x_0} f(h, \ldots, h)\| = o(\|h\|^n).$$

Démonstration : Elle se fait par récurrence sur n. La formule est par définition de la différentiablité en x_0 vraie pour $n = 1$.
Supposons-la vraie pour $n - 1$ et montrons qu'elle est vraie pour n. Posons,

$$\begin{aligned} g(t) &= f(x_0 + th) - f(x_0) - D_{x_0} f(th) - \cdots - \frac{1}{n!} D^n_{x_0} f(th, \ldots, th) \\ &= f(x_0 + th) - f(x_0) - tD_{x_0} f(h) - \cdots - \frac{1}{n!} t^n D^n_{x_0} f(h, \ldots, h) \end{aligned}$$

La fonction g est alors dérivable et on a,

$$\begin{aligned} g'(t) &= D_{x_0+th} f(h) - D_{x_0} f(h) - \cdots - \frac{1}{(n-1)!} t^{n-1} D^n_{x_0} f(h, \ldots, h) \\ &= [D_{x_0+th} f - D_{x_0} f - \cdots - \frac{1}{(n-1)!} D^{n-1}_{x_0} (Df)(th, \ldots, th)](h), \end{aligned}$$

et l'hypothèse de récurrence s'applique à Df qui est $(n-1)$ fois différentiable en x_0 :

$$\|D_{x_0+th} f - D_{x_0} f - \cdots - \frac{1}{(n-1)!} D^{n-1}_{x_0} (Df)(th, \ldots, th)\| = o(\|th\|^{n-1}).$$

Par conséquent, $\|g'(t)\| = o(\|h\|^{n-1})\|h\| = o(\|h\|^n)$. L'inégalité des accroissements finis appliquée sur $[0,1]$ à g montre que,

$$\|g(1) - g(0)\| \leqslant \sup_{t \in [0,1]} \|g'(t)\|,$$

ce qui est la conclusion recherchée (car $g(0) = 0$).

□

Théorème 4.4.4. *Supposons que U soit un ouvert connexe et que $f : U \to F$ soit une application $n+1$ fois dérivable sur U et qu'il existe $M \geqslant 0$ tel que,*

$$\sup_{x \in U} \|D^n_y f\| \leqslant M.$$

Alors pour tout $a, b \in U$,

$$\|f(b) - f(a) - D_a f(b-a) - \cdots - \frac{1}{n!} D_a f^{(n)}(b-a, \ldots, b-a)\| \leqslant M \frac{\|b-a\|^{n+1}}{(n+1)!}.$$

Remarque 4.4.5. *La différence avec le théorème précédent est que l'on a fait une hypothèse plus forte sur f, et qu'en retour on a obtenu en particulier le résultat précédent avec un $\varepsilon(b-a)$ qui non seulement tend vers zéro mais est lipschitzien.*

Démonstration : Faisons la démonstration dans le cas où U est convexe en procédant comme précédemment par récurrence sur n et en introduisant,

$$g(t) = f(x_0 + th) - f(x_0) - D_{x_0} f(th) - \cdots - \frac{1}{n!} D^n_{x_0} f(th, \ldots, th),$$

où $h = b - a$. On voit, en appliquant l' hypothèse de récurrence à Df, que pour tout $t \in [0,1]$,

$$\|g'(t)\| \leqslant M\|b-a\|.\|b-a\|^n \frac{t^n}{n!};$$

Intégrons cette dernière expression par rapport à $t \in [0,1]$:

$$\|g(1) - g(0)\| \leqslant M\|b-a\|^{n+1} \int_0^1 \frac{t^n}{n!} dt \leqslant M\|b-a\|^{n+1}. \frac{1}{(n+1)!},$$

ce qui donne le résultat voulu.

□

Théorème 4.4.6 (Taylor avec reste intégral)**.** *Supposons que U soit un ouvert convexe et que $f : U \to F$ soit de classe C^{n+1}. Alors,*

$$f(b) - f(a) - D_a f(b-a) - \cdots - \frac{1}{n!} D_a f^{(n)}(b-a, \ldots, b-a) = \int_0^1 \frac{(1-t)^n}{n!} D^{(n+1)}_{a+t(b-a)} f(b-a, \ldots, b-a) dt.$$

Démonstration : La démonstration de ce théorème s'effectue de la manière habituelle en intégrant par partie.

