Cours d'Analyse Réelle

MM003

Jean SAINT RAYMOND

Université Pierre et Marie Curie

Avant-propos

Ce texte a été rédigé pour servir de support écrit à un cours de Master 1 de l'Université Pierre-et-Marie Curie. Les chapitres 1, 2, 3 et 5, sauf quelques sections, font partie du programme de la Licence et sont donnés ici surtout à titre de référence. Le contenu essentiel du cours réside donc dans les chapitres 4, 6, 7, 8 et 9.

On a ajouté à la fin un appendice consacré à l'Axiome du Choix et au théorème de Zorn, qui sont utilisés à plusieurs reprises dans cet ouvrage (théorème de Tychonoff, théorème de Hahn-Banach) et rarement étudiés, tant dans le cursus de la Licence que dans celui du Master.

Ma gratitude va aux étudiants qui ont suivi ce cours durant l'année écoulée, et dont les réactions m'ont permis d'améliorer la présentation. Il va de soi que des erreurs se sont sûrement glissées dans le texte qui suit, et que je remercie toute personne qui pourra m'en signaler, et ainsi m'aider à accroître la valeur de ce document.

J. Saint Raymond

Mis à jour le 12 Novembre 2010

1

TOPOLOGIE

1.1 Espaces métriques

Soit E un ensemble. On veut définir sur E une notion de "proximité" qui permette de donner un sens à la convergence des suites de points de E.

Définition 1.1.1. *On appelle* distance *sur un ensemble E une fonction d de $E \times E$ dans $\mathbb{R}^+$ satisfaisant pour tout x, tout y et tout z de E*

$$d(x,y) = 0 \iff x = y \tag{i}$$

$$d(x,y) = d(y,x) \tag{ii}$$

$$d(x,z) \leq d(x,y) + d(y,z) \tag{iii}$$

Définition 1.1.2. *On appelle* espace métrique *un couple (E,d) où E est un ensemble et d une distance sur E.*

Définition 1.1.3. *Dans un espace métrique (E,d), on appelle* boule ouverte (resp. fermée) de centre x et de rayon r *l'ensemble des points y de E dont la distance à x est strictement inférieure (resp. inférieure ou égale) à r.*

Exemple 1.1.4.

Sur la droite réelle $\mathbb{R}$, la fonction $d : (x,y) \mapsto |x-y|$ est une distance, pour laquelle les boules ouvertes sont des intervalles ouverts et les boules fermées des intervalles fermés.
Sur le plan complexe $\mathbb{C}$ la fonction $d : (z,w) \mapsto |z-w|$ est une distance, pour laquelle les boules ouvertes sont des disques ouverts et les boules fermées des disques fermés.

Définition 1.1.5. *Deux distances d_1 et d_2 sur l'ensemble E sont dites* équivalentes *si pour tout $x \in E$ et tout $r > 0$, il existe $r_1 > 0$ et $r_2 > 0$ tels que la boule $B_1(x,r_1)$ de centre x et de rayon r_1 pour d_1 soit contenue dans la boule $B_2(x,r)$ de centre x et de rayon r pour d_2 et que la boule $B_2(x,r_2)$ soit contenue dans la boule $B_1(x,r)$.*

Définition 1.1.6. *Si X est une partie d'un espace métrique E, on appelle diamètre de X la borne supérieure (dans $[0,+\infty]$) des distances de deux points de X :*

$$\mathrm{diam}(X) = \sup_{x,y \in X} d(x,y) \ .$$

1.2 Ouverts

Définition 1.2.1. *Une partie U d'un espace métrique est dite* ouverte *si, pour tout point x de U, il existe une boule ouverte de centre x (et de rayon > 0) contenue dans U.*
Une partie F de E est dite fermée *si son complémentaire est ouvert.*

En d'autres termes, U est ouvert s'il est réunion de boules ouvertes. Il est clair que $\emptyset$ et E sont des parties ouvertes (et fermées) de E.

Théorème 1.2.2. *Toute boule ouverte de l'espace métrique E est ouverte. Toute boule fermée de l'espace métrique E est fermée.*

Démonstration : Soit y un point de la boule ouverte $B(x,r)$ de centre x et de rayon r. On a $d(x,y) < r$. Si on note $\rho = r - d(x,y) > 0$, on a $B(y,\rho) \subset B(x,r)$: en effet, si $z \in B(y,\rho)$,

$$d(x,z) \leqslant d(x,y) + d(y,z) < d(x,y) + \rho = r \ ,$$

ce qui montre que $z \in B(x,r)$.
De même, si y n'appartient pas à la boule fermée $\tilde{B}$ de centre x et de rayon r, on a $\rho = d(x,y) - r > 0$. Et la boule ouverte $B(y,\rho)$ est disjointe de $\tilde{B}$: en effet, si $z \in B(y,\rho) \cap \tilde{B}$, on doit avoir

$$d(x,y) \leqslant d(x,z) + d(z,y) < r + \rho = d(x,y) \ ,$$

ce qui est absurde. Ceci prouve que $E \setminus \tilde{B}$ est ouvert, donc que $\tilde{B}$ est fermé. ■

Théorème 1.2.3. *L'intersection de deux ouverts est un ouvert. La réunion d'une famille quelconque d'ouverts est ouverte.*
La réunion de deux fermés est fermée. L'intersection d'une famille quelconque de fermés est fermée.

Démonstration : Soient U et V deux ouverts. Si x est un point de $U \cap V$, il existe $r > 0$ et $r' > 0$ tels que $B(x,r) \subset U$ et $B(x,r') \subset V$. Alors $r'' = \min(r,r') > 0$ et $B(x,r'') \subset U \cap V$.
Si $(U_i)_{i\in I}$ est une famille de parties ouvertes de E, si $U = \bigcup_{i\in I} U_i$ et si $x \in U$, il existe $i \in I$ tel que $x \in U_i$, donc $r > 0$ tel que $B(x,r) \subset U_i$. Alors $B(x,r) \subset U$, ce qui prouve que U est ouvert.
Les propositions analogues concernant les ensembles fermés s'en déduisent par passage au complémentaire. ■

Théorème 1.2.4. *Si d_1 et d_2 sont des distances équivalentes sur E, elles définissent les mêmes ouverts. Inversement, si d_1 et d_2 définissent les mêmes ouverts, elles sont équivalentes.*

Démonstration : Supposons les deux distances équivalentes. Si B est une boule ouverte pour d_1, pour tout $x \in B$, il existe $r > 0$ tel que la boule $B_1(x,r)$ soit contenue dans B. Mais alors , il existe $r_2 > 0$ tel que $B_2(x,r_2) \subset B_1(x,r) \subset B$. Donc B est réunion de boules ouvertes pour d_2, c'est-à-dire ouverte pour d_2. Et tout ouvert pour d_1, qui est réunion de boules ouvertes pour d_1 est donc réunion d'ouverts pour d_2, donc ouvert pour d_2. De même tout ouvert pour d_2 est ouvert pour d_1.
Inversement, si d_1 et d_2 définissent les mêmes ouverts, la boule de centre x et de rayon r pour d_2 est ouverte pour d_1 et contient x. Il existe donc un r_1 tel que $B_1(x,r_1)$ soit incluse dans $B_2(x,r)$. En inversant les rôles de d_1 et d_2, on obtient l'équivalence des distances. ■

Définition 1.2.5. *Si F est une partie fermée non vide de l'espace métrique E, on appelle* distance d'un point x de E au fermé F *le nombre positif ou nul*

$$d(x, F) = \inf_{z \in F} d(x, z) \ .$$

Il convient de remarquer que cette borne inférieure n'est pas nécessairement atteinte, et que, dans le cas général, il peut n'exister aucun point z de F tel que $d(x, z) = d(x, F)$.

Théorème 1.2.6. *La distance d'un point x au fermé F est nulle si et seulement si x appartient à F. De plus, si x et y sont des points de E, on a*

$$|d(x, F) - d(y, F)| \leqslant d(x, y) \ .$$

Démonstration : Si $x \in F$, on a clairement $0 \leqslant d(x, F) \leqslant d(x, x) = 0$. Inversement, si $x \notin F$, il existe une boule ouverte $B(x, r)$ disjointe de F, ce qui montre que, pour tout y de F, $d(x, y) \geqslant r$, donc que $d(x, F) \geqslant r > 0$.
Si z est un point quelconque de F, on a $d(x, z) \leqslant d(x, y) + d(y, z)$, donc, en passant à la borne inférieure $d(x, F) \leqslant d(x, y) + d(y, z)$ pour tout z de F. Et passant à nouveau à la borne inférieure $d(x, F) \leqslant d(x, y) + d(y, F)$, c'est-à-dire $d(x, F) - d(y, F) \leqslant d(x, y)$. En intervertissant x et y, on obtient $d(y, F) - d(x, F) \leqslant d(x, y)$, d'où l'inégalité cherchée. ■

1.3 Espaces topologiques

Plus généralement, on peut considérer la notion d'espace topologique. On appelle *espace topologique* un couple $(E, \mathscr{O})$ formé d'un ensemble E et d'une topologie $\mathscr{O}$ sur E, c'est-à-dire un ensemble de parties de E qu'on appelle les *ouverts* de E, satisfaisant les propriétés suivantes, dont nous avons vu qu'elles sont satisfaites pour les espaces métriques.

i) E et $\emptyset$ appartiennent à $\mathscr{O}$,
ii) $\mathscr{O}$ est stable par intersection finie ,
iii) $\mathscr{O}$ est stable par union quelconque.

Les espaces topologiques les plus intéressants vérifient en outre la condition suivante :

Définition 1.3.1. *Un espace topologique $(E, \mathscr{O})$ est dit* séparé *si, pour tout couple (x, y) de points distincts, il existe deux ouverts disjoints U et V contenant respectivement x et y.*

Théorème 1.3.2. *Tout espace métrique est séparé.*

Démonstration : Soient en effet x et y deux points distincts de l'espace métrique E. Si on note $\delta = d(x, y) > 0$, $U = B(x, \delta/2)$ et $V = B(y, \delta/2)$, U et V sont des ouverts disjoints : en effet si z appartenait à $U \cap V$, on aurait

$$\delta = d(x, y) \leqslant d(x, z) + d(y, z) < \delta/2 + \delta/2 = \delta \ ,$$

ce qui est absurde. ■

Définition 1.3.3. *Un point a d'un espace topologique E est dit* isolé *si le singleton $\{a\}$ est ouvert dans E.*

Définition 1.3.4. *Sur un ensemble E on appelle* topologie discrète *la topologie pour laquelle toutes les parties de E sont ouvertes.*

Un espace topologique est discret si et seulement si tous ses points sont isolés. Il est clair qu'un espace discret (c'est-à-dire muni de la topologie discrète) est séparé.

Définition 1.3.5. *Dans un espace topologique, une partie V est appelée* voisinage *d'un point x s'il existe un ouvert U contenant x et contenu dans V.*

Dans un espace métrique, V est un voisinage de x s'il existe une boule de rayon non nul centrée en x et contenue dans V.

Théorème 1.3.6. *L'intersection de deux voisinages du point x est un voisinage de x.*

Démonstration : Soient V_1 et V_2 deux voisinages de x. Il existe donc deux ouverts U_1 et U_2 contenant x et contenus respectivement dans V_1 et V_2. Alors $U_1 \cap U_2$ est un ouvert contenant x et contenu dans $V_1 \cap V_2$., ce qui montre que $V_1 \cap V_2$ est un voisinage de x. ■

Théorème 1.3.7. *Dans un espace topologique, une partie X de E est ouverte si et seulement si elle est un voisinage de chacun de ses points.*

Démonstration : Par définition, un ouvert est voisinage de chacun de ses points. Inversement, si X est voisinage de chacun de ses points, il existe pour tout $x \in X$ un ouvert U_x contenant x et inclus dans X. Alors $X = \bigcup_{x \in X} U_x$ est ouvert, comme réunion d'une famille d'ouverts. ■

Définition 1.3.8. *Une famille $\mathscr{V}$ de parties d'un espace topologique E est appelée* base de voisinages *d'un point a si elle est formée de voisinages de a et si tout voisinage de a contient un élément V de $\mathscr{V}$.*

Les boules centrées en a dans un espace métrique forment une base de voisinages de a.
On peut définir une topologie $\mathscr{O}$ sur un ensemble E en associant à chaque point a de E une famille $\mathscr{V}_a$ de parties de E en sorte que $\mathscr{V}_a$ soit une base de voisinages de a. Les ouverts pour $\mathscr{O}$ seront alors les ensembles U tels que, pour tout x de U, il existe un $V \in \mathscr{V}_x$ tel que $V \subset U$, c'est-à-dire les ensembles qui sont voisinages de chacun de leurs points. Pour cela, on doit avoir les deux conditions suivantes :

i) $\forall a \in E,\ \forall V \in \mathscr{V}_a \quad a \in V,$

ii) $\forall a \in E,\ \forall V \in \mathscr{V}_a,\ \exists W \in \mathscr{V}_a,\ \forall x \in W,\ \exists V_1 \in \mathscr{V}_x \quad V_1 \subset V.$

La première condition signifie qu'un voisinage de a contient a, et la seconde qu'un voisinage de a est voisinage aussi de tous les points x d'un voisinage de a.

Intérieur et adhérence.

Définition 1.3.9. *Un point x est dit* adhérent *à une partie X de l'espace topologique E si tout voisinage de x rencontre X. La partie X est dite* partout dense *dans E (ou simplement* dense *dans E) si tout point de E est adhérent à X.*

Théorème 1.3.10. *Si X est une partie de l'espace topologique E, il existe un plus grand ouvert contenu dans X, qu'on appelle* l'intérieur *de X, et un plus petit fermé contenant X, qu'on appelle* l'adhérence *de X.*

Démonstration : Si on note $\overset{\circ}{X}$ l'ensemble des points de E dont X est un voisinage, il est clair que $\overset{\circ}{X}$ est la réunion de tous les ouverts contenus dans X, donc que $\overset{\circ}{X}$ est ouvert, contenu dans X, et qu'il contient tout ouvert contenu dans X. C'est donc le plus grand ouvert contenu dans X.
De même, on voit que l'ensemble $\bar{X}$ des points adhérents à X est un fermé contenant X, et que tout fermé contenant X contient $\bar{X}$. ■

Sous-espaces et produits.

Définition 1.3.11. *Si X est une partie d'un espace topologique E, l'ensemble $\mathscr{O}_X$ des parties de X de la forme $X \cap U$ avec $U \in \mathscr{O}$, est une topologie sur X. L'ensemble X muni de cette topologie est appelé* sous-espace topologique *de E.*

Dans le cas où (E, d) est un espace métrique, le sous-espace X est l'espace topologique associé à l'espace métrique obtenu en munissant X de la distance d restreinte à $X \times X$.

Il est clair que l'ensemble $\mathscr{O}_X$ est une topologie sur X puisque $(X \cap U) \cap (X \cap V) = X \cap (U \cap V)$ et que $\bigcup_{i \in I}(X \cap U_i) = X \cap \bigcup_{i \in I} U_i$.
Si d_X est la restriction de la distance d à la partie X de l'espace métrique E, il est clair que d_X est une distance sur X, et que la boule ouverte de centre x et de rayon r pour d_X est la trace sur X de la boule ouverte de centre x et de rayon r pour d, donc que les ouverts de X, qui sont les réunions de boules ouvertes pour d_X sont les traces sur X des réunions de boules ouvertes pour d, c'est-à-dire d'ouverts de E.

Définition 1.3.12. *Soient (E_1, d_1) et (E_2, d_2) deux espaces métriques. Le* produit *de ces deux espaces est l'ensemble $E_1 \times E_2$ muni de la distance δ définie par*

$$\delta((x_1, x_2), (y_1, y_2)) = \max(d(x_1, y_1), d(x_2, y_2)) \ .$$

On peut remarquer que δ est équivalente à la distance δ_1 définie par

$$\delta_1((x_1, x_2), (y_1, y_2)) = d(x_1, y_1) + d(x_2, y_2) \ ,$$

puisque $\delta \leqslant \delta_1 \leqslant 2\,\delta$.
La boule de centre (x_1, x_2) et de rayon r est le produit $B_1 \times B_2$, où B_i est la boule de centre x_i et de rayon r pour d_i dans E_i ($i = 1, 2$). Si U_1 et U_2 sont deux ouverts de E_1 et E_2 respectivement, $U_1 \times U_2$ est ouvert dans $E_1 \times E_2$: en effet, si $(x_1, x_2) \in U_1 \times U_2$, il existe $r_1 > 0$ et $r_2 > 0$ tels que $B_1(x_1, r_1) \subset U_1$ et $B_2(x_2, r_2) \subset U_2$. Alors, si $r = \min(r_1, r_2) > 0$, la boule de centre (x_1, x_2) et de rayon r pour δ est incluse dans $U_1 \times U_2$.
On en déduit qu'une partie X de $E_1 \times E_2$ est ouverte si elle est réunion d'ouverts "élémentaires" de la forme $U_1 \times U_2$. On peut généraliser ce qui précède à un nombre fini quelconque d'espaces métriques (E_1, d_1), (E_2, d_2), $\ldots$, (E_n, d_n). On définit la distance δ par

$$\delta((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) = \max_{1 \leq i \leq n} d_i(x_i, y_i)$$

et on obtient à nouveau qu'un ensemble est ouvert s'il est réunion d'ouverts "élémentaires" de la forme $U_1 \times U_2 \times \cdots \times U_n$, où les U_i sont ouverts dans E_i.
Plus généralement encore, on définit le produit d'une suite d'espaces métriques $(E_n, d_n)_{n \in \mathbb{N}}$.

Définition 1.3.13. *Soit $(E_n, d_n)_{n \in \mathbb{N}}$ une suite d'espaces métriques. L'ensemble produit $E = \prod_{n \in \mathbb{N}} E_n$ est l'ensemble des suites $(x_n)_{n \in \mathbb{N}}$ où x_n appartient à E_n pour tout n. On munit E de la distance δ définie par*

$$\delta((x_n), (y_n)) = \max_{n \in \mathbb{N}}(\min(2^{-n}, d_n(x_n, y_n)))$$

En fait, on utilise très rarement une distance explicite sur l'espace produit, et toute distance équivalente à δ convient aussi bien.

Théorème 1.3.14. *Un ensemble X est ouvert dans le produit des E_n s'il est réunion d'ouverts élémentaires de la forme $\prod_{n\in\mathbb{N}} U_n$, où les U_n sont ouverts dans E_n et égaux à E_n pour tout n assez grand.*

Démonstration : Si X est ouvert dans $E = \prod_{n\in\mathbb{N}} E_n$, et si $x = (x_n)$ appartient à X, il existe $r > 0$ tel que, pour tout $y = (y_n)$, $\delta(x, y) < r \Longrightarrow y \in X$. Si $2^{-m} < r$ et si on définit

$$U_n = \begin{cases} \{y_n \in E_n : d_n(y_n, x_n) < r\} & \text{si } n \leqslant m \\ E_n & \text{si } n > m \end{cases} ,$$

alors, pour tout $y = (y_n) \in \prod_{n\in\mathbb{N}} U_n$, $\delta(x, y) < \max(r, 2^{-m}) = r$, donc $y \in X$, c'est-à-dire que $x \in \prod_{n\in\mathbb{N}} U_n \subset X$.
Inversement, si X est réunion d'ouverts élémentaires, pour tout $x = (x_n)$ de X, il existe des ouverts U_n contenant les x_n tels que $x \in \prod_{n\in\mathbb{N}} U_n \subset X$ et que $U_n = E_n$ si $n > m$. On choisit alors, pour $n \leqslant m$ un $r_n > 0$ tel que la boule de centre x_n et de rayon r_n dans E_n soit contenue dans U_n. Si on pose alors

$$r = \min(2^{-m}, \min_{n\leqslant m} r_n) ,$$

on a $r > 0$ et si $\delta(x, y) < r$, on a $y_n \in U_n$ pour tout $n \in \mathbb{N}$, donc $y \in X$, c'est-à-dire que X contient la boule de centre x et de rayon r, donc que X est ouvert. ■

Généralisation. Si (E_i) est une famille d'espaces topologiques, on définit sur l'ensemble produit $\prod_{i\in I} E_i$ une topologie pour laquelle les ouverts sont les réunions d'ouverts élémentaires, c'est-à-dire de produits $\prod_{i\in I} U_i$, où les U_i sont ouverts dans E_i et tous égaux à E_i *sauf un nombre fini* d'entre eux.

Théorème 1.3.15. *Un espace topologique E est séparé si et seulement si la diagonale $\Delta = \{(x, y) \in E \times E : x = y\}$ de E est fermée dans $E \times E$.*

Démonstration : Supposons E séparé, et $(x, y) \in E \times E \setminus \Delta$. Alors $x \neq y$ et il existe deux ouverts disjoints U et V dans E contenant respectivement x et y. Alors $U \times V$ est un voisinage ouvert de (x, y) dans $E \times E$, qui est disjoint de Δ.
Inversement, si Δ est fermé et si $x \neq y$, le point (x, y) n'appartient pas à Δ, et il existe un voisinage ouvert W de (x, y) qui est disjoint de Δ. Par définition de la topologie de $E \times E$, W contient un ouvert élémentaire $U \times V$, qui lui-même contient (x, y). Alors U et V sont des ouverts disjoints de E, contenant respectivement x et y. Donc E est séparé. ■

Suites convergentes.

Rappelons qu'une suite dans un ensemble E est une application : $n \mapsto x_n$ de $\mathbb{N}$ dans E.

Définitions 1.3.16. *On dit que la suite (x_n)* converge *vers x dans l'espace topologique E si pour tout voisinage V de x il existe un entier m tel que $x_n \in V$ pour $n > m$ (c'est-à-dire que les x_n appartiennent tous à V sauf un nombre fini). On dit alors que x est* limite *de la suite (x_n).*

On dit que x est point d'accumulation *ou* valeur d'adhérence *de la suite (x_n) si tout voisinage V de x contient une infinité de termes de la suite.*

Théorème 1.3.17. *Si X est une partie de E et si (x_n) est une suite de points de X, toute valeur d'adhérence, et en particulier toute limite, de la suite (x_n) appartient à l'adhérence de X.*

DÉMONSTRATION : En effet, si x est une valeur d'adhérence de (x_n), tout voisinage de x contient des points de la suite, donc des points de X. ■

Théorème 1.3.18. *Dans un espace topologique séparé (en particulier dans un espace métrique), une suite a au plus une limite.*

DÉMONSTRATION : Si x et x' sont deux limites distinctes de la suite (x_n), il doit exister deux voisinages disjoints V et V' de x et x' respectivement, donc deux entiers m et m' tels que x_n appartienne à V pour $n > m$ et à V' pour $n > m'$. Alors pour $n > \max(m, m')$, x_n doit appartenir à $V \cap V' = \emptyset$, ce qui est absurde. ■

Théorème 1.3.19. *Dans un espace métrique, tout point adhérent à une partie X est limite d'une suite d'éléments de X.*

DÉMONSTRATION : Soit en effet $x \in \overline{X}$. Pour tout entier n, il existe un point x_n de X dans la boule $B(x, 2^{-n})$, qui est un voisinage ouvert de x. Alors la suite (x_n) converge vers x. En effet, tout voisinage V de x contient une boule ouverte $B(x, r)$, donc tous les points x_n pour $n > m$, si m vérifie $2^{-m} \leq r$. ■

Définition 1.3.20. *Soient (x_n) et (y_n) deux suites dans E. On dit que la suite (y_n) est* extraite *de la suite (x_n) (ou que (y_n) est une* sous-suite *de la suite (x_n)) s'il existe une injection croissante $j : k \mapsto n_k$ de $\mathbb{N}$ dans $\mathbb{N}$ telle que $y_k = x_{n_k}$ pour tout entier k.*

Si H est une partie infinie de $\mathbb{N}$, il existe une seule injection croissante j de $\mathbb{N}$ dans lui-même telle que $H = j(\mathbb{N})$. Une sous-suite de la suite (x_n) est donc entièrement définie par l'ensemble H des indices n tels que x_n soit un terme de cette sous-suite. Si la suite extraite correspondante est convergente, on notera alors $\lim_{n\to\infty, n\in H} x_n$ sa limite.

Théorème 1.3.21. *Si x est une valeur d'accumulation d'une suite (x_n) dans un espace métrique, il existe une sous-suite de la suite (x_n) qui converge vers x.*

DÉMONSTRATION : Si on définit par récurrence sur k, $n_0 = 0$ et $n_{k+1} = \min\{n > n_k : d(x_n, x) < 2^{-k}\}$ pour $k \geqslant 0$, l'entier n_k est défini pour tout k si x est valeur d'accumulation de (x_n), et on a clairement $n_{k+1} > n_k$. Enfin la suite $(y_k) = (x_{n_k})$ converge vers x. ■

Proposition 1.3.22. *Une partie X d'un espace métrique est fermée si et seulement si elle contient la limite de chacune de ses suites convergentes.*

DÉMONSTRATION : Si X est fermé et si (x_n) est une suite de points de X qui converge vers x, $x \in \bar{X} = X$. Inversement, si X contient les limites de ses suites convergentes et si x est un point adhérent à X, il existe une suite de points de X qui converge vers x ; il en résulte que $x \in X$. Donc X contient tous ses points adhérents, et $X = \bar{X}$ est fermé. ■

Théorème 1.3.23. *Une suite (x^n) dans un produit d'espaces topologiques $(E_i)_{i\in I}$ converge vers $x = (x_i)$ si et seulement si, pour tout $i \in I$, la suite (x_i^n) converge vers x_i.*

Ceci se déduit aisément de la définition des voisinages ouverts élémentaires de x.
L'énoncé suivant sera utilisé ultérieurement à plusieurs reprises dans l'étude de la compacité.

Lemme 1.3.24. (Lemme diagonal de Cantor) *Soit (H_n) une suite décroissante de parties infinies de $\mathbb{N}$. Alors il existe une partie infinie H de $\mathbb{N}$ qui est presque incluse dans chacune des H_n, c'est-à-dire que tous les éléments de H sauf un nombre fini appartiennent à H_n (ou que les ensembles $(H \setminus H_n)$ sont des ensembles finis).*

Démonstration : Si on définit n_k comme le $k^{\text{ème}}$ élément de H_k (pour l'ordre usuel des entiers), et H comme l'ensemble des n_k pour $k \in \mathbb{N}$, il suffit de vérifier — que la suite (n_k) est strictement croissante : le $k+1^{\text{ème}}$ élément de H_{k+1} est supérieur au $k+1^{\text{ème}}$ élément de H_k puisque $H_{k+1} \subset H_k$, donc strictement supérieur au $k^{\text{ème}}$ élément de H_k — et que n_k appartient à H_n si $k \geqslant n$ puisque alors $H_k \subset H_n$. ∎

1.4 Applications continues

Définition 1.4.1. *Soit f une application d'un espace topologique E dans un espace topologique F. On dit que f est* continue en a *si pour tout voisinage W de $f(a)$, il existe un voisinage V de a tel que $f(V) \subset W$.*

Intuitivement, si f est continue en a les images des points "assez proches" de a sont proches de $f(a)$. Si f est continue en a et si W est un voisinage de $f(a)$, $f^{-1}(W)$, qui contient le voisinage V de a est lui-même un voisinage de a. Inversement, si $f^{-1}(W)$ est un voisinage de a, on a $f(V) \subset W$, pour le voisinage $V = f^{-1}(W)$ de a.

Théorème 1.4.2. *Soient E et F deux espaces topologiques, et f une application de E dans F. Si U est un voisinage de $a \in E$ et si la restriction $f|_U : U \to F$ de f à U est continue en a, f est continue en a.*

Démonstration : Si W est un voisinage de $f(a)$, $U \cap f^{-1}(W) = (f|_U)^{-1}(W)$ est un voisinage de a dans U, donc la trace sur U d'un voisinage V de a dans E. Et puisque $U \cap V$, intersection de deux voisinages de a dans E, est un voisinage de a dans E, on a

$$U \cap V = U \cap f^{-1}(W) \subset f^{-1}(W) ,$$

ce qui montre que $f^{-1}(W)$ est un voisinage de a dans E, donc que f est continue en a. ∎

Définition 1.4.3. *Soit f une application d'un espace topologique E dans un espace topologique F. On dit que f est continue sur E si elle est continue en chaque point de E.*

L'énoncé suivant est fondamental pour caractériser les fonctions continues.

Théorème 1.4.4. *Pour une fonction f de E dans F les trois propriétés suivantes sont équivalentes :*

i) *f est continue.*

ii) *Pour tout ouvert U de F, $f^{-1}(U)$ est ouvert dans E.*

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

Démonstration : Si f est continue et si U est ouvert dans F, $f^{-1}(U)$ est voisinage de chacun de ses points : si $a \in f^{-1}(U)$, $f(a) \in U$ et U est voisinage de $f(a)$, donc $f^{-1}(U)$ est voisinage de a.

Si l'image réciproque de tout ouvert est ouverte, $E \setminus f^{-1}(A) = f^{-1}(F \setminus A)$ est ouvert si A est fermé dans F, puisque $F \setminus A$ l'est ; donc $f^{-1}(A)$ est fermé.
Enfin, si l'image réciproque de tout fermé est fermée, si $a \in E$ et si W est un voisinage de $f(a)$, le point $f(a)$ n'appartient pas à l'adhérence A de $F \setminus W$. Donc a n'appartient pas au fermé $f^{-1}(A)$, et il existe un voisinage V de a qui est disjoint de $f^{-1}(A)$. Alors $f(V)$ est disjoint de A, donc contenu dans W. ■

Corollaire 1.4.5. *Soient E et F deux espaces topologiques, f une fonction de E dans F, A et B deux fermés de E tels que $E = A \cup B$. Si les restrictions de f à A et à B sont continues, alors f est continue.*

Démonstration : Il suffit de montrer que $f^{-1}(H)$ est fermé dans E pour tout fermé H de F. Or

$$f^{-1}(H) = f^{-1}(H) \cap (A \cup B) = \Big(A \cap f^{-1}(H)\Big) \cup \Big(B \cap f^{-1}(H)\Big)$$

et puisque $A_1 = A \cap f^{-1}(H)$ est l'image réciproque de H par $f_{|A}$, A_1 est fermé dans A, donc dans E puisque A est un fermé de E. Pour la même raison $B_1 = A \cap f^{-1}(H)$ est fermé dans E, et $f^{-1}(H) = A_1 \cup B_1$ est fermé dans E. ■

Théorème 1.4.6. *Si E et F sont des espaces métriques, si f est une application de E dans F et a un point de E, f est continue en a si et seulement si toute suite (x_n) qui converge vers a a une image par f qui converge vers $f(a)$.*

Démonstration : Si (x_n) converge vers a et si f est continue en a, il existe pour tout voisinage W de $f(a)$ un voisinage V de a tel que $f(V) \subset W$, donc un entier m tel que $x_n \in V$ pour tout $n > m$. Alors, pour tout $n > m$, $f(x_n) \in W$, c'est-à-dire que la suite $(f(x_n))$ converge vers $f(a)$.
Inversement, si f est discontinue en a, il existe un voisinage W de $f(a)$ tel que $X = f^{-1}(W)$ ne soit pas un voisinage de a. Pour tout entier n, la boule de centre a et de rayon 2^{-n} n'est pas incluse dans X. On peut donc trouver un point x_n dans cette boule qui n'appartienne pas à X. Puisque $d(a, x_n) < 2^{-n}$, la suite (x_n) converge vers a, mais la suite $(f(x_n))$ dont aucun terme n'appartient à W ne converge pas vers $f(a)$. ■

Théorème 1.4.7. *Si E, F et G sont trois espaces topologiques, f et g des applications continues de E dans F et de F dans G respectivement, alors $g \circ f$ est continue de E dans G.*

Démonstration : En effet, si U est un ouvert de G, $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$. L'ensemble $V = g^{-1}(U)$ est ouvert dans F puisque g est continue, et $f^{-1}(V)$ est ouvert dans E puisque f est continue. ■

Théorème 1.4.8. *Soit f une application de E dans F. Si X est une partie de F qui contient $f(E)$, f est continue de E dans F si et seulement si elle est continue de E dans le sous-espace X de F.*

Démonstration : Si f est continue de E dans F, et si U est un ouvert de X, il existe un ouvert V de F tel que $U = X \cap V$. Alors $f^{-1}(U) = f^{-1}(V)$ est ouvert dans E.
Inversement si f est continue de E dans X et si V est un ouvert de F, $f^{-1}(V) = f^{-1}(V \cap X)$ qui est ouvert dans E puisque $V \cap X$ est un ouvert de X. ■

Si f est une application de E dans un espace produit $\prod_{i \in I} F_i$, f est déterminée par ses applications coordonnées (f_i) de E dans F_i :

$$f(x) = (f_i(x))_{i \in I}$$

Théorème 1.4.9. *Une application f de E dans un espace produit est continue si et seulement si ses applications coordonnées sont continues.*

Démonstration : Remarquons d'abord que les projections $\pi_j : \prod_{i\in I} F_i \to F_j$ sont continues. En effet, si U_j est ouvert dans F_j, l'ensemble $\pi_j^{-1}(U_j)$ est un ouvert élémentaire $U_j \times \prod_{i\neq j} F_i$. Il en résulte que si f est continue, $f_j = \pi_j \circ f$ est continue.
Inversement, si les f_i sont toutes continues et si $U = \prod U_i$ est un ouvert élémentaire, on a $f^{-1}(U) = \bigcap_{i\in I} f_i^{-1}(U_i)$. Cet ensemble est une intersection d'ouverts tous égaux à E sauf pour les i appartenant à une partie finie J de I. Alors $f^{-1}(U) = \bigcap_{i\in J} f_i^{-1}(U_i)$ qui est ouvert. Et puisque un ouvert de $\prod F_i$ est réunion d'ouverts élémentaires, son image réciproque par f est un ouvert de E. ■

Théorème 1.4.10. *La somme et le produit sont des applications continues de $\mathbb{R}\times\mathbb{R}$ dans $\mathbb{R}$.*

Démonstration : Si (a,b) et (x,y) sont deux points de $\mathbb{R}\times\mathbb{R}$, on a

$$\begin{aligned} d((x+y),(a+b)) &= |(x+y)-(a+b)| = |(x-a)+(y-b)| \leqslant |x-a|+|y-b| \\ &\leqslant d(x,a)+d(y,b) \leqslant 2\max(d(x,a),d(y,b)) = 2\,d((x,y),(a,b)) \end{aligned}$$

Il suffit donc que la distance de (x,y) à (a,b) dans $\mathbb{R}\times\mathbb{R}$ soit inférieure à $r/2$ pour que la distance de $(x+y)$ à $(a+b)$ dans $\mathbb{R}$ soit inférieure à r, ce qui prouve la continuité de la somme en (a,b).
De même, si (a,b) et (x,y) sont deux points de $\mathbb{R}\times\mathbb{R}$, on a

$$\begin{aligned} d(xy,ab) = |xy-ab| &= |(x-a)(y-b)+a(y-b)+(x-a)b| \\ &\leqslant |x-a|\,|y-b|+|a|\,|y-b|+|b|\,|x-a| \end{aligned}$$

Si $r>0$ et si $\delta = \min(1, \dfrac{r}{1+|a|+|b|})$, on a pour (x,y) dans la boule de centre (a,b) et de rayon δ : $|x-a|<1$, $|x-a|<\delta$ et $|y-b|<\delta$. Donc

$$d(xy,ab) \leqslant \delta(1+|a|+|b|) \leqslant r$$

ce qui achève de prouver la continuité du produit en (a,b). ■

Théorème 1.4.11. *L'inversion $x\mapsto 1/x$ est continue de $\mathbb{R}^*$ dans $\mathbb{R}$.*

Démonstration : Si x et a sont non nuls, on a

$$\left|\frac{1}{x}-\frac{1}{a}\right| = \left|\frac{x-a}{ax}\right| \leqslant \frac{|x-a|}{|a|\,(|a|-|x-a|)}$$

Donc si $r>0$, et si $\delta = \min(\dfrac{|a|}{2}, \dfrac{ra^2}{2})$, on a, pour x dans la boule de centre a et de rayon δ, $|a|-|x-a| \geqslant \dfrac{|a|}{2}$, donc $d(1/x,1/a)<r$, ce qui prouve la continuité en a de l'inversion. ■

Théorème 1.4.12. *La somme et le produit sont continus de $\mathbb{C}\times\mathbb{C}$ dans $\mathbb{C}$ et l'inversion est continue de $\mathbb{C}^*$ dans $\mathbb{C}$.*

La démonstration utilise exactement les mêmes inégalités que pour le cas réel.

Corollaire 1.4.13. *Si f et g sont deux fonctions continues de l'espace topologique E dans $\mathbb{R}$ ou dans $\mathbb{C}$, $f+g$ et fg sont continues. De plus, si f ne s'annule pas, $1/f$ est continue.*

DÉMONSTRATION : La fonction $h : E \to \mathbb{R} \times \mathbb{R}$ (resp. $E \to \mathbb{C} \times \mathbb{C}$) dont les fonctions coordonnées sont f et g est continue d'après le théorème 1.4.9. Il suffit donc de voir que $f+g$ et fg sont les composées de h avec la somme et le produit, et que $1/f$ est la composée de f et de l'inversion de $\mathbb{R}^*$ (resp $\mathbb{C}^*$). ■

Définition 1.4.14. *Une application f d'un espace métrique E dans un espace métrique F est dite k-*lipschitzienne *(avec $k \in \mathbb{R}^+$) si pour tout x et tout y de E on a*

$$d(f(x), f(y)) \leqslant k.\, d(x, y)$$

Théorème 1.4.15. *Si f est lipschitzienne (c'est-à-dire k-lipschitzienne pour un certain k), f est continue.*

DÉMONSTRATION : En effet, la boule de centre x et de rayon r/k est contenue dans l'image réciproque par f de la boule de centre $f(x)$ et de rayon r. ■

Corollaire 1.4.16. *La fonction valeur absolue est continue de $\mathbb{R}$ dans $\mathbb{R}^+$. La fonction module est continue de $\mathbb{C}$ dans $\mathbb{R}^+$.*

DÉMONSTRATION : L'inégalité

$$\big|\, |x| - |y| \,\big| \leqslant |x - y|$$

montre que cette fonction est 1-lipschitzienne, donc continue. On en déduit comme plus haut que la valeur absolue d'une fonction réelle continue (ou le module d'une fonction complexe continue) est continue. ■

Théorème 1.4.17. *Si E est un espace métrique, la distance est une fonction continue de $E \times E$ dans $\mathbb{R}$.*

DÉMONSTRATION : Soient (x, y) et (x', y') deux points de $E\times$E. On a

$$d(x', y') \leqslant d(x, x') + d(x, y) + d(y, y') \ ,$$

donc

$$d(x', y') - d(x, y) \leqslant d(x, x') + d(y, y') \leqslant 2\, \max(d((x, x'), d(y, y')) = 2\, d((x, y), (x', y')) \ ,$$

et par symétrie

$$|d(x', y') - d(x, y)| \leqslant 2\, d((x, y), (x', y')) \ ,$$

ce qui montre que la fonction d est 2-lipschitzienne donc continue. ■

Théorème 1.4.18. *Si E est un espace métrique, et F un fermé non vide de E, la fonction distance à F est continue.*

DÉMONSTRATION : On a prouvé au théorème 1.2.6 l'inégalité $|d(x, F) - d(y, F)| \leqslant d(x, y)$. Il en résulte que la fonction distance à F est 1-lipschitzienne, donc continue. ■

Théorème 1.4.19. *Soient E un espace topologique, a un point de E, F un espace métrique et f une application de E dans F. On suppose que, pour tout $\varepsilon > 0$, il existe une fonction f_ε de E dans F continue en a, vérifiant pour tout x de E : $d(f(x), f_\varepsilon(x)) \leqslant \varepsilon$. Alors f est continue en a.*

Démonstration : Soit $r > 0$. On cherche un voisinage W de a tel que $f(W) \subset B(f(a), r)$. Prenons $\varepsilon = r/3$, et choisissons une fonction f_ε vérifiant la condition ci-dessus. Puisque f_ε est continue en a, il existe un voisinage W de a tel que $f_\varepsilon(W) \subset B(f_\varepsilon(a), \varepsilon)$. Alors, pour tout x de W on a

$$d(f(x), f(a)) \leqslant d(f(x), f_\varepsilon(x)) + d(f_\varepsilon(x), f_\varepsilon(a)) + d(f_\varepsilon(a), f(a)) < \varepsilon + \varepsilon + \varepsilon = r \ ,$$

puisque $d(f_\varepsilon(x), f_\varepsilon(a)) < \varepsilon$, ce qui entraîne que $f(x)$ appartient à la boule ouverte de centre $f(a)$ et de rayon r. ■

Définition 1.4.20. *On dit qu'une suite (f_n) de fonctions de l'ensemble X dans l'espace métrique E* converge uniformément *vers une fonction f si, pour tout $\varepsilon > 0$, il existe un entier m tel que, pour tout $x \in X$ et tout $n > m$, on ait $d(f_n(x), f(x)) < \varepsilon$.*

Théorème 1.4.21. *Si la suite (f_n) de fonctions continues de l'espace topologique X dans l'espace métrique E converge uniformément vers f, la fonction f est continue.*

Démonstration : Ceci résulte immédiatement du théorème 1.4.19, appliqué à tout point a de X. ■

Homéomorphismes.

Si f est une bijection continue de l'espace E sur l'espace F, il n'est pas toujours vrai que f^{-1} soit continue de F sur E. Par exemple, si $\mathbb{N}$ est le sous-espace de $\mathbb{R}$ formé des entiers naturels, et si f est définie par

$$f(n) = \frac{n}{1+n^2}$$

l'espace F étant le sous-espace $f(\mathbb{N})$ de $\mathbb{R}$, l'application f est continue : puisque pour tout $n \in \mathbb{N}$, la boule ouverte de $\mathbb{N}$ de centre n et de rayon 1 est égale à $\{n\}$, toute partie de $\mathbb{N}$ est ouverte dans $\mathbb{N}$; l'image réciproque par f de tout ouvert de F est donc ouverte dans $\mathbb{N}$. Cependant la suite $(f(n))_{n\in\mathbb{N}}$ converge vers $f(0) = 0$, alors que la suite (n) ne converge pas vers 0, ce qui montre que f^{-1} n'est pas continue.

Définition 1.4.22. *On dit que f est un* homéomorphisme *de E sur F si f est bijective et continue de E sur F et si f^{-1} est continue de F sur E.*

On dit que E et F sont homéomorphes *s'il existe un homéomorphisme de E sur F.*

Définition 1.4.23. *Un espace topologique E est dit* métrisable *s'il est homéomorphe à un espace métrique.*

Ceci revient à dire qu'il existe sur E une distance qui définit la topologie. Naturellement, une topologie métrisable peut être définie par plusieurs distances. On a déjà vu que deux distances définissent la même topologie si et seulement si elles sont équivalentes. On peut en donner un autre énoncé.

Théorème 1.4.24. *Deux distances d_1 et d_2 sur un ensemble E définissent la même topologie si et seulement si elles ont les mêmes suites convergentes.*

Démonstration : Si d_1 et d_2 définissent la même topologie $\mathscr{O}$, une suite (x_n) converge vers x pour d_1 si elle converge pour la topologie $\mathscr{O}$, et de même pour d_2.
Inversement, si toute suite qui converge pour d_1 converge pour d_2, la limite est la même : si la suite (x_n) converge vers x pour d_1, la suite (y_n) définie par $y_{2n} = x_n$ et $y_{2n+1} = x$ converge vers x pour d_1 donc aussi pour d_2 ; La limite pour d_2 de la suite (y_{2n}) est donc la même que la limite pour d_2 de la suite constante (y_{2n+1}), c'est-à-dire x. Et puisque un ensemble X est fermé pour d_1 si toute suite convergente de points de X a sa limite dans X, on voit que les ensembles fermés pour d_1 et pour d_2 coïncident. Donc d_1 et d_2 définissent la même topologie. ■

Définition 1.4.25. *Une application f d'un espace métrique E dans un espace métrique F est appelée une* isométrie *si, pour tout x et tout y de E on a*

$$d(f(x), f(y)) = d(x, y) \ .$$

Il est clair qu'une isométrie de E dans F est un homéomorphisme de E sur le sous-espace $f(E)$ de F : f et f^{-1} sont alors 1-lipschitziennes.

Continuité uniforme.

Si f est une fonction continue de l'espace métrique E dans l'espace métrique F, il existe pour tout $r > 0$ et tout x de E un δ tel que $f(B(x, \delta)) \subset B(f(x), r)$. En général, pour un r donné le δ va dépendre de x, et il se peut qu'aucun δ ne puisse être valable simultanément pour tous les x. La propriété qu'on va introduire maintenant est donc plus forte que la continuité. On verra néanmoins ultérieurement que dans certains cas, elle est équivalente à la continuité.

Définition 1.4.26. *On dit que la fonction f de l'espace métrique E dans l'espace métrique F est* uniformément continue *si, pour tout $r > 0$, on peut trouver un $\delta > 0$ tel que, pour tout x et tout y de E*

$$d(x, y) < \delta \Longrightarrow d(f(x), f(y)) < r \ .$$

Il est clair que si f est k-lipschitzienne, elle est uniformément continue : il suffit de prendre $\delta = r/k$, avec les notations précédentes.

Espaces métriques séparables.

Définition 1.4.27. *Un espace topologique est dit* séparable *s'il contient une partie dénombrable partout dense.*

Par exemple, $\mathbb{R}$, qui contient le sous-ensemble dénombrable dense $\mathbb{Q}$ des rationnels est séparable.

Théorème 1.4.28. *Si l'espace métrique E est séparable, il existe dans E une famille dénombrable $(U_n)_{n\in\mathbb{N}}$ d'ouverts telle que tout ouvert de E soit la réunion d'une sous-famille de $(U_n)_{n\in\mathbb{N}}$. Une telle famille est appelée* base de la topologie.

Démonstration : Soit D une partie dénombrable dense dans E. La famille $\mathscr{B}$ de toutes les boules ouvertes $B(x, r)$ où x appartient à D et r est rationnel est dénombrable, donc peut être énumérée en une suite (U_n). Si O est un ouvert de E et si $H = \{n \in \mathbb{N} : U_n \subset O\}$, on a clairement

$$\bigcup_{n \in H} U_n \subset O \ .$$

De plus, si $x \in O$, il existe $\rho >$ tel que $B(x, \rho) \subset O$, donc $r > 0$ rationnel tel que $2r < \rho$ et $a \in D$ tel que $d(x, a) < r$. Alors $B(a, r) \in \mathscr{B}$, $x \in B(a, r)$ et

$$B(a, r) \subset B(x, 2r) \subset B(x, \rho) \subset O$$

Donc il existe un entier n tel que $U_n = B(a, r)$ et $n \in H$, d'où l'on déduit que $x \in \bigcup_{n \in H} U_n$ et que $O = \bigcup_{n \in H} U_n$. ■

Théorème 1.4.29. *Si la topologie de l'espace métrique E possède une base dénombrable, E est séparable.*

Démonstration : Soit $(U_n)_{n \in \mathbb{N}}$ une telle base. Pour chaque entier n tel que $U_n \neq \emptyset$, on choisit un point x_n dans U_n. Alors l'ensemble $D = \{x_n : n \in \mathbb{N}\}$ est dense dans E : soient en effet $a \in E$ et $\varepsilon > 0$. Alors l'ouvert $B(a, \varepsilon)$, réunion d'une sous-famille des (U_n), contient un U_{n_0} tel que $a \in U_{n_0}$. Et le point x_{n_0} appartient à D et vérifie $d(x_{n_0}, a) < \varepsilon$ puisque $x_{n_0} \in U_{n_0} \subset B(a, \varepsilon)$. ■

Corollaire 1.4.30. *Si E est un espace métrique séparable et F un sous-espace de E, alors F est séparable.*

Démonstration : Puisque E est séparable, il possède une base dénombrable $(U_n)_{n \in \mathbb{N}}$. Il est alors clair que la famille $(U_n \cap F)_{n \in \mathbb{N}}$ est une base dénombrable de F, donc que F est séparable. ■

Théorème 1.4.31. *Soit E un espace métrique séparable. Si $\mathscr{A}$ est une famille d'ouverts de E non vides et deux-à-deux disjoints, cette famille est dénombrable.*

Démonstration : Soit (x_n) une suite de points de E partout dense dans E. A chaque $A \in \mathscr{A}$ on peut associer un entier $n(A)$ tel que $x_{n(A)} \in A$. Et puisque les éléments de $\mathscr{A}$ sont deux-à-deux disjoints, l'application $A \mapsto n(A)$ est injective de $\mathscr{A}$ dans $\mathbb{N}$. Donc $\mathscr{A}$ est dénombrable. ■

1.5 Espaces compacts

La propriété de Borel-Lebesgue.

Définitions 1.5.1. *Une famille de parties $(O_i)_{i \in I}$ d'un ensemble E est appelée* recouvrement *de E si E est la réunion de cette famille, c'est-à-dire si tout point de E appartient à l'un au moins des O_i.*

Si $\mathscr{R} = (O_i)_{i \in I}$ est un recouvrement de E, on appelle sous-recouvrement *de $\mathscr{R}$ une sous-famille $(O_i)_{i \in J}$, avec $J \subset I$, qui est un recouvrement de E.*

On appelle recouvrement ouvert *d'un espace topologique E toute famille d'ouverts de E qui est un recouvrement de E.*

Remarque. On notera qu'un recouvrement de E n'est pas une partie de E, mais une famille de parties de E.

Définition 1.5.2. *Un espace topologique E est dit* compact *s'il est séparé et si tout recouvrement ouvert de E contient un sous-recouvrement fini.*

Remarque. Quel que soit l'espace topologique E, $\{E\}$ est toujours un recouvrement ouvert fini (à un élément !) de E. La compacité de E n'est donc pas l'existence de recouvrements ouverts finis de E.

Proposition 1.5.3. *Un espace topologique discret est compact si et seulement s'il est fini.*

DÉMONSTRATION : Si E est fini et si (U_i) est un recouvrement ouvert, il suffit de choisir pour chaque point de E un U_i qui le contienne pour obtenir un sous-recouvrement fini. Inversement, il suffit de remarquer que la famille de tous les singletons constitue alors un recouvrement ouvert. ■

Théorème 1.5.4. *Un espace topologique séparé est compact si et seulement si toute famille de fermés non vides, qui est stable par intersection finie, possède une intersection non vide.*

DÉMONSTRATION : Si E est compact, et si $(F_i)_{i\in I}$ est une famille de fermés non vides stable par intersection finie et d'intersection vide, la famille des complémentaires $O_i = E\setminus F_i$ est un recouvrement ouvert de E. Il doit donc exister une partie finie J de I telle que $E = \bigcup_{i\in J} O_i$, c'est-à-dire $\bigcap_{i\in J} F_i = \emptyset$. Mais si la famille (F_i) est stable par intersection finie, il existe un $j \in I$ tel que $F_j = \bigcap_{i\in J} F_i = \emptyset$, ce qui contredit l'hypothèse.
Inversement, si E n'est pas compact, il existe un recouvrement ouvert $(O_i)_{i\in I}$ sans sous-recouvrement fini. Si on définit, pour toute partie finie J de I, $F_J = E \setminus \bigcup_{i\in J} O_i$, la famille des (F_J) est formée de fermés non vides, est stable par intersection finie, et possède une intersection vide. ■

Théorème 1.5.5. *Si E est compact et si F est un fermé de E, F est compact.*

DÉMONSTRATION : Notons d'abord que F est séparé : si x et y sont deux points distincts de F, il existe deux voisinages ouverts disjoints U et V de x et y respectivement dans E. Alors, $U \cap F$ et $V \cap F$ sont des voisinages ouverts disjoints de x et de y dans F.
Si $(A_i)_{i\in I}$ est une famille de fermés non vides de F, stable par intersection finie, chaque A_i, intersection de F avec un fermé de E est lui-même fermé dans E. Et par compacité de E la famille (A_i) doit avoir une intersection non vide. Ceci prouve la compacité de F. ■

Définition 1.5.6. *Une partie X d'un espace topologique séparé E est dite* relativement compacte *si elle est contenue dans une partie compacte de E.*

Il est équivalent de dire que l'adhérence $\overline{X}$ de X est une partie compacte de E.

Théorème 1.5.7. *Si E est séparé, et si le sous-espace X de E est compact, X est fermé dans E.*

DÉMONSTRATION : Soit a un point de $E \setminus X$. On veut prouver qu'il existe un voisinage de a disjoint de X. Pour tout x de X, $x \neq a$. Il existe donc des ouverts disjoints U_x et V_x contenant respectivement a et x. Ceci entraîne que $X \subset \bigcup_{x\in X} V_x$, donc que les $(X \cap V_x)_{x\in X}$ forment un recouvrement ouvert de X. Il doit donc exister une partie finie Y de X telle que

$\bigcup_{x\in Y}(X\cap V_x)=X$. Mais alors $U=\bigcap_{x\in Y}U_x$, intersection finie d'ouverts contenant a, est un voisinage ouvert de a. Et puisque

$$U\cap X\subset\bigcup_{x\in Y}(U\cap V_x)\subset\bigcup_{x\in Y}(U_x\cap V_x)=\emptyset\ ,$$

U est le voisinage cherché de a. ∎

Théorème 1.5.8. *Si E est un espace séparé, K_1 et K_2 deux parties compactes de E, $K_1\cup K_2$ est compact.*

DÉMONSTRATION : Soit $(U_i)_{i\in I}$ un recouvrement ouvert de $K_1\cup K_2$. Puisque les ouverts $U_i\cap K_1$ de K_1 recouvrent le compact K_1, il existe une partie finie J_1 de I telle que $K_1\subset\bigcup_{i\in J_1}U_i$. Il existe de même une partie finie J_2 de I telle que $K_2\subset\bigcup_{i\in J_2}U_i$. Alors $J=J_1\cup J_2$ est finie et $K_1\cup K_2\subset\bigcup_{i\in J}U_i$. ∎

Théorème 1.5.9. *Soient E un espace compact et f une surjection continue de E sur un espace séparé F. Alors F est lui-même compact.*

DÉMONSTRATION : Si $(O_i)_{i\in I}$ est un recouvrement ouvert de F, et si on pose $U_i=f^{-1}(O_i)$, les (U_i) sont ouverts dans E et

$$\bigcup_{i\in I}U_i=\bigcup_{i\in I}f^{-1}(O_i)=f^{-1}(\bigcup_{i\in I}O_i)=f^{-1}(F)=E\ ,$$

ce qui montre que (U_i) est un recouvrement ouvert de E. Et si J est une partie finie de I telle que $E=\bigcup_{i\in J}U_i$, on a

$$F=f(E)=f(\bigcup_{i\in J}U_i)=\bigcup_{i\in J}f(U_i)=\bigcup_{i\in J}O_i\ ,$$

ce qui montre que $(O_i)_{i\in J}$ est un sous-recouvrement fini de F. ∎

Corollaire 1.5.10. *Un espace homéomorphe à un espace compact est lui-même compact.*

DÉMONSTRATION : Si h est un homéomorphisme de l'espace compact E sur F, on remarque d'abord que F est séparé. En effet, si x et y sont deux points distincts de F, $x'=h^{-1}(x)$ et $y'=h^{-1}(y)$ qui sont distincts ont des voisinages ouverts U' et V' disjoints dans E. Et puisque h^{-1} est continue, $h(U)$ et $h(V)$ sont ouverts et disjoints dans F, contenant respectivement x et y.
Et puisque h est continue, $F=h(E)$ est compact. ∎

Corollaire 1.5.11. *Si f est bijective et continue de l'espace compact E sur l'espace séparé F, f est un homéomorphisme.*

DÉMONSTRATION : Il suffit de montrer que f^{-1} est continue. Soit donc X un fermé de E : l'image réciproque de X par f^{-1} est $f(X)$. Or X, fermé dans le compact E, est compact, et son image continue $f(X)$ dans l'espace séparé F est compacte, donc fermée dans F. Tout fermé de E a donc une image réciproque par f^{-1} qui est fermée, et f^{-1} est continue. ∎

Théorème 1.5.12. *Soient E un espace topologique, K un compact et $\pi : E \times K \to E$ la première projection. Alors si H est un fermé de $E \times K$, $\pi(H)$ est une partie fermée de E.*

Démonstration : On va montrer que si a est un point de E n'appartenant pas à $P := \pi(H)$, alors $E \setminus P$ est un voisinage de a.
Pour tout $y \in K$, on a en effet : $(a, y) \notin H$; donc $(E \times K) \setminus H$ est un voisinage de (a, y) et il existe un ouvert U_y de E et un ouvert V_y de K tels que

$$(a, y) \in U_y \times V_y \subset (E \times K) \setminus H\,.$$

La famille $(V_y)_{y \in K}$ est alors un recouvrement ouvert du compact K, et il existe une partie finie J de K telle que $K = \bigcup_{y \in J} V_y$. Et si on pose $U = \bigcap_{y \in J} U_y$, l'ensemble U est un voisinage ouvert de a, puisque J est finie et que chacun des U_y est un ouvert contenant a.

Alors U est contenu dans $E \setminus P$: en effet si x était un point de $U \cap P$, il existerait un $z \in K$ tel que $(x, z) \in H$, donc un $y \in J$ tel que $z \in V_y$, mais on aurait

$$(x, z) \in U \times V_y \subset U_y \times V_y \subset (E \times K) \setminus H\ ,$$

contrairement au fait que $(x, z) \in H$.
Et puisque $E \setminus P$ est voisinage de chacun de ses points, il est ouvert, ce qui signifie que P est fermé. ■

Le corollaire suivant est quelquefois d'un usage commode.

Corollaire 1.5.13. *Soient E un espace topologique, K un espace compact et f une application de E dans K. Alors f est continue si et seulement si son graphe est une partie fermée de $E \times K$.*

Démonstration : Supposons d'abord f continue. Puisque K est séparé, sa diagonale $\Delta := \{(z, y) \in K \times K : z = y\}$ est fermée dans $K \times K$ d'après le théorème 1.3.15. Alors le graphe $G = \{(x, y) \in E \times K : y = f(x)\}$ de f est l'image réciproque de Δ par la fonction $F : (x, y) \mapsto (f(x), y)$, dont les fonctions coordonnées $(x, y) \mapsto f(x)$ et $(x, y) \mapsto y$ sont continues. Il résulte alors du théorème 1.4.9 que F est continue, donc que $G = F^{-1}(\Delta)$ est fermé.

Inversement, supposons le graphe G de f fermé et notons $\pi : E \times K \to E$ la première projection. Si F est une partie fermée de K, l'ensemble $E \times F$ est fermé dans $E \times K$, ainsi que $H := (E \times F) \cap G$. Si $x \in f^{-1}(F)$, le point $(x, f(x))$ appartient à H, donc $x \in \pi(H)$. Inversement, si $x \in \pi(H)$, il existe $(u, v) \in H$ tel que $x = \pi(u, v)$, c'est-à-dire $x = u$, $v = f(u)$ et $v \in F$; ceci montre que $f(x) = v \in F$, donc que $x \in f^{-1}(F)$. Il en résulte que $f^{-1}(F) = \pi(H)$, et puisque H est fermé, $f^{-1}(H)$ est fermé d'après le théorème 1.5.12. L'image réciproque par f de tout fermé de K est fermée dans E. Donc f est continue. ■

Compacts métrisables.

On s'intéresse maintenant au cas des espaces métrisables. On va voir que la compacité peut s'exprimer en termes de suites. Si E est un espace métrisable, on choisira, sans toujours la préciser, une distance d sur E qui définit la topologie.

Théorème 1.5.14. *Soit E un espace métrique. Parmi les propriétés suivantes,*

i) *E est compact.*

ii) *Pour tout recouvrement ouvert $(O_i)_{i\in I}$ de E, il existe un nombre $\rho > 0$ tel que toute boule ouverte de rayon ρ centrée en un point de E soit contenue dans l'un au moins des O_i.*

iii) *Toute suite de points de E possède dans E une valeur d'adhérence.*

iv) *Toute suite de points de E possède une sous-suite qui converge dans E.*

les propriétés i), iii) et iv) sont équivalentes, et elles entraînent la propriété ii).

Démonstration : Si la suite (x_n) a une sous-suite qui converge vers x, le point x est valeur d'adhérence de la suite (x_n). Donc la propriété *iv)* entraîne la propriété *iii)*.
Supposons que toute suite de E ait une valeur d'adhérence, et que la propriété *ii*) ne soit pas vérifiée. Soit donc (O_i) un recouvrement ouvert tel que pour tout entier n, il existe un point x_n tel que la boule $B(x_n, 2^{-n})$ ne soit contenue dans aucun des O_i. Il existerait alors une valeur d'adhérence x de la suite (x_n), et x appartiendrait à un certain O_j. Soit $\varepsilon > 0$ tel que $O_j \supset B(x, \varepsilon)$. Il existe m tel que $2^{-m} < \varepsilon$, et $n > m$ tel que $d(x, x_n) < \varepsilon/2$. Alors, pour tout y appartenant à $B(x_n, 2^{-n})$, on a

$$d(x,y) \leqslant d(x,x_n) + d(x_n,y) < \varepsilon/2 + 2^{-n} \leqslant \varepsilon/2 + \frac{1}{2}2^{-m} < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

d'où $B(x_n, 2^{-n}) \subset B(x, \varepsilon) \subset O_j$, contrairement au choix de x_n.
Si la propriété *iii)* est vérifiée, la propriété *ii)* l'est aussi. Nous montrons par l'absurde que E est compact. Soit $(O_i)_{i\in I}$ un recouvrement ouvert de E ne possédant aucun sous-recouvrement fini. Il existe alors un $\rho > 0$ tel que toute boule de rayon ρ soit contenue dans l'un au moins des O_i. On construit alors par récurrence une suite (i_n) d'éléments de I et une suite (y_n) de points de E telles que $y_n \notin \bigcup_{k<n} O_{i_k}$ et $B(y_n, \rho) \subset O_{i_n}$. En effet, la réunion $\bigcup_{k<n} O_{i_k}$ n'est pas égale à E, donc ne contient pas un certain y_n, et la boule de centre y_n et de rayon ρ est incluse dans l'un des O_i, soit O_{i_n}. La suite (y_n) doit donc avoir une valeur d'adhérence y, et il existe au moins deux entiers distincts m et n tels que $d(y_m, y) < \rho/2$ et $d(y_n, y) < \rho/2$. On peut supposer $m < n$. Alors, puisque $d(y_m, y_n) < \rho$, $y_n \in B(y_m, \rho) \subset O_{i_m}$, contrairement au choix de y_n.
Enfin, si E est compact, et si (x_n) était une suite de points de E sans sous-suite convergente, tout point x de E possèderait un voisinage U_x qui ne rencontre qu'un nombre fini de points de la suite. Sinon, x serait une valeur d'accumulation de la suite (x_n), et serait limite d'une sous-suite de (x_n) : on pourrait en effet construire une suite strictement croissante d'entiers (n_k) telle que $d(x, x_{n_k}) < 2^{-k}$. Alors $(U_x)_{x\in E}$ serait un recouvrement ouvert du compact E, et si Y était une partie finie de E telle que $\bigcup_{x\in Y} U_x = E$, E ne pourrait contenir qu'un nombre fini de points de la suite infinie (x_n), ce qui est absurde. ∎

Théorème 1.5.15. *Si E est un espace métrique compact et (x_n) une suite de points de E qui possède une seule valeur d'adhérence a, la suite (x_n) converge vers a.*

Démonstration : Soit V un voisinage ouvert de a. On veut démontrer que tous les termes de la suite (x_n) sont dans V sauf un nombre fini. Dans le cas contraire, l'ensemble $H = \{n \in \mathbb{N} : x_n \notin V\}$ serait infini, et la suite $(x_n)_{n\in H}$ serait une suite dans le compact E sans valeur d'adhérence : en effet, a ne peut être une valeur d'adhérence puisqu'aucun terme de cette sous-suite n'appartient à V, et une valeur d'adhérence de la sous-suite serait une valeur d'adhérence de (x_n). ∎

Définition 1.5.16. *Si E est un espace métrique, on dit que E est* précompact *si, pour tout $\varepsilon > 0$, il existe un recouvrement de E par un nombre fini de parties de diamètre inférieur à ε.*

Théorème 1.5.17. *Tout espace métrique compact est précompact.*

DÉMONSTRATION : Soit $\varepsilon > 0$. La famille de toutes les boules ouvertes de rayon $\varepsilon/2$ est un recouvrement ouvert de E par des parties de diamètre inférieur à ε : en effet, si y et z appartiennent à $B(x, \varepsilon/2)$, $d(y, z) \leqslant d(x, y) + d(x, z) < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Et puisque E est compact, ce recouvrement admet un sous-recouvrement fini. ■

Théorème 1.5.18. *Si E est un espace métrique compact, il est séparable.*

DÉMONSTRATION : D'après le théorème précédent, il existe pour tout entier n une partie finie J_n de E telle que $E = \bigcup_{x \in J_n} B(x, 2^{-n})$. Alors $D = \bigcup_{n \in \mathbb{N}} J_n$ est dénombrable, et il existe une énumération (x_n) de D. De plus, pour tout y de E et tout $\varepsilon > 0$, il existe n tel que $2^{-n} < \varepsilon$, et $x \in J_n$ tel que $y \in B(x, 2^{-n})$, donc que $d(x, y) < \varepsilon$. Tout point y de E est donc adhérent à D, ce qui montre que E est séparable. ■

Produit de compacts métrisables.

Théorème 1.5.19. *Si E et F sont deux compacts métrisables, le produit $E \times F$ est compact.*

DÉMONSTRATION : Il suffit de démontrer que de toute suite de $E \times F$ on peut extraire une sous-suite convergente. Soit donc $(z_n)_{n \in \mathbb{N}}$ une suite dans $E \times F$. Alors $z_n = (x_n, y_n)$, avec $x_n \in E$ et $y_n \in F$. Puisque E est compact, la suite (x_n) possède une sous-suite convergente, de limite x. Il existe donc une partie infinie H de $\mathbb{N}$ telle que $x = \lim_{n \to \infty, n \in H} x_n$. Puisque F est compact, la suite $(y_n)_{n \in H}$ possède aussi une sous-suite convergente, et il existe un $y \in F$ et une partie infinie H' de H telle que $y = \lim_{n \to \infty, n \in H'} y_n$. Alors la suite $(x_n)_{n \in H'}$ extraite de la suite $(x_n)_{n \in H}$ converge vers x. Donc la suite $(z_n)_{n \in H'}$ converge vers $(x, y) \in E \times F$, ce qui achève la démonstration. ■

La même méthode permet de démontrer plus généralement :

Théorème 1.5.20. *Le produit d'une famille finie ou dénombrable d'espaces compacts métrisables est compact.*

DÉMONSTRATION : Pour une famille finie, il suffit de faire une démonstration par récurrence sur le nombre de facteurs, en utilisant le théorème précédent.
Si maintenant E_n est un espace métrique compact pour tout entier n, et si (x^k) est une suite de points du produit $E = \prod_{n \in \mathbb{N}} E_n$, on va montrer que la suite (x^k) possède une sous-suite convergente. Si on note x_n^k la $n^{\text{ème}}$ coordonnée du point x^k, la suite $(x_0^k)_{k \in \mathbb{N}}$ est une suite du compact métrique E_0. Il existe donc une partie infinie H_0 de $\mathbb{N}$ telle que la sous-suite $(x_0^k)_{k \in H_0}$ converge dans E_0 vers un point a_0.
La suite $(x_1^k)_{k \in H_0}$ est une suite dans le compact E_1. Il existe donc une partie infinie H_1 de H_0 telle que la suite $(x_1^k)_{k \in H_1}$ converge vers un point a_1 de E_1.
Répétant cette opération, on construit une suite décroissante (H_j) de parties infinies de $\mathbb{N}$ telles que la suite $(x_j^k)_{k \in H_j}$ converge dans E_j vers un point a_j. On voit alors, puisque $H_j \subset H_i$ si $j \geqslant i$, que la suite $(x_i^k)_{k \in H_j}$ converge vers a_i si $j \geqslant i$. Appliquant le lemme 1.3.24, on trouve une partie infinie H de $\mathbb{N}$ qui est presque incluse dans chacune des parties H_i. On en déduit que la suite $(x_i^k)_{k \in H}$ converge vers a_i pour tout entier i, c'est-à-dire que

la suite $(x^k)_{k\in H}$ converge dans E vers le point a de coordonnées (a_i), ce qui prouve que la suite (x^k) a une sous-suite convergente. Donc E est compact. ■

On verra dans la section 1.6 une démonstration de la généralisation suivante :

Théorème 1.5.21. (Tychonoff) *Tout produit d'espaces compacts est un espace compact.*

Parties compactes de la droite réelle.

Théorème 1.5.22. (Borel-Lebesgue) *L'intervalle* $[0,1]$ *de* $\mathbb{R}$ *est compact.*

Démonstration : Soit $(U_i)_{i\in I}$ un recouvrement ouvert de $[0,1]$. On va démontrer que ce recouvrement possède un sous-recouvrement fini. On note A l'ensemble des $x \in [0,1]$ tels que l'intervalle $[0,x]$ possède un recouvrement par un nombre fini des U_i, et on veut prouver que $1 \in A$. Puisqu'il existe un $k \in I$ tel que $0 \in U_k$ il existe $r > 0$ tel que $[0,r] \subset U_k$, A n'est pas vide et contient $[0,r]$. Puisque A est borné par 1, il possède une borne supérieure $\alpha \leqslant 1$. Il existe un $j \in I$ tel que $\alpha \in U_j$, donc un $\delta > 0$ tel que U_j contienne l'intersection de $[0,1]$ et de l'intervalle $]\alpha-\delta,\alpha+\delta[$. Par définition de α, il existe un x dans $A\cap]\alpha-\delta,\alpha]$. Puisqu'il existe une partie finie J de I tel que $[0,x]$ soit recouvert par $\bigcup_{i\in J} U_i$, $[0,\alpha]$ est recouvert par $\bigcup_{i\in J'} U_i$, où $J' = J \cup \{j\}$. Et puisque J' est fini, $\alpha \in A$. Et si α était strictement inférieur à 1, il existerait un y dans $[0,1]\cap]\alpha,\alpha+\delta]$. Alors l'intervalle $[0,y]$ serait aussi contenu dans $\bigcup_{i\in J'} U_i$, ce qui prouverait que $y \in A$, contrairement à la définition de α. ■

Théorème 1.5.23. *Les parties compactes de* $\mathbb{R}$ *sont les parties fermées et bornées.*

Démonstration : Soit K une partie compacte de $\mathbb{R}$. Puique l'espace métrique $\mathbb{R}$ est séparé, K est fermé. Et puisque la famille décroissante de fermés $F_n = K \setminus]-n,n[$, de K est d'intersection vide, l'un d'entre eux doit être vide, ce qui signifie que K est contenu dans un intervalle $]-n,n[$, donc est borné.
Inversement, pour tout entier n, l'application $x \mapsto n(2x-1)$ est continue de $[0,1]$ sur $[-n,n]$. Il en résulte que $[-n,n]$ est compact. Enfin, si K est une partie fermée et bornée de $\mathbb{R}$, il existe un entier n tel que $K \subset [-n,n]$. Alors $K = K \cap [-n,n]$ est fermé dans le compact $[-n,n]$, donc lui-même compact. ■

Corollaire 1.5.24. *Si* K *est un compact non vide de* $\mathbb{R}$*,* K *contient un plus grand point et un plus petit point.*

Démonstration : Puisque K est borné et non vide, il possède une borne supérieure α et une borne inférieure β. Puisque, pour tout $\varepsilon > 0$, K rencontre $]\alpha-\varepsilon,\alpha]$, α est adhérent à l'ensemble fermé K, donc appartient à K, dont il est le plus grand point. On voit de même que β est le plus petit point de K. ■

Théorème 1.5.25. *Les parties compactes de* $\mathbb{R}^n$ *sont les parties fermées et bornées.*

Démonstration : Si K est une partie compacte de l'espace métrisable $\mathbb{R}^n$, elle est fermée dans $\mathbb{R}^n$. Et si on note P_k le pavé ouvert $\{x = (x_1, x_2, \ldots, x_n) : \max|x_i| < k\}$ (qui est la boule ouverte centrée à l'origine et de rayon k), la famille décroissante de fermés $(K \setminus P_k)$ du compact K a une intersection vide : l'un d'entre eux est donc vide, c'est-à-dire que K est contenu dans l'un des P_k, donc est borné.
Inversement, si K est fermé et borné dans $\mathbb{R}^n$, il existe un entier k tel que, pour tout x de K, chacune des n coordonnées de x soit majorée en valeur absolue par k. Puisque $[-k,k]$ est compact dans $\mathbb{R}$, $P = [-k,k]^n$ est compact dans $\mathbb{R}^n$, et $K = K \cap P$ est fermé dans le compact P, donc compact. ■

Fonctions continues sur un compact.

Théorème 1.5.26. *Si f est une fonction continue du compact K non vide dans $\mathbb{R}$, f est bornée et atteint sur K sa borne supérieure et sa borne inférieure.*

Démonstration : Puisque f est continue, $f(K)$ est une partie compacte non vide de $\mathbb{R}$; elle est donc bornée et contient un plus grand point α et un plus petit point β. Si x et y sont des points de K tels que $f(x) = \alpha$ et $f(y) = \beta$, f atteint sa borne supérieure α en x et sa borne inférieure β en y. ■

Théorème 1.5.27. *Soient K un espace compact et F un espace métrique. Sur l'ensemble $\mathscr{C}(K, F)$ des fonctions continues de K dans F, on peut définir une distance, appelée* distance de la convergence uniforme, *par*

$$d(f, g) = \sup_{x \in K} d(f(x), g(x)) \ .$$

Cette borne supérieure est atteinte en au moins un point de K.

Démonstration : Si f et g sont continues de K dans F, la fonction $\varphi = f \times g : K \to F \times F$ dont les fonctions coordonnées sont f et g est continue de K dans $F \times F$. Il en résulte, puisque la distance est continue de $F \times F$ dans $\mathbb{R}$, que la fonction $x \mapsto d(f(x), g(x))$ est continue de K dans $\mathbb{R}$, donc est bornée et atteint sa borne supérieure.
On a clairement $d(f, f) = 0$, et $d(f, g) = d(g, f)$. Enfin, si f, g et h sont trois fonctions continues de K dans F, on a pour tout x de K

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

donc, en passant à la borne supérieure, on a pour tout x

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

ce qui fournit l'inégalité cherchée $d(f, h) \leqslant d(f, g) + d(g, h)$. ■

Théorème 1.5.28. (Heine) *Si f est une application continue de l'espace métrique compact E dans l'espace métrique F, f est uniformément continue.*

Démonstration : Soit $\varepsilon > 0$. Pour tout x de E, il existe un voisinage ouvert U_x du point x tel que $f(U_x) \subset B(f(x), \varepsilon/2)$. La famille $(U_x)_{x \in E}$ est un recouvrement ouvert de E. D'après le théorème 1.5.14, il existe un nombre $\rho > 0$ tel que toute boule ouverte de E de rayon ρ soit contenue dans l'un des U_x.
Si y et z sont deux points de E tels que $d(y, z) < \rho$, $z \in B(y, \rho)$. Il existe donc un $x \in E$ tel que $B(y, \rho) \subset U_x$. On en déduit que y et z appartiennent à U_x, donc que $d(f(y), f(x)) < \varepsilon/2$ et $d(f(z), f(x)) < \varepsilon/2$, et enfin que

$$d(f(y), f(z)) \leqslant d(f(y), f(x)) + d(f(z), f(x)) < \varepsilon/2 + \varepsilon/2 = \varepsilon \ ,$$

ce qui achève de montrer que f est uniformément continue. ■
On peut de la même manière démontrer une généralisation de ce théorème, qui aura plus loin des applications.

Théorème 1.5.29. *Soient E et F deux espaces métriques, f une fonction continue de E dans F et K une partie compacte de E. Alors, pour tout $\varepsilon > 0$, il existe un $\eta > 0$ tel que, si $x \in K$, $y \in E$ et $d(x,y) < \eta$ on ait $d(f(x), f(y)) < \varepsilon$.*

Démonstration : Soit $\varepsilon > 0$. Pour tout z de K, il existe un $\rho_z > 0$ tel que

$$f(B(z, 2\rho_z)) \subset B(f(z), \varepsilon/2) \ .$$

Alors la famille $(B(z,\rho_z) \cap K)$ est un recouvrement ouvert du compact K, et il existe un sous-recouvrement fini $(B(z,\rho_z) \cap K)_{z \in J}$ de K extrait de ce recouvrement. Si on désigne par η le nombre strictement positif

$$\eta = \inf_{z \in J} \rho_x \ ,$$

on a, pour $x \in K$, $y \in E$ et $d(x,y) < \eta$, l'existence d'un $z \in J$ tel que $x \in B(z,\rho_z) \cap K$. Alors, puisque $\eta \leqslant \rho_z$, on a

$$d(y,z) \leqslant d(y,x) + d(x,z) < \eta + \rho_z \leqslant 2\rho_z \ .$$

On en déduit que x et y appartiennent à $B(z, 2\rho_z)$ donc que

$$d(f(x), f(y)) \leqslant d(f(x), f(z)) + d(f(y), f(z)) < \varepsilon/2 + \varepsilon/2 = \varepsilon \ ,$$

ce qui achève la démonstration. ■

Théorème 1.5.30. (Dini) *Si (f_n) est une suite croissante de fonctions continues de l'espace compact K dans $\mathbb{R}$, qui converge en tout point de K vers une fonction continue f, la convergence est uniforme.*

Démonstration : Soit $\varepsilon > 0$. L'ensemble

$$U_n = \{x \in K : f_n(x) > f(x) - \varepsilon\} = \{x \in K : (f - f_n)(x) < \varepsilon\}$$

est ouvert puisque $f - f_n$ est continue. De plus $U_n \subset U_{n+1}$ puisque la suite (f_n) est croissante. Et K est réunion des ouverts (U_n) puisque $f(x) = \lim_{n\to\infty} f_n(x)$ pour tout x. Par compacité de K, on voit qu'il existe un m tel que U_m soit égal à K ce qui signifie que, pour tout $x \in K$ et tout $n \geqslant m$

$$f(x) - \varepsilon < f_m(x) \leqslant f_n(x) \leqslant f(x) \ ,$$

donc que la suite (f_n) converge uniformément. ■

Définition 1.5.31. *On dit qu'un ensemble V de fonctions sur l'espace K* sépare les points *de K si, pour tout couple (x,y) de points distincts dans K, il existe $f \in V$ telle que $f(x) \neq f(y)$.*

Théorème 1.5.32. *Soit K un espace compact. Alors K est métrisable si et seulement s'il existe une suite de fonctions réelles continues sur K qui sépare les points de K.*

Démonstration : Si K est compact métrisable, il est séparable. Il existe donc une suite (x_n) partout dense. Et si d est une distance sur K qui définit la topologie, la suite (f_n) de fonctions définie par $f_n(x) = d(x, x_n)$ sépare les points : en effet, si $x \neq y$, il existe n tel que $d(x,x_n) < \dfrac{1}{2} d(x,y)$ et on a alors $f_n(x) < \dfrac{1}{2} d(x,y) < f_n(y)$.

Inversement, si la suite $(f_n)_{n\in\mathbb{N}}$ de fonctions continues sépare les points de K, la fonction $F : K \to \mathbb{R}^{\mathbb{N}}$ dont les fonctions coordonnées sont les (f_n) est continue et injective. Alors $H = F(K)$ est une partie compacte de l'espace métrisable $\mathbb{R}^{\mathbb{N}}$, et il résulte du corollaire1.5.11 que F est un homéomorphisme de K sur H, donc que K est métrisable. ■

Espaces localement compacts.

Définition 1.5.33. *Un espace topologique séparé E est dit* localement compact *si tout point de E possède un voisinage compact.*

Proposition 1.5.34. *Si K est une partie compacte de l'espace localement compact E, il existe un voisinage compact de K, c'est-à-dire un compact qui soit voisinage de chaque point de K.*

Démonstration : Pour tout x de K il existe un voisinage compact K_x de x. La famille des intérieurs $(\overset{\circ}{K_x})$ forme donc un recouvrement ouvert du compact K. Alors, il existe une partie finie J de K telle que

$$K \subset U = \bigcup_{x \in J} \overset{\circ}{K_x}$$

et on a $\overline{U} \subset \bigcup_{x \in J} K_x$, ce qui prouve que $\overline{U}$, qui est fermé dans une réunion finie de compacts, est lui-même compact, et voisinage de tout point de K puisqu'il contient U. ■

Définition 1.5.35. *Un espace topologique localement compact E est dit* dénombrable à l'infini *s'il existe une suite de parties compactes de E qui recouvre E.*

Théorème 1.5.36. *Si E est un espace localement compact dénombrable à l'infini, il existe une* suite exhaustive de compacts *de E, c'est-à-dire une suite croissante (K_n) de compacts de E telle que tout compact de E soit contenu dans l'un des K_n.*

Démonstration : Puisque E est dénombrable à l'infini, il existe une suite (L_n) de compacts qui recouvre E. On construit alors par récurrence une suite (K_n) de compacts en posant $K_0 = L_0$ et en choisissant, pour $n \geqslant 1$, un voisinage compact K_n de $K_{n-1} \cup L_n$. Alors les K_n recouvrent E puisqu'ils contiennent les L_n, et si K est un compact de E, K est recouvert par les ouverts $U_n = \overset{\circ}{K_n}$ puisque $K_n \subset U_{n+1}$. Il est donc recouvert par un nombre fini d'entre eux. Et puisque ceux-ci sont croissants, il existe un m tel que $K \subset U_m \subset K_m$. ■

Théorème 1.5.37. *Si U est un ouvert de l'espace métrique compact E, U est localement compact dénombrable à l'infini.*

Démonstration : Si on note F le fermé $E \setminus U$, et φ la fonction continue : $x \mapsto d(x, F)$, U est l'ensemble des points de E où φ est strictement positive. Alors si $x \in U$ et $\alpha = \varphi(x) > 0$, l'ensemble $V = \{y \in E : \varphi(y) \geqslant \alpha/2\}$ est un voisinage fermé de x dans E, donc est compact et il est contenu dans U. De plus, chaque ensemble

$$K_n = \{y \in E : \varphi(y) \geqslant 2^{-n}\}$$

est compact, contenu dans U, et $U = \bigcup_{n \in \mathbb{N}} K_n$. ■

Théorème 1.5.38. *Si L est un espace localement compact, il existe une unique topologie d'espace compact sur $L^\bullet = L \cup \{\infty\}$, pour laquelle L est un sous-espace de $L^\bullet$. L'espace $L^\bullet$ est appelé le* compactifié d'Alexandroff *de L. Une fonction f de $L^\bullet$ dans l'espace E est continue si et seulement si sa restriction à L est continue et si $f(x)$ tend vers $f(\infty)$ lorsque x tend vers l'infini, c'est-à-dire si, pour tout voisinage W de $f(\infty)$ dans E, l'ensemble $f^{-1}(E \setminus W)$ est relativement compact dans L.*
Si L est métrisable et dénombrable à l'infini, alors $L^\bullet$ est un compact métrisable.

Démonstration : On considère sur $L^\bullet = L \cup \{\infty\}$ la famille de parties $\mathscr{O}$ formée des ouverts de L et des complémentaires dans $L^\bullet$ des compacts de L. On vérifie sans peine la

stabilité par intersections finies et par unions quelconques de $\mathscr{O}$. Et puisque, clairement, $\emptyset$ et $L^\bullet$ appartiennent à $\mathscr{O}$, on a ainsi défini une topologie sur $L^\bullet$, pour laquelle L est un sous-espace topolopgique.
Puisque L est séparé, on voit que deux points de L ont, dans $L^\bullet$, des voisinages ouverts disjoints (et contenus dans L). Et si $a \in L$, il existe un voisinage compact K de a dans L : alors $L^\bullet \setminus K$ est un voisinage de ∞ disjoint du voisinage K de a. Il en résulte que $L^\bullet$ est séparé.
Enfin, si $(U_i)_{i\in I}$ est un recouvrement ouvert de $L^\bullet$, il existe un $i_0 \in I$ tel que $\infty \in U_{i_0}$. Alors $K = L^\bullet \setminus U_{i_0}$ est fermé dans L et relativement compact, donc compact dans L. Et puisque $(U_i \cap K)_{i\in I}$ est un recouvrement ouvert de K, on trouve une partie finie J' de I telle que $K \subset \bigcup_{i\in J'} U_i$. Et on a clairement $L^\bullet = \bigcup_{i\in J} U_i$, si on prend $J = J' \cup \{i_0\}$. Ceci achève de montrer la compacité de $L^\bullet$.

Si $\mathscr{O}'$ est une topologie compacte sur $L^\bullet$ pour laquelle L est un sous-espace, $\{\infty\}$ est fermé dans $L^\bullet$ et L est ouvert dans $L^\bullet$. Tout ouvert de L est donc dans $\mathscr{O}'$; et tout élément de $\mathscr{O}'$ ne contenant pas ∞ est ouvert dans L. Si U est un ouvert de $L^\bullet$ qui contient ∞, il a dans $L^\bullet$ un complémentaire fermé, donc compact, et contenu dans L : $L \setminus U$ est un compact de L. Et inversement, si K est un compact de L, il est compact pour $\mathscr{O}'$, donc fermé, et $L^\bullet \setminus K$ est dans $\mathscr{O}'$. Il en résulte que $\mathscr{O}' = \mathscr{O}$, d'où l'unicité de la topologie de l'espace $L^\bullet$.

De plus, si $f : L^\bullet \to E$ est continue, sa restriction à L est continue, et f est continue en ∞ : pour tout voisinage W de $f(\infty)$ dans E , $f^{-1}(W)$ est un voisinage de ∞ dans $L^\bullet$, c'est-à-dire que $f^{-1}(E \setminus W)$ est relativement compact dans L. Et si f est continue sur L, avec une limite $a \in E$ à l'infini, le prolongement $f^\bullet$ de f défini sur $L^\bullet$ par $f(\infty) = a$ est continu en tout point x de L, et continu en ∞, ce qui montre la continuité de $f^\bullet$.

Enfin, si L est métrisable et dénombrable à l'infini, il existe une suite exhaustive (K_p) de compacts, ainsi qu'une distance d, bornée par 1, qui définit la topologie de L. Si U_p est l'intérieur de K_p, la fonction $g : x \mapsto \sum_p 2^{-p} d(x, U_p^c)$ est continue et strictement positive sur L, et tend vers 0 à l'infini. Puisque chacun des K_p est séparable, il en est de même de $L = \bigcup_p K_p$, et si (x_n) est une suite dense dans L, les fonctions $f_n : x \mapsto \mathrm{e}^{-d(x,x_n)}.g(x)$ sont continues et strictement positives sur L, et tendent vers 0 à l'infini, donc se prolongent en fonctions continues $f_n^\bullet$ sur $L^\bullet$ nulles en ∞. Pour montrer que le compact $L^\bullet$ est métrisable, il suffit, en vertu du théorème 1.5.32, de montrer que la famille $(f_n^\bullet)$ sépare les points de $L^\bullet$. Puisque $f_n^\bullet(\infty) = 0 \neq f_n^\bullet(x)$ pour $x \in L$, il suffit même de montrer que la famille (f_n) sépare les points de L.
Soient donc x et y distincts. Quitte à les intervertir, on peut supposer $g(y) \leqslant g(x)$. Et on peut trouver n tel que $d(x, x_n) < \dfrac{1}{2} d(x, y)$. On a alors $d(x, x_n) < \dfrac{1}{2} d(x, y) < d(y, x_n)$, donc $f_n(y) = \mathrm{e}^{-d(y,x_n)}.g(y) \leqslant \mathrm{e}^{-d(y,x_n)}.g(x) < \mathrm{e}^{-d(x,x_n)}.g(x) = f_n(x)$. ■

Applications propres.

Définition 1.5.39. *Soient X et Y deux espaces métrisables. Une application continue $f : X \to Y$ est dite* propre *si l'image réciproque par f de tout compact de Y est un compact de X.*

Lemme 1.5.40. *Si $f : X \to Y$ est propre et si F est fermé dans X, alors $f(F)$ est fermé dans Y.*

Démonstration : Soit b un point adhérent à $f(F)$; on veut montrer que $b \in f(F)$. Il existe une suite (y_n) dans $f(F)$ qui converge vers b, donc une suite (x_n) dans F telle que

$f(x_n) = y_n$. Alors l'ensemble $K = \{b\} \cup \{y_n : n \in \mathbb{N}\}$ est compact dans Y et $L = f^{-1}(K)$ est compact dans X. La suite (x_n) dans le compact $L \cap F$ possède donc une sous-suite (x_{n_k}) qui converge vers un point $a \in L \cap F$. Alors $y_{n_k} = f(x_{n_k}) \to f(a)$, et on en déduit que $b = f(a) \in f(F)$. ■

En particulier, une bijection continue est propre si et seulement si c'est un homéomorphisme.

Théorème 1.5.41. *Soient X et Y deux espaces métrisables et $f : X \to Y$ une application continue. Alors f est propre si et seulement si f est fermée (c'est-à-dire que l'image directe de tout fermé de X est un fermé de Y) et si l'image réciproque de chaque point de Y est compacte dans X.*

Démonstration : Il résulte du lemme ci-dessus que, si f est propre, elle est fermée. Et puisque tout singleton $\{y\}$ de Y est compact, il est clair que, si f est propre, $f^{-1}(y)$ est compact pour tout $y \in Y$.

Inversement, si f est fermée, si $f^{-1}(y)$ est compact pour tout y de Y et si K est compact dans Y, on doit montrer que $L = f^{-1}(K)$ est compact. Supposons, par l'absurde, que (x_n) soit une suite dans L sans valeur d'adhérence. Alors la suite $(y_n) = (f(x_n))$ possède dans K une sous-suite (y_{n_k}) qui converge vers un point b. Puisque la suite (x_{n_k}) n'a pas de valeur d'adhérence, elle ne peut avoir qu'un nombre fini de termes dans le compact $f^{-1}(b)$, et il existe k_0 tel que $f(x_{n_k}) \neq b$ si $k \geqslant k_0$. Alors l'ensemble $F = \{x_{n_k} : k \geqslant k_0\}$ est fermé dans X puisque la suite (x_{n_k}) n'a pas de valeur d'adhérence, mais l'ensemble $f(F)$ contient les $(y_{n_k})_{k \geqslant k_0}$ sans contenir leur limite b, ce qui montre que $f(F)$ ne peut être fermé et fournit la contradiction cherchée. ■

1.6 Filtres et ultrafiltres. Théorème de Tychonoff

On a vu précédemment que, dans le cas des espaces métriques, il existe de bonnes caractérisations de la compacité et de la continuité en termes de suites convergentes. Pour des espaces topologiques plus généraux, on est amené à introduire la notion de filtre.

Définition 1.6.1. *Un* filtre *sur un ensemble E est un ensemble $\mathscr{F}$ non vide de parties* non vides *de E satisfaisant les propriétés :*

i) $A \in \mathscr{F}$ et $A \subset B \subset E \Longrightarrow B \in \mathscr{F}$.

ii) $A \in \mathscr{F}$ et $B \in \mathscr{F} \Longrightarrow A \cap B \in \mathscr{F}$.

Un filtre $\mathscr{F}$ est dit libre *si $\bigcap\{A : A \in \mathscr{F}\} = \emptyset$, c'est-à-dire si, pour tout $a \in E$, on peut trouver un $A \in \mathscr{F}$ tel que $a \notin A$.*

Il est clair que si $\mathscr{F}$ est un filtre sur E, alors $E \in \mathscr{F}$. Il faut bien noter aussi que $\mathscr{F} \in \mathscr{P}(\mathscr{P}(E))$.

Exemples.

1) L'ensemble $\mathscr{N}$ des parties A de $\mathbb{N}$ telles que $\mathbb{N} \setminus A$ soit fini est un filtre sur $\mathbb{N}$ appelé *filtre de Fréchet*.

2) Si X est un espace topologique, a un point de X et E une partie de X à laquelle a est adhérent, l'ensemble $\mathscr{F} = \{V \cap E : V \in \mathscr{V}(a)\}$ des voisinages de a dans E est un filtre sur E.

Image d'un filtre.

Si E et F sont deux ensembles, $f : E \to F$ une application et $\mathscr{E}$ un filtre sur E, l'ensemble $\mathscr{F} = \{A \subset F : f^{-1}(A) \in \mathscr{E}\}$ est un filtre sur F, noté $f(\mathscr{E})$ et appelé *image de* $\mathscr{E}$.

Comparaison des filtres.

Soient $\mathscr{E}$ et $\mathscr{F}$ deux filtres sur le même ensemble E. On dit que $\mathscr{F}$ est *plus fin que* $\mathscr{E}$ (ou que $\mathscr{F}$ *raffine* $\mathscr{E}$) si tout élément de $\mathscr{E}$ appartient à $\mathscr{F}$, c'est-à-dire si $\mathscr{E} \subset \mathscr{F}$.

Convergence des filtres.

On suppose maintenant que E est un espace topologique, que $a \in E$ et que $\mathscr{F}$ est un filtre sur E.
On dit que $\mathscr{F}$ *converge vers* a (ou que a est limite de $\mathscr{F}$) si $\mathscr{F}$ est plus fin que le filtre des voisinages de a, c'est-à-dire si tout ouvert de E contenant a appartient à $\mathscr{F}$. Il est clair que si E est séparé, un filtre a au plus une limite.

On dit que a est une *valeur d'adhérence de* $\mathscr{F}$ si a est adhérent à chaque élément de $\mathscr{F}$, c'est-à-dire si $V \cap A \neq \emptyset$ chaque fois que V est un ouvert contenant a et que $A \in \mathscr{F}$. Il est clair que si a est limite de $\mathscr{F}$, il est valeur d'adhérence de $\mathscr{F}$.

Lemme 1.6.2. *Soient E un espace topologique, a un point de E, $\mathscr{E}$ et $\mathscr{F}$ deux filtres sur E tels que $\mathscr{F}$ soit plus fin que $\mathscr{E}$. Alors, si $\mathscr{E}$ converge vers a, il en est de même de $\mathscr{F}$. Et si a est valeur d'adhérence de $\mathscr{F}$, a est aussi valeur d'adhérence de $\mathscr{E}$.*

Démonstration : Supposons d'abord que $\mathscr{E}$ converge vers a. Si W est un voisinage de a, on a $W \in \mathscr{E} \subset F$, donc $W \in \mathscr{F}$.
Inversement, si a est valeur d'adhérence de $\mathscr{F}$, W voisinage de a et $A \in \mathscr{E}$, on a alors $A \in \mathscr{F}$, donc $A \cap W \neq \emptyset$, ce qui montre que a est valeur d'adhérence de $\mathscr{E}$. ■

Définition 1.6.3. *Soient E un ensemble, X un espace topologique, f une application de E dans X et $\mathscr{F}$ un filtre sur E. On dit que f* converge vers $a \in X$ suivant $\mathscr{F}$ *si le filtre $f(\mathscr{F})$ converge vers a, c'est-à-dire si, pour tout voisinage W de a dans X, $f^{-1}(W) \in \mathscr{F}$.*

La proposition suivante fait le lien entre convergence des suites et convergence des filtres.

Proposition 1.6.4. *Soient X un espace topologique, a un point de X et (x_n) une suite dans X, c'est-à-dire une application $x : \mathbb{N} \to X$. Alors, la suite (x_n) converge vers a si et seulement si l'application x converge vers a selon le filtre de Fréchet, et a est valeur d'adhérence de la suite (x_n) si et seulement si a est valeur d'adhérence du filtre $x(\mathscr{N})$.*

Dire que (x_n) converge vers a signifie que, pour tout voisinage W de a dans X, l'ensemble des n tels que $x_n \notin W$ est fini, c'est-à-dire que $x^{-1}(W) \in \mathscr{N}$ et que $W \in x(\mathscr{N})$.
On voit de même que a est valeur d'adhérence de (x_n) si a est valeur d'adhérence du filtre $x(\mathscr{N})$. ■

La notion de sous-suite est alors à remplacer par celle de filtre plus fin : si $x : \mathbb{N} \to X$ est une suite dans X, une sous-suite de x est une suite $y = x \circ \varphi$, où φ est une injection croissante de $\mathbb{N}$ dans $\mathbb{N}$. Alors le filtre $\mathscr{N}_\varphi = \varphi(\mathscr{N})$ est un filtre sur $\mathbb{N}$ plus fin que $\mathscr{N}$ et dire que la sous-suite y converge vers a dans X revient à dire que la fonction x converge vers a selon le filtre $\mathscr{N}_\varphi$.
Rappelons qu'un espace métrique est compact si toute suite possède une sous-suite convergente. On va voir qu'un espace séparé est compact si tout filtre peut être raffiné en un filtre convergent.

Ultrafiltres.

Définition 1.6.5. *Un filtre $\mathscr{F}$ sur un ensemble E est appelé* ultrafiltre *si, pour toute partie A de E, on a soit $A \in \mathscr{F}$, soit $E \setminus A \in \mathscr{F}$.*

Bien entendu, si $A \subset E$, on ne peut avoir simultanément $A \in \mathscr{F}$ et $E \setminus A \in \mathscr{F}$, puisque $A \cap (E \setminus A) = \emptyset \notin \mathscr{F}$.

Théorème 1.6.6. *Si $\mathscr{E}$ est un ultrafiltre sur E et $f : E \to F$ une application, alors $\mathscr{F} = f(\mathscr{E})$ est un ultrafiltre sur F.*

Démonstration : Si B est une partie de F et si on pose $A = f^{-1}(B)$, on a — soit $A \in \mathscr{E}$, donc $B \in \mathscr{F}$, — soit $E \setminus A = f^{-1}(F \setminus B) \in \mathscr{E}$, donc $F \setminus B \in \mathscr{F}$. ■

L'utilité des ultrafiltres vient de l'énoncé suivant.

Théorème 1.6.7. *Si E est un espace topologique, $\mathscr{F}$ un ultrafiltre sur E et a une valeur d'adhérence de $\mathscr{F}$, alors $\mathscr{F}$ converge vers a.*

Démonstration : Supposons en effet que $\mathscr{F}$ ne converge pas vers a. Il existerait alors un voisinage W de a tel que $W \notin \mathscr{F}$. Et puisque $\mathscr{F}$ est un ultrafiltre, on devrait alors avoir $E \setminus W \in \mathscr{F}$. Et comme a est valeur d'adhérence de $\mathscr{F}$ et W voisinage de a, on aurait alors $\emptyset = W \cap (E \setminus W) \in \mathscr{F}$, ce qui est absurde. ■

Théorème 1.6.8. (AC) *Si $\mathscr{E}$ est un filtre sur l'ensemble E, il existe sur E un ultrafiltre $\mathscr{F}$ plus fin que $\mathscr{E}$.*

Démonstration : Notons $\mathscr{O}$ l'ensemble des filtres sur E, ordonné par la relation d'inclusion $\subset$ (c'est-à-dire de finesse). On va montrer que cet ensemble ordonné est inductif, déduire du théorème de Zorn (d'où l'hypothèse sur l'axiome du choix) que $\mathscr{E}$ est moins fin qu'un élément maximal et finalement montrer qu'un élément maximal est un ultrafiltre.

Soit donc $(\mathscr{E}_i)_{i \in I}$ une famille totalement ordonnée de filtres sur E, et posons $\mathscr{E} = \bigcup_{i \in I} \mathscr{E}_i$. Puisque $\emptyset$ n'appartient à aucun des $\mathscr{E}_i$, il ne peut appartenir à $\mathscr{E}$.
Si $A \in \mathscr{E}$ et $A \subset B$, il existe $i \in I$ tel que $A \in \mathscr{E}_i$; alors $B \in \mathscr{E}_i \subset \mathscr{E}$. Enfin, si A et B appartiennent à $\mathscr{E}$, il existe i et j dans I tels que $A \in \mathscr{E}_i$ et $B \in \mathscr{E}_j$; notant alors $k = \max(i, j)$, on a $A \in \mathscr{E}_k$ et $B \in \mathscr{E}_k$, donc $A \cap B \in \mathscr{E}_k \subset \mathscr{E}$. Et ceci montre que $\mathscr{E}$ est un filtre, clairement plus fin que chacun des $\mathscr{E}_i$.

Reste à voir que si $\mathscr{F}$ est un filtre maximal, c'est un ultrafiltre. Soit donc $A_1 \subset E$ et $A_2 = E \setminus A_1$. S'il existait $B_1 \in \mathscr{F}$ tel que $A_1 \cap B_1 = \emptyset$ et $B_2 \in \mathscr{F}$ tel que $A_2 \cap B_2 = \emptyset$, on aurait

$$B_1 \cap B_2 = (A_1 \cup (A_2) \cap (B_1 \cap B_2) \subset (A_1 \cap B_1) \cup (A_2 \cap B_2) = \emptyset \ ,$$

en contradiction avec le fait que $B_1 \cap B_2 \in \mathscr{F}$. Il existe donc $j \in \{1, 2\}$ tel que $A_j \cap B \neq \emptyset$ quel que soit $B \in \mathscr{F}$. Si on pose alors $\mathscr{F}' = \{C \subset E : \exists B \in \mathscr{F} \quad C \supset A_j \cap B\}$, on vérifie sans peine que $\mathscr{F}'$ est un filtre plus fin que $\mathscr{F}$ et que $A_j = A_j \cap E \in \mathscr{F}'$.
Si $\mathscr{F}$ est un ultrafiltre, on a nécessairement $\mathscr{F}' = \mathscr{F}$, donc $A_j \in \mathscr{F}$. Et on a bien montré que $A_1 \in \mathscr{F}$ ou $E \setminus A_1 = A_2 \in \mathscr{F}$. ■

Théorème 1.6.9. *Soit X un espace topologique séparé. Les propriétés suivantes sont équivalentes :*

i) X est compact.

ii) Tout filtre sur X possède une valeur d'adhérence.

iii) Tout ultrafiltre sur X est convergent.

DÉMONSTRATION : Il résulte du théorème 1.6.7 que $ii) \Longrightarrow iii)$.
Si tout ultrafiltre sur X converge et si $\mathscr{E}$ est un filtre sur X, il existe par le théorème 1.6.8, un ultrafiltre $\mathscr{F}$ plus fin que $\mathscr{E}$, qui converge vers a. Alors a est valeur d'adhérence de $\mathscr{F}$, et a fortiori de $\mathscr{E}$ d'après le lemme 1.6.2. Et ceci montre que $iii) \Longrightarrow ii)$.
Si X est compact et si $\mathscr{E}$ est un filtre sur X, la famille de fermés $(\overline{A})_{A\in\mathscr{E}}$ possède la propriété d'intersection finie : si $(A_1, A_2, \ldots, A_m)$ est une famille finie d'éléments de $\mathscr{E}$, on a

$$\bigcap_{1\leqslant j\leqslant m} \overline{A_j} \supset \overline{\bigcap_{1\leqslant j\leqslant m} A_j} \neq \emptyset \ ,$$

puisque $\bigcap_{1\leqslant j\leqslant m} A_j \in \mathscr{E}$. Par compacité de X on doit donc avoir $K := \bigcap_{A\in\mathscr{E}} \overline{A} \neq \emptyset$. Alors tout point $a \in K$ est valeur d'adhérence de $\mathscr{E}$, ce qui montre que $i) \Longrightarrow ii)$.
Enfin, si la condition $i)$ n'est pas réalisée et si $(O_j)_{j\in J}$ est un recouvrement ouvert de X ne possédant aucun sous-recouvrement fini, l'ensemble

$$\mathscr{E} = \{A \subset X : \exists m \in \mathbb{N}\ \exists (j_1, j_2, \ldots, j_m) \in J^m \quad A \supset X \setminus \bigcup_{1\leqslant k\leqslant m} O_{j_k}\}$$

est un filtre sur X, comme on le vérifie sans peine. De plus, tout point a de X est contenu dans un ouvert O_j, ce qui montre que $X \setminus O_j$ est un élément de $\mathscr{E}$ disjoint du voisinage O_j de a ; et a n'est pas valeur d'adhérence de $\mathscr{E}$, ce qui prouve que $ii)$ n'est pas réalisée. Ceci achève la démonstration. ∎

Théorème 1.6.10. (Tychonoff) *Tout produit d'espaces compacts est compact.*

DÉMONSTRATION : Soient $(E_j)_{j\in J}$ une famille d'espaces topologiques compacts, et E l'espace produit de cette famille. On peut clairement supposer que chacun des E_j est non vide, puisque sinon $E = \emptyset$ est compact. Pour prouver la compacité de E, on va utiliser le théorème 1.6.9 - $iii) \Longrightarrow i)$: soit donc $\mathscr{F}$ un ultrafiltre sur E.
Pour chacune des applications de projection $\pi_j : E \to E_j$, le filtre image $\mathscr{F}_j = \pi_j(\mathscr{F})$ est, en vertu du théorème 1.6.6, un ultrafiltre sur E_j qui converge sur le compact E_j vers un point a_j d'après le théorème 1.6.9. Soit $a = (a_j)_{j\in J} \in E$; on va montrer que $\mathscr{F}$ converge vers a. En effet, si W est un ouvert de E contenant a, il existe, par définition de la topologie de E, une partie finie J_0 de J et des ouverts $O_j \subset E_j$ tels que $\prod_{j\in J} O_j \subset W$ et $O_j = E_j$ si $j \notin J_0$. Alors $O_j \in \mathscr{F}_j$, c'est-à-dire $\pi_j^{-1}(O_j) \in \mathscr{F}$, pour tout j (avec $\pi_j^{-1}(O_j) = E$ si $j \notin J_0$), et on a

$$W \supset \prod_{j\in J} O_j = \bigcap_{j\in J} \pi_j^{-1}(O_j) = \bigcap_{j\in J_0} \pi_j^{-1}(O_j) \in \mathscr{F} \ ,$$

ce qui montre que $W \in \mathscr{F}$ et que $\mathscr{F}$ converge vers a. ∎

1.7 Partitions de l'unité

Définition 1.7.1. *Soit E un espace topologique. Une famille $(\varphi_j)_{j\in J}$ de fonctions continues de E dans $\mathbb{R}$ est dite* localement finie *si chaque point a de E possède un voisinage W sur lequel toutes les fonctions φ_j sauf un nombre fini sont identiquement nulles, c'est-à-dire*

$$\forall a \in E\ \exists W \in \mathscr{V}(a)\ \exists J_0 \text{ partie finie de } J\ \forall x \in W\ \forall j \in J \setminus J_0 \quad \varphi_j(x) = 0\ .$$

En particulier, la somme $\sum_{j\in J}\varphi_j$ est bien définie et continue, puisque, au voisinage de chaque point, elle coïncide avec la somme d'une sous-famille finie d'entre elles.

Définition 1.7.2. *Soient E un espace topologique et $\mathscr{U} = (U_j)_{j\in J}$ un recouvrement ouvert de E. Une famille localement finie $(\varphi_j)_{j\in J}$ de fonctions continues de E dans $\mathbb{R}^+$ est appelée* partition de l'unité subordonnée à $\mathscr{U}$ *si*

i) $\forall j \in J\ \forall x \in E \quad \varphi_j(x) > 0 \Longrightarrow x \in U_j$.

ii) $\sum_{j\in J}\varphi_j = 1$.

On va montrer ici que pour tout espace métrique séparable E et tout recouvrement ouvert $\mathscr{U}$ de E, il existe une partition de l'unité subordonnée à $\mathscr{U}$. On peut en fait montrer, en utilisant l'axiome du choix (cf. 10.5.3), qu'on peut omettre l'hypothèse de séparabilité. Néanmoins, l'existence de partitions de l'unité sera surtout utilisée dans le cas où E est un sous-espace de l'espace euclidien $\mathbb{R}^d$.

Lemme 1.7.3. *Soient E un espace métrique séparable et $\mathscr{U} = (U_j)_{j\in J}$ un recouvrement ouvert de E. Alors il existe un sous-recouvrement dénombrable de $\mathscr{U}$, c'est-à-dire qu'il existe une partie dénombrable D de J telle que $(U_j)_{j\in D}$ soit un recouvrement de E.*

Démonstration : Puisque E est séparable, il existe une base dénombrable $(O_n)_{n\in\mathbb{N}}$ de la topologie de E. Alors l'ensemble $\Delta = \{n \in \mathbb{N} : \exists j \in J \quad O_n \subset U_j\}$ est dénombrable, et on peut trouver une fonction $\psi : \Delta \to J$ telle que, pour tout $n \in \Delta$, on ait $O_n \subset U_{\psi(n)}$. Alors l'ensemble $D = \psi(\Delta)$ est une partie dénombrable de J.
De plus, si x est un point quelconque de E, il existe, puisque $\mathscr{U}$ est un recouvrement de E, un indice $j \in J$ tel que $x \in U_j$; et puisque U_j est un voisinage de x, il existe un ouvert de base O_m tel que $x \in O_m \subset U_j$. Ceci montre que $m \in \Delta$ et que $x \in O_m \subset U_{\psi(m)}$ (même si $\psi(m) \neq j$). On en conclut que $x \in \bigcup_{j\in D} U_j$, donc que $(U_j)_{j\in D}$ est un sous-recouvrement dénombrable de $\mathscr{U}$. ■

Théorème 1.7.4. *Soient E un espace métrique séparable et $\mathscr{U} = (U_j)_{j\in J}$ un recouvrement ouvert de E. Alors il existe une partition de l'unité subordonnée à $\mathscr{U}$.*

Il résulte du lemme précédent qu'on peut supposer le recouvrement $\mathscr{U}$ dénombrable, et même supposer que J est une partie de $\mathbb{N}$.
Alors, pour tout $n \in J$, la fonction $\varphi_n : x \mapsto \inf(d(x, E \setminus U_n), 2^{-n})$ est continue sur E, nulle hors de U_n et strictement positive sur U_n.
La fonction $\varphi : x \mapsto \sup_{n\in J}\varphi_n(x)$ est alors strictement positive sur E puisque $\mathscr{U}$ est un recouvrement et que $\varphi(x) \geqslant \varphi_n(x) > 0$ si $x \in U_n$. De plus φ est continue sur E. En effet, soit $a \in E$: puisque $\varphi(a) > 0$, on peut choisir $m \in J$ tel que $\varphi_m(a) > \dfrac{1}{2}\varphi(a)$ puis n_0 tel que

$2^{-n_0} < \varphi(a)$ et trouver un voisinage W de a tel que $\varphi_m(x) > \frac{1}{2}\varphi(a)$ en tout point $x \in W$. Alors, pour $n \in J$ et $x \in W$, on a, lorsque $n > n_0$:

$$\varphi_n(x) \leqslant 2^{-n} \leqslant 2^{-n_0-1} < \frac{1}{2}\varphi(a) < \varphi_m(x) \leqslant \varphi(x) \ ,$$

donc $\varphi(x) = \sup_{n\in J, n\leqslant n_0} \varphi_n(x)$ en tout point $x \in W$, ce qui montre que φ coïncide sur W avec la fonction continue $\sup_{n\in J, n\leqslant n_0} \varphi_n$.

