Topologie et théorie de la mesure, avancement du CM [i]

Georges Tomanov

Université Lyon I
Semestre d'automne 2021-2022

i. Parmi les plusieurs sources, on utilise de manière essentielle : le CM 2018-2019 enseigné par T.Blossier, M.Carrizosa et J.Melleray, le CM 2019-2020 enseigné par Y.Dabrowski et G.T., le CM 2020-2021 enseigné par G.T.,le livre "Probabilité", ch.1 et 2, par Ph.Barbe et M.Ledoux, le livre classique "Measure theory" par P.Halmos, le livre "Elements of the theory of functions and functional analysis" par A.N.Kolmogorov et S.V.Fomin, etc.

Chapitre 1

Rappels et premières définitions

1.1 Limites, sous-ensembles, dénombrabilité

....

Le cours du 15/09/2021

....

On introduit les "nombres" $-\infty$ et ∞ sur la droite réelle étandue $\overline{\mathbb{R}} \overset{def}{=} \mathbb{R} \cup \{-\infty, \infty\}$ et/ou sur la demi-droite réelle étandu $\overline{\mathbb{R}}^+ \overset{def}{=} \mathbb{R}^+ \cup \{\infty\}$. Si $x \in \mathbb{R}$ alors $-\infty < x < \infty$ et $-\infty < \infty$. La somme $x + y$ avec $x, y \in \overline{\mathbb{R}}$ est définie à l'exception du cas où $x = \pm\infty$ et $y = -x$. Le produit tx, $t \in \mathbb{R}, x \in \overline{\mathbb{R}}$ est défini sauf si $t = 0$ et $x = \pm\infty$.

Notations. Si $\emptyset \neq A \subset \overline{\mathbb{R}}$ alors

$$\sup A \overset{def}{=} \text{le plus petit majorant } M \text{ de } A,$$

$$\inf A \overset{def}{=} \text{le plus grand minorant } m \text{ de } A.$$

Si $x, y \in \overline{\mathbb{R}}$ alors on écrit souvent

$$x \vee y \overset{def}{=} \max\{x, y\},$$

$$x \wedge y \overset{def}{=} \min\{x, y\}.$$

Définition 1.1.
1. Un ensemble A est *dénombrable* s'il est en correspondance bijective avec $\mathbb{N}$, autrement dit, s'il existe une application *bijective* $x : \mathbb{N} \to A$ qui permet d'énumérer A (i.e. $A = \{x_n, n \in \mathbb{N}\}$.)
2. Un ensemble est *au plus dénombrables* (a.p.d.) s'il est soit fini soit dénombrable.

Définition 1.2. Soit (u_n) une suite de E. On appelle *suite extraite* ou *sous-suite* une suite de la forme $v_n = u_{\phi(n)}$, pour $\phi : \mathbb{N} \to \mathbb{N}$ une application strictement croissante.

Définition 1.3. On appelle *valeur d'adhérence* d'une suite réelle (u_n) toute limite d'une suite extraite convergente.

Vu en L1 :

Proposition 1.4. *Toute suite extraite d'une suite convergente converge vers la même limite. (Autrement dit, toute suite convergente n'a qu'une seule valeur d'adhérence, sa limite.) Réciproquement, une suite réelle converge si et seulement si elle n'a qu'une seule valeur d'adhérence.*

Définition 1.5. Si $(x_n) \subset \mathbb{R}$, alors $\limsup x_n = \lim_{n\to\infty} \sup_{k\geq n} x_k$ et $\liminf x_n = \lim_{n\to\infty} \inf_{k\geq n} x_k$.

Il est facile à voir que

$$\liminf_{n\to\infty} x_n = -\limsup_{n\to\infty} -x_n.$$

Proposition 1.6. *On a toujours, pour toute valeur d'adhérence l de (x_n) :*

$$\limsup_{n\to\infty} x_n \geq l \geq \liminf_{n\to\infty} x_n.$$

La suite (x_n) converge si et seulement si :

$$\limsup_{n\to\infty} x_n = \liminf_{n\to\infty} x_n.$$

Démonstration. Le deuxième point est une conséquence du premier point et de la Proposition 1.4. En passant à la suite des opposées il faut prouver que $\limsup_{n\to\infty} x_n \geq l$. La preuve est facile si $l = \infty$. Supposons que $l \in \mathbb{R}$. Si $(x_{\phi(n)})$ est extraite de (x_n) de limite l, on a

$$\sup_{k\geq n} x_k \geq \sup_{k\geq n} x_{\phi(k)}$$

car $\phi(k) \geq k$ et $\{\phi(k) : k \geq n\} \subset \{k \geq n\}$ pour chaque n. En passant à la limite, on obtient :

$$\limsup_{n\to\infty} x_n \geq \limsup_{n\to\infty} x_{\phi(n)}.$$

Donc, il suffit de montrer que si $l = \lim_{n\to\infty} x_n$ alors $l = \limsup_{n\to\infty} x_n$. Pour chaque $\epsilon > 0$ il existe $n(\epsilon) > 0$ tel que $l - \epsilon < x_n < l + \epsilon$ si $n > n(\epsilon)$. Donc,

$$l - \epsilon < \sup_{n>m} x_n < l + \epsilon \text{ pour tous } m \geq n(\epsilon) \Rightarrow \; l - \epsilon \leq \limsup_{n\to\infty} x_n < l + \epsilon.$$

Par conséquence, $l = \limsup_{n\to\infty} x_n$. □ □

Pour des fonctions réelles f_n, on définit les limsup et liminf ponctuellement

$$(\limsup_{n\to\infty} f_n)(x) = \limsup_{n\to\infty} f_n(x),$$

$$(\liminf_{n\to\infty} f_n)(x) = \liminf_{n\to\infty} f_n(x).$$

Notations.
1. $\mathcal{P}(\Omega)$ désigne l'ensemble de toutes les parties d'un ensemble Ω.
2. $A\Delta B = (A \setminus B) \cup (B \setminus A)$ est *la difference symétrique* de $A, B \in \mathcal{P}(\Omega)$.
3. Si $A \subset \Omega$ alors $A^c = \Omega \setminus A$ est *le complément de A (dans Ω).*

Rappel 1.7. On rappelle que $\emptyset^c = \Omega, \Omega^c = \emptyset, (A^c)^c = A$, $A \cup A^c = \Omega, A \cap A^c = \emptyset$. Pour une famille $(A_i)_{i\in I} \in \mathcal{P}(\Omega)$ on a :

$$\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c, \tag{1.1}$$

$$\left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} A_i^c. \tag{1.2}$$

Rappel 1.8. Soit $f : \Omega \to E$ une application. On rappelle les propriétés des images réciproques :

$$\begin{aligned}(f^{-1}(B))^c &= f^{-1}(B^c),\\ f^{-1}\Big(\bigcup_{i\in I} B_i\Big) &= \bigcup_{i\in I} f^{-1}(B_i),\\ f^{-1}\Big(\bigcap_{i\in I} B_i\Big) &= \bigcap_{i\in I} f^{-1}(B_i).\end{aligned} \tag{1.3}$$

Exercice 1.9. Soit $A, B, C \in \mathcal{P}(\Omega)$. Montrer que

1. $A\Delta B = (A \cup B) \setminus (A \cap B)$;
2. $(A\Delta B) \cap C = (A \cap C)\Delta(B \cap C)$.

Définition 1.10. Une suite $(A_n) \in \mathcal{P}(\Omega)$ est

1. croissante (dans $\mathcal{P}(\Omega)$) si $A_n \subset A_{n+1}$ pour tout n ;
2. décroissante (dans $\mathcal{P}(\Omega)$) si $A_n \supset A_{n+1}$ pour tout n ;
3. la suite (A_n) est une suite d.d.d. (acronyme pour "deux à deux disjointes") si $A_n \cap A_m = \emptyset$ pour $n \neq m$.

Notations.

1. $A_n \nearrow A$ signifie que la suite (A_n) est croissante et $A = \cup A_n$;
2. $A_n \searrow A$ signifie que la suite (A_n) est décroissante et $A = \cap A_n$;
3. si $(B_i)_{i\in I}$ est une famille d'ensembles d.d.d. on utilisera souvant la notation $\sqcup_{i\in I} B_i$ au lieu de $\cup_{i\in I} B_i$.

Exercice 1.11.

1. Soit A_1 et $A_2 \in \mathcal{P}(\Omega)$. Alors $A_1 \cup A_2 = (A_1 \cap A_2) \sqcup (A_1\Delta A_2)$ et $A_1\Delta A_2 = (A_1 \setminus A_2) \sqcup (A_2 \setminus A_1)$.
2. Soit $A_1, A_2, \cdots A_n \in \mathcal{P}(\Omega), n \geq 2$. Alors $A_1 \cup \cdots \cup A_n = B_1 \sqcup \cdots \sqcup B_m$ et chaque A_i est union de sous-ensembles parmi les $\{B_1, \cdots, B_m\}$. (Indication : recurrence sur n.)

Définition 1.12. Soient $n > 1$, $A_i \in \mathcal{P}(\Omega_i), 1 \leq i \leq n$. Alors $A_1 \times \cdots \times A_n$ est *un pavé* dans $\Omega_1 \times \cdots \times \Omega_n$.

Exercice 1.13.

1. L'union de deux pavés dans $\Omega_1 \times \Omega_2$ est égale à l'union d'au plus 5 pavés d.d.d.
2. Soit P_1 et P_2 des pavés dans $\Omega_1 \times \Omega_2$. Alors chacun des ensembles $P_1^c, P_1 \cap P_2, P_1 \setminus P_2, P_1\Delta P_2$ est union de pavés d.d.d.
3. L'union d'un nombre fini de pavés dans $\Omega_1 \times \Omega_2$ est égale à l'union d'un nombre fini de pavés d.d.d. (Indication : recurrence sur le nombre de pavés.)
4. Généraliser 2 et 3 pour les pavés dans $\Omega_1 \times \cdots \times \Omega_n, n \geq 2$.
5. Soit $\Omega_i = \mathbb{R}$ pour tous i. Le pavé $A_1 \times \cdots \times A_n$ est dite *connexe* si chaque A_i est un interval (fermé, ouvert ou semi-ouvert) dans $\mathbb{R}$. Alors l'union d'un nombre fini de pavés connexes dans $\mathbb{R}^n$ est égale à l'union d'un nombre fini de pavés connexes d.d.d. (Rappelons que $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$.)

Notation. Soit $A \subset \Omega$. *La fonction indicatrice* 1_A est définie par

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

Exercice 1.14. Soit $(A_n) \in \mathcal{P}(\Omega)$. On pose

$$\liminf A_n \overset{def}{=} \bigcup_n \bigcap_{k \geq n} A_k, \tag{1.4}$$

$$\limsup A_n \overset{def}{=} \bigcap_n \bigcup_{k \geq n} A_k. \tag{1.5}$$

1. Montrer que
$$\liminf A_n = \{\omega \in \Omega : \omega \in A_n \text{ pour n assez grand}\},$$
$$\limsup A_n = \{\omega \in \Omega : \omega \in A_n \text{ pour un nombre infini de n}\}.$$
2. Montrer que la limite supérieur des fonctions indicatrices de A_n est donnée par la fonction indicatrice de la limite supérieur des ensembles A_n :
$$\limsup_n 1_{A_n} = 1_{\limsup A_n}. \tag{1.6}$$
3. De même, montrer que :
$$\liminf_n 1_{A_n} = 1_{\liminf A_n}. \tag{1.7}$$

1.2 Compléments (facultatifs) sur les ensembles dénombrables

1.2.1 Ensembles finis et infinis

On va définir les ensembles finis comme les ensembles en bijection avec une partie de la forme

$$[\![1, n]\!] := \{k \in \mathbb{N} : 1 \leq k \leq n\}$$

pour $n \in \mathbb{N}$[i]. On remarquera que pour $n = 0$, on a $[\![1, n]\!] = \emptyset$.

On rappelle que $\mathfrak{S}_n$ est l'ensemble des bijections de $[\![1, n]\!]$ dans lui-même.

On commence par une proposition qui dit que ces ensembles ne sont pas en bijection.

Proposition 1.15. *Soient $m, n \in \mathbb{N}$. Si $[\![1, n]\!]$ est en bijection avec $[\![1, m]\!]$ alors $m = n$.*

Démonstration. On montre la propriété par récurrence sur n. Pour $n = 0$, l'énoncé est vrai car l'ensemble vide n'est en bijection qu'avec l'ensemble vide.

Supposons le résultat vrai au rang $n \geq 0$. Soit $h : [\![1, n+1]\!] \to [\![1, m]\!]$ une bijection (donc forcément $m \geq 1$ comme le premier ensemble n'est pas vide). On distingue 2 cas :

Le cas simple est le cas $h(n+1) = m$, alors on définit $g : [\![1, n]\!] \to [\![1, m-1]\!]$ par $g(k) = h(k)$. g est bien défini car h est une bijection donc pour $k \neq n+1$, $h(k) \neq h(n+1) = m$, donc $h(k) < m$. g est injective car la restriction d'une application injective est encore injective. Soit $l \in [\![1, m-1]\!]$ il existe $k \in [\![1, n+1]\!]$ tel que $h(k) = l$ par surjectivité de h. Mais $l \neq n+1$ par injectivité comme ci-dessus, donc $k \in [\![1, n]\!]$ et donc $g(k) = l$. Cela montre la surjectivité de g. Donc, par l'hypothèse de récurrence, $n = m-1$ et donc $m = n+1$.

Le second cas est le cas $h(n+1) < m$. On a vu au premier semestre l'existence de la transposition $\tau = (h(n+1), m)$ de sorte que $\tau \circ h$ est encore une bijection par composée de bijection et elle vérifie le premier cas. D'où l'étape suivante de la récurrence dans tous les cas. $\square$

Définition 1.16. On dit qu'un ensemble A est *fini* s'il existe un entier $n \in \mathbb{N}$, qui est alors unique (par le lemme précédent), et une bijection $f : [\![1, n]\!] \to A$. On écrit alors $Card(A) = n$ ou $|A| = n$. Un ensemble qui n'est pas fini est dit *infini*.

i. La section 1.2 est rédigée par Yoann Dabrowski.

1.2.2 Ensembles au plus dénombrables

On peut représenter les éléments d'un ensemble dénombrable A à l'aide d'une suite infinie en écrivant $A = \{x_n; n \geq 1\}$.

Proposition 1.17. *Les ensembles au plus dénombrables sont soit finis, soit dénombrables. De plus, pour une partie infinie $P \subset \mathbb{N}$, il existe une bijection strictement croissante et une seule de $\mathbb{N} \to P$.*

Démonstration. Les ensembles au plus dénombrables sont par définition en bijection avec les parties de $\mathbb{N}$. Dans le cas infini, il suffit de voir le second point pour obtenir la bijection avec $\mathbb{N}$. On définit par récurrence la bijection $f : \mathbb{N} \to P$. Plus précisément, on construit par récurrence sur n une application strictement croissante $f_n : [\![1, n]\!] \to P$ telle que pour tout $x \in Im(f_n), y \in P - Im(f_n)$, $x < y$ et $f_n|_{[\![1,k]\!]} = f_k$. Comme P, infini, il est non-vide donc admet un élément $a_0 = \min(P)$ On pose $f_0(0) = a_0$ d'où l'initialisation.

On suppose construit f_n, et on prend $a_{n+1} = min(P - Im(f_n))$ qui existe car cette partie est infinie de $\mathbb{N}$ donc non vide (si elle n'était pas infinie, P serait finie comme union finie de parties finies). On pose $f_{n+1}(k) = f_n(k), k \leq n, f_{n+1}(n+1) = a_{n+1}$ de sorte que par l'hyp de rec sur f_n, $a_{n+1} > f_n(k), k \leq n$ ce qui donne la stricte croissance de f_{n+1} en combinant avec celle de f_n. Enfin, si $y \in P - Im(f_{n+1}) \subset P - Im(f_n)$ on a par hyp de rec $y > f_n(k)k \leq n$ et $y > a_{n+1}$ car c'est le min donc $\geq$ et on a $y \neq a_{n+1}$ par construction. Donc la relation demandée à l'étape suivante est vérifiée.

On obtient f strictement croissante donc injective en rassemblant les valeurs des f_n qui s'accordent $(f(n) = f_n(n) = f_m(n), m \geq n)$.

Pour voir que f bijective, par l'absurde, sinon il existe $b \in P - Im(f)$ mais par stricte croissance d'entiers $f(n) \to \infty$ donc il existe n minimal tel que $b < f(n) = f_n(n)$ ce qui impose par minimalité $b > f(n-1)$ et contredit $f_n(n) = Min(P - Im(f_{n-1}))$ vu $b \in P - Im(f_{n-1})$.

Pour l'unicité, si g est une autre telle bijection $g^{-1} \circ f$ est une bijection strictement croissante de $\mathbb{N} \to \mathbb{N}$ ainsi que sa réciproque et le lemme suivant donne donc $g^{-1} \circ f(n) \geq n, f^{-1} \circ g(n) \geq n$. D'où par croissance de g, f appliquée encore à ces relations : $f = g$. □

Proposition 1.18. *Une application strictement croissante $f : \mathbb{N} \to \mathbb{N}$ (resp. $f : [\![0, n]\!] \to \mathbb{N}$) vérifie $f(p) \geq p$ pour tout p dans son domaine.*

Démonstration. Il suffit de voir le deuxième cas (en restreignant aux segments initiaux), on le montre par récurrence sur n. Si $n = 0$, $f(0) \in \mathbb{N}$ donc c'est évident. En supposant l'hypothèse vraie au rang n, on considère $f : [\![0, n+1]\!] \to \mathbb{N}$, la restriction à $[\![0, n]\!]$ vérifie l'hypothèse de récurrence, donc $f(p) \geq p$ pour $p \leq n$ et $f(n+1) > f(n) \geq n$ mais dans $\mathbb{N}$ cela implique $f(n+1) \geq n+1$ et conclut l'étape d'induction. □ □

On trouve l'équivalence avec la définition vu à la section 1.

Proposition 1.19. *Un ensemble P est au plus dénombrable si et seulement si il existe un surjection $f : \mathbb{N} \to P$.*

Démonstration. Pour l'implication directe, si P est dénombrable, la bijection de la définition convient, si P est fini, en bijection avec $[\![0, n-1]\!]$ alors le reste modulo n donne la surjection $\mathbb{N} \to [\![0, n-1]\!]$ qui composée à la bijection donne la surjection cherchée. Réciproquement, l'ensemble $f^{-1}(p), p \in P$ est une partie de $\mathbb{N}$ qui a un plus petit élément a_p : $a : P \to \mathbb{N}$ est l'injection cherchée. □

On va obtenir des exemples d'ensembles dénombrables les plus courants. Pour cela on a besoin de quelques méthodes de constructions.

Proposition 1.20. *1. La réunion d'une suite $(X_n)_{n\geq 0}$ d'ensembles finis d.d.d. est au plus dénombrable.*

2. *Un ensemble X est au plus dénombrable si et seulement si il admet une* suite exhaustive *de parties finies, c'est à dire une suite croissante de parties finies dont l'union est X.*
3. *Le produit cartésien d'un nombre fini d'ensembles au plus dénombrables est au plus dénombrable.*

Démonstration. 1. Soit $a_n = Card(X_n)$ et $A_n = \sum_{k=0}^{n} a_k$ $(A_{-1} = 0)$. On a des bijections h_n : $[\![A_{n-1}+1, A_n]\!] \to [\![1, a_n]\!] \to X_n$ qui induisent une application $h : \mathbb{N}^* \to \cup_n X_n$ dès qu'un nombre infini de X_i n'est pas vide, ou $h : [\![1, A_p]\!] \to \cup_n X_n$ qui est par construction surjective. L'injectivité des h_n et le fait que les X_n sont disjoints donne l'injectivité de h.

2. Si X est fini, on prend la suite constante, sinon, pour une bijection $h : \mathbb{N} \to X$ on prend $X_n = h([\![0, n]\!])$ comme suite croissante cherchée. Réciproquement, la suite croissante X_n donne une suite disjointe X_0, $X_{n+1} - X_n$ de parties finies, donc le point 1. donne que l'union est au plus dénombrable.

3. Une récurrence triviale ramène au cas du produit de 2 ensembles A, B. Soit $h : \mathbb{N} \to A$, $g : \mathbb{N} \to B$ des surjections données par la proposition 1.19. $f = h \times g : \mathbb{N}^2 \to A \times B$ est une surjection qui ramène au cas $\mathbb{N}^2$ qui admet pour suite exhaustive d'ensembles finis $[\![0, n]\!]^2$. □

Proposition 1.21. *Les ensembles $\mathbb{N}^k, k \in \mathbb{N}^*; \mathbb{Z}$ et $\mathbb{Q}, \mathbb{Q}^k$ sont infinis dénombrables.*

Démonstration. On a vu le cas du produit $\mathbb{N}^k$ au lemme précédent. $[\![-n, n]\!]$ est une suite exhaustive d'ensembles finis pour $\mathbb{Z}$ qui est donc au plus dénombrable par la proposition précédente, il est infini car il contient $\mathbb{N}$. Enfin $(p, q) \mapsto p/q$ est une surjection de $\mathbb{Z} \times \mathbb{N}^* \to \mathbb{Q}$, donc par la proposition 1.19 $\mathbb{Q}$ est au plus dénombrable, et infini car contient $\mathbb{N}$. □

Enfin, on améliore le lemme précédent.

Proposition 1.22. *Une réunion au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable.*

Démonstration. Soit $(X_n)_{n \geq 0}$ une suite d'ensembles dénombrables (si la suite est finie, on peut la prolonger en une suite infinie.). Soit $f_n : \mathbb{N} \to X_n$ une surjection donnée par la proposition 1.19. (Petite subtilité : passer de l'existence de chaque surjection à n fixé, à la suite de surjections n'est pas complètement anodin et utilise l'axiome du choix dénombrable). On considère la fonction $f : \mathbb{N}^2 \to \bigcup_{n \in \mathbb{N}} X_n$ définie par $f(n, p) = f_n(p)$ et en composant avec une surjection $\mathbb{N} \to \mathbb{N}^2$, on obtient le résultat par la réciproque dans la proposition juste citée. □

Les ensembles au plus dénombrables serviront de base aux probabilités discrètes.

1.2.3 Ensembles infinis non dénombrables

Les ensembles qui n'appartiennent pas aux catégories précédentes (finis ou infinis dénombrables) sont dits infinis non dénombrables. On va voir que par exemple, $\mathbb{R}$ et $\mathbb{C}$, $[a, b]$, $a < b$ sont infinis non dénombrables.

Le résultat clé est toujours un argument diagonal :

Proposition 1.23. *(Théorème de Cantor) Il n'existe pas de surjection $h : E \to \mathcal{P}(E)$ entre un ensemble E et l'ensemble de ses parties.*

Démonstration. En effet une application $h : E \to \mathcal{P}(E)$ permet de considérer l'ensemble $A = \{x \in E : x \notin h(x)\}$. Il n'existe pas de y tel que $h(y) = A$ car par l'absurde, si il existait, soit $y \in A$ et alors $y \notin h(y) = A$ une contradiction, soit $y \notin A$ et alors $y \in h(y) = A$ encore une contradiction. □

Remarque 1.24. En conséquence de ce lemme et de la proposition 1.19, $\mathcal{P}(\mathbb{N})$ n'est pas dénombrable (il est infini à cause de l'injection $x \mapsto \{x\}$ défini sur $\mathbb{N}$), car sinon on aurait une surjection de $\mathbb{N} \to \mathcal{P}(\mathbb{N})$. En conséquence $\{0,1\}^{\mathbb{N}}$, en bijection avec la fonction indicatrice n'est pas non-plus dénombrable.

Théorème 1.25. $[0,1]$ *et* $\mathbb{R}$ *ne sont pas dénombrables.*

En conséquence un intervalle quelconque $[a,b]$, en bijection avec $[0,1]$ ne l'est pas non plus. et un intervalle quelconque contenant au moins deux points (qui contient donc aussi un $[a,b]$) est aussi non-dénombrable.

Démonstration. On construit une injection $\varphi : \{0,1\}^{\mathbb{N}} \to [0,1]$ (le cas $\mathbb{R}$ s'en déduit. (l'image de cette injection va être l'ensemble triadique de Cantor). On fixe $a = (a_n) \in \{0,1\}^{\mathbb{N}}$ on définit une suite de segments emboîtés, on pose $J_0 = [0,1]$ et si $J_n = [x_n, y_n]$ alors on découpe l'intervalle en trois en posant $u_n = (2x_n + y_n)/3$ et $v_n = (x_n + 2y_n)/3$. Si $a_n = 0$, on pose $J_{n+1} = [x_n, u_n]$, et si $a_n = 1$, on pose $J_{n+1} = [v_n, y_n]$. On obtient par construction une suite de segments emboîtés, x_n, y_n sont des suites adjacentes et $y_n - x_n \leq 1/3^n$ (récurrence facile) donc l'intersection est un singleton $\cap_n J_n = \{\varphi(a)\}$.

Pour voir que φ est injective on note que si $a \neq a'$ sont deux suites et n le premier indice avec $a_n \neq a'_n$, alors $J_n \cap J'_n = \emptyset$ et les images sont donc distinctes. $\square$

Remarque 1.26. L'ensemble triadique de Cantor a plein de propriétés intéressantes. Topologiquement il est fermé, totalement disconnecté (les composantes connexes sont les singletons). Il est de longueur nulle (car inclus dans l'union sur tous les cas possibles des J_n dont la longueur perd un facteur $2/3$ à chaque n).

Exemple 1. L'ensemble des nombres irrationnels $\mathbb{R} \setminus \mathbb{Q}$ est non-dénombrable, car sinon son union avec $\mathbb{Q}$ à savoir $\mathbb{R}$ serait dénombrable, ce qui n'est pas le cas.

Chapitre 2

Notions basiques de la topologie : espaces métriques

2.1 Distances

La notion de *distance* a été introduite pour formaliser les propriétés d'une façon de mesurer l'écart entre des éléments d'un même ensemble ; ces propriétés sont modelées sur celles de la longueur d'un vecteur dans $\mathbb{R}^2$ ou $\mathbb{R}^3$.

Définition 2.1. Soit X un ensemble. Une *distance* sur X est une fonction $d\colon X \times X \to [0, +\infty[$ satisfaisant les propriétés suivantes :

1. $\forall x, y \in X \quad d(x, y) = 0 \Leftrightarrow x = y$.
2. $\forall x, y \in X \quad d(x, y) = d(y, x)$.
3. $\forall x, y, z \in X \quad d(x, z) \leq d(x, y) + d(y, z)$.

On dit alors que (X, d) est un *espace métrique*.

La première des trois propriétés ci-dessus est appelée *axiome de séparation* : elle dit en particulier que deux points distincts sont nécessairement à distance strictement positive. La deuxième propriété est l'*axiome de symétrie*. Enfin, la troisième est appelée *inégalité triangulaire*. C'est peut-être la moins intuitive ; dans $\mathbb{R}^2$, muni de sa notion usuelle de distance, elle correspond au fait que la longueur d'un côté d'un triangle est toujours inférieure à la somme des longueurs des deux autres côtés.

Propriété 2.2. Dans un espace métrique (X, d), pour tout $x, y, z \in X$, on a $|d(x, z) - d(y, z)| \leq d(x, y)$. (Preuve en exercice.)

Exemple 2. L'exemple le plus important d'espace métrique est $\mathbb{R}$ muni de la distance $d(x, y) = |x - y|$.

Exercice 2.3. Définir deux structures différentes d'espaces métriques sur la sphère de dimension deux.

Presque toutes les distances que nous rencontrerons proviennent d'une *norme*, dont nous rappelons la définition maintenant.

Définition 2.4. Soit X un *espace vectoriel* sur $\mathbb{R}$ (resp. sur $\mathbb{C}$). Une *norme* sur X est une fonction $\|\cdot\|\colon X \to [0, +\infty[$ satisfaisant les conditions suivantes :

1. $\forall x \in X \quad \|x\| = 0 \Leftrightarrow x = 0$.

2. $\forall x, y \in X \quad \|x+y\| \leq \|x\| + \|y\|$.
3. $\forall \lambda \in \mathbb{R}$ (resp. $\in \mathbb{C}$) $\forall x \in X \quad \|\lambda x\| = |\lambda| \cdot \|x\|$.

On dit alors que $(X, \|\cdot\|)$ est un *espace normé.*

Exemple 3. La norme associée au produit scalaire d'un espace euclidien est appelée norme euclidienne.

Exemple 4. Pour $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, on pose

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \text{ et } \|x\|_\infty = \max(|x_1|, \ldots, |x_n|).$$

Les fonctions $\|\cdot\|_1$ et $\|\cdot\|_\infty$ sont des normes sur $\mathbb{R}^n$. (Preuve en exercice.)

Propriété 2.5. Si $(X, \|\cdot\|)$ est un espace normé, alors la fonction $d\colon X \times X \to [0, +\infty[$ définie par $d(x, y) = \|x - y\|$ est une distance sur X. (Preuve en exercice.)Donc, les espaces normés sont des espaces métriques.

On dira que d définie comme ci-dessus est la distance induite par $\|\cdot\|$. Ce sont les distances avec lesquelles nous travaillerons principalement ; l'ensemble X ne sera pas forcément un espace vectoriel tout entier, c'est pourquoi la remarque suivante sera importante : si (X, d) un espace métrique, et $A \subseteq X$, alors la restriction de d à A munit A d'une structure d'espace métrique.

Exercice 2.6. Soit X un ensemble. On définit une fonction $d\colon X \times X \to [0, +\infty[$ en posant

$$\forall x, y \in X \quad d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{sinon} \end{cases}.$$

Montrer que d est une distance sur X, appelée *distance discrète*.

Définition 2.7. Soit (X, d) un espace métrique, $x \in X$ et $r \in \mathbb{R}$. On définit :
- La *boule ouverte* de centre x et de rayon r par

$$B(x, r) = \{y \in X \colon d(x, y) < r\}.$$

- La *boule fermée* de centre x et de rayon r par

$$\overline{B}(x, r) = \{y \in X \colon d(x, y) \leq r\}.$$

Remarque 2.8. Notons que $B(x, r) = \emptyset$ si $r \leq 0$ et $B(x, \infty) = X$.

Exercice 2.9. Montrer que pour $\mathbb{R}$ muni de la distance usuelle, $B(x, r) =]x - r, x + r[$ et $\overline{B}(x, r) = [x - r, x + r]$.

Exercice 2.10. Représenter les boules ouvertes/fermées dans $\mathbb{R}^2$ pour les distances associées aux normes $\|\cdot\|_1$, $\|\cdot\|_2$ et $\|\cdot\|_\infty$ de centre $(1, 0)$ et de rayon 1. (Rappelons que $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ où $x = (x_1, x_2)$.)

Exercice 2.11. Déterminer les boules ouvertes et les boules fermées d'un ensemble X muni de la distance discrète.

Proposition 2.12. *Soit (X, d_X) et (Y, d_Y) deux espaces métriques. Alors d_1 et d_∞ définies pour (x_1, y_1) et (x_2, y_2) dans $X \times Y$ par*

$$d_1((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2) \qquad \textit{et}$$

$$d_\infty((x_1, y_1), (x_2, y_2)) = \max(d_X(x_1, x_2), d_Y(y_1, y_2))$$

sont des distances sur $X \times Y$. On appellera d_∞ la distance produit *de d_X et d_Y.*

Démonstration. Les applications d_1 et d_∞ sont à valeurs dans $[0, +\infty[$ et satisfont de manière évidente les deux premières propriétés (axiome de séparation et axiome de symétrie).

Vérifions l'inégalité triangulaire : soit (x_1, y_1), (x_2, y_2) et (x_3, y_3) dans $X \times Y$. Alors,

$$\begin{aligned} d_1((x_1,y_1),(x_3,y_3)) = d_X(x_1,x_3) + d_Y(y_1,y_3) &\leq d_X(x_1,x_2) + d_X(x_2,x_3) + d_Y(y_1,y_2) + d_Y(y_1,y_3) \\ &= d_1((x_1,y_1),(x_2,y_2)) + d_1((x_2,y_2),(x_3,y_3)). \end{aligned}$$

On a

$$d_X(x_1,x_3) \leq d_X(x_1,x_2) + d_X(x_2,x_3) \leq \max(d_X(x_1,x_2), d_Y(y_1,y_2)) + \max(d_X(x_2,x_3), d_Y(y_2,y_3)) \ ,$$

d'où

$$d_X(x_1,x_3) \leq d_\infty((x_1,y_1),(x_2,y_2)) + d_\infty((x_2,y_2),(x_3,y_3)).$$

De même,

$$d_Y(y_1,y_3) \leq d_\infty((x_1,y_1),(x_2,y_2)) + d_\infty((x_2,y_2),(x_3,y_3)) \ .$$

Ainsi,

$$d_\infty((x_1,y_1),(x_3,y_3)) = \max(d_X(x_1,x_3), d_Y(y_1,y_3)) \leq d_\infty((x_1,y_1),(x_2,y_2)) + d_\infty((x_2,y_2),(x_3,y_3)) \ .$$

□

2.2 Suites dans un espace métrique

L'intérêt principal de la notion de distance, pour nous, est de pouvoir formaliser la notion de suites convergentes : une suite $(x_n)_{n\geq 0}$ converge vers x si et seulement si la distance $d(x_n, x)$ tend vers 0 quand n tend vers l'infini. Avec des quantificateurs, on obtient la définition suivante.

Définition 2.13. Soit (X, d) un espace métrique, $(x_n)_{n\geq 0}$ une suite d'éléments de X et $x \in X$. On dit que la suite $(x_n)_{n\geq 0}$ converge vers x si :

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N \ d(x_n, x) \leq \varepsilon \ ,$$

c'est-à-dire si :

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

Propriété 2.14. Soit (X, d) un espace métrique et (x_n) une suite d'éléments de X, qui converge à la fois vers x et x'. Alors $x = x'$, autrement dit, la limite d'une suite convergente est unique. (Preuve en exercice.)

Proposition 2.15. *Soit (X, d_X) et (Y, d_Y) deux espaces métriques. On munit $X \times Y$ de la distance produit d_∞ définie par la Proposition 2.12. Alors une suite $(x_n, y_n)_{n\geq 0}$ d'éléments de $X \times Y$ converge vers (x, y) si, et seulement si, $(x_n)_{n\geq 0}$ converge vers x et $(y_n)_{n\geq 0}$ converge vers y.* (Preuve en exercice.)

Ce résultat reste vrai si on remplace d_∞ par d_1. (Preuve en exercice.)

Ainsi, dans $\mathbb{R}^n$ muni de la distance induite par $\|\cdot\|_\infty$, on retrouve le fait qu'une suite est convergente si et seulement si chaque suite de coordonnées converge.

A priori, la notion de convergence dépend de la distance considérée sur X ; mais il arrive que des distances différentes aient les mêmes suites convergentes.

Définition 2.16. Soit X un espace vectoriel, et $\|\cdot\|_1, \|\cdot\|_2$ deux normes sur X. On dit que $\|\cdot\|_1$ et $\|\cdot\|_2$ sont *équivalentes* s'il existe des constantes m, M strictement positives et telles que :

$$\forall x \in X \quad m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1 \ .$$

Exercice 2.17. Montrer que la norme $\|\cdot\|_\infty$ et la norme euclidienne sur $\mathbb{R}^n$ sont équivalentes.

On reverra plus tard le théorème, normalement déjà connu et qu'il est en tout cas sans doute utile d'avoir en tête, selon lequel toutes les normes sur $\mathbb{R}^n$ sont équivalentes.

Cette définition a un analogue pour les distances :

Définition 2.18. Soit X un ensemble. Deux distances d_1, d_2 sur X sont dites *équivalentes* s'il existe des constantes m, M strictement positives et telles que :

$$\forall x, y \in X \quad md_1(x,y) \leq d_2(x,y) \leq Md_1(x,y) \ .$$

Bien sûr, si d_1, d_2 sont les distances associées à des normes équivalentes, alors elles sont elles-mêmes équivalentes. Cette définition est importante pour nous à cause de la propriété suivante.

Proposition 2.19. *Soit X un ensemble et d_1, d_2 deux distances sur X. Si d_1 et d_2 sont équivalentes, alors une suite $(x_n)_{n\geq 0}$ converge dans (X, d_1) si et seulement si elle converge dans (X, d_2).*

Autrement dit : deux distances équivalentes ont les mêmes suites convergentes.

Démonstration. Soit deux constantes m, M strictement positives telles que :

$$\forall x, y \in X \quad md_1(x,y) \leq d_2(x,y) \leq Md_1(x,y) \ .$$

Soit (x_n) une suite et x un élément de X. On a alors pour tout $n \in \mathbb{N}$,

$$md_1(x_n, x) \leq d_2(x_n, x) \leq Md_1(x_n, y) \ ,$$

ce qui entraîne que :

$$d_1(x_n, x) \underset{n\to+\infty}{\longrightarrow} 0 \text{ si et seulement si } d_2(x_n, x) \underset{n\to+\infty}{\longrightarrow} 0 \ .$$

□

Définition 2.20. Soit X un ensemble, et $(x_n)_{n\geq 0}$ une suite d'éléments de X. Une *suite extraite* de $(x_n)_{n\geq 0}$ est une *sous-suite* de la forme $(x_{n_k})_{k\geq 0}$ où $(n_k)_{k\geq 0}$ est une suite strictement croissante d'entiers, ou, de manière équivalente, de la forme $(x_{\varphi(k)})_{k\geq 0}$, où $\varphi\colon \mathbb{N} \to \mathbb{N}$ est une fonction *strictement croissante* (il suffit de poser pour $k \geq 0$, $\varphi(k) = n_k$).

Intuitivement : une suite extraite est obtenue en ne gardant que certains termes de la suite $(x_n)_{n\geq 0}$ et en oubliant les autres ; par exemple, la suite $(x_{2k})_{k\geq 0}$ est une suite extraite de la suite $(x_n)_{n\geq 0}$.

Propriété 2.21. Soit $\varphi\colon \mathbb{N} \to \mathbb{N}$ une fonction strictement croissante. Alors pour tout $k \in \mathbb{N}$ on a $\varphi(k) \geq k$ (et en particulier $\varphi(k)$ tend vers $+\infty$!). (Preuve en exercice.)

Proposition 2.22. *Soit (X, d) un espace métrique. Si $(x_n)_{n\geq 0}$ est une suite convergente alors toute suite extraite $(x_{\varphi(k)})_{k\geq 0}$ est convergente.* (Preuve en exercice.)

Exercice 2.23. Soit (X, d) un espace métrique et (x_n) une suite d'éléments de X. Supposons d'abord que les suites (x_{2k}) et (x_{2k+1}) convergent vers la même limite. Montrer que (x_n) est convergente. Ce résultat reste-t-il vrai si l'on ne suppose pas que les limites de (x_{2k}) et (x_{2k+1}) sont égales ? Montrer que si (x_{2k}), (x_{2k+1}) et (x_{3k}) sont toutes les trois convergentes alors (x_n) est convergente.

Définition 2.24. Une suite (x_n) est dite monotone si $a_n \geq a_{n+1}$ pour tous n ou $a_n \leq a_{n+1}$ pour tous n.

Proposition 2.25. *Toute suite de réels admet une sous-suite monotone.*

Démonstration. Soit (x_n) une suite de réels. Pour cette preuve, on dira qu'un entier n est un pic, si pour tout $m > n$, $x_n > x_m$. S'il y a une infinité de pics $n_0 < n_1 < n_2 < \ldots < n_k < \ldots$, alors la suite extraite (x_{n_k}) est strictement décroissante.

Sinon, il existe un entier N tel qu'aucun entier $n \geq N$ n'est un pic. On construit alors par récurrence une suite strictement croissante d'entiers (n_k) tel que la sous-suite (x_{n_k}) soit croissante : pour l'initialisation, on pose $n_0 = N$. Supposons que l'on a choisit $N = n_0 < n_1 < \ldots < n_k$, tel que $x_{n_0} \leq \ldots \leq x_{n_k}$ (condition vide si $k = 0$). Comme $n_k \geq N$, n_k n'est pas un pic et par conséquent il existe $n_{k+1} > n_k$ tel que $x_{n_k} \leq x_{n_{k+1}}$. □

2.3 Ouverts et fermés

Si on est intéressé plus par la notion de "proximité" induite par une distance que par les valeurs exactes de la distance, on est amené à la notion d'ouvert : $A \subseteq X$ est ouvert ssi, dès que $a \in A$, tout point suffisamment proche de a appartient à A. Autrement dit, A contient une petite boule non vide autour de chacun des ses points.

Définition 2.26. Soit (X, d) un espace métrique. Une partie A de X est un *ouvert* si

$$\forall a \in A \ \exists r > 0 \ B(a, r) \subseteq A.$$

Une boule ouverte est bien un ouvert ! L'ensemble X tout entier et l'ensemble vide $\emptyset$ sont des ouverts. Dans $\mathbb{R}$, muni de la distance usuelle, tout intervalle ouvert est un ouvert (intervalles de la forme $]a, b[$, $]-\infty, a[$, $]b, +\infty[$ et $]-\infty, +\infty[$, avec $a, b \in \mathbb{R}$).

Exercice 2.27. Vérifier qu'une partie A d'un espace métrique est un ouvert si et seulement si

$$\forall a \in A \ \exists n \in \mathbb{N}^* \ \overline{B}(a, 1/n) \subseteq A.$$

Exercice 2.28. Vérifier qu'une partie A d'un espace métrique est un ouvert si et seulement si

$$\forall a \in A \ \exists r > 0 \ \overline{B}(a, r) \subseteq A.$$

La notion d'ouvert dépend de la distance sur X ; par contre pour deux distances équivalentes les ouverts sont les mêmes.

Exercice 2.29.

1. Donner un exemple de deux distances sur $\mathbb{R}$ qui ne définissent pas les mêmes ouverts.
2. Soit X un ensemble et d_1, d_2 deux distances équivalentes sur X. Montrer qu'une partie A de X est un ouvert de (X, d_1) si et seulement si elle est un ouvert de (X, d_2).

Théorème 2.30. *Soit A une partie d'un espace métrique (X, d). Alors, A est un ouvert si et seulement pour toute suite (x_n) d'éléments de X qui converge vers un élément de A, alors x_n appartient à A pour tout n suffisamment grand.*

Démonstration. Supposons A ouvert et considérons une suite (x_n) d'éléments de X qui converge vers $a \in A$. Comme A est ouvert, il existe $r > 0$ tel que $B(a, r) \subseteq A$. La convergence de la suite vers a, entraîne qu'il existe un entier N tel que pour tout $n \geq N$, $d(x_n, a) < r$, c'est-à-dire tel que pour tout $n \geq N$, $x_n \in B(a, r) \subseteq A$.

Réciproquement, si A n'est pas ouvert, il existe $a \in A$ tel que pour tout $n \in \mathbb{N}$, $B(a, \frac{1}{n+1}) \not\subseteq A$. On choisit ainsi pour tout n, un élément $x_n \in B(a, \frac{1}{n+1}) \setminus A$. Alors aucun élément de la suite x_n n'appartient à A mais elle converge vers a qui appartient à A. □

Définition 2.31. Une partie A d'un espace métrique est un *fermé* si son complémentaire $X \setminus A$ est ouvert.

Une boule fermée est bien un fermé ! L'ensemble X tout entier et l'ensemble vide sont des fermés. Dans $\mathbb{R}$, muni de la distance usuelle, tout intervalle fermé est un fermé.