□

Remarque 4.4.7. *Les hypothèses sont ici encore plus fortes, et de nouveau ce théorème contient comme cas particulier les deux précédents. On fait cette hypothèse de régularité supplémentaire pour pouvoir faire l'intégration par parties sans se soucier d'avoir à la justifier (les fonctions intervenant sont au pire continues). En pratique, cette formule de Taylor est celle pour laquelle il est le plus simple d'estimer finement le reste lorsque cela est nécessaire.*

Remarque 4.4.8. *Remarquons que nous n'avons cependant pas défini l'intégrale des fonctions continues d'un intervalle de $\mathbb{R}$ dans un espace de Banach F et que nous n'avons pas montré la formule d'intégration par partie :*

$$\int_0^1 \lambda'(t) v(t) dt + \int_0^1 \lambda(t) v'(t) dt = [\lambda(t) v(t)]_0^1,$$

où $\lambda : [0,1] \to \mathbb{R}$ est à valeurs scalaires et $v : [0,1] \to F$ à valeurs dans le Banach F. Quand F est en fait de dimension finie, il suffit de définir l'intégrale via les applications coordonnées. Lorsque F est un Banach quelconque, on décide que $f \in \mathcal{F}(\mathbb{R}, F)$ est intégrable si $\int \|f(x)\|_F dx < \infty$, et on peut ensuite donner un sens à $\int f(x)dx$. Si l'on se limite à l'intégrale de Riemann, il suffit de passer à la limite dans les sommes de Riemann et comme ici on intègre des fonctions continues, l'intégrale de Riemann est suffisante.

4.4.4 Application aux problèmes d'extrema

Définition 4.4.9. *Soit f une fonction de Ω dans $\mathbb{R}$, où Ω est un ouvert d'un espace de Banach E. On dit que f admet un maximum local en $x_0 \in \Omega$ lorsqu'il existe un voisinage V de x_0 tel que pour tout $x \in V \cap \Omega$, $f(x) \leqslant f(x_0)$. On définit de même un minimum local avec l'inégalité dans l'autre sens, et les notions de maximum et minimum stricts ou globaux. Un extremum est un point qui est minimum ou maximum.*

Théorème 4.4.10. *Soit f une fonction Ω dans $\mathbb{R}$, où Ω est un ouvert d'un esapce de Banach E. On suppose f différentiable en $x_0 \in \Omega$. Une condition nécessaire pour que f admette un extremum local en x_0 est que x_0 soit un point critique de f,* i.e., *$D_{x_0} f = 0$.*

Démonstration : On considère $\phi_h(t) = f(x_0 + th)$, où $[x_0, x_0 + h] \subset \Omega$. La fonction ϕ_h est dérivable, et $\phi'(0) = D_{x_0} f(h)$, ceci quelque soit h. Or, si f a un extremum en x_0, alors $\phi_h(t)$ aussi, donc $\phi'(0) = 0$.

□

Remarque 4.4.11. *Le théorème est faux si Ω n'est pas un ouvert et que x_0 n'est pas à l'intérieur de Ω. Par ailleurs, cette condition n'est bien sûr pas suffisante (le cas du point selle est un contre-exemple).*

On va donner une condition suffisante pour avoir un extremum local qui fait intervenir la différentielle seconde. L'ordre 1 de différentiabilité n'est pas suffisant pour caractériser un extremum.

Théorème 4.4.12. *Considérons f une fonction de classe C^2 de Ω dans $\mathbb{R}$, x_0 un point critique, et notons $\phi(h) = D^2_{x_0} f(h,h)$ la forme quadratique associée à la différentielle seconde. Alors,*

1. *si ϕ est définie positive, f admet un minimum en x_0 ;*
2. *si ϕ est définie négative, f admet un maximum en x_0 ;*
3. *si ϕ est non-définie et non-dégénérée, f n'admet pas d'extremum en x_0 (c'est le cas du point-selle) ;*
4. *si ϕ est dégénérée, on ne peut pas conclure (il faudrait pousser le développement de Taylor plus loin).*

La forme quadratique ϕ est appelée Hessienne de f en x_0.

Démonstration : On utilise la formule de Taylor- Young :

$$f(x_0 + h) - f(x_0) = \frac{1}{2} D^2_{x_0} f(h,h) + \|h\|^2 \varepsilon(h) = \frac{1}{2}\phi(h) + \|h\|^2 \varepsilon(h) = \frac{1}{2}\|h\|^2 \left(\phi(\tfrac{h}{\|h\|}) + \varepsilon(h)\right).$$

Si ϕ est définie positive, alors $y \mapsto \phi(y)$ est strictement positive sur la sphère $S(0,1) = \{\|y\| = 1\}$, et elle y est continue, donc il existe $A > 0$ tel que $\phi(y) \geqslant A$ pour $y \in S(0,1)$. On choisit alors η tel que si $\|h\| < \eta$ alors $|\varepsilon(h)| < \frac{A}{2}$. On en déduit que

$$f(x_0 + h) - f(x_0) > \frac{\|h\|^2}{2} \frac{A}{2},$$

donc que l'on a un minimum (strict). Le cas ϕ définie négative se traite de la même façon.