On pose alors $\varphi'_n(x) = \max(\varphi_n(x) - \frac{1}{2}\varphi(x), 0)$. On voit immédiatement que φ'_n est continue et nulle hors de U_n. De plus, il résulte de ce qui précède que pour tout $a \in E$ il existe un voisinage W de a et un n_0 tel que $\varphi_n(x) \leqslant \frac{1}{2}\varphi(x)$, donc $\varphi'_n(x) = 0$, si $n \in J$ est supérieur à n_0 et $x \in W$. Ceci montre que la famille $(\varphi'_n)_{n\in J}$ est localement finie. Et puisque, pour tout $x \in E$ on a $\varphi(x) > 0$, il existe un $n \in J$ tel que $\varphi_n(x) > \frac{1}{2}\varphi(x)$, donc tel que $\varphi'_n(x) > 0$.
Il en résulte que la fonction $\varphi' = \sum_{n\in J} \varphi'_n$ est continue et strictement positive sur E, que $\chi_n : x \mapsto \dfrac{\varphi'_n(x)}{\varphi'(x)}$ est continue sur E et nulle hors de U_n, et que la famille $(\chi_n)_{n\in J}$ est localement finie. Et puisque l'on a pour tout x : $\sum_{n\in J} \chi_n(x) = \dfrac{1}{\varphi'(x)} \displaystyle\sum_{n\in J} \varphi'_n(x) = 1$, ceci achève de prouver que $(\chi_n)_{n\in J}$ est une partition de l'unité subordonnée à $\mathscr{U}$. ■

1.8 Espaces métriques complets

Suites de Cauchy.

On va maintenant étudier une condition un peu plus large que la compacité, qui permette d'obtenir qu'une suite est convergente, sans connaître a priori sa limite.

Définition 1.8.1. *Une suite (x_n) dans un espace métrique est appelée* suite de Cauchy *si, pour tout $\varepsilon > 0$, il existe un entier m tel que la distance de deux termes quelconques de la suite, d'indices supérieurs à m, soit inférieure à ε.*

Ceci revient à dire que, en notant Q_m l'ensemble $\{x_n : n > m\}$, on a $\lim_{m\to\infty} \operatorname{diam}(Q_m) = 0$.

Proposition 1.8.2. *Si (x_n) est une suite de Cauchy, et (y_n) une suite extraite de (x_n), la suite (y_n) est une suite de Cauchy.*

Démonstration : Il existe une suite croissante (n_k) d'entiers telle que $y_k = x_{n_k}$. Soit $\varepsilon > 0$. Il existe m tel que, pour n et p supérieurs à m, on ait $d(x_n, x_p) < \varepsilon$. Alors, si k et ℓ sont supérieurs à m, on a $n_k \geqslant k > m$ et $n_\ell \geqslant \ell > m$, donc $d(y_k, y_\ell) = d(x_{n_k}, x_{n_\ell}) < \varepsilon$, ce qui achève la démonstration. ■

Théorème 1.8.3. *Toute suite convergente dans un espace métrique est une suite de Cauchy.*

Démonstration : Supposons que la suite (x_n) converge vers a dans E. Alors, si $\varepsilon > 0$ est donné, il existe un entier m tel que $d(x_n, a) < \varepsilon/2$ pour tout $n > m$. Si n et p sont alors deux entiers supérieurs à m, on a

$$d(x_n, x_p) \leqslant d(x_n, a) + d(x_p, a) < \varepsilon/2 + \varepsilon/2 = \varepsilon \ ,$$

ce qui prouve que la suite (x_n) est de Cauchy. ■

On obtient ainsi une propriété valable pour toute suite convergente, et qui ne fait pas référence à la limite. Malheureusement, cette propriété ne caractérise pas, en général, les suites convergentes. Si E est le sous-espace $]0, 1[$ de $\mathbb{R}$, la suite définie par $x_n = 2^{-n}$ est une suite de Cauchy, puisqu'elle converge dans $\mathbb{R}$ vers 0, mais elle ne converge pas dans E.

Théorème 1.8.4. *Une suite de Cauchy est convergente si et seulement si elle a une valeur d'adhérence.*

Démonstration : Il est clair qu'une suite convergente a une valeur d'adhérence, sa limite. Inversement, si la suite de Cauchy (x_n) a une valeur d'adhérence a, nous montrons qu'elle converge vers a.
Soit $\varepsilon > 0$. Il existe un entier m tel que, pour n et p supérieurs à m, on ait $d(x_n, x_p) < \varepsilon/2$. Puisque a est une valeur d'adhérence de la suite, il existe un $p > m$ tel que $d(x_p, a) < \varepsilon/2$. Alors, pour $n > m$ on a

$$d(x_n, a) \leqslant d(x_n, x_p) + d(x_p, a) < \varepsilon/2 + \varepsilon/2 = \varepsilon \ ,$$

ce qui montre que la suite (x_n) converge vers a. ■

Proposition 1.8.5. *Si f est uniformément continue de l'espace métrique E dans l'espace métrique F, toute suite de Cauchy de E est transformée par f en une suite de Cauchy de F.*

Démonstration : Soient (x_n) une suite de Cauchy de E, $(y_n) = (f(x_n))$ la suite image par f, et $\varepsilon > 0$. Puisque f est uniformément continue, il existe $\delta > 0$ tel que

$$\forall x, x' \in E \quad d(x, x') < \delta \Longrightarrow d(f(x), f(x')) < \varepsilon \ ;$$

alors, puisque (x_n) est une suite de Cauchy, il existe $m \in \mathbb{N}$ tel que si n et p sont supérieurs à m, on ait $d(x_n, x_p) < \delta$. Pour de tels n et p, on a $d(y_n, y_p) < \varepsilon$, ce qui montre que (y_n) est une suite de Cauchy. ■

Définition 1.8.6. *Un espace métrique E est dit* complet *si toute suite de Cauchy de E est convergente.*

On a alors un critère de convergence pour une suite, qui ne demande pas la connaissance a priori de la limite. On peut remarquer néanmoins que ce critère n'est pas topologique, c'est-à-dire invariant par homéomorphie. On verra un peu plus loin que $\mathbb{R}$ est complet, alors que l'intervalle $]0, 1[$, qui lui est homéomorphe, n'est pas complet, comme le montre l'exemple précédent de suite de Cauchy non convergente.

Théorème 1.8.7. *Un sous-espace complet F d'un espace métrique E est fermé dans E.*

Démonstration : Soit a un point de E adhérent à F. Il existe alors une suite (x_n) de points de F qui converge vers a. Cette suite convergente de E est donc une suite de Cauchy dans F. Et si F est complet, elle converge vers un point b de F. Par unicité de la limite, on a $a = b$. Donc $a \in F$, ce qui prouve que F contient chacun de ses points adhérents, donc est fermé. ■

Théorème 1.8.8. *Tout sous-espace fermé d'un espace métrique complet est complet.*

Démonstration : Soit F un sous-espace fermé de E. Si (x_n) est une suite de Cauchy de F, c'est en particulier une suite de Cauchy de E, donc une suite convergente dans E si celui-ci est complet. La limite a de cette suite est alors un point adhérent à F, donc un point de F puisque celui-ci est fermé. On en conclut que la suite (x_n) est convergente dans F, donc que F est complet. ■

Théorème 1.8.9. *Soient E un espace métrique complet, et (F_n) une suite décroissante de fermés non vides dont les diamètres tendent vers 0. Alors l'intersection des (F_n) contient un point et un seul.*

Démonstration : Si a et b étaient deux points distincts dans l'intersection des (F_n), on aurait pour tout n : $\operatorname{diam}(F_n) \geqslant d(a,b) > 0$ et le diamètre des (F_n) serait minoré par $d(a,b)$. Si on choisit pour tout n un point x_n dans F_n, on a, pour $n < p$, $x_n \in F_n$ et $x_p \in F_p \subset F_n$. Donc $d(x_n, x_p) \leqslant \operatorname{diam}(F_n)$, ce qui prouve que la suite (x_n) est une suite de Cauchy, donc qu'elle converge vers un point a de E. Comme les x_p sont tous dans le fermé F_n pour $p \geqslant n$, la limite a est dans F_n. Ceci montre que $a \in \bigcap_{n\in\mathbb{N}} F_n$. ■

Théorème 1.8.10. *Un produit fini ou dénombrable d'espaces complets est complet.*

Démonstration : Soit (E_n) une suite finie ou infinie d'espaces métriques complets. On munit le produit E des (E_n) de la distance $d : (x,y) \mapsto \sup_n \inf(2^{-n}, d(x_n,y_n))$ ou d'une ditance uniformément équivalente.
Si (x^k) est une suite de Cauchy dans E, on voit que, pour tout n, on a $d(x_n^k, x_n^\ell) \leqslant d(x^k, x^\ell)$ pour k et ℓ assez grands, (en désignant par x_n^k la $n^{\text{ème}}$ coordonnée du terme d'indice k de la suite). On en déduit que la suite $(x_n^k)_{k\in\mathbb{N}}$ est une suite de Cauchy de E_n, donc une suite qui converge vers un a_n de E_n puisque E_n est complet. Ceci prouve que la suite (x^k) converge vers $a = (a_n)$. Donc E est complet. ■

Théorème 1.8.11. *La droite réelle $\mathbb{R}$ est un espace complet.*

Démonstration : Soit (x_n) une suite de Cauchy dans $\mathbb{R}$. La suite (x_n) est bornée : en effet, il existe un m tel que, pour n et p supérieurs à m on ait $|x_n - x_p| < 1$. On en déduit que, pour tout n on a

$$|x_n| \leqslant M = 1 + \max_{j\leqslant m+1} |x_j| < +\infty ,$$

et puisque l'intervalle $[-M, +M]$ est compact, la suite de Cauchy (x_n) a au moins une valeur d'adhérence, donc est convergente (cf. 1.8.4). ■

Corollaire 1.8.12. *Pour tout entier n, l'espace $\mathbb{R}^n$ est complet.*

Théorème 1.8.13. *Si K est un espace compact et F un espace métrique complet, l'espace $\mathscr{C}(K,F)$ muni de la distance de la convergence uniforme est complet.*

Démonstration : Soit (f_n) une suite de Cauchy dans l'espace $\mathscr{C}(K,F)$. Pour chaque $x \in K$, on a $d(f_n(x), f_p(x)) \leqslant d(f_n, f_p)$. Il en résulte que l'application : $f \mapsto f(x)$ de $\mathscr{C}(K,F)$ dans F est 1-lipschitzienne, donc uniformément continue, et que la suite $(f_n(x))$ est une suite de Cauchy dans F. On en déduit que cette suite converge vers un certain point $f(x)$ de F. De plus, si $\varepsilon > 0$ est donné, il existe un $m \in \mathbb{N}$ tel que pour n et p supérieurs à m on ait $d(f_n, f_p) \leqslant \varepsilon$. On a donc, si $m < n \leqslant p$

$$d(f_n(x), f_p(x)) \leqslant d(f_n, f_p) \leqslant \varepsilon ,$$

et puisque $f(x) = \lim_{p\to\infty} f_p(x)$, on obtient, pour $n > m$ et $x \in K$:

$$d(f_n(x), f(x)) \leqslant \varepsilon \ ,$$

ce qui montre que la suite (f_n) converge uniformément vers f. Donc f est continue, et $d(f_n, f) = \sup_{x\in K} d(f_n(x), f(x)) \leqslant \varepsilon$ si $n > m$. Ceci montre que la suite (f_n) converge vers f dans $\mathscr{C}(K, F)$. ■

Compacité et complétude.

Théorème 1.8.14. *Tout espace métrique compact est complet.*

Démonstration : Soit E un espace métrique compact. Si (x_n) est une suite de Cauchy de E, elle a au moins une valeur d'adhérence dans E, donc elle est convergente. ■

Théorème 1.8.15. *Un espace métrique est compact si et seulement s'il est précompact et complet.*

Démonstration : On a déjà vu qu'un espace métrique compact est précompact et qu'il est complet.
Inversement, si E est précompact, il existe pour tout entier n un recouvrement fini $(A_i^{(n)})_{i\in J_n}$ par des parties de diamètre au plus 2^{-n}. Alors le produit $K = \prod_n J_n$ des espaces discrets finis J_n est compact. On notera π_p la projection de K sur J_p. On choisit pour tout p et tout $j \in J_p$ un point $a_j^{(p)}$ dans $A_j^{(p)}$. Alors le sous-espace K_0 de K défini par

$$K_0 = \{(j_n) \in K : \forall p \ \forall q \quad d(a_{j_p}^{(p)}, a_{j_q}^{(q)}) \leqslant 2^{-p} + 2^{-q}\}$$

est clairement fermé dans K, donc lui aussi compact. On définit alors une fonction φ_p sur K par $\varphi_p(z) = a_{\pi_p(z)}^{(p)}$. Puisque J_p est discret et que π_p est continue, cette fonction est continue sur K, et a fortiori sur K_0. De plus, il résulte de la définition de K_0 que $d(\varphi_p, \varphi_q) \leqslant 2^{-p}+2^{-q}$, donc que la suite (φ_p) est une suite de Cauchy dans l'espace métrique complet $\mathscr{C}(K_0, E)$. Cette suite converge donc uniformément vers une fonction continue $\varphi : K_0 \to E$. Donc $\varphi(K_0)$ est compact, et il suffit de montrer que φ est surjective.
Soit donc $x \in E$. Si, pour tout entier p on choisit un $j_p \in J_p$ tel que $x \in A_{j_p}^{(p)}$ et si on pose $z = (j_p) \in K$, on a pour p et q :

$$d(a_{j_p}^{(p)}, a_{j_q}^{(q)}) \leqslant d(a_{j_p}^{(p)}, x) + d(x, a_{j_q}^{(q)}) \leqslant \operatorname{diam}(A_{j_p}^{(p)}) + \operatorname{diam}(A_{j_q}^{(q)}) \leqslant 2^{-p} + 2^{-q} \ ,$$

ce qui montre que $z \in K_0$. Et puisque, pour tout p, on a $d(\varphi_p(z), x) \leqslant 2^{-p}$, on voit que $\varphi_p(z) \to x$, donc que $\varphi(z) = x$, ce qui achève la preuve. ■

Prolongement d'une application uniformément continue.

Théorème 1.8.16. *Soient E et F deux espaces métriques, F étant complet. On suppose que X est une partie partout dense de E et f une application uniformément continue de X dans F. Il existe alors une unique application continue $\tilde f$ de E dans F qui prolonge f.*

Démonstration : Si g_1 et g_2 sont deux prolongements continus de f à E, l'application $g = g_1 \times g_2 : E \to F \times F$ définie par $g(x) = (g_1(x), g_2(x))$ est continue. L'ensemble A des

points de E où g_1 et g_2 coïncident est l'image réciproque par g de la diagonale de F, donc est fermé, et contient X puisque les restrictions de g_1 et g_2 à X sont égales à f. On en déduit que $A = E$ puisque X est dense, c'est-à-dire que $g_1 = g_2$, ce qui prouve l'unicité du prolongement.
Soient a un point de E et (x_n) une suite de points de X qui converge vers a (une telle suite existe puisque X est dense dans E). Si $\tilde{f}$ est une fonction continue sur E qui prolonge f, on doit avoir

$$\tilde{f}(a) = \lim_{n\to\infty} \tilde{f}(x_n) = \lim_{n\to\infty} f(x_n) \ .$$

Pour prouver l'existence de $\tilde{f}$ on doit donc prouver que toute suite (x_n) de X qui converge vers un point a de E a une image par f qui converge dans F, et que cette limite ne dépend que de a. En effet, si la suite (x_n) converge vers a, c'est une suite de Cauchy, qui est transformée par l'application uniformément continue f en une suite de Cauchy de l'espace complet F, donc en une suite convergente. Et si (x_n) et (y_n) sont deux suites de X qui convergent vers le même point a de E, la suite (z_n) définie par

$$z_{2n} = x_n \qquad \text{et} \qquad z_{2n+1} = y_n$$

converge elle aussi vers a. Alors, si b est la limite de la suite $(f(z_n))$, on a $f(x_n) = f(z_{2n}) \to b$ et $f(y_n) = f(z_{2n+1}) \to b$, ce qui montre que les suites $(f(x_n))$ et $(f(y_n))$ ont même limite, qu'on notera $\tilde{f}(a)$.
Pour prouver la continuité de la fonction $\tilde{f}$ ainsi définie, on va montrer que $\tilde{f}$ est uniformément continue. Soit donc $\varepsilon > 0$. Il existe un $\delta > 0$ tel que si x et y appartiennent à X et sont à distance strictement inférieure à δ, $f(x)$ et $f(y)$ sont à distance inférieure à ε. Soient alors a et b dans E avec $d(a,b) < \delta$, et (x_n) et (y_n) deux suites de X convergeant respectivement vers a et b. Puisque

$$d(x_n, y_n) \leqslant d(x_n, a) + d(a,b) + d(b, y_n) \to d(a,b) < \delta \ ,$$

on a $d(x_n, y_n) < \delta$ pour tout n assez grand, donc $d(f(x_n), f(y_n)) \leqslant \varepsilon$ pour n assez grand. Et puisque $d(f(x_n), f(y_n)) \to d(\tilde{f}(a), \tilde{f}(b))$, on obtient que $d(\tilde{f}(a), \tilde{f}(b)) \leqslant \varepsilon$. Ceci montre que $\tilde{f}$ est uniformément continue. ∎

Points fixes des contractions.

Définition 1.8.17. *On dit que l'application f de l'espace métrique E dans l'espace métrique F est une* contraction *si elle est q-lipschitzienne pour un q strictement inférieur à* 1.

Théorème 1.8.18. *Si f est une contraction de l'espace métrique complet E dans lui-même, il existe un unique* point fixe *de f, c'est-à-dire un point a tel que $a = f(a)$. De plus, pour tout $x_0 \in E$, la suite (x_n) définie par récurrence par $x_{n+1} = f(x_n)$ converge vers le point fixe.*

Démonstration : Soit x_0 un point quelconque de E. On définit par récurrence la suite (x_n) en posant $x_{n+1} = f(x_n)$. Puisque f est q-lipschitzienne, on a

$$d(x_{n+1}, x_{n+2}) = d(f(x_n), f(x_{n+1})) \leqslant q\, d(x_n, x_{n+1}) \ ,$$

donc $d(x_n, x_{n+1}) \leqslant q^n d(x_0, x_1)$. Et, pour $n < p$

$$d(x_n, x_p) \leqslant \sum_{j=n}^{p-1} d(x_j, x_{j+1}) \leqslant \sum_{j=n}^{p-1} q^j \, d(x_0, x_1) = \frac{q^n - q^p}{1-q} \, d(x_0, x_1) < q^n \, \frac{d(x_0, x_1)}{1-q} \ ,$$

ce qui montre que la suite (x_n) est une suite de Cauchy puisque $\lim_{n\to\infty} q^n = 0$.
La suite (x_n) est donc convergente, et si a est sa limite, on doit avoir $x_{n+1} \to a$ et $x_{n+1} = f(x_n) \to f(a)$, donc $a = f(a)$. Donc a est un point fixe de f.
Enfin, si a et b sont deux points fixes distincts de f, on doit avoir :

$$d(a,b) = d(f(a), f(b)) \leqslant q \, d(a,b) < d(a,b) \ ,$$

ce qui est absurde. On en déduit l'unicité du point fixe. ■

1.9 Le théorème de Baire

Définition 1.9.1. *Une partie X d'un espace topologique est dite* rare *si elle est contenue dans un fermé d'intérieur vide (c'est-à-dire si son adhérence est d'intérieur vide).*

Une partie d'un espace topologique est dite maigre *si elle est réunion dénombrable de parties rares.*

Théorème 1.9.2. *Si E est un espace métrique complet, et M une partie maigre de E, alors $E \setminus M$ est partout dense dans E. En particulier, si $E \neq \emptyset$ est réunion dénombrable de fermés (F_n), l'un d'entre eux au moins est d'intérieur non vide.*

Démonstration : Si M est maigre, il existe une suite (R_n) de fermés rares de E qui recouvre M. Il suffit donc de démontrer que $G = E \setminus \bigcup_{n\in\mathbb{N}} R_n$ est partout dense. Soient $x_0 \in E$ et $\varepsilon > 0$. On va montrer que G rencontre la boule de centre x_0 et de rayon ε. On construit pour cela une suite (x_n) de points de E et une suite (r_n) de nombres strictement positifs telles que

i) $r_0 = \varepsilon$,

ii) $r_{n+1} \leqslant r_n/2$,

iii) $\tilde{B}(x_{n+1}, r_{n+1}) \subset B(x_n, r_n) \setminus R_n$,

en notant $\tilde{B}(x, r)$ la boule fermée de centre x et de rayon r. Pour cela, si x_n et r_n sont déterminés, le fermé rare R_n ne peut contenir la boule ouverte $B(x_n, r_n)$: il serait sinon d'intérieur non vide. On peut donc trouver un point x_{n+1} dans $B(x_n, r_n) \setminus R_n$. Et puisque $B(x_n, r_n) \setminus R_n$ est ouvert, il existe un r_{n+1}, qu'on peut supposer inférieur à $r_n/2$, tel que $\tilde{B}(x_{n+1}, r_{n+1}) \subset B(x_n, r_n) \setminus R_n$.
Alors, puisque les boules $\tilde{B}(x_n, r_n)$ forment une suite décroissante, on a, pour $n < p$, $x_p \in \tilde{B}(x_n, r_n)$, donc $d(x_n, x_p) \leqslant r_n \leqslant r_0 \, 2^{-n}$, en vertu de ii). On en déduit que la suite (x_n) est une suite de Cauchy dans E, donc qu'elle converge vers un x de E. Puisque $\tilde{B}(x_{n+1}, r_{n+1})$ est fermée et contient tous les (x_p) pour $p > n$, elle contient x ; donc $x \notin R_n$, et puisque ceci est valable pour tout n, $x \in G$. On en conclut que $B(x_0, \varepsilon) \cap G$ est non vide.

En particulier, E ne peut être réunion dénombrable de fermés rares sans être vide. ■

1.10 Espaces fonctionnels

Ensembles compacts de fonctions continues.

Soient K un espace topologique compact et F un espace métrique complet. On a vu au théorème 1.8.13 que l'espace métrique $\mathscr{C}(K,F)$ des fonctions continues de K dans F, muni de la distance de la convergence uniforme, est complet. On veut maintenant déterminer quelles sont les parties relativement compactes de cet espace.

Proposition 1.10.1. *Si H est une partie compacte de $\mathscr{C}(K,F)$ et $x \in K$, l'ensemble $H(x)$ des valeurs prises en x par les éléments de H est une partie compacte de F.*

Démonstration : Si f et g sont deux éléments de $\mathscr{C}(K,F)$, on a

$$d(f(x),g(x)) \leqslant \sup_{y \in K} d(f(y),g(y)) = d(f,g)\,,$$

ce qui entraîne que l'application $V_x : f \mapsto f(x)$ est 1-lipschitzienne de $\mathscr{C}(K,F)$ dans F, donc continue. Il en résulte que $H(x) = V_x(H)$ est compact. ■

Définition 1.10.2. *Une partie H de $\mathscr{C}(K,F)$ est dite* équicontinue *au point x de K si, pour tout $\varepsilon > 0$ il existe un voisinage V de x dans K tel que, pour tout f de H et tout y de V on ait $d(f(x),f(y)) < \varepsilon$.*

Dans le cas où K est métrisable, on peut donner une définition en apparence plus restrictive.

Définition 1.10.3. *Une partie H de $\mathscr{C}(K,F)$ est dite* uniformément équicontinue *sur K si, pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que, pour tout f de H, tout x et tout y de V on ait $d(f(x),f(y)) < \varepsilon$ dès que $d(x,y) < \delta$.*

En fait, une démonstration analogue à celle du théorème de Heine (théorème 1.5.28), montre qu'une partie H équicontinue en chaque point du compact métrisable K est nécessairement uniformément équicontinue.

Théorème 1.10.4. *Si H est une partie compacte de $\mathscr{C}(K,F)$, H est équicontinue en tout point de K.*

Démonstration : Soient $x \in K$ et $\varepsilon > 0$. Pour chaque voisinage V de x , on considère l'ensemble $A(V)$ de toutes les fonctions continues f de K dans F qui vérifient $\sup_{y \in V} d(f(y),f(x)) < \varepsilon$.
Pour toute fonction f continue en x, il existe un voisinage V de x tel que $f \in A(V)$. Il en résulte que, notant $\mathscr{V}$ l'ensemble des voisinages de x, les $(A(V))_{V \in \mathscr{V}}$ forment un recouvrement de $\mathscr{C}(K,F)$. De plus, chaque $A(V)$ est une partie ouverte de $\mathscr{C}(K,F)$. En effet, si $f \in A(V)$ et si on pose

$$\sup_{y \in V} d(f(y),f(x)) = \varepsilon - \delta\,,$$

on a $\delta > 0$. Pour $g \in \mathscr{C}(K,F)$ telle que $d(f,g) < \delta/3$, et $y \in V$ on obtient :

$$\begin{aligned} d(g(y),g(x)) &\leqslant d(g(y),f(y)) + d(f(y),f(x)) + d(f(x),g(x)) \\ &\leqslant 2\,d(f,g) + (\varepsilon - \delta) < \varepsilon - \delta/3\,. \end{aligned}$$

Donc $\sup_{y \in V} d(g(y), g(x)) \leqslant \varepsilon - \delta/3 < \varepsilon$, et g appartient à $A(V)$. Le recouvrement ouvert $(A(V))_{V \in \mathscr{V}}$ du compact H possède donc un sous-recouvrement fini. Il existe donc V_1, V_2, ..., V_p dans $\mathscr{V}$ tels que $H \subset \bigcup_{1 \leqslant j \leqslant p} A(V_j)$. Si on pose $V = \bigcap_{1 \leqslant j \leqslant p} V_j$, V est un voisinage de x. De plus, pour tout $f \in H$ et tout $y \in V$, il existe un j tel que $f \in A(V_j)$ et $y \in V_j$, donc on a $d(f(y), f(x)) < \varepsilon$, et l'équicontinuité de H en x est prouvée. ∎

Corollaire 1.10.5. *Si H est une partie relativement compacte de $\mathscr{C}(K, F)$, l'ensemble $H(x) = \{f(x) : f \in H\}$ est relativement compact pour tout x de K, et H est équicontinue en chaque point de K.*

Démonstration : Si $\overline{H}$ est l'adhérence de H dans $\mathscr{C}(K, F)$, $H(x)$ est contenue pour tout x dans le compact $\overline{H}(x)$. Et puisque $\overline{H}$ est équicontinue en chaque point, H l'est a fortiori. ∎

On va voir que ces propriétés caractérisent en fait les parties relativement compactes de $\mathscr{C}(K, F)$. Nous ne le montrerons que dans le cas où K est un espace métrique compact (cas auquel on peut en fait toujours se ramener).

Théorème 1.10.6. (Ascoli) *Soient K un espace métrique compact, F un espace métrique complet, H une partie de $\mathscr{C}(K, F)$. On suppose que $H(x)$ est relativement compact pour tout x de K et que H est équicontinu en chaque point de K. Alors H est relativement compact dans $\mathscr{C}(K, F)$.*

Démonstration : D'après le théorème 1.5.18, il existe une suite (x_k) de points de K qui est partout dense dans K. Pour chaque entier k, $H(x_k)$ est relativement compact, donc contenu dans une partie compacte F_k de F. D'après le théorème 1.5.20, le produit $T = \prod_{k \in \mathbb{N}} F_k$ est compact. Pour montrer que $\overline{H}$ est compact, on va montrer que toute suite (f_n) dans $\overline{H}$ possède une sous-suite qui converge dans $\mathscr{C}(K, F)$ pour la distance de la convergence uniforme. On choisit pour tout n une fonction g_n dans H telle que $d(f_n, g_n) < 2^{-n}$, ce qui est possible puisque f_n est adhérente à H.
On considère l'application $\Phi : H \to T$ définie par $f \mapsto (f(x_k))_{k \in \mathbb{N}}$. Alors la suite $(\Phi(g_n))_{n \in \mathbb{N}}$ possède une sous-suite $(\Phi(g_{n_j}))_{j \in \mathbb{N}}$ qui converge dans le compact T. Donc, quel que soit k, la suite $(g_{n_j}(x_k))_j$ est convergente. On va montrer que la suite $(g_{n_j})_j$ est une suite de Cauchy dans $\mathscr{C}(K, F)$, ce qui entraînera qu'elle converge, puisque $\mathscr{C}(K, F)$ est complet. On sait que H est équicontinue en chaque point, donc uniformément équicontinue. Soit donc $\varepsilon > 0$. Il existe un $\delta > 0$ tel que pour $f \in H$, x et y dans K avec $d(x, y) < \delta$, on ait $d(f(x), f(y)) < \varepsilon/3$. Puisque toute boule de rayon δ dans K contient au moins l'un des (x_k), les boules de centre x_k et de rayon δ forment un recouvrement ouvert de K. Il en résulte qu'existe un ℓ tel que $K = \bigcup_{k \leq \ell} B(x_k, \delta)$. On peut donc trouver un m tel que si $i > m$ et $j > m$ on ait $d(g_{n_j}(x_k), g_{n_i}(x_k)) < \varepsilon/3$ pour $k = 0, 1, 2, \ldots, \ell$.
Alors, pour x quelconque dans K, i et j supérieurs à m, on peut, en choisissant k tel que $d(x, x_k) < \delta$, majorer $d(g_{n_i}(x), g_{n_j}(x))$ par

$$\begin{aligned} d(g_{n_i}(x), g_{n_i}(x_k)) + d(g_{n_i}(x_k), g_{n_j}(x_k)) + d(g_{n_j}(x_k), g_{n_j}(x)) \\ < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon\,, \end{aligned}$$

d'où $d(g_{n_i}, g_{n_j}) \leqslant \varepsilon$, pour i et j supérieurs à m. Ceci montre que la suite (g_{n_j}) est une suite de Cauchy dans l'espace complet $\mathscr{C}(K, F)$, qui converge donc uniformément vers une fonction continue g. Puisque les (g_n) appartiennent à H, $g \in \overline{H}$, et puisque $(d(f_{n_j}, g_{n_j}))$ tend vers 0, la suite (f_{n_j}) tend également vers g. Ceci achève de prouver la compacité de $\overline{H}$. ■

Ensembles denses de fonctions continues.

On suppose maintenant que K est un espace compact, et on cherche des critères permettant de montrer qu'une partie de $\mathscr{C}(K,\mathbb{R})$ ou de $\mathscr{C}(K,\mathbb{C})$ est partout dense.

Lemme 1.10.7. *Soit $a > 0$. Il existe une suite (P_n) de polynômes à coefficients réels qui converge uniformément sur $[-a,a]$ vers la fonction $t \mapsto |t|$.*

Démonstration : Posons $P_0(t) = 0$, et définissons par récurrence les polynômes (P_n) par

$$P_{n+1}(t) = P_n(t) + \frac{1}{2a}\left[t^2 - P_n^2(t)\right] \ . \qquad (*)$$

On montre par récurrence que,pour tout $n \in \mathbb{N}$ et tout $t \in [-a,a]$, on a :

$$0 \leqslant P_n(t) \leqslant P_{n+1}(t) \leqslant |t| \ .$$

Supposons en effet que $0 \leqslant P_n(t) \leqslant |t|$ pour $|t| \leqslant a$. On a alors, pour $t \in [-a,a]$: $t^2 - P_n^2(t) \geqslant 0$ donc $P_{n+1}(t) \geqslant P_n(t)$.
De plus,

$$\begin{aligned} t^2 - P_{n+1}^2(t) &= \big(|t| - P_{n+1}(t)\big)\big(|t| + P_{n+1}(t)\big) \\ &= \big(|t| + P_{n+1}(t)\big)\left[|t| - P_n(t) - \frac{1}{2a}\big(|t| - P_n(t)\big)\big(|t| + P_n(t)\big)\right] \\ &= \big(|t| + P_{n+1}(t)\big)\big(|t| - P_n(t)\big)\left[1 - \frac{1}{2a}\big(|t| + P_n(t)\big)\right] \ , \end{aligned}$$

et il suffit de remarquer que $|t| + P_n(t) \leqslant 2\,|t| \leqslant 2a$ pour voir que les trois facteurs ci-dessus sont positifs. Il en résulte que $t^2 - P_{n+1}(t)^2 \geqslant 0$, donc que $P_{n+1}(t) \leqslant |t|$.
La suite $(P_n(t))$ est donc, pour tout $t \in [-a,a]$, une suite croissante et majorée par $|t|$. Donc elle converge vers un $\varphi(t) \geqslant 0$. On doit donc avoir, en utilisant $(*)$,

$$\varphi(t) = \varphi(t) + \frac{1}{2a}(t^2 - \varphi(t)^2)\,,$$

donc $\varphi(t)^2 = t^2$, c'est-à-dire $\varphi(t) = |t|$ puisque $\varphi(t) \geqslant 0$. Alors le théorème de Dini (théorème 1.5.30) montre que la convergence est uniforme, puisque la fonction limite est continue. ■

Lemme 1.10.8. *Soit V un sous-ensemble de $\mathscr{C}(K,\mathbb{R})$ tel que :*

i) V contient les fonctions constantes.

ii) Pour tout couple (x,y) de points distincts de K, tout couple (u,v) de réels et tout $\varepsilon > 0$, il existe $f \in V$ telle que $|f(x) - u| < \varepsilon$ et $|f(y) - v| < \varepsilon$.

iii) Chaque fois que V contient deux fonctions f et g, il contient aussi $\max(f,g)$ *et* $\min(f,g)$.

Alors V est partout dense dans $\mathscr{C}(K,\mathbb{R})$.

Démonstration : Soient $g \in \mathscr{C}(K,\mathbb{R})$ et $\varepsilon > 0$. Prenons d'abord un point a de K. Pour tout x de K, il existe une fonction $f_x \in V$ telle que $f_x(a) \in]g(a)-\varepsilon, g(a)+\varepsilon[$ et $f_x(x) \in]g(x)-\varepsilon, g(x)+\varepsilon[$: pour $x = a$, il suffit de prendre la fonction constante de valeur $g(a)$, et pour $x \neq a$, c'est l'hypothèse (ii). Et pour chaque $x \in K$, l'ensemble

$$U_x = \{y \in K : f_x(y) < g(y) + \varepsilon\}$$

est un voisinage ouvert de x. La famille $(U_x)_{x\in K}$ est un recouvrement ouvert du compact K, dont on peut donc extraire un sous-recouvrement fini. Il existe donc une partie finie J de K telle que $K = \bigcup_{x\in J} U_x$. Alors la fonction

$$h_a = \min_{x\in J} f_x$$

appartient à V d'après iii), est partout inférieure à $g+\varepsilon$ et vérifie $h_a(a) > g(a) - \varepsilon$.
À toute fonction h de V inférieure à $g+\varepsilon$, on peut associer l'ouvert

$$W_h = \{y \in K : h(y) > g(y) - \varepsilon\}.$$

Ce qui précède montre que les ouverts (W_h) forment un recouvrement ouvert du compact K. Il existe donc une famille finie $h_1, h_2, \ldots, h_p$ de fonctions de V inférieures à $g+\varepsilon$ telle que

$$K = \bigcup_{1\leqslant i\leqslant p} W_{h_i}.$$

Alors la fonction $f = \max_{1\leqslant i\leqslant p} h_i$ appartient à V d'après iii), est inférieure à $g+\varepsilon$, et est partout supérieure à $g-\varepsilon$. Il en résulte que $d(f,g) \leqslant \varepsilon$. Ceci montre que V est partout dense dans $\mathscr{C}(K,\mathbb{R})$. ∎

Lemme 1.10.9. *Soit V un sous-espace vectoriel de $\mathscr{C}(K,\mathbb{R})$ contenant les fonctions constantes et séparant les points de K. On suppose de plus que, pour toute fonction f de V, la fonction $|f|$ est dans V. Alors V est partout dense dans $\mathscr{C}(K,\mathbb{R})$.*

Démonstration : Il suffit de montrer que V vérifie les hypothèses du lemme précédent. Si $x \neq y$ et si u et v sont deux réels, il existe, puisque V sépare les points, une fonction $g \in V$ telle que $g(x) \neq g(y)$. Alors la fonction définie par

$$f(z) = u + (v-u)\frac{g(z)-g(x)}{g(y)-g(x)}$$

appartient à l'espace vectoriel V, vaut u en x et v en y.
Enfin, si g et h sont deux fonctions dans V, on a

$$\max(g,h) = \frac{1}{2}\big(g+h+|g-h|\big)$$
$$\min(g,h) = \frac{1}{2}\big(g+h-|g-h|\big),$$

ce qui montre que $\max(g,h)$ et $\min(g,h)$ appartiennent à V. ∎

Théorème 1.10.10. (Stone-Weierstrass) *Soit A une sous-algèbre de $\mathscr{C}(K,\mathbb{R})$ contenant les constantes et séparant les points. Alors A est partout dense dans $\mathscr{C}(K,\mathbb{R})$.*

DÉMONSTRATION : En général, A n'est pas stable par valeurs absolues, mais il suffit de démontrer que l'adhérence $\overline{A}$ de A dans $\mathscr{C}(K,\mathbb{R})$ vérifie les conditions du lemme précédent. En effet, on aura alors que $\overline{A}$ est dense dans $\mathscr{C}(K,\mathbb{R})$, donc égale à $\mathscr{C}(K,\mathbb{R})$ puisqu'elle est fermée.
Clairement $\overline{A}$ contient les constantes et sépare les points de K. Pour voir que $\overline{A}$ est une algèbre, prenons f et g dans $\overline{A}$, et $\lambda \in \mathbb{R}$. Il existe deux suites (f_n) et (g_n) de A qui convergent uniformément vers f et g respectivement. Alors, les fonctions $(f_n + \lambda g_n)$ appartiennent à A et convergent uniformément vers $f + \lambda g$. Donc $f + \lambda g \in \overline{A}$. De même, la suite $(f_n g_n)$ d'éléments de A converge uniformément vers fg, et fg appartient à $\overline{A}$.
Enfin, si $f \in \overline{A}$, nous montrons que $|f| \in \overline{A}$. Puisque f est continue sur le compact K, elle est bornée, et il existe un $a > 0$ tel que $|f(x)| \leqslant a$ pour tout x de K. Il existe, d'après le lemme 1.10.7, une suite de polynômes (P_n) à coefficients réels qui converge uniformément sur $[-a,a]$ vers la fonction $t \mapsto |t|$. Alors les fonctions $(P_n \circ f)$ appartiennent à $\overline{A}$ et convergent vers $|f|$. Donc $|f|$ appartient à $\overline{A}$. ■

Pour les algèbres de fonctions complexes, les conditions du théorème de Stone-Weierstrass ci-dessus ne suffisent pas à assurer la densité. On a néanmoins le résultat suivant.

Théorème 1.10.11. *Soit A une sous-algèbre de $\mathscr{C}(K,\mathbb{C})$ qui contient les constantes et sépare les points de K. Si, de plus, la sous-algèbre A est involutive, c'est-à-dire contient la fonction conjuguée $\bar{f}$ de chaque fonction f de A, alors A est partout dense dans $\mathscr{C}(K,\mathbb{C})$.*

DÉMONSTRATION : Si on note $\Re e(A)$ l'ensemble des parties réelles $\Re e(f)$ des éléments f de A, on a $\Re e(A) = A \cap \mathscr{C}(K,\mathbb{R})$ puisque $\Re e(f) = \dfrac{f+\bar{f}}{2} \in A$ si $f \in A$. Alors $\Re e(A)$ est une algèbre de fonctions réelles sur K qui contient les constantes et sépare les points de K : en effet si $f \in A$ avec $f(x) \neq f(y)$, on a $\Re e(f(x)) \neq \Re e(f(y))$ ou $\Im m(f(x)) \neq \Im m(f(y))$, c'est-à-dire $\Re e(if(x)) \neq \Re e(if(y))$. Donc $\Re e(A)$ est dense dans $\mathscr{C}(K,\mathbb{R})$ et si $g \in \mathscr{C}(K,\mathbb{C})$, on a, pour $\varepsilon > 0$, des fonctions f_1 et f_2 de A telles que $d(\Re e(g), \Re e(f_1)) < \varepsilon/2$ et $d(\Im m(g), \Im m(f_1)) < \varepsilon/2$. Alors, en posant $f = \dfrac{f_1 + \bar{f}_1 + if_2 - i\bar{f}_2}{2}$, on a $f \in A$ et $d(f,g) < \varepsilon$. ■

2

ESPACES DE BANACH

Les espaces vectoriels considérés dans ce chapitre seront toujours des espaces vectoriels sur $\mathbb{R}$ ou sur $\mathbb{C}$. On notera $\mathbb{K}$ le corps de base pour désigner indifféremment $\mathbb{R}$ ou $\mathbb{C}$. Pour λ dans $\mathbb{K}$, la notation $|\lambda|$ désignera la valeur absolue de λ si $\mathbb{K} = \mathbb{R}$, et le module de λ si $\mathbb{K} = \mathbb{C}$

2.1 Espaces vectoriels normés

Définition 2.1.1. *On appelle* norme *sur un espace vectoriel une application* $\|.\|$ *de* E *dans* $\mathbb{R}^+$ *vérifiant les conditions :*

$$\|x\| = 0 \iff x = 0 \ . \tag{i}$$

$$\|\lambda x\| = |\lambda|\,\|x\| \qquad \text{pour tout } x \in E \text{ et tout } \lambda \in \mathbb{K} \ . \tag{ii}$$

$$\|x + y\| \leqslant \|x\| + \|y\| \qquad \text{pour tout } x \text{ et tout } y \text{ de } E \ . \tag{iii}$$

Proposition 2.1.2. *Si* $\|.\|$ *est une norme sur l'espace vectoriel* E*, alors la fonction réelle* $d : (x, y) \mapsto \|x - y\|$ *est une distance sur* E*.*

Démonstration : On a $d(x, y) = 0 \iff x - y = 0 \iff x = y$,
$d(x, y) = \|x - y\| = \|(-1).(x - y\|) = |-1|\,.\,\|y - x\| = \|y - x\| = d(y, x)$ et

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

ce qui achève la preuve. ■

Un espace vectoriel normé sera muni de la distance et de la topologie associées à la norme.

Définition 2.1.3. *On appelle* boule unité *d'un espace normé la boule de centre* 0 *et de rayon 1.*

Définition 2.1.4. *On appelle* espace de Banach *tout espace normé complet pour la distance associée à la norme.*

Définition 2.1.5. *Deux normes* $\|.\|$ *et* $|||.|||$ *sur un espace vectoriel* E *sont dites* équivalentes *s'il existe deux constantes* α *et* β *strictement positives telles que, pour tout* x *de* E*,*

$$\alpha\,\|x\| \leqslant |||x||| \leqslant \beta\,\|x\| \ .$$

Théorème 2.1.6. *Deux normes $\|.\|$ et $|||.|||$ définissent sur E la même topologie si et seulement si elles sont équivalentes.*

DÉMONSTRATION : Si les deux normes sont équivalentes, il est clair que les distances associées sur E sont équivalentes, donc que les topologies associées coïncident.
Inversement, si elles ne sont pas équivalentes, on a soit $\sup_{x\in E, x\neq 0} \dfrac{|||x|||}{\|x\|} = +\infty$, soit $\sup_{x\in E, x\neq 0} \dfrac{\|x\|}{|||x|||} = +\infty$. Quitte à les intervertir, supposons que $\sup_{x\in E, x\neq 0} \dfrac{|||x|||}{\|x\|} = +\infty$. Il existe alors une suite (x_n) dans E telle que $\dfrac{|||x_n|||}{\|x_n\|} > 2^n$. Si on pose alors $y_n = \dfrac{1}{|||x_n|||} x_n$, on a $|||y_n||| = 1$ et $\|y_n\| = \dfrac{\|x_n\|}{|||x_n|||} < 2^{-n}$. Donc la suite (y_n) converge vers 0 pour la distance associée à la norme $\|.\|$, et pas pour l'autre. Ceci montre que les topologies associées à ces deux normes sont différentes. ■

Théorème 2.1.7. *Si E est un espace vectoriel normé, les applications $(x, y) \mapsto x + y$ de $E \times E$ dans E et $(\lambda, x) \mapsto \lambda.x$ de $\mathbb{K} \times E$ dans E sont continues.*

DÉMONSTRATION : On a

$$d((x+y),(a+b)) = \|(x+y)-(a+b)\| = \|(x-a)+(y-b)\| \leqslant \|x-a\| + \|y-b\| \ ,$$

d'où la continuité de la somme, et

$$\begin{aligned} d(\lambda.x, \mu.a) &= \|\lambda.x - \mu a\| \leqslant \|\lambda(x-a\|) + \|(\lambda-\mu)a\| \\ &\leqslant |\lambda| \|x-a\| + |\lambda-\mu| \|a\| \leqslant |\lambda-\mu| \|x-a\| + |\mu| \|x-a\| + |\lambda-\mu| \|a\| \ , \end{aligned}$$

quantité qui peut être rendue inférieure à $\varepsilon > 0$ en prenant $\|x-a\|$ inférieure à $\min(1, \dfrac{\varepsilon}{3|\mu|})$ et $|\lambda-\mu|$ inférieur à $\min(\dfrac{\varepsilon}{3}, \dfrac{\varepsilon}{3\|a\|})$. ■

Proposition 2.1.8. *Les boules d'un espace normé sont convexes, donc connexes par arc.*

DÉMONSTRATION : Soit B la boule fermée de centre a et de rayon r. Si u et v appartiennent à B, $0 \leqslant t \leqslant 1$ et $w = tu + (1-t)v$, on a

$$\begin{aligned} \|w-a\| &= \|t(u-a) + (1-t)(v-a)\| \leqslant \|t(u-a)\| + \|(1-t)(v-a)\| \\ &\leqslant t\|u-a\| + (1-t)\|v-a\| \leqslant tr + (1-t)r = r \ , \end{aligned}$$

d'où on déduit que $w \in B$. Ceci prouve la convexité de B. Il résulte du théorème précédent que l'application $\gamma : t \mapsto tv + (1-t)u$ est continue de $[0, 1]$ dans B, donc est un arc qui joint u à v dans B. ■

Corollaire 2.1.9. *Tout espace normé est localement connexe.*

Définition 2.1.10. *Si E est un espace normé, une série (x_n) de points de E est dite* normalement convergente *si la série $\sum_{n=0}^{\infty} \|x_n\|$ est convergente.*

Théorème 2.1.11. *Dans un espace de Banach E, toute série normalement convergente est convergente.*

Démonstration : Soit $S_n = \sum_{j=0}^{n} x_j$. On a, pour $n < p$

$$\|S_p - S_n\| = \left\| \sum_{j=n+1}^{p} x_j \right\| \leqslant \sum_{j=n+1}^{p} \|x_j\| \leqslant \sum_{j=n+1}^{\infty} \|x_j\|$$

et puisque le reste de la série numérique convergente $\sum \|x_n\|$ tend vers 0, la suite (S_n) est une suite de Cauchy dans E, donc converge vers un élément x de E. ∎

Théorème 2.1.12. *Un espace normé E est complet si et seulement si toute série normalement convergente est convergente.*

Démonstration : Compte tenu de ce qui précède, il reste à montrer que, si toute série normalement convergente est convergente, alors E est complet.
Soit donc (x_n) une suite de Cauchy ; on peut trouver une suite strictement croissante n_k d'entiers telle que $\|x_n - x_p\| \leqslant 2^{-k}$ si $n_k \leqslant n \leqslant p$. On voit alors que $y_k = x_{n_{k+1}} - x_{n_k}$ vérifie $\|y_k\| \leqslant 2^{-k}$, donc que la série (y_k) est normalement convergente. Si $y = \sum_{k=0}^{\infty} y_k$, on a : $x_{n_k} = x_{n_0} + \sum_{j=0}^{k-1} y_j \to u = x_{n_0} + y$. Puisque la suite (x_{n_k}) converge vers u, la suite de Cauchy (x_n) possède u comme valeur d'adhérence, donc converge vers u (cf. 1.8.4). ∎

Espaces normés de dimension finie.

Théorème 2.1.13. *Sur l'espace vectoriel $\mathbb{K}^n$, toutes les normes sont équivalentes.*

Démonstration : Soit $\|.\|$ la norme sur $\mathbb{K}^n$ définie par

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

pour $x = (x_1, x_2, \dots, x_n)$. Si on note S la sphère unité de cette norme, c'est-à-dire l'ensemble des points de $\mathbb{K}^n$ de norme 1, S est fermé puisque la distance est continue, et borné puisque chaque coordonnée est majorée par 1 en valeur absolue sur S. Donc S est compacte.
Si $|||.|||$ est une autre norme sur $\mathbb{K}^n$, et si on désigne par $(e_1, e_2, \dots, e_n)$ la base canonique de $\mathbb{K}^n$, on a pour tout x

$$|||x||| = \left|\left|\left| \sum_{i=1}^{n} x_i e_i \right|\right|\right| \leqslant \sum_{i=1}^{n} |x_i| \,.|||e_i||| \leqslant \max_{1 \leqslant i \leqslant n} |x_i| \,. \sum_{i=1}^{n} |||e_i|||$$

donc, en notant $M = \sum_{i=1}^{n} |||e_i|||$,

$$\big| |||x||| - |||y||| \big| \leqslant |||x - y||| \leqslant M \, \|x - y\| = M \, d(x, y) \ ,$$

ce qui montre que la fonction $|||.|||$ est M-lipschitzienne, donc continue sur S. Elle atteint donc sur S sa borne inférieure α, qui est strictement positive puisque $|||.|||$ ne s'annule pas sur S et sa borne supérieure β, qui est finie.
Pour $x \neq 0$ dans $\mathbb{K}^n$, on a $y = \dfrac{x}{\|x\|} \in S$, et, puisque $\dfrac{|||x|||}{\|x\|} = \left|\left|\left| \dfrac{x}{\|x\|} \right|\right|\right| = |||y|||$, on en déduit que

$$0 < \alpha \leqslant \frac{|||x|||}{\|x\|} \leqslant \beta < +\infty \ ,$$

ce qui est l'inégalité cherchée. ∎

Théorème 2.1.14. *Si E est un espace normé de dimension finie n et si $(a_1, a_2, \ldots, a_n)$ est une base de E, l'application linéaire T de $\mathbb{K}^n$ sur E définie par $T(\lambda) = \sum_{i=1}^n \lambda_i a_i$ pour $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ est un homéomorphisme de $\mathbb{K}^n$ sur E.*

Démonstration : Notons $|||.|||$ la norme sur $\mathbb{K}^n$ définie par $|||\lambda||| = \|T(\lambda)\|$. D'après le théorème précédent, la norme $|||.|||$ est équivalente à la norme $\|.\|$. Puisque T est une isométrie de $(\mathbb{K}^n, |||.|||)$ sur E, c'est un homéomorphisme, et puisque les normes $\|.\|$ et $|||.|||$ sur $\mathbb{K}^n$ sont équivalentes, elles définissent sur $\mathbb{K}^n$ la même topologie. Donc T est un homéomorphisme de $\mathbb{K}^n$ sur E. ■

Corollaire 2.1.15. *Toutes les normes sont équivalentes sur un espace de dimension finie.*

Théorème 2.1.16. *Soit E un espace normé. Alors, pour tout entier $n \geqslant 1$, l'ensemble L_n des éléments $(u_1, u_2, \ldots, u_n)$ de E^n tels que u_1, u_2,..., u_n soient linéairement indépendants est ouvert dans E^n.*

Démonstration : Soit $u = (u_1, u_2, \ldots, u_n)$ un point de L_n ; on veut montrer que L_n est un voisinage de u. Si on note T l'application linéaire de $\mathbb{K}^n$ dans E définie par $T(\lambda) = \sum_{i=1}^n \lambda_i u_i$, T est injective et la fonction $|||.|||$ définie par $|||\lambda||| = \|T(\lambda)\|$ est une norme sur $\mathbb{K}^n$ équivalente à la norme $\lambda \mapsto \|\lambda\| = \sup_i |\lambda_i|$. Il existe donc $\delta > 0$ tel que, pour tout λ, on ait $|||\lambda||| \geqslant \delta \|\lambda\|$. Alors, pour tout $(v_1, v_2, \ldots, v_n) \in E^n$ tel que $\sup_i \|v_i - u_i\| < \dfrac{\delta}{2n}$, et pour tout $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbb{K}^n$, on a :

$$\begin{aligned}\left\|\sum_{i=1}^n \lambda_i v_i\right\| = \left\|\sum_{i=1}^n \lambda_i u_i + \sum_{i=1}^n \lambda_i (v_i - u_i)\right\| &\geqslant \left\|\sum_{i=1}^n \lambda_i u_i\right\| - \left\|\sum_{i=1}^n \lambda_i (v_i - u_i)\right\| \\ &\geqslant \|T(\lambda)\| - \sum_{i=1}^n |\lambda_i| \, \|v_i - u_i\| \geqslant |||\lambda||| - n \sup_i |\lambda_i| \, . \sup_i \|v_i - u_i\| \\ &\geqslant \delta \|\lambda\| - n \|\lambda\| . \frac{\delta}{2n} = \frac{\delta}{2} \|\lambda\| \ .\end{aligned}$$

Il en résulte que la combinaison linéaire $\sum_{i=1}^n \lambda_i v_i$ne peut s'annuler que si $\|\lambda\| = 0$, c'est-à-dire si tous les λ_i sont nuls ; et ceci signifie que les v_i sont linéairement indépendants, donc que $v = (v_1, v_2, \ldots, v_n) \in L_n$. Et puisque L_n contient la boule de centre u et de rayon $\dfrac{\delta}{2n}$ dans E^n, L_n est un voisinage de u. ■

Théorème 2.1.17. *Dans un espace normé, tout sous-espace de dimension finie est fermé.*

Démonstration : Soit F un sous-espace de dimension finie n d'un espace normé E. Si $(a_1, a_2, \ldots, a_n)$ est une base de F, l'application linéaire T de $\mathbb{K}^n$ sur F définie par $T(\lambda) = \sum_{i=1}^n \lambda_i a_i$ est un homéomorphisme. De plus, si (x_k) est une suite de Cauchy dans F, on a

$$\left\|T^{-1} x_k - T^{-1} x_\ell\right\| = \left\|T^{-1}(x_k - x_\ell)\right\| \leqslant M \|x_k - x_\ell\|$$

pour une certaine constante M en vertu de l'équivalence des normes. Il en résulte que la suite $(\lambda^k) = (T^{-1}(x_k))$ est une suite de Cauchy dans $\mathbb{K}^n$, donc qu'elle converge vers un $\lambda \in \mathbb{K}^n$. Alors la suite (x_k) converge vers $T(\lambda) \in F$. Il en résulte que F est complet dans E, donc fermé. ■

Théorème 2.1.18. *Tout espace normé de dimension finie est localement compact.*

Démonstration : Puisque toute boule fermée de $\mathbb{K}^n$ est compacte, comme partie fermée bornée de $\mathbb{R}^n$ (si $\mathbb{K} = \mathbb{R}$) ou de $\mathbb{R}^{2n}$ (si $\mathbb{K} = \mathbb{C}$), il est clair que $\mathbb{K}^n$ est localement compact. Et puisque tout espace normé de dimension finie est homéomorphe à $\mathbb{K}^n$, il est également localement compact.

Lemme 2.1.19. *Soient E un espace normé et F un sous-espace vectoriel fermé de E. Si $F \neq E$, il existe pour tout $\varepsilon > 0$ un point x de E de norme 1 dont la distance à F est supérieure à $1 - \varepsilon$.*

Démonstration : Puisque $F \neq E$, il existe dans E un $a \notin F$. La distance δ de a à F est alors strictement positive. Il existe donc un $y \in F$ tel que $\|a - y\| = d(a, y) < \delta + \delta\varepsilon$. Puisque $y \neq a$, on peut poser $x = \dfrac{a-y}{\|a-y\|}\cdot$ On a alors clairement $\|x\| = 1$. De plus, pour tout $z \in F$, le point $u := y + \|a - y\|\, z$ appartient à F et on a

$$x - z = \frac{1}{\|a-y\|}\Big(a - (y + \|a-y\|\, z)\Big) = \frac{a-u}{\|a-y\|}\ ,$$

donc

$$\|x - z\| = \frac{\|a-u\|}{\|a-y\|} \geqslant \frac{\delta}{\delta(1+\varepsilon)} > 1 - \varepsilon\,,$$

et, puisque z est quelconque dans F, on obtient $d(x, F) \geqslant 1 - \varepsilon$. ■

Lorsque V est de dimension finie, on a un résultat plus précis.

Proposition 2.1.20. *Si E est un espace vectoriel normé, V un sous-espace vectoriel de E de dimension finie et x un point de E, alors il existe un point v de V tel que $\|x - v\| = d(x, V)$.*

Démonstration : L'ensemble $K := \{w \in V : \|w\| \leqslant 2\,\|x\|\,\}$ est une partie fermée bornée de l'espace de dimension finie V, donc une partie compacte. Alors la fonction continue $w \mapsto \|w - x\|$ atteint sur K sa borne inférieure δ en un point v. Et, puisque $0 \in K$, on a $\delta = \inf_{w\in K} \|w - x\| \leqslant \|x\|$. Pour $w \in V\setminus K$, on a $\|w - x\| \geqslant \|w\| - \|x\| > 2\,\|x\| - \|x\| = \|x\|$. Il en résulte que $\|x - v\| = \delta = \inf_{v\in V} \|w - x\| = d(x, V)$. ■

Théorème 2.1.21. (Riesz) *Tout espace normé localement compact est de dimension finie.*

Démonstration : Puisque 0 possède dans E un voisinage compact, il existe une boule B de E, centrée en 0, et compacte. Puisque les homothéties, $x \mapsto \lambda x$, sont des homéomorphismes de E pour $\lambda \neq 0$, on peut supposer que B est la boule unité de E.
Le compact B est recouvert par la famille des boules ouvertes $\big(B(x,1/2)\big)_{x\in B}$, et il existe donc des points $(x_1, x_2, \ldots, x_n)$ de B en nombre fini tels que $B \subset \bigcup_{i=1}^n B(x_i, 1/2)$. Le sous-espace vectoriel F de E engendré par les x_i est de dimension finie (au plus n), donc fermé dans E d'après le théorème 2.1.17. Et si F n'était pas égal à E, il existerait par le lemme 2.1.19 un x de B dont la distance à F serait supérieure à $2/3$. Mais alors, il existerait un $i \leqslant n$ tel que $x \in B(x_i, 1/2)$, et on aurait

$$2/3 \leqslant d(x, F) \leqslant d(x, x_i) < 1/2\,,$$

et cette contradiction montre que $F = E$, donc que E est de dimension finie. ■

Exemples d'espaces normés.

Si K est un espace compact, on peut définir, pour tout élément f de l'espace vectoriel $\mathscr{C}(K,\mathbb{K})$ des fonctions continues de K dans $\mathbb{K}$

$$\|f\| = \sup_{x\in K} |f(x)| \in \mathbb{R}^+ \ .$$

On vérifie sans peine que cette fonction est une norme, et que la distance associée est la distance de la convergence uniforme. On appelle cette norme la *norme de la convergence uniforme*. Il résulte du théorème 1.8.13 que $\mathscr{C}(K,\mathbb{K})$ est un espace de Banach sur $\mathbb{K}$.

On note c_0 l'ensemble de toutes les suites de nombres complexes qui tendent vers 0. Il est clair que, pour l'addition terme-à-terme et la multiplication par un nombre complexe, c_0 est un espace vectoriel. On pose, pour $x = (x_n)$ dans c_0

$$\|x\| = \sup_{n\in\mathbb{N}} |x_n| \ ,$$

et on vérifie sans peine que ceci définit une norme sur c_0.

Théorème 2.1.22. *L'espace c_0 est un espace de Banach.*

Démonstration : Soit (x^k) une suite de Cauchy dans c_0. On a, pour tout k et tout ℓ, $\left|x_n^k - x_n^\ell\right| \leqslant \left\|x^k - x^\ell\right\|$. Il en résulte que la suite numérique $(x_n^k)_{k\in\mathbb{N}}$ est, pour tout n, une suite de Cauchy dans $\mathbb{C}$, donc converge vers un nombre complexe x_n. Il faut démontrer que $x = (x_n)$ est dans c_0 et est la limite dans c_0 de la suite (x^k).
Soit donc $\varepsilon > 0$. Il existe un m tel que si k et ℓ sont supérieurs à m on ait $\|x_k - x_\ell\| \leqslant \varepsilon/2$. Puisque x_m appartient à c_0, il existe un p tel que $|x_n^m| \leqslant \varepsilon/2$ pour $n > p$. Alors, pour $n > p$ et $k \geqslant m$ on a

$$\left|x_n^k\right| \leqslant |x_n^m| + \left\|x^k - x^m\right\| \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon \ ,$$

et en passant à la limite, $|x_n| \leqslant \varepsilon$ pour $n > p$. Ceci signifie que $x \in c_0$.
De plus, puisque $\left|x_n^k - x_n^\ell\right| \leqslant \varepsilon/2$ pour k et ℓ supérieurs à m, on a, en passant à la limite $\left|x_n - x_n^\ell\right| \leqslant \varepsilon/2$ pour $\ell > m$, c'est-à-dire $\left\|x - x^\ell\right\| \leqslant \varepsilon/2$ pour $\ell > m$. Donc x est limite dans c_0 de la suite (x^k). ■

On note ℓ^1 l'ensemble de toutes les suites de nombres complexes $x = (x_n)$ telles que la série (x_n) converge absolument. Il est clair que ℓ^1 est un espace vectoriel sur lequel la fonction $x \mapsto \sum_{n=0}^{\infty} |x_n|$ est une norme.

Théorème 2.1.23. *L'espace ℓ^1 est un espace de Banach.*

Démonstration : Soit (x^k) une suite de Cauchy dans ℓ^1. Pour tout n, tout k et tout ℓ, on a $\left|x_n^k - x_n^\ell\right| \leqslant \left\|x^k - x^\ell\right\|$. Il en résulte que la suite numérique $(x_n^k)_{k\in\mathbb{N}}$ est une suite de Cauchy dans $\mathbb{C}$, donc converge vers un nombre x_n. Il reste à démontrer que $x = (x_n)$ appartient à ℓ^1 et que x est limite dans ℓ^1 de la suite (x^k). Si $\varepsilon > 0$ est donné, il existe m tel que si k et ℓ sont supérieurs à m, $\left\|x^k - x^\ell\right\| \leqslant \varepsilon$. On a alors, pour n fixé

$$\sum_{j=0}^{n} \left|x_j^k\right| \leqslant \sum_{j=0}^{n} \left|x_j^m\right| + \sum_{j=0}^{n} \left|x_j^k - x_j^m\right| \leqslant \sum_{j=0}^{\infty} \left|x_j^m\right| + \varepsilon \ .$$

Et en passant à la limite quand k tend vers l'infini, on obtient

$$\sum_{j=0}^{n} |x_j| \leqslant \|x_m\| + \varepsilon \ .$$

Et comme cette dernière quantité ne dépend pas de n, on en conclut que $x \in \ell^1$.
De plus, pour n fixé, $k \geqslant m$ et $\ell \geqslant m$, on a

$$\sum_{j=0}^{n} \left|x_j^k - x_j^\ell\right| \leqslant \left\|x^k - x^\ell\right\| \leqslant \varepsilon \ ,$$

donc en passant à la limite quand $k \to \infty$,

$$\sum_{j=0}^{n} \left|x_j - x_j^\ell\right| \leqslant \varepsilon \ .$$

Et puisque cette inégalité est valable pour tout n, on obtient $\left\|x - x^\ell\right\| \leqslant \varepsilon$ pour $\ell \geqslant m$. Donc la suite (x^k) converge vers x dans ℓ^1. ■

2.2 Applications linéaires continues

Théorème 2.2.1. *Soit f une application linéaire de l'espace normé E dans l'espace normé F. Les propriétés suivantes sont équivalentes :*

- *i)* *f est continue.*
- *ii)* *f est continue en 0.*
- *iii)* *f est uniformément continue.*
- *iv)* *f est lipschitzienne.*
- *v)* *Il existe une constante M telle que pour tout x de E, on ait $\|f(x)\| \leqslant M \ \|x\|$.*

Démonstration : Il est clair que *iv)* $\Longrightarrow$ *iii)* $\Longrightarrow$ *i)* $\Longrightarrow$ *ii)*. Si la condition *v)* est satisfaite, on a pour tout x et tout y de E

$$d(f(x), f(y)) = \|f(x) - f(y)\| = \|f(x-y)\| \leqslant M \ \|x-y\| = M \, d(x,y) \ ,$$

ce qui prouve que f est M-lipschitzienne, donc que *v)* $\Longrightarrow$ *iv)*.
Enfin, si f est continue en 0, il existe un $\delta > 0$ tel que $\|x\| \leqslant \delta \Longrightarrow \|f(x)\| \leqslant 1$, puisque la boule unité de F est un voisinage de 0. Soit $x \in E$. Si $x = 0$ on a $f(x) = 0$, donc $\|f(x)\| \leqslant \delta^{-1} \ \|x\|$. Et si $x \neq 0$, on pose $\lambda = \dfrac{\delta}{\|x\|}$. Alors $y = \lambda x$ vérifie $\|y\| = \lambda \, \|x\| = \delta$, d'où

$$\|f(y)\| = \|f(\lambda x)\| = \|\lambda f(x)\| = \lambda \, \|f(x)\| \leqslant 1 \ ,$$

et on en conclut que $\|f(x)\| \leqslant \dfrac{1}{\lambda} = \delta^{-1} \ \|x\|$. Et ceci achève de prouver *ii)* $\Longrightarrow$ *v)* avec $M = \delta^{-1}$. ■

Définition 2.2.2. *Si f est une application linéaire continue de l'espace normé E dans l'espace normé F, on appelle* norme de f *la plus petite constante M telle que, pour tout $x \in E$ on ait $\|f(x)\| \leqslant M \|x\|$. On la note $\|f\|$.*

On a donc $\|f(x)\| \leqslant \|f\| . \|x\|$ pour tout x de E et toute application linéaire continue f de E dans F.

Proposition 2.2.3. *La norme d'une application linéaire continue f est égale à*

$$\|f\| = \sup_{x \in B} \|f(x)\| \ ,$$

où B désigne la boule unité de E.

Démonstration : Puisqu'on a pour tout x, $\|f(x)\| \leqslant \|f\| . \|x\|$, on a nécessairement $\sup_{x \in B} \|f(x)\| \leqslant \|f\|$. Et si $0 < \rho < \|f\|$ il existe un $x \neq 0$ dans E tel que $\|f(x)\| \geqslant \rho \|x\|$. Alors le point $y = \dfrac{x}{\|x\|}$ appartient à B, et on a $\|f(y)\| > \rho$. ■

On notera $\mathscr{L}(E, F)$ l'espace vectoriel des applications linéaires continues de l'espace normé E dans l'espace normé F.

Théorème 2.2.4. *Si f est une application linéaire de l'espace normé E de dimension finie dans l'espace normé F, f est continue.*

Démonstration : Posons, pour $x \in E$

$$|||x||| = \|x\| + \|f(x)\| \ .$$

Il est clair que $|||.|||$ est une norme sur E. Cette norme est équivalente à la norme initiale $\|.\|$. Il existe donc un α tel que $|||x||| \leqslant \alpha \|x\|$ pour tout x de E. On en déduit que $\|f(x)\| \leqslant \alpha \ \|x\|$, ce qui prouve la continuité de f. ■

Théorème 2.2.5. *Muni de l'application $f \mapsto \|f\|$, l'espace vectoriel $\mathscr{L}(E, F)$ est un espace normé. De plus, si F est complet, $\mathscr{L}(E, F)$ est un espace de Banach.*

Démonstration : Notons B la boule unité de E. Il est clair que l'application nulle a 0 comme norme. Et si $\|f\| = 0$, on a $\|f(x)\| \leqslant \|f\| . \|x\| = 0$, donc f est l'application nulle. Si $\lambda \in \mathbb{K}$, on a, pour tout $x \in B$

$$\|(\lambda f)(x)\| = |\lambda| \, \|f(x)\| \ ,$$

d'où, en passant à la borne supérieure, $\|\lambda f\| = |\lambda| \, \|f\|$.
Enfin, si f et g sont deux applications linéaires continues de E dans F, on a pour tout $x \in B$

$$\|(f + g)(x)\| = \|f(x) + g(x)\| \leqslant \|f(x)\| + \|g(x)\| \leqslant \|f\| + \|g\| \ ,$$

donc $\|f + g\| = \sup_{x \in B} \|(f + g)(x)\| \leqslant \|f\| + \|g\|$.
Si F est complet et si (f_n) est une suite de Cauchy dans $\mathscr{L}(E, F)$, on a pour tout $x \in E$, tout n et tout p :

$$\|f_n(x) - f_p(x)\| = \|(f_n - f_p)(x)\| \leqslant \|f_n - f_p\| . \|x\| \ ,$$

d'où l'on déduit que la suite $(f_n(x))$ est une suite de Cauchy dans F. Par conséquent, il existe une application f de E dans F telle que, pour tout x de E, $f(x) = \lim_{n\to\infty} f_n(x)$. On a clairement

$$f(x+\lambda y) = \lim_{n\to\infty}(f_n(x) + \lambda f_n(y)) = \lim_{n\to\infty} f_n(x) + \lambda \lim_{n\to\infty} f_n(y) = f(x) + \lambda f(y) \ ,$$

ce qui montre que f est linéaire.
Puisque la suite (f_n) est une suite de Cauchy, il existe un m tel que, pour $n \geqslant m$ on ait $\|f_n - f_m\| \leqslant 1$. On en déduit que, pour $n \geqslant m$, $\|f_n\| \leqslant 1 + \|f_m\|$, donc que pour $x \in E$ et $n \geqslant m$, $\|f_n(x)\| \leqslant (1 + \|f_m\|)\,\|x\|$ et on en déduit que $f(x)$ appartient à la boule fermée de centre 0 et de rayon $(1 + \|f_m\|)\,\|x\|$. Ceci montre que pour tout x

$$\|f(x)\| \leqslant (1 + \|f_m\|)\,\|x\| \ ,$$

c'est-à-dire que l'application linéaire f est continue. Enfin, si $\varepsilon > 0$, il existe un m tel que, pour n et p supérieurs à m, on ait $\|f_n - f_p\| \leqslant \varepsilon$. On en déduit que si $x \in E$ et si n et p sont supérieurs à m, on a

$$\|f_n(x) - f_p(x)\| \leqslant \varepsilon\,\|x\| \ ,$$

donc que $f_p(x)$ appartient à la boule fermée de centre $f_n(x)$ et de rayon $\varepsilon\,\|x\|$. La limite $f(x)$ de la suite $(f_p(x))$ appartient donc à cette même boule et on en déduit que

$$\|f_n(x) - f(x)\| \leqslant \varepsilon\,\|x\| \ ,$$

donc que $\|f - f_n\| \leqslant \varepsilon$ pour $n > m$, c'est-à-dire que f est limite de la suite (f_n). La suite de Cauchy (f_n) est donc convergente dans $\mathscr{L}(E,F)$. ■

Définition 2.2.6. *Une partie X de l'espace normé E est dite* totale *si le sous-espace vectoriel qu'elle engendre est partout dense dans E, c'est-à-dire si tout point de E est limite de combinaisons linéaires d'éléments de X.*

Proposition 2.2.7. *Si f et g sont deux applications linéaires continues de E dans F qui coïncident sur une partie totale X de E, alors f et g sont égales.*

Démonstration : Le noyau de l'application linéaire continue $f - g$ est un sous-espace vectoriel de E qui contient X, donc aussi l'espace vectoriel engendré par X, qui est dense. Et puisque $\ker(f-g)$ est dense et fermé, il est égal à E, ce qui signifie que $f - g = 0$. ■

Définition 2.2.8. *Si E est un espace normé, on appelle* espace dual *de E l'espace $E' = \mathscr{L}(E,\mathbb{K})$ des formes linéaires continues sur E.*

Puisque $\mathbb{K}$ est complet, il résulte du théorème précédent que E' est un espace de Banach.

Théorème 2.2.9. *Si E, F et G sont trois espaces normés, f une application linéaire continue de E dans F et g une application linéaire continue de F dans G, l'application linéaire continue $g\circ f$ a une norme au plus égale à $\|f\|\,.\,\|g\|$.*

Démonstration : Pour $x \in E$ et $y \in F$ on a $\|f(x)\| \leqslant \|f\|\,.\,\|x\|$ et $\|g(y)\| \leqslant \|g\|\,.\,\|y\|$. Donc

$$\|g\circ f(x)\| \leqslant \|g\|\,.\,\|f(x)\| \leqslant \|g\|\,.\,\|f\|\,.\,\|x\|$$

ce qui montre que $\|g\circ f\| \leqslant \|g\|\,.\,\|f\|$. ■

Théorème 2.2.10. *Si E est un espace de Banach, l'espace $\mathscr{L}(E) = \mathscr{L}(E,E)$ est une algèbre de Banach, c'est-à-dire un espace de Banach muni d'une multiplication associative et distributive par rapport à la somme et vérifiant $\|fg\| \leqslant \|f\| \,.\, \|g\|$ pour tout f et tout g.*

Ceci résulte immédiatement de ce qui précède. On notera I l'application identique de E, qui est l'unité de l'algèbre $\mathscr{L}(E)$.

Lemme 2.2.11. (Neumann) *Si E est un espace de Banach, et si $f \in \mathscr{L}(E)$ vérifie $\|f\| < 1$, l'application $I - f$ est inversible, c'est-à-dire qu'il existe un élément g de $\mathscr{L}(E)$ tel que $(I-f)\circ g = g\circ(I-f) = I$. De plus $\left\|(I-f)^{-1} - I\right\| \leqslant \dfrac{\|f\|}{1-\|f\|}$.*

Démonstration : Notons S_n l'élément de $\mathscr{L}(E)$ défini comme la somme $\sum_{j=0}^{n} f^j$, où $f^0 = I$ et f^j désigne, pour $j \geqslant 1$, le produit de composition de j applications égales à f (on a donc $f^j\circ f^k = f^{j+k}$ pour j et k entiers). On a, pour $n < p$:

$$\|S_p - S_n\| = \left\|\sum_{j=n+1}^{p} f^j\right\| \leqslant \sum_{j=n+1}^{p} \left\|f^j\right\| \leqslant \sum_{j=n+1}^{p} \|f\|^j \leqslant \frac{\|f\|^{n+1}}{1-\|f\|} \ ,$$

car $\|f\| < 1$. On en déduit que si n et p sont supérieurs à m, $\|S_n - S_p\| \leqslant \dfrac{\|f\|^{m+1}}{1-\|f\|}$, donc que la suite (S_n) est une suite de Cauchy dans $\mathscr{L}(E)$, puisque $\lim_{m\to\infty} \|f\|^m = 0$, et que (S_n) converge vers un élément g. De plus on a

$$(I-f)\circ S_n = S_n\circ(I-f) = \sum_{j=0}^{n}(f^j - f^{j+1}) = I - f^{n+1} \ ,$$

et puisque $\left\|f^{n+1}\right\| \leqslant \|f\|^{n+1} \to 0$, on voit que $(I - f^{n+1})$ converge vers I. On conclut que

$$I = \lim_{n\to\infty}(I-f)\circ S_n = (I-f)\circ g = \lim_{n\to\infty} S_n\circ(I-f) = g\circ(I-f) \ ,$$

et que

$$\left\|(I-f)^{-1} - I\right\| = \left\|\sum_{n=1}^{\infty} f^n\right\| \leqslant \sum_{n=1}^{\infty} \|f\|^n = \frac{\|f\|}{1-\|f\|} \ .$$

Ceci achève la démonstration. ■

Théorème 2.2.12. *Si E est un espace de Banach, l'ensemble $\mathscr{G}$ des éléments inversibles de l'algèbre $\mathscr{L}(E)$ est ouvert, et l'application $g \mapsto g^{-1}$ est continue sur $\mathscr{G}$.*

Démonstration : Soit $g \in \mathscr{G}$. Pour $h \in \mathscr{L}(E)$, on a $g + h = g\circ(I + g^{-1}\circ h)$. Donc si $\|h\| < \rho = \dfrac{1}{\|g\|^{-1}}$, on peut poser $f = -g^{-1}\circ h$. On a alors $\|f\| \leqslant \left\|g^{-1}\right\| \,.\, \|h\| < 1$. On en déduit par le théorème précédent que $I - f$ est inversible, d'inverse u. Alors, puisque $g + h = g\circ(I-f)$, on vérifie que $u\circ g^{-1}$ est l'inverse de $g + h$.
Il en résulte que la boule ouverte de centre g et de rayon ρ est contenue dans $\mathscr{G}$, donc que $\mathscr{G}$ est ouvert dans $\mathscr{L}(E)$. De plus

$$\left\|(g+h)^{-1} - g^{-1}\right\| = \left\|(u-I)\circ g^{-1}\right\| \leqslant \frac{\left\|g^{-1}\right\| \|h\|}{1 - \left\|g^{-1}\right\| \|h\|} \left\|g^{-1}\right\| \ ,$$

qui tend vers 0 avec $\|h\|$.

■

Il est clair qu'une application linéaire bijective f entre deux espaces vectoriels normés E et F possède une application réciproque linéaire. Par contre, même si de plus f est supposée continue, l'application f^{-1} 'est pas nécessairement continue. On dira que l'application linéaire f est un *isomorphisme* de E sur F si elle est bijective, continue et si f^{-1} est continue. Signalons néanmoins le théorème suivant, dû à S. Banach, dont on trouvera la démonstration dans un cadre plus général en 4.7.1.

Théorème 2.2.13. *Soient E et F deux espaces de Banach. Si f est une application linéaire continue et bijective de E sur F, alors f est un isomorphisme.*

2.3 Spectre d'un opérateur

Dans toute cette section, E désignera un $\mathbb{C}$-espace de Banach non réduit à $\{0\}$, et T un opérateur sur E, c'est-à-dire une application linéaire continue de E dans lui-même.

Théorème 2.3.1. *L'ensemble X des nombres complexes λ tels que $T - \lambda I$ ne soit pas inversible dans $\mathscr{L}(E)$ est une partie compacte de $\mathbb{C}$ appelée* spectre *de T. L'application $\lambda \mapsto (T - \lambda I)^{-1}$ définie sur $\mathbb{C} \setminus X$, appelée* résolvante *de T, est continue de $\mathbb{C} \setminus X$ dans $\mathscr{L}(E)$.*

Démonstration : L'application $f : \lambda \mapsto T - \lambda I$ est continue de $\mathbb{C}$ dans $\mathscr{L}(E)$. Il résulte immédiatement du théorème 2.2.12 que $f^{-1}(\mathscr{G}(E))$ est un ouvert du plan complexe, donc que X est fermé, et que la résolvante est continue. Pour montrer la compacité du spectre, il reste uniquement à montrer que X est borné. Mais puisque, pour $|\lambda| > \|T\|$, on a $\left\|\lambda^{-1}T\right\| < 1$ et

$$T - \lambda I = -\lambda(I - \lambda^{-1}T) \ ,$$

il résulte du Lemme 2.2.11 que l'on alors $I - \lambda^{-1}T \in \mathscr{G}(E)$, donc aussi $T - \lambda I \in \mathscr{G}(E)$, c'est-à-dire $\lambda \notin X$. On conclut que X est contenu dans le disque fermé de centre 0 et de rayon $\|T\|$. ■

Théorème 2.3.2. *Le spectre de T n'est pas vide.*

Démonstration : La résolvante $R(\lambda) = (T - \lambda I)^{-1}$ vérifie, pour λ et μ hors du spectre X de T :

$$R(\lambda) - R(\mu) = (T - \lambda I)^{-1}\Big((T - \mu I) - (T - \lambda I)\Big)(T - \mu I)^{-1} = (\lambda - \mu)R(\lambda)R(\mu)$$

Pour $\mu \notin X$, la continuité en μ de la résolvante assure que $\lim_{\lambda\to\mu} \dfrac{R(\lambda) - R(\mu)}{\lambda - \mu} = R(\mu)^2$.
La fonction R est donc holomorphe de $\mathbb{C} \setminus X$ dans $\mathscr{L}(E)$. De plus, lorsque $|\lambda| > \|T\|$, on a $R(\lambda) = -\lambda^{-1}(I - \lambda^{-1}T)^{-1}$ ce qui montre que $R(\lambda) \to 0$ lorsque $\lambda \to \infty$.
Si le spectre X était vide, cette fonction serait entière et bornée sur $\mathbb{C}$. Et il résulterait du théorème de Liouville que R serait constante, ce qui est impossible puisque $T - \lambda I = R(\lambda)^{-1}$ n'est pas constante. ■

2.4 Le théorème de Banach-Steinhaus

Lemme 2.4.1. *Soient E un espace de Banach et C une partie convexe symétrique fermée et* absorbante *de E (c'est-à-dire que $\bigcup_{n \in \mathbb{N}} nC = E$). Alors C est un voisinage de 0.*

Démonstration : L'application $h_n : x \mapsto \dfrac{1}{n} x$ est continue, et l'ensemble $nC = h^{-1}(C)$ est donc fermé pour tout n. Et puisque l'espace métrique complet E est réunion de la famille dénombrable de fermés $(nC)_{n \in \mathbb{N}}$, il résulte du théorème de Baire (théorème 1.9.2) que l'un de ces fermés est d'intérieur non vide.
Soient donc $n_0 \in \mathbb{N}$, $a \in E$ et $r > 0$ tels que $\tilde{B}(a, r) \subset n_0 C$. Alors, par symétrie, on a aussi $\tilde{B}(-a, r) \subset n_0 C$. Et pour tout $x \in E$ tel que $\|x\| \leqslant r$, on a $a + x \in \tilde{B}(a, r) \subset C$ et $-a + r \in \tilde{B}(-a, r) \subset C$. On en déduit que

$$x = \frac{1}{2}((a + x) + (-a + x)) \in \frac{1}{2}(C + C) \subset C \ ,$$

c'est-à-dire que $\tilde{B}(a, r) \subset C$. Et ceci montre que C est un voisinage de 0 dans E. ■

Théorème 2.4.2. *Soient E et F deux espaces de Banach et H une partie de $\mathscr{L}(E, F)$. Si pour tout $x \in E$, l'ensemble $H(x) = \{Tx : T \in H\}$ est borné dans F, alors H est borné dans $\mathscr{L}(E, F)$.*

Démonstration : Notons B_F la boule unité fermée de F et considérons l'ensemble

$$C = \{x \in E : \sup_{T \in H} \|Tx\| \leqslant 1\} = \bigcap_{T \in H} T^{-1}(B_F) \ .$$

Puisque chaque $T \in H$ est linéaire et continu, l'ensemble $T^{-1}(B_F)$ est convexe fermé et symétrique. On en déduit que l'intersection C de ces ensembles est aussi convexe fermée et symétrique.
De plus, si x est un point de E, il résulte de l'hypothèse que $H(x)$ est borné, qu'il existe un entier n tel que $n \geqslant \sup_{T \in H} \|Tx\|$. Alors $\dfrac{x}{n} \in C$, c'est-à-dire $x \in nC$.
On déduit alors du lemme 2.4.1 que C est un voisinage de 0, donc qu'il existe $r > 0$ tel que $\tilde{B}(0, r) \subset C$. Pour tout $x \neq 0$ dans E, on a $y := r\dfrac{x}{\|x\|} \in \tilde{B}(0, r) \subset C$, donc $\|Ty\| = \dfrac{r}{\|x\|} \|Tx\| \leqslant 1$, ou encore $\|Tx\| \leqslant \dfrac{1}{r} \|x\|$ pour tout $T \in H$. Cette inégalité est encore clairement vérifiée pour $x = 0$. Et ceci signifie que $\sup_{T \in H} \|T\| \leqslant \dfrac{1}{r} < +\infty$. ■

Corollaire 2.4.3. *Soient E et F des espaces de Banach, et (T_n) une suite dans $\mathscr{L}(E, F)$. Si, pour tout x de E, la suite $(T_n x)$ converge dans F, alors la suite (T_n) est bornée dans $\mathscr{L}(E, F)$ et l'application $T : x \mapsto \lim_n T_n x$ est linéaire continue de E dans F.*

Démonstration : Pour tout x de E la suite convergente $(T_n x)$ est bornée dans F. Il résulte alors du théorème précédent, appliqué à $H = \{T_n : n \in \mathbb{N}\}$, que la suite (T_n) est bornée dans $\mathscr{L}(E, F)$. Et si $M = \sup_n \|T_n\|$, on a $\|Tx\| \leqslant M \|x\|$ pour tout x. Et puisque T est clairement linéaire, ceci montre que $T \in \mathscr{L}(E, F)$ et que $\|T\| \leqslant M$. ■

2.5 Opérateurs compacts

Définition 2.5.1. *Une application linéaire u d'un espace de Banach E dans un espace de Banach F est dite* compacte *si $u(B)$ est une partie relativement compacte de F dès que B est une partie bornée de E.*

Il revient au même de dire que l'image par u de la boule unité de E est relativement compacte dans F. Et puisque toute partie relativement compacte est bornée, une application linéaire compacte est nécessairement continue.

Théorème 2.5.2. *Si u est un isomorphisme compact de l'espace de Banach E, alors E est de dimension finie.*

DÉMONSTRATION : L'image par u de la boule unité de E est alors un voisinage de 0 relativement compact. Il en résulte que E est localement compact, donc de dimension finie (cf. 2.1.21). ∎

Proposition 2.5.3. *Si u est une application linéaire continue et de rang fini de l'espace de Banach E dans l'espace de Banach F, alors u est compacte.*

DÉMONSTRATION : Si B est la boule unité de E, et si V désigne l'espace image $u(E)$, qui est de dimension finie, alors $u(B)$ est une partie de V, bornée puisque u est continue, donc relativement compacte dans V, et a fortiori dans F. ∎

Lemme 2.5.4. *Si E, F, G et H sont des espaces de Banach, $u \in \mathscr{L}(E,F)$, $v \in \mathscr{L}(F,G)$ et $w \in \mathscr{L}(G,H)$ et si v est compacte, alors $w \circ v \circ u$ est compacte.*

DÉMONSTRATION : Si B est une partie bornée de E, alors $u(B)$ est une partie bornée de F. Il en résulte que $X = v(u(B))$ est relativement compacte dans G, c'est-à-dire que $\overline{X}$ est compacte. On en déduit que $w(\overline{X})$ est une partie compacte de H, qui contient $w \circ v \circ u(B)$ ∎

Théorème 2.5.5. *L'ensemble $\mathscr{K}(E,F)$ des opérateurs compacts de E dans F est un espace vectoriel.*

DÉMONSTRATION : Soient u et v deux opérateurs compacts de E dans F et $\lambda \in \mathbb{K}$. Si B est la boule unité de E, les ensembles $K = \overline{u(B)}$ et $L = \overline{v(B)}$ sont compacts dans F ; donc $K \times L$ est compact dans $F \times F$. Et puisque la fonction $s : (y,z) \mapsto y + \lambda z$ est continue de $F \times F$ dans F, l'ensemble $K + \lambda L = s(K \times L)$ est compact dans F. Alors on a $(u + \lambda v)(B) \subset K + \lambda L$, et ceci montre que $u + \lambda v$ est compact. ∎

Théorème 2.5.6. *L'ensemble $\mathscr{K}(E,F)$ des opérateurs compacts de E dans F est fermé dans $\mathscr{L}(E,F)$.*

DÉMONSTRATION : Désignons par B_E la boule unité de E, et soit u adhérent à $\mathscr{K}(E,F)$. Pour montrer que $u(B_E)$ est relativement compact dans F, il suffit de montrer que $u(B_E)$ est précompact dans l'espace complet F. Soit donc $\varepsilon > 0$. Il existe par hypothèse $v \in \mathscr{K}(E,F)$ tel que $\|u - v\| < \dfrac{\varepsilon}{3}$. Et puisque $\overline{v(B_E)}$ est compact et recouvert par les boules ouvertes $B(v(x), \dfrac{\varepsilon}{3})$ pour $x \in B_E$, il existe un nombre fini $(x_1, x_2, \ldots, x_m)$ de points de B_E tels que $\overline{v(B_E)} \subset \bigcup_{j=1}^{m} B(v(x_j), \dfrac{\varepsilon}{3})$. On va montrer que $u(B_E) \subset \bigcup_{j=1}^{m} B(u(x_j), \varepsilon)$. En effet, si

$y \in u(B_E)$, il existe $x \in B_E$ tel que $y = u(x)$, et un $j \leqslant m$ tel que $\|v(x) - v(x_j)\| < \dfrac{\varepsilon}{3}$. On a donc

$$\begin{aligned}\|u(x) - u(x_j)\| &\leqslant \|u(x) - v(x)\| + \|v(x) - v(x_j)\| + \|v(x_j) - u(x_j)\| \\ &\leqslant \|u - v\| \left(\|x\| + \|x_j\|\right) + \|u(x_j) - v(x_j)\| < \varepsilon \ ,\end{aligned}$$

et ceci montre la précompacité de $u(B_E)$ et achève la démonstration. ■

Corollaire 2.5.7. *Soient E et F deux espaces de Banach et (T_n) une suite d'opérateurs continus de rang fini de E dans F. Si la suite (T_n) converge dans $\mathscr{L}(E, F)$ vers un opérateur T, celui-ci est compact.*

Démonstration : Ceci résulte immédiatement de 2.5.6 et 2.5.3. ■

3

ESPACES DE HILBERT

Dans tout ce chapitre, les espaces vectoriels considérés seront encore des espaces vectoriels sur $\mathbb{R}$ ou sur $\mathbb{C}$. Mais dans le cas complexe, la conjugaison $\lambda \mapsto \bar{\lambda}$ jouera un rôle important. Les énoncés seront donc toujours donnés avec $\mathbb{C}$ comme corps de base, et on désignera toujours, pour un nombre λ, ses parties réelle et imaginaire par $\Re e\, \lambda$ et $\Im m\, \lambda$.
Dans le cas où le corps de base est $\mathbb{R}$, il suffira de convenir que pour $\lambda \in \mathbb{R}$, on a $\Im m\, \lambda = 0$ et que $\bar{\lambda} = \lambda = \Re e\, \lambda$.

3.1 Orthogonalité

Produit scalaire.

Définition 3.1.1. *Si E est un espace vectoriel sur $\mathbb{C}$ on appelle* produit scalaire hermitien *ou simplement* produit scalaire *sur E une application $(x, y) \mapsto \langle x, y\rangle$ de $E \times E$ dans $\mathbb{C}$ vérifiant les propriétés suivantes :*

i) $\forall y \in E$ *l'application $x \mapsto \langle x, y\rangle$ est linéaire.*

ii) $\forall x, y \in E \quad \langle y, x\rangle = \overline{\langle x, y\rangle}$.

iii) $\forall x \in E \quad \langle x, x\rangle \in \mathbb{R}^+$.

iv) $\forall x \in E \quad \langle x, x\rangle = 0 \Longrightarrow x = 0$.

On peut noter que les propriétés (i) et (ii) entraînent que pour x fixé, l'application $y \mapsto \langle x, y\rangle$ est semi-linéaire, c'est-à-dire vérifie

$$\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle \ ,$$
$$\langle x, \lambda y\rangle = \overline{\lambda}\, \langle x, y\rangle \ .$$

Théorème 3.1.2. (Cauchy-Schwarz) *Si $\langle \cdot, \cdot\rangle$ est un produit scalaire sur E, on a pour tout x et tout y de E l'inégalité*

$$|\langle x, y\rangle|^2 \leqslant \langle x, x\rangle \cdot \langle y, y\rangle \ .$$

De plus, si on a l'égalité $|\langle x,y\rangle|^2 = \langle x,x\rangle \cdot \langle y,y\rangle$*, les vecteurs* x *et* y *sont proportionnels.*

DÉMONSTRATION : Pour $\lambda \in \mathbb{C}$, x et y dans E on a

$$\begin{aligned}\langle x-\lambda y, x-\lambda y\rangle &= \langle x, x-\lambda y\rangle - \lambda \langle y, x-\lambda y\rangle \\ &= \langle x,x\rangle - \overline{\lambda}\,\langle x,y\rangle - \lambda\,\overline{\langle x,y\rangle} + \lambda\overline{\lambda}\,\langle y,y\rangle \\ &= \langle x,x\rangle - 2\,\Re e\big(\overline{\lambda}\,\langle x,y\rangle\big) + |\lambda|^2\,\langle y,y\rangle\,.\end{aligned}$$

Il existe $\rho \in \mathbb{R}^+$ et $\theta \in \mathbb{R}$ tels que $\langle x,y\rangle = \rho\,\mathrm{e}^{i\theta}$. Et si, pour $t \in \mathbb{R}$, on choisit $\lambda = t\,\mathrm{e}^{i\theta}$, on obtient que, pour tout $t \in \mathbb{R}$,

$$P(t) = \langle x,x\rangle - 2t\rho + t^2\langle y,y\rangle \geqslant 0\,.$$

Le polynôme du second degré P ci-dessus a donc un discriminant négatif ou nul. On en déduit

$$\Delta' = \rho^2 - \langle x,x\rangle \cdot \langle y,y\rangle \leqslant 0\ ,$$

c'est-à-dire

$$|\langle x,y\rangle|^2 \leqslant \langle x,x\rangle \cdot \langle y,y\rangle\,.$$

Si on a égalité dans la formule précédente, le discriminant Δ' est nul. Ou bien $y = 0$, auquel cas $y = 0 \cdot x$, ou bien P a une racine double $t_0 = \dfrac{\langle x,y\rangle\,\mathrm{e}^{-i\theta}}{\langle y,y\rangle}$, ce qui signifie que $P(t_0) = \left\|x - t_0\,\mathrm{e}^{i\theta}\,y\right\|^2 = 0$, donc que $x = t_0\,\mathrm{e}^{i\theta}\,y$. ■

Théorème 3.1.3. *Si* $\langle \cdot,\cdot\rangle$ *est un produit scalaire sur* E*, l'application* $x \mapsto \|x\| = \sqrt{\langle x,x\rangle}$ *est une norme sur* E*.*

DÉMONSTRATION : Il résulte immédiatement de la définition d'un produit scalaire que le vecteur x est nul si et seulement si $\langle x,x\rangle = 0$, et que $\langle \lambda x, \lambda x\rangle = \lambda\bar{\lambda}\langle x,x\rangle = |\lambda|^2\,\langle x,x\rangle$, c'est-à-dire que $\|x\| = 0 \iff x = 0$ et que $\|\lambda x\| = |\lambda|\,\|x\|$.
Si maintenant x et y sont deux vecteurs de E, on a

$$\begin{aligned}\|x+y\|^2 = \langle x+y, x+y\rangle &= \langle x,x\rangle + \langle y,y\rangle + \langle x,y\rangle + \langle y,x\rangle \\ &= \|x\|^2 + \|y\|^2 + 2\,\Re e(\langle x,y\rangle)\ ,\end{aligned}$$

et d'après l'inégalité de Cauchy-Schwarz

$$\begin{aligned}\|x+y\|^2 &\leqslant \|x\|^2 + \|y\|^2 + 2\,|\langle x,y\rangle| \leqslant \|x\|^2 + \|y\|^2 + 2\sqrt{\langle x,x\rangle \cdot \langle y,y\rangle} \\ &\leqslant \|x\|^2 + \|y\|^2 + 2\,\|x\| \cdot \|y\| = \big(\|x\| + \|y\|\big)^2\ ,\end{aligned}$$

donc $\|x+y\| \leqslant \|x\| + \|y\|$. ■

Définition 3.1.4. *On appelle* espace préhilbertien *un espace vectoriel normé muni d'un produit scalaire tel que* $\|x\|^2 = \langle x,x\rangle$*.*

Définition 3.1.5. *Un vecteur* x *de l'espace préhilbertien* E *est dit orthogonal à un vecteur* y *si leur produit scalaire est nul. On notera alors* $x \perp y$*.*

Cette relation est clairement symétrique. Le seul vecteur de E qui soit orthogonal à lui-même est 0.

Proposition 3.1.6. *Si deux vecteurs x et y sont orthogonaux, on a*

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \ .$$

En effet, on a en général $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\Re e(\langle x,y\rangle)$. ■

Théorème 3.1.7. *Si E est un espace préhilbertien, le produit scalaire est continu de $E \times E$ dans $\mathbb{C}$.*

Démonstration : On a, pour a, b, x, y dans E

$$\langle x,y\rangle - \langle a,b\rangle = \langle x,y-b\rangle + \langle x-a,b\rangle = \langle a,y-b\rangle + \langle x-a,b\rangle + \langle x-a,y-b\rangle\,,$$

donc, d'après l'inégalité de Cauchy-Schwarz,

$$|\langle x,y\rangle - \langle a,b\rangle| \leqslant \|a\|\cdot\|y-b\| + \|b\|\cdot\|x-a\| + \|x-a\|\cdot\|y-b\| \ ,$$

et si $\varepsilon > 0$ est donné, il suffit de prendre $\|x-a\| < \min(1, \dfrac{\varepsilon}{3\,\|b\|})$ et $\|y-b\| < \min(\dfrac{\varepsilon}{3}, \dfrac{\varepsilon}{3\,\|a\|})$ pour que chacun des trois termes ci-dessus soit inférieur à $\varepsilon/3$, donc que $|\langle x,y\rangle - \langle a,b\rangle| \leqslant \varepsilon$, ce qui montre la continuité du produit scalaire. ■

Théorème 3.1.8. *Si E est un espace préhilbertien, le produit scalaire de deux vecteurs x et y est donné par*

$$\langle x,y\rangle = \frac{1}{4}\big(\|x+y\|^2 + i\,\|x+iy\|^2 - \|x-y\|^2 - i\,\|x-iy\|^2\big)$$

Démonstration : On a, pour $k = 0,1,2$ ou 3,

$$\begin{aligned}
\left\|x+i^k y\right\|^2 &= \langle x,x\rangle + \langle y,y\rangle + i^{-k}\langle x,y\rangle + i^k\langle y,x\rangle \ ,\\
i^k\left\|x+i^k y\right\|^2 &= i^k\,\|x\|^2 + i^k\,\|y\|^2 + \langle x,y\rangle + (-1)^k\langle y,x\rangle \ ,\\
\sum_{k=0}^{3} i^k\left\|x+i^k y\right\|^2 &= \|x\|^2\,(1+i+i^2+i^3) + \|y\|^2\,(1+i+i^2+i^3)\\
&\qquad + \langle x,y\rangle(1+1+1+1) + \langle y,x\rangle(1-1+1-1) = 4\langle x,y\rangle,
\end{aligned}$$

ce qui prouve l'égalité cherchée. ■

Théorème 3.1.9. *Si x, y et z sont trois points d'un espace préhilbertien, et $u = \dfrac{x+y}{2}$ le milieu du segment $[x,y]$, on a*

$$\|z-x\|^2 + \|z-y\|^2 = 2\,\|z-u\|^2 + \frac{1}{2}\,\|x-y\|^2 \,.$$

Démonstration : Si on pose $v = \dfrac{y-x}{2}$, on a $z-x = (z-u)-v$ et $z-y = (z-u)+v$, donc

$$\begin{aligned}
\|z-x\|^2 + \|z-y\|^2 &= \|z-u\|^2 + \|v\|^2 + 2\,\Re e(\langle z-u,v\rangle)\\
&\qquad + \|z-u\|^2 + \|v\|^2 - 2\,\Re e(\langle z-u,v\rangle)\\
&= 2\,\|z-u\|^2 + 2\,\|v\|^2 = 2\,\|z-u\|^2 + \frac{1}{2}\,\|x-y\|^2 \ .
\end{aligned}$$ ■

Définition 3.1.10. *On appelle* espace hilbertien *ou* espace de Hilbert *un espace préhilbertien complet.*

Exemple 3.1.11. *L'espace vectoriel ℓ^2 des suites (x_n) de nombres complexes telles que la série $\sum_{n=0}^{\infty} |x_n|^2$ converge, muni du produit scalaire défini par*

$$\langle x, y \rangle = \sum_{n=0}^{\infty} x_n \bar{y}_n \,, \tag{$*$}$$

si $x = (x_n)$ et $y = (y_n)$, est un espace de Hilbert.

Il est clair que si x appartient à ℓ^2, il en est de même de λx. Et puisque pour a et b dans $\mathbb{C}$ on a $|a+b|^2 \leqslant 2(|a|^2 + |b|^2)$, on voit que si x et y sont dans ℓ^2

$$\sum_{n=0}^{\infty} |x_n + y_n|^2 \leqslant 2\Big(\sum_{n=0}^{\infty} |x_n|^2 + \sum_{n=0}^{\infty} |y_n|^2 \Big) < +\infty \,,$$

ce qui montre que $x + y$ appartient à ℓ^2.
L'inégalité de Cauchy-Schwarz montre que, pour tout m,

$$\begin{aligned} \sum_{n=0}^{m} |x_n \bar{y}_n| &\leqslant \Big(\sum_{n=0}^{m} |x_n|^2 \Big)^{1/2} \Big(\sum_{n=0}^{m} |y_n|^2 \Big)^{1/2} \\ &\leqslant \Big(\sum_{n=0}^{\infty} |x_n|^2 \Big)^{1/2} \Big(\sum_{n=0}^{\infty} |y_n|^2 \Big)^{1/2} < +\infty \,, \end{aligned}$$

ce qui entraîne la convergence absolue de la série de terme général $(x_n \bar{y}_n)$, et permet de définir $\langle x, y \rangle$. On vérifie sans peine que la formule $(*)$ définit bien un produit scalaire sur ℓ^2, et que la norme associée est définie par

$$\|x\| = \Big(\sum_{n=0}^{\infty} |x_n|^2 \Big)^{1/2} .$$

Il ne reste plus qu'à montrer que ℓ^2 est complet. Il suffit pour cela de reprendre la méthode utilisée dans le théorème 2.1.23 pour démontrer que ℓ^1 est complet. ■

Projection orthogonale.

Théorème 3.1.12. *Soient E un espace préhilbertien et A une partie convexe complète non vide de E. Alors, pour tout x de E, il existe un unique point y de A, appelé* projection orthogonale de x sur A, *tel que $d(x, A) = \|x - y\|$.*
Ce point est caractérisé par la propriété

$$\forall z \in A \quad \Re e(\langle x - y, z - y \rangle) \leqslant 0 \,.$$

Démonstration : Si $x \in A$, il est clair que $y = x$ est le seul point de A qui minimise la distance à x. De plus, on a $\langle x - x, z - x \rangle = 0$ pour tout $z \in A$. Et si pour un certain $y \in A$

on a $\Re e(\langle x-y, z-y\rangle) \leqslant 0$ pour tout $z \in A$, on a $\|x-y\|^2 = \Re e(\langle x-y, x-y\rangle) \leqslant 0$, d'où $x = y$.
Si, au contraire, $x \notin A$, on note δ la distance de x à A. Pour tout $n \in \mathbb{N}$, on pose

$$C_n = \{z \in A : \|z-x\|^2 \leqslant \delta^2 + 2^{-2n}\}.$$

Alors la suite (C_n) est une suite décroissante de fermés non vides de l'espace complet A. En vertu du théorème 1.8.9, il suffit de montrer que le diamètre de C_n tend vers 0 pour montrer que l'intersection des (C_n) est un singleton $\{y\}$.
Soient donc z et w deux points de C_n. Puisque A est convexe, $u = \dfrac{z+w}{2}$ appartient aussi à A, d'où on déduit que $\|x-u\| \geqslant \delta$. Et le théorème 3.1.9 donne alors

$$\begin{aligned}2\delta^2 + \frac{1}{2}\|z-w\|^2 \leqslant 2\|x-u\|^2 + \frac{1}{2}\|z-w\|^2 &= \|x-z\|^2 + \|x-w\|^2 \\ &\leqslant \delta^2 + 2^{-2n} + \delta^2 + 2^{-2n}.\end{aligned}$$

d'où on déduit que $\|z-w\|^2 \leqslant 4{\cdot}2^{-2n}$, c'est-à-dire $\operatorname{diam}(C_n) \leqslant 2^{1-n}$. Ceci achève de prouver l'existence d'un point $y \in A$ tel que $\bigcap_{n\in\mathbb{N}} C_n = \{y\}$. Le point y est donc le seul point de A tel que $\|x-y\| = d(x, A)$.
Soit z un point de A. Par convexité de A, on a $z_t := y + t(z-y) \in A$ pour tout $t \in]0,1]$. Donc

$$\begin{aligned}\delta^2 \leqslant \|x-z_t\|^2 &= \langle x-y-t(z-y), x-y-t(z-y)\rangle \\ &= \|x-y\|^2 + t^2\|z-y\|^2 - 2t\,\Re e(\langle x-y, z-y\rangle) \\ &= \delta^2 + t^2\|z-y\|^2 - 2t\,\Re e(\langle x-y, z-y\rangle).\end{aligned}$$

On en déduit que $2\,\Re e(\langle x-y, z-y\rangle) \leqslant t\|z-y\|^2$, et puisque t peut être pris arbitrairement petit, on obtient que $\Re e(\langle x-y, z-y\rangle) \leqslant 0$.
Si maintenant un point y' de A vérifie cette relation pour tout $z \in A$, on a pour $z = y$,

$$\begin{aligned}0 \geqslant \Re e(\langle x-y', y-y'\rangle) &= \Re e(\langle x-y, y-y'\rangle) + \Re e(\langle y-y', y-y'\rangle) \\ &= \|y-y'\|^2 - \Re e(\langle x-y, y'-y\rangle) \geqslant \|y-y'\|^2,\end{aligned}$$

puisque y est la projection de x sur A. Ceci entraîne $\|y-y'\| = 0$ donc $y = y'$. ∎

Théorème 3.1.13. *Soient E un espace hilbertien et A une partie convexe fermée non vide de E. Alors la projection orthogonale sur A est une application 1-lipschitzienne sur E.*

Démonstration : Soient x et x' deux points de E, y et y' leurs projections respectives sur A. On a alors

$$\begin{aligned}d(y,y')^2 &= \Re e(\langle y-y', y-y'\rangle) \\ &\leqslant \Re e(\langle x-y, y-y'\rangle) + \Re e(\langle y-y', y-y'\rangle) + \Re e(\langle y'-x', y-y'\rangle) \\ &= \Re e(\langle x-x', y-y'\rangle) \leqslant |\langle x-x', y-y'\rangle| \\ &\leqslant \|x-x'\| \,.\, \|y-y'\| = d(x,x'){\cdot}d(y,y'),\end{aligned}$$

donc finalement $d(y,y') \leqslant d(x,x')$, ce qui achève la démonstration. ∎

Définition 3.1.14. *Si A est une partie d'un espace préhilbertien E, on appelle orthogonal de A et on note $A^\perp$ l'ensemble des vecteurs y de E orthogonaux à tout élément de A.*

Pour tout $x \in A$, $x^\perp = \{y : \langle y, x\rangle = 0\}$ est le noyau de la forme linéaire continue (voir le théorème 3.1.7) $y \mapsto \langle y, x\rangle$, donc est un sous-espace vectoriel fermé de E. Il en résulte que

$$A^\perp = \bigcap_{x\in A} x^\perp$$

est un sous-espace vectoriel fermé de A.