Les fermés sont caractérisés par la propriété suivante : la limite de toute suite convergente d'éléments d'un fermé est dans ce fermé. Notons que toute propriété des ouverts se traduit, par passage au complémentaire, en une propriété des fermés.

Proposition 2.32. *Soit A une partie d'un espace métrique (X, d). Alors, A est un fermé si et seulement pour toute suite (x_n) d'éléments de A qui converge vers un élément x de X on a que x appartient à A.*

Démonstration. Supposons A fermé. Soit (x_n) une suite d'éléments de A qui converge vers un élément x. Comme aucun x_n n'est dans l'ouvert $X \setminus A$, (encore moins pour tout n suffisamment grand), d'après le théorème 2.30, x ne peut appartenir à l'ouvert $X \setminus A$ et donc appartient à A.

Réciproquement, si A n'est pas fermé, c'est-à-dire si $X \setminus A$ n'est pas ouvert, alors le théorème 2.30 garantit l'existence d'une suite (x_n) d'éléments n'appartenant pas à $X \setminus A$, c'est-à-dire appartenant à A, qui converge vers un élément de $X \setminus A$. □ □

Proposition 2.33. *Dans un espace métrique (X, d) donné,*

1. *si $(O_i)_{i \in I}$ est une famille d'ouverts, alors $\bigcup_{i \in I} O_i$ est également un ouvert (une* **union quelconque** *d'ouverts est un ouvert) ;*
2. *si $O_1, \ldots, O_n$ sont des ouverts, alors $O_1 \cap \ldots \cap O_n$ est également un ouvert (une* **intersection finie** *d'ouverts est un ouvert) ;*
3. *si $(F_i)_{i \in I}$ est une famille de fermés, alors $\bigcap_{i \in I} F_i$ est également un fermé (une* **intersection quelconque** *de fermés est un fermé) ;*
4. *si $F_1, \ldots, F_n$ sont des fermés, alors $F_1 \cup \ldots \cup F_n$ est également un fermé (une* **union finie** *de fermés est un fermé).*

Démonstration. (1) Soit $a \in \bigcup_{i \in I} O_i$, alors il existe $i_0 \in I$ tel que $a \in O_{i_0}$. Comme O_{i_0} est ouvert, il existe $r > 0$, tel que $B(a, r) \subseteq O_{i_0} \subseteq \bigcup_{i \in I} O_i$, ce qui permet de conclure.

(2) Soit $a \in O_1 \cap \ldots \cap O_n$, alors pour chaque $i \in \{1, \ldots, n\}$, on a $a \in O_i$ qui est un ouvert. Ainsi, pour chaque $i \in \{1, \ldots, n\}$, il existe $r_i > 0$ tel que $B(a, r_i) \subseteq O_i$. En posant $r = \min(r_1, \ldots, r_n)$, on en déduit que la boule ouverte $B(a, r)$ est incluse dans chacun des O_i et donc que $B(a, r) \subseteq O_1 \cap \ldots \cap O_n$, ce qui permet de conclure.

(3) On a $X \setminus \bigcap_{i \in I} F_i = \bigcup_{i \in I} (X \setminus F_i)$, qui est ouvert par (1). Ainsi, $\bigcap_{i \in I} F_i$ est fermé.

(4) On a $X \setminus (F_1 \cup \ldots \cup F_n) = (X \setminus F_1) \cap \ldots \cap (X \setminus F_n)$, qui est ouvert par (2), et donc $F_1 \cup \ldots \cup F_n$ est fermé. □

Exercice 2.34.
1. Donner un exemple d'une intersection d'ouverts qui n'est pas un ouvert.
2. Donner un exemple d'une union de fermés qui n'est pas un fermé.

Définition 2.35. Un espace topologique est un couple $(X, \mathcal{O})$ où $\mathcal{O} \subset \mathcal{P}(X)$ et

1. $\emptyset$ et $X \in \mathcal{O}$;
2. Pour chaque famille (dénonbrable ou non-dénombrable) $(U_i)_{i \in I}, U_i \in \mathcal{O}$, on a $\bigcup_{i \in I} U_i \in \mathcal{O}$;
3. Pour chaque famille finie $U_1, \cdots, U_n \in \mathcal{O}$ on a $\bigcap_{1 \leq i \leq n} U_i \in \mathcal{O}$. (Les éléments de $\mathcal{O}$ sont les *ouverts* et les éléments de $\{U^c : U \in \mathcal{O}\}$ sont les *fermés* de l'espace topologique $(X, \mathcal{O})$.)

Exemple 5. 1. Si $\mathcal{O} = \mathcal{P}(X)$ alors $(X, \mathcal{O})$ est un espace topologique avec topologie discrète.

2. Si $X = \mathbb{R}^n$ ou $\mathbb{C}^n$ alors $\mathcal{O}$ consiste des ouverts dans le sens habituel.

3. Par Proposition 2.33 chaque espace métrique est un espace topologique.

Proposition 2.36. *Soient (X, d) et (X, D) des espaces métriques tels que les distances d et D sur X sont équivalentes, i.e. ils existent $m > 0$ et $M > 0$ tels que $md(x, y) \leq D(x, y) \leq Md(x, y)$ pour tous $x, y \in X$. Alors la topologie de (X, d) coincide avec la topologie de (X, D).*

Démonstration. Soit $r > 0$. Alors $d(a, x) < r/M$ implique $D(a, x) \leq Md(a, x) < r$. Donc, $B_d(a, r/M) \subset B_D(a, r)$. Cela montre que chaque ouvert de (X, D) et ouvert de (X, d). Par le même argument chaque ouvert de (X, d) et ouvert de (X, D). □

Définition 2.37. Étant donné un espace métrique (X, d) (ou, plus généralement, un espace topologique X), et une partie $A \subseteq X$, l'*intérieur* de A, dénoté $\mathring{A}$, est la réunion de tous les ouverts contenus dans A, c'est-à-dire

$$\mathring{A} = \bigcup_{\substack{O \text{ ouvert} \\ O \subseteq A}} O.$$

Par définition, $\mathring{A}$ est un ouvert, est contenu dans A, et il contient tous les autres ouverts contenus dans A : c'est le plus grand ouvert contenu dans A. Attention, $\mathring{A}$ peut tout à fait être vide même si A ne l'est pas !

Propriété 2.38. Soit A une partie d'un espace métrique (X, d). Alors pour tout $x \in X$, on a

$$x \in \mathring{A} \iff \exists r > 0 \text{ tel que } B(x, r) \subseteq A.$$

(Preuve en exercice.)

Définition 2.39. Étant donné un espace métrique (X, d) (ou, plus généralement, un espace topologique X), et une partie $A \subseteq X$, l'*adhérence* de A, notée $\overline{A}$, est l'intersection de tous les fermés contenant A, c'est-à-dire

$$\overline{A} = \bigcap_{\substack{F \text{ fermé} \\ F \supseteq A}} F.$$

Cette fois, $\overline{A}$ est un fermé, contient A, et est contenu dans tous les autres fermés qui contiennent A : c'est le plus petit de tous les fermés contenant A.

Propriété 2.40. Soit X un espace topologique et $A \subseteq X$. On a $\overline{X \setminus A} = X \setminus \mathring{A}$ et $\overbrace{X \setminus A}^{\circ} = X \setminus \overline{A}$. (Preuve en exercice.)

Le cours du 22/09/2021

Propriété 2.41. Soit (X, d) un espace métrique, $A \subseteq X$ et $x \in X$. Il existe une suite d'éléments de A qui converge vers x si, et seulement si, $B(x, r) \cap A \neq \emptyset$ pour tout $r > 0$. Comment généraliser la propriété pour les espaces topologique ? (Preuve en exercice.)

Ceci nous permet d'établir la caractérisation suivante de l'adhérence, qui est fondamentale.

Proposition 2.42. *Soit (X, d) un espace métrique, A une partie de X et $x \in X$. Alors $x \in \overline{A}$ si, et seulement si, il existe une suite d'éléments de A qui converge vers x.*

Démonstration. Supposons qu'il existe une suite d'éléments de A qui converge vers $x \in X$. Alors, pour tout fermé F contenant A, il existe aussi une suite (la même !) d'éléments de F qui converge vers x ; par conséquent, x appartient à F. Donc x appartient à tous les fermés qui contiennent A : $x \in \overline{A}$.

Réciproquement, supposons qu'il n'existe pas de suite d'éléments de A qui converge vers $x \in X$. Alors, d'après la propriété 2.41, il doit exister $r > 0$ tel que $B(x,r) \cap A = \emptyset$. Autrement dit, $X \setminus B(x,r)$ est un fermé de X, contenant A, mais auquel x n'appartient pas : $x \notin \overline{A}$. $\square$

Exercice 2.43. Déterminer l'intérieur et l'adhérence de $\mathbb{Q}, [0,1] \cap \mathbb{Q},]0,1[\cap \mathbb{Q}$ (dans $\mathbb{R}$ muni de sa distance usuelle).

Proposition 2.44. *Soit $(E, \|\cdot\|)$ un espace vectoriel normé (on considère E muni de la distance induite par sa norme). Soit $a \in E$ et $r > 0$.*

1. *L'intérieur de la boule fermée $\overline{B}(a,r)$ est la boule ouverte $B(a,r)$.*
2. *L'adhérence de la boule ouverte $B(a,r)$ est la boule fermée $\overline{B}(a,r)$.*

(Preuve en exercice.)

Exercice 2.45. Les deux propriétés de la proposition ci-dessus sont elles vraies pour tout espace métrique ? (Indication : on peut considérer la distance discrète.)

Définition 2.46. Soit (X,d) un espace métrique, et $A \subseteq X$. On dit que A est *dense* dans X si $\overline{A} = X$.

Exercice 2.47. On considère $\mathbb{R}$ avec la topologie (la métrique) euclidienne. Montrer que $\mathbb{Q}$ et $\mathbb{R} \setminus \mathbb{Q}$ sont tous les deux denses dans $\mathbb{R}$.

Chapitre 3

Fonctions continues entre espaces métriques

3.1 Fonctions continues et uniformément continues

Maintenant qu'on sait ce qu'est une distance, on peut définir la continuité pour des fonctions entre espaces métriques, plutôt que de $\mathbb{R}$ dans $\mathbb{R}$; c'est essentiellement la même chose, en remplaçant $|x-y|$ (qui n'a apriori pas de sens dans un espace métrique) par $d(x,y)$.

Définition 3.1. Soit (X,d) et (Y,D) deux espaces métriques, $f\colon X \to Y$ et $x \in X$. On dit que f est *continue en* x si :

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x' \in X \ d(x,x') < \delta \Rightarrow D(f(x),f(x')) < \varepsilon \ .$$

On dit que f est *continue sur* X si elle est continue en x pour tout $x \in X$, autrement dit :

$$\forall x \in X \ \forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x' \in X \ d(x,x') < \delta \Rightarrow D(f(x),f(x')) < \varepsilon \ .$$

Ou encore (l'ordre dans lequel on écrit les deux $\forall$ ne change pas le sens de l'énoncé) :

$$\forall \varepsilon > 0 \ \forall x \in X \ \exists \delta > 0 \ \forall x' \in X \ d(x,x') < \delta \Rightarrow D(f(x),f(x')) < \varepsilon \ .$$

Remarque 3.2. Il faut bien comprendre que, ci-dessus, δ dépend de ε *et du point* x *où l'on se place.* Une définition plus forte imposerait que le même δ fonctionne pour tous les $x \in X$ simultanément ; dans ce cas, on dit que f est *uniformément continue.*

Définition 3.3. Soit (X,d) et (Y,D) deux espaces métriques, et $f\colon X \to Y$. On dit que f est *uniformément continue sur* X si

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x \in X \ \forall x' \in X \ d(x,x') < \delta \Rightarrow D(f(x),f(x')) < \varepsilon \ .$$

Par rapport à la définition de la continuité, on a remplacé "$\forall x \in X \ \exists \delta > 0$" par "$\exists \delta > 0 \ \forall x \in X$" : δ dépend toujours de ε, mais ne dépend plus de x. Toute fonction uniformément continue est continue, mais la réciproque est fausse.

Exercice 3.4.
1. Montrer que la fonction $x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$.
2. Montrer qu'une fonction réelle, périodique et continue est uniformément continue.

Définition 3.5. Soit (X,d) et (Y,D) deux espaces métriques. On dit que $f\colon X \to Y$ est *lipschitzienne* s'il existe $K > 0$ tel que

$$\forall x,x' \in X \ D(f(x),f(x')) \leq K d(x,x') \ .$$

Propriété 3.6. Toute fonction lipschitzienne est uniformément continue. (Preuve en exercice.)

Exercice 3.7. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction dérivable, et telle qu'il existe M satisfaisant $|f'(x)| \leq M$ pour tout $x \in \mathbb{R}$. Montrer que f est lipschitzienne.

Exercice 3.8. Soit $f : [0, \frac{1}{2}] \colon \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} 1/\ln x & \text{si } x \in]0, 1/2] \\ 0 & \text{si } x = 0 \end{cases}$. Montrer que f est continue mais NON-lipschitzienne. Selon le théorème de Heine démontré plus loin (voir Théorème 4.13) la fonction f est uniformément continue !

Théorème 3.9. *Soit (X, d) et (Y, D) deux espaces métriques, $f : X \to Y$ une fonction et $x \in X$. La fonction f est continue en x si et seulement si pour toute suite (x_n) d'éléments de X qui converge vers x la suite $(f(x_n))$ converge vers $f(x)$.*

Démonstration. $\Rightarrow$) Supposons tout d'abord que f est continue en x, et fixons une suite (x_n) qui converge vers x ainsi que $\varepsilon > 0$. D'une part il existe $\delta > 0$ tel que $D(f(x), f(x')) \leq \varepsilon$ dès que $d(x, x') < \delta$; d'autre part il existe $N \in \mathbb{N}$ tel que $d(x_n, x) \leq \delta$ pour tout $n \geq N$. Alors, pour tout $n \geq N$, on a $d(f(x_n), f(x)) \leq \varepsilon$, ce qui prouve que $(f(x_n))$ converge vers $f(x)$.

$\Leftarrow$) Réciproquement, supposons que f ne soit pas continue en x :

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists y \in X \; d(x, y) < \delta \text{ et } d(f(x), f(y)) \geq \varepsilon \; .$$

Fixons $\varepsilon > 0$ comme ci-dessus, et appliquons la propriété pour $\delta = \frac{1}{n}$: ceci nous donne une suite (y_n) telle que $d(x, y_n) < \frac{1}{n}$ pour tout $n \in \mathbb{N}^*$ (en particulier, (y_n) converge vers x) mais $d(f(y_n), f(x)) \geq \varepsilon$ (par conséquent, $f(y_n)$ ne converge pas vers $f(x)$). $\square$

Théorème 3.10. *Soit (X, d) et (Y, D) deux espaces métriques et $f : X \to Y$ une fonction. Les propriétés suivantes sont équivalentes :*

1. *f est continue.*
2. *Pour tout ouvert O de Y, $f^{-1}(O)$ est un ouvert de X.*
3. *Pour tout fermé F de Y, $f^{-1}(F)$ est un fermé de X.*

On rappelle que $f^{-1}(A) = \{x \in X : f(x) \in A\}$ désigne l'*image inverse* de A par f.

Démonstration. Supposons que f est continue, et soit O un ouvert de Y. Fixons $x \in f^{-1}(O)$, et considérons une suite (x_n) qui tend vers x. Alors $f(x_n)$ tend vers $f(x)$ puisque f est continue, donc $f(x_n)$ appartient à O pour n suffisamment grand puisque $f(x) \in O$ et O est ouvert. Par conséquent, $x_n \in f^{-1}(O)$ pour n suffisamment grand, ce qui nous montre que $f^{-1}(O)$ est ouvert, et on a montré que (1) $\Rightarrow$ (2).

Montrons que (2) $\Rightarrow$ (1) ; supposons donc de nouveau que (2) soit vérifié, et considérons $x \in X$ et $\varepsilon > 0$. Puisque $B(f(x), \varepsilon)$ est un ouvert contenant $f(x)$, son image inverse est par hypothèse un ouvert contenant x, par conséquent il existe $\delta > 0$ tel que $B(x, \delta) \subseteq f^{-1}(B(f(x), \varepsilon))$, c'est-à-dire :

$$\forall x' \in X \; d(x, x') < \delta \to D(f(x), f(x')) < \varepsilon \; .$$

On a bien montré que f est continue.

Si (2) est vrai et F est fermé dans Y, alors $Y \setminus F$ est ouvert et par hypothèse on obtient que $f^{-1}(Y \setminus F) = X \setminus f^{-1}(F)$ est ouvert dans X, autrement dit $f^{-1}(F)$ est fermé dans X. Ceci établit l'implication (2) $\Rightarrow$ (3), et en fait le même argument de passage au complémentaire donne l'implication réciproque (3) $\Rightarrow$ (2). $\square$

On voit dans cette preuve qu'il vaut mieux être à l'aise avec les propriétés de l'image inverse par une fonction... (Ce très important dans la partie du cours consacrée à la théorie de la mesure.)

Exercice 3.11. Soit X, Y deux ensembles, $f\colon X \to Y$ une fonction. Montrer que, pour tout $A, B \subseteq Y$ on a $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$ et $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$.

Exercice 3.12. Soit l'application de $\mathbb{R}$ dans $\mathbb{R}$, $f\colon x \mapsto x^2$.

1. Déterminer les ensembles suivants : $f([-3,-1])$, $f([-2,1])$, $f([-3,-1]\cup[-2,1])$ et $f([-3,-1]\cap[-2,1])$. Les comparer.
2. Mêmes questions avec les ensembles $f^{-1}(]-\infty,2])$, $f^{-1}([1,+\infty[)$, $f^{-1}(]-\infty,2]\cup[1,+\infty[)$ et $f^{-1}(]-\infty,2]\cap[1,+\infty[)$.

Proposition 3.13. *Soit (X, d_X), (Y, d_Y) et (Z, d_Z) trois espaces métriques, ainsi que $f\colon Y \to Z$ et $g\colon X \to Y$ deux fonctions continues. Alors $f \circ g\colon X \to Z$ est continue.*

Démonstration. Fixons $\varepsilon > 0$. Comme f est continue, il existe $\delta_1 > 0$ tel que pour tout $y, y' \in Y$ on ait $d_Y(y, y') < \delta_1 \Rightarrow d_Z(f(y), f(y')) < \varepsilon$. Puis, comme g est continue, il existe δ_2 tel que pour tout $x, x' \in X$ on ait $d_X(x, x') < \delta_2 \Rightarrow d_Y(g(x), g(x')) < \delta_1$. On a alors, pour tout $x, x' \in X$:

$$d_X(x, x') < \delta_2 \Rightarrow d_Y(g(x), g(x')) < \delta_1 \Rightarrow d_Z(f(g(x)), f(g(x')) < \varepsilon \ .$$

On vient de prouver que $f \circ g$ est continue. □ □

Exercice 3.14. On munit $\mathbb{R}^2$ de la distance induite par $\|\cdot\|_\infty$, et $\mathbb{R}$ de sa distance usuelle. Montrer que les fonctions $(x, y) \mapsto x + y$ et $(x, y) \mapsto xy$ sont continues.

Exercice 3.15. Soit (X, d) un espace métrique et $f, g\colon X \to \mathbb{R}$ deux fonctions continues. Montrer que la somme $f + g$ et le produit fg sont également des fonctions continues.

Exercice 3.16. Soit (X, d_X), (Y, d_Y) et (Z, d_Z) trois espaces métriques, ainsi que $f\colon Y \to Z$ et $g\colon X \to Y$ deux fonctions uniformément continues. Montrer que $f \circ g\colon X \to Z$ est uniformément continue.

3.2 Suites de fonctions

Tout comme la continuité, les notions de convergence simple/uniforme de suites de fonctions qu'on connaît pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ s'étendent sans difficultés aux fonctions entre espaces métriques.

Définition 3.17. Soit (X, d), (Y, D) deux espaces métriques, et (f_n) une suite de fonctions de X dans Y. On dit que (f_n) converge *simplement* vers une fonction $f\colon X \to Y$ si pour tout $x \in X$ la suite $(f_n(x))$ converge vers $f(x)$; autrement dit :

$$\forall x \in X \ \forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall n \geq N \ D(f_n(x), f(x)) < \varepsilon \ .$$

Ci-dessus, N dépend à la fois de ε et de x ; comme dans la définition de la continuité, on pourrait demander que N ne dépende que de ε, et on obtient ainsi la définition de la convergence *uniforme*.

Définition 3.18. Soit (X, d), (Y, D) deux espaces métriques, et (f_n) une suite de fonctions de X dans Y. On dit que (f_n) converge *uniformément* vers une fonction $f\colon X \to Y$ si

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall x \in X \ \forall n \geq N \ D(f_n(x), f(x)) < \varepsilon \ .$$

Bien entendu, la convergence uniforme entraîne la convergence simple. La réciproque est fausse, comme le montre l'exercice suivant.

Exercice 3.19. Pour tout $n \in \mathbb{N}^*$, on définit $f_n : [0,1] \to [0,1]$ en posant

$$f_n(x) = \begin{cases} 0 & \text{si } x \geq \frac{1}{n} \\ 1 - nx & \text{si } 0 \leq x \leq \frac{1}{n} \end{cases} .$$

Pour $n \in \mathbb{N}^*$, représenter le graphe de la fonction f_n, puis montrer que (f_n) converge simplement vers une fonction f que l'on déterminera. La convergence est-ele uniforme ? f est-elle continue ?

On voit donc que la convergence simple ne préserve pas la continuité (ce qui est une bonne raison pour travailler avec des fonctions *mesurables* plutôt que des fonctions continues).

Proposition 3.20. *Soit (X, d) et (Y, D) deux espaces métriques, et (f_n) une suite de fonctions continues de X dans Y. Si (f_n) converge uniformément vers $f : X \to Y$ alors f est continue.*

Démonstration. Fixons $\varepsilon > 0$ et $x \in X$. Il existe $N \in \mathbb{N}$ tel que, pour tout $n \geq N$, on ait $D(f_n(x), f(x)) \leq \varepsilon$ pour tout $x \in X$.

Fixons un tel N ; comme f_N est continue, il existe $\delta > 0$ tel que pour tout $x' \in X$,

$$d(x, x') < \delta \Rightarrow D(f_N(x), f_N(x')) < \varepsilon.$$

Alors on a, pour tout $x' \in X$ tel que $d(x, x') < \delta$:

$$\begin{aligned} D(f(x), f(x')) &\leq D(f(x), f_N(x)) + D(f_N(x), f_N(x')) + D(f_N(x'), f(x')) \\ &\leq \varepsilon + \varepsilon + \varepsilon \\ &= 3\varepsilon . \end{aligned}$$

Comme ε était quelconque, ceci suffit à démontrer que f est continue en x. □ □

Propriété 3.21. Une limite uniforme de fonctions uniformément continues est uniformément continue. (Preuve en exercice.)

Chapitre 4

Compacité

Définition 4.1. Soit X un espace topologique et $Y \subset X$. *Un recouvrement de* Y *par des ouverts* est une famille d'ouverts $(O_i)_{i\in I}$ telle que $Y \subset \underset{i\in I}{\cup} O_i$. Si $J \subset I$ et $Y \subset \underset{i\in J}{\cup} O_i$ alors $(O_i)_{i\in J}$ est *un sous-recouvrement de* Y. Le sous-recouvrement $(O_i)_{i\in J}$ est fini si J est fini.

Définition 4.2. Soit X un espace topologique. On dit que X est *compact* si chaque recouvrement de X par des ouverts contient un sous-recouvrement fini.

Le théorème suivant est fondamental.

Théorème 4.3. *Soit* (X, d) *un espace métrique. Alors* (X, d) *est compact si et seulement si pour toute suite* (x_n) *d'éléments de* (X, d) *il existe une sous-suite* (x_{n_i}) *qui converge dans* (X, d).

Démonstration. (*facultative*) $\Rightarrow$) Soit $(x_n) \in X$ une suite. On peut supposer que $x_n \neq x_m$ si $n \neq m$. Supposons par l'absurd que (x_n) ne contient pas une sous-suite convergente. Alors pour tout $x \in X$ il existe un ouvert U_x qui contient x tel que le nombre de x_n dans U_x est fini. La famille $U_x, x \in X$, et un recouvrement de X. Par l'hypothèse, il existe un sous-recouvrement fini $U_1, \cdots, U_m$ de X. Donc, la suite (x_n) est finie. Absurd.

$\Leftarrow$) *Pas* 1 Soit $\varepsilon > 0$. Montrons qu'il existe un recouvrement fini $X = \overset{k}{\underset{i=1}{\cup}} B(x_i, \varepsilon)$. Supposons par l'absurd qu'un tel recouvrement n'existe pas. Soit $a_1 \in X$. Alors $B(a_1, \varepsilon) \neq X$. Si $a_2 \in B(a_1, \varepsilon)^c$ alors $B(a_1, \varepsilon) \cup B(a_2, \varepsilon) \neq X$ et $d(a_1, a_2) \geq \varepsilon$. Supposons par récurrence que $B(a_1, \varepsilon) \cup \cdots \cup B(a_m, \varepsilon) \neq X$ et $d(a_i, a_j) \geq \varepsilon$ si $i \neq j$. Soit $a_{m+1} \in (\overset{m}{\underset{i=1}{\cup}} B(a_i, \varepsilon))^c$. Alors $\overset{m+1}{\underset{i=1}{\cup}} B(a_i, \varepsilon) \neq X$ et $d(a_i, a_j) \geq \varepsilon$ si $i \neq j$. Par l'hypothèse, la suite (a_n) contient une sous-suite (a_{l_i}) qui converge vers a. Il existe i_0 tel que $d(a_{l_i}, a) < \frac{\varepsilon}{3}$ si $i > i_0$. Soit $i_0 < i < j$. Alors

$$d(a_{l_i}, a_{l_j}) \leq d(a_{l_i}, a_{l_{i_0}}) + d(a_{l_{i_0}}, a_{l_j}) < \frac{2\varepsilon}{3} < \varepsilon.$$

Absurd.

Pas 2 Montrons que X admet un recouvrement dénombrable par des ouverts $U_i, i \in \mathbb{N}^*$, tel que pour chaque $x \in X$ et chaque ouvert V qui contient x (c.à.d. pour chaque *voisinage ouvert* V de x) il existe i tel que $x \in U_i \subset V$.

Par Pas 1, pour chaque $n \in \mathbb{N}^*$ il existe un nombre fini d'éléments $a_{n,1}, \cdots, a_{n,s_n} \in X$ tels que $\overset{s_n}{\underset{i=1}{\cup}} B(a_{n,i}, \frac{1}{n}) = X$. Soit V un voisinage de $x \in X$. Il existe $n \in \mathbb{N}^*$ tel que $B(x, \frac{1}{n}) \subset V$. Soit $x \in B(a_{3n,i}, \frac{1}{3n})$. Si $y \in B(a_{3n,i}, \frac{1}{3n})$ alors

$$d(x, y) \leq d(x, a_{3n,i}) + d(a_{3n,i}, y) \leq \frac{1}{3n} + \frac{1}{3n} < \frac{1}{n}.$$

Donc, $B(a_{3n,i}, \frac{1}{3n}) \subset B(x, \frac{1}{n}) \subset V$.

Pas 3 Soit $\{V_j : j \in J\}$ un recouvrement de X par des ouverts. Montrons l'existence d'un sous-recouvrement fini de ce recouvrement.

Par Pas 2, on peut supposer que $\{V_j : j \in J\}$ est dénombrable et, donc, $J = \mathbb{N}$. Supposons par l'absurd qu'un sous-recouvrement fini n'existe pas. Alors pour chaque $n \in \mathbb{N}$ il existe $x_n \notin \bigcup_{i=0}^{n} V_i$. Par l'hypothèse, la suite (x_n) contient une sous-suite (x_{n_i}) qui convege vers un x. Soit $x \in V_{j_0}$. Puisque V_{j_0} est un voisinage de x, il existe un constant $C > 0$ tel que $x_{n_i} \in V_{j_0}$ si $n_i > C$. Mais $x_{n_i} \notin V_{j_0}$ si $n_i > j_0$. Absurd. □

Un exemple fondamental d'espace compact est donné par un intervalle fermé borné (un *segment*) de $\mathbb{R}$ ou, plus généralement n'importe quelle partie fermée bornée de $\mathbb{R}$. On verra plus loin (les Théorèmes 4.11 et 4.14) qu'en fait les compacts de $\mathbb{R}^n$ sont exactement ses parties fermées bornées.

Théorème 4.4. *Toute partie fermée bornée de $\mathbb{R}$, en particulier tout segment, est compacte.*

Démonstration. Considérons une suite (x_n) d'éléments d'une partie F fermée bornée de $\mathbb{R}$. Par Proposition 2.25, il existe une suite extraite (x_{n_k}) qui est monotone. Comme cette sous-suite est bornée et monotone, elle converge. Sa limite est dans F car F est fermé. □

Proposition 4.5. *Soit (X, d), (Y, D) deux espaces métriques compacts. Alors $X \times Y$, muni de la distance produit, est encore un espace métrique compact.*

Démonstration. Soit (x_n, y_n) une suite d'éléments de $X \times Y$. Par compacité de (X, d), on peut extraire une sous-suite $(x_{\varphi_1(k)})$ qui converge vers $x \in X$. ($\varphi_1 : \mathbb{N} \to \mathbb{N}$ désigne une fonction strictement croissante.) Et par compacité de (Y, D), on peut extraire de $(y_{\varphi_1(k)})$ une nouvelle sous-suite $y_{\varphi_1(\varphi_2(l))}$ qui converge vers $y \in Y$.

Notons $\psi = \varphi_1 \circ \varphi_2$. Alors la suite $(x_{\psi(l)}, y_{\psi(l)})$ est telle que $x_{\psi(l)}$ converge vers x et $y_{\psi(l)}$ converge vers y, autrement dit, cette suite est une suite extraite de (x_n, y_n) qui converge vers (x, y). □

Corollaire 4.6. Pour tout $n \in \mathbb{N}^*$, et tout $a_1, \ldots, a_n, b_1, \ldots, b_n$, l'ensemble $\prod [a_i, b_i]$ est un compact de $\mathbb{R}^n$ muni de la distance induite par $\| \cdot \|_\infty$.

Démonstration. Chacun des espaces $[a_i, b_i]$ est compact, et par récurrence on montre facilement à partir de la proposition précédente qu'un produit fini d'espaces métriques compacts est compact. □

Théorème 4.7. *Soit (X, d) un espace métrique, et A un sous-ensemble compact de X (c'est-à-dire que (A, d) est un espace métrique compact). Alors A est fermé dans X.*

Démonstration. Soit (x_n) une suite d'éléments de A qui converge vers $x \in X$; on doit prouver que $x \in A$. Comme A est compact, il existe une sous-suite (x_{n_k}) qui converge vers $a \in A$; et (x_{n_k}) converge toujours vers x. Par unicité de la limite, $x = a \in A$. □

Réciproquement, on a le résultat suivant.

Proposition 4.8. *Soit (X, d) un espace métrique compact et A une partie fermée de X. Alors (A, d) est un espace métrique compact.*

Première preuve. Soit (a_n) une suite d'éléments de A. Comme X est compact et (a_n) est aussi une suite d'éléments de X, il existe une sous-suite (a_{n_k}) qui converge vers $x \in X$; comme A est fermé, $x \in A$, ce qui montre que (a_n) a une sous-suite convergente dans A : (A, d) est compact. □

Deuxième preuve. Soit $A = \bigcup_{i \in I} O_i$ un recouvrement ouvert de A. Alors $O_i = \widetilde{O}_i \cap A$ où $\widetilde{O}_i$ est ouvert dans X. Donc, $X = A^c \cup (\bigcup_{i \in I} \widetilde{O}_i)$ et il existe un $J \subset I$ fini tel que $X = A^c \cup (\bigcup_{i \in J} \widetilde{O}_i)$. D'où, $A = \underset{i \in L}{O_i}$, c.à.d., A est compact. □

La compacité est importante pour nous en particulier parce que les fonctions continues sur les espaces compacts ont des propriétés très fortes.

Théorème 4.9. *Soit (X, d) un espace métrique compact, et $f: X \to \mathbb{R}$ une fonction continue. Alors f est bornée et atteint ses bornes.*

Démonstration. Montrons que f atteint sa borne supérieure $M = \sup(\{f(x): x \in X\})$. (L'argument pour la borne inférieure est symétrique, ou découle du résultat pour la borne supérieure appliqué à $-f$.) Par définition d'une borne supérieure, il existe une suite (x_n) d'éléments de X telle que $f(x_n)$ converge vers M. Comme (X, d) est compact, (x_n) admet une sous-suite convergente (x_{n_k}) ; appelons x sa limite. Alors $f(x_{n_k})$ converge à la fois vers M (c'est une sous-suite d'une suite qui converge vers M) et vers $f(x)$ (par continuité de f). Par conséquent, $f(x) = M$. $\square$

La proposition suivante est une conséquence facile de ce théorème.

Proposition 4.10. *Tout espace métrique compact (X, d) est* borné*, c'est-à-dire qu'il existe M tel que pour tout $x, x' \in X$ on ait $d(x, x') \leq M$.*

Première preuve. La fonction $d: X \times X \to \mathbb{R}$ est continue lorsqu'on munit $X \times X$ de la distance produit D ; en effet, elle est lipschitzienne :

$$\begin{aligned}
|d(x_1, x_2) - d(x_1', x_2')| &= |d(x_1, x_2) - d(x_1, x_2') + d(x_1, x_2') - d(x_1', x_2')| \\
&\leq |d(x_1, x_2) - d(x_1, x_2')| + |d(x_1, x_2') - d(x_1', x_2')| \\
&\leq d(x_2, x_2') + d(x_1, x_1') \leq \max\{d(x_2, x_2'), d(x_1, x_1')\} \\
&\leq 2D((x_1, x_2), (x_1', x_2')).
\end{aligned}$$

Comme $(X \times X, D)$ est compact, la proposition précédente nous permet de conclure que d est bornée, ce qui revient à dire que (X, d) est un espace métrique borné. $\square$

Deuxième preuve. Soit $x_0 \in X$. Il suffit de montrer qu'il existe $C > 0$ tel que $d(x, x_0) \leq C$ pour tout x. En effet, dans ce cas $d(x_1, x_2) \leq d(x_1, x_0) + d(x_0, x_2) \leq 2C$ qui montre la proposition. Supposons par l'absurd que pour chaque $n \in \mathbb{N}^*$ il existe x_n tel que $d(x_0, x_n) > n$. Alors il existe une sous-suite $x_{n_k} \to x'$. Mais $d(x_0, x_{n_k}) \leq d(x_0, x') + d(x', x_{n_k})$ et la partie à droite converge vers $d(x_0, x')$. Absurd. $\square$

Théorème 4.11. *Dans $\mathbb{R}^n$ muni de la distance d_∞ induite par $\|\cdot\|_\infty$, les compacts sont les fermés bornés.*

Démonstration. Si $A \subseteq \mathbb{R}^n$ est tel que (A, d) soit compact, alors on sait que (A, d) doit être fermé dans $\mathbb{R}^n$, et borné d'après la proposition précédente.

Réciproquement, si A est fermé borné dans $\mathbb{R}^n$, alors il existe M tel que A soit contenu dans $[-M, M]^n$; on a vu que cet ensemble est compact, et A y est fermé, donc (A, d) est compact. $\square$

Théorème 4.12. *Soit (X, d) un espace métrique compact, (Y, D) un espace métrique, et $f: X \to Y$ une fonction continue. Alors $f(X)$ est un sous-ensemble compact de Y.*

Notons que, une fois qu'on sait que les sous-ensembles compacts de $\mathbb{R}$ sont les fermés bornés, ce résultat généralise le théorème 4.9.

Première preuve. Soit (y_n) une suite d'éléments de $f(X)$. Pour tout n on peut choisir x_n tel que $f(x_n) = y_n$, et ensuite on peut extraire une sous-suite convergente (x_{n_k}) de la suite (x_n). Par continuité de f, la suite $y_{n_k} = f(x_{n_k})$ converge vers $y = f(x)$. $\square$

Deuxième preuve. Soit $f(X) \subset \bigcup_{i \in I} O_i$ un recouvrement ouvert de $f(X)$. Alors $X = \bigcup_{i \in I} f^{-1}(O_i)$ est un recouverement ouvert de X. Par Théorème 4.3, il existe $J \subset I$ fini tel que $X = \bigcup_{i \in J} f^{-1}(O_i)$. D'ou, $f(X) \subset \bigcup_{i \in J} O_i$ et $f(X)$ est compact (Théorème 4.3). $\square$

Théorème 4.13 (Théorème de Heine). *Soit (X,d) un espace métrique compact, (Y,D) un espace métrique, et $f\colon X\to Y$ une fonction continue. Alors f est uniformément continue.*

Première preuve. On va montrer la contraposée : si f n'est pas uniformément continue, alors f n'est pas continue. Supposons donc que f ne soit pas uniformément continue, c'est-à-dire qu'il est faux que

$$\forall\varepsilon>0\ \exists\delta>0\ \forall x,x'\in X\ d(x,x')<\delta\Rightarrow D(f(x),f(x'))<\varepsilon\ .$$

Autrement dit, notre hypothèse est que

$$\exists\varepsilon>0\ \forall\delta>0\ \exists x,x'\in X\ d(x,x')<\delta \text{ et } D(f(x),f(x'))\geq\varepsilon\ .$$

Fixons $\varepsilon>0$ comme ci-dessus. Pour tout $n\in\mathbb{N}^*$, on sait qu'il existe x_n,x'_n tels que $d(x_n,x'_n)<\frac{1}{n}$ et $D(f(x_n),f(x'_n))\geq\varepsilon$.

Comme l'espace produit $X\times X$ est compact, on peut extraire une sous-suite convergente (x_{n_k},x'_{n_k}) de la suite (x_n,x'_n) ; la suite x_{n_k} converge vers $x\in X$, et la suite x'_{n_k} converge vers $x'\in X$.

Puisque $n_k\to+\infty$ et $d(x_{n_k},x'_{n_k})\leq\frac{1}{n_k}$, on doit avoir $x=x'$. Mais on a aussi $D(f(x_{n_k}),f(x'_{n_k}))\geq\varepsilon$ pour tout k : il est impossible que les deux suites $f(x_{n_k})$ et $f(x'_{n_k})$ convergent vers $f(x)$, par conséquent f n'est pas continue en x. □

Deuxième preuve. On fixe $\varepsilon>0$. Pour chaque $x\in X$ il existe $\delta(x)>0$ tel que $f(x')\in B(f(x),\frac{\varepsilon}{2})$ si $x'\in B(x,\delta(x))$. Puisque $X=\bigcup\limits_{x\in X}B(x,\frac{\delta(x)}{2})$ il existe un sous-recouvrement fini $X=\bigcup\limits_{i=1}^{m}B(x_i,\frac{\delta(x_i)}{2})$.

Soit $\delta=\min\limits_i\delta(x_i)$ et soit $d(x,x')<\frac{\delta}{2}$. Il existe i tel que $x\in B(x_i,\frac{\delta(x_i)}{2})$. Alors

$$d(x_i,x')\leq d(x_i,x)+d(x,x')<\frac{\delta(x_i)}{2}+\frac{\delta}{2}\leq\delta(x_i).$$

D'où,

$$D(f(x),f(x'))\leq D(f(x),f(x_i))+D(f(x_i),f(x'))<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

□

Le cours du 29/09/2021

Le théorème suivant est fondamental. Il est faux dans le cas des e.v.n. de dimension infinie. (Voir, par exemple, l'Exercice 8 de la Feuille II.)

Théorème 4.14. *Toutes les normes sur $\mathbb{R}^n$ sont équivalentes.*

Démonstration. Soit N une norme sur $\mathbb{R}^n$. On va montrer que N est équivalente à $\|\cdot\|_\infty$ ($\equiv$ la distance produit) ; alors toutes les normes sont équivalentes à $\|\cdot\|_\infty$, donc elles sont toutes équivalentes.

Notons $e_1,\dots,e_n$ la base canonique de $\mathbb{R}^n$, et $M=\max(\{N(e_i)\colon 1\leq i\leq n\})$. Alors on a, pour $x=(x_1,\dots,x_n)$:

$$\begin{aligned}N(x) &= N(x_1e_1+\ldots+x_ne_n)\\ &\leq |x_1|N(e_1)+\ldots+|x_n|N(e_n)\\ &\leq M(|x_1|+\ldots+|x_n|)\\ &\leq M.n.\|x\|_\infty\ .\end{aligned}$$

Cela nous donne une des deux inégalités nécessaires pour montrer que N est équivalente à $\|\cdot\|_\infty$; cette inégalité implique aussi que $N\colon (\mathbb{R}^n, d_\infty) \to \mathbb{R}$ est lipschitzienne :

$$\forall x, x' \in \mathbb{R}^n \quad |N(x) - N(x')| \leq N(x - x') \leq Mn\|x - x'\|_\infty \ .$$

Donc, l'application $N\colon (\mathbb{R}^n, d_\infty) \to \mathbb{R}$ est continue par rapport à la topologie sur $\mathbb{R}^n$ induite par $\|\cdot\|_\infty$.

On a vu que les fermés bornés de $(\mathbb{R}^n, d_\infty)$ sont compacts ; par conséquent, l'ensemble

$$S = \{x \in [-1,1]^n \colon \exists k,\ x_k = \pm 1\} = \{x \in \mathbb{R}^n : \|x\|_\infty = 1\}$$

(la sphère unité pour $\|\cdot\|_\infty$) est compact, et comme N y est continue, elle y atteint son minimum m ; notons que, comme $0 \notin S$, $m > 0$. Soit maintenant $x \in \mathbb{R}^n$ un vecteur non nul ; le vecteur $\frac{x}{\|x\|_\infty}$ appartient à S, et on a donc $N(\frac{x}{\|x\|_\infty}) \geq m$.

Autrement dit, pour tout $x \in \mathbb{R}^n$ on a $m\|x\|_\infty \leq N(x)$ (cette inégalité est bien sûr aussi vraie pour $x = 0$), et on a fini de prouver que N et $\|\cdot\|_\infty$ sont équivalentes. □

Corollaire 4.15.
1. La topologie sur $\mathbb{R}^n$, $n \in \mathbb{N}^*$, induite par une norme sur $\mathbb{R}^n$ coincide avec la topologie standarde. En particulier, les compacts dans $\mathbb{R}^n$ sont les fermés bornés.
2. Chaque forme linéaire $\phi : \mathbb{R}^n \to \mathbb{R}$ est continue pour la topologie standarde sur $\mathbb{R}^n$.

Démonstration. La première partie est cas particulier de la Proposition 2.36 et du Théorème 4.11.

Soit ϕ une forme linéaire non-zéro sur $\mathbb{R}^n$. Soit $H = \ker(\phi)$. Il existe une base $v_1, \cdots, v_n$ de $\mathbb{R}^n$ telle que H est engendré par $v_1, \cdots, v_{n-1}$ et $\phi(v_n) = 1$. Alors $\phi(\sum\limits_i x_i v_i) = x_n$. On choisit la norme $\|\sum\limits_i x_i v_i\|_\infty = \max\limits_i\{|x_i|\}$. On voit facilement que ϕ est continue par rapport à $\|\cdot\|_\infty$. Donc, ϕ est aussi continue par rapport à la topologie standarde. □

Le Corollaire 4.15 n'est plus valable en dimension infinie.