Lorsque ϕ est non définie et non dégénérée, il existe $y, z \in S(0,1)$ tels que $\phi(y) > 0$ et $\phi(z) < 0$. On aura alors que $f(x_0 + \lambda y) - f(x_0) > 0$ pour λ assez petit, et d'autre part $f(x_0 + \lambda z) - f(x_0) < 0$ pour λ assez petit, donc il ne peut y avoir d'extremum.

Finalement, si ϕ est dégénérée, l'exemple suivant montre que l'on ne peut pas conclure. Soit $f(x,y) = x^2 - y^3$, fonction de $\mathbb{R}^2 \to \mathbb{R}$. Au voisinage de $(0,0)$, $f(x,y) = x^2 + o(\|(x,y)\|^2)$, donc $\phi(x,y) = x^2$ est dégénérée positive, mais $f(0,y)$ est négative si $y > 0$ alors que $f(x,0) \geqslant 0$. Par contre, la fonction $g(x,y) = x^2 + y^4$ a le même ϕ en $(0,0)$, mais elle admet un minimum.

□

Considérons le cas particulier où l'espace de départ est $E = \mathbb{R}^2$. En utilisant les dérivées partielles, la forme quadratique s'écrit

$$\phi(h_1, h_2) = h_1^2 \frac{\partial^2 f}{\partial x_1^2}(x_0) + h_2^2 \frac{\partial^2 f}{\partial x_2^2}(x_0) + 2h_1 h_2 \frac{\partial^2 f}{\partial x_1 x_2}(x_0).$$

On voit donc que la forme quadratique est associée à l'endomorphisme sur $\mathbb{R}^2$ de matrice

$$A = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_1 x_2}(x_0) \\ \frac{\partial^2 f}{\partial x_1 x_2}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) \end{pmatrix},$$

qui est une matrice définie (positive ou négative) si et seulement si son déterminant est strictement positif, soit

$$\frac{\partial^2 f}{\partial x_1^2}(x_0) \frac{\partial^2 f}{\partial x_2^2}(x_0) - \left(\frac{\partial^2 f}{\partial x_1 x_2}(x_0) \right)^2 > 0.$$

Pour étudier un extremum, on calcul donc A en un point critique x_0 puis on étudie les signes de ses valeurs propres.

4.4.5 Différentielle d' un produit

Il est parfois pratique de savoir calculer les différentielles n-ièmes d'un produit de fonctions de classe C^n.

Théorème 4.4.13 (Formule de Leibniz). *Soient $B(\cdot,\cdot)$ une forme bilinéaire continue sur $E \times E$ et soient $f, g : I \to E$ deux fonctions de classe C^n sur I (I est un intervalle de $\mathbb{R}$). Alors, la fonction $h : I \to \mathbb{R}$ définie par $h(x) = B(f(x), g(x))$ est également de classe C^n sur I et sa différentielle k-ième (pour $0 \leqslant k \leqslant n$) est donnée par*

$$D_x^k h = \sum_{l=0}^{k} \binom{k}{l} B(D_x^l f, D_x^{k-l} g).$$

On remarquera la similitude de cette expression avec la formule du binôme de Newton. La démonstration de ce résultat peut se faire par récurrence sur n.

4.5 Inversion locale et fonctions implicites

4.5.1 Difféomorphismes

Rappelons qu'un homéomorphisme $f : U \to V$ entre deux ouverts $U \subset E$ et $V \subset F$ est une application continue de U dans F qui établit une bijection entre U et V et telle que son inverse $f^{-1} : V \to U$ est continue.

Définition 4.5.1. *Nous dirons qu'un homéomorphisme $f : U \to V$ entre deux ouverts $U \subset E$ et $V \subset F$ est un C^k-difféomorphisme si $f : U \to V$ et $f^{-1} : V \to U$ sont de classe C^k.*

Donnons le critère suivant pour déterminer si un homéomorphisme est un difféomorphisme.

Proposition 4.5.2. *Avec les notations précédentes, un homéomorphisme $f : U \to V$ est un C^k-difféomorphisme si et seulement si f est de classe C^k et si, pour tout $x \in U$, $D_x f \in \mathcal{L}_c(E, F)$ est une application linéaire continue inversible* (i.e *dont l'inverse existe et est continu). On a alors,*

$$\forall y \in V,\, D_y(f^{-1}) = \left(D_{f^{-1}(y)} f\right)^{-1}.$$

Démonstration : Posons, pour $y = f(x) \in V$ et $h \in F$ suffisamment petit, $x_h = f^{-1}(y + h)$. On a alors,

$$h = f(x_h) - f(x) = D_x f(x_h - x) + \|x_h - x\|_E \varepsilon(x_h - x),$$

où $\lim_{z\to 0} \varepsilon(z) = 0$ et,

$$x_h - x = (D_x f)^{-1}(h) - \|x_h - x\|_E (D_x f)^{-1}(\varepsilon(x_h - x)).$$