Théorème 3.1.15. *Si F est un sous-espace vectoriel* fermé *de l'espace hilbertien E, il existe pour tout x de E un unique couple (y,z) avec $y \in F$ et $z \in F^\perp$ tel que $x = y + z$. On a alors $\|y\| \leqslant \|x\|$ et $\|z\| \leqslant \|x\|$.*

Démonstration : Un sous-espace vectoriel fermé d'un espace complet est convexe et complet. Il résulte donc du théorème 3.1.12 que pour tout x de E il existe un unique $y \in F$ tel que $\|x - y\| = d(x, F)$. On a alors, pour tout $w \in F$, $y + w \in F$, $y - w \in F$, $y + iw \in F$ et $y - iw \in F$ donc

$$\begin{aligned} \Re e(\langle x-y, w\rangle) &= \Re e(\langle x-y, (y+w)-y\rangle) \leqslant 0 \\ -\Re e(\langle x-y, w\rangle) &= \Re e(\langle x-y, (y-w)-y\rangle) \leqslant 0 \\ \Im m(\langle x-y, w\rangle) &= \Re e(\langle x-y, (y+iw)-y\rangle) \leqslant 0 \\ -\Im m(\langle x-y, w\rangle) &= \Re e(\langle x-y, (y-iw)-y\rangle) \leqslant 0\,, \end{aligned}$$

d'où $\langle x - y, w\rangle = 0$, c'est-à-dire que $z = x - y \in F^\perp$. Puisque y et z sont orthogonaux, on a $\|x\|^2 = \|y + z\|^2 = \|x\|^2 + \|z\|^2$, donc $\|y\| \leqslant \|x\|$ et $\|z\| \leqslant \|x\|$.
Enfin, si on a une autre décomposition de x en somme d'un vecteur y' de F et d'un vecteur z' de $F^\perp$, on a $y + z = y' + z'$, donc $y - y' = z' - z$. Le vecteur $y - y'$ appartient à F et $z' - z \in F^\perp$. Donc $y - y'$ est orthogonal à lui-même, c'est-à-dire est nul. On conclut que $y = y'$ et $z = z'$. ■

Théorème 3.1.16. *Si F est un sous-espace vectoriel fermé d'un espace hilbertien, alors $F^{\perp\perp} = F$.*

Démonstration : Puisque tout vecteur de F est orthogonal à tout vecteur de $F^\perp$, on a $F \subset F^{\perp\perp}$.
Inversement, si x est un vecteur de $F^{\perp\perp}$, x se décompose en $y + z$, avec $y \in F$ et $z \in F^\perp$. Puisque x et y sont orthogonaux à $F^\perp$, $z = x - y$ appartient à $F^\perp$ et est orthogonal à $F^\perp$; il est donc orthogonal à lui-même, c'est-à-dire nul. Donc $x = y \in F$. ■

Théorème 3.1.17. *Soit X un sous-ensemble de l'espace hilbertien E. Alors X est total si et seulement si $X^\perp = \{0\}$.*

Démonstration : Si l'espace vectoriel V engendré par X est dense et si $y \in X^\perp$, on a $V \subset y^\perp$. Et puisque $y^\perp$ est fermé, il contient l'adhérence de V, c'est-à-dire E ; en particulier $y \in y^\perp$, d'où $y = 0$. Ceci montre que $X^\perp = \{0\}$.
Inversement, si X n'est pas total, l'adhérence du sous-espace V engendré par X est un sous-espace vectoriel fermé F distinct de E, et tout vecteur z de $F^\perp$ appartient à $X^\perp$. Si on avait $F^\perp = \{0\}$, on aurait $F = F^{\perp\perp} = \{0\}^\perp = E$, contrairement à l'hypothèse. Donc $X^\perp \neq \{0\}$. ■

Théorème 3.1.18. *Si F est un sous-espace vectoriel* fermé *de l'espace hilbertien E, il existe une unique application linéaire continue P de E dans E de norme 1, telle que, pour tout x de E, $P(x)$ appartienne à F et, pour tout x de F, $P(x) = x$. Cette application est appelée* projecteur orthogonal *sur F.*

DÉMONSTRATION : Pour tout x de E, il existe un unique couple (y, z) dans $F \times F^\perp$ tel que $x = y + z$. Posons $P(x) = y$. On a $P(x) \in F$, et si $x \in F$ on a $y = x$ et $z = 0$, donc $P(x) = x$.
Si x et x' sont deux vecteurs de E, et $\lambda \in \mathbb{C}$, si $x = y + z$ et $x' = y' + z'$, avec y et y' dans F, z et z' dans $F^\perp$, on a

$$x + \lambda x' = (y + \lambda y') + (z + \lambda z')\,,$$

et puisque $y + \lambda y' \in F$ et $z + \lambda z' \in F^\perp$, on conclut que

$$P(x + \lambda x') = y + \lambda y' = P(x) + \lambda P(x')\,,$$

donc que P est linéaire.
Enfin, si P_1 vérifie les mêmes conditions, soient $z \in F^\perp$ et $y = P_1(z) \in F$. Alors, pour tout $t \in \mathbb{R}$ on a $P_1(z + ty) = P_1(z) + tP_1(y) = (1 + t)y$. On doit avoir, pour tout t, $\|z + ty\|^2 - \|P_1(z + ty)\|^2 \geqslant 0$. Et

$$\begin{aligned}\|z + ty\|^2 - \|P_1(z + ty)\|^2 &= \|z\|^2 + t^2\|y\|^2 - (1 + t)^2\|y\|^2 \\ &= \|z\|^2 - (1 + 2t)\|y\|^2\,,\end{aligned}$$

qui ne peut être positif pour tout t que si $\|y\| = 0$. Il en résulte que P_1 est nul sur $F^\perp$. Et si $x = y + z$ avec $y \in F$ et $z \in F^\perp$, on a

$$P_1(x) = P_1(y) + P_1(z) = P_1(y) = y = P(x)\,,$$

ce qui montre que P_1 coïncide avec P. ■

Théorème 3.1.19. (Riesz) *Si f est une forme linéaire continue sur l'espace de Hilbert E, il existe un unique $y \in E$ tel que, pour tout x de E, on ait $f(x) = \langle x, y\rangle$. De plus, on a $\|y\| = \|f\|$.*
L'application de E dans E' qui à y associe la forme linéaire $x \mapsto \langle x, y\rangle$ est un isomorphisme isométrique et semi-linéaire de E sur E'.

DÉMONSTRATION : Notons, pour $y \in E$, f_y la forme linéaire : $x \mapsto \langle x, y\rangle$. D'après la définition d'un produit scalaire, l'application $y \mapsto f_y$ est semi-linéaire de E dans E'. D'après l'inégalité de Cauchy-Schwarz, on a

$$|f_y(x)| = |\langle x, y\rangle| \leqslant \|x\|\cdot\|y\|$$

d'où $\|f_y\| \leqslant \|y\|$. De plus, puisque $f_y(y) = \langle y, y\rangle = \|y\|^2$, on a $\|f_y\| \geqslant \|y\|$. D'où $\|f_y\| = \|y\|$.
Si $f_y = f_z$, on a $f_{y-z} = f_y - f_z = 0$, donc $\|y - z\| = \|f_{y-z}\| = 0$, c'est-à-dire $y = z$.
Il reste uniquement à démontrer que toute forme linéaire continue f sur E est de la forme f_y pour un certain y. Si $f = 0$, il est clair que $y = 0$ convient. Si $f \neq 0$, le noyau H de f est un sous-espace vectoriel fermé de E. Soit $a \notin H$. Posons $\alpha = f(a)$ et notons b la projection

orthogonale de a sur H ; alors le vecteur $y = \dfrac{\bar{\alpha}}{\|a-b\|^2}(a-b)$ est orthogonal à H puisque $a-b$ l'est. En particulier, $\langle b, a-b\rangle = 0$, donc : $\langle a, a-b\rangle = \langle a-b, a-b\rangle = \|a-b\|^2$. De plus,

$$\langle a, y\rangle = \frac{\alpha}{\|a-b\|^2}\langle a, a-b\rangle = \alpha\,.$$

Il en résulte que $f_y(a) = f(a)$. Donc f et f_y coïncident sur a, et sur H puisqu'elles y sont nulles toutes deux.
Soit $x \in E$. Le vecteur $y = x - \alpha^{-1} f(x).a$ vérifie $f(y) = 0$, c'est-à-dire $y \in H$. Tout vecteur de E est donc somme d'un vecteur de H et d'un multiple de a ; on en déduit que f et f_y coïncident sur E, c'est-à-dire $f = f_y$. ■

Définition 3.1.20. *Soit E un espace de Hilbert. Une fonction $f : E \times E \to \mathbb{C}$ est appelée* forme sequilinéaire *si la fonction $x \mapsto f(x,y)$ est linéaire pour tout y et si la fonction $y \mapsto f(x,y)$ est semi-linéaire pour tout x.*

Théorème 3.1.21. *Si E est un espace de Hilbert et $f : E \times E \to \mathbb{C}$ une forme sequilinéaire continue, il existe une unique application linéaire continue $T \in \mathscr{L}(E)$ telle que $f(x,y) = \langle T(x), y\rangle$ pour tout x et tout y de E.*

Démonstration : Pour tout x fixé, l'application $g_x : y \mapsto \bar{f}(x,y)$ est une forme linéaire continue sur E. Il existe donc, d'après le théorème de Riesz, un unique vecteur $T(x) \in E$ tel que $g_x(y) = \langle y, T(x)\rangle$, c'est-à-dire $f(x,y) = \langle T(x), y\rangle$. Il résulte aisément de la linéarité de f par rapport à x que $f(x+\lambda x', y) = \langle T(x) + \lambda T(x'), y\rangle$, donc que $T(x+\lambda x') = T(x) + \lambda T(x')$, ce qui signifie que T est linéaire de E dans E. La continuité de f entraîne l'existence d'une constante M telle que $|f(x,y)| \leqslant M.\|x\|.\|y\|$; on en déduit que $\|T(x)\| = \|g_x\| \leqslant M.\|x\|$, ce qui prouve la continuité de T. ■

Inversement, il est clair que si T est un opérateur linéaire continu sur E, l'application $f : (x,y) \mapsto \langle T(x), y\rangle$ est une forme sequilinéaire continue sur E.

3.2 Séries orthogonales

Théorème 3.2.1. *Soient E un espace de Hilbert et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments deux-à-deux orthogonaux de E. Alors :*
i) si $\sum_n \|x_n\|^2 < +\infty$, la série de terme général (x_n) converge commutativement (c'est-à-dire que, pour toute bijection σ de $\mathbb{N}$ sur $\mathbb{N}$, on a $\sum_{n=0}^{\infty} x_{\sigma(n)} = \sum_{n=0}^{\infty} x_n$) vers un élément u de E tel que $\|u\|^2 = \sum_n \|x_n\|^2$.
ii) si $\sum_n \|x_n\|^2 = +\infty$, on a $\lim_{n\to\infty} \|\sum_{m=0}^{n} x_m\| = +\infty$.

Démonstration : Puisque les x_m sont deux-à-deux orthogonaux, il résulte du théorème de Pythagore que $\|\sum_{m=0}^{n} x_m\|^2 = \sum_{m=0}^{n} \|x_m\|^2$. On voit donc que si la série réelle de terme général $(\|x_n\|^2)$ est divergente, la suite $(\|\sum_{m=0}^{n} x_m\|)_n$ tend vers l'infini.

Inversement, si $\sum_n \|x_n\|^2 < +\infty$, on va montrer que la famille (x_n) est *sommable*, c'est-à-dire qu'il existe un $u \in E$ tel que

$$\forall \varepsilon > 0 \quad \exists J_0 \text{ partie finie de } \mathbb{N} \quad \forall J \text{ partie finie de } \mathbb{N} \quad J \supset J_0 \Longrightarrow \left\| u - \sum_{n\in J} x_n \right\| < \varepsilon\ .$$

Notons d'abord que la suite $s_n = \sum_{m=0}^{n} x_m$ est une suite de Cauchy dans E . On a en effet : $\|s_{n+p} - s_n\|^2 = \left\|\sum_{m=n+1}^{n+p} x_m\right\|^2 = \sum_{m=n+1}^{n+p} \|x_m\|^2$, et cette quantité tend vers 0 lorsque n tend vers l'infini. La complétude de E assure alors l'existence d'un $u \in E$ tel que $s_n \to u$. Alors, puisque $\|u\|^2 = \lim_n \|\sum_{m=0}^{n} x_m\|^2 = \lim_n \sum_{m=0}^{n} \|x_m\|^2$, on a $\|u\|^2 = \sum_n \|x_n\|^2$.
Soit maintenant $\varepsilon > 0$. Il existe n_0 tel que $\sum_{m=0}^{n_0} \|x_n\|^2 > \sum_{m=0}^{\infty} \|x_n\|^2 - \varepsilon^2$. On prend alors $J_0 = \{0, 1, \dots, n_0\}$. Et si J est une partie finie de $\mathbb{N}$ qui contient J_0, il existe p tel que $J \subset \{0, 1, \dots, n_0 + p\}$. On a alors

$$\left\|\sum_{m\in J} x_m - \sum_{m\in J_0} x_m\right\|^2 = \left\|\sum_{m\in J\setminus J_0} x_m\right\|^2 = \sum_{m\in J\setminus J_0} \|x_m\|^2 \leqslant \sum_{m=n_0+1}^{n+p} \|x_m\|^2$$
$$\leqslant \sum_{m=0}^{\infty} \|x_n\|^2 - \sum_{m=0}^{n} \|x_n\|^2 < \varepsilon^2 ,$$

donc $\left\|\sum_{m\in J} x_m - \sum_{m\in J_0} x_m\right\| < \varepsilon$. Puisque ceci est valable quel que soit $J \supset J_0$, on peut prendre $J = \{0, 1, \dots, n + p\}$ et faire tendre p vers l'infini : on en déduit que $\left\|u - \sum_{m\in J_0} x_m\right\| \leqslant \varepsilon$, donc que $\left\|u - \sum_{m\in J} x_m\right\| < 2\varepsilon$ dès que J est une partie finie de $\mathbb{N}$ contenant J_0.

Enfin, si σ est une bijection de $\mathbb{N}$ sur lui-même, et si $\varepsilon > 0$ est donné, la sommabilité de la famille (x_n) permet de trouver un J_0 possédant la propriété ci-dessus. L'ensemble fini $\sigma^{-1}(J_0)$ possède alors un plus grand élément n_1, et pour $n \geqslant n_1$, on a $\sigma(\{0, 1, \dots, n\}) \supset J_0$, donc $\left\|u - \sum_{m=0}^{n} x_{\sigma(m)}\right\| < \varepsilon$, ce qui montre que la série de terme général $(x_{\sigma(n)})$ converge vers u. ■

Théorème 3.2.2. *Soient E un espace de Hilbert et (e_n) une suite orthonormale de E, c'est-à-dire une suite d'éléments de norme 1 deux-à-deux orthogonaux. Alors, pour toute suite (λ_n) de scalaires, la série $\sum_n \lambda_n e_n$ converge si et seulement si $\sum_n |\lambda_n|^2 < +\infty$.*

De plus, pour tout élément x du sous-espace vectoriel fermé engendré par les (e_n), il existe une unique suite (λ_n) telle que $x = \sum_n \lambda_n e_n$. De façon précise, $\lambda_n = \langle x, e_n\rangle$.

Démonstration : Puisque $\|\lambda_n e_n\|^2 = |\lambda_n|^2$, la première partie de l'énoncé résulte immédiatement du théorème précédent. Si on considère alors l'application linéaire φ de ℓ^2 dans E définie par $\varphi(\lambda) = \sum_n \lambda_n e_n$ pour $\lambda = (\lambda_n) \in \ell^2$, le même théorème montre que φ est une isométrie sur son image V. Cette image est alors un sous-espace vectoriel complet, donc fermé, de E. Il est clair que chacun des e_n appartient à V (prendre les λ_n tous nuls sauf un qui vaut 1), donc que V contient le sous-espace vectoriel fermé engendré par les e_n. Inversement, chaque somme finie $\sum_{m=0}^{n} \lambda_m e_m$ appartient au sous-espace vectoriel W engendré par les (e_n), et les limites de telles sommes sont dans $\overline{W}$. On en conclut que $V = \overline{W}$.

Enfin, si $x = \sum_n \lambda_n e_n$, on a nécessairement pour tout entier p :

$$\langle x, e_p\rangle = \lim_{n\to\infty} \sum_{m=0}^{n} \langle \lambda_m e_m, e_p\rangle$$

et, par orthonormalité, cette dernière quantité est égale à λ_p dès que $n \geqslant p$. On a donc $\lambda_p = \langle x, e_p\rangle$, ce qui garantit l'unicité de la suite (λ_n). ■

Définition 3.2.3. *On appelle* base hilbertienne *de l'espace de Hilbert séparable E toute famille orthonormale (e_n) totale (c'est-à-dire engendrant un sous-espace vectoriel dense de E).*

Il résulte alors immédiatement du théorème précédent que si (e_n) est une base hilbertienne de E, l'application $\lambda \mapsto \sum_n \lambda_n e_n$ est un isomorphisme isométrique de ℓ^2 sur E.

Et si E est un espace de Hilbert séparable de dimension infinie, on peut trouver dans E une suite dense (u_n). La méthode d'orthogonalisation de Hilbert-Schmidt fournit, à partir de (u_n), une suite orthonormale (e_n) telle que, pour tout n, u_n soit combinaison linéaire des $(e_p)_{p \leqslant n}$. Il en résulte que (e_n) est alors une base hilbertienne, et que tout espace hilbertien séparable est isométriquement isomorphe à ℓ^2.

Exemple : séries de Fourier.

Si E désigne l'espace hilbertien $L^2([0, 2\pi])$, muni du produit scalaire $\langle f, g \rangle = \int_0^{2\pi} f(t)\bar{g}(t)\, dt$ (voir 5.12.5), la famille $(e_n)_{n \in \mathbb{Z}}$ définie par $e_n(t) = \dfrac{1}{\sqrt{2\pi}}\, \mathrm{e}^{int}$ est orthonormale : en effet

$$\langle e_n, e_p \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{e}^{i(n-p)t}\, dt \ ,$$

et cette quantité vaut 1 si $n = p$ et 0 sinon.
De plus, si $f \in E$ est orthogonale à tous les (e_n), elle est orthogonale à tous les polynômes trigonométriques $\sum_{-n}^{n} c_n\, \mathrm{e}^{int}$. Il résulte du théorème de Stone-Weierstrass (1.10.11) que ces polynômes trigonométriques forment une partie dense dans l'espace des fonctions continues 2π-périodiques pour la topologie de la convergence uniforme. On en déduit, par densité, que f est alors orthogonale à $L^2(\,[0, 2\pi]\,)$, donc nulle. Le théorème 3.1.17 montre alors que la suite orthonormale (e_n) est totale, donc est une base hilbertienne de E.
On voit alors que, pour tout $f \in L^2([0, 2\pi])$, on a le développement $f(t) = \sum_{n \in \mathbb{Z}} c_n\, \mathrm{e}^{int}$ (au sens de la convergence dans L^2), avec les coefficients $c_n = \dfrac{1}{\sqrt{2\pi}} \langle f, e_n \rangle = \dfrac{1}{2\pi} \displaystyle\int_0^{2\pi} f(t)\, \mathrm{e}^{int}\, dt$.

3.3 Adjoint d'un opérateur

Définition 3.3.1. *On appelle* opérateur *sur l'espace de Hilbert E toute application linéaire continue de E dans E.*

Si T est un opérateur sur E, il est d'usage de noter Tx l'image de x par T.

Théorème 3.3.2. *Si T est un opérateur sur l'espace de Hilbert E, il existe un opérateur T^* sur E, appelé* l'adjoint *de T tel que, pour tout x et tout y de E, on ait*

$$\langle Tx, y \rangle = \langle x, T^* y \rangle \, .$$

De plus, on a $T^{**} = T$ et $\|T\| = \|T^*\|$.

Démonstration : Pour tout y fixé, l'application $x \mapsto \langle Tx, y \rangle$, composée de la forme linéaire continue $f_y : x \mapsto \langle x, y \rangle$ et de T, est une forme linéaire continue. Il résulte alors du théorème 3.1.19 qu'il existe un $z = \varphi(y)$ tel que cette forme linéaire continue soit égale à f_z. On a donc, pour tout x et tout y

$$\langle Tx, y\rangle = \langle x, \varphi(y)\rangle .$$

On a donc pour tout x tout y et tout y'

$$\begin{aligned}\langle x, \varphi(y+y')\rangle &= \langle Tx, y+y'\rangle = \langle Tx, y\rangle + \langle Tx, y'\rangle \\ &= \langle x, \varphi(y)\rangle + \langle x, \varphi(y')\rangle ,\end{aligned}$$

donc, pour tout x,

$$\langle x, \varphi(y+y') - \varphi(y) - \varphi(y')\rangle = 0 ,$$

d'où $\varphi(y+y') - \varphi(y) - \varphi(y') = 0$ (il suffit de prendre $x = \varphi(y+y') - \varphi(y) - \varphi(y')$ pour obtenir $\|x\|^2 = 0$). De même, pour tout x,

$$\langle x, \varphi(\lambda y)\rangle = \langle Tx, \lambda y\rangle = \bar\lambda\langle Tx, y\rangle = \bar\lambda\langle x, \varphi(y)\rangle = \langle x, \lambda\varphi(y)\rangle \ ,$$

d'où l'on déduit que $\varphi(\lambda y) = \lambda\varphi(y)$.
Ceci montre que φ est linéaire de E dans E. Enfin, pour montrer que φ est continue, nous prenons $x = \varphi(y)$. On a alors

$$\begin{aligned}\|\varphi(y)\|^2 &= \langle \varphi(y), \varphi(y)\rangle = \langle T\varphi(y), y\rangle \\ &\leqslant \|T\varphi(y)\|\cdot\|y\| \leqslant \|T\|\cdot\|\varphi(y)\|\cdot\|y\| \ ,\end{aligned}$$

d'où l'on déduit $\|\varphi(y)\| \leqslant \|T\|\cdot\|y\|$, ce qui montre que l'application linéaire φ est continue de norme au plus $\|T\|$. Donc φ est un opérateur que l'on notera T^*. Et on a $\|T^*\| \leqslant \|T\|$. Enfin, puisque pour tout x et tout y on a

$$\langle Tx, y\rangle = \langle x, T^*y\rangle ,$$

on a

$$\langle y, Tx\rangle = \langle T^*y, x\rangle = \langle y, T^{**}x\rangle ,$$

ce qui montre que l'adjoint de T^* est T. On a donc $\|T\| = \|T^{**}\| \leqslant \|T^*\| \leqslant \|T\|$. Et ceci prouve que $\|T\| = \|T^*\|$. ■

Théorème 3.3.3. *Si S et T sont deux opérateurs sur E, et $\lambda \in \mathbb{C}$, on a $(S+T)^* = S^* + T^*$, $(\lambda S)^* = \bar\lambda S^*$ et $(ST)^* = T^*S^*$.*

Démonstration : On a pour tout x et tout y

$$\begin{aligned}\langle x, (S+T)^*y\rangle &= \langle (S+T)x, y\rangle = \langle Sx, y\rangle + \langle Tx, y\rangle \\ &= \langle x, S^*y\rangle + \langle x, T^*y\rangle = \langle x, (S^*+T^*)y\rangle ,\end{aligned}$$

donc $(S+T)^* = S^* + T^*$. De même,

$$\begin{aligned}\langle x, (\lambda S)^*y\rangle &= \langle \lambda Sx, y\rangle = \lambda\langle Sx, y\rangle \\ &= \lambda\langle x, S^*y\rangle = \langle x, \bar\lambda S^*y\rangle ,\end{aligned}$$

donc $(\lambda S)^* = \bar\lambda S^*$. Enfin,

$$\begin{aligned}\langle x, (ST)^*y\rangle &= \langle STx, y\rangle = \langle Tx, S^*y\rangle \\ &= \langle x, T^*S^*y\rangle ,\end{aligned}$$

d'où $(ST)^* = T^*S^*$. ■

Théorème 3.3.4. *Si T est un opérateur compact sur E, son adjoint T^* est compact.*

DÉMONSTRATION : Soit $\varepsilon > 0$. Puisque $T(B)$ est relativement compact, donc précompact, il existe un nombre fini de points $(x_1, x_2, \ldots, x_k)$ de B tels que $T(B) \subset \bigcup_{1 \leqslant j \leqslant k} B(Tx_j, \varepsilon)$. Si V désigne le sous-espace de E de dimension finie engendré par Tx_1, Tx_2, …, Tx_k, et P le projecteur orthogonal sur V, on a $\|PT - T\| \leqslant \varepsilon$: en effet, si $x \in B$, il existe $j \leqslant k$ tel que $\|Tx - Tx_j\| \leqslant \varepsilon$, donc que

$$\|Tx - PTx\| = d(Tx, V) \leqslant \|Tx - Tx_j\| \leqslant \varepsilon \ ,$$

donc $\|T - PT\| = \sup_{x \in B} \|Tx - PTx\| \leqslant \varepsilon$. Et puisque $P^* = P$, on a $\|T^* - T^*P\| \leqslant \varepsilon$. Comme P est de rang fini, il en est de même de T^*P. Il résulte alors du théorème 2.5.6 que T^* est compact. ■

Définition 3.3.5. *Soient E un espace de Hilbert et T un opérateur sur E. On dit que T est* hermitien *(ou* auto-adjoint*) si T est égal à son adjoint T^*, et que T est* normal *s'il commute avec son adjoint.*

Il est clair qu'un opérateur auto-adjoint est normal.

3.4 Compacité faible

On a vu que la boule unité d'un espace normé de dimension infinie n'est pas compacte et que, par conséquent, il existe des suites bornées qui n'ont aucune sous-suite convergente. On va voir néanmoins, dans le cas des espaces hilbertiens, que toute suite bornée possède des sous-suites qui convergent en un sens faible.

Définition 3.4.1. *On dit que la suite (x_n) dans l'espace de Hilbert E* converge faiblement *vers a si, pour tout y de E, on a*

$$\lim_{n \to \infty} \langle x_n, y \rangle = \langle a, y \rangle \, .$$

On note alors $(x_n) \rightharpoonup a$.

Théorème 3.4.2. *Toute suite faiblement convergente est bornée.*

C'est une conséquence immédiate du théorème de Banach-Steinhaus (corollaire 2.4.3), puisque toute suite convergente dans $\mathbb{C}$ est bornée. ■

Théorème 3.4.3. *Si (x_n) est une suite dans la boule unité qui converge faiblement vers un point a de norme 1, alors (x_n) converge en norme vers a.*

DÉMONSTRATION : On a

$$\|x_n - a\|^2 = \langle x_n - a, x_n - a \rangle = \|x_n\|^2 + \|a\|^2 - 2\,\Re e(\langle x_n, a \rangle) \leqslant 2 - 2\,\Re e(\langle x_n, a \rangle) \leqslant 2\,|1 - \langle x_n, a \rangle|$$

et puisque $\langle x_n, a \rangle \to \langle a, a \rangle = \|a\|^2 = 1$, il existe pour tout $\varepsilon > 0$ un rang N tel que $|1 - \langle x_n, a \rangle| < \dfrac{\varepsilon^2}{2}$ si $n \geqslant N$. On a alors $\|x_n - a\| < \varepsilon$ si $n \geqslant N$. ■

Lemme 3.4.4. *Si (x_n) est une suite bornée dans l'espace hilbertien E, si D est une partie dense de E et si, pour tout élément y de D la suite numérique $(\langle x_n, y\rangle)$ est de Cauchy dans $\mathbb{C}$, alors la suite (x_n) est faiblement convergente.*

Démonstration : Soit $M = \sup_{n\in\mathbb{N}} \|x_n\|$. Notons V l'ensemble de tous les z de E tels que la suite $(\langle x_n, z\rangle)$ soit une suite de Cauchy. Par hypothèse, D est contenu dans V.
De plus, V est un sous-espace vectoriel de E. En effet, si z et z' appartiennent à V et si $\lambda \in \mathbb{C}$ on a

$$\langle x_n, z + \lambda z'\rangle = \langle x_n, z\rangle + \bar{\lambda}\langle x_n, z'\rangle\,,$$

ce qui montre que la suite $(\langle x_n, z + \lambda z'\rangle)$ est une suite de Cauchy, donc que $z + \lambda z' \in V$.
Enfin V est fermé dans E. En effet, si w est adhérent à V et si $\varepsilon > 0$, il existe un $z \in V$ tel que $\|z - w\| < \dfrac{\varepsilon}{3M}$, et un m tel que, pour n et p supérieurs à m, on ait

$$|\langle x_n, z\rangle - \langle x_p, z\rangle| < \frac{\varepsilon}{3}\,.$$

Alors, pour n et p supérieurs à m on peut majorer $|\langle x_n, w\rangle - \langle x_p, w\rangle|$ par

$$\begin{aligned}
&|\langle x_n, w\rangle - \langle x_n, z\rangle| + |\langle x_n, z\rangle - \langle x_p, z\rangle| + |\langle x_p, z\rangle - \langle x_p, w\rangle| \\
&\quad \leqslant \|x_n\|\,\|w - z\| + \varepsilon/3 + \|x_p\|\,\|w - z\| \\
&\quad \leqslant 2M\,\|w - z\| + \varepsilon/3 < \varepsilon\,,
\end{aligned}$$

ce qui montre que $w \in V$.
Il en résulte que $V = E$, donc que la suite $(\langle x_n, z\rangle)$ converge vers un nombre $f(z)$ pour tout z de E. On vérifie sans peine que $z \mapsto \bar{f}(z)$ est une forme linéaire sur E et que

$$\left|\bar{f}(z\right| \leqslant \sup_{n\in\mathbb{N}} |\langle x_n, z\rangle| \leqslant \sup_{n\in\mathbb{N}} \|x_n\|\,\|z\| \leqslant M\ \|z\|\ ,$$

donc que $\bar{f}$ est continue. D'après le théorème 3.1.19, il existe un x tel que $\bar{f}(z) = \langle z, x\rangle$, c'est-à-dire $f(z) = \langle x, z\rangle$. Ceci signifie exactement que $x_n \rightharpoonup x$. ∎

Lemme 3.4.5. *Si A est une partie convexe fermée de l'espace hilbertien E, et si (x_n) est une suite de points de A qui converge faiblement vers un point x, ce point x appartient à A.*

Démonstration : Supposons que x n'appartienne pas à A. Alors, si a désigne la projection orthogonale de x sur A, on a, $\Re e(\langle x - a, z - a\rangle) \leqslant 0$ pour tout z de A. Donc, pour tout n, $\Re e(\langle x - a, x_n - a\rangle) \leqslant 0$. Mais la suite de terme général $\Re e(\langle x - a, x_n - a\rangle)$ tend vers $\|x - a\|^2 > 0$; on obtient ainsi une contradiction. ∎

Théorème 3.4.6. *Si A est une partie convexe fermée et bornée de l'espace de Hilbert E, toute suite (x_n) de A possède une sous-suite qui converge faiblement vers un point de A.*

Démonstration : Soient $V = \{y_k : k \in \mathbb{N}\}$ l'ensemble (dénombrable) des combinaisons linéaires à coefficients rationnels des x_n et des ix_n, et $M = \sup_{x\in A} \|x\| < +\infty$. Pour tout x de A et tout entier k on a $|\langle x, y_k\rangle| \leqslant \|x\| . \|y_k\| \leqslant \rho_k = M. \|y_k\|$. Le nombre $\langle x, y_k\rangle$ appartient donc au disque compact Δ_k de rayon ρ_k. On considère l'application Φ de A dans $\prod_{k\in\mathbb{N}} \Delta_k$ définie par

$$\Phi(x) = (\langle x, y_k\rangle)_{k\in\mathbb{N}}\,.$$

Puisque le produit des (Δ_k) est un espace métrique compact, la suite $(\Phi(x_n))_{n\in\mathbb{N}}$ possède une sous-suite $(\Phi(x_n))_{n\in H}$ convergente. Alors, la suite $(\langle x_n, y_k\rangle)_{n\in H}$ est convergente pour tout k. L'ensemble $\overline{V}$ est le sous-espace vectoriel fermé engendré par les x_n. Si $y \in V^\perp$, la suite $(\langle x_n, y\rangle)$ est identiquement nulle, donc converge. Il en résulte que l'ensemble $D = V + V^\perp$ est dense et que la suite $(\langle x_n, y\rangle)$ converge pour tout $y \in D$.
On déduit alors des deux lemmes précédents que la suite $(x_n)_{n\in H}$ converge faiblement vers un point de A. ∎

Corollaire 3.4.7. *Si T est un opérateur sur l'espace hilbertien E, l'image par T de la boule unité fermée B est fermée dans E.*

Démonstration : Soit a un point adhérent à $T(B)$. Il existe donc une suite (x_n) dans B telle que la suite (Tx_n) converge vers a. La suite (x_n) possède donc une sous-suite $(x_n)_{n\in H}$ qui converge faiblement vers un point b du convexe fermé borné B. Alors la suite (Tx_n) converge faiblement vers Tb. En effet, pour $y \in E$

$$\langle Tx_n, y\rangle = \langle x_n, T^*y\rangle \to \langle b, T^*y\rangle = \langle Tb, y\rangle.$$

Par ailleurs, puisque (Tx_n) converge vers a, on a $\langle Tx_n, y\rangle \to \langle a, y\rangle$, ce qui montre que, pour tout y, $\langle Tb, y\rangle = \langle a, y\rangle$, donc que $a = Tb \in T(B)$. Ceci achève de prouver que $T(B)$ est fermée. ∎

Théorème 3.4.8. *Soient E un espace de Hilbert et $T \in \mathscr{L}(E)$ un opérateur compact. Alors, pour toute suite (x_n) qui converge faiblement vers un point $a \in E$, la suite (Tx_n) converge en norme vers Ta.*

Démonstration : Il résulte du théorème 3.4.2 que la suite (x_n) est bornée. Puisque T est compact, l'ensemble $X = \{Tx_n : n \in \mathbb{N}\}$ est relativement compact dans E, et son adhérence $\overline{X}$ est compacte. La suite (Tx_n) possède donc dans $\overline{X}$ au moins une valeur d'adhérence. De plus, si u est une valeur d'adhérence de (Tx_n), on a pour tout $y \in E$:

$$\langle Tx_n - Ta, y\rangle = \langle T(x_n - a), y\rangle = \langle x_n - a, T^*y\rangle \to \langle a, T^*y\rangle - \langle a, T^*y\rangle = 0$$

donc $\langle u - Ta, y\rangle = 0$. En particulier, pour $y = u - Ta$, ceci montre que $u = Ta$. Et puisque, dans le compact $\overline{X}$, la suite (Tx_n) possède la seule valeur d'adhérence Ta, elle converge vers Ta. ∎

3.5 Le théorème spectral pour les opérateurs normaux compacts

Lemme 3.5.1. *Soit T un opérateur normal sur l'espace de Hilbert E. Si x est un vecteur propre de T pour la valeur propre λ, alors x est vecteur propre de T^* pour $\bar\lambda$.*

Démonstration : Puisque $T - \lambda I$ commute avec $T^* - \bar\lambda I$ on a, si $Tx = \lambda x$,

$$\begin{aligned} 0 = \|(T-\lambda I)x\|^2 &= \langle (T-\lambda I)x, (T-\lambda I)x\rangle = \langle (T^*-\bar\lambda I)(T-\lambda I)x, x\rangle \\ &= \langle (T-\lambda I)(T^*-\bar\lambda I)x, x\rangle = \langle (T^*-\bar\lambda I)x, (T^*-\bar\lambda I)x\rangle = \left\|(T^*-\bar\lambda I)x\right\|^2 \ , \end{aligned}$$

donc $\left\|(T^*-\bar\lambda I)x\right\| = 0$. ∎

Lemme 3.5.2. *Si T est un opérateur normal sur l'espace de Hilbert E, les sous-espaces propres de T sont deux-à-deux orthogonaux.*

DÉMONSTRATION : Soient λ et μ deux valeurs propres distinctes, $x \in \ker(T-\lambda I)$ et $y \in \ker(T-\mu I)$. D'après le lemme précédent, on a $T^*y = \bar{\mu}y$. Donc :

$$\lambda\langle x,y\rangle = \langle \lambda x,y\rangle = \langle Tx,y\rangle = \langle x,T^*y\rangle = \langle x,\bar{\mu}y\rangle = \mu\langle x,y\rangle \ ,$$

c'est-à-dire $(\lambda-\mu)\langle x,y\rangle = 0$, donc $x \perp y$ puisque $\lambda-\mu \neq 0$. ∎

Lemme 3.5.3. *Soit T un opérateur normal compact sur l'espace de Hilbert E. Alors les sous-espaces propres de T correspondant aux valeurs propres non nulles sont de dimension finie, et, pour tout $\varepsilon > 0$, l'ensemble des valeurs propres de module $\geqslant \varepsilon$ est fini.*

DÉMONSTRATION : Soit λ une valeur propre non nulle et V_λ le sous-espace propre associé. Alors V_λ est un sous-espace fermé, sur lequel la restriction de T est égale à λI. On a donc $I_{|V_\lambda} = \lambda^{-1}T_{|V_\lambda}$, ce qui montre que l'identité est sur V_λ un opérateur compact. L'espace V_λ est donc de dimension finie, en vertu du théorème 2.5.2.

S'il existait une infinité de valeurs propres de module supérieur à ε, on pourrait trouver une suite (λ_n) de telles valeurs propres deux-à-deux distinctes, et une suite (x_n) de vecteurs propres associés de norme 1. Comme x_n et x_m sont orthogonaux d'après le lemme 3.5.2,

$$\|Tx_n - Tx_m\|^2 = \|\lambda_n x_n - \lambda_m x_m\|^2 = |\lambda_n|^2\,\|x_n\|^2 + |\lambda_m|^2\,\|x_m\|^2 \geqslant 2\varepsilon^2 \ .$$

Mais puisque T est compact, il devrait exister une suite extraite (x_{n_k}) telle que (Tx_{n_k}) converge vers un élément a de E. Et ceci est impossible puisque $\|Tx_{n_k} - Tx_{n_\ell}\| \geqslant \varepsilon\sqrt{2}$. ∎

Lemme 3.5.4. *Si T est un opérateur normal compact sur l'espace de Hilbert E, il existe une valeur propre de T dont le module est égal à $\|T\|$.*

DÉMONSTRATION : On peut supposer que $T \neq 0$. On peut alors trouver dans la boule unité de E une suite (x_n) telle que $\|Tx_n\| \to \|T\|$. Quitte à remplacer x_n par $y_n = \dfrac{x_n}{\|x_n\|}$, on peut supposer $\|x_n\| = 1$: en effet

$$\|Tx_n\| \leqslant \|Ty_n\| \leqslant \|T\| \ ,$$

ce qui montre que $\|Ty_n\| \to \|T\|$. La suite (x_n) possède une sous-suite (x_{n_k}) qui converge faiblement vers un point a de la boule unité de E. Puisque T est compact, il résulte du théorème 3.4.8 que (Tx_{n_k}) converge vers Ta, et on a $\|Ta\| = \lim \|Tx_{n_k}\| = \|T\|$.
Si on avait $\|a\| < 1$, on aurait $\|Ta\| \leqslant \|T\|\,.\,\|a\| < \|T\|$. Donc $\|a\| = 1$, et il résulte du théorème 3.4.3 que $x_{n_k} \to a$, ce qui montre que $\|Ta\| = \|T\|$. La fonction $x \mapsto \|Tx\|^2$ atteint donc en a son maximum sur la sphère $\{x : \|x\|^2 = 1\}$. On en déduit qu'il existe un multiplicateur de Lagrange λ tel que les différentielles en a des fonctions $x \mapsto \|Tx\|^2$ et $x \mapsto \|x\|^2$ soient proportionnelles dans le rapport λ.
On vérifie sans peine que ces différentielles sont les formes $\mathbb{R}$-linéaires $h \mapsto 2\,\Re e(\langle T^*Ta,h\rangle)$ et $h \mapsto 2\,\Re e(\langle a,h\rangle)$. On en déduit que, pour tout $h \in E$ on doit avoir

$$\Re e\Big(\langle T^*Ta - \lambda a, h\rangle\Big) = 0 \ ,$$

donc $T^*Ta = \lambda a$ (en prenant $h = T^*Ta - \lambda a$), ce qui montre que a est un vecteur propre de T^*T pour la valeur propre λ. De plus, on a

$$\lambda = \lambda\,\|a\|^2 = \langle \lambda a, a\rangle = \langle T^*Ta, a\rangle = \langle Ta, Ta\rangle = \|Ta\|^2 = \|T\|^2 \ .$$

Puisque T^*T est compact et que λ n'est pas nul, le sous-espace propre F de T^*T pour la valeur propre λ est de dimension finie d'après le théorème 2.5.2, et stable par T puisque T commute avec T^*T. Il en résulte que $T_{|F}$ possède une valeur propre μ et un vecteur propre u associé à μ. On a alors $Tu = \mu u$, ainsi que $T^*u = \bar{\mu}u$ d'après le lemme 3.5.1. Il en résulte que $\|T\|^2 u = T^*Tu = T^*(\mu u) = \mu T^*u = \mu\bar{\mu}.u$. On conclut que u est un vecteur propre de T pour une valeur propre μ satisfaisant $|\mu|^2 = \|T\|^2$, c'est-à-dire $|\mu| = \|T\|$. ■

Lemme 3.5.5. *Soit T un opérateur normal compact sur l'espace de Hilbert E. Alors la somme des sous-espaces propres de T est dense dans E.*

Démonstration : Soit S l'ensemble des valeurs propres de T. Si on définit W comme la somme des sous-espaces propres $V_\lambda = \ker(T-\lambda I)$ pour λ dans S, on a clairement $T(W) \subset W$ et $T^*(W) \subset W$ puisque, si $x_j \in V_{\lambda_j}$,

$$T(\sum_{j=1}^m x_j) = \sum_1^m \lambda_j x_j \in W \quad \text{et} \quad T^*(\sum_{j=1}^m x_j) = \sum_1^m \bar{\lambda}_j x_j \in W\ .$$

Alors $T(\overline{W}) \subset \overline{W}$ et $T^*(\overline{W}) \subset \overline{W}$. Si $x \in W^\perp$, on a, pour $y \in W$: $\langle Tx, y\rangle = \langle x, T^*y\rangle = 0$ et $\langle T^*x, y\rangle = \langle x, Ty\rangle = 0$, ce qui montre que $Tx \in W^\perp$ et $T^*x \in W^\perp$.
Le sous-espace $W^\perp$ est donc un espace de Hilbert et $T_{|W^\perp}$ est un opérateur normal compact sur $W^\perp$. Par ailleurs, si u était un vecteur propre de $T_{|W^\perp}$ pour une valeur propre μ, ce serait un vecteur propre de T et on aurait $u \in V_\mu \subset W$, donc $u \in W \cap W^\perp = \{0\}$. Ceci contredit le lemme 3.5.4 si $W^\perp \neq \{0\}$. Finalement $W^\perp = \{0\}$ et W est dense. ■

Théorème 3.5.6. *Soit T un opérateur normal compact sur l'espace de Hilbert E de dimension infinie. Il existe une suite $(V_n)_{n\geqslant 1}$ de sous-espaces vectoriels de dimension finie deux-à-deux orthogonaux et une suite $(\lambda_n)_{n\geqslant 1}$ convergeant vers 0 (éventuellement finie) dans $\mathbb{C}$ telles que T soit limite en norme dans $\mathscr{L}(E)$ des sommes $\sum_{j=0}^n \lambda_j P_j$, où P_j désigne le projecteur orthogonal sur V_j.*
Le spectre de T est alors $\{\lambda_n : n \in \mathbb{N}\} \cup \{0\}$.

Démonstration : Il résulte du lemme 3.5.3 que les valeurs propres non nulles forment une suite $(\lambda_n)_{n\geqslant 1}$ convergeant vers 0 (ou finie), et que les sous-espaces propres correspondants V_n sont de dimension finie et deux-à-deux orthogonaux. Notons P_n le projecteur orthogonal sur V_n. On a en particulier $P_m.P_n = P_n.P_m = 0$ si $m \neq n$.
Si on note $S_n = \sum_{j=1}^n \lambda_j P_j$, on a $S_{n+p} - S_n = \sum_{j=n+1}^{n+p} \lambda_j P_j$. Pour $x \in E$ tel que $\|x\| \leqslant 1$, on a

$$\|(S_{n+p} - S_n)x\|^2 = \sum_{j=n+1}^{n+p} |\lambda_j|^2 \|P_j x\|^2 \leqslant \sup_{j\geqslant n} |\lambda_j|^2 . \sum_{j\leqslant n+p} \|P_j x\|^2 = \sup_{j\geqslant n} |\lambda_j|^2 . \left\| \sum_{j\leqslant n+p} P_j x \right\|^2$$

On vérifie sans peine que $\sum_{j\leqslant n+p} P_j$ est le projecteur orthogonal sur le sous-espace (de dimension finie) $\bigoplus_{1\leqslant j\leqslant n+p} V_j$; il en résulte que $\left\|\sum_{1\leqslant j\leqslant n+p} P_j\right\| \leqslant 1$. Et puisque la suite (λ_n) tend vers 0, il existe pour tout ε un N tel que $|\lambda_j| \leqslant \varepsilon$ si $j \geqslant N$. On en déduit que, si $n \geqslant N$, on a $\|(S_{n+p} - S_n)x\| \leqslant \varepsilon \|x\|$. La suite (S_n) d'opérateurs de rang fini est donc une suite de Cauchy dans $\mathscr{L}(E)$, qui converge vers un opérateur compact S.

Pour tout m et tout $x \in V_m$ on a $P_j x = 0$ si $j \neq m$: il en résulte que $S_n x \to \lambda_m P_m x = \lambda_m x$, c'est-à-dire que $Sx = \lambda_m x = Tx$. Les opérateurs S et T coïncident donc sur l'espace vectoriel engendré par les V_j.
Et si $x \in \ker T$, on a $P_j x = 0$ pour tout x, donc $Sx = 0$. Il en résulte que S et T coïncident sur la somme des sous-espaces propres, qui est dense en vertu du lemme précédent. Donc $S = T$.

Si E est de dimension infinie, l'opérateur compact T ne peut être inversible : donc 0 appartient au spectre X de T. Et chacune des valeurs propres est également dans X. Il reste à montrer que tout nombre $\mu \neq 0$ qui n'est pas valeur propre est dans $\mathbb{C} \setminus X$. L'ensemble $X_0 = \{\lambda_n : n \geqslant 1\} \cup \{0\}$ est compact. Si donc $\mu \notin X_0$, il existe $\eta > 0$ tel que $|\mu - \lambda_n| \geqslant \eta$ pour tout n. Notons $V_0 = \ker T$ et P_0 le projecteur orthogonal sur $\ker T$. Alors $\sum_{0 \leqslant j \leqslant n} P_j$ est le projecteur orthogonal sur $W_n = \bigoplus_{0 \leqslant j \leqslant n} V_j$; il en résulte que, pour tout x de E on a

$$\sum_{j \leqslant n} \|P_j x\|^2 \leqslant \|x\|^2$$

De plus, puisque la somme des sous-espaces propres est dense, il existe pour tout x et tout ε un n tel que $d(x, W_n) < \varepsilon$: on a alors

$$\|x\|^2 - \sum_{j \leqslant n} \|P_j x\|^2 = d(x, W_n)^2 < \varepsilon^2$$

On en déduit que $\sum_{j=0}^{\infty} \|P_j x\|^2 = \|x\|^2$ et que $\left\| x - \sum_{j=0}^{n} P_j x \right\|^2 = \|x\|^2 - \sum_{j=0}^{n} \|P_j x\|^2 \to 0$.
Soit alors $y \in E$. Si on pose $z_j = \dfrac{1}{\lambda_j - \mu} P_j y$, on vérifie sans peine que $(T - \mu I) z_j = P_j y$.
Alors, puisque les (z_j) sont deux-à-deux orthogonaux et que

$$\sum_j \|z_j\|^2 \leqslant \sum_j \eta^{-2} \|P_j y\|^2 \leqslant \eta^{-2} \|y\|^2 < +\infty \ ,$$

on voit que la série (z_j) converge vers un élément z de E et on a

$$(T - \mu I) z = \lim_{n \to \infty} \sum_{j=0}^{n} P_j y = y$$

Il résulte de ce qui précède que $T - \mu I$ est injectif puisque μ n'est pas valeur propre, surjectif puisque z est solution de $(T - \mu I) z = y$, et que $\left\|(T - \mu I)^{-1}\right\| \leqslant \eta^{-1}$ puisque $\|z\|^2 = \sum_j \|z_j\|^2 \leqslant \eta^{-2} \|y\|^2$. Ainsi $\mu \notin X$, et $X = X_0$. ■

Théorème 3.5.7. *Soit T un opérateur normal compact sur l'espace de Hilbert E. Pour tout nombre complexe $\lambda \neq 0$, il y a équivalence entre*

i) $T - \lambda I$ *est injectif,*

ii) $T - \lambda I$ *est surjectif,*

iii) $T^* - \bar{\lambda} I$ *est injectif,* *iv)* $T^* - \bar{\lambda} I$ *est surjectif.*

Démonstration : Il résulte du théorème 3.5.6 que, pour $\lambda \neq 0$, $T - \lambda I$ est inversible si et seulement si λ n'est pas valeur propre, c'est-à-dire si $T - \lambda I$ est injectif. Donc

$$T - \lambda I \text{ injectif} \Longrightarrow T - \lambda I \text{ bijectif} \Longrightarrow T - \lambda I \text{ surjectif}$$

On en déduit que *i)* $\Longrightarrow$ *ii)*, et, puisque T^* est aussi normal et compact (cf. Théorème 3.3.4), que *iii)* $\Longrightarrow$ *iv)*.

De plus, si $T - \lambda I$ est surjectif, $T^* - \bar{\lambda} I$ est injectif : en effet, si $(T^* - \bar{\lambda} I)x = 0$, on a pour tout y : $\langle Ty - \lambda y, x \rangle = \langle y, T^*x - \bar{\lambda} x \rangle = 0$, en particulier pour un y tel que $(T - \lambda I)y = x$, ce qui montre que $x = 0$.

Ceci montre que *ii)* $\Longrightarrow$ *iii)*, et par application à T^*, que *iv)* $\Longrightarrow$ *i)*. ■

4

ESPACES LOCALEMENT CONVEXES

Il s'avère que, pour étudier les notions de convergence dans nombre d'espaces vectoriels qui s'introduisent naturellement en Analyse, et tout particulièrement dans l'étude des distributions qui possède une place centrale dans ce cours, le cadre des espaces normés est insuffisant. On va donc présenter maintenant une généralisation de ce cadre.

Comme dans le chapitre précédent, $\mathbb{K}$ désignera selon le contexte le corps $\mathbb{R}$ des réels ou celui $\mathbb{C}$ des complexes. Pour un élément λ de $\mathbb{K}$, la notation $|\lambda|$ désignera dans le premier cas la valeur absolue de λ, et dans le second son module.

Définition. *Soit E un $\mathbb{K}$-espace vectoriel. On appelle* semi-norme sur E *une fonction $p : E \to \mathbb{R}^+$ satisfaisant les deux conditions :*

(i) $\forall x \in E\ \forall \lambda \in \mathbb{K} \quad p(\lambda x) = |\lambda|\, p(x)$

(ii) $\forall x \in E\ \forall y \in E \quad p(x+y) \leqslant p(x) + p(y)$.

Une norme est donc une semi-norme satisfaisant pour tout x de E : $p(x) = 0 \Longrightarrow x = 0$.

4.1 Topologie définie par une famille de semi-normes.

Soient E un $\mathbb{K}$-espace vectoriel et P un ensemble de semi-normes sur E. On appellera P-boule de E centrée en $a \in E$ tout ensemble de la forme

$$W(a, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k) := \{x \in E : \forall j \leqslant k \quad p_j(x-a) < r_j\} ,$$

où les p_j appartiennent à P, et les r_j sont strictement positifs. Lorsque P est un ensemble constitué d'une seule norme, ces P-boules centrées en a sont exactement les boules ouvertes de centre a pour cette norme.

Théorème 4.1.1. *Si P est une famille de semi-normes sur E, il existe une unique topologie sur E (qu'on appellera la P-*topologie*) pour laquelle, pour tout $a \in E$, les P-boules de centre a forment une base de voisinages de a.*

Pour cette topologie, la fonction : $(x, y) \mapsto x + y$ est continue de $E \times E$ dans E, et la fonction $(\lambda, x) \mapsto \lambda x$ continue de $\mathbb{K} \times E$ dans E (on dira que c'est une topologie d'espace vectoriel*).*

DÉMONSTRATION : Il est clair que l'intersection de deux P-boules de centre a est encore une P-boule de centre a. De plus, si $W = W(a, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$ est une P-boule de centre a contenant b, les nombres $s_i = r_i - p_i(b - a)$ sont strictement positifs et on vérifie immédiatement que $W(b, p_1, p_2, \ldots, p_k, s_1, s_2, \ldots, s_k) \subset W(a, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$, donc que W est un voisinage de b. Il en résulte que ceci définit une topologie sur E.
Si $W = W(a + b, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$ est un voisinage de $a + b$, on a lorsque $x \in W(a, p_1, p_2, \ldots, p_k, r_1/2, r_2/2, \ldots, r_k/2)$ et $y \in W(b, p_1, p_2, \ldots, p_k, r_1/2, r_2/2, \ldots, r_k/2)$: $p_j(x + y - a - b) \leqslant p_j(x - a) + p_j(y - b) < r_j/2 + r_j/2 = r_j$, donc $x + y \in W$: ceci montre la continuité en (a, b) de la fonction $(x, y) \mapsto x + y$.
De même, puisque

$$p_j(\lambda x - \mu a) \leqslant |\lambda|\, p_j(x - a) + |\lambda - \mu|\, p_j(a) \leqslant |\mu|\, p_j(x - a) + |\lambda - \mu|\, p_j(a) + |\lambda - \mu|\, p_j(x - a) \;,$$

on voit qu'en assurant $p_j(x - a) < \inf(1, \dfrac{r_j}{3\mu})$ et $|\lambda - \mu| < \inf(\dfrac{r_j}{3}, \dfrac{r_j}{p_j(a)})$, on assure $p_j(\lambda x - \mu a) < r_j$. On en déduit la continuité en (μ, a) de la fonction $(\lambda, x) \mapsto \lambda x$. ■

Proposition 4.1.2. *Si P est une famille de semi-normes sur E, la P-topologie est séparée si et seulement si, pour tout $x \neq 0$ de E, il existe une semi-norme $p \in P$ telle que $p(x) > 0$.*

DÉMONSTRATION : Si la P-topologie est séparée et si x est un vecteur non nul dans E, il existe une P-boule $W(0, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$ de centre 0 ne contenant pas x : il existe donc un $j \leqslant k$ tel que $p_j(x) \geqslant r_j > 0$.
Inversement, si $x \neq y$ et s'il existe $p \in P$ telle que $p(x - y) = r > 0$, les P-boules $W_1 = W(x, p, r/2)$ et $W_2 = W(y, p, r/2)$ sont des voisinages disjoints de x et y : en effet, si on avait $z \in W_1 \cap W_2$, on aurait $r = p(x - y) \leqslant p(x - z) + p(y - z) < \dfrac{r}{2} + \dfrac{r}{2} = r$. ■

Définition 4.1.3. *Une partie A de l'espace vectoriel E est dite* équilibrée *si λx appartient à A pour tout x de A et tout λ tel que $|\lambda| \leqslant 1$.*

Théorème 4.1.4. *Si P est une famille de semi-normes sur E, chacune des fonctions $p \in P$ est continue pour la P-topologie, et toute P-boule est un ensemble convexe ouvert, équilibré si son centre est en 0.*

DÉMONSTRATION : On a en effet, pour $p \in P$:

$$p(x + h) \leqslant p(x) + p(h)$$

et $p(x) \leqslant p(x + h) + p(-h) = p(x + h) + p(h)$, donc $|p(x + h) - p(x)| \leqslant p(h)$. Il s'en suit que $|p(x + h) - p(x)| < \varepsilon$ dès que $x + h \in W(x, p, \varepsilon)$.
Soit $p \in P$. Puisque p est continue, l'ensemble $\{x : p(x - a) < r\}$ est ouvert. Donc toute P-boule, intersection finie de tels ouverts, est ouverte.
De plus, si $p \in P$, $p(x - a) < r$, $p(y - a) < r$ et $0 \leqslant t \leqslant 1$, on a

$$p((tx + (1 - t)y) - a) = p(t(x - a) + (1 - t)(y - a)) \leqslant tp(x - a) + (1 - t)p(y - a) < tr + (1 - t)r = r$$

et si $|\lambda| \leqslant 1$, on a $p(\lambda x) \leqslant p(x)$ pour tout x. Il en résulte que l'ensemble $\{x = p(x - a) < r\}$ est convexe, et équilibré si $a = 0$; toute P-boule, intersection de convexes, est donc convexe, et équilibrée si elle est centrée en 0. ■

Théorème 4.1.5. *Si P est une famille de semi-normes sur E, et q une semi-norme sur E. Alors il y a équivalence entre*

(i) q *est continue pour la P-topologie.*

(ii) *il existe un nombre fini $(p_1, p_2, \ldots, p_k)$ d'éléments de P et un nombre M tels que $q(x) \leqslant M.\sup_{j \leqslant k} p_j(x)$ pour tout x de E.*

(iii) q *est bornée sur un voisinage de 0.*

Démonstration : Si q est une semi-norme continue, il existe un voisinage V de 0 pour la P-topologie pour lequel on a $|q(x) - q(0)| < 1$ pour tout x de V. Il existe donc une P-boule $W(0, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$, de centre 0 sur laquelle q est majorée par 1.

Si maintenant q est bornée par μ sur un voisinage V de 0, il existe dans V une P-boule $W = W(0, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$, de centre 0 sur laquelle q est majorée par μ. Soit alors $x \in E$: si $0 < t < \inf_j \dfrac{r_j}{p_j(x)}$, on a $p_j(tx) = tp_j(x) < r_j$ pour tout j, donc $tx \in W$, et $q(tx) = tq(x) < \mu$, ou encore $q(x) < \dfrac{\mu}{t}$, c'est-à-dire $q(x) \leqslant \mu \sup_j \dfrac{1}{r_j} p_j(x)$. On peut donc prendre $M = \sup_j \dfrac{\mu}{r_j}$.

Enfin, si q vérifie *(ii)*, puisque l'on a

$$|q(y) - q(x)| \leqslant q(x-y) \leqslant M \sup_j p_j(x-y)$$

on a $|q(y) - q(x)| < \varepsilon$ en tout point y de $W(x, p_1, p_2, \ldots, p_k, \varepsilon/M, \varepsilon/M, \ldots, \varepsilon/M)$, ce qui montre la continuité de q en x. ■

Corollaire 4.1.6. *Soient P une famille de semi-normes sur E, Q une famille de semi-normes sur F et $f : E \to F$ une application linéaire. Alors f est continue de E muni de la P-topologie dans F muni de la Q-topologie si et seulement si $q \circ f$ est une semi-norme continue sur E pour tout $q \in Q$.*

Démonstration : Si f est continue et si $q \in Q$, alors q est continue sur F et $q \circ f$ est continue sur E. Inversement, si $W = W(f(a), q_1, q_2, \ldots q_k, r_1, r_2, \ldots, r_k$ est une Q-boule de E, on a $f^{-1}(W) = W(a, q_1 \circ f, \ldots q_k \circ f, r_1, \ldots r_k)$, qui est un voisinage de a dans E pour la P-topologie si chacune des semi-normes $q_j \circ f$ est continue. ■

On notera $\mathscr{L}(E, F)$ l'espace vectoriel des applications linéaires continues de E dans F.

Corollaire 4.1.7. *Si P est une famille de semi-normes sur E, et $f : E \to \mathbb{K}$ une forme linéaire, alors f est continue pour la P-topologie si et seulement s'il existe un nombre fini $(p_1, p_2, \ldots, p_k)$ d'éléments de P et un nombre M tels que, pour tout $x \in E$, on ait $|f(x)| \leqslant M.\sup_{j \leqslant k} p_j(x)$.*

Démonstration : Il suffit de remarquer que, lorsque f est une forme linéaire, la fonction $q : x \mapsto |f(x)|$ est une semi-norme, que q est continue si f l'est, et que si q est continue, on a

$$|f(x) - f(y)| = |f(x-y)| = q(x-y) < \varepsilon$$

en tout point y d'un voisinage de x. ■

Si E est un $\mathbb{K}$-espace vectoriel muni d'une P topologie, on appelle *dual* de E l'espace vectoriel $E' = \mathscr{L}(E, \mathbb{K})$ des formes linéaires continues sur E.

Proposition 4.1.8. *Soient P une famille de semi-normes sur E, Q une famille de semi-normes sur F et $u : E \to F$ une application linéaire continue de E muni de la P-topologie dans F muni de la Q-topologie. Alors il existe une application linéaire ${}^t u$ du dual F' de F dans le dual E' de E, appelée* transposée *de u telle que, pour tout $x \in E$ et tout $f \in F'$ on ait*

$$\langle {}^t u(f), x\rangle = \langle f, u(x)\rangle \ .$$

Démonstration : Il suffit de remarquer que si $f \in F'$ la forme linéaire $f \circ u$ est continue sur E, et que l'application $f \mapsto f \circ u$ est linéaire de F' dans E'. ■

Théorème 4.1.9. *Si P est une famille de semi-normes sur E, alors E est localement compact pour la P-topologie si et seulement s'il est séparé et de dimension finie.*

Démonstration : Ceci résulte des deux lemmes suivants.

Lemme 4.1.10. *Si P est une famille de semi-normes sur un espace E de dimension finie pour laquelle la P-topologie est séparée, et si $(a_1, a_2, \ldots, a_n)$ est une base de E, l'application $\varphi : (x_1, x_2, \ldots, x_n) \mapsto \sum x_j a_j$ est un homéomorphisme de $\mathbb{K}^n$ sur E.*

Démonstration : Puisque chaque $p \circ \varphi$, pour $p \in P$, est une semi-norme sur $\mathbb{K}^n$, on se ramène immédiatement au cas où $E = \mathbb{K}^n$, qu'on munit de la norme $\|x\| = \sum_j |x_j|$ si $x = (x_1, x_2, \ldots, x_n)$.
On remarque d'abord que toute semi-norme p sur $\mathbb{K}^n$ est continue : si $(e_1, e_2, \ldots, e_n)$ est la base canonique de $\mathbb{K}^n$, on a

$$p(x) = p(\sum_{j=1}^{n} x_j e_j) \leqslant \sum_j |x_j|\, p(e_j) \leqslant M\, \|x\|$$

où on a posé $M = \sum_j p(e_j)$, donc $|p(y) - p(x)| \leqslant p(y - x) \leqslant M\, \|y - x\|$. Il en résulte que toute P-boule est ouverte, pour la topologie définie par la norme.
Inversement, pour toute semi-norme p, l'ensemble $p^{-1}(0)$ est un sous-espace vectoriel. Si la P-topologie est séparée, on peut donc construire par récurrence une suite strictement décroissante (V_k) de sous-espaces vectoriels de E et des p_k dans P telle que $V_0 = E$ et $V_{k+1} = V_k \cap p_k^{-1}(0)$ où $p_k(x_k) > 0$ pour un $x_k \in V_k$ si $V_k \neq \{0\}$. Cette suite est finie, de longueur $\ell \leqslant n$, Alors $\sup_{k \leqslant \ell} p_k$ est une semi-norme ne s'annulant qu'en 0, donc une norme, qui est équivalente à la norme $\|.\|$ puisque la dimension est finie : on a $\|x\| \leqslant M \sup_{k \leqslant \ell} p_k(x)$, ce qui montre que la norme $\|.\|$ est continue pour la P-topologie, c'est-à-dire que cette dernière est plus fine que la topologie usuelle sur $\mathbb{K}^n$. ■

Lemme 4.1.11. *Si E est localement compact pour la P-topologie, E est de dimension finie.*

Démonstration : Supposons que E soit localement compact : il existe alors une P-boule $W(0, p_1, p_2, \ldots, p_k, r_1, r_2, \ldots, r_k)$ qui est relativement compacte dans E. Soit $a \neq 0$ un point de cette P-boule W. Il existe, puisque E est séparé une $p \in P$ telle que $p(a) > 0$; comme p est continue, elle est bornée sur le compact $\overline{W}$ par un nombre M, ce qui montre que $na \notin \overline{W}$ si $n > \dfrac{M}{p(a)}$, donc qu'il existe $j \leqslant k$ telle que $np_j(a) = p_j(na) \geqslant r_j > 0$, et enfin que $p_j(a) > 0$. Il en résulte que $q = \sup_{j \leqslant k} p_j$ est une norme sur E, et que W est un voisinage de 0 dans l'espace E normé par q. La topologie définie par q est moins fine que la P-topologie, et séparée : $\overline{W}$ est donc compact dans cet espace normé. Celui-ci est donc localement compact, donc de dimension finie (cf. 2.1.21). ■

Théorème 4.1.12. *Si P est une famille dénombrable de semi-normes sur E et si la P-topologie est séparée, cette topologie peut être définie par une distance d invariante par translation, c'est-à-dire vérifiant pour tout x, tout y et tout z de E : $d(x+z, y+z) = d(x, y)$, pour laquelle les boules de centre 0 sont convexes et équilibrées.*

Inversement, si P est une famille de semi-normes et si la P-topologie est métrisable, il existe une suite croissante (p_k) de semi-normes pour laquelle la P-topologie coïncide avec la topologie définie par la famille des (p_k).
Une suite (x_n) de E converge alors vers $a \in E$ si et seulement si $p_k(x_n - a) \to 0$ pour tout k. Si la distance est invariante par translation, une suite (x_n) est de Cauchy si et seulement si, pour tout k, $p_k(x_n - x_m) \to 0$, c'est-à-dire si $(x_n - x_m)$ tend vers 0 lorsque n et m tendent vers l'infini.

Démonstration : Si $(p_k)_{k\geqslant 1}$ est une énumération de P, on définit une distance d sur E en posant :

$$d(x,y) = \sup_{k\geqslant 1}\bigl(\inf(\frac{1}{k}, p_k(x-y)\bigr)$$

(noter que ceci est une distance car $p_k(x-y)$ n'est nul pour tout k que si $x - y = 0$) et cette distance est clairement invariante par translation. On peut remarquer également que si $\lambda \in \mathbb{K}$ vérifie $|\lambda| \leqslant 1$, on a $d(\lambda x, 0) \leqslant d(x, 0)$. On a, pour cette distance

$$B(a,r) = \{x : \forall k \leqslant \frac{1}{r} \quad p_k(x-a) < r\}$$

qui est une P-boule centrée en a, ce qui montre que la P-topologie est plus fine que la topologie associée à d et que les boules de d sont convexes, et équilibrées si elles sont centrées en 0. Inversement, chacune des semi-normes p_k est continue pour la distance d, puisque, pour $\varepsilon < \dfrac{1}{k}$, on a, si $d(x,y) < \varepsilon$:

$$|p_k(y) - p_k(x)| \leqslant p_k(y-x) = \inf(\frac{1}{k}, p_k(x-y)) \leqslant d(x,y) < \varepsilon\ .$$

Si, maintenant, la P-topologie est métrisable, définie par une distance d, chaque boule $B_d(0, 2^{-k})$ est un voisinage de 0, donc contient une P-boule ; il existe donc une partie finie J_k de P et un ε_k tels que $B_d(0, 2^{-k}) \supset \{x : \forall p \in J_k \quad p(x) < \varepsilon_k\}$. Si (q_m) est une énumération de l'ensemble dénombrable $P_0 = \bigcup_k J_k$, on remarque que chaque q_m est continue, et que la P_0-topologie est plus fine que la P-topologie initiale. Il suffit alors de prendre $p_k = \sum_{m\leqslant k} q_m$ pour avoir la suite croissante cherchée.

Il résulte de ce qui précède qu'une suite (x_n) converge alors vers a si $p_k(x_n - a) \to 0$ pour tout k. Si d est invariante par translation, on a $d(x_m, x_n) = d(x_m - x_n, 0)$: une suite (x_n) est alors de Cauchy pour d si et seulement si, pour tout k et tout $\varepsilon > 0$, il existe un rang N tel que $p_k(x_m - x_n) < \varepsilon$ pour m et n supérieurs à N, ou encore que $(x_m - x_n)$ tende vers 0 quand m et n tendent vers l'infini. ■

On remarque ainsi que la complétude de E ne dépend alors que de la topologie, et pas de la distance invariante par translation choisie.

Définition 4.1.13. *On appelle* espace de Fréchet *un espace vectoriel muni d'une P-topologie métrisable, et complet pour cette topologie.*

4.2 Topologies faibles

Soient E un espace vectoriel sur $\mathbb{K}$ et F un espace vectoriel de formes linéaires sur E. Pour tout $f \in F$, la fonction $p_f : x \mapsto |f(x)|$ est une semi-norme sur E. On définit alors la topologie faible $\sigma(E, F)$ sur E comme la P-topologie où $P = \{p_f : f \in F\}$. On notera souvent $\langle f, x\rangle$ le nombre $f(x)$. L'application $(f, x) \mapsto \langle f, x\rangle$ est clairement bilinéaire de $E \times F$ dans $\mathbb{K}$.

Théorème 4.2.1. *Une forme linéaire f sur E est continue pour la topologie $\sigma(E, F)$ si et seulement si $f \in F$.*

Démonstration : Il résulte immédiatement de la définition de la semi-norme p_f que si $f \in F$, on a $|f(x)| \leqslant p_f(x)$ pour tout $x \in E$, donc que f est continue pour $\sigma(E, F)$.
Inversement, si g est une forme linéaire continue pour $\sigma(E, F)$, il existe f_1, f_2, $\dots f_k$ dans F et M tels que, pour tout x, on ait $|g(x)| \leqslant M. \sup_j p_{f_j}(x)$. En particulier, si $x \in \bigcap_j \ker(f_j)$, on a nécessairement $g(x) = 0$. Si on note Φ l'application linéaire de E dans $\mathbb{K}^k$ définie par $\Phi(x) = (f_j(x))_{j\leqslant k}$, et par G le sous-espace $\Phi(E)$ de $\mathbb{K}^k$, on remarque que si $\Phi(x) = \Phi(y)$, on a $f_j(x-y) = 0$ pour tout j, donc $g(x-y) = 0$. Il en résulte qu'existe une unique fonction φ de G dans $\mathbb{K}$ telle que $g(x) = \varphi(\Phi(x))$, dont on voit aisément qu'elle est linéaire. Si H est un supplémentaire de G dans $\mathbb{K}^k$, et π la projection sur G parallèlement à H, $\psi = \varphi\circ\pi$ est une forme linéaire sur $\mathbb{K}^k$ qui prolonge φ. On a alors

$$\psi(u_1, u_2, \dots, u_k) = \sum_{j\leqslant k} \alpha_j u_j$$

avec des coefficients α_j dans $\mathbb{K}$. On en déduit que, si $x \in E$, $g(x) = \psi\circ\Phi(x) = \sum_{j\leqslant k} \alpha_j f_j(x)$, donc que $g = \sum_{j\leqslant k} \alpha_j f_j \in F$. ∎

Théorème 4.2.2. *Si E est un espace vectoriel muni d'une P-topologie et E' son dual, la topologie faible $\sigma(E, E')$ (appelée* topologie affaiblie de E*) est moins fine que la P-topologie initiale.*

Démonstration : Puisque chaque élément de E' est continu pour la P-topologie initiale $\mathscr{O}$, on voit immédiatement que chaque semi-norme p_f pour $f \in E'$ est $\mathscr{O}$-continue, donc que toute boule ouverte pour la topologie $\sigma(E, E')$ est ouverte pour $\mathscr{O}$, c'est-à-dire que la topologie affaiblie est moins fine que $\mathscr{O}$. ∎

Théorème 4.2.3. *Si P est une famille de semi-normes sur E, Q une famille de semi-normes sur F et u une application linéaire continue de E muni de la P-topologie dans F muni de la Q-topologie, l'application ${}^t u$ transposée de u est continue de F' muni de $\sigma(F', F)$ dans E' muni de $\sigma(E', E)$.*

Démonstration : Il suffit pour cela de montrer que pour tout $x \in E$ la semi-norme p_x sur E' définie par $p_x(g) = |\langle g, x\rangle|$ satisfait que $p_x\circ{}^t u$ est continue sur F' pour $\sigma(F', F)$. Or $p_x\circ{}^t u(f) = q_y(f)$, où $y = u(x)$ et q_y est la semi-norme $f \mapsto |\langle f, y\rangle|$, qui est continue par définition de la topologie $\sigma(F', F)$. ∎

4.3 Jauge d'un ensemble convexe

Définition 4.3.1. *Une fonction $p : E \to \mathbb{R}$ est dite* sous-linéaire *si elle vérifie les deux conditions :*
(i) $\forall x \in E\ \forall t \in \mathbb{R}^+ \quad p(tx) = tp(x)$
(ii) $\forall x \in E\ \forall y \in E \quad p(x+y) \leqslant p(x) + p(y)$.

Une semi-norme est donc une fonction sous-linéaire satisfaisant en plus $p(\lambda x) = p(x)$ pour tout $\lambda \in \mathbb{K}$ tel que $|\lambda| = 1$.

Lemme 4.3.2. *Une fonction sous-linéaire est convexe.*

Démonstration : Si x et y sont des vecteurs de E , $t \in [0,1]$ et $z = tx + (1-t)y$, on a

$$p(z) \leqslant p(tx) + p((1-t)y) = tp(x) + (1-t)p(y)$$

ce qui montre le résultat. ■

Théorème 4.3.3. *Soient E un $\mathbb{K}$-espace vectoriel et C une partie convexe de E contenant l'origine. On suppose que C est absorbant, c'est-à-dire que pour tout x de E on peut trouver un $t > 0$ tel que $tx \in C$. Alors il existe une fonction sous-linéaire $p : E \to \mathbb{R}^+$ appelée* jauge *de C telle que, pour tout x de E, on ait*

$$p(x) < 1 \Longrightarrow x \in C \Longrightarrow p(x) \leqslant 1 \ .$$

Démonstration : On pose, pour $x \in E$:

$$p(x) = \inf\{s > 0 : s^{-1}x \in C\} \ ,$$

qui est bien défini dans $\mathbb{R}^+$ puisqu'il existe des $s > 0$ tels que $s^{-1}x \in C$. Si $\lambda > 0$ et $y = \lambda x$, on a $t^{-1}y \in C \iff (t^{-1}\lambda)x \in C$, donc $p(y) = \lambda p(x)$.
Soient x et y dans E, avec $p(x) = \alpha$ et $p(y) = \beta$. On veut montrer que $p(x+y) \leqslant \alpha + \beta$. Soit donc $s > \alpha + \beta$. On veut trouver $s_1 < s$ que $s_1^{-1}(x+y) \in C$; il existe $u > \alpha$ et $v > \beta$ tels que $s = u + v$, donc $u_1 < u$ tel que $x' = u_1^{-1}x \in C$ et $v_1 < v$ tel que $y' = v_1^{-1}y \in C$. Alors, en posant $s_1 = u_1 + v_1$,

$$s_1^{-1}x = \frac{1}{u_1 + v_1}(u_1x' + v_1y') \in C \ ,$$

ce qui achève la preuve.

Si $x \in C$, le nombre 1 appartient à $\{s : s^{-1}x \in C\}$. Donc $p(x) \leqslant 1$. Et si $p(x) < 1$, il existe $s \in]p(x), 1[$ tel que $s^{-1}x \in C$. Alors $x = s.(s^{-1}x) + (1-s).0 \in C$, puisque C est convexe et contient 0 et $s^{-1}x$. ■

Définition 4.3.4. *On appelle* espace vectoriel topologique *un $\mathbb{K}$-espace vectoriel muni d'une topologie $\mathscr{O}$ pour laquelle la fonction : $(x, y) \mapsto x + y$ est continue de $E \times E$ dans E, et la fonction $(\lambda, x) \mapsto \lambda x$ continue de $\mathbb{K} \times E$ dans E.*

Théorème 4.3.5. *Soient E un espace vectoriel topologique. On suppose que, pour tout voisinage V de 0 dans E, il existe un voisinage* convexe *W de 0 contenu dans V. Alors,*

il existe une famille P de semi-normes sur E telle que la topologie de E soit égale à la P-topologie.
Un tel espace est appelé espace localement convexe.

Démonstration : On définit P comme l'ensemble de toutes les semi-normes continues sur $(E, \mathscr{O})$. Les P-boules sont alors ouvertes pour $\mathscr{O}$, c'est-à-dire que la P-topologie est moins fine que $\mathscr{O}$.
Inversement, si V est un voisinage de 0 pour $\mathscr{O}$, il existe par hypothèse un voisinage convexe V_0 de 0 pour $\mathscr{O}$, contenu dans V. On définit alors $W = \bigcap_{|\lambda|=1} \lambda V_0$, qui est une partie convexe de E contenant 0. Pour $a \in E$, la continuité de l'application $\mu \mapsto \mu a$ entraîne l'existence d'un $r > 0$ tel que $\mu a \in V_0$ si $|\mu| \leqslant r$, donc que $ra \in W$. On note alors p la jauge de W. Puisque $\lambda W = W$ pour $|\lambda| = 1$, on voit que $s^{-1}(\lambda x) \in W \iff s^{-1}x \in W$, donc que $p(\lambda x) = p(x)$ si $x \in E$ et $|\lambda| = 1$. Ceci montre que p est une semi-norme.
On vérifie maintenant que p est continue : il suffit pour cela de montrer que W est un voisinage de 0. En effet, par continuité de l'application $(\lambda, x) \mapsto \lambda x$ en $(0, 0)$, il existe un voisinage V_1 de 0 et $\rho > 0$ tels que $\lambda x \in V_0$ si $x \in V_1$ et $|\lambda| \leqslant \rho$. L'homothétie $x \mapsto \rho x$ est alors un homéomorphisme de E sur lui-même, et $V_2 = \rho V_1$ est un voisinage de 0. On a alors, pour $x \in V_2$ et $|\lambda| = 1$: $\lambda x \in V_0$, ce qui montre que $W \supset V_2$ est un voisinage de 0, que $p \in P$ et que V est un voisinage de 0 pour la P-topologie. ■

4.4 Le théorème de Hahn-Banach

Théorème 4.4.1. *Soient E un $\mathbb{R}$-espace vectoriel, p une fonction sous-linéaire sur E, V un sous-espace vectoriel de E et f une forme linéaire sur V satisfaisant $f(x) \leqslant p(x)$ pour tout x de V. Alors il existe une forme linéaire g sur E qui prolonge f et qui satisfait $f(x) \leqslant p(x)$ pour tout x de E.*

Démonstration : Désignons par $\mathscr{V}$ l'ensemble des couples (M, φ), où M est un sous-espace vectoriel de E et φ une forme linéaire sur M vérifiant $\varphi(x) \leqslant p(x)$ pour tout x de M. On munit $\mathscr{V}$ de l'ordre $\leqslant$ défini par $(M, \varphi) \leqslant (N, \psi) \iff M \subset N$ et $\psi_{|M} = \varphi$.
On montre d'abord que $\mathscr{V}$ est inductif pour $\leqslant$: soit en effet $(M_i, \varphi_i)_{i \in I}$ une famille totalement ordonnée. Si on pose $M = \bigcup_{i \in I} M_i$, il est aisé de vérifier que M est un sous-espace vectoriel de E, qu'il existe une fonction $\varphi : M \to \mathbb{R}$ qui prolonge chacune des φ_i et que φ est linéaire. De plus, pour tout $x \in M$, il existe un $i \in I$ tel que $x \in M_i$ et on a $\varphi(x) = \varphi_i(x) \leqslant p(x)$. Donc (M, φ) appartient à $\mathscr{V}$ et est dans $\mathscr{V}$ la borne supérieure de la famille $(M_i, \varphi_i)_{i \in I}$.

D'après le théorème de Zorn, l'élément (V, f) de $\mathscr{V}$ est majoré dans $\mathscr{V}$ par un élément maximal (M, g). Il suffit donc de démontrer que cet élément maximal satisfait $M = E$ pour achever la démonstration. Supposons donc par l'absurde qu'existe un $a \in E \setminus M$. On va construire alors un élément (N, φ) de $\mathscr{V}$ qui majore strictement (M, g). On pose pour cela $N = M \oplus \mathbb{R}.a$, on choisit $\alpha \in \mathbb{R}$ et on définit la forme linéaire φ sur N par $\varphi(x + ta) = g(x) + t\alpha$. Il suffit donc de montrer qu'on peut choisir α de sorte que l'on ait $g(x) + t\alpha \leqslant p(x + ta)$ pour tout x de M et tout $t \in \mathbb{R}$.
La condition précédente est satisfaite pour $t = 0$ quel que soit $x \in M$. Pour $t > 0$ cette condition est équivalente à la condition $g(\dfrac{x}{t}) + \alpha \leqslant p(\dfrac{x}{t} + a)$, et puisque $\dfrac{x}{t} \in M$, équivalente à $\alpha \leqslant p(y + a) - g(y)$ pour tout y de M. Enfin, pour $t < 0$, la condition précédente est

équivalente à $g(-\frac{x}{t}) - \alpha \leqslant p(-\frac{x}{t} - a)$, c'est-à-dire à $\alpha \geqslant -p(y-a) + g(y)$ pour tout y de M puisque $-\frac{x}{t} \in M$. On doit donc choisir α tel que

$$\sup_{y \in M} g(y) - p(y-a) \leqslant \alpha \leqslant \inf_{x \in M} p(x+a) - g(x) \ ,$$

ce qui est possible si, pour tout y et tout x de M, on a $g(y) - p(y-a) \leqslant p(x+a) - g(x)$. Or, pour x et y dans M,

$$\begin{aligned}(p(x+a) - g(x)) - (g(y) - p(y-a)) &= p(x+a) + p(y-a) - g(x+y) \\ &\geqslant p(x+a+y-a) - g(x+y) \\ &= p(x+y) - g(x+y) \geqslant 0\end{aligned}$$

puisque $x + y \in M$. Ceci montre que le choix de α est possible, et que (N, φ) majore strictement (M, g), contrairement à la maximalité de ce dernier. Donc $M = E$ et g est une forme linéaire sur E. ■

Corollaire 4.4.2. *Soient E un $\mathbb{K}$-espace vectoriel, p une semi-norme sur E, V un sous-espace vectoriel de E et f une forme linéaire sur V satisfaisant $|f(x)| \leqslant p(x)$ pour tout x de V. Alors il existe une forme linéaire g sur E prolongeant f et satisfaisant $|g(x)| \leqslant p(x)$ pour tout x de E.*

Démonstration : Si $\mathbb{K} = \mathbb{R}$, il suffit de remarquer que la semi-norme p est sous-linéaire et d'appliquer le théorème précédent pour obtenir une forme linéaire g sur E prolongeant f et vérifiant $g(x) \leqslant p(x)$ pour tout x de E. Mais on a alors aussi $-g(x) = g(-x) \leqslant p(-x) = p(x)$, dont on déduit $|g(x)| \leqslant p(x)$ pour $x \in E$.
Si $\mathbb{K} = \mathbb{C}$, la fonction $\tilde{f} : x \mapsto \Re e(f(x))$ est $\mathbb{R}$-linéaire sur V et vérifie, pour tout x de V : $\tilde{f}(x) \leqslant |f(x)| \leqslant p(x)$. Le théorème précédent permet alors de trouver une forme $\mathbb{R}$-linéaire φ sur E prolongeant $\tilde{f}$ et vérifiant $\varphi(x) \leqslant p(x)$ pour tout x de E. On pose alors

$$g(x) = \varphi(x) - i\varphi(ix) \ ,$$

ce qui définit une application $\mathbb{R}$-linéaire de E dans $\mathbb{C}$ telle que $\Re e(g) = \varphi$. Mais on a

$$g(ix) = \varphi(ix) - i\varphi(-x) = i\varphi(x) + i(-i\varphi(ix)) = ig(x) \ ,$$

ce qui montre que g est en fait $\mathbb{C}$-linéaire. Pour $v \in V$, on a $iv \in V$, donc

$$g(v) = \varphi(v) - i\varphi(iv) = \tilde{f}(v) - i\tilde{f}(iv) = \Re e(f(v)) - i\,\Re e(if(v)) = \Re e(f(v)) + i\,\Im m(f(v)) = f(v)$$

ce qui montre que g prolonge f. Enfin, si $x \in E$ et $g(x) = |g(x)|\,\mathrm{e}^{i\theta}$, on a

$$|g(x)| = g(x)\,\mathrm{e}^{-i\theta} = \Re e(g(x)\,\mathrm{e}^{-i\theta}) = \Re e(g(x\,\mathrm{e}^{-i\theta})) = \varphi(x\,\mathrm{e}^{-i\theta}) \leqslant p(x\,\mathrm{e}^{-i\theta}) = p(x) \ ,$$

ce qui achève la démonstration. ■

Corollaire 4.4.3. *Soient E un espace normé et $a \neq 0$ un point de E. Alors il existe une forme linéaire g sur E de norme 1 telle que $g(a) = \|a\|$.*

Démonstration : Si on considère le sous-espace $V = \mathbb{K}.a$ et la forme linéaire f sur V définie par $f(\lambda a) = \lambda\,\|a\|$, on a $|f(x)| \leqslant \|x\|$ pour tout x de V. Il existe donc, d'après le corollaire 4.4.2 une forme linéaire g sur E prolongeant f et vérifiant $|g(x)| \leqslant \|x\|$ pour tout x de E, c'est-à-dire $\|g\| \leqslant 1$. Puisque $g(a) = \|a\|$, on a aussi $\|g\| \geqslant 1$. Donc $\|g\| = 1$. ■

Théorème 4.4.4. *Soient E un $\mathbb{K}$-espace vectoriel muni d'une P-topologie, C une partie convexe fermée non vide de E et a un point de E n'appartenant pas à C. Alors il existe une forme linéaire continue g sur E et un nombre réel $\alpha < \Re e(g(a))$ tels que $\Re e(g(x)) \leqslant \alpha$ pour tout x de C.*

Démonstration : Puisque a n'est pas adhérent à C, on peut trouver une P-boule $W = W(0, p_1, \dots, p_k, r_1, \dots r_k)$ telle que $a + W = W(a, p_1, \dots, p_k, r_1, \dots r_k)$ soit disjointe de C. Alors, puisque $W = -W$ et que l'ensemble convexe

$$C' = C - W = \bigcup_{c \in C} c - W = \bigcup_{c \in C} W(c, p_1, \dots, p_k, r_1, \dots r_k)$$

est une réunion d'ouverts de E, on voit que C' est ouvert. Le point a n'appartient pas à C' : sinon il existerait $c \in C$ et $w \in W$ tels que $a = c - w$, donc que $c = a + w \in C \cap (a + W) = \emptyset$.

Soit b un point de C. On a $b + W \subset C'$, donc $W \subset -b + C'$ et $b \in C'$, donc $a - b \notin -b + C'$. La jauge p de $C'' = -b + C'$ est une fonction sous-linéaire qui vérifie donc $p(x) \leqslant 1$ pour $x \in W$ et $p(a-b) \geqslant 1$. En appliquant le théorème 4.4.1 avec $M = \mathbb{R}.(a-b)$ et f la forme $\mathbb{R}$-linéaire sur M définie par $f(t(b-a)) = t$, on trouve donc une forme $\mathbb{R}$-linéaire φ sur E vérifiant $\varphi(b-a) = 1$ et $\varphi(x) \leqslant p(x)$ pour tout $x \in E$. Si $\mathbb{K} = \mathbb{R}$, on prend $g = \varphi$. Si $\mathbb{K} = \mathbb{C}$, on voit comme dans la démonstration du corollaire 4.4.2 que la fonction $g : x \mapsto \varphi(x) - i\varphi(ix)$ est alors une forme $\mathbb{C}$-linéaire vérifiant $\Re e(g) = \varphi$. Et on a dans tous les cas $\Re e(g(a-b)) = 1$.

Puisque $\varphi(x) \leqslant p(x) \leqslant 1$ pour $x \in W$, la fonction φ est continue sur E, de même que la fonction $x \mapsto i\varphi(ix)$ dans le cas $\mathbb{K} = \mathbb{C}$. Il en résulte que g est continue.
Puisque la fonction $t \mapsto t(a-b)$ est continue, il existe $\varepsilon > 0$ tel que $\varepsilon(a-b) \in W$. Pour tout $x \in C$, on a alors $x + \varepsilon(a-b) \in C'$, donc $x - b + \varepsilon(a-b) \in C''$. Il en résulte que

$$\Re e(g(x) - \Re e(g(b)) + \varepsilon = \Re e(g(x) - \Re e(g(b)) + \varepsilon \, \Re e(g(a-b)) = \Re e(g(x - b + \varepsilon(a-b))) \leqslant 1$$

ainsi que $\Re e(g(a)) - \Re e(g(b)) = 1$. Donc, avec $\alpha = \Re e(g(b)) + 1 - \varepsilon$, on a $\Re e(g(x)) \leqslant \alpha$ pour $x \in C$ et $\Re e(g(a)) > \alpha$. ■

Corollaire 4.4.5. *Soient E un $\mathbb{K}$-espace vectoriel muni d'une P-topologie, C une partie convexe fermée équilibrée de E (c'est-à-dire que $\lambda x \in C$ chaque fois que $x \in C$ et $|\lambda| \leqslant 1$) et a un point de E n'appartenant pas à C. Alors il existe une forme linéaire continue f sur E telle que $|f(a)| > \sup_{x \in C} |f(x)|$.*

En particulier, si V est un sous-espace vectoriel de E qui n'est pas partout dense, il existe une forme linéaire continue non nulle sur E, qui s'annule sur V.

Démonstration : Si on applique le théorème précédent, on trouve une forme linéaire continue f sur E et un $\alpha \in \mathbb{R}$ tels que $\Re e(f(x)) \leqslant \alpha < \Re e(f(a)) \leqslant |f(a)|$ pour tout $x \in C$. On a alors $\lambda x \in C$, pour $x \in C$ et $\lambda \in \mathbb{K}$ avec $|\lambda| = 1$; donc $\Re e(\lambda f(x)) \leqslant \alpha < |f(a)|$. Et puisque $|f(x)| = \sup_{|\lambda|=1} \Re e(\lambda f(x))$, on obtient le résultat cherché.

Enfin, si V est un sous-espace qui n'est pas partout dense, $\overline{V}$ est un convexe fermé équilibré, pour lequel existe $a \notin \overline{V}$. Alors si f est une forme linéaire continue telle que $\sup_{x \in \overline{V}} |f(x)| < |f(a)|$, $f(V)$ est un sous-espace vectoriel borné de $\mathbb{C}$, donc $\{0\}$, alors que $f(a) \neq 0$, donc $f \neq 0$. ■

4.5 Ensembles bornés

Définition 4.5.1. *Soit E un espace localement convexe. Une partie H de E est dite* bornée *si, pour tout voisinage V de 0 dans E, il existe un entier n tel que $H \subset nV$.*

Lorsque E est un espace normé, les parties bornées de E sont les parties dont le diamètre est fini.

Lemme 4.5.2. *Pour une partie H de l'espace localement convexe E, il y a équivalence entre :*
i) H est borné.
ii) toute semi-norme continue p sur E est bornée sur H.
iii) pour toute suite (x_n) dans H et toute suite (ε_n) convergeant vers 0 dans $\mathbb{K}$, la suite $(\varepsilon_n x_n)$ tend vers 0 dans E.

Démonstration : Si H est borné et si p est une semi-norme continue, l'ensemble $V = \{x : p(x) < 1\}$ est un voisinage de 0. Il existe donc un entier n tel que $H \subset nV$: on a alors $p(x) \leqslant n$ pour tout x de H, et p est bornée sur H.
Si toute semi-norme est bornée sur H, si (x_n) est une suite dans l'ensemble borné H et si $\varepsilon_n \to 0$, on a pour toute semi-norme p sur E : $p(\varepsilon_n x_n) \leqslant \varepsilon_n . \sup_H p$, donc $p(\varepsilon_n x_n) \to 0$. Et ceci montre que $(\varepsilon_n x_n)$ tend vers 0.
Enfin, si H n'est pas borné, il existe un voisinage V de 0 tel que $H \not\subset nV$ pour tout n. Si on prend alors $x_n \in H \setminus nV$, on a $\frac{1}{n} x_n \notin V$, donc $(\frac{1}{n} x_n) \not\to 0$. ■

On va maintenant donner une version générale du théorème de Banach-Steinhaus.

Théorème 4.5.3. *Une partie H de de l'espace localement convexe E est bornée si et seulement si toute forme linéaire continue sur E est bornée sur H.*

Démonstration : Si H est bornée et si $f \in E'$, l'application $p_f : x \mapsto |f(x)|$ est une semi-norme continue, donc bornée sur H.
Inversement, si toute forme linéaire continue sur E est bornée sur H, et si p est une semi-norme continue sur E, on veut montrer que p est bornée sur H. Si on note F l'espace vectoriel des formes linéaires continues ψ sur E pour lesquelles existe un M tel que $|\psi(x)| \leqslant M.p(x)$ pour tout $x \in E$, il est clair que F est inclus dans le dual E' de E. Et on vérifie sans peine que $|||\psi||| := \sup_{p(x) \leqslant 1} |\psi(x)|$ est une norme sur F. De plus, si (ψ_k) est une suite dans F vérifiant $\sum_{k=0}^{\infty} |||\psi_k||| < \infty$, on a pour tout $x \in E$: $\sum_{k=0}^{\infty} |\langle \psi_k, x\rangle| \leqslant \sum_k |||\psi_k|||.p(x) < \infty$, ce qui montre la convergence absolue de la série de terme général $(\langle \psi_k, x\rangle)$. Alors la forme linéaire ψ définie par $\langle \psi, x\rangle = \sum_k \langle \psi_k, x\rangle$ appartient à F. Et on a clairement $\left|\left|\left|\psi - \sum_{k=0}^{N} \psi_k\right|\right|\right| \leqslant \sum_{k=N+1}^{\infty} |||\psi_k|||$, ce qui montre que la série normalement convergente (ψ_k) converge vers ψ dans F. Et F est un espace de Banach.
Alors l'ensemble $A = \{\psi \in F : \sup_{x \in H} |\langle \psi, x\rangle| \leqslant 1\}$ est une partie convexe symétrique et fermée de F. De plus A est absorbant, puisque, si $\psi \in F \subset E'$, il existe un entier n tel que $\sup_{f \in H} |\langle \psi, x\rangle| \leqslant n$. On déduit alors du lemme 2.4.1 que A est un voisinage de 0 dans F, c'est-à-dire qu'il existe $\delta > 0$ pour lequel $|||\psi||| \leqslant \delta \Longrightarrow \psi \in A$.
On va montrer alors que p est bornée par δ^{-1} sur H. Si cela n'était pas vrai, il existerait un $x \in H$ tel que $p(x) > \delta^{-1}$. Et le théorème de Hahn-Banach (corollaire 4.4.5), appliqué à x et à $C = \{y : p(y) \leqslant \delta^{-1}\}$ montrerait l'existence d'une forme linéaire ψ sur E, bornée en module par 1 sur C telle que $|\psi(x)| > 1$. Il en résulterait que $\psi \in F$ avec $|||\psi||| \leqslant \delta$, mais que $|\psi(x)| > 1$, donc que $\psi \notin A$. Et cette contradiction achève de montrer que H est borné. ■

4.6 Dualité des espaces normés

Rappelons ce qui a été vu en 2.2.5 et 2.2.8 :

Théorème 4.6.1. *Si E est un $\mathbb{K}$-espace vectoriel normé, la fonction $f \mapsto \sup_{x \in E, \|x\| \leqslant 1} |f(x)|$ est une norme (appelée norme duale) sur le dual E' de E, pour laquelle E' est complet.*

Théorème 4.6.2. *Soient E un espace normé et E' son dual muni de la norme duale. Alors, pour tout $x \in E$, la fonction $\varphi_x : f \mapsto f(x)$ est une forme linéaire continue sur E', et l'application $j : x \mapsto \varphi_x$ est une application linéaire isométrique de E dans le bidual de E, c'est-à-dire le dual E'' de E'.*
L'espace normé E est dit réflexif *si j est surjective.*

Démonstration : Par définition des opérations sur E', on a $(f_1 + f_2)(x) = f_1(x) + f_2(x)$ et $(\lambda f)(x) = \lambda.f(x)$. Ceci montre que l'application $\varphi_x : f \mapsto f(x)$ est linéaire de E' dans $\mathbb{K}$. Puisque $|\varphi_x(f)| = |f(x)| \leqslant \|f\| . \|x\|$, on voit que $j(x) = \varphi_x$ est une forme linéaire continue sur E' et même que $\|j(x)\| \leqslant \|x\|$. Par ailleurs, par linéarité des éléments de E', on a :

$$j(x + \lambda y)(f) = f(x + \lambda y) = f(x) + \lambda f(y) = (j(x) + \lambda j(y))(f) \ ,$$

c'est-à-dire que j est linéaire. On en conclut que j est une application linéaire continue de E dans E'' de norme au plus 1.
Enfin, si $a \in E$, il résulte du corollaire 4.4.3 qu'il existe $g \in E'$ telle que $\|g\| = 1$ et $g(a) = \|a\|$. On a ainsi $j(a)(g) = \|a\|$, donc $\|j(a)\| \geqslant \|a\|$. On conclut de ce qui précède que $\|j(a)\| = \|a\|$, donc que j est isométrique de E dans E''. En particulier j est injective. Lorsque E est un espace de Banach, $j(E)$ est complet, donc fermé dans E''. ■

Définitions 4.6.3. *Soit E un espace vectoriel normé. On appelle* topologie faible *sur E la topologie affaiblie $\sigma(E, E')$.*
On appelle topologie préfaible *(ou $*$-faible) sur E' la topologie faible $\sigma(E', E)$, où on a identifié par l'application j ci-dessus l'espace E à un espace vectoriel de formes linéaires sur E'.*

Si l'espace E est réflexif, la topologie faible $\sigma(E', E'')$ sur E' coïncide avec la topologie préfaible $\sigma(E', E)$.

Compte tenu du théorème de Riesz, qui permet d'identifier le dual d'un espace de Hilbert à cet espace, le résultat suivant peut être vu comme une généralisation du théorème 3.4.6.

Théorème 4.6.4. (Banach-Alaoglu) *Si E est un espace normé, la boule unité de son dual E' est compacte pour la topologie préfaible $\sigma(E', E)$.*

Démonstration : Notons B' la boule unité de E'. Pour tout $f \in B'$ on a $|f(x)| \leqslant \|x\|$. On peut donc définir une application Φ de B' dans le compact $K = \prod_{x \in E} D_x$, où $D_x = \{\lambda \in \mathbb{K} : |\lambda| \leqslant \|x\|\}$ par $\Phi(f) = (f(x))_{x \in E}$. Par définition de la topologie préfaible sur E', chaque fonction coordonnée $f \mapsto f(x)$ est continue, ce qui montre que Φ est continue.
L'image $\Phi(B')$ est formée des fonctions $f \in K$ satisfaisant pour tout $x \in E$, tout $y \in E$ et tout $\lambda \in \mathbb{K}$: $f(x + \lambda y) - f(x) - \lambda f(y) = 0$. Donc $\Phi(B')$ est l'ensemble fermé

$$H = \bigcap_{x,y,\lambda} \{f \in K : f(x + \lambda y) - f(x) - \lambda f(y) = 0\} \ ,$$

qui est compact dans K. Pour chaque semi-norme $p_x : f \mapsto |f(x)|$ sur E', l'application $p_x \circ \Phi^{-1}$ est continue sur H puisque elle y coïncide avec l'application $f \mapsto |f(x)|$ qui est continue pour la topologie produit. On en déduit que Φ^{-1} est continue de H dans E' pour la topologie préfaible, donc que Φ est un homéomorphisme de B' sur le compact H. Ceci montre que B' est préfaiblement compact. ■

Corollaire 4.6.5. *Si E est un espace de Banach séparable, toute suite bornée de E' possède une sous-suite préfaiblement convergente.*

Démonstration : Soit $R = \sup_n \|f_n\| < +\infty$. Quitte à remplacer (f_n) par $(\dfrac{f_n}{R})$, on se ramène au cas où la suite (f_n) est dans la boule unité B' de E'. Si D est une partie dénombrable dense de E, l'application $\psi : B' \to \mathbb{K}^D$ définie par $\psi(f) = (f(x))_{x\in D}$ est continue à valeurs dans l'espace métrisable $\mathbb{K}^D$ (produit dénombrable d'espaces métriques) quand on munit B' de la topologie préfaible. De plus, l'application ψ est injective : en effet, si $\psi(f) = \psi(g)$, la forme linéaire continue $f - g$ s'annule en tout point de l'ensemble dense D, donc est identiquement nulle. Puisque B' est compact, ceci montre que ψ est un homéomorphisme, donc que B' est un compact métrisable. La suite (f_n) possède donc une sous-suite convergente, en vertu du théorème 1.5.14. ■

Théorème 4.6.6. *Soient E un espace normé, E' son dual et E'' le bidual de E. Alors la boule unité de E, identifiée par j à une partie de E'', est dense dans la boule unité de E'' pour la topologie préfaible $\sigma(E'', E')$.*

Démonstration : Notons B la boule unité de E et B'' celle de E''. On a $j(B) \subset B''$. Si $j(B)$ n'était pas dense dans B'' pour la topologie préfaible $\sigma(E'', E')$ son adhérence serait un convexe fermé équilibré $\tilde{B}$ ne contenant pas B''. On pourrait donc trouver un $\xi \in B'' \setminus \tilde{B}$. D'après le corollaire 4.4.5, il existerait alors une forme linéaire g sur E'', continue pour $\sigma(E'', E')$ telle que $|g(\xi)| > \sup_{y\in\tilde{B}} |g(y)| \geqslant \sup_{x\in B} |g\circ j(x)|$. D'après le théorème 4.2.1, il existe $f \in E'$ tel que $g(z) = z(f)$ pour $z \in E''$.
On a donc $g\circ j(x) = f(x)$ et $\sup_{x\in B} |g\circ j(x)| = \sup_{x\in B} |f(x)| = \|f\|$. Mais, puisque $\xi \in B''$, $|g(\xi)| = |\xi(f)| \leqslant \|\xi\| \, \|f\| \leqslant \|f\|$. Et cette contradiction achève de prouver la densité préfaible de $j(B)$ dans B''. ■

Théorème 4.6.7. *Soit E un espace normé. Les conditions suivantes sont équivalentes :*
(i) E est réflexif
(ii) E' est réflexif.
(iii) La boule unité de E est faiblement compacte.

Démonstration : Si E est réflexif, le dual de E' est égal à E : le bidual de E' est donc égal au dual de E, c'est-à-dire E'. Et ceci montre que E' est réflexif.

Supposons que E' soit réflexif et qu'il existe un élément a de E'' n'appartenant pas à E. Alors $j(E)$ est un sous-espace vectoriel fermé de E'' ne contenant pas a, et il existe par le corollaire 4.4.5 une forme linéaire f sur E'' satisfaisant $|f(a)| > \sup_{z\in j(E)} |f(z)|$. S'il existait un $z \in j(E)$ tel que $f(z) \neq 0$, il existerait un entier n tel que $|f(nz)| = n\,|f(z)| > |f(a)|$. Et puisque $nz \in j(E)$, on obtient une contradiction qui montre que $f = 0$ sur $j(E)$. Et comme E' est réflexif, le dual de E'' est E' : donc $f \in E'$. On a ainsi $f(x) = 0$ pour tout $x \in E$, c'est-à-dire $f = 0$, contrairement à $|f(a)| > 0$. Donc E est réflexif.

Si E est réflexif, la boule de E munie de la topologie faible $\sigma(E, E')$ est égale à la boule de E'' munie de la topologie préfaible $\sigma(E'', E')$, qui est compacte d'après le théorème de Banach-Alaoglu (théorème 4.6.4).

Enfin, si la boule B de E est faiblement compacte, $j(B)$ est compacte pour la topologie $\sigma(E'', E')$, donc fermée dans B''. Et puisqu'elle est dense dans B'' d'après le théorème 4.6.6, on a $j(B) = B''$, donc $j(E) = E''$. Et E est réflexif. ■

Définition 4.6.8. *Une norme $\|.\|$ sur un espace E est dite* uniformément convexe *si, pour tout $\varepsilon > 0$, il existe un $\alpha > 0$ tel que, pour tout x et tout y de la boule unité, on ait $\|x - y\| < \varepsilon$ dès que leur milieu $\dfrac{x+y}{2}$ a une norme supérieure à $1 - \alpha$.*

Il résulte aisément du théorème 3.1.9 que les espaces de Hilbert sont uniformément convexes. Et le théorème suivant utilise des arguments analogues à ceux qui servent à démontrer le théorème de la projection orthogonale dans les espaces de Hilbert, qui lui-même, via le théorème de Riesz, a permis de prouver la réflexivité des espaces de Hilbert.

Théorème 4.6.9. *Un espace dont la norme est uniformément convexe est réflexif.*

Démonstration : Il suffit de montrer que si a est un point de la boule unité B'' de E'', alors $a \in E$. Pour toute famille finie f_1, f_2, …, f_k d'éléments de E' on considère l'ensemble convexe fermé et borné de E :

$$H_{f_1,f_2,\ldots,f_k} = \{x \in B : |\langle f_i, x\rangle - \langle a, f_i\rangle| \leqslant 1 \quad , \ i = 1, 2, \ldots, k\} \ .$$

Supposons, par l'absurde, que $a \notin j(E)$. Puisque a est adhérent à B pour la topologie $\sigma(E'', E')$, chacun de ces ensembles est non vide. Et l'intersection de deux ensembles de la famille $\mathscr{H}$ des $H_{f_1,f_2,\ldots,f_k}$ est encore dans $\mathscr{H}$. Si on considère $\delta = \sup_{H\in\mathscr{H}} d(0, H) \leqslant 1$, la boule de centre 0 et de rayon $\delta + 2^{-k}$ dans E rencontre chaque $H \in \mathscr{H}$. Alors, si $H^{(k)}$ est un élément de $\mathscr{H}$ tel que $d(0, H^{(k)}) \geqslant \delta - 2^{-k}$, et si x et y appartiennent à $H_0^{(k)} = H^{(k)} \cap B(0, \delta + 2^{-k})$, on a $\left\|\dfrac{x+y}{2}\right\| \geqslant \delta - 2^{-k}$ puisque $H_0^{(k)}$ est convexe. La condition de convexité uniforme montre alors que le diamètre des $H_0^{(k)}$ tend vers 0. Comme ces ensembles se coupent deux-à-deux, si on choisit un point z_k dans $H_0^{(k)}$, la suite (z_k) est une suite de Cauchy dans E, donc converge vers un point $b \in E$.
Puisque $a \neq j(b)$, il existe $f \in E'$ tel que $|\langle a - j(b), f\rangle| > 1$, ce qui montre que $d(b, H_f) > 0$, et que H_f ne peut rencontrer tous les $H_0^{(k)}$. Il existe donc un k tel que $H_k \cap H^{(k)}$ ne rencontre pas $B(0, \delta + 2^{-k})$, en contradiction avec la définition de δ, puisque $H_k \cap H^{(k)} \in \mathscr{H}$. Il en résulte que $a \in E$, donc que E est réflexif. ■

On peut noter que cette condition suffisante de réflexivité n'est en fait pas nécessaire : il existe des espaces réflexifs dont la norme n'est pas uniformément convexe.

Définition 4.6.10. *Soient E et F deux espaces de Banach, et $\varphi : (x, y) \mapsto \langle x, y\rangle$ une forme bilinéaire continue sur $E \times F$. On dira que φ* identifie *F au dual de E si l'application linéaire $y \mapsto \varphi_y$, où φ_y est la forme linéaire $x \mapsto \langle x, y\rangle$, est un isomorphisme de F sur E'.*

On a pour $x \in E$ et $y \in F$, $|\langle x, y\rangle| = |\varphi(x, y)| \leqslant \|\varphi\| \, \|x\| \, \|y\|$. La forme linéaire φ_y vérifie donc $|\varphi_y(x)| \leqslant \|\varphi\| \, \|x\| \, \|y\|$, c'est-à-dire $\|\varphi_y\| \leqslant \|\varphi\| \, \|y\|$. Et cette dernière inégalité montre que l'application linéaire $\Phi : y \mapsto \varphi_y$ est continue de F à valeurs dans E'.
Lorsque E est réflexif, le produit $(f, x) \mapsto \langle f, x\rangle = f(x)$ défini sur $E' \times E$ identifie ainsi E au dual E'' de E'. Dans ce cas l'application Φ définie ci-dessus coïncide avec le plongement canonique j de E dans E''.
On peut remarquer, dans le cas général, que d'après le théorème de l'application ouverte (corollaire 4.7.2), il suffit que Φ soit bijective de F sur E', pour que φ identifie F à E'.

4.7 Le théorème de l'application ouverte pour les espaces de Fréchet

Théorème 4.7.1. *Soient E et F deux espaces de Fréchet et u une application linéaire continue de E dans F. Si $u(E)$ n'est pas maigre dans F, l'application u est surjective et ouverte.*

Démonstration : On peut choisir sur E et sur F des distances invariantes par translation pour lesquelles les boules de centre 0 sont convexes et équilibrées, et qui définissent respectivement la topologie de E et de F, rendant ainsi complets E et F.
Fixons $r > 0$ et $\varepsilon \in]0,1[$, puis posons, pour $n \in \mathbb{N}$: $B_n = \{x \in E : d(x,0) < r\varepsilon^n\}$. Puisque $u(F) \subset \bigcup_{k\in\mathbb{N}} \overline{u(k.B_n)}$ n'est pas maigre, l'un au moins des $\overline{u(k.B_n)}$ n'est pas rare. Et puisque l'homothétie de rapport k est un homéomorphisme, $\overline{u(B_n)}$ n'est pas rare non plus. Il existe donc un $a \in F$ et un $\rho_n > 0$ tel que $B_F(a,\rho_n) \subset \overline{u(B_n)}$. Par symétrie, $B_F(-a,\rho_n) \subset \overline{u(B_n)}$, et par convexité $B_F(0,\rho_n) \subset \dfrac{1}{2}(B_F(a,\rho_n) + B_F(-a,\rho_n)) \subset \overline{u(B_n)}$.
On va montrer qu'alors $u(B_E(0, \dfrac{r}{1-\varepsilon}))$ contient $B_F(0,\rho_0)$. Quitte à diminuer ρ_n, on peut supposer que $\rho_n \leqslant \rho_0.2^{-n}$.
Soit donc $y \in B_F(0,\rho_0)$. On va construire par récurrence une suite (x_n) dans E telle que $x_n \in B_n$ et que $y - \sum_{j=0}^{n-1} u(x_j) \in B_F(0,\rho_n)$, ce qui est possible pour $n = 0$. Si la construction est faite pour $n < m$, on a $y_m := y - \sum_{j=0}^{m-1} u(x_j) \in B_F(0,\rho_m) \subset \overline{u(B_m)}$. Il existe donc un $x_m \in B_m$ tel que $y_m - u(x_m) \in B_F(0,\rho_{m+1})$, ce qui achève la construction par récurrence.
Alors on a $\sum_{n=0}^{\infty} d(0,x_n) < \sum_0^{\infty} r.\varepsilon^n < +\infty$, ce qui montre la convergence de la série (x_n). Et puisque $d(y, u(\sum_0^{n-1} x_j)) \leqslant \rho_n \to 0$, on voit que $x := \lim_{n\to\infty} \sum_{j=0}^{n} x_j$ est un point de E tel que $u(x) = y$. Comme on a $d(0,x) < \sum_{j=0}^{\infty} r.\varepsilon^j = \dfrac{r}{1-\varepsilon}$, on voit que $y \in u(B_E(0, \dfrac{r}{1-\varepsilon}))$.
Etr ceci achève de montrer que u est surjective et ouverte. En particulier, si u est injective, c'est un isomorphisme. ■

Corollaire 4.7.2. *Soient E et F deux espaces de Fréchet et u une application linéaire continue surjective de E dans F. Alors l'application u est ouverte. Si, de plus, u est bijective, c'est un isomorphisme.*

Démonstration : En vertu du théorème de Baire, F n'est pas maigre dans F : si donc u est surjective, $u(E)$ n'est pas maigre dans F. Et si u est bijective et ouverte, u^{-1} est continue. ■

Théorème 4.7.3. *Soient E et F deux espaces de Fréchet, et u une application linéaire de E dans F. Si le graphe de u est fermé dans $E \times F$, alors u est continue.*

Démonstration : Notons G le graphe de $E \times F$, qui est un espace de Fréchet puisque fermé dans l'espace complet $E \times F$, π_1 et π_2 les restrictions à G des projections de $E \times F$ sur E et F respectivement. Par définition de la topologie de $E \times F$, π_1 et π_2 sont continues. Puisque π_1 est bijective, il résulte du théorème précédent que π_1 est un isomorphisme, et puisque $u = \pi_2 \circ \pi_1^{-1}$, ceci montre la continuité de u. ■

4.8 Limites inductives d'espaces de Fréchet

Théorème 4.8.1. *Soient E un espace de Fréchet et (F_n) une suite croissante de sous-espaces vectoriels fermés de E. alors il existe sur l'espace vectoriel $F = \bigcup_n F_n$ une unique topologie localement convexe $\mathscr{O}$ appelée* limite inductive *de la suite (F_n) qui vérifie :*
(i) la topologie $\mathscr{O}$ coïncide sur chaque F_n avec la topologie induite par E.
(ii) Une application linéaire f de F dans un espace normé G (en particulier une forme linéaire) est continue pour $\mathscr{O}$ si et seulement si la restriction de f à chaque F_n est continue.