Exercice 4.16. Soit $l^1(\mathbb{N})$ l'ensemble de suites réelles $(x_i) = (x_1, x_2, \cdots, x_n, \cdots)$ telles que $\sum_i |x_i| < \infty$. Si $x = (x_i) \in l^1(\mathbb{N})$ on pose $\|x\| = \sum_i |x_i|$.

1. Montrer que $l^1(\mathbb{N})$ est un espace vectoriel normé de dimension infinie.
2. Montrer que la boule fermée $\overline{B}(0,1)$ de $l^1(\mathbb{N})$ est bornée sans d'être compacte.
3. Soit V le sous-espace de $l^1(\mathbb{N})$ engendré par les $e_n = (0, \cdots, \underset{n}{1}, 0, 0, \cdots)$, $n \in \mathbb{N}$. Montrer que $V \subsetneq l^1(\mathbb{N})$ et $\overline{V} = l^1(\mathbb{N})$. En déduire l'existance de formes linéaires non-continues sur $l^1(\mathbb{N})$.

Exercice 4.17. Soit (X, d) un espace métrique, et (x_n) une suite d'éléments de X qui converge vers $x \in X$. Montrer que l'ensemble $\{x_n \colon n \in \mathbb{N}\} \cup \{x\}$ est compact.

Définition 4.18. Soit (X, d), (Y, D) deux espaces métriques. Une fonction $f\colon X \to Y$ est un *homéomorphisme* si f est une bijection telle que les fonctions f et f^{-1} soient continues.

Proposition 4.19. *Soit (X, d), (Y, D) deux espaces métriques compacts, et $f\colon X \to Y$ une bijection continue. Alors f^{-1} est nécessairement continue, autrement dit, f est un homéomorphisme.*

Démonstration. On doit montrer que la fonction $g = f^{-1}$ est continue. Soit F un fermé de X ; il nous suffit de montrer que $g^{-1}(F)$ est fermé dans Y. Pour cela, notons que

$$g^{-1}(F) = \{y \in Y \colon g(y) \in F\} = \{y \in Y \colon f^{-1}(y) \in F\} = f(F) \ .$$

D'après le théorème 4.12, $f(F)$ est compact ; comme les compacts sont fermés, on en déduit bien que $g^{-1}(F) = f(F)$ est fermé, ce qu'il fallait démontrer. □

Chapitre 5

Espaces mesurables

5.1 Clans, tribus, classes monotonnes, mesures

Rappel 5.1.
1. $\mathcal{P}(\Omega)$ désigne l'ensemble de toutes les parties d'un ensemble Ω.
2. $A\Delta B = (A \setminus B) \cup (B \setminus A)$ est la difference symétrique de $A, B \in \Omega$.
3. $A^c = \Omega \setminus A$.

Définition 5.2. Un sous-ensemble $\mathcal{C}$ de $\mathcal{P}(\Omega)$ est un *clan* (ou une *algèbre* ou une *algèbre de Boole*) sur Ω si :
1. $\Omega \in \mathcal{C}$,
2. (stabilité par complémentaire) $A \in \mathcal{C}$ entraîne que $A^c \in \mathcal{C}$,
3. (stabilité par réunion finie) si $A_1, \cdots A_n \in \mathcal{C}$ alors $\cup_{n=1}^N A_n \in \mathcal{C}$.

Exercice 5.3.
1. L'ensemble $\mathfrak{C}_1$ des unions finies d'intervales de $\mathbb{R}$ est un clan.
2. Rappelons qu'un *pavé connexe* de $\mathbb{R}^n$ est un ensemble de la forme $I_1 \times \cdots \times I_n$ avec I_k intervales de $\mathbb{R}$. L'ensemble $\mathfrak{C}_n$ des unions finies de pavés connexes de $\mathbb{R}^n$ est un clan.
3. Tout élément de $\mathfrak{C}_n$ est union finie de pavés connexes de $\mathbb{R}^n$ deux à deux disjoints.

Exercice 5.4. Soit $\mathcal{C}$ un clan.
1. Si $A_1, \cdots A_n \in \mathcal{C}$ alors $\cap_i A_i \in \mathcal{C}$.
2. Si $A, B \in \mathcal{C}$ alors $A \setminus B \in \mathcal{C}$.

Pour faire de l'analyse on a besoin aussi d'unions dénombrables, d'où la définition que l'on utilisera le plus souvent.

Définition 5.5. Un sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$ est une *tribu* (ou une σ-algèbre) sur Ω si :
1. $\Omega \in \mathcal{T}$,
2. (stabilité par complémentaire) Si $A \in \mathcal{T}$ alors $A^c \in \mathcal{T}$.
3. (stabilité par union dénombrable) Pour toute suite $(A_n)_{n\in\mathbb{N}}$ de parties de $\mathcal{T}$, alors l'union
$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}.$$

Le couple $(\Omega, \mathcal{T})$ sera appelé un *espace mesurable*. Les éléments de $\mathcal{T}$ sont appelés *ensembles mesurables*.

Remarque 5.6. Chaque tribu est un clan.

Exercice 5.7. Est-ce que le clan de l'exercice 5.3(1) est une tribu ?

Exemple. 1. $\mathcal{P}(\Omega)$ est une tribu ;

2. Le sous-ensemble $\{\emptyset, \Omega\}$ de $\mathcal{P}(\Omega)$ est un clan et une tribu, appelée *clan ou tribu triviale.*
3. $\Omega = \{0,1,2\}$, $\mathcal{C}_1 = \{\emptyset, \{0,1,2\}, \{0\}, \{1,2\}\}$ et $\mathcal{C}_2 = \{\emptyset, \{0,1,2\}, \{1\}, \{0,2\}\}$. Alors $\mathcal{C}_1$ et $\mathcal{C}_2$ sont des algèbres mais $\mathcal{C}_1 \cup \mathcal{C}_2$ n'est pas une algèbre (car $\{0,1\} \notin \mathcal{C}_1 \cup \mathcal{C}_2$).

Exercice 5.8. 1. Soit Ω un ensemble infini et $\mathcal{C}$ la famille de sous-ensembles de Ω finis ou de compléments finis. Alors $\mathcal{C}$ est un clan qui n'est pas une tribu ;

2. Soit Ω un espace topologique et $\mathcal{C}$ l'ensemble des ouverts (resp. des fermés) de Ω. Est-ce que $\mathcal{C}$ un clan ? une tribu ?

Exercice 5.9. Soit $\mathcal{T}$ une tribu et (A_n) une suite dans $\mathcal{T}$. Alors $\cap_i A_i \in \mathcal{T}$.

Rappel 5.10. 1. Une suite (A_n) de parties de Ω est

(a) croissante (dans $\mathcal{P}(\Omega)$) si $A_n \subset A_{n+1}$ pour tout n ;
(b) décroissante (dans $\mathcal{P}(\Omega)$) si $A_n \supset A_{n+1}$ pour tout n ;
(c) la suite (A_n) est une suite d.d.d. (acronyme pour "deux à deux disjointes") si $A_n \cap A_m = \emptyset$ pour $n \neq m$.

2. $A_n \nearrow A$ signifie que la suite (A_n) est croissante et $A = \cup A_n$;
3. $A_n \searrow A$ signifie que la suite (A_n) est décroissante et $A = \cap A_n$;
4. si $(B_i)_{i \in I}$ est une famille d'ensembles d.d.d. on écrit $\sqcup_{i \in I} B_i$ au lieu de $\cup_{i \in I} B_i$.

Définition 5.11. Une famille $\mathcal{M}$ dans $\mathcal{P}(\Omega)$ est une *classe monotonne* sur Ω si :

1. $\Omega \in \mathcal{M}$,
2. si $A, B \in \mathcal{M}$ et $B \subset A$ alors $A \setminus B \in \mathcal{M}$,
3. si (A_n) est une suite *croissante* alors $\cup_n A_n \in \mathcal{M}$.

Exemple 6. 1. Chaque tribu est une classe monotonne.

2. Soit $\mathcal{A}$ une tribu (resp. un clan, resp. une classe monotonne) sur Ω et soit $X \subset \Omega$. Alors $\mathcal{A}_X \stackrel{def}{=} \{A \cap X : A \in \mathcal{A}\}$ est une tribu (resp.un clan, resp. une classe monotonne) sur X appelé la tribu induite (resp. le clan induit, resp. la classe monotonne induite) par $\mathcal{A}$ sur X.
3. Si $X \in \mathcal{A}$ alors $\mathcal{A}_X = \{A \in \mathcal{A} : A \subset X\}$.

Définition 5.12. Soit $(\Omega, \mathcal{T})$ un espace mesurable. Une application $\mu : \mathcal{T} \to \overline{\mathbb{R}}^+$ est une *mesure* (positive) si les conditions suivantes sont satisfaites :

1. $\mu(\emptyset) = 0$.
2. (σ-additivité) Si $(A_n)_{n \geq 1} \subset \mathcal{T}$ est une suite d.d.d. alors

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

Un triplet $(\Omega, \mathcal{T}, \mu)$ où μ est une mesure sur un espace mesurable $(\Omega, \mathcal{T})$ est appelé *espace mesuré.* Si $\mu(\Omega) < \infty$ on dit que μ est *une mesure finie* et si $\mu(\Omega) = 1$ on dit que μ est *une mesure de probabilité* ou simplement une *probabilité.* La mesure μ est *σ-finie* s'il existe une famille dénombrable $A_n \in \mathcal{T}, n \in \mathbb{N}$, telle que $\Omega = \bigcup\limits_{n \in \mathbb{N}} A_n$ et $\mu(A_n) < \infty$ pour tout n.

5.2 Quelques propriétés de base et exemples

Proposition 5.13. *Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré. On suppose que A, B et $A_n, n \in \mathbb{N}$, sont des ensembles mesurables.*

1. *Si $A \subset B$ alors $\mu(B) = \mu(A) + \mu(B \setminus A)$. En particulier, si $A \subset B$ alors $\mu(A) \leq \mu(B)$.*
2. *Si $A \cap B = \emptyset$, alors $\mu(A \cup B) = \mu(A) + \mu(B)$. Dans le cas général, si $\mu(A \cap B) < \infty$ alors*
$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B).$$
3. *Si (A_n) est une suite croissante, alors*
$$\mu(\underset{n\geq 1}{\cup} A_n) = \lim_{n\to\infty} \mu(A_n) = \sup_{n\geq 1} \mu(A_n).$$
4. *Pour une suite (A_n) d'ensembles non nécessairement disjoints :*
$$\mu\Big(\underset{n\geq 1}{\cup} A_n\Big) \leq \sum_{n\geq 1} \mu(A_n).$$
5. *Si (A_n) est une suite décroissante et s'il existe n_0 tel que $\mu(A_{n_0}) < \infty$, alors*
$$\mu(\underset{n\geq 1}{\cap} A_n) = \lim_{n\to\infty} \mu(A_n).$$

Démonstration. 1. Par la Définition 5.5(2), $A^c \in \mathcal{T}$ et par l'exercice 5.4 $B \setminus A = B \cap A^c \in \mathcal{T}$. Puisque $B = A \sqcup (B \setminus A)$, $\mu(B) = \mu(A) + \mu(B \setminus A)$.

2. Si $A \cap B = \emptyset$ on unilise la Définition 5.12(2) avec $A_1 = A$, $A_2 = B$ et $A_i = \emptyset, i > 2$. En général, $A \cup B = (A \setminus B) \sqcup (A \cap B) \sqcup (B \setminus A)$. D'où

$$A \cup B = (\mu(A \setminus B) + \mu(A \cap B)) + (\mu(B \setminus A) + \mu(A \cap B)) - \mu(A \cap B) = \mu(A) + \mu(B) - \mu(A \cap B).$$

3. Soit $A_1 = B_1$ et $B_{k+1} = A_{k+1} \setminus A_k, k \geq 1$. Alors $B_i \cap B_j = \emptyset$ si $i \neq j$. Comme

$$A_n = \underset{1\leq k\leq n}{\sqcup} B_k,$$

on obtient

$$\underset{n}{\cup} A_n = \underset{k}{\sqcup} B_k,$$

et par la σ-additivité

$$\mu(\underset{n}{\cup} A_n) = \sum_{k\geq 1} \mu(B_k) = \lim_{n\to\infty} \sum_{1\leq k\leq n} \mu(B_k) = \lim_{n\to\infty} \mu(A_n).$$

4. La suite $C_n \overset{def}{=} \underset{1\leq i\leq n}{\cup} A_i$ est croissante et $\underset{n}{\cup} A_n = \underset{n}{\cup} C_n$. En utilisant 1, on obtient par recurrence $\mu(C_n) \leq \sum_{1\leq i\leq n} \mu(A_i)$. Par 3,

$$\mu\Big(\underset{i\geq 1}{\cup} A_i\Big) = \lim_{n} \mu(C_n) \leq \sum_{i\geq 1} \mu(A_i).$$

5. On peut supposer que $\mu(A_1) < \infty$. Soit $B_i \overset{def}{=} A_1 \setminus A_i, i \geq 1$. Alors $B_i \nearrow A_1 \setminus \underset{i}{\cap} A_i$, $\underset{i}{\cup} B_i = A_1 \setminus \underset{i}{\cap} A_i$, et $A_1 = A_i \sqcup B_i$. Par 2 et 3,

$$\mu(\underset{i}{\cup} B_i) = \mu(A_1) - \mu(\underset{i}{\cap} A_i) = \lim_{i} \mu(B_i) = \lim_{i} (\mu(A_1) - \mu(A_i)) = \mu(A_1) - \lim_{i} \mu(A_i).$$

Donc,

$$\mu(\bigcap_i A_i) = \lim_i \mu(A_i).$$

□

Exemple 7. Si (μ_n) est une suite de mesures sur l'espace mesurable $(\Omega, \mathcal{T})$ alors $\sum\limits_n \mu_n$ est une mesure sur $(\Omega, \mathcal{T})$.

Exemple 8. (Jeu de dé) Soit $\Omega = \{1, 2, 3, 4, 5, 6\}$ les 6 faces possibles d'un dé et soit $\mathcal{T} = \mathcal{P}(\Omega)$. On définie la probabilité $\mu(A) \overset{def}{=} \text{card}(A)/6$. Alors $\mu(A)$ *représente la probabilité que* A *survienne* c.à.d. c'est le nombre de faces qui provoquent A divisé par le nombre total de faces du dé.

Exemple 9. (Mesure de Dirac) Soit $(\Omega, \mathcal{T})$ un espace mesurable *quelconque* et soit $x \in \Omega$. Pour tout $A \in \mathcal{T}$ on pose

$$\delta_x(A) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$$

Alors δ_x est une mesure (en effet, une probabilité). Pour montrer la σ-additivité, il suffit de noter que dans une union disjointe $x \in A_n$ pour un seul n, donc si $A = \bigsqcup_n A_n$, $\delta_x(A) = \delta_x(A_n) = \sum_k \delta_x(A_k)$.

Exemple 10. (Mesure de comptage) Soit $(\Omega, \mathcal{T})$ un espace mesurable. Pour chaque $A \in \mathcal{T}$ on pose

$$\nu(A) = \begin{cases} Card(A) & \text{si } A \text{ fini} \\ +\infty & \text{sinon} \end{cases}$$

Alors ν est une mesure. Pour montrer la σ-additivité, il suffit de observer que si les A_n sont disjoints alors $Card(\bigsqcup_n A_n) = \sum_n Card(A_n)$. *Notons que si* Ω *est dénombrable alors* $\nu = \sum\limits_{x \in \Omega} \delta_x$.

Exemple 11. (Mesure sur la tribu induite) Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré et $A \in \mathcal{T}$. Alors $\mu_A(B) \overset{def}{=} \mu(A \cap B)$ est une mesure sur $(\Omega, \mathcal{T})$. On peut bien sûr remplacer la tribu $\mathcal{T}$ par la tribu induite $\mathcal{T}_A \overset{def}{=} \{A \cap B : B \in \mathcal{T}\}$. Evidemment, $(\Omega, \mathcal{T}, \mu_A)$ et $(A, \mathcal{T}_A, \mu_A)$ sont des espaces mesurés.

Exemple 12. Si $(\Omega, \mathcal{T}, \mu)$ est un espace mesuré et $\lambda \in \mathbb{R}^+$ alors $(\lambda\mu)(A) \overset{def}{=} \lambda\mu(A), A \in \mathcal{T}$, définie une mesure sur $\mathcal{T}$. En particulier, si $0 < \mu(A) < \infty$ alors $\mu_A(\cdot)/\mu(A)$ est une probabilité (appelée *la probabilité conditionnelle sachant* A).

5.2.1 Tribus engendrées, tribu borélienne

Remarque 5.14. On sait rarement décrire explicitement tous les ensembles d'une tribu. On travaille donc avec un ensemble plus petit de "générateurs" de la tribu. Ceci est basé sur la proposition suivante.

Proposition 5.15. *Si* $(\mathcal{T}_i)_{i \in I}$ *est une famille de tribus (resp. clans, resp. classes monotones) sur* Ω*, alors l'intersection* $\cap_{i \in I} \mathcal{T}_i$ *est aussi une tribu (resp. clan, resp. classe monotonne).*

Si $\mathcal{E} \subset \mathcal{P}(\Omega)$*, alors l'intersection de toutes les tribus (resp. clans, resp. classes monotones) contenant* $\mathcal{E}$ *est une tribu (resp. clan, resp. classe monotone). C'est la plus petite tribu (resp. clan, resp. classe monotone) contenant* $\mathcal{E}$ *(pour l'inclusion), appelée tribu (resp. clan, resp. classe monotone) engendrée par* $\mathcal{E}$.

Notation. Si $\mathcal{E} \subset \mathcal{P}(\Omega)$ on désigne par $\mathcal{T}(\mathcal{E})$ (resp. $\mathcal{C}(\mathcal{E})$, resp. $\mathcal{M}(\mathcal{E})$) la tribu (resp. clan, resp. classe monotonne) engendrée par $\mathcal{E}$.

Démonstration. On fait la preuve pour les tribus ; preuve identique pour les clans et les classes monotones.

Soit $\{\mathcal{T}_i : i \in I\}$ des tribus sur Ω.

Comme $\Omega \in \mathcal{T}_i$ pour tout $i \in I$, on déduit que $\Omega \in \cap_{i\in I}\mathcal{T}_i$. Si $A \in \cap_{i\in I}\mathcal{T}_i$ alors $A \in \mathcal{T}_i$ pour tout $i \in I$ donc $A^c \in \mathcal{T}_i$ (car $\mathcal{T}_i$ est une tribu donc est stable par complémentaire) pour tout $i \in I$ donc $A^c \in \cap_{i\in I}\mathcal{T}_i$. Si $A_n \in \cap_{i\in I}\mathcal{T}_i$ pour tout $n \geq 0$ alors pour tout $i \in I$ et tout $n \geq 0$, $A_n \in \mathcal{T}_i$. Comme $\mathcal{T}_i$ est une tribu $\cup_{n\geq 0}A_n \in \mathcal{T}_i$ et ce pour tout i. Donc $\cup_{n\geq 0}A_n \in \cap_{i\in I}\mathcal{T}_i$. On en déduit la dernière stabilité requise, donc $\cap_{i\in I}\mathcal{T}_i$ est bien une tribu.

On vient de voir que comme intersection de tribus, $\mathcal{T}(\mathcal{E})$ est une tribu qui contient clairement $\mathcal{E}$. Réciproquement, une tribu $\mathcal{T}$ contenant $\mathcal{E}$ est dans l'indice de l'intersection, donc $\mathcal{T}(\mathcal{E}) \subset \mathcal{T}$, donc c'est bien la plus petite tribu contenant $\mathcal{E}$ (c'est-à-dire elle est contenue dans toute tribu contenant $\mathcal{E}$). □

Le cours du 6/10/2021

Exercice 5.16. Soient $(\Omega, \mathcal{T})$ un espace mesurable et $\mathcal{T} = \mathcal{T}(\mathcal{E})$. Alors $\mathcal{T} = \mathcal{T}(\mathcal{E}^c)$ où $\mathcal{E}^c \stackrel{def}{=} \{A^c : A \in \mathcal{E}\}$.

Exercice 5.17. Montrer que $\mathcal{T}(\{A\}) = \{\emptyset, A, A^c, \Omega\}$.

Exercice 5.18. Si Ω est dénombrable, montrer que $\mathcal{T}(\{\omega\} \in \mathcal{P}(\Omega)) = \mathcal{P}(\Omega)$.

Exercice 5.19. Soit $\mathcal{E} = (A_i)_{i\in I}$ une partition dénombrable de Ω (c'est à dire les A_i sont deux à deux disjoints et leur union est Ω). Montrer que $\mathcal{T}(\mathcal{E}) = \{\cup_{i\in J}A_i : J \subset I\}$.

Définition 5.20. Soient $(\Omega_i, \mathcal{T}_i), i = 1, 2$, deux espaces mesurables. On appelle *ensemble élémentaire* de $\Omega = \Omega_1 \times \Omega_2$ une réunion fini de *pavés* $A_1 \times A_2$ avec $A_i \in \mathcal{T}_i, i = 1, 2$. *La tribu produit* $\mathcal{T}_1 \otimes \mathcal{T}_2$ sur Ω est la tribu engendrée par les ensembles élémentaires (c.à.d. par les pavés).

Notation. Dons la formulation du théorème suivant on utilisera la notation $\mathcal{E}_1 \times \mathcal{E}_2 := \{A \times B : A \in \mathcal{E}_1, B \in \mathcal{E}_2\}$ où $\mathcal{E}_1 \subset \mathcal{P}(\Omega_1)$ et $\mathcal{E}_2 \subset \mathcal{P}(\Omega_2)$.

Théorème 5.21. *Soient* $\Omega_1 \in \mathcal{E}_1 \subset \mathcal{P}(\Omega_1)$ *et* $\Omega_2 \in \mathcal{E}_2 \subset \mathcal{P}(\Omega_2)$. *Alors* $\mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2) = \mathcal{T}(\mathcal{E}_1) \otimes \mathcal{T}(\mathcal{E}_2)$.

Démonstration. On a $\mathcal{E}_1 \times \mathcal{E}_2 \subset \mathcal{T}(\mathcal{E}_1) \otimes \mathcal{T}(\mathcal{E}_2)$. D'où $\mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2) \subset \mathcal{T}(\mathcal{E}_1) \otimes \mathcal{T}(\mathcal{E}_2)$.

Pour montrer que $\mathcal{T}(\mathcal{E}_1) \otimes \mathcal{T}(\mathcal{E}_2) \subset \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)$ il suffit de montrer que si $M_1 \in \mathcal{T}(\mathcal{E}_1)$ et $M_2 \in \mathcal{T}(\mathcal{E}_2)$ alors $M_1 \times M_2 \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)$.

Fixons $E_2 \in \mathcal{E}_2$. On considère

$$\mathcal{L}_1 = \{A : A \in \mathcal{P}(\Omega_1), A \times E_2 \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)\}.$$

Alors $\Omega_1 \in \mathcal{L}_1$. Si $A \in \mathcal{L}_1$ alors $A \times E_2 \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)$ et

$$A^c \times E_2 = (A \times E_2)^c \cap (\Omega_1 \times E_2) \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2),$$

i.e. $A^c \in \mathcal{L}_1$. En plus, si $A_1, A_2, \cdots \in \mathcal{L}_1$ alors

$$(\cup_i A_i) \times E_2 = \cup_i (A_i \times E_2) \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2).$$

Donc, $\cup_i A_i \in \mathcal{L}_1$ et, par conséquant, $\mathcal{L}_1$ est une tribu qui contient $\mathcal{E}_1$. Ainsi, nous avons démontré que pour tous $M_1 \in \mathcal{T}(\mathcal{E}_1)$ et $E_2 \in \mathcal{E}_2$ on a $M_1 \times E_2 \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)$.

Maintenant on fixe $M_1 \in \mathcal{T}(\mathcal{E}_1)$ et on considère

$$\mathcal{L}_2 = \{B : B \in \mathcal{P}(\Omega_2), M_1 \times B \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)\}.$$

De la même manière comme ci-dessus on montre que $\mathcal{L}_2$ est une tribu. En effet, on note d'abord que $\Omega_2 \in \mathcal{L}_2$. Si $B \in \mathcal{L}_2$ alors $M_1 \times B \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)$ et

$$M_1 \times B^c = (M_1 \times B)^c \cap (M_1 \times \Omega_2) \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2),$$

i.e. $B^c \in \mathcal{L}_2$. Si $B_1, B_2, \cdots \in \mathcal{L}_2$ alors

$$M_1 \times (\cup_i B_i) = \cup_i (M_1 \times \times B_i) \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2).$$

Donc, $\cup_i B_i \in \mathcal{L}_2$ et, par conséquant, $\mathcal{L}_2$ est une tribu qui contient $\mathcal{E}_2$. D'où, $\mathcal{T}(\mathcal{E}_2) \subset \mathcal{L}_2$. Nous avons démontré que pour tous $M_1 \in \mathcal{T}(\mathcal{E}_1)$ et $M_2 \in \mathcal{T}(\mathcal{E}_2)$ on a que $M_1 \times M_2 \in \mathcal{T}(\mathcal{E}_1 \times \mathcal{E}_2)$. □

Rappelons la définition suivante :

Définition 5.22. Un espace topologique est un couple $(E, \mathcal{O})$ (si E n'est pas implicite on utilisera aussi la notation $\mathcal{O}_E$ au lieu de $\mathcal{O}$) où $\mathcal{O} \subset \mathcal{P}(E)$ et, en plus,

1. $\emptyset$ et $E \in \mathcal{O}$,
2. Pour chaque famille (dénonbrable ou non-dénombrable) $(U_i)_{i \in I}, U_i \in \mathcal{O}$, on a $\bigcup\limits_{i \in I} U_i \in \mathcal{O}$,
3. Pour chaque famille finie $U_1, \cdots, U_n \in \mathcal{O}$ on a $\bigcap\limits_{1 \leq i \leq n} U_i \in \mathcal{O}$. (Les éléments de $\mathcal{O}$ sont les *ouverts* et les éléments de $\{U^c : U \in \mathcal{O}\}$ sont les *fermés* de l'espace topologique $(E, \mathcal{O})$.)

Remarque 5.23. Si $\mathcal{O}$ est implicite on écrit souvent E au lieu de $(E, \mathcal{O})$.

Exemple 13.

1. Si $\mathcal{O} = \mathcal{P}(E)$ alors $(E, \mathcal{O})$ est un espace topologique avec topologie discrète.
2. Si $E = \mathbb{R}^n$ ou $\mathbb{C}^n$ alors $\mathcal{O}$ consiste des ouverts dans le sens habituel.

Définition 5.24. Si $(E, \mathcal{O})$ est une espace topologique (par exemple, si $E = \mathbb{R}^n$), on appelle *tribu borélienne*, noté $\mathcal{B}(E)$, la tribu engendrée par les ouverts de E, c.à.d. $\mathcal{B}(E) = \mathcal{T}(\mathcal{O})$. *Un borélien* est un ensemble appartenant à $\mathcal{B}(E)$.

Définition 5.25.

1. Soit $(E, \mathcal{O})$ un espace topologique et $\mathcal{U} \subset \mathcal{O}$ une famille des ouverts. Alors $\mathcal{U}$ est *une base* de l'espace topologique $(E, \mathcal{O})$ si chaque ouvert est union d'éléments de $\mathcal{U}$.
2. Soient $(E_1, \mathcal{O}_1)$ et $(E_2, \mathcal{O}_2)$ deux espaces topologiques avec de bases $\mathcal{U}_1$ et $\mathcal{U}_2$, respectivement. Alors on a une structure d'espace topologique sur $E_1 \times E_2$ telle que $\mathcal{U} \overset{def}{=} \{A_1 \times A_2 : A_1 \in \mathcal{U}_1, A_2 \in \mathcal{U}_2\}$ est sa base, i.e. chaque ouvert de l'ensemble des ouverts $\mathcal{O}$ de cet espace topologique est une union (quelconque) des éléments de $\mathcal{U}$. L'espace topologique $(E_1 \times E_2, \mathcal{O})$ est *le produit des espaces topologiques* $(E_1, \mathcal{O}_1)$ et $(E_2, \mathcal{O}_2)$.

Exemple 14. L'espace topologique $\mathbb{R}^2$ est le produit de deux copies de l'espace topologique $\mathbb{R}$. Par recurrence, $\mathbb{R}^n, n \geq 3$, est le produit des espaces topologiques $\mathbb{R}^{n-1}$ et $\mathbb{R}$.

Proposition 5.26. *Si $\mathcal{U}$ est une base dénombrable de l'espace topologique E alors $\mathcal{B}(E) = \mathcal{T}(\mathcal{U})$.*

Démonstration. Il est evident que $\mathcal{B}(E) \supset \mathcal{T}(\mathcal{U})$. D'autre part, chaque ouvert de E est union dénombrable des ouverts de $\mathcal{U}$. Donc, $\mathcal{B}(E) \subset \mathcal{T}(\mathcal{U})$. □

Corollaire 5.27. Soient E_1 et E_2 des espaces topologiques qui admettent des bases dénombrables. Alors, $\mathcal{B}(E_1 \times E_2) = \mathcal{B}(E_1) \otimes \mathcal{B}(E_1))$.

Démonstration. Sient $\mathcal{U}_i, i = 1, 2$, des bases dénombrables de $E_i, i = 1, 2$, respectivement. Alors $\mathcal{U} = \{A \times B : A \in \mathcal{U}_1, B \in \mathcal{U}_2\}$ est une base dénombrable de $E_1 \times E_2$. On peut supposer sans restriction que $E_i \in \mathcal{U}_i, i = 1, 2$. Par Proposition 5.26 $\mathcal{B}(E_1) = \mathcal{T}(\mathcal{U}_1)$, $\mathcal{B}(E_2) = \mathcal{T}(\mathcal{U}_2)$ et $\mathcal{B}(E_1 \times E_2) = \mathcal{T}(\mathcal{U})$. Par Théorème 5.21 $\mathcal{B}(E_1 \times E_2) = \mathcal{B}(E_1) \otimes \mathcal{B}(E_1))$. □

Corollaire 5.28. La tribu $\mathcal{B}(\mathbb{R}^n)$ est engendrée par les pavés connexes fermés $\{[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] : a_i \le b_i\}$ et $\mathcal{B}(\mathbb{R}^n) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{n}$.

Démonstration. Notons que $\mathbb{R}^n$ est le produit de n copies de l'espace topologique $\mathbb{R}$ et les intervalles ouverts (a, b) où a et $b \in \mathbb{Q}$ constituent une base de $\mathbb{R}$. Par recurrence (en utilisant Corollaire 5.27), $\mathcal{B}(\mathbb{R}^n) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{n}$. Chaque pavé connexe fermé se représente comme intersection d'une suite de pavés connexes ouverts. Donc chaque pavé connexe fermé appartient à $\mathcal{B}(\mathbb{R}^n)$. D'autre part, chaque pavé connexe ouvert est union d'une suite de pavés connexes fermés. Mais les pavés connexes ouverts engendrent $\mathcal{B}(\mathbb{R}^n)$. D'où le résultat. □

Exercice 5.29.

$$\mathcal{B}(\mathbb{R}) = \mathcal{T}\Big(\big\{]-\infty, a], a \in \mathbb{R}\big\}\Big) = \mathcal{T}\Big(\big\{[a, +\infty[, a \in \mathbb{R}\big\}\Big).$$

5.3 Applications mesurables

Rappel 5.30. Si f est une application de Ω dans E et si $B \in \mathcal{P}(E)$ alors

$$f^{-1}(B) = \{\omega \in \Omega : f(\omega) \in B\} \in \mathcal{F}$$

est l'image réciproque de B.

Notation. Si $\mathcal{B}$ est une famille de parties de E, on notera

$$f^{-1}(\mathcal{B}) = \{f^{-1}(B) : B \in \mathcal{B}\}.$$

Les formules (1.3) impliqent :

Proposition 5.31. *Soit Ω un ensemble et $(E, \mathcal{B})$ un espace mesurable. Pour toute application $f : \Omega \to E$, l'ensemble $f^{-1}(\mathcal{B})$ est une tribu sur Ω.*

Définition 5.32.
1. Soient $(\Omega, \mathcal{T}), (E, \mathcal{B})$ des espaces mesurables. Une application $f : \Omega \to E$ est dite *mesurable* (pour $\mathcal{T}$ et $\mathcal{B}$) si $f^{-1}(\mathcal{B}) \subset \mathcal{T}$, c.à.d. $f^{-1}(B) \in \mathcal{T}$ pour tout $B \in \mathcal{B}$.
2. Une application mesurable de $(\Omega, \mathcal{T})$ dans un espace topologique E muni de sa tribu borélienne $\mathcal{B}$ est dite *borélienne.* Si $(E, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ on parle d'une fonction borélienne.
3. Si f est une application d'un ensemble Ω dans un espace mesurable $(E, \mathcal{B})$, on appelle *tribu engendrée par* f, notée $\mathcal{T}(f)$, la plus petite tribu (sur Ω) qui rend f mesurable ; autrement dit, $\mathcal{T}(f) = f^{-1}(\mathcal{B})$ (voir la Proposition 5.31).

Remarque 5.33. Etant donnée une application $f : (\Omega, \mathcal{T}) \to (E, \mathcal{B})$, dire que f est mesurable de $(\Omega, \mathcal{T})$ dans $(E, \mathcal{B})$ revient à dire que $\mathcal{T}(f) \subset \mathcal{T}$.

Exercice 5.34. Soit $(\Omega, \mathcal{T})$ un espace mesurable. Vérifier que $A \subset \Omega$ est mesurable si et seulement si sa fonction indicatrice 1_A est borélienne.

Définition 5.35. Soit $(\Omega, \mathcal{T})$ un espace mesurable et soit $f : \Omega \to \overline{\mathbb{R}}$. Alors f est dite fonction borélienne (à valeurs dans $\overline{\mathbb{R}}$) si $f^{-1}(\infty)$, $f^{-1}(-\infty)$ et $f^{-1}(\mathcal{B}), B \in \mathcal{B}(\mathbb{R})$, sont tous ensembles mesurables (i.e. appartiennent à $\mathcal{T}$).

5.3.1 Propriétés des applications mesurables

Théorème 5.36. *Soient Ω un ensemble, $(E, \mathcal{B})$ un espace mesurable et $\mathcal{E}$ une famille dans $\mathcal{B}$ telle que $\mathcal{B} = \mathcal{T}(\mathcal{E})$. Soit $f : \Omega \to E$. Alors $\mathcal{T}(f) = \mathcal{T}(f^{-1}(\mathcal{E}))$. En particulier, si $(\Omega, \mathcal{T})$ est un espace mesurable alors pour qu'une application f de $(\Omega, \mathcal{T})$ dans $(E, \mathcal{B})$ soit mesurable il faut et il suffit que $f^{-1}(\mathcal{E})$ soit inclus dans $\mathcal{T}$.*

Démonstration. Rappelons que $\mathcal{T}(f) = f^{-1}(\mathcal{B})$. Puisque $\mathcal{E} \subset \mathcal{B}$ on a $\mathcal{T}(f^{-1}(\mathcal{E})) \subset \mathcal{T}(f)$. Pour montrer l'inclusion réciproque on considère

$$\mathcal{A} = \{B \subset E : f^{-1}(B) \subset \mathcal{T}(f^{-1}(\mathcal{E}))\}.$$

En utilisant les formules de commutations (1.3) de l'image réciproque avec les réunions et compléments on voit que $\mathcal{A}$ est une tribu. Mais $\mathcal{E} \subset \mathcal{A}$. Donc, $\mathcal{T}(\mathcal{E}) = \mathcal{B} \subset \mathcal{A}$. Par la definition de $\mathcal{A}$, $f^{-1}(\mathcal{A}) \subset \mathcal{T}(f^{-1}(\mathcal{E}))$. D'où, $\mathcal{T}(f) = f^{-1}(\mathcal{B}) \subset f^{-1}(\mathcal{A}) \subset \mathcal{T}(f^{-1}(\mathcal{E}))$.

$\square$

Corollaire 5.37. Soient X et Y des espaces topologiques munis de leurs tribus boréliennes. Alors chaque application continue $F : X \to Y$ est borélienne. En particulier, chaque application continue $f : R^n \to R^m$ est borélienne.

Proposition 5.38. *Si $(E, \mathcal{E}), (F, \mathcal{F})$ et $(G, \mathcal{G})$ sont des espaces mesurables et $f : E \to F$ et $g : F \to G$ sont des applications mesurables alors la composée $g \circ f : E \to G$ est mesurable.*

Démonstration. Pour $A \subset G$, on remarque que $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$. Donc, si $A \in \mathcal{G}$, alors $g^{-1}(A) \in \mathcal{F}$ et $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A)) \in \mathcal{E}$. Ceci conclut la preuve de la mesurabilité de la composée. $\square$

Corollaire 5.39. Soient $(\Omega, \mathcal{T})$ un espace mesurable et $f_i : \Omega \to \mathbb{R}$, $i = 1, \cdots, n$. Alors l'application $(f_1, ..., f_n) : \Omega \to \mathbb{R}^n, a \mapsto (f_1(a), ..., f_n(a))$, est borélienne si et seulement si f_i est borélienne pour tout i.

Démonstration. Si $(f_1, ..., f_n) : \Omega \to \mathbb{R}^n$ borélienne, $f_i = p_i \circ (f_1, ..., f_n)$ avec $p_i(x_1, ..., x_n) = x_i$ la i-ème projection qui est continue. Donc la composée est mesurable.

Réciproquement, par Théorème 5.36 et la définition de la tribu $\mathcal{B}(\mathbb{R}^n)$ il suffit de vérifier que $(f_1, ..., f_n)^{-1}(\prod_{i=1}^n [a_i, b_i])$ est borélienne.

Mais on a :

$$(f_1, ..., f_n)^{-1}(\prod_{i=1}^{n} [a_i, b_i]) = \bigcap_{i=1}^{n} f_i^{-1}([a_i, b_i]).$$

Donc par intersection finie d'images inverses d'intervalles par des fonctions boréliennes, cet ensemble est bien mesurable.

$\square$

Corollaire 5.40. L'espace des fonctions boréliennes de $(\Omega, \mathcal{T})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est stable pour les operations de multiplication par une constante $(\alpha f)(\omega) = \alpha f(\omega)$, d'addition $(f + g)(\omega) = f(\omega) + g(\omega)$, de multiplication $(fg)(\omega) = f(\omega)g(\omega)$, du maximum $(f \vee g)(\omega) = f(\omega) \vee g(\omega)$ et du minimum $(f \wedge g)(\omega) = f(\omega) \wedge g(\omega)$.

Démonstration. La fonction $\omega \mapsto \alpha f(\omega)$ est la composition de la fonction mesurable f et de la fonction continue $x \mapsto \alpha x$. De même, $f + g$ (resp.fg, resp.$f \vee g$, resp.$f \wedge g$) est la composée de la fonction mesurable $\omega \mapsto (f(\omega), g(\omega))$ (en vertu de la Proposition 5.39) et de la fonction continue $(x, y) \mapsto x + y$ (resp.$(x, y) \mapsto xy$, resp.$(x, y) \mapsto x \vee y$, resp.$(x, y) \mapsto x \wedge y$).

$\square$

Exercice 5.41. Soit $(\Omega, \mathcal{T})$ un espace mesurable. $f : \Omega \mapsto \mathbb{R}$ est borélienne si et seulement si on a l'une des conditions équivalentes :

1. Pour tout $a \le b$, $f^{-1}([a,b]) \in \mathcal{T}$.
2. Pour tout $a \in \mathbb{R}$, $f^{-1}(\{y \le a\}) \in \mathcal{T}$.
3. Pour tout $a \in \mathbb{R}$, $f^{-1}(\{y \ge a\}) \in \mathcal{T}$.

5.3.2 Suites de fonctions mesurables

Théorème 5.42. *Soit (f_n) une suite d'applications mesurables d'un espace mesurable $(\Omega, \mathcal{T})$ dans $\mathbb{R}^n$ muni de sa tribu borélienne. Si (f_n) converge ponctuellement vers une application $f : \Omega \to \mathbb{R}^n$ (i.e. pour tout $\omega \in \Omega$, $\lim_{n\to\infty} f_n(\omega) = f(\omega)$), alors f est mesurable.*

Démonstration. Notons que chaque ouvert dans $\mathbb{R}^n$ est union de boules ouvertes. Donc, la tribu borélienne sur $\mathbb{R}^n$ est engendrée par les boules ouvertes dans $\mathbb{R}^n$. Il suffit de montrer que si B est une boule ouverte alors $f^{-1}(B) \in \mathcal{T}$. Soit B la boule ouverte centrée dans $a \in \mathbb{R}^n$ et de rayon r. Pour chaque $s \in \mathbb{N}$, $r > \frac{1}{s}$, notons par B_s la boule ouverte centrée dans a et de rayon $r - \frac{1}{s}$. Puisque $\mathcal{T}$ est fermée pour les réunions et les intersections dénombrables, il suffit de montrer que

$$f^{-1}(B) = \bigcup_{s,m} \bigcap_{n \ge m} f_n^{-1}(B_s).$$

Si $x \in \bigcap_{n\ge m} f_n^{-1}(B_s)$ alors $f_n(x) \in B_s$ pour tous $n \ge m$ qui entraine que $f(x) = \lim f_n(x) \in \overline{B_s} \subset B$. D'où $f^{-1}(B) \supset \bigcup_{s,m} \bigcap_{n\ge m} f_n^{-1}(B_s)$. D'autre part, $x \in f^{-1}(B) \Leftrightarrow f(x) \in B \Rightarrow \exists s > 1$ tel que $f(x) \in B_s \Rightarrow \exists m$ t.q. $f_n(x) \in B_s$ pour $\forall n > m \Leftrightarrow x \in \bigcap_{n\ge m} f_n^{-1}(B_s)$. Donc, $f^{-1}(B) \subset \bigcup_{s,m} \bigcap_{n\ge m} f_n^{-1}(B_s)$. $\square$

Exercice 5.43. Dans le théorème ci-dessus $\mathbb{R}^n$ est consideré comme un espace normé, donc, métrique. Le théorème reste vrai (avec presque la même preuve) si on remplace $\mathbb{R}^n$ par un espace métrique quelconque. Démontrer cette généralisation du théorème.

Rappel 5.44. Si (a_n) est une suite dans $\mathbb{R}$ alors $\lim_n a_n = \infty$ (resp. $\lim_n a_n = -\infty$) si pour chaque $m \in \mathbb{N}$ il existe $n_\circ(m)$ tel que $a_n > m$ (resp. $a_n < -m$) si $n > n_\circ(m)$.

Corollaire 5.45. Soit $f_n : \Omega \to \mathbb{R}$ une suite de fonctions boréliennes et soit $f : \Omega \to \overline{\mathbb{R}}$ une fonction. Supposons que $\lim_n f_n(\omega) = f(\omega)$ pour tout $\omega \in \Omega$. Alors f est aussi borélienne.

Démonstration. Pour tous $m, l \in \mathbb{N}^*$ on pose $A_{m,l} \overset{def}{=} \bigcap_{n\ge l} f_n^{-1}(\{y > m\})$. Donc, $A_{m,l} = \bigcap_{n\ge l} \{x \in \Omega : f_n(x) > m$ pour tous $n \ge l\}$. Alors $\bigcap_m (\bigcup_l A_{m,l}) = \{x \in \Omega : f(x) = \lim_n f_n(x) = \infty\} = f^{-1}(\infty)$. Donc, $f^{-1}(\infty)$ est mesurable. De manière analogique on montre que $f^{-1}(-\infty)$ est aussi mesurable. Par Théorème 5.42 utilisé pour les restrictions de f_n et f sur $\Omega \setminus (f^{-1}(\infty) \cup f^{-1}(-\infty))$, $f^{-1}(B) \in \mathcal{T}$ pour tout $B \in \mathcal{B}(\mathbb{R})$. Donc, f est messurable. (Voir la Définition 5.35).