Comme x_h tend vers x quand h tend vers 0 (puisque f^{-1} est continue), on a en particulier, pour h suffsamment petit

$$\|\varepsilon(x_h - x)\| \leqslant \frac{1}{2} |||(D_x f)^{-1}|||^{-1},$$

et,

$$\begin{aligned}\|x_h - x\| &\leqslant |||(D_x f)^{-1}||| \cdot \|h\| + \|x_h - x\| \cdot |||(D_x f)^{-1}||| \|\varepsilon(x_h - x)\| \\ &\leqslant |||(D_x f)^{-1}||| \cdot \|h\| + \frac{1}{2}\|x_h - x\|.\end{aligned}$$

Finalement,

$$\|x_h - x\| \leqslant 2|||(D_x f)^{-1}||| \cdot \|h\|,$$

et en revenant à la formule donnant $x_h - x$, on a

$$f^{-1}(y + h) - f^{-1}(y) = x_h - x = (D_x f)^{-1}(h) + o(h),$$

ce qui montre que f^{-1} est différentiable en y et que,

$$D_y(f^{-1}) = (D_x f)^{-1} = \left(D_{f^{-1}(y)} f\right)^{-1}.$$

Cette dernière égalité établit (en utilisant le théorème de composition et celui sur l'inversion) que f^{-1} est de classe C^k.

□

4.5.2 Inversion locale

Nous allons d'abord revenir sur le théorème de point fixe de Picard et en montrer une version améliorée, à paramètre.

Théorème 4.5.3. *Soient X un espace métrique complet et Λ un espace métrique. Soit $f : X \times \Lambda \to X$. On suppose que pour tout $x \in X$, l'application $\lambda \mapsto f(x, \lambda)$ est continue en λ_0 et que, pour tout $\lambda \in \Lambda$, l'application $x \mapsto f(x, \lambda)$ est lipschitzienne de rapport $k < 1$ indépendant de λ. Alors l'application qui à λ associe le point fixe $x(\lambda)$ solution de $x = f(x, \lambda)$ est continue en $\lambda = \lambda_0$.*

Démonstration : Considérons $x = x(\lambda_0)$ le point fixe de $x = f(x, \lambda_0)$. Par continuité de $\lambda \mapsto f(x(\lambda_0), \lambda)$,

$$\forall \varepsilon > 0,\ \exists V \in \mathcal{V}(\lambda_0),\ \lambda \in V \Longrightarrow d(f(x(\lambda_0), \lambda), f(x(\lambda_0), \lambda_0)) < \varepsilon.$$

On écrit alors, pour $\lambda \in V$

$$\begin{aligned} d(x(\lambda), x(\lambda_0)) \leqslant & d(f(x(\lambda), \lambda), f(x(\lambda_0), \lambda_0)) \\ \leqslant & d(f(x(\lambda), \lambda), f(x(\lambda_0), \lambda)) + d(f(x(\lambda_0), \lambda), f(x(\lambda_0), \lambda_0)) \\ \leqslant & k d(x(\lambda), x(\lambda_0)) + \varepsilon. \end{aligned}$$

Finalement $d(x(\lambda), x(\lambda_0)) \leqslant \frac{\varepsilon}{1-k}$, et $x(\lambda)$ est continue en λ_0.

□

Cette version du théorème du point fixe de Picard permet de démontrer le théorème fondamental suivant.

Théorème 4.5.4 (Inversion locale)**.** *Soient E et F deux espaces de Banach et soit $f : E \to F$ une application de classe C^k, ($k \geqslant 1$) définie sur un voisinage de $x_0 \in E$ et telle que $f(x_0) = y_0 \in F$. Supposons que $D_{x_0} f \in \mathcal{L}_c(E, F)$ soit inversible. Alors f est un difféomorphisme local d'un voisinage x_0 sur un voisinage de y_0.*

Démonstration : Puisque $(D_{x_0} f)^{-1}$ existe et est continue, f réalisera un difféomorphisme d'un voisinage de x_0 sur un voisinage de y_0 si et seulement si,

$$f_0(\cdot) = (D_{x_0} f)^{-1} (f(x_0 + \cdot) - y_0),$$

réalise un difféomorphisme d'un voisinage de $0 \in E$ sur un voisinage de $0 \in E$ (remarquer l'astuce qui permet de se ramener au même espace de départ et d'arrivée). Notons que f_0 est de classe C^1 et que l'on a,