De plus, cette topologie vérifie la propriété :
(iii) Toute suite convergente (x_k) est contenue dans l'un des F_n.

Démonstration : Désignons par P la famille de toutes les semi-normes p sur F telles que la restriction de p à chaque F_n soit continue pour la topologie d'espace de Fréchet sur F_n. On va montrer que cette P-topologie, $\mathscr{O}$, a les propriétés cherchées. Il existe sur E, d'après le théorème 4.1.12, une distance d invariante par translation qui définit la topologie de E et pour laquelle les boules sont convexes et équilibrées, distance elle-même définie à partir d'une suite croissante (p_k) de semi-normes sur E.
Puisque toute semi-norme dans P est continue sur F_n muni de d, on voit que la topologie induite par $\mathscr{O}$ sur F_n est moins fine que la topologie induite par d. Inversement, si V est un voisinage de 0 dans F_n, il existe un k et un $r > 0$ tels que

$$F_n \cap \{x \in E : p_k(x) < r\} \subset V$$

et puisque p_k est continue sur chaque F_n, on a $p_k \in P$, ce qui montre que V est un voisinage de 0 dans F_n pour la topologie induite par $\mathscr{O}$. On en déduit que $\mathscr{O}$ induit sur F_n la même topologie que E.

Si G est un espace normé et f une application linéaire de F dans G, continue sur chaque F_n, la fonction $q : x \mapsto \|f(x)\|$ est une semi-norme sur F, dont la restriction à F_n est continue pour tout n : il en résulte que $q \in P$, ce qui montre que f est continue pour la topologie $\mathscr{O}$.

On vérifie maintenant que si $\mathscr{O}'$ est la Q-topologie définie par une famille de semi-normes sur E avec les propriétés (i) et (ii), les semi-normes continues pour $\mathscr{O}$ et $\mathscr{O}'$ sont les-mêmes, c'est-à-dire que $\mathscr{O} = \mathscr{O}'$, ce qui est le résultat d'unicité annoncé.
Soit en effet q une semi-norme continue pour $\mathscr{O}'$: alors q est continue sur chaque F_n, donc appartient à P, et est continue pour $\mathscr{O}$. Inversement, si $p \in P$, l'ensemble $\{x \in F : p(x) = 0\}$ est un sous-espace vectoriel H de F. Si x et y sont deux points de F tels que $x - y \in H$, on a $p(x) \leqslant p(y) + p(x - y) = p(y)$ et de même $p(y) \leqslant p(x)$ donc $p(x) = p(y)$.
Sur l'espace vectoriel quotient $G = F/H$, il existe donc une fonction $q : E/H \to \mathbb{R}^+$ telle que $q(\pi(x)) = p(x)$ si π est la projection canonique de F sur G. On vérifie immédiatement que q est une semi-norme sur G, et même une norme puisque , si $q(\pi(x)) = 0$, on a $p(x) = 0$ donc $x \in H$ et $\pi(x) = 0$. Alors l'application linéaire $\pi : F \to G$ est continue sur chaque F_n, donc continue pour $\mathscr{O}'$, ce qui signifie que $q \circ \pi$ est une semi-norme continue pour $\mathscr{O}'$. Et ceci montre l'unicité de la topologie $\mathscr{O}$.

Pour montrer le point (iii), on va montrer d'abord un lemme permettant de construire les voisinages de 0 dans $\mathscr{O}$.

Lemme 4.8.2. *On suppose que (n_k) est une suite strictement croissante d'entiers, (ε_k) une suite de nombres réels telle que $0 < \varepsilon_{k+1} < \varepsilon_k/2$. Et on définit les ensembles convexes C_k par*

$$C_k = \sum_{j=0}^{k} F_{n_j} \cap B(0, \varepsilon_j) = \Big\{ \sum_{j=0}^{k} x_j : x_j \in F_{n_j} \ , \ d(x_j, 0) \leqslant \varepsilon_j \Big\}$$

où $B(0,r)$ désigne la boule de centre 0 et de rayon r pour la distance d, puis $C = \bigcup_{k=0}^{\infty} C_k$. Alors C est un voisinage de 0 pour $\mathscr{O}$. De plus, si $x \in C$, on a $d(x, F_{n_k}) < \varepsilon_k$ pour tout k. De plus, tout voisinage de 0 dans F pour $\mathscr{O}$ contient un tel voisinage.

Démonstration : Si on a construit la distance d comme en 4.1.12, on a $\lambda C_k = C_k$ pour tout $\lambda \in \mathbb{K}$ tel que $|\lambda| = 1$. Il en résulte que C_k est un convexe équilibré. La suite (C_k) est clairement croissante. Si $x \in F$, il existe un k tel que $x \in F_{n_k}$, donc un $t > 0$ tel que $tx \in F_{n_k} \cap B(0, \varepsilon_k) \subset C_k \subset C$. On peut alors considérer la jauge p de C, qui est une semi-norme.
Puisque $C \cap F_{n_k} \supset F_{n_k} \cap B(0, \varepsilon_k)$, $C \cap F_{n_k}$ est un voisinage de 0 dans F_{n_k} sur lequel $p \leqslant 1$. Il en résulte que la restriction de p à F_{n_k} est continue, et a fortiori la restriction de p à F_n si $n \leqslant n_k$. On en déduit que $p \in P$, et que C est un voisinage de 0 pour $\mathscr{O}$.

Soient $x \in C$ et k entier. Il existe, par définition de C, un ℓ tel que $x \in C_\ell$, donc des $(x_i)_{i \leqslant \ell}$ tels que $x = \sum_{i=0}^{\ell} x_i$ et $x_i \in F_{n_i} \cap B(0, \varepsilon_i)$. Alors $y = \sum_{i=0}^{k} x_i \in F_{n_k}$ et $d(x, F_{n_k}) \leqslant d(x,y) \leqslant \sum_{i=k+1}^{\ell} d(0, x_i) \leqslant \sum_{i=k+1}^{\ell} \varepsilon_i$. Pour $i \geqslant k$, on a $\varepsilon_i < 2^{k-i} \varepsilon_k$, donc $d(x, F_{n_k}) \leqslant \varepsilon_k \sum_{i=k+1}^{\ell} 2^{k-i} < \varepsilon_k$.
Inversement, si $p \in P$ et $r > 0$, l'ensemble $W_n = \{x \in F_n : p(x) < r.2^{-n-1}\}$ est un voisinage de 0 dans F_n, et on peut construire par récurrence une suite (ε_n) telle que que $\varepsilon_n < \varepsilon_{n-1}$ et que W_n contienne une boule $F_n \cap B(0, \varepsilon_n)$ de F_n. Alors, si $x = \sum_{j \leqslant k} x_j$ avec $x_j \in F_j \cap B(0, \varepsilon_j)$, on a $p(x) \leqslant \sum_{j \leqslant k} p(x_j) \leqslant \sum_{j \leqslant k} r.2^{-j-1} < r$, ce qui montre que $C = \bigcup_k \sum_{j \leqslant k} F_j \cap B(0, \varepsilon_j) \subset \{x : p(x) < r\}$, donc que tout voisinage de 0 pour $\mathscr{O}$ contient un voisinage du type décrit ci-dessus. ■

On achève maintenant la démonstration de *(iii)*. Si (x_n) était une suite de F qui converge vers a pour $\mathscr{O}$ sans que les (x_n) soient tous dans un même F_k, la suite $(x_n - a)$ convergerait vers 0 sans que tous ses termes soient dans un même F_k. On se ramène ainsi au cas où $a = 0$. Puisque, pour tout ℓ, il existe une infinité de termes de la suite situés hors de F_ℓ, on peut construire par récurrence deux suites strictement croissantes d'entiers (n_k) et (m_k) telles que : $x_{m_k} \in F_{n_k}$ et $x_{m_{k+1}} \notin F_{n_k}$. ainsi qu'une suite (ε_k) de réels telle que $0 < \varepsilon_{k+1} < \dfrac{1}{2}\varepsilon_k$ et $\varepsilon_k < d(x_{m_{k+1}}, F_{n_k})$. Le voisinage C de 0 construit dans le lemme précédent ne peut contenir alors aucun des points x_{m_k} pour $k \geqslant 1$, ce qui contredit la convergence de la suite (x_n) vers 0, et achève la démonstration. ■

Corollaire 4.8.3. *Soient E un espace de Fréchet et (F_n) une suite croissante de sous-espaces vectoriels fermés de E. Si on munit l'espace vectoriel $F = \bigcup_n F_n$ de la topologie limite inductive des (F_n), les suites convergentes (x_m) de F sont les suites dont tous les termes appartiennent à un même F_n et qui convergent dans cet espace F_n.*

Démonstration : Puisque F_n est un sous-espace fermé de F, toute suite (x_m) à termes dans F_n qui converge vers a dans F_n converge aussi vers a dans F.

Inversement, si (x_m) converge vers a dans F, il résulte du théorème 4.8.1 *iii*) que a ainsi que tous les termes de cette suite sont dans un même F_n. Et puisque F_n est un sous-espace topologique de F, la suite (x_m) converge vers a dans F_n. ■

Le théorème du graphe fermé.

On va maintenant étendre aux limites inductives de suites d'espaces de Fréchet les théorèmes de l'application ouverte et du graphe fermé, qui ont été montrés plus haut pour les espaces de Fréchet.

Théorème 4.8.4. *Soient E et F deux espaces de Fréchet, (E_n) une suite croissante de sous-espaces fermés de E, et (F_n) une suite croissante de sous-espaces fermés de F. On munit $E_\infty = \bigcup_n E_n$ et $F_\infty = \bigcup_n F_n$ de la topologie limite inductive. Alors, si u est une application linéaire continue surjective de E_∞ dans F_∞, l'application u est ouverte.*

Démonstration : Pour tout n et tout p , le sous-espace $H_n = u^{-1}(F_n)$ est fermé dans E_∞ et la restriction de u à $E_p \cap H_n$ est continue de H_n dans F_n. De plus $\bigcup_p u(E_p \cap H_n) = F_n$, ce qui montre que l'un des $u(E_p \cap H_n)$ n'est pas maigre dans F_n. Alors la restriction de u à $E_p \cap H_n$ est surjective et ouverte de $E_p \cap H_n$ sur F_n. Et si W est un voisinage convexe de 0 dans E_∞, on en déduit que $W \cap E_p \cap H_n$ est un voisinage de 0 dans $E_p \cap H_n$ et que $u(W) \cap F_n$ est un voisinage de 0 dans F_n. Et par définition de la topologie de F_∞, le convexe $u(W)$ est alors un voisinage de 0, ce qui montre que u est ouverte.
Et si u est bijective, elle est alors un isomorphisme. ■

Théorème 4.8.5. *Soient E et F deux espaces de Fréchet, (E_n) une suite croissante de sous-espaces fermés de E, et (F_n) une suite croissante de sous-espaces fermés de F. On munit $E_\infty = \bigcup_n E_n$ et $F_\infty = \bigcup_n F_n$ de la topologie limite inductive. Alors, si u est une application linéaire de E_∞ dans F_∞ dont le graphe est fermé dans $E_\infty \times F_\infty$, l'application u est continue.*

Démonstration : Si on note π_1 et π_2 les restrictions au graphe G de u des projections de $E_\infty \times F_\infty$ sur E_∞ et F_∞ respectivement, on remarque d'abord que sur G la topologie limite inductive G_∞ des sous-espaces fermés $G \cap (E_n \times F_n)$ est plus fine que la topologie de sous-espace de $E_\infty \times F_\infty$: en effet tout voisinage convexe de 0 dans $E_\infty \times F_\infty$ coupe $G \cap (E_n \times F_n)$ suivant un voisinage de 0 dans l'espace de Fréchet $G \cap (E_n \times F_n)$, donc est un voisinage de 0 dans G_∞. Il en résulte que π_1, qui est linéaire continue et bijective de G_0 sur E_∞, est un isomorphisme. Donc π_1^{-1} est continue de E_∞ sur G_∞ et a fortiori sur G. Donc $u = \pi_2 \circ \pi_1^{-1}$ est continue. ■

5

INTÉGRATION

5.1 Fonctions mesurables

Tribus.

Définition 5.1.1. *Soit X un ensemble. On appelle* algèbre de parties *de X un ensemble non vide $\mathfrak{A}$ de parties de X tel que si A et B appartiennent à $\mathfrak{A}$, on ait aussi $A^c := X \setminus A \in \mathfrak{A}$ et $A \cup B \in \mathfrak{A}$.*

Clairement, si $\mathfrak{A}$ est une algèbre de parties de X, alors $\mathfrak{A}$ est stable par intersections. Et si $A \in \mathfrak{A}$, on a $\emptyset = A \cap A^c \in \mathfrak{A}$.

Définition 5.1.2. *Soit X un ensemble. On appelle* tribu *(ou σ-algèbre) sur X un ensemble $\mathfrak{S}$ de parties de X satisfaisant les propriétés :*

i) $\emptyset \in \mathfrak{S}$

ii) si $A \in \mathfrak{S}$, alors $A^c = X \setminus A \in \mathfrak{S}$

iii) si (A_n) est une suite d'éléments de $\mathfrak{S}$, alors $\bigcup_{n\in\mathbb{N}} A_n \in \mathfrak{S}$.

Il résulte clairement de *ii*) et *iii*) que si $\mathfrak{S}$ est une tribu, et si (A_n) est une suite d'éléments de $\mathfrak{S}$, alors $\bigcap_n A_n \in \mathfrak{S}$. Et si A et B appartiennent à la tribu $\mathfrak{S}$, alors $A \setminus B = A \cap B^c$ appartient aussi à $\mathfrak{S}$.
Il est immédiat que, pour tout ensemble X, $\{\emptyset, X\}$ est une tribu sur X (la plus petite), et que $\mathscr{P}(X)$ est une tribu sur X (la plus grosse).

Si $\mathscr{H}$ est un ensemble de parties de X, l'intersection de toutes les tribus sur X contenant $\mathscr{H}$ est encore une tribu sur X contenant $\mathscr{H}$, et c'est la plus petite tribu sur X contenant $\mathscr{H}$. On l'appelle *la tribu engendrée par $\mathscr{H}$*.

En particulier, si E est un espace métrique séparable, la tribu $\mathfrak{B}(E)$ engendrée sur E par les ouverts est appelée *tribu de Borel*, ou tribu borélienne de X. Un sous-ensemble de E est dit *borélien* s'il appartient à $\mathfrak{B}(E)$.

Il sera nécessaire à plusieurs reprises, dans la suite, de considérer des ensembles de parties de X qui sont seulement stables par réunions dénombrables *croissantes* et intersections dénombrables *décroissantes*. Le théorème suivant sera donc utile.

Théorème 5.1.3. *Soient $\mathfrak{A}$ une algèbre de parties de X, et $\mathscr{H}$ un ensemble de parties de X contenant $\mathfrak{A}$ et stable par unions dénombrables croissantes et intersections dénombrables décroissantes. Alors $\mathscr{H}$ contient la tribu engendrée par $\mathfrak{A}$.*

Démonstration : Notons $\mathfrak{C}$ la plus petite famille de parties de X contenant $\mathfrak{A}$ et stable par unions dénombrables croissantes et intersections dénombrables décroissantes. On a $\mathfrak{C} \subset \mathscr{H}$ par hypothèse. La preuve utilise plusieurs lemmes.

Lemme 5.1.4. *Si $A \in \mathfrak{A}$, alors $A \cap C \in \mathfrak{C}$ pour tout $C \in \mathfrak{C}$.*

En effet, l'ensemble $\mathfrak{C}_A = \{Z \subset X : A \cap Z \in \mathfrak{C}\}$ contient $\mathfrak{A}$ puisque $\mathfrak{A}$ est une algèbre. De plus, si (Z_n) est une suite croissante dans $\mathfrak{C}_A$ de réunion Z, la suite $(Z_n \cap A)$ d'éléments de $\mathfrak{C}$ est croissante et de réunion $Z \cap A$: il en résulte que $Z \in \mathfrak{C}_A$. De même, si (Z_n) est une suite décroissante dans $\mathfrak{C}_A$, d'intersection Z, la suite $(Z_n \cap A)$ est décroissante dans $\mathfrak{C}$, d'intersection $Z \cap A$; alors $Z \cap A \in \mathfrak{C}$, et $Z \in \mathfrak{C}_A$. On conclut que $\mathfrak{C}_A$ est stable par unions dénombrables croissantes et intersections dénombrables décroissantes, donc contient $\mathfrak{C}$. ■

Lemme 5.1.5. *L'intersection de deux éléments de $\mathfrak{C}$ est dans $\mathfrak{C}$.*

Notons $\mathfrak{C}_2$ l'ensemble $\{Z \subset X : \forall C \in \mathfrak{C} \quad Z \cap C \in \mathfrak{C}\}$. Il résulte du lemme précédent que $\mathfrak{A} \subset \mathfrak{C}_2$. Et comme ci-dessus, si (Z_n) est une suite croissante d'union Z dans $\mathfrak{C}_2$, $C \cap Z$ est la réunion croissante des $(C \cap Z_n)$ pour tout $C \in \mathfrak{C}$; il en résulte que $Z \in \mathfrak{C}_2$. Et de même, si (Z_n) est une suite décroissante dans $\mathfrak{C}_2$ d'intersection Z, on voit que $Z \in \mathfrak{C}_2$. Il en résulte que $\mathfrak{C}_2$ est stable par unions dénombrables croissantes et intersections dénombrables décroissantes. Et à nouveau $\mathfrak{C} \subset \mathfrak{C}_2$. ■

Lemme 5.1.6. *Si $C \in \mathfrak{C}$, alors $C^c \in \mathfrak{C}$.*

On pose maintenant $\widetilde{\mathfrak{C}} = \{Z : Z^c \in \mathfrak{C}\}$. Puisque $\mathfrak{A}$ est une algèbre, on a $A^c \in \mathfrak{A} \subset \mathfrak{C}$ pour tout $A \in \mathfrak{A}$, donc $\mathfrak{A} \subset \widetilde{\mathfrak{C}}$. Si (Z_n) est une suite croissante dans $\widetilde{\mathfrak{C}}$, de réunion Z, la suite (Z_n^c) est décroissante, d'intersection Z^c, dans $\mathfrak{C}$; on en déduit que $Z^c \in \mathfrak{C}$, donc que $Z \in \widetilde{\mathfrak{C}}$. Et de même $\widetilde{\mathfrak{C}}$ est stable par intersections dénombrables décroissantes. Donc $\mathfrak{C} \subset \widetilde{\mathfrak{C}}$. ■

Lemme 5.1.7. *L'ensemble $\mathfrak{C}$ est une tribu.*

Il résulte des lemmes précédents que $\mathfrak{C}$ est stable par complémentation et par intersection binaire. Si (C_n) est une suite dans $\mathfrak{C}$, et si on pose $C'_n = \bigcap_{m \leqslant n} C_m$, on a $C'_{n+1} = C'_n \cap C_{n+1}$ et on voit par récurrence que la suite décroissante (C'_n) est dans $\mathfrak{C}$. Alors $\bigcap_n C_n = \bigcap_n C'_n \in \mathfrak{C}$. Il en résulte que $\mathfrak{C}$ est stable par intersections dénombrables, ainsi que par unions dénombrables puisque $\bigcup_n C_n = X \setminus \bigcap_n C_n^c$. Enfin $\emptyset \in \mathfrak{A} \subset \mathfrak{C}$. ■

Démonstration : Puisque $\mathfrak{C}$ est une tribu qui contient $\mathfrak{A}$, elle contient la tribu $\mathfrak{S}$ engendrée par $\mathfrak{A}$, et $\mathscr{H}$ qui contient $\mathfrak{C}$, contient aussi $\mathfrak{S}$. ■

Fonctions mesurables.

Définition 5.1.8. *On appellera* espace mesurable *un couple $(X, \mathfrak{S})$, où X est un ensemble et $\mathfrak{S}$ est une tribu sur X.*

Définition 5.1.9. *Soient $(X, \mathfrak{S})$ un espace mesurable et E un espace métrique séparable. Une partie A de X est dite* mesurable *si elle appartient à $\mathfrak{S}$. La fonction $f : X \to E$ est dite* mesurable *si l'image réciproque par f de tout ouvert U de E appartient à $\mathfrak{S}$.*

La topologie de E possède une base dénombrable (U_n), et tout ouvert de E est réunion (nécessairement dénombrable) d'ouverts de la base. Il en résulte que si $f^{-1}(U_n) \in \mathfrak{S}$ pour tout n, la fonction f est mesurable.
Si $f : X \to E$ est mesurable, l'ensemble $\mathfrak{T} = \{Z \subset E : f^{-1}(Z) \in \mathfrak{S}\}$ est une tribu, comme on le vérifie sans peine, et les ouverts de E appartiennent à $\mathfrak{T}$ par hypothèse. Il en résulte que $\mathfrak{B}(E) \subset \mathfrak{T}$, autrement dit : l'image réciproque par f de toute partie borélienne de E est dans la tribu $\mathfrak{S}$.

Théorème 5.1.10. *Soient $(X, \mathfrak{S})$ un espace mesurable, E et F deux espaces métriques séparables, $f : X \to E$ une fonction mesurable, et $\varphi : E \to F$ une fonction continue. Alors $\varphi \circ f$ est mesurable de X dans F.*

Démonstration : Soit U un ouvert de F. Alors $\varphi^{-1}(U)$ est un ouvert de E, et $(\varphi \circ f)^{-1}(U) = f^{-1}(\varphi^{-1}(U)) \in \mathfrak{S}$ puisque f est mesurable. ■

Théorème 5.1.11. *Soient $(X, \mathfrak{S})$ un espace mesurable, (E_n) une suite d'espaces métriques séparables, et (f_n) une suite d'applications mesurables : $X \to E_n$. Alors l'application $f : X \to E = \prod_n E_n$ dont les applications coordonnées sont les (f_n) est mesurable.*

Démonstration : Il suffit de montrer que $f^{-1}(U) \in \mathfrak{S}$ pour les ouverts d'une base de la topologie. Or E possède une base formée d'ensembles de la forme $U = \prod_n U_n$, où les U_n sont ouverts (et presque tous égaux à E_n). Alors, pour un tel ensemble,

$$f^{-1}(U) = \bigcap_n f_n^{-1}(U_n) \in \mathfrak{S} \ .$$

■

Corollaire 5.1.12. *Soient $(X, \mathfrak{S})$ un espace mesuré, f et g deux fonctions mesurables de X dans $\mathbb{R}$ ou $\mathbb{C}$. Alors $\bar{f}$, $f + g$ et fg sont mesurables. Il en est de même de $\dfrac{f}{g}$ si g ne s'annule pas.*

Démonstration : La fonction $z \mapsto \bar{z}$ est continue de $\mathbb{C}$ dans $\mathbb{C}$, les fonctions $(u, v) \mapsto u + v$ et $(u, v) \mapsto uv$ sont continues de $\mathbb{R} \times \mathbb{R}$ dans $\mathbb{R}$ et de $\mathbb{C} \times \mathbb{C}$ dans $\mathbb{C}$. La mesurabilité de $\bar{f}$, de $f + g$ et de fg résulte donc immédiatement des énoncés précédents. Et puisque la fonction $z \mapsto \dfrac{1}{z}$ est continue de $\mathbb{R}^*$ dans $\mathbb{R}$ et de $\mathbb{C}^*$ dans $\mathbb{C}$, on voit de même que $1/g$ est mesurable si g est mesurable et ne s'annule pas, donc que $\dfrac{f}{g} = f.\dfrac{1}{g}$ est mesurable sous les mêmes hypothèses. ■

Théorème 5.1.13. *Soient $(X, \mathfrak{S})$ un espace mesurable et E un espace métrique séparable. On suppose que (f_n) est une suite de fonctions mesurables de X dans E, et que la suite (f_n) converge simplement vers f. Alors f est mesurable.*

Démonstration : Soit $(U_m)_{m \in \mathbb{N}}$ une base de la topologie de E. Si V est un ouvert de E, on définit l'ensemble dénombrable $\Delta(V) = \{m : \overline{U_m} \subset V\}$.
Si $f(x) \in V$, il existe un $m \in \Delta(V)$ tel que $f(x) \in U_m$. Et puisque la suite $f_n(x)$ converge vers un point de U_m, il existe un p tel que $f_q(x) \in U_m$ si $q \geqslant p$. Inversement, si $f_q(x) \in U_m$ pour tout $q \geqslant p$, la limite $f(x)$ appartient à $\overline{U_m} \subset V$. On en déduit que

$$f^{-1}(V) = \bigcup_{m \in \Delta(V)} \bigcup_{p \in \mathbb{N}} \bigcap_{q \geqslant p} f_q^{-1}(U_m)$$

et cet ensemble appartient à $\mathfrak{S}$, puisque $f_q^{-1}(U_m) \in \mathfrak{S}$ et que $\mathfrak{S}$ est stable par unions et intersections dénombrables. ■

Lemme 5.1.14. *Soient $(X, \mathfrak{S})$ un espace mesurable, f et g deux fonctions mesurables de X dans $\mathbb{R}$. Alors $\sup(f, g)$ et $\inf(f, g)$ sont mesurables.*

Démonstration : Puisque la topologie de $\mathbb{R}$ possède une base formée des intervalles ouverts $]\alpha, \beta[=]-\infty, \beta[\cap]\alpha, +\infty[$, il suffit, pour montrer la mesurabilité d'une fonction h, de démontrer que les ensembles $h^{-1}(]-\infty, \beta[)$ (qu'on notera souvent $\{h < \beta\}$) et les ensembles $h^{-1}(]\alpha, +\infty[)$ (qu'on notera souvent $\{h > \alpha\}$ sont dans $\mathfrak{S}$.
De plus, si les ensembles $\{h > \alpha\}$ sont dans $\mathfrak{S}$, on a $\{h \leqslant \alpha\} = \{h > \alpha\}^c \in \mathfrak{S}$ et $\{h < \beta\} = \bigcup_m \{h \leqslant \beta - 2^{-m}\} \in \mathfrak{S}$. Il suffit donc de montrer que les ensembles $\{h > \alpha\}$ sont mesurables pour tout $\alpha \in \mathbb{R}$ (ou même pour tout $\alpha \in \mathbb{Q}$).
Or, pour $h = \sup(f, g)$, on a $\{h > \alpha\} = \{f > \alpha\} \cup \{g > \alpha\}$. Et pour $h = \inf(f, g)$, on a $\{h > \alpha\} = \{f > \alpha\} \cap \{g > \alpha\}$. ■

Théorème 5.1.15. *Soient $(X, \mathfrak{S})$ un espace mesurable et (f_n) une suite de fonctions mesurables de X dans $\mathbb{R}$. Alors $\sup_n f_n$ et $\inf_n f_n$ sont mesurables de X dans $\overline{\mathbb{R}}$. Et ces fonctions sont mesurables de X dans $\mathbb{R}$ si elles sont partout finies.*

Démonstration : Il résulte du lemme précédent par une récurrence simple que les fonctions $g_n = \sup_{m \leqslant n} f_m$ et $h_n = \inf_{m \leqslant n} f_m$ sont mesurables. Et puisque l'on a dans $\overline{\mathbb{R}}$: $\sup_n f_n = \lim g_n$ et $\inf f_n = \lim_n h_n$, la conclusion en découle.
On voit de même que $\sup_n f_n$ (resp. $\inf_n f_n$) est mesurable de X dans $\mathbb{R}$ si elle est partout finie. ■

Fonctions élémentaires.

Définition 5.1.16. *Soit $(X, \mathfrak{S})$ un espace mesurable. Une fonction mesurable f sur X à valeurs dans $\mathbb{R}$ ou $\mathbb{C}$ sera dite* élémentaire *si elle ne prend qu'un nombre fini de valeurs.*

Proposition 5.1.17. *L'ensemble des fonctions élémentaires est l'espace vectoriel de fonctions sur X engendré par les fonctions caractéristiques $\mathbb{1}_A$ des éléments A de $\mathfrak{S}$.*

Démonstration : Si g est une fonction mesurable, le singleton $\{\lambda\}$ est fermé dans $\mathbb{R}$ ou $\mathbb{C}$. Il en résulte que son image réciproque $g^{-1}(\lambda)$ appartient à $\mathfrak{S}$.
Si f est une fonction élémentaire, avec $f(X) = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, l'ensemble $A_j = f^{-1}(\lambda_j)$ appartient à $\mathfrak{S}$ et on a clairement $f = \sum_{j=1}^m \lambda_j \mathbb{1}_{A_j}$, puisque les A_j forment une partition de X.
Inversement, si $f = \sum_1^m \lambda_j \mathbb{1}_{B_j}$, avec les $B_j \in \mathfrak{S}$, il résulte des théorèmes précédents que f est somme de fonctions mesurables, donc elle-même mesurable, et que les valeurs de f appartiennent à l'ensemble $\{\sum_{j \in J} \lambda_j : J \subset \{1, 2, \ldots, m\}\ \}$ qui possède au plus 2^m éléments.
Donc f est élémentaire. ■

Théorème 5.1.18. *Soient $(X, \mathfrak{S})$ un espace mesurable, et $f : X \to \mathbb{R}$ une fonction mesurable. Alors il existe une suite (f_n) de fonctions mesurables élémentaires qui converge vers f.*

Démonstration : Posons pour $n \in \mathbb{N}$ et p entier dans $[-2^{2n}, +2^{2n}]$:

$$A_{n,p} = \{x \in X : p.2^{-n} \leqslant f(x) < (p+1).2^{-n}\} \ .$$

Ces ensembles sont tous mesurables. La fonction $f_n = \sum_{p=-2^{2n}}^{2^{2n}} \dfrac{p}{2^n}.\mathbb{1}_{A_{n,p}}$ est alors une fonction élémentaire. Et on a pour tout x et tout n :

$$|f(x) - f_n(x)| < 2^{-n} \text{ ou } |f(x)| > 2^n \ ,$$

ce dont on conclut que $f_n(x) \to f(x)$. ■

Théorème 5.1.19. *Soient $(X, \mathfrak{S})$ un espace mesurable, et $f : X \to \mathbb{C}$ une fonction mesurable. Alors il existe une suite (f_n) de fonctions mesurables élémentaires qui converge vers f.*

DÉMONSTRATION : Puisque les fonctions $\Re e$ et $\Im m$ sont continues de $\mathbb{C}$ dans $\mathbb{R}$, les fonctions $g = \Re e(f)$ et $h = \Im m(f)$ sont mesurables réelles, donc toutes deux limites de suites de fonctions élémentaires (g_n) et (h_n). Il est clair alors que la suite $(g_n + ih_n)$ est une suite de fonctions élémentaires qui converge vers f. ■

5.2 Mesures positives sur un espace mesurable

Définition 5.2.1. *On appelle* mesure *sur l'espace mesurable $(X, \mathfrak{S})$ toute fonction μ définie sur $\mathfrak{S}$ à valeurs dans $\overline{\mathbb{R}}_+ = [0, +\infty]$ qui satisfait $\mu(\emptyset) = 0$ et $\mu(A \cup B) = \mu(A) + \mu(B)$ lorsque $A \cap B = \emptyset$.*

Lorsque $A \subset B$, si on pose $C = B \setminus A$, il résulte de la définition ci-dessus qu'on doit avoir $\mu(B) = \mu(A) + \mu(C) \geqslant \mu(A)$.

Définition 5.2.2. *Une mesure μ sur l'espace mesurable $(X, \mathfrak{S})$ est dite σ-*finie *s'il existe une suite (A_n) d'éléments de $\mathfrak{S}$ telle que $X = \bigcup_n A_n$ et $\mu(A_n) < +\infty$ pour tout n.*

Définition 5.2.3. *Une mesure μ sur l'espace mesurable $(X, \mathfrak{S})$ est dite σ-*additive *si, pour toute suite croissante (A_n) d'éléments de $\mathfrak{S}$ on a $\mu(\bigcup_n A_n) = \sup_n \mu(A_n)$.*

On doit remarquer que l'égalité ci-dessus a bien un sens puisque la réunion des A_n appartient à la tribu $\mathfrak{S}$.

Lemme 5.2.4. *Soit μ une mesure sur l'espace mesurable $(X, \mathfrak{S})$. Il y a équivalence entre :*

i) μ est σ-additive.

ii) si (A_n) est une suite décroissante de parties mesurables d'intersection vide et si $\mu(A_0) < +\infty$, alors $\mu(A_n) \to 0$.

iii) pour toute suite (A_n) de parties mesurables, on a $\mu(\bigcup_n A_n) \leqslant \sum_n \mu(A_n)$.

DÉMONSTRATION : *i)* $\Longrightarrow$ *ii)*. Supposons μ σ-additive et la suite (A_n) décroissante d'intersection vide. Alors la suite $B_n = A_0 \setminus A_n$ est croissante et de réunion A_0. Puisque $\mu(A_0) < \infty$, on a $\mu(B_n) + \mu(A_n) = \mu(A_0)$, donc $\mu(A_n) = \mu(A_0) - \mu(B_n)$.
Et puisque $\mu(A_0) = \sup \mu(B_n) = \lim \mu(B_n)$, on conclut que $\mu(A_n) \to 0$.

iii) $\Longrightarrow$ *i)*. Supposons la suite (A_n) croissante. Puisque $\mu(A_m) \leqslant \mu(\bigcup_n A_n)$ pour tout m, on a l'égalité $\mu(\bigcup_n A_n) = \sup_n \mu(A_n)$ dès que $\sup_n \mu(A_n) = \infty$. Supposons donc que $\sup_n \mu(A_n) < \infty$. Si on pose $B_0 = A_0$ et $B_{n+1} = A_{n+1} \setminus A_n$ pour $n \geqslant 0$, on a alors $\bigcup_n B_n = \bigcup_n A_n$ et

$$\sum_{m \leqslant n} \mu(B_m) = \mu(A_0) + \sum_{m=0}^{n} \mu(A_{m+1}) - \mu(A_m) = \mu(A_{n+1})$$

donc $\mu(\bigcup_n A_n) = \mu(\bigcup_n B_n) \leqslant \sum_n \mu(B_n) = \sup_n \mu(A_n) \leqslant \mu(\bigcup_n A_n)$, ce qui montre la σ-additivité de μ.

ii) $\Longrightarrow$ *iii)*. On considère donc une suite décroissante (A_n) d'intersection vide telle que $\mu(A_0) < \infty$. Et on pose $B_n = A_n \setminus A_{n+1}$. Puisque les B_n sont deux-à-deux disjoints, et que $A_n \setminus A_{n+p} = \bigcup_{m=n}^{p+n-1} B_m$, on a

$$\sum_{m=n}^{p+n-1} \mu(B_m) = \mu(A_n) - \mu(A_{n+p}) \leqslant \mu(A_n) \leqslant \mu(A_0)$$

Il en résulte que la série de terme général $r_m = \mu(B_m)$ est convergente. Alors on a $\bigcup_{m \geqslant n} B_m = A_n$, donc $\mu(A_m) \leqslant \sum_{m \geqslant n} r_m$. Et cette quantité, reste d'une série convergente, tend vers 0 quand $n \to \infty$. Donc $\mu(A_n) \to 0$. ■

Définition 5.2.5. *On appellera* espace mesuré *un triplet $(X, \mathfrak{S}, \mu)$, où $(X, \mathfrak{S})$ est un espace mesurable, et μ une mesure σ-additive sur $(X, \mathfrak{S})$.*

Dans tout ce qui suit, la mesure μ sera toujours supposée σ-finie.

Définitions 5.2.6. *Soit $(X, \mathfrak{S}, \mu)$ un espace mesuré. Une partie mesurable N de X sera dite* négligeable *si $\mu(N) = 0$.*
Une propriété $P(x)$ d'un point x de X sera dite vraie presque partout si l'ensemble des points x pour lesquels P n'est pas vraie est négligeable.

Il est clair que si N est négligeable, si $N' \in \mathfrak{S}$ et si $N' \subset N$, alors N' est aussi négligeable. De plus, si (N_k) est une suite de parties négligeables, la réunion $N = \bigcup_k N_k$ est encore négligeable, puisque $\mu(N) \leqslant \sum_k \mu(N_k) = 0$.

Théorème 5.2.7. *Soient $(X, \mathfrak{S})$ un espace mesuré et $\mathfrak{A}$ une algèbre de parties de X telle que $\mathfrak{S}$ soit la tribu engendrée par $\mathfrak{A}$. Soient μ et ν deux mesures σ-additives sur $(X, \mathfrak{S})$ qui coïncident sur $\mathfrak{A}$. S'il existe une suite (A_n) d'éléments de $\mathfrak{A}$ de réunion X tels que $\mu(A_n) < \infty$ pour tout n, alors $\mu = \nu$.*

Démonstration : Quitte à remplacer A_k par $\bigcup_{m \leqslant k} A_m$, on peut supposer la suite (A_k) croissante. Notons $\mathscr{H}_k$ la famille des parties C de X appartenant à $\mathfrak{S}$ telles que $\mu(A_k \cap C) = \nu(A_k \cap C)$. Par hypothèse, on a $\mathfrak{A} \subset \mathscr{H}_k$ puisque $A_k \cap C \in \mathfrak{A}$ quand $C \in \mathfrak{A}$.
La famille $\mathscr{H}_k$ est stable par réunions dénombrables croissantes : en effet, si (Z_n) est une suite d'éléments de $\mathscr{H}_k$ de réunion Z, on a

$$\mu(A_k \cap Z) = \sup_n \mu(A_k \cap Z_n) = \sup_n \nu(A_k \cap Z_n) = \nu(A_k \cap Z)\ .$$

Et de même, si (Z_n) est une suite décroissante dans $\mathscr{H}_k$, d'intersection Z, il résulte du fait que $\mu(A_k \cap Z_0) \leqslant \mu(A_k) < \infty$ et que $\nu(A_k \cap Z_0) \leqslant \nu(A_k) = \mu(A_k) < \infty$ que l'on a $\mu(A_k \cap Z) = \inf_n \mu(A_k \cap Z_n)$ et $\nu(A_k \cap Z) = \inf_n \nu(A_k \cap Z_n)$, donc

$$\mu(A_k \cap Z) = \inf_n \mu(A_k \cap Z_n) = \inf_n \nu(A_k \cap Z_n) = \nu(A_k \cap Z)\ ,$$

ce qui montre que $\mathscr{H}_k$ est stable par intersections dénombrables décroissantes. On déduit alors du théorème 5.1.3 que $\mathfrak{S} \subset \mathscr{H}_k$. Pour tout $Z \in \mathfrak{S}$, on a donc $\mu(Z \cap A_k) = \nu(Z \cap A_k)$ pour tout k, donc $\mu(Z) = \sup_k \mu(Z \cap A_k) = \sup_k \nu(Z \cap A_k) = \nu(Z)$. ■

5.3 Intégrale des fonctions élémentaires positives

On suppose désormais fixé un espace mesuré $(X, \mathfrak{S}, \mu)$.

Lemme 5.3.1. *Si $(A_j)_{1 \leqslant j \leqslant m}$ est une famille finie d'éléments de $\mathfrak{S}$, il existe une partition finie $(C_i)_{i \in I}$ de X en éléments de $\mathfrak{S}$ et des parties $(I_1, I_2, \ldots, I_m)$ de I telles que, pour $1 \leqslant j \leqslant m$, on ait $A_j = \bigcup_{i \in I_j} C_i$. On a alors pour tout j : $\mu(A_j) = \sum_{i \in I_j} \mu(C_i)$.*

Démonstration : On procède par récurrence sur l'entier m. Si $m = 1$, on prend $C_0 = A_1$ et $C_1 = A_1^c$. Et on a $A_1 = \bigcup\{C_j : j = 0\}$.

Si la propriété est vraie pour m et si $(A_j)_{1 \leqslant j \leqslant m+1}$ est une famille de parties mesurables, il existe par l'hypothèse de récurrence une famille $(C_i)_{i \in I}$ et des parties $(I_1, I_2, \ldots, I_m)$ de I avec la propriété ci-dessus pour $(A_1, A_2, \ldots, A_m)$. On pose alors $I' = I \times \{0, 1\}$, $C'_{i,0} = C_i \cap A_{m+1}$ et $C'_{i,1} = C_i \cap A^c_{m+1}$. Quitte à éliminer ceux des $C'_{i,\alpha}$ qui sont vides, on a une partition de X. De plus, pour $k \leqslant m$, on a $A_k = \bigcup\{C'_{i,\alpha} : i \in I_k, \alpha \in \{0, 1\}\}$ et $A_{m+1} = \bigcup\{C'_{i,0} : i \in I\}$.

Enfin, puisque les $(C_i)_{i \in I_j}$ forment une partition de A_j en ensembles mesurables, on a $\mu(A_j) = \sum_{i \in I_j} \mu(C_i)$. ■

Lemme 5.3.2. *Si $(A_j)_{1 \leqslant j \leqslant m}$ est une famille finie de parties mesurables de mesure finie et si la fonction élémentaire $f = \sum_{j=1}^m \lambda_j \mathbb{1}_{A_j}$ est nulle, alors $\sum_j \lambda_j \mu(A_j) = 0$.*

En appliquant le lemme précédent, on trouve une partition finie $(C_i)_{i \in I}$ de X et des parties I_j de I telles que $A_j = \bigcup_{i \in I_j} C_i$. Si i appartient à l'un des I_j, on a nécessairement $\mu(C_i) \leqslant \mu(A_j) < \infty$.
Pour tout $i^* \in I$ tel que $C_{i^*} \neq \emptyset$, on peut trouver $x \in C_{i^*}$. On a alors $x \notin C_i$ pour $i \neq i^*$, et

$$0 = f(x) = \sum_j \lambda_j \mathbb{1}_{A_j}(x) = \sum_j \lambda_j \sum_{i \in I_j} \mathbb{1}_{C_i}(x) = \sum \{\lambda_j : i^* \in I_j\} \ .$$

On a donc soit $\sum_j \{\lambda_j : i^* \in I_j\} = 0$, soit $C_{i^*} = \emptyset$, donc $\mu(C_{i^*}) = 0$. Et puisque

$$\sum_j \lambda_j \mu(A_j) = \sum_j \lambda_j \Big(\sum_{i \in I_j} \mu(C_i) \Big) = \sum_{i \in I} \mu(C_i). \sum \{\lambda_j : i \in I_j\} \ ,$$

on conclut que $\sum_j \lambda_j \mu(A_j) = 0$. ■

Notons Γ l'ensemble des fonctions élémentaires à valeurs dans $\mathbb{R}^+$.

Théorème 5.3.3. *Il existe sur Γ une fonction Φ à valeurs dans $\overline{\mathbb{R}}_+$ vérifiant les propriétés suivantes :*

i) $\Phi(\mathbb{1}_A) = \mu(A)$ *si* $A \in \mathfrak{S}$.

ii) $\Phi(f + g) = \Phi(f) + \Phi(g)$, *et* $\Phi(\lambda f) = \lambda \Phi(f)$ *si* $\lambda > 0$

iii) $\Phi(f) \leqslant \Phi(g)$ *si* $f \leqslant g$.

Démonstration : Pour $f \in \Gamma$, $f(X)$ est un ensemble fini $\{\lambda_1, \lambda_2, \ldots, \lambda_m\} \subset \mathbb{R}^+$ et si on pose $A_j = f^{-1}(\lambda_j)$, on a $f = \sum_j \lambda_j \mathbb{1}_{A_j}$. On pose alors $\Phi(f) = \sum_j \lambda_j \mu(A_j) \in \overline{\mathbb{R}}_+$. Et cette quantité est finie si et seulement si les $\mu(A_j)$ sont finis chaque fois que $\lambda_j \neq 0$.

Si on a $f = \sum_j \lambda_j \mathbb{1}_{A_j} = \sum_k \nu_k \mathbb{1}_{B_k}$ avec $\mu(A_j) < \infty$ et $\mu(B_k) < \infty$, $\lambda_j \in \mathbb{R}$ et $\nu_k \in \mathbb{R}$, on a $\sum_j \lambda_j \mathbb{1}_{A_j} - \sum_k \nu_k \mathbb{1}_{B_k} = 0$ et il résulte du lemme précédent que

$$\sum_j \lambda_j \mu(A_j) - \sum_k \nu_k \mu(B_k) = 0 \ ,$$

donc que la quantité $\sum_j \lambda_j \mu(A_j)$ est égale à $\Phi(f)$ dès que $f = \sum_j \lambda_j \mathbb{1}_{A_j}$. En particulier, on a $\Phi(\mathbb{1}_A) = \mu(A)$.
Si $f = \sum_j \lambda_j \mathbb{1}_{A_j}$ et $g = \sum_k \nu_k \mathbb{1}_{B_k}$, on a $f + g = \sum_j \lambda_j \mathbb{1}_{A_j} + \sum_k \nu_k \mathbb{1}_{B_k}$, donc

$$\Phi(f+g) = \sum_j \lambda_j \mu(A_j) + \sum_k \nu_k \mu(B_k) = \Phi(f) + \Phi(g) \ .$$

De même $\lambda f = \sum_j (\lambda.\lambda_j) \mathbb{1}_{A_j}$ et $\Phi(\lambda f) = \sum_j (\lambda.\lambda_j)\mu(A_j) = \lambda. \sum_j \lambda_j \mu(A_j) = \lambda\Phi(f)$.
Enfin, la propriété *iii)* découle clairement de la propriété *ii)* puisque si $f \leqslant g$, la fonction $h = g - f$ est encore dans Γ et que $\Phi(g) = \Phi(f) + \Phi(h) \geqslant \Phi(f)$. ■

On notera désormais $\int f(x)\, d\mu(x)$ et on appellera *intégrale de la fonction f* la quantité $\Phi(f)$.

Lemme 5.3.4. *Si (g_m) est une suite décroissante de fonctions élémentaires positives telle que $\inf_n g_n = 0$ et $\int g_0(x)\, d\mu(x) < \infty$, alors $\int g_n(x)\, d\mu(x) \to 0$.*

Démonstration : Puisque g_0 ne prend qu'un nombre fini de valeurs réelles positives, il existe $M = \sup_{x \in X} g_0(x) < \infty$.
Pour $0 < \varepsilon < 1$, notons $A_{n,\varepsilon}$ l'ensemble mesurable $\{x : g_n(x) > \varepsilon g_0(x)\}$. Puisque la suite (g_n) est décroissante, on a $A_{n+1,\varepsilon} \subset A_{n,\varepsilon}$. Et puisque $g_n(x) = 0$ en tout point où $g_0(x) = 0$, on voit que $\bigcap_n A_{n,\varepsilon} = \emptyset$.
L'ensemble $A_{0,\varepsilon}$ est égal à l'ensemble $\{g_0 > 0\}$ qui est réunion finie d'ensembles de mesure finie puisque $\int g_0(x)\, d\mu(x) < \infty$. Donc $\mu(A_{0,\varepsilon}) < \infty$. Il résulte donc de la σ-additivité de la mesure que $\mu(A_{n,\varepsilon}) \to 0$ pour tout $\varepsilon \in]0, 1[$.
Enfin on a pour tout n : $g_n \leqslant \varepsilon.g_0 + M.\mathbb{1}_{A_{n,\varepsilon}}$. En effet, pour $x \in X$, ou bien $g_n(x) \leqslant \varepsilon g_0(x)$, ou bien $x \in A_{n,\varepsilon}$ et on a alors $g_n(x) \leqslant g_0(x) \leqslant M = M.\mathbb{1}_{A_{n,\varepsilon}}(x)$. Il résulte de cette inégalité que

$$\int g_n(x)\, d\mu(x) \leqslant \varepsilon \int g_0(x)\, d\mu(x) + M\mu(A_{n,\varepsilon}) \ .$$

Alors si $\eta > 0$ est donné, on peut trouver $\varepsilon \in]0, 1[$ tel que $\varepsilon \int g_0(x)\, d\mu(x) < \eta/2$, puis p tel que $\mu(A_{n,\varepsilon}) < \eta/2$ si $n \geqslant p$. On aura alors $\int g_n(x)\, d\mu(x) < \eta$ si $n \geqslant p$. ■

5.4 Intégrale des fonctions positives

Lemme 5.4.1. *Si $f : X \to \overline{\mathbb{R}}_+$ est mesurable, il existe une suite croissante (f_n) de fonctions positives élémentaires qui tend vers f.*

Démonstration : La preuve est semblable à celle du théorème 5.1.18. Pour $n \in \mathbb{N}$ et $p \in \mathbb{N}$ tel que $p < 2^{2n}$, on pose $A_{n,p} = \{x : p.2^{-n} < f(x) \leqslant (p+1).2^{-n}\}$, $A_{n,2^{2n}} = \{x : f(x) > 2^n\}$ et

$$f_n(x) = \sum_{p=0}^{2^{2n}} p.2^{-n}.\mathbb{1}_{A_{n,p}} \ .$$

La fonction f_n est clairement une fonction élémentaire. Puisque les $(A_{n,p})$ sont deux-à-deux disjoints pour un n fixé, la fonction f_n vaut, en un point x,

- 0 si x n'appartient à aucun des $A_{n,p}$, c'est-à-dire si $f(x) = 0$,
- 2^n si $f(x) > 2^n$,
- $\dfrac{p}{2^n}$ si $\dfrac{p}{2^n} < f(x) \leqslant \dfrac{p+1}{2^n}$.

En particulier, on a $f_n \leqslant f$. Et on a $f_{n+1}(x) = f_n(x)$ si $f(x) = 0$, ou si $f(x) > 2^{n+1}$ ou $\dfrac{2p}{2^{n+1}} < f(x) \leqslant \dfrac{2p+1}{2^{n+1}}$, alors que $f_n(x) < f_{n+1}(x)$ sinon.
Si $f(x) = \infty$, on a $f_n(x) = 2^n$ pour tout n donc $f_n(x) \to f(x)$.
Et sinon, on a $f(x) > 2^n$ ou $0 \leqslant f(x) - f_n(x) < 2^{-n}$, donc $f_n(x) \to f(x)$. ■

Théorème 5.4.2. *Si (f_n) et (g_m) sont deux suites croissantes de fonctions élémentaires positives telles que $\sup_n f_n = \sup_m g_m$, alors $\lim_n \int f_n(x)\,d\mu(x) = \lim_m \int g_m(x)\,d\mu(x)$.*

Démonstration : Les suites $\left(\int f_n(x)\,d\mu(x)\right)$ et $\left(\int g_m(x)\,d\mu(x)\right)$ sont croissantes, donc convergent dans $\overline{\mathbb{R}}_+$. Fixons $m \in \mathbb{N}$.
Si $\int g_m(x)\,d\mu(x) = \infty$, il existe un $A \in \mathfrak{S}$ et un $\lambda > 0$ tels que $\mu(A) = \infty$ et $g_m = \lambda$ sur A. Alors les ensembles $A_n = \{x \in A : f_n(x) > \lambda/2\}$ forment une suite croissante de réunion A. Et puisque l'on a $\mu(A) = \sup \mu(A_n)$, on a $\mu(A_n) \to \infty$ et $\int f_n(x)\,d\mu(x) \geqslant \dfrac{\lambda}{2}\mu(A_n)$, donc $\int f_n(x)\,d\mu(x) \to \infty$ et $\sup_n \int f_n(x)\,d\mu(x) = \infty \geqslant \sup_k \int g_k(x)\,d\mu(x)$.
Dans le cas contraire, $\int g_m(x)\,d\mu(x) < \infty$. Alors la suite (h_n) de fonctions élémentaires positives définie par $h_n(x) = \sup(g_m(x) - f_n(x), 0)$ est décroissante et converge en tout point vers 0, puisque $\sup_n f_n \geqslant g_m$. Il résulte alors du lemme 5.3.4 que $\int h_n(x)\,d\mu(x) \to 0$, donc que

$$\int f_n(x)\,d\mu(x) \geqslant \int (g_m - h_n)(x)\,d\mu(x) = \int g_m(x)\,d\mu(x) - \int h_n(x)\,d\mu(x) \to \int g_m(x)\,d\mu(x)\ ,$$

et finalement que $\sup_n \int f_n(x)\,d\mu(x) = \lim_n \int f_n(x)\,d\mu(x) \geqslant \int g_m(x)\,d\mu(x)$. On en déduit que, dans les deux cas, $\sup_m \int g_m(x)\,d\mu(x) \leqslant \sup_n \int f_n(x)\,d\mu(x)$.
On montre de même que $\sup_m \int f_n(x)\,d\mu(x) \leqslant \sup_n \int g_m(x)\,d\mu(x)$ en échangeant les suites (f_n) et (g_m). ■

Définition de l'intégrale des fonctions positives.

Il résulte de ce qui précède que toute fonction mesurable f à valeurs dans $\overline{\mathbb{R}}_+$ est limite d'une suite croissante (f_n) de fonctions de Γ, et que la limite des intégrales de ces fonctions ne dépend que de f. On définit donc, pour une telle fonction,

$$\int f(x)\,d\mu(x) = \lim \int f_n(x)\,d\mu(x)$$

pour toute suite croissante de fonctions élémentaires positives de limite f.

Théorème 5.4.3. *Si f et g sont des fonctions mesurables à valeurs dans $\overline{\mathbb{R}}_+$, on a*

$$\int (f+g)(x)\,d\mu(x) = \int f(x)\,d\mu(x) + \int g(x)\,d\mu(x)\ .$$

En particulier, si $f \leqslant g$, on a $\int f(x)\,d\mu(x) \leqslant \int g(x)\,d\mu(x)$. Et pour $\lambda > 0$, on a $\int \lambda f(x)\,d\mu(x) = \lambda \int f(x)\,d\mu(x)$.

Démonstration : Si (f_n) et (g_n) sont deux suites croissantes de fonctions positives élémentaires qui convergent respectivement vers f et g, la suite (λf_n) converge en croissant vers f, et la suite $(f_n + g_n)$ converge en croissant vers $f + g$. Il en résulte que

$$\int \lambda f(x)\,d\mu(x) = \lim_n \int \lambda f_n(x)\,d\mu(x) = \lambda \lim_n \int f_n(x)\,d\mu(x) = \lambda \int f(x)\,d\mu(x)$$

et que

$$\begin{aligned}\int (f+g)(x)\,d\mu(x) &= \lim_n \int (f_n+g_n)(x)\,d\mu(x)\\ &= \lim_n \int f_n(x)\,d\mu(x) + \lim_n \int g_n(x)\,d\mu(x) = \int f(x)\,d\mu(x) + \int g(x)\,d\mu(x)\ .\end{aligned}$$

Et si $f \leqslant g$, il existe $h \geqslant 0$ telle que $g = f + h$. On a alors

$$\int g(x)\,d\mu(x) = \int f(x)\,d\mu(x) + \int h(x)\,d\mu(x) \geqslant \int f(x)\,d\mu(x)\ . \qquad \blacksquare$$

Définition 5.4.4. *Une fonction mesurable positive $f : X \to \overline{\mathbb{R}}_+$ est dite* intégrable *si $\int f(x)\,d\mu(x) < \infty$.*

Définition 5.4.5. *Si f est une fonction mesurable positive sur X et A un élément de $\mathfrak{S}$, on appelle* intégrale de f sur A *et on note $\int_A f(x)\,d\mu(x)$ l'intégrale $\int \mathbb{1}_A(x).f(x)\,d\mu(x)$.*

Le théorème de convergence monotone.

Ce théorème, appelé encore théorème de Beppo-Levi, est, avec le théorème de convergence dominée de Lebesgue qui s'en déduit, le théorème fondamental de la théorie de l'intégration.

Théorème 5.4.6. (Convergence monotone) *Si (f_n) est une suite croissante de fonctions mesurables positives, on a*

$$\int \sup_n f_n(x)\,d\mu(x) = \sup \int f_n(x)\,d\mu(x)\ .$$

Démonstration : Posons $f = \sup_n f_n$ Puisque $f \geqslant f_n$ pour tout n, on a nécessairement $\int f(x)\,d\mu(x) \geqslant \int f_n(x)\,d\mu(x)$, donc $\int f(x)\,d\mu(x) \geqslant \sup_n \int f_n(x)\,d\mu(x)$.

Si, pour chaque n, on choisit une suite croissante $(f_{n,k})_k$ de fonctions élémentaires positives telle que $f_n = \sup_k f_{n,k}$, puis si on pose $g_n(x) = \sup\{f_{m,k}(x) : m \leqslant n,\ k \leqslant n\}$, on construit une suite croissante (g_n) de fonctions élémentaires positives telle que

$$\sup_n g_n(x) = \sup_{m,k} f_{m,k}(x) = \sup_m \big(\sup_k f_{m,k}(x)\big) = \sup_m f_m(x) = f(x)\ .$$

Il en résulte que $\int f(x)\,d\mu(x) = \sup_n \int g_n(x)\,d\mu(x)$. Et puisque, pour $m \leqslant n$ et $k \in \mathbb{N}$, on a $f_{m,k}(x) \leqslant f_m(x) \leqslant f_n(x)$, on a $g_n \leqslant f_n$, donc

$$\sup_n \int f_n(x)\,d\mu(x) \leqslant \int f(x)\,d\mu(x) = \sup_n \int g_n(x) \leqslant \sup_n \int f_n(x)\,d\mu(x)\ .$$

Et ceci achève la démonstration. $\blacksquare$

Corollaire 5.4.7. *Soit (f_n) une suite de fonctions mesurables positives qui décroît partout vers 0. Si f_0 est intégrable, on a $\int f_n(x)\,d\mu(x) \to 0$.*

Démonstration : Il suffit d'appliquer le théorème précédent à la suite $g_n = f_0 - f_n$ qui croît vers f_0. Alors $f_0 = f_n + g_n$ et on a $\int f_0(x)\,d\mu(x) = \int f_n(x)\,d\mu(x) + \int g_n(x)\,d\mu(x)$, ainsi que $\int g_n(x)\,d\mu(x) \to \int f_0(x)\,d\mu(x)$. On en conclut que $\int f_n(x)\,d\mu(x) \to 0$. ■

Lemme 5.4.8. *Soit $f : X \to \overline{\mathbb{R}}_+$ une fonction mesurable. Alors $\int f(x)\,d\mu(x) = 0$ si et seulement si f est nulle presque partout.*

Démonstration : Si f est nulle presque partout, l'ensemble $N = \{f(x) > 0\}$ est négligeable. Alors la fonction $g_n = n.\mathbb{1}_N$ est élémentaire et on a $\int g_n(x)\,d\mu(x) = n\mu(N) = 0$. Donc $\sup_n \int g_n(x)\,d\mu(x) = 0$. Puisque l'on a $f \leqslant \sup_n g_n$, il résulte de ce qui précède que

$$\int f(x)\,d\mu(x) \leqslant \int \sup_n g_n(x)\,d\mu(x) = \sup_n \int g_n(x)\,d\mu(x) = 0\ .$$

Inversement, si $\int f(x)\,d\mu(x) = 0$ et si $N_k := \{f > 2^{-k}\}$, on a

$$2^{-k}\mu(N_k) = \int 2^{-k}\mathbb{1}_{N_k}(x)\,d\mu(x) \leqslant \int f(x)\,d\mu(x) = 0\ ,$$

donc $\mu(N_k) = 0$, et $\mu(\{f > 0\}) = \mu(\bigcup_k N_k) \leqslant \sum_k \mu(N_k) = 0$. Ceci montre que f est nulle presque partout. ■

Lemme 5.4.9. *Soit $f : X \to \overline{\mathbb{R}}_+$ une fonction intégrable. Alors f est finie presque partout.*

Démonstration : Soit A_k l'ensemble $\{f > 2^k\}$. On a $f \geqslant g_k := 2^k.\mathbb{1}_{A_k}$, donc

$$2^k.\mu(A_k) = \int g_k(x)\,d\mu(x) \leqslant \int f(x)\,d\mu(x) = M < \infty\ .$$

Il en résulte que $\mu(A_k) \leqslant 2^{-k}M$. Et puisque, pour tout k, $\mu(\bigcap_n A_n) \leqslant \mu(A_k)$, l'ensemble N des points où f est infinie, égal à $\bigcap_n A_n$, vérifie $\mu(N) = 0$. On conclut que f est finie presque partout. ■

Théorème 5.4.10. *Soient f et g deux fonctions mesurables de X dans $\overline{\mathbb{R}}_+$. Si f et g sont égales presque partout, alors $\int f(x)\,d\mu(x) = \int g(x)\,d\mu(x)$.*

Démonstration : Soit h la fonction qui vaut ∞ en tout point x où $f(x) \neq g(x)$ et 0 ailleurs. Alors h est nulle presque partout, donc d'intégrale nulle. De plus, on a $f \leqslant g + h$ et $g \leqslant f + h$. Donc

$$\begin{aligned}\int f(x)\,d\mu(x) \leqslant \int g(x)\,d\mu(x) + \int h(x)\,d\mu(x) = \int g(x)\,d\mu(x)\\ \leqslant \int f(x)\,d\mu(x) + \int h(x)\,d\mu(x) = \int f(x)\,d\mu(x)\ ,\end{aligned}$$

ce qui achève la démonstration. ■

Lemme 5.4.11. (Fatou) *Soit (f_n) une suite de fonctions mesurables positives. On a*

$$\int \liminf_{n\to\infty} f_n(x)\, d\mu(x) \leqslant \liminf_{n\to\infty} \int f_n(x)\, d\mu(x)\ .$$

Démonstration : Posons $g_n(x) = \inf_{p\geqslant n} f_p(x)$. La fonction g_n est mesurable et positive. De plus, la suite (g_n) converge en croissant vers $\liminf_{n\to\infty} f_n$. Puisque l'on a $g_n \leqslant f_p$ pour tout $p \geqslant n$, on a $\int g_n(x)\, d\mu(x) \leqslant \inf_{p\geqslant n} \int f_p(x)\, d\mu(x)$. On en déduit, par convergence monotone, que

$$\begin{aligned}\int \liminf_{n\to\infty} f_n(x)\, d\mu(x) &= \int \sup_n g_n(x)\, d\mu(x) = \sup_n \int g_n(x)\, d\mu(x) \\ &\leqslant \sup_n \inf_{p\geqslant n} \int f_p(x)\, d\mu(x) = \liminf_{n\to\infty} \int f_n(x)\, d\mu(x)\ . \quad \blacksquare\end{aligned}$$

Lemme 5.4.12. (Borel-Cantelli) *Soit (A_n) une suite de parties mesurables de X. Si la série de terme général $\mu(A_n)$ est convergente, alors presque tout point de X n'appartient qu'à un nombre fini des (A_n).*

Démonstration : Si s_n désigne la fonction $\sum_{p=0}^n \mathbb{1}_{A_n}$, on a $\int s_n(x)\, d\mu(x) = \sum_{p=0}^n \mu(A_p)$. Et puisque la suite (s_n) est croissante, il résulte du théorème de convergence monotone que

$$\int \sup_n s_n(x)\, d\mu(x) = \sup_n \int s_n(x)\, d\mu(x) = \sup_n \sum_{p=0}^n \mu(A_p) = \sum_0^\infty \mu(A_n) < \infty\ .$$

Et puisque la fonction $s = \sup_n s_n$ est intégrable, elle est finie presque partout. Alors si $s(x) \leqslant k$, le point x appartient à au plus k parmi les ensembles A_n. $\blacksquare$

5.5 Fonctions intégrables

Définition 5.5.1. *Une fonction mesurable f définie sur X à valeurs réelles ou complexes est dite* intégrable *si la fonction positive $|f|$ est intégrable, c'est-à-dire si $\int |f(x)|\, d\mu(x) < \infty$.*

Il résulte immédiatement du lemme 5.4.10 que si g est une fonction mesurable égale presque partout à la fonction intégrable f, alors g est intégrable elle aussi.
Si $f : X \to \mathbb{K}$ est mesurable et définie presque partout, c'est-à-dire qu'il existe un ensemble négligeable $N \in \mathfrak{S}$ tel que f soit définie sur $X \setminus N$ et que $f^{-1}(U) \in \mathfrak{S}$ pour tout ouvert U de $\mathbb{K}$, alors il existe des prolongements mesurables $g : X \to \mathbb{K}$ partout définis (par exemple en posant $g = 0$ hors du domaine de de f). Deux tels prolongements sont égaux presque partout, donc tous deux intégrables ou tous deux non intégrables. On peut donc ainsi définir la notion d'intégrabilité pour une fonction définie presque partout.

Lemme 5.5.2. *L'ensemble $\mathscr{L}^1(X,\mathbb{K})$ des fonctions intégrables de X dans $\mathbb{K}$ est un espace vectoriel sur $\mathbb{K}$.*

Démonstration : Si f est intégrable, et si $\lambda \in \mathbb{K}$, alors $|\lambda f| = |\lambda|\,.\,|f|$. Puisque la fonction positive $|f|$ est intégrable et que $|\lambda| \in \mathbb{R}^+$, on a

$$\int |\lambda f(x)|\ d\mu(x) = |\lambda|\,.\int |f(x)|\ d\mu(x) < \infty\ .$$

De plus, si f et g sont intégrables, les fonctions $|f|$ et $|g|$ sont intégrables et on a $|f+g| \leqslant |f| + |g|$. On a donc

$$\int |(f+g)(x)|\ d\mu(x) \leqslant \int (|f(x)| + |g(x)|)\, d\mu(x) = \int |f(x)|\ d\mu(x) + \int |g(x)|\ d\mu(x) < \infty\ ,$$

ce qui montre que $f+g$ est intégrable. ■

Lemme 5.5.3. *Soient f_1, f_2, g_1 et g_2 quatre fonctions intégrables positives. Si $f_1 - f_2$ et $g_1 - g_2$ sont égales,on a*

$$\int f_1(x)\, d\mu(x) - \int f_2(x)\, d\mu(x) = \int g_1(x)\, d\mu(x) - \int g_2(x)\, d\mu(x)\ .$$

Démonstration : Puisque $f_1 + g_2 = f_2 + g_1$, on a

$$\int f_1(x)\, d\mu(x) + \int g_2(x)\, d\mu(x) = \int g_1(x)\, d\mu(x) + \int f_2(x)\, d\mu(x)\ ,$$

et puisque ces nombres sont tous finis, on en déduit la conclusion cherchée. ■

Théorème 5.5.4. *A toute fonction intégrable $f : X \to \mathbb{K}$ on peut associer un élément de $\mathbb{K}$ appelé* intégrale de f, *et noté $\int f(x)\, d\mu(x)$, avec les propriétés suivantes ;*

i) si f est réelle positive, l'intégrale de f coïncide avec celle qui a été définie dans la section précédente.

ii) l'application $f \mapsto \int f(x)\, d\mu(x)$ est $\mathbb{K}$-linéaire sur $\mathscr{L}^1(X, \mathbb{K})$.

iii) $\left|\int f(x)\, d\mu(x)\right| \leqslant \int |f(x)|\ d\mu(x)$ pour toute $f \in \mathscr{L}^1(X, \mathbb{K})$.

Démonstration : Supposons d'abord $\mathbb{K} = \mathbb{R}$. Pour toute fonction mesurable réelle f, on définit les fonctions mesurables positives f_+ et f_- par

$$f_+(x) = \sup(f(x), 0) \qquad \text{et} \qquad f_-(x) = \sup(-f(x), 0)$$

On a alors $f(x) = f_+(x) - f_-(x)$ et $|f(x)| = f_+(x) + f_-(x)$. Puisque $f_+ \leqslant |f|$ et $f_- \leqslant |f|$, ces deux fonctions sont intégrables, et on peut définir

$$\int f(x)\, d\mu(x) = \int f_+(x)\, d\mu(x) - \int f_-(x)\, d\mu(x) \in \mathbb{R}\ .$$

Clairement, si f est positive, on a $f_+ = f$ et $f_- = 0$, et on en déduit $i)$.

Il résulte du lemme précédent que, pour tout couple (g, h) de fonctions positives telles que $f = g - h = f_+ - f_-$, on a $\int f(x)\, d\mu(x) = \int g(x)\, d\mu(x) - \int h(x)\, d\mu(x)$. En particulier, si f et f' appartiennent à $\mathscr{L}^1(X, \mathbb{R})$, on a $f + f' = (f_+ + f'_+) - (f_- + f'_-)$, dont on déduit que

$$\begin{aligned}\int (f+f')(x)\, d\mu(x) &= \int f_+(x)\, d\mu(x) + \int f'_+(x)\, d\mu(x) - \int f_-(x)\, d\mu(x) - \int f'_-(x)\, d\mu(x) \\ &= \int f(x)\, d\mu(x) + \int f'(x)\, d\mu(x)\ .\end{aligned}$$

De même, pour $f \in \mathscr{L}^1(X, \mathbb{R})$ et $\lambda \in \mathbb{R}$, on a $\lambda f = \lambda f_+ - \lambda f_-$ si $\lambda \geqslant 0$ et $\lambda f = |\lambda|\, f_- - |\lambda|\, f_+$ si $\lambda \leqslant 0$, dont on déduit que $\int \lambda f(x)\, d\mu(x) = \lambda \int f(x)\, d\mu(x)$. Et ceci montre $ii)$.

Enfin, on a

$$\left|\int f(x)\,d\mu(x)\right| = \left|\int f_+(x)\,d\mu(x) - \int f_-(x)\,d\mu(x)\right| \leqslant \int f_+(x)\,d\mu(x) + \int f_-(x)\,d\mu(x)$$
$$= \int (f_+(x) + f_-(x))\,d\mu(x) = \int |f(x)|\;d\mu(x)\;.$$

ce qui montre *iii*).

Maintenant, si $\mathbb{K} = \mathbb{C}$, et si $f \in \mathscr{L}^1(X, \mathbb{C})$, on a $|\Re e(f)| \leqslant |f|$ et $|\Im m(f)| \leqslant |f|$. Il en résulte que les fonctions réelles $\Re e(f)$ et $\Im m(f)$ sont intégrables ; on définit alors

$$\int f(x)\,d\mu(x) = \int \Re e(f(x))\,d\mu(x) + i \int \Im m(f(x))\,d\mu(x) \in \mathbb{C}\;.$$

Ceci définit clairement une application $\mathbb{R}$-linéaire sur $\mathscr{L}^1(X, \mathbb{C})$ qui prolonge l'intégrale déjà définie sur $\mathscr{L}^1(X, \mathbb{R})$. On vérifie sans peine, puisque $\Re e(if) = -\Im m(f)$ et $\Im m(if) = \Re e(f)$ que $\int if(x)\,d\mu(x) = i \int f(x)\,d\mu(x)$, dont on déduit la $\mathbb{C}$-linéarité de l'intégrale. Les points *i*) et *ii*) sont donc vérifiés.

Enfin, si on pose $\rho = |f(x)\,d\mu(x)|$ et si on choisit $\theta \in \mathbb{R}$ tel que $\int f(x)\,d\mu(x) = \rho\,\mathrm{e}^{i\theta}$, on a

$$\rho = \Re e\Big(\int f(x)\,\mathrm{e}^{-i\theta}\;d\mu(x)\Big) = \int \Re e(f(x)\,\mathrm{e}^{-i\theta})\,d\mu(x) \leqslant \int \left|\Re e(f(x)\,\mathrm{e}^{-i\theta})\right|\;d\mu(x)$$

et puisque $\left|\Re e(f(x)\,\mathrm{e}^{-i\theta})\right| \leqslant |f(x)|$, on conclut que $\left|\int f(x)\,d\mu(x)\right| = \rho \leqslant \int |f(x)|\;d\mu(x)$, ce qui prouve *iii*). ■

Il résulte de ce qui précède que si f et g sont intégrables et égales presque partout, on a

$$\left|\int f(x)\,d\mu(x) - \int g(x)\,d\mu(x)\right| = \left|\int (f(x) - g(x))\,d\mu(x)\right| \leqslant \int |f(x) - g(x)|\;d\mu(x) = 0\;,$$

c'est-à-dire que f et g ont même intégrale. On considère alors la relation d'équivalence sur $\mathscr{L}^1(X, \mathbb{K})$ définie par $f \equiv g$ si $f = g$ presque partout.

L'ensemble $\mathscr{N}(X)$ des fonctions nulles presque partout est un sous-espace vectoriel de $\mathscr{L}^1(X, \mathbb{K})$, et on notera $L^1(X, \mathbb{K})$ l'espace vectoriel quotient $\mathscr{L}^1(X, \mathbb{K})/\mathscr{N}(X)$ des classes de fonctions intégrables modulo la relation $\equiv$. En particulier l'intégrale d'une fonction ne dépend que de sa classe, ce qui définit une application linéaire de $L^1(X, \mathbb{K})$ dans $\mathbb{K}$.

Théorème 5.5.5. *L'application $f \mapsto \|f\|_1 = \int |f(x)|\;d\mu(x)$ est une norme sur l'espace vectoriel $L^1(X, \mathbb{K})$.*

Démonstration : On voit aisément que si $f \equiv g$, alors $|f| \equiv |g|$; on en déduit que $\|f\|_1$ ne dépend que de la classe de f, c'est-à-dire que $\|.\|_1$ est bien une application sur $L^1(X, \mathbb{K})$ (à valeurs dans $\mathbb{R}^+$).

On a clairement $\|\lambda f\|_1 = |\lambda|\,.\,\|f\|_1$ puisque $|\lambda f| = |\lambda|\,.\,|f|$. Et puisque $|f + g| \leqslant |f| + |g|$, on voit immédiatement que $\|f + g\|_1 \leqslant \int \big(|f(x)| + |g(x)|\big)\,d\mu(x) = \|f\|_1 + \|g\|_1$.

Enfin, si f est la classe d'une fonction g et si $\|f\|_1 = 0$, il résulte du lemme 5.4.8 que $|g(x)| = 0$ presque partout, donc que $g \in \mathscr{N}(X)$, c'est-à-dire que la classe de g est la classe nulle, ou encore que $f = 0$ dans l'espace $L^1(X, \mathbb{K})$. ■

On va maintenant montrer le théorème de convergence dominée de Lebesgue, la pierre angulaire de la théorie de l'intégration.

Théorème 5.5.6. (Convergence dominée) *Soient (f_n) une suite de fonctions mesurables de X dans $\mathbb{K}$ et f une fonction mesurable. On suppose que*

i) la suite $(f_n(x))$ converge vers $f(x)$ pour presque tout x,

ii) il existe une fonction intégrable positive g telle que $|f_n(x)| \leqslant g(x)$ pour tout x.

Alors la suite des intégrales $\left(\int f_n(x)\, d\mu(x)\right)$ converge vers $\int f(x)\, d\mu(x)$.

Démonstration : Si on pose $h_n(x) = \sup_{p \geqslant n} |f_p(x) - f(x)|$, on définit ainsi une suite décroissante de fonctions mesurables de X dans $\overline{\mathbb{R}}_+$. Quitte à remplacer $f(x)$ et chacune des $f_n(x)$ par 0 en tout point de l'ensemble négligeable N où la suite $(f_n(x))$ ne converge pas vers f, on a $h_n(x) \leqslant 2g(x)$ pour tout x, et la suite $(h_n(x))$ décroît partout vers 0. Il résulte alors du corollaire 5.4.7 que $\int h_n(x)\, d\mu(x) \to 0$, et puisque

$$\left| \int f_n(x)\, d\mu(x) - \int f(x)\, d\mu(x) \right| \leqslant \int |f_n(x) - f(x)|\ d\mu(x) \leqslant \int h_n(x)\, d\mu(x)\ ,$$

on obtient la conclusion désirée. ■

On remarque que la fonction intégrable g n'intervient pas dans la conclusion. Mais *l'existence d'une telle fonction* est la condition de validité du théorème.

Lemme 5.5.7. *Soit (u_n) une suite de fonctions mesurables. Si la série $\sum_n \int |u_n(x)|\ d\mu(x)$ converge, alors la série de terme général $(u_n(x))$ converge absolument pour presque tout x vers une fonction intégrable s, et on a*

$$\int \Big| s(x) - \sum_{p \leqslant n} u_p(x) \Big|\, d\mu(x) \leqslant \sum_{p=n+1}^{\infty} \int |u_p(x)|\ d\mu(x)\ .$$

Démonstration : Si on pose $s_n^*(x) = \sum_{p=0}^{n} |u_p(x)|$, on définit une suite croissante de fonctions positives sur X. Il résulte alors du théorème de convergence monotone que $s^* = \sup_n s_n^* = \sum_{n=0}^{\infty} |u_p(x)|$ vérifie

$$\int s^*(x)\, d\mu(x) = \sup_n \int s_n^*(x)\, d\mu(x) = \sup_n \sum_{p=0}^{n} \int |u_p(x)|\ d\mu(x) = \sum_{n=0}^{\infty} \int |u_n(x)|\ d\mu(x) < \infty\ ,$$

donc que s^* est intégrable, et en particulier presque partout finie. Alors, en chaque point de l'ensemble $Z = \{s^* < \infty\}$, la série $(u_n(x))$ converge absolument.

Si on pose maintenant $s_n(x) = \sum_{p=0}^{n} u_p(x)$, la suite $s_n(x)$ converge en tout point de Z vers $s(x) = \sum_{n=0}^{\infty} u_n(x)$. On a pour tout $q > n$:

$$|s_q(x) - s_n(x)| = \left| \sum_{p=n+1}^{q} u_p(x) \right| \leqslant \sum_{p=n+1}^{q} |u_p(x)| \leqslant \sum_{p=n+1}^{\infty} |u_p(x)| \leqslant s^*(x)\ ,$$

donc, par convergence dominée de $|s_q - s_n|$ vers $|s - s_n|$,

$$\begin{aligned} \int |s(x) - s_n(x)|\ d\mu(x) &= \lim_{q \to \infty} \int |s_q(x) - s_n(x)|\ d\mu(x) \\ &\leqslant \lim_{q \to \infty} \int \sum_{p=n+1}^{q} |u_p(x)|\ d\mu(x) = \sum_{p=n+1}^{\infty} \int |u_p(x)|\ d\mu(x)\ , \end{aligned}$$

ce qui achève la démonstration. ■

Théorème 5.5.8. *L'espace normé $L^1(X, \mathfrak{S}, \mu)$ est complet.*

DÉMONSTRATION : En vertu du théorème 2.1.12, il suffit de montrer que toute série normalement convergente dans L^1 est convergente. Si (u_n) est une série normalement convergente, on a

$$\sum_{n=0}^{\infty} \int |u_n(x)| \, d\mu(x) = \sum \|u_n\|_1 < \infty \; .$$

Il résulte alors du lemme précédent que, pour tout n,

$$\int |s(x) - s_n(x)| \, d\mu(x) \leqslant \sum_{p=n+1}^{\infty} \|u_p\|_1 \; ,$$

ce qui montre que $\|s - s_n\|_1 \leqslant \sum_{p=n+1}^{\infty} \|u_p\|_1$, donc que (s_n) converge vers s dans L^1 puisque le reste de la série de terme général $(\|u_n\|_1)$ tend vers 0. Et ceci montre que s est, dans L^1, la somme de la série (u_n). ■

Théorème 5.5.9. *Si la suite (f_n) converge dans L^1 vers une fonction f, il existe une sous-suite (f_{n_k}) qui converge presque partout vers f.*

DÉMONSTRATION : On peut trouver une suite strictement croissante (n_k) d'entiers telle que $\|f - f_{n_k}\|_1 < 2^{-k}$. On a alors

$$\left\|f_{n_{k+1}} - f_{n_k}\right\|_1 \leqslant \left\|f_{n_{k+1}} - f\right\|_1 + \|f - f_{n_k}\|_1 \leqslant 2^{-k} + 2^{-k-1} < 2^{1-k} \; ,$$

ce qui montre que la série de terme général $g_k = f_{n_{k+1}} - f_{n_k}$ est normalement convergente. Il résulte alors du lemme 5.5.7 que la série de terme général (g_k) converge absolument presque partout vers une fonction s, et que, si $r_k = \sum_{n=k+1}^{\infty} g_n$, on a

$$\int \left| s(x) - \sum_{n \leqslant k} g_k(x) \right| d\mu(x) = \int |r_k(x)| \, d\mu(x) \leqslant \sum_{k+1}^{\infty} \int |g_n| \, d\mu(x) \leqslant \sum_{k+1}^{\infty} 2^{1-n} = 2^{1-k} \to 0 \; ,$$

c'est-à-dire $\left\| s - \sum_{n \leqslant k} g_k \right\|_1 \to 0$. Puisque $\sum_{n \leqslant k} g_k(x) = f_{n_{k+1}}(x) - f_{n_0}(x)$, on a aussi $\left\| \sum_{n \leqslant k} g_k - (f - f_{n_0}) \right\|_1 \to 0$. On en conclut que $s = f - f_{n_0}$ dans L^1. Il en résulte que, pour presque tout x, on a $f_{n_{k+1}}(x) = f_{n_0}(x) + \sum_{n \leqslant k} g_k(x) \to f_{n_0}(x) + s(x) = f(x)$.
Et la suite (f_{n_k}) converge presque partout vers f. ■

Définition 5.5.10. *Si f est une fonction mesurable sur X et A un élément de $\mathfrak{S}$, on dira que f* est intégrable sur *A si $f.\mathbb{1}_A$ est intégrable. On appellera* intégrale de f sur A *et on notera $\int_A f(x) \, d\mu(x)$ l'intégrale de $f.\mathbb{1}_A$.*

5.6 Fonctions définies par des intégrales

On suppose maintenant que $(X, \mathfrak{S}, \mu)$ est un espace mesuré, que T est un espace métrique et que f est une fonction réelle ou complexe définie sur $X \times T$ telle que $f(\,.\,, t)$ soit intégrable pour tout $t \in T$. On considère alors la fonction F définie sur T par $F(t) = \int f(x,t)\, d\mu(x)$.

Théorème 5.6.1. *On suppose que la fonction $x \mapsto f(x,t)$ est mesurable pour tout t, que la fonction $t \mapsto f(x,t)$ est continue pour presque tout x et qu'il existe une fonction intégrable g de X dans $\overline{\mathbb{R}}_+$ telle que $|f(x,t)| \leqslant g(x)$ pour tout t. Alors la fonction $F : t \mapsto \int f(x,t)\, d\mu(x)$ est continue sur T.*

Démonstration : Puisque T est un espace métrique, il suffit de montrer que, pour toute suite (t_n) convergeant vers t, on a $F(t_n) \to F(t)$. Si on considère alors les fonctions $h_n : x \mapsto f(x, t_n)$ et $h : x \mapsto f(x,t)$, on a $h_n(x) \to h(x)$ pour presque tout x et $|h_n(x)| \leqslant g(x)$. Le théorème de convergence dominée montre alors que $\int h_n(x)\, d\mu(x) \to \int h(x)\, d\mu(x)$, c'est-à-dire $F(t_n) \to F(t)$. ∎

Lemme 5.6.2. *On suppose maintenant que T est un intervalle ouvert $]a,b[$ de $\mathbb{R}$, que f est définie sur $X \times T$, que l'application $x \mapsto f(x,t)$ est mesurable pour tout $t \in T$ et que l'application $t \mapsto f(x,t)$ est de classe $\mathscr{C}^1$ pour (presque) tout x. Alors, s'il existe une fonction intégrable $g : X \to \overline{\mathbb{R}}_+$ telle que $\left|\dfrac{\partial f}{\partial t}(x,t)\right| \leqslant g(x)$ pour tout t, la fonction $F : t \mapsto \int f(x,t)\, d\mu(x)$ est de classe $\mathscr{C}^1$sur T, et on a*

$$\frac{dF}{dt}(t_0) = \int \frac{\partial f}{\partial t}(x, t_0)\, d\mu(x)\ .$$

Démonstration : On observe d'abord que, pour tout $t_0 \in T$, on a

$$\frac{\partial f}{\partial t}(x,t_0) = \lim_{k \to \infty} k.\Big(f(x, t_0 + \frac{1}{k}) - f(x,t_0)\Big)\ ,$$

dont on déduit que $x \mapsto \dfrac{\partial f}{\partial t}(x,t_0)$ est mesurable pour tout t_0. Si on fixe $t_0 \in T$ et si on définit

$$h(x,t) = \begin{cases} \dfrac{\partial f}{\partial t}(x,t_0) & \text{si } t = t_0 \\[2mm] \dfrac{f(x,t) - f(x,t_0)}{t - t_0} & \text{si } t \neq t_0 \end{cases}\ ,$$

on vérifie sans peine que $x \mapsto h(x,t)$ est mesurable pour tout t, que $t \mapsto h(x,t)$ est continue sur T pour presque tout x (dès que $t \mapsto f(x,t)$ est $\mathscr{C}^1$), et enfin que, par la formule des accroissements finis, $|h(x,t)| \leqslant g(x)$. Il résulte alors du théorème 5.6.1 que $H : t \mapsto \int h(x,t)\, d\mu(x)$ est continue sur T, et en particulier en t_0. Donc, lorsque t tend vers t_0, on a

$$\frac{F(t) - F(t_0)}{t - t_0} = H(t) \to H(t_0) = \int \frac{\partial f}{\partial t}(x,t_0)\, d\mu(x)\ .$$

On en déduit que F est dérivable en t_0, avec $F'(t_0) = \displaystyle\int \frac{\partial f}{\partial t}(x,t_0)\, d\mu(x)$. De plus, les hypothèses faites sur $\dfrac{\partial f}{\partial t}$ permettent de lui appliquer le théorème 5.6.1 et d'obtenir la continuité de F' sur T. Donc F est de classe $\mathscr{C}^1$. ∎

Théorème 5.6.3. (Différentiation sous le signe somme) *On suppose que* T *est un ouvert de* $\mathbb{R}^d$*, que* f *est définie sur* $X \times T$*, que* $x \mapsto f(x,t)$ *est mesurable pour tout* $t \in T$*, que* $t \mapsto f(x,t)$ *est de classe* $\mathscr{C}^1$ *sur* T *pour presque tout* x *et qu'il existe des fonctions intégrables* $g_j : X \to \overline{\mathbb{R}}_+$ *telles que* $\left|\dfrac{\partial f}{\partial t_j}(x,t)\right| \leqslant g_j(x)$ *pour tout* $t \in T$*.*

Alors la fonction $F : t \mapsto \int f(x,t)\, d\mu(x)$ *est de classe* $\mathscr{C}^1$ *sur* T *et on a pour tout* $j \leqslant d$ *:*

$$\frac{\partial F}{\partial t_j}(t) = \int \frac{\partial f}{\partial t_j}(x,t)\, d\mu(x) \ .$$

Démonstration : Pour montrer que F est de classe $\mathscr{C}^1$, il suffit de montrer que F possède des dérivées partielles continues. En fixant toutes les coordonnées de t sauf la $j^{\text{ème}}$ au voisinage d'un point $t^* \in T$, on déduit du lemme précédent que F possède en t^* une dérivée partielle par rapport à la $j^{\text{ème}}$ variable, donnée par la formule ci-dessus. Et il résulte des hypothèses faites que le théorème 5.6.1 s'applique à $\dfrac{\partial f}{\partial t_j}$, donc que $\dfrac{\partial F}{\partial t_j}$ est continue sur T. ■

5.7 Produits d'espaces mesurables

Soient $(X, \mathfrak{S})$ et $(Y, \mathfrak{T})$ deux espaces mesurables. On considère sur $X \times Y$ la tribu $\mathfrak{S} \otimes \mathfrak{T}$ engendrée par la famille $\{A \times B : A \in \mathfrak{S}, B \in \mathfrak{T}\}$.

Lemme 5.7.1. *La famille* $\mathfrak{A}$ *des réunions finies d'ensembles* $A \times B$ *deux-à-deux disjoints, où* $A \in \mathfrak{S}$ *et* $B \in \mathfrak{T}$*, est une algèbre de parties de* $X \times Y$*.*

Démonstration : Si $C = \bigcup A_j \times B_j$ et $C' = \bigcup A'_k \times B'_k$, on a

$$C \cap C' = \bigcup_{j,k} (A_j \cap A'_k) \times (B_j \cap B'_k) \ ,$$

et ces ensembles sont deux-à-deux disjoints si les $(A_j \times B_j)$ et les $(A'_k \times B'_k)$ l'étaient. Ceci montre la stabilité de $\mathfrak{A}$ par intersection finie. Et puisque $C^c = \bigcap (A_j \times B_j)^c$, il suffit de montrer que $(A \times B)^c \in \mathfrak{A}$, ce qui découle de l'égalité $(A \times B)^c = (A^c \times Y) \cup (A \times B^c)$. ■

Lemme 5.7.2. *Si* $A \in \mathfrak{S} \otimes \mathfrak{T}$*, l'ensemble* $A_y = \{x : (x,y) \in A\}$ *appartient à* $\mathfrak{S}$ *pour tout* y*, et l'ensemble* $A^x = \{y : (x,y) \in A\}$ *appartient à* $\mathfrak{T}$ *pour tout* x*.*

Démonstration : La famille $\mathscr{H}$ des parties $Z \subset X \times Y$ telles que Z_y appartient à $\mathfrak{S}$ pour tout y est de façon claire une tribu. De plus, si $Z = A \times B$, on a $Z_y = A$ ou $Z_y = \emptyset$ selon que $y \in B$ ou $y \notin B$. Il en résulte que $A \times B \in \mathscr{H}$ si $A \in \mathfrak{S}$ et $B \in \mathfrak{T}$. Donc $\mathfrak{S} \otimes \mathfrak{T} \subset \mathscr{H}$, ce qui montre que si $A \in \mathfrak{S} \otimes \mathfrak{T}$, les coupes A_y de A sont toutes dans $\mathfrak{S}$.
L'autre partie du lemme s'obtient en échangeant le rôle des facteurs X et Y. ■

Lemme 5.7.3. *Si* $f : X \times Y \to E$ *est* $\mathfrak{S} \otimes \mathfrak{T}$*-mesurable,* $f_y : x \mapsto f(x,y)$ *est* $\mathfrak{S}$*-mesurable pour tout* $y \in Y$*, et* $f^x : y \mapsto f(x,y)$ *est* $\mathfrak{T}$*-mesurable pour tout* x*.*

Démonstration : Soient U un ouvert de E et $A = f^{-1}(U)$. On a alors $f_y^{-1}(U) = A_y$. Il en résulte que $f_y^{-1}(U)$ appartient à $\mathfrak{S}$ pour tout y. On en conclut que f_y est $\mathfrak{S}$-mesurable pour tout y. Le même raisonnement montre que f^x est $\mathfrak{T}$-mesurable pour tout x. ■

Lemme 5.7.4. *Si $Z \in \mathfrak{S} \otimes \mathfrak{T}$, et si μ est une mesure σ-additive et σ-finie sur $\mathfrak{S}$, la fonction $y \mapsto \mu(Z_y)$ est $\mathfrak{T}$-mesurable.*

Démonstration : Il existe une suite croissante (A_k) dans $\mathfrak{S}$ qui recouvre X et telle que $\mu(A_k) < \infty$ pour tout k. On considère la famille $\mathscr{H}_k$ des parties Z de $X \times Y$ telles que la fonction $y \mapsto \mu(A_k \cap Z_y)$ soit $\mathfrak{T}$-mesurable, et on va montrer qu'elle satisfait les hypothèses du théorème 5.1.3.
Si $(Z^{(n)})$ est une suite croissante dans $\mathscr{H}_k$, de réunion Z, on a $\mu(A_k \cap Z_y) = \sup_n \mu(A_k \cap Z_y^{(n)})$, d'où la mesurabilité de la fonction $y \mapsto \mu(A_k \cap Z_y)$. Et de même, si $(Z^{(n)})$ est une suite décroissante dans $\mathscr{H}_k$, d'intersection Z, on a $\mu(A_k \cap Z_y) = \inf_n \mu(A_k \cap Z_y^{(n)})$ puisque $\mu(A_k) < \infty$, d'où la mesurabilité de la fonction $y \mapsto \mu(A_k \cap Z_y)$.
Enfin, si Z est réunion disjointe de la famille finie $(B_j \times C_j)$, on a, pour $y \in Y$: $\mu(Z_y) = \sum_j \mu(B_j).\mathbb{1}_{C_j}(y)$, ce qui montre que la fonction $y \mapsto \mu(A_k \cap Z_y)$ est élémentaire ; il en résulte que $\mathfrak{A} \subset \mathscr{H}_k$.

On conclut alors du théorème 5.1.3 que $\mathscr{H}_k$ contient la tribu $\mathfrak{S} \otimes \mathfrak{T}$. En particulier, si $Z \in \mathfrak{S} \otimes \mathfrak{T}$, la fonction $\varphi_k : y \mapsto \mu(A_k \cap Z^y)$ est mesurable pour tout k, et, puisque $\mu(Z^y) = \sup_k \mu(A_k \cap Z^y)$, on conclut que $\sup_k \varphi_k = \varphi : y \mapsto \mu(Z^y)$ est $\mathfrak{T}$-mesurable. ∎

Théorème 5.7.5. *Si μ (resp. ν) est une mesure σ-finie sur $(X, \mathfrak{S})$ (resp. $(Y, \mathfrak{T})$), il existe une unique mesure σ-finie $\mu \otimes \nu$ sur $\mathfrak{S} \otimes T$ telle que $(\mu \otimes \nu)(A \times B) = \mu(A).\nu(B)$ pour $A \in \mathfrak{S}$ et $B \in \mathfrak{T}$, et on a, pour tout $Z \in \mathfrak{S} \otimes \mathfrak{T}$:*

$$(\mu \otimes \nu)(Z) = \int \mu(Z_y)\, d\nu(y)\ .$$

Démonstration : On démontre d'abord l'unicité. Supposons que ϖ_1 et ϖ_2 soient des mesures σ-additives sur $\mathfrak{S} \otimes \mathfrak{T}$ telles que $\varpi_1(A \times B) = \varpi(_2 A \times B) = \mu(A).\nu(B)$.
Par l'hypothèse de σ-finitude de μ et de ν, on peut trouver une suite croissante (X_k) dans $\mathfrak{S}$ et une suite croissante (Y_k) dans $\mathfrak{T}$ telles que $X = \bigcup_k X_k$, $Y = \bigcup_k Y_k$, $\mu(X_k) < \infty$ et $\nu(Y_k) < \infty$. On posera alors $Q_k = X_k \times Y_k \in \mathfrak{A}$.
Soit $Z \in \mathfrak{A}$, réunion finie disjointe des $A_j \times B_j$; on a alors

$$\varpi_1(Z) = \sum_j \mu(A_j).\nu(B_j) = \varpi_2(Z)\ ,$$

ce qui montre que ϖ_1 et ϖ_2 coïncident sur $\mathfrak{A}$. On déduit alors du théorème 5.2.7 que $\varpi_1 = \varpi_2$.

Pour $Z \in \mathfrak{S} \otimes \mathfrak{T}$, on pose maintenant $\varpi(Z) = \displaystyle\int \mu(Z_y)\, d\nu(y) \in \overline{\mathbb{R}}_+$, qui est bien définie, puisque, d'après le lemme 5.7.3, la fonction $y \mapsto \mu(Z_y)$ est $\mathfrak{T}$-mesurable à valeurs dans $\overline{\mathbb{R}}_+$. Et on va montrer que ϖ est une mesure σ-additive et σ-finie sur $\mathfrak{S} \otimes \mathfrak{T}$.
Si $Z = A \times B$, on a $\mu(Z_y) = \mu(A).\mathbb{1}_B(y)$, donc $\varpi(Z) = \mu(A).\nu(B)$. On voit de même que $\varpi(Q_k) = \mu(X_k).\nu(Y_k) < \infty$, et que $X \times Y = \bigcup_k Q_k$. Si Z' et Z'' sont disjoints , il en est de même de Z'_y et Z''_y pour tout y. Si $Z = Z' \cup Z''$, on a donc $\mu(Z_y) = \mu(Z'_y) + \mu(Z''_y)$ pour tout y, donc

$$\begin{aligned}\varpi(Z) = \int \mu(Z_y)\, d\nu(y) &= \int \Big(\mu(Z'_y) + \mu(Z''_y)\Big)\, d\nu(y) \\ &= \int \mu(Z'_y)\, d\nu(y) + \int \mu(Z''_y)\, d\nu(y) = \varpi(Z') + \varpi(Z'')\ .\end{aligned}$$

Enfin, si $(Z^{(n)})$ est une suite croissante d'éléments de $\mathfrak{S} \otimes \mathfrak{T}$, de réunion Z, on a pour tout $y \in Y$: $Z_y = \bigcup_n Z_y^{(n)}$, donc $\mu(Z_y) = \sup_n \mu(Z_y^{(n)})$, et par convergence monotone,

$$\varpi(Z) = \int \mu(Z_y)\, d\nu(y) = \sup_n \int \mu(Z_y^{(n)})\, d\nu(y) = \sup_n \varpi(Z^{(n)}) \ .$$

Il en résulte que ϖ est une mesure σ-additive et σ-finie sur $\mathfrak{S} \otimes \mathfrak{T}$, qu'on notera $\mu \otimes \nu$. ■

Théorème 5.7.6. *Si N est un élément de la tribu $\mathfrak{S}\otimes\mathfrak{T}$, l'ensemble N est $\mu\otimes\nu$-négligeable si et seulement ses coupes N_y sont ν-presque toutes négligeables pour μ.*

Démonstration : Ceci découle immédiatement de la formule de l'énoncé précédent : $(\mu \otimes \nu)(N) = \int \mu(N_y)\, d\nu(y)$. ■

Lemme 5.7.7. *Si $f : X \times Y \to \overline{\mathbb{R}}_+$ est une fonction élémentaire $\mathfrak{S} \otimes \mathfrak{T}$-mesurable, la fonction $g : y \mapsto \int f(x,y)\, d\mu(x)$ est mesurable positive, et l'intégrale de f par rapport à $\mu \otimes \nu$ est égale à $\displaystyle\int g(y)\, d\nu(y)$.*

Démonstration : Il existe des $\lambda_j \geqslant 0$ et des $Z^{(j)} \in \mathfrak{S} \otimes \mathfrak{T}$ en nombre fini tels que $f = \sum_j \lambda_j \mathbb{1}_{Z^{(j)}}$. Alors on a

$$\int f(x,y)\, d(\mu\otimes\nu)(x,y) = \sum_j \lambda_j.\mu\otimes\nu(Z_j) = \sum_j \lambda_j \int \mu(Z_y^{(j)})\, d\nu(y) = \int \sum_j \lambda_j \mu(Z_y^{(j)})\, d\nu(y)$$

et puisque, pour tout y on a $f_y(x) := f(x,y) = \sum_j \lambda_j \mathbb{1}_{Z_y^{(j)}}(x)$, on a

$$g(y) = \int f_y(x)\, d\mu(x) = \sum \lambda_j \mu(Z_y^{(j)}) \ ,$$

donc g, combinaison linéaire de fonctions mesurables, est mesurable, et

$$\int f(x,y)\, d(\mu \otimes \nu)(x,y) = \int g(y)\, d\nu(y) \ .$$ ■

Théorème 5.7.8. (Fubini) *Si $f : X \times Y \to \overline{\mathbb{R}}_+$ est $\mathfrak{S} \otimes \mathfrak{T}$-mesurable, les fonctions J_1 définie sur Y par $J_1(y) = \int f(x,y)\, d\mu(x)$ et J_2 définie sur X par $J_2(x) = \int f(x,y)\, d\nu(y)$ sont respectivement $\mathfrak{T}$-mesurable et $\mathfrak{S}$-mesurable à valeurs dans $\overline{\mathbb{R}^+}$, et on a*

$$\int f(x,y)\, d(\mu \otimes \nu)(x,y) = \int J_1(y)\, d\nu(y) = \int J_2(x)\, d\mu(x) \ .$$

Démonstration : Si f est $\mathfrak{S}\otimes\mathfrak{T}$-mesurable à valeurs dans $\overline{\mathbb{R}}_+$, il existe une suite croissante (f_n) de fonctions élémentaires positives qui converge vers f. Et on a, d'après le théorème précédent, en notant $g_n(y) = \int f_n(x,y)\, d\mu(x)$,

$$\int f(x,y)\, d(\mu \otimes \nu)(x,y) = \sup_n \int f_n(x,y)\, d(\mu \otimes \nu)(x,y) = \sup_n \int g_n(y)\, d\nu(y) \ .$$

Pour chaque $y \in Y$, le théorème de convergence monotone (pour (X, μ)) montre que

$$J_1(y) = \int f(x,y)\, d\mu(x) = \sup_n \int f_n(x,y)\, d\mu(x) = \sup_n g_n(y)$$

et, en particulier, montre que J_1 est mesurable. Puis, appliqué à (Y, ν), le théorème de convergence monotone montre que

$$\int J_1(y)\, d\nu(y) = \sup_n \int g_n(y)\, d\nu(y) = \int f(x,y)\, d(\mu \otimes \nu)(x,y)\ .$$

L'unicité de la mesure σ-additive ϖ telle que $\varpi(A \times B) = \mu(A).\nu(B)$ montre que si on échange les rôles de X et de Y dans la construction de $\mu \otimes \nu$, en définissant, pour $Z \in \mathfrak{S} \otimes \mathfrak{T}$, $\varpi'(Z) = \int \nu(Z^x)\, d\mu(x)$, on obtient encore $\varpi' = \mu \otimes \nu$. Il en résulte que J_2 est μ-mesurable et que $\int J_2(x)\, d\mu(x) = \int f(x,y)\, d(\mu \otimes \nu)(x,y)$. ■

On notera désormais cette intégrale par $\displaystyle\int f(x,y)\, d\mu(x)\, d\nu(y)$. Pour mieux mettre en évidence les deux intégrations successives par rapport à μ et par rapport à ν, on la note aussi parfois $\displaystyle\iint f(x,y)\, d\mu(x)\, d\nu(y)$.

Théorème 5.7.9. (Fubini-Tonelli) *Si* $f : X \times Y \to \mathbb{C}$ *est intégrable sur* $(X \times Y, \mathfrak{S} \otimes \mathfrak{T}, \mu \otimes \nu)$, *la fonction* $f_y : x \mapsto f(x,y)$ *est* μ*-intégrable pour* ν*-presque tout* y *et on a*

$$\int f(x,y)\, d(\mu \otimes \nu)(x,y) = \int \left(\int f_y(x)\, d\mu(x) \right) d\nu(y)\ .$$

Et de même la fonction $f^x : y \mapsto f(x,y)$ *est* ν*-intégrable pour* μ*-presque tout* x *et on a*

$$\int f(x,y)\, d(\mu \otimes \nu)(x,y) = \int \left(\int f^x(y)\, d\nu(y) \right) d\mu(x)\ .$$

Démonstration : Supposons la fonction f intégrable. Alors, par définition, la fonction $|f|$ est intégrable positive, et il résulte du théorème de Fubini ci-dessus que l'on a

$$\int \left(\int |f_y(x)|\ d\mu(x) \right) d\nu(y) = \int |f(x,y)|\ d\mu(x)\, d\nu(y) < \infty\ .$$

La fonction mesurable positive $J_1 : y \mapsto \displaystyle\int |f_y(x)|\ d\mu(x)$ est donc intégrable. Il en résulte que l'ensemble $N = \{J_1 = \infty\}$ est ν-négligeable (cf. Lemme 5.4.9). Pour tout $y \notin N$, la fonction mesurable f_y est donc intégrable. De plus, puisque f est intégrable, il en est de même de $\Re e(f)$ et de $\Im m(f)$: chacune de ces deux fonctions réelles intégrables est donc différence de deux fonctions positives intégrables. Il existe donc g_1, g_2, h_1 et h_2 intégrables positives telles que $f = g_1 - g_2 + i(h_1 - h_2)$. On a $g_1 + g_2 + h_1 + h_2 \leqslant 2\,|f|$; il en résulte que, en tout point $y \notin N$, chacune des fonctions $(g_1)_y$, $(g_2)_y$, $(h_1)_y$ et $(h_2)_y$ est μ-intégrable. On a alors

$$\int f_y(x)\, d\mu(x) = \int (g_1)_y(x)\, d\mu(x) - \int (g_2)_y(x)\, d\mu(x) + i \int (h_1)_y(x)\, d\mu(x) - i \int (h_2)_y(x)\, d\mu(x)$$

et puisque

$$\int f(x,y)\,d\mu(x)\,d\nu(y) = \int g_1(x,y)\,d\mu(x)\,d\nu(y) - \int g_2(x,y)\,d\mu(x)\,d\nu(y)$$
$$+i\int h_1(x,y)\,d\mu(x)\,d\nu(y) - i\int h_2(x,y)\,d\mu(x)\,d\nu(y)\ ,$$

en appliquant le théorème de Fubini à nouveau à chacune des quatre fonctions, on obtient :

$$\begin{aligned}\int f(x,y)\,d\mu(x)\,d\nu(y) &= \int\left(\int\Big((g_1)_y - (g_2)_y + i(h_1)_y - i(h_2)_y\Big)\,d\mu(x)\right)\,d\nu(y)\\ &= \int\left(\int f_y(x)\,d\mu(x)\right)\,d\nu(y)\ .\end{aligned}$$

Le reste s'obtient en intervertissant les rôles de (X,μ) et de (Y,ν). ■

5.8 Construction d'une mesure borélienne

Proposition 5.8.1. *Soient X un espace métrique séparable et μ une mesure σ-additive et localement finie (c'est-à-dire que tout point possède un voisinage de mesure finie) sur la tribu borélienne $\mathfrak{B}(X)$. Alors la restriction m de μ à la famille des compacts de X possède les propriétés suivantes :*

i) Pour tout K compact et tout $\varepsilon > 0$, il existe un voisinage ouvert U de K tel que $m(L) < m(K) + \varepsilon$ si $L \subset U$.

ii) $m(K \cup L) \leqslant m(K) + m(L)$.

iii) Si $K \cap L = \emptyset$, alors $m(K \cup L) = m(K) + m(L)$.

Démonstration : Les propriétés *ii)* et *iii)* sont vérifiées pour toute mesure μ, ainsi que pour toutes parties K et L de $\mathfrak{B}(X)$.
Si K est compact, il possède un recouvrement fini par des ouverts $(V_j)_{j\in J}$ de mesure finie. Alors l'ouvert $V = \bigcup_{j\in J} V_k$ est un voisinage de K, et la fonction continue $x \mapsto d(x, V^c)$ atteint sur K un minimum $\eta > 0$.
Alors, la suite décroissante d'ouverts $U_k = \{x : d(x,K) < \eta.2^{-k}\}$ est formée d'ensembles contenus dans V, donc de mesure finie, et puisque K est l'intersection des U_k, on a

$$m(K) = \mu(K) = \inf_k \mu(U_k)\ .$$

Pour $\varepsilon > 0$, on peut donc trouver un k tel que $\mu(U_k) < m(K) + \varepsilon$. Et si L est un compact contenu dans U_k, on a $m(L) = \mu(L) \leqslant \mu(U_k) < m(K) + \varepsilon$. ■

Théorème 5.8.2. *Soient X un espace métrique séparable et localement compact, et m une fonction positive sur les compacts de X satisfaisant :*

i) Pour tout K compact et tout $\varepsilon > 0$, il existe un voisinage ouvert U de K tel que $m(L) < m(K) + \varepsilon$ si $L \subset U$.

ii) $m(K \cup L) \leqslant m(K) + m(L)$.

iii) *Si $K \cap L = \emptyset$, alors $m(K \cup L) \geqslant m(K) + m(L)$.*

Il existe alors une unique mesure σ-additive μ sur la tribu borélienne $\mathfrak{B}(X)$ telle que $\mu(K) = m(K)$ pour tout compact K.

Démonstration : On considère la famille $\mathfrak{S}$ des parties de X telles que, pour tout compact T et tout $\varepsilon > 0$, il existe deux compacts, $L \subset A \cap T$ et $L' \subset T \setminus A$, satisfaisant $m(L) + m(L') > m(T) - \varepsilon$. On va, au moyen des lemmes qui suivent, montrer que $\mathfrak{S}$ est une tribu sur X contenant la tribu borélienne, et construire une mesure μ, σ-additive et σ-finie sur $\mathfrak{S}$, coïncidant avec m sur les compacts.