$\square$

Exercice 5.46. Avec les notations et les hypothèses du Corollaire 5.45, montrer que $f^{-1}(-\infty) \in \mathcal{T}$.

Proposition 5.47. *Soit $(\Omega, \mathcal{T})$ un espace mesurable et $f_n : \Omega \to \mathbb{R}$ une suite de fonctions boréliennes. Alors*

1. *les fonctions* $\sup_n f_n$ *et* $\inf_n f_n$ *sont boréliennes ;*
2. *les fonctions* $\limsup_n f_n$ *et* $\liminf_n f_n$ *sont boréliennes ;*

3. la partie $A = \{x \in \Omega : la\ suite\ (f_n(x))\ converge\}$ est mesurable ;

4. si $h : \Omega \to \mathbb{R}$ est une fonction borélienne alors $\{x \in \Omega : la\ \ suite\ \ (f_n(x))\ converge\ \ et \lim_n f_n(x) = h(x)\}$ est mesurable.

Démonstration. 1. La fonction $g_n = f_1 \vee \cdots \vee f_n$ est borélienne en vertu du Corollaire 5.40. On a $g_{n+1} \geq g_n$ et $\sup_n f_n = \lim_n g_n$. D'où, $\sup_n f_n$ est mesurable par Corollaire 5.45. La fonction $\inf_n f_n$ est borélienne parce que $\inf_n f_n = -\sup_n (-f_n)$.

2. Considérons $\liminf$.

La fonction $h_n = \inf_{m \geq n} f_m$ est borélienne en vertu de 1. Mais $\liminf_n f_n = \lim_n h_n = \sup_n h_n$. En utilisant 1, on obtient que $\liminf_n f_n$ est borélienne. La fonction $\limsup f_n$ est borélienne car $\limsup_n f_n = -\liminf_n (-f_n)$.

3. A est mesurable en vertu de 2 parce que $A = \{\omega \in \Omega : \liminf\ f_n(\omega) = \limsup\ f_n(\omega)\}$.

4. Notons $f = \liminf\ f_n$ et $g = \limsup\ f_n$. Alors $\{\omega \in \Omega : (f-g)(\omega) = (f-h)(\omega) = 0\}$ est mesurable parce que f, g et h sont des fonctions mesurables.

$\square$

5.3.3 Fonctions étagées

Définition 5.48. Soit $(\Omega, \mathcal{T})$ un espace mesurable. On appelle fonction étagée (à valeurs dans $\mathbb{R}$) une fonction de la forme $f(\omega) = \sum_{1 \leq i \leq k} a_i\ 1_{A_i}(\omega)$ où $A_i \in \mathcal{T}$ sont d.d.d., $a_i \in \mathbb{R}$ et $k \in \mathbb{N}^*$.

Théorème 5.49. *Toute fonction borélienne de $(\Omega, \mathcal{T})$ dans $\mathbb{R}$ est limit simple de fonctions étagées. Si f est positive, la limite peut être choisie croissante.*

Démonstration. Prenons d'abord f positive. Pour tous $n, k \in \mathbb{N}$, on définit

$$A_{n,k} = \{\omega : \frac{k-1}{2^n} \leq f(\omega) < \frac{k}{2^n}\}.$$

Alors les parties $A_{n,k} \in \mathcal{T}$ et elles sont d.d.d. La suite

$$f_n(\omega) = \sum_{1 \leq k \leq 2^{n^2}} \frac{k-1}{2^n} 1_{A_{n,k}}(\omega)$$

est croissante et converge vers $f(\omega)$.

Si f est quelconque on peut écrire $f = f^+ - f^-$ où $f^+ \stackrel{def}{=} f \vee 0$ et $f^- \stackrel{def}{=} (-f) \vee 0$ et, puis, approximer les fonctions positives f^+ et f^- comme ci-dessus. Si $\lim_n f_n^+ = f^+$ et $\lim_n f_n^- = f^-$ alors $f = \lim_n (f_n^+ - f_n^-)$. $\square$

Le cours du 13/10/2020

Corollaire 5.50. Soit $f : \Omega \to \overline{\mathbb{R}}$ une fonction borélienne. Alors f est limit simple de fonctions étagées. Si f est positive, la limite peut être choisie croissante.

Démonstration. En utilisant la formule $f = f^+ - f^-$, le cas général se réduit au cas $f \geq 0$. Supposons que $f \geq 0$ et soit $A = f^{-1}(\infty)$. Alors $A \in \mathcal{T}$. Soit $f' = f_{|\Omega \setminus A}$. Alors f' est borélienne à valeurs dans $\mathbb{R}^+$. Soit (f_n') une suite de fonctions étagées positives sur $\Omega \setminus A$ (Théorème 5.49) telle que $\lim_n f_n'(\omega) = f'(\omega)$ pour tous $\omega \in \Omega \setminus A$. Soit

$$f_n(\omega) = \begin{cases} f_n'(\omega) & \text{si } \omega \in \Omega \setminus A \\ n & \text{si } \omega \in A \end{cases}$$

Alors f_n est étagée et $\lim_n f_n = f$.

□

5.3.4 Le théorème de classes monotones et le théorème de prolongement

Rappel 5.51. Une tribu est une classe monotone.

Proposition 5.52. *Une classe monotone $\mathcal{M}$ stable par intersection finie est une tribu.*

Démonstration. Soit $A_1, A_2 \in \mathcal{M}$. Montrons que $A_1 \cup A_2 \in \mathcal{M}$. Supposons que $A_1 \cap A_2 = \emptyset$. Alors $A_1^c \supset A_2 \Rightarrow A_1^c \setminus A_2 = (A_1 \cup A_2)^c \in \mathcal{M} \Rightarrow A_1 \cup A_2 \in \mathcal{M}$. Soit $A_1, A_2 \in \Omega$ quelconques. Par l'hypothèse $A_1 \cap A_2 \in \mathcal{M}$ et, donc, $A_1 \setminus (A_1 \cap A_2) \in \mathcal{M}$. Mais $A_1 \setminus (A_1 \cap A_2)$ et A_2 sont disjoints et leur réunion est $A_1 \cup A_2$. Par le précédent, $A_1 \cup A_2 \in \mathcal{M}$. En générale, si A_i est une suite dans $\mathcal{M}$ on obtient par recurrence que $B_n = \bigcup_{1 \le i \le n} A_i \in \mathcal{M}$. Mais $B_n \subset B_{n+1}$ et $\bigcup_n B_n = \bigcup_i A_i$. D'où, $\bigcup_i A_i \in \mathcal{M}$, i.e., $\mathcal{M}$ est une tribu.

□

Théorème 5.53. ***(des clases monotones)*** *Soit $\mathcal{E}$ une famille de parties de Ω stable par intersection finie. Alors $\mathcal{M}(\mathcal{E}) = \mathcal{T}(\mathcal{E})$.*

Démonstration. On a $\mathcal{M}(\mathcal{E}) \subset \mathcal{T}(\mathcal{E})$. Pour démontrer l'inclusion inverse, nous montrons que $\mathcal{M}(\mathcal{E})$ est stable par intersection finie (Proposition 5.52). Soit

$$\mathcal{M}_1 = \{A \in \mathcal{M}(\mathcal{E}) : \forall\, B \in \mathcal{E}, A \cap B \in \mathcal{M}(\mathcal{E})\}.$$

Si $A_1, A_2 \in \mathcal{M}_1$, $A_1 \subset A_2$ et $B \in \mathcal{E}$ alors $(A_2 \setminus A_1) \cap B = (A_2 \cap B) \setminus (A_1 \cap B) \in \mathcal{M}(\mathcal{E})$. Donc, $A_2 \setminus A_1 \in \mathcal{M}_1$. Encore, si $(A_n) \in \mathcal{M}_1$, $A_n \subset A_{n+1}$ et $B \in \mathcal{E}$ alors $(\bigcup_n A_n) \cap B = \bigcup_n (A_n \cap B) \in \mathcal{M}(\mathcal{E})$. Donc, $\bigcup_n A_n \in \mathcal{M}_1$. Nous avons démontré que $\mathcal{M}_1$ est une classe monotone qui contient $\mathcal{E}$ et est contenue dans $\mathcal{M}(\mathcal{E})$. D'où, $\mathcal{M}_1 = \mathcal{M}(\mathcal{E})$.

Soit

$$\mathcal{M}_2 = \{B \in \mathcal{M}(\mathcal{E}) : \forall\, C \in \mathcal{M}(\mathcal{E}), B \cap C \in \mathcal{M}(\mathcal{E})\}.$$

On montre comme ci-dessus que $\mathcal{M}_2$ est une classe monotone. Par le précédent, $\mathcal{M}_2$ contient $\mathcal{E}$ et, par définition, $\mathcal{M}_2 \subset \mathcal{M}(\mathcal{E})$. Donc, $\mathcal{M}_2 = \mathcal{M}(\mathcal{E})$. Par conséquence, $\mathcal{M}(\mathcal{E})$ est stable par intersection finie. Proposition 5.52 implique que $\mathcal{M}(\mathcal{E})$ est une tribu.

□

Corollaire 5.54. Soient μ et ν des mesures sur un espace mesurable $(\Omega, \mathcal{T})$. Soit $\mathcal{C}$ un clan qui engendre $\mathcal{T}$. Si μ et ν coïncident sur $\mathcal{C}$ et sont finies (c.à.d. $\mu(\Omega) = \nu(\Omega) < \infty$) alors elles sont égales.

Démonstration. Soit $\mathcal{M} = \{A \in \mathcal{T} : \mu(A) = \nu(A)\}$. Il est facile à voir que $\mathcal{M}$ est une classe monotone contenante $\mathcal{C}$. Donc, $\mathcal{M}(\mathcal{C}) \subset \mathcal{M} \subset \mathcal{T}$. Mais $\mathcal{C}$ est stable par intersection finie. Par Théorème 5.53, $\mathcal{M}(\mathcal{C}) = \mathcal{T}(\mathcal{C}) = \mathcal{T}$. D'où, $\mathcal{M} = \mathcal{T}$.

□

Exercice 5.55. Est-ce que la preuve du Corollaire 5.54 marche pour les mesures infinies ?

La démonstration du théorème suivant est admise.

Théorème 5.56. ***(de prolongement)*** *Soit μ une fonction additive d'ensembles (c.à.d. $\mu(A \sqcup B) = \mu(A) + \mu(B)$), positive, définie sur un clan $\mathcal{C}$ de parties de Ω. Si pour toute suite croissante (A_n) d'éléments de $\mathcal{C}$ de réunion $A \in \mathcal{C}$, $\lim_{n\to\infty} \mu(A_n) = \mu(A)$, alors μ se prolonge à une mesure sur $(\Omega, \mathcal{T}(\mathcal{C}))$. Si μ est σ-finie sur $\mathcal{C}$ (c.à.d. s'il existe une suite (B_n) d'éléments de $\mathcal{C}$ telle que $B_n \nearrow \Omega$ et $\mu(B_n) < \infty$ pour tous n) alors le prolongement est unique.*

Exercice 5.57. En utilisant Corollaire 5.54 démontrer l'unicité de prolongement de μ dans la formulation du Théorème 5.56.

Théorème 5.56 implique :

Théorème 5.58. *Soit $n \geq 1$ un entier. Il existe une unique mesure sur $\mathcal{B}(\mathbb{R}^n)$, notée λ_n (et appeléé* ***mesure de Lebesque****), telle que :*

$$\lambda_n([a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]) = \prod_{i=1}^{n} (b_i - a_i).$$

Démonstration. Soit $\mathcal{C}$ le clan engendré par les pavés connexes fermés. Par Corollaire 5.28 la tribu borélienne est engendrée par ce clan. La fonction λ_n sur $\mathcal{C}$ est additive en vertu de l'execice 5.3. Par Théorème 5.56 λ_n se prolonge à une mesure sur $\mathbb{R}^n$.

$\square$

Remarque 5.59. Pour $n = 1$ on retrouve la notion de longueur du segment $[a, b]$, pour $n = 2$ celle d'aire du $[a, b] \times [c, d]$, etc.

Exercice 5.60. Montrer que λ_n est invariante par rapport aux translations.

Exercice 5.61. *(Exemple d'un sous-ensemble de $\mathbb{R}$ non-borélien)* Si $x, y \in [0, 1]$ on écrit $x \sim y$ si $x - y \in \mathbb{Q}$.

1. Montrer que $\sim$ est une relation d'équivalence.
   Donc, on peut presenter $[0, 1]$ comme union des classes d'équivalence d.d.d. : $[0, 1] = \bigsqcup_{i \in I} C_i$.
   Puis, pour chaque $i \in I$ on choisit un élément et un seul $x_i \in C_i$ et on pose $A = \{x_i : i \in I\}$.
   Posons $A_q = q + A$, $\forall q \in \mathbb{Q} \cap [-1, 1]$.
2. Montrer que $A_q \cap A_r = \emptyset$ si $q \neq r$.
3. Montrer que $[0, 1] \subset \bigcup_{q \in \mathbb{Q} \cap [-1,1]} A_q \subset [-1, 2]$.
4. Soit λ la mesure de Lebesque sur $\mathbb{R}$. En supposant A borélien montrer que $\lambda(A_q) = \lambda(A)$.
5. En déduire que $1 \leq \infty \cdot \lambda(A) \leq 3$.
6. Conclure que A ne pourrait pas être borélien.

5.3.5 Ensembles négligeables et propriétés vraies μ-presque partout

On se fixe un espace mesuré $(\Omega, \mathcal{T}, \mu)$.

Définition 5.62.

1. On dit que $A \in \mathcal{P}(\Omega)$ est *μ-négligeable* s'il existe $B \in \mathcal{T}$ tel que $A \subset B$ et $\mu(B) = 0$.
2. Une fonction mesurable f vérifie une propriété P *μ-presque partout* (*μ-p.p.*) si l'ensemble $\{x \in \Omega : f(x) \text{ ne vérifie pas } P\}$ est μ-négligeable.

Exemple 15.

1. On considère l'espace mesuré $(\Omega, \mathcal{T}, \mu)$ avec $\Omega = \{0, 1, 2\}$, $\mathcal{T} = \mathcal{P}(\Omega)$, $\mu(\{0\}) = \mu(\{1\}) = 1$ et $\mu(\{2\}) = 0$. Soit $f : \Omega \to \mathbb{R}$ telle que $f(0) = f(1) = 7$ et $f(2) = 0$. Alors f est constante et égale à 7 μ-p.p.

2. On considère l'espace mesuré $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ où λ est la mesure de Lebesque. Soit $f : \mathbb{R} \to \mathbb{R}$ telle que $f(x) = 1$ si $x \in \mathbb{Q}$ et $f(x) = 0$ si $x \in \mathbb{R} \setminus \mathbb{Q}$. Alors $f = 0$ λ-p.p.
3. Soit f une fonction borélienne sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ telle que $f(0) = 0$, $f(x) = 1$ si $x > 0$ et $f(x) = -1$ si $x < 0$. Alors f est continue λ-p.p.

5.3.6 Mesure image

Définition 5.63. Soit f une application mesurable d'un espace mesuré $(\Omega, \mathcal{T}, \mu)$ dans un espace mesurable $(E, \mathcal{E})$. On appelle **mesure image de** μ **par** f la mesure sur $(E, \mathcal{E})$, notée μ_f, définie par

$$\mu_f(B) = \mu(f^{-1}(B)), B \in \mathcal{E}.$$

Exercice 5.64. Montrer que μ_f est une mesure sur l'espace mesurable $(E, \mathcal{E})$. En plus, si μ est une probabilité, μ_f est aussi une probabilité.

Exemple 16. Considérons le jeu de dé avec $\Omega = \{1, 2, 3, 4, 5, 6\}$ et μ probabilité définie par $\mu(A) \overset{def}{=} \text{card}(A)/6$. Soit $f : \Omega \to \{0, 1\}$ définie par $f(\omega) = 1$ si ω est pair, et 0 si ω est impaire. On voit que

$$\mu_f(\{0\}) = \mu_f(\{1\}) = 1/2,$$

i.e. on a une chance sur deux d'obtenir un chiffre pair en jouant au dé. L'exemple montre que le formalisme utilisé dans la définition de μ_f coincide avec l'intuition que l'on peut avoir du hazard.

Chapitre 6

Integration

6.1 Intégrale de fonctions mesurables positives

Dans ce chapitre on considère des fonctions boréliennes d'un espace mesuré $(\Omega, \mathcal{T}, \mu)$ à valeurs dans $\overline{\mathbb{R}}$.

Définition 6.1. Soient $B \in \mathcal{T}$ et f une fonction étagées, positive et borélienne, c.à.d. $f(\omega) = \sum\limits_{1 \leq i \leq n} a_i 1_{A_i}(\omega)$ où $a_i \geq 0$ et A_i sont des ensembles mesurables d.d.d. *L'intégrale de f sur B par rapport à μ* est défini par

$$\int\limits_B f d\mu = \sum_{1 \leq i \leq n} a_i \mu(A_i \cap B) = \sum_{1 \leq i \leq n} a_i \int\limits_B 1_{A_i} d\mu.$$

Si $B = \Omega$ on parle tout simplement *d'intégrale de f par rapport à μ.*

Exercice 6.2. La valeur de $\int\limits_B f d\mu$ ne dépend pas de la décomposition de f en somme d'indicatrices (i.e., en somme de $a_i 1_{A_i}$).

Définition 6.3. Soit $f : \Omega \to \overline{\mathbb{R}}^+$ une fonction borélienne. La définition de son intégrale par rapport à μ sur l'ensemble mesurable B est donnée par

$$\int\limits_B f d\mu = \int\limits_B f(\omega) d\mu(\omega) = \sup\{\int\limits_B g d\mu : g \text{ étagée positive}, g \leq f\}.$$

Si $B = \Omega$ on écrit $\int f d\mu$ au lieu de $\int\limits_\Omega f d\mu$. La fonction f est *intégrable sur B* si $\int\limits_B f d\mu < \infty$.

Proposition 6.4. *Soient f et g des fonctions boréliennes positives et $A, B \in \mathcal{T}$.*

1. $f \leq g \Rightarrow 0 \leq \int\limits_B f d\mu \leq \int\limits_B g d\mu$ *;*
2. $A \subset B \Rightarrow \int\limits_B f d\mu = \int\limits_A f d\mu + \int\limits_{B \setminus A} f d\mu$. *En particulier,* $\int\limits_A f d\mu \leq \int\limits_B f d\mu$ *;*
3. $c \geq 0 \Rightarrow \int\limits_B c f d\mu = c \int\limits_B f d\mu$ *;*
4. $f = 0 \Rightarrow \int f d\mu = 0$ *;*
5. $\int\limits_B f d\mu = \int 1_B f d\mu$.

Démonstration. Idée de la preuve (facile) est la suivante : d'abord on montre la proposition pour les fonctions étagées positives et puis on passe au supremum pour les fonctions mesurables positives. $\square$

Exercice 6.5. Démontrer la proposition ci-dessus en utilisant l'indication.

Proposition 6.6. ***(inégalité de Markov)*** *Soit $f : \Omega \to \overline{\mathbb{R}}^+$ mesurable. Alors pour tout $a > 0$ on a*

$$\mu(\{x \in \Omega : f(x) \geq a\}) \leq \frac{1}{a}\int_\Omega f d\mu.$$

Démonstration. Soit $A = \{x \in \Omega : f(x) \geq a\}$. Alors $\int_\Omega f d\mu \geq \int_A f d\mu \geq a\mu(A)$. $\square$

Corollaire 6.7. Soit $f : \Omega \to \overline{\mathbb{R}}^+$ mesurable. Alors $\int_\Omega f d\mu = 0$ si et seulement si $f = 0$ μ-p.p.

Démonstration. Supposons que $\int_\Omega f d\mu = 0$. Soit $\Omega_n = \{x \in \Omega : f(x) \geq 1/n\}$ où $n \in \mathbb{N}^*$. Alors $\Omega_n \subset \Omega_{n+1}$ et par l'inégalité de Markov $\mu(\Omega_n) = 0$. Donc, $\mu(\bigcup_n \Omega_n) = 0$. Mais $\bigcup_n \Omega_n = \{x \in \Omega : f(x) \neq 0\}$. Donc, $f = 0$ μ-p.p.

Soit $f = 0$ μ-p.p. et h une fonction étagée positive et telle que $h \leq g$. Si $h = \sum_{j=1}^m a_j 1_{A_j}$ avec $a_j > 0$ alors $A_j \subset \{x \in \Omega : f(x) \geq a_j\}$. Donc, $\mu(A_j) = 0$ et, par conséquant, $\int_\Omega h d\mu = 0$. Par la Définition 6.3, $\int_\Omega f d\mu = 0$. $\square$

Proposition 6.8. *Soient $f, g : \Omega \to \overline{\mathbb{R}}^+$ mesurables et telles que $f = g$ $\mu-$ p.p. Alors $\int_\Omega f d\mu = \int_\Omega g d\mu$. En particulier, f est intégrable si et seulement si g l'est.*

Démonstration. Soit $A = \{x \in \Omega : f(x) = g(x)\}$. Alors A est mesurable et $\mu(A^c) = 0$. Donc, $\int_{A^c} f d\mu = \int_{A^c} g d\mu = 0$. D'où, $\int_\Omega f d\mu = \int_A f d\mu = \int_A g d\mu = \int_\Omega g d\mu$. $\square$

Proposition 6.9. *Si $f : \Omega \to \overline{\mathbb{R}}^+$ est intégrable alors $\mu(\{x \in \Omega : f(x) = \infty\}) = 0$.*

Démonstration. Soit $A = \{x \in \Omega : f(x) = +\infty\}$. Pour tout $\alpha > 0$ on a $\alpha \cdot 1_A \leq f$. Donc,

$$\int_\Omega \alpha \cdot 1_A d\mu = \alpha\mu(A) \leq \int_\Omega f d\mu < \infty.$$

Si on avait $\mu(A) > 0$ on obtiendrait une contradiction en faisant tendre α vers $+\infty$. $\square$

Le resultat suivant est laissé en exercice. Il représente un corollaire facile des propriétés basiques de la mesure et de l'intégrale.

Proposition 6.10. *Soient $(\Omega, \mathcal{T}, \mu)$ un espace mesuré et $f : \Omega \to [0, +\infty]$ une fonction mesurable. On définit une application $\nu : \mathcal{T} \to [0, +\infty]$ par*

$$\nu(A) = \int_A f d\mu.$$

Alors ν est une mesure sur l'espace mesurable $(\Omega, \mathcal{T})$ appeleé mesure à densité par rapport à μ *ou* mesure de densité f. *Pour une mesure à densité ν par rapport à μ si $\mu(A) = 0$ alors $\nu(A) = 0$.*

Remarque 6.11. Un resultat beaucoup plus profond (le théorème de Radon-Nikodym) dit que si μ et ν sont des mesures de probabilités sur $(\Omega, \mathcal{T})$ telles que $\mu(A) = 0$ implique $\nu(A) = 0$ alors ν est mesure à densité par rapport à μ. Ce resultat sort du programme de l'UE.

Exemple 17. On définit une mesure de probabilité sur $\mathcal{B}(\mathbb{R})$ en posant

$$\mu(A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-\frac{x^2}{2}} d\lambda(x).$$

C'est la mesure *gaussienne* importante dans la théorie de probabilité. Le fait qu'il s'agit d'une mesure de proba se vérifie par changement de variables expliqué plus loin. (Voir Exemple 28.)

6.2 Intégrale de fonctions mesurables et les théorèmes de convergence

Théorème 6.12. ***(théorème de convergence monotone de Beppo Levi)*** *Soit $f_n : \Omega \to \overline{\mathbb{R}}^+$ une suite croissante de fonctions mesurables convergeante ponctuellement vers f. Alors f est mesurable et*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Démonstration. Par Corollaire 5.45 la fonction f est mesurable. En vertu de la Proposition 6.4(1) la suite $\int f_n d\mu$ est positive, croissante et, donc, converge vers un $\alpha \in \overline{\mathbb{R}}^+$. Puisque $f_n \leq f$, $\alpha \leq \int f d\mu$. Soit g une fonction positive, étagée, et telle que $f \geq g$. Fixons un $0 < c < 1$. Soit $A_n = \{x \in \Omega : f_n(x) \geq c \cdot g(x)\}$. Alors $A_n \nearrow \Omega$ et

$$\int\limits_{A_n} f_n d\mu + \int\limits_{A_n^c} cg d\mu \geq c \int g d\mu. \tag{6.1}$$

Il existe $m > 0$ tel que $cg(x) \leq m$ pour tout x. Par conséquence, $\int\limits_{A_n^c} cg d\mu \leq m\mu(A_n^c)$. Mais $A_n^c \searrow \emptyset$. Donc, $\mu(A_n^c) \to 0$, entraînant que $\int\limits_{A_n^c} cg d\mu \to 0$. Par la définition de α et la Proposition 6.4(2) $\alpha \geq \int\limits_{A_n} f_n d\mu$. En passant vers la limite dans (6.1) on obtient

$$\alpha \geq c \int g d\mu.$$

Par la définition de l'intégrale $\int f d\mu$, $\alpha \geq c \int f d\mu$ pour tout $0 < c < 1$. D'où $\alpha \geq \int f d\mu$, i.e. $\alpha = \int f d\mu$. $\square$

Le cours du 20/10/2021

Exercice 6.13. Montrer que la fonction f dans la formulation du Théorème 6.12 est intégable si et seulement si la suite $\int f_n d\mu$ est majorée et si c'est le cas alors $\int f d\mu = \sup_n \int f_n d\mu$.

Exemple 18. (**d'application**) Soit $f_n : [0,1] \to \mathbb{R}, x \mapsto n \cdot x \cdot 1_{[0,\frac{1}{n}]}(x) + 1_{]\frac{1}{n},1]}(x)$. C'est une suite croissante de fonctions mesurables sur $[0,1]$. Donc, $\int_0^1 f_n d\lambda \to \int_0^1 f d\lambda$ puisque $f = \lim(f_n)$ est égale à 1 sur $]0,1]$.

Les exemples suivants montrent que, en général, on ne peut pas passer à la limite en l'absence de monotonie.

Exemples 1. 1. Soit $f_n = 1_{[n,n+1]}$. Alors $\lim_n f_n(x) = f(x) = 0$ pour tout $x \in \mathbb{R}$ et $\int_{\mathbb{R}} f_n d\lambda = 1 \neq 0$.

2. Soit $f_n = n1_{[0,\frac{1}{n}]}$. Alors $\lim_n f_n(x) = (\infty)1_{\{0\}}$. On a $\int_{\mathbb{R}} f_n d\lambda = 1 \neq \int_{\mathbb{R}} (\lim f_n) d\lambda = 0$. *La suite* (f_n) *n'est ni croissante ne décroissante* parce que

$$f_{n+1}(1/n) = 0 < f_{n-1}(1/n) = n - 1 < f_n(1/n) = n.$$

Corollaire 6.14. Soit (f_n) une suite de fonctions mesurables positives et soit $f = \sum_{n\in\mathbb{N}} f_n$. Alors

$$\int f d\mu = \sum_{n\in\mathbb{N}} \int f_n d\mu.$$

Démonstration. On applique Théorème 6.12 pour la suite $g_n = \sum_{0\leq m\leq n} f_n$ qui est croissante et converge vers f. □

Corollaire 6.15. Soit f et $g : \Omega \to \overline{\mathbb{R}}^+$ et α et $\beta \geq 0$. Alors

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu.$$

Démonstration. La preuve est facile à démontrer si f et g sont des fonctions étagées (laissée en exercice). Dans le cas général, d'après le Corollaire 5.50 $f = \lim_n f_n$ et $g = \lim_n g_n$ où (f_n) et (g_n) sont des suites croissantes de fonctions étagées et positives. Alors la suite $(\alpha f_n + \beta g_n)$ converge en croissant vers $\alpha f + \beta g$ et le resultat se déduit du théorème de convergence monotone (Théorème 6.12) et de sa validité pour les fonctions étagées. □

Corollaire 6.16. Soit $f_n : \Omega \to [0, +\infty[$ une suite **décroissante** de fonctions mesurables et soit $f = \lim(f_n)$. Alors si f_0 est intégrable f est aussi intégrable et

$$\int f d\mu = \lim_n \int f_n d\mu.$$

Démonstration. On a $f_0 = (f_0 - f_n) + f_n = (f_0 - f) + f$. Par Corolaire 6.15,

$$\int f_0 d\mu = \int (f_0 - f_n) d\mu + \int f_n d\mu = \int (f_0 - f) d\mu + \int f d\mu.$$

Par Théorème 6.12, $\int (f_0 - f) d\mu = \lim_n \int (f_0 - f_n) d\mu$. Donc, $\int f d\mu = \lim_n \int f_n d\mu$. □

Exemple 19. Soit $f_n = 1_{[n,+\infty[}$. C'est une suite décroissante de fonctions mesurables positives qui converge vers la fonction $f = 0$. On a que $\int f d\mu = 0$ malgré que $\int f_n d\mu = \lim_n \int f_n d\mu = +\infty$. Donc, le corollaire ci-dessus *n'est plus valable* si on omet l'hypothèse que f_0 est intégrable.

Proposition 6.17. *Soit* $f : \Omega \to \overline{\mathbb{R}}^+$ *mesurable et* ν *la mesure de densité* f. *Pour toute fonction* $g : \Omega \to \overline{\mathbb{R}}^+$ *mesurable on a*

$$\int_\Omega g(x) d\nu(x) = \int_\Omega g(x) f(x) d\mu(x).$$

Démonstration. L'égalité est vraie par définition de ν si g est une fonction indicatrice d'un ensemble mesurable. On en déduit qu'elle est vraie également si $g : \Omega \to [0, +\infty[$ est étagée mesurable car si $g = \sum_{j=1}^{m} a_j 1_{A_j}$ alors

$$\int_\Omega g(x) d\nu(x) = \sum_{j=1}^{m} a_j \nu(A_j) = \sum_{j=1}^{m} a_j \int_\Omega 1_{A_j} f d\mu = \int_\Omega g \cdot f d\mu.$$

De façon général, il existe une suite croissante de fonctions étagées mesurables $g_n : \Omega \to [0, +\infty[$ tendant vers g (Corollaire 5.50). Mais la suite de fonctions étagées $(g_n f)$ est croissante, tend vers gf et, d'après ce qui précède, on a

$$\int_\Omega g_n d\nu = \int_\Omega g_n \cdot f d\mu.$$

En appliquant le théorème de convergence monotone aux deux membres, on obtient le resultat. □

Théorème 6.18. ***(lemme de Fatou)*** *Soit (f_n) une suite de fonctions mesurables positives. Alors*

$$\int \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

Démonstration. Soit $g_n = \inf_{m \geq n} f_m$. Alors chaque g_n est mesurable (voir Proposition 5.47), la suite (g_n) est croissante et $\lim_{n\to\infty} g_n = \liminf_{n\to\infty} f_n d\mu$ est mesurable aussi (Proposition 5.47). Par Théorème 6.12, $\lim_{n\to\infty} \int g_n d\mu = \int \liminf_{n\to\infty} f_n d\mu$. Mais, $g_n \leq f_m$ pour tous $m \geq n$. Donc $\int g_n d\mu \leq \inf_{m\geq n} \int f_m d\mu$. D'où

$$\int \liminf_{n\to\infty} f_n d\mu \leq \lim_{n\to\infty} \inf_{m\geq n} \int f_m d\mu = \liminf_{n\to\infty} \int f_n d\mu.$$

□

Rappel 6.19. On a $f = f^+ - f^-$ et $|f| = f^+ + f^-$ où $f^+ = f \vee 0$ et $f^- = (-f) \vee 0 = -(f \wedge 0)$.

Définition 6.20. Soit f une fonction mesurable et $B \in \mathcal{T}$. On dit que f est *μ-intégrable sur B au sens de Lebesgue* (ou brièvement *intégrable sur B* si μ est implicite) si $\int_B |f| d\mu < \infty$. Si f est μ-intégrable sur B alors $\int_B f d\mu \overset{def}{=} \int_B f^+ d\mu - \int_B f^- d\mu$ est son intégrale par rapport à μ sur B.

Si $B = \Omega$, on dit que f est *μ-intégrable* et note souvent $\int f d\mu$ au lieu de $\int_\Omega f d\mu$. L'ensemble de toutes fonctions intégrables est noté par $L^1(\Omega, \mathcal{T}, \mu)$.

Exercice 6.21. Soit $(\Omega, \mathcal{T}, \mu)$ où μ est σ-finie. On suppose que $\mathcal{T} = \mathcal{T}(\mathcal{C})$ où $\mathcal{C}$ est un clan. Soit $f : \Omega \to \mathbb{R}$ une fonction intégrable telle que $\int_B f d\mu = 0$ pour tout $B \in \mathcal{C}$. On se propose de démontrer que $f = 0$ μ-p.p.

1. On pose $\mu^+(A) = \int_A f^+ d\mu$ et $\mu^-(A) = \int_A f^- d\mu$ pour tout $A \in \mathcal{T}$. Montrer que μ^+ et μ^- sont des mesures sur $(\Omega, \mathcal{T})$.
2. Montrer que $\mu^+ = \mu^-$. (Indication : On pourra utiliser Théorème 5.56.)
3. Montrer que $\int f^+ d\mu = \int f^- d\mu = 0$.
4. En déduire que $f^+ = f^- = 0$ μ-p.p. (Indication : Voir l'exercice 6.7.)
5. Conclure.

Proposition 6.22. *Soit f et g intégrables et $B \in \mathcal{T}$.*

1. *Si $\alpha, \beta \in \mathbb{R}$ alors $\alpha f + \beta g$ est intégrable sur B et*

$$\int_B (\alpha f + \beta g) d\mu = \alpha \int_B f d\mu + \beta \int_B g d\mu.$$

En particulier, $L^1(\Omega, \mathcal{T}, \mu)$ est un espace vectoriel sur $\mathbb{R}$.

2. *Si $f \leq g$ alors*

$$\int_B f d\mu \leq \int_B g d\mu.$$

Démonstration. 1. On a $\int_B |f| d\mu \leq \int_\Omega |f| d\mu$. Donc, si f est intégrable elle est aussi intégrable sur B et il suffit de démontrer la proposition pour $B = \Omega$. Si $f, g \geq 0$ et $\alpha, \beta \geq 0$ le résultat est démontré par le Corolaire 6.15.

Maintenant, soient $f, g \in L^1(\Omega, \mathcal{T}, \mu)$ et $\alpha, \beta \in \mathbb{R}$. Comme $|\alpha f + \beta g| \leq |\alpha| \cdot |f| + |\beta| \cdot |g|$ le précédent implique que $L^1(\Omega, \mathcal{T}, \mu)$ est un espace vectoriel. De plus, $(f+g)^+ - (f+g)^- = f + g = f^+ - f^- + g^+ - g^-$, d'où $(f+g)^+ + f^- + g^- = (f+g)^- + f^+ + g^+$. Par l'additivité de l'intégrale des fonctions positives (Corolaire 6.15) on obtient

$$\int (f+g)^+ d\mu + \int f^- \mu + \int g^- \mu = \int (f+g)^- \mu + \int f^+ \mu + \int g^+ \mu.$$

D'où,

$$\int (f+g)^+ d\mu - \int (f+g)^- \mu = \int f^+ \mu - \int f^- \mu + \int g^+ \mu - \int g^- \mu$$

c.à.d. $\int (f+g) d\mu = \int f \mu + \int g \mu$.

Il reste de montrer que $\int (\lambda f) d\mu = \lambda \int f d\mu$ si $f \in L^1(\Omega, \mathcal{T}, \mu)$ et $\lambda \in \mathbb{R}$. Si $\lambda \geq 0$ alors $(\lambda f)^+ = \lambda f^+$, $(\lambda f)^- = \lambda f^-$. Grace à l'homogénéité de l'intégrale des fonctions positives, on a

$$\int (\lambda f) d\mu = \lambda \int f^+ d\mu - \lambda \int f^- d\mu = \lambda \int f d\mu.$$

Si maintenant, $\lambda < 0$ alors $\lambda f^+ = -(\lambda f)^-$ et $\lambda f^- = -(\lambda f)^+$. D'où

$$\int (\lambda f) d\mu = (-\lambda) \int f^- d\mu - (-\lambda) \int f^+ d\mu = \lambda \int f d\mu.$$

2. Si $f \leq g$ alors $g - f \geq 0$ et $\int (g - f) d\mu \geq 0$ d'après la Proposition 6.4 (1). Maintenant il suffit d'utiliser la linéarité de l'intégale montrée ci-dessus. □

Corollaire 6.23. (inégalité "triangulaire") Si $f \in L^1(\Omega, \mathcal{T}, \mu)$ (i.e. si f est intégrable) alors

$$\left| \int f d\mu \right| \leq \int | f | \, d\mu.$$

Démonstration. En effet, si $f \in L^1(\Omega, \mathcal{T}, \mu)$ alors $|f|$ et $\pm f$ sont intégrables et $|f| \geq \pm f$. Par la Proposition 6.22(2)

$$\int | f | \, d\mu \geq \pm \int f d\mu,$$

autrement dit, $\int | f | \, d\mu \geq \left| \int f d\mu \right|$. □

Le lemme de Fatou implique :

Corollaire 6.24. Soient $g \in L^1(\Omega, \mathcal{T}, \mu)$ et $f_n \in L^1(\Omega, \mathcal{T}, \mu), n \in \mathbb{N}$. Alors

1. si $g \leq f_n$ alors $\int \liminf\limits_{n\to\infty} f_n d\mu \leq \liminf\limits_{n\to\infty} \int f_n d\mu$;
2. si $f_n \leq g$ alors $\limsup\limits_{n\to\infty} \int f_n d\mu \leq \int \limsup\limits_{n\to\infty} f_n d\mu$.

Démonstration. 1. D'après le lemme de Fatou (Théorème 6.18), on a

$$\int \liminf_{n\to\infty}(f_n - g)d\mu \leq \liminf_{n\to\infty} \int (f_n - g)d\mu.$$

Pour conclure, il suffit d'ajouter $\int g d\mu$ de deux cotés d'inégalité ci-dessus, d'utiliser la définition de lim inf et d'appliquer la Proposition 6.22.

2. Par le lemme de Fatou

$$\int \liminf_{n\to\infty}(g - f_n)d\mu \leq \liminf_{n\to\infty} \int (g - f_n)d\mu.$$

Pour completer la preuve on utilise que $\liminf\limits_{n\to\infty}(a_n) = -\limsup\limits_{n\to\infty}(-a_n)$ pout chaque suite (a_n). □

Théorème 6.25. ***(de convergence dominée de Lebesque)*** *Soit (f_n) une suite de fonctions mesurables telle que $|f_n| \leq g$ où g est intégrable et f_n converge ponctuellement vers f. Alors f est intégrable et*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Démonstration. La fonction f est mesurable en vertu du Théorème 5.42 et elle est intégrable parce que $|f| \leq g$ et g est intégrable. Notons que $-g \leq f \leq g$. Par Corollaire 6.24 et l'hypothèse $\lim\limits_{n} f_n = f$, on obtient

$$\limsup_{n\to\infty} \int f_n d\mu \leq \int \limsup_{n\to\infty} f_n d\mu = \int f d\mu = \int \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu \leq \limsup_{n\to\infty} \int f_n d\mu.$$

D'où,

$$\liminf_{n\to\infty} \int f_n d\mu = \limsup_{n\to\infty} \int f_n d\mu = \lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

□

6.3 Théorème de transfert

Rappelons la définition d'une mesure image :

Définition 6.26. (voir *Définition* 5.63) Soit ϕ une application mesurable d'un espace mesuré $(\Omega, \mathcal{T}, \mu)$ dans un espace mesurable $(E, \mathcal{E})$. On appelle **mesure image de μ par ϕ** la mesure sur $(E, \mathcal{E})$, notée $\phi_*\mu$, définie par

$$\phi_*\mu(B) = \mu(\phi^{-1}(B)), B \in \mathcal{E}.$$

Théorème 6.27 (Théorème de transfert)**.** *Soit $\phi : (\Omega, \mathcal{T}, \mu) \to (E, \mathcal{E})$ une application mesurable et $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une fonction mesurable. Alors, si f est à valeur positive :*

$$\int_\Omega f \circ \phi d\mu = \int_E f d(\phi_*\mu).$$

Dans le cas général (c.à.d. si f est à valeur quelconque) $f \circ \phi \in L^1(\Omega, \mathcal{T}, \mu)$ si et seulement si $f \in L^1(E, \mathcal{E}, \phi_*\mu)$ et on a encore $\int_\Omega (f \circ \phi)d\mu = \int_E f d(\phi_*\mu)$.

Démonstration. On procède comme pour la construction de l'intégrale. Si $f = 1_B$ avec $B \in \mathcal{E}$, $f \circ \phi = 1_{\phi^{-1}(B)}$ et donc

$$\int f \circ \phi d\mu = \mu(\phi^{-1}(B)) = \phi_* \mu(B) = \int f d(\phi_* \mu).$$

Si f est étagée alors $f \circ \phi$ et aussi étagée et par linéarité, on obtient la preuve du théorème si f est étagée. Si f positive, f est la limite croissante d'une suite de fonctions étagées f_n. (Voir Théorème 5.49.) Comme $f_n(x) \to f(x)$, on applique le théorème de convergence monotone aux deux mesures :

$$\int f \circ \phi d\mu = \lim_{n\to\infty} \int f_n \circ \phi d\mu = \lim_{n\to\infty} \int f_n d(\phi_* \mu) = \int f d(\phi_* \mu).$$

Le dernier résultat du cas intégrable s'ensuit facilement du cas de fonctions positives (considéré plus haut) et des rélations $(f \circ \phi)^+ = f^+ \circ \phi$ et $(f \circ \phi)^- = f^- \circ \phi$. Pour démontrer l'égalité on utilise la linéarité de l'intégrale. □

Remarque 6.28.

1. Si $E = \mathbb{R}$ le théorème ramène une intégrale sur Ω à une intégrale sur $\mathbb{R}$:

$$\int_\Omega f \circ \phi d\mu = \int_{\mathbb{R}} f d(\phi_* \mu).$$

2. Le théorème de transfère est à la base du théorème de changement de variables que l'on verra plus tard au chapitre sur les intégrales de fonctions de plusieurs variables.

Le cours du 3/11/2021

6.4 Intégrales de Riemann et Lebesgue

Dans ce section $a < b$, $I = [a, b]$ et $f : I \to \mathbb{R}$. On suppose que f est bornée, i.e. il existe $C > 0$ tel que $-C < f(x) < C$ pour tous $x \in I$.