$$f(0) = 0 \quad \text{et} \quad D_0 f_0 = \mathrm{Id}.$$

Posons alors pour u, v dans un voisinage de $0 \in E$,

$$\tilde{f}_v(u) = v + (u - f_0(u)),$$

et observons que, $f_0(u) = v$ si et seulement si $\tilde{f}_v(u) = u$, c'est-à-dire si et seulement si f_v admet u pour point fixe. Vérifions donc que f_v est contractante dans un voisinage de 0 pour v suffsamment petit. Soient $\delta > 0$ suffisamment petit et u_1, u_2 dans la boule fermée $B_f(0, \delta)$ de centre 0 et de rayon δ. Alors,

$$\|\tilde{f}(u_1) - \tilde{f}(u_2)\| = \|(\mathrm{Id} - f_0)(u_1) - (\mathrm{Id} - f_0)(u_2)\|,$$

et d'après le théorème des accroissements finis,

$$\|(\mathrm{Id}-f_0)(u_1)-(\mathrm{Id}-f_0)(u_2)\| \leqslant \sup_{w\in B(0,\delta)} |||D_w(\mathrm{Id}-f_0)||| \cdot \|u_1-u_2\|.$$

Mais, comme Df_0 est continue sur un voisinage de 0 et que $D_0f_0 = \mathrm{Id}$, on a, pourvu que δ soit assez petit,

$$\sup_{w\in B_f(0,\delta)} |||D_w(\mathrm{Id}-f_0)||| \leqslant \frac{1}{2},$$

et l'application $\tilde{f}_v$ est $\frac{1}{2}$-contractante sur $B_f(0,\delta)$. On a en particulier, si on prend $u_2 = 0$,

$$\|\tilde{f}(u_1)-v\| \leqslant \frac{1}{2}\|u_1\|, \text{ et } \quad \|\tilde{f}(u_1)\| \leqslant \|v\| + \frac{1}{2}\delta.$$

Ceci prouve que si $\|v\| \leqslant \frac{\delta}{2}$, $\tilde{f}_v$ envoie $B_f(0,\delta)$ dans elle-même. Les conditions d'application du théorème du point fixe sont vérifiées et $\tilde{f}_v$ admet donc un unique point fixe u_v dans $B_f(0,\delta)$.

En outre comme $\tilde{f}_v(\cdot)$ est continue en v, les hypothèses du théorème du point fixe à paramètre sont vérifiées et on en déduit que l'unique point fixe u_v obtenu précédemment dépend continûment de v.

Tout ceci montre que f réalise un homéomorphisme d'un voisinage de x_0 sur un voisinage de y_0 et d'après la caractérisation des C^k-difféomorphismes vue précédemment, comme D_xf est inversible dans ce voisinage de x_0, f est un C^k-difféomorphisme d'un voisinage de x_0 sur un voisinage de y_0.

□

4.5.3 Fonctions implicites

Théorème 4.5.5 (Fonctions implicites)**.** *Soient $f : E\times F \to E$ de classe C^k et $(x_0,\lambda_0) \in E\times F$. Si on a $f(x_0,\lambda_0) = 0$ et si $D_xf(x_0,\lambda_0) \in \mathcal{L}_c(E)$ est inversible, alors l' ensemble des solutions de $f(x,\lambda) = 0$ est, dans un voisinage de (x_0,λ_0), de la forme $(x(\lambda),\lambda)$ où $\lambda \to x(\lambda)$ est C^k. On a alors $\partial_\lambda x(\lambda) = -(D_xf(x,\lambda))^{-1}\circ D_\lambda f$.*

Démonstration : On applique le théorème d'inversion locale à l'application définie, sur un voisinage de $(x_0,\lambda_0) \in E\times F$ et à valeurs dans un voisinage de $(0,\lambda_0) \in E\times F$, par $\phi(x,\lambda) = (f(x,\lambda),\lambda)$. Cette application est de classe C^k.

Calculons sa différentielle en (x_0,λ_0), $D_{(x_0,\lambda_0)}\phi \in \mathcal{L}_c(E\times F)$. On a pour tout $(\Delta x,\Delta\lambda) \in E\times F$ (nous utilisons une notation matricielle) :

$$D_{(x_0,\lambda_0)}\phi(\Delta x,\Delta\lambda) = \begin{pmatrix} D_xf(x_0,\lambda_0) & D_\lambda f(x_0,\lambda_0) \\ 0 & \mathrm{Id}_F \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix},$$

qui est triangulaire supérieure et qui est inversible puisque par hypothèse, on sait que $D_xf(x_0,\lambda_0) \in \mathcal{L}_c(E)$ l'est. On peut donc appliquer le théorème d'inversion locale : ϕ réalise un difféomorphisme de classe C^k d'un voisinage de (x_0,λ_0) dans un voisinage de $(0,\lambda_0)$. Au vu la forme de ϕ, le difféomorphisme inverse ϕ^{-1} est de la forme,

$$\phi^{-1}(y,\lambda) = (g(y,\lambda),\lambda),$$

où g est de classe C^k d'un voisinage de 0 à valeurs dans un voisinage de x_0. On a donc,

$$(x,\lambda) = (g(f(x,\lambda),\lambda),\lambda),$$

pour tout (x, λ) dans un voisinage de (x_0, λ_0) et par conséquent pour (x, λ) dans ce voisinage, $f(x, \lambda) = 0$ si et seulement si, $x = g(0, \lambda)$. Ceci termine la preuve du théorème des fonctions implicites.