Lemme 5.8.3. *Tout compact de X appartient à $\mathfrak{S}$.*

Démonstration : Soient K et T compacts de X, et $\varepsilon > 0$. Il existe, par la condition *i)*, un ouvert U contenant $K \cap T$ tel que, pour tout compact $K' \subset U$, on ait $m(K') < m(K \cap T) + \varepsilon$. Il existe alors un voisinage ouvert V de $K \cap T$ tel que $\overline{V} \subset U$. On pose alors $L = K \cap T$ et $L' = T \setminus V$; L est un compact contenu dans $K \cap T$ (et même égal), et L' est un compact contenu dans $T \setminus K$. De plus, on a $T = (T \cap \overline{V}) \cup (T \setminus V)$, donc $m(T \cap \overline{V}) + m(L') \geqslant m(T)$ d'après *ii)*. Et puisque $T \cap \overline{V} \subset U$, on a $m(T \cap \overline{V}) < m(K \cap T) - \varepsilon$. On en déduit que $m(L') > m(T) - m(K \cap T) - \varepsilon$, ou encore que $m(L) + m(L') > m(T) - \varepsilon$. ■

Lemme 5.8.4. *Si $A \in \mathfrak{S}$, alors $A^c \in \mathfrak{S}$.*

Démonstration : Soient T compact et $\varepsilon > 0$. Puisque $A \in \mathfrak{S}$, il existe L et L' compacts contenus respectivement dans $A \cap T$ et $A^c \cap T$ tels que $m(L) + m(L') > m(T) - \varepsilon$. Il suffit d'échanger L et L' pour voir que $A^c \in \mathfrak{S}$. ■

Lemme 5.8.5. *Si A et B appartiennent à $\mathfrak{S}$, alors $A \cup B$ appartient à $\mathfrak{S}$.*

Démonstration : Soient T compact et $\varepsilon > 0$. Il existe L et L' compacts contenus respectivement dans $A \cap T$ et $A^c \cap T$ tels que $m(L) + m(L') > m(T) - \dfrac{\varepsilon}{2}$. Il existe alors L_1 et L'_1 compacts contenus respectivement dans $L' \cap B$ et $L' \cap B^c$ tels que $m(L_1) + m(L'_1) > m(L) - \dfrac{\varepsilon}{2}$. Si on pose $K = L \cup L_1$ et $K' = L'_1$, on a $K \subset T \cap (A \cup B)$ et $K' \subset T \setminus (A \cup B)$. De plus, puisque L et L_1 sont disjoints, on a

$$m(K) + m(K') \geqslant m(L) + m(L_1) + m(L'_1) > m(L) + m(L') - \frac{\varepsilon}{2} > m(T) - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = m(T) - \varepsilon\ ,$$

et ceci montre que $A \cup B \in \mathfrak{S}$. ■

Lemme 5.8.6. *Si (A_n) est une suite croissante d'éléments de $\mathfrak{S}$, la réunion A des (A_n) est dans $\mathfrak{S}$.*

Démonstration : Soient T compact et $\varepsilon > 0$. On construit par récurrence deux suites $(L_n)_{n \geqslant 0}$ et $(L'_n)_{n \geqslant -1}$ de compacts de T de sorte que $L'_{-1} = T$, que $L_n \subset L'_{n-1} \cap A_n$ et $L'_n \subset L'_{n-1} \setminus A_n$ et $m(L_n) + m(L'_n) > m(L'_{n-1}) - \varepsilon.2^{-n-2}$, ce qui est possible puisque $A_n \in \mathfrak{S}$. Il est clair que $L_n \subset A_n \subset A$ pour tout n.
On vérifie alors par récurrence sur n que

$$m(L'_n) + \sum_{k=0}^{n} m(L_k) > m(T) - \sum_{k=0}^{n} \varepsilon.2^{-2-k} > m(T) - \frac{\varepsilon}{2}\ .$$

Si on pose $L' = \bigcap_n L'_n$, on a $L' \cap A_n = \emptyset$ pour tout n, donc $L' \subset T \setminus A$. Et il existe un ouvert U contenant L' tel que $m(K) < m(L') + \dfrac{\varepsilon}{2}$ pour tout $K \subset U$. La suite décroissante de compacts $(L'_n \setminus U)$ a une intersection vide ; il existe donc un p tel que $L'_p \setminus U = \emptyset$, et $m(L') > m(L'_p) - \dfrac{\varepsilon}{2}$. On voit alors que, puisque les L_k sont deux-à-deux disjoints,

$$m(L') + m(\bigcup_{k\leqslant p} L_k) \geqslant m(L') + \sum_{k\leqslant p} m(L_k) > m(L'_p) + \sum_{k\leqslant p} m(L_k) - \frac{\varepsilon}{2} > m(T) - \varepsilon \ .$$

Il suffit alors de poser $L = \bigcup_{k\leqslant p} L_k \subset T \cap A$ pour obtenir les deux compacts L et L' contenus respectivement dans $T \cap A$ et $T \setminus A$. ■

Lemme 5.8.7. *$\mathfrak{S}$ est une tribu contenant la tribu borélienne de X.*

Démonstration : Il résulte des lemmes précédents que $\mathfrak{S}$ est stable par complémentation et par réunions dénombrables, donc que $\mathfrak{S}$ est une tribu. De plus $\mathfrak{S}$ contient les compacts. Alors, si U est un ouvert de X, chaque point x de U possède un voisinage ouvert V_x relativement compact dans U. Et puisque X est séparable, le recouvrement $(V_x)_{x\in U}$ possède un sous-recouvrement dénombrable $(V_x)_{x\in D}$; on a alors $U = \bigcup_{x\in D} \overline{V_x}$, ce qui montre que U, réunion dénombrable de compacts, appartient à $\mathfrak{S}$. Et la tribu $\mathfrak{S}$ qui contient les ouverts, contient toute la tribu borélienne $\mathfrak{B}(X)$. ■

On pose alors, pour tout $A \in \mathfrak{S}$,

$$\mu(A) = \sup\{m(K) : K \text{ compact }, K \subset A\} \ .$$

Il est clair que $\mu(K) = m(K)$ si K est compact, et il résulte de ce qui précède que X, qui est ouvert, est réunion d'une suite (K_n) de compacts : ceci montre que μ est σ-finie, puisque $\mu(K_n) = m(K_n) < \infty$.
On va montrer que μ est une mesure σ-additive sur $\mathfrak{S}$.

Lemme 5.8.8. *Si A et B appartiennent à $\mathfrak{S}$ et sont disjoints, on a $\mu(A\cup B) = \mu(A)+\mu(B)$.*

Démonstration : Si K est un compact contenu dans A, avec $m(K) > \mu(A) - \dfrac{\varepsilon}{2}$ et L un compact contenu dans B, avec $m(L) > \mu(B) - \dfrac{\varepsilon}{2}$, alors $K \cap L = \emptyset$ et $K \cup L$ est un compact contenu dans $A \cup B$: on a donc $\mu(A\cup B) \geqslant m(K\cup L) \geqslant m(K) + m(L) > \mu(A) + \mu(B) - \varepsilon$.
Inversement, si K est un compact contenu dans $A\cup B$, avec $m(K) > \mu(A\cup B) - \dfrac{\varepsilon}{2}$, il existe, puisque $A \in \mathfrak{S}$, des compacts $L \subset K \cap A$ et $L' \subset K \setminus A$ tels que

$$m(L) + m(L') > m(K) - \frac{\varepsilon}{2} \ .$$

Puisque $L' \subset K \setminus A \subset (A \cup B) \setminus A = B$, on voit alors que

$$\mu(A) + \mu(B) \geqslant m(L) + m(L') > m(K) - \frac{\varepsilon}{2} > \mu(A\cup B) - \varepsilon \ .$$

Et, puisque ε est arbitraire, il résulte de ces inégalités que $\mu(A\cup B) = \mu(A) + \mu(B)$. ■

Lemme 5.8.9. *Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante dans $\mathfrak{S}$, et si $A = \bigcup_{n\in\mathbb{N}} A_n$, on a $\mu(A) = \sup_n \mu(A_n)$.*

Démonstration : On a nécessairement $\mu(A) \geqslant \mu(A_n)$ pour tout entier n, et donc $\mu(A) \geqslant \sup_n \mu(A_n)$. En particulier, si $\sup_n \mu(A_n) = +\infty$, on a l'égalité cherchée.
Supposons donc que $M = \sup_n \mu(A_n) < \infty$ et que $\mu(A) > M$. On pourrait alors trouver $\eta > 0$ et K compact contenu dans A tels que $m(K) > M + \eta$. Puisque chaque A_n est dans $\mathfrak{S}$, on pourrait construire par récurrence une suite décroissante de compacts (L'_n) et une suite de compacts (L_n) telles que $L'_0 = K$, $L'_{n+1} \subset L'_n \setminus A_n$, $L_{n+1} \subset L'_n \cap A_n$, $m(L_{n+1}) + m(L'_{n+1}) > m(L'_n) - \eta.2^{-n-1}$. On voit alors par récurrence, comme plus haut, que

$$m(L'_{n+1})+m(\bigcup_1^{n+1} L_k) \geqslant m(L'_{n+1})+\sum_{k=1}^{n+1} m(L_k) > m(L'_0)-\sum_1^m \eta.2^{-k-1} > m(K)-\frac{\eta}{2} > M+\frac{\eta}{2}$$

et puisque $\bigcup_1^{n+1} L_k$ est un compact contenu dans A_n, on a $m(\bigcup_1^{n+1} L_k) \leqslant M$, donc $m(L'_{n+1}) \geqslant \dfrac{\eta}{2}$. En particulier, on a $L'_{n+1} \neq \emptyset$, et l'intersection L' de la suite décroissante (L'_n) de compacts non vides est un compact non vide, contenu dans $K \setminus \bigcup_n A_n = \emptyset$. Cette contradiction achève la démonstration du lemme, et termine de montrer que μ est σ-additive. ■

Lemme 5.8.10. *Il existe au plus une mesure σ-additive μ sur $\mathfrak{B}(X)$ telle que $\mu(K) = m(K)$ pour tout compact.*

Démonstration : Si on note $\mathfrak{A}$ l'ensemble des parties Z de X qui sont, ainsi que leurs complémentaires, réunions dénombrables de compacts, on vérifie sans peine que $\mathfrak{A}$ est une algèbre de parties de X qui contient les ouverts. Il en résulte que la tribu engendrée par $\mathfrak{A}$ est la tribu borélienne $\mathfrak{B}(X)$.
Si μ et ν sont des mesures σ-additives sur $\mathfrak{B}(X)$ telles que $\mu(K) = \nu(K) = m(K)$ pour tout compact, on a $\mu(Z) = \nu(Z)$ pour tout $Z \in \mathfrak{A}$: en effet, si $Z \in \mathfrak{A}$, il existe une suite croissante (K_n) de compacts telle que $Z = \bigcup_n K_n$; et on a alors $\mu(Z) = \sup_n m(K_n) = \nu(Z)$.
Enfin, X est réunion d'une suite (L_n) de compacts ; et on a alors $L_n \in \mathfrak{A}$ ainsi que $\mu(L_n) = m(L_n) < \infty$. On déduit alors du théorème 5.2.7 que $\mu = \nu$. ■

Et ceci achève la démonstration du théorème 5.8.2. ■

Corollaire 5.8.11. *Si X est un espace métrique séparable localement compact, μ et ν deux mesures σ-additives sur $\mathfrak{B}(X)$ localement finies qui coïncident sur les compacts de X, alors μ et ν sont égales.*

Démonstration : Il résulte de la proposition 5.8.1 que la restriction m de μ aux compacts de X vérifie les conditions du théorème 5.8.2. Alors, il résulte du lemme 5.8.10 que ν, qui est égale à m sur les compacts, coïncide avec μ sur $\mathfrak{B}(X)$. ■

5.9 La mesure de Lebesgue sur l'espace euclidien

Théorème 5.9.1. *Il existe sur $\mathbb{R}^d$, muni de sa tribu borélienne, une unique mesure σ-additive μ, appelée* mesure de Lebesgue, *telle que $\mu(P) = \prod_{j=1}^d (b_\ell - a_\ell)$ pour tout pavé compact $P = \prod_\ell [a_\ell, b_\ell]$. Cette mesure est invariante par translation : $\mu(A+t) = \mu(A)$ si A est borélien, $t \in \mathbb{R}^d$ et $A + t = \{a + t : a \in A\}$.*

On va construire cette mesure en utilisant le résultat de la section 5.8. On définit pour cela une fonction m sur l'ensemble des compacts de $\mathbb{R}^d$ par

$$m(K) = \inf\Big\{\sum_{j=1}^p v(P_j) : K \text{ est recouvert par les cubes ouverts } (P_j)_{1 \leqslant j \leqslant p}\Big\} ,$$

où on note $v(P) = h^d$, le "volume" du cube ouvert $P = \prod_1^d]a_\ell, a_\ell + h[$. Plus généralement, on notera $v(P) = \prod_1^d (b_\ell - a_\ell)$ le volume du pavé $P = \prod_1^d]a_\ell, b_\ell[$.
Pour simplifier les démonstrations, on introduit aussi, pour $\eta > 0$, une variante m_η de cette fonction, en restreignant les cubes recouvrant K à être de diamètre au plus η :

$$m_\eta(K) = \inf\{\sum_{j=1}^p v(Q_j) : K \subset \bigcup_{j=1}^p v(Q_j) \ , \quad \operatorname{diam}(Q_j) < \eta\} \ .$$

Lemme 5.9.2. *Soient E un espace métrique, K un compact de E, et $(U_j)_{j \leqslant p}$ un recouvrement ouvert de K. Alors il existe des compacts $(K_j)_{j \leqslant p}$ tels que $K_j \subset U_j$ et $K = \bigcup_{j \leqslant p} K_j$.*

Démonstration : Pour chaque $j \leqslant p$, la fonction $h_j : x \mapsto d(x, U_j^c)$ est continue et strictement positive sur U_j. Il en résulte que $h = \max_{j \leqslant p} h_j$ est continue et strictement positive sur K. Elle y admet donc un minimum strictement positif η. Alors si on pose $K_j = \{x \in K : h_j(x) \geqslant \eta\}$, on voit que $K = \bigcup_{j \leqslant p} K_j$ et que $K_j \subset U_j$. ■

Lemme 5.9.3. *Soient K un compact de $\mathbb{R}^d$ contenu dans un cube ouvert Q de côté h, et $p \geqslant 1$. Alors il existe p^d cubes ouverts de côté $\dfrac{h}{p}$ qui recouvrent K.*

Démonstration : Si K est une partie compacte du cube ouvert $Q = \prod_{\ell=1}^d]a_\ell, a_\ell + h[$ de côté h il existe $\alpha > 0$ tel que $K \subset \prod_{\ell=1}^d]a_\ell, a_\ell + h - \alpha[$.
Pour $w = (j_1, j_2, \ldots, j_p) \in \{1, 2, \ldots, p\}^d$, on définit le cube de côté $\dfrac{h}{p}$:

$$Q_w = \prod_{\ell=1}^d \Big]a_\ell - \frac{j_\ell}{p}\alpha + \frac{j_\ell - 1}{p} h \, , \, a_\ell - \frac{j_\ell}{p}\alpha + \frac{j_\ell}{p} h\Big[$$

et on vérifie immédiatement que $\bigcup_w Q_j = \prod_1^d]a_\ell, a_\ell + h - \alpha[\supset K$. ■

Lemme 5.9.4. *Pour tout $\eta > 0$ et tout compact $K \subset \mathbb{R}^d$, on a $m_\eta(K) = m(K)$.*

Démonstration : Il est clair que $m_\eta(K) \geqslant m(K)$. Inversement, si K est recouvert par un nombre fini de cubes ouverts $(Q_j)_{j \in J}$, il existe, par le lemme 5.9.2, des compacts $(K_j)_{j \in J}$

de réunion K, contenus chacun dans un Q_j de côté h_j. En appliquant le lemme 5.9.3 avec p assez grand pour que $\dfrac{h_j}{p} < \eta' = \dfrac{\eta}{\sqrt{d}}$, on peut remplacer Q_j par une famille de cubes $Q_{j,k}$ de côté inférieur à η', recouvrant K_j, et dont la somme $\sum_k v(Q_{j,k})$ est égale à $v(Q_j)$. On voit alors que la famille $(Q_{j,k})_{j,k}$ recouvre K, est formée de cubes de diamètre inférieur à η et que la somme des volumes de ces cubes est égale à la somme des volumes des (Q_j). Donc $m_\eta(K) \leqslant m(K)$. ■

Lemme 5.9.5. *Si K est un compact de $\mathbb{R}^d$ et $\varepsilon > 0$, il existe un voisinage ouvert U de K tel que $m(L) < m(K) + \varepsilon$ si L est un compact contenu dans U.*

Démonstration : Il existe par définition une famille finie de cubes ouverts $(Q_j)_{j\in J}$ telle que $\sum_j v(Q_j) < m(K) + \varepsilon$ et qui recouvre K. Si on pose $U = \bigcup_j Q_j$, il est clair que tout compact L contenu dans U est recouvert par les (Q_j), donc vérifie $m(L) < m(K) + \varepsilon$. ■

Lemme 5.9.6. *Si K et L sont deux compacts de $\mathbb{R}^d$, on a $m(K \cup L) \leqslant m(K) + m(L)$.*

Démonstration : Soit $\varepsilon > 0$. Il existe par définition des familles finies $(Q_j)_{j\in J}$ et $(Q_j)_{j\in J'}$ de cubes ouverts recouvrant respectivement K et L, telles que $\sum_{j\in J} v(Q_j) < m(K) + \dfrac{\varepsilon}{2}$ et $\sum_{j\in J'} v(Q_j) < m(L) + \dfrac{\varepsilon}{2}$. Alors la famille $(Q_j)_{j\in J\cup J'}$ recouvre $K \cup L$ et on a

$$\sum_{j\in J\cup J'} v(Q_j) \leqslant \sum_{j\in J} v(Q_j) + \sum_{j\in J'} v(Q_j) < m(K) + m(L) + \varepsilon\ ,$$

ce qui montre que $m(K \cup L) \leqslant m(K) + m(L)$. ■

Lemme 5.9.7. *Soient K et L deux compacts dans $\mathbb{R}^d$. Si K et L sont disjoints, on a $m(K \cup L) \geqslant m(K) + m(L)$.*

Démonstration : Puisque $K \cap L = \emptyset$, la fonction continue $(x, y) \mapsto \|x - y\|$ atteint sur le compact $K \times L$ un minimum strictement positif η. Si $\varepsilon > 0$ est donné, il existe, puisque $m(K \cup L) = m_\eta(K \cup L)$, un recouvrement fini $(Q_j)_{j\in J}$ de $K \cup L$ par des cubes ouverts de diamètre $< \eta$ tel que $\sum_{j\in J} v(Q_j) < m(K \cup L) + \varepsilon$. On remarque alors qu'aucun des Q_j ne peut rencontrer simultanément K et L, donc que $J' = \{j : Q_j \cap K \neq \emptyset\}$ et $J'' = \{j : Q_j \cap L \neq \emptyset\}$ sont disjoints. On a donc $K \subset \bigcup_{j\in J'} Q_j$ et $L \subset \bigcup_{j\in J''} Q_j$, donc

$$m(K) + m(L) \leqslant \sum_{j\in J'} v(Q_j) + \sum_{j\in J''} v(Q_j) = \sum_{j\in J'\cup J''} v(Q_j) \leqslant \sum_{j\in J} v(Q_j) < m(K \cup L) + \varepsilon\ .$$

Et ceci prouve l'inégalité cherchée. ■

Lemme 5.9.8. *Si P est le pavé compact $\prod_1^d [a_\ell, b_\ell]$, on a $m(P) = v(P) = \prod_1^d (b_\ell - a_\ell)$.*

Démonstration :
Si P est recouvert par une famille finie de cubes ouverts $(Q_j)_{j\in J}$, si on note M le nombre d'éléments de J et h_j le côté de Q_j, et si $\eta > 0$, on remarque que pour $0 \leqslant p < \dfrac{b_\ell - a_\ell}{\eta}$, le point $a_\ell + p\eta$ appartient à $[a_\ell, b_\ell]$. Donc si w appartient à l'ensemble W_η des $(p_1, p_2, \ldots, p_d)$ de $\mathbb{N}^d$ avec $p_\ell < \dfrac{b_\ell - a_\ell}{\eta}$ pour $1 \leqslant \ell \leqslant d$, le point $x_w = (a_1 + p_1\eta, a_2 + p_2\eta, \ldots, a_d + p_d\eta)$ appartient à P, donc à l'un des U_j. On pose alors $F_j = \{w : x_w \in U_j\}$, et on a

$\bigcup_{j\in J} F_j = W_\eta$. Il en résulte que, si on note N_j le nombre d'éléments de F_j, on a $\sum_{j\in J} N_j \geqslant \prod_1^d \dfrac{b_\ell - a_\ell}{\eta} = \eta^{-d}.v(P)$.
De plus, si un intervalle ouvert contient N points de la forme $a + p\eta$, il est de longueur au moins $(N-1)\eta$: il existe donc, pour tout ℓ, au plus $1 + \dfrac{h_j}{\eta}$ valeurs possibles de p_ℓ pour que le point x_w appartienne à Q_j : c'est-à-dire que $N_j \leqslant (1 + \dfrac{h_j}{\eta})^d$.

On en déduit que $\eta^{-d}.v(P) \leqslant \sum_j (1 + \dfrac{h_j}{\eta})^d$, ou encore $v(P) \leqslant \sum_j (h_j + \eta)^d$. Alors, si on choisit $\eta < \varepsilon.\inf_j h_j$, on obtient $v(P) \leqslant (1+\varepsilon)^d \sum h_j^d = (1+\varepsilon)^d \sum v(Q_j)$, donc $m(P) \geqslant (1+\varepsilon)^{-d} v(P)$.
Et puisque ε est arbitraire, on conclut que $m(P) \geqslant v(P)$.

Il est clair que, pour tout $\varepsilon > 0$, un cube compact Q de côté h est contenu dans un cube ouvert de côté $h + \varepsilon$, donc que $m(Q) \leqslant (h+\varepsilon)^d$, c'est-à-dire que $m(Q) \leqslant h^d = v(Q)$.
Et si $h > 0$, si, pour $1 \leqslant \ell \leqslant d$, on a $(p_\ell - 1)h < b_\ell - a_\ell \leqslant p_\ell h$, l'intervalle $[a_\ell, b_\ell]$ est recouvert par p_ℓ intervalles fermés de longueur h. Et par produits, on trouve $\prod_1^d p_\ell$ cubes compacts de volume h^d recouvrant P. En vertu du lemme 5.9.6, on voit alors que $m(P) \leqslant h^d \prod_1^d p_\ell \leqslant \prod_1^d (b_\ell - a_\ell + h)$. Et puisque $h > 0$ est arbitraire, on conclut que $m(P) \leqslant v(P)$. ■

Les lemmes qui précèdent montrent que les conditions du théorème 5.8.2 sont vérifiées par cette fonction m. On en déduit l'existence et l'unicité de la mesure σ-additive μ sur $\mathfrak{B}(\mathbb{R}^d)$ telle que $\mu(K) = m(K)$ pour tout compact K.
Pour achever la démonstration du théorème 5.9.1, il reste uniquement à prouver l'invariance de μ par translation. Soit $t \in \mathbb{R}^d$. La translation $\tau : x \mapsto x - t$ est un homéomorphisme de $\mathbb{R}^d$: il en résulte que l'ensemble $\{A \subset \mathbb{R} : \tau^{-1}(A) \in \mathfrak{B}(\mathbb{R}^d)\}$ est une tribu contenant les ouverts, donc contenant $\mathfrak{B}(\mathbb{R}^d)$. Alors la mesure σ-additive μ_t définie sur $\mathfrak{B}(\mathbb{R}^d)$ par $\mu_t(A) = \mu(\tau^{-1}(A))$ vérifie pour K compact : $\mu_t(K) = m(\tau^{-1}(K))$. Mais puisque τ conserve la longueur des intervalles, donc le volume des cubes, on voit clairement que $m(\tau^{-1}(K)) = m(K)$. donc que $\mu_t(K) = \mu(K)$. Et le théorème d'unicité 5.8.11 montre alors que $\mu_t = \mu$, c'est-à-dire que $\mu(A + t) = \mu(A)$ pour tout $A \in \mathfrak{B}(\mathbb{R}^d)$. ■

Lemme 5.9.9. *Soient U un ouvert de $\mathbb{R}^d$ et μ et ν deux mesures σ-additives localement finies sur $\mathfrak{B}(U)$ qui coïncident sur les pavés compacts contenus dans U. Alors $\mu = \nu$.*

DÉMONSTRATION : On montre d'abord par récurrence sur l'entier p que μ et ν coïncident sur les réunions d'au plus p pavés compacts contenus dans U. Ceci est clair pour $p = 1$.
Si c'est vrai pour p et si K est réunion de $p+1$ pavés compacts $(P_j)_{1\leqslant j\leqslant p+1}$, on pose $L = \bigcup_{1\leqslant j\leqslant p} P_j$, et on a $K = L \cup P_{p+1}$ donc $\mu(K) + \mu(L \cap P_{p+1}) = \mu(L) + \mu(P_{p+1})$ et de même $\nu(K) + \nu(L \cap P_{p+1}) = \nu(L) + \nu(P_{p+1})$. Comme $L \cap P_{p+1} = \bigcup_{1\leqslant j\leqslant p} P_j \cap P_{p+1}$ est, comme L, réunion d'au plus p pavés compacts, on a

$$\mu(K) = \mu(L) + \mu(P_{p+1}) - \mu(L \cap P_{p+1}) = \nu(L) + \nu(P_{p+1}) - \nu(L \cap P_{p+1}) = \nu(K) ,$$

ce qui montre que la propriété est vraie pour $p + 1$.

Soit maintenant K un compact contenu dans U. Pour k entier assez grand, l'ouvert $U_k = \{x : d(x, K) < 2^{-k}\}$ est relativement compact dans U, donc de μ-mesure et de ν-mesure finies. Et puisque la suite (U_k) est décroissante, on a $\mu(K) = \inf_k \mu(U_k)$ et $\nu(K) = \inf_k \nu(U_k)$.

Alors, si $2^{-n}\sqrt{d} < 2^{-k}$, tout cube compact de côté 2^{-n} qui rencontre K est contenu dans U_k. Et si on note W_n l'ensemble des $(q_1, q_2, \ldots, q_d) \in \mathbb{Z}^d$ tels que le cube $Q_w = \prod_1^d [\frac{q_\ell}{2^n}, \frac{q_\ell+1}{2^n}]$ rencontre K, on a $K \subset K_n = \bigcup_{w\in W_n} Q_w \subset U_k$.
On en déduit que $\mu(K) \leqslant \mu(K_n) = \nu(K_n) \leqslant \nu(U_k)$ et $\nu(K) \leqslant \nu(K_n) = \mu(K_n) \leqslant \mu(U_k)$. Donc $\mu(K) \leqslant \nu(K)$ et $\nu(K) \leqslant \mu(K)$. On déduit alors du corollaire 5.8.11 que $\mu = \nu$. ■

Théorème 5.9.10. *Pour tout entier d, il existe sur l'espace mesurable $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ une unique mesure σ-additive μ invariante par translation telle que $\mu([0,1]^d) = 1$.*

Démonstration : Le théorème 5.9.1 prouve que la mesure de Lebesgue est l'unique mesure σ-additive μ sur $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ telle que $\mu(K) = m(K)$ (où m est définie par la formule de 5.9.1), et que cette mesure est invariante par translation.
Supposons maintenant que ν soit une mesure sur $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ invariante par translation et telle que $\nu([0,1]^d) = 1$. Notons z_n la ν-mesure d'un cube compact de côté 2^{-n} (en raison de l'invariance par translation, cette mesure ne dépend pas du cube choisi). En coupant en deux chacune des arêtes d'un cube Q de côté 2^{-n}, on obtient un recouvrement de Q par 2^d cubes de côté 2^{-n-1}, ce qui montre que $2^d.z_{n+1} \geqslant z_n$, et donc que $z_{n+p} \geqslant 2^{-pd} z_n$.
D'autre part, il est aisé de trouver $2^p - 1$ intervalles compacts de longueur 2^{-np} deux-à-deux disjoints dans un intervalle de longueur 2^{-n}, et par produit, de trouver $(2^p-1)^d$ cubes compacts de côté 2^{-n-p} deux-à-deux disjoints dans un cube de côté 2^{-n} ; il en résulte que $(2^p-1)^d z_{n+p} \leqslant z_n$. On en déduit que

$$(2^p-1)^d.2^{(1-p)d} z_{n+1} \leqslant (2^p-1)^d z_{n+p} \leqslant z_n \leqslant 2^d z_{n+1} \ ,$$

donc $(\frac{2^p-1}{2^{p-1}})^d \leqslant \frac{z_n}{z_{n+1}} \leqslant 2^d$. Puisque p est arbitraire on en déduit que $z_n = 2^d z_{n+1}$, puis que $z_n = 2^{-nd}$.
De même, si P est le pavé compact $\prod_1^d [a_\ell, b_\ell]$ et si, pour tout $\ell \leqslant d$, on a

$$p_\ell < 2^n(b_\ell - a_\ell) \leqslant p_\ell + 1 \ ,$$

on peut – d'une part trouver p_ℓ intervalles compacts de longueur 2^{-n} deux-à-deux disjoints dans $[a_\ell, b_\ell]$, donc $\prod_1^d p_\ell$ cubes deux-à-deux disjoints de côté 2^{-n} dans P – d'autre part recouvrir $[a_\ell, b_\ell]$ par $p_\ell + 1$ intervalles de longueur 2^{-n}, donc recouvrir P par $\prod_1^d (p_\ell + 1)$ cubes de côté 2^{-n}. On en déduit que

$$\prod_1^d \frac{p_\ell}{2^n} = \prod_1^d p_\ell . z_n \leqslant \nu(P) \leqslant z_n . \prod_1^d (p_\ell + 1) = \prod_1^d \frac{p_\ell + 1}{2^n}$$

et en passant à la limite quand $n \to \infty$, $\nu(P) = v(P) = \mu(P) = \prod_1^d (b_\ell - a_\ell)$.
Alors μ et ν coïncident sur les pavés compacts, et il résulte du lemme 5.9.9 que $\mu = \nu$. ■

Théorème 5.9.11. *Si μ_n désigne la mesure de Lebesgue sur $\mathbb{R}^n$, on a $\mu_d \otimes \mu_\ell = \mu_{d+\ell}$, en identifiant $\mathbb{R}^{d+\ell}$ à $\mathbb{R}^d \times \mathbb{R}^\ell$.*

Démonstration : On voit aisément que la mesure $\mu_d \otimes \mu_\ell$ sur $\mathbb{R}^{d+\ell}$ est invariante par translation. Et si I_n désigne le cube unité $[0,1]^n$ dans $\mathbb{R}^n$, on a $I_{d+\ell} = I_d \times I_\ell$, donc

$$(\mu_d \otimes \mu_\ell)(I_{d+\ell}) = \mu_d(I_d).\mu_\ell(I_\ell) = 1$$

et ceci permet d'identifier $\mu_d \otimes \mu_\ell$ à la mesure de Lebesgue sur $\mathbb{R}^{d+\ell}$. ■

Lemme 5.9.12. *Soit f une fonction mesurable positive sur l'espace mesuré $(X, \mathfrak{S}, \mu)$. Alors, si μ_1 est la mesure de Lebesgue sur $\mathbb{R}$, on a*

$$\int_X f(x)\, d\mu(x) = \int_{\mathbb{R}^+} \mu(\{f > t\})\, d\mu_1(t)\ .$$

Démonstration : La fonction $(x, t) \mapsto f(x) - t$ est $\mathfrak{S} \otimes \mathfrak{B}(\mathbb{R}^+)$ mesurable sur $X \times \mathbb{R}^+$. Il en résulte que l'ensemble $A = \{(x, t) : f(x) - t > 0\}$ est mesurable. On déduit alors de la définition de $\mu \otimes \mu_1$ et du théorème de Fubini (théorème 5.7.8), que

$$\mu \otimes \mu_1(A) = \int \mathbb{1}_A(x, t)\, d(\mu \otimes \mu_1)(x, t) = \int_X \mu_1(A^x)\, d\mu(x) = \int_{\mathbb{R}^+} \mu(A_t)\, d\mu_1(t)\ .$$

Et puisque $A^x = [0, f(x)[$, donc $\mu_1(A^x) = f(x)$, et $A_t = \{f > t\}$, ceci achève la preuve. ∎

5.10 Primitives

Un des premiers objectifs du calcul intégral, qui a motivé l'introduction de la mesure de Lebesgue sur $\mathbb{R}$ est le calcul des primitives, c'est-à-dire la recherche des fonctions dont on connaît la dérivée.

Théorème 5.10.1. *Si f est une fonction réelle ou complexe continue sur un intervalle $]a, b[$ de $\mathbb{R}$, elle possède une primitive sur cet intervalle, c'est-à-dire qu'il existe une fonction F de classe $\mathscr{C}^1$ dont elle est la dérivée. De plus, si $a < u \leqslant v < b$, on a : $F(v) - F(u) = \int_{[u,v]} f(x)\, dx$.*

Démonstration : Fixons $c \in]a, b[$. Puisque f est continue, elle est mesurable, et bornée, donc intégrable, sur tout compact de $]a, b[$. Si on définit alors la fonction F sur $]a, b[$ par

$$F(x) = \begin{cases} 0 & \text{si } x = c \\ \int_{[c,x]} f(t)\, dt & \text{si } x \geqslant c \\ -\int_{[x,c]} f(t)\, dt & \text{si } x \leqslant c \end{cases}$$

on vérifie immédiatement qu'on a, pour $u < v$, $F(v) - F(u) = \int_{[u,v]} f(t)\, dt$. Pour x_0 fixé dans $]a, b[$ et $\varepsilon > 0$, la continuité de f en x_0 permet de trouver $\eta > 0$ el que $[x_0 - \eta, x_0 + \eta] \subset]a, b[$ et que $|f(t) - f(x_0)| < \varepsilon$ si $|t - x_0| \leqslant \eta$. Il en résulte que si $x_0 - \eta \leqslant u < v \leqslant x_0 + \eta$, on a

$$|F(v) - F(u) - (v - u)f(x_0)| \leqslant \int_{[u,v]} |f(t) - f(x_0)|\, dt \leqslant \varepsilon(v - u)\ ,$$

ou encore $\left|\dfrac{F(v) - F(u)}{v - u} - f(x_0)\right| \leqslant \varepsilon$, puis, en prenant soit $u = x_0 < v \leqslant x_0 + \eta$, soit $x_0 - \eta \leqslant u < v = x_0$ qu'on a $f(x_0) = \lim_{h \to 0, h \neq 0} \dfrac{F(x_0 + h) - F(x_0)}{h}$, c'est-à-dire $f(x_0) = F'(x_0)$, et que F est bien une primitive de f. ∎

Théorème 5.10.2. *Si g est une fonction bornée sur l'intervalle $]a,b[$ de $\mathbb{R}$ et si g possède une primitive f (c'est-à-dire si f est partout dérivable de dérivée g), alors on a, pour $u \leqslant v$,*

$$f(v) - f(u) = \int_{[u,v]} g(t)\,dt\ .$$

DÉMONSTRATION : Si $[u,v] \subset]a,b[$, on a, pour n assez grand, $v + 2^{-n} < b$, donc

$$\begin{aligned}\int_{[u,v]} 2^n\big(f(t+2^{-n}) - f(t)\big)\,dt &= 2^n\Big(\int_{[u+2^{-n},v+2^{-n}]} f(t)\,dt - \int_{[u,v]} f(t)\,dt\Big)\\ &= 2^n \int_{[v,v+2^n]} f(t)\,dt - 2^n \int_{[u,u+2^n]} f(t)\,dt\ .\end{aligned}$$

Alors la fonction g_n définie sur $[u,v]$ par $g_n(t) = 2^n(f(t+2^{-n}) - f(t))$ est, par la formule des accroissements finis, partout majorée en module par $M = \sup_{t\in]a,b[} |f'(t)| < \infty$, et on a $g_n(t) \to f'(t) = g(t)$ pour tout t. On déduit alors du théorème de convergence dominée que $\int_{[u,v]} g_n(t)\,dt \to \int_{[u,v]} g(t)\,dt$. Comme la fonction f est dérivable, donc continue, il résulte du théorème précédent que $2^n \int_{[x,x+2^{-n}]} f(t)\,dt \to f(x)$ pour tout $x \in]a,b[$ (en particulier pour $x = u$ et pour $x = v$), donc que $f(v) - f(u) = \int_{[u,v]} g(t)\,dt$, ce qui permet de retrouver une primitive de g par intégration. ∎

Remarque 5.10.3. On ne peut, dans l'énoncé précédent, affaiblir l'hypothèse en supposant seulement f continue et dérivable presque partout, avec dérivée $g \in L^\infty$. Il existe en effet une fonction croissante et continue f sur $\mathbb{R}$, la “fonction de Lebesgue”, qui vaut 0 en 0 et 1 en 1, et qui est localement constante sur le complémentaire U d'un compact négligeable K : cette fonction est dérivable (avec dérivée nulle) en tout point de U, donc presque partout. En particulier f' est dans L^∞, mais néanmoins $f(1) - f(0) = 1 \neq \int_0^1 f'(t)\,dt = 0$.

Pour cela, on construit par récurrence une suite (f_n) de fonctions continues croissantes sur $[0,1]$ telles que $f_0(x) = x$ et que f_n soit affine sur chaque intervalle $J_{n,p} = [p.3^{-n}, (p+1).3^{-n}]$. La fonction f_{n+1} coïncide avec f_n en chacun des points $p.3^{-n}$ pour $p = 0, 1, \ldots, 3^n$ et est constante sur le tiers central $J_{n+1,3p+1}$ de $J_{n,p}$ avec la valeur $\frac{1}{2}\big(f_n(p.3^{-n}) + f_n((p+1).3^{-n})\big)$. On vérifie alors qu'en chaque point x de $J_{n,p}$, on a

$$|f_{n+1}(x) - f_n(x)| \leqslant \frac{1}{2}\big(f_n((p+1).3^{-n}) - f_n(p.3^{-n})\big) \leqslant 2^{-n}\ ,$$

donc que la série de terme général $(f_{n+1} - f_n)$ converge normalement sur $[0,1]$. Donc la suite (f_n) converge vers une fonction f continue et croissante. De plus, lorsque f_n est constante sur un intervalle $J_{n,p}$, les fonctions f_m pour $m \geqslant n$ coïncident avec f_n sur cet intervalle, et la fonction f aussi. La réunion K_n des 2^n intervalles $J_{n,p}$ sur lesquels f_n n'est pas constante a donc une mesure $(2/3)^n$, et f est localement constante sur $U = [0,1] \setminus \bigcap_n K_n$, qui est un ouvert de mesure 1, puisque $K = \bigcap_n K_n$ est un compact négligeable.

5.11 La formule du changement de variables

Théorème 5.11.1. *Si Φ est une application linéaire bijective de $\mathbb{R}^d$ sur $\mathbb{R}^d$, et si μ est la mesure de Lebesgue sur $\mathbb{R}^d$, on a, pour tout $A \in \mathfrak{B}(\mathbb{R}^d)$: $\mu(\Phi(A)) = \mu(A).|\det(\Phi)|$.*

Si on pose, pour $A \in \mathfrak{B}(\mathbb{R}^d)$: $\nu(A) = \mu(\Phi(A))$, on vérifie immédiatement que $\Phi(\mu) := \nu$ est une mesure σ-additive finie sur les compacts. De plus, pour $t \in \mathbb{R}^d$ et $A \in \mathfrak{B}(\mathbb{R}^d)$, on a $\Phi(A+t) = \Phi(A) + u$, où $u = \Phi(t)$. Alors l'invariance de μ par translation assure que $\nu(A) = \nu(A+t)$, donc que ν elle-même est invariante par translation. Et si I est le cube $[0,1]^d$, la mesure $\nu_0 = \dfrac{1}{\nu(I)}.\nu$ est invariante par translation et vaut 1 sur I : on a donc $\nu_0 = \mu$, c'est-à-dire $\Phi(\mu) = \nu(I).\mu$. Il existe donc une fonction positive ψ sur le groupe linéaire $GL(d,\mathbb{R})$ telle que $\psi(\Phi) = \mu(\Phi(I))$, donc que $\Phi(\mu) = \psi(\Phi).\mu$, c'est-à-dire $\mu(\Phi(A)) = \psi(\Phi).\mu(A)$, pour $A \in \mathfrak{B}(\mathbb{R}^d)$. Et on a clairement $\psi(\varphi_1 \circ \varphi_2) = \psi(\varphi_1).\psi(\varphi_2)$.
Si Ψ est une isométrie de $\mathbb{R}^n$, et si B désigne la boule unité, B est compacte et contient un cube de côté $d^{-1/2}$: on a donc $0 < \mu(B) < \infty$. De plus, $\Psi(B) = B$, et on en déduit que $\mu(B) = \mu(\Psi(B)) = \psi(\Psi).\mu(B)$, donc que $\psi(\Psi) = 1 = |\det(\Psi)|$.
Si Φ a une matrice diagonale, de valeurs propres $(\lambda_1, \lambda_2, \dots, \lambda_d)$, l'image du cube unité par Φ est un pavé de côtés $\lambda_1, \lambda_2, \dots, \lambda_n$, dont le volume est $\psi(\Phi) = \prod_j |\lambda_j| = |\det(\Phi)|$.
Enfin, tout élément Φ de $GL(d,\mathbb{R})$ peut se mettre sous la forme $\Phi = \Psi_1 \circ \Phi_0 \circ \Psi_2$, où Ψ_1 et Ψ_2 sont des isométries, et Φ_0 est diagonale : il en résulte que

$$\psi(\Phi) = \psi(\Psi_1).\psi(\Phi_0).\psi(\Psi_2) = \psi(\Phi_0) = |\det(\Phi_0)| = |\det(\Psi_1).\det(\Phi_0).\det(\Psi_2)| = |\det(\Phi)| \ ,$$

et ceci achève la démonstration. ■

Lemme 5.11.2. *Soient U et V deux ouverts de $\mathbb{R}^d$, φ un $\mathscr{C}^1$-difféomorphisme de U sur V, a un point de U, $b = \varphi(a) \in V$ et $\Phi = \varphi'(a) \in GL(d,\mathbb{R})$. Si K est un voisinage compact symétrique de 0 dans $\mathbb{R}^d$, il existe pour tout $\varepsilon > 0$ un entier n_0 tel que, si $n \geqslant n_0$ et $a \in x + 2^{-n}K$, il existe y tel que*

$$y + 2^{-n}(1-\varepsilon)\Phi(K) \subset \varphi(x + 2^{-n}K) \subset y + (1+\varepsilon).2^{-n}\Phi(K) \ .$$

Démonstration : On peut supposer $\varepsilon < \dfrac{1}{2}$. Notons B la boule unité de $\mathbb{R}^d$. Puisque K est un voisinage de 0, et que Φ est un homéomorphisme de $\mathbb{R}^d$ sur $\mathbb{R}^d$, il existe $r > 0$ tel que $\Phi(K)$ contienne la boule de rayon $r.\operatorname{diam}(K)$. Et puisque φ est différentiable en a, avec $\varphi'(a) = \Phi$, on a $f(z) = f(a) + \Phi.(z-a) + o(z-a)$.
Il existe donc $\eta > 0$ tel que $\|f(z) - b - \Phi.(z-a)\| < r\varepsilon \|z-a\|$ lorsque $\|z-a\| < \eta$. Donc si on choisit n_0 tel que $2^{1-n_0}.\operatorname{diam}(K) < \eta$, on a $z - a < \eta$ si $z \in x + 2^{-n}K$ lorsque $n \geqslant n_0$ et $a \in x + 2^{1-n}K$. De plus $z = x + 2^{-n}h$, avec $h \in K$, donc

$$\left\|\varphi(z) - b - \Phi(x-a) - 2^{-n}\Phi.h\right\| \leqslant r\varepsilon \|z-a\| \leqslant r\varepsilon 2^{-n}\operatorname{diam}(K) \ ,$$

donc, en posant $y = b + \Phi.(x-a)$, on obtient $\varphi(z) - y - 2^{-n}\Phi.h \in \varepsilon.2^{-n}\Phi(K)$, et finalement $\varphi(z) \in y + 2^{-n}(1+\varepsilon)\Phi(K)$, donc $\varphi(x + 2^{-n}K) \subset y + (1+\varepsilon).2^{-n}.\Phi(K)$.
Le même raisonnement appliqué à φ^{-1}, dont la différentielle en b est Φ^{-1}, et au convexe compact symétrique $(1-\varepsilon)\Phi(K)$ montre que pour n asssez grand, et $x = a + \Phi^{-1}.b$, on a

$$\varphi^{-1}(y + (1-\varepsilon)2^{-n}.\Phi(K)) \subset x + (1+\varepsilon)(1-\varepsilon).2^{-n}K \subset x + 2^{-n}.K$$

lorsque $b \in y + 2^{1-n}(1-\varepsilon)\Phi(K)$. c'est-à-dire $\Phi.(x-a) \in 2^{1-n}(1-\varepsilon).\Phi(K)$, ou encore $a \in x + (1-\varepsilon)2^{1-n}K$. Il en résulte que si $a \in x + 2^{-n}.K \subset x + (1-\varepsilon)2^{1-n}K$ et n est assez grand, on a $y + 2^{-n}(1-\varepsilon).\Phi(K) \subset \varphi(x + 2^{-n}K)$. ■

Lemme 5.11.3. *Soient U un ouvert de $\mathbb{R}^d$, μ_1 et μ_2 deux mesures σ-additives et localement finies sur $(U, \mathfrak{B}(U))$. On suppose que Q_0 est un pavé compact d'intérieur non vide contenu dans U, que $\mu_2(Q' \cap Q'') = 0$ si Q' et Q'' sont deux pavés contenus dans Q_0 dont les intérieurs sont disjoints, et que $\mu_1(Q_0) \geqslant (1+\varepsilon).\mu_2(Q_0)$. Alors on peut trouver une suite décroissante (Q_n) de pavés compacts compacts homothétiques à Q_0 dans le rapport 2^{-n}, tels que $\mu_1(Q_n) \geqslant (1+\varepsilon)\mu_2(Q_n)$.*

Démonstration : Si le pavé Q_n est défini, on détermine 2^d pavés $(Q_{n,j})$ homothétiques à Q_n dans le rapport $\dfrac{1}{2}$ en coupant en deux parties égales les arêtes de Q_n. Ces cubes ont deux-à-deux des intérieurs disjoints, et on a donc $\mu_2(Q_j \cap Q_{j'}) = 0$ si $j \neq j'$. Il en résulte que $\mu_1(Q_n) \leqslant \sum_j \mu(Q_{n,j})$ et $\nu(Q_n) = \sum_j \nu(Q_{n,j})$. Si on avait $\mu(Q_{n,j}) < (1+\varepsilon)\nu(Q_{n,j})$ pour tout j, on aurait alors

$$\mu_1(Q_n) \leqslant \sum_j \mu_1(Q_{n,j}) < (1+\varepsilon)\sum_j \mu_2(Q_{n,j}) = (1+\varepsilon)\mu_2(Q_n) \ ,$$

contrairement à l'hypothèse de récurrence.
Il existe donc un j_0 tel que $\mu_1(Q_{n,j}) \geqslant (1+\varepsilon)\mu_2(Q_{n,j_0})$, et on peut choisir $Q_{n+1} = Q_{n,j_0}$. ■

Théorème 5.11.4. *Soient U et V deux ouverts de $\mathbb{R}^d$ et φ un $\mathscr{C}^1$-difféomorphisme de U sur V. Alors, pour toute fonction mesurable positive f sur V, on a*

$$\int_V f(x)\, d\mu(x) = \int_U f{\circ}\varphi(y)\, |J_\varphi(y|\; d\mu(y) \ ,$$

où $J_\varphi(y)$ désigne le déterminant jacobien de φ en y et μ la mesure de Lebesgue.

Démonstration : On considère sur $(U, \mathfrak{B}(U))$ les deux mesures ν_1 et ν_2 définies par : $\nu_1(A) = \mu(\varphi(A))$ et $\nu_2(A) = \int_A |J_\varphi(x)|\, d\mu(x)$ si $A \in \mathfrak{B}(U)$, dont on vérifie aisément qu'elles sont σ-additives et localement finies (noter que $\varphi(A) \in \mathfrak{B}(V)$ puisque φ^{-1} est continue). On veut montrer que $\nu_1 = \nu_2$: d'après le lemme 5.9.9, il suffit pour cela de montrer que, pour tout pavé compact $Q \subset U$, on a $\nu_1(Q) = \nu_2(Q)$. Et puisque tout pavé compact est intersection d'une suite décroissante de pavés compacts d'intérieur non vide, il suffit de le faire pour les pavés compacts d'intérieur non vide.
Supposons qu'il existe un pavé compact Q d'intérieur non vide tel que $\nu_1(Q) \neq \nu_2(Q)$; il existerait alors $\alpha > 0$ tel que $\nu_1(Q) \geqslant (1+\alpha)\nu_2(Q)$ (ou $\nu_2(Q) \geqslant (1+\alpha)\nu_1(Q)$).
Si Q' et Q'' sont des pavés dont les intérieurs sont disjoints, alors $Q' \cap Q''$ est un pavé μ-négligeable, et $\nu_2(Q' \cap Q'') = \int_{Q' \cap Q''} |J_\varphi(x)|\, d\mu(x) = 0$.
Si Q' et Q'' sont deux pavés compacts d'intérieurs disjoints, et si $\nu_1(Q' \cap Q'') > 0$, alors $Q' \cap Q''$ est un pavé d'intérieur vide, intersection d'une suite décroissante (P_k) de pavés compacts contenus dans U et d'intérieur non vide. On a alors $\nu_2(P_k) \to \nu_2(Q' \cap Q'') = 0$ et $\nu_1(P_k) \geqslant \nu_1(Q' \cap Q'') > 0$. On en déduit que, pour un entier k, on a $\nu_1(P_k) \geqslant (1+\alpha).\nu_2(P_k)$.

Appliquant le lemme précédent selon le cas à $\mu_1 = \nu_1$ et $\mu_2 = \nu_2$, ou à $\mu_1 = \nu_2$, $\mu_2 = \nu_1$ et en remplaçant Q par P_k, on peut donc supposer que $\mu_2(Q' \cap Q'') = 0$ lorsque Q' et Q'' sont des pavés d'intérieurs disjoints et que $\mu_1(Q) \geqslant (1+\alpha)\mu_2(Q)$.
On peut alors trouver une suite décroissante (Q_n) de pavés homothétiques à Q dans le rapport 2^{-n} tels que $\mu_1(Q_n) \geqslant (1+\alpha)\mu_2(Q_n)$). Et il existe un $a \in U$ tel que $\bigcap_n Q_n = \{a\}$. Il existe un voisinage convexe compact symétrique K de 0 dont Q est un translaté ;

alors $\mu(Q_n) = 2^{-nd}\mu(Q)$. Pour tout $\varepsilon > 0$, on a $\sup_{x\in Q_n} |J_\varphi(x)| < (1+\varepsilon)\,|J_\varphi(a)|$ et $\inf_{x\in Q_n} |J_\varphi(x)| > (1-\varepsilon)\,|J_\varphi(a)|$ pour n assez grand, donc

$$(1-\varepsilon)2^{-nd}\,|J_\varphi(a)|\,\mu(Q) < \nu_2(Q_n) < (1+\varepsilon)2^{-nd}\,|J_\varphi(a)|\,\mu(Q)\ ,$$

donc $\nu_2(Q_n) \simeq 2^{-nd}\,|J_\varphi(a)|\,\mu(Q) = 2^{-nd}\,|\det(\Phi)|\,\mu(Q)$ lorsque $n \to \infty$ (avec $\Phi = \varphi'(a)$).
De même, pour $\varepsilon > 0$, il résulte du lemme 5.11.2 que, pour n assez grand, $\varphi(Q_n)$ est contenu dans un homothétique de rapport $(1+\varepsilon).2^{-n}$ du convexe $\Phi(K)$, lui-même translaté de $\Phi(Q)$, et que $\varphi(Q_n)$ contient un homothétique de rapport $2^{-n}(1-\varepsilon)$ du convexe $\Phi(K)$. On en déduit que

$$\nu_1(Q_n) \leqslant (1+\varepsilon)^d 2^{-nd}\mu(\Phi(K)) = (1+\varepsilon)^d 2^{-nd}\,|\det(\Phi)|\,\mu(Q)$$

et que

$$\nu_1(Q_n) \geqslant (1-\varepsilon)^d 2^{-nd}\mu(\Phi(K)) = (1-\varepsilon)^d 2^{-nd}\,|\det(\Phi)|\,\mu(Q)\ ,$$

donc $\nu_1(Q_n) \simeq 2^{-nd}\,|\det(\Phi)|\,\mu(Q)$ lorsque $n \to \infty$. Il en résulte que $\dfrac{\nu_2(Q_n)}{\nu_1(Q_n)} \to 1$, donc $\dfrac{\mu_2(Q_n)}{\mu_1(Q_n)} \to 1$, ce qui contredit le choix de Q_n, et achève de prouver que $\nu_1 = \nu_2$.
Si f est une fonction élémentaire sur V : $f = \sum_j \lambda_j \mathbb{1}_{E_j}$, avec $E_j \in \mathfrak{B}(V)$, on a, en posant $A_j = \varphi^{-1}(E_j) \in \mathfrak{B}(U)$,

$$\int_U f\circ\varphi(x)\,d\nu_1(x) = \int_V \sum_j \lambda_j \mathbb{1}_{A_j}(x)\,d\nu_1(y) = \sum_j \lambda_j \mu(\varphi(A_j)) = \sum_j \lambda_j \mu(E_j) = \int f(y)\,d\mu(y)$$

et

$$\int f\circ\varphi(x)\,d\nu_2(x) = \int_U \sum_j \lambda_j \mathbb{1}_{A_j}(x).\,|J_\varphi(x)|\,d\mu(x) = \int_U f\circ\varphi(x)\,|J_\varphi(x)|\,d\mu(x)\ .$$

Alors, pour une fonction mesurable positive f, il existe une suite croissante de fonctions élémentaires positives (f_n) qui converge vers f, et il résulte du théorème de convergence monotone, appliqué à (f_n) et à $(f_n\circ\varphi.\,|J_\varphi|)$ que

$$\int f(y)\,d\mu(y) = \sup_n \int f_n(y)\,d\mu(y) = \sup_n \int f_n\circ\varphi(x)\,|J_\varphi(x)|\,d\mu(x) = \int f\circ\varphi(x)\,|J_\varphi(x)|\,d\mu(x)$$

et achève la démonstration. ∎

Théorème 5.11.5. *Soient U et V deux ouverts de $\mathbb{R}^d$ et φ un $\mathscr{C}^1$-difféomorphisme de U sur V. Alors, pour toute fonction mesurable f sur V, la fonction $f\circ\varphi.J_\varphi$ est intégrable sur U si et seulement si f est intégrable sur V, et on a*

$$\int_V f(x)\,d\mu(x) = \int_U f\circ\varphi(y)\,|J_\varphi(y|\,d\mu(y)\ ,$$

où $J_\varphi(y)$ désigne le déterminant jacobien de φ en y.

Démonstration : Cet énoncé se déduit aisément du précédent : en appliquant à la fonction positive $|f|$ le théorème 5.11.4, on voit que $|f\circ\varphi|\,.\,|J_\varphi|$ est intégrable sur U si et seulement si $|f|$ est intégrable sur V.
Si f est intégrable, en l'écrivant sous la forme $f_1 - f_2 + ig_1 - ig_2$ avec des fonctions positives toutes quatre intégrables et en appliquant le théorème 5.11.4 à chacune d'entre elles, on obtient l'égalité souhaitée. ∎

5.12 Inégalités de convexité

Lemme 5.12.1. *Soient p, q et r positifs, tels que $\dfrac{1}{p}+\dfrac{1}{q}+\dfrac{1}{r}=1$. On a alors, pour u, v et w dans $\mathbb{R}^+$:*

$$uvw \leqslant \frac{u^p}{p}+\frac{v^q}{q}+\frac{w^r}{r}\ .$$

De même, si $\dfrac{1}{p}+\dfrac{1}{p'}=1$ (on dit alors que p et p' sont des exposants conjugués*), on a, pour u et v réels : $uv \leqslant \dfrac{|u|^p}{p}+\dfrac{|v|^{p'}}{p'}$.*

Démonstration : Sur le compact $T=\{(x,y,z)\in(\mathbb{R}^+)^3 : \dfrac{x}{p}+\dfrac{y}{q}+\dfrac{z}{r}=1\}$ la fonction de classe $\mathscr{C}^1$ $f:(x,y,z)\mapsto x^{1/p}.y^{1/q}.z^{1/r}$ atteint son maximum m, qui est strictement positif, en un point (a,b,c) tel que $abc>0$. Il existe donc un multiplicateur de Lagrange λ tel que

$$\begin{cases} \dfrac{1}{p}a^{1/p-1}b^{1/q}c^{1/r} = \lambda\dfrac{1}{p} \\ \dfrac{1}{q}a^{1/p}b^{1/q-1}c^{1/r} = \lambda\dfrac{1}{q} \\ \dfrac{1}{r}a^{1/p}b^{1/q}c^{1/r-1} = \lambda\dfrac{1}{p} \end{cases}$$

donc $a^{1/p}b^{1/q}c^{1/r}=\lambda a=\lambda b=\lambda c$. On a nécessairement $\lambda>0$, donc $a=b=c=1$, et $m=1$. La fonction f étant positivement homogène, on en déduit que $f(x,y,z)\leqslant \dfrac{x}{p}+\dfrac{y}{q}+\dfrac{z}{r}$ pour tous x, y et z réels positifs. En prenant alors $x=u^p$, $y=v^q$ et $w=z^r$, on obtient la première inégalité cherchée.
En prenant $w=1$ et en faisant tendre r vers $+\infty$, l'application de la première inégalité à $|u|$ et $|v|$ donne la seconde. ■

Lemme 5.12.2. *Soient p, q et r supérieurs à 1 et vérifiant $\dfrac{1}{p}+\dfrac{1}{q}+\dfrac{1}{r}=2$. On a alors, pour u, v et w réels positifs :*

$$uvw \leqslant \frac{v^q w^r}{p'}+\frac{w^r u^p}{q'}+\frac{u^p v^q}{r'}\ ,$$

où p', q' et r' sont définis par $\dfrac{1}{p}+\dfrac{1}{p'}=\dfrac{1}{q}+\dfrac{1}{q'}=\dfrac{1}{r}+\dfrac{1}{r'}=1$.

Démonstration : L'inégalité est claire si $uvw=0$. On remarque que $\dfrac{1}{p'}+\dfrac{1}{q'}+\dfrac{1}{r'}=1\,$; il résulte alors du lemme 5.12.1 que pour x,y et $z>0$, on a

$$x^{1/p'}y^{1/q'}z^{1/r'} \leqslant \frac{x}{p'}+\frac{y}{q'}+\frac{z}{r'}\ ,$$

donc, en divisant par xyz :

$$x^{-1/p}y^{-1/q}z^{-1/r} \leqslant \frac{1}{p'yz}+\frac{1}{q'zx}+\frac{1}{r'xy}\ ,$$

et remplaçant x par u^{-p}, y par v^{-q} et z par w^{-r}, on obtient l'inégalité cherchée. ■

Théorème 5.12.3. *(Inégalité de Hölder). Soient $(X, \mathfrak{S}, \mu)$ un espace mesuré, f et g deux fonctions complexes mesurables sur X, p et p' deux exposants conjugués. On a*

$$\int_X |f(x)g(x)|\, d\mu(x) \leqslant \left(\int_X |f(x)|^p\, d\mu(x)\right)^{1/p} \left(\int_X |g(x)|^{p'}\, d\mu(x)\right)^{1/p'} .$$

En particulier, si $|f|^p$ et $|g|^{p'}$ sont dans $L^1(X, \mu)$, il en est de même de fg et on a

$$\left|\int_X f(x)g(x)\, d\mu(x)\right| \leqslant \left(\int_X |f(x)|^p\, d\mu(x)\right)^{1/p} \left(\int_X |g(x)|^{p'}\, d\mu(x)\right)^{1/p'} .$$

Démonstration : L'inégalité est claire si l'un des facteurs de droite est nul ou infini. Sinon, on pose $\alpha = \left(\int_X |f(x)|^p\, d\mu(x)\right)^{1/p}$, $\beta = \left(\int_X |g(x)|^{p'}\, d\mu(x)\right)^{1/p'}$, $f_1(x) = \dfrac{1}{\alpha} f(x)$, $g_1(x) = \dfrac{1}{\beta} g(x)$. On a alors, pour tout x,

$$|f(x)g(x)| = \alpha\beta\, |f_1(x)|\, |g_1(x)| \leqslant \alpha\beta \left(\frac{1}{p} |f_1(x)|^p + \frac{1}{p'} |g_1(x)|^{p'}\right)$$

et en intégrant cette inégalité, on trouve

$$\int_X |f(x)g(x)|\, d\mu(x) \leqslant \alpha\beta(\frac{1}{p} + \frac{1}{p'}) = \alpha\beta ,$$

ce qui est l'inégalité annoncée. Si, de plus, $\alpha < +\infty$ et $\beta < +\infty$, on conclut que $|fg|$ est intégrable et que

$$\left|\int_X f(x)g(x)\, d\mu(x)\right| \leqslant \int_X |f(x)g(x)|\, d\mu(x) \leqslant \alpha\beta(\frac{1}{p} + \frac{1}{p'}) = \alpha\beta ,$$

■

Corollaire 5.12.4. *Soient f une fonction complexe mesurable sur X, p et p' deux exposants conjugués. On a alors*

$$\left(\int_X |f(x)|^p\, d\mu(x)\right)^{1/p} = \sup\left\{\left|\int_X f(x)g(x)\, d\mu(x)\right| : \int_X |g(x)|^{p'}\, d\mu(x) \leqslant 1\right\} .$$

Démonstration : Supposons d'abord $|f|^p$ intégrable. Compte tenu de l'inégalité de Hölder, il suffit de trouver une fonction g réalisant l'égalité. On prend $\alpha = p - 2$ et $g_1(x) = \bar{f}(x). |f(x)|^\alpha$ si $f(x) \neq 0$, ainsi que $g_1(x) = 0$ si $f(x) = 0$.
On a alors $f(x)g_1(x) = |f(x)|^{\alpha+2} = |f(x)|^p$ et $|g_1(x)|^{p'} = |f(x)|^{(\alpha+1)p'} = |f(x)|^p$.
On prend alors $g(x) = \lambda g_1(x)$, où $\lambda = \left(\int_X |f(x)|^p\, d\mu(x)\right)^{-1/p'}$. Il en résulte que $\int_X |g(x)|^{p'}\, d\mu(x) = 1$ et que

$$\begin{aligned}\int_X f(x)g(x)\, d\mu(x) = \lambda \int_X |f(x)|^p\, d\mu(x) &= \left(\int_X |f(x)|^p\, d\mu(x)\right)^{1-1/p'} \\ &= \left(\int_X |f(x)|^p\, d\mu(x)\right)^{1/p} .\end{aligned}$$

Enfin, si $|f|^p$ n'est pas intégrable, l'hypothèse de σ-finitude de μ permet de trouver dans $\mathfrak{S}$ une suite croissante (A_n) de parties de mesure finie telles que $X = \bigcup_n A_n$. On a alors

$$\int_X |f(x)|^p \, d\mu(x) = \sup_n \int_{A_n} \big(\inf(n, |f(x)|)\big)^p \, d\mu(x) = +\infty \ .$$

En appliquant à $f_n := \inf(n, |f(x)|) . \dfrac{f(x)}{|f(x)|}$ ce qui précède, on trouve une fonction g_n telle que $\int_X |g_n(x)|^{p'} \, d\mu(x) = 1$ et que $\left|\int_X f(x) g_n(x) \, d\mu(x)\right| = \left(\int_X |f_n(x)|^p \, d\mu(x)\right)^{1/p} \to +\infty$. ■

Théorème 5.12.5. *Soit $p \in]1, \infty[$. L'ensemble $L^p(X, \mathfrak{S}, \mu)$ des (classes de) fonctions complexes mesurables f telles que $|f|^p$ soit intégrable est un espace vectoriel. Sur cet espace vectoriel la fonction $\|.\|_p : f \mapsto \left(\int |f(x)|^p \, d\mu(x)\right)^{1/p}$ est une norme pour laquelle L^P est un espace de Banach.*
En particulier, muni du produit scalaire $\langle f, g\rangle = \int f(x)\bar{g}(x) \, d\mu(x)$, $L^2(X, \mathfrak{S}, \mu)$ est un espace de Hilbert.

Démonstration : Il est clair que si $f \in L^P(X, \mathfrak{S}, \mu)$, il en est de même de λf pour $\lambda \in \mathbb{C}$. Si f et g sont dans L^p, on a pour tout x

$$|f(x) + g(x)|^p \leqslant 2^p \sup(|f(x)|^p, |g(x)|^p) \leqslant 2^p |f(x)|^p + 2^p |g(x)|^p \ ,$$

ce dont on déduit que $f + g \in L^p(X, \mathfrak{S}, \mu)$ puisque $\int_X |f(x) + g(x)|^p \, d\mu(x) < +\infty$.
Si on note $p' = \dfrac{p}{p-1}$ et B' l'ensemble des fonctions g telles que $\int_X |g(x)|^{p'} \, d\mu(x) \leqslant 1$, il résulte du corollaire 5.12.4 que, pour toute $f \in L^p$, $\|f\|_p := \left(\int |f(x)|^p \, d\mu(x)\right)^{1/p}$ est égal à $\sup_{g \in B'} \left|\int f(x) g(x) \, d\mu(x)\right|$. Et puisque, pour $g \in B'$,

$$\left|\int (f_1 + f_2)(x) g(x) \, d\mu(x)\right| \leqslant \left|\int f_1(x) g(x) \, d\mu(x)\right| + \left|\int f_2(x) g(x) \, d\mu(x)\right| \leqslant \|f\|_p + \|g\|_p \ ,$$

on vérifie sans peine que $\|f_1 + f_2\|_p \leqslant \|f_1\|_p + \|f_2\|_p$, et que $\|.\|_p$ est une semi-norme sur L^p. De plus si $\|f\|_p = 0$, la fonction positive $|f|^p$ est d'intégrale nulle, donc nulle presque partout, ce qui signifie que $f = 0$ dans L^p.
Pour montrer que L^p est complet, on va montrer que toute série normalement convergente est convergente. Soit donc (f_n) une suite telle que $\alpha := \sum_n \|f_n\|_p < +\infty$. Si on pose $f'_j(x) = |f_j(x)|$ et $\sigma_n(x) = \sum_{j=0}^m f'_j(x)$, on a $\sigma_n \leqslant \sigma_{n+1}$ et

$$\int \sigma_n(x)^p \, d\mu(x) = \left(\left\|\sum\nolimits_{j=0}^n f'_j\right\|_p\right)^p \leqslant \left(\sum_{j=0}^n \|f'_j\|_p\right)^p = \left(\sum_{j=0}^n \|f_j\|_p\right)^p \leqslant \alpha^p \ .$$

Il résulte alors du théorème de convergence monotone que $\sigma = \lim_n \sigma_n$ vérifie

$$\int \sigma(x)^p \, d\mu(x) \leqslant \alpha^p < \infty \ .$$

Donc $\sigma(x) < \infty$ pour presque tout x, ce qui signifie que la série de terme général $f_j(x)$ est absolument convergente pour presque tout x. En notant alors $s_n(x) = \sum_{j=0}^n f_j(x)$ et

$s(x) = \sum_{j=0}^{\infty} f_j(x)$, on a $|s_n(x) - s(x)|^p \to 0$ presque partout et $|s_n(x) - s(x)|^p \leqslant \sigma(x)^p$. Alors le théorème de convergence dominée montre que $\|s_n - s\|_p^p \to 0$, c'est-à-dire que $s_n \to s$ dans L^p, et que la série converge.

Si $p = 2$, la fonction $(f, g) \mapsto \langle f, g\rangle = \int f(x)\bar{g}(x)\, d\mu(x)$ définit un produit scalaire hermitien sur L^2 puisque, en vertu de l'inégalité de Hölder, la fonction $f\bar{g}$ est intégrable. Et on a clairement, pour $f \in L^2$: $\|f\|_2^2 = \langle f, f\rangle$. ■

Corollaire 5.12.6. *Si $1 \leqslant p < r < q$, on a $L^p \cap L^q \subset L^r$. Et pour $f \in L^p \cap L^q$ on a $\|f\|_r \leqslant \|f\|_p^{\alpha} \|f\|_q^{1-\alpha}$ où $\alpha = \frac{p(q-r)}{r(q-p)}$.*

Démonstration : Il est clair que, pour tout $t \in \mathbb{R}^+$, on a $t^r \leqslant \max(t^p, t^q) \leqslant t^p + t^q$, donc que, si $f \in L^p \cap L^q$:

$$\int |f(x)|^r \, d\mu(x) \leqslant \int |f(x)|^p \, d\mu(x) + \int |f(x)|^q \, d\mu(x) < +\infty \ .$$

Si α est choisi tel que $\dfrac{1}{r} = \alpha\dfrac{1}{p} + (1-\alpha)\dfrac{1}{q}$, c'est-à-dire $\alpha = \frac{p(q-r)}{r(q-p)}$, et si on pose $g = |f|^{\alpha r}$ et $h = |f|^{(1-\alpha)r}$, on a $|f|^r = g.h$, et on déduit de l'inégalité de Hôlder que

$$\begin{aligned}
\|f\|_r^r = \int |f(x)|^r \, d\mu(x) &= \int g(x)h(x)\, d\mu(x) \\
&\leqslant (\|g\|_{p/r\alpha})^{r\alpha/p} \,\|(\|\, h_{q/r(1-\alpha)})^{r(1-\alpha)/q} \\
&= \|f\|_p^{r\alpha} \, \|f\|_q^{(1-\alpha)r} \quad ,
\end{aligned}$$

d'où l'inégalité $\|f\|_r \leqslant \|f\|_p^{\alpha} \|f\|_q^{1-\alpha}$. ■

Corollaire 5.12.7. *Si la mesure μ est finie, on a, pour $1 \leqslant p < q \leqslant \infty$, $L^q \subset L^p$. Et, pour $f \in L^q$, on a $\|f\|_p \leqslant \|f\|_q .\mu(X)^{\frac{1}{p}-\frac{1}{q}}$.*

Démonstration : On a en effet $g = |f|^p \in L^{q/p}$ et :

$$\|f\|_p^p = \int g(x)\mathbb{1}_X(x)\, d\mu(x) \leqslant \|g\|_{q/p} .\Big(\int \mathbb{1}_X(x)\, d\mu(x)\Big)^{1-p/q} = \|f\|_q^p \, \mu(X)^{1-p/q} \ ,$$

d'où $\|f\|_p \leqslant \|f\|_q .\mu(X)^{\frac{1}{p}-\frac{1}{q}}$. ■

Théorème 5.12.8. *Soient $(X, \mathfrak{S}, \mu)$ un espace mesuré de mesure finie, $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction convexe et f une fonction réelle intégrable sur X. Si $\varphi\circ f$ est intégrable, on a*

$$\int_X \varphi\circ f(x)\, d\mu(x) \geqslant \mu(X).\varphi\Big(\frac{\int_X f(x)\, d\mu(x)}{\mu(X)}\Big) \ .$$

Démonstration : Soit $a = \dfrac{1}{\mu(X)} \displaystyle\int_X f(x)\, d\mu(x) \in \mathbb{R}$. Si c désigne la dérivée à droite en a de φ, on a pour tout $t \in \mathbb{R}$: $\varphi(t) \geqslant \varphi(a) + c(t - a)$. On a donc, pour tout x : $\varphi\circ f(x) \geqslant \varphi(a) + c(f(x) - a)$, et en intégrant :

$$\int_X \varphi\circ f(x)\, d\mu(x) \geqslant \varphi(a)\mu(X) + c\Big(\int_X f(x)\, d\mu(x) - a\mu(X)\Big) = \varphi(a)\mu(X) \ ,$$

d'où le résultat. ■

5.13 Inégalités de Clarkson

Lemme 5.13.1. *Soit $p \in]1,+\infty[$. Alors la fonction $z \mapsto |z|^p$ est convexe sur $\mathbb{C}$, et on a : $\liminf_{z\to 0} \dfrac{|1+z|^p+|1-z|^p-2}{|z|^2} \geqslant p-1$. On a de plus $|1+z|^p+|1-z|^p > 2$ pour tout $z \neq 0$.*

Démonstration : La fonction $z \mapsto |z|$ est convexe, de $\mathbb{C}$ dans $\mathbb{R}^+$ et la fonction $t \mapsto t^p$ est convexe croissante sur $\mathbb{R}^+$, d'où la convexité de la fonction $z \mapsto |z|^p$. Pour $z = x+iy$, on a $|1+z|^p = |(1+z)(1+\bar{z})|^{p/2} = (1+2x+|z|^2)^{p/2}$, donc

$$\begin{aligned}|1+z|^p+|1-z|^p &= 2+\frac{p}{2}(2x+|z|^2)+\frac{p(p-2)}{2}x^2+\frac{p}{2}(-2x+|z|^2)+\frac{p(p-2)}{2}x^2+o(z^2)\\ &= 2+p\,|z|^2+p(p-2)x^2+o(z^2) = 2+p(p-1)x^2+py^2+o(z^2)\\ &\geqslant 2+(p-1)x^2+py^2+o(z^2) \geqslant 2+(p-1)\,|z|^2+o(z^2)\ ,\end{aligned}$$

ce dont découle le premier résultat.
La fonction convexe $f : z \mapsto |1+z|^p+|1-z|^p-2$ est paire et nulle en 0, donc partout positive sur $\mathbb{C}$. Par convexité, si elle s'annulait en $z_0 \neq 0$, on aurait $f(tz_0) \leqslant 0$ pour $t \in [0,1]$, et on aurait $0 < p-1 \leqslant \liminf_{t\to 0} \dfrac{|1+tz_0|^p+|1-tz_0|^p-2}{t^2\,|z_0|^2} \leqslant 0$. ∎

Lemme 5.13.2. *Soit $p \geqslant 2$. Il existe une constante $C_p \in]0,2]$ telle que, pour tout $z \in \mathbb{C}$, on ait*

$$|1+z|^p+|1-z|^p \geqslant 2+C_p\,|z|^p\ .$$

Démonstration : La fonction $\varphi_p : z \mapsto \dfrac{|1+z|^p+|1-z|^p-2}{|z|^p}$ est continue sur $\mathbb{C}^*$. Il résulte du lemme 5.13.1 que, pour tout $z \in \mathbb{C}^*$, on a $\varphi_p(z) > 0$ et que $\liminf_{z\to 0}\varphi_p(z) \geqslant p-1$. On a aussi $\lim_{|z|\to\infty}\varphi_p(z) = 2$.
Un argument de compacité montre alors que $C_p = \inf_{z\in\mathbb{C}^*}\varphi_p(z) > 0$. ∎

Lemme 5.13.3. *Soit $p \in]1,2]$. Il existe une constante $C_p > 0$ telle que, pour tout $z \in \mathbb{C}$, on ait*

$$|1+z|^p+|1-z|^p-2 \geqslant C_p\frac{|z|^2}{1+|z|^{2-p}}\ .$$

Démonstration : Comme ci-dessus, $\varphi_p : z \mapsto (1+|z|^{2-p})\dfrac{|1+z|^p+|1-z|^p-2}{|z|^2}$ est une fonction continue et strictement positive sur $\mathbb{C}^*$ qui vérifie $\lim_{|z|\to\infty}\varphi_p(z) = 2$ et $\liminf_{z\to 0}\varphi_p(z) \geqslant p-1 > 0$. On a donc de même $C_p = \inf_z \varphi_p(z) > 0$. ∎

Théorème 5.13.4. *Pour $p \in]1,+\infty[$, l'espace L^p est uniformément convexe, donc réflexif.*

Démonstration : Soient f et h dans L^p. On définit $X' = \{x : f(x) \neq 0\}$ et, pour $x \in X'$, $g(x) = \dfrac{h(x)}{f(x)}$.
Dans le cas $p \geqslant 2$, on a pour tout $x \in X'$:

$$\left|1+\frac{h(x)}{f(x)}\right|^p+\left|1-\frac{h(x)}{f(x)}\right|^p \geqslant 2+C_p\left|\frac{h(x)}{f(x)}\right|^p\ ,$$

donc
$$|f(x)+h(x)|^p + |f(x)-h(x)|^p \geqslant 2\,|f(x)|^p + C_p\,|h(x)|^p \ ,$$
inégalité encore valable si $f(x)=0$. En intégrant on trouve alors
$$\|f+h\|_p^p + \|f-h\|_p^p \geqslant 2\,\|f\|_p^p + C_p\,\|h\|_p^p \ ,$$
ce qui montre la convexité uniforme et la réflexivité de L^p.

Et si $1<p<2$, la fonction $\psi : s \mapsto \dfrac{s^{2/p}}{1+s^{2/p-1}}$ est convexe sur $\mathbb{R}^+$, comme le montre le calcul de la dérivée seconde (avec $\alpha = \dfrac{2}{p}-1 \in]0,1[$) :
$$\psi''(s) = \alpha s^{\alpha-1}\frac{\alpha+1+(1-\alpha)s^\alpha}{(1+s^\alpha)^3} > 0 \ .$$
En appliquant à la mesure $d\nu(x) = |f(x)|^p\ d\mu(x)$ et à la fonction $k=|g|^p$ le théorème 5.12.8, on obtient, puisque $k\in L^1(\nu)$ et que $0\leqslant \psi{\circ}k \leqslant k$:
$$\int |f(x)|^p\,\frac{|g(x)|^2}{1+|g(x)|^{p-2}}\,d\mu(x) = \int \psi{\circ}k(x)\,d\nu(x) \geqslant \int |f(x)|^p\ d\mu(x).\psi(\frac{\int k(x)\,d\nu(x)}{\nu(X)}) \ ,$$
donc :
$$\|f+h\|_p^p + \|f-h\|_p^p - 2\,\|f\|_p^p \geqslant C_p\,\|f\|_p^p\,\psi(\frac{\|h\|_p^p}{\|f\|_p^p}) = C_p\frac{\|h\|_p^2}{\|f\|_p^{2-p}+\|h\|_p^{2-p}} \ ,$$
dont on déduit sans peine, comme plus haut, la convexité uniforme et la réflexivité de L^p. ■

Théorème 5.13.5. *Pour $1\leqslant p<\infty$, la forme bilinéaire $(f,g)\mapsto \langle f,g\rangle = \int_X f(x)g(x)\,d\mu(x)$ identifie le dual de L^p à $L^{p'}$.*

Démonstration : Supposons d'abord $p>1$. Il résulte de l'inégalité de Hölder et du corollaire 5.12.4 que cette forme bilinéaire définit une isométrie j de $L^{p'}$ sur un sous-espace, nécessairement complet donc fermé, du dual de L^p. Si ce sous-espace fermé $j(L^{p'})$ de $(L^p)'$ était distinct de $(L^p)'$, il existerait une forme linéaire continue non nulle $\xi\in(L^p)''$ nulle sur l'image de $L^{p'}$. Mais puisque $(L^p)''=L^p$, il existerait $g\in L^p$ telle que $\langle\xi,\varphi\rangle = \langle\varphi,g\rangle$ pour $\varphi\in(L^p)'$. Pour toute $f\in L^{p'}$, on aurait alors $\langle j(f),g\rangle = \int fg\,d\mu = 0$, donc $g=0$ (prendre $f=\bar g\,|g|^\alpha$ avec $\alpha=\frac{2-p}{p-1}$). Donc $\xi=0$, et cette contradiction achève la démonstration.

Si maintenant $p=1$, et $\varphi\in(L^1)'$, on considère la forme sesquilinéaire définie sur l'espace de Hilbert L^2 par $\rho(f,g)=\langle\varphi,f\bar g\rangle$, qui est bien définie et continue puisque $\|f\bar g\|_1\leqslant\|f\|_2\,\|g\|_2$ et donc que $|\rho(f,g)|\leqslant\|\varphi\|\,\|f\|_2\,\|g\|_2$. Il existe donc (cf. 3.1.21) une application linéaire $T\in\mathscr{L}(L^2)$ telle que $\rho(f,g)=\langle Tf,g\rangle = \int Tf(x)\bar g(x)\,d\mu(x)$. Pour toute partie mesurable A de X on a
$$\int_A Tf(x)\bar g(x)\,d\mu(x) = \rho(f,g\mathbb{1}_A) = \varphi(f\bar g\mathbb{1}_A) = \rho(\bar g,\bar f\mathbb{1}_A) = \int_A T\bar g(x)f(x)\,d\mu(x) \ ,$$

donc $\langle Tf.\bar{g} - T\bar{g}.f, \mathbb{1}_A\rangle = 0$ pour tout A, dont on déduit que la fonction $Tf.\bar{g} - T\bar{g}.f$ est nulle presque partout. Sous l'hypothèse de σ-finitude de μ, on peut trouver une suite (A_n) de parties de mesure finie recouvrant X. Alors la fonction $g_0 = \sum_n 2^{-n}\mu(A_n)^{-1/2}\mathbb{1}_{A_n}$ est dans L^2 et partout non nulle : il en résulte que, en posant $h = \dfrac{Tg_0}{g_0}$, on a pour tout $f \in L^2$: $Tf = h.f$. Et puisque $\|Tf\|_2 \leqslant \|\varphi\| . \|f\|_2$, on a pour tout $f \in L^2$,

$$\int_X (\|\varphi\|^2 - |h(x)|^2).\,|f(x)|^2\; d\mu(x) \geqslant 0\;.$$

En prenant $f = g_0.\mathbb{1}_B$, où $B = \{|h| > \|\varphi\|\}$, on voit que $\mu(B) = 0$, c'est-à-dire que $h \in L^\infty$, et que, pour $f_1 \in L^1$, en posant $g = |f_1|^{1/2} \in L^2$ et $f = f.\,|f_1|^{-1/2} \in L^2$, on a $\langle\varphi, f_1\rangle = \langle Tf, g\rangle = \int_X h(x).f_1(x)\,d\mu(x)$. Et ceci identifie φ à l'élément h de L^∞. ∎

5.14 Le théorème de représentation de Riesz

Théorème 5.14.1. *Soient K un espace compact métrisable et Φ une forme linéaire sur l'espace de Banach $\mathscr{C}(K,\mathbb{R})$. On suppose que Φ est* positive, *c'est-à-dire que $\Phi(f) \geqslant 0$ pour toute fonction continue positive f. Alors il existe une mesure positive μ, finie et σ-additive, sur la tribu borélienne de K telle que, pour toute fonction continue $f : K \to \mathbb{R}$, on ait*

$$\Phi(f) = \int f(x)\,d\mu(x)\;.$$

Démonstration : Pour tout fermé L du compact K et toute fonction continue $\varphi \geqslant \mathbb{1}_L$, on a $\Phi(\varphi) \geqslant \Phi(0) = 0$, et on peut définir

$$m(L) = \inf\{\Phi(\varphi) : \varphi \in \mathscr{C}(K,\mathbb{R}), \varphi \geqslant 1_L\} \in \mathbb{R}^+\;,$$

et puisque $\mathbb{1}_L \leqslant \mathbb{1}$, on a $0 \leqslant m(L) \leqslant \Phi(\mathbb{1}) = m(K)$. On va montrer que cette fonction m satisfait les hypothèses du théorème 5.8.2 pour construire une mesure σ-additive μ sur $\mathfrak{B}(K)$.

Lemme 5.14.2. *Si L est un compact de K, il existe pour tout $\varepsilon > 0$ un voisinage ouvert U de L tel que $m(L') < m(L) + \varepsilon$ si L' est compact contenu dans U.*

Démonstration : Il existe par hypothèse une fonction continue $\varphi \geqslant \mathbb{1}_L$ sur K telle que $\Phi(\varphi) < m(L) + \dfrac{\varepsilon}{2}$. On choisit alors $\theta < 1$ et on définit $U = \{x \in K : \varphi(x) > \theta\}$, qui est un ouvert contenant L. Si L' est un compact contenu dans U et si on pose $\psi = \theta^{-1}.\varphi$, on a $\psi > 1$ sur U, donc $\psi \geqslant \mathbb{1}_{L'}$. De plus $\Phi(\psi) = \theta^{-1}\Phi(\varphi) < \dfrac{m(L)+\varepsilon/2}{\theta}$, donc $\Phi(\psi) < m(L)+\varepsilon$ si $\dfrac{m(L)+\varepsilon/2}{m(L)+\varepsilon} < \theta < 1$. Dans ce cas, on a $m(L') \leqslant \Phi(\psi) < m(L) + \varepsilon$ pour tout compact $L' \subset U$. ∎

Lemme 5.14.3. *Si L_1 et L_2 sont deux compacts de K, on a $m(L_1 \cup L_2) \leqslant m(L_1) + m(L_2)$.*

DÉMONSTRATION : Soit $\varepsilon > 0$. Il existe deux fonctions continues φ_1 et φ_2 telles que $\varphi_1 \geqslant \mathbb{1}_{L_1}$ et $\varphi_2 \geqslant \mathbb{1}_{L_2}$ et que $\Phi(\varphi_1) < m(L_1) + \varepsilon/2$ et $\Phi(\varphi_2) < m(L_2) + \varepsilon/2$. Alors $\varphi_1 + \varphi_2 \geqslant \mathbb{1}_{L_1} + \mathbb{1}_{L_2} \geqslant \mathbb{1}_{L_1 \cup L_2}$. Donc

$$m(L_1 \cup L_2) \leqslant \Phi(\varphi_1 + \varphi_2) = \Phi(\varphi_1) + \Phi(\varphi_2) < m(L_1) + m(L_2) + \varepsilon \ ,$$

d'où l'on déduit l'inégalité cherchée. ■

Lemme 5.14.4. *Si L_1 et L_2 sont deux compacts de K, on a $m(L_1 \cup L_2) \geqslant m(L_1) + m(L_2)$ si $L_1 \cap L_2 = \emptyset$.*

DÉMONSTRATION : Soit $\varepsilon > 0$. Il existe une fonction continue $\varphi \geqslant \mathbb{1}_{L_1 \cup L_2}$ telle que $\Phi(\varphi) < m(L_1 \cup L_2) + \varepsilon$. On considère alors les fonctions continues $h_1 : x \mapsto d(x, L_2)$ et $h_2 : x \mapsto d(x, L_1)$. Puisque $L_1 \cap L_2 = \emptyset$, on a $h_1 + h_2 > 0$ sur K. Posant alors $\psi_1 = \dfrac{h_1}{h_1 + h_2}$ et $\psi_2 = \dfrac{h_2}{h_1 + h_2}$, on a $\psi_1 + \psi_2 = 1$, $\psi_1 = 1$ sur L_1 et $\psi_2 = 1$ sur L_2. On en déduit que $\psi_1 \varphi \geqslant \mathbb{1}_{L_1}$ et $\psi_2 \varphi \geqslant \mathbb{1}_{L_2}$. Donc

$$m(L_1) + m(L_2) \leqslant \Phi(\psi_1 \varphi) + \Phi(\psi_2 \varphi) = \Phi((\psi_1 + \psi_2)\varphi) = \Psi(\varphi) < m(L_1 \cup L_2) + \varepsilon \ .$$

Et ceci montre l'inégalité recherchée, puisque ε est arbitraire. ■

Il résulte alors du théorème 5.8.2 qu'existe une mesure σ-additive μ sur $(K, \mathfrak{B}(K))$ telle que $\mu(L) = m(L)$ pour tout compact L de K. En particulier, la mesure μ est finie puisque, pour tout $A \in \mathfrak{B}(K)$, on a $\mu(A) \leqslant \mu(K) = m(K) = \Phi(\mathbb{1})$. Il reste à montrer que $\Phi(f) = \int f(x)\, d\mu(x)$ pour toute fonction continue f sur K.

Supposons d'abord que f soit à valeurs dans $\mathbb{R}^+$ et posons $M = \sup_x f(x)$. Pour n et p dans $\mathbb{N}$, on considère les ensembles compacts $L_{n,p} = f^{-1}([p.2^{-n}, \infty[)$. Soit $x \in K$: si $q.2^{-n} \leqslant f(x) < (q+1).2^{-n}$, on a $\mathbb{1}_{L_{n,p}}(x) = 1$ si et seulement si $p \leqslant q$, et en particulier $L_{n,p} = \emptyset$ si $p > M.2^n$. Il en résulte que $\sum_{p=1}^{\infty} 2^{-n} \mathbb{1}_{L_{n,p}}(x) = q.2^{-n}$.

On a donc $f - 2^{-n} \leqslant \sum_{p=1}^{\infty} 2^{-n} \mathbb{1}_{L_{n,p}} \leqslant f$. Alors, pour chaque p, il existe une fonction continue φ_p, nulle si $L_{n,p} = \emptyset$, telle que $\mathbb{1}_{L_{n,p}} \leqslant \varphi_p$ et $\Phi(\varphi_p) \leqslant m(L_{n,p}) + 2^{-p} = \mu(L_{n,p}) + 2^{-p}$, donc

$$\int f(x)\, d\mu(x) \geqslant \sum_{p=1}^{\infty} 2^{-n} \mu(L_{n,p}) \geqslant \sum_{p=1}^{\infty} (2^{-n} \Phi(\varphi_p) - 2^{-n-p}) \geqslant \Phi(\sum_{1}^{\infty} 2^{-n} \varphi_p) - 2^{-n} \ .$$

Et puisque $\sum_1^{\infty} 2^{-n} \varphi_p \geqslant \sum_1^{\infty} 2^{-n} \mathbb{1}_{L_{n,p}} \geqslant f - 2^{-n}$, on obtient

$$\int f(x)\, d\mu(x) \geqslant \Phi(f) - 2^{-n} \Phi(\mathbb{1}_K) - 2^{-n} = \Phi(f) - 2^{-n}(1 + m(K))$$

et en passant à la limite lorsque $n \to \infty$: $\Phi(f) \leqslant \int f(x)\, d\mu(x)$. La fonction $M.\mathbb{1} - f$ vérifie de même :

$$\begin{aligned} \Phi(M.\mathbb{1} - f) &= M.\Phi(\mathbb{1}) - \Phi(f) = M.m(K) - \Phi(f) \\ &\leqslant \int (M - f(x))\, d\mu(x) = M.m(K) - \int f(x)\, d\mu(x) \ , \end{aligned}$$

donc $\Phi(f) \geqslant \int f(x)\,d\mu(x)$, et on conclut que $\Phi(f) = \int f(x)\,d\mu(x)$.
Et puisque toute fonction continue f est différence de deux fonctions continues positives $g = f_+$ et $h = f_-$, on voit que

$$\Phi(f) = \Phi(g) - \Phi(h) = \int g(x)\,d\mu(x) - \int h(x)\,d\mu(x) = \int (g-h)(x)\,d\mu(x) = \int f(x)\,d\mu(x) \ ,$$

et ceci termine la démonstration du théorème 5.14.1. ■

On cherche maintenant à montrer une représentation analogue des formes linéaires continues sur $\mathscr{C}(K)$.

Lemme 5.14.5. *Soient f_1, f_2 et g trois fonctions continues réelles positives sur l'espace topologique X telles que $g \leqslant f_1 + f_2$. Alors il existe g_1 et g_2 continues positives telles que $g_1 \leqslant f_1$, $g_2 \leqslant f_2$ et $g = g_1 + g_2$.*

Démonstration : La fonction g_1 doit vérifier $0 \leqslant g_1 \leqslant f_1$ et $g \geqslant g_1 = g - g_2 \geqslant g - f_2$. Si on pose $g_1 = \inf(f_1, g)$, on a bien $g_1 \geqslant 0$, $g_1 \leqslant f_1$ et $g_1 \leqslant g$. Il reste à voir que $g_1 \geqslant g - f_2$, c'est-à-dire $g_1 + f_2 \geqslant g$. Or $g_1 + f_2 = \inf(f_1 + f_2, g + f_2) \geqslant \inf(g, g + f_2) = g$. ■

Lemme 5.14.6. *Soient K un espace métrique compact et Φ une forme linéaire continue sur l'espace $\mathscr{C}(K, \mathbb{C})$. Alors il existe une forme linéaire positive Ψ sur $\mathscr{C}(K, \mathbb{R})$ (c'est-à-dire $\Psi(f) \geqslant 0$ pour toute f positive) telle que $|\Phi(f)| \leqslant \Psi(|f|)$.*

Démonstration : Posons, pour f continue positive sur K :

$$\Psi(f) = \sup\{\Re e(\Phi(g)) : g \in \mathscr{C}(K, \mathbb{C}), |g| \leqslant f\} \ .$$

On a clairement $\Psi(f) \geqslant 0$ et $\Psi(\lambda f) = \lambda \Psi(f)$ pour $\lambda > 0$. Si f_1 et f_2 sont continues positives et $\varepsilon > 0$, il existe g_1 et g_2 continues avec $|g_1| \leqslant f_1$, $|g_2| \leqslant f_2$, $\Re e(\Phi(g_1)) \geqslant \Psi(f_1) - \varepsilon/2$ et $\Re e(\Phi(g_2)) \geqslant \Psi(f_2) - \varepsilon/2$.
On a alors $|g_1 + g_2| \leqslant |g_1| + |g_2| \leqslant f_1 + f_2$, donc

$$\Psi(f_1) + \Psi(f_2) - \varepsilon < \Re e(\Phi(g_1)) + \Re e(\Phi(g_2)) = \Re e(\Phi(g_1 + g_2)) \leqslant \Psi(f_1 + f_2) \ .$$

Et on en déduit que $\Psi(f_1) + \Psi(f_2) \leqslant \Psi(f_1 + f_2)$.
Inversement, pour $\varepsilon > 0$, il existe une fonction continue g telle que $|g| \leqslant f_1 + f_2$ et $\Re e(\Phi(g)) \geqslant \Psi(f_1 + f_2) - \varepsilon$. Il résulte du lemme 5.14.5 qu'existent alors deux fonctions positives g'_1 et g'_2 telles que $g'_1 + g'_2 = |g|$, $g'_1 \leqslant f_1$ et $g'_2 \leqslant f_2$. La fonction $g_1 = g'_1 . \dfrac{g}{|g|}$ est continue en tout point où $g \neq 0$, et tend vers 0 en tout point où g s'annule puisque $|g_1| = g'_1 \leqslant |g|$. De même $g_2 = g'_2 . \dfrac{g}{|g|}$ est continue. On a donc $g = g_1 + g_2$, $|g_1| \leqslant f_1$ et $|g_2| \leqslant f_2$. On en déduit que

$$\Psi(f_1 + f_2) - \varepsilon < \Re e(\Phi(g)) = \Re e(\Phi(g_1)) + \Re e(\Phi(g_2)) \leqslant \Psi(f_1) + \Psi(f_2) \ ,$$

donc que $\Psi(f_1) + \Psi(f_2) \geqslant \Psi(f_1 + f_2)$ puisque ε est arbitraire.
Si f_1, f_2, f'_1 et f'_2 sont des fonctions continues réelles positives telles que $f_1 - f_2 = f'_1 - f'_2$, on a $\Psi(f_1) + \Psi(f'_2) = \Psi(f_2) + \Psi(f'_1)$, donc $\Psi(f_1) - \Psi(f_2) = \Psi(f'_1) - \Psi(f'_2)$.
Il en résulte que Ψ se prolonge en une forme linéaire sur $\mathscr{C}(K, \mathbb{R})$.
Enfin si f est continue, $|f|$ est continue, et on a, pour tout $\theta \in \mathbb{R}$, $\Psi(|f|) \geqslant \Re e(\Phi(f)\,\mathrm{e}^{-i\theta})$, puisque $\left|f\,\mathrm{e}^{-i\theta}\right| \leqslant |f|$. On a donc $|\Phi(f)| \leqslant \Psi(|f|)$. ■

Théorème 5.14.7. *Soient K un espace compact métrisable et Φ une forme linéaire continue sur l'espace de Banach $\mathscr{C}(K, \mathbb{C})$. Alors il existe une mesure positive μ, finie et σ-additive, sur la tribu borélienne de K et une fonction mesurable bornée ω sur K telles que, pour toute fonction continue $f : K \to \mathbb{C}$, on ait*

$$\Phi(f) = \int f(x)\omega(x)\, d\mu(x) \ .$$

Démonstration : Il résulte du théorème 5.14.6 qu'il existe une forme linéaire positive Ψ sur $\mathscr{C}(K, \mathbb{R})$ telle que $|\Phi(f)| \leqslant \Psi(|f|)$. Il résulte ensuite du théorème 5.14.1 qu'il existe une mesure σ-additive finie μ sur $(K, \mathfrak{B}(K))$ telle que $\Psi(f) = \int f(x)\, d\mu(x)$ pour toute fonction réelle continue sur K. En particulier, pour f complexe continue, la norme de f dans l'espace $L^1(\mu)$ est $\|f\|_1 = \displaystyle\int |f(x)|\, d\mu(x) = \Psi(|f|)$. On en déduit que $|\Phi(f)| \leqslant \|f\|_1$. Alors, par le théorème de Hahn-Banach, la forme linéaire Φ se prolonge en une forme linéaire continue $\tilde{\Phi}$ de norme au plus 1 sur $L^1(K, \mathfrak{B}(K), \mu)$. Il existe donc (cf. 5.13.5) un élément ω dans $L^\infty(K, \mathfrak{B}, \mu)$ tel que, pour $g \in L^1(\mu)$, on ait $\langle \tilde{\Phi}, g\rangle = \int \omega(x) g(x)\, d\mu(x)$. Et finalement, pour $f \in \mathscr{C}(K, \mathbb{C})$ on a

$$\Phi(f) = \tilde{\Phi}(f) = \int f(x).\omega(x)\, d\mu(x) \ . \qquad \blacksquare$$

Inversement, si μ est une mesure finie sur $(K, \mathfrak{B}(K))$ et $\omega \in L^\infty(K, \mathfrak{B}(K), \mu)$, toute fonction continue f sur K est bornée, et $f.\omega$ est dans $L^\infty(\mu) \subset L^1(\mu)$. L'application linéaire $\Phi : f \mapsto \int f(x).\omega(x)\, d\mu(x)$ est donc bien définie sur $\mathscr{C}(K)$, et continue pour la norme de la convergence uniforme puisque $|\Phi(f)| \leqslant \|f\| . \|\omega\|_\infty . \mu(K)$.

6

CONVOLUTION ET TRANSFORMATION DE FOURIER

6.1 Densité des fonctions continues à support compact

Rappelons la définition du support d'une fonction numérique.

Définition 6.1.1. *Si f est une fonction numérique continue sur un espace topologique X, on appelle* support *de f le plus petit fermé de X en dehors duquel f est identiquement nulle.*

Lorsque X est un espace localement compact, il est clair que l'ensemble des fonctions numériques continues à support compact est un sous-espace vectoriel de l'espace $\mathscr{C}_0(X)$ des fonctions continues qui tendent vers 0 à l'infini.

Dans ce chapitre, X désigne (un ouvert de) $\mathbb{R}^d$, $\mathfrak{S}$ la tribu borélienne de X et μ la (restriction de la) mesure de Lebesgue. Alors, pour tout $p \geqslant 1$, l'espace $\mathscr{C}_c(X)$ des fonctions complexes continues à support compact sur X est contenu dans $L^p(X)$.

Théorème 6.1.2. *L'espace $\mathscr{C}_c(X)$ est dense dans $\mathscr{C}_0(X)$ pour la norme uniforme.*

Démonstration : Soient $f \in \mathscr{C}_0(X)$ et $\varepsilon > 0$. L'ensemble $K = \{x \in X : |f(x)| \geqslant \varepsilon\}$ est compact dans X et possède un voisinage ouvert W relativement compact. Alors la fonction $g : x \mapsto \dfrac{d(x, W^c)}{d(x, K) + d(x, W^c)}$ est continue, à valeurs dans $[0, 1]$, égale à 1 sur K et à support dans le compact $\overline{W}$. Alors $f.g$ appartient à $\mathscr{C}_c(X)$, est nulle sur K, et vérifie, pour $x \notin K$:

$$|f(x) - f.g(x)| = |f(x)| .(1 - g(x)) \leqslant |f(x)| < \varepsilon \ ,$$

donc $\|f - f.g\|_\infty \leqslant \varepsilon$. ■

Lemme 6.1.3. *Soit $A \in \mathfrak{S}$. Pour tout compact $T \subset X$ et tout $\varepsilon > 0$, il existe un compact K et un ouvert U de X tels que $K \cap T \subset A \cap T \subset U \cap T$ et $\mu(U \setminus K) < \varepsilon$.*

Démonstration : Compte tenu de la définition de la mesure de Lebesgue en 5.9.1, ceci est exactement l'énoncé du lemme 5.8.7. ■

Lemme 6.1.4. *Soient $p \in [1, \infty]$ et $f \in L^p(X)$. Si, pour tout compact $K \subset X$ on a $\int_K f(x)\, d\mu(x) = 0$, alors $f = 0$.*

Démonstration : Soient (T_m) une suite exhaustive de compacts de X et $A \in \mathfrak{S}$. Il existe pour chaque m, d'après le lemme précédent, une suite (K_n) de compacts de X telle que

$K_n \subset T_m \cap A$ et $\mu(T_m \cap (A \setminus K_n)) < 2^{-n}$. La suite $(f.\mathbb{1}_{K_n \cap T_m})_n$ converge donc presque partout vers $f.\mathbb{1}_{A \cap T_m}$, en restant majorée en module par la fonction intégrable $|f|.\mathbb{1}_{T_m}$. On en déduit que $\int_{A \cap T_m} f(x)\,d\mu(x) = \lim_{n\to\infty} \int_{K_n \cap T_m} f(x)\,d\mu(x) = 0$.
En appliquant ceci pour tout entier q à $A = \{\Re e(f) \geqslant 2^{-q}\}$, $A = \{\Re e(f) \leqslant -2^{-q}\}$, $A = \{\Im m(f) \geqslant 2^{-q}\}$ et $A = \{\Im m(f) \leqslant -2^{-q}\}$, on voit que $\Re e(f)$ et $\Im m(f)$ sont nulles presque partout sur T_m, donc que $f = 0$ presque partout sur X. ■

Théorème 6.1.5. *Soit $p \in [1, \infty[$. Alors $\mathscr{C}_c(X)$ est dense dans $L^p(X)$.*

Démonstration : Si ce n'était pas le cas, il existerait, en vertu du théorème de Hahn-Banach, une fonction f non nulle dans $L^{p'}(X)$ telle que $\int_X f(x)g(x)\,d\mu(x) = 0$, pour toute fonction continue g à support compact.
Soient alors K un compact de X et W un voisinage ouvert de K relativement compact dans X. Alors $\mathbb{1}_W \in L^p(X)$ et, si on pose $W_n = W \cap \{x : d(x,K) < 2^{-n}\}$, la fonction continue $g_n : x \mapsto \dfrac{d(x, W_n^c)}{d(x, W_n^c) + d(x,K)}$ est nulle hors de W et vaut 1 sur K. La suite (g_n) converge partout vers $\mathbb{1}_K$ en restant majorée par $\mathbb{1}_W$. On a donc, par convergence dominée, $\int_X f(x).\mathbb{1}_K(x)\,d\mu(x) = \lim_{n\to\infty} \int_X f(x).g_n(x)\,d\mu(x) = 0$. Il résulte alors du lemme précédent que $f = 0$. Et cette contradiction achève la preuve. ■

Théorème 6.1.6. *Soit $p \in [1, \infty[$. Alors l'espace $\mathscr{D}(X)$ des fonctions indéfiniment différentiables à support compact dans X est dense dans $L^p(X)$. Et $\mathscr{D}(X)$ est dense dans $\mathscr{C}_0(X)$.*

Démonstration : La même démonstration que pour les théorèmes 6.1.5 et 6.1.2 fournit le résultat, si on sait trouver, pour tout compact K et tout voisinage ouvert W de K une fonction $\mathscr{C}^\infty$, g, à valeurs dans $[0,1]$ valant 1 sur K et nulle hors de W. Ceci résulte immédiatement des lemmes suivants, qui seront aussi utiles ultérieurement pour l'étude des distributions. ■

Lemme 6.1.7. *Soient K un compact de $\mathbb{R}^d$ et W un voisinage ouvert de K. Il existe une fonction $\mathscr{C}^\infty$, g, à valeurs positives, nulle hors de W et strictement positive sur K.*

Démonstration : Un calcul facile montre que la fonction ρ définie sur $\mathbb{R}$ par $\rho(t) = 0$ si $t \leqslant 0$ et $\rho(t) = \mathrm{e}^{-1/t}$ si $t > 0$ est de classe $\mathscr{C}^\infty$. Il en résulte que la fonction $\varphi_{a,r}$ définie sur $\mathbb{R}^d$ par $\varphi_{a,r}(x) = \rho(r^2 - \|x-a\|^2)$ est $\mathscr{C}^\infty$, strictement positive sur $B(a,r)$ et nulle en dehors. Pour tout $a \in K$, il existe une boule ouverte $B(a,r)$ contenue dans W, et K est recouvert par un nombre fini $(B(a_j, r_j)_{1\leqslant j \leqslant m})$ d'entre elles. Alors $g = \sum_{j=1}^m \varphi_{a_j, r_j}$ satisfait l'énoncé du lemme. ■

Lemme 6.1.8. *Soient K un compact de $\mathbb{R}^d$ et W un voisinage ouvert de K. Il existe une fonction $\mathscr{C}^\infty$ à valeurs dans $[0,1]$ nulle hors de W et valant 1 sur K.*

Démonstration : La fonction $x \mapsto d(x, W^c)$ est continue et strictement positive sur le compact K, donc y atteint son minimum $r > 0$. Si on définit $V_1 = \{x : d(x,K) < \dfrac{r}{3}\}$ et $V_2 = \{x : d(x,K) < \dfrac{2r}{3}\}$, V_1 et V_2 sont des voisinages ouverts bornés de K, donc relativement compacts. Et on a $\overline{V_1} \subset V_2 \subset \overline{V_2} \subset W$. Il existe donc d'après le lemme précédent une fonction $\mathscr{C}^\infty$ positive g strictement positive sur $\overline{V_1}$ et nulle hors de V_2, ainsi qu'une fonction $\mathscr{C}^\infty$ positive h strictement positive sur $\overline{V_2} \setminus V_1$ et nulle sur K. Alors $g + h$ est strictement positive au voisinage de $\overline{V_2}$, et la fonction $\dfrac{g}{g+h}$ satisfait les exigences de l'énoncé. ■

6.2 Convolution

Théorème 6.2.1. *Soient f et g dans $L^1(\mathbb{R}^d)$. Alors pour presque tout $x \in \mathbb{R}^d$ la fonction $y \mapsto f(y)g(x-y)$ est intégrable. La fonction $f*g$ définie alors presque partout par $f*g(x) = \int f(y)g(x-y)\,d\mu(y)$ est dans L^1 et on a $\|f*g\|_1 \leqslant \|f\|_1\,\|g\|_1$. De plus $f*g = g*f$. On appelle* convolée de f et de g *la fonction $f*g$.*

Démonstration : Par le théorème de Fubini (5.7.8), on a

$$\begin{aligned}\int\left(\int |f(y)g(x-y)|\;d\mu(y)\right)\,d\mu(x) &= \iint |f(y)|\,|g(x-y)|\;d\mu(x)\,d\mu(y)\\ &= \int |f(y)|\left(\int |g(x-y)|\;d\mu(x)\right)\,d\mu(y)\\ &= \int |f(y)|\,\|g\|_1\;d\mu(y) = \|f\|_1\,\|g\|_1 < +\infty\ ,\end{aligned}$$

ce qui montre que, pour presque tout x, la fonction $|f|*|g|$ est définie en x. Si $|f|*|g|$ est définie en x, alors la fonction $y \mapsto f(y)g(x-y)$ est dans L^1, et on a $|f*g(x)| \leqslant |f|*|g|\,(x)$, ce qui montre que $f*g$ est définie en x, que $f*g \in L^1$ et que $\|f*g\|_1 \leqslant \|f\|_1\,\|g\|_1$.
Si $|f|*|g|$ et $|g|*|f|$ sont définis en x, le changement de variable $y = x-z$ dans l'intégrale montre que

$$f*g(x) = \int f(y)g(x-y)\,d\mu(y) = \int f(x-z)g(z)\,d\mu(z) = g*f(x)$$

donc que $f*g = g*f$ dans L^1. ■

D'une façon générale, si f et g sont deux fonctions mesurables, on notera $f*g$ la fonction définie par $f*g(x) = \int f(y)g(x-y)\,d\mu(x)$ en tout point où la fonction $y \mapsto f(y).g(x-y)$ est intégrable.

Proposition 6.2.2. *Si f et g sont dans $\mathscr{C}_c(\mathbb{R}^d)$, alors $f*g \in \mathscr{C}_c(\mathbb{R}^d)$. De plus le support de $f*g$ est contenu dans* $\mathrm{supp}(f)+\mathrm{supp}(g)$.

Démonstration : La fonction g est uniformément continue. Pour $\varepsilon > 0$ il existe donc un nombre $\delta > 0$ tel que $\|z-z'\| < \delta \Longrightarrow |g(z)-g(z')| < \varepsilon$. Pour $\|x-x'\| < \delta$ on a donc

$$|f*g(x) - f*g(x'))| \leqslant \int |f(y)|\,.\,|g(x-y)-g(x'-y)|\;d\mu(y) \leqslant \varepsilon\,\|f\|_1$$

ce qui montre la continuité (uniforme) de g. L'application $(x,y) \mapsto x+y$ est continue de $\mathbb{R}^d\times\mathbb{R}^d$ dans $\mathbb{R}^d$. Il en résulte que $\mathrm{supp}(f)+\mathrm{supp}(g)$, image du compact $\mathrm{supp}(f)\times\mathrm{supp}(g)$, est compact. Alors, pour tout $x \notin \mathrm{supp}(f)+\mathrm{supp}(g)$, on a pour tout y : $y \notin \mathrm{supp}(f)$ ou $x-y \notin \mathrm{supp}(g)$, donc $f(y)g(x-y) = 0$. Il en résulte que $f*g(x) = 0$, et que $\mathrm{supp}(f*g) \subset \mathrm{supp}(f)+\mathrm{supp}(g)$. ■

Théorème 6.2.3. *Soient $f \in L^1(\mathbb{R}^d)$ et $g \in \mathscr{C}_c(\mathbb{R}^d)$. Alors la fonction $f*g$ est dans $\mathscr{C}_0(\mathbb{R}^d)$ et $\sup_x |f*g(x)| \leqslant \|f\|_1\,.\,\|g\|_\infty = \|f\|_1\,.\sup_x |g(x)|$.*

Démonstration : On a $|f*g(x)| \leqslant \int |f(y)|\,\|g\|_\infty\;d\mu(y) = \|f\|_1\,.\,\|g\|_\infty$.