Soit σ *une division de I*, i.e. $\sigma = (a = x_0 < x_1 < \cdots < x_n = b)$. Alors $\delta(\sigma) \overset{def}{=} \max\{x_i - x_{i-1},\ i \geq 1\}$ est *le pas de* σ. On pose $m_i = \inf\limits_{x_{i-1} \leq x \leq x_i} f(x)$ et $M_i = \sup\limits_{x_{i-1} \leq x \leq x_i} f(x)$. Rappelons que

$$S^-(f, \sigma) = \sum_{i=1}^n m_i (x_i - x_{i-1})$$

est la somme inférieure de Darboux associée à σ et

$$S^+(f, \sigma) = \sum_{i=1}^n M_i (x_i - x_{i-1})$$

est la somme supérieure de Darboux associée à σ. Il est evident que $S^-(f, \sigma) \leq S^+(f, \sigma)$. Introduisons les fonctions étagées suivantes :

$$f_\sigma^-(x) = \begin{cases} m_1 & \text{si } x \in [x_0, x_1] \\ m_i & \text{si } x \in (x_{i-1}, x_i] \text{ et } 2 \leq i \leq n, \end{cases}$$

et

$$f_\sigma^+(x) = \begin{cases} M_1 & \text{si } x \in [x_0, x_1] \\ M_i & \text{si } x \in (x_{i-1}, x_i] \text{ et } 2 \leq i \leq n. \end{cases}$$

Alors

$$\int_I f_\sigma^- d\lambda = S^-(f,\sigma) \text{ et } \int_I f_\sigma^+ d\lambda = S^+(f,\sigma), \tag{6.2}$$

où λ est la mesure de Lebesgue.

On dit que la division $\sigma' = (a = y_0 < y_1 < \cdots < y_m = b)$ est plus petite que la division $\sigma = (a = x_0 < x_1 < \cdots < x_n = b)$, on note $\sigma' \leq \sigma$, si $\{y_0, y_1, \cdots, y_m\} \supset \{x_0, x_1, \cdots, x_n\}$. Si (σ_k) est une suite de divisions de I alors $\sigma_k \to 0$ si $\sigma_k \geq \sigma_{k+1}$ pour tous k et $\lim\limits_n \delta(\sigma_k) = 0$. Etant donnée une telle suite on pose $f_i^- = f_{\sigma_i}^-$ et $f_i^+ = f_{\sigma_i}^+$. Alors $f_k^- \leq f_{k+1}^- \leq f \leq f_{k+1}^+ \leq f_k^+$. Soit $\lim\limits_k f_k^- = f^-$ et $\lim\limits_k f_k^+ = f^+$. Les fonctions f^- et f^+ sont boréliennes, par le théorème de convergence dominée de Lébesque intégrables sur I, $f^- \leq f \leq f^+$ et

$$\int_I f^- d\lambda = \lim_k \int_I f_k^- d\lambda \leq \lim_k \int_I f_k^+ d\lambda = \int_I f^+ d\lambda. \tag{6.3}$$

Rappelons la définition de l'intégrale de Reimann.

Définition 6.29. On dit que la fonction f est *intégrable au sens de Riemann (ou Riemann intégrable)* s'il existe une suite de divisions (σ_k) telle que $\sigma_k \to 0$ et $\lim\limits_{k\to\infty} S^-(f,\sigma_k) = \lim\limits_{k\to\infty} S^+(f,\sigma_k)$. Dans ce cas $\lim\limits_{k\to\infty} S^{\pm}(f,\sigma_k)$ est noté par $\int_a^b f(x)dx$ est appelé intégrale de Riemann.

Desormais on suppose que f est Riemann intégrable et (σ_k) est comme dans la définition ci-dessus. D'après (6.2) et (6.3), on obtient

$$\int_I f^- d\lambda = \int_a^b f(x)dx = \int_I f^+ d\lambda.$$

Soit $A = \{x \in I : f^+(x) = f(x) = f^-(x)\}$. On a $f^+ - f^- \geq 0$ et $\int_I (f^+ - f^-)d\lambda = 0$. Donc,

$$\lambda(A) = \lambda(I) = b - a.$$

On peut montrer que chaque $\alpha \in A$ est un point de continuité pour f, c.à.d. si (α_i) est une suite dans I convergeante vers α alors $(f(\alpha_i))$ converge vers $f(\alpha)$. En effet, supposons par l'absurd que $(f(\alpha_i))$ ne converge pas vers $f(\alpha)$. En remplaçant (α_i) par sa sous-suite on peut supposer que $(f(\alpha_i))$ converge vers $\beta \neq f(\alpha)$ et que soit $\alpha_i < \alpha$ pour tous i soit $\alpha_i > \alpha$ pour tous i. Supposons que $\alpha_i < \alpha$ pour tous i et $\beta < f(\alpha)$. (Les cas restants se traitent de manière analogique.) Il existe $\varepsilon > 0$ tel que $\beta < f(\alpha) - \varepsilon$. Pour chaque k ils existent deux points voisins p_k et q_k de la division σ_k tels que $p_k < \alpha \leq q_k$. Donc, il existe n_0 tel que $p_k < \alpha_n < \alpha \leq q_k$ pour tous $n > n_0$. D'où, $f_k^-(\alpha) = f_k^-(\alpha_n) \leq \inf f(\alpha_n) = \beta < f(\alpha) - \varepsilon$ si $n > n_0$. On obtient que $\lim\limits_k f_k^-(\alpha) = f^-(\alpha) \leq f(\alpha) - \varepsilon$. Absurd.

On a démontré le théorème suivant.

Théorème 6.30. *Soit f Riemann intégrable. Alors f est continue λ-p.p. et il existe une fonction borélienne intégrable g sur I telle que $f = g$ λ-p.p. et $\int_a^b f(x)dx = \int_I g d\lambda$.*

Remarque 6.31. Le théorème implique que les théorèmes généraux qu'on a vu pour les intégrales de Lebesgue s'appliquent en particulier aux intégrales de Riemann.

Exercice 6.32. Soit $f = 1_{\mathbb{Q}\cap[0,1]}$. Alors f est Lebesgue intégrable et $\int f d\mu = 0$ mais f n'est pas Riemann intégrable.

Chapitre 7

Intégrales de fonctions de plusieurs variables

7.1 Intégrales dépendant d'un paramètre

Définition 7.1. Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré. Si $f : \Omega \to \mathbb{R}^m$ alors $f = (f_1, \cdots, f_m)$ où $f_j : \Omega \to \mathbb{R}$. La fonction f est mesurable (respectivement, intégrable) si et seulement si chaque f_j est mesurable (respectivement, intégrable). Si f est intégrable $\int f d\mu := (\int f_1 d\mu, \cdots, \int f_m d\mu)$. Puisque $\mathbb{C} = \mathbb{R} + i\mathbb{R}$ on peut identifier $\mathbb{C}$ avec $\mathbb{R}^2$ via $a + ib \mapsto (a, b)$. Donc, si $f : \Omega \to \mathbb{C}$, en utilisant cette identification, on peut définir la notion d'une fonction complexe intégrable. On peut introduire cette notion directement en utilisant les fonctions étagées complexes comme dans le chapitre 6 on a utilisé les fonctions étagées réelles pour introduire la notion d'une fonction réelle intégrable.

Définition 7.2. Soient $(\Omega, \mathcal{T}, \mu)$ un espace mesuré, A un espace métrique et $f : A \times \Omega \to \mathbb{R}^m$. On suppose que pour tout $x \in A$, $t \mapsto f(x,t)$, est intégrable. On pose $F(x) = \int_\Omega f(x,t) d\mu(t)$. La fonction $F : A \to \mathbb{R}^m$ est *une intégrale dépendant de paramètre x.*

Dans cette section on propose de conditions suffisantes pour que l'intégrale $F(x) = \int_\Omega f(x,t) d\mu(t)$ de paramètre x soit continue et/ou dérivable. Nos preuves sont basées sur le théorème de convergence dominée de Lebesque (Théorème 6.25).

Théorème 7.3. *(Théorème de continuité avec hypothèse de domination) Soit $f : A \times \Omega \to \mathbb{R}^m$. On suppose :*

1. *pour tout $x \in A$, la fonction $t \mapsto f(x,t)$, est mesurable sur Ω ;*
2. *pour presque tout $t \in \Omega$, la fonction $x \mapsto f(x,t)$, est continue ;*
3. *(hypothèse de domination) il existe une fonction $g \in L^1(\Omega, \mathcal{T}, \mu)$ telle que pour tout $x \in A$ et pour presque tout $t \in \Omega$ on a*
$$|f(x,t)| \leq g(t).$$

Alors la fonction $x \mapsto F(x) = \int_\Omega f(x,t) d\mu(t)$ est continue.

Démonstration. On se ramène facilement au cas $m = 1$. En ecartant de Ω un sous-ensemble mesurable de mesure 0 on se ramène au cas quand la fonction $x \mapsto f(x,t)$ est continue pour tout t et $|f(x,t)| \leq g(t)$ pour tout t aussi. L'hypothèse de domination garantit que $t \mapsto f(x,t)$ est intégrable. Soit $x_n \in A$ tel que $x_n \to a$. Par la continuité de $x \mapsto f(x,t)$, $f(x_n,t) \to f(a,t)$. On peut donc appliquer le théorème de convergence dominée de Lebesque (avec domination par g) pour conclure

$$\lim_{n\to\infty}\int_\Omega f(x_n,t)d\mu(t)=\int_\Omega f(a,t)d\mu(t)\Leftrightarrow\lim_{n\to\infty}F(x_n)=F(a).$$

□

Remarque 7.4. Il est important de remarquer que dans l'hypothèse de domination, la fonction g ne dépend pas du paramètre x.

Remarque 7.5. Dans la pratique l'espace métrique A dans la formulation du théorème est souvent une partie non-vide d'un evn E, par exemple, $E=\mathbb{R}^n$.

Exemple 20. Soit $f:\mathbb{R}\to\mathbb{C}$ intégrable sur $\mathbb{R}$. Sa *transformée de Fourier* est définie par :

$$\hat{f}(x)=\int_{\mathbb{R}} f(t)e^{itx}dt.$$

En vertu de Théorème 7.3, $\hat{f}$ est continue parce que $|f(t)e^{itx}|=|f(t)|$ est intégrable et pour tout t la fonction $x\mapsto f(t)e^{itx}$ est continue.

Exemple 21. Considérons la fonction Γ définie pour tout $z\in\mathbb{C}$ tel que $\mathfrak{Re}(z)>0$ par la formule $\Gamma(z)=\int\limits_0^\infty t^{z-1}e^{-t}dt$. Montrons la continuité de Γ. Pour ce faire vérifions les conditions $1-3$ du théorème. Les conditions 1 et 2 sont faciles à vérifier. Vérifions 3. Si $z=x+iy$ alors $|t^{z-1}e^{-t}|=t^{x-1}e^{-t}$. Donc, il suffit de trouver une fonction intégrable g telle que $t^{x-1}e^{-t}\le g(t)$ pour tous $t>0$ et $x\in[x_0,x_1]$ où $0<x_0<1<x_1<+\infty$. Soit d'abord $t\ge 1$. On a $0\le e^{-t}t^{x-1}\le e^{-t}t^{x_1-1}$. On sait que $\frac{t^{x_1-1}}{e^{\frac{t}{2}}}\underset{t\to\infty}{\to}0$. Donc il existe $C_1>0$ tel que $t^{x_1-1}\le C_1e^{\frac{t}{2}}$ pour tout $t\ge 1$. D'où, $e^{-t}t^{x_1-1}\le C_1e^{-\frac{t}{2}}$ et $e^{-t}t^{x-1}\le C_1e^{-\frac{t}{2}}$ si $t\ge 1$. Maintenant, soit $t\in]0,1[$. Alors $e^{-t}t^{x-1}\le t^{x_0-1}\Leftrightarrow e^{-t}\le t^{x_0-x}$ pour tout $x\in[x_0,x_1]$. Donc, $e^{-t}t^{x-1}\le C_1e^{-\frac{t}{2}}\mathbf{1}_{[1,+\infty[}(t)+t^{x_0-1}\mathbf{1}_{]0,1[}(t)$ et $g(t)=C_1e^{-\frac{t}{2}}\mathbf{1}_{[1,+\infty[}(t)+t^{x_0-1}\mathbf{1}_{]0,1[}(t)$ convient.

Rappel 7.6. Le théorème des accroissements finis affirme que si $\varphi:[a,b]\to\mathbb{R}$ est une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$ alors il existe $c\in]a,b[$ tel que $\varphi(b)-\varphi(a)=\varphi'(c)(b-a)$.

Théorème 7.7. *(Théorème de dérivabilité avec hypothèse de domination) Soit $f:U\times\Omega\to\mathbb{R}^m$ avec $U\subset\mathbb{R}^n$ un ouvert. On fixe un entier $1\le i\le n$ et on suppose :*

1. *pour tout $x\in U$, $t\mapsto f(x,t)$, est intégrable sur Ω ;*
2. *il existe $N\in\mathcal{T}$ avec $\mu(N^c)=0$, tel que pour tout $t\in N$, la fonction $x\mapsto f(x,t)$ admet une i-ème dérivée partielle sur U.*
3. *(hypothèse de domination) pour tout compact $K\subset U$, il existe une fonction $g_K\in L^1(\Omega,\mathcal{T},\mu)$ telle que*
$$\forall t\in N,\forall x\in K,\quad\left\|\frac{\partial f}{\partial x_i}(x,t)\right\|\le g_K(t).$$

Alors la fonction $x\mapsto F(x)=\int_\Omega f(x,t)d\mu(t)$ admet une i-ème dérivée partielle sur U, $\frac{\partial f}{\partial x_i}\in L^1(\Omega,\mathcal{T},\mu)$ et :

$$\frac{\partial F}{\partial x_i}(x)=\int_\Omega\frac{\partial f}{\partial x_i}(x,t)d\mu(t).$$

Démonstration. Il suffit de démontrer le théorème pour $m = 1$. On fixe $a = (a_j) \in U$. En plus, puisque l'intégrale de Lebesgue d'une fonction intégrable coïncide avec son intégrale sur $N \in \mathcal{T}$ si $\mu(N^c) = 0$, en remplaçant Ω par N on se ramène au cas $N = \Omega$.

L'ensemble U étant ouvert il existe $\varepsilon > 0$ tel que $K := \{x = (x_j) \in \mathbb{R}^n | x_j = a_j \text{ si } j \neq i \text{ et } x_i \in [a_i - \varepsilon, a_i - \varepsilon]\} \subset U$. Notons que K est compact. On définit $h : K \times \Omega \to \mathbb{R}$:

$$h(x,t) = \frac{f(x,t) - f(a,t)}{x_i - a_i} \quad \text{si} \quad x \neq a \qquad \text{et} \qquad h(a,t) = \frac{\partial f}{\partial x_i}(a,t) \text{ si } x = a.$$

Par l'hyphothèse 2, f est x_i-dérivable pour tout $t \in \Omega$. Le théorème des accroissements finis implique

$$|h(x,t)| \leq \sup_{u \in K} \left\| \frac{\partial f}{\partial x_i}(u,t) \right\|.$$

Par l'hypothèse de domination, $\left\| \frac{\partial f}{\partial x_i}(u,t) \right\| \leq g_K(t)$. D'où, $|h(x,t)| \leq g_K(t)$. Mais $h(x,t)$ est continue en a pour tout t (parce que $f(x,t)$ est x_i-dérivable en a pour tout t). Le théorème de convergence dominée implique que $\frac{\partial f}{\partial x_i}(a,t)$ est intégrable et $\int \frac{\partial f}{\partial x_i}(a,t) d\mu(t) = \lim_{x \to a} \int h(x,t) d\mu(t)$. Mais $\int h(x,t)dt = \frac{F(x)-F(a)}{x_i - a_i}$ si $x \neq a$. D'où,

$$\lim_{x \to a} \int h(x,t) d\mu(t) = \frac{\partial F}{\partial x_i}(a,t) = \int \frac{\partial f}{\partial x_i}(a,t) d\mu(t)$$

pour tout $a \in U$. □

Le théorème ci-dessus implique :

Corollaire 7.8. (Théorème de dérivation successive) Soit $f : U \times I \to \mathbb{R}^m$, avec $U \subset \mathbb{R}^n$ un ouvert et I un intervalle de $\mathbb{R}$, une fonction $\mathcal{C}^k$ ($k \in \mathbb{N} \cup \{+\infty\}$) par rapport au paramètre $x \in U$. Soit λ la mesure de Lebesgue sur I. On suppose qu'il existe $\phi_0, \phi_1, ..., \phi_k$ intégrables sur I telles que

$$\forall (i_1, ..., i_n), i_1 + ... + i_n = p \leq k, \forall x \in U, \forall t \in I, \left\| \frac{\partial^p f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x,t) \right\| \leq \phi_p(t).$$

Alors la fonction $x \mapsto F(x) = \int_\Omega f(x,t) d\lambda(t)$ est de classe $\mathcal{C}^k$ sur U et pour $p = i_1 + ... + i_n \leq k$:

$$\frac{\partial^p F}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x) = \int_I \frac{\partial^p f}{\partial x_1^{i_1} ... \partial x_n^{i_n}}(x,t) d\lambda(t).$$

Exemple 22. Soit $L^1(\mathbb{R}^+)$ l'espace de fonctions Lebesgue intégrables sur $\mathbb{R}^+ := \{t > 0 : t \in \mathbb{R}\}$. On définie *la transformation de Laplace* $(\mathcal{L}u)(x) = \int_0^\infty e^{-xt} u(t) dt$, où $u \in L^1(\mathbb{R}^+)$ et $x \in \mathbb{R}^+$. Montrons que $\mathcal{L}u$ est dérivable sur $]0, +\infty[$. Il suffit de vérifier les conditions 1-3 dans la formulation du Théorème 7.7 pour la fonction $f(x,t) = e^{-xt}u(t)$ et $x \in]x_0, +\infty[$ avec $x_0 > 0$. La condition 1 est satisfaite parce que $e^{-xt}u(t) \leq u(t)$ et $u(t)$ est intégrable. La condition 2 est satisfaite et $\frac{\partial f}{\partial x}(x,t) = -te^{-xt}u(t)$. Vérifions l'hypothèse de domination (la condition 3). On a que $| \frac{\partial f}{\partial x}(x,t) | \leq te^{-x_0 t}|u(t)|$ pour tout $x \in]x_0, +\infty[$. On sait que $\frac{t}{e^{x_0 t}} \underset{t \to \infty}{\to} 0$. Donc la fonction $t \mapsto \frac{t}{e^{x_0 t}}$ est bornée et par conséquent $g(t) = te^{-x_0 t}|u(t)|$ est intégrable, c.à.d. l'hypothèse de domination est satisfaite. Par Théorème 7.7

$$(\mathcal{L}u)'(x) = -\int_0^\infty te^{-xt} u(t) dt.$$

Exemple 23. Montrons que la fonction Γ est dérivable sur $]0,+\infty[$. Pour ça il suffit de montrer la dérivabilité de Γ sur tout intervalle $]x_0,x_1[$ avec $0<x_0<1<x_1<+\infty$. Donc, on suppose que $x\in]x_0,x_1[$ avec $0<x_0<1<x_1<+\infty$. Vérifions les conditions 1-3 dans la formulation du Théorème 7.7 pour la fonction $f(x,t)=e^{-t}t^{x-1}=e^{-t}e^{(\ln t)(x-1)}$. La fonction $t\mapsto f(x,t)$ est intégrable (voir Exemple 21). Elle est dérivable par rapport à x et $\frac{\partial f}{\partial x}(x,t)=e^{-t}(\ln(t))t^{x-1}$. Il reste à trouver $g_K(t)$. Soit $\varepsilon\in]0,x_0[$. Il existe $C_\varepsilon>0$ tel que $|\ln(t)|\leq C_\varepsilon t^{-\varepsilon}\mathbf{1}_{]0,1[}(t)+C_\varepsilon t^{\varepsilon}\mathbf{1}_{[1,+\infty[}(t)$. (On utilise que $\frac{s}{e^s}\underset{s\to+\infty}{\to}0$ et $\frac{\ln(s)}{s^\varepsilon}\underset{s\to+\infty}{\to}0$.) D'où $\frac{\partial f}{\partial x}(x,t)\leq C_\varepsilon(e^{-t}t^{x-\varepsilon-1}\mathbf{1}_{]0,1[}(t)+e^{-t}t^{x+\varepsilon-1}\mathbf{1}_{[1,+\infty[}(t))\leq C_\varepsilon(t^{x_0-\varepsilon-1}\mathbf{1}_{]0,1[}(t)+e^{-t}t^{x_1+\varepsilon-1}\mathbf{1}_{[1,+\infty[}(t))$. On a $\int_0^1 t^{x_0-\varepsilon-1}dt<\infty$ car $x_0-\varepsilon>0$ et $\int_1^{+\infty}\frac{t^{x_1+\varepsilon-1}}{e^t}dt<\infty$ car $\frac{t^{x_1+\varepsilon-1}}{e^{\frac{t}{2}}}\underset{t\to+\infty}{\to}0$ et $\int_{\frac{1}{2}}^{+\infty}\frac{dt}{e^{\frac{t}{2}}}<\infty$. Donc,

$$g_K(t)=g(t)=C_\varepsilon(t^{x_0-\varepsilon-1}\mathbf{1}_{]0,1[}(t)+e^{-t}t^{x_1+\varepsilon-1}\mathbf{1}_{[1,+\infty[}(t))$$

pour tout compact $K\subset\mathbb{R}^+$.

7.2 Mesure produit et théorèmes de Fubini

7.2.1 Mesure produit

Dans ce sous-section on introduit sur la tribu produit dans le cadre de deux espaces mesurés σ-finis. Commençons avec quelques rappels.

Définition 7.9. Soit $(\Omega,\mathcal{T},\mu)$ un espace mesuré. On dit que $(\Omega,\mathcal{T},\mu)$ est *σ-fini* s'il existe une suite de parties mesurables $(A_n)_{n\in\mathbb{N}}$ telle que $\mu(A_n)<+\infty$ pour tout n, et $\Omega=\bigcup\limits_n A_n$.

Cette hypothèse est par exemple vérifiée quand $\mu(\Omega)<+\infty$ (donc en particulier quand μ est une mesure de probabilité), quand $\Omega=\mathbb{N}$ muni de la mesure de comptage, ou quand $\Omega=\mathbb{R}^n$ muni de la mesure de Lebesgue.

Définition 7.10. ($\equiv$ *Définition* 5.20) Soient $(\Omega_i,\mathcal{T}_i), i=1,2,$ deux espaces mesurables. On appelle *ensemble élémentaire* de $\Omega=\Omega_1\times\Omega_2$ une réunion fini de *pavés* $A_1\times A_2$ avec $A_i\in\mathcal{T}_i, i=1,2.$ *La tribu produit* $\mathcal{T}_1\otimes\mathcal{T}_2$ sur Ω est la tribu engendrée par les ensembles élémentaires (donc, par les pavés).

Nous avons démontré que $\mathcal{B}(\mathbb{R}^n)=\underbrace{\mathcal{B}(\mathbb{R})\otimes\cdots\otimes\mathcal{B}(\mathbb{R})}_{n}$ (Corollaire 5.28). De sa part, en vertu de Théorème 5.58, la tribu $\mathcal{B}(\mathbb{R}^n)$ est munie de la mesure de Lebesque λ_n. Théorème 5.58 est un corollaire facile du théorème important suivant (admis) :

Théorème 7.11. *($\equiv$Théorème 5.56) Soit μ une fonction additive d'ensembles (c.à.d. $\mu(A\sqcup B)=\mu(A)+\mu(B)$), positive, définie sur un clan $\mathcal{C}$ de parties de Ω.*

1. *Si pour toute suite croissante (A_n) d'éléments de $\mathcal{C}$ de réunion $A\in\mathcal{C}$, $\lim\limits_{n\to\infty}\mu(A_n)=\mu(A)$, alors μ se prolonge à une mesure sur $(\Omega,\mathcal{T}(\mathcal{C}))$.*
2. *Si μ est σ-finie sur $\mathcal{C}$ (c.à.d. s'il existe une suite (B_n) d'éléments de $\mathcal{C}$ telle que $B_n\nearrow\Omega$ et $\mu(B_n)<\infty$ pour tous n) alors le prolongement est unique.*

On se servira de nouveau du Théorème 7.11 pour montrer le résultat suivant plus général définissant la mesure produit.

Théorème 7.12 (définissant la mesure produit). *Soient* $(\Omega_1, \mathcal{T}_1, \mu_1)$ *et* $(\Omega_2, \mathcal{T}_2, \mu_2)$ *deux espaces mesurés* σ*-finis. Alors il existe une unique mesure* ν *sur* $\mathcal{T}_1 \otimes \mathcal{T}_2$ *vérifiant*

$$\nu(A \times B) = \mu_1(A)\mu_2(B)$$

pour tout $A \in \mathcal{T}_1$ *et tout* $B \in \mathcal{T}_2$ *(avec la convention usuelle* $0.(+\infty) = 0$*). Cette mesure est notée* $\mu_1 \otimes \mu_2 = \nu$*, et est* σ*-finie.*

Démonstration. Soit $\mathcal{C}$ la famille de tous ensembles élémentaires de $\Omega_1 \times \Omega_2$. (Rappelons que chaque ensemble élémentaire est une union disjointe d'un nombre finie de pavés $A_1 \times A_2$ avec $A_i \in \mathcal{T}_i, i = 1, 2$.) En utilisant $(A_1 \times A_2)^c = (A_1^c \times A_2) \sqcup (A_1 \times A_2^c)$ on voit facilement que $\mathcal{C}$ est un clan. Pour chaque pavé $A_1 \times A_2$ on pose $\nu(A_1 \times A_2) = \mu(A_1) \cdot \mu(A_1)$. Puisque chaque ensemble élémentaire se represente comme union disjointe d'un nombre finie de pavés, ν se prolonge à une fonction additive sur $\mathcal{C}$. Par Théorème 7.11(1) ν se prolonge à une mesure notée $\mu_1 \otimes \mu_2$ sur l'espace mesurable $(\Omega \times \Omega, \mathcal{T}_1 \otimes \mathcal{T}_2)$. D'autre part, les espaces mesurés $(\Omega_1, \mathcal{T}_1, \mu_1)$ et $(\Omega_2, \mathcal{T}_2, \mu_2)$ sont σ-finis. Soient $X_i \nearrow \Omega_1$ et $Y_i \nearrow \Omega_2$ où $X_i \in \mathcal{T}_1$, $Y_i \in \mathcal{T}_2$, $\mu_1(X_i) < \infty$ et $\mu_2(Y_i) < \infty$ pour tous i. Alors $X_i \times Y_i \nearrow \Omega \times \Omega$ et $\mu_1 \otimes \mu_2(X_i \times Y_i) < \infty$. Par Théorème 7.11(2) le prolongement $\mu_1 \otimes \mu_2$ est unique. □

7.2.2 Théorème de Fubini-Tonelli et Fubini (admis)

La mesure produit $\mu_1 \otimes \mu_2$ étant définie à partir de μ_1 et μ_2, on s'attend à ce qu'il en soit de même de l'intégrale d'une fonction mesurable relativement à $\mu_1 \otimes \mu_2$.En effet, c'est effectivement le contenu des théorèmes de Fubini. On commence par le cas positif.

Théorème 7.13 (Fubini–Tonelli). *Soient* $(\Omega_1, \mathcal{T}_1, \mu_1)$ *et* $(\Omega_2, \mathcal{T}_2, \mu_2)$ *deux espaces mesurés* σ*-finis. Soit* $f : \Omega_1 \times \Omega_2 \to [0, +\infty]$ *une fonction* $\mathcal{T}_1 \otimes \mathcal{T}_2$*-mesurable. Alors :*

1. $y \mapsto f(x, y)$ *est une fonction mesurable (sur* $(\Omega_2, \mathcal{T}_2)$ *dans* $[0, +\infty]$ *) pour tout* $x \in \Omega_1$*, et* $x \mapsto \displaystyle\int_{\Omega_2} f(x, y)d\mu_2(y)$ *est une fonction mesurable (sur* $(\Omega_1, \mathcal{T}_1)$*).*
2. $x \mapsto f(x, y)$ *est une fonction mesurable (sur* $(\Omega_1, \mathcal{T}_1)$ *dans* $[0, +\infty]$*) pour tout* $y \in \Omega_2$*, et* $y \mapsto \displaystyle\int_{\Omega_1} f(x, y)d\mu_1(x)$ *est une fonction mesurable (sur* $(\Omega_2, \mathcal{T}_2)$*).*
3. *On a*

$$\int_{\Omega_1 \times \Omega_2} f(x, y)d(\mu_1 \otimes \mu_2)(x, y) = \int_{\Omega_1} \left(\int_{\Omega_2} f(x, y)d\mu_2(y) \right) d\mu_1(x) = \int_{\Omega_2} \left(\int_{\Omega_1} f(x, y)d\mu_1(x) \right) d\mu_2(y) \,.$$

Exercice 7.14. Calculer l'aire du disque unité $D = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

Comme dans le cas des fonctions définies sur $\mathbb{R}^n$, on en déduit facilement un théorème qui s'applique à toutes les fonctions intégrables (et pour vérifier qu'une fonction est intégrable, on peut commencer par appliquer le théorème de Fubini–Tonelli à $|f|$).

Théorème 7.15 (Fubini). *Soient* $(\Omega_1, \mathcal{T}_1, \mu_1)$ *et* $(\Omega_2, \mathcal{T}_2, \mu_2)$ *deux espaces mesurés* σ*-finis. Soit* $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ *une fonction intégrable. Alors :*

1. $y \mapsto f(x, y)$ *est une fonction intégrable (sur* Ω_2*) pour presque tout* $x \in \Omega_1$*, et* $x \mapsto \displaystyle\int_{\Omega_2} f(x, y)d\mu_2(y)$ *est une fonction intégrable (sur* Ω_1*).*
2. $x \mapsto f(x, y)$ *est une fonction intégrable (sur* Ω_1*) pour presque tout* $y \in \Omega_2$*, et* $y \mapsto \displaystyle\int_{\Omega_1} f(x, y)d\mu_1(x)$ *est une fonction intégrable (sur* Ω_2*)*

3. On a

$$\int_{\Omega_1\times\Omega_2} f(x,y)d(\mu_1\otimes\mu_2)(x,y) = \int_{\Omega_1}\Big(\int_{\Omega_2} f(x,y)d\mu_2(y)\Big)\,d\mu_1(x) = \int_{\Omega_2}\Big(\int_{\Omega_1} f(x,y)d\mu_1(x)\Big)\,d\mu_2(y)\,.$$

Le théorème de Fubini est un utile puissant pour intervertir l'ordre d'intégration sur un espace produit. Cependant, comme l'exemple suivant le montre, avant d'utiliser ce théorème il faut bien vérifier que la fonction $f\colon \Omega_1\times\Omega_2\to\mathbb{R}$ est *intégrable*.

Exemple 24. Soit $f\,:]0,1[^2\to\mathbb{R}, (x,y)\mapsto \frac{x^2-y^2}{(x^2+y^2)^2}$. La fonction est continue donc mesurable. On a

$$\int_{]0,1[} f(x,y)d\lambda_1(y) = \Big[\frac{y}{x^2+y^2}\Big]_0^1 = \frac{1}{1+x^2}$$

et

$$\int_{]0,1[}\Big(\int_{]0,1[} f(x,y)d\lambda_1(y)\Big)d\lambda_1(x) = \big[\arctan x\big]_0^1 = \frac{\pi}{4}.$$

D'autre part, comme $f(x,y)=-f(y,x)$, on a

$$\int_{]0,1[}\Big(\int_{]0,1[} f(x,y)d\lambda_1(x)\Big)d\lambda_1(y) = -\frac{\pi}{4}.$$

Donc,

$$\int_{]0,1[}\Big(\int_{]0,1[} f(x,y)d\lambda_1(y)\Big)d\lambda_1(x) \neq \int_{]0,1[}\Big(\int_{]0,1[} f(x,y)d\lambda_1(x)\Big)d\lambda_1(y).$$

La raison est que f n'est pas intégrable. En effet, d'après le théorème de Fubini-Tonelli

$$\begin{aligned}\int_{]0,1[^2} f_+(x,y)d(\lambda_1\otimes\lambda_1)(x,y) &= \int_{]0,1[}\Big(\int_{]0,x[}\frac{x^2-y^2}{(x^2+y^2)^2}d\lambda_1(y)\Big)d\lambda_1(x)\\ &= \int_{]0,1[}\Big[\frac{y}{x^2+y^2}\Big]_0^x d\lambda_1(x) = \int_{]0,1[}\frac{d\lambda_1(x)}{2x} = +\infty.\end{aligned}$$

Exercice 7.16. Soit f,g des fonctions mesurables positives sur $\mathbb{R}$, on définit la *convolution* de f,g par :

$$f*g(x) = \int_{\mathbb{R}} f(x-y)g(y)d\lambda(y)\in[0,\infty].$$

On rappelle que

$$||f||_1 = \int_{\mathbb{R}}|f(x)|d\lambda(x).$$

1. Montrer que $f*g$ est mesurable et que
$$||f*g||_1 = ||f||_1||g||_1.$$
2. Montrer que la définition de $f*g$ s'étend pour presque tout x au $f,g\in L^1(\mathbb{R},d\lambda)$ et que $f*g\in L^1(\mathbb{R},d\lambda)$.
3. Montrer que pour f,g,h toutes mesurables positives ou toutes intégrables, alors
$$f*(g*h) = (f*g)*h.$$

7.3 Théorème de changement de variables

En pratique, pour calculer une intégrale multiple, on est souvent amené à faire un changement de variables pour se ramener à un domaine plus simple sur lequel appliquer le théorème de Fubini. On énonce le théorème dans le cadre le plus courant où les fonctions que l'on peut utiliser pour faire un changement de variables sont les *difféomorphismes de classe $\mathcal{C}^1$*. On commence notre étude avec le cas le plus simple de transformations affines linéaires.

7.3.1 Cas linéaire affine

Une application $\phi : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Ax + b$, où $A \in M_n(\mathbb{R})$ et $b \in \mathbb{R}^n$ s'appèle application affine (linéaire). Ici les éléments de $\mathbb{R}^n$ sont des colonnes (c.à.d. des matrices $n \times 1$) et Ax est le produit de la matrice A avec la matrice x. L'ensemble des matrices inversibles est noté par $GL_n(\mathbb{R})$. L'application ϕ est un isomorphisme de $\mathbb{R}^n$ si et seulement si $A \in GL_n(\mathbb{R})$. Dans ce cas $\phi^{-1}(y) = A^{-1}y - A^{-1}b$. (En exercice.)

La caractérisation suivante de la mesure de Lebesgue est essentielle pour la suite.

Théorème 7.17. *(admis) La mesure de Lebesgue sur R^n est invariante par translation, au sens où pour tout $M \in \mathcal{B}(\mathbb{R}^n)$ et tout $x \in \mathbb{R}^n$ on a $\lambda_n(x+M) = \lambda_n(M)$ avec $x + M := \{x + a, a \in M\}$. Inversement, si μ est une mesure sur $(R^n, \mathcal{B}(\mathbb{R}^n))$ finie sur les parties bornées et invariante par translation, alors il existe une constante $c \geq 0$ telle que $\mu = c\lambda_n$.*

Le CM du 10/11/2021

Exercice 7.18. Montrer la première partie du théorème que $\lambda_n(x + M) = \lambda_n(M)$ pour tous $M \in \mathcal{B}(\mathbb{R}^n)$ et $x \in \mathbb{R}^n$. (Indication : on pourra utiliser Théorème 7.11.)

Lemme 7.19. *Soit $\phi : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Ax + b$, où $A \in M_n(\mathbb{R})$ et $b \in \mathbb{R}^n$, une application affine. Alors pour tout borélien M de $\mathbb{R}^d$, on a :*

$$\lambda_n(\phi(M)) = |det(A)|\lambda_n(M).$$

Ainsi, si $det(A) \neq 0$, la mesure image de λ_n par ϕ est $\phi_*\lambda_n := \lambda_n \circ \phi^{-1} = |det(A)|^{-1}\lambda_n$.

Exercice 7.20. Si A n'est pas inversible montrer que $\lambda(\phi(\mathbb{R}^n)) = 0$. (Indication : on pourra montrer que $\phi(\mathbb{R}^n)$ est inclus dans un hyperplan affine, i.e. inclu dans $a + L$ où $a \in \mathbb{R}^n$ et L est un s.e.v. de dimension $n - 1$.)

Démonstration du lemme. Soit $\mu = \lambda_n \circ \phi$. Alors μ est une mesure sur $(R^n, \mathcal{B}(\mathbb{R}^n))$ finie sur les parties bornées et invariante par translation. Par Théorème 7.17, $\mu = c\lambda_n$ où $c \geq 0$. Il faut montrer que $c = |\det(A)|$. Comme λ_n est invariante par translation, on se ramène au cas $b = 0$.

Si $\lambda_n(M) > 0$ alors $c = \frac{\lambda_n(\phi(M))}{\lambda_n(M)}$. Donc, il suffit de calculer $\lambda_n(\phi(M))$ pour un M tel que $\lambda_n(M) > 0$. Considérons quelques cas particuliers pour A. Soit A orthogonale. Si $v_1, \cdots, v_n$ est une base orthogonale de $\mathbb{R}^n$ alors $\phi(v_1), \cdots, \phi(v_n)$ est une base orthogonale de $\mathbb{R}^n$ et $\| v_i \| = \| \phi(v_i) \|$ pour tout i. Soit $M = \{\sum_i \alpha_i v_i : 0 \leq \alpha_i \leq 1\}$. Alors M est un parallélépipède droit dans $\mathbb{R}^n$ de volume $\lambda_n(M) = \| v_1 \| \cdots \| v_n \|$. Notons que $\lambda_n(M) = \lambda_n(\phi(M))$. Mais $|det(A)| = 1$. Donc, si A est une matrice orthogonale alors $c = |det(A)|$.

Soit A diagonale de valeurs propres $d_1, \cdots, d_n$. On considère le parallélépipède droit $M_0 = \{\sum_i \alpha_i e_i : 0 \leq \alpha_i \leq 1\}$ où $e_1, \cdots, e_n$ est la base orthonormée canonique de $\mathbb{R}^n$. Alors $\lambda_n(M_0) = 1$ et $\lambda_n(\phi(M_0)) = |d_1 \cdots d_n| = |det(A)| \cdot \lambda_n(M_0)$. Donc, si A est une matrice diagonale alors $c = |det(A)|$.

Maintenant soit $A \in GL_n(\mathbb{R})$ quelconque. Par la décomposition polaire, $A = O \cdot S$ où O est une matrice orthogonale et S est une matrice symétrique définie positive. On sait que S se diagonalise dans une base orthonormée, i.e. $S = O' \cdot D \cdot O'^{-1}$ où O' est une matrice orthogonale et D est une matrice diagonale. Donc, $A = O_1 \cdot D \cdot O_2$ où O_1 et O_2 sont des matrices orthogonales. Avec M_0 comme ci-dessus,

$$\lambda_n(\phi(M_0)) = \lambda_n(O_1(DO_2(M_0))) = \lambda_n(D(O_2(M_0))) = |\det(D)| \cdot \lambda_n(O_2(M_0)) = |\det(A)| \cdot \lambda_n(M_0).$$

(On utilise que $|\det A| = |\det D|$.) Donc, $c = |\det(A)|$. $\square$

Remarque 7.21. Dans la preuve ci-dessus nous avons démontré que chaque $A \in GL_n(\mathbb{R})$ s'écrit sous la forme $A = O_1 \cdot D \cdot O_2$ où O_1 et O_2 sont des matrices orthogonales et D est une matrice diagonale. La décomposition $A = O_1 \cdot D \cdot O_2$ est très importante et s'appèle *décomposition de Cartan.*

Théorème 7.22 (Théorème de changement de variables linéaires)**.** *Soient $\phi : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Ax + b$, un isomorphisme de $\mathbb{R}^n$, U un ouvert de $\mathbb{R}^n$ et $V = \phi(U)$. Si la fonction $f : V \to \mathbb{R}$ est intégrable alors $f \circ \phi : U \to \mathbb{R}$ est intégrable et*

$$\int_V f(y) d\lambda_n(y) = \int_U (f \circ \phi)(x) |\det(A)| d\lambda_n(x).$$

Démonstration. L'application ϕ étant un isomorphisme de $\mathbb{R}^n$, V est un ouvert de $\mathbb{R}^n$ et $\det(A) \neq 0$. L'application $(U, \mathcal{B}(U), \lambda_n) \to (V, \mathcal{B}(V)), x \mapsto \phi(x)$, est une application mesurable. Par le théorème de transfert (Théorème 6.27) :

$$\int_U (f \circ \phi)(x) d\lambda_n(x) = \int_V f(y) d(\phi_* \lambda_n)(y).$$

D'après Lemme 7.19, $\phi_* \lambda_n = |det(A)|^{-1} \lambda_n$ i.e. $\phi_* \lambda_n$ est une mesure à densité $|det(A)|^{-1}$ par rapport à λ_n (voir Proposition 6.10). Par Proposition 6.17

$$\int_V f(y) d(\phi_* \lambda_n)(y) = |det(A)|^{-1} \int_V f(y) d\lambda_n(y).$$

Donc,

$$\int_V f(y) d\lambda_n(y) = |\det(A)| \int_U (f \circ \phi)(x) d\lambda_n(x) = \int_U (f \circ \phi)(x) |\det(A)| d\lambda_n(x).$$

$\square$

Exemple 25. Soit $\phi : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x + y, x - y)$. C'est un isomorphisme linéaire de réciproque $\phi^{-1} : \mathbb{R}^2 \to \mathbb{R}^2, (s, t) \mapsto (\frac{s+t}{2}, \frac{s-t}{2})$ et $|\det(\phi)| = 2$. Soit $U =]0, 1[^2$ et $V = \phi(U)$. Alors $V = \{(s, t) : 0 < s + t < 2, -1 < s - t < 1\}$. Soit $f : V \to \mathbb{R}$ une fonction intégrable. Alors en utilisant Théorème 7.22 d'abord et le théorème de Fubini après, on obtient :

$$\int_V f(s, t) d\lambda_2(s, t) = \int_U (f \circ \phi)(x, y) |\det(\phi)| d\lambda_2(x, y) = 2 \int_{]0,1[} \Big(\int_{]0,1[} (f \circ \phi)(x, y) d\lambda_1(x) \Big) d\lambda_1(y).$$

En général, le dernier intégrale est beaucoup plus facile à calculer. (Notons qu'on utilise souvent les notations $d\lambda_2(x, y) := dxdy$ et $d\lambda_1(x) := dx$.) Remarquons que $\lambda_2(\overline{V} \setminus V) = 0$. Par conséquent, $\int_{\overline{V}} f(s, t) dsdt = \int_V f(s, t) dsdt$.

7.3.2 Rappel (de L2) sur les difféomorphismes

Si E, F sont des espaces vectoriels on note par $L(E, F)$ l'espace vectoriel des applications linéaires de L dans F. Comme d'habitude, $\| \cdot \|$ est la norme standarde sur $\mathbb{R}^n$ (ou $\mathbb{R}^p$).

Définition 7.23. Si U est un ouvert de $\mathbb{R}^n$, alors $f : U \to \mathbb{R}^p$ est *différentiable en* $x \in U$ s'il existe un $T \in L(\mathbb{R}^n, \mathbb{R}^p)$, notée $df(x)$, telle que

$$||f(x + h) - f(x) - df(x)(h)|| = \circ(||h||), \text{ si } ||h|| \to 0.$$

La fonction f est C^1-*différentiable* (ou *continuement différentiable*) sur U si f est différentiable en tout $x \in U$ et $df : U \to L(\mathbb{R}^n, \mathbb{R}^p), x \mapsto df(x)$, est continue. On note aussi $D_h f(x) := df(x)(h)$.

Définition 7.24. Soient $U \subset \mathbb{R}^n$ et $V \subset \mathbb{R}^p$ des ouverts. Soit $f : U \to V$ une application différentiable. Alors f est un *difféomorphisme* si f est bijective et f^{-1} est différentiable.