$\square$

Nous terminons en donnant un énoncé plus simple, en deux variables, du théorème des fonctions implicites.

Théorème 4.5.6 (Théorème des fonctions implicites à deux variables)**.** *Soit f une fonction de classe C^k définie sur un ouvert Ω de $\mathbb{R}^2$, à valeurs dans R. Soit $(a, b) \in \Omega$ tel que $f(a, b) = 0$.*

On suppose que la dérivée partielle de f par rapport à la seconde variable est non nulle en (a, b) : $\frac{\partial f}{\partial y}(a, b) \neq 0$.

Alors il existe un voisinage ouvert U de (a, b) dans $\mathbb{R}^2$, un intervalle ouvert I de $\mathbb{R}$ contenant a et une fonction $g : I \to \mathbb{R}$ de classe C^k telle que, pour tout $(x, y) \in U$ on ait :

$$f(x, y) = 0 \iff y = g(x).$$

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

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

Annexe A

Quelques notations

Soit A un ensemble et $B \subset A$. La *fonction indicatrice*, ou *fonction caractéristique* de B, est la fonction $\mathbb{1}_B : A \to \{0,1\}$ définie par

$$\begin{cases} \mathbb{1}_B(x) = 1 & \text{si } x \in B \\ \mathbb{1}_B(x) = 0 & \text{si } x \notin B. \end{cases}$$

On note $A \backslash B$ le complémentaire de B dans A :

$$A \backslash B = \{x \in A \mid x \notin B\}.$$

Lorsqu'il n'y a pas d'ambiguïté sur A, on notera parfois cet ensemble B^c.
Soit E et F deux sous-ensembles de $\mathbb{R}^d$. On note $d(E,F)$ leur distance :

$$d(E,F) = \inf\{|x-y| \ : \ x \in E, \ y \in E\},$$

et $E+F$ leur somme

$$E+F = \Big\{x+y \ : \ x \in E, \ y \in F\Big\}.$$

Si a et b sont deux éléments de $\mathbb{R}^d$, $[a,b]$ est le segment :

$$[a,b] = \{ta + (1-t)b, \ t \in [0,1]\}.$$

Quand $d = 1$ et $a < b$, on retrouve la notation usuelle $[a,b] = \{x \in \mathbb{R}, \ a \leqslant x \leqslant b\}$. Mais on pourra aussi employer la même notation pour $a > b$ auquel cas $[a,b] = \{x \in \mathbb{R}, \ b \leqslant x \leqslant a\}$.

On termine cet appendice par l'alphabet grec, qu'il est très utile de connaître pour lire et écrire des mathématiques.

Minuscule	Majuscule	Nom	Minuscule	Majuscule	Nom
α	A	alpha	ν	N	nu
β	B	bêta	ξ	Ξ	xi
γ	Γ	gamma	o	O	omicron
δ	Δ	delta	π	Π	pi
ε	E	epsilon	ρ	P	rhô
ζ	Z	zêta	σ	Σ	sigma
η	H	êta	τ	T	tau
θ	Θ	thêta	υ	Υ	upsilon
ι	I	iota	ϕ	Φ	phi
κ	K	kappa	χ	X	khi
λ	Λ	lambda	ψ	Ψ	psi
μ	M	mu	ω	Ω	oméga

Annexe B

Distance à une partie

Soient (E, d) un espace métrique et A une partie non vide de E.

B.1 Généralités

Tout d'abord, nous allons donner la définition de la distance à une partie.

Définition B.1.1. *Soit $x \in E$. On définit la distance de x à A par :*

$$d_A(x) = \inf_{y \in A} d(x, y)$$

L'application $d_A \ : \ x \to d_A(x)$ est alors la distance à la partie A.

Nous allons à présent donner une caractérisation de la distance à une partie.

Lemme B.1.2. *On a l'équivalence : $a \leqslant b \Leftrightarrow \forall \varepsilon > 0, \ a \leqslant b + \varepsilon$.*

Démonstration : Dans le sens direct, on a : $a \leqslant b \leqslant b + \varepsilon$. Pour la réciproque, on raisonne par contraposée : si $b < a$, alors avec $\varepsilon = \frac{a-b}{2}$, on a : $b + \varepsilon < a$.