Soit $\varepsilon > 0$. Par continuité uniforme de g, il existe un nombre δ strictement positif tel que $\|z - z'\| < \delta \Longrightarrow |g(z) - g(z')| < \varepsilon$. Alors, si $\|x - x'\| < \delta$, on a

$$|f * g(x) - f * g(x')| \leqslant \int |f(y)| \,.\, |g(x-y) - g(x'-y)| \, d\mu(y) \leqslant \varepsilon \, \|f\|_1 \ ,$$

ce qui montre la continuité uniforme de $f * g$. De plus, si (f_n) est une suite dans $\mathscr{C}_c(\mathbb{R}^d)$ qui converge vers f dans L^1, les $(f_n * g)$ sont continues à support compact et convergent uniformément vers $f * g$. Donc $f * g \in \mathscr{C}_0(\mathbb{R}^d)$. ∎

Théorème 6.2.4. *Soient $p \in]1, \infty[$ et p' l'exposant conjugué. Alors, si $f \in L^p$ et $g \in L^{p'}$, la fonction $f * g : x \mapsto \int f(y) g(x-y) \, d\mu(y)$ est définie en tout point de $\mathbb{R}^d$. De plus, la fonction $f * g$ est continue sur $\mathbb{R}^d$, nulle à l'infini et on a $\|f * g\|_\infty \leqslant \|f\|_p \|g\|_{p'}$.*

Démonstration : Pour $x \in \mathbb{R}^d$, notons g_x la fonction $y \mapsto g(x-y)$.
On a $\int |g_x(y)|^{p'} \, d\mu(y) = \int |g(z)|^{p'} \, d\mu(z) = \|g\|_{p'}^{p'} < \infty$, d'où $g_x \in L^{p'}$. On a donc $|f * g(x)| = |\langle f, g_x \rangle| \leqslant \|f\|_p \|g_x\|_{p'} = \|f\|_p \|g\|_{p'}$. Et l'application bilinéaire $(f, g) \mapsto f * g$ est continue de $L^p \times L^{p'}$ dans L^∞.
Il existe deux suites (f_n) et (g_n) dans $\mathscr{C}_c(\mathbb{R}^d)$ qui convergent respectivement vers f dans L^p et vers g dans $L^{p'}$. Alors la suite $(f_n * g_n)$ est une suite de fonctions continues à supports compacts qui converge uniformément vers $f * g$. On en déduit que $f * g \in \mathscr{C}_0(\mathbb{R}^d)$. ∎

Théorème 6.2.5. *Si $f \in L^1$ et $g \in L^\infty$, alors la fonction $f * g$ est continue bornée, et on a $\|f * g\|_\infty \leqslant \|f\|_1 \|g\|_\infty$.*

Démonstration : On a pour tout $x \in \mathbb{R}^d$: $f * g(x) = \langle f, g_x \rangle$ en posant $g_x(y) = g(x-y)$, donc $|f * g(x)| \leqslant \|f\|_1 \|g_x\|_\infty = \|f\|_1 \|g\|_\infty$. Si K est un compact de $\mathbb{R}^d$ et si $\alpha > 0$ est donné, on peut trouver $h \in \mathscr{C}_c(\mathbb{R}^d)$ telle que $\|h - f\|_1 \|g\|_\infty < \alpha$. Alors, en notant $R = \sup_{h(x) \neq 0} \|x\|$, l'ensemble $K + \tilde{B}(0, R)$ est un compact K' et $g' := g \mathbb{1}_{K'} \in L^1$. De plus, pour $x \in K$ on a $h * g(x) = h * g'(x)$ puisque pour tout y tel que $h(y) \neq 0$ on a $x - y \in K'$, donc $g(x-y) = g'(x-y)$.
Puisque $h \in \mathscr{C}_c(\mathbb{R}^d)$ et $g' \in L^1$, on a $h * g' \in \mathscr{C}_0(\mathbb{R}^d)$. Donc $h * g$ est continue sur K. Et puisque $\|h * g - f * g\|_\infty \leqslant \|f - h\|_1 \|g\|_\infty < \alpha$, on voit que $f * g$ est limite uniforme sur K de fonctions continues, donc elle-même continue sur K. Et puisque ceci est vrai pour tout compact K, la fonction $f * g$ est continue sur $\mathbb{R}^d$. ∎

Théorème 6.2.6. *Si p, q et r sont supérieurs à 1 et vérifient $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$, si $f \in L^P$ et $g \in L^q$ alors la fonction $y \mapsto f(y) g(x-y)$ est dans L^1 pour presque tout x, et la fonction $f * g : x \mapsto \int f(y) g(x-y) \, d\mu(y)$ appartient à L^r. On a alors $\|f * g\|_r \leqslant \|f\|_p \|g\|_q$, c'est-à-dire que la convolution définit une application bilinéaire continue de $L^p \times L^q$ dans L^r.*

Démonstration : Si $r = \infty$, le cas est déjà traité ci-dessus, de même que si $r = 1$, car alors $p = q = 1$.

Notons p', q' et r' les exposants conjugués de p, q et r. On a $\dfrac{1}{p} + \dfrac{1}{q} + \dfrac{1}{r'} = 2$. Quitte à diviser f par $\|f\|_p$ et g par $\|g\|_q$, on peut supposer $\|f\|_p = \|g\|_q = 1$. Si $h \in L^{r'}$, avec $\|h\|_{r'} = 1$, on va montrer que la fonction $(x, y) \mapsto |f(y) g(x-y) h(x)|$ est dans $L^1(\mu \otimes \mu)$.
Il résulte du Lemme 5.12.2 que, pour tout (x, y), on a

$$|f(y) g(x-y) h(x)| \leqslant \frac{1}{p'} |g(x-y)|^q |h(x)|^{r'} + \frac{1}{q'} |h(x)|^{r'} |f(y)|^p + \frac{1}{r} |f(y)|^p |g(x-y)|^q$$

On déduit du théorème de Fubini que

$$\begin{aligned}\iint |g(x-y)|^q \, |h(x)|^{r'} \, d\mu(x)\, d\mu(y) &= \iint |h(x)|^{r'} \, |f(y)|^p \, d\mu(x)\, d\mu(y) \\ &= \iint |f(y)|^p \, |g(x-y)|^q \, d\mu(x)\, d\mu(y) = 1\end{aligned}$$

donc que $\displaystyle\iint |f(y)g(x-y)h(x)| \, d\mu(x)\, d\mu(y) \leqslant \frac{1}{p'} + \frac{1}{q'} + \frac{1}{r} = 1$. Une nouvelle application du théorème de Fubini montre que l'intégrale $\int |f(y)g(x-y))| \, d\mu(y)$ est finie en presque tout point de $\mathbb{R}^d$ où $h \neq 0$. Il en résulte que la fonction $f * g$ est presque partout définie par cette intégrale, et que $|\langle f * g, h\rangle| \leqslant 1$ dès que $\|h\|_{r'} = 1$. Donc $f * g \in L^r$ et $\|f * g\|_r \leqslant 1$. ■

6.3 Régularisation

Lemme 6.3.1. *Soient $p \in [1, \infty[$, $f \in L^p(\mathbb{R}^d)$ et g une fonction de classe $\mathscr{C}^k$ à support compact sur $\mathbb{R}^d$. Alors $f * g$ est de classe $\mathscr{C}^k$ sur $\mathbb{R}^d$ et appartient à $L^p(\mathbb{R}^d)$, avec $\|f * g\|_p \leqslant \|f\|_p \, \|g\|_1$.*

Démonstration : Puisque $g \in L^1$, on a $f * g \in L^p$ et $\|f * g\|_p \leqslant \|f\|_p \, \|g\|_1$ d'après le théorème 6.2.6.
Si $k = 0$, la continuité de $f * g$ découle alors de 6.2.4 puisque $g \in L^{p'}(\mathbb{R}^d)$. Si l'énoncé est vrai pour k, il suffit, pour le montrer pour $k+1$, de vérifier que les dérivées partielles de $f * g$ existent et sont de classe $\mathscr{C}^k$. Mais une application du théorème de différentiation sous le signe somme (théorème 5.6.3) montre que

$$\frac{\partial}{\partial x_i}(f * g) = \frac{\partial}{\partial x_i} \int f(y) g(x-y) \, d\mu(y) = \int f(y) \frac{\partial g}{\partial x_i}(x-y) \, d\mu(y) = f * \frac{\partial g}{\partial x_i}$$

Et l'hypothèse de récurrence appliquée aux $\dfrac{\partial g}{\partial x_i}$, qui sont $\mathscr{C}^k$ si g est $\mathscr{C}^{k+1}$, montre que $f * g$ est $\mathscr{C}^{k+1}$ si g l'est. ■

Théorème 6.3.2. *Soit $g \in L^1(\mathbb{R}^d)$, telle que $\int g(x) \, d\mu(x) = 1$. Pour $\varepsilon > 0$, on définit $g_\varepsilon(x) = \varepsilon^{-d} g(\frac{x}{\varepsilon})$. Si $p \in [1, \infty[$ et $f \in L^p(\mathbb{R}^d)$, on a $\lim_{\varepsilon \to 0} \|f - f * g_\varepsilon\|_p = 0$.*

Si $f \in \mathscr{C}_0(\mathbb{R}^d)$, on a $\lim_{\varepsilon \to 0} \|f - f * g_\varepsilon\|_\infty = 0$.

Démonstration : On remarque pour commencer que, par le changement de variable $y = \dfrac{x}{\varepsilon}$, on a $\|g_\varepsilon\|_1 = \|g\|_1$ et $\int g_\varepsilon(x) \, d\mu(x) = 1$.
Supposons d'abord $f \in \mathscr{C}_c(\mathbb{R}^d)$. Pour $0 < \alpha < 1$, il existe $h \in \mathscr{C}_c(\mathbb{R}^d)$ telle que $\|g - h\|_1 < \alpha$. Si on pose $\beta = \int h(x) \, d\mu(x)$, on a donc $|\beta - 1| < \alpha$. Alors, pour $h' = \beta^{-1}.h$, on obtient $\int h'(x) \, d\mu(x) = 1$ et $\|g - h'\|_1 < \alpha + (1 - \beta^{-1}) \|h\|_1 < \alpha + \dfrac{\alpha}{1-\alpha}(\|g\|_1 + \alpha) = (1 + \|g\|_1)\dfrac{\alpha}{1-\alpha}$.
Et quitte à remplacer h par h' et α par $\alpha' = (1 + \|g\|_1)\dfrac{\alpha}{1-\alpha}$, on se ramène au cas où $\int h(x) \, d\mu(x) = 1$.

On pose $R = \sup\{\|x\| : x \in \operatorname{supp}(h)\}$. La fonction $f * h_\varepsilon$ est dans $\mathscr{C}_c(\mathbb{R}^d)$. Par continuité uniforme de f, il existe un $\eta > 0$ tel que $\|y - y'\| < \eta R \Longrightarrow |f(y) - f(y')| < \alpha$. En particulier, pour $\varepsilon \leqslant \eta$, on a $h_\varepsilon(x-y) \neq 0 \Longrightarrow \|x - y\| \leqslant \eta R \Longrightarrow |f(x) - f(y)| < \alpha$. Alors, si $\varepsilon \leqslant \eta$, on a

$$f(x) - f * h_\varepsilon(x) = \int (f(x) - f(y)).h_\varepsilon(x-y)\, d\mu(y)$$

donc

$$|f(x) - f * h_\varepsilon(x)| \leqslant \sup_{h_\varepsilon(x-y)\neq 0} |f(x) - f(y)| \,. \int |h_\varepsilon(x-y)|\, d\mu(y) \leqslant \alpha \|h_\varepsilon\|_1 = \alpha \|h\|_1$$

ce qui montre la convergence uniforme de $f * h_\varepsilon$ vers f. Pour $\varepsilon \leqslant 1$, le support de $f * h_\varepsilon$ est contenu dans le compact fixe $K = \{x : d(x, \operatorname{supp}(f)) \leqslant R\}$. On en déduit que $\|f - f * h_\varepsilon\|_p \to 0$ lorsque ε tend vers 0. Fixons maintenant $\delta > 0$. Si on choisit α tel que $\alpha \|f\|_p < \delta$ puis η tel que $\|f - f * h_\varepsilon\|_p < \delta$ quand $\varepsilon \leqslant \eta$, on peut noter que

$$\|f * g_\varepsilon - f * h_\varepsilon\|_p = \|f * (g-h)_\varepsilon\|_p \leqslant \|f\|_p \,.\, \|(g-h)_\varepsilon\|_1 = \|f\|_p \,.\, \|g-h\|_1 \leqslant \alpha \|f\|_p$$

et on en déduit que $\|f - f * g_\varepsilon\|_p \leqslant \|f - f * h_\varepsilon\|_p + \alpha \|f\|_p < 2\delta$, donc $\lim_{\varepsilon \to 0} f * g_\varepsilon = f$. Les mêmes inégalités, avec $p = \infty$, montrent que $f * g_\varepsilon$ tend uniformément vers f lorsque ε tend vers 0.

Enfin, pour f quelconque dans L^p, il existe une suite (f_n) dans $\mathscr{C}_c(\mathbb{R}^d)$ qui converge vers f dans L^p. On a alors pour tout n :

$$\begin{aligned}\|f - f * g_\varepsilon\|_p &\leqslant \|f - f_n\|_p + \|f_n - f_n * g_\varepsilon\|_p + \|(f_n - f) * g_\varepsilon\| \\ &\leqslant (1 + \|g_\varepsilon\|_1) \|f - f_n\|_p + \|f_n - f_n * g_\varepsilon\|_p \\ &= (1 + \|g\|_1) \|f - f_n\|_p + \|f_n - f_n * g_\varepsilon\|_p\end{aligned}$$

et cette quantité peut être rendue arbitrairement petite, en choisissant d'abord n assez grand, puis ε assez petit pour un n fixé. Le même argument de densité de $\mathscr{C}_c(\mathbb{R}^d)$, avec $p = \infty$, montre la convergence uniforme de $f * g_\varepsilon$ vers f si $f \in \mathscr{C}_0(\mathbb{R}^d)$. ■

6.4 Compacité des opérateurs de convolution

Lemme 6.4.1. *Soit $f \in \mathscr{C}_c(\mathbb{R}^d)$. Alors l'ensemble $\{f * g : g \in L^1(\mathbb{R}^d), \|g\|_1 \leqslant 1\}$ est une partie équicontinue de $\mathscr{C}_0(\mathbb{R}^d)$.*

Démonstration : La fonction f est uniformément continue. Pour tout $\varepsilon > 0$, on peut donc trouver $\eta > 0$ tel que $\|x - y\| < \eta \Longrightarrow |f(x) - f(y)| < \varepsilon$.
Alors si $g \in B_1 := \{h \in L^1(\mathbb{R}^d) : \|h\|_1 \leqslant 1\}$ on a $f * g \in \mathscr{C}_0(\mathbb{R}^d)$ par le théorème 6.2.3. De plus si, x et y appartiennent à $\mathbb{R}^d$ avec $\|x - y\| < \eta$, on a $|f(x-z) - f(y-z)| < \varepsilon$ quel que soit $z \in \mathbb{R}^d$, donc :

$$|f * g(x) - f * g(y)| \leqslant \int |f(x-z) - f(y-z)| \,.\, |g(z)|\, d\mu(z) < \int \varepsilon. |g(z)|\, d\mu(z) \leqslant \varepsilon$$

ce qui achève la démonstration du lemme. ■

Théorème 6.4.2. *Soient $f \in \mathscr{C}_c(\mathbb{R}^d)$ et K un compact de $\mathbb{R}^d$. Alors, si $L^1(K)$ désigne le sous-espace de $L^1(\mathbb{R}^d)$ formé des fonctions nulles hors de K, l'opérateur $C_f : g \mapsto f * g$ est compact de $L^1(K)$ dans $\mathscr{C}_0(\mathbb{R}^d)$.*

Démonstration : Si T désigne le support de f, qui est compact, $K' = K+T$ est compact dans $\mathbb{R}^d$, et $f*g$ est une fonction continue à support dans K' pour tout $g \in L^1(K)$. Alors, si g appartient à la boule unité $B_1(K)$ de $L^1(K)$, on a $\|f*g\|_\infty \leqslant \|f\|_\infty \cdot \|g\|_1 \leqslant \|f\|_\infty$, ce qui montre que $C_f(B_1(K))$ est uniformément borné. En vertu du lemme précédent, cet ensemble est équicontinu. Il résulte alors du théorème d'Ascoli (théorème 1.10.6) que $\{(f*g)_{|K'} : g \in B_1(K)\}$ est une partie relativement compacte de $\mathscr{C}(K')$, donc, par isométrie, de $\mathscr{C}_0(\mathbb{R}^d)$. ∎

Théorème 6.4.3. *Soient $p > 1$, p' son exposant conjugué, $f \in L^{p'}(\mathbb{R}^d)$ et K compact de $\mathbb{R}^d$. Alors, si $L^p(K)$ désigne le sous-espace de $L^p(\mathbb{R}^d)$ formé des fonctions nulles hors de K, l'opérateur $C_f : g \mapsto f * g$ est compact de $L^p(K)$ dans $\mathscr{C}_0(\mathbb{R}^d)$.*

Démonstration : Il résulte du théorème 6.2.4 que C_f envoie continuement $L^p(\mathbb{R}^d)$ dans $\mathscr{C}_0(\mathbb{R}^d)$. Et de plus, si $g \in B_p(K) := \{h \in L^p(K) : \|h\|_p \leqslant 1\}$, il résulte de l'inégalité de Hölder que $\|g\|_1 \leqslant \|g\|_p . \mu(K)^{1/p'} \leqslant \mu(K)^{1/p'}$, donc que $B_p(K)$ est une partie bornée de $L^1(K)$. Si $\varphi \in \mathscr{C}_c(\mathbb{R}^d)$, le théorème précédent montre que $C_\varphi(B_p(K))$ est relativement compact dans $\mathscr{C}_0(\mathbb{R}^d)$. Et le théorème 6.2.4 montre que $\|C_f - C_\varphi\| \leqslant \|f-\varphi\|_{p'}$. Puisque $p' < \infty$, donc que $\mathscr{C}_c(\mathbb{R}^d)$ est dense dans $L^{p'}(\mathbb{R}^d)$, on conclut que C_f est limite en norme d'opérateurs compacts, donc lui-même compact (cf. théorème 2.5.6). ∎

On peut remarquer que l'énoncé devient faux si on prend $p = 1$ et $p' = \infty$: en effet, si $d = 1$, $f = \mathbb{1}_{[0,1[}$ et $g_n = 2^n . \mathbb{1}_{[0,2^{-n}]}$, la suite (g_n) est bornée dans L^1, mais la suite de fonctions continues $(f*g_n)$, qui tend simplement vers f, n'a aucune sous-suite qui converge uniformément dans $\mathscr{C}_0(\mathbb{R})$, puisque sa limite n'est pas continue.

Théorème 6.4.4. *Soient p, q et r dans $[1,\infty[$ vérifiant $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$, K un compact de $\mathbb{R}^d$ et $f \in L^q(\mathbb{R}^d)$. Alors, si $L^p(K)$ désigne le sous-espace de $L^p(\mathbb{R}^d)$ formé des fonctions nulles hors de K, l'opérateur $C_f : g \mapsto f * g$ est compact de $L^p(K)$ dans $L^r(\mathbb{R}^d)$.*

Démonstration : Il résulte du théorème 6.2.6 que C_f envoie continuement $L^p(\mathbb{R}^d)$ dans $L^r(\mathbb{R}^d)$ et que $\|C_f\| \leqslant \|f\|_q$. Si $\varphi \in \mathscr{C}_c(\mathbb{R}^d)$, si $T = \operatorname{supp}(\varphi)$ et $K' = K + T$, alors K' est un compact de $\mathbb{R}^d$. On a démontré au théorème précédent que C_φ est un opérateur compact de $L^p(K)$ dans $\mathscr{C}_0(\mathbb{R}^d)$. De plus, pour tout $g \in L^p(K)$, on a $\operatorname{supp}(f*g) \subset K'$, donc $\|\varphi * g\|_r \leqslant \|\varphi * g\|_\infty \mu(K')^{1/r}$. Il en résulte que l'injection dans $L^r(\mathbb{R}^d)$ du sous-espace $\mathscr{C}_{K'}(\mathbb{R}^d)$ de $\mathscr{C}_0(\mathbb{R}^d)$ formé des fonctions à support dans K' est continue, donc, par composition, que C_φ est un opérateur compact de $L^p(K)$ dans $L^r(\mathbb{R}^d)$.
A nouveau, par le théorème 6.2.6, on voit que $\|C_f - C_\varphi\| \leqslant \|f-\varphi\|_q$. Et par densité de $\mathscr{C}_c(\mathbb{R}^d)$ dans $L^q(\mathbb{R}^d)$, on conclut que C_f est limite en norme d'opérateurs compacts, donc lui-même compact (cf. théorème 2.5.6). ∎

6.5 Fonctions maximales

Lemme 6.5.1. (Vitali) *Soient K un compact de $\mathbb{R}^d$ et $(B(x_j,r_j))_{1\leqslant j\leqslant m}$ un recouvrement fini de K par des boules ouvertes. Alors il existe une sous-famille $(B(x_j,r_j))_{j\in J}$ formée de boules deux-à-deux disjointes telle que $\sum_{j\in J}\mu(B(x_j,r_j))\geqslant 3^{-d}\mu(K)$.*

On peut supposer les r_j rangés par ordre décroissant, c'est-à-dire $r_j\geqslant r_{j+1}$ pour $1\leqslant j<m$. Et on définit J comme la partie de $[1,m]$ maximisant $\Phi(J):=\sum_{j\in J}2^{-j}$ parmi celles pour lesquelles les boules sont deux-à-deux disjointes. Alors, si $k\in[1,m]\setminus J$, il existe $j<k$ dans J tel que $B(x_j,r_j)\cap B(x_k,r_k)\neq\emptyset$: on pourrait sinon considérer $J'=(J\cap[1,k])\cup\{k\}$, qui vérifie $\Phi(J')>\Phi(J)$ puisque $2^{-k}>\sum_{n=k+1}^m 2^{-n}\geqslant\sum_{n\in J\setminus J'}2^{-n}$. En particulier $B(x_k,r_k)\subset B(x_j,r_j+2r_k)\subset B(x_j,3r_j)$. Donc $K\subset\bigcup_{k\leqslant m}B(x_k,r_k)\subset\bigcup_{j\in J}B(x_j,3r_j)$; et on a

$$\mu(K)\leqslant\sum_{j\in J}\mu(B(x_j,3r_j))=3^d\sum_{j\in J}\mu(B(x_j,r_j))\ . \qquad\blacksquare$$

Définition 6.5.2. *Une fonction f sur un ouvert U de $\mathbb{R}^d$ est dite* localement intégrable *si tout point x de U possède un voisinage V tel que $f.\mathbb{1}_V$ soit intégrable. On note $L^1_{\rm loc}(U)$ l'ensemble des fonctions localement intégrables sur U.*
Il est équivalent de dire que $f.\mathbb{1}_K$ est intégrable pour tout compact K de U.

Pour toute fonction $f\in L^1_{\rm loc}$, on définit la *fonction maximale* f^* par

$$f^*(x)=\sup_{B(y,r)\ni x}\frac{1}{\mu(B(y,r))}\int_{B(y,r)}|f(z)|\ d\mu(z)$$

Cette fonction est clairement semi-continue inférieurement (c'est-à-dire que les ensembles $\{x:f^*(x)>\lambda\}$ sont ouverts pour tout λ) à valeurs dans $\overline{\mathbb{R}}_+$.

Lemme 6.5.3. *Pour tout $\lambda>0$, on a : $\lambda\mu(E^*_\lambda)\leqslant 3^d\int_{E^*_\lambda}|f(x)|\ d\mu(x)$, en notant E^*_λ l'ensemble $\{f^*>\lambda\}$.*

Démonstration : Soit K un compact contenu dans E^*_λ. Alors K est, par définition, recouvert par des boules ouvertes $B(x,r)$ telles que $\int_{B(x,r)}|f(y)|\ d\mu(y)>\lambda\mu(B(x,r))$. Chacune de ces boules est, par définition, contenue dans E^*_λ. En extrayant de cette famille un sous-recouvrement fini, puis en appliquant le lemme précédent, on trouve une famille disjointe de ces boules $(B(x_j,r_j))_{j\in J}$ telle que $C:=\bigcup_{j\in J}B(x_j,r_j)\subset E^*_\lambda$ et

$$\begin{aligned}\lambda\mu(K)&\leqslant 3^d\sum_{j\in J}\lambda.\mu(B(x_j,r_j))<3^d\sum_{j\in J}\int_{B(x_j,r_j)}|f(x)|\ d\mu(x)=3^d\int_C|f(x)|\ d\mu(x)\\&\leqslant 3^d\int_{E^*_\lambda}|f(x)|\ d\mu(x)\end{aligned}$$

d'où le résultat, puisque $\mu(E^*_\lambda)=\sup\{\mu(K):K\subset E^*_\lambda, K\text{ compact}\}$ $\blacksquare$

Lemme 6.5.4. *Si f est une fonction mesurable positive sur (X,μ) , et si $1 \leqslant p < \infty$, on a*

$$\int_X f(x)^p \, d\mu(x) = p \int_{\mathbb{R}^+} t^{p-1} \mu(\{f > t\} \, dt \ .$$

Démonstration : On voit, comme dans la démonstration du lemme 5.9.12, que l'ensemble $A = \{(x,t) : f(x) > t\}$ est mesurable, et qu'en appliquant le théorème de Fubini à la fonction $f : (x,t) \mapsto pt^{p-1} \mathbb{1}_A(x,t)$, on obtient $A_t = \{f > t\}$ et $A^x = [0, f(x)[$, d'où

$$\iint f(x,t) \, d\mu(x) \, dt = p \int_{\mathbb{R}^+} t^{p-1} \mu(A_t) \, dt = \int_X \Big(\int_{A^x} pt^{p-1} \, dt \Big) \, d\mu(x) = \int_X f(x)^p \, d\mu(x) \ . \quad \blacksquare$$

Lemme 6.5.5. *Soit $p \in]1,\infty[$. Si $f \in L^p$, alors $f^* \in L^p$, et on a $\|f^*\|_p \leqslant \dfrac{3^d.p}{p-1} \|f\|_p$.*

Démonstration : Supposons d'abord f bornée par M et à support contenu dans la boule $B(0,R)$. Alors la fonction f^* est bornée par M et vérifie $f^*(x) \leqslant \dfrac{\|f\|_1}{\mu(B(a,r))}$ pour toute boule $B(a,r)$ contenant x et rencontrant le support de f, donc $f^*(x) \leqslant \dfrac{C}{(\|x\| - R)^d}$ si $\|x\| > R$, ce dont on déduit que $f^* \in L^p$. De plus, en vertu des lemmes précédents,

$$\begin{aligned}
\int f^*(x)^p \, d\mu(x) = p \int_0^\infty t^{p-1} \mu(E_t^*) \, dt &\leqslant 3^d.p \int_0^\infty t^{p-2} \int_{t \leqslant f^*(x)} |f(x)| \, d\mu(x) \\
&= 3^d.p \int |f(x)| \, d\mu(x) \int_0^{f^*(x)} t^{p-2} \, dt \leqslant 3^d \frac{p}{p-1} \int |f(x)| \, f^*(x)^{p-1} \, d\mu(x) \\
&\leqslant 3^d \frac{p}{p-1} \|f\|_p \left\|(f^*)^{p-1}\right\|_{p'} \ .
\end{aligned}$$

Et puisque

$$\left\|(f^*)^{p-1}\right\|_{p'} = \Big(\int f^*(x)^{p'(p-1)} \, d\mu(x) \Big)^{1/p'} = \Big(\int f^*(x)^p \, d\mu(x) \Big)^{1/p'} = \|f^*\|_p^{p/p'} \ ,$$

on obtient

$$\|f^*\|_p^p \leqslant 3^d \frac{p}{p-1} \|f\|_p \|f^*\|_p^{p/p'} \ ,$$

donc $\|f^*\|_p = \|f^*\|_p^{p - \frac{p}{p'}} \leqslant 3^d \dfrac{p}{p-1} \|f\|_p$.

Dans le cas général, si $f \in L^p$, et si on note $f_n = \inf(n, |f| . \mathbb{1}_{B_n})$, où B_n est la boule de centre 0 et de rayon n, on trouve que f^* est la limite croissante des $(f_n)^*$, donc que

$$\Big(\int |f^*(x)|^p \, dx \Big)^{1/p} = \sup_n \Big(\int |f_n^*(x)|^p \, dx \Big)^{1/p} \leqslant 3^d \frac{p}{p-1} \sup_n \|f_n\|_p = 3^d \frac{p}{p-1} \|f\|_p \ ,$$

ce qui montre que f^* est dans L^p et prouve l'inégalité cherchée. $\blacksquare$

Lemme 6.5.6. *Soient $\alpha > 1$, A et $B \geqslant 0$. On a alors*

$$\int_0^\infty \min(A, Bt^{-\alpha})\, dt = \frac{\alpha}{\alpha-1} A^{1-\frac{1}{\alpha}} B^{\frac{1}{\alpha}}$$

Démonstration : Soit en effet θ tel que $A = B\theta^{-\alpha}$, c'est-à-dire $\theta = \big(\frac{B}{A}\big)^{1/\alpha}$. On a

$$\int_0^\infty \min(A, Bt^{-\alpha})\, dt = \int_0^\theta A\, dt + \int_\theta^\infty Bt^{-\alpha}\, dt = A\theta + B\frac{\theta^{1-\alpha}}{\alpha-1}$$
$$= A^{1-1/\alpha}B^{1/\alpha} + \frac{1}{\alpha-1}A^{1-1/\alpha}B^{1/\alpha} = \frac{\alpha}{\alpha-1}A^{1-1/\alpha}B^{1/\alpha}\ ,$$

ce qui achève la preuve. ∎

Théorème 6.5.7. *Si on note ψ la fonction $x \mapsto \|x\|^{1-d}$ sur $\mathbb{R}^d$, on a $f * \psi_1 \in L^q(\mathbb{R}^d)$ lorsque $f \in L^p(\mathbb{R}^d)$, $|\psi_1| \leqslant \psi$, $1 < p < d$ et $\frac{1}{q} = \frac{1}{p} - \frac{1}{d}$. Et il existe une constante $C(p,d)$ telle que $\|f * \psi_1\|_q \leqslant C(p,d)\, \|f\|_p$.*

Démonstration : Notons f^* la fonction maximale de f. Si on a $1 < p < d$, on a, pour tout a tel que $f^*(a) < +\infty$:

$$|f * \psi_1(a)| \leqslant \int |f(x))|\, \|x-a\|^{1-d}\ dx = \frac{1}{d-1} \iint_{t > \|x-a\|} |f(x)|\, t^{-d}\, dt\, dx$$
$$= \frac{1}{d-1}\int_0^\infty t^{-d} \int_{B(a,t)} |f(x)|\ dx\ .$$

Pour t fixé, on a $I_a(t) := \displaystyle\int_{B(a,t)} |f(x)|\ dx \leqslant f^*(a) v_d . t^d$ par définition de $f^*(a)$. On a aussi :

$$I_a(t) \leqslant \|f\|_p . \left\|\mathbb{1}_{B(a,t)}\right\|_{p'} = \|f\|_p\, t^{d/p'} v_d^{1/p'}$$

donc, en vertu du lemme précédent,

$$\int_0^\infty I_a(t)\, t^{-d}\, dt \leqslant \int_0^\infty \min\big(f^*(a)v_d, \|f\|_p\, v_d^{1/p'} t^{-d/p}\big)\, dt$$
$$\leqslant \frac{d}{d-p} f^*(a)^{1-\frac{p}{d}}\, \|f\|_p^{\frac{p}{d}}\, v_d^{1-\frac{p}{d}+\frac{p}{dp'}}$$
$$= \frac{dv_d^{1-\frac{1}{d}}}{d-p} f^*(a)^{1-\frac{p}{d}}\, \|f\|_p^{\frac{p}{d}} = \frac{dv_d^{1-\frac{1}{d}}}{d-p}\, \|f\|_p^{\frac{p}{d}}\, f^*(a)^{\frac{p}{q}}$$

et $|f * \psi_1(a)| \leqslant \dfrac{dv_d^{1-\frac{1}{d}}}{(d-1)(d-p)}\, \|f\|_p^{\frac{p}{d}}\, f^*(a)^{\frac{p}{q}}$. On en déduit que

$$\|f * \psi_1\|_q^q \leqslant \Big(\frac{dv_d^{1-\frac{1}{d}}}{(d-1)(d-p)}\, \|f\|_p^{\frac{p}{d}}\Big)^q \int f^*(x)^p\, dx \leqslant \Big(\frac{dv_d^{1-\frac{1}{d}}}{(d-1)(d-p)}\Big)^q \Big(\frac{3^d p}{p-1}\Big)^p \|f\|_p^{p+\frac{pq}{d}}\ ,$$

c'est-à-dire $\|f * \psi_1\|_q \leqslant C(p,d) . \|f\|_p < +\infty$, avec $C(p,d) = \dfrac{dv_d^{1-\frac{1}{d}}}{(d-1)(d-p)}\Big(\dfrac{3^d p}{p-1}\Big)^{1-\frac{p}{d}}$ puisque $p + \dfrac{pq}{d} = q$ et que $\dfrac{p}{q} = 1 - \dfrac{p}{d}$. ∎

6.6 Transformation de Fourier des fonctions

Pour x et ξ dans l'espace euclidien $\mathbb{R}^d$, on notera maintenant $x \cdot \xi$ le produit scalaire des vecteurs x et ξ.

Définition 6.6.1. *Pour $f \in L^1(\mathbb{R}^d)$ et $\xi \in \mathbb{R}^d$, on notera $\hat{f}(\xi)$ l'intégrale $\int f(x)\,\mathrm{e}^{-ix\cdot\xi}\,d\mu(x)$, qui est bien définie puisque $\left|f(x)\,\mathrm{e}^{-ix\cdot\xi}\right| = |f(x)|$. On appelle* transformée de Fourier *de f la fonction $\hat{f}$.*

Théorème 6.6.2. *Si $f \in L^1(\mathbb{R}^d)$, sa transformée de Fourier est continue bornée et on a $\left\|\hat{f}\right\|_\infty \leqslant \|f\|_1$.*

Démonstration : On a pour tout ξ : $\left|\int f(x)\,\mathrm{e}^{-ix\cdot\xi}\right| d\mu(x) \leqslant \int |f(x)|\,d\mu(x) = \|f\|_1$, d'où : $\left\|\hat{f}\right\|_\infty \leqslant \|f\|_1$. Et la continuité de $\hat{f}$ découle immédiatement du théorème 5.6.1 ■

Lemme 6.6.3. *Si f est de classe $\mathscr{C}^1$ à support compact et $1 \leqslant k \leqslant d$, on a $\widehat{\partial_k f}(\xi) = i\xi_k\hat{f}(\xi)$.*

Démonstration : Pour tout $x' = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_d) \in \mathbb{R}^{d-1}$ on notera, par abus, $f(x', t)$ pour $f(x_1, \ldots, x_{k-1}, t, x_{k+1}, \ldots, x_d)$.
Alors la fonction $f_{x'} : t \mapsto f(x', t)\,\mathrm{e}^{-ix'\cdot\xi' - it\xi_k}$ est $\mathscr{C}^1$ à support compact sur $\mathbb{R}$. On a donc $\displaystyle\int_{-\infty}^{+\infty} \frac{d}{dt} f_{x'}(t)\,dt = 0$, c'est-à-dire

$$\int_{-\infty}^{+\infty} \left(\frac{\partial f}{\partial x_k}(x', t)\,\mathrm{e}^{-ix\cdot\xi} - i\xi_k f(x', t)\,\mathrm{e}^{-ix\cdot\xi}\right) dt\ ,$$

ou encore

$$\int_{-\infty}^{+\infty} \frac{\partial f}{\partial x_k}(x', t)\,\mathrm{e}^{-ix\cdot\xi}\,dt = i\xi_k \int_{-\infty}^{+\infty} f(x', t)\,\mathrm{e}^{-ix\cdot\xi}\,dt\ ,$$

et en intégrant sur $\mathbb{R}^{d-1}$ par rapport à x', on obtient l'égalité cherchée. ■

Corollaire 6.6.4. *Si f est dans $L^1(\mathbb{R}^d)$, alors $\hat{f}$ est dans $\mathscr{C}_0(\mathbb{R}^d)$.*

Démonstration : Si f est $\mathscr{C}^1$ à support compact, il résulte du lemme 6.6.3 ci-dessus que $\left|\xi_j.\hat{f}(\xi)\right| \leqslant \left\|\dfrac{\partial f}{\partial x_j}\right\|_1$ et que $\|\xi\| . \left|\hat{f}(\xi)\right| \leqslant \sum_j |\xi_j| . \left|\hat{f}(\xi)\right| \leqslant \sum_j \left\|\dfrac{\partial f}{\partial x_j}\right\|_1$. Ceci montre que $\hat{f}(\xi) = O(\|\xi\|^{-1})$ et que la fonction continue $\hat{f}$ tend vers 0 à l'infini lorsque $f \in \mathscr{C}_c^1(\mathbb{R}^d)$.
Enfin, si f est quelconque dans L^1, il existe une suite (f_n) dans $\mathscr{C}_c^1(\mathbb{R}^d)$ qui converge vers f dans L^1. Alors la suite $(\hat{f}_n)$ converge uniformément vers $\hat{f}$, ce qui montre que $\hat{f}$ tend vers 0 à l'infini. ■

Lemme 6.6.5. *Si $f \in \mathscr{C}^1(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$ et si la différentielle de f est intégrable, on a*

$$\widehat{\partial_k f}(\xi) = i\xi_k.\hat{f}(\xi)$$

Démonstration : On choisit une fonction φ à support compact sur $\mathbb{R}^d$, de classe $\mathscr{C}^1$, à valeurs dans $[0, 1]$ et égale à 1 sur la boule unité, puis on pose, pour tout $\varepsilon > 0$, $f^{(\varepsilon)}(x) = f(x)\varphi(\varepsilon x)$.

La fonction $f^{(\varepsilon)}$ est $\mathscr{C}^1$ à support compact, vérifie $\partial_k f^{(\varepsilon)}(x) = \partial_k f(x)\varphi(\varepsilon x) + \varepsilon f(x)\partial_k\varphi(\varepsilon x)$, et la formule précédente montre que

$$\int \partial_k f(x)\varphi(\varepsilon x)\,\mathrm{e}^{-ix\cdot\xi}\ d\mu(x) + \varepsilon \int f(x)\partial_k\varphi(\varepsilon x)\,\mathrm{e}^{-ix\cdot\xi}\ d\mu(x) = i\xi_k \int f(x)\varphi(\varepsilon x)\,\mathrm{e}^{-ix\cdot\xi}\ d\mu(x)\ .$$

Par convergence dominée, on voit que la première intégrale tend vers $\widehat{\partial_k f}(\xi)$ et la troisième vers $\hat{f}(\xi)$ lorsque ε tend vers 0. Et puisque

$$\left|\varepsilon \int f(x)\partial_k\varphi(\varepsilon x)\,\mathrm{e}^{-ix\cdot\xi}\ d\mu(x)\right| \leqslant \varepsilon \left\|f\right\|_1 . \sup_y |\partial_k\varphi(y)| \to 0$$

lorsque ε tend vers 0, on aboutit à la conclusion cherchée. ■

Lemme 6.6.6. *Si $f \in L^1(\mathbb{R}^d)$ et si la fonction $g : x \mapsto \|x\| . f(x)$ est intégrable, la fonction $\hat{f}$ est de classe $\mathscr{C}^1$ et, pour $1 \leqslant k \leqslant d$, on a $\partial_k \hat{f}(\xi) = -i \int f(x) x_k\,\mathrm{e}^{-ix\cdot\xi}\ d\mu(x)$.*

Démonstration : Puisque la fonction $g_k : x \mapsto x_k f(x)$ est dans L^1, sa transformée de Fourier $\widehat{g_k}$ est continue et la formule de différentiation d'une intégrale (5.6.3) donne

$$\partial_k \hat{f}(\xi) = \int f(x) \frac{\partial}{\partial \xi_k}(\mathrm{e}^{-ix\cdot\xi})\, d\mu(x) = \int f(x) . (-ix_k\,\mathrm{e}^{-ix\cdot\xi})\, d\mu(x) = -i\widehat{g_k}(\xi)\ ,$$

d'où la continuité des dérivées partielles et la formule annoncée. ■

Théorème 6.6.7. *Soient f et g dans $L^1(\mathbb{R}^d)$. Alors la transformée de Fourier de $f * g$ est le produit $\hat{f}.\hat{g}$ des transformées de Fourier de f et de g.*

Démonstration : On a, par théorème de Fubini et changement de variables :

$$\begin{aligned}\widehat{f * g}(\xi) &= \int \left(\int f(y) g(x-y)\, d\mu(y)\right) \mathrm{e}^{-ix\cdot\xi}\ d\mu(x) = \iint f(y) g(z)\,\mathrm{e}^{-i(y+z)\xi}\ d\mu(y)\, d\mu(z)\\ &= \int f(y)\,\mathrm{e}^{-iy\xi}\ d\mu(y) . \int g(z)\,\mathrm{e}^{-iz\xi}\ d\mu(z) = \hat{f}(\xi).\hat{g}(\xi)\ ,\end{aligned}$$

c'est-à-dire $\widehat{f * g} = \hat{f}.\hat{g}$. ■

Lemme 6.6.8. *Pour tout $t \in \mathbb{R}$, on a*

$$\int_{\mathbb{R}} \mathrm{e}^{-x^2/2}\ \mathrm{e}^{-ixt}\ dx = \sqrt{2\pi}\ \mathrm{e}^{-t^2/2}\ .$$

Démonstration : On voit que la fonction $g(t) = \int \mathrm{e}^{-x^2/2}\ \mathrm{e}^{-itx}\ dx$ vérifie $g(0) \geqslant 0$ et

$$g(0)^2 = \iint \mathrm{e}^{-(x^2+y^2)/2}\ dx\, dy = \iint r\,\mathrm{e}^{-r^2/2}\ dr\, d\theta = 2\pi \int_0^\infty \mathrm{e}^{-r^2/2}\ dr = 2\pi$$

et, par dérivation sous le signe $\int$ et intégration par parties :

$$g'(t) = \int -ix\,\mathrm{e}^{-x^2/2}\,\mathrm{e}^{-itx}\ dx = \left[i\,\mathrm{e}^{-x^2/2}\,\mathrm{e}^{-itx}\right]_{-\infty}^{+\infty} - \int i\,\mathrm{e}^{-x^2/2}(-it\,\mathrm{e}^{-itx})\, dx = -tg(t)$$

donc $g(t) = \lambda\,\mathrm{e}^{-t^2/2}$, pour une constante λ. Alors $\lambda = g(0) = \sqrt{2\pi}$. ■

Lemme 6.6.9. *Soient u un isomorphisme linéaire de $\mathbb{R}^d$ et ${}^t u$ son transposé. Si $f \in L^1(\mathbb{R}^d)$, on a $\widehat{f\circ u^{-1}}(\xi) = |\det(u)|\,.\hat{f}\circ{}^t u(\xi)$. En particulier, si u est l'homothétie de rapport $\varepsilon > 0$, on a $\widehat{f_\varepsilon}(\xi) = \hat{f}(\varepsilon.\xi)$.*

Démonstration : On a $\widehat{f\circ u^{-1}}(\xi) = \int f\circ u^{-1}(x)\,\mathrm{e}^{-i\langle x,\xi\rangle}\,d\mu(x)$. Par le changement de variable $x = u.y$, on trouve

$$\begin{aligned}\int f\circ u^{-1}(x)\,\mathrm{e}^{-i\langle x,\xi\rangle}\,d\mu(x) &= \int f(y)\,\mathrm{e}^{-i\langle u.y,\xi\rangle}\,|\det(u)|\,d\mu(y)\\ &= \det(u)\int f(y)\,\mathrm{e}^{-i\langle y,{}^t u.\xi\rangle}\,d\mu(y) = |\det(u)|\,.\hat{f}({}^t u\xi)\ .\end{aligned}$$

Et si u est l'homothétie de rapport $\varepsilon > 0$, on a $\det(u) = \varepsilon^d$ et ${}^t u = u$. ■

On notera désormais γ la fonction de classe $\mathscr{C}^\infty$ définie pour $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ par

$$\gamma(x) = (2\pi)^{-d/2}\ \mathrm{e}^{-\|x\|^2/2} = \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x_j^2/2}$$

On vérifie sans peine que γ est dans $L^p(\mathbb{R}^d)$ pour tout $p \in [1, +\infty[$, ainsi que dans $\mathscr{C}_0(\mathbb{R}^d)$.

Lemme 6.6.10. *On a $\hat{\gamma}(\xi) = (2\pi)^{d/2}\gamma(\xi)$. En particulier $\int \gamma(x)\,d\mu(x) = \hat{\gamma}(0) = \|\gamma\|_1 = 1$.*

Démonstration : Ceci se déduit immédiatement du lemme 6.6.8. ■

Lemme 6.6.11. *Soient $f \in L^1(\mathbb{R}^d)$, $\varepsilon > 0$ et $a \in \mathbb{R}^d$. Alors on a*

$$\int \widehat{f*\gamma_\varepsilon}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi) = (2\pi)^d.f*\gamma_\varepsilon(a)$$

Démonstration : La fonction $h : (x,\xi) \mapsto f(x)\widehat{\gamma_\varepsilon}(\xi)\,\mathrm{e}^{i(a-x)\cdot\xi}$ est intégrable sur $\mathbb{R}^d \times \mathbb{R}^d$ puisque $f \in L^1(\mathbb{R}^d)$, que $\widehat{\gamma_\varepsilon}(\xi) = (2\pi)^{d/2}\gamma(\varepsilon\xi)$ et que

$$\int \gamma(\varepsilon\xi)\,d\mu(\xi) = \varepsilon^{-d}\int \gamma(\xi)\,d\mu(\xi) = \varepsilon^{-d} < \infty\ .$$

En appliquant le théorème de Fubini et en utilisant la parité de γ, on trouve alors

$$\begin{aligned}\iint h(x,\xi)\,d\mu(x)\,d\mu(\xi) &= \int f(x)\left(\int (2\pi)^{d/2}\gamma(\varepsilon\xi)\,\mathrm{e}^{i(a-x)\cdot\xi}\,d\mu(\xi)\right)d\mu(x)\\ &= (2\pi)^{d/2}\int f(x)\varepsilon^{-d}\hat{\gamma}(\frac{x-a}{\varepsilon})\,d\mu(x)\\ &= (2\pi)^d\int f(x)\varepsilon^{-d}\gamma(\frac{x-a}{\varepsilon})\,d\mu(x) = (2\pi)^d\int f(x)\gamma_\varepsilon(a-x)\,d\mu(x)\\ &= (2\pi)^d f*\gamma_\varepsilon(a)\ .\end{aligned}$$

On trouve aussi

$$\begin{aligned}\iint h(x,\xi)\,d\mu(x)\,d\mu(\xi) &= \int \mathrm{e}^{ia\cdot\xi}\ \widehat{\gamma_\varepsilon}(\xi)\left(\int f(x)\,\mathrm{e}^{-ix\cdot\xi}\,d\mu(x)\right)d\mu(\xi)\\ &= \int \mathrm{e}^{ia\cdot\xi}\,\hat{f}(\xi)\,\widehat{\gamma_\varepsilon}(\xi)\,d\mu(\xi) = \int \widehat{f*\gamma_\varepsilon}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi)\ ,\end{aligned}$$

ce qui achève la démonstration du lemme. ■

Théorème 6.6.12. (Formule d'inversion) *Si f et $\hat{f}$ sont intégrables, on a pour presque tout $a \in \mathbb{R}^d$:*

$$f(a) = (2\pi)^{-d} \int \hat{f}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi)\ .$$

Démonstration : Puisque $\left|\widehat{f * \gamma_\varepsilon}(\xi)\ \mathrm{e}^{ia\cdot\xi}\right| = \left|\hat{f}(\xi)\right| |\widehat{\gamma_\varepsilon}(\xi)| \leqslant \left|\hat{f}(\xi)\right|$ et que $\lim_{\varepsilon\to 0} \widehat{\gamma_\varepsilon}(\xi) = 1$ pour tout ξ, il résulte du théorème de convergence dominée que

$$\lim_{\varepsilon\to 0} \int \widehat{f * \gamma_\varepsilon}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi) = \int \hat{f}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi)\ .$$

De plus, on sait (cf. théorème 6.3.2) que $\lim_{\varepsilon\to 0} \|f * \gamma_\varepsilon - f\|_1 = 0$. On peut donc (cf. théorème 5.5.9) trouver une suite (ε_k) tendant vers 0 telle que la suite $(f * \gamma_{\varepsilon_k})(a)$ converge vers $f(a)$ pour presque tout a.
Il résulte alors du lemme 6.6.11 que, pour presque tout a, on a

$$f(a) = (2\pi)^{-d} \int \hat{f}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi)\ .$$

On peut remarquer qu'alors, puisque $\hat{f}$ est dans L^1, sa transformée de Fourier est dans $\mathscr{C}_0(\mathbb{R}^d)$, donc que la fonction $a \mapsto (2\pi)^{-d} \int \hat{f}(\xi)\ \mathrm{e}^{-ia\cdot\xi}\ d\mu(\xi)$, qui est presque partout égale à $f(-a)$, est continue. ■

Corollaire 6.6.13. *La transformation de Fourier est injective sur L^1.*

Démonstration : Soit f une fonction dans $L^1(\mathbb{R}^d)$ telle que $\hat{f} = 0$. On a alors $\hat{f} \in L^1$, et la formule d'inversion ci-dessus montre que $f = 0$ dans L^1. ■

Lemme 6.6.14. *Si la fonction f est dans $L^1 \cap L^2$, on a*

$$\int \left|\hat{f}(\xi)\right|^2\ d\mu(\xi) = (2\pi)^d \int |f(x)|^2\ d\mu(x)\ .$$

Démonstration : Notons f^* la fonction définie par $f^*(x) = \bar{f}(-x)$, qui est, comme f, dans $L^1 \cap L^2$. On a alors :

$$\widehat{f^*}(\xi) = \int \bar{f}(-x)\,\mathrm{e}^{-ix\cdot\xi}\ d\mu(x) = \int \bar{f}(x)\,\mathrm{e}^{ix\cdot\xi}\ d\mu(x) = \overline{\hat{f}(\xi)}\ .$$

En appliquant à la fonction $g = f * f^*$ le lemme 6.6.11, on trouve, pour $\varepsilon > 0$,

$$\int \hat{g}(\xi)\,\widehat{\gamma_\varepsilon}(\xi)\ \mathrm{e}^{ia\cdot\xi}\ d\mu(\xi) = (2\pi)^d g * \gamma_\varepsilon(a)\ ,$$

donc, pour $a = 0$,

$$(2\pi)^{d/2} \int \hat{f}(\xi)\,\overline{\hat{f}(\xi)}\,\gamma(\varepsilon\xi)\,d\mu(\xi) = (2\pi)^d g * \gamma_\varepsilon(0)\ ,$$

c'est-à-dire

$$\int \left|\hat{f}(\xi)\right|^2 \gamma(\varepsilon\xi)\,d\mu(\xi) = (2\pi)^{d/2} g * \gamma_\varepsilon(0)\ .$$

Alors, puisque $f * f^*$, convolée de deux fonctions de L^2, appartient à $\mathscr{C}_0(\mathbb{R}^d)$, $g * \gamma_\varepsilon$ tend uniformément vers g lorsque ε tend vers 0. De plus, si (ε_k) est une suite qui décroît vers 0, la suite $\gamma(\varepsilon_k\xi)$ tend en croissant vers $\gamma(0) = (2\pi)^{-d/2}$. Le théorème de convergence monotone montre alors que $\lim_k \int \left|\hat{f}(\xi)\right|^2 \gamma(\varepsilon_k\xi)\,d\mu(\xi) = (2\pi)^{-d/2} \int \left|\hat{f}(\xi)\right|^2\ d\mu(\xi)$. On conclut que

$$\int \left|\hat{f}(\xi)\right|^2\ d\mu(\xi) = (2\pi)^d g(0) = (2\pi)^d \int f(x) f^*(-x)\,d\mu(x) = (2\pi)^d \int |f(x)|^2\ d\mu(x)\ ,$$

ce qui achève la preuve. ■

Théorème 6.6.15. (Formule de Plancherel) *L'application* $\mathscr{F} : f \mapsto (2\pi)^{-d/2}\hat{f}$ *définie sur* $L^1 \cap L^2$ *se prolonge en une isométrie surjective de* L^2 *sur* L^2. *L'application réciproque* $\overline{\mathscr{F}}$ *est définie sur* $L^1 \cap L^2$ *par* $\overline{\mathscr{F}}f(x) = (2\pi)^{-d/2}\hat{f}(-x)$.

DÉMONSTRATION : L'application $\mathscr{F}$ est clairement linéaire de $L^1 \cap L^2$ dans $\mathscr{C}_0(\mathbb{R}^d)$, et le lemme précédent montre d'une part que $\mathscr{F}(L^1 \cap L^2) \subset L^2$ et d'autre part que

$$\|\mathscr{F}(f)\|_2^2 = (2\pi)^{-d} \int \left|\hat{f}(\xi)\right|^2 d\mu(\xi) = \int |f(x)|^2 \, d\mu(x) = \|f\|_2^2 \ ,$$

c'est-à-dire que $\mathscr{F}$ est isométrique. Puisque $L^1 \cap L^2$ contient $\mathscr{C}_c(\mathbb{R}^d)$, donc est dense dans L^2, $\mathscr{F}$ se prolonge par continuité en une isométrie linéaire de L^2 dans L^2.

Pour les mêmes raisons, l'application linéaire $\overline{\mathscr{F}}$, définie sur $L^1 \cap L^2$ par

$$\overline{\mathscr{F}}f(x) = (2\pi)^{-d/2}\hat{f}(-x)$$

se prolonge en une isométrie de L^2 dans L^2. Et il résulte du lemme 6.6.12, appliqué à f^*, que si f et $\hat{f}$ sont dans $L^1 \cap L^2$, on a $\mathscr{F} \circ \overline{\mathscr{F}}(f) = f$. Il suffit donc de montrer que l'ensemble des f telles que f et $\hat{f}$ soient dans $L^1 \cap L^2$ est dense dans L^2 pour montrer que $\mathscr{F} \circ \overline{\mathscr{F}}$ est l'identité de L^2, donc que $\mathscr{F}$ est surjective.
Soient donc $h \in L^2$ et $\eta > 0$. Puisque $\mathscr{C}_c(\mathbb{R}^d)$ est dense dans L^2, il existe $g \in \mathscr{C}_c(\mathbb{R}^d)$ telle que $\|h - g\|_2 < \eta/2$, et un $\varepsilon > 0$ tel que $\|g - g * \gamma_\varepsilon\|_2 < \eta/2$. On a alors, avec $f = g * \gamma_\varepsilon$, $\|h - f\|_2 < \eta$. Puisque $g \in L^1$, il en est de même de f, et puisque $g \in L^2$, f aussi est dans L^2. Comme on a $\hat{f} = \hat{g}.\widehat{\gamma_\varepsilon}$ et que $\|\hat{g}\|_\infty \leqslant \|g\|_1$, il suffit de constater que $\widehat{\gamma_\varepsilon}$ est dans $L^1 \cap L^2$, ce qui est clair puisque $\widehat{\gamma_\varepsilon}(\xi) = (2\pi)^{d/2}\gamma(\varepsilon\xi)$, pour voir que $\hat{f} \in L^1 \cap L^2$. ∎

7

DISTRIBUTIONS

7.1 L'espace des fonctions de test

Définition 7.1.1. *Soit Ω un ouvert de l'espace euclidien $\mathbb{R}^d$. On appelle* fonction de test *sur Ω toute fonction complexe indéfiniment différentiable à support compact sur Ω. On note $\mathscr{D}(\Omega)$ l'espace vectoriel des fonctions de test sur Ω.*

Si (K_n) est une suite exhaustive de compacts de Ω et si on désigne par $\mathscr{E}(\Omega) = \mathscr{C}^\infty(\Omega)$ l'espace de Fréchet des fonctions indéfiniment différentiables sur Ω muni de la topologie de la convergence compacte de toutes les dérivées, le sous-espace $\mathscr{E}_n(\Omega)$ des fonctions nulles hors de K_n est fermé et $\mathscr{D}(\Omega) = \bigcup_n \mathscr{E}_n$. On munit $\mathscr{D}(\Omega)$ de la topologie limite inductive des $\mathscr{E}_n$, dont il est aisé de voir qu'elle ne dépend pas de la suite (K_n) choisie. Pour $\alpha \in \mathbb{N}^d$ et K compact contenu dans Ω, on notera p_α (resp. $p_{\alpha,K}$) la semi-norme $\varphi \mapsto \sup_x |\partial_\alpha \varphi(x)|$ (resp. $\sup_{x\in K} |\partial_\alpha \varphi(x)|$). On notera aussi $|\alpha| = \sum_{j=1}^d \alpha(j)$. Et on notera $\mathscr{D}_K(\Omega)$ le sous-espace des $\varphi \in \mathscr{D}(\Omega)$ dont le support est contenu dans K.

Théorème 7.1.2. *Les suites convergentes dans $\mathscr{D}(\Omega)$ sont les suites de fonctions (φ_n) dont toutes les dérivées convergent uniformément, et dont les supports restent dans un compact fixe de Ω.*

Démonstration : Il est montré au corollaire 4.8.3 que dans une limite inductive d'espaces de Fréchet (F_n), les suites convergentes sont les suites dont tous les termes appartiennent à un même F_n et qui convergent dans cet espace. ■

Théorème 7.1.3. *Les compacts de $\mathscr{D}(\Omega)$ sont les parties fermées et bornées.*

Démonstration : Si H est compact, il est fermé ; et pour tout voisinage ouvert convexe V de 0 dans $\mathscr{D}(\Omega)$, la suite croissante d'ouverts (nV) recouvre H : l'un d'entre eux contient donc K, ce qui montre que K est borné.

Inversement, si H est borné, il résulte du théorème 7.1.2 que tous les éléments de H ont leur support dans un même compact K de Ω : si (K_n) est une suite exhaustive de compacts de Ω et s'il existait pour tout n dans H une fonction φ_n dont le support n'est pas contenu dans K_n, la suite $(2^{-n}.\varphi_n)$ ne pourrait converger vers 0. Chacune des semi-normes continues p_α est alors bornée sur H par un $M_\alpha < +\infty$.
Si $\varphi \in H$, φ est $\mathscr{C}^\infty$ sur $\mathbb{R}^d$ et ses dérivées partielles sont nulles hors de K. Il résulte alors du théorème des accroissements finis, appliqué à $\partial_\alpha\varphi$ entre les points x et y de K, que, pour

$\varphi \in H$,
$$|\partial_\alpha \varphi(x) - \partial_\alpha \varphi(y)| \leqslant \|x-y\| \left(\sum_{j=1}^{d} M_{\alpha+\tau_j}^2\right)^{1/2},$$
où τ_j désigne l'élément de $\mathbb{N}^d$ dont toutes les coordonnées sont toutes nulles sauf la $j^{\text{ème}}$ qui vaut 1. On en déduit que l'ensemble $H_\alpha := \{\partial_\alpha \varphi : \varphi \in H\}$, qui est uniformément borné par M_α, est équicontinu dans $\mathscr{C}(K)$. Il résulte alors du théorème d'Ascoli (théorème 1.10.6) que H_α est relativement compact dans $\mathscr{C}(K)$. Par définition de la topologie de $\mathscr{D}_K(\Omega)$, l'application $\Psi : H \to T = \prod_{\alpha\in\mathbb{N}^d} \overline{H_\alpha}$ est un homéomorphisme de H sur le sous-espace $\Psi(H)$ du compact métrisable T. Alors, si (φ_n) est une suite dans H, la suite $(\Psi(\varphi_n))$ possède une sous-suite $(\Psi(\varphi_{n_k}))$ convergente dans T, ce qui signifie que les suites $(\partial_\alpha \varphi_{n_k})$ convergent uniformément pour chaque α. Alors les fonctions $h_\alpha = \lim_k \partial_\alpha \varphi_{n_k}$ sont toutes de classe $\mathscr{C}^1$ et vérifient $\dfrac{\partial}{\partial x_j} h_\alpha = h_{\alpha+\tau_j}$. Il en résulte que $h \in \mathscr{C}^\infty$, que $\partial_\alpha h = h_\alpha$ pour tout α et que $\operatorname{supp}(h) \subset K$, donc que $h \in \mathscr{D}_K(\Omega)$ et que $\varphi_{n_k} \to h$ dans $\mathscr{D}_K(\Omega)$. Et puisque toute suite de H possède une sous-suite qui converge dans l'espace de Fréchet $\mathscr{D}_K(\Omega)$, H est relativement compact dans $\mathscr{D}_K(\Omega)$, et a fortiori dans $\mathscr{D}(\Omega)$. L'ensemble H est donc compact si, de plus, il est fermé dans $\mathscr{D}(\Omega)$. ■

Théorème 7.1.4. *Soit (φ_n) une suite de fonctions de $\mathscr{D}(\Omega)$. On suppose qu'il existe un compact K de Ω tel que* $\operatorname{supp}(\varphi_n) \subset K$ *pour tout n, que la suite (φ_n) converge simplement vers φ et que, pour tout $\alpha \in \mathbb{N}^d$, $\sup_n \sup_x |\partial_\alpha \varphi_n(x)| < +\infty$. Alors $\varphi \in \mathscr{D}(\Omega)$ et (φ_n) converge vers φ dans $\mathscr{D}(\Omega)$.*

Démonstration : Il résulte du théorème 7.1.3 que l'ensemble $\{\varphi_n : n \in \mathbb{N}\}$ est relativement compact dans $\mathscr{D}(\Omega)$. De plus, si h est une valeur d'adhérence de (φ_n) dans $\mathscr{D}(\Omega)$, on a nécessairement $h(x) = \varphi(x)$ pour tout $x \in \Omega$. Ceci montre que $\varphi = h \in \mathscr{D}(\Omega)$ et que cette valeur d'adhérence est unique, donc est la limite de la suite. ■

Le lemme suivant permet d'interpréter la différentiabilité d'une fonction réelle ou complexe en termes de continuité, et sera commode pour manipuler une notion de "différentiabilité faible" d'une fonction φ définie sur un ouvert Ω de $\mathbb{R}^d$ et à valeurs dans un espace localement convexe E (dans le sens que $\xi\circ\varphi$ est de classe $\mathscr{C}^1$ sur Ω pour tout $\xi \in E'$), en pratique l'espace $\mathscr{D}(\Omega')$ dans la section 7.9 de ce chapitre, ou l'espace $\mathscr{S}(\mathbb{R}^\ell)$ dans la section 8.5 du chapitre suivant.

Lemme 7.1.5. *Soient Ω un ouvert de $\mathbb{R}^d$, $f : \Omega \to \mathbb{C}$ et $g = (g_1, g_2, \ldots, g_d) : \Omega \to \mathbb{C}^d$ deux fonctions. On définit la fonction $\Theta(f,g)$ sur Ω^2 par*
$$\Theta(f,g) : (x,y) = \begin{cases} 0 & \text{si } x = y, \\ \dfrac{f(y) - f(x) - \langle g(x), y-x\rangle}{\|y-x\|} & \text{si } y \neq x. \end{cases}$$
avec $\langle g(x), h\rangle = \sum_j g_j(x).h_j$, pour $h = (h_1, h_2, \ldots, h_d) \in \mathbb{R}^d$. Alors f est de classe $\mathscr{C}^1$ sur Ω et y admet les g_j pour dérivées partielles si et seulement si $\Theta(f,g)$ est continue sur Ω^2.

Démonstration : Si f est de classe $\mathscr{C}^1$ et $g_j = \partial_j f$, il est clair que $\Theta(f,g)$ est continue en tout point (x,y) tel que $x \neq y$. De plus, si $a \in \Omega$ et si $B(a,r) \subset \Omega$, la fonction $f_1 : x \mapsto f(x) - \langle g(a), x\rangle$ est différentiable en a avec une différentielle nulle : il résulte donc du théorème des accroissements finis que $|f_1(x) - f_1(y)| \leqslant \|x - y\| \,.\, \sup_{z \in B(a,r)} \|f_1'(z)\|$ si x et y appartiennent à $B(a, r)$. Et puisque

$$\begin{aligned} |\Theta(f, g)(x, y)| &= \frac{1}{\|x - y\|} \, |f_1(y) - f_1(x) - \langle g(x) - g(y), x - y \rangle| \\ &\leqslant \|g(y) - g(x)\| + \sup_{z \in B(a,r)} \|f_1'(z)\| \end{aligned}$$

on voit que $\Theta(f, g)(x, y) \to 0 = \Theta(f, g)(a, a)$ lorsque $(x, y) \to (a, a)$, d'où la continuité de Θ en tout point de Ω^2.

Inversement, si $\Theta(f, g)$ est continue sur Ω^2, on voit, en fixant arbitrairement x, que f est continue sur $\Omega \setminus \{x\}$, donc continue en tout point de Ω. On voit ensuite, en fixant $h \neq 0$ et prenant $y = x + h$, que g est continue en tout point de Ω. Alors la continuité en a de $x \mapsto \Theta(f, g)(x, a)$ montre que f est différentiable en a, avec les $g_j(a)$ pour dérivées partielles, et la continuité de g achève de prouver que f est de classe $\mathscr{C}^1$. ■

Lemme 7.1.6. *Soient Ω un ouvert de $\mathbb{R}^d$, Ω' un ouvert de $\mathbb{R}^\ell$ et $\varphi \in \mathscr{D}(\Omega \times \Omega')$. Alors, si φ_x désigne la fonction $y \mapsto \varphi(x, y)$ sur Ω', la fonction $x \mapsto \varphi_x$ est continue de Ω dans $\mathscr{D}(\Omega')$. De plus, si $\partial\varphi$ désigne la différentielle partielle de φ par rapport à la variable $x \in \Omega$, la fonction $\psi = \Theta(\varphi_x, (\partial\varphi)_x)$ est continue sur Ω^2.*

Démonstration : Identifiant $\mathbb{N}^{d+\ell}$ à $\mathbb{N}^d \times \mathbb{N}^\ell$, on notera, pour $\alpha \in \mathbb{N}^d$ et $\beta \in \mathbb{N}^\ell$: $\partial_{\alpha,\beta}\varphi$ la dérivée partielle de φ obtenue en dérivant "α fois" par rapport à la première variable (dans $\mathbb{R}^d$) et "β fois" par rapport à la seconde (dans $\mathbb{R}^\ell$). Puisque chaque $\partial_{\alpha,\beta}\varphi$ est continue sur le compact $K = \operatorname{supp}(\varphi)$, elle y est bornée en module par un nombre $M_{\alpha,\beta}$. De plus, la projection sur $\mathbb{R}^\ell$ de K est un compact L contenu dans Ω'. Si $(x^{(j)})$ est une suite dans Ω qui converge vers x, on a $\varphi(x^{(j)}, y) \to \varphi(x, y)$ pour tout y de Ω'. Il résulte donc du théorème 7.1.4 que $\varphi_{x^{(j)}}$ converge vers φ_x dans $\mathscr{D}(\Omega')$, c'est à-dire que la fonction $x \mapsto \varphi_x$ est continue de Ω dans $\mathscr{D}(\Omega')$.

On voit comme plus haut que les fonctions $\psi(x, u) \in \mathscr{D}(\Omega')$ ont leur support dans K et leurs dérivées partielles uniformément majorées pour $(x, u) \in \Omega^2$. Et puisque, pour tout $y \in \Omega'$, la fonction $x \mapsto \varphi(x, y)$ est de classe $\mathscr{C}^1$, la fonction $(x, u) \mapsto \psi(x, u)(y)$ est continue sur Ω^2. Il résulte donc encore du théorème 7.1.4 que ψ est continue de Ω^2 dans $\mathscr{D}(\Omega')$. ■

Théorème 7.1.7. *Soient $\varphi \in \mathscr{D}(\mathbb{R}^d)$ et $\rho \in L^1(\mathbb{R}^d)$ d'intégrale 1 à support compact. Alors $\varphi * \rho_\varepsilon$ tend vers φ dans $\mathscr{D}$ lorsque ε tend vers 0.*

Démonstration : Soit $R = \sup_{x \in \operatorname{supp}(\rho)} \|x\|$. Si on note $H = \{\varphi * \rho_\varepsilon \,:\, 0 < \varepsilon \leqslant 1\}$, le support de tout élément de H est contenu dans le compact $K = \operatorname{supp}(\varphi) + \tilde{B}(0, R)$. De plus, on a

$$p_\alpha(\varphi * \rho_\varepsilon) = \sup_x |\partial_\alpha\varphi * \rho_\varepsilon(x)| \leqslant p_\alpha(\varphi). \, \|\rho\|_1 \;,$$

ce qui montre que H est relativement compact. Et puisque $\varphi * \rho_\varepsilon$ converge uniformément vers φ quand ε tend vers 0, on déduit du théorème précédent la convergence dans $\mathscr{D}(\mathbb{R}^d)$. ■

Lemme 7.1.8. *Si A est une partie convexe fermée symétrique absorbante de $\mathscr{D}(\Omega)$, alors A est un voisinage de 0 dans $\mathscr{D}(\Omega)$.*

Démonstration : Cet énoncé est une généralisation du lemme 2.4.1, et se démontre de façon semblable. Par définition de la topologie de $\mathscr{D}(\Omega)$, il suffit de montrer que, pour tout compact K de Ω, $A_K := A \cap \mathscr{D}_K(\Omega)$ est un voisinage de 0 dans l'espace de Fréchet $\mathscr{D}_K(\Omega)$. Puisque $\mathscr{D}_K(\Omega) = \bigcup_{n\in\mathbb{N}} nA_K$, il résulte du théorème de Baire que l'un des fermés nA_K n'est pas rare, donc contient un point intérieur φ ; et il existe alors un voisinage convexe symétrique W de 0 dans $\mathscr{D}_K(\Omega)$ tel que $\varphi + W \subset nA$. Puisque nA est convexe et symétrique, on a alors $-\varphi + W \subset nA_K$, donc $\dfrac{1}{n}W \subset \dfrac{1}{2n}\big((\varphi+W)+(-\varphi+W)\big) \subset \dfrac{1}{2n}(nA_K + nA_K) = A_K$, ce qui montre que $A \cap \mathscr{D}_K(\Omega)$ est un voisinage de 0 dans $\mathscr{D}_K(\Omega)$, puisque $\dfrac{1}{n}W$ l'est. ■

Théorème 7.1.9. *Si u est une application linéaire de $\mathscr{D}(\Omega)$ dans $\mathscr{D}(\Omega')$ et si le graphe de u est fermé dans $\mathscr{D}(\Omega) \times \mathscr{D}(\Omega')$, alors u est continue.*

Démonstration : Ceci se déduit immédiatement du théorème 4.8.5, compte tenu de la définition 7.1.1 de la topologie de $\mathscr{D}(\Omega)$ comme limite inductive. ■

7.2 Partitions différentiables de l'unité

On va montrer dans cette section un analogue du théorème 1.7.4 en construisant des partitions de l'unité en fonctions de classe $\mathscr{C}^\infty$.

Théorème 7.2.1. *Soient K un compact de Ω et $(V_j)_{1\leqslant j\leqslant m}$ un recouvrement ouvert fini de K. Alors il existe des φ_j positives dans $\mathscr{D}(\Omega)$ telles que* $\operatorname{supp}(\varphi_j) \in V_j$, $\sum_1^m \varphi_j \leqslant 1$ *et* $\sum_1^m \varphi_j = 1$ *au voisinage de K.*

Démonstration : On a montré en 6.1.7 que si V est un ouvert de $\mathbb{R}^d$ et K un compact de V, il existe une fonction positive $\varphi \in \mathscr{D}(\mathbb{R}^d)$, strictement positive sur K et nulle hors de V. Quitte à diminuer les V_j, on peut supposer que $V_j \subset \Omega$ pour tout j. La fonction $h : x \mapsto \sup_{j\leqslant m} d(x, V_j^c)$ est continue sur $\mathbb{R}^d$ et strictement positive sur K. Elle y atteint donc un minimum $\eta > 0$.
On pose alors $K_j = \{x \in K : d(x, V_j^c) \geqslant \eta\}$. Chaque K_j est une partie compacte de K contenue dans $V_j' = \{x : d(x, V_j^c) > \eta/2\}$, et on a $K = \bigcup_j K_j$. On choisit pour chaque j une fonction $\psi_j \geqslant 0$ dans $\mathscr{D}(\mathbb{R}^d)$ nulle hors de V_j' et strictement positive sur K_j. Alors la fonction $\psi = \sum_{1\leqslant j\leqslant m} \psi_j$ est strictement positive sur K, ce qui montre que $\theta = \inf_{x\in K} \psi(x) > 0$. Et $K_0 = \{x \in \bigcup_{1\leqslant j\leqslant m} \overline{V_j'} : \psi(x) \leqslant \theta/3\}$ est compact disjoint de K et contenu dans l'ouvert $V_0 = \{x \in \Omega : \psi(x) < 2\theta/3\}$. Il existe donc une fonction positive $\psi_0 \in \mathscr{D}(\mathbb{R}^d)$, nulle hors de V_0 et strictement positive sur K_0. Alors $\psi + \psi_0$ est $\mathscr{C}^\infty$ et strictement positive sur $\bigcup_{1\leqslant j\leqslant m} \overline{V_j'}$. On en déduit que, pour $1 \leqslant j \leqslant m$, la fonction $\varphi_j = \dfrac{\psi_j}{\psi + \psi_0}$ est $\mathscr{C}^\infty$, à support dans $\overline{V_j'} \subset V_j \subset \Omega$, donc que $\varphi_j \in \mathscr{D}(\Omega)$, et enfin que, sur le voisinage $\{x : \psi(x) > 2\theta/3\}$ de K, on a $\psi_0(x) = 0$, donc $\sum_{1\leqslant j\leqslant m} \varphi_j(x) = 1$. ■

Théorème 7.2.2. *Soit $(V_j)_{j\in J}$ un recouvrement de Ω par des ouverts relativement compacts. Il existe une famille localement finie $(\varphi_j)_{j\in J}$ de fonctions positives dans $\mathscr{D}(\Omega)$ vérifiant* $\operatorname{supp}(\varphi_j) \subset V_j$ *pour tout j et* $\sum_{j\in J} \varphi_j = 1$.

Une telle famille est une partition de l'unité subordonnée au recouvrement $(V_j)_{j\in J}$ (cf. 1.7.2), formée de fonctions de classe $\mathscr{C}^\infty$.

Démonstration : Puisque Ω est réunion d'une suite (T_n) de compacts, et que chacun d'entre eux est recouvert par une sous-famille finie $(V_j)_{j\in J_n}$, on peut remplacer J par $\bigcup_n J_n$, c'est-à-dire supposer que J est dénombrable, et en choisir une énumération $j \mapsto n_j$.
On définit alors une fonction 1- lipschitzienne, h, sur Ω en posant $h(x) = \sup_j h_j(x)$ où $h_j(x) = \inf(2^{-n_j}, d(x, V_j^c))$. Puisque $(V_j)_{j\in J}$ recouvre Ω, on a $h(x) > 0$ en tout point $x \in \Omega$. Alors, pour tout j, l'ensemble $W_j = \{x : h_j(x) > h(x)/2\}$ est un ouvert contenu dans V_j. Et l'ensemble $K_j = \{x : h_j(x) = h(x)\}$ est un fermé contenu dans $W_j \subset V_j \subset \overline{V_j}$, donc compact.
Pour tout x et tout $\varepsilon > 0$, l'ensemble $\{j : h_j(x) \geqslant \varepsilon\}$ est fini. Il en résulte qu'il existe un j tel que $h(x) = h_j(x)$, donc que $x \in K_j$. D'autre part si m est choisi tel que $2^{-m} < h(x)$, alors le voisinage $W = \{x, h(x) > 2^{-m}\}$ de x ne peut rencontrer W_j en un point y que si $2^{-m-1} < h(y)/2 < h_j(y) \leqslant 2^{-n_j}$, c'est-à-dire si $n_j < m+1$. On en déduit que chaque point x de Ω possède un voisinage ne rencontrant qu'un nombre fini de W_j.
On peut alors pour tout j trouver une fonction positive $\psi_j \in \mathscr{D}(\Omega)$ nulle hors de W_j et strictement positive sur K_j. On a $\operatorname{supp}(\psi_j) \subset \overline{W_j} \subset V_j$. La somme $\psi = \sum_j \psi_j$ est localement finie, donc est $\mathscr{C}^\infty$ sur Ω et partout non nulle puisque $\bigcup_j K_j = \Omega$. Il suffit alors de poser $\varphi_j = \dfrac{\psi_j}{\psi}$, pour obtenir la famille cherchée. ∎

7.3 Distributions sur un ouvert

Définition 7.3.1. *On appelle* distribution *sur Ω toute forme linéaire continue sur $\mathscr{D}(\Omega)$.*

Théorème 7.3.2. *Soit T une forme linéaire sur $\mathscr{D}(\Omega)$. Il y a équivalence entre*

(i) T est continue.

(ii) T est bornée sur un voisinage de 0 dans $\mathscr{D}(\Omega)$

(iii) pour toute suite (φ_n) convergeant dans $\mathscr{D}(\Omega)$ vers φ, on a $\langle T, \varphi_n\rangle \to \langle T, \varphi\rangle$.

(iv) pour tout compact $K \subset \Omega$, il existe un $m \in \mathbb{N}$ et une constante C tels que $|\langle T, \varphi\rangle| \leqslant C \sup_{|\alpha|\leqslant m} \sup_{x\in K} |\partial_\alpha \varphi(x)|$ pour toute fonction de test φ à support dans K.

(v) T est bornée sur chaque partie bornée de $\mathscr{D}(\Omega)$.

Démonstration : Il est clair que $(i) \Longrightarrow (iii)$. Par définition de la topologie de $\mathscr{D}(\Omega)$, T est continue si et seulement si sa restriction à $\mathscr{D}_K(\Omega)$ est continue pour tout compact $K \subset \Omega$, et la continuité de $T_{|\mathscr{D}_K(\Omega)}$ s'exprime par la condition (iv), d'où l'équivalence de (i) et (iv).
Si T est continue, $W := \{\varphi \in \mathscr{D}(\Omega) : |\langle T, \varphi\rangle| < 1\}$ est un voisinage de 0 sur lequel T est bornée par 1. Et si T est bornée par M sur un voisinage W de 0, on a $|\langle T, \psi\rangle - \langle T, \varphi\rangle| \leqslant M\varepsilon$ pour tout ψ dans le voisinage $\varphi + \varepsilon W$ de φ, ce qui prouve l'équivalence de (i) et (ii).
Si on a (iii), la restriction de T à chaque espace de Fréchet $\mathscr{D}_K(\Omega)$ est continue, puisque $\mathscr{D}_K(\Omega)$ est métrisable. Et ceci suffit pour assurer la continuité de T sur $\mathscr{D}(\Omega)$. Donc $(iii) \Longrightarrow (i)$.
Si B était un borné de $\mathscr{D}(\Omega)$ sur lequel T n'est pas bornée, on pourrait trouver $\varphi_n \in B$ telle que $|\langle T, \varphi_n\rangle| \geqslant 2^{2n}$. Alors la suite $(2^{-n}\varphi_n)$ convergerait vers 0, alors que $|\langle T, 2^{-n}\varphi_n\rangle| \to \infty$. Donc $(iii) \Longrightarrow (v)$.

Enfin, si (iii) n'est pas vérifié, il existe une suite (φ_n) dans $\mathscr{D}(\Omega)$, convergeant vers φ et $\eta > 0$ tels que $|\langle T', \varphi_n - \varphi\rangle| \geqslant \eta$. Les φ_n ont toutes leur support dans un même compact K. Si q est une distance définissant la topologie de l'espace de Fréchet $\mathscr{D}_K(\Omega)$, la suite $(k(\varphi_n - \varphi))_n$ converge vers 0 pour tout entier k. On peut donc trouver une suite croissante n_k d'entiers telle que $q(0, k(\varphi_n - \varphi)) \leqslant 2^{-k}$ pour $n \geqslant n_k$. On définit alors la suite (ε_n) par $\varepsilon_n = \dfrac{1}{k}$ si $n_k \leqslant n < n_{k+1}$. Il est clair que $\varepsilon_n \to 0$ et que $q(0, \psi_n) \leqslant 2^{-k}$, si $n \geqslant n_k$ et $\psi_n = \dfrac{1}{\varepsilon_n}(\varphi_n - \varphi)$. Donc la suite (ψ_n) converge vers 0, ce qui montre que $B = \{\psi_n : n \in \mathbb{N}\}$ est borné, alors que $|\langle T, \psi_n\rangle| \geqslant \dfrac{\eta}{\varepsilon_n}$, c'est-à-dire $\sup_{\psi \in B} |\langle T, \psi\rangle| = +\infty$. Et $(v) \Longrightarrow (iii)$. ∎

On munit l'espace $\mathscr{D}'(\Omega)$ des distributions sur Ω de la topologie faible $\sigma(\mathscr{D}', \mathscr{D})$.

Proposition 7.3.3. *Une suite (T_n) de distributions sur Ω converge vers une distribution T si $\langle T_n, \varphi\rangle \to \langle T, \varphi\rangle$ pour tout φ dans $\mathscr{D}(\Omega)$.*

Démonstration : Ceci découle immédiatement de la définition de la topologie faible. ∎

Théorème 7.3.4. *Soit (T_n) une suite de distributions sur Ω. On suppose que, pour toute $\varphi \in \mathscr{D}(\Omega)$, la limite $S(\varphi) := \lim_n \langle T_n, \varphi\rangle$ existe. Alors S est une distribution.*

Démonstration : L'ensemble $A = \{\varphi \in \mathscr{D}(\Omega) : \sup_n |\langle T_n, \varphi\rangle| \leqslant 1\}$ est convexe fermé équilibré et absorbant, puisque $\sup_n |\langle T_n, \varphi\rangle| < +\infty$ pour tout φ. Il résulte du théorème 7.1.8 que A est un voisinage de 0 dans $\mathscr{D}(\Omega)$ et que la forme linéaire S est bornée sur ce voisinage, donc continue. ∎

Théorème 7.3.5. *(Restriction d'une distribution) Soient Ω un ouvert de $\mathbb{R}^d$ et Ω' un sous-ouvert de Ω. Si T est une distribution sur Ω, la restriction de T à $\mathscr{D}(\Omega') \subset \mathscr{D}(\Omega)$ est une distribution sur Ω', qu'on appelle la* restriction de T à Ω.

Démonstration : Si (φ_n) est une suite qui converge vers 0 dans $\mathscr{D}(\Omega')$, les (φ_n) ont toutes leur support dans un même compact $K \subset \Omega' \subset \Omega$ et les dérivées partielles $\partial_\alpha \varphi_n$ convergent uniformément vers $\partial_\alpha \varphi$ pour tout $\alpha \in \mathbb{N}^d$. Il en résulte que (φ_n) tend vers φ dans $\mathscr{D}(\Omega)$, donc que $\langle T, \varphi_n\rangle \to \langle T, \varphi\rangle$. Ceci montre la continuité de la restriction de T à Ω. ∎

Définition 7.3.6. *On dit qu'une distribution $T \in \mathscr{D}'(\Omega)$ est* d'ordre m *si, pour tout compact $K \subset \Omega$, il existe une constante C_K telle que, pour toute $\varphi \in \mathscr{D}(\Omega)$ à support dans K, on ait $|\langle T, \varphi\rangle| \leqslant C_K \sup_{|\alpha| \leqslant m} p_{\alpha,K}(\varphi)$.*
Il est équivalent de dire que T se prolonge en une forme linéaire continue sur l'espace $\mathscr{C}_c^m(\Omega)$ des fonctions de classe $\mathscr{C}^m$ à support compact sur Ω, muni de la topologie limite inductive des $\mathscr{C}_K^m(\Omega)$.

Démonstration : Si T est la restriction à $\mathscr{D}(\Omega)$ d'une forme linéaire continue sur $\mathscr{C}_c^m(\Omega)$, elle vérifie, par définition de la continuité, une inégalité du type recherché. Inversement, par densité de $\mathscr{D}_K(\Omega)$ dans $\mathscr{C}_K^m(\Omega)$, toute forme linéaire T sur $\mathscr{D}(\Omega)$ qui vérifie l'inégalité ci-dessus se prolonge de manière unique par continuité en une forme linéaire continue Θ_K sur $\mathscr{C}_K^m(\Omega)$; alors, par définition de la topologie limite inductive, il existe une forme linéaire continue Θ sur $\mathscr{C}_c^m(\Omega)$ telle que $\Theta_{|\mathscr{C}_K^m(\Omega)} = \Theta_K$, donc que $\Theta_{|\mathscr{D}(\Omega)} = T$. ∎

7.4 Exemples de distributions

FONCTIONS

Théorème 7.4.1. *Soit f une fonction localement intégrable sur Ω. Alors l'application $T_f : \varphi \mapsto \int f(x)\varphi(x)\,d\mu(x)$ est une distribution sur Ω. On dit qu'une distribution T est* une fonction *s'il existe une $f \in L^1_{\text{loc}}(\Omega)$ telle que $T = T_f$.*

DÉMONSTRATION : Si le support de φ est contenu dans le compact $K \subset \Omega$, on sait que $f.\mathbb{1}_K \in L^1$ donc que $M_K = \int_K |f(x)|\,d\mu(x) < \infty$.
Alors $|\langle T_f, \varphi\rangle| \leqslant M_K \sup|\varphi(x)| = M_K.p_0(\varphi)$, ce qui montre la continuité de T_f. Cette inégalité montre aussi que T_f est d'ordre 0. ■

Théorème 7.4.2. *L'application $f \mapsto T_f$ est injective et d'image dense de $L^1_{\text{loc}}(\Omega)$ dans $\mathscr{D}'(\Omega)$.*

DÉMONSTRATION : Si cette application n'était pas injective, il existerait f non nulle dans $L^1_{\text{loc}}(\Omega)$ telle que $\int f(x)\varphi(x)\,d\mu(x) = 0$ pour toute fonction $\varphi \in \mathscr{D}(\Omega)$. Pour tout compact K de Ω et tout voisinage relativement compact W de K dans Ω, on sait trouver une suite (φ_n) de fonctions $\mathscr{C}^\infty$ à valeurs dans $[0,1]$ nulles hors de W et convergeant simplement vers $\mathbb{1}_K$. Alors, puisque $\mathbb{1}_W.f \in L^1$, on déduit du théorème de convergence dominée que$\int_K f(x)\,d\mu(x) = \lim \int f(x)\varphi_n(x)\,d\mu(x) = 0$. Et ceci entraîne que $f = 0$ presque partout.
Si le sous-espace $\{T_f : f \in L^1_{\text{loc}}(\Omega)\}$ n'était pas dense dans $\mathscr{D}'(\Omega)$, il existerait, par Hahn-Banach, une forme linéaire continue sur $\mathscr{D}'(\Omega)$ non identiquement nulle s'annulant sur toutes les T_f, donc une fonction de test φ non nulle telle que $\int f(x)\varphi(x)\,d\mu(x) = 0$ pour toute $f \in L^1_{\text{loc}}(\Omega)$. Mais la fonction $\bar\varphi$ serait alors dans L^1_{loc} ; on devrait donc avoir $\int |\varphi(x)|^2\,d\mu(x) = 0$, c'est-à-dire $\varphi = 0$ presque partout sur Ω, donc $\varphi = 0$ par continuité. Et cette contradiction achève la preuve. ■

MESURES DE DIRAC

Pour tout point $a \in \Omega$, l'application $\delta_a : \varphi \mapsto \varphi(a)$ est une forme linéaire sur $\mathscr{D}(\Omega)$ telle que $|\langle \delta_a, \varphi\rangle| \leqslant p_0(\varphi) = \sup_x |\varphi(x)|$. C'est donc une distribution d'ordre 0, appelée *mesure de Dirac en a*.

VALEUR PRINCIPALE DE $1/x$

Théorème 7.4.3. *Si $\varphi \in \mathscr{D}(\mathbb{R})$, l'intégrale $\int_{|x|\geqslant\varepsilon} \varphi(x)\dfrac{dx}{x}$ possède une limite lorsque ε tend vers 0. Et l'application $\text{vp}(\dfrac{1}{x}) : \varphi \mapsto \lim\limits_{\varepsilon\to 0}\displaystyle\int_{|x|\geqslant\varepsilon} \varphi(x)\frac{dx}{x}$ est une distribution appelée* valeur principale de $\dfrac{1}{x}$.

DÉMONSTRATION : La fonction g, définie par $g(0) = \varphi'(0)$ et $g(x) = \dfrac{\varphi(x) - \varphi(0)}{x}$ sinon, est continue. On a donc, si φ est nulle hors de $[-R, R]$,

$$\int_{|x|\geqslant\varepsilon} \varphi(x)\frac{dx}{x} = \int_{\varepsilon\leqslant|x|\leqslant R} g(x)\,dx + \varphi(0)\int_{\varepsilon\leqslant|x|\leqslant R} \frac{dx}{x} = \int_{\varepsilon\leqslant|x|\leqslant R} g(x)\,dx \to \int_{-R}^{R} g(x)\,dx$$

d'où l'existence de la limite. Puisque $|g(x)| \leqslant \sup_t |\varphi'(t)|$, on a $\left|\langle \text{vp}(\dfrac{1}{x}), \varphi\rangle\right| \leqslant 2R\sup_x |\varphi'(x)|$, ce qui montre que $\text{vp}(\dfrac{1}{x})$ est une distribution d'ordre 1 sur $\mathbb{R}$. ■

7.5 Produit d'une distribution par une fonction lisse

Si f est une fonction localement intégrable sur Ω et si $\psi \in \mathscr{C}^\infty(\Omega)$, la fonction $g = f\psi$ est encore localement intégrable sur Ω puisque ψ y est localement bornée. Il en résulte que les distributions associées T_f et T_g sont liées par la relation

$$\langle T_g, \varphi\rangle = \int \big(f(x)\psi(x)\big)\varphi(x)\,d\mu(x) = \int f(x)\big(\psi(x)\varphi(x)\big)\,d\mu(x) = \langle T_f, \psi\varphi\rangle$$

qui a un sens puisque $\psi\varphi$ est $\mathscr{C}^\infty$ avec un support contenu dans celui de φ.
D'une façon générale, on va définir le produit $\psi.T$ d'une fonction $\mathscr{C}^\infty$ et d'une distribution T par une formule analogue.

Théorème 7.5.1. *Soient $\psi \in \mathscr{C}^\infty(\Omega)$ et $T \in \mathscr{D}'(\Omega)$, il existe une distribution S, notée $\psi.T$ et appelée* produit de ψ et de T *telle que, pour tout $\varphi \in \mathscr{D}(\Omega)$, on ait $\langle S, \varphi\rangle = \langle T, \psi\varphi\rangle$. De plus, l'application $T \mapsto \psi.T$ est continue de $\mathscr{D}'(\Omega)$ dans $\mathscr{D}'(\Omega)$.*

Démonstration : Puisque $\psi\varphi$ est $\mathscr{C}^\infty$, avec support contenu dans le compact $\operatorname{supp}(\varphi)$, la formule ci-dessus définit bien une forme linéaire sur $\mathscr{D}(\Omega)$. Pour voir que S est une distribution, il faut montrer sa continuité. L'application linéaire $M_\psi : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega)$ définie par $M_\psi(\varphi) = \psi\varphi$ a clairement un graphe fermé, puisque

$$\varphi_1 = M_\psi(\varphi) \iff \forall x \in \Omega \quad \varphi_1(x) = \psi(x)\varphi(x)$$

On en déduit la continuité de M_ψ. Il en résulte que $S = {}^tM_\psi(T) \in \mathscr{D}'(\Omega)$ et que ${}^tM_\psi$ est continue de $\mathscr{D}'(\Omega)$ dans lui-même. ∎

Exemples.

Calculons le produit $x.\delta_0$ de la fonction $\mathscr{C}^\infty$ sur $\mathbb{R}$: $x \mapsto x$ et de la mesure de Dirac à l'origine. Si $\varphi \in \mathscr{D}(\mathbb{R})$, on a $\langle x.\delta_0, \varphi\rangle = \langle \delta_0, x.\varphi\rangle = 0$ puisque la fonction $x\varphi$ s'annule en 0, c'est-à-dire $x.\delta_0 = 0$.

De même, pour évaluer le produit $x.\operatorname{vp}(\frac{1}{x})$, prenons $\varphi \in \mathscr{D}(\mathbb{R})$: on a alors

$$\begin{aligned}\langle x.\operatorname{vp}(\frac{1}{x}), \varphi\rangle = \langle \operatorname{vp}(\frac{1}{x}), x\varphi\rangle &= \lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon} x\varphi(x)\,\frac{dx}{x}\\ &= \lim_{\varepsilon\to 0}\int_{|x|\geqslant\varepsilon}\varphi(x)\,dx = \int \varphi(x)\,dx = \langle \mathbb{1}, \varphi\rangle\end{aligned}$$

ce qui montre que le produit $x\operatorname{vp}(\frac{1}{x})$ est la fonction constante, donc L^1_{loc}, $\mathbb{1}$. On écrira $x.\operatorname{vp}(\frac{1}{x}) = 1$.

7.6 Support d'une distribution

Théorème 7.6.1. *Si T est une distribution sur l'ouvert Ω de $\mathbb{R}^d$, il existe un plus grand ouvert U de Ω tel que la restriction de T à U soit nulle. Le complémentaire de U dans Ω est un fermé de Ω, non vide si $T \neq 0$, appelé le* support de T.

Démonstration : Si on note U l'ensemble des points $a \in \Omega$ possédant un voisinage ouvert V_a sur lequel la restriction de T est nulle, on va montrer que $T_{|U} = 0$. En effet, U est ouvert et, si $\varphi \in \mathscr{D}(\Omega)$ a son support K contenu dans U, K posède un recouvrement fini par des ouverts $(V_j)_{1\leqslant j\leqslant m}$ tels que $T_{|V_j} = 0$. Il existe alors, d'après le théorème 7.2.1, des fonctions positives $(\varphi_j)_{1\leqslant j\leqslant m}$ dans $\mathscr{D}(\Omega)$, avec : $\text{supp}(\varphi_j) \subset V_j$ et $\sum_1^m \varphi_j = 1$ au voisinage de K. Alors, pour tout j, on a : $\varphi_j.T = 0$. En effet, pour tout $\psi \in \mathscr{D}(\Omega)$, $\varphi_j.\psi \in \mathscr{D}(V_j)$, donc $\langle \varphi_j.T, \psi\rangle = \langle T, \varphi_j.\psi\rangle = \langle T_{|V_j}, \varphi_j.\psi\rangle = 0$. Il en résulte, puisque $(\sum_j \varphi_j).\varphi = \varphi$, que

$$\langle T, \varphi\rangle = \langle T, (\sum_j \varphi_j).\varphi\rangle = \langle (\sum\nolimits_j \varphi_j).T, \varphi\rangle = \sum\nolimits_j \langle \varphi_j.T, \varphi\rangle = 0 \ ,$$

ce qui montre que $T_{|U} = 0$. Enfin si $U = \Omega$, on a $T_{|\Omega} = T = 0$. ■

Théorème 7.6.2. *Si $(\Omega_j)_{j\in J}$ est un recouvrement ouvert de l'ouvert $\Omega \subset \mathbb{R}^d$ et si $(T_j)_{j\in J}$ est une famille de distributions avec $T_j \in \mathscr{D}'(\Omega_j)$ et $T_{j|\Omega_j\cap\Omega_k} = T_{k|\Omega_j\cap\Omega_k}$ pour j et k dans J, alors il existe une unique distribution $T \in \mathscr{D}'(\Omega)$ telle que $T_j = T_{|\Omega_j}$.*

Démonstration : L'unicité résulte du théorème précédent : en effet, si S et T vérifient : $T_{|\Omega_j} = T_j = S_{|\Omega_j}$ pour tout $j \in J$, on a $(S-T)_{|\Omega_j} = 0$ pour tout j, donc $\text{supp}(S-T)\cap\Omega_j = \emptyset$ pour tout j, et $\text{supp}(S-T) = \emptyset$, donc $S - T = 0$.

Inversement, on peut se ramener au cas où J est dénombrable, et où les (Ω_j) sont relativement compacts dans Ω. Il résulte du théorème 7.2.2 qu'existe alors une partition $\mathscr{C}^\infty$ localement finie, $(\varphi_j)_{j\in J}$, subordonnée au recouvrement $(\Omega_j)_{j\in J}$. On définit alors $T = \sum_{j\in J} \varphi_j.T_j$. On remarque d'abord que, pour tout $\varphi \in \mathscr{D}(\Omega)$, il n'existe qu'un nombre fini de $j \in J$ tels que $\varphi_j.\varphi \neq 0$, donc que la somme $\sum_j \langle \varphi_j.T, \varphi\rangle$ a un sens pour tout $\varphi \in \mathscr{D}(\Omega)$. Si on fixe une énumération $n \mapsto j_n$ de J, on a pour tout $\varphi \in \mathscr{D}(\Omega)$: $\langle T, \varphi\rangle = \lim_{m\to\infty} \sum_{k=0}^m \langle \varphi_{j_k}.T, \varphi\rangle$. On déduit alors du théorème 7.1.8 que T est une distribution.

Il reste à montrer que $T_{|\Omega_j} = T_j$. Soit alors $\varphi \in \mathscr{D}(\Omega_j)$. On a $\langle T, \varphi\rangle = \sum_{k\in J} \langle T_k, \varphi_k.\varphi\rangle$, et puisque $\text{supp}(\varphi_k.\varphi) \subset \Omega_j \cap \Omega_k$, on a

$$\langle T_k, \varphi_k.\varphi\rangle = \langle T_{k|\Omega_j\cap\Omega_k}, \varphi_k.\varphi\rangle = \langle T_{j|\Omega_j\cap\Omega_k}, \varphi_k.\varphi\rangle = \langle T_j, \varphi_k.\varphi\rangle$$

et $\langle T, \varphi\rangle = \sum_{k\in J} \langle T_j, \varphi_k.\varphi\rangle = \langle T_j, (\sum_k \varphi_k)\varphi\rangle = \langle T_j, \varphi\rangle$. Donc $T_{|\Omega_j} = T_j$. ■

Distributions à support compact.

Théorème 7.6.3. *Si T est une distribution sur Ω à support compact, T se prolonge de façon unique en une forme linéaire continue sur l'espace de Fréchet $\mathscr{E}(\Omega) = \mathscr{C}^\infty(\Omega)$.*

Inversement, la restriction à $\mathscr{D}(\Omega)$ de toute forme linéaire continue sur $\mathscr{E}(\Omega)$ est une distribution sur Ω dont le support est compact. On identifiera donc le dual $\mathscr{E}'(\Omega)$ de $\mathscr{E}(\Omega)$ à l'espace des distributions à support compact.

Démonstration : Soit $T \in \mathscr{D}'(\Omega)$ dont le support est un compact K de Ω. Il existe une fonction $\psi \in \mathscr{D}(\Omega)$ égale à 1 au voisinage de K. Alors, pour toute fonction $\varphi \in \mathscr{E}(\Omega)$, la fonction $\psi\varphi$ est dans $\mathscr{D}(\Omega)$, et on peut définir sur $\mathscr{E}(\Omega)$ une forme linéaire Θ par $\langle\Theta, \varphi\rangle = \langle T, \psi.\varphi\rangle$. L'application $M_\psi : \mathscr{E}(\Omega) \to \mathscr{D}(\Omega)$ définie par $M_\psi(\varphi) = \psi.\varphi$ a clairement un graphe fermé, donc est continue, ce qui montre que $\Theta = T \circ M_\psi$ est une forme linéaire continue sur $\mathscr{E}(\Omega)$. La topologie de $\mathscr{E}(\Omega)$ est définie par la famille P des semi-normes $p_{\alpha,K} : \varphi \mapsto \sup_{x\in K} |\partial_\alpha\varphi(x)|$. On en déduit que toute P-boule B satisfait la propriété : "*il existe un compact $L \subset \Omega$ tel que si $\varphi \in B$ et $\psi_{|L} = \varphi_{|L}$, on a $\psi \in B$* " dont on déduit la densité de $\mathscr{D}(\Omega)$ dans $\mathscr{E}(\Omega)$ puisque, en multipliant une fonction dans B par une fonction de $\mathscr{D}(\Omega)$ égale à 1 sur L, on obtient une fonction de $B \cap \mathscr{D}(\Omega)$. Et cette densité assure l'unicité du prolongement de T à $\mathscr{E}(\Omega)$.

Inversement, si Θ est une forme linéaire continue sur $\mathscr{E}(\Omega)$, il existe une P-boule sur laquelle Θ est bornée, donc un compact K, un entier m et une constante M tels que, pour $\varphi \in \mathscr{E}(\Omega)$, on ait $|\langle\Theta, \varphi\rangle| \leqslant M \sup_{|\alpha|\leqslant m} p_{\alpha,K}(\varphi)$, Il en résulte que la restriction de Θ à $\mathscr{D}(\Omega)$ est une distribution T d'ordre $\leqslant m$ et que $\langle\Theta, \varphi\rangle = 0$ si $\mathrm{supp}(\varphi) \cap K = \emptyset$. Donc $T_{|\Omega\setminus K} = 0$ et $\mathrm{supp}(T) \subset K$. ■

Corollaire 7.6.4. *Toute distribution à support compact est d'ordre fini.*

Démonstration : Ceci résulte immédiatement de ce qui précède. ■

Théorème 7.6.5. *Soient T une distribution d'ordre m sur l'ouvert $\Omega \subset \mathbb{R}^d$ et $\varphi \in \mathscr{D}(\Omega)$ nulle ainsi que toutes ses dérivées partielles d'ordre $\leqslant m$ sur le support de T. Alors $\langle T, \varphi\rangle = 0$.*

Démonstration : Soit $L \subset \Omega$ un voisinage compact du support de φ. On définit $\eta = \inf_{x\in L} d(x, \Omega^c) > 0$, $K = L \cap \mathrm{supp}(T)$ et $K_\varepsilon = \{x : d(x, K) \leqslant 2\varepsilon\}$. Alors K_ε est compact dans Ω si $\varepsilon < \eta/3$. On choisit une fonction positive $\rho \in \mathscr{D}(\mathbb{R}^d)$ à support dans la boule unité et d'intégrale 1, et, pour $\varepsilon < \eta/3$, la fonction $\psi^{(\varepsilon)} = \mathbb{1}_{K_\varepsilon} * \rho_\varepsilon$ appartient à $\mathscr{D}(\Omega)$, est à valeurs dans $[0, 1]$ et vaut 1 sur $\{x : d(x, K) \leqslant \varepsilon\}$.
Pour $\alpha \in \mathbb{N}^d$, on a

$$\left|\partial_\alpha\psi^{(\varepsilon)}(x)\right| = |\mathbb{1}_{K_\varepsilon} * \partial_\alpha\rho_\varepsilon(x)| \leqslant \mu(K_\varepsilon).\sup_y |\partial_\alpha\rho_\varepsilon(y)| = \varepsilon^{-|\alpha|}\mu(K_\varepsilon).\sup_y |\partial_\alpha\rho(y)| \leqslant C_\alpha\varepsilon^{-|\alpha|}$$

La fonction $\varphi.(1 - \psi^{(\varepsilon)})$ a son support contenu dans L et disjoint de K, donc disjoint de $\mathrm{supp}(T)$. Par définition du support de T, on a donc $\langle T, \varphi.(1 - \psi^{(\varepsilon)})\rangle = 0$, donc $\langle T, \varphi\rangle = \langle T, \varphi.\psi^{(\varepsilon)}\rangle$. Puisque T est d'ordre m et que $\mathrm{supp}(\varphi.\psi^{(\varepsilon)}) \subset K$, il existe une constante M telle que $\left|\langle T, \varphi.\psi^{(\varepsilon)}\rangle\right| \leqslant M \sup_{|\alpha|\leqslant m} \sup_x \left|\partial_\alpha(\varphi.\psi^{(\varepsilon)})(x)\right|$. On va montrer que, pour $|\alpha| \leqslant m$, on a $\sup_x \left|\partial_\alpha(\varphi.\psi^{(\varepsilon)})(x)\right| = o(\varepsilon^{m-|\alpha|})$, ce qui montrera que $\langle T, \varphi\rangle = 0$.
Par continuité uniforme de la différentielle d'ordre m de φ, il existe pour tout $r > 0$, un $\eta' < \eta$ tel que $\left\|\varphi^{(m)}(x)\right\| < r$ si $d(x, K) < \eta'$. Il résulte alors de la formule de Taylor appliquée à φ en un point $a \in K$ tel que $\|x - a\| = d(x, K)$, que $|\varphi(x)| \leqslant \dfrac{r}{m!} \|x - a\|^m$. Le même raisonnement appliqué à $\partial_\beta\varphi$, dont les dérivées partielles s'annulent sur K jusqu'à l'ordre $m - |\beta|$, pour $|\beta| \leqslant m$, montre que $|\partial_\beta\varphi(x)| \leqslant \dfrac{r}{(m - |\beta|)!} \cdot d(x, K)^{m-|\beta|}$ si $d(x, K) < \eta'$.
Par la formule de Leibniz, il existe des constantes $c_{\alpha,\beta}$ telles que

$$\partial_\alpha(\varphi.\psi^{(\varepsilon)})(x) = \sum_{\beta\leqslant\alpha} c_{\alpha,\beta}\, \partial_\beta\varphi(x)\partial_{\alpha-\beta}\psi^{(\varepsilon)}(x)$$

donc, si $|\alpha| \leqslant m$ et $d(x,K) < \eta'$,

$$\begin{aligned}\left|\partial_\alpha(\varphi.\psi^{(\varepsilon)})(x)\right| &\leqslant \sum_{\beta\leqslant\alpha} c_{\alpha,\beta} C_{\alpha-\beta}\varepsilon^{|\beta|-|\alpha|} \sup_{d(x,K)\leqslant 3\varepsilon} |\partial_\beta\varphi(x)| \\ &\leqslant r \sum_{\beta\leqslant\alpha} c_{\alpha,\beta} C_{\alpha-\beta} \frac{\varepsilon^{|\beta|-|\alpha|}(3\varepsilon)^{m-|\beta|}}{(m-|\beta|)!}\end{aligned}$$

ce qui signifie que $\sup_x \left|\partial_\alpha(\varphi.\psi^{(\varepsilon)})(x)\right| = o(\varepsilon^{m-|\alpha|})$, et achève la démonstration. ■

Corollaire 7.6.6. *Soient T une distribution sur l'ouvert $\Omega \subset \mathbb{R}^d$ et $\varphi \in \mathscr{D}(\Omega)$ nulle sur le support de T, ainsi que ses dérivées partielles de tout ordre. Alors $\langle T, \varphi\rangle = 0$.*

Démonstration : Soient (K_n) une suite exhaustive de compacts de Ω, et, pour tout n, $\varphi_n \in \mathscr{D}(\Omega)$ égale à 1 au voisinage de K_n. Alors $\mathrm{supp}(\varphi_n.T) \subset K_n \cap \mathrm{supp}(T)$. Il en résulte que $\varphi_n.T$ est une distribution à support compact, donc d'ordre m_n fini, et que φ et toutes ses dérivées partielles d'ordre m_n sont nulles sur $\mathrm{supp}(\varphi_n.T)$. Il résulte du théorème précédent que $\langle \varphi_n.T, \varphi\rangle = 0$ pour tout n. Et puisqu'il existe un n tel que $\mathrm{supp}(\varphi) \subset K_n$, donc $\varphi_n\varphi = \varphi$, on a $\langle T, \varphi\rangle = (T, \varphi.\varphi_n\rangle = \langle \varphi_n.T, \varphi\rangle = 0$. ■

Support singulier

De façon analogue, on montre que si $T \in \mathscr{D}(\Omega)$, il existe un plus grand ouvert U tel que $T_{|U}$ soit une fonction $\mathscr{C}^\infty$. En effet, si $T_{|V}$ est une fonction $\mathscr{C}^\infty$ et si $\psi \in \mathscr{D}(V)$, le produit $\psi.T$ est, en tant que distribution sur Ω, une fonction $\mathscr{C}^\infty$. Et, si $(\psi_j)_{j\in J}$ est une partition $\mathscr{C}^\infty$ localement finie de l'unité sur U avec $\mathrm{supp}(\psi_j) \subset V_j$, et $T_{|V_j} \in \mathscr{C}^\infty$, on voit immédiatement que $T = \sum_j \psi_j.T$ est une fonction $\mathscr{C}^\infty$.
On définit donc le *support singulier* de T, $\mathrm{singsupp}(T)$, comme le fermé complémentaire dans Ω de cet ouvert U. Il est alors clair que la distribution T sur Ω est une fonction dans $\mathscr{C}^\infty(\Omega)$ si et seulement si son support singulier est vide.
A titre d'exemple, on vérifie immédiatement que, dans $\mathscr{D}(\mathbb{R})$, le support singulier de la mesure de Dirac δ_a est le singleton $\{a\}$, et que $\mathrm{singsupp}(\mathrm{vp}(\frac{1}{x})) = \{0\}$.

7.7 Dérivation des distributions

Si f est une fonction de classe $\mathscr{C}^1$ sur $\mathbb{R}^d$, $\varphi \in \mathscr{D}(\mathbb{R}^d)$ et $1 \leqslant k \leqslant d$, on a, pour tout $x' = (x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_d) \in \mathbb{R}^{d-1}$,

$$\int_{-\infty}^{+\infty} \frac{d}{dt}(f.\varphi)(x_1, x_2, \dots, x_{k-1}, t, x_{k+1}, \dots, x_d)\, dt = 0$$

donc

$$\int_{\mathbb{R}^{d-1}} \left(\int_{\mathbb{R}} (f\frac{\partial\varphi}{\partial x_k} + \varphi\frac{\partial f}{\partial x_k})(x_1, x_2, \dots, x_{k-1}, t, x_{k+1}, \dots, x_d)\, dt\right) dx' = 0$$

c'est-à-dire $\int(f\partial_k\varphi + \varphi\partial_k f)\, dx = 0$, ou encore que la distribution associée à la fonction $\partial_k f$ est la forme linéaire $\varphi \mapsto -\int f(x)\partial_k\varphi(x)\, dx$. On peut ainsi généraliser à toutes les distributions en définissant la dérivée partielle de T par rapport à la $k^{\text{ième}}$ variable par $\langle \partial_k T, \varphi\rangle = -\langle T, \partial_k\varphi\rangle$.