On dit que f est un $\mathcal{C}^1$-*difféomorphisme* si de plus f et f^{-1} sont de classe $\mathcal{C}^1$.

Proposition 7.25. *Soit $f : U \to V$ un difféomorphisme. Alors $\forall x \in U$, $df(x) : \mathbb{R}^n \to \mathbb{R}^p$ est un isomorphisme linéaire (en particulier nécessairement $n = p$) et on a :*

$$(df(x))^{-1} = df^{-1}(f(x)).$$

Remarque 7.26.
1. La proposition montre que la dimension est invariante par difféomorphisme. De même des ouverts de $\mathbb{R}^n$ et $\mathbb{R}^p$ ne peuvent être homéomorphes que si $n = p$ mais c'est beaucoup plus dur (Théorème d'invariance du domaine de Brouwer). Par contre, il existe des applications continues surjectives de $[0, 1]$ dans $[0, 1]^2$.
2. Le théorème d'inversion locale va donner des conditions pour la réciproque de la proposition précédente.

Démonstration. Comme $f^{-1} \circ f(y) = y$, en différenciant $f^{-1} \circ f$ par le théorème des fonctions composées en x, on obtient : $df^{-1}(f(x)) \circ df(x) = id$.

De même en différenciant $f \circ f^{-1}(y) = y$ en $y = f(x)$ on obtient : $df(x) \circ df^{-1}(y) = Id$. Donc $df(x)$ et $df^{-1}(f(x))$ sont inverses l'une de l'autre, ce qui conclut. □

Définition 7.27. Soient $U \subset R^n$ un ouvert et $f : U \to \mathbb{R}^p$ où $f(x) = (f_1(x), ..., f_p(x))$, une application différentiable. La matrice de l'application linéaire $df(x)$ dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ est appelée, *matrice jacobienne* de f et notée $J(f)(x)$:

$$(J(f)(x))_{ij} = (\frac{\partial f_i}{\partial x_j}(x)).$$

Remarque 7.28. Le théorème de dérivation des fonctions composées donne donc :

$$J(g \circ f)(x) = J(g)(f(x))J(f)(x),$$

et le résultat pour les inverses de la proposition précédente s'écrit :

$$[J(f)(x)]^{-1} = J(f^{-1})(y),$$

où $y = f(x)$.

Le théorème suivant (admis) avec $k = 1$ permettra de vérifier l'hypothèse du théorème de changement de variables ci-dessous (Théorème 7.31).

Théorème 7.29. ***(d'inversion globale)*** *Soit $f : U \to \mathbb{R}^n$ une application de classe $\mathcal{C}^1$ injective et telle que pour tout $x \in U$, $df(x) : \mathbb{R}^n \to \mathbb{R}^n$ est un isomorphisme linéaire. Alors $f(U)$ est un ouvert de $\mathbb{R}^n$ et $f : U \to f(U)$ est un $\mathcal{C}^1$-difféomorphisme.*

Remarque 7.30. $df(x)$ est un isomorphisme si et seulement si $det((Jf)(x)) \neq 0$.

7.3.3 Cas général (admis)

Nous pouvons maintenant énoncer le théorème de changement de variables.

Théorème 7.31 (Théorème de changement de variables)**.** *Soient U, V deux ouverts de $\mathbb{R}^n$, et $\varphi \colon U \to V$ un difféomorphisme de classe $\mathcal{C}^1$. Rappelons qu'on note λ_n la mesure de Lebesgue sur $\mathbb{R}^n$. Alors on a :*

1. *Pour toute partie B borélienne de U, $\lambda_n(\varphi(B)) = \int_B |\det(J\varphi(x))| d\lambda_n(x)$.*
2. *Si $f \colon V \to [0, +\infty]$ est borélienne, alors*

$$\int_V f(y) d\lambda_n(y) = \int_U f \circ \varphi(x) |\det(J\varphi(x))| d\lambda_n(x) \ .$$

3. Si $f\colon V \to \mathbb{R}$ est intégrable, alors $x \mapsto f \circ \varphi(x)|\det(J\varphi(x))|$ est intégrable sur U et on a

$$\int_V f(y)d\lambda_n(y) = \int_U f \circ \varphi(x)|\det(J\varphi(x))|d\lambda_n(x) \ .$$

Remarque 7.32. Théorème 7.22 est un cas très particulier du Théorème 7.31.

Exemple 26 (changement de variables en polaires)**.** On considère l'application $\phi : U =]0,+\infty[\times]0,2\pi[\to \mathbb{R}^2$ définie par $\phi(r,\theta) = (r\cos\theta, r\sin\theta)$.

Alors, la matrice jacobienne de ϕ est $\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$, de déterminant $r > 0$.

De plus, ϕ est injective et $\phi(U) = \mathbb{R}^2 \setminus [0,+\infty[= V$.

D'après Théorème 7.29, ϕ est un $\mathcal{C}^1$-difféomorphisme de U sur V. Comme $\lambda_2(\mathbb{R}^2 \setminus V) = 0$, c'est-à-dire $\mathbb{R}^2 \setminus V$ est négligeable, il n'est pas gênant que ϕ ne soit pas un difféomorphisme de U sur $\mathbb{R}^2$ tout entier. Donc, si $f : \mathbb{R}^2 \to \mathbb{R}$ est intégrable, on a

$$\int_{\mathbb{R}^2} f(x,y)dxdy = \int_0^{+\infty}\int_{-\pi}^{\pi} f(r\cos\theta, r\sin\theta)rdrd\theta = \int_0^{+\infty}\int_0^{2\pi} f(r\cos\theta, r\sin\theta)rdrd\theta.$$

Exemple 27. Calculons

$$I = \int_D (x+y)^2dxdy, \text{ où } D = \{(x,y)\colon x^2+y^2 < 1\}.$$

En utilisant le théorème de changement de variables avec les coordonnées polaires (et le théorème de Fubini), on obtient $\phi^{-1}(D) =]0,1[\times]0,2\pi[$ et

$$\begin{aligned} I &= \int_{\phi^{-1}(D)} (r\cos\theta + r\sin\theta)^2 rdrd\theta \\ &= \int_0^1 \left(\int_0^{2\pi} r^3(\cos^2\theta + \sin^2\theta + 2\cos\theta\sin\theta)d\theta\right) dr \\ &= \int_0^1 r^3 \left(\int_0^{2\pi} (1+\sin 2\theta)d\theta\right) dr \\ &= \int_0^1 2\pi r^3 dr \\ &= \frac{\pi}{2} \,. \end{aligned}$$

Exemple 28. Montrons que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Rappelons que $\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}dt$. Donc, $\Gamma(\frac{1}{2}) = \int_0^{\infty} t^{-1/2}e^{-t}dt$.

En faisant le changement de variable $u^2 = t$, $dt = 2udu$, et en utilisant le théorème de Fubini-Tonelli on obtient

$$\Gamma(\frac{1}{2}) = 2\int_0^{+\infty} e^{-u^2}du = \int_{-\infty}^{+\infty} e^{-u^2}du = \left(\int_{-\infty}^{+\infty} e^{-x^2}dx \int_{-\infty}^{+\infty} e^{-y^2}dy\right)^{1/2} = \left(\int_{\mathbb{R}^2} e^{-x^2-y^2}d\lambda(x,y)\right)^{1/2}.$$

Puis en utilisant le changement de variables en polaires, on conclut

$$\Gamma(\frac{1}{2}) = \left(\int_0^{2\pi} d\theta \int_0^{+\infty} e^{-r^2}2r/2dr\right)^{1/2} = \sqrt{\pi}.$$

Chapitre 8

Convexité

8.1 Ensemles convexes

Dans ce chapitre E est un espace vectoriel normé (e.v.n.) sur $\mathbb{R}$ de dimension finie ou infinie.

Définition 8.1. Soit $x, y \in E$. Alors

$$[\,x, y\,] = \{(1-\lambda)x + \lambda y : 0 \leq \lambda \leq 1\}$$

est un segment d'extrémités x et y.

Les intervales finies et fermés sur $\mathbb{R}$ sont des segements.

Définition 8.2. Un ensemble $C \subset E$ est dit *convexe* si $\forall\ x,\ y, [x, y] \subset C$. Par convention $C = \emptyset$ est convexe aussi.

Exercice 8.3. Si E est un espace vectoriel normé (e.v.n.) alors les boules (ouvertes et fermées) sont des connexes.

Le resultat suivant est laissé en exercice.

Proposition 8.4.
1. *Soit E un e.v.n. et C un sous-ensemble convexe de E. L'adhérence $\overline{C}$ et l'intérieur $\overset{\circ}{C}$ de C sont convexes.*
2. *L'intersection (finie ou infinie) d'ensembles convexes est convexe.*
3. *Si $C_1 \subset E_1$ et $C_2 \subset E_2$ (où E_1 et E_2 sont des espaces vectoriels sur $\mathbb{R}$) alors $C_1 \times C_2$ est convexe dans $E_1 \times E_2$.*

Rappel 8.5. Si $A, B \subset E$, $C \subset \mathbb{R}$ et $x \in E$ alors $A + B = \{a + b : a \in A, b \in B\}$, $CA = \{ca : c \in C, a \in A\}$, $A - x = \{a - x : a \in A\}$ et $x + A = \{a + x : a \in A\}$. Aussi, $\mathbb{R}_+^* = \{x > 0 : x \in \mathbb{R}\}$.

Définition 8.6. Le *cône tangent du convexe* $S \subset E$ au point $x \in S$ est

$$T_S(x) = \overline{\{\frac{u-x}{s} : u \in S, s \in \mathbb{R}_+^*\}} = \overline{\mathbb{R}_+(S-x)}.$$

Soit E un espace euclidien (par exemple, $E = \mathbb{R}^n$ avec le produit scalaire $< \cdot >$ habituel). Le *cône normal du convexe* $S \subset E$ au point $x \in S$ est le polaire de $T_S(x)$, c.à.d. le cône convexe fermé définit par

$$N_S(x) = \{y \in E : \forall u \in S, < y, u - x > \leq 0\} = \{y \in E : \forall v \in T_S(x), < y, v > \leq 0\}.$$

(Dessins en CM.)

La proposition suivante permet de se ramener à des cas plus simples.

Proposition 8.7. *Soient A, B des convexes de E.*

1. *Si $A \subset B$ et $x \in A$ alors $T_A(x) \subset T_B(x)$.*
2. *Si $a \in \mathring{A}$, $T_A(a) = E$.*
3. *Pour $x \in A$, $A \subset x + T_A(x)$.*

(Dessins en CM.)

Démonstration. (1) $\mathbb{R}_+(A-x) \subset \mathbb{R}_+(B-x)$. Par la monotomie de l'adhérence on obtient $T_A(x) \subset T_B(x)$.

(2) Il existe un $r > 0$ tel que $B(a,r) \subset A$. Par (1), $T_A(a) \supset T_{B(a,r)}(a) \supset \mathbb{R}_+(B(a,r) - a) = \mathbb{R}_+B(0,r) = E$.

(3) On a $x + T_A(x) = x + \overline{\mathbb{R}_+(A - x)} \supset x + (A - x) = A$.

□

8.2 Fonctions convexes

On commence par la définition générale d'une fonction convexe f. Nous verrons plus tard des caractérisations plus simples dans le cas où f est différentiable ou doublement différentiable.

Soit $f : E \to \mathbb{R}$. Les propriétés que l'on considère vont être déterminées par l'ensemble des valeurs au dessus du graphe de f appelé **épigraphe de** f :

$$\mathrm{Epi}(f) = \{(x, \lambda) \in E \times \mathbb{R} : f(x) \leq \lambda\}.$$

(Dessin en CM.)

Définition 8.8. Soient C un ensemble convexe et $f : C \to \mathbb{R}$.

1. La fonction f est **convexe** si pour tous $x, y \in C$ et tout $\lambda \in [0,1]$
$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$
2. La fonction f est **strictement convexe** si pour tous $x, y \in C, x \neq y$, et tout $\lambda \in]0,1[$
$$f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y).$$
3. La fonction f est **concave** (respectivement, **strictement concave**) si $-f$ est convexe (respectivement, strictement convexe).

(Dessins en CM.)

Exemples 2. La fonction $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$, est strictement convexe. La fonction *affine* $f : \mathbb{R}^n \to \mathbb{R}, (x_1, \cdots, x_n) \mapsto a_1x_1 + \cdots + a_nx_n + b$, est à la fois convexe et concave. Une norme sur E est convexe. (En exercice.)

Proposition 8.9. *Soit C un ensemble convexe.*

1. *$f : C \to \mathbb{R}$ est convexe si et seulement si Epi(f) est convexe ;*
2. *Si $\mu > 0$, $\nu > 0$ et f et g sont des fonctions convexes (respectivement, une des fonctions f ou g est strictement convexe) sur C alors $\mu f + \nu g$ est convexe (respectivement, strictement convexe) sur C.*
3. *Si $f_i, i \in I$, est une famille de fonctions convexes sur C alors $f(x) = \sup\limits_{i \in I} f_i(x)$ est convexe.*
4. *Si f est strictement convexe alors f a au plus un minimum sur C.*

Démonstration. 1. Soit $(x, t_1), (y, t_2) \in \mathrm{Epi}(f)$. Donc, $f(x) \leq t_1$ et $f(y) \leq t_2$ et $(1-\lambda)f(x) + \lambda f(y) \leq (1-\lambda)t_1 + \lambda t_2$ pour tout $\lambda \in [0,1]$. Si f est convexe alors $f((1-\lambda)x+\lambda y) \leq (1-\lambda)f(x) + \lambda f(y) \leq (1-\lambda)t_1 + \lambda t_2$ et $(1-\lambda)(x, t_1) + \lambda(y, t_2) \in \mathrm{Epi}(f)$ qui montre la convexité de $\mathrm{Epi}(f)$.

Reciproquement, soit $\mathrm{Epi}(f)$ convexe. Si $x, y \in C$ alors $(x, f(x)), (y, f(y)) \in \mathrm{Epi}(f)$ et, par la convexité de $\mathrm{Epi}(f)$, $(1-\lambda)(x, f(x)) + \lambda(y, f(y)) = ((1-\lambda)x + \lambda y, (1-\lambda)f(x) + \lambda f(y)) \in \mathrm{Epi}(f) \Leftrightarrow f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$.

2. Les fonctions μf et νg sont evidement convexes. Donc,

$$\begin{aligned}
&(\mu f + \nu g)((1-\lambda)x + \lambda y) \leq \\
&((1-\lambda)\mu f(x) + \lambda \mu f(y)) + ((1-\lambda)\nu g(x) + \lambda \nu g(y)) = \\
&(1-\lambda)(\mu f + \nu g)(x) + \lambda(\mu f + \nu g)(y)
\end{aligned}$$

qui prouve la convexité. La preuve de la convexité stricte est analogique.

3. Vient de la partie 1 de la proposition, de l'égalité $\bigcap_{i\in I} \mathrm{Epi}(f_i) = \mathrm{Epi}(f)$ et de la stabilité des ensembles convexes par intersection.

4. Si $x \neq y$ sont deux points atteignant le minima alors $f(\frac{x+y}{2}) < \frac{f(x)+f(y)}{2}$ contredisant la minimalité. $\square$

La proposition suivante caractérise les fonctions convexes en terme d'accroissements.

Proposition 8.10. *Soit $f : C \to \mathbb{R}$ où C est convexe. Alors f est convexe si et seulement si pour tous $x, y \in C$ la fonction de $]0,1]$ dans $\mathbb{R}$ définie par $\lambda \mapsto \frac{f(x+\lambda(y-x))-f(x)}{\lambda} (= \frac{f((1-\lambda)x+\lambda y)-f(x)}{\lambda})$ est croissante.*

Démonstration. $\Rightarrow$) Soit $0 < \lambda \leq \lambda' \leq 1$. La convexité de f imlique

$$\begin{aligned}
&f(x + \lambda(y-x)) = f(\frac{\lambda}{\lambda'}(x + \lambda'(y-x)) + (1 - \frac{\lambda}{\lambda'})x) \leq \\
&\frac{\lambda}{\lambda'} f(x + \lambda'(y-x)) + (1 - \frac{\lambda}{\lambda'})f(x).
\end{aligned}$$

Après la soustraction de $f(x)$ de deux côtés de l'inégalité ci-dessus et sa division par λ, on obtient

$$\frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \leq \frac{f(x+\lambda'(y-x)) - f(x)}{\lambda'}.$$

La nécessité est démontrée.

$\Leftarrow$) Soit $0 < \lambda \leq \lambda' = 1$. Par l'hyphothèse,

$$\frac{f(x+\lambda(y-x)) - f(x)}{\lambda} \leq f(y) - f(x).$$

L'inégalité ci-dessus est équivalente à la suivante :

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y).$$

Donc, f est convexe. $\square$

Corollaire 8.11. Soient $f : C \to \mathbb{R}$ une fonction convexe et $x \in C$. Soit encore $h \in E, h \neq 0$, tel que $x + h \in C$. Alors *la dérivée directionnelle* existe au sens où la limite suivante existe et vaut :

$$D'_h f(x) := \lim_{t \to 0^+} \frac{f(x+th) - f(x)}{t} = \inf_{1 \geq t > 0} \frac{f(x+th) - f(x)}{t}.$$

Démonstration. Par Proposition 8.10, $g(t) = \frac{f(x+th)-f(x)}{t}$ est une fonction croissante donc admet une limite pour $t \to 0^+$ qui coïncide avec l'infimum. (Dessin en CM.) $\square$

Le cours du 17/11/2021

8.3 Fonctions convexes sur $\mathbb{R}$

Dans cette section I est un intervalle de $\mathbb{R}$.

Les inégalités de la propossition suivantes (appelées des *inégalités de pentes*) caractérisent les fonctions convexes sur I.

Proposition 8.12. *Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si*

$$\forall a, b, c \in I, a < b < c \Rightarrow \frac{f(b) - f(a)}{b - a} \le \frac{f(c) - f(a)}{c - a} \le \frac{f(c) - f(b)}{c - b}.$$

Démonstration. $\Rightarrow$) Si $x, y \in I$ et $0 < \lambda \le 1$, la Proposition 8.10 implique

$$\frac{f(x + \lambda(y - x)) - f(x)}{\lambda} \le f(y) - f(x).$$

Pour obtenir la prmière inégalité de la formulation on pose $x = a$, $y = c$ et $\lambda = \frac{b-a}{c-a}$ et pour la deuxième on pose $x = c$, $y = a$ et $\lambda = \frac{c-b}{c-a}$.

$\Leftarrow$) Il suffit de montrer que si $x, y \in I$, $x < y$ et $0 < \lambda < 1$ alors $f((1 - \lambda)x + \lambda y) \le (1 - \lambda)f(x) + \lambda f(y)$. (L'inégalité est triviale si $x = y$ ou $\lambda = 0$ ou $\lambda = 1$.)

On pose $a = x$, $c = y$ et $b = (1 - \lambda)x + \lambda y$. Alors, $a < b < c$ et

$$\frac{f(b) - f(a)}{b - a} = \frac{f((1 - \lambda)x + \lambda y) - f(x)}{\lambda(y - x)} \le \frac{f(c) - f(a)}{c - a} = \frac{f(y) - f(x)}{y - x}.$$

D'où, $f((1 - \lambda)x + \lambda y) \le (1 - \lambda)f(x) + \lambda f(y)$. □

Rappelons que pour une fonction $f : I \to \mathbb{R}$ et $a \in I$ *le taux d'accroissement de f en a*, noté $\Delta_a f$, est définie par $\Delta_a f(x) = \frac{f(x)-f(a)}{x-a}$ pour tout $x \in I \setminus \{a\}$. Ci-dessous une autre caractérisation de la convexité en termes de $\Delta_a f$.

Proposition 8.13. *Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si pour tout $a \in I$ la fonction $\Delta_a f$ est croissante sur $I \setminus \{a\}$ c.à.d. $\Delta_a f(x) \le \Delta_a f(y)$ si x, y et $x < y$. En plus, $\Delta_a f(x) < \Delta_a f(y)$ si $x < y$ et f est strictement convexe.*

Démonstration. La proposition est un corollaire facile de la Proposition 8.12. (Détails en exercice.) □

Théorème 8.14. *Soit I un intervalle ouvert de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une fonction convexe. Alors pour tout $a \in I$, f admet des dérivées à droite et à gauche en a. On a pour tout $x \in I$: $f(x) \ge f'_d(a)(x - a) + f(a)$ et $f(x) \ge f'_g(a)(x - a) + f(a)$. En particulier, il existe une fonction affine $g(x)$ telle que $g(a) = f(a)$ et $g(x) \le f(x)$ pour tout $x \in I$. De plus, si $a < b$ sont dans I, on a $f'_g(a) \le f'_d(a) \le f'_g(b)$. Si f est strictement convexe alors $f'_d(a) < f'_g(b)$.*

Démonstration. Soient $\varepsilon_1 > 0, \varepsilon_2 > 0$ tels que $a - \varepsilon_1$ et $a + \varepsilon_2 \in I$. Par l'inégalité de pentes

$$\frac{f(a - \varepsilon_1) - f(a)}{-\varepsilon_1} \le \frac{f(a + \varepsilon_2) - f(a)}{\varepsilon_2}.$$

Si $\varepsilon_1 \to 0$ et $\varepsilon_2 \to 0$ on obtient $f'_g(a) \le f'_d(a)$.

Soit $a < b$. Choisissons $\varepsilon_1 > 0$ et $\varepsilon_2 > 0$ tels que $a + \varepsilon_1 < b - \varepsilon_2$. Par l'inégalité de pentes de nouveau

$$\frac{f(a + \varepsilon_1) - f(a)}{\varepsilon_1} \le \frac{f(a + \varepsilon_1) - f(b - \varepsilon_2)}{a - b + \varepsilon_1 - \varepsilon_2} \le \frac{f(b) - f(b - \varepsilon_2)}{\varepsilon_2}.$$

Si $\varepsilon_1 \to 0$ et $\varepsilon_2 \to 0$ on obtient $f'_d(a) \leq f'_g(b)$. Si f est strictement convexe alors les inégalités ci-dessus sont strictes, $\frac{f(a+\varepsilon_1)-f(a)}{\varepsilon_1}$ décroît quand $\varepsilon_1 \to 0$ et $\frac{f(b)-f(b-\varepsilon_2)}{\varepsilon_2}$ croît quand $\varepsilon_2 \to 0$. Donc, $f'_d(a) < f'_g(b)$.

Soit $a < x, x \in I$, et $\varepsilon > 0$ tel que $a + \varepsilon < x$. Alors

$$\frac{f(a+\varepsilon) - f(a)}{\varepsilon} \leq \frac{f(x) - f(a)}{x - a}.$$

Si $\varepsilon \to 0$, on obtient $f(x) \geq f'_d(a)(x-a) + f(a)$ pour tout $x > a$. De manière analogique on montre que $f(x) \geq f'_g(a)(x-a) + f(a)$ pour tout $x < a$.

Puisque $f'_g(a) \leq f'_d(a)$, si $x < a$ on obtient

$$f'_g(a)(x-a) \geq f'_d(a)(x-a) \text{ si } x < a.$$

D'où $f(x) \geq f'_g(a)(x-a) + f(a) \geq f'_d(a)(x-a) + f(a)$ si $x < a$. Ainsi, nous avons demontré que $f(x) \geq f'_d(a)(x-a) + f(a)$ pour tout $x \in I$. De manière analogique on montre que $f(x) \geq f'_g(a)(x-a) + f(a)$ pour tout $x \in I$. (Dessins en CM.) □

Le fait que $f'_g(a)$ et $f'_d(a) \in \mathbb{R}$ implique :

Corollaire 8.15. Si $f : I \to \mathbb{R}$ est une fonction convexe alors f est continue.

Démonstration. L'existence de f'_d (resp. f'_g) entraîne la continuité de f à droite (resp. à gauche). D'où la continuité de f. □

On obtient facilement le critère suivant pour les fonctions dérvable sur $\mathbb{R}$:

Corollaire 8.16. Si $f : I \to \mathbb{R}$ est une fonction dérivable alors f est convexe (respectivement, stictement convexe) si et seulement si f' est croissante (respectivement, stictement croissante). En particulier, si f est deux fois dérivable alors f est convexe (respectivement, stictement convexe) si f'' est positive (respectivement, stictement positive).

Démonstration. $\Rightarrow$) Notons que $f' = f'_d = f'_g$. Soit $a, b \in I$ et $a \leq b$. Le Théorème 8.14 implique que $f'(a) \leq f'(b)$ si f est convexe et $f'(a) < f'(b)$ si f est strictement convexe.

$\Leftarrow$) Soient $a, b \in I$, $a < b$ et $0 < \lambda < 1$. Par l'égalité des accroissements finis, ils existent $\xi_1 \in]a, a + \lambda(b-a)[$ et $\xi_2 \in]a + \lambda(b-a), b[$ tels que

$$f'(\xi_1) = \frac{f(a + \lambda(b-a)) - f(a)}{\lambda(b-a)} \text{ et } f'(\xi_2) = \frac{f(b) - f(a + \lambda(b-a))}{(1-\lambda)(b-a)}.$$

Mais f' est croissante (resp. strictement croissannte). Donc,

$$\frac{f(a + \lambda(b-a)) - f(a)}{\lambda(b-a)} \leq \frac{f(b) - f(a + \lambda(b-a))}{(1-\lambda)(b-a)}$$

où $'' \leq''$ est remplacé par "$<$" si f est strictement croissante. Un calcul simple montre que la dennière inégalité est équivalente à

$$f((1-\lambda)a + \lambda b) \leq (1-\lambda)f(a) + \lambda f(b)$$

où $'' \leq''$ devient "$<$" si f est strictement croissante. □

8.4 Propriétés différentielles des fonctions convexes

8.4.1 Rappels

Si E, F sont des e.v.n. alors l'ensemble $L(E,F)$ des applications linéaires continues est un e.v.n. avec la norme $|||f||| := \sup_{||x||_E=1} ||f(x)||_F$.

Définition 8.17. Si U est un ouvert de E, alors $f : U \to F$ est *différentiable en* $x \in U$ s'il existe un $T \in L(E,F)$, notée $df(x)$, telle que

$$||f(x+h) - f(x) - df(x)(h)|| = \circ(||h||), \text{ si } ||h|| \to 0.$$

La fonction f est C^1-*différentiable* (ou *continuement différentiable*) sur U si f est différentiable en tout $x \in U$ et $df : U \to L(E,F), x \mapsto df(x)$, est continue. On note aussi $D_h f(x) := df(x)(h)$.

La fonction f est C^2-*différentiable* si f et df sont C^1-différentiables. On a $d^2 f : U \to L(E, L(E,F)), x \mapsto d^2 f(x)$. Facile à voir que $L(E, L(E,F)) = L^2(E \times E, F)$ ($\equiv$ l'ensemble des applications bi-linéaires). On note $d^2 f(x)(h,k) = D_k(D_h f)(x)$.

Dans le cas le plus fréquent pour nous où $E = \mathbb{R}^n$ et $F = \mathbb{R}$, si f est C^1 elle admet des dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ en $a \in U$, *le gradient de* f en a est noté $\nabla f(a) = (\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a))$ et

$$df(a)(h) =< \nabla f(a), h >= \sum_i \frac{\partial f}{\partial x_i}(a) h_i.$$

Si f est C^2 alors

$$d^2 f(a)(h,k) = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i k_j.$$

Donc, $d^2 f(a)$ est une forme bilinéaire qui est positive (c.à.d. $d^2 f(a)(h,h) \geq 0$ pour tout $h \in E$) ssi son *hessienne* $Hf(a) := (\frac{\partial^2 f}{\partial x_i \partial x_j}(a))$ est positive (c.à.d. les mineurs principaux dominants de la matrice $Hf(a)$ sont tous positifs.)

8.4.2 Caratérisations différentielles des fonctions convexes pour $\dim(E) < \infty$

Théorème 8.18. *Soient U un ouvert convexe de $E = \mathbb{R}^n$ et $f : U \to \mathbb{R}$ une fonction différentiable en tout point de U.*

1. *f est convexe si et seulement si pour tous $u, v \in U$:*

$$f(u) - f(v) \geq df(v)(u - v).$$

2. *f est convexe si et seulement si pour tous $u, v \in U$:*

$$(df(u) - df(v))(u - v) \geq 0.$$

3. *Si, en plus, f est C^2, f est convexe si et seulement si la forme bilinéaire symétrique $d^2 f(x)$ est positive pour tout $x \in U$. De plus, si $d^2 f(x)$ est strictement positive (c.à.d. $d^2 f(x)(h,h) > 0$ si $h \neq 0$) alors f est strictement convexe.*

Démonstration. 1. Si f est convexe, par Corollaire 8.11 appliqué pour $h = u - v$, on a

$$df(v)(u-v) = df(v)(h) = \inf_{1 \geq t > 0} \frac{f(v+th) - f(v)}{t} \leq f(v+h) - f(v) = f(u) - f(v).$$

Reciproquement, si $z = (1-t)v + tu$ où $1 \leq t \leq 1$, on obtient que $z \in U$ et

$$f(v) - f(z) \geq df(z)(v-z) \text{ et } f(u) - f(z) \geq df(z)(u-z).$$

D'où

$$(1-t)f(v) + tf(u) - f(z) \geq df(z)((1-t)(v-z) + t(u-z)) = df(z)(0) = 0$$

qui donne l'inégalité de convexité.

2. Si f est convexe, la partie 1 du théorème donne

$$df(v)(v-u) \geq f(v) - f(u) \text{ et } df(u)(u-v) \geq f(u) - f(v).$$

En sommant, on obtient l'inégalité voulue $(df(u) - df(v))(u-v) \geq 0$.

Pour montrer la réciproque, il suffit de justifier que la fonction $\varphi(t) = f((1-t)u+tv)$, $t \in [0,1]$, est convexe. On a

$$\varphi'(t) = df((1-t)u + tv)(v-u).$$

Or, si $t < s$

$$\begin{aligned} & \varphi'(s) - \varphi'(t) = (df((1-s)u + sv) - df((1-t)u + tv))(v-u) \\ = \quad & \frac{1}{s-t}(df(u + s(v-u)) - df(u + t(v-u))(u + s(v-u) - (u + t(v-u))) \geq 0. \end{aligned}$$

Donc, φ' est croissante et par Corollaire 8.16 φ est convexe.

3. Soient f convexe, $x \in U$ et $h \in E$. Puisque U est ouvert, il existe $\varepsilon > 0$ tel que $x - \varepsilon h$ et $x + \varepsilon h \in U$. La fonction $g :]-1,1[\to \mathbb{R}, t \mapsto f(x + t\varepsilon h)$, est convexe et deux fois dérivable. Par Corollaire 8.16 $g''(t) \geq 0$. Mais

$$g''(t) = (d^2 f)(x)(t\varepsilon h, t\varepsilon h) = (t\varepsilon)^2 (d^2 f)(x)(h,h) \geq 0.$$

Donc, $(d^2 f)(x)(h,h) \geq 0$ pour tout $h \in E$.

Réciproquement, l'inégalité $(d^2 f)(x)(h,h) \geq 0$ implique (en vertu du Corollaire 8.16) que la restriction de f sur chaque segment $[u,v]$ où $u, v \in U$, est convexe. Donc, f est convexe. Il est facile à voir que si $(d^2 f)(x)(h,h) > 0$, $h \neq 0$, alors f est strictement convexe (laissé en exercice.) $\square$

Remarque 8.19. Dans les preuves des parties 1 et 2 du théorème nous n'avons pas utilisé que $\dim E < \infty$. Donc, elles restent valables pour un e.v.n. quelconque.

8.5 Critère d'extremum global des fonctions convexes

Proposition 8.20. *Si f est une fonction convexe de class $\mathcal{C}^1$ sur un ouvert convexe $U \subset E$ alors $a \in U$ est un point critique de f (c.à.d. $df(a) = 0$) si et seulement si a est un minimum global de f.*

Démonstration. Soit $df(a) = 0$. Par Théorème 8.18(1), $f(u) - f(a) \geq 0$. Donc, $f(u) \geq f(a)$ pour tout $u \in U$ et a est un minimum glogal de f.

Réciproquement, soit a un minimum glogal de f. Supposons par l'absurd que $df(a) \neq 0$. Il existe $u \in U$ tel que $df(a)(u-a) > 0$. Mais f est de class $\mathcal{C}^1$. Donc, il existe $b \in U$ tel que $df(b)(b-a) > 0$. Par Corolaire 8.11

$$0 < df(b)(b-a) = \inf_{0<t\leq 1} \frac{f(b + t(a-b)) - f(b)}{t} \leq f(a) - f(b).$$

D'où, $f(b) < f(a)$. Absurd. $\square$

Théorème 8.21. *Soit C un convexe de $\mathbb{R}^n$ avec $C \subset U$ où U est un ouvert de $\mathbb{R}^n$. Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$, convexe sur C. Alors $a \in C$ est un minimum global de f sur C si et seulement si*

$$\forall c \in C, df(a)(c-a) =< \bigtriangledown f(a), c-a >\geq 0.$$

Démonstration. Si a est un minimum global sur le convexe C, $f(a) \leq f((c-a)t+a)$ pour tous $c \in C$ et $t \in [0,1]$. Donc,

$$\lim_{t \to 0^+} \frac{f((c-a)t+a) - f(a)}{t} = df(a)(c-a) \geq 0.$$

Réciproquement, si $df(a)(c-a) \geq 0$ pour tout $c \in C$, en vertu du Corollaire 8.11

$$0 \leq df(a)(c-a) = \inf_{0<t\leq 1} \frac{f((c-a)t+a) - f(a)}{t} \leq f(c) - f(a).$$

D'où, $f(c) \geq f(a)$ pour tout $c \in C$. □

Exemple 29. Soient $U = \mathbb{R}^n$, $u \in \mathbb{R}^n$ et $f(x) =\| u - x \|^2$, $x \in \mathbb{R}^n$. Alors $\bigtriangledown f(x) = -2(u-x)$ et le hessienne $Hf(x)$ est deux fois la matrice identité. En vertu du Théorème 8.18 (3), f est strictement convexe. Si C est un convexe de $\mathbb{R}^n$ alors la restriction de f sur C (notée aussi par f) est strictement convexe. Théorème 8.21 imlique que le point a dans sa formulation minimise la distance de u à C si et seulement si :

$$\forall c \in C, < u-a, c-a >\leq 0.$$

Le théorème suivant montre l'existence et l'unicité d'un tel point a. Le théorème analogique (mais avec une démonstration différente) est valable dans le cadre plus général des espaces Hilbert (e.v.n. qu'on va introduire plus tard).

Théorème 8.22. *(théorème de projection sur un convexe fermé de $\mathbb{R}^n$) Soit $C \subset \mathbb{R}^n$ un convexe fermé non-vide et $\| \cdot \|$ la norme euclidienne. Pour tout $u \in \mathbb{R}^n$ il existe un unique $a \in C$, noté $P_C(u)$ et appelé projection de u sur C, tel que*

$$\| u - a \|= \underset{c \in C}{inf} \| u - c \| .$$

C'est l'unique vecteur $a \in C$ tel que

$$\forall c \in C, < u-a, c-a >\leq 0.$$

De plus, pour tout $b \in C$, $b + N_c(b) = P_C^{-1}(b)$ et $b + N_c(b), b \in C$, forment une partition de $\mathbb{R}^n$.

Démonstration. Comme C est non-vide $r = \inf_{c \in C} \| u - c \|< \infty$. Soit $D := C \cap \overline{B}(u, r+1)$, où $\overline{B}(u, r+1)$ est la boule fermée de rayon $r+1$ centrée dans u. Alors D est fermé et borné, donc c'est un compact de $\mathbb{R}^n$. Evidemment, $D \cap C \neq \emptyset$ et D contient les éléments $c \in C$ tels que $\| c - u \|\leq r+1$. Donc,

$$r = \inf_{c \in D} \| u - c \| .$$

Or $c \mapsto\| u - c \|$ est continue sur le compact D, donc atteint son infimum en $a \in D$. Par croissance du carré, c'est aussi le point où $\| u - c \|^2$ atteint son minimum. La hessienne de $x \mapsto\| x - u \|^2$ est deux fois l'indentité, donc cette application est strictement convexe. Donc le minimum a est unique par Proposition 8.9(4).

En fin,

$$P_C^{-1}(b) = \{u \in E | \forall c \in C, < u-b, c-b >\leq 0\} = \{u \in E | u - b \in N_C(b)\} = b + N_C(b).$$

Le fait que $P_C : E \to C$ est une application surjective (vu que $P_C(c) = c$ pour tout $c \in C$) implique le rélultat de partition. □

8.6 Inégalités de convexité

La convexité (ou la concavité) est souvent utilisée pour établir des inégalités. Citons un exemple important et relativement simple.

Proposition 8.23. *Soit $f\colon [0,+\infty[\to [0,+\infty[$ une fonction concave. Alors pour tout $x,y \geq 0$ on a $f(x+y) \leq f(x)+f(y)$.*

Démonstration. Fixons $y \geq 0$ et considérons la fonction $g\colon [0,+\infty[\to \mathbb{R}$ définie par $g(x) = f(x) + f(y) - f(x+y)$.

Alors, pour tous $a,b \in [0,+\infty[, a < b$, on a

$$\frac{g(b)-g(a)}{b-a} = \frac{f(b)-f(a)}{b-a} - \frac{f(b+y)-f(a+y)}{b-a} \; .$$

Puisque $\frac{f(b+y)-f(a+y)}{b-a} = \frac{f(b+y)-f(a+y)}{(b+y)-(a+y)}$ est le taux d'accroissement de f entre $(b+y)$ et $(a+y)$, l'inégalité des pentes nous donne donc que $\frac{g(b)-g(a)}{b-a} \geq 0$, autrement dit g est croissante.

Par conséquent, on a pour tout x que $g(x) \geq g(0) = f(0)$, et donc $f(x)+f(y)-f(x+y) \geq f(0) \geq 0$, ce qu'on voulait démontrer. □

Voyons maintenant l'inégalité la plus importante de notre cours.

Théorème 8.24 (Inégalité de Jensen)**.** *Soit $(X,\mathcal{A},\mu)$ un espace de probabilité, g une fonction μ-intégrable à valeurs dans un intervalle (quelconque) I, et $\varphi\colon I \to \mathbb{R}$ une fonction convexe. Alors $\int_X g d\mu \in I$ et*

$$\varphi\left(\int_X g d\mu\right) \leq \int_X \varphi \circ g d\mu.$$

(L'intégrale de droite peut être égale à $+\infty$!)

Démonstration. Par Corollaire 8.15 la fonction φ est continue sur I, donc borélienne sur I, donc la composée $\varphi \circ g$ est borélienne aussi.

Posons $m = \int_X g d\mu$. (Rappelons que $m \in \mathbb{R}$.) Soit $m \notin \overline{I}$. Alors il existe $\varepsilon > 0$ tel que soit $m \geq g(x)+\varepsilon$ pour tout $x \in X$ soit $m \leq g(x)-\varepsilon$ pour tout $x \in X$. En utilisant que $\mu(X) = 1$, dans le premier cas $m \geq \int_X g d\mu + \varepsilon = m+\varepsilon$ et dans le deuxième $m \leq \int_X g d\mu - \varepsilon = m - \varepsilon$. Absurd. Donc, $m \in \overline{I}$. Soit m le minimum de I. Alors on a $\int_X (g-m) d\mu = 0$ et $g - m \geq 0$, donc $g - m = 0$ μ-p.p. Mais $\mathrm{Im}(g) \subset I$. Donc, $m \in I$. Par conséquent on a

$$\int_X \varphi \circ g d\mu = \int_X \varphi(m) d\mu = \varphi(m) = \varphi\left(\int_X g d\mu\right) \; .$$

Le cas où m est le maximum de I est traité de façon similaire.

Finalement, le cas qui nous reste est celui où m appartient à l'intérieur de I.

Le Théorème 8.14 donne que

$$\forall t \in I, \quad \varphi(t) - \varphi(m) \geq \alpha(t-m)$$

où $\alpha = \varphi'_g(m)$ ou $\varphi'_d(m)$. Mais $\mathrm{im}(g) \subset I$. En particulier, pour tout $x \in X$ on a $\varphi(g(x)) \geq \varphi(m) + \alpha(g(x)-m)$. La fonction à droite est intégrable. Donc, l'inégalité est conservée après l'intégration et on obtient

$$\int_X \varphi \circ g d\mu \geq \int_X \varphi(m) d\mu + \alpha \int_X (g-m) d\mu = \varphi(m) + \alpha(\int_X g d\mu - m) = \varphi(m) \; .$$

□

Une analyse de la preuve ci-dessus permet de préciser et renforcer le Théorème 8.24 dans certains cas.

Proposition 8.25. *Avec les notations et les hypothèses du Théorème 8.24, supposons en plus que φ est strictement convexe. Alors*

$$\varphi\left(\int_X g d\mu\right) = \int_X \varphi \circ g d\mu$$

si et seulement si g est constant μ-p.p.

Démonstration. $\Leftarrow$) Si $g = m \in \mathbb{R}$ μ-p.p. alors $\int_X g d\mu = m$ et $\varphi \circ g(x) = \varphi(m)$ μ-p.p. Donc, $\int_X \varphi \circ g d\mu = \varphi(m) = \varphi\left(\int_X g d\mu\right)$.

$\Rightarrow$) Au cours de la preuve du Théorème 8.24 nous avons établi que $m = \int_X g d\mu \in I$ et, par la convexité de φ, que

$$\forall t \in I, \quad \varphi(t) - \varphi(m) \geq \alpha(t - m)$$

où $\alpha = \varphi'_g(m)$ ou $\varphi'_d(m)$. En particulier,

$$\forall x \in X, \varphi \circ g(x) - \varphi(m) - \alpha(g(x) - m) \geq 0.$$

En intégrant cette inégalité (et en utilisant linéarité de l'intégrale de fonctions intégrables et l'hypothèse $\varphi(m) = \int_X \varphi \circ g d\mu$), on obtient que

$$\int_X (\varphi \circ g(x) - \varphi(m) - \alpha(g(x) - m)) d\mu = 0.$$

D'où, $\varphi \circ g(x) - \varphi(m) = \alpha(g(x) - m)$ μ-p.p. Supposons que g n'est pas égal à m μ-p.p. Soit $X' = \{x \in X | g(x) \neq m\}$. Alors $\mu(X') > 0$. Supposons que $g(x) = m'$ pour tout $x \in X'$. Alors, $m = \int_X g d\mu = \mu(X')m' + (1 - \mu(X'))m = m + \mu(X')(m' - m) \neq m$. Absurd. Il reste à considerer le cas quand X' contient des éléments x_1 et x_2 tels que $g(x_1) > g(x_2)$, $\varphi \circ g(x_1) - \varphi(m) = \alpha(g(x_1) - m)$ et $\varphi \circ g(x_2) - \varphi(m) = \alpha(g(x_2) - m)$. Donc, $\frac{\varphi \circ g(x_1) - \varphi(m)}{g(x_1) - m} = \frac{\varphi \circ g(x_2) - \varphi(m)}{g(x_2) - m}$. Mais φ est *strictement* convexe et $g(x_1) > g(x_2)$. En vertu de la Proposition 8.13, $\frac{\varphi \circ g(x_1) - \varphi(m)}{g(x_1) - m} > \frac{\varphi \circ g(x_2) - \varphi(m)}{g(x_2) - m}$. Absurd. □

Proposition 8.26. *Avec les notations et les hypothèses du Théorème 8.24, supposons en plus qu'il existe $Y \in \mathcal{A}$ tel que $0 < \mu(Y) < 1$. Alors $\varphi\left(\int_X g d\mu\right) = \int_X \varphi \circ g d\mu$ pour* **toute** *fonction g sur X, μ-intégrable et à valeurs dans I si et seulement si φ est linéaire, i.e. $\varphi(x) = \alpha x + \beta$ où $\alpha, \beta \in \mathbb{R}$.*

Démonstration. $\Leftarrow$) Si $\varphi(x) = \alpha x + \beta$ alors $\varphi(\int_X g d\mu) = \alpha \int_X g d\mu + \beta = \int_X (\alpha g + \beta) d\mu = \int_X \varphi \circ g d\mu$.