□

Proposition B.1.3. *Soit $x \in E$. On a : $d_A(x) = \inf\{r > 0 \mid B(x, r) \cap A \neq \emptyset\}$.*

Démonstration : On note : $\lambda = d_A(x)$ et $\mu = \inf\{r > 0 \mid B(x, r) \cap A \neq \emptyset\}$.

Montrons tout d'abord que : $\lambda \geqslant \mu$. Pour cela montrons que : $\forall \varepsilon > 0, \ \lambda + \varepsilon \geqslant \mu$. Soit $\varepsilon > 0$. Par définition de λ, on a : $\exists y \in A, \ \lambda \leqslant d(x, y) < \lambda + \varepsilon$. On pose $r = \lambda + \varepsilon$. Alors : $B(x, r) \cap A \neq \emptyset$. D'où : $\lambda + \varepsilon \geqslant \mu$. Et donc, en vertu de lemme ci-dessus : $\mu \leqslant \lambda$.

Réciproquement, montrons que : $\lambda \leqslant \mu$. Pour cela, montrons : $\forall \varepsilon > 0, \ \lambda \leqslant \mu + \varepsilon$. Soit $\varepsilon > 0$. Par définition de μ, on a : $\exists r > 0, \mu \leqslant r < \mu + \varepsilon$ et $B(x, r) \cap A \neq \emptyset$. Soit alors $y \in B(x, r) \cap A$. On a : $(y \in B(x, r)) \Leftrightarrow (\exists r', \ \mu \leqslant r' < r, \ d(x, y) = r')$. Comme de plus $y \in A$, on a : $r' \in \{d(x, y) \mid y \in A\}$. Donc : $\lambda \leqslant r'$. On a donc :

$$\lambda \leqslant r' < r < \mu + \varepsilon.$$

Et en particulier : $\lambda \leqslant \mu + \varepsilon$ et donc par le lemme : $\lambda \leqslant \mu$. On a donc bien : $\lambda = \mu$.

□

Nous allons maintenant passer à quelques premiers résultats.

Proposition B.1.4. *Soit B une autre partie non vide de E. On a alors :*

$$A \subset B \Rightarrow d_B \leqslant d_A.$$

Démonstration : Soit $x \in E$. Comme $A \subset B$, on a :

$$B(x,r) \cap A \neq \varnothing \Rightarrow B(x,r) \cap B \neq \varnothing$$

D'où :

$$\{r > 0 \mid B(x,r) \cap A \neq \varnothing\} \subset \{r > 0 \mid B(x,r) \cap B \neq \varnothing\}$$

et

$$\inf\{r > 0 \mid B(x,r) \cap B \neq \varnothing\} \leqslant \inf\{r > 0 \mid B(x,r) \cap A \neq \varnothing\}$$

Par la caractérisation que nous venons de donner, il vient : $d_B(x) \leqslant d_A(x)$. Or, ceci est vrai pour tout x dans E donc : $d_B \leqslant d_A$.

□

Proposition B.1.5. *On a : $\overline{A} = \{x \in E \mid d_A(x) = 0\}$.*

Démonstration : On a :

$$\begin{aligned} x \in \overline{A} &\Leftrightarrow \forall V \in V(x),\ V \cap A \neq \varnothing \ \Leftrightarrow\ \forall \varepsilon > 0,\ A \cap B(x,\varepsilon) \neq \varnothing \\ &\Leftrightarrow\ \forall \varepsilon > 0,\ \exists y \in A \cap B(x,\varepsilon)\ \Leftrightarrow \forall \varepsilon > 0,\ \exists y \in A \text{ et } d(x,y) \leqslant \varepsilon \\ &\Leftrightarrow\ \inf\{d(x,y) \mid y \in A\} = 0\ \Leftrightarrow\ d_A(x) = 0. \end{aligned}$$

□

Corollaire B.1.6. *Si A est fermée, alors : $x \in A = \overline{A} \Leftrightarrow d_A(x) = 0$.*

Proposition B.1.7. *On a : $d_A = d_{\overline{A}}$.*

Démonstration : Comme $A \subset \overline{A}$, on a : $d_{\overline{A}} \leqslant d_A$.

Soit $x \in E$. Soit $y \in \overline{A}$ tel que : $d_{\overline{A}} \leqslant d(x,y)$. On a :

$$\begin{aligned} y \in \overline{A} \quad &\Leftrightarrow \quad (\exists (z_n) \in A^{\mathbb{N}})\ (\exists \lim_{n\to+\infty} z_n = y) \\ &\Rightarrow \quad (\exists (z_n) \in A^{\mathbb{N}})\ (d(x,y) = \lim_{n\to+\infty} d(x,z_n)) \end{aligned}$$

Or : $\forall n \in \mathbb{N},\ d(x,z_n) \geqslant d_A(x)$. D'où par passage à la limite : $d(x,y) \geqslant d_A(x)$. De là par passage à l'inf : $d_{\overline{A}}(x) \geqslant d_A(x)$. On a donc : $d_{\overline{A}} \geqslant d_A$. Et enfin : $d_A = d_{\overline{A}}$.