Théorème 7.7.1. *Pour tout $\alpha \in \mathbb{N}^d$ et toute distribution T sur l'ouvert Ω de $\mathbb{R}^d$, la forme linéaire $\varphi \mapsto (-1)^{|\alpha|}\langle T, \partial_\alpha \varphi\rangle$ sur $\mathscr{D}(\Omega)$ est une distribution, notée $\partial_\alpha T$. On a $\partial_\alpha(\partial_\beta T) = \partial_{\alpha+\beta} T$. De plus, l'application linéaire $D'_\alpha : T \mapsto \partial_\alpha T$ est continue de $\mathscr{D}'(\Omega)$ dans $\mathscr{D}'(\Omega)$.*

Démonstration : L'application D'_α est la transposée de l'application linéaire de dérivation $D_\alpha : \varphi \mapsto (-1)^{|\alpha|}\partial_\alpha \varphi$ de $\mathscr{D}(\Omega)$ dans $\mathscr{D}(\Omega)$. Il suffit donc, par composition, de montrer que chaque dérivation partielle ∂_k est linéaire continue. Et ceci résulte immédiatement de l'égalité : $p_{\alpha,K}(\partial_k \varphi) = p_{\alpha+\tau_k,K}(\varphi)$. Et puisque l'on a $\partial_\beta(\partial_\alpha \varphi) = \partial_{\alpha+\beta}\varphi$ pour toute $\varphi \in \mathscr{D}(\Omega)$, on vérifie immédiatement que $\partial_\alpha(\partial_\beta T) = \partial_{\alpha+\beta} T$. ■

Lemme 7.7.2. *Soit $\varphi \in \mathscr{D}(\mathbb{R}^d)$. Si $\int \varphi(x)\,dx = 0$, il existe d fonctions $(\varphi_1, \varphi_2, \ldots, \varphi_d)$ dans $\mathscr{D}(\mathbb{R}^d)$ telles que $\varphi = \sum_{j=1}^d \partial_j \varphi_j$.*

Démonstration : On va montrer ceci par récurrence sur d. Supposons d'abord $d = 1$ et soit $\varphi \in \mathscr{D}(\mathbb{R})$ telle que $\int \varphi(x)\,dx = 0$. Il en résulte que la primitive $\psi : x \mapsto \int_{-\infty}^x \varphi(t)\,dt$ est une fonction $\mathscr{C}^\infty$, localement constante hors de $\operatorname{supp}(\varphi)$, et nulle en $-\infty$ et en $+\infty$. On en déduit que $\psi \in \mathscr{D}(\mathbb{R})$ et que $\varphi = \psi'$.

Supposons maintenant le lemme vrai pour $d-1$ et soit $\varphi \in \mathscr{D}(\mathbb{R}^d)$ d'intégrale nulle. Si on pose $\Phi(x_1, x_2, \ldots, x_{d-1}) := \int_{-\infty}^{+\infty} \varphi(x_1, x_2, \ldots, x_{d-1}, t)\,dt$, on vérifie sans peine par dérivation sous le signe $\int$, que Φ est $\mathscr{C}^\infty$ sur $\mathbb{R}^{d-1}$. De plus Φ est nulle hors de la projection K sur $\mathbb{R}^{d-1}$ du support de φ, qui est un compact de $\mathbb{R}^{d-1}$. Il en résulte que $\Phi \in \mathscr{D}(\mathbb{R}^{d-1})$. De plus, par Fubini, $\int_{\mathbb{R}^{d-1}} \Phi(x)\,dx = \int_{\mathbb{R}^d} \varphi(y)\,dy = 0$. Par l'hypothèse de récurrence, il existe $d-1$ fonctions $(\psi_1, \psi_2, \ldots, \psi_{d-1})$ dans $\mathscr{D}(\mathbb{R}^{d-1})$ telles que $\Phi = \sum_{j=1}^{d-1} \partial_j \psi_j$.
On choisit alors une fonction $\rho \in \mathscr{D}(\mathbb{R})$ d'intégrale 1 et on pose

$$\Psi(x_1, x_2, \ldots, x_d) = \int_{-\infty}^{x_d} \Big(\varphi(x_1, x_2, \ldots, x_{d-1}, t) - \Phi(x_1, x_2, \ldots, x_{d-1})\Big)\rho(t)\,dt$$

et on vérifie aisément que Ψ est $\mathscr{C}^\infty$, nulle hors de $K \times \mathbb{R}$, et que $\Psi(x_1, x_2, \ldots, x_d) = 0$ si $x_d \notin H$, où H est le compact de $\mathbb{R}$ projection du support de φ. Donc $\Psi \in \mathscr{D}(\mathbb{R}^d)$ et on a

$$\partial_d \Psi(x_1, x_2, \ldots, x_d) = \varphi(x_1, x_2, \ldots, x_d) - \Phi(x_1, x_2, \ldots, x_{d-1})\rho(x_d)$$

donc $\varphi = \partial_d \Psi + \sum_{j=1}^{d-1} \partial_j(\psi_j \otimes \rho)$, ce qui achève la démonstration par récurrence puisque les fonctions $\psi_j \otimes \rho : (x_1, x_2, \ldots, x_d) \mapsto \psi_j(x_1, x_2, \ldots, x_{d-1}).\rho(x_d)$ sont dans $\mathscr{D}(\mathbb{R}^d)$. ■

Proposition 7.7.3. *Si T une distribution sur $\mathbb{R}^d$ telle que $\partial_k T = 0$ pour $k = 1, 2, \ldots, d$. alors T est une fonction constante.*

Démonstration : Soit ρ une fonction positive dans $\mathscr{D}(\mathbb{R}^d)$ d'intégrale 1. On pose $\lambda = \langle T, \rho\rangle$. Alors la fonction $\varphi_1 := \varphi - \langle \mathbb{1}, \varphi\rangle.\rho$ a une intégrale nulle. En vertu du lemme précédent, il existe des $\varphi_j \in \mathscr{D}(\mathbb{R}^d)$ telles que $\varphi_1 = \sum_{j=1}^d \partial_j \varphi_j$. On a alors

$$\langle T - \lambda.\mathbb{1}, \varphi\rangle = \langle T, \varphi\rangle - \langle \mathbb{1}, \varphi\rangle\langle T, \rho\rangle = \langle T, \varphi_1\rangle = \sum_{j=1}^d \langle T, \partial_j \varphi_j\rangle = \sum_{j=1}^d -\langle \partial_j T, \varphi_j\rangle = 0\ ,$$

d'où le résultat : $T = \lambda.\mathbb{1}$. ■

Proposition 7.7.4. *Si $T \in \mathscr{D}'(\Omega)$ est une distribution d'ordre m, alors $\partial_k T$ est d'ordre au plus $m+1$.*

Démonstration : S'il existe pour tout compact K de Ω une constante $C(K)$ telle que, pour toute $\varphi \in \mathscr{D}(\Omega)$ à support dans K, on ait $|\langle T, \varphi\rangle| \leqslant C(K).\sup_{|\alpha|\leqslant m} p_{\alpha,K}(\varphi)$, on a, lorsque $\operatorname{supp}(\varphi) \subset K$:

$$\begin{aligned}|\langle \partial_k T, \varphi\rangle| = |\langle T, \partial_k \varphi\rangle| &\leqslant C(K).\sup_{|\alpha|\leqslant m} p_{\alpha,K}(\partial_k\varphi) = C(K).\sup_{|\alpha|\leqslant m} p_{\alpha+\tau_k,K}(\varphi)\\ &\leqslant C(K).\sup_{|\alpha|\leqslant m+1} p_{\alpha,K}(\varphi)\end{aligned}$$

ce qui est le résultat cherché. ∎

Théorème 7.7.5. *(Dérivées partielles d'un produit) Soient Ω un ouvert de $\mathbb{R}^d$, $\varphi \in \mathscr{D}(\Omega)$ et T une distribution sur Ω. Alors, pour $1 \leqslant k \leqslant d$, on a $\partial_k(\varphi.T) = \partial_k\varphi.T + \varphi.\partial_k T$.*

Démonstration : En effet, si $\psi \in \mathscr{D}(\Omega)$, on a

$$\langle \partial_k(\varphi.T), \psi\rangle = -\langle \varphi.T, \partial_k\psi\rangle = -\langle T, \varphi.\partial_k\psi\rangle$$

et

$$\begin{aligned}\langle \varphi.\partial_k T, \psi\rangle = \langle \partial_k T, \varphi.\psi\rangle &= -\langle T, \partial_k(\varphi.\psi)\rangle = -\langle T, \varphi.\partial_k\psi\rangle - \langle T, \psi.\partial_k\varphi\rangle\\ &= -\langle T, \varphi.\partial_k\psi\rangle - \langle \partial_k\varphi.T, \psi\rangle = \langle \partial_k(\varphi.T), \psi\rangle - \langle \partial_k\varphi.T, \psi\rangle\end{aligned}$$

donc

$$\langle \partial_k\varphi.T + \varphi.\partial_k T, \psi\rangle = \langle \partial_k(\varphi.T), \psi\rangle$$

ce qui est l'égalité cherchée.
En itérant, on obtient une généralisation des formules de Leibniz pour le produit d'une fonction de test et d'une distribution. ∎

Dérivées de la mesure de Dirac.

Pour tout $\alpha \in \mathbb{N}^d$, la dérivée d'ordre α de la mesure de Dirac δ_a en $a \in \mathbb{R}^d$ est la distribution $\varphi \mapsto \langle \partial_\alpha\delta_a, \varphi\rangle = (-1)^{|\alpha|}\langle \delta_a, \partial_\alpha\varphi\rangle = (-1)^{|\alpha|}\,\partial_\alpha\varphi(a)$.

Théorème 7.7.6. *Dans $\mathscr{D}'(\mathbb{R})$, pour $\ell \leqslant m$, on a $x^\ell.\delta_0^{(m)} = (-1)^\ell \dfrac{m!}{(m-\ell)!}\delta_0^{(m-\ell)}$. Et pour tout $\ell > m$ on a $x^\ell.\delta_0^{(m)} = 0$.*

Démonstration : En effet, pour $\varphi \in \mathscr{D}(\mathbb{R})$, on a

$$\langle x^\ell.\delta_0^{(m)}, \varphi\rangle = \langle \delta_0^{(m)}, x^\ell\varphi\rangle = (-1)^m \frac{d^m}{dx^m}(x^\ell\varphi)(0) = (-1)^m \sum_{j=0}^{m} \mathrm{C}_m^j \frac{d^j}{dx^j}(x^\ell)(0)\varphi^{(m-j)}(0)$$

Le seul terme éventuellement non nul de cette somme est celui où $j = \ell$. On en déduit la nullité de $x^\ell.\delta_0^{(m)}$ lorsque $\ell > m$, et sinon $\langle x^\ell.\delta_0^{(m)}, \varphi\rangle = (-1)^m\mathrm{C}_m^\ell \ell!\langle (-1)^{m-\ell}\delta_0^{(m-\ell)}, \varphi\rangle$. ∎

Proposition 7.7.7. *Les distributions $(\partial_\alpha \delta_0)_{\alpha\in\mathbb{N}^d}$ sont linéairement indépendantes dans $\mathscr{D}'(\mathbb{R}^d)$.*

Démonstration : Si χ est une fonction de $\mathscr{D}(\mathbb{R}^d)$ égale à 1 au voisinage de 0, on définit, pour tout $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$, la fonction $\psi_\alpha \in \mathscr{D}(\mathbb{R}^d)$ par $\psi_\alpha(x) = \chi(x).\prod_{j=1}^d \dfrac{x_j^{\alpha_j}}{\alpha_j!}$. Alors on vérifie en utilisant le théorème précédent que $\langle \partial_\alpha\delta_0, \psi_\alpha\rangle = (-1)^{|\alpha|}$ et que $\langle \partial_\alpha\delta_0, \psi_\beta\rangle = 0$ si $\alpha \neq \beta$.
On en déduit que, si $\sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha \delta_0 = 0$, on a

$$0 = \langle \sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha\delta_0, \psi_\beta\rangle = c_\beta(-1)^{|\beta|} \ ,$$

d'où $c_\beta = 0$ pour tout β. ∎

7.7.8. Exemples.

La fonction définie sur $\mathbb{R}$ par $x \mapsto \log \dfrac{1}{|x|}$ est dans L^1_{loc}, donc définit une distribution L sur $\mathbb{R}$. Pour $\varphi \in \mathscr{D}(\mathbb{R})$, on a

$$\langle \partial L, \varphi\rangle = -\langle L, \varphi'\rangle = -\int \log\frac{1}{|x|}\varphi'(x)\,dx = -\lim_{\varepsilon\to 0}\int_\varepsilon^{+\infty}(\varphi'(x)+\varphi'(-x)).\log\frac{1}{x}\,dx$$

Et une intégration par parties donne

$$\begin{aligned}\int_\varepsilon^{+\infty}(\varphi'(x)+\varphi'(-x)).\log\frac{1}{x}\,dx &= \Big[\big(\varphi(x)-\varphi(-x)\big)\log\frac{1}{x}\Big]_\varepsilon^{+\infty} + \int_\varepsilon^{+\infty}\frac{\varphi(x)-\varphi(-x)}{x}\,dx \\ &= \big(\varphi(-\varepsilon)-\varphi(\varepsilon)\big)\log\varepsilon + \int_{|x|\geqslant\varepsilon}\varphi(x)\,\frac{dx}{x} \to \langle \mathrm{vp}(\frac{1}{x}), \varphi\rangle \ ,\end{aligned}$$

puisque $\varphi(-\varepsilon)-\varphi(\varepsilon) = O(\varepsilon)$, d'où le résultat $\partial L = -\,\mathrm{vp}(\dfrac{1}{x})$.

La fonction f définie sur $\mathbb{R}^2$ par $f(z) = \log\dfrac{1}{\|z\|}$ (avec $z = (x,y)$) est dans L^1_{loc}. On va calculer les distributions $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$ et $\Delta f = \dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2}$. Les dérivées partielles $\dfrac{\partial f}{\partial x} = -\dfrac{x}{\|z\|^2}$ et $\dfrac{\partial f}{\partial y} = -\dfrac{y}{\|z\|^2}$ sont aussi dans L^1_{loc}. De plus, pour toute $\varphi \in \mathscr{D}(\mathbb{R}^2)$ et tout $y \neq 0$, la fonction $x \mapsto f(x,y).\varphi(x,y)$ est $\mathscr{C}^1$ à support compact : on a donc

$$0 = \int_{-\infty}^{+\infty}\frac{d}{dx}\Big(f(x,y)\varphi(x,y)\Big)\,dx = \int_{-\infty}^{+\infty}\Big(f(x,y)\frac{\partial\varphi}{\partial x}(x,y) - \frac{x}{\|z\|^2}\varphi(x,y)\Big)\,dx$$

et puisque les fonctions $f\dfrac{\partial\varphi}{\partial x}$ et $\dfrac{x\varphi}{\|z\|^2}$ sont dans L^1, on obtient par Fubini :

$$\langle f, \partial_x\varphi\rangle - \langle\frac{x}{\|z\|^2}, \varphi\rangle = \iint f(x,y)\frac{\partial\varphi}{\partial x}(x,y)\,dx\,dy - \iint\frac{x}{\|z\|^2}\varphi(x,y)\,dx\,dy = 0$$

c'est-à-dire que, en tant que distribution, la dérivée $\partial_x f$ est égale à la fonction $\dfrac{-x}{\|z\|^2}$. Et de même, la dérivée $\partial_y f$ est la fonction $\dfrac{-y}{\|z\|^2}$.

On a donc pour $\varphi \in \mathscr{D}(\mathbb{R}^2)$,

$$\begin{aligned}\langle \Delta f, \varphi\rangle &= \langle \partial_x(\partial_x f), \varphi\rangle + \langle \partial_y(\partial_y f), \varphi\rangle = -\langle \partial_x f, \partial_x \varphi\rangle - \langle \partial_y f, \partial_y \varphi\rangle \\ &= \iint \Big(\frac{x}{\|z\|^2}\partial_x\varphi + \frac{y}{\|z\|^2}\partial_y\varphi\Big)\,dx\,dy\end{aligned}$$

et, en passant en coordonnées polaires (r, θ), puisque $\dfrac{\partial\varphi}{\partial r} = \dfrac{x}{\|z\|}\varphi + \dfrac{y}{\|z\|}\partial_y\varphi$,

$$\langle \Delta f, \varphi\rangle = \iint \frac{1}{r}\frac{\partial\varphi}{\partial r}(r,\theta) r\,dr\,d\theta = \int_0^{2\pi}\left(\int_0^{\infty}\frac{\partial\varphi}{\partial r}(r,\theta)\,dr\right)d\theta = \int_0^{2\pi}(-\varphi(0))\,d\theta = -2\pi\varphi(0)$$

dont on déduit que $\Delta f = -2\pi\delta_0$.

La formule des sauts.

Soit f une fonction sur $\mathbb{R}$. On suppose qu'il existe un nombre fini $a_1 < a_2 < \cdots < a_m$ de points de $\mathbb{R}$ tels que f soit de classe $\mathscr{C}^1$ en dehors de a_j et possède ainsi que f' une limite à droite et une limite à gauche en chaque a_j. Alors la dérivée de la distribution T_f associée à f est la somme de la fonction f' et de $\sum_{j=1}^m (f(a_j^+) - f(a_j^-)).\delta_{a_j}$.
En effet, si $\varphi \in \mathscr{D}(\mathbb{R})$, on a

$$\langle \partial T_f, \varphi\rangle = -\langle T_f, \varphi'\rangle = -\int f(x)\varphi'(x)\,dx$$

et en choisissant $a_0 < a_1$ et $a_{m+1} > a_m$ tels que $a_0 < \inf(\mathrm{supp}(\varphi))$ et $a_{m+1} > \sup(\mathrm{supp}(\varphi))$,

$$\begin{aligned}\langle \partial T_f, \varphi\rangle &= -\sum_{j=0}^m \int_{a_j}^{a_{j+1}} f(x)\varphi'(x)\,dx = -\sum_{j=0}^m\left([f(x)\varphi(x)]_{a_j}^{a_{j+1}} - \int_{a_j}^{a_{j+1}} f'(x)\varphi(x)\,dx\right) \\ &= \int_{a_0}^{a_{m+1}} f'(x)\varphi(x)\,dx + \sum_{j=1}^m \varphi(a_j)\Big(f(a_j^+) - f(a_j^-)\Big)\end{aligned}$$

c'est-à-dire $\partial T_f = T_{f'} + \sum_{j=1}^m (f(a_j^+) - f(a_j^-)).\delta_{a_j}$.

Distributions à support singleton.

Théorème 7.7.9. *Si T est une distribution sur $\mathbb{R}^d$ dont le support est le singleton $\{0\}$, alors T est combinaison linéaire finie de dérivées de la mesure de Dirac δ_0.*

Démonstration : Puisque T est à support compact, elle est d'ordre fini m. En vertu du théorème 7.6.5, T s'annule sur l'intersection (finie) des noyaux des distributions $\partial_\alpha\delta_0$ pour $|\alpha| \leqslant m$. Il en résulte, comme dans la démonstration du théorème 4.2.1, que T est combinaison linéaire de ces distributions. ∎

Distributions satisfaisant $x^m.T = 1$.

Théorème 7.7.10. *Pour tout entier m, il existe une distribution T_m sur $\mathbb{R}$ telle que $x^m.T_m = 1$. Les solutions de l'équation $x^m.T = 1$ dans $\mathscr{D}'(\mathbb{R})$ sont les distributions de la forme $T_m + \sum_{j=0}^{m-1} \lambda_j \dfrac{d^j}{dx^j}\delta_0$.*

Démonstration : En appliquant à la fonction $\varphi \in \mathscr{D}(\mathbb{R})$ la formule de Taylor avec reste intégral, on a

$$\varphi(x) = \sum_{j=0}^{m-1} \frac{x^j}{j!}\varphi^{(j)}(0) + x^m \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!}\varphi^{(m)}(tx)\,dt\ .$$

On vérifie aisément à l'aide du théorème de dérivation sous le signe $\int$, que la fonction $\varphi_m : x \mapsto \displaystyle\int_0^1 \frac{(1-t)^{m-1}}{(m-1)!}\varphi^{(m)}(tx)\,dt$ est $\mathscr{C}^\infty$. Alors, si χ est une fonction de $\mathscr{D}(\mathbb{R})$ égale à 1 au voisinage de 0, la fonction $\varphi(x) - \sum_{j=0}^{m-1} \dfrac{\varphi^{(j)}(0)}{j!}x^j\chi(x)$ est dans $\mathscr{D}(\mathbb{R})$ et égale au voisinage de 0 à $x^m\varphi_m(x)$. Et la fonction $\psi_m : x \mapsto x^{-m}\Big(\varphi(x) - \sum_{j=0}^{m-1} \dfrac{\varphi^{(j)}(0)}{j!}x^j\chi(x)\Big)$ est dans $\mathscr{D}(\mathbb{R})$. L'application linéaire $\varphi \mapsto \psi_m$ est continue de $\mathscr{D}(\mathbb{R})$ dans lui-même, et on en déduit que $T_m : \varphi \mapsto \int \psi_m(x)\,dx$ est une distribution sur $\mathbb{R}$. Et, puisque, pour $j < m$, on a $(x^m\varphi)^{(j)}(0) = 0$, on obtient

$$\langle x^m.T_m, \varphi\rangle = \langle S_m, x^m\varphi\rangle = \int x^{-m}(x^m\varphi(x))\,dx = \langle 1, \varphi\rangle\ ,$$

ce qui montre que $x^m.T_m = 1$.

Et si S est une distribution telle que $x^m.S = 1$, on a nécessairement $x^m.(S - T_m) = 0$. Puisque la fonction x^{-m} est $\mathscr{C}^\infty$ sur $\mathbb{R}^* = \mathbb{R}\setminus\{0\}$, on doit avoir alors $(S - T_m)_{|\mathbb{R}^*} = 0$, donc $\operatorname{supp}(S - T_m) \subset \{0\}$. Il résulte alors du théorème 7.7.9 que $S - T_m$ est une combinaison linéaire de dérivées de la mesure de Dirac δ_0 : $S - T_m = \sum_{j=0}^{k} \lambda_j\delta_0^{(j)}$. Alors, si $k \geqslant m$ et $\lambda_k \neq 0$,

$$0 = x^{k-m}.x^m(S - T_m) = \sum_{j=0}^{k} \lambda_j x^k \delta_0^{(j)} = \lambda_k(-1)^k k!\,\delta_0$$

ce qui force $\lambda_j = 0$ pour tout $j \geqslant m$. On en conclut que $S - T_m = \sum_{j=0}^{m-1} \lambda_j\delta_0^{(j)}$. ■

7.8 Composition par un difféomorphisme

Soient Ω et Ω' deux ouverts de $\mathbb{R}^d$, et Φ un $\mathscr{C}^\infty$-difféomorphisme de Ω sur Ω'. Si $f \in L^1_{\mathrm{loc}}(\Omega)$, alors $g = f\circ\Phi^{-1} \in L^1_{\mathrm{loc}}(\Omega')$, et la formule de changement de variables dans l'intégrale (théorème 5.11.5) donne, pour $\varphi \in \mathscr{D}(\Omega')$:

$$\int_{\Omega'} g(x).\varphi(x)\,dx = \int_{\Omega'} f\circ\Phi^{-1}(x).\varphi(x)\,dx = \int_{\Omega} f(y).\varphi\circ\Phi(y)\,|J_\Phi(y)|\,dy$$

où $J_\Phi(y)$ est le déterminant jacobien de Φ en y, et puisque $\varphi\circ\Phi \in \mathscr{D}(\Omega)$ et $|J_\Phi| \in \mathscr{C}^\infty(\Omega)$ (noter que J_Φ est $\mathscr{C}^\infty$ et ne s'annule pas sur Ω), on trouve $\langle T_g, \varphi\rangle = \langle |J_\Phi|.T_f, \varphi\circ\Phi\rangle$. On peut donc définir de même pour toute distribution T sur Ω une forme linéaire $\Phi^*(T)$ sur $\mathscr{D}(\Omega')$ en posant $\langle \Phi^*(T), \varphi\rangle = \langle |J_\Phi|.T, \varphi\circ\Phi\rangle$.

Théorème 7.8.1. *La forme linéaire $\Phi^*(T)$ est une distribution sur Ω' et l'application linéaire Φ^* est continue de $\mathscr{D}'(\Omega)$ dans $\mathscr{D}'(\Omega')$.*

DÉMONSTRATION : A nouveau, il suffit de montrer que l'application $C_\Phi : \varphi \mapsto \varphi\circ\Phi$ est continue de $\mathscr{D}(\Omega')$ dans $\mathscr{D}(\Omega)$. Et ceci est une conséquence immédiate du théorème du graphe fermé, puisque

$$\psi = \varphi\circ\Phi \iff \forall x \in \Omega \quad \psi(x) = \varphi(\Phi(x)) \ ,$$

ce qui montre que le graphe de C_Φ est fermé dans $\mathscr{D}(\Omega') \times \mathscr{D}(\Omega)$. ■

Distributions homogènes

Une fonction f sur $\mathbb{R}^d$ est dite homogène de degré k si $f(\lambda x) = \lambda^k f(x)$ pour tout $x \in \mathbb{R}^d$ et tout $\lambda > 0$, c'est-à-dire si $\Phi_\lambda^*(T_f) = \lambda^k T_f$ pour tout $\lambda > 0$ lorsque Φ_λ est l'homothétie $x \mapsto \lambda^{-1}x$. On dira de même qu'une distribution T sur $\mathbb{R}^d$ est *homogène de degré k* si cette même condition $\Phi_\lambda^*(T) = \lambda^k T$ est vérifiée pour tout $\lambda > 0$. Et puisque $J_{\Phi_\lambda}(y) = \lambda^{-d}$ pour tout y, cette condition peut encore s'écrire $\lambda^k\langle T, \varphi(x)\rangle = \lambda^{-d}\langle T, \varphi(\frac{x}{\lambda})\rangle$, ou encore $\langle T, \varphi(x)\rangle = \lambda^{-d-k}\langle T, \varphi(\frac{x}{\lambda})\rangle$. Il en résulte en particulier que δ_0 est homogène de degré $-d$.

Théorème 7.8.2. *Si T est une distribution homogène de degré k, et $\alpha \in \mathbb{N}^d$, la distribution $\partial_\alpha T$ est homogène de degré $k - |\alpha|$.*

DÉMONSTRATION : Il suffit, pour le montrer par récurrence sur $|\alpha|$, de vérifier que $\partial_j T$ est homogène de degré $k-1$ si T est homogène de degré k. Si $\varphi \in \mathscr{D}(\mathbb{R}^d)$ et si on note φ_λ la fonction $x \mapsto \varphi(\frac{x}{\lambda})$, on a $\partial_j\varphi_\lambda(x) = \frac{1}{\lambda}\partial_j\varphi(\frac{x}{\lambda})$, donc

$$\langle\partial_j T, \varphi\rangle = -\langle T, \partial_j\varphi\rangle = -\lambda^{-k-d}\langle T, \partial_j\varphi(\frac{x}{\lambda})\rangle = -\lambda^{-k-d}\langle T, \lambda\partial_j\varphi_\lambda\rangle = \lambda^{-k-d+1}\langle\partial_j T, \varphi_\lambda\rangle$$

d'où le résultat. ■

Théorème 7.8.3. *Si T est une distribution homogène de degré $-d$ sur $\mathbb{R}^d$ à support compact, alors T est proportionnelle à la mesure de Dirac en 0.*

DÉMONSTRATION : On vérifie sans peine que le support de $\Phi_\lambda^*(T)$ est égal à $\Phi_\lambda(\mathrm{supp}(T))$. Donc si T est homogène, on doit avoir $\mathrm{supp}(T) = \Phi_\lambda(\mathrm{supp}(T))$. Et ceci n'est possible, lorsque $\mathrm{supp}(T)$ est compact, que si $\mathrm{supp}(T) = \{0\}$. On en déduit que T doit être une combinaison linéaire de dérivées de δ_0 : $T = \sum_{|\alpha|\leqslant m} c_\alpha.\partial_\alpha\delta_0$. On a alors

$$\sum_{|\alpha|\leqslant m} c_\alpha.\lambda^{-d}\partial_\alpha\delta_0 = \lambda^{-d}T = \Phi_\lambda^*(T) = \sum_{|\alpha|\leqslant m} c_\alpha\Phi_\lambda^*(\partial_\alpha\delta_0) = \sum_{|\alpha|\leqslant m} c_\alpha\lambda^{-d-|\alpha|}\partial_\alpha\delta_0$$

donc $\sum_{0<|\alpha|\leqslant m} c_\alpha\lambda^{-|\alpha|}\partial_\alpha\delta_0 = 0$. On déduit alors de 7.7.7 que $c_\alpha = 0$ dès que $|\alpha| \neq 0$. Donc $T = c\delta_0$. ■

7.9 Convolution d'une distribution et d'une fonction de test

Soient f une fonction dans $L^1_{\mathrm{loc}}(\mathbb{R}^d)$, et $\psi \in \mathscr{D}(\mathbb{R}^d)$. On vérifie sans peine que la fonction $g = f*\psi$ est une fonction $\mathscr{C}^\infty$, donc localement intégrable. De plus, si $\varphi \in \mathscr{D}(\mathbb{R}^d)$, la fonction f est intégrable sur le compact $K = \{x - z : x \in \mathrm{supp}(\varphi), z \in \mathrm{supp}(\psi)\} \subset \mathbb{R}^d$, tandis que φ et ψ sont bornées : on a donc, d'après le théorème de Fubini,

$$\begin{aligned}\langle T_g, \varphi\rangle = \int g(x)\varphi(x)\,d\mu(x) &= \iint f(y)\psi(x-y)\varphi(x)\,d\mu(x)\,d\mu(y)\\ &= \int f(y)\Big(\int \varphi(x)\psi(x-y)\,d\mu(x)\Big)\,d\mu(y)\\ &= \int f(y)\Big(\int \varphi(x)\tilde{\psi}(y-x)\,d\mu(x)\Big)\,d\mu(y)\\ &= \int f(y)(\tilde{\psi}*\varphi(y))\,d\mu(y) = \langle T_f, \tilde{\psi}*\varphi\rangle\end{aligned}$$

si on note $\tilde{\psi}$ la fonction $y \mapsto \psi(-y)$. On va généraliser cette opération à toutes les distributions.

Théorème 7.9.1. *Soient $T \in \mathscr{D}'(\mathbb{R}^d)$ et $\psi \in \mathscr{D}(\mathbb{R}^d)$. Il existe une distribution S sur $\mathbb{R}^d$, notée $T*\psi$ et appelée convolée de T et de ψ telle que, pour toute $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on ait $\langle S,\varphi\rangle = \langle T, \tilde{\psi}*\varphi\rangle$, où $\tilde{\psi}(x) = \psi(-x)$. De plus l'application $T \mapsto T*\psi$ est continue de $\mathscr{D}'(\mathbb{R}^d)$ dans $\mathscr{D}'(\mathbb{R}^d)$.*

Démonstration : L'application linéaire C_ψ sur $\mathscr{D}(\mathbb{R}^d)$ à valeurs dans $\mathscr{D}(\mathbb{R}^d)$ définie par $C_\psi(\varphi) = \tilde{\psi}*\varphi$ est continue puisque son graphe est fermé. En effet

$$\varphi_1 = \tilde{\psi}*\varphi \iff \forall x \in \mathbb{R}^d \quad \varphi_1(x) = \int \psi(y-x)\varphi(y)\,d\mu(y)$$

et puisque la fonction $f_x : y \mapsto \psi(y-x)$ est localement intégrable pour x fixé, la fonction $\varphi \mapsto \langle T_{f_x}, \varphi\rangle$ est continue sur $\mathscr{D}(\mathbb{R}^d)$, comme l'est, pour x fixé, l'application $\delta_x : \varphi_1 \mapsto \varphi_1(x)$. Alors la transposée de C_ψ est une application linéaire continue de $\mathscr{D}'(\mathbb{R}^d)$ dans lui-même, et $S = {}^tC_\psi(T)$ est la distribution cherchée. ■

Théorème 7.9.2. *Soient Ω un ouvert de $\mathbb{R}^d$, Ω' un ouvert de $\mathbb{R}^\ell$, T une distribution sur Ω' et $\varphi \in \mathscr{D}(\Omega\times\Omega')$. Alors, si on note φ_x la fonction $y \mapsto \varphi(x,y)$, qui appartient à $\mathscr{D}(\Omega')$, la fonction $\varphi^T : x \mapsto \langle T, \varphi_x\rangle$ appartient à $\mathscr{D}(\Omega)$. De plus l'application linéaire $\pi_T : \varphi \mapsto \varphi^T$ est continue de $\mathscr{D}(\Omega\times\Omega')$ dans $\mathscr{D}(\Omega)$.*

Démonstration : Le support de φ est un compact de $\Omega\times\Omega'$ dont les projections sont respectivement un compact K de Ω et un compact K' de Ω'. Alors, pour tout $x \in \Omega$, on a $\mathrm{supp}(\varphi_x) \subset K'$ et $\varphi_x = 0$ si $x \notin K$. Il résulte alors du lemme 7.1.6 que la fonction $x \mapsto \varphi_x$ est continue sur Ω à support dans K, donc que $\varphi^T : x \mapsto \langle T, \varphi_x\rangle$ est continue à support dans K. Il en résulte aussi, en vertu du lemme 7.1.5, que φ^T est de classe $\mathscr{C}^1$ sur Ω avec $\partial_j(\varphi^T) = (\partial_j\varphi)^T$. Et en appliquant successivement le même raisonnement aux fonctions $\partial_{\alpha,0}\varphi$, on montre par récurrence sur $|\alpha|$ que φ^T est de classe $\mathscr{C}^\infty$ à support dans K, donc que φ^T appartient à $\mathscr{D}(\Omega)$.

Enfin, pour montrer la continuité de $\pi_T : \varphi \mapsto \varphi^T$ de $\mathscr{D}(\Omega \times \Omega')$ dans $\mathscr{D}(\Omega)$, on utilise le théorème du graphe fermé : en effet, pour $(\varphi, \psi) \in \mathscr{D}(\Omega \times \Omega') \times \mathscr{D}(\Omega)$,

$$\psi = \pi_T(\varphi) \iff \forall x \in \Omega \quad \psi(x) = \langle T, \varphi_x \rangle$$

et pour tout x fixé dans Ω, les applications linéaires $\delta_x : \psi \mapsto \psi(x)$ et $R_x : \varphi \mapsto \varphi_x$ sont continues. La continuité de R_x peut aussi se montrer en utilisant le théorème du graphe fermé, puisque le graphe de R_x est $\{(\varphi, \psi) \in \mathscr{D}(\Omega \times \Omega') \times \mathscr{D}(\Omega') : \forall y \quad \varphi(x, y) = \psi(y)\}$. ■

Théorème 7.9.3. *Soient L un compact de $\mathbb{R}^\ell$, μ la mesure de Lebesgue sur L, Ω un ouvert de $\mathbb{R}^d$ et F une application continue de L dans $\mathscr{D}(\Omega)$. Alors il existe un élément φ_F de $\mathscr{D}(\Omega)$ tel que, pour toute distribution $T \in \mathscr{D}'(\Omega)$, on ait : $\int_L \langle T, F(x) \rangle \, d\mu(x) = \langle T, \varphi_F \rangle$.*

Démonstration : L'ensemble $F(L)$ est compact dans $\mathscr{D}(\Omega)$. Il existe donc pour tout $\alpha \in \mathbb{N}^d$, un M_α tel que $p_\alpha(F(x)) \leqslant M_\alpha$ pour tout $x \in L$. Et il existe un compact $K \subset \Omega$ tel que $\mathrm{supp}(F(x)) \subset K$ pour tout $x \in L$.
Alors l'ensemble $H = \{\varphi \in \mathscr{D}(\Omega) : \mathrm{supp}(\varphi) \subset K$ et $\forall \alpha \in \mathbb{N}^d \ p_\alpha(\varphi) \leqslant M_\alpha\}$ est convexe équilibré compact dans $\mathscr{D}(\Omega)$ et contient tous les points de $F(L)$.
Soit Δ un ensemble dénombrable de distributions sur Ω contenant les mesures de Dirac $(\delta_y)_{y \in Y}$ pour un ensemble dense Y de points de K. On définit une application linéaire continue Ψ_Δ de $\mathscr{D}(\Omega)$ dans $\mathbb{C}^\Delta$ par $\Psi_\Delta(\varphi) = \big(\langle T, \varphi \rangle\big)_{T \in \Delta}$. Et on définit un élément $\theta = (\theta_T)_{T \in \Delta}$ de $\mathbb{C}^\Delta$ en posant pour tout $T \in \Delta$: $\theta_T = \dfrac{1}{\mu(L)} \displaystyle\int_L \langle T, F(x) \rangle \, d\mu(x)$. On veut montrer que $\theta \in \widehat{H} := \Psi_\Delta(H)$.
Puisque $\widehat{H}$ est convexe, équilibré et compact, donc fermé dans $\mathbb{C}^\Delta$, il résulte du théorème de Hahn-Banach (corollaire 4.4.5) que si θ n'appartenait pas à $\widehat{H}$, il existerait une forme linéaire continue ζ sur $\mathbb{C}^\Delta$ telle que $|\langle \zeta, \theta \rangle| > \sup_{\lambda \in \widehat{H}} |\langle \zeta, \lambda \rangle| \geqslant \sup_{x \in L} |\langle \zeta, \Psi_\Delta \circ F(x) \rangle|$.
Puisque toute forme linéaire continue sur $\mathbb{C}^\Delta$ est combinaison linéaire finie de fonctions coordonnées, il existe un entier m, $(T_1, T_2, \ldots, T_m)$ dans Δ et $(\alpha_1, \alpha_2, \ldots, \alpha_m)$ dans $\mathbb{C}$ tels que $\langle \zeta, \lambda \rangle = \sum_{j=1}^m \alpha_j \lambda_{T_j}$, pour $\lambda = (\lambda_T)_{T \in \Delta}$. Et si on pose $T = \sum_j \alpha_j T_j \in \mathscr{D}'(\Omega)$, on a

$$|\langle \zeta, \theta \rangle| = \frac{1}{\mu(L)} \left| \sum_j \alpha_j \int_L \langle T_j, F(x) \rangle \, d\mu(x) \right| = \frac{1}{\mu(L)} \left| \int \langle T, F(x) \rangle \, d\mu(x) \right| \leqslant \sup_{x \in L} |\langle T, F(x) \rangle|$$

alors que $\langle \zeta, \Psi_\Delta \circ F(x) \rangle = \sum_j \alpha_j \langle T_j, F(x) \rangle = \langle T, F(x) \rangle$. Cette contradiction montre que $\theta \in \widehat{H}$, donc qu'il existe $\varphi \in H$ telle que $\theta = \Psi_\Delta(\varphi)$, ou encore que, pour tout $T \in \Delta$, en posant $\varphi_F = \mu(L)\varphi$, on a $\int_L \langle T, F(x) \rangle \, d\mu(x) = \langle T, \mu(L)\varphi \rangle = \langle T, \varphi_F \rangle$. On a en particulier $\varphi_F(y) = \int_L \langle \delta_y, F(x) \rangle \, d\mu(x)$ pour tout y dans une partie dense de K. La fonction φ_F est donc entièrement déterminée, indépendamment du choix de Δ. Il en résulte que, quelle que soit $T \in \mathscr{D}'(\Omega)$, si on choisit $\Delta = \{\delta_y : y \in Y\} \cup \{T\}$, on a l'égalité cherchée. ■

Théorème 7.9.4. *Si ψ appartient à $\mathscr{D}(\mathbb{R}^d)$ et si T est une distribution sur $\mathbb{R}^d$, la distribution $T * \psi$ est une fonction $\mathscr{C}^\infty$ sur $\mathbb{R}^d$. De plus l'application linéaire $\psi \mapsto T * \psi$ est continue de $\mathscr{D}(\mathbb{R}^d)$ dans $\mathscr{E}(\mathbb{R}^d)$.*

Démonstration : Pour tout $x \in \mathbb{R}^d$, la fonction $\widetilde{\tau_x \psi} : y \mapsto \psi(x - y)$ est dans $\mathscr{D}(\mathbb{R}^d)$, de sorte que la fonction θ définie par $\theta(x) = \langle T, \widetilde{\tau_x \psi} \rangle$ est bien définie. On va montrer que θ est $\mathscr{C}^\infty$ et que la distribution associée à la fonction θ est égale à $T * \psi$.

Soit $a \in \mathbb{R}^d$. Il existe une fonction χ dans $\mathscr{D}(\mathbb{R}^d)$ qui est égale à 1 en tout point d'un voisinage W de a. Alors la fonction ρ définie sur $\mathbb{R}^d \times \mathbb{R}^d$ par $\rho(x,y) = \chi(x)\psi(x-y)$ est $\mathscr{C}^\infty$ et à support compact : en effet $\rho(x,y) = 0$ en tout (x,y) qui n'appartient pas à $\mathrm{supp}(\chi) \times (\mathrm{supp}(\chi) - \mathrm{supp}(\psi))$. On déduit alors du théorème 7.9.2 appliqué à T et ρ que la fonction $x \mapsto \chi(x)\theta(x)$ est dans $\mathscr{D}$, donc que θ est $\mathscr{C}^\infty$ sur W, donc au voisinage de tout point de $\mathbb{R}^d$.
Si, maintenant, $\varphi \in \mathscr{D}(\mathbb{R}^d)$, et si le compact L est son support, on considère la fonction $\rho : (x,y) \mapsto \varphi(x)\psi(x-y)$ qui appartient à $\mathscr{D}(\mathbb{R}^d \times \mathbb{R}^d)$. Il résulte alors du lemme 7.1.6 que $F : x \mapsto \rho_x$ est continue de L dans $\mathscr{D}(\mathbb{R}^d)$. L'application du théorème 7.9.3 donne alors l'existence d'une fonction $\varphi_F \in \mathscr{D}(\mathbb{R}^d)$ telle que l'on ait : $\int_L \langle S, F(x)\rangle \, d\mu(x) = \langle S, \varphi_F\rangle$ pour toute distribution S. En particulier, pour $a \in \mathbb{R}^d$ et $S = \delta_a$, on trouve

$$\varphi_F(a) = \int_L \varphi(x).\psi(x-a)\, d\mu(x) = \tilde{\psi} * \varphi(a) \ .$$

On en déduit que $\varphi_F = \tilde{\psi} * \varphi$. Et pour $S = T$, on obtient

$$\int_L \varphi(x)\theta(x)\, d\mu(x) = \langle T, \varphi_F\rangle = \langle T, \tilde{\psi} * \varphi\rangle = \langle T * \psi, \varphi\rangle \ ,$$

ce qui montre que $T * \psi$ est la distribution associée à la fonction θ de classe $\mathscr{C}^\infty$.

Enfin, pour montrer la continuté de l'application linéaire $C_T : \psi \mapsto T * \psi$, il suffit de vérifier que son graphe est fermé dans $\mathscr{D}(\mathbb{R}^d) \times \mathscr{E}(\mathbb{R}^d)$. Et puisque “$\theta = T * \psi$” est équivalent à “ $\theta(y) = \langle T, \widetilde{\tau_y \psi}\rangle$ pour tout y de $\mathbb{R}^d$ ”, ceci résulte immédiatement de la continuité, pour y fixé, de l'application $\psi \mapsto \widetilde{\tau_y \psi}$ de $\mathscr{D}(\mathbb{R}^d)$ dans lui-même. ■

Corollaire 7.9.5. *Soit Ω un ouvert de $\mathbb{R}^d$. Toute distribution sur Ω est limite dans $\mathscr{D}'$ d'une suite de fonctions.*

Démonstration : Supposons d'abord que $\Omega = \mathbb{R}^d$ et soit $T \in \mathscr{D}'(\mathbb{R}^d)$. Choisissons une fonction positive $\rho \in \mathscr{D}(\mathbb{R}^d)$ à support dans la boule unité et d'intégrale 1. Alors, quelle que soit la fonction de test $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on sait que $\tilde{\rho}_\varepsilon * \varphi$ converge dans $\mathscr{D}(\mathbb{R}^d)$ vers φ quand $\varepsilon \to 0$. Il en résulte que $\langle T * \rho_\varepsilon, \varphi\rangle = \langle T, \tilde{\rho}_\varepsilon * \varphi\rangle \to \langle T, \varphi\rangle$ pour tout φ, c'est-à-dire que les fonctions $T * \rho_\varepsilon$ convergent vers T dans $\mathscr{D}'(\mathbb{R}^d)$.
Dans le cas général, il existe une suite exhaustive (K_n) de compacts de Ω, et on peut trouver pour chaque n une fonction $\psi_n \in \mathscr{D}(\Omega)$ égale à 1 au voisinage de K_n. Si on choisit une suite (ε_n) strictement positive convergeant vers 0 telle que $\varepsilon_n < \inf_{x \in \mathrm{supp}(\psi_n)} d(x, \Omega^c)$, la convolée $\theta_n = (\psi_n.T) * \rho_{\varepsilon_n}$ est une fonction $\mathscr{C}^\infty$. Et on va montrer que $T = \lim_n \theta_n$ dans $\mathscr{D}'(\Omega)$. Soit en effet $\varphi \in \mathscr{D}(\Omega) \subset \mathscr{D}(\mathbb{R}^d)$. Alors, il résulte de ce qui précède que $T * \rho_{\varepsilon_n} \to T$, c'est-à-dire que $\langle T * \rho_{\varepsilon_n}, \varphi\rangle \to \langle T, \varphi\rangle$. Pour $n \geqslant n_0$, on a $\varepsilon_n < \inf_{\mathrm{supp}(\varphi)} d(., \Omega^c)$. Le support de $\tilde{\rho}_{\varepsilon_n} * \varphi$ est alors contenu dans Ω, donc dans un K_ℓ, et si on prend alors $n \geqslant \max(n_0, \ell)$, on a $\psi_n = 1$ au voisinage du support de $\tilde{\rho}_{\varepsilon_n} * \varphi$, donc

$$\langle \theta_n, \varphi\rangle = \langle (\psi_n.T) * \rho_{\varepsilon_n}, \varphi\rangle = \langle \psi_n.T, \tilde{\rho}_{\varepsilon_n} * \varphi\rangle = \langle T, \tilde{\rho}_{\varepsilon_n} * \varphi\rangle = \langle T * \rho_{\varepsilon_n}, \varphi\rangle \to \langle T, \varphi\rangle \ ,$$

ce qui achève la démonstration si on vérifie que le support de θ_n est contenu dans Ω. Et ceci est assuré par le choix de ε_n. ■

7.10 Produit tensoriel de distributions

Soient Ω un ouvert de $\mathbb{R}^d$ et Ω' un ouvert de $\mathbb{R}^\ell$. Si φ est une fonction complexe sur Ω et ψ une fonction complexe sur Ω', on note $\varphi \otimes \psi$ la fonction $(x, y) \mapsto \varphi(x).\psi(y)$ sur $\Omega \times \Omega'$. Il est clair que si $\varphi \in \mathscr{D}(\Omega)$ et $\psi \in \mathscr{D}(\Omega')$, alors $\varphi \otimes \psi \in \mathscr{D}(\Omega \times \Omega')$.

Théorème 7.10.1. *Le sous-espace de $\mathscr{D}(\Omega \times \Omega')$ engendré par les fonctions $\varphi \otimes \psi$, avec $(\varphi, \psi) \in \mathscr{D}(\Omega) \times \mathscr{D}(\Omega')$ est dense.*

En vertu du théorème de Hahn-Banach, il suffit de montrer que si une distribution T sur $\Omega \times \Omega'$ vérifie $\langle T, \varphi \otimes \psi\rangle = 0$ pour toute $\varphi \in \mathscr{D}(\Omega)$ et toute $\psi \in \mathscr{D}(\Omega')$, alors $T = 0$. On peut d'abord se ramener au cas où T est à support compact dans $\Omega \times \Omega'$. En effet, si on prend une partition localement finie (φ_j) de l'unité sur Ω en éléments de $\mathscr{D}(\Omega)$, et une partition localement finie (ψ_k) de l'unité sur Ω' en éléments de $\mathscr{D}(\Omega')$, chacune des distributions à support compact $T_{j,k} = (\varphi_j \otimes \psi_k).T$ sera nulle sur $\mathscr{D}(\Omega) \otimes \mathscr{D}(\Omega')$ puisque $\langle T_{j,k}, \varphi \otimes \psi\rangle = \langle T, (\varphi.\varphi_j) \otimes (\psi\psi_k)\rangle = 0$. Et si on prouve que chaque $T_{j,k}$ est nulle, on aura $T = \sum_{j,k} T_{j,k} = 0$.
Choisissons une fonction positive $\rho_1 \in \mathscr{D}(\mathbb{R}^d)$ d'intégrale 1 et une fonction positive $\rho_2 \in \mathscr{D}(\mathbb{R}^\ell)$ d'intégrale 1. Alors la fonction positive $\sigma = \rho_1 \otimes \rho_2$ appartient à $\mathscr{D}(\mathbb{R}^{d+\ell})$ et est d'intégrale 1. La convolée $T * \sigma_\varepsilon$ est une distribution sur $\Omega \times \Omega'$ pour ε assez petit, et cette distribution est la fonction h de classe $\mathscr{C}^\infty$ définie par $h(x, y) = \varepsilon^{-d-\ell}\langle T, \chi_{1,x} \otimes \chi_{2,y}\rangle$ où $\chi_{1,x}(u) = \rho_1(x - \dfrac{u}{\varepsilon})$ et $\chi_{2,x}(v) = \rho_2(y - \dfrac{v}{\varepsilon})$. Il en résulte que $h = 0$, donc que $T * \sigma_\varepsilon = 0$. Et puisque, dans $\mathscr{D}'(\Omega \times \Omega')$, on a $T = \lim_{\varepsilon\to 0} T * \sigma_\varepsilon$, on a bien $T = 0$. ■

Théorème 7.10.2. *Soient Ω un ouvert de $\mathbb{R}^d$, Ω' un ouvert de $\mathbb{R}^\ell$, S une distribution sur Ω et T une distribution sur Ω'. Alors il existe sur $\Omega \times \Omega'$ une unique distribution, notée $S \otimes T$, satisfaisant*

$$\langle S \otimes T, \varphi \otimes \psi\rangle = \langle S, \varphi\rangle.\langle T, \psi\rangle$$

pour tout $(\varphi, \psi) \in \mathscr{D}(\Omega) \times \mathscr{D}(\Omega')$.

Démonstration : L'unicité résulte du théorème 7.10.1 : si U_1 et U_2 avaient cette propriété, la distribution $U_1 - U_2$ serait nulle sur le sous-espace fermé de $\mathscr{D}(\Omega \times \Omega')$ engendré par les $\varphi \otimes \psi$ pour $\varphi \in \mathscr{D}(\Omega)$ et $\psi \in \mathscr{D}(\Omega')$, qui est égal à $\mathscr{D}(\Omega \times \Omega')$ en vertu du théorème 7.10.1. Donc $U_1 - U_2 = 0$.

Inversement, si S appartient à $\mathscr{D}(\Omega)$, l'application $U : \varphi \mapsto \langle S, \varphi^T\rangle$ (avec les notations du théorème 7.9.2) est continue sur $\mathscr{D}(\Omega \times \Omega')$. C'est donc une distribution.
Et si $(\psi_1, \psi_2) \in \mathscr{D}(\Omega) \times \mathscr{D}(\Omega')$, on a $(\psi_1 \otimes \psi_2)_x = \psi_1(x).\psi_2$ donc $\pi_T(\psi_1 \otimes \psi_2) = \langle T, \psi_2\rangle.\psi_1$ et $\langle U, \psi_1 \otimes \psi_2\rangle = \langle S, \psi_1\rangle.\langle T, \psi_2\rangle$, ce qui prouve l'existence.
On peut remarquer qu'en échangeant les deux facteurs dans le théorème 7.9.2, la formule $\varphi \mapsto \langle T, \langle S, \varphi(., y)\rangle$ construirait de manière analogue une distribution U' satisfaisant $\langle U', \psi_1 \otimes \psi_2\rangle = \langle S, \psi_1\rangle.\langle T, \psi_2\rangle$, et le résultat d'unicité montre que $U' = U$. ■

7.11 Convolution des distributions

Si $f \in L^1_{\text{loc}}(\mathbb{R}^d)$, si $g \in L^1(\mathbb{R}^d)$ a un support compact, alors $f * g$ est dans L^1_{loc}, et on a, pour toute $\varphi \in \mathscr{D}(\mathbb{R}^d)$, $\varphi * \tilde{g} \in \mathscr{D}(\mathbb{R}^d)$ et

$$\langle T_{f*g}, \varphi\rangle = \iint f(y)g(x-y)\varphi(x)\,dx\,dy = \iint f(y)\varphi(x)\tilde{g}(y-x)\,dy\,dx = \langle T_f, \varphi * \tilde{g}\rangle\ .$$

On va généraliser ceci pour définir la convolution d'une distribution et d'une distribution à support compact.

Rappelons qu'une application continue f entre deux espaces métriques X et Y est dite *propre* si l'image réciproque par f de tout compact de Y est compacte dans X (cf. 1.5.39).

Lemme 7.11.1. *Soient Ω un ouvert de $\mathbb{R}^d$, Ω' un ouvert de $\mathbb{R}^\ell$ et Φ une application linéaire de $\mathbb{R}^d$ dans $\mathbb{R}^\ell$ telle que $\Phi(\Omega) \subset \Omega'$. On suppose que T est une distribution sur Ω et que la restriction de Φ à $\operatorname{supp}(T)$ est propre de $\operatorname{supp}(T)$ dans Ω'. Alors il existe une distribution $U = \Phi(T)$ sur Ω' telle que $\langle U, \varphi\rangle = \langle T, \psi.\varphi\circ\Phi\rangle$, pour toute $\varphi \in \mathscr{D}(\Omega')$ et toute $\psi \in \mathscr{D}(\Omega)$ égale à 1 au voisinage du compact $\operatorname{supp}(T) \cap \Phi^{-1}(\operatorname{supp}(\varphi))$.*

Démonstration : Si $\varphi \in \mathscr{D}(\Omega')$, l'ensemble $L = \Phi^{-1}(\operatorname{supp}(\varphi))$ est fermé et l'ensemble $K = \operatorname{supp}(T) \cap L = (\Phi_{|\operatorname{supp}(T)})^{-1}(\operatorname{supp}(\varphi))$ est compact par hypothèse. Si $\psi \in \mathscr{D}(\Omega)$ est égale à 1 au voisinage de K, la fonction $\varphi_1 = \psi.\varphi\circ\Phi$ est $\mathscr{C}^\infty$ sur Ω, avec un support contenu dans celui de ψ. Donc $\varphi_1 \in \mathscr{D}(\Omega)$. Et si χ est une autre fonction de $\mathscr{D}(\Omega)$ égale à 1 au voisinage de K, on a $\psi - \chi = 0$ au voisinage de K. On a donc $\operatorname{supp}(\psi-\chi) \cap L \cap \operatorname{supp}(T) = \emptyset$, et puisque $\operatorname{supp}((\psi-\chi).\varphi\circ\Phi) \subset L \cap \operatorname{supp}(\psi-\chi)$ est un compact contenu dans $\Omega \setminus \operatorname{supp}(T)$, on a $(\psi-\chi).\varphi\circ\Phi = 0$ au voisinage de $\operatorname{supp}(T)$.
Il en résulte que $\langle T, (\psi-\chi).\varphi\circ\Phi\rangle = 0$, c'est-à-dire que la forme linéaire U sur $\mathscr{D}(\Omega')$ est bien définie, indépendamment du choix de ψ. Sa continuité résulte de celle de l'application linéaire $\Psi : \mathscr{D}(\Omega) \to \mathscr{D}(\Omega')$ définie par $\Psi(\varphi) = \psi.\varphi\circ\Phi$, qui se déduit elle-même du théorème du graphe fermé puisque

$$\varphi_1 = \psi.\varphi\circ\Phi \iff \forall (x,y) \in \text{graphe}(\Phi) \quad \varphi_1(x) = \psi(x).\varphi(y)\ ,$$

ce qui achève la démonstration. ■

En particulier, l'application $\Phi : x \mapsto -x$ est propre (c'est même un homéomorphisme) de $\mathbb{R}^d$ dans $\mathbb{R}^d$. Pour $S \in \mathscr{D}'(\mathbb{R}^d)$, la distribution $\tilde{S} = \Phi(S)$ vérifie $\langle \tilde{S}, \varphi\rangle = \langle S, \tilde{\varphi}\rangle$ pour toute $\varphi \in \mathscr{D}(\mathbb{R}^d)$, ainsi que $\operatorname{supp}(\tilde{S}) = -\operatorname{supp}(S)$. On a pour tout $\alpha \in \mathbb{N}^d$ et $\varphi \in \mathscr{D}(\mathbb{R}^d)$, $\partial_\alpha(\tilde{\varphi})(x) = (-1)^{|\alpha|}\partial_\alpha\varphi(-x)$, donc $\widetilde{\partial_\alpha S} = (-1)^{|\alpha|}\partial_\alpha\tilde{S}$.

Théorème 7.11.2. *Soient S une distribution à support compact sur $\mathbb{R}^d$, $\tilde{S}$ la distribution définie par $\langle \tilde{S}, \varphi\rangle = \langle S, \tilde{\varphi}\rangle$ et $T \in \mathscr{D}'(\mathbb{R}^d)$. Alors, pour $\varphi \in \mathscr{D}(\mathbb{R}^d)$, $\tilde{S} * \varphi$ est dans $\mathscr{D}(\mathbb{R}^d)$ et l'application U définie sur $\mathscr{D}(\mathbb{R}^d)$ par $U(\varphi) = \langle T, \tilde{S} * \varphi\rangle$ est une distribution notée $T * S$.*

Démonstration : On sait que la convolée de la distribution $\tilde{S}$ et de la fonction de test φ est $\mathscr{C}^\infty$ et que son support est contenu dans $\operatorname{supp}(\varphi) + \operatorname{supp}(\tilde{S}) = \operatorname{supp}(\varphi) - \operatorname{supp}(S)$. Et ce dernier ensemble est compact puisque $\operatorname{supp}(\varphi)$ et $\operatorname{supp}(S)$ le sont ; ceci donne un sens à $U(\varphi)$. Il est clair alors que U est une forme linéaire sur $\mathscr{D}(\Omega)$, dont la continuité résulte du théorème 7.9.4. ■

Théorème 7.11.3. *Si $T \in \mathscr{D}'(\mathbb{R}^d)$ et $\alpha \in \mathbb{N}^d$, on a $T * \partial_\alpha \delta_0 = \partial_\alpha T$.*

Démonstration : La distribution $\partial_\alpha \delta_0$ étant à support compact, il résulte de ce qui précède que, pour $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on a $\langle T * \partial_\alpha \delta_0, \varphi \rangle = \langle T, \widetilde{\partial_\alpha \delta_0} * \varphi \rangle$. Puisque $\widetilde{\partial_\alpha \delta_0}$ est défini par

$$\langle \widetilde{\partial_\alpha \delta_0}, \psi \rangle = \langle \partial_\alpha \delta_0, \tilde{\psi} \rangle = (-1)^{|\alpha|} \langle \delta_0, \partial_\alpha (\tilde{\psi}) \rangle = (-1)^{|\alpha|} \langle \delta_0, (-1)^{|\alpha|} \widetilde{\partial_\alpha \psi} \rangle = \langle \delta_0, \partial_\alpha \psi \rangle$$

on voit que $\widetilde{\partial_\alpha \delta_0} * \varphi(x) = \langle \widetilde{\partial_\alpha \delta_0}, \widetilde{\tau_x \varphi} \rangle = \langle \partial_\alpha \delta_0, \tau_x \varphi \rangle = (-1)^{|\alpha|} \langle \delta_0, \partial_\alpha (\tau_x \varphi) \rangle = (-1)^{|\alpha|} \partial_\alpha \varphi(x)$, donc que $\langle T * \partial_\alpha \delta_0, \varphi \rangle = (-1)^{|\alpha|} \langle T, \partial_\alpha \varphi \rangle = \langle \partial_\alpha T, \varphi \rangle$, c'est-à-dire $T * \partial_\alpha \delta_0 = \partial_\alpha T$. ∎

Si Φ désigne maintenant l'application $(x, y) \mapsto x + y$ de $\mathbb{R}^d \times \mathbb{R}^d$ dans $\mathbb{R}^d$, pour $S \in \mathscr{E}'(\mathbb{R}^d)$ et $T \in \mathscr{D}'(\mathbb{R}^d)$, le support de $S \otimes T$ est $H = \mathrm{supp}(S) \times \mathrm{supp}(T)$, et on voit sans peine que la restriction de Φ à H est propre : en effet, si $K \subset \mathbb{R}^d$ est compact, $\Phi^{-1}(K) \cap H$ est un fermé contenu dans le compact $\mathrm{supp}(S) \times (K - \mathrm{supp}(S))$.

Proposition 7.11.4. *Sous les hypothèses précédentes, $T * S = \Phi(T \otimes S)$.*

Démonstration : Si $S \in \mathscr{E}'(\mathbb{R}^d)$, $T \in \mathscr{E}'(\mathbb{R}^d)$ et $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on a défini $\langle T * S, \varphi \rangle$ par $\langle T, \tilde{S} * \varphi \rangle$ et, pour $x \in \mathbb{R}^d$, $\tilde{S} * \varphi(x) = \langle \tilde{S}, \widetilde{\tau_x \varphi} \rangle = \langle S, \tau_{-x} \varphi \rangle$. Alors la fonction $\psi : (x, y) \mapsto \varphi(x + y)$ est $\mathscr{C}^\infty$, la distribution $T \otimes S$ est à support compact, et on a

$$\langle \Phi(T \otimes S), \varphi \rangle = \langle T \otimes S, \varphi \circ \Phi \rangle = \langle T, \chi \rangle$$

où $\chi(x) = \langle S, \psi_x \rangle = \tilde{S} * \varphi(x)$. C'est-à-dire que $\Phi(T \otimes S) = T * S$.
Et si $T \in \mathscr{D}'(\mathbb{R}^d)$, avec $\rho \in \mathscr{D}(\mathbb{R}^d)$ égale à 1 au voisinage de $\mathrm{supp}(\varphi) - \mathrm{supp}(S)$, on a défini $\langle \Phi(T \otimes S), \varphi \rangle$ comme $\langle T \otimes S, \varphi_1 \rangle$, où $\varphi_1(x, y) = \rho(x) \varphi(x + y)$, c'est-à-dire $\langle (\rho.T) \otimes S, \varphi \circ \Phi \rangle$, ou encore $(\rho.T) * S, \varphi \rangle = \langle \rho.T, \tilde{S} * \varphi \rangle = \langle T, \rho.(\tilde{S} * \varphi) \rangle = \langle T, \tilde{S} * \varphi \rangle$ puisque $\rho = 1$ au voisinage du support de la fonction $\tilde{S} * \varphi$, et que $\rho.T \in \mathscr{E}'(\mathbb{R}^d)$. On conclut que $\langle \Phi(T \otimes S), \varphi \rangle = \langle T * S, \varphi \rangle$, c'est-à-dire $T * S = \Phi(T \otimes S)$. ∎

Théorème 7.11.5. *La convolution est associative et commutative sur $\mathscr{E}'(\mathbb{R}^d)$, et possède la mesure de Dirac en 0 comme unité.*

Démonstration : L'application linéaire $\Sigma : (x, y) \mapsto (y, x)$ de $\mathbb{R}^d \times \mathbb{R}^d$ dans lui-même transforme clairement $S \otimes T$ en $T \otimes S$. Et puisque $\Phi \circ \Sigma = \Phi$, on en déduit que

$$T * S = \Phi(T \otimes S) = \Phi \circ \Sigma(S \otimes T) = \Phi(S \otimes T) = S * T$$

ce qui montre la commutativité.
Si on identifie $\mathbb{R}^d \times (\mathbb{R}^\ell \times \mathbb{R}^m)$ et $(\mathbb{R}^d \times \mathbb{R}^\ell) \times \mathbb{R}^m$, à $\mathbb{R}^{d+\ell+m}$, on a $\varphi \otimes (\psi \otimes \chi) = (\varphi \times \psi) \otimes \chi$, si $\varphi \in \mathscr{D}(\mathbb{R}^d)$, $\psi \in \mathscr{D}(\mathbb{R}^\ell)$ et $\chi \in \mathscr{D}(\mathbb{R}^m)$. Il en résulte que si $T \in \mathscr{D}'(\mathbb{R}^d)$, $S \in \mathscr{D}'(\mathbb{R}^\ell)$ et $U \in \mathscr{D}'(\mathbb{R}^m)$, on a $T \otimes (S \otimes U) = ((T \otimes S) \otimes U$ en tant que distributions sur $\mathbb{R}^{d+\ell+m}$.
Alors, si T, S et U sont des distributions à support compact sur $\mathbb{R}^d$, et si Φ est l'application linéaire $(x, y, z) \mapsto x + y + z$ de $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \simeq \mathbb{R}^{3d}$ dans $\mathbb{R}^d$, on voit que

$$T * (S * U) = \Phi(T \otimes (S \otimes U)) = \Phi((T \otimes S) \otimes U) = (T * S) * U \ ,$$

ce qui montre l'associativité.
Enfin, pour $T \in \mathscr{D}'(\mathbb{R}^d)$ et $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on a $\langle T * \delta_0, \varphi \rangle = \langle T, \tilde{\delta}_0 * \varphi \rangle = \langle T, \delta_0 * \varphi \rangle = \langle T, \varphi \rangle$, donc $T * \delta_0 = T$. ∎

On peut aussi définir la convolution de deux distributions S et T lorsque l'application "somme" est propre sur $\mathrm{supp}(S) \times \mathrm{supp}(T)$.

Théorème 7.11.6. *Si S et T sont des distributions sur $\mathbb{R}^d$ et si l'application $(x, y) \mapsto x+y$ est propre de* $\mathrm{supp}(S) \times \mathrm{supp}(T)$ *dans $\mathbb{R}^d$, alors on peut définir une distribution $S * T$ à support dans le fermé* $\mathrm{supp}(S) + \mathrm{supp}(T)$ *telle que $\langle S * T, \varphi\rangle = \langle \psi.T, \tilde{S} * \varphi\rangle$ pour toute $\varphi \in \mathscr{D}(\mathbb{R}^d)$ et toute $\psi \in \mathscr{D}(\mathbb{R}^d)$ égale à 1 au voisinage de* $\mathrm{supp}(T) \cap (\mathrm{supp}(\varphi) - \mathrm{supp}(S))$.

Démonstration : La distribution $S \otimes T$ est définie sur $\mathbb{R}^d \times \mathbb{R}^d$, et son support est contenu dans $\mathrm{supp}(S) \times \mathrm{supp}(T)$. Si la restriction p à $\mathrm{supp}(S) \times \mathrm{supp}(T)$ de l'application somme : $(x, y) \mapsto x + y$ est propre, l'ensemble $\mathrm{supp}(S) + \mathrm{supp}(T) = p(\mathrm{supp}(S) \times \mathrm{supp}(T))$ est fermé en vertu du lemme 1.5.40.
Et l'ensemble $\mathrm{supp}(T) \cap (\mathrm{supp}(\varphi) - \mathrm{supp}(S)) = \{y : \exists x \quad (x, y) \in p^{-1}(\mathrm{supp}(\varphi))\}$ est la deuxième projection de l'ensemble compact $p^{-1}(\mathrm{supp}(\varphi))$, donc est une partie compacte de $\mathrm{supp}(T)$. Il existe donc des fonctions ψ dans $\mathscr{D}(\mathbb{R}^d)$ égales à 1 au voisinage de ce compact. Et l'application du lemme 7.11.1 permet alors de définir une distribution $S * T$ ayant les propriétés cherchées.

Dans le cas où S est à support compact, cette définition rejoint celle du théorème 7.11.2 en vertu de la proposition 7.11.4. ■

Convolution des distributions à support dans $\mathbb{R}^+$.

Il est immédiat que la somme est une application propre de $\mathbb{R}^+ \times \mathbb{R}^+$ dans $\mathbb{R}^+$. Il en résulte qu'on peut définir la convolée de deux distributions sur $\mathbb{R}$ à support dans $\mathbb{R}^+$, et que cette convolée est elle-même à support dans $\mathbb{R}^+$.
A titre d'exemple, on va calculer les "puissances de convolution" successives Y_k de la "fonction de Heaviside" Y définie par $Y(x) = 1$ si $x \geqslant 0$ et $Y(x) = 0$ si $x < 0$. Il découle immédiatement de la formule des sauts que la dérivée $Y' = \delta_0' * Y$ de Y est la mesure de Dirac δ_0 en 0.
Si on a $Y_{k+1} = Y * Y_k$, on doit avoir $\delta_0' * Y_{k+1} = (\delta_0' * Y) * Y_k = \delta_0 * Y_k = Y_k$. On va montrer par récurrence sur $k \geqslant 0$ que $Y_{k+1} = \dfrac{x^k}{k!} Y$. C'est clair si $k = 0$. Et si c'est vrai pour $k - 1$, on a, avec $k \geqslant 1$,

$$\delta_0' * (\frac{x^k}{k!}.Y) = (\frac{x^k}{k!}.Y)' = \frac{x^{k-1}}{(k-1)!}.Y + \frac{1}{k!}x^k.Y' = Y_k + \frac{1}{k!}x^k.\delta_0 = Y_k$$

On en déduit que la dérivée de la distribution $Z_k = \dfrac{x^k}{k!}.Y - Y_{k+1}$ est nulle, donc que cette distribution est une constante, nulle pour $x < 0$ puisque le support de Z_k est contenu dans $\mathbb{R}^+$.

7.12 Quelques solutions fondamentales

Définition 7.12.1. *Soit P un opérateur différentiel à coefficients constants, c'est-à-dire une combinaison linéaire $\sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha$. On dit qu'une distribution E est une* solution fondamentale *de P si $P.E = \delta_0$.*

On identifiera P à l'application linéaire de $\mathscr{D}(\Omega)$ dans $\mathscr{D}(\Omega)$: $\varphi \mapsto P\varphi = \sum_{|\alpha|\leqslant m} \partial_\alpha \varphi$. En vertu du théorème 7.7.9, , il est équivalent de dire qu'il existe une distribution T à support dans $\{0\}$ telle que $P\varphi = T * \varphi$. Alors une solution fondamentale de P est une distribution E telle que $E * T = \delta_0$.

Proposition 7.12.2. *Si P est un opérateur différentiel à coefficients constants et E une solution fondamentale de P, alors P est injectif de $\mathscr{E}'(\mathbb{R}^d)$ dans $\mathscr{E}'(\mathbb{R}^d)$, et pour toute distribution à support compact S, la distribution $E * S$ est solution de l'équation $PT = S$.*

Démonstration : En effet, si $S \in \mathscr{E}'(\mathbb{R}^d)$ et $PS = 0$, on doit avoir

$$\begin{aligned} S = \delta_0 * S = (PE) * S = \sum_{|\alpha|\leqslant m} c_\alpha(\partial_\alpha E) * S &= \sum_{|\alpha|\leqslant m} c_\alpha(\partial_\alpha \delta_0) * E * S \\ &= E * \sum_{|\alpha|\leqslant m} c_\alpha(\partial_\alpha \delta_0) * S = E * \sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha S \\ &= E * (PS) = 0 \ , \end{aligned}$$

ce qui prouve l'injectivité de P sur $\mathscr{E}'(\mathbb{R}^d)$. D'autre part,

$$\begin{aligned} P(E * S) &= \sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha (E * S) = \sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha \delta_0 * (E * S) \\ &= \sum_{|\alpha|\leqslant m} c_\alpha (\partial_\alpha \delta_0 * E) * S = \Big(\sum_{|\alpha|\leqslant m} c_\alpha \partial_\alpha E\Big) * S = \delta_0 * S = S \ , \end{aligned}$$

ce qui montre que $E * S$ est bien solution de l'équation. ■

Théorème 7.12.3. *Soient S une distribution à support compact sur $\mathbb{R}^d$ et T une distribution sur $\mathbb{R}^d$. On suppose que le support singulier de T est compact. Alors*

$$\operatorname{singsupp}(S * T) \subset \operatorname{singsupp}(S) + \operatorname{singsupp}(T) \ .$$

Démonstration : Notons $K = \operatorname{singsupp}(S)$ et $L = \operatorname{singsupp}(T)$. Puisque $K \subset \operatorname{supp}(S)$, K et L sont compacts par hypothèse.
Soit alors $\varepsilon > 0$. On notera K_ε l'ouvert $\{x : d(x, K) < \varepsilon\}$ et L_ε l'ouvert $\{x : d(x, L) < \varepsilon\}$. Il existe des fonctions $\psi \in \mathscr{D}(K_\varepsilon)$ et $\theta \in \mathscr{D}(L_\varepsilon)$ égales respectivement à 1 au voisinage de K et de L. On pose alors $S_1 = \psi.S$ et $T_1 = \theta.T$. Par définition des supports singuliers, les distributions $S - S_1 = (1-\psi).S$ et $T - T_1 = (1-\theta).T$ sont des fonctions $\mathscr{C}^\infty$. Et on a $S * T = S_1 * T_1 + (S - S_1) * T_1 + S * (T - T_1)$. Puisque $S - S_1 \in \mathscr{D}(\mathbb{R}^d)$, on a $(S - S_1) * T_1 \in \mathscr{C}^\infty(\mathbb{R}^d)$, et puisque $S \in \mathscr{E}'(\mathbb{R}^d)$, on a $S * (T - T_1) \in \mathscr{C}^\infty(\mathbb{R}^d)$. Donc $S * T$ est $\mathscr{C}^\infty$ hors du support de $S_1 * T_1$. Et le support de $S_1 * T_1$ est contenu dans la somme des supports de S_1 et T_1, donc dans $K + \tilde{B}(0,\varepsilon) + L + \tilde{B}(0,\varepsilon) \subset K + L + \tilde{B}(0, 2\varepsilon)$.
On en déduit que $\operatorname{singsupp}(S*T) \subset K + L + \tilde{B}(0, 2\varepsilon)$, et puisque ε est arbitraire, ceci achève la démonstration. ■

Théorème 7.12.4. *Si l'opérateur différentiel P à coefficients constants sur $\mathbb{R}^d$ possède une solution fondamentale E dont le support singulier est $\{0\}$, et si S est une distribution telle que PS soit $\mathscr{C}^\infty$ sur un ouvert $\omega \subset \mathbb{R}^d$, alors S elle-même est $\mathscr{C}^\infty$ sur ω.*

Démonstration : Soit a un point de ω. On veut montrer que S est $\mathscr{C}^\infty$ au voisinage de a. Soient donc ω' un voisinage ouvert relativement compact de a dans ω et $\theta \in \mathscr{D}(\omega)$ égale à 1 au voisinage de $\overline{\omega'}$. Alors la distribution $\theta.S$ possède un support compact, et $\theta.S$ est égale à S au voisinage de tout point de ω'. Il en résulte que $P(\theta.S)$ est égale à PS au voisinage de tout point de ω', donc est $\mathscr{C}^\infty$ sur ω', puisque PS l'est.
De plus, $\mathrm{supp}(P(\theta.S)) \subset \mathrm{supp}(\theta.S) \subset \mathrm{supp}(\theta)$, ce qui montre que $P(\theta.S) \in \mathscr{E}'(\mathbb{R}^d)$. Il résulte alors de la proposition 7.12.2, que $E * P(\theta.S)$ est solution de l'équation $PT = P(\theta.S)$, et que c'est l'unique solution à support compact de cette équation. On en déduit que $\theta.S = E * P(\theta.S)$. Alors

$$\mathrm{singsupp}(\theta.S) \subset \mathrm{singsupp}(E) + \mathrm{singsupp}(P(\theta.S)) = \mathrm{singsupp}(P(\theta.S)) \subset \mathbb{R}^d \setminus \omega' ,$$

dont on déduit que $\theta.S$ est $\mathscr{C}^\infty$ sur ω', et donc aussi S, ce qui montre que S est $\mathscr{C}^\infty$ au voisinage de tout point a de ω. ■

La dérivation dans $\mathbb{R}$.

Si P est l'opérateur différentiel δ_0' sur $\mathbb{R}$, une solution fondamentale de P est la fonction de Heaviside Y, puisque $Y' = \delta_0$ comme il ressort aisément de la "formule des sauts". Plus généralement, si P est l'opérateur différentiel $\delta_0' - \lambda\delta_0$, avec $\lambda \in \mathbb{C}$, on vérifie immédiatement que $Y(t)\,\mathrm{e}^{\lambda t}$ est une solution fondamentale de P à support dans $\mathbb{R}^+$.

Le laplacien dans $\mathbb{R}^2$

Comme il a été vu en 7.7.8, la fonction $f : x \mapsto -\dfrac{1}{2\pi}\log\dfrac{1}{\|x\|}$ sur $\mathbb{R}^2$ vérifie $\Delta f = \delta_0$. C'est donc une solution fondamentale du laplacien en dimension 2.

Le laplacien dans $\mathbb{R}^3$

On cherche une solution fondamentale du laplacien dans $\mathbb{R}^3$, c'est-à-dire de l'opérateur différentiel $\Delta = \dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2} + \dfrac{\partial^2}{\partial z^2}$. Soit f la fonction : $u = (x,y,z) \mapsto \dfrac{1}{\|u\|}$. Puisque f est localement intégrable sur $\mathbb{R}^3$ et $\mathscr{C}^\infty$ sur l'ouvert $\mathbb{R}^3 \setminus \{0\}$, la distribution Δf est sur $\mathbb{R}^3 \setminus \{0\}$ la fonction $\mathscr{C}^\infty$: $\dfrac{\partial^2 f}{\partial x^2} + \dfrac{\partial^2 f}{\partial y^2} + \dfrac{\partial^2 f}{\partial z^2}$.
Pour une fonction g de classe $\mathscr{C}^2$ sur $\mathbb{R}$, la fonction $g_r : u \mapsto g(\|u\|)$ sur $\mathbb{R}^3$ vérifie $\dfrac{\partial g_r}{\partial x}(u) = g'(\|u\|)\dfrac{x}{\|u\|}$, $\dfrac{\partial g_r}{\partial y}(u) = g'(\|u\|)\dfrac{y}{\|u\|}$ et $\dfrac{\partial g_r}{\partial z}(u) = g'(\|u\|)\dfrac{z}{\|u\|}$, donc

$$\frac{\partial^2 g_r}{\partial x^2}(u) = g''(\|u\|)\frac{x^2}{\|u\|^2} + \frac{g'(\|u\|)(\|u\|^2 - x^2)}{\|u\|^3}$$

et

$$\Delta g_r(u) = g''(\|u\|)\frac{x^2+y^2+z^2}{\|u\|^2} + \frac{g'(\|u\|)}{\|u\|^3}(3\,\|u\|^2 - x^2 - y^2 - z^2) = g''(\|u\|) + 2\frac{g'(\|u\|)}{\|u\|} .$$

Lorsque $g(t) = \dfrac{1}{t}$, on trouve $\Delta g_r = 0$. Il en résulte que le support de la distribution Δf est $\{0\}$. De plus, la fonction f est homogène de degré -1. Il en résulte que Δf est une distribution homogène de degré -3 et le théorème 7.8.3 montre alors que $\Delta f = c.\delta_0$. Alors en prenant une fonction radiale $\varphi(u) = g(\|u\|)$ à support compact, on trouve

$$\begin{aligned} c\varphi(0) &= \int \Big(g''(\|u\|) + 2\frac{g'(\|u\|)}{\|u\|}\Big)\frac{du}{\|u\|} = \int (g''(r) + 2\frac{g'(r)}{r})\frac{r^2 \cos\beta \, dr\, d\beta\, d\theta}{r} \\ &= 2\pi \int_{-\pi/2}^{\pi/2} cos\beta \, d\beta \int_0^{\infty} (rg''(r) + 2g'(r))\, dr \\ &= 4\pi \int_0^{\infty} (rg'(r) + g(r))'\, dr = -4\pi g(0) = -4\pi\varphi(0) \ . \end{aligned}$$

On en déduit finalement que $c = -4\pi$, et que la fonction $u \mapsto -\dfrac{1}{4\pi\,\|u\|}$ est une solution fondamentale de Δ.

L'OPÉRATEUR $\bar{\partial}$ DE CAUCHY-RIEMANN DANS $\mathbb{C}$

On définit sur $\mathbb{C} \simeq \mathbb{R}^2$ les opérateurs différentiels $\dfrac{\partial}{\partial z} = \dfrac{1}{2}(\dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y})$ et $\dfrac{\partial}{\partial \bar{z}} = \dfrac{1}{2}(\dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y})$. Pour simplifier, on notera aussi $\bar{\partial}$ le second. On vérifie aisément que $\dfrac{\partial}{\partial \bar{z}} * \dfrac{\partial}{\partial z} = \dfrac{1}{4}\Delta$. Et si E est une solution fondamentale de Δ, on doit avoir

$$4\frac{\partial}{\partial \bar{z}} * (\frac{\partial E}{\partial z}) = \Delta * E = \delta_0 \ ,$$

ce qui entraîne que $4\dfrac{\partial E}{\partial z}$ est une solution fondamentale de $\bar{\partial}$. On a vu plus haut que la fonction $h : z \mapsto -\dfrac{1}{2\pi}\log\dfrac{1}{|z|}$ était une solution fondamentale de Δ. Donc, puisque

$$\frac{\partial h}{\partial z} = -\frac{1}{4\pi}(\frac{-x}{x^2+y^2} - i\frac{-y}{x^2+y^2}) = \frac{1}{4\pi(x+iy)} = \frac{1}{4\pi z} \ ,$$

on obtient la fonction localement intégrable $\dfrac{1}{\pi z}$ comme solution fondamentale de $\bar{\partial}$.

Pour chacun de ces exemples, on peut remarquer que le support singulier de la solution fondamentale trouvée est $\{0\}$, donc que le théorème 7.12.4 s'applique.

8

TRANSFORMATION DE FOURIER DES DISTRIBUTIONS

8.1 Fonctions à décroissance rapide

Définition 8.1.1. *Une fonction f sur $\mathbb{R}^d$ est dite* à décroissance rapide *si elle est indéfiniment différentiable et si chacune de ses dérivées partielles est dominée à l'infini par toute puissance négative de la norme, c'est-à-dire si, pour tout $\alpha \in \mathbb{N}^d$ et tout k,*

$$\sup_{x \in \mathbb{R}^d} (1 + \|x\|)^k . \partial_\alpha f(x) < \infty \ .$$

On note $\mathscr{S}(\mathbb{R}^d)$ (ou simplement $\mathscr{S}$ s'il n'y a pas dambigüité sur la dimension d) l'espace des fonctions à décroissance rapide sur $\mathbb{R}^d$, et on le munit de la famille (dénombrable) des semi-normes $p_{\alpha,k}$ définies, pour $\alpha \in \mathbb{N}^d$ et $k \in \mathbb{N}$, par

$$p_{\alpha,k}(f) = \sup_{x \in \mathbb{R}^d} (1 + \|x\|)^k . \left|\partial_\alpha f(x)\right|$$

Proposition 8.1.2. *L'espace $\mathscr{S}(\mathbb{R}^d)$ est un espace de Fréchet.*

Démonstration : Il suffit de montrer que $\mathscr{S}(\mathbb{R}^d)$ est complet. Si (f_n) est une suite de Cauchy dans $\mathscr{S}$, la suite $(\partial_\alpha f_n)$ converge uniformément pour tout $\alpha \in \mathbb{N}^d$ vers une fonction continue g_α sur $\mathbb{R}^d$. On voit alors que $\partial_k(g_\alpha) = \lim_n \partial_{\alpha+\tau_k} f_n = g_{\alpha+\tau_k}$, donc, par récurrence sur $|\alpha|$, que $g_\alpha = \partial_\alpha g_0$ et que g_0 est $\mathscr{C}^\infty$. Enfin, puisque pour tout k, la suite $((1+\|x\|)^k \partial_\alpha f_n)$ converge uniformément vers $(1 + \|x\|)^k \partial_\alpha g_0$, on conclut que la fonction g_0 est limite dans $\mathscr{S}$ de (f_n). ■

Théorème 8.1.3. *Si f appartient à $\mathscr{S}(\mathbb{R}^d)$, alors f est dans $\mathscr{C}_0(\mathbb{R}^d)$, ainsi que dans $L^p(\mathbb{R}^d)$ pour tout $p \in [1, \infty]$.*

Démonstration : Toute fonction $f \in \mathscr{S}(\mathbb{R}^d)$ est $\mathscr{C}^\infty$, et en particulier continue. De plus, puisque $\sup_x (1 + \|x\|)^{d+1} |f(x)| < \infty$, on voit que $f(x) \to 0$ lorsque $\|x\| \to \infty$, et que $f \in L^1(\mathbb{R}^d)$. Et puisque $f \in \mathscr{C}_0(\mathbb{R}^d) \cap L^1(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, on a $f \in L^p(\mathbb{R}^d)$ pour tout $p \in [1, \infty]$. ■

Théorème 8.1.4. *L'espace $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\mathscr{S}(\mathbb{R}^d)$.*

Démonstration : Soient $f \in \mathscr{S}(\mathbb{R}^d)$, $q \in \mathbb{N}$ et $\varepsilon > 0$. On va montrer l'existence d'une fonction $\varphi \in \mathscr{D}(\mathbb{R}^d)$ telle que $p_{\alpha,q}(f-\varphi) < \varepsilon$, pour tout $\alpha \in \mathbb{N}^d$ tel que $|\alpha| \leqslant q$. Il existe en effet une fonction positive $\chi \in \mathscr{D}(\mathbb{R}^d)$ telle que $\chi(x) = 1$ si $\|x\| \leqslant 1$ et $\chi(x) = 0$ si $\|x\| \geqslant 2$. Alors la fonction $\chi_n : x \mapsto \chi(2^{-n}x)$ est dans $\mathscr{D}(\mathbb{R}^d)$, ainsi que $f_n : x \mapsto f(x).\chi_n(x)$ pour tout n. Et pour tout α, il existe des coefficients $c_{\alpha,\beta}$ tels que $c_{\alpha,\alpha} = 1$ et

$$\partial_\alpha(f - f_n)(x) = \sum_{\beta \leqslant \alpha} c_{\alpha,\beta}.\partial_\beta f(x).\partial_{\alpha-\beta}(1-\chi_n)(x)$$

Il en résulte que

$$\begin{aligned} p_{\alpha,q}(f-f_n) &\leqslant \sup_{\|x\| \geqslant 2^n} (1+\|x\|)^q \, |\partial_\alpha f(x)| + \sum_{\beta < \alpha} c_{\alpha,\beta}.p_{\beta,q}(f).2^{-n(|\alpha|-|\beta|)} \sup_y |\partial_{\alpha-\beta}\chi(y)| \\ &\leqslant 2^{-n} p_{\alpha,q+1}(f) + 2^{-n} \sum_{\beta<\alpha} c_{\alpha,\beta}.m_{\alpha-\beta}.p_{\beta,q}(f) \ , \end{aligned}$$

si l'on note $m_\gamma = \sup_y |\partial_\gamma \chi(y)|$. Et on conclut que $f_n \to f$ dans $\mathscr{S}(\mathbb{R}^d)$. ∎

Théorème 8.1.5. *Si $f \in \mathscr{S}(\mathbb{R}^d)$ et $\alpha \in \mathbb{N}^d$, alors $\partial_\alpha f \in \mathscr{S}(\mathbb{R}^d)$.*

Démonstration : Il suffit de remarquer que $\partial_\alpha f$ est $\mathscr{C}^\infty$ et que $p_{\beta,q}(\partial_\alpha f) = p_{\alpha+\beta,q}(f)$ est fini pour tout β et tout q. ∎

Proposition 8.1.6. *Si $f \in \mathscr{S}(\mathbb{R}^d)$ et $a \in \mathbb{R}^d$, la fonction $f_a = x \mapsto f(x-a)$ est dans $\mathscr{S}(\mathbb{R}^d)$.*

Démonstration : On a, pour tout $x \in \mathbb{R}^d$:

$$(1+\|a\|)(1+\|x-a\|) \geqslant 1 + \|a\| + \|x-a\| \geqslant 1 + \|x\|$$

donc $\dfrac{1+\|x\|}{1+\|x-a\|} \leqslant 1 + \|a\|$. On voit alors que $p_{\alpha,q}(f_a) \leqslant (1+\|a\|)^q p_{\alpha,q}(f)$, et on en déduit que la fonction f_a, qui est $\mathscr{C}^\infty$, est dans $\mathscr{S}(\mathbb{R}^d)$. ∎

Théorème 8.1.7. *Les parties relativement compactes de $\mathscr{S}(\mathbb{R}^d)$ sont les parties bornées.*

Démonstration : Si H est une partie relativement compacte de $\mathscr{S}(\mathbb{R}^d)$ et p une semi-norme continue sur $\mathscr{S}(\mathbb{R}^d)$, les ensembles $V_n = \{x : p(x) < n\}$ sont ouverts et la suite croissante d'ouverts (V_n) recouvre $\overline{H}$: l'un d'entre eux contient donc $\overline{H}$; et ceci montre que p est bornée sur H. Donc H est borné.

Inversement, si H est borné, chacune des semi-normes $p_{\alpha,k}$, pour $k \in \mathbb{N}$ et $\alpha \in \mathbb{N}^d$, est uniformément bornée sur H par un nombre $M_{\alpha,k}$. On considère alors l'espace compact $\mathbb{R}^\bullet_d = \mathbb{R}^d \cup \{\infty\}$, compactifié d'Alexandroff de $\mathbb{R}^d$ (cf. 1.5.38), qui est homéomorphe à la sphère unité de $\mathbb{R}^{d+1}$.
On définit sur $\mathscr{S}(\mathbb{R}^d)$ l'application Φ à valeurs dans $Q = \prod_{\alpha \in \mathbb{N}^d, k \in \mathbb{N}} \mathscr{C}(\mathbb{R}^\bullet_d, \mathbb{C})$ par ses fonctions coordonnées $\varphi_{\alpha,k} : f \mapsto (1+\|x\|)^q \partial_\alpha f$ (avec $\varphi_{\alpha,k}(f)(\infty) = 0$). Chacune des fonctions coordonnées de Φ est continue, et Φ est donc continue. On voit également que si une suite (f_n) dans $\mathscr{S}$ vérifie $\Phi(f_n) \to \theta \in Q$, alors la suite (f_n) converge dans $\mathscr{S}$ vers une

fonction f telle que $\Phi(f) = \theta$: il en résulte que Φ est un homéomorphisme de $\mathscr{S}$ sur le fermé $\Phi(\mathscr{S})$ de Q.
On vérifie ensuite, comme dans la démonstration du théorème 7.1.3 que $\varphi_{\alpha,k}(H)$ est équicontinu en chaque point x de $\mathbb{R}_d^\bullet$: ceci résulte de la formule des accroissements finis pour $x \in \mathbb{R}^d$ (en utilisant que les semi-normes $p_{\alpha+\tau_j,k}$ sont bornées sur H) et , pour $x = \infty$, du fait que $p_{\alpha,k+1}$ est bornée sur H. Il résulte alors du théorème d'Ascoli (1.10.6) que $\varphi_{\alpha,k}(H)$ est relativement compact dans $\mathscr{C}(\mathbb{R}_d^\bullet, \mathbb{C})$, puis du théorème de Tychonoff que $\Phi(H)$ est relativement compact dans Q. Donc $\overline{\Phi(H)}$ est compact dans $\overline{\Phi(\mathscr{S})} = \Phi(\mathscr{S})$, et H est contenu dans le compact $\Phi^{-1}(\overline{\Phi(H)})$. ■

Théorème 8.1.8. *Une partie H de $\mathscr{S}(\mathbb{R}^d)$ est relativement compacte dans $\mathscr{S}(\mathbb{R}^d)$ si et seulement si toute forme linéaire continue sur $\mathscr{S}$ est bornée sur H.*

Démonstration : Si H est relativement compacte dans $\mathscr{S}$, toute forme linéaire continue sur $\mathscr{S}$ est bornée sur le compact $\overline{H}$, et a fortiori sur H.
Inversement, si toute forme linéaire continue sur $\mathscr{S}$ est bornée sur H, il résulte du théorème 4.5.3 que H est borné dans $\mathscr{S}$. La conclusion découle alors du théorème 8.1.7. ■

On va maintenant présenter quelques opérations préservant l'espace $\mathscr{S}(\mathbb{R}^d)$.

Définitions 8.1.9. *On notera $\mathscr{O}_M(\mathbb{R}^d)$ l'ensemble des fonctions f de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^d$ à* croissance lente, *c'est-à-dire vérifiant pour tout α : $|\partial_\alpha f(x)| \leqslant C.(1+\|x\|)^q$ pour un entier q et une constante C convenables.*
On notera $L^1_S(\mathbb{R}^d)$ l'ensemble des fonctions mesurables f sur $\mathbb{R}^d$ telles que $(1+\|x\|)^q.f$ soit intégrable pour tout entier q.

Il est clair que tout polynôme appartient à $\mathscr{O}_M$ et aussi que, pour tout $\varepsilon > 0$, la fonction $x \mapsto \mathrm{e}^{-\varepsilon\|x\|^2}$ appartient à L^1_S .

Théorème 8.1.10. *Si $f \in \mathscr{S}(\mathbb{R}^d)$ et $\varphi \in \mathscr{O}_M(\mathbb{R}^d)$, le produit $f.\varphi$ est dans $\mathscr{S}(\mathbb{R}^d)$. De plus, pour $\varphi \in \mathscr{O}_M$, l'application linéaire $f \mapsto f.\varphi$ est continue de $\mathscr{S}$ dans $\mathscr{S}$.*

Démonstration : Soient $\varphi \in \mathscr{O}_M$, $q \in \mathbb{N}$ et $\alpha \in \mathbb{N}^d$. On voit d'abord que $f.\varphi$ est $\mathscr{C}^\infty$. Il résulte ensuite de la formule de Leibniz qu'existent des constantes positives $c_{\alpha,\beta}$ telles que :

$$(1+\|x\|)^q \partial_\alpha(f.\varphi)(x) = \sum_{\beta \leqslant \alpha} c_{\alpha,\beta}(1+\|x\|)^q.\partial_\beta f(x).\partial_{\alpha-\beta}\varphi(x) \ .$$

Alors, si pour $\beta \leqslant \alpha$, on a $|\partial_{\alpha-\beta}\varphi(x)| \leqslant C_\beta(1+\|x\|)^{q_\beta}$, on voit que

$$p_{\alpha,q}(f.\varphi) \leqslant \sum_{\beta \leqslant \alpha} c_{\alpha,\beta}.C_\beta.p_{\beta,q+q_\beta}(f) < \infty \ ,$$

ce qui montre que $f.\varphi \in \mathscr{S}(\mathbb{R}^d)$, et que l'application linéaire $M_\varphi : f \mapsto f.\varphi$ est continue de $\mathscr{S}$ dans $\mathscr{S}$. ■

Théorème 8.1.11. *Si $f \in \mathscr{S}(\mathbb{R}^d)$ et $g \in L^1_S(\mathbb{R}^d)$, alors $f * g \in \mathscr{S}(\mathbb{R}^d)$. De plus, pour $g \in L^1_S$, l'application linéaire $f \mapsto f * g$ est continue de $\mathscr{S}$ dans $\mathscr{S}$.*

Démonstration : Comme f est $\mathscr{C}^\infty$ et intégrable, la fonction $f * g$ est $\mathscr{C}^\infty$. Puisque $\partial_\alpha(f * g) = \partial_\alpha f * g$, il suffit de montrer que $(1+\|x\|)^q.f * g$ est bornée pour toute $f \in \mathscr{S}$. Pour cela on remarque que

$$(1+\|x\|)^q\,|f * g(x)| \leqslant \int |f(x-y)|\,.\,|g(y)|\,.(1+\|x-y\|+\|y\|)^q\,dy$$

et puisque

$$(1 + \|x-y\| + \|y\|)^q \leqslant \Big(2\max(1+\|x-y\|\,, 1+\|y\|)\Big)^q \leqslant 2^q\big((1+\|x-y\|)^q + (1+\|y\|)^q\big)\ ,$$

on voit que $(1+\|x\|)^q\,|f*g(x)| \leqslant 2^q(f_1*|g|\,(x)+|f|*g_1(x))$, où $f_1(x) = (1+\|x\|)^q\,|f(x)|$ et $g_1(x) = (1+\|x\|)^q.\,|g(x)|$. Par hypothèse, g et g_1 sont dans $L^1(\mathbb{R}^d)$. Et puisque $(1+\|x\|)^q.f$ est bornée, on voit que f et f_1 sont dans L^∞. On en déduit que $f*g_1$ et f_1*g sont continues et bornées, donc que $(1+\|x\|^q).f*g$ est bornée.

Les inégalités précédentes fournissent une majoration uniforme de $(1+\|x\|^q).\partial_\alpha(f*g)$ en fonction des $p_{\alpha,r}(f)$, et prouvent la continuité de l'application linéaire $f\mapsto f*g$ de $\mathscr{S}$ dans $\mathscr{S}$. ■

La principale propriété de $\mathscr{S}(\mathbb{R}^d)$, que ne partage pas l'espace $\mathscr{D}(\mathbb{R}^d)$ des fonctions de test, est sa stabilité par transformation de Fourier.

Lemme 8.1.12. *Si $f\in\mathscr{S}(\mathbb{R}^d)$, alors, pour tout entier q, la fonction $\xi\mapsto(1+\|\xi\|)^q\hat{f}(\xi)$ est bornée sur $\mathbb{R}^d$.*

Démonstration : Il suffit pour cela de montrer que la fonction $g:\xi\mapsto(1+\|\xi\|^2)^q\hat{f}(\xi)$ est bornée : en effet $1+\|\xi\|\leqslant 2.(1+\|\xi\|^2)$. Mais puisque, en vertu du lemme 6.6.5, g est la transformée de Fourier de $f_q=(I-\Delta)^q f$, il suffit de remarquer que f_q appartient à $\mathscr{S}(\mathbb{R}^d)$ d'après le théorème 8.1.5, donc que $f_q\in L^1(\mathbb{R}^d)$ et que $\|g\|_\infty = \left\|\hat{f}_q\right\|_\infty \leqslant \|f_q\|_1 < +\infty$. ■

Théorème 8.1.13. *Si f est dans $\mathscr{S}(\mathbb{R}^d)$, sa transformée de Fourier $\mathscr{F}f$ est dans $\mathscr{S}$. La transformation de Fourier est un isomorphisme de $\mathscr{S}$ sur $\mathscr{S}$ et, si on définit $\overline{\mathscr{F}}$ par $\overline{\mathscr{F}}f=\mathscr{F}(\tilde{f})$, on a $\overline{\mathscr{F}}\mathscr{F}f(x)=\mathscr{F}\overline{\mathscr{F}}f(x)=(2\pi)^d f(x)$.*

Démonstration : D'après le lemme précédent, pour montrer que $\mathscr{F}f=\hat{f}$ est dans $\mathscr{S}(\mathbb{R}^d)$, il suffit de montrer que $\partial_\alpha\hat{f}$ est, pour tout $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_d)\in\mathbb{N}^d$, la transformée de Fourier d'une fonction $g_\alpha\in\mathscr{S}(\mathbb{R}^d)$. Mais le lemme 6.6.6 montre que c'est bien le cas avec $g_\alpha(x)=f(x).\prod_{j=1}^d(ix_j)^{\alpha_j}$, car il résulte du théorème 8.1.10 que $g_\alpha\in\mathscr{S}(\mathbb{R}^d)$.

On montre de même que $\overline{\mathscr{F}}f\in\mathscr{S}(\mathbb{R}^d)$ si $f\in\mathscr{S}(\mathbb{R}^d)$. Et puisque le lemme 8.1.3 prouve que $\mathscr{S}(\mathbb{R}^d)\subset L^1(\mathbb{R}^d)$, il résulte du théorème d'inversion 6.6.12 que $\overline{\mathscr{F}}\mathscr{F}f(x)=(2\pi)^d f(x)$. Ce même théorème d'inversion, appliqué à $\tilde{f}$, montre que $\mathscr{F}\overline{\mathscr{F}}f(x)=(2\pi)^d f(x)$.
Et ceci entraîne que $\mathscr{F}$ est bijective de $\mathscr{S}(\mathbb{R}^d)$ sur lui-même, et que $\mathscr{F}^{-1}g$ est la fonction $x\mapsto(2\pi)^{-d}\overline{\mathscr{F}}g(x)$. Enfin la continuité de $\mathscr{F}$ sur $\mathscr{S}(\mathbb{R}^d)$ résulte immédiatement du théorème du graphe fermé : en effet, pour $(f,g)\in\mathscr{S}\times\mathscr{S}$, on a

$$g=\mathscr{F}f \iff \forall\xi\in\mathbb{R}^d \quad g(\xi)=\langle f,e_\xi\rangle\ ,$$

où e_ξ est la fonction continue bornée : $x\mapsto \mathrm{e}^{-ix\cdot\xi}$. On a $|g(\xi)|\leqslant p_{0,0}(g)$; et puisque $\mathscr{S}$ s'injecte continuement dans L^1, la fonction $f\mapsto\langle f,e_\xi\rangle$ est continue sur $\mathscr{S}$. Le même argument prouve la continuité de $\overline{\mathscr{F}}$, et donc celle de $\mathscr{F}^{-1}$. ■

Lemme 8.1.14. *Soient φ et ψ deux éléments de $\mathscr{S}(\mathbb{R}^d)$. Alors $\mathscr{F}(\varphi.\psi)=(2\pi)^{-d}\mathscr{F}\varphi*\mathscr{F}\psi$.*

Démonstration : Les fonctions $\mathscr{F}\varphi$ et $\mathscr{F}\psi$ appartiennent à $\mathscr{S}$, on a donc $\mathscr{F}\varphi*\mathscr{F}\psi\in\mathscr{S}$ et $\overline{\mathscr{F}}(\mathscr{F}\varphi*\mathscr{F}\psi)=\overline{\mathscr{F}}\mathscr{F}\varphi.\overline{\mathscr{F}}\mathscr{F}\psi=(2\pi)^d\varphi.(2\pi)^d\psi=(2\pi)^{2d}\varphi.\psi\in\mathscr{S}$. Il en résulte que

$$(2\pi)^d(\mathscr{F}\varphi*\mathscr{F}\psi)=(\mathscr{F}\overline{\mathscr{F}})(\mathscr{F}\varphi*\mathscr{F}\psi)=\mathscr{F}\Big(\overline{\mathscr{F}}(\mathscr{F}\varphi*\mathscr{F}\psi)\Big)=(2\pi)^{2d}\mathscr{F}(\varphi\psi)\ ,$$

et le résultat cherché en découle. ■

8.2 Distributions tempérées

Définition 8.2.1. *On appelle* distribution tempérée sur $\mathbb{R}^d$ *toute forme linéaire continue sur $\mathscr{S}(\mathbb{R}^d)$. Et on note $\mathscr{S}'(\mathbb{R}^d)$ l'espace des distributions tempérées, qu'on munit de la topologie faible $\sigma(\mathscr{S}'(\mathbb{R}^d), \mathscr{S}(\mathbb{R}^d))$.*

Théorème 8.2.2. *Si f appartient à $\mathscr{S}(\mathbb{R}^d)$, elle est localement intégrable et la distribution associée est une distribution tempérée.*

Démonstration : Si $f \in \mathscr{S}(\mathbb{R}^d)$, elle est continue et bornée, donc localement intégrable (et même intégrable). De plus, pour $\varphi \in \mathscr{S}(\mathbb{R}^d)$, on a

$$\begin{aligned} \left| \int f(x).\varphi(x)\, dx \right| &\leqslant \sup_x |f(x)| \,.\, \sup_y (1+\|x\|)^{d+1} \,|\varphi(x)| \,.\, \int \frac{dx}{(1+\|x\|)^{d+1}} \\ &\leqslant C.p_{0,0}(f).p_{0,d+1}(\varphi) \end{aligned}$$

ce qui montre la continuité de $\varphi \mapsto \langle T_f, \varphi \rangle$ sur $\mathscr{S}(\mathbb{R}^d)$. ■

Théorème 8.2.3. *Si S est une distribution tempérée sur $\mathbb{R}^d$, la restriction T de S à $\mathscr{D}(\mathbb{R}^d)$ est une distribution sur $\mathbb{R}^d$ vérifiant pour toute fonction de test φ :*

$$|\langle T, \varphi \rangle| \leqslant C. \sup_{x \in \mathbb{R}^d, |\alpha| \leqslant q} (1+\|x\|)^q \,|\partial_\alpha \varphi(x)| \tag{$*$}$$

pour une constante C et un entier q. Inversement, une distribution T satisfaisant la condition précédente se prolonge de façon unique en une distribution tempérée.

Démonstration : Compte tenu de la définition de la topologie de $\mathscr{S}(\mathbb{R}^d)$, toute forme linéaire continue S sur $\mathscr{S}(\mathbb{R}^d)$ vérifie la condition $(*)$ pour une certaine constante C, un certain entier q et toute fonction $\varphi \in \mathscr{S}(\mathbb{R}^d)$. La restriction T de S à $\mathscr{D}(\mathbb{R}^d)$ vérifie donc a fortiori cette inégalité.
Inversement, si T est une distribution satisfaisant l'inégalité $(*)$, elle définit une forme linéaire continue sur $\mathscr{D}(\mathbb{R}^d)$ muni de la topologie induite par $\mathscr{S}(\mathbb{R}^d)$, et se prolonge de façon unique par continuité uniforme à $\mathscr{S}(\mathbb{R}^d)$, puisque $\mathscr{D}(\mathbb{R}^d)$ est dense dans $\mathscr{S}(\mathbb{R}^d)$ (cf. 8.1.4). ■

On identifiera donc désormais les distributions tempérées avec les distributions satisfaisant la condition $(*)$.

Théorème 8.2.4. *Toute distribution à support compact est tempérée. Toute distribution tempérée est d'ordre fini.*

Démonstration : Si $T \in \mathscr{E}'(\mathbb{R}^d)$, il existe C, q et R tels que, pour tout $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on ait $|\langle T, \varphi \rangle| \leqslant C. \sup_{\|x\| \leqslant R, |\alpha| \leqslant q} |\partial_\alpha \varphi(x)| \leqslant C. \sup_{x \in \mathbb{R}^d, |\alpha| \leqslant q} (1+\|x\|)^q \,|\partial_\alpha \varphi(x)|$. Et ceci montre que T satisfait la condition $(*)$.
Et si $T \in \mathscr{S}'(\mathbb{R}^d)$, elle satisfait la condition $(*)$ avec des constantes C et q : on voit alors que T est d'ordre au plus q. ■

Théorème 8.2.5. *Toute distribution homogène sur $\mathbb{R}^d$ est tempérée.*

Démonstration : Soit T une distribution sur $\mathbb{R}^d$ homogène de degré k. Pour $n \in \mathbb{N}$, notons Φ_n l'homothétie $x \mapsto 2^n.x$.

Soit alors $\chi \in \mathscr{D}(\mathbb{R}^d)$ égale à 1 au voisinage de la boule $B(0,1)$ et à support dans $B(0,2)$. La distribution $\chi.T$ est à support compact, donc est tempérée.
On pose également $\chi_n(x) = \chi(2^{-n-1}x) - \chi(2^{-n}x)$. On a donc $\chi_n(x) = \chi_0(2^{-n}x)$, en particulier. Si $\varphi \in \mathscr{D}(\mathbb{R}^d)$ a son support dans $B(0,2^m)$, on a $\chi(2^{-n}x).\varphi(x) = \varphi(x)$ et $\chi_n.\varphi = 0$ si $n \geqslant m$. Il en résulte que $\varphi = \chi.\varphi + \sum_{n=0}^{m} \chi_n.\varphi = \chi.\varphi + \sum_{n=0}^{\infty} \chi_n.\varphi$, et que

$$\langle (1-\chi).T, \varphi\rangle = \sum_{n=0}^{\infty} \langle \chi_n.T, \varphi\rangle \ .$$

Puisque T est homogène de degré k et que $\chi_n \circ \Phi_n = \chi_0$, , on a

$$\langle T, \chi_n.\varphi\rangle = 2^{n(d+k)}\langle T, (\chi_n.\varphi)\circ\Phi_n\rangle = 2^{n(d+k)}\langle T, \chi_0.\varphi\circ\Phi_n\rangle = 2^{n(k+d)}\langle \chi_0.T, \varphi\circ\Phi_n\rangle \ .$$

Et puisque $\chi_0.T$ est une distribution dont le support est un compact disjoint de la boule fermée $\tilde{B}(0,1)$, il existe C et $m \in \mathbb{N}$ tels que $|\langle \chi_0.T, \varphi\rangle| \leqslant C.\sup_{\|x\|\geqslant 1, |\alpha|\leqslant m} |\partial_\alpha \varphi(x)|$. On en déduit que

$$|\langle \chi_0.T, \varphi\circ\Phi_n\rangle| \leqslant C. \sup_{\|x\|\geqslant 1, |\alpha|\leqslant m} |\partial_\alpha(\varphi\circ\Phi_n)(x)| = C. \sup_{\|x\|\geqslant 2^n, |\alpha|\leqslant m} 2^{n|\alpha|}\,|\partial_\alpha \varphi(x)| \ .$$

Et puisque $2^{nq}.\sup_{\|x\|\geqslant 2^n} |\partial_\alpha\varphi(x)| \leqslant \sup_x (1+\|x\|)^q\,|\partial_\alpha \varphi(x)|$, on obtient

$$|\langle \chi_n.T, \varphi\rangle| \leqslant C.2^{n(k+d)}.2^{nm}.2^{-nq}\sup_x (1+\|x\|)^q\,|\partial_\alpha\varphi(x)| \ ,$$

ce qui montre que, si $q \geqslant m+k+d+1$, on a

$$|\langle (1-\chi).T, \varphi\rangle| \leqslant C.\Big(\sum_{n=0}^{\infty} 2^{n(m+k+d-q)}\Big).\sup_x (1+\|x\|)^q\,|\partial_\alpha\varphi(x)| \leqslant 2C\cdot p_{\alpha,q}(\varphi) \ ,$$

et finalement que $(1-\chi).T$ est tempérée. Ceci achève la démonstration puisque $\chi.T$ est tempérée. ■

Opérations sur les distributions tempérées.