$\Rightarrow$) Soit $\lambda = \mu(Y)$. Pour a et $b \in I$, $a < b$, on définit la fonction $g_{a,b} = a 1_{Y^c} + b 1_Y$. La fonction $g_{a,b}$ est μ-intégrable, à valeurs dans I et $\int_X g_{a,b} d\mu = (1-\lambda)a + \lambda b$. Alors $\varphi \circ g_{a,b} = \varphi(a) 1_{Y^c} + \varphi(b) 1_Y$ et, donc,

$$\int_X \varphi \circ g_{a,b} d\mu = (1 - \lambda)\varphi(a) + \lambda\varphi(b) = \varphi\left(\int_X g_{a,b} d\mu\right) = \varphi((1 - \lambda)a + \lambda b).$$

Les inégalités de pentes (8.12) impliquent que la restriction de φ sur $[a, b]$ est linéaire pour tous a et $b \in I$. (Faites un dessin.) Donc, φ est linéaire sur I, i.e. $\varphi(x) = \alpha x + \beta$ où $\alpha, \beta \in \mathbb{R}$. □

Le corollaire suivant est un cas très particulier de l'inégalité de Jensen, qui peut se montrer facielment sans théorie de la mesure en utilisant la définition d'une fonction convexe.

Corollaire 8.27. Soit I un intervalle de $\mathbb{R}$, $\alpha_1, \ldots, \alpha_n$ des réels positifs tels que $\sum_{i=1}^n \alpha_i = 1$, et φ une fonction convexe sur I. Alors, pour tout $x_1, \ldots, x_n \in I$ on a

$$\varphi\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i \varphi(x_i)\ .$$

Démonstration. On considère l'espace de probabilité $(X, \mathcal{P}(X), \mu)$ où $X = \{x_1, \ldots, x_n\}$ et $\mu = \sum_{i=1}^n \alpha_i \delta_{x_i}$, où δ_{x_i} désigne la mesure de Dirac en x_i. Pour toute fonction $g \colon X \to \mathbb{R}$ on a

$$\int_X g d\mu = \sum_{i=1}^n \alpha_i g(x_i)\ .$$

En considérant pour g la fonction identité, on a donc $\displaystyle\int_X \varphi \circ g d\mu = \sum_{i=1}^n \alpha_i \varphi(x_i)$, et $\displaystyle\int_X g d\mu = \sum_{i=1}^n \alpha_i x_i$. L'inégalité de Jensen nous donne comme attendu

$$\varphi\left(\sum_{i=1}^n \alpha_i x_i\right) \leq \sum_{i=1}^n \alpha_i \varphi(x_i)\ .$$

□

Remarque 8.28. Dans le corollaire ci-dessus, le cas $n = 2$ correspond exactement à la définition de la convexité. En particulier, une application φ qui satisfait l'inégalité de Jensen pour toute fonction intégrable sur un espace de probabilité, est nécessairement convexe.

Exercice 8.29. Donner une démonstration du corollaire ci-dessus qui n'utilise que la définition d'une fonction convexe (et le principe de récurrence).

Le corollaire 8.27 permet d'établir plusieurs inégalités classiques, comme par exemples celles de l'exercice suivant.

Exercice 8.30.

1. Montrer que la fonction $x \mapsto -\ln(x)$ est strictement convexe sur $]0, +\infty[$. En déduire l'*inégalité arithmético-géométrique* : pour tout $x_1, \ldots, x_n$ positifs on a

$$\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}} \leq \frac{1}{n}\sum_{i=1}^n x_i\ .$$

2. Montrer que pour $p \geq 1$ la fonction $x \mapsto x^p$ est convexe sur $[0, +\infty[$ et en déduire que pour tout $x_1, \ldots, x_n$ positifs on a

$$\frac{1}{n}\sum_{k=1}^n x_k \leq \left(\frac{1}{n}\sum_{k=1}^n x_k^p\right)^{\frac{1}{p}}\ .$$

(Indication : Pour montrer que $x \mapsto x^p$ est convexe sur $[0, +\infty[$, on peut montrer qu'elle est convexe sur $]0, +\infty[$ et continue sur $[0, +\infty[$, conditions qui impliquent la convexité sur $[0, +\infty[$.)

3. Que peut-on dire si " $\leq$ " dans les parties 1 et 2 de l'exercice est remplacé par " $=$ " ?

Proposition 8.31. *Supposons que $0 < p < q < +\infty$ et que $(X, \mathcal{A}, \mu)$ est un espace de probabilité. Alors, pour toute fonction positive mesurable f, on a*

$$\left(\int_X f^p d\mu\right)^{\frac{1}{p}} \leq \left(\int_X f^q d\mu\right)^{\frac{1}{q}}$$

Démonstration. Posons $r = \frac{q}{p}$ et $g = f^p$. Comme $x \mapsto x^r$ est convexe sur $[0, +\infty[$, l'inégalité de Jensen nous dit que quand g est intégrable on a

$$\left(\int_X g d\mu\right)^r \leq \int_X g^r d\mu \ .$$

Alors $\left(\int_X f^p d\mu\right)^{\frac{q}{p}} \leq \int_X f^q d\mu$, et cette inégalité est équivalente à celle qu'on souhaitait établir.

Reste à traiter le cas où g n'est pas intégrable ; dans ce cas, on peut par exemple poser $f_n = \min(f, n)$, et noter que les suites (f_n^p) et (f_n^q) sont croissantes vers f^p et f^q respectivement ; on peut donc appliquer le théorème de convergence monotone pour conclure que

$$\int_X f^p d\mu = \lim_{n \to +\infty} \int_X f_n^p d\mu \quad \text{et} \int_X f^q d\mu = \lim_{n \to +\infty} \int_X f_n^q d\mu \ .$$

Comme chaque f_n^p est intégrable, la première étape de notre raisonnement nous donne $\left(\int_X f_n^p d\mu\right)^{\frac{1}{p}} \leq \left(\int_X f_n^q d\mu\right)^{\frac{1}{q}}$, et l'inégalité souhaitée s'en déduit par passage à la limite. □

Notation. Si $(X, \mathcal{A}, \mu)$ est un espace mesuré, f est une fonction mesurable sur X et $1 \leq p < +\infty$, on notera

$$\|f\|_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}} \ .$$

Remarque 8.32. La notation $\|f\|_p$ est introduite **uniquement** dans le cas $p \geq 1$. Dans ce cas est valable l'inégalité de Minkowski démontrée dans le chapitre suivant qui entraîne que l'ensemble de fonctions avec $\|f\|_p < \infty$ constitue un espace vectoriel normé.

La Proposition 8.31 consiste à dire que, sur un espace de probabilité, on a pour toutes fonctions f, g mesurables positives l'inégalité $\|f\|_p \leq \|f\|_q$ dès que $1 \leq p < q < +\infty$.

Définition 8.33. Soit $p, q \in [1, +\infty]$. On dit que p et q sont des *exposants conjugués* si on a $\dfrac{1}{p} + \dfrac{1}{q} = 1$. On dit qu'alors q est l'exposant conjugué de p.

Notons que, si q est l'exposant conjugué de p, alors p est l'exposant conjugué de q ; 1 est l'exposant conjugué de $+\infty$, tandis que l'exposant conjugué de 2 est 2.

Cette notion est importante en grande partie à cause de l'inégalité suivante.

Théorème 8.34 (Inégalité de Hölder). *Soit $(X, \mathcal{A}, \mu)$ un espace mesuré, $p, q \in]1, +\infty[$ deux exposants conjugués et f, g deux fonctions mesurables à valeurs réelles telles que $|f|^p$ et $|g|^q$ sont intégrables. Alors le produit fg est intégrable, et on a*

$$\int_X |fg| d\mu \leq \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}} \left(\int_X |g|^q d\mu\right)^{\frac{1}{q}} \ .$$

De manière plus condensée : $\|fg\|_1 \leq \|f\|_p \|g\|_q$.

Démonstration. Comme seules les valeurs absolues de f, g interviennent dans l'énoncé, on peut supposer f, g à valeurs positives. Notons ensuite que, si f ou g est nulle presque partout alors il en va de même du produit fg et l'inégalité désirée est vraie. On peut donc supposer que g n'est pas nulle presque partout, auquel cas $\|g\|_q > 0$. Considérons l'espace mesurable $(X, \mathcal{A})$ muni de *la mesure à densité* ν définie par $\nu(A) = \displaystyle\int_A \frac{g^q}{\int_X g^q d\mu} d\mu$. C'est un espace de probabilité.

Considérons la fonction $h = fg^{1-q}$. Alors en utilisant la Proposition 8.31 et le fait que, comme p et q sont conjugués, on a $p(1-q)+q=0$, on obtient :

$$\begin{aligned}
\int_X fg d\mu &= \left(\int_X g^q d\mu\right)\left(\int_X fg^{1-q}\frac{g^q}{\int_X g^q d\mu}d\mu\right) \\
&= \left(\int_X g^q d\mu\right)\int_X h d\nu \\
&\leq \left(\int_X g^q d\mu\right)\left(\int_X h^p d\nu\right)^{\frac{1}{p}} \\
&= \left(\int_X g^q d\mu\right)\left(\int_X f^p g^{p(1-q)}\frac{g^q}{\int_X g^q d\mu}d\mu\right)^{\frac{1}{p}} \\
&= \left(\int_X g^q d\mu\right)\left(\int_X \frac{f^p}{\int_X g^q d\mu}d\mu\right)^{\frac{1}{p}} \\
&= \left(\int_X g^q d\mu\right)^{1-\frac{1}{p}}\left(\int_X f^p d\mu\right)^{\frac{1}{p}} \\
&= \|g\|_q \|f\|_p \ .
\end{aligned}$$

□

L'inégalité de Hölder est une des inégalités fondamentales concernant les espaces L^p, qui constituent la matière du chapitre suivant. Dans le cas particulier où $p = q = 2$, on retrouve *l'inégalité de Cauchy-Schwarz*, valable dès que f^2 et g^2 sont intégrables :

$$\int_X |fg| d\mu \leq \left(\int_X f^2 d\mu\right)^{\frac{1}{2}}\left(\int_X g^2 d\mu\right)^{\frac{1}{2}}.$$

Le CM du 24/11/2021

Chapitre 9

Introduction aux espaces L^p

Dans ce chapitre, $(\Omega, \mathcal{T}, \mu)$ désignera un espace mesuré. Il est facile à voir que la rélation $f = g$ μ-p.p. est une rélation d'équivalence. Par la suite chaque fonction mesurable sera idintifiée avec sa classe d'équivalence. *Autrement dit, on va travailler en identifiant les fonctions si elles coïncident μ-presque partout et on écrira $f = g$ quand $\mu(\{x : f(x) \neq g(x)\}) = 0$.* En particulier, $f = 0$ signifiera que f vaut 0 presque partout. Par exemple, si f est la fonction caractéristique de $\mathbb{Q}$, on pourra écrire $f = 0$.

9.1 L'espace L^∞

Définition 9.1. Soit $f : \Omega \to [0, +\infty]$ une fonction mesurable. On dit que $M \in [0, +\infty[$ est un *majorant essentiel* de f si $\mu(\{x : f(x) > M\}) = 0$, autrement dit, si $f \leq M$ presque partout. On dit que $+\infty$ est un *majorant essentiel* de f si $\mu(\{x : f(x) > M\}) > 0$ pour tout $M \in \mathbb{R}^+$.

Si $f : \Omega \to \overline{\mathbb{R}}$ est une fonction mesurable, on définit $\|f\|_\infty$ comme le plus petit majorant essentiel de $|f|$.

Notation. Si désigne par $L^\infty(\Omega, \mathcal{T}, \mu)$, ou simplement $L^\infty(\Omega)$ quand il n'y a pas de risque de confusion, l'ensemble formé par toutes les fonctions f telles que $\|f\|_\infty < +\infty$ [i].

Proposition 9.2. *$(L^\infty(\Omega), \|\cdot\|_\infty)$ est un e.v.n. (espace vectoriel normé).*

Démonstration. Commençons par vérifier l'axiome de séparation : $\|f\|_\infty = 0$ est équivalent à dire que $\mu(\{x : |f(x)| > 0\}) = 0$, autrement dit que $f = 0$ (presque partout).

Ensuite, notons que si $a \in \mathbb{R}$, alors M est un majorant essentiel de $f \in L^\infty(\Omega)$ si et seulement si $|a|M$ est un majorant essentiel de af. Il suit que $af \in L^\infty(\Omega)$ et $\|af\|_\infty = |a|\|f\|_\infty$.

Vérifions l'inégalité triangulaire : soit $f, g \in L^\infty(\Omega)$. Alors on a

$$\mu(\{x : |f(x)| > \|f\|_\infty \text{ ou } |g(x)| > \|g\|_\infty\}) = 0$$

puisque cet ensemble est la réunion de deux ensembles de mesure nulle. Soit $|f(x) + g(x)| \geq \|f\|_\infty + \|g\|_\infty$. Mais $|f(x)| + |g(x)| \geq |f(x) + g(x)|$. Donc, $|f(x)| > \|f\|_\infty$ or $|g(x)| > \|g\|_\infty$. D'où,

$$\mu(\{x : |f(x) + g(x)| > \|f\|_\infty + \|g\|_\infty\}) = 0.$$

On vient de montrer que $\|f\|_\infty + \|g\|_\infty$ est un majorant essentiel de $f + g$, ce qui revient à dire que $f + g \in L^\infty(\Omega)$ et $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. En particulier, si f et $g \in L^\infty(\Omega)$ alors $f + g \in L^\infty(\Omega)$. □

i. Répétons pour la dernière fois que deux fonctions sont identifiées si elles coïncident presque partout ; notons que si $f = g$ presque partout alors $\|f\|_\infty = \|g\|_\infty$.

Rappel 9.3. Rappelons quelques notations et resultats de la sous-section 7.6.

1. Si $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré, f est une fonction mesurable sur Ω et $1 \leq p < +\infty$, on note
$$\|f\|_p = \left(\int_X |f|^p d\mu\right)^{\frac{1}{p}}.$$
2. La Proposition 8.31 montre que, sur un espace de probabilité, on a pour toute fonction f mesurable positive l'inégalité $\|f\|_p \leq \|f\|_q$ dès que $1 \leq p < q < +\infty$.
3. Soit $p, q \in [1, +\infty]$. On dit que p et q sont des *exposants conjugués* si on a $\dfrac{1}{p} + \dfrac{1}{q} = 1$. On dit qu'alors q est l'exposant conjugué de p (ou que p est l'exposant conjugué de q).
4. L'inégalité de Hölder (voir Théorème 8.34) dit que si $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré, $p, q \in]1, +\infty[$ sont deux exposants conjugués et f, g deux fonctions mesurables à valeurs réelles telles que $|f|^p$ et $|g|^q$ sont intégrables alors le produit fg est intégrable, et on a
$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

La proposition suivante montre que l'inégalité de Hölder reste vraie pour les exposants conjugués 1 et $+\infty$.

Proposition 9.4. *Si f, g sont mesurables, $\|f\|_1 < +\infty$ et $\|g\|_\infty < +\infty$, alors fg est intégrable et $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$.*

Démonstration. Il suffit de noter que, μ-presque partout, on a $|g(x)| \leq \|g\|_\infty$, et donc $|f(x)g(x)| \leq |f(x)|\|g\|_\infty$. En intégrant cette inégalité et en utilisant la Proposition 6.8, on obtient bien
$$\|fg\|_1 = \int_\Omega |f(x)g(x)| d\mu \leq \int_\Omega |f(x)|\|g\|_\infty d\mu = \|f\|_1 \|g\|_\infty .$$
□

9.2 Les espaces $L^p(\Omega)$

La notation $L^\infty(\Omega, \mathcal{T}, \mu)$ introduite plus haut se généralise naturellement comme suit :

Notation. Si $p \in [1, +\infty]$ alors $L^p(\Omega, \mathcal{T}, \mu)$, noté $L^p(\Omega, \mu)$ ou encore $L^p(\Omega)$ quand il n'y a pas de risque de confusion, est l'ensemble des fonctions f telles que $\|f\|_p < +\infty$.

On voudrait montrer que $(L^p(\Omega), \|\cdot\|_p)$ est un e.v.n. L'axiome de séparation n'est pas difficile à montrer : on a bien $\|0\|_p = 0$; et réciproquement, si $\|f\|_p = 0$ alors $\int_\Omega |f(x)|^p d\mu = 0$, ce qui n'est possible (comme $|f(x)|^p \geq 0$ pour tout x) que si $|f(x)|^p = 0$ μ-presque partout, c'est-à-dire si $f = 0$ (presque partout). (Voir Corollaire 6.7.)

L'axiome d'homogénéité se vérifie également très facilement : pour $f \in L^p(\Omega)$ et $\lambda \in \mathbb{R}$,
$$|af|_p = \left(\int_\Omega |af(x)|^p d\mu\right)^{\frac{1}{p}} = \left(|a|^p \int_\Omega |f(x)|^p d\mu\right)^{\frac{1}{p}} = |a|\|f\|_p.$$

Pour completer la preuve que $(L^p(\Omega), \|\cdot\|_p)$ est un e.v.n. il suffit de démontrer l'inégalité triangulaire établit par le théorème suivant.

Théorème 9.5 (Inégalité de Minkowski)**.** *Soit $p \in [1, +\infty]$ et $f, g \in L^p(\Omega)$. Alors $f + g \in L^p(\Omega)$ et $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.*

Démonstration. On a déjà traité le cas $p = +\infty$ (Proposition 9.2) et le cas $p = 1$ s'ensuit simplement de l'inégalité triangulaire habituelle. Supposons donc $p \in]1, +\infty[$ et $f, g \in L^p(\Omega)$.

Commençons par montrer que $\|f + g\|_p < +\infty$. Comme la fonction $x \mapsto x^p$ est convexe, on a pour tout x que

$$\left|\frac{1}{2}f(x) + \frac{1}{2}g(x)\right|^p \leq \left(\frac{1}{2}|f(x)| + \frac{1}{2}|g(x)|\right)^p \leq \frac{1}{2}|f(x)|^p + \frac{1}{2}|g(x)|^p .$$

En intégrant cette inégalité, on obtient que

$$\frac{1}{2^p}\|f + g\|_p^p \leq \frac{1}{2}(\|f\|_p^p + \|g\|_p^p)) .$$

Ceci nous prouve que $\|f + g\|_p < +\infty$.

Maintenant, notons que $q = \dfrac{p}{p-1}$ est l'exposant conjugué de p. Notons que $(f + g)^{(p-1)q} = (f + g)^p$. Donc, $(f + g)^{p-1} \in L^q(\Omega)$. Ci-dessous, on utilisera l'inégalité de Hölder pour la paire de fonctions f et $(f + g)^{p-1}$ et, aussi, pour la paire de fonctions g et $(f + g)^{p-1}$. On a

$$\begin{aligned}
\|f + g\|_p^p &= \int_\Omega |f + g|^p d\mu \\
&\leq \int_\Omega (|f| + |g|)|f + g|^{p-1} d\mu \\
&= \int_\Omega |f||f + g|^{p-1} d\mu + \int_\Omega |g||f + g|^{p-1} d\mu \\
&\leq \|f\|_p \cdot \left\|(f + g)^{p-1}\right\|_q + \|g\|_p \cdot \left\|(f + g)^{p-1}\right\|_q \\
&= (\|f\|_p + \|g\|_p) \cdot \left\|(f + g)^{p-1}\right\|_q \\
&= (\|f\|_p + \|g\|_p)\|f + g\|_p^{p-1}
\end{aligned}$$

Si $\|f+g\|_p = 0$ on n'a rien à démontrer. Sinon, en divisant les deux côtés par $\|f+g\|_p^{p-1}$ on obtient $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. □

On se pose la question naturelle comment les espaces $L^p(\Omega, \mathcal{T}, \mu)$, $p \geq 1$, sont liés entre eux.

Proposition 9.6. *Si $\mu(\Omega) < \infty$ et $1 \leq p < q$ alors $L^q(\Omega, \mathcal{T}, \mu) \subset L^p(\Omega, \mathcal{T}, \mu)$.*

Première preuve. Si $f \in L^q(\Omega, \mathcal{T}, \mu)$ et $p < q$ alors $\int_\Omega |f|^p d\mu = \int\limits_{\{x\in\Omega:|f(x)|\leq 1\}} |f|^p d\mu + \int\limits_{\{x\in\Omega:|f(x)|>1\}} |f|^p d\mu \leq \mu(\Omega) + \int_\Omega |f|^q d\mu < \infty$. □

Deuxième preuve. Rien à démontrer si $\mu(\Omega) = 0$. Soit $\mu(\Omega) > 0$ et $\nu = \frac{\mu}{\mu(\Omega)}$. Alors $L^p(\Omega, \mathcal{T}, \mu) = L^p(\Omega, \mathcal{T}, \nu)$ et $L^q(\Omega, \mathcal{T}, \mu) = L^q(\Omega, \mathcal{T}, \nu)$. Puisque $(\Omega, \mathcal{T}, \nu)$ est un espace de probabilité, la Proposition 8.31 implique que $L^q(\Omega, \mathcal{T}, \nu) \subset L^p(\Omega, \mathcal{T}, \nu)$. □

Exemple 30. On a $L^1(]0, 1[) \supsetneqq L^2(]0, 1[)$ parce que la fonction $]0, 1[\to \mathbb{R}$, $x \mapsto \frac{1}{\sqrt{x}}$, appartient à $L^1(]0, 1[)$ mais pas à $L^2(]0, 1[)$. (Vérification en exercice.) Donc, en général, si $p < q$ et $\mu(\Omega) < \infty$ alors $L^q(\Omega, \mathcal{T}, \mu) \subsetneqq L^p(\Omega, \mathcal{T}, \mu)$.

Exemple 31. Lorsque $\mathcal{T} = \mathcal{P}(\Omega)$ et μ est la mesure de comptage sur Ω, on note $l^p(\Omega)$ au lieu de $L^p(\Omega, \mathcal{P}(\Omega), \mu)$. En particulier, si $\Omega = \mathbb{N}$, $l^p(\mathbb{N})$ est l'espace des suites de puissance p-ième sommable, c.à.d. $(a_n) \in l^p(\mathbb{N})$ ssi $\sum\limits_{n\in\mathbb{N}} |a_n|^p < \infty$.

Proposition 9.7. *Si $p < q$ alors, $l^q(\mathbb{N}) \supset l^p(\mathbb{N})$.*

Démonstration. Si $f \in l^p(\mathbb{N})$ alors $\sum\limits_{n\geq 0} |u_n|^p < \infty$ où $u_n := f(n)$. Donc la suite (u_n) tend vers 0. Par suite, si $p < q$ on a $|u_n|^q = O(|u_n|^p)$. D'où, $\sum\limits_{n\geq 0} |u_n|^q < \infty$ par comparaison. Ceci montre que $f \in l^q(\mathbb{N})$. □

Exemple 32. L'inclusions dans la Proposition 9.7 est stricte. Par exemple, $f : n \mapsto \frac{1}{n+1}$, appartient à $l^2(\mathbb{N})$ mais pas à $l^1(\mathbb{N})$. (Vérification en exercice.)

Exemple 33. Si $f : \mathbb{R}^+ \to \mathbb{R}$, $x \mapsto \frac{1}{1+x}$, alors $f \in L^2(\mathbb{R}^+) \setminus L^1(\mathbb{R}^+)$. (Vérification en exercice.) D'autre part, si $g : \mathbb{R}^+ \to \mathbb{R}, x \mapsto 1_{]0,1[}(x)\frac{1}{\sqrt{x}}$, l'exemple 30 nous montre que $g \in L^1(\mathbb{R}^+) \setminus L^2(\mathbb{R}^+)$.

Remarque 9.8. Les exemples et les propositions ci-dessus montrent que, en général, il n'y a pas d'ordre entre les e.v.n. $L^p(\Omega, \mathcal{T}, \mu)$ mais pour certain classes des espaces mesurés $(\Omega, \mathcal{T}, \mu)$ des ordres existent. Donc, il faut faire beaucoup d'attention en travaillant avec les e.v.n. $L^p(\Omega, \mathcal{T}, \mu)$.

Si $0 < p < q < +\infty$ et $\mu(\Omega) = 1$ selon la Proposition 8.31 on a pour toute fonction mesurable positive f l'inégalité $\|f\|_p \leq \|f\|_q$. Il est trés facile à voir que $\|f\|_q \leq \|f\|_\infty$. (En exercice.) La proposition suivante nous dit encore plus.

Proposition 9.9. *Soit* $(\Omega, \mathcal{T}, \mu)$ *un espace de probabilité et* $f : \Omega \to \overline{\mathbb{R}}$ *une fonction mesurable. Alors*

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

Démonstration. Commençons par remarquer que l'on a toujours

$$\|f\|_p^p = \int_\Omega |f|^p d\mu \leq \|f\|_\infty^p \mu(\Omega) = \|f\|_\infty^p \ .$$

Par conséquent, $\lim_{p\to+\infty} \|f\|_p \leq \|f\|_\infty$.

Pour voir la réciproque, notons que pour $t < \|f\|_\infty$ fixé, l'ensemble $A_t = \{x : |f(x)| > t\}$ est de mesure strictement positive. Par conséquent

$$\|f\|_p \geq (t^p \mu(A_t))^{\frac{1}{p}} = t\mu(A_t)^{\frac{1}{p}} \text{ et } \mu(A_t)^{\frac{1}{p}} \to 1 \text{ quand } p \to +\infty \ .$$

Donc, $\lim_{p\to+\infty} \|f\|_p \geq t$ si $t < \|f\|_\infty$. D'où, $\|f\|_\infty = \lim_{p\to+\infty} \|f\|_p$. □

Rappelons la définition suivante.

Définition 9.10. Une suite (u_n) dans un e.v.n. $(E, \|\cdot\|)$ est une suite de Cauchy si quelque soit $\varepsilon > 0$ il existe un $N > 0$ tel que pour tous $n, m \in \mathbb{N}$, $n > N$ et $m > N$, on a $\|u_n - u_m\| < \varepsilon$.

Définition 9.11. Un e.v.n. $(E, \|\cdot\|)$ est un espace de Banach si toute suite de Cauchy dans $(E, \|\cdot\|)$ est convergente.

Remarque 9.12. Il est facile à voir que chaque e.v.n. est un espace de Banach. Par contre, en dimension infinie il y a des e.v.n. qui ne sont pas Banach. (Voir le chapitre suivant ou la feuille TD.)

Le théorème suivant est admis.

Théorème 9.13. *(de Riesz-Fisher) Pour tout* $p \in [1, +\infty]$, $L^p(\Omega, \mathcal{T}, \mu)$ *est un espace de Banach.*

Définition 9.14. Soit $(\Omega, \mathcal{T}, \mu)$ un espace de probabilité, et (f_n) une suite de fonctions mesurables. On dit que f_n *converge en probabilité* vers une fonction mesurable f si, pour tout $\varepsilon > 0$, on a

$$\lim_{n\to+\infty} \mu\left(\{x : |f_n(x) - f(x)| \geq \varepsilon\}\right) = 0 \ .$$

Théorème 9.15. *Soit $p \in [1, +\infty]$, $(\Omega, \mathcal{T}, \mu)$ un espace de probabilité, et (f_n) une suite de fonctions telles que $f_n \in L^p(\Omega)$ pour tout n et la suite (f_n) converge vers f dans $L^p(\Omega)$. Alors (f_n) converge vers f en probabilité.*

Démonstration. Si $p = +\infty$ le résultat est immédiat puisque quel que soit $\varepsilon > 0$, pour n suffisamment grand on a $\|f_n - f\|_\infty < \varepsilon$ et donc $\mu(\{x\colon |f_n(x) - f(x)| \geq \varepsilon\}) = 0$.

Supposons donc $p < +\infty$, fixons $\varepsilon > 0$ et considérons l'ensemble $A_n = \{x\colon |f_n(x) - f(x)| \geq \varepsilon\}$. Sur A_n, on a $|f_n - f|^p \geq \varepsilon^p$, et donc

$$\varepsilon^p \mu(A_n) \leq \int_\Omega |f_n - f|^p d\mu \ .$$

Autrement dit, $\mu(A_n) \leq \dfrac{\|f_n - f\|_p^p}{\varepsilon^p}$, et le terme de droite tend vers 0 quand n tend vers $+\infty$ puisque $\|f_n - f\|_p$ tend vers 0. □

On pourrait aussi se demander quel est le rapport entre convergence L^p et convergence presque partout. La proposition suivante est une conséquence facile du théorème de convergence dominée.

Proposition 9.16. *Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré, $p \in [1, +\infty[$, et (f_n) une suite de fonctions mesurables telles que $f_n(x)$ converge simplement vers une limite $f(x)$ (presque partout), et supposons qu'il existe une fonction $g \in L^p(\Omega)$ telle que pour tout n on ait $|f_n(x)| \leq g(x)$ presque partout. Alors $f \in L^p(\Omega)$ et $\|f_n - f\|_p \to 0$ quand n tend vers $+\infty$.*

Démonstration. En écartant un sous-ensemble de mesure 0 on se ramène au cas $f_n(x) \to f(x)$ et $|f_n(x)| \leq g(x)$ pour tout x. La fonction $f(x)$ est mesurable comme limit de fonctions mesurables et $|f(x)| \leq g(x)$. Donc, $|f(x)|^p \leq g(x)^p$ pour tout x. Par consequant, $\int |f(x)|^p d\mu \leq \int g(x)^p d\mu$ et $f \in L^p(\Omega)$. En plus, $|f(x) - f_n(x)| \leq 2g(x)$ et $|f(x) - f_n(x)|^p \leq 2^p g(x)^p$. Puique g^p est intégrable, le théorème de convergence dominée implique que $\int |f(x) - f_n(x)|^p d\mu \to 0$. Ainsi, nous avons demontré que $\|f_n - f\|_p \to 0$ quand n tend vers $+\infty$. □

Réciproquement, on a le résultat suivant, qu'on se contente de mentionner ici sans démonstration.

Théorème 9.17. *Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré, $p \in [1, +\infty[$, et (f_n) une suite d'éléments de $L^p(\Omega)$ telle que $\|f_n - f\|_p \to 0$ quand n tend vers $+\infty$. Alors il existe une suite extraite (f_{n_k}) telle que $(f_{n_k}(x))$ tend vers $f(x)$ presque partout.*

Exercice 9.18. Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré, et (f_n) une suite d'éléments de $L^\infty(\Omega)$ qui converge vers f dans $(L^\infty(\Omega), \|\cdot\|_\infty)$.

1. Montrer que $(f_n(x))$ converge vers $f(x)$ presque partout.
2. Montrer que la réciproque est fausse.

Chapitre 10

Espaces de Hilbert ; bases hilbertiennes

Dans tout ce chapitre, E désignera un $\mathbb{R}$-espace vectoriel *de dimension quelconque.* On commence par des rappels.

10.1 Quelques résultats sur les espaces euclidiens

Définition 10.1. 1. Un *produit scalaire* sur E est une application $\langle \cdot, \cdot \rangle : E \times E \to \mathbb{R}$ qui est :

- *symétrique*, c'est-à-dire que pour tout $x, y \in E$ on a $\langle x, y \rangle = \langle y, x \rangle$;
- *bilinéaire*, c'est-à-dire que pour tout $x, y, z \in E$ et tout $\lambda \in \mathbb{R}$ on a $\langle \lambda x + y, z \rangle = \lambda \langle x, z \rangle + \langle y, z \rangle$ (et donc aussi $\langle x, \lambda y + z \rangle = \lambda \langle x, y \rangle + \langle y, z \rangle$ par symétrie) ;
- *définie positive*, c'est-à-dire que pour tout $x \in E$ on a $\langle x, x \rangle \geq 0$ et $\langle x, x \rangle = 0 \Leftrightarrow x = 0_E$.

2. Si E est mune d'un produit scalaire alors E est un **espace euclidien**.

Exercice 10.2. Montrer que, si $(X, \mathcal{A}, \mu)$ est un espace mesuré, alors $\langle \cdot, \cdot \rangle$ défini par

$$\langle f, g \rangle = \int_X f g d\mu$$

est un produit scalaire sur $L^2(X)$, donc, $L^2(X)$ est un espace euclidien.

Dans la suite, E désigne un espace euclidien avec produit scalaire $\langle \cdot, \cdot \rangle$.

Notons qu'ici on travaillera avec des espaces vectoriels réels ; on peut aussi définir un produit scalaire pour des espaces vectoriels complexes mais dans ce cas-là on doit demander qu'il soit *sesquilinéaire* plutôt que bilinéaire.

Pour nous habituer aux calculs avec les produits scalaires, notons deux conséquences faciles. La première est qu'on connaît les valeurs de $\langle x, y \rangle$ pour tout x, y dès qu'on connaît la valeur de $\langle x, x \rangle$ pour tout $x \in E$.

Proposition 10.3 (Identité de polarisation)**.** *Pour tout $x, y \in E$ on a*

$$4\langle x, y \rangle = \langle x + y, x + y \rangle - \langle x - y, x - y \rangle \ .$$

Démonstration. Par bilinéarité et symétrie, on a

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

Un calcul similaire (ou le résultat précédent appliqué à $y' = -y$) donne

$$\langle x-y, x-y\rangle = \langle x,x\rangle + \langle y,y\rangle - 2\langle x,y\rangle$$

d'où le résultat. $\square$

Proposition 10.4 (Identité du parallélogramme). *Pour tout $x, y \in E$ on a*

$$\langle x+y, x+y\rangle + \langle x-y, x-y\rangle = 2\langle x,x\rangle + 2\langle y,y\rangle\ .$$

La preuve est laissée en exercice.

Théorème 10.5 (Inégalité de Cauchy-Schwarz). *Pour tout $x, y \in E$ on a*

$$\langle x,y\rangle^2 \leq \langle x,x\rangle\langle y,y\rangle$$

avec égalité si et seulement si x et y sont liés, c.à.d. $x \in \mathbb{R}y$ ou $y \in \mathbb{R}x$

Avant de donner la preuve de l'inégalité de Cauchy-Schwarz, notons tout de suite une conséquence fondamentale.

Corollaire 10.6. L'application $x \mapsto \sqrt{\langle x,x\rangle}$ est une norme sur E qu'on notera dans la suite $\|x\|$. Donc, E est un e.v.n.

Démonstration. L'axiome de séparation se déduit immédiatement du fait que le produit scalaire est défini positif. Celui d'homogénéité suit de sa bilinéarité. Il reste à vérifier l'inégalité triangulaire : soit $x, y \in E$. On a

$$\begin{aligned}\langle x+y, x+y\rangle &= \langle x,x\rangle + \langle y,y\rangle + 2\langle x,y\rangle \\ &\leq \langle x,x\rangle + \langle y,y\rangle + 2\sqrt{\langle x,x\rangle}\sqrt{\langle y,y\rangle} && \text{(Cauchy-Schwarz)} \\ &= \left(\sqrt{\langle x,x\rangle} + \sqrt{\langle y,y\rangle}\right)^2\end{aligned}$$

Comme toutes les quantités en jeu sont positives, on peut passer à la racine carrée de chaque côté de l'inégalité et obtenir comme espéré

$$\sqrt{\langle x+y, x+y\rangle} \leq \sqrt{\langle x,x\rangle} + \sqrt{\langle y,y\rangle}\ .$$

$\square$

Remarque 10.7. En utilisant la notation $\|x\| = \sqrt{\langle x,x\rangle}$ l'inégalité de Cauchy-Schwarz s'écrit

$$|\langle x,y\rangle| \leq \|x\| \cdot \|y\|$$

et les identités de polarisation et du parallélogramme s'écrivent ainsi :

$$\begin{aligned}&\forall x, y \in E \quad 4\langle x,y\rangle = \|x+y\|^2 - \|x-y\|^2 && \text{(polarisation)} \\ &\forall x, y \in E \quad 2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2 && \text{(parallélogramme)}\end{aligned}$$

Preuve de l'inégalité de Cauchy-Schwarz. Fixons $x, y \in E$, et considérons l'application définie sur $\mathbb{R}$ par

$$t \mapsto \langle x+ty, x+ty\rangle = \langle x,x\rangle + 2t\langle x,y\rangle + t^2\langle y,y\rangle.$$

Il s'agit d'une application polynomiale, de degré 2, à valeurs positives. Le polynôme en question ne peut avoir qu'une racine réelle au plus (sans quoi il changerait de signe) : son discriminant doit donc être négatif ou nul. Autrement dit, on doit avoir

$$4\langle x,y\rangle^2 - 4\langle x,x\rangle\langle y,y\rangle \leq 0 \ \Leftrightarrow\ |\langle x,y\rangle| \leq \|x\| \cdot \|y\|.$$

Il est facile à voir que si x et y sont liés alors $|\langle x,y\rangle| = \|x\|\cdot\|y\|$. Supposons que $|\langle x,y\rangle| = \|x\|\cdot\|y\| \neq 0$. Alors $t = -\frac{<x,y>}{<y,y>} \neq 0$ est la seule racine du polynôme ci-dessus. Donc, $\langle x+ty, x+ty\rangle = 0$ et $x + ty = 0$, i.e. x et y dont liés. $\square$

Corollaire 10.8. Si $a \in E$ alors l'application $E \to \mathbb{R}$, $x \mapsto \langle a, x\rangle$, est continue.

Démonstration. Soit $x_i \to x$ dans E, c.à.d. $\|x_i - x\| \to 0$. Par l'inégalité de Cauchy-Schwarz,

$$|\langle a, x - x_i\rangle| \leq \|a\| \cdot \|x - x_i\|.$$

D'où,

$$|\langle a, x - x_i\rangle| = |\langle a, x\rangle - \langle a, x_i\rangle| \to 0.$$

□

Définition 10.9. Soit A un sous-ensemble de l'espace euclidien E. On définit

$$A^{\perp} = \{x \in E \colon \forall a \in A \ \langle a, x\rangle = 0\} .$$

On appelle $A^{\perp}$ l'*orthogonal* de A.

Exercice 10.10. Montrer que, pour $A \subseteq E$:

1. $A^{\perp}$ est un sous-espace vectoriel de E.
2. $A^{\perp}$ est fermé dans E.
3. $A^{\perp} = (\overline{A})^{\perp}$.

Le CM du 8/12/2021

Définition 10.11. Une famille $(a_i)_{i\in I}$ de vecteurs non nuls de l'espace euclidien E est dite *orthogonale* si pour tout $i \neq j$ on a $\langle a_i, a_j\rangle = 0$.

La famille est dite *orthonormale* si elle est orthogonale et de plus $\|a_i\| = 1$ pour tout $i \in I$.

Proposition 10.12. *Toute famille orthogonale est une famille libre.*

Démonstration. Soit $(a_i)_{i\in I}$ une famille orthogonale et $a_{i_1}, \ldots, a_{i_n}$, $\lambda_1, \ldots, \lambda_n$ tels que $\displaystyle\sum_{k=1}^{n} \lambda_k a_{i_k} = 0$.

Observons que pour tout $j \in \{1, \ldots, n\}$ on a

$$\begin{aligned}
\left\langle \sum_{k=1}^{n} \lambda_k a_{i_k}, a_{i_j} \right\rangle &= \sum_{k=1}^{n} \lambda_k \langle a_{ik}, a_{i_j}\rangle \\
&= \sum_{k\neq j} \lambda_k \langle a_{ik}, a_{i_j}\rangle + \lambda_j \|a_{i_j}\|^2 \\
&= \lambda_j \|a_{i_j}\|^2
\end{aligned}$$

On en déduit que $\lambda_j \|a_{i_j}\|^2 = 0$ pour tout j puis, puisque tous les a_{i_j} sont non nuls, que $\lambda_j = 0$ pour tout j : la famille est bien libre. □

Exercice 10.13. Montrer que, si $(a_i)_{i\in I}$ est une famille orthogonale de vecteurs de E on a pour tout $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ et tout $i_1, \ldots, i_n \in I$

$$\|\sum_{k=1}^{n} \lambda_k a_{i_k}\|^2 = \sum_{k=1}^{n} \lambda_k^2 \|a_{i_k}\|^2 .$$

Exercice 10.14. On se place dans l'espace $L^2([0,1], \mathcal{B}([0,1]), \lambda)$ muni du produit scalaire $\langle f, g\rangle = \displaystyle\int_0^1 f(x)g(x)dx$. On considère la famille $(u_k)_{k\geq 1}$ définie par $u_k(x) = \sin(k\pi x)$ pour tout $x \in [0,1]$ et tout $k \geq 1$. Montrer que c'est une famille orthogonale.

10.2 Espaces métriques complets

La notion d'une suite de Cauchy, déjà introduite dans le cadre des e.v.n. (Définition 9.10), se généralise naturellement pour les espaces métriques.

Définition 10.15. Soit (X, d) un espace métrique, et (x_n) une suite d'éléments de X. On dit que (x_n) est une *suite de Cauchy* si la propriété suivante est vérifiée :

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

Autrement dit : à partir d'un certain rang, les termes de la suite deviennent arbitrairement proches les uns des autres.

Exercice 10.16. Soit (X, d) un espace métrique. Montrer que toute suite convergente est de Cauchy, et que la réciproque est fausse en général (on pourra par exemple considérer la suite définie par $x_n = 2^{-n}$ dans l'espace $X =]0, +\infty[$ muni de sa distance usuelle).

Exercice 10.17. Soit (X, d) un espace métrique, et (x_n) une suite de Cauchy d'éléments de X. Montrer que (x_n) est bornée.

Définition 10.18. Soit (X, d) un espace métrique. On dit que (X, d) est *complet* si toute suite de Cauchy d'éléments de X est convergente.

Remarque 10.19. D'après la Définition 9.11 un espace de Banach $(E, \|\cdot\|)$ muni de la distance induite par la norme $\|\cdot\|$ est un espace métrique complet.

Proposition 10.20. *Soit (X, d) un espace métrique complet et $F \subseteq X$ un sous-ensemble. Alors F est fermé dans X si et seulement si (F, d) est complet.*

Démonstration. Supposons que F est fermé. Soit (x_n) une suite de Cauchy d'éléments de F. C'est en particulier une suite de Cauchy d'éléments de X, elle converge donc vers un certain $x \in X$ puisque (X, d) est complet. Comme F est fermé dans X, on doit avoir $x \in F$. On vient de montrer que toute suite de Cauchy d'éléments de F converge dans F : (F, d) est complet.