□

B.2 Continuité

Nous allons tout d'abord montrer que d_A est continue ce que l'on déduit immédiatement de la proposition suivante :

Proposition B.2.1. *d_A est 1-lipschitzienne.*

Démonstration : Soit $(x,x') \in E^2$. Soit $y \in A$. On a : $d(x',y) \leqslant d(x',x) + d(x,y)$. Or : $d_A(x') \leqslant d(x',y)$. D'où : $d_A(x') \leqslant d(x',x) + d(x,y)$. D'où : $d_A(x') - d(x',x) \leqslant d(x,y)$. Alors $d_A(x') - d(x',x)$ est un minorant de : $\{d(x,y) \mid y \in A\}$. D'où : $d_A(x') - d(x',x) \leqslant d_A(x)$. On a donc : $d_A(x') - d_A(x) \leqslant d(x,x')$. Par symétrie des rôles joués par x et x', on a aussi : $d_A(x) - d_A(x') \leqslant d(x,x')$. On a donc bien le caractère 1-lipschitzien de d_A :

$$\forall (x,x') \in E^2, \mid d_A(x) - d_A(x') \mid \leqslant d(x,x').$$

□

Nous allons maintenant donner deux cas où pour x fixé dans E, le réel $d_A(x)$ est atteint.

Proposition B.2.2. *Soit $x \in E$. Lorsque A est compacte, $d_A(x)$ est atteint.*

Démonstration : On a : $y \mapsto d(x,y)$ est continue sur A compacte donc elle est bornée et y atteint ses bornes.

□

Proposition B.2.3. *Soit A une partie fermée de $\mathbb{R}^n$. Soit $x \in \mathbb{R}^n$. Alors $d_A(x)$ est atteinte.*

Démonstration : L'idée de cette preuve est très importante. Nous allons voir comment on se ramène au cas où A est compacte.

On pose : $R = d_A(x) + 1$. Alors : $B_f(x,R) \cap A = K \neq \emptyset$. Or, K est fermée comme intersection de deux fermés. De plus comme $K \subset B_f(x,R)$, K est bornée. Donc, comme on est en dimension finie, K est compacte. Alors, par la proposition ci-dessus, $d_K(x)$ est atteinte, mettons en a. Il nous reste à montrer que : $d(x,a) = d_A(x)$. Soit $y \in A$.

Si $y \in K : d(x,y) \geqslant d(x,a)$ par définition.

Si $y \in A - K : d(x,y) \geqslant R \geqslant d_K(x) = d(x,a)$.

Dans les deux cas, on a bien : $d(x,y) \geqslant d(x,a)$. Or, cela est valable pour tout y dans A. Donc : $d(x,a) = \inf_{y \in A} d(x,y) = d_A(x)$.

□

B.3 Applications

Nous allons maintenant donner deux applications de la distance à une partie.

Proposition B.3.1. *Si A est fermée, il existe une famille dénombrable d'ouverts $(\Omega_n)_{n\in\mathbb{N}}$ de E telle que :*

$$\forall n \in \mathbb{N},\ \Omega_{n+1} \subset \mathring{\Omega}_n, \ \ et\ \ A = \bigcap_{n\in\mathbb{N}} \Omega_n.$$

Démonstration : Posons pour $n \in \mathbb{N} : \Omega_n = \{x \in E \mid d_A(x) < \frac{1}{n+1}\}$. Alors, comme d_A est continue, Ω_n est un ouvert de E. De plus, on a :

$$\bigcap_{n\in\mathbb{N}} \Omega_n = \{x \in E \mid d_A(x) = 0\} = \overline{A} = A.$$

D'où le résultat voulu.

□

Proposition B.3.2. *Soient A et B deux fermés disjoints non vides. Alors, il existe deux ouverts disjoints U et V tels que : $A \subset U$ et $B \subset V$.*

Démonstration : Soit f définie sur E par : $f(x) = d_B(x) - d_A(x)$. Alors f est continue sur E comme somme de deux fonctions continues sur E. Si $x \in A$, alors $f(x) = d_B(x) > 0$. En effet, si $d_B(x) = 0$, alors comme B est fermé, $x \in B$. Mais dans ce cas, x est dans l'intersection de A et B ce qui est exclu. De là, posons : $U = \{x \in E \mid f(x) > 0\}$ et $V = \{x \in E \mid f(x) < 0\}$. Alors, U est un ouvert contenant A, V est un ouvert contenant B (on le montre comme pour U et A) et bien évidemment : $U \cap V = \varnothing$.

□