Montrons la réciproque, i.e. on supposons que (F, d) est complet. Soit (x_n) une suite d'éléments de F qui converge vers $x \in X$. Alors (x_n) est de Cauchy puisque toute suite convergente est de Cauchy ; comme F est complet (x_n) converge dans F, ce qui montre que $x \in F$. $\square$

Proposition 10.21. *Soit (X, d) un espace métrique, et (x_n) une suite de Cauchy d'éléments de X telle que (x_n) admette une sous-suite (x_{n_k}) convergente. Alors (x_n) converge.*

Démonstration. Soit (x_n) une suite de Cauchy d'éléments de X, et $x \in X$ tel qu'il existe une sous-suite (x_{n_k}) qui converge vers x. Fixons $\varepsilon > 0$. Alors on sait qu'il existe N tel que $d(x_n, x_m) \leq \varepsilon$ pour tout $n, m \geq N$. De plus, il existe aussi K tel que $d(x_{n_k}, x) \leq \varepsilon$ pour tout $k \geq K$. Comme n_k tend vers $+\infty$, on peut trouver k_0 tel qu'on ait à la fois $k_0 \geq K$ et $n_{k_0} \geq N$. Par conséquent, pour tout $n \geq N$ on a à la fois $d(x_n, x_{n_{k_0}}) \leq \varepsilon$ et $d(x_{n_{k_0}}, x) \leq \varepsilon$, donc aussi $d(x_n, x) \leq 2\varepsilon$. Ceci prouve que (x_n) converge vers x. $\square$

Corollaire 10.22. Tout espace métrique compact est complet.

Démonstration. Soit (X, d) un espace métrique compact, et (x_n) une suite de Cauchy d'éléments de X. Par définition de la compacité, on peut extraire une sous-suite de (x_n) qui converge vers $x \in X$. La proposition précédente nous permet donc de conclure que (x_n) converge vers x. $\square$

Ceci nous fournit des exemples simples d'espaces complets.

Théorème 10.23. *Soit $n \in \mathbb{N}^*$ et $\|\cdot\|$ une norme sur $\mathbb{R}^n$. Alors $(\mathbb{R}^n, \|\cdot\|)$ est complet.*

Démonstration. Soit (x_n) une suite de Cauchy d'élements de $(\mathbb{R}^n, \|\cdot\|)$. Alors (x_n) doit être bornée et, donc, l'adhérence $\overline{\{x_n : n \in \mathbb{N}\}}$ est compact. Le corollaire 10.22 implique la convergence de (x_n). $\square$

Exercice 10.24. Soit $(E, \|\cdot\|)$ un $\mathbb{R}$-espace vectoriel, et $F \subseteq E$ un sous-espace de dimension finie. Montrer que F est fermé dans E.

Remarque 10.25. Par le théorème de Riesz-Fisher (Théorème 9.13) si $(X, \mathcal{A}, \mu)$ est un espace mesuré et $p \in [1, +\infty]$ alors l'espace $L^p(X)$ est un espace de Banach, donc, un espace métrique complet. Si $p = 2$ alors $L^2(X)$ est, *en plus*, un espace euclidien. (Voir l'inégalité de Hölder.) En effet, les espaces $L^2(X)$ procurent d'exemples dans le cas de dimension infinie des espaces de Hilbert définis ci-dessous.

Définition 10.26. Un espace euclidien *complet* s'appelle **espace de Hilbert**.

Remarque 10.27. On sait qu'un espace euclidien de dimension fini et toujours complet. La notion d'espace de Hilbert fait de sens dans le cas des espaces euclidiens de dimension infini (souvent des espaces de fonctions) qui sont rarement complets.

Exemple 34. Donons un exemple d'espace euclidien de dimension infini qui n'est pas complet, donc, n'est pas un espace de Hilbert. Soit $C([-1,1])$ l'ensemble des fonctions continues sur $[-1,1]$. Alors $C([-1,1])$ est un sous-espace de $L^2([-1,1], \mathcal{B}(\mathbb{R}), \lambda_1)$ (noté aussi par $L^2([-1,1])$) et le produit scalaire sur $L^2([-1,1])$ induit un produit scalaire sur $C([-1,1])$. Par conséquant, $C([-1,1])$ est un espace euclidien de dimension infinie. Il reste de montrer que $C([-1,1])$ n'est pas complet.

On considères la suite de fonctions

$$\phi_n(t) = \begin{cases} -1 & \text{si } -1 \le t \le \frac{-1}{n} \\ nt & \text{si } \frac{-1}{n} \le t \le \frac{1}{n} \\ 1 & \text{si } \frac{1}{n} \le t \le 1 \end{cases}$$

C'est une suite de Cauchy dans $C([-1,1])$ parce que

$$\|\phi_n - \phi_m\|_2^2 := \int_{-1}^{1} (\phi_n(t) - \phi_m(t))^2 dt \le \frac{2}{\min(n,m)}.$$

Supposons par l'absurd que $\phi_n \to f$ où $f \in C([-1,1])$. Soit ψ la fonction *discontinue* égale à -1 si $-1 \le t < 0$ et égale à $+1$ si $0 \le t \le 1$. Par l'inégalité de Minkowski (valable dans $L^2([-1,1])$)

$$\|f - \psi\|_2 \le \|f - \phi_n\|_2 + \|\phi_n - \psi\|_2.$$

Par la continuité de f et la discontinuité de ψ, $\|f - \psi\|_2 > 0$. On voit facilement que $\|\phi_n - \psi\|_2 \to 0$ et, par l'hyphothèse, $\|\phi_n - f\|_2 \to 0$. D'où, $\|f - \psi\|_2 = 0$. Absurd.

Remarque 10.28. Selon l'exercice 1 de la feuille d'exercices VIII chaque sous-espace de dimension finie d'un e.v.n. est fermé. Par contre, l'exemple ci-dessus montre que le sous espace $C([-1,1])$ de $L^2([-1,1], \mathcal{B}(\mathbb{R}), \lambda_1)$ n'est pas fermé.

10.3 Projection sur un convexe fermé

On désigne par E un espace euclidien. Rappelons que E est un espace de Hilbert si E est complet.

Définition 10.29. Soit A une partie non vide de E et $x \in E$. On dit que $a_0 \in A$ est une *projection* de x sur A si on a

$$\|x - a_0\| = \inf\{\|x - a\| \colon a \in A\} \ .$$

Remarque 10.30. Il n'existe pas toujours une projection ; et même quand elle existe, elle n'est pas nécessairement unique. Les exercices suivants donnent des exemples de ce phénomène.

Exercice 10.31. Soit E un $\mathbb{R}$-espace vectoriel muni d'un produit scalaire possédant une famille orthonormale $(e_n)_{n>0}$. On considère $F = \left\{(1+\frac{1}{n})e_n \colon n > 0\right\}$.

1. Montrer que 0_E n'a pas de projection sur F.
2. Montrer que pour tout $n \neq m$, $\|(1+\frac{1}{n})e_n - (1+\frac{1}{m})e_m\| \geq 1$. En déduire que F est fermé. (Plus précisement, il faut montrer que chaque point de F est fermé c.à.d. que F est *discret.*)
3. Expliciter un tel exemple pour l'espace d'Hilbert $L^2([0,1])$.

Exercice 10.32. On considère un triangle équilatéral dans $\mathbb{R}^2$. Montrer que son centre de gravité a trois projections.

Exercice 10.33. Soit S^{n-1} la sphère unité dans $\mathbb{R}^n$ centrée dans $0 \in \mathbb{R}^n$. Trouver les projections de 0 sur S^{n-1}.

Les deux théorèmes suivants ont été déjà établis pour l'espace euclidien $\mathbb{R}^n$ (voir Théorème 8.22).

Théorème 10.34. *Soit E un espace de Hilbert, $A \subseteq E$ un sous-ensemble convexe et fermé. Alors tout point $x \in E$ admet une unique projection $p_A(x)$ sur A.*

Démonstration. Fixons $x \in E$, et notons $\delta = \inf_{a \in A} \|x - a\|$. On peut trouver une suite (a_n) d'éléments de A tels que $\|x - a_n\|$ converge vers δ. On va prouver que (a_n) est une suite de Cauchy ; alors on pourra conclure (comme E est complet) que (a_n) converge vers a, qui appartient à A puisque A est fermé. Par continuité de la norme on aura alors $\|x-a\| = \lim_{n\to+\infty} \|x-a_n\| = \delta$. Ceci montrera l'existence d'une projection (mais pas encore son unicité).

Pour justifier que (a_n) est de Cauchy, on va utiliser la convexité de A et utiliser le fait que $\frac{a+b}{2} \in A$ pour tout $a, b \in A$. En particulier, pour tout $a, b \in A$ on doit avoir

$$\|x - \frac{a+b}{2}\| \geq \delta \ .$$

Fixons $n, m \in \mathbb{N}$ et appliquons l'identité du parallélogramme à $x - a_n$ et $x - a_m$:

$$\|(x - a_n) + (x - a_m)\|^2 + \|(x - a_n) - (x - a_m)\|^2 = 2\|x - a_n\|^2 + 2\|x - a_m\|^2 \ .$$

On en déduit

$$\begin{aligned}
\|a_n - a_m\|^2 &= 2\|x - a_n\|^2 + 2\|x - a_m\|^2 - \|2x - a_n - a_m)\|^2 \\
&= 2\|x - a_n\|^2 + 2\|x - a_m\|^2 - 4\|x - \frac{a_n + a_m}{2}\|^2 \\
&\leq 2\|x - a_n\|^2 + 2\|x - a_m\|^2 - 4\delta^2
\end{aligned}$$

Fixons $\varepsilon > 0$. On sait qu'il existe N tel que pour tous $n, m \geq N$ on a $\|x - a_n\|^2 \leq \delta^2 + \varepsilon$ et $\|x - a_m\|^2 \leq \delta^2 + \varepsilon$, et donc aussi $\|a_n - a_m\|^2 \leq 4\varepsilon$. On vient de montrer que (a_n) est de Cauchy.

Pour montrer que le projeté est unique, considérons $a, b \in A$ tels que $\|x - a\| = \|x - b\| = \delta$. Par un calcul similaire que ci-dessus on obtient

$$\|a - b\|^2 = 2\|x - a\|^2 + 2\|x - b\|^2 - \|2x - (a + b)\|^2 \leq 4\delta^2 - 4\delta^2 = 0.$$

Ceci n'est possible que si $a = b$. $\square$

Théorème 10.35. *Soient E un espace euclidien (de dimension fini ou infini) et A un convexe fermé de E.*

(a) Soit $x \in E$ et $p_A(x)$ le projeté de x sur A. Alors $p_A(x)$ est l'unique $a \in A$ tel que

$$\forall b \in A \quad \langle x-a, b-a\rangle \leq 0 \ .$$

(b) Supposons que $p_A(x)$ existe pour tout $x \in E$ (qui est le cas si E est de dimension finie ou E est un espace de Hilbert). Alors l'application $x \mapsto p_A(x)$ est 1-lipschitzienne (Définition 3.5).

Première preuve de (*a*). Notons $a = p_A(x)$. Pour tout $b \in A$, on a $\|x-a\|^2 \leq \|x-b\|^2$ par définition du projeté. En utilisant la relation

$$\begin{aligned}\|x-b\|^2 &= \|(x-a)+(a-b)\|^2 \\ &= \|x-a\|^2 + \|a-b\|^2 + 2\langle x-a, a-b\rangle,\end{aligned}$$

on en déduit que pour tout $b \in A$

$$\begin{aligned}2\langle x-a, b-a\rangle &= \|x-a\|^2 - \|x-b\|^2 + \|a-b\|^2 \\ &\leq \|a-b\|^2 \ .\end{aligned}$$

Fixons maintenant $b \in A$, différent de a ; pour tout $t \in]0,1]$ on peut appliquer l'inégalité précédente à $b_t = tb + (1-t)a$, qui appartient à A, et obtenir en utilisant le fait que $b_t - a = t(b-a)$

$$2\langle x-a, t(b-a)\rangle \leq t^2\|b-a\|^{\ 2}.$$

En divisant des deux côtés par t et en faisant tendre t vers 0, on voit que ceci n'est possible que si $\langle x-a, b-a\rangle \leq 0$, ce qu'on voulait démontrer.

Supposons maintenant qu'on a $a' \in A$ tel que $\langle x-a', b-a'\rangle \leq 0$ pour tout $b \in A$. Avec un calcul similaire au précédent, on a

$$\|x-a'\|^2 = \|x-a\|^2 + 2\langle x-a', a-a'\rangle - \|a-a'\|^2 \ .$$

Puisque $\langle x-a', a-a'\rangle \leq 0$ par hypothèse sur a', on a donc que $\|x-a'\|^2 \leq \|x-a\|^2$, ce qui n'est possible que si $a = a'$ par définition d'un projeté. □

Deuxième preuve de (*a*). Comme ci-dessus, $a = p_A(x)$. Supposons par la contraposée qu'il existe $b \in A$ tel que $\langle x-a, b-a\rangle > 0$. Soit V le sous espace de E engendré par $x-a$ et $b-a$. Le produit scalaire sur E induit un produit scalaire sur V. D'où, V est un plan euclidien qu'on peut identifié au plan $\mathbb{R}^2$. On considère le triangle de sommets $0, x-a, b-a$. L'angle entre $x-a$ et $b-a$ est $< \frac{\pi}{2}$ car $\langle x-a, b-a\rangle > 0$. Soit $0 < \lambda < 1$. Alors si λ est suffisament proche de 0, $a + \lambda(b-a) \in A$ et $\|x-a\| > \|x-(a-\lambda(b-a))\|$. Contradiction.

Si $a' \in A$ et tel que $\langle x-a', b-a'\rangle \leq 0$ pour tout $b \in A$ alors $\langle x-a', a-a'\rangle \leq 0$ et $\langle x-a, a'-a\rangle = \langle -x+a, a-a'\rangle \leq 0$. D'où, $\langle a-a', a-a'\rangle = \|a-a'\|^2 \leq 0$ et $a = a'$. □

Preuve de (*b*). Montrons que $x \mapsto p_A(x)$ est 1-lipschitzienne. Soit $x, x' \in X$. D'après (a), on a :

$$\langle x-p_A(x), p_A(x')-p_A(x)\rangle = \langle -x+p_A(x), p_A(x)-p_A(x')\rangle \leq 0 \quad \text{et} \quad \langle x'-p_A(x'), p_A(x)-p_A(x')\rangle \leq 0 \, .$$

Notons $r = x - x' + p_A(x') - p_A(x)$. En sommant les deux inégalités précédentes on obtient que

$$\langle p_A(x) - p_A(x'), x' - x + p_A(x) - p_A(x')\rangle \leq 0 \ ,$$

autrement dit, $\langle p_A(x) - p_A(x'), r\rangle \geq 0$. Finalement,

$$\begin{aligned}\|x-x'\|^2 &= \|p_A(x) - p_A(x') + r\|^2 \\ &= \|p_A(x) - p_A(x')\|^2 + \|r\|^2 + 2\langle p_A(x) - p_A(x'), r\rangle \\ &\geq \|p_A(x) - p_A(x')\|^2 \ .\end{aligned}$$

On a bien montré que p_A est 1-lipschitzienne. □

Remarque 10.36. Reprenons les notations du théorème précédent, et supposons de plus que $x \notin A$. L'ensemble $H = \{z \in E\colon \langle x-a, z-a\rangle = 0\}$ est un sous-espace affine de E (égal à $a + (\mathrm{Vect}(x-a))^{\perp}$), qui admet un supplémentaire de dimension 1 (la droite de vecteur directeur $x-a$) ; on dit que H est un *hyperplan* (affine). L'hyperplan H "coupe" E en deux demi-espaces : l'ensemble H^-des z tels que $\langle x-a, z-a\rangle \leq 0$ et l'ensemble H^+ des z tels que $\langle x-a, z-a\rangle > 0$. Le théorème ci-dessus nous dit que A est entièrement contenu dans H^-, tandis que x appartient à H^+. On dit alors que H *sépare* x et A. L'existence d'un hyperplan séparant un point d'un convexe fermé est une conséquence importante du théorème de projection. Par exemple, elle nous permet de voir que tout convexe fermé dans un espace de Hilbert est une intersection de demi-espaces fermés.

Un cas particulier très important est celui de la projection sur un sous-espace vectoriel fermé de E.

Théorème 10.37. *Soit E un espace de Hilbert et $F \subseteq E$ un sous-espace vectoriel fermé. Alors, pour tout $x \in E$, le projeté $p_F(x)$ est l'unique élément de F tel que $x - p_F(x)$ soit orthogonal à F. On le nomme* projeté orthogonal *de x sur F.*

La projection orthogonale sur F est une application linéaire continue.

Démonstration. Soit $x \in E$. Par définition, $p_F(x)$ appartient à F, et on a vu que $p_F(x)$ est l'unique élément de F tel que $\langle x - p_F(x), f - p_F(x)\rangle \leq 0$ pour tout $f \in F$. Mais comme $f + p_F(x) \in F$, on en déduit que $\langle x - p_F(x), f\rangle \leq 0$ pour tout $f \in F$ et, donc, $\langle x - p_F(x), f\rangle = 0$ pour tout $f \in F$. Réciproquement, si $a \in F$ est tel que $\langle x - a, f\rangle = 0$ pour tout $f \in F$, alors on a en particulier $\langle x-a, a\rangle = 0$ et donc $\langle x-a, f-a\rangle = 0$ pour tout $f \in F$, et d'après le théorème précédent cette propriété impose que $a = p_F(x)$.

Reste à montrer que p_F est linéaire (on sait déjà qu'elle est continue puisqu'elle est 1-lipschitzienne). Considérons $x, y \in E$ et $\lambda \in \mathbb{R}$. Alors pour tout $f \in F$ on a

$$\begin{aligned}\langle x + \lambda y - p_F(x) - \lambda p_F(y), f\rangle &= \langle x - p_F(x), f\rangle + \langle \lambda y - \lambda p_F(y), f\rangle \\ &= \langle x - p_F(x), f\rangle + \lambda\langle y - p_F(y), f\rangle \\ &= 0\end{aligned}$$

Par conséquent, $x + \lambda y - (p_F(x) + \lambda p_F(y))$ est orthogonal à tous les éléments de F, ce qui montre que $p_F(x + \lambda y) = p_F(x) + \lambda p_F(y)$. □

Corollaire 10.38. Soit E un espace de Hilbert, et F un sous-espace vectoriel fermé. Alors on a $E = F \oplus F^{\perp}$. De plus, $F^{\perp}$ est fermé et c'est l'unique supplémentaire de F orthogonal à F c.à.d. si $E = F \oplus G$ et G est orthogonal à F alors $G = F^{\perp}$.

Démonstration. On sait déjà que $F^{\perp}$ est un sous-espace vectoriel (fermé) de E ; de plus, si $x \in F \cap F^{\perp}$ alors $\|x\|^2 = \langle x, x\rangle = 0$. Enfin, pour tout $x \in E$ on a $x = (x - p_F(x)) + p_F(x)$, ce qui prouve que $x \in F^{\perp} + F$.

Soit G un supplémentaire de F orthogonal à F, c'est-à-dire $E = F \oplus G$ et $G \subset F^{\perp}$. Prenons $x \in F^{\perp}$. Alors, $x = f + g$ avec $f \in F$ et $g \in G$, d'où $f = x - g \in F \cap F^{\perp} = \{0_E\}$ et donc $x \in G$. Le fait que $F^{\perp}$ est fermé se déduit facilement du Corollaire 10.8 (en exercice). □

Exercice 10.39. Montrer qu'un sous-espace vectoriel F de E est dense dans E si, et seulement si, $F^{\perp} = \{0_E\}$.

Exercice 10.40. Soit $(a_i)_{i \in I}$ une famille de vecteurs de E. Montrer que le sous-espace vectoriel engendré par les $(a_i)_{i \in I}$ est dense dans E si, et seulement si,

$$\forall x \in E \quad (\forall i \in I\ \langle x, a_i\rangle = 0) \Leftrightarrow x = 0_E\ .$$

Exercice 10.41. Soit F un sous-espace vectoriel de E. Montrer que $(F^{\perp})^{\perp} = \overline{F}$.

Remarque 10.42. Soit $a \in E$ où E est un espace euclidien. Par le Corollaire 10.8 l'application $E \to \mathbb{R}, x \mapsto < a, x >$, est une forme linéaire continue. Lr théorème suivant montre que dans le cas d'espaces de Hilbert chaque forme linéaire *continue* est de ce type. Rappelons que dans le cas des espaces normés de dimension infinie ils existent de formes linéaires non-continues. (Voir Exemple 4.16.)

Théorème 10.43. *Soit E un espace de Hilbert et $\phi : E \to \mathbb{R}$ une forme linéaire continue. Alors il existe un unique $a \in E$ tel que $\phi(x) =< a, x >$ pour tour $x \in E$. De plus, $\|\phi\| = \|a\|$.*

Démonstration. Montrons l'existance de a. Si $\phi = 0$ on prend $a = 0$. Soit $\phi \neq 0$ et $F = \{x \in E | \phi(x) = 0\}$. Alors F est un sous-espace fermé, $E = F \oplus F^{\perp}$ et $F^{\perp} \neq (0)$. Donc, il existe $v \in F^{\perp}$ tel que $\phi(v) \neq 0$. Montrons que $F + \mathbb{R}v = E$. En effet, si $x \in E$ alors $\phi(x - \frac{\phi(x)}{\phi(v)} \cdot v) = 0$, i.e. $x - \frac{\phi(x)}{\phi(v)} \cdot v \in F$ qui montre que $F + \mathbb{R}v = E$.

On pose $a = \frac{\phi(v)}{<v,v>} \cdot v$. Alors $\phi(v) =< a, v >$ et $\phi(x) =< a, x >= 0$ si $x \in F$. Donc, $\phi(x) =< a, x >$ pour tour $x \in E$.

Pour montrer l'unicité de a, supposons que $< a, x >=< a', x >$ pour tous $x \in E$. Alors $< a - a', x >=< a - a', a - a' >= 0$, i.e. $\|a - a'\| = 0$ et $a = a'$.

Rappelons que $\|\phi\| = \sup_{\|v\|=1} \phi(v)$. Donc, $\|\phi\| = \sup_{\|v\|=1} < a, v >$. Par l'inégalité de Cauchy-Schwarz (Théorème 10.5) si $\|v\| = 1$ alors $| < a, v > | \leq \|a\|$ avec égalité si $v = \frac{a}{\|a\|}$. D'où, $\|\phi\| = \|a\|$. □

Définition 10.44. Soit $(a_i)_{i \in I}$ une famille de vecteurs de l'espace de Hilbert E. On dit que $(a_i)_{i \in I}$ est une *base hilbertienne* de E si la famille $(a_i)_{i \in I}$ est *orthonormale et engendre un sous-espace vectoriel dense.*

On rappelle que la famille $(a_i)_{i \in I}$ engendre un sous-espace vectoriel dense si, et seulement si le seul vecteur qui est orthogonal à chacun des a_i est le vecteur nul. Attention, une base hilbertienne d'un espace de Hilbert E n'est pas nécessairement une base algébrique ! Donc, un espace de Hilbert pourrait contenir un sous-espace propre et dense. Par contre, dans un espace euclidien E de dimension finie aucun sous-espace vectoriel propre n'est dense dans E (ceci suit du fait qu'en dimension finie, tout sous-espace vectoriel est fermé).

Exercice 10.45.

1. Soit E un espace euclidien de dimension finie et soit $\varphi : E \to \mathbb{R}$ une application linéaire. Montrer que φ est continue.
2. Soit E un espace de Hilbert. Supposons que E contient un sous-espace vectoriel qui est à la fois propre et dense. (Par exemple, si $E = L^2([-1, 1], \mathcal{B}(\mathbb{R}), \lambda_1)$ alors un théorème classique de Weierstrass implique que le sous espace des fonctions polynomiales à coefficients rationnels est propre et dense.) En déduire qu'il existe une application linéaire $\varphi : E \to \mathbb{R}$ qui n'est pas continue.

10.4 Procédé de Gram-Schmidt, inégalité de Bessel, identités de Parseval

Notons tout d'abord que la projection d'un vecteur sur un sous-espace vectoriel F de dimension finie se calcule facilement à l'aide d'une base orthonormale de F (preuve laissée en exercice) :

Proposition 10.46. *Soit E un espace de Hilbert et F un sous-espace vectoriel de dimension finie avec $(e_1, \ldots, e_n)$ une base orthonormale de F. Alors pour tout $x \in E$, on a*

$$p_F(x) = \sum_{i=1}^{n} \langle x, e_i \rangle e_i.$$

Rappelons que le procédé de Gram-Schmidt permet de calculer une base orthonormale d'un espace euclidien à partir d'une base donnée :

Proposition 10.47 (Procédé de Gram-Schmidt)**.** *Soit E un espace euclidien de dimension finie et $(e_1, \ldots, e_n)$ une base de E. Pour chaque $0 < i < n$, notons F_i le sous-espace vectoriel* $\mathrm{Vec}(e_1, \ldots, e_i)$ *engendré par $e_1, \ldots, e_i$. Alors, la famille $(e'_1, \ldots, e'_n)$ définie de la manière suivante est une base orthonormale de E :*

$$\begin{aligned} e'_1 &= \frac{e_1}{\|e_1\|} \\ e'_i &= \frac{e_i - p_{F_{i-1}}(e_i)}{\|e_i - p_{F_{i-1}}(e_i)\|} = \frac{e_i - \displaystyle\sum_{k=1}^{i-1}\langle e_i, e'_k\rangle e'_k}{\|e_i - \displaystyle\sum_{k=1}^{i-1}\langle e_i, e'_k\rangle e'_k\|} \quad \textit{pour } 1 < i \leq n \end{aligned}$$

Exercice 10.48. Vérifier que les vecteurs $e_1 = (1, 1, 1)$, $e_2 = (1, 1, -1)$ et $e_3 = (0, 1, 2)$ forment une base de $\mathbb{R}^3$. Utiliser le procédé de Gram-Schmidt sur cette base pour obtenir une base orthonormale.

Maintenant, étant donnée une base orthonormale d'un espace euclidien, on a les identités suivantes (à vérifier en exercice) :

Proposition 10.49. *Soit E un espace euclidien de dimension finie et $(e_1, \ldots, e_n)$ une base orthonormale de E. Alors pour tous x, y de E, on a :*

$$\begin{aligned} x &= \sum_{i=1}^{n}\langle x, e_i\rangle e_i \\ \|x\|^2 &= \sum_{i=1}^{n}|\langle x, e_i\rangle|^2 \\ \langle x, y\rangle &= \sum_{i=1}^{n}\langle x, e_i\rangle\langle y, e_i\rangle \end{aligned}$$

Plus loin dans cette partie (voir Théorème 10.55), nous allons voir que ces identités se généralisent aux espaces de Hilbert ayant une base hilbertienne dénombrable.

Définition 10.50. Un espace métrique est dit *séparable* s'il contient une partie dénombrable dense.

Exercice 10.51. Un théorème d'approximation de Weierstrass montre que toute fonction réelle définie et continue sur $[-1, 1]$ est limite uniforme de fonctions polynomiales à coefficients rationnels. Notons que l'ensemble des fonctions polynomiales à coefficients rationnels est dénombrable. En déduire que l'espace de Hilbert $L^2([-1, 1], \mathcal{B}(\mathbb{R}), \lambda_1)$ (noté souvant par $L^2([-1, 1])$) est séparable.

Proposition 10.52. *Un espace de Hilbert séparable possède une base hilbertienne dénombrable $(e_i)_{i\in\mathbb{N}}$.*

Démonstration. La preuve est laissée en exercice. Indication : on considère une énumération $(a_n)_{n\in\mathbb{N}}$ d'une partie dénombrable A dense, puis on en extrait une base du sous-espace vectoriel engendré par A, et enfin, en utilisant le procédé de Gram-Schmidt, on construit (par récurrence) une base orthonormale de ce sous-espace vectoriel. □

Exercice 10.53 (Polynômes de Legendre)**.** On se place dans l'espace de Hilbert $L^2([-1, 1])$.

1. En utilisant l'exercice précédent et le procédé de Gram-Schmidt, montrer qu'il existe une base hilbertienne $(p_n)_{n\in\mathbb{N}}$, telle que pour chaque $n \in \mathbb{N}$, la fonction p_n est une fonction polynomiale de degré n.

2. Pour $n \in \mathbb{N}$, on considère la fonction polynomiale l_n définie par $l_n(t) = \dfrac{1}{2^n n!}\dfrac{d^n}{dt^n}(t^2-1)^n$ pour $t \in [-1,1]$ (polynômes de Legendre).
   (a) Montrer que la famille $(l_n)_{n\in\mathbb{N}}$ est une famille orthogonale.
   (b) Déterminer le degré de l_n pour chaque n.
   (c) En déduire que pour chaque n on a $p_n = \dfrac{l_n}{\|l_n\|}$.

Théorème 10.54 (Inégalité de Bessel)**.** *Soit E un espace de Hilbert et $(e_i)_{i\in\mathbb{N}}$ une famille orthonormale. Alors pour tout $x \in E$,*

$$\sum_{i=0}^{+\infty}\langle x, e_i\rangle^2 \leq \|x\|^2.$$

Démonstration. Soit $n \geq 0$. La projection de x sur l'espace vectoriel engendré par $e_0, \ldots, e_n$ est égale à $\sum_{i=0}^{n}\langle x, e_i\rangle e_i$ et est orthogonale à $x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i$.

Cette orthogonalité entraîne que :

$$\|x\|^2 = \left\|\sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 + \left\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 = \sum_{i=0}^{n}\langle x, e_i\rangle^2 + \left\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\right\|^2 \geq \sum_{i=0}^{n}\langle x, e_i\rangle^2.$$

On conclut en passant à la limite. □

Théorème 10.55 (Identités de Parseval)**.** *Soit E un espace de Hilbert possédant une base hilbertienne dénombrable $(e_i)_{i\in\mathbb{N}}$. Alors, pour tous x, y de E, on a :*

$$\begin{aligned} x &= \sum_{i=0}^{+\infty}\langle x, e_i\rangle e_i \\ \|x\|^2 &= \sum_{i=0}^{+\infty}\langle x, e_i\rangle^2 \\ \langle x, y\rangle &= \sum_{i=0}^{+\infty}\langle x, e_i\rangle\langle y, e_i\rangle \end{aligned}$$

Démonstration. Notons F le sous-espace vectoriel engendré par la famille $(e_i)_{i\in\mathbb{N}}$ et pour chaque n, notons F_n le sous-espace vectoriel engendré par $e_0, \ldots, e_n$. Donc, $F = \underset{n}{\cup} F_n$. Comme F est dense dans E, il existe une suite (x_n) d'éléments de F convergeant vers x. Quitte à passer à sous-suite, à ajouter de termes nuls et à répéter des termes, on peut supposer que pour chaque n, le terme x_n appartient à F_n. Comme $\sum_{i=0}^{n}\langle x, e_i\rangle e_i$ est la projection de x sur F_n et x_n coincide avec sa projection sur F_n, on obtient ainsi l'inégalité

$$\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\| \leq \|x - x_n\|,$$

ce qui entraîne que $\|x - \sum_{i=0}^{n}\langle x, e_i\rangle e_i\|$ converge vers 0 et, par conséquant, $\sum_{i=0}^{n}\langle x, e_i\rangle e_i$ converge vers x. D'où, $x = \sum_{i=0}^{+\infty}\langle x, e_i\rangle e_i$.

En passant à la limite dans l'égalité

$$\|x\|^2 = \left\|\sum_{i=0}^{n} \langle x, e_i\rangle e_i\right\|^2 + \left\|x - \sum_{i=0}^{n} \langle x, e_i\rangle e_i\right\|^2 = \sum_{i=0}^{n} |\langle x, e_i\rangle|^2 + \left\|x - \sum_{i=0}^{n} \langle x, e_i\rangle e_i\right\|^2$$

on obtient la seconde identité du théorème.

Notons que par bilinéarité du produit scalaire on a,

$$\left\langle \sum_{i=0}^{n} \langle x, e_i\rangle e_i, \sum_{i=0}^{n} \langle y, e_i\rangle e_i \right\rangle = \sum_{i=0}^{n} \langle x, e_i\rangle \langle y, e_i\rangle.$$

Par continuité du produit scalaire, le terme de gauche converge vers $\langle x, y\rangle$ ce qui donne la dernière identité.

□

Remarque 10.56. Le théorème ci-dessus est valide quel que soit l'ordre de l'énumération de la base de Hilbert et donc indépendamment de l'ordre avec lequel l'on somme les séries.

Les identités de Parseval permettent de montrer réciproquement que si un espace de Hilbert a une base hilbertienne dénombrable alors il est séparable.

Exercice 10.57. Soit E un espace de Hilbert possédant une base hilbertienne dénombrable $(e_i)_{i\in\mathbb{N}}$. On considère la partie A des combinaisons linéaires finies de $(e_i)_{i\in\mathbb{N}}$ à coefficients rationnels. Notons que A est dénombrable. Montrer que A est dense dans E

Terminons cette partie en montrant l'unicité du développement en série dans le théorème 10.55. On étudiera un exemple important dans la partie suivante, les séries de Fourier.

Proposition 10.58 (unicité du développement)**.** *Soit $(e_i)_{i\in\mathbb{N}}$ une base hilbertienne d'un espace de Hilbert E. Soient $x \in E$ et $(\lambda_i)_{i\in\mathbb{N}}$ une famille de scalaires dans $\mathbb{R}$ tels que la série $\sum_{i=0}^{+\infty} \lambda_i e_i$ converge vers x. Alors pour tout $i \in \mathbb{N}$, on a $\lambda_i = \langle x, e_i\rangle$. On appellera ces scalaires, les coefficients de x dans la base $(e_i)_{i\in\mathbb{N}}$.*

Démonstration. Pour chaque n, posons $x_n = \sum_{i=0}^{n} \lambda_i e_i$. Par hypothèse, on a (x_n) qui converge vers x. Notons que pour $i \leq n$, on a $\langle x_n, e_i\rangle = \lambda_i$ et par continuité du produit scalaire (Corollaire 10.8), $\langle x_n, e_i\rangle$ converge vers $\langle x, e_i\rangle$ On obtient $\langle x, e_i\rangle = \lambda_i$.

□

10.5 Une application : les séries de Fourier

On munit l'espace $L^2([-\pi, \pi])$ du produit scalaire défini par

$$\langle f, g\rangle = \frac{1}{2\pi} \int_{[-\pi,\pi]} fg d\lambda\ .$$

Pour nous simplifier un peu les notations ci-dessous, notons $s_n(t) = \sin(nt)$ pour $n \geq 1$, et $c_n(t) = \cos(nt)$ pour $n \geq 0$ (en particulier c_0 est la fonction constante égale à 1). On a déjà vu (Exercice 10.14) que les fonctions (s_n) forment une famille orthogonale ; de plus le changement de variable $u = \frac{\pi}{2} - x$ permet de voir que pour $n \geq 1$ on a $\|s_n\| = \|c_n\|$ (la norme étant celle associée à notre produit scalaire : la constante multiplicative est là pour que $\|1\| = 1$), et comme $s_n^2 + c_n^2 = 1$, on déduit que pour $n \geq 1$ on a $\displaystyle\int_{-\pi}^{\pi} s_n^2 d\mu = \pi$, et donc $\|s_n\| = \frac{\sqrt{2}}{2}$, et de même pour c_n.

Le même type de calculs que ceux effectués pour la famille (s_n) permet de voir que la famille (c_m) est orthogonale, et aussi que pour tout n, m on a $\langle s_n, c_m \rangle = 0$. A titre d'exemple, montrons cette dernière égalité : pour $n \geq 1$ et $m \geq 0$ on a

$$\sin(nt)\cos(mt) = \frac{\sin((n+m)t) + \sin((n-m)t)}{2} \ .$$

De plus, $\displaystyle\int_{-\pi}^{\pi} \sin(kt) = 0$ pour tout k, ce qui prouve que $\displaystyle\int_{-\pi}^{\pi} \sin(nt)\cos(mt)dt = 0$ pour tout n, m.

Il ressort des calculs précédents que la famille de fonctions $((s_n)_{n\geq 1}, (c_m)_{m\geq 0}$ est une famille orthogonale ; cette famille n'est pas orthonormale mais le serait si pour $n \geq 1$ on remplaçait (s_n) par $\sqrt{2}s_n$ et c_n par $\sqrt{2}c_n$. L'espace vectoriel engendré par cette famille est appelé l'ensemble des ***polynômes trigonométriques***. Les formules de linéarisation nous permettent de vérifier qu'un produit de polynômes trigonométriques est encore un polynôme trigonométrique.

Théorème 10.59. *L'espace des polynômes trigonométriques est dense dans* $L^2([-\pi, \pi])$.

Démonstration. On doit montrer que, si $f \in L^2([-\pi, \pi]$ est telle que $\langle f, c_n \rangle = 0$ pour tout $n \geq 0$ et $\langle f, s_n \rangle = 0$ pour tout $n \geq 1$ alors f est la fonction nulle.

Commençons par supposer que f est une fonction continue et non nulle ; on souhaite montrer que f n'est pas orthogonale à tous les polynômes trigonométriques. Pour simplifier les notations, on va se contenter de traiter le cas où il existe $h > 0$ tel que $f(x) > 0$ pour tout $x \in]-h, h[$. (En remplaçant f par $\alpha \pm f$ pour un $\alpha \in \mathbb{R}$ bien choisi, la preuve ci-dessous s'adapte facilement au cas général.) Considérons alors le polynôme trigonométrique

$$P_n \colon x \mapsto (1 + \cos(x) - \cos(h))^n \ .$$

Pour $x \in]-h, h[$ on a que $P_n(x)$ croît (à x fixé) vers $+\infty$ quand $n \to +\infty$ car $\cos(x) - \cos(h) > 0$, tandis que pour $x \in [-\pi, -h[\cup]h, \pi[$ on a $P_n(x) \to 0$ et $|P_n(x)| < 1$. Mais alors :

- Sur $]-h, h[$, on peut appliquer le théorème de convergence monotone pour conclure que $\displaystyle\int_{]-h,h[} fP_n d\lambda$ tend vers $+\infty$.
- Sur $]-\pi, -h[\cup]h, \pi[$, on peut utiliser le fait que $|f(x)P_n(x)| \leq |f(x)| \leq \|f\|_\infty$ pour conclure, à l'aide du théorème de convergence dominée, que $\displaystyle\int_{]-\pi,-h[\cup]h,\pi[} fP_n d\lambda$ tend vers 0.

Tout ceci nous dit que $\displaystyle\int_{]-\pi,\pi]} fP_n d\lambda$ tend vers $+\infty$; en particulier on ne peut avoir $\langle f, P_n \rangle = 0$ pour tout n, par conséquent f n'est pas orthogonale à l'espace des polynômes trigonométriques.

Si maintenant on ne suppose plus f continue, mais seulement L^2, et que f est orthogonale aux polynômes trigonométriques, remarquons d'abord que $x \mapsto F(x) = \int_{-\pi}^{x} f d\lambda$ est bien définie (d'après Hölder, f est intégrable puisque la fonction constante 1 appartient à $L^2([\pi, \pi])$ et continue : pour tout $x, y \in [-\pi, \pi]$ on a

$$|F(x) - F(y)| = \left| \int_y^x f d\lambda \right| \leq \sqrt{\int_y^x f^2 d\lambda} \sqrt{\int_y^x 1 d\lambda} \leq \|f\|_2 \sqrt{y - x} \ .$$

Soit maintenant P un polynôme trigonométrique, et Q une primitive de P (qui est de nouveau un polynôme trigonométrique). Alors

$$\begin{aligned}
\langle F, P\rangle &= \frac{1}{2\pi}\int_{-\pi}^{\pi} F(x)P(x)dt \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{-\pi}^{x} f(t)dt\right)P(x)dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\int_{t}^{\pi} P(x)dx\right)f(t)dt \text{ (par Fubini)} \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}(Q(\pi)-Q(t))f(t)dt \\
&= \langle Q(\pi)-Q, f\rangle \\
&= 0\ .
\end{aligned}$$

Ci-dessus, l'application de Fubini est justifiée par le fait que $f(t)P(x)$ est intégrable sur l'ensemble compact $\{(x,t)\in[-\pi,\pi]^2\colon t\le x\}$; et la dernière égalité vient du fait que f est supposée orthogonale à tous les polynômes trigonométriques. Comme F est continue, la première partie de notre raisonnement nous permet de conclure que $F=0$. On vient donc de conclure que, pour toute fonction $f\in L^2([-\pi,\pi])$ qui est orthogonale aux polynômes trigonométriques, on a pour tout $x,y\in[-\pi,\pi]$ que $\displaystyle\int_{[x,y]} f d\lambda = 0$. Il reste de déduire que $f=0$ λ-p.p. Soit C le clan engendré par les intervalles $[x,y]$. Alors $\int_A f d\lambda = 0$ pour tout $A\in C$. On sait que la tribue borélienne est engendrée par C. En veru de l'exercice 6.21, $f=0$ λ-p.p.

□

Une autre façon d'énoncer le théorème précédent est de dire que la famille $(c_0=1,(\sqrt{2}s_n)_{n\ge1},(\sqrt{2}c_n)_{n\ge1})$ est une base hilbertienne de $L^2([-\pi,\pi])$. Pour $f\in L^2([-\pi,\pi])$, introduisons ses *coefficients de Fourier* :

$$\begin{aligned}
a_0(f) = \langle f, c_0\rangle &= \frac{1}{2\pi}\int_{[-\pi,\pi]} f d\lambda \\
\forall n\ge 1,\ a_n(f) = 2\langle f, c_n\rangle &= \frac{1}{\pi}\int_{[-\pi,\pi]} f(t)\cos(nt)d\lambda(t) \\
\forall n\ge 1,\ b_n(f) = 2\langle f, s_n\rangle &= \frac{1}{\pi}\int_{[-\pi,\pi]} f(t)\sin(nt)d\lambda(t)
\end{aligned}$$

Puisque la famille $(c_0,(\sqrt{2}s_n)_{n\ge1},(\sqrt{2}c_n)_{n\ge1})$ est une base hilbertienne, on sait (Théorème 10.55) que pour tout $f\in[-\pi,\pi]$ on peut écrire :

$$f = \langle f, c_0\rangle + \sum_{n\ge1}\left(\langle f,\sqrt{2}c_n\rangle\sqrt{2}c_n + \langle f,\sqrt{2}s_n\rangle\sqrt{2}s_n\right)$$

$$\|f\|^2 = \langle f, c_0\rangle^2 + \sum_{n\ge1}(\langle f,\sqrt{2}c_n\rangle)^2 + (\langle f,\sqrt{2}s_n\rangle)^2\ .$$

En termes de coefficients de Fourier, on a donc, pour toute fonction $f\in L^2([-\pi,\pi])$ et presque tout $t\in[-\pi,\pi]$:

1. $f(t) = a_0(f) + \sum_{n\ge1} a_n(f)\cos(nt) + b_n(f)\sin(nt)$;

2. $\displaystyle \frac{1}{2\pi}\int_{-\pi}^{\pi} f^2 d\lambda = a_0(f)^2 + \frac{1}{2}\sum_{n\geq 1}(a_n(f))^2 + (b_n(f))^2.$

La série apparaissant dans la première égalité ci-dessus est appelée *série de Fourier* de f. Remarquons que, même si f est continue, on n'a justifié la première égalité ci-dessus que presque partout (c'est une égalité au sens des fonctions L^2) : il peut a priori y avoir des points où on n'a pas égalité entre la fonction et sa série de Fourier ; effectivement, on a besoin d'hypothèses supplémentaires (par exemple, si f est C^1 par morceaux) pour conclure que la série de Fourier de f converge en tout point vers f.

La deuxième égalité ci-dessus, souvent appelée égalité de Parseval, est exactement la deuxième identité du Théorème 10.55.

Exercice 10.60. Calculer les coefficients de Fourier de la fonction f définie sur $[-\pi, \pi]$ par $f(x) = x$. A l'aide de l'égalité de Parseval, en déduire l'égalité :

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6} .$$