Comme pour les distributions, on peut définir un certain nombre d'opérations sur $\mathscr{S}'(\mathbb{R}^d)$ par transposition d'applications linéaires continues de $\mathscr{S}(\mathbb{R}^d)$ dans $\mathscr{S}(\mathbb{R}^d)$.

Théorème 8.2.6. *Pour $\alpha \in \mathbb{N}^d$ et $T \in \mathscr{S}'(\mathbb{R}^d)$, la distribution $\partial_\alpha T$ est une distribution tempérée. De plus, l'application $T \mapsto \partial_\alpha T$ est continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même.*

Démonstration : La distribution $\partial_\alpha T$ est définie par $\langle \partial_\alpha T, \varphi\rangle = (-1)^{|\alpha|}\langle T, \partial_\alpha f\rangle$. On a vu plus haut que l'application $D_\alpha : \varphi \mapsto (-1)^{|\alpha|}\partial_\alpha\varphi$ est linéaire continue de $\mathscr{S}(\mathbb{R}^d)$ dans lui-même. Sa transposée ${}^tD_\alpha$ est donc linéaire continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même, et on vérifie que $\langle {}^tD_\alpha T, \varphi\rangle = (-1)^{|\alpha|}\langle T, \partial_\alpha f\rangle = \langle \partial_\alpha T, \varphi\rangle$. ■

Théorème 8.2.7. *Si f appartient à $\mathscr{O}_m(\mathbb{R}^d)$ et $T \in \mathscr{S}'(\mathbb{R}^d)$, l'application $\varphi \mapsto \langle T, f.\varphi\rangle$ est une distribution tempérée, appelée* produit de T par f *et notée $f.T$. De plus l'application $T \mapsto f.T$ est continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même.*

Démonstration : On a montré plus haut que l'application $M_f : \varphi \mapsto f.\varphi$ est linéaire et continue de $\mathscr{S}(\mathbb{R}^d)$ dans lui-même. La transposée ${}^tM_f$ est linéaire continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même, et on vérifie immédiatement que $\langle {}^tM_f(T), \varphi\rangle = \langle T, f.\varphi\rangle$. ■

Théorème 8.2.8. *Si $g \in L^1_S(\mathbb{R}^d)$ et $T \in \mathscr{S}'(\mathbb{R}^d)$, l'application définie sur $\mathscr{S}(\mathbb{R}^d)$ par $\varphi \mapsto \langle T, \tilde{g} * \varphi\rangle$ est une distribution tempérée notée $g * T$. De plus, l'application $T \mapsto g * T$ est continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même.*

Démonstration : La fonction $\tilde{g} : x \mapsto g(-x)$ est dans $L^1_S(\mathbb{R}^d)$. Il en résulte que l'application $C_g : \varphi \mapsto \tilde{g} * \varphi$ est continue de $\mathscr{S}(\mathbb{R}^d)$ dans lui-même. Et la transposée ${}^tC_g$ est linéaire continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même. On a $\langle {}^tC_gT, \varphi\rangle = \langle T, \tilde{g} * \varphi\rangle$. En particulier, si $g \in \mathscr{D}(\mathbb{R}^d) \subset L^1_S(\mathbb{R}^d)$, on retrouve la définition de la convolée $g * T$ donnée au chapitre précédent. ■

8.3 Transformation de Fourier des distributions tempérées

Théorème 8.3.1. *Si $f \in L^1(\mathbb{R}^d)$ et $\varphi \in \mathscr{D}(\mathbb{R}^d)$, la distribution $T_{\hat{f}}$ associée à la fonction $\hat{f}$ est liée à la distribution T_f associée à f par $\langle T_{\hat{f}}, \varphi\rangle = \langle T_f, \hat{\varphi}\rangle$.*

Démonstration : La fonction $(x, \xi) \mapsto f(x)\varphi(\xi)\,\mathrm{e}^{-ix\cdot\xi}$ est intégrable sur $\mathbb{R}^d \times \mathbb{R}^d$. On a, par Fubini :

$$\begin{aligned}\langle T_f, \hat{\varphi}\rangle = \int f(x)\hat{\varphi}(x)\,dx &= \int f(x)\Big(\int \varphi(\xi)\,\mathrm{e}^{-ix\cdot\xi}\,d\xi\Big)\,dx\\ &= \int \varphi(\xi)\Big(\int f(x)\,\mathrm{e}^{-ix\cdot\xi}\,dx\Big)\,d\xi = \int \hat{f}(\xi)\varphi(\xi)\,d\xi = \langle T_{\hat{f}}, \varphi\rangle\ ,\end{aligned}$$

ce qui achève la démonstration. ■

Il apparaît donc comme tentant de définir de façon générale la transformée d'une distribution T par la formule $\langle \mathscr{F}T, \varphi\rangle = \langle T, \hat{\varphi}\rangle$. Malheureusement, la fonction $\hat{\varphi}$ n'est en général pas dans $\mathscr{D}(\mathbb{R}^d)$ lorsque $\varphi \in \mathscr{D}(\mathbb{R}^d)$, et $\langle T, \hat{\varphi}\rangle$ n'est pas défini.
De fait, lorsque $\varphi \in \mathscr{D}(\mathbb{R}^d)$, la fonction $\hat{\varphi}$ se prolonge en une fonction holomorphe sur $\mathbb{C}^d$ et ne peut s'annuler sur un ouvert non vide de $\mathbb{R}^d$ sans être identiquement nulle : il en résulte que si φ et $\hat{\varphi}$ sont dans $\mathscr{D}(\mathbb{R}^d)$, alors $\varphi = \hat{\varphi} = 0$.
Néanmoins, si T est une distribution tempérée, la formule ci-dessus prend sens pour $\varphi \in \mathscr{S}(\mathbb{R}^d)$, puisque alors $\hat{\varphi} \in \mathscr{S}(\mathbb{R}^d)$.

Théorème 8.3.2. *Si T est une distribution tempérée sur $\mathbb{R}^d$, l'application définie sur $\mathscr{S}(\mathbb{R}^d)$ par $\varphi \mapsto \langle T, \mathscr{F}\varphi\rangle$ est une distribution tempérée, qu'on note $\mathscr{F}T$, et qu'on appelle* transformée de Fourier *de T. De même, l'application $\varphi \mapsto \langle T, \overline{\mathscr{F}}\varphi\rangle$ est une distribution tempérée notée $\overline{\mathscr{F}}T$. De plus, les applications $T \mapsto \mathscr{F}T$ et $T \mapsto \overline{\mathscr{F}}T$ sont des isomorphismes de $\mathscr{S}'(\mathbb{R}^d)$ sur lui-même, et on a $\mathscr{F}\overline{\mathscr{F}}T = \overline{\mathscr{F}}\mathscr{F}T = (2\pi)^d.T$.*

Démonstration : Il résulte du théorème 8.1.13 que $\mathscr{F}$ et $\overline{\mathscr{F}}$ sont des applications linéaires continues de $\mathscr{S}(\mathbb{R}^d)$ dans lui-même. Les transposées de ces applications sont les applications linéaires continues $\mathscr{F}$ et $\overline{\mathscr{F}}$ définies ci-dessus de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même.
De plus, si $T \in \mathscr{S}'(\mathbb{R}^d)$ et $\varphi \in \mathscr{S}(\mathbb{R}^d)$, on a

$$\langle \overline{\mathscr{F}}\mathscr{F}T, \varphi\rangle = \langle \mathscr{F}T, \overline{\mathscr{F}}\varphi\rangle = \langle T, \mathscr{F}\overline{\mathscr{F}}\varphi\rangle = \langle T, \overline{\mathscr{F}}\mathscr{F}\varphi\rangle = \langle \overline{\mathscr{F}}T, \mathscr{F}\varphi\rangle = \langle \mathscr{F}\overline{\mathscr{F}}T, \varphi\rangle\ .$$

Et comme $\mathscr{F}\overline{\mathscr{F}}\varphi = \overline{\mathscr{F}}\mathscr{F}\varphi = (2\pi)^d\varphi$, on obtient $\langle \overline{\mathscr{F}}\mathscr{F}T, \varphi\rangle = \langle \mathscr{F}\overline{\mathscr{F}}T, \varphi\rangle = (2\pi)^d\langle T, \varphi\rangle$, ce qui montre que $\overline{\mathscr{F}}\mathscr{F}T = \mathscr{F}\overline{\mathscr{F}}T = (2\pi)^d.T$. ■

Remarque 8.3.3. On a montré plus haut que, si f est dans $\mathscr{S}(\mathbb{R}^d)$, la distribution associée T_f est une distribution tempérée, qu'on identifie à f. On doit remarquer que la transformée de Fourier de f en tant que fonction de $\mathscr{S}$ coïncide avec la transformée de Fourier de T_f en tant que distribution tempérée, c'est-à-dire que pour toute $\varphi \in \mathscr{S}$, on a $\langle \mathscr{F}f, \varphi\rangle = \langle f, \mathscr{F}\varphi\rangle$. Et ceci résulte immédiatement de la démonstration du théorème 8.3.1, en remplaçant l'hypothèse " $\varphi \in \mathscr{D}$ " par l'hypothèse " $\varphi \in \mathscr{S}$ ".

Théorème 8.3.4. *Si T est une distribution tempérée sur $\mathbb{R}^d$ et $1 \leqslant k \leqslant d$, on a $\mathscr{F}(\partial_k T) = i\xi_k.\mathscr{F}T$. On a aussi $\partial_k(\mathscr{F}T) = \mathscr{F}(-ix_k.T)$.*

Démonstration : Soit $\varphi \in \mathscr{S}(\mathbb{R}^d)$. On a, par définition :

$$\langle \mathscr{F}(\partial_k T), \varphi\rangle = \langle \partial_k T, \mathscr{F}\varphi\rangle = -\langle T, \partial_k(\mathscr{F}\varphi)\rangle$$

et puisque, d'après 6.6.6, $\partial_k(\mathscr{F}\varphi)(\xi) = \mathscr{F}(-i\xi_k\varphi)(\xi)$, on voit que

$$\langle \mathscr{F}(\partial_k T), \varphi\rangle = \langle T, \mathscr{F}(i\xi_k\varphi)\rangle = \langle \mathscr{F}T, i\xi_k.\varphi\rangle = \langle i\xi_k\mathscr{F}T, \varphi\rangle \ ,$$

d'où l'égalité $\mathscr{F}(\partial_k T) = i\xi_k.\mathscr{F}T$. De même, en utilisant 6.6.5 :

$$\begin{aligned}\langle \partial_k(\mathscr{F}T), \varphi\rangle &= \langle \mathscr{F}T, -\partial_k\varphi\rangle = \langle T, -\mathscr{F}(\partial_k\varphi)\rangle \\ &= \langle T, -ix_k.\mathscr{F}\varphi\rangle = \langle -ix_k.T, \mathscr{F}\varphi\rangle = \langle \mathscr{F}(-ix_k.T), \varphi\rangle \ ,\end{aligned}$$

d'où on tire l'égalité $\partial_k(\mathscr{F}T) = \mathscr{F}(-ix_k.T)$. ■

Théorème 8.3.5. *Si S est une distribution tempérée sur $\mathbb{R}^d$ et si φ appartient à $\mathscr{S}(\mathbb{R}^d)$, alors $\varphi * S$ est une distribution tempérée et $\mathscr{F}(\varphi * S) = \hat{\varphi}.\mathscr{F}S$.*

Démonstration : Puisque $\tilde{\varphi}$ est dans $\mathscr{S}(\mathbb{R}^d)$, donc dans $L^1_S(\mathbb{R}^d)$, pour toute $\psi \in \mathscr{S}(\mathbb{R}^d)$ on a $\tilde{\varphi} * \psi \in \mathscr{S}(\mathbb{R}^d)$, et l'application $\psi \mapsto \tilde{\varphi} * \psi$ est continue de $\mathscr{S}$ dans $\mathscr{S}$: il en résulte que l'application $\psi \mapsto \langle S, \tilde{\varphi} * \psi\rangle$ est une forme linéaire continue sur $\mathscr{S}$, donc une distribution tempérée, qu'on notera $\varphi * S$ comme dans le cas où $\varphi \in \mathscr{D}(\mathbb{R}^d)$ et $S \in \mathscr{D}'(\mathbb{R}^d)$.
Alors, si $\chi \in \mathscr{S}(\mathbb{R}^d)$, on a

$$\langle \mathscr{F}(\varphi * S), \chi\rangle = \langle \varphi * S, \hat{\chi}\rangle = \langle S, \tilde{\varphi} * \hat{\chi}\rangle = \langle \mathscr{F}S, \mathscr{F}^{-1}(\tilde{\varphi} * \hat{\chi})\rangle \ .$$

Et puisque

$$\begin{aligned}\mathscr{F}^{-1}(\tilde{\varphi} * \hat{\chi}) &= (2\pi)^{-d}\overline{\mathscr{F}}(\tilde{\varphi} * \hat{\chi}) = (2\pi)^{-d}\overline{\mathscr{F}}(\tilde{\varphi} * \hat{\chi}) \\ &= (2\pi)^{-d}\mathscr{F}(\varphi).\overline{\mathscr{F}}(\hat{\chi}) = \mathscr{F}(\varphi)\mathscr{F}^{-1}(\hat{\chi}) \\ &= \mathscr{F}(\varphi).\mathscr{F}^{-1}(\hat{\chi}) = \hat{\varphi}.\chi \ ,\end{aligned}$$

on obtient

$$\langle \mathscr{F}(\varphi * S), \chi\rangle = \langle \mathscr{F}S, \hat{\varphi}.\chi\rangle = \langle \hat{\varphi}.\mathscr{F}S, \chi\rangle \ ,$$

ce qui montre que $\mathscr{F}(\varphi * S) = \hat{\varphi}.\mathscr{F}S$. ■

Lemme 8.3.6. *Si T est une distribution à support compact sur $\mathbb{R}^d$, alors $\mathscr{F}T$ est une fonction $\mathscr{C}^\infty$ sur $\mathbb{R}^d$, et on a pour tout ξ : $\mathscr{F}T(\xi) = \langle T, e_\xi\rangle$, où e_ξ désigne la fonction $\mathscr{C}^\infty$: $x \mapsto \mathrm{e}^{-ix\cdot\xi}$.*

Démonstration : Soient ρ une fonction $\mathscr{C}^\infty$ paire, positive, d'intégrale 1 et à support dans la boule unité, et, pour $0 < \varepsilon < 1$, ρ_ε la fonction $x \mapsto \varepsilon^{-d}\rho(\dfrac{x}{\varepsilon})$. Alors $T * \rho_\varepsilon$ appartient

à $\mathscr{D}(\mathbb{R}^d)$. Il résulte du théorème 8.3.5 que $\mathscr{F}(T*\rho_\varepsilon) = \mathscr{F}T.\widehat{\rho_\varepsilon}$. De plus, on a $\widehat{\rho_\varepsilon}(\xi) = \hat{\rho}(\varepsilon\xi)$, d'après 6.6.9.
Puisque $\hat{\rho}$ est $\mathscr{C}^\infty$ et que $\hat{\rho}(0) = \int \rho(x)\,dx = 1$, il existe un $\delta > 0$ tel que $\hat{\rho}(\xi) \neq 0$ si $\|\xi\| < \delta$. Alors, pour tout ξ_0, on peut trouver $\varepsilon > 0$ tel que $\varepsilon\,\|\xi_0\| < \delta$, et on voit que, pour ξ voisin de ξ_0, on a $\mathscr{F}T(\xi) = \dfrac{\mathscr{F}(T*\rho_\varepsilon)(\xi)}{\hat{\rho}(\varepsilon\xi)}$, et on en déduit que $\mathscr{F}T$ est, au voisinage de ξ_0, une fonction $\mathscr{C}^\infty$.
Puisque e_ξ est une fonction $\mathscr{C}^\infty$ et T une distribution à support compact, $T*\rho_\varepsilon$ est dans $\mathscr{D}(\mathbb{R}^d)$ et on a

$$\mathscr{F}(T*\rho_\varepsilon)(\xi) = \langle T*\rho_\varepsilon, e_\xi\rangle = \langle T, \widetilde{\rho_\varepsilon} * e_\xi\rangle$$

et on a $\widetilde{\rho_\varepsilon} * e_\xi(a) = \displaystyle\int \rho_\varepsilon(-x)e_\xi(a-x)\,dx = e_\xi(a)\int \rho_\varepsilon(-x)e_\xi(-x)\,dx = \widehat{\rho_\varepsilon}(\xi)e_\xi(a)$, donc

$$\mathscr{F}T(\xi).\widehat{\rho_\varepsilon}(\xi) = \mathscr{F}(T*\rho_\varepsilon)(\xi) = \langle T, \widetilde{\rho_\varepsilon} * e_\xi\rangle = \widehat{\rho_\varepsilon}(\xi)\langle T, e_\xi\rangle\ ,$$

dont on conclut que $\mathscr{F}T(\xi) = \langle T, e_\xi\rangle$ puisque, pour tout ξ, on peut trouver un $\varepsilon > 0$ tel que $\widehat{\rho_\varepsilon}(\xi) = \hat{\rho}(\varepsilon\xi) \neq 0$. ■

Lemme 8.3.7. *Si T est une distribution à support compact sur $\mathbb{R}^d$, alors $\mathscr{F}T$ appartient à $\mathscr{O}_M(\mathbb{R}^d)$.*

Démonstration : Puisque T est une distribution à support compact, il existe une constante M et un entier m tels que, pour toute fonction φ de classe $\mathscr{C}^\infty$, on ait $|\langle T,\varphi\rangle| \leqslant M\sup_{|\alpha|\leqslant m, \|x\|\leqslant R}|\partial_\alpha\varphi(x)|$. En particulier, pour la fonction $e_\xi : x \mapsto \mathrm{e}^{-ix\cdot\xi}$, on obtient

$$|\mathscr{F}T(\xi)| = |\langle T, e_\xi\rangle| \leqslant M \sup_{|\alpha|\leqslant m, \|x\|\leqslant R} |\partial_\alpha e_\xi(x)|\ .$$

Et puisque, pour $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d)$, on a

$$|\partial_\alpha e_\xi(x)| = \left|\prod_{j=1}^d (-i\xi_j)^{\alpha_j}.e_\xi(x)\right| \leqslant \prod_{j=1}^d \|\xi\|^{\alpha_j} = \|\xi\|^{|\alpha|}\ ,$$

on obtient $|\mathscr{F}T(\xi)| \leqslant M\sup_{|\alpha|\leqslant m}\|\xi\|^{|\alpha|} \leqslant M(1+\|\xi\|)^m$.
Pour montrer que $\mathscr{F}T$ appartient à $\mathscr{O}_M$, il reste à montrer que chacune des dérivées partielles de $\mathscr{F}T$ satisfait une majoration similaire. Mais puisque $\partial_\beta\mathscr{F}T = \mathscr{F}(T_\beta)$ où $T_\beta = \prod_{j=1}^d(-ix_j)^{\beta_j}.T$, on remarque que T_β est, comme T, une distribution à support compact, donc que $|\mathscr{F}(T_\beta)(\xi)|$ admet une majoration par un polynôme en $\|\xi\|$. ■

Théorème 8.3.8. *Si T est une distribution à support compact sur $\mathbb{R}^d$ et $\varphi \in \mathscr{S}(\mathbb{R}^d)$, alors $\varphi * T$ est dans $\mathscr{S}(\mathbb{R}^d)$. Et l'application linéaire $\varphi \mapsto \varphi * T$ est continue de $\mathscr{S}$ dans $\mathscr{S}$.*

Démonstration : Puisque $\varphi \in \mathscr{S}$ et $T \in \mathscr{E}' \subset \mathscr{S}'$, il résulte du théorème 8.3.5 que $\varphi * T$ appartient à $\mathscr{S}'$ et que $\mathscr{F}(\varphi * T) = \hat{\varphi}.\mathscr{F}T$. Puisque $\hat{\varphi} \in \mathscr{S}$ et que $\mathscr{F}T \in \mathscr{O}_M$ d'après le lemme précédent, on déduit du théorème 8.1.10 que $S = \mathscr{F}(\varphi * T)$ appartient à $\mathscr{S}$. Il existe donc un élément $S_1 = \mathscr{F}^{-1}(S)$ de $\mathscr{S}$ tel que $\mathscr{F}(S_1) = \mathscr{F}(\varphi * T)$. Et puisque $\mathscr{F}$ est un isomorphisme de $\mathscr{S}'$ sur $\mathscr{S}'$, on conclut que $\varphi * T = S_1 \in \mathscr{S}$.
La continuité de l'application linéaire $\varphi \mapsto \varphi * T$ de l'espace de Fréchet $\mathscr{S}$ dans lui-même résulte du théorème du graphe fermé : en effet, l'égalité $\psi = \varphi * T$ équivaut à la propriété :

$$\forall \xi \in \mathbb{R}^d \quad \hat{\psi}(\xi) = \hat{\varphi}(\xi)\cdot\mathscr{F}T(\xi)\ ,$$

et il est clair que, pour ξ fixé, les fonctions $\varphi \mapsto \hat{\varphi}(\xi)$ et $\psi \mapsto \hat{\psi}(\xi)$ sont continues sur $\mathscr{S}$. ■

Théorème 8.3.9. *Si T est une distribution à support compact sur $\mathbb{R}^d$ et S une distribution tempérée sur $\mathbb{R}^d$, alors $T * S$ est une distribution tempérée. De plus, l'application $S \mapsto T * S$ est continue de $\mathscr{S}'(\mathbb{R}^d)$ dans lui-même.*

DÉMONSTRATION : On a défini, pour $T \in \mathscr{E}'(\mathbb{R}^d)$ et $S \in \mathscr{D}'(\mathbb{R}^d)$ la convolution $T * S$ par l'égalité $\langle T * S, \varphi\rangle = \langle S, \varphi * \tilde{T}\rangle$ pour toute $\varphi \in \mathscr{D}(\mathbb{R}^d)$.
Il résulte du théorème précédent que, pour $T \in \mathscr{E}'(\mathbb{R}^d)$, l'application linéaire $C_T : \varphi \mapsto \tilde{T} * \varphi$ est continue de $\mathscr{S}$ dans $\mathscr{S}$. Alors sa transposée ${}^tC_T$ est linéaire continue de $\mathscr{S}'$ dans $\mathscr{S}'$. En particulier, pour $S \in \mathscr{S}'$ et $\varphi \in \mathscr{D} \subset \mathscr{S}$, on a

$$\langle {}^tC_T S, \varphi\rangle = \langle S, C_T\varphi\rangle = \langle S, \varphi * \tilde{T}\rangle = \langle T * S, \varphi\rangle \ ,$$

et ceci montre que la distribution $T * S$ est égale à la distribution tempérée ${}^tC_T S$. On voit aussi que l'application $S \mapsto T * S$ coïncide avec l'application continue ${}^tC_T$. ∎

Théorème 8.3.10. *Si T est une distribution homogène de degré k sur $\mathbb{R}^d$, alors $\mathscr{F}T$ est homogène de degré $-d-k$.*

DÉMONSTRATION : On a montré plus haut que si T est une distribution homogène, alors T est tempérée : on peut donc considérer la distribution tempérée $\mathscr{F}T$. Soit $\varphi \in \mathscr{S}(\mathbb{R}^d)$. Pour $\lambda > 0$, on considère la fonction $\varphi_\lambda : x \mapsto \varphi(\dfrac{x}{\lambda})$. On a

$$\widehat{\varphi_\lambda}(\xi) = \int \varphi(\frac{x}{\lambda})\,\mathrm{e}^{-ix\cdot\xi}\,dx = \lambda^d \int \varphi(u)\,\mathrm{e}^{-i\lambda u\cdot\xi}\,du = \lambda^d\hat{\varphi}(\lambda\xi) \ ,$$

donc, avec $h = -d-k$,

$$\begin{aligned}\langle \mathscr{F}T, \varphi_\lambda\rangle = \langle T, \mathscr{F}(\varphi_\lambda)\rangle = \lambda^d\langle T, (\mathscr{F}\varphi)_{1/\lambda}\rangle &= \lambda^d.(\frac{1}{\lambda})^{d+k}\langle T, \mathscr{F}\varphi\rangle \\ &= \lambda^{-k}\langle \mathscr{F}T, \varphi\rangle = \lambda^{d+h}\langle \mathscr{F}T, \varphi\rangle \ ,\end{aligned}$$

ce qui montre que $\mathscr{F}T$ est homogène de degré h. ∎

Théorème 8.3.11. *Si f et g sont dans $L^2(\mathbb{R}^d)$, on a $fg \in L^1(\mathbb{R}^d)$ et $\mathscr{F}(f.g) = (2\pi)^{-d}\hat{f} * \hat{g}$.*

DÉMONSTRATION : Si f et g appartiennent à $\mathscr{S}(\mathbb{R}^d)$, ceci est le résultat du Lemme 8.1.14. La fonction $(f,g) \mapsto fg$ est continue de $L^2 \times L^2$ dans L^1, et $\mathscr{F}$ est continue de L^1 dans $\mathscr{C}_0$. De même la fonction $f \mapsto \hat{f}$ est continue de L^2 dans L^2, dont il résulte que la fonction $(f,g) \mapsto (\hat{f}, \hat{g})$ est continue de $L^2 \times L^2$ dans $L^2 \times L^2$. Enfin, l'application $(u,v) \mapsto u * v$ est continue de $L^2 \times L^2$ dans $\mathscr{C}_0$. Par composition, on obtient que la fonction $\Phi : (f,g) \mapsto \mathscr{F}(f.g) - (2\pi)^{-d}\hat{f} * \hat{g}$ est continue de $L^2 \times L^2$ dans $\mathscr{C}_0$, et nulle sur le sous-espace dense $\mathscr{S} \times \mathscr{S}$. Il en résulte que Φ est identiquement nulle, et que l'égalité cherchée est vérifiée pour toute $f \in L^2$ et toute $g \in L^2$. ∎

8.4 Exemples

Théorème 8.4.1. *Dans $\mathscr{S}'(\mathbb{R}^d)$, on a $\mathscr{F}(\delta_0) = \mathbb{1}$ et $\mathscr{F}(\mathbb{1}) = (2\pi)^d.\delta_0$.*

DÉMONSTRATION : Soit $\varphi \in \mathscr{S}(\mathbb{R}^d)$. On a

$$\langle \mathscr{F}(\delta_0), \varphi \rangle = \langle \delta_0, \mathscr{F}\varphi \rangle = \hat{\varphi}(0) = \int \varphi(x)\, dx = \langle \mathbb{1}, \varphi \rangle \ ,$$

d'où $\mathscr{F}(\delta_0) = \mathbb{1}$.
Le même calcul montre que $\overline{\mathscr{F}}(\delta_0) = \mathbb{1}$. On a donc $\mathscr{F}(\mathbb{1}) = \mathscr{F}\overline{\mathscr{F}}(\delta_0) = (2\pi)^d\delta_0$. ■

Théorème 8.4.2. *Soit S une distribution tempérée sur $\mathbb{R}^d$. Alors il existe une distribution tempérée T telle que $(I - \Delta)T = S$.*

DÉMONSTRATION : Si T est solution de cette équation, on doit avoir $\mathscr{F}S = \mathscr{F}T - \mathscr{F}(\Delta T)$. Et, d'après le théorème 8.3.4, $\mathscr{F}(\Delta T) = -(\sum_{j=1}^d \xi_j^2)\mathscr{F}T$.
Il en résulte que $\mathscr{F}S = (1 + \|\xi\|^2)\mathscr{F}T$, c'est-à-dire $\mathscr{F}T = \dfrac{1}{1+\|\xi\|^2}.\mathscr{F}S$. Et puisque la fonction $\xi \mapsto (1 + \|\xi\|^2)^{-1}$ est clairement dans $\mathscr{O}_M(\mathbb{R}^d)$ car ses dérivées partielles de tout ordre sont dans $\mathscr{C}_0(\mathbb{R}^d)$, le produit $T_1 = (1 + \|\xi\|^2)^{-1}\mathscr{F}S$ est une distribution tempérée. Il suffit alors de prendre $T = \mathscr{F}^{-1}(T_1)$ pour obtenir une distribution tempérée T satisfaisant $\mathscr{F}T = T_1$, donc $(1 + \|\xi\|^2)\mathscr{F}T = \mathscr{F}S$, ou encore $\mathscr{F}((I - \Delta)T) = \mathscr{F}S$. Et puisque $\mathscr{F}$ est un isomorphisme de $\mathscr{S}'(\mathbb{R}^d)$, ceci entraîne l'égalité cherchée : $(I - \Delta)T = S$. ■

En particulier, lorsque $S = \delta_0$, ceci fournit une solution fondamentale de l'opérateur différentiel $I - \Delta$.

Théorème 8.4.3. *La distribution $S = \mathrm{vp}(\dfrac{1}{x})$ sur $\mathbb{R}$ est tempérée et $\mathscr{F}S$ est la fonction impaire sur $\mathbb{R}$ égale à $-i\pi$ sur $\mathbb{R}^+$.*

DÉMONSTRATION : Si $\varphi \in \mathscr{D}(\mathbb{R})$, $\lambda > 0$ et $\varphi_\lambda(x) = \varphi(\dfrac{x}{\lambda})$, on a

$$\langle S, \varphi_\lambda \rangle = \lim_{\varepsilon \to 0} \int_{|x|>\varepsilon} \varphi(\frac{x}{\lambda})\frac{dx}{x} = \lim_{\varepsilon \to 0} \int_{|u|>\varepsilon/\lambda} \varphi(u)\frac{du}{u} = \langle S, \varphi \rangle \ ,$$

d'où l'on tire que S est homogène de degré -1, et donc tempérée. On voit de plus que $\mathscr{F}S$ est homogène de degré 0 (cf. 8.3.10).
Puisque $x.S = \mathbb{1}$, on a $\mathscr{F}(x.S) = \mathscr{F}\mathbb{1} = 2\pi\delta_0$. Et puisque $(\mathscr{F}S)' = \mathscr{F}(-ix.S) = -2i\pi\delta_0$, on remarque que $(\mathscr{F}S + 2i\pi Y)' = 0$. Il résulte alors de la proposition 7.7.3 que $\mathscr{F}S + 2i\pi Y$ est une fonction constante θ. On remarque enfin que S est impaire, c'est-à-dire $\tilde{S} = -S$; et puisque $\widetilde{\mathscr{F}S} = \mathscr{F}(\tilde{S}) = -\mathscr{F}S$, on a donc $\theta = \widetilde{\mathscr{F}S} + 2i\pi\tilde{Y} = -\mathscr{F}S + 2i\pi(1 - Y) = 2i\pi - \theta$, dont on conclut que $\theta = i\pi$, et que $\mathscr{F}S$ vaut $i\pi$ sur $]-\infty, 0[$ et $-i\pi$ sur $]0, +\infty[$. ■

Corollaire 8.4.4. *L'application définie sur $\mathscr{S}(\mathbb{R})$ par $\varphi \mapsto \varphi * S$, où $S = \mathrm{vp}(\dfrac{1}{x})$, se prolonge en une application linéaire continue de $L^2(\mathbb{R})$ dans lui-même de norme π.*

DÉMONSTRATION : Pour $\psi \in \mathscr{S}(\mathbb{R})$, on a $\psi * S \in \mathscr{S}'(\mathbb{R})$, puisque S est tempérée, en vertu du théorème 8.3.5, ainsi que $\mathscr{F}(\psi * S) = \hat{\psi}.\mathscr{F}S$. Puisque $\hat{\psi} \in \mathscr{S}(\mathbb{R}) \subset L^2(\mathbb{R})$ et que $\mathscr{F}S$

est, en tant que distribution, une fonction bornée, le produit $\hat{\psi}.\mathscr{F}S$ est une fonction de L^2, et on a $\displaystyle\int \left|\hat{\psi}(\xi).\mathscr{F}S(\xi)\right|^2 d\xi = \pi^2 \int \left|\hat{\psi}(\xi)\right|^2 d\xi$. Il résulte alors de la formule de Plancherel que $\displaystyle\int \left|\hat{\psi}(\xi)\right|^2 d\xi = 2\pi \int |\psi(x)|^2 \, dx$.
Et puisque la transformation de Fourier est un isomorphisme de L^2 sur lui-même et que $\mathscr{F}(\psi * S) \in L^2$, on voit que $\psi * S$ appartient à L^2. Alors, la formule de Plancherel donne

$$2\pi \int |\psi * S(x)|^2 \, dx = \int |\mathscr{F}(\psi * S)(\xi)|^2 \, d\xi = \pi^2 \int \left|\hat{\psi}(\xi)\right|^2 d\xi = 2\pi^3 \int |\psi(x)|^2 \, dx \; .$$

On en tire que $\|\psi * S\|_2^2 = \pi^2 \|\psi\|_2^2$, donc $\|\psi * S\|_2 = \pi \|\psi\|_2$. Et, puisque $\mathscr{D}(\mathbb{R})$, et a fortiori $\mathscr{S}(\mathbb{R})$, est dense dans L^2, l'application linéaire $\psi \mapsto \psi * S$ se prolonge en une application linéaire continue H de norme π de L^2 dans lui-même.
On peut même remarquer que $\|Hf\|_2 = \pi \|f\|_2$ pour toute $f \in L^2(\mathbb{R})$, et que, puisque $\mathscr{F}(H^2\psi) = (\mathscr{F}S)^2\hat{\psi} = -\pi^2\hat{\psi}$, pour tout $\psi \in \mathscr{S}(\mathbb{R})$, on a nécessairement $H^2\psi = -\pi^2\psi$ pour $\psi \in \mathscr{S}$, donc aussi par densité : $H^2 f = -\pi^2 f$ pour toute $f \in L^2$. ■

8.5 Transformation de Fourier partielle

On identifie ici $\mathbb{R}^{d+\ell}$ à $\mathbb{R}^d \times \mathbb{R}^\ell$. Si f appartient à $\mathscr{S}(\mathbb{R}^{d+\ell})$, chacune des fonctions $f_x : y \mapsto f(x,y)$, pour $x \in \mathbb{R}^d$, est clairement dans $\mathscr{S}(\mathbb{R}^\ell)$. On va maintenant donner une caractérisation des fonctions f de $\mathscr{S}(\mathbb{R}^{d+\ell})$ utilisant la famille des fonctions partielles f_x.

Lemme 8.5.1. *Soit f une fonction sur $\mathbb{R}^{d+\ell}$. Alors f est dans $\mathscr{S}(\mathbb{R}^{d+\ell})$ si et seulement si chacune des fonctions partielles f_x est dans $\mathscr{S}(\mathbb{R}^\ell)$ et si, pour toute distribution tempérée S sur $\mathbb{R}^\ell$, la fonction $f^S : x \mapsto \langle S, f_x\rangle$ est dans $\mathscr{S}(\mathbb{R}^d)$.*

Démonstration : Comme dans la démonstration du lemme 7.1.6, on identifiera dans la suite $\mathbb{N}^{d+\ell}$ avec $\mathbb{N}^d \times \mathbb{N}^\ell$. Et pour f de classe $\mathscr{C}^\infty$ sur $\mathbb{R}^{d+\ell}$ on notera $\partial_{\alpha,\beta}$ la dérivée partielle de f obtenue en dérivant "α fois" par rapport à la variable dans $\mathbb{R}^d$ et "β fois" par rapport à la variable dans $\mathbb{R}^\ell$.

Supposons d'abord que f appartienne à $\mathscr{S}(\mathbb{R}^{d+\ell})$. On va montrer que $f^S \in \mathscr{S}(\mathbb{R}^d)$ pour chaque $S \in \mathscr{S}'(\mathbb{R}^\ell)$. La démonstration est similaire à celle du théorème 7.9.2.
Pour $\beta \in \mathbb{N}^\ell$, k et q dans $\mathbb{N}$, on a : $\sup_{x\in\mathbb{R}^d, y\in\mathbb{R}^\ell}(1+\|x\|+\|y\|)^{k+q}\,|\partial_{0,\beta}f(x,y)| < \infty$, donc

$$\sup_{x\in\mathbb{R}^d}(1+\|x\|)^q p_{\beta,k}(f_x) = \sup_{x\in\mathbb{R}^d}(1+\|x\|)^q \sup_{y\in\mathbb{R}^\ell}(1+\|y\|)^k\,|\partial_\beta f_x(y)| < \infty \; .$$

On en déduit que, pour toute semi-norme continue p sur $\mathscr{S}(\mathbb{R}^\ell)$ et tout entier q, on a $p(f_x) = O(\|x\|^{-q})$ et $\sup_x p(f_x) < \infty$. Il en résulte en particulier que l'ensemble $\{f_x : x \in \mathbb{R}^d\}$ est borné dans $\mathscr{S}(\mathbb{R}^\ell)$, donc relativement compact (cf. 8.1.7). Si la suite (x_n) converge vers x dans $\mathbb{R}^d$, on a $f(x_n,y) \to f(x,y)$ pour tout $y \in \mathbb{R}^\ell$, ce qui montre que la seule valeur d'adhérence possible de la suite (f_{x_n}) dans $\mathscr{S}(\mathbb{R}^\ell)$ est f_x : et la suite (f_{x_n}) converge vers f_x, ce qui montre la continuité de l'application $x \mapsto f_x$ de $\mathbb{R}^d$ dans $\mathscr{S}(\mathbb{R}^\ell)$. Alors, si S est

dans $\mathscr{S}'(\mathbb{R}^\ell)$, l'application f^S est continue. Et puisque $p : g \mapsto |\langle S, g\rangle|$ est une semi-norme continue sur $\mathscr{S}(\mathbb{R}^\ell)$ on a $f^S(x) = O(\|x\|^{-q})$ pour tout entier q.
Pour tout $\alpha \in \mathbb{N}^d$, la fonction $f_\alpha = \partial_{\alpha,0} f$ est dans $\mathscr{S}(\mathbb{R}^{d+\ell})$. Il résulte donc de ce qui précède qu'on a aussi la compacité relative de l'ensemble des $(f_\alpha)_x$, la continuité de $x \mapsto (f_\alpha)_x$ et le fait que $(f_\alpha)^S(x) = O(\|x\|)^{-q})$ pour tout $S \in \mathscr{S}'(\mathbb{R}^\ell)$ et tout entier q.
Pour x et u dans $\mathbb{R}^d$, la fonction

$$g_{x,u} : y \mapsto \frac{1}{\|x-u\|}\Big(f(x,y) - f(u,y) - \sum_j \partial_j f(u,y).(x_j - u_j)\Big)$$

vérifie

$$p_{\beta,k}(g_{x,u}) \leqslant \sup_x (1+\|y\|)^k \sum_j \big|\partial_{\tau_j,\beta} f(x,y)\big| \leqslant \sum_j p_{\tau_j,\beta,k}(f)\ .$$

Il en résulte à nouveau que l'ensemble des $g_{x,u}$ est relativement compact dans $\mathscr{S}(\mathbb{R}^\ell)$, et que $(x,u) \mapsto g_{x,u}$ est continue de $\mathbb{R}^d \times \mathbb{R}^d$ dans $\mathscr{S}(\mathbb{R}^\ell)$. On en déduit que, pour $S \in \mathscr{S}'(\mathbb{R}^\ell)$, $(x,u) \mapsto g_{x,u}^S$ est continue, ce qui montre, d'après le lemme 7.1.5, que f^S est de classe $\mathscr{C}^1$ et que $\partial_j(f^S) = (\partial_j f)^S$. Et puisque les dérivées partielles de f^S sont définies de la même manière que f^S, on montre par récurrence sur $|\alpha|$ que f^S est $\mathscr{C}^\infty$, avec $\partial_\alpha(f^S) = (\partial_{\alpha,0} f)^S$. On a donc $\partial_\alpha(f^S)(x) = O(\|x\|^{-q})$ pour tout $\alpha \in \mathbb{N}^d$ et tout q, c'est-à-dire que $f^S \in \mathscr{S}(\mathbb{R}^d)$.

Supposons maintenant que les f_x soient dans $\mathscr{S}(\mathbb{R}^\ell)$ pour tout $x \in \mathbb{R}^d$ et les f^S dans $\mathscr{S}(\mathbb{R}^d)$ pour tout $S \in \mathscr{S}'(\mathbb{R}^\ell)$. Il résulte du théorème 8.1.8 que, pour tout k, l'ensemble des $(1+\|x\|)^k f_x$ est relativement compact dans $\mathscr{S}(\mathbb{R}^\ell)$. Pour chaque $y \in \mathbb{R}^\ell$ et chaque $\beta \in \mathbb{N}^d$, la forme linéaire $S : g \mapsto \partial_\beta g(y)$ est continue sur $\mathscr{S}(\mathbb{R}^\ell)$: l'hypothèse que f^S appartient à $\mathscr{S}(\mathbb{R}^d)$ entraîne en particulier que, si une suite (x_n) converge dans $\mathbb{R}^d$ vers x, les valeurs d'adhérence g de $(\partial_\beta f_{x_n})$ vérifient $g(y) = \partial_\beta f(x,y)$ pour tout y, donc $g = \partial_\beta f_x$. Et l'unicité de cette valeur d'adhérence montre la continuité de $x \mapsto \partial_\beta f_x$ de $\mathbb{R}^d$ dans $\mathscr{S}(\mathbb{R}^\ell)$. Et puisque

$$\begin{aligned}|\partial_{0,\beta} f(x',y') - \partial_{0,\beta} f(x,y)| &\leqslant |\partial_{0,\beta} f(x',y') - \partial_{0,\beta} f(x,y')| + |\partial_{0,\beta} f(x,y') - \partial_{0,\beta} f(x,y)| \\ &\leqslant p(f_{x'} - f_x) + |\partial_{0,\beta} f(x,y') - \partial_{0,\beta} f(x,y)|\ ,\end{aligned}$$

où $p = p_{\beta,0}$ est la semi-norme continue : $g \mapsto \sup_y |\partial_\beta g(y)|$, on voit que $\partial_{0,\beta} f(x',y')$ tend vers $\partial_{0,\beta} f(x,y)$ lorsque (x',y') tend vers (x,y) donc que $\partial_{0,\beta} f$ est continue sur $\mathbb{R}^{d+\ell}$.
On voit de même que, pour tout $\alpha \in \mathbb{N}^d$ et tout $k \in \mathbb{N}$, l'ensemble des $(1+\|x\|)^k(\partial_\alpha f)_x$ est relativement compact dans $\mathscr{S}(\mathbb{R}^\ell)$, et on en déduit que la fonction $\partial_{\alpha,\beta} f$ est continue sur $\mathbb{R}^{d+\ell}$. Ceci montre que f est $\mathscr{C}^\infty$ sur $\mathbb{R}^{d+\ell}$. Enfin, puisque l'ensemble des $(1+\|x\|)^k(\partial_\alpha f)_x$ est borné, on voit que pour tout $\beta \in \mathbb{N}^\ell$ et tout entier q, on a

$$\sup_{x\in\mathbb{R}^d} p_{\beta,q}((1+\|x\|)^k \partial_\alpha f_x) = \sup_{(x,y)\in\mathbb{R}^{d+\ell}} (1+\|x\|)^k (1+\|y\|)^q\, |\partial_{\alpha,\beta} f(x,y)| < \infty\ ,$$

et puisque $(1+\|(x,y)\|)^k \leqslant (1+\|x\|+\|y\|)^k \leqslant (1+\|x\|)^k(1+\|y\|)^k$, on conclut que $f \in \mathscr{S}(\mathbb{R}^{d+\ell})$. ∎

Théorème 8.5.2. *Si S est une distribution tempérée sur $\mathbb{R}^d$ et T une distribution tempérée sur $\mathbb{R}^\ell$, alors $S \otimes T$ est une distribution tempérée sur $\mathbb{R}^{d+\ell}$.*

Démonstration : Si φ appartient à $\mathscr{D}(\mathbb{R}^{d+\ell})$, on a, par définition de la distribution $S \otimes T$, $\langle S\otimes T, f\rangle = \langle S, \varphi^T\rangle$, où $\varphi^T(x) = \langle T, \varphi_x\rangle$. Il suffit de montrer que la formule similaire, lorsque φ est dans $\mathscr{S}(\mathbb{R}^{d+\ell})$, définit une forme linéaire continue sur $\mathscr{S}(\mathbb{R}^{d+\ell})$.

En utlisant à deux reprises le théorème du graphe fermé, on montre ainsi, — d'une part, que pour x fixé dans $\mathbb{R}^d$, l'application $f \mapsto f_x$ est continue de $\mathscr{S}(\mathbb{R}^{d+\ell})$ dans $\mathscr{S}(\mathbb{R}^\ell)$: en effet, l'égalité $g = f_x$ équivaut à " $\forall y \quad g(y) = f(x,y)$ ", — d'autre part que, pour T fixé dans $\mathscr{S}'(\mathbb{R}^\ell)$, la fonction $f \mapsto f^T$ est continue : en effet l'égalité $g = f^T$ équivaut à " $\forall x \quad g(x) = \langle T, f_x\rangle$ ". Alors l'application $S \otimes T : f \mapsto \langle S, f^T\rangle$ est une forme linéaire continue sur $\mathscr{S}(\mathbb{R}^{d+\ell})$, c'est-à-dire une distribution tempérée. ■

On s'intéresse maintenant à l'opération suivante : étant donné une fonction f sur $\mathbb{R}^d \times \mathbb{R}^\ell$, on définit la fonction g sur $\mathbb{R}^d \times \mathbb{R}^\ell$ par $g(x,\xi) = \mathscr{F}(f_x)(\xi)$, c'est-à-dire

$$g(x,\xi) = \int f(x,y)\,\mathrm{e}^{-iy\cdot\xi}\,dy\ ,$$

ce qui correspond à "prendre la transformée de Fourier partielle" de f selon la variable y. On notera $g = \mathscr{F}_{\mathrm{part}}(f)$ cette fonction.

Théorème 8.5.3. *L'application $\mathscr{F}_{\mathrm{part}}$ est un isomorphisme de $\mathscr{S}(\mathbb{R}^{d+\ell})$ sur $\mathscr{S}(\mathbb{R}^{d+\ell})$.*

Démonstration : Soient $f \in \mathscr{S}(\mathbb{R}^{d+\ell})$ et $g = \mathscr{F}_{\mathrm{part}}(f)$. Pour tout $x \in \mathbb{R}^d$, on notera f_x la fonction $y \mapsto f(x,y)$ et g_x la fonction $\xi \mapsto g(x,\xi)$. Pour montrer que g est dans $\mathscr{S}$, il suffit de montrer que g_x est dans $\mathscr{S}(\mathbb{R}^\ell)$ pour tout x et que, pour toute distribution tempérée S, la fonction $g^S = x \mapsto \langle S, g_x\rangle$ est dans $\mathscr{S}(\mathbb{R}^d)$. Mais puisque $f_x \in \mathscr{S}$, on a $g_x = \mathscr{F}(f_x) \in \mathscr{S}$ et, puisque $\langle S, g_x\rangle = \langle S, \mathscr{F}(f_x)\rangle = \langle \mathscr{F}(S), f_x\rangle$, on a $g^S = f^T$, où T est la distribution tempérée $\mathscr{F}(S)$. Et puisque $f^T \in \mathscr{S}$, ceci achève de prouver que $g \in \mathscr{S}(\mathbb{R}^{d+\ell})$.
Pour prouver la continuité de l'application linéaire $\mathscr{F}_{\mathrm{part}}$, on peut utiliser le théorème du graphe fermé : en effet, dire que $g = \mathscr{F}_{\mathrm{part}}(f)$ est équivalent à l'assertion :

$$\text{“ } \forall x \in \mathbb{R}^d\ \forall \xi \in \mathbb{R}^\ell \quad g(x,\xi) = \int f(x,y)\,\mathrm{e}^{-iy\cdot\xi}\,dy \text{ ”}$$

et on vérifie immédiatement que, pour x et ξ fixés, les applications $f \mapsto \int f(x,y)\,\mathrm{e}^{-iy\cdot\xi}\,dy$ et $g \mapsto g(x,\xi)$ sont continues de $\mathscr{S}(\mathbb{R}^{d+\ell})$ dans $\mathbb{C}$.
Enfin, il résulte clairement de la formule d'inversion : $\mathscr{F}^{-1}g_x = (2\pi)^{-d}.\overline{\mathscr{F}}(f_x)$, que $\mathscr{F}_{\mathrm{part}}$ est un isomorphisme. ■

Théorème 8.5.4. *Si S est une distribution tempérée sur $\mathbb{R}^{d+\ell}$, l'application linéaire définie sur $\mathscr{S}(\mathbb{R}^{d+\ell})$ par $\varphi \mapsto \langle S, \mathscr{F}_{\mathrm{part}}(\varphi)\rangle$ est une distribution tempérée qu'on notera $\mathscr{F}_{\mathrm{part}}(S)$. De plus, l'application $\mathscr{F}_{\mathrm{part}}$ est un isomorphisme de $\mathscr{S}'(\mathbb{R}^{d+\ell})$ sur lui-même.*

Démonstration : L'application $\mathscr{F}_{\mathrm{part}}$ sur $\mathscr{S}'$ est clairement la transposée de l'application $\mathscr{F}_{\mathrm{part}}$ sur $\mathscr{S}$. Le résultat annoncé découle donc immédiatement du théorème 8.5.3. ■

Une solution fondamentale de l'opérateur de la chaleur.

On va utiliser le résultat précédent pour trouver une solution fondamentale de l'opérateur différentiel à coefficients constants $H = \partial_t - \Delta_x = \dfrac{\partial}{\partial t} - \displaystyle\sum_{k=1}^{d} \dfrac{\partial^2}{\partial x_j^2}$ sur $\mathbb{R} \times \mathbb{R}^d$.

Si E est une distribution tempérée solution fondamentale de H, et si $\varphi \in \mathscr{S}$, on doit avoir $HE = \delta_0(t,x) = \partial_0(t) \otimes \delta_0(x)$, c'est-à-dire $\langle \partial_t E - \Delta_x E, \varphi\rangle = \langle E, -\partial_t\varphi - \Delta_x\varphi\rangle = \varphi(0)$. Et si F_{part} désigne ici la transformation de Fourier partielle relative à $x = (x_1, x_2, \ldots, x_d)$, et

si on prend $\varphi = \mathscr{F}_{\text{part}}(\psi)$, pour ψ quelconque dans $\mathscr{S}$, on aura $\partial_t(\mathscr{F}_{\text{part}}(\psi)) = \mathscr{F}_{\text{part}}(\partial_t\psi)$, et $\Delta_x(\mathscr{F}_{\text{part}}(\psi)) = \mathscr{F}_{\text{part}}(-\|\xi\|^2\psi)$.
On doit donc avoir pour tout $\psi \in \mathscr{S}$: $\langle E, \mathscr{F}_{\text{part}}(-\partial_t\psi) + \mathscr{F}_{\text{part}}(\|\xi\|^2\psi)\rangle = \mathscr{F}_{\text{part}}(\psi)(0)$, ou encore

$$\langle \mathscr{F}_{\text{part}}(E), \|\xi\|^2\psi - \partial_t\psi\rangle = \langle \|\xi\|^2\mathscr{F}_{\text{part}}(E) + \partial_t\mathscr{F}_{\text{part}}(E), \psi\rangle = \langle \delta_0(t)\otimes \mathbb{1}(\xi), \psi\rangle \ ,$$

c'est-à-dire, en notant $\widehat{E} = \mathscr{F}_{\text{part}}(E)$: $\partial_t\widehat{E} + \|\xi\|^2\widehat{E} = \delta_0(t)\otimes\mathbb{1}(\xi)$. On voit alors que $\widehat{E}(t,\xi) = Y(t)\,\mathrm{e}^{-t\|\xi\|^2}$ est une distribution tempérée qui satisfait cette équation. Il suffit donc de prendre pour E la transformée de Fourier partielle inverse de $\widehat{E}$. Et, compte tenu des lemmes 6.6.8 et 6.6.9, ceci donne

$$E(x,t) = (2\pi t)^{-d/2}.Y(t).\,\mathrm{e}^{-\|x\|^2/4t} \ .$$

On vérifie alors sans peine que cette distribution est dans $L^1_{\text{loc}}(\mathbb{R}^{d+1})$ et que le support singulier de E est le singleton $\{0\}$. Et, à nouveau, le théorème 7.12.4 peut s'appliquer.

9

ESPACES DE SOBOLEV

9.1 Définition

Soit Ω un ouvert de $\mathbb{R}^d$. Pour $s \in \mathbb{N}$ et $1 \leqslant p \leqslant \infty$, on notera $W^{s,p}(\Omega)$ l'espace vectoriel des (classes de) fonctions f sur Ω qui appartiennent à $L^p(\Omega)$, ainsi que toutes leurs dérivées partielles $\partial_\alpha f$ au sens des distributions, pour $|\alpha| \leqslant s$. On définit alors une norme sur $W^{s,p}$ en posant

$$\|f\|_{W^{s,p}} = \Big(\sum_{|\alpha|\leqslant s} \|\partial_\alpha f\|_p^p\Big)^{1/p}$$

Théorème 9.1.1. *L'espace $W^{s,p}(\Omega)$ est un espace de Banach.*

Démonstration : Si on note $E_s = \{\alpha \in \mathbb{N}^d : |\alpha| \leqslant s\}$, l'espace $W^{s,p}$ s'identifie à un sous-espace de l'espace de Banach $L^p(\Omega)^{E_s}$ par l'injection linéaire $j : f \mapsto (\partial_\alpha f)_{\alpha\in E_s}$. Il suffit donc de voir que $j(W^{s,p})$ est fermé, ce qui résulte aisément de l'équivalence :

$$(g_\alpha)_{\alpha\in E} \in j(W^{s,p}) \iff \forall\alpha \in E_s\ \forall\varphi \in \mathscr{D}(\Omega) \quad \int_\Omega (g_0\partial_\alpha\varphi - (-1)^{|\alpha|}g_\alpha\varphi)\,dx = 0\ ,$$

puisque φ et $\partial_\alpha\varphi$ appartiennent à $L^{p'}(\Omega)$. ∎

Théorème 9.1.2. *Lorsque $p = 2$, l'espace $W^{s,2}(\Omega)$ est un espace de Hilbert, souvent noté $H^s(\Omega)$.*

Démonstration : Il suffit de remarquer que la norme de $W^{s,2}(\Omega)$ est associée au produit scalaire défini par

$$\langle f, g\rangle = \sum_{|\alpha|\leqslant s} \int_\Omega \partial_\alpha f(x).\partial_\alpha \bar{g}(x)\,dx\ ,$$

puisque la complétude résulte du théorème précédent. ∎

Théorème 9.1.3. *Soient Ω et Ω' deux ouverts de $\mathbb{R}^d$, Φ un $\mathscr{C}^s$-difféomorphisme de Ω sur Ω', ω un sous-ouvert relativement compact de Ω et $\omega' = \Phi(\omega)$. L'application $\Phi^* : f \mapsto f\circ\Phi^{-1}$ envoie continuement $W^{s,p}(\omega)$ dans $W^{s,p}(\omega')$.*

Démonstration : On montre par récurrence sur $|\alpha| \leqslant s$ l'existence de fonctions $c_{\alpha,\beta}$ de classe $\mathscr{C}^{s-|\alpha|}$ sur Ω' telles que $\partial_\alpha(f\circ\Phi^{-1})(x) = \sum_{\beta\leqslant\alpha} c_{\alpha,\beta}(x)\partial_\beta f\circ\Phi^{-1}(x)$. Puisque

$\omega' \subset \Phi(\overline{\omega})$ est relativement compact dans Ω', les fonctions $c_{\alpha,\beta}$ sont toutes uniformément bornées sur ω', et le jacobien J_Φ de Φ est borné sur ω. On a

$$\int_{\omega'} \left|\partial_\beta f\circ\Phi^{-1}(x)\right|^p \, dx = \int_\omega |\partial_\beta f(y)|^p \, |J_\Phi(y)| \, dy < +\infty \ ,$$

dont on déduit que $\partial_\beta f\circ\Phi^{-1} \in L^p(\omega')$, et puisque chaque $c_{\alpha,\beta}$ est bornée sur ω', on conclut que $\partial_\alpha(f\circ\Phi^{-1}) \in L^p(\omega')$ pour $|\alpha| \leqslant s$, donc que $f\circ\Phi^{-1} \in W^{s,p}(\omega')$. Les majorations de J_Φ et des $c_{\alpha,\beta}$ ne dépendant que de Φ et de ω, on voit alors qu'il existe une constante M telle que $\left\|f\circ\Phi^{-1}\right\|_{W^{s,p}} \leqslant M \left\|f\right\|_{W^{s,p}}$, ce qui montre la continuité de Φ^*. ■

Théorème 9.1.4. *Si $\varphi \in \mathscr{D}(\Omega)$, l'application $M_\varphi : f \mapsto \varphi.f$ envoie continuement $W^{s,p}(\Omega)$ dans lui-même.*

Démonstration : Par la formule de Leibniz, il existe pour tout α des coefficients $(c_{\alpha,\beta})_{\beta\leqslant\alpha}$ tels que $\partial_\alpha(\varphi.f)(x) = \sum_{\beta\leqslant\alpha} c_{\alpha,\beta}\partial_\beta\varphi(x).\partial_{\alpha-\beta}f(x)$. Par continuité sur le support compact de φ, les $\partial_\beta\varphi$ sont uniformément bornées sur Ω. On en déduit aisément que $\partial_\alpha(\varphi.f)$ est dans $L^p(\Omega)$ pour tout $\alpha \in E_s$, donc que $\varphi.f \in W^{s,p}(\Omega)$ et qu'on peut trouver une constante M telle que $\|\varphi.f)\|_{W^{s,p}} \leqslant M. \|f\|_{W^{s,p}}$, ce qui prouve la continuité de l'application linéaire M_φ. ■

9.2 Injections de Sobolev

On va maintenant montrer que, sous des hypothèses convenables sur l'ouvert Ω, les espaces $W^{s,p}$ s'injectent les uns dans les autres. On notera désormais H_d le demi-espace $\{y = (y_1, y_2, \ldots, y_d) \in \mathbb{R}^d : y_1 < 0\}$ de $\mathbb{R}^d$.

Lemme 9.2.1. *Soit s un entier. On suppose que l'ouvert Ω est borné dans $\mathbb{R}^d$ et que, pour tout point $a \in \partial\Omega$, il existe un $\mathscr{C}^s$-difféomorphisme Φ_a d'un voisinage V_a de a sur un voisinage U_a de 0 dans $\mathbb{R}^d$ tel que $\Phi_a(\Omega \cap V_a) = U_a \cap H_d$. Alors il existe un recouvrement ouvert fini $(V_j)_{j\in J}$ de $\overline{\Omega}$, des fonctions $\psi_j \in \mathscr{D}(V_j)$ telles que $\sum_{j\in J}\psi_j = 1$ au voisinage de $\overline{\Omega}$, et des $\mathscr{C}^s$-difféomorphismes Φ_j de V_j sur des ouverts bornés U_j de $\mathbb{R}^d$ tels que $\Phi_j(\Omega\cap V_j) = U_j\cap H_d$.*

Démonstration : Par hypothèse, tout point de la frontière de Ω possède un voisinage ouvert $\mathscr{C}^s$-difféomorphe à un voisinage de 0 dans $\mathbb{R}^d$ par une fonction transformant la trace de Ω en celle de H_d. Et tout point de Ω possède un voisinage ouvert $\mathscr{C}^s$-difféomorphe à un sous-ouvert de H_d (par exemple par une translation). Par compacité de $\overline{\Omega}$, on peut donc trouver un sous-recouvrement fini $(V_j)_{j\in J}$ extrait de ce recouvrement, puis une partition $\mathscr{C}^\infty$ de l'unité, $(\psi_j)_{j\in J}$, subordonnée à ce recouvrement, c'est-à-dire des fonctions positives $\psi_j \in \mathscr{D}(V_j)$ telles que $\sum_j \psi_j = 1$ au voisinage du compact $\overline{\Omega}$. ■

Alors si $f \in W^{s,p}(\Omega)$, chaque fonction $f_j := \Phi_j^*(\psi_j.f)$ est dans $W^{s,p}(H_d)$ avec un support compact contenu dans $\Phi_j(V_j)$. Inversement, si (f_j) est une famille de fonctions de $W^{s,p}(H_d)$ telles que le support de f_j soit un compact contenu dans $\Phi_j(V_j)$, la somme $\sum_j(\Phi_j^*)^{-1}(f_j)$ est dans $W^{s,p}(\Omega)$. Ceci permet de transposer à $W^{s,p}(\Omega)$ des résultats prouvés pour $W^{s,p}(H_d)$.

Si K est un compact de $\mathbb{R}^d$, on notera encore $W^{s,p}_K(\Omega)$ le sous-espace (fermé) de $W^{s,p}(\Omega)$ formé des fonctions nulles hors de K.

9.3 Le cas s=1 pour le demi-espace

Dans tout ce qui suit, on se place sur l'espace euclidien $\mathbb{R}^d$. On note B la boule unité de $\mathbb{R}^d$ et $B^+ = B \cap (-H_d)$, $v_d = \dfrac{\pi^{d/2}}{\Gamma(1+\frac{d}{2})}$ le volume d-dimensionnel de la boule unité et $\sigma_d = dv_d$ le volume $(d-1)$-dimensionnel de la sphère unité.
On considère, pour $1 \leqslant j \leqslant d$, les distributions T_j associées aux fonctions

$$\chi_j : x \mapsto \mathbb{1}_{B^+}(x).x_j \frac{1-\|x\|^d}{d\,\|x\|^d} \ .$$

Lemme 9.3.1. *Chaque T_j est dans L^p pour $1 \leqslant p < \dfrac{d}{d-1}$, et on a*

$$\sum_{j=1}^{d} \partial_j T_j = \frac{v_d}{2}\delta_0 - \mathbb{1}_{B^+} \ .$$

Démonstration : On a pour tout j : $|x_j| \leqslant \|x\|$, donc $|\chi_j(x)| \leqslant \|x\|^{1-d}.\mathbb{1}_{B^+}(x)$, et

$$\int |\chi_j(x)|^p \, dx \leqslant \frac{\sigma_d}{2}\int_0^1 r^{p(1-d)} r^{d-1}\, dr = \frac{\sigma_d}{2q} \ ,$$

où $q = 1-(p-1)(d-1) = p+d-pd$, pourvu que $p(d-1) < d$,
Si $\varphi \in \mathscr{D}(\mathbb{R}^d)$, on a pour $y = (y_1, y_2, \ldots, y_d) \in \mathbb{R}^d$: $\varphi(y) - \varphi(0) = \displaystyle\int_0^1 \sum_j y_j \partial_j \varphi(ty)\, dt$, donc, en faisant $x = ty$, puis $s = \dfrac{\|x\|}{t}$,

$$\begin{aligned}
\int_{B^+} (\varphi(y)-\varphi(0))\, dy &= \sum_j \int_0^1 dt \int y_j \partial_j \varphi(ty) \mathbb{1}_{B^+}(y)\, dy \\
&= \sum_j \int \partial_j \varphi(x)\, dx \int_0^1 \mathbb{1}_{B^+}(\frac{x}{t}) \frac{x_j}{t^{d+1}}\, dt \\
&= \sum_j \int_{B^+} x_j \partial_j \varphi(x)\, dx \int_{\|x\|}^1 \|x\|^{-d} s^{d-1}\, ds \\
&= \sum_j \int_{B^+} \frac{x_j.(1-\|x\|^d)}{d\,\|x\|^d} \partial_j \varphi(x)\, dx = \sum_j \langle T_j, \partial_j \varphi\rangle = -\sum_j \langle \partial_j T_j, \varphi\rangle \ ,
\end{aligned}$$

c'est-à-dire $\langle \mathbb{1}_{B^+} - c.\delta_0, \varphi\rangle = -\langle \sum_j \partial_j T_j, \varphi\rangle$, en posant $c = \mu(B^+) = \dfrac{v_d}{2}$, d'où le résultat. ■

Théorème 9.3.2. *Soit $p > d$. Si g est dans $W^{1,p}(H_d)$, alors g est sur H_d presque partout égale à la restriction d'une fonction de $\mathscr{C}_0(\overline{H_d})$ et l'injection de $W^{1,p}(H_d)$ dans $\mathscr{C}_0(\overline{H_d})$ est continue.*
De plus, si K est un compact de $\mathbb{R}^d$, l'injection de $W_K^{1,p}(H_d)$ dans $\mathscr{C}_0(\overline{H_d})$ est compacte.

Démonstration : Si $p > d$, on a $p' < \dfrac{d}{d-1}$. Il en résulte que les χ_j et $\mathbb{1}_{B^+}$ appartiennent à $L^{p'}(\mathbb{R}^d)$, et que $g * \mathbb{1}_{B^+}$ appartient à $\mathscr{C}_0(\mathbb{R}^d)$, ainsi que les $\chi_j * \partial_j g$ pour $1 \leqslant j \leqslant d$. Pour $x \in H_d$, on a $x - B^+ \subset H_d$. Il en résulte que si $\varphi \in \mathscr{D}(H_d)$, on a

$$\langle g * \partial_j \chi_j, \varphi \rangle = \langle g, \varphi * (\partial_j \chi_j)\check{} \rangle = -\langle g, \varphi * \partial_j \tilde{\chi}_j \rangle = -\langle g, \partial_j(\varphi * \tilde{\chi}_j) \rangle ,$$

en notant, comme dans le chapitre précédent, $\tilde{f}(x) = f(-x)$. Alors $\partial_j(\varphi * \tilde{\chi}_j) \in \mathscr{D}(H_d)$, et la quantité précédente est égale à $\langle \partial_j g, \varphi * \tilde{\chi}_j \rangle = \langle \partial_j g * \chi_j, \varphi \rangle$, et on a donc dans $\mathscr{D}'(H_d) : \partial_j g * \chi_j = g * \partial_j \chi_j$. Et puisque les distributions $\mathbb{1}_{B^+}$ et χ_j sont à support compact on a dans $\mathscr{D}'(H_d)$:

$$g = \delta_0 * g = \frac{1}{c}(\mathbb{1}_{B^+} * g + \sum_j \partial_j \chi_j * g) = \frac{1}{c}(\mathbb{1}_{B^+} * g + \sum_j \chi_j * \partial_j g) ,$$

et cette dernière distribution est dans $\mathscr{C}_0(\mathbb{R}^d)$. Sa restriction à $\overline{H_d}$ est dans $\mathscr{C}_0(\overline{H_d})$ et coïncide sur H_d avec g en tant que distributions, donc presque partout en tant que fonctions. De plus on a

$$\|g\|_\infty \leqslant \frac{1}{c}\Big(\|\mathbb{1}_{B^+}\|_{p'} \cdot \|g\|_p + \sum_j \|\chi_j\|_{p'} \cdot \|\partial_j g\|_p\Big) \leqslant \frac{1}{c}\Big(\|\mathbb{1}_{B^+}\|_{p'} + \sum_j \|\chi_j\|_{p'}\Big) \cdot \|g\|_{W^{1,p}} ,$$

ce qui prouve la continuité de l'injection de $W^{1,p}(H_d)$ dans $\mathscr{C}_0(\overline{H_d})$.
Et si K est un compact de H_d, il résulte du théorème de compacité des opérateurs de convolution (théorème 6.4.3) que l'injection de $W_K^{1,p}(H_d)$ dans $\mathscr{C}_0(\overline{H_d})$ est compacte. ■

Théorème 9.3.3. *Soient $p < d$ et r tel que $\dfrac{1}{r} > \dfrac{1}{p} - \dfrac{1}{d}$. Si g est dans $W^{1,p}(H_d)$, alors g est dans $L^r(H_d)$ et l'injection de $W^{1,p}(H_d)$ dans $L^r(H_d)$ est continue. De plus, si K est un compact de $\mathbb{R}^d$, l'injection de $W_K^{1,p}(H_d)$ dans $L^r(H_d)$ est compacte.*

Démonstration : Soit q tel que $\dfrac{1}{p} + \dfrac{1}{q} = 1 + \dfrac{1}{r}$. On a $\dfrac{1}{q} = 1 - \dfrac{1}{p} + \dfrac{1}{r} > 1 - \dfrac{1}{d}$. Il en résulte que $\mathbb{1}_{B^+}$ et les χ_j sont dans $L^q(\mathbb{R}^d)$, donc que $f * \mathbb{1}_{B^+}$ et $\partial_j g * \chi_j$ appartiennent à $L^r(\mathbb{R}^d)$. Et comme ci-dessus, on a dans $\mathscr{D}'(H_d)$:

$$g = \delta_0 * g = \frac{1}{c}(\mathbb{1}_{B^+} * g + \sum_j \partial_j \chi_j * g) = \frac{1}{c}(\mathbb{1}_{B^+} * g + \sum_j \chi_j * \partial_j g) ,$$

et cette dernière distribution est dans $L^r(\mathbb{R}^d)$. Sa restriction à H_d est dans $L^r(H_d)$ et coïncide sur H_d avec g en tant que distributions, donc presque partout en tant que fonctions.
Et on a

$$\|g\|_r \leqslant \frac{1}{c}\Big(\|\mathbb{1}_{B^+}\|_q \cdot \|g\|_p + \sum_j \|\chi_j\|_q \, \|\partial_j g\|_p\Big) \leqslant \frac{1}{c}\Big(\|\mathbb{1}_{B^+}\|_q + \sum_j \|\chi_j\|_q\Big) \|g\|_{W^{1,p}} ,$$

ce qui montre la continuité de l'injection de $W^{1,p}(H_d)$ dans $L^r(H_d)$.
Enfin, si K est un compact de H_d, il résulte du théorème de compacité des opérateurs de convolution (théorème 6.4.4) que l'injection de $W_K^{1,p}(H_d)$ dans $L^r(H_d)$ est compacte. ■

9.4 Le cas général

Théorème 9.4.1. *Soient $s \geqslant 1$ et $p < d$. Si g est dans $W^{s,p}(H_d)$, alors g est dans $W^{s-1,q}(H_d)$ pour q tel que $\dfrac{1}{q} > \dfrac{1}{p} - \dfrac{1}{d}$ et l'injection de $W^{s,p}(H_d)$ dans $W^{s-1,q}(H_d)$ est continue. De plus, si K est un compact de $\mathbb{R}^d$, l'injection de $W_K^{s,p}(H_d)$ dans $W^{s-1,q}(H_d)$ est compacte.*

Démonstration : Pour $\alpha \in \mathbb{N}^d$ tel que $|\alpha| < s$, on a $\partial_\alpha f \in W^{1,p}(H_d)$, avec la majoration $\|\partial_\alpha f\|_{W^{1,p}} \leqslant \|\chi\|_{W^{s,p}}$. Il résulte alors du théorème 9.3.3 que chaque $\partial_\alpha f$ appartient à $L^q(H_d)$, puisque $\dfrac{1}{q} > \dfrac{1}{p} - \dfrac{1}{d}$ et que $\|\partial_\alpha f\|_q \leqslant C(q,d).\, \|\partial_\alpha f\|_{W^{1,p}} \leqslant C(q,d)\, \|\chi\|_{W^{s,p}}$. Ceci montre que f appartient à $W^{s-1,q}(H_d)$ et que l'injection de $W^{s,p}(H_d)$ dans $W^{s-1,q}(H_d)$ est continue.

Si de plus, K est un compact de $\mathbb{R}^d$, le théorème de compacité des opérateurs de convolution (théorème 6.4.4) montre que l'injection de $W_K^{s,p}(H_d)$ dans $W^{s-1,q}(H_d)$ est compacte. ■

Théorème 9.4.2. *Soient $s \geqslant 1$ et $p > d$. Si g est dans $W^{s,p}(H_d)$, alors g est la restriction à $\overline{H_d}$ d'un élément de $\mathscr{C}_0^{s-1}(\mathbb{R}^d)$, et l'injection de $W^{s,p}(H_d)$ dans $\mathscr{C}_0^{s-1}(\overline{H_d})$ est continue. De plus, si K est un compact de $\mathbb{R}^d$, l'injection de $W_K^{s,p}(H_d)$ dans $\mathscr{C}_0^{s-1}(\overline{H_d})$ est compacte.*

Démonstration : A nouveau, pour tout $\alpha \in \mathbb{N}^d$ tel que $|\alpha| < s$, il résulte du théorème 9.3.2 que $\partial_\alpha g$ est la restriction à H_d d'une fonction de $\mathscr{C}_0(\mathbb{R}^d)$. On en déduit que g est $\mathscr{C}^{s-1}$ sur H_d et que toutes ses dérivées partielles d'ordre $\leqslant s-1$ sont continues sur $\overline{H_d}$. Et on voit comme précédemment que l'injection de $W^{s,p}$ dans $\mathscr{C}_0^{s-1}$ est continue, et même compacte sur $W_K^{s,p}(H_d)$. ■

Théorème 9.4.3. *Soient $s \geqslant 1$, $p > \dfrac{d}{s}$ et $k \in \mathbb{N}$ tel que $k < s - \dfrac{d}{p}$. Si g est dans $W^{s,p}(H_d)$, alors g est la restriction à $\overline{H_d}$ d'un élément de $\mathscr{C}_0^k(\mathbb{R}^d)$, et l'injection de $W^{s,p}(H_d)$ dans $\mathscr{C}_0^k(\overline{H_d})$ est continue. De plus, si K est un compact de $\mathbb{R}^d$, l'injection de $W_K^{s,p}(H_d)$ dans $\mathscr{C}_0^k(\overline{H_d})$ est compacte.*

Démonstration : Soit $m < s$ le plus grand entier tel que $\dfrac{1}{p} - \dfrac{m}{d} \geqslant 0$. On a alors $\dfrac{s-k}{d} > \dfrac{1}{p} \geqslant \dfrac{m}{d}$, donc $s-m-1 \geqslant k$, et $\dfrac{m+1}{d} > \dfrac{1}{p}$. Prenant alors q_j tel que $\dfrac{1}{q_j} = \dfrac{m+1-j}{p(m+1)}$, on obtient $q_0 = p < q_1 < q_2 < \cdots q_m$ tels que $\dfrac{1}{q_{j+1}} > \dfrac{1}{q_j} - \dfrac{1}{d}$ pour $0 \leqslant j < m$ et $\dfrac{1}{q_m} < \dfrac{1}{d}$. Il résulte alors du théorème 9.4.1 que, pour $0 \leqslant j < m$, $W^{s-j,q_j}(H_d)$ s'injecte continuement dans $W^{s-j-1,q_{j+1}}(H_d)$, donc que $W^{s,p}(H_d)$ s'injecte continuement dans $W^{s-m,q_m}(H_d)$, et du théorème précédent que ce dernier espace s'injecte continuement dans $\mathscr{C}_0^k(\overline{H_d})$ puisque $s-m-1 \geqslant k$.

Et de même, si K est un compact de $\mathbb{R}^d$, par composition d'injections compactes, on trouve que l'injection de $W_K^{s,p}(H_d)$ dans $\mathscr{C}_0^k(\overline{H_d})$ est compacte. ■

Soit s entier non nul. On suppose maintenant que l'ouvert Ω est borné dans $\mathbb{R}^d$ et que, pour tout point $a \in \partial\Omega$, il existe un $\mathscr{C}^s$-difféomorphisme Φ_a d'un voisinage V_a de a sur un voisinage U_a de 0 dans $\mathbb{R}^d$ tel que $\Phi_a(\Omega \cap V_a) = U_a \cap H_d$.
Le recouvrement $(V_j)_{j\in J}$, les ouverts bornés U_j de $\mathbb{R}^d$, les difféomorphismes Φ_j et la partition $\mathscr{C}^\infty$ de l'unité, $(\psi_j)_{j\in J}$, fournis par le lemme 9.2.1 permettent de construire des opérateurs linéaires continus $\Psi_{s,p} : W^{s,p}(\Omega) \to \prod_{j\in J} W^{s,p}(H_d)$ et $\Theta_{s,p} : \prod_{j\in J} W^{s,p}(H_d) \to W^{s,p}(\Omega)$ tels que $\Psi_{s,p}(f) = (\Phi_j^*(f.\psi_j))_{j\in J}$ et $\Theta_{s,p}((g_j)_{j\in J}) = \sum_j (\Phi_j^*)^{-1}(g_j)$, vérifiant $\Theta_{s,p}\circ\Psi_{s,p} = Id$. De plus, lorsque $f \in W^{s,p}(\Omega)$, la fonction $\Phi_j^*(f.\psi_j)$ appartient à $W^{s,p}_{K_j}(H_d)$ pour $K_j = \overline{U_j}$.

Théorème 9.4.4. *Soient $s \geqslant 1$ et $p < d$. Si g est dans $W^{s,p}(\Omega)$, alors g est dans $W^{s-1,q}(\Omega)$ pour q tel que $\dfrac{1}{q} > \dfrac{1}{p} - \dfrac{1}{d}$ et l'injection de $W^{s,p}(\Omega)$ dans $W^{s-1,q}(\Omega)$ est compacte.*

Démonstration : Soit $g \in W^{s,p}(\Omega)$. On a $\Phi_j^*(g.\psi_j) \in W^{s,p}_{K_j}(H_d) \subset W^{s-1,q}_{K_j}(H_d)$ pour tout $j \in J$. On en déduit que $g.\psi_j = (\Phi_j^*)^{-1}(\Phi_j^*(g.\psi_j)) \in W^{s-1,q}(\Omega)$ et enfin que $g = \sum_j g.\psi_j \in W^{s-1,q}(\Omega)$. Et puisque chaque injection τ_j de $W^{s,p}_{K_j}(H_d)$ dans $W^{s-1,q}_{K_j}(H_d)$ est compacte, l'injection $\tau = \Theta_{s-1,q}\circ(\prod_j \tau_j)\circ\Phi_{s,p}$ de $W^{s,p}(\Omega)$ dans $W^{s-1,q}(\Omega)$ est compacte. ∎

Le résultat précédent concernant la compacité de l'injection de $W^{1,p}(\Omega)$ dans $L^q(\Omega)$ est appelé *théorème de Rellich-Kondrachov*.

Théorème 9.4.5. *Soient $s \geqslant 1$, $p > \dfrac{d}{s}$ et $k \in \mathbb{N}$ tel que $k < s - \dfrac{d}{p}$. Si g est dans $W^{s,p}(\Omega)$, alors g est dans $\mathscr{C}^k(\overline{\Omega})$, et l'injection de $W^{s,p}(\Omega)$ dans $\mathscr{C}^k(\overline{\Omega})$ est compacte.*

Démonstration : Soit $g \in W^{s,p}(\Omega)$. On a $\Phi_j^*(g.\psi_j) \in W^{s,p}_{K_j}(H_d) \subset \mathscr{C}^k_0(\overline{H_d})$, comme précédemment, pour tout $j \in J$ et par composition avec le difféomorphisme Φ_j, on voit que $g.\psi_j$ est la la restriction à Ω d'une fonction de classe $\mathscr{C}^k$ sur $\mathbb{R}^d$, donc que g est la restriction à Ω d'une fonction de classe $\mathscr{C}^k$ sur $\mathbb{R}^d$. Et puisque chaque injection τ_j de $W^{s,p}_{K_j}(H_d)$ dans $\mathscr{C}^k_0(\overline{H_d})$ est compacte, l'injection $\tau = \Theta_k\circ(\prod_j \tau_j)\circ\Phi_{s,p}$ de $W^{s,p}(\Omega)$ dans $\mathscr{C}^k(\overline{\Omega})$ est compacte, où Θ_k est l'application linéaire continue de $\prod_j \mathscr{C}^k_0(\overline{H_d})$ dans $\mathscr{C}^k(\overline{\Omega})$ définie par $\Theta_k((g_j)_{j\in J}) = \sum_j g_j\circ\Phi_j$. ∎

9.5 Le cas limite

De fait, si $p < d$ et $\dfrac{1}{q} = \dfrac{1}{p} - \dfrac{1}{d}$, on peut montrer que $W^{1,p}(H_d)$ se plonge encore dans $L^q(H_d)$. Néanmoins, la compacité de l'opérateur de plongement est perdue, comme le montre l'exemple suivant, où $d = 2$, $p = 1$ et $q = 2$.

Exemple 9.5.1. *Il existe une suite bornée (g_k) de fonctions dans $W^{1,1}(H_2)$, dont les supports restent dans un même compact, mais dont aucune sous-suite ne converge dans $L^2(\mathbb{R}^2)$.*

Démonstration : Soit ρ une fonction de classe $\mathscr{C}^1$ positive à support dans $]-1, -\dfrac{1}{2}[$. On définit la fonction $g_k \in \mathscr{C}^1_c(H_2)$ par $g_k(x,y) = 2^k g(2^k x).g(2^k y)$. Clairement le support de

g_k est un compact contenu dans $]-1,0[\times]-1,0[$, donc dans H_2. On vérifie immédiatement que $\|g\|_1 = 2^{-k}\|\rho\|_1^2$ et que $\|\partial_x g\|_1 = \|\partial_y g\|_1 = \|\rho\|_1 . \|\rho'\|_1$, ce qui montre que la suite (g_k) est bornée dans $W^{1,1}(H_2)$. Enfin, on a

$$\|g_k\|_2^2 = \int 2^{2k}\rho^2(2^k x)\rho^2(2^k y)\,dx\,dy = \Big(\int_{-1}^{0}\rho^2(t)\,dt\Big)^2 = \|\rho\|_2^4 \ ,$$

et

$$\langle g_k, g_\ell\rangle = 2^{k+\ell}\int \rho(2^k x)\rho(2^\ell x)\rho(2^k y)\rho(2^\ell y)\,dx\,dy = 2^{k+\ell}\Big(\int \rho(2^k x)\rho(2^\ell x)\,dx\Big)^2 .$$

Cette quantité est nulle si $k < \ell$ puisque $\rho(2^k x) = 0$ si $x \leqslant -2^{-k}$ et $\rho(2^\ell x) = 0$ si $x \geqslant -2^{1-\ell} \geqslant -2^{-k}$. Il en résulte que $\|g_k - g_\ell\|_2^2 = \|g_k\|_2^2 + \|g_\ell\|_2^2 = 2\|\rho\|_2^4$ si $k \neq \ell$, et que la suite (g_k) ne peut avoir de sous-suite qui converge dans L^2. ■

Théorème 9.5.2. *Si f est dans $W^{1,1}(H_d)$, alors $f \in L^q(H_d)$, pour $q = \dfrac{d}{d-1}$, et on a $\|f\|_q \leqslant \|f\|_{W^{1,1}}$.*

Démonstration : On va établir ce résultat par récurrence sur d. Supposons d'abord $d = 1$. On a pour $a < 0$:

$$\int_{a-1}^{a}\Big((x-a+1)f'(x) + f(x)\Big)\,dx = \Big[(x-a+1)f(x)\Big]_{a-1}^{a} = f(a) \ ,$$

donc

$$|f(a)| \leqslant \int_{a-1}^{a}|f'(x)|\,dx + \int_{a-1}^{a}|f(x)|\,dx \leqslant \|f'\|_1 + \|f\|_1 = \|f\|_{W^{1,1}} \ .$$

Alors $f \in \mathscr{C}_0(\overline{H_1})$ et $\|f\|_\infty \leqslant \|f\|_{W^{1,1}}$. La même démonstration prouve que si $f \in W^{1,1}(\mathbb{R})$, on a aussi $f \in \mathscr{C}_0(\mathbb{R})$ et $\|f\|_\infty \leqslant \|f\|_{W^{1,1}}$.

Supposons maintenant le résultat établi pour $d-1 \geqslant 1$, et soit $f \in W^{1,1}(H_d)$. Pour $t \in \mathbb{R}$ et $y \in H_{d-1}$, le point $x = (y,t)$ appartient à H_d. On va alors noter f_t la fonction $y \mapsto f(y,t)$ sur H_{d-1} et f^y la fonction $t \mapsto f(y,t)$ sur $\mathbb{R}$. On a :

$$\int_{\mathbb{R}}\|f_t\|_{W^{1,1}}\,dt = \int_{\mathbb{R}}\Big(\int |f_t(y)|\,dy + \sum_{j=1}^{d-1}\int|\partial_j f_t(y)|\,dy\Big)\,dt \ ,$$

et en vertu du théorème de Fubini

$$\int_{\mathbb{R}}dt\int_{H_{d-1}}\Big(|f_t(y)| + \sum_{j=1}^{d-1}|\partial_j f_t(y)|\Big)\,dy = \|f\|_1 + \sum_{j=1}^{d-1}\|\partial_j f\|_1 \leqslant \|f\|_1 + \sum_{j=1}^{d}\|\partial_j f\|_1 = \|f\|_{W^{1,1}} \ .$$

Il en résulte que la fonction g définie sur $\mathbb{R}$ par $g(t) = \|f_t\|_{W^{1,1}}$ est dans $L^1(\mathbb{R})$ avec $\|g\|_1 \leqslant \|f\|_{W^{1,1}}$.

De même, on a

$$\int_{H_{d-1}}dy\int_{\mathbb{R}}\Big(|f^y(t)| + |\partial f^y(t)|\Big)\,dt \leqslant \|f\|_1 + \|\partial_d f\|_1 \leqslant \|f\|_{W^{1,1}} \ ,$$

ce qui montre que la fonction h définie sur H_{d-1} par $h(y) = \|f^y\|_{W^{1,1}}$ est dans $L^1(H_{d-1})$, avec $\|h\|_1 \leqslant \|f\|_{W^{1,1}}$.

Pour t fixé tel que $g(t) < +\infty$, il résulte de l'hypothèse de récurrence que $f_t \in L^r(H_{d-1})$, où $\dfrac{1}{r} = 1 - \dfrac{1}{d-1}$, avec $\|f_t\|_r \leqslant g(t)$. Et puisque $\|f_y\|_\infty \leqslant \|f\|_{W^{1,1}}$, on a

$$|f(y,t)| = |f^y(t)| \leqslant \|f^y\|_\infty \leqslant \|f^y\|_{W^{1,1}} = h(y) \ .$$

On en déduit que $|f_t(y)|^q \leqslant |f_t(y)|\, h(y)^{q-1}$ et que, par l'inégalité de Hölder,

$$\|f_t\|_q^q = \int_{H_{d-1}} |f_t(y)|^q \ dy \leqslant \Big(\int |f_t(y)|^r \ dy\Big)^{1/r} \Big(\int h(y)^\beta \, dy\Big)^{1-\frac{1}{r}} \ ,$$

où $\beta = \dfrac{q-1}{1-\frac{1}{r}} = \dfrac{\frac{1}{d-1}}{\frac{1}{d-1}} = 1$. Il en résulte que

$$\|f_t\|_q^q \leqslant \|f_t\|_r \, \|h\|_1^{q-1} \leqslant \|h\|_1^{q-1} .g(t) \ ,$$

et finalement que

$$\|f\|_q^q = \int_{\mathbb{R}} \|f_t\|_q^q \ dt \leqslant \|f\|_{W^{1,1}}^{1-\frac{1}{r}} . \int_{\mathbb{R}} g(t)\, dt \leqslant \|f\|_{W^{1,1}}^{q-1} \|g\|_1 \leqslant \|f\|_{W^{1,1}}^{q-1} \|f\|_{W^{1,1}} = \|f\|_{W^{1,1}}^q \ ,$$

d'où $\|f\|_q \leqslant \|f\|_{W^{1,1}}$. ■

Cette démonstration ne s'étend pas aisément à $W^{1,p}$ pour $p > 1$. On va donc employer, pour le cas $p > 1$, une autre méthode, qui, elle, ne fonctionne pas lorsque $p = 1$.

Théorème 9.5.3. *Si $s \geqslant 1$ est entier, $1 < p < d$ et si q est défini par $\dfrac{1}{q} = \dfrac{1}{p} - \dfrac{1}{d}$, on a $W^{s,p}(H_d) \subset W^{s-1,q}(H_d)$, et cette injection est continue.*

Démonstration : On procède comme pour la démonstration du théorème 9.3.3 en remarquant que chacune des fonctions $\mathbb{1}_{B^+}$ et χ_j est majorée par la fonction $x \mapsto \|x\|^{1-d}$. L'application du théorème 6.5.7 montre donc que, si $g \in W^{s,p}$ et $|\alpha| < s$, la fonction $\partial_\alpha g * \chi_j$ est définie presque partout et appartient à $L^q(H_d)$, avec $\|\partial_\alpha g * \chi_j\|_q \leqslant C(p,d) \|\partial_\alpha g\|_p$, c'est-à-dire que

$$\partial_\alpha g = \delta_0 * \partial_\alpha g = \frac{1}{c}(\mathbb{1}_{B^+} * \partial_\alpha g + \sum_j \chi_j * \partial_j \partial_\alpha g) \in L^q(H_d) \ ,$$

ou encore que $g \in W^{s-1,q}(H_d)$, avec $\|g\|_{W^{s-1,q}} \leqslant C \|g\|_{W^{s,p}}$ pour une constante C ne dépendant que de p et d. ■

En utilisant les méthodes de la section 9.4, on montrerait de même que, sous réserve que Ω vérifie les conditions convenables, on a $W^{s,p}(\Omega) \subset W^{s-1,q}(\Omega)$. Et on remarque que cette inclusion peut ne pas être vérifiée si on n'impose aucune condition sur Ω.

Exemple 9.5.4. *Soit Ω l'ouvert borné de $\mathbb{R}^2$ défini par $\Omega = \{(x,y) : 0 < x < 1 \text{ et } |y| < x^5\}$. Alors la fonction $g : (x,y) \mapsto x^{-4}$ est dans $W^{1,p}(\Omega)$ pour $1 \leqslant p < \frac{6}{5}$, mais pas dans $L^q(\Omega)$ pour $\frac{3}{2} \leqslant q$. En particulier, pour $p = 1$ et $q = \frac{3}{2}$, on a $\frac{1}{q} > \frac{1}{p} - \frac{1}{2}$, mais $W^{1,p}(\Omega) \not\subset L^q(\Omega)$.*

Démonstration : On a

$$\int_\Omega g(x,y)^p \, dx\, dy = \int_0^1 x^{-4p} \, dx \int_{-x^5}^{x^5} dy = 2 \int_0^1 x^{5-4p} \, dx < +\infty \text{ si } p < \frac{3}{2} \ ,$$

$$\int_\Omega |\partial_x g|^p \ dx\, dy = \int_0^1 4^p x^{-5p} \, dx \int_{-x^5}^{x^5} dy = 2^{2p+1} \int_0^1 x^{5-5p} \, dx < +\infty \text{ si } p < \frac{6}{5} \ ,$$

$$\int_\Omega |\partial_y g|^p \ dx\, dy = 0 \ ,$$

mais $\int_\Omega g(x,y)^q \, dx\, dy = +\infty$ si si $q \geqslant \frac{3}{2}$. ■

Au contraire du cas où $p < d$ et $\frac{1}{q} = \frac{1}{p} - \frac{1}{d}$ pour lequel $W^{1,p}$ se plonge dans L^q, l'espace $W^{1,p}(\Omega)$ ne se plonge pas dans $L^\infty(\Omega)$ lorsque $p = d \geqslant 2$, même pour un ouvert borné comme la boule unité.

Exemple 9.5.5. *Soit D le disque unité du plan $\mathbb{R}^2$. La fonction $f : x \mapsto \log^{1/3}(2/\|x\|)$ est dans $W^{1,2}(D)$ sans être localement bornée.*

Il est clair que f n'est bornée sur aucun voisinage de 0. Néanmoins, on a

$$\int_D |f(x,y)|^2 \ dx\, dy = \int_0^{2\pi} d\theta \int_0^1 \log^{2/3}(2/r)\, r\, dr \leqslant 2\pi \sup_{0<r\leqslant 1} r \log^{2/3}(2/r) < +\infty$$

et

$$\begin{aligned} \int |\partial_x f(x,y)|^2 \ dx\, dy = \int |\partial_y f(x,y)|^2 \ dx\, dy &= \int_0^{2\pi} \cos^2\theta \, d\theta \int_0^1 \frac{1}{9r^2 \log^{4/3}(2/r)} \, r\, dr \\ &= \frac{\pi}{9} \int_0^1 \frac{dr}{r \log^{4/3}(2/r)} = \frac{\pi}{9} \Big[3 \log^{-1/3}(2/r) \Big]_0^1 = \frac{\pi \log^{-1/3}(2)}{3} < \infty \ , \end{aligned}$$

ce qui montre que $f \in W^{1,2}(D)$. ■

Appendice

LE THÉORÈME DE ZORN ET L'AXIOME DU CHOIX

10.1 L'axiome du choix

On adjoint souvent aux axiomes usuels de la Théorie des Ensembles (ZF = les axiomes de Zermelo-Fraenkel) un axiome indépendant, appelé *axiome du choix* (AC). Cet axiome, qui a l'air d'un "axiome de bon sens", est souvent très utile (et utilisé à plusieurs reprises dans ce cours : cf. le théorème de Tychonoff et le théorème de Hahn-Banach). Il permet aussi d'obtenir des réultats paradoxaux, comme l'existence d'ensembles non mesurables ou de formes linéaires discontinues sur un espace de Banach. L'énoncé de cet axiome qui semble le plus "couler de source" est le suivant :

Tout produit d'ensembles non vides est non vide.

c'est-à-dire que si I est un ensemble, et $(E_i)_{i\in I}$ une famille de sous-ensembles tous non vides d'un ensemble E, on peut "choisir un élément dans chaque E_i" : de façon précise, il existe un $\xi \in \prod_{i\in I} E_i$. Un tel ξ est *une fonction* de I dans E telle que $\xi(i) \in E_i$ pour tout i, c'est-à-dire une partie de $I \times E$ contenant pour tout $i \in I$ *exactement un* élément (i, x) pour $x \in E$, avec de plus $x \in E_i$.

10.2 Le théorème de Zorn

Définitions 10.2.1. *Soit $(\mathscr{O}, \leqslant)$ un ensemble ordonné. On dit qu'une partie H de $\mathscr{O}$ est une* chaîne *si elle est totalement ordonnée par $\leqslant$ (c'est-à-dire que pour tout f et tout g dans H, on a $f \leqslant g$ ou $g \leqslant f$).*
On dit que $h_0 \in \mathscr{O}$ est la borne supérieure *de la chaîne H s'il en est le plus petit majorant, c'est-à-dire si* $\quad (\forall h \in H \quad h \leqslant f) \iff h_0 \leqslant f$.
On dit enfin que l'ordre $\leqslant$ est inductif *sur $\mathscr{O}$ si toute chaîne H de $\mathscr{O}$ possède une borne supérieure.*

Définition 10.2.2. *Soit $(\mathscr{O}, \leqslant)$ un ensemble ordonné. Un élément a de $\mathscr{O}$ est dit* maximal *s'il n'est strictement majoré par aucun élément de $\mathscr{O}$.*

Il est clair que si a est le plus grand élément de $\mathscr{O}$, il est maximal, mais lorsque l'ordre n'est pas total, il peut exister plusieurs éléments maximaux ; ces éléments maximaux sont alors incomparables.

Théorème 10.2.3. (Zorn) *Soit $(\mathscr{O}, \leqslant)$ un ensemble ordonné inductif. S'il existe une fonction $f : \mathscr{O} \to \mathscr{O}$ vérifiant $f(x) \geqslant x$ pour tout $x \in \mathscr{O}$ et $f(x) \neq x$ pour x non maximal, alors tout élément $a \in \mathscr{O}$ est majoré par un élément maximal.*

Clairement une telle fonction existe pour tout ensemble ordonné, sous l'axiome du choix : il suffit de considérer, pour $x \in \mathscr{O}$, l'ensemble

$$P_x = \begin{cases} \{x\} & \text{si } x \text{ est maximal,} \\ \{y \in \mathscr{O} : x \leqslant y \text{ et } x \neq y\} & \text{sinon,} \end{cases}$$

et de choisir $f \in \prod_{x \in \mathscr{O}} P_x$. Le théorème de Zorn est donc démontrable dans ZFC, le système d'aximes obtenu en ajoutant l'axiome du choix à ceux de Zemelo-Fraenkel.

Inversement, si $(E_i)_{i \in I}$ est une famille de sous-ensembles non vides de E, on peut considérer l'ensemble $\mathscr{O}$ des fonctions partielles f de domaine $J \subset I$ à valeurs dans E satisfaisant $f(j) \in E_j$ pour tout $j \in J$, et l'ordonner par la relation d'ordre "$f \leqslant g \iff g$ prolonge f". Cet ordre est clairement inductif. Et un élément maximal ne peut être qu'une fonction à domaine total : si f maximale avait un domaine $J \neq I$, on pourrait trouver $i \in I \setminus J$ et $x \in E_i$ et il suffirait de prolonger f en lui donnant en i la valeur x pour contredire la maximalité de f.
On en conclut que l'énoncé du théorème de Zorn est équivalent dans ZF à l'axiome du choix.

Démonstration : Soit $\mathscr{X}$ l'ensemble des parties X de $\mathscr{O}$ contenant a, stables par f et inductives. Il est clair que l'intersection X_0 de tous les éléments de $\mathscr{X}$ est un élément de $\mathscr{X}$, donc le plus petit élément de $\mathscr{X}$. On va montrer que X_0 est une chaîne, donc contient sa borne supérieure m : il en résultera que $f(m) \in X_0$, donc que $f(m) \leqslant \sup(X_0) = m \leqslant f(m)$, et enfin que m est maximal.
On pose pour cela

$$X_1 := \{x \in X_0 : \forall y \in X_0 \quad x \leqslant y \text{ ou } f(y) \leqslant x\}$$

et

$$X_2 := \{z \in X_0 : \forall x \in X_1 \quad x \leqslant z \Longrightarrow (x = z \text{ ou } f(x) \leqslant z)\} \ .$$

On remarque immédiatement que tout élément x de X_1 est comparable avec tout élément y de X_0, puisque $f(y) \leqslant x$ entraîne $y \leqslant f(y) \leqslant x$. On voit aussi que $A = \{y : y \geqslant a\}$ appartient à $\mathscr{X}$ donc que $X_0 \subset A$, c'est-à-dire que a est le plus petit élément de X_0. Ceci entraîne que $a \in X_1$ et $a \in X_2$.

Lemme 10.2.4. *X_1 et X_2 sont inductifs.*

Démonstration : Soient Y une chaîne de X_1 de borne supérieure x^* et $y \in X_0$. Si tout $x \in Y$ est au plus égal à y, on a $x^* \leqslant y$. Dans le cas contraire, il existe $x' \in Y$ tel que $f(y) \leqslant x' \leqslant x^*$. Ceci montre que $x^* \in X_1$.
De même, si Z est une chaîne de X_2 de borne supérieure z^* et $x \in X_1$ avec $x \leqslant z^*$, on a pour tout $z \in Z$: $x \leqslant z$ ou $f(z) \leqslant x$. Si, pour tout $z \in Z$ on a $z \leqslant f(z) \leqslant x$, on a $z^* \leqslant x \leqslant z^*$, donc $x = z^*$. Si au contraire, il existe $z' \in Z$ tel que $x < z'$, on a $f(x) \leqslant z' \leqslant z^*$, donc $z^* \in X_2$. ■

Lemme 10.2.5. *L'ensemble X_2 est stable par f.*

DÉMONSTRATION : Soient $y \in X_2$ et $z = f(y)$. On veut montrer que $z \in X_2$. Soit donc $x \in X_1$ tel que $x \leqslant z$ et $x \neq z$. Il faut prouver que $f(x) \leqslant z$.
Alors, puisque $x \in X_1$, ou bien $x \leqslant y$, ou bien $f(y) \leqslant x \leqslant z = f(y)$. Dans le second cas, on a $x = z$, contrairement à l'hypothèse. Dans le premier, puisque $y \in X_2$, on conclut que $x = y$, c'est-à-dire $f(x) = f(y) = z$, ou $f(x) \leqslant y \leqslant f(y) = z$. Donc $z \in X_2$. ■

Lemme 10.2.6. *L'ensemble X_1 est stable par f.*

DÉMONSTRATION : On conclut du lemme précédent que $X_2 \in \mathscr{X}$, donc que $X_0 \subset X_2 \subset X_0$. Et puisque $X_2 = X_0$, on a pour tout $x \in X_1$ et tout $y \in X_0$ tels que $x \leqslant y$, soit $x = y$ soit $f(x) \leqslant y$.
Soient $x \in X_1$, $x' = f(x)$ et $y \in X_0$. Il faut montrer que $x' \leqslant y$ ou que $f(y) \leqslant x'$. Si on a $x \leqslant y$, on a soit $x' = f(x) \leqslant y$, soit $x = y$ c'est-à-dire $f(y) = f(x) = x'$. Et dans le cas contraire on a $f(y) \leqslant x \leqslant f(x) = x'$. Donc $x' \in X_1$.
Il en résulte que X_1 est stable par f. ■

DÉMONSTRATION : On conclut du Lemme 10.2.6 que $X_1 \in \mathscr{X}$, donc que $X_0 \subset X_1 \subset X_0$, c'est-à-dire $X_1 = X_0$, donc que deux éléments quelconques de X_0 sont toujours comparables, ou encore que X_0 est totalement ordonné et que la borne supérieure m de X_0 est un élément maximal supérieur à a. Et ceci achève la démonstration du théorème 10.2.3. ■

On va donner dans ce qui suit quelques résultats utilisant l'Axiome du Choix.

10.3 Une version ordonnée du théorème de Hahn-Banach

Soit E un espace vectoriel sur $\mathbb{R}$. Un *ordre compatible avec la structure d'espace vectoriel* est une relation d'ordre $\leqslant$ sur E vérifiant les deux propriétés :

i) si $x \leqslant y$ et $z \in E$, on a $x + z \leqslant y + z$

ii) si $x \leqslant y$ et $\lambda \in \mathbb{R}^+$, on a $\lambda.x \leqslant \lambda.y$.

Une telle relation d'ordre est entièrement définie par le cône convexe $E_+ = \{x \in E : x \geqslant 0\}$ des éléments positifs, puisque $x \leqslant y \iff y - x \in E_+$.

Une forme linéaire f sur E est dite *positive* si $f(x) \geqslant 0$ pour tout $x \geqslant 0$, ou, ce qui revient au même, si $f(x) \leqslant f(y)$ lorsque $x \leqslant y$.

Théorème 10.3.1. (AC) *Soient E un $\mathbb{R}$-espace vectoriel ordonné, V un sous-espace vectoriel cofinal de E (c'est-à-dire que tout élément de E est majoré par un élément de V) et $f : V \to \mathbb{R}$ une forme linéaire positive sur V. Alors il existe une forme linéaire positive $\tilde{f}$ sur E qui prolonge f.*

DÉMONSTRATION : On considère l'ensemble $\mathscr{O}$ des couples (M, g) où M est un sous-espace vectoriel cofinal et $g : M \to \mathbb{R}$ un prolongement linéaire positif de f. On munit $\mathscr{O}$ de l'ordre : $(M, g) \leqslant (M', g') \iff M \subset M'$ et $g'_{|M} = g$. On vérifie aisément, comme dans la démonstration du théorème 4.4.1, que cet ordre est inductif, et on conclut que (V, f) est majoré par un élément maximal (M^*, g^*). Il suffit de montrer que si M^* n'était pas égal à E, (M^*, g^*) ne pourrait être maximal.

Supposons donc qu'existe $a \in E \setminus M^*$, posons $M' = M^* \oplus \mathbb{R}.a$ et définissons, comme dans la preuve de 4.4.1, une forme linéaire g sur M' par $g(m+ta) = g^*(m) + t\alpha$ pour $m \in M^*$ et $t \in \mathbb{R}$. Pour que g soit positive, il faut que $\alpha \geqslant -g^*(\frac{m}{t})$ si $m+ta \geqslant 0$ et $t > 0$ et que $\alpha \leqslant -g^*(\frac{m}{t})$ si $m+ta \geqslant 0$ et $t < 0$, c'est-à-dire $\alpha \geqslant g^*(m)$ si $a \geqslant m$ et $\alpha \leqslant g^*(m)$ si $a \leqslant m$. Comme dans la preuve de 4.4.1, il suffit de prouver que les ensembles $P_1 = \{g^*(m) : m \in M^*, m \leqslant a\}$ et $P_2 = \{g^*(m) : m \in M^*, m \geqslant a\}$ sont non vides et que tout élément de P_1 est inférieur à tout élément de P_2.
Il résulte de l'hypothèse de cofinalité qu'existe $v \in V \subset M^*$ tel que $a \leqslant v$ (donc $g^*(v) \in P_2$ et $P_2 \neq \emptyset$) et qu'existe $v' \in V \subset M^*$ tel que $-a \leqslant v'$ (donc $a \geqslant -v'$, $-g^*(v') \in P_1$ et $P_1 \neq \emptyset$). Enfin, si $s \in P_1$ et $t \in P_2$, il existe m et m' dans M^* tels que $m \leqslant a \leqslant m'$, $s = g^*(m)$ et $t = g^*(m')$. Mais puisque $m' - m \geqslant 0$, on a alors $t - s = g^*(m' - m) \geqslant 0$, d'où $s \leqslant t$. On a donc $-\infty < \sup(P_1) \leqslant \inf(P_2) < +\infty$, et on peut prendre $\alpha = \sup(P_1)$. Ceci achève de montrer que (M^*, g^*) n'est pas maximal si $M^* \neq E$. ■

Remarque 10.3.2. On peut en fait déduire le théorème 4.4.1 de cet énoncé. On munit $E_1 = E \times \mathbb{R}$ de l'ordre défini par $(x,t) \geqslant 0 \iff t \geqslant p(x)$. Si V est un sous-espace de E et f une forme linéaire sur V majorée par p, la forme linéaire F définie sur $V_1 = V \times \mathbb{R}$ par $F(x,t) = t - f(x)$ est positive et V_1 est cofinal (en effet $L = \{0\} \times \mathbb{R}$ est cofinal puisque $(x,t) \leqslant (0, p(-x) + t)$ et $V_1 \supset L$). Alors, si G est une forme linéaire positive sur E_1 prolongeant F, on voit que

$$G(x,t) = G(x,0) + G(0,t) = G(x,0) + F(0,t) = t - g(x) ,$$

où $g : x \mapsto -G(x,0)$ est une forme linéaire sur E telle que $p(x) - g(x) = G(p(x), x) \geqslant 0$, puisque $(p(x), x) \geqslant 0$, c'est-à-dire que $g \leqslant p$ et que g prolonge f.

10.4 Bases algébriques d'un espace vectoriel

Théorème 10.4.1. (AC) *Soient $\mathbb{K}$ un corps commutatif, E un espace vectoriel sur $\mathbb{K}$ et $a \in E$ non nul. Alors il existe une base de E, c'est-à-dire une partie génératrice et libre, qui contient a.*

Démonstration : On considère l'ensemble $\mathscr{O}$ des parties de E linéairement indépendantes sur $\mathbb{K}$, qu'on ordonne par inclusion. Alors cet ensemble ordonné est inductif : si $\mathscr{H}$ est une chaîne de $\mathscr{O}$ et si on pose $X = \bigcup \mathscr{H} = \{x \in E : \exists H \in \mathscr{H} \quad x \in H\}$, on définit clairement un ensemble X tel que $H \subset X$ pour tout $H \in \mathscr{H}$: X est alors un majorant de la chaîne $\mathscr{H}$. Et si Y est un majorant de $\mathscr{H}$, on a aussi clairement $X \subset Y$, c'est-à-dire que X est la borne supérieure de $\mathscr{H}$.
De plus, la partie X est linéairement indépendante dans E. En effet, si on a une combinaison linéaire nulle d'éléments de X : $\sum_{j=1}^m \lambda_j x_j = 0$, avec $\lambda_j \in \mathbb{K}$ et $x_j \in X$, il existe puisque $\mathscr{H}$ est une chaîne, un $H_0 \in \mathscr{H}$ qui contient tous les x_j. Alors l'indépendance linéaire de H_0 montre que les coefficients λ_j sont tous nuls.

On vérifie également sans peine que si X est un élément maximal de $\mathscr{O}$, c'est une partie génératrice : si V est l'espace engendré par X, c'est-à-dire l'ensemble des combinaisons linéaires *finies* d'éléments de X, il faut prouver que $V = E$. En effet, s'il existait un $u \in E \setminus V$, on voit que $Y := X \cup \{u\}$ appartiendrait à $\mathscr{O}$ et majorerait strictement X, en contradiction avec la maximalité de ce dernier.

On conclut alors du théorème de Zorn que la partie $\{a\}$, qui appartient à $\mathscr{O}$, est contenue dans une partie maximale de $\mathscr{O}$, qui est une base. ■

Corollaire 10.4.2. (AC) *Il existe dans* $\mathbb{R}$ *une partie qui n'est pas mesurable pour la mesure de Lebesgue.*

DÉMONSTRATION : Puisque $\mathbb{R}$ est un espace vectoriel sur le corps $\mathbb{Q}$ des rationnels, il existe une base algébrique B de $\mathbb{R}$ sur $\mathbb{Q}$ contenant l'élément 1. Il existe alors une unique forme $\mathbb{Q}$-linéaire ψ sur $\mathbb{R}$ telle que $\psi(1) = 1$ et $\psi(b) = 0$ pour tout $b \in B \setminus \{1\}$. Alors $\psi(\mathbb{R}) = \mathbb{Q}$ et on pose $V = \ker \psi$.

On va montrer que V ne peut être mesurable pour la mesure de Lebesgue μ. Supposons par l'absurde qu'il le soit. Alors, pour m et n entiers, les ensembles $V_{m,n} = ([-m,m] \cap V) + 1 - 2^{-n}$ seraient tous mesurables. Comme on a $\psi(v) = 0$ pour tout $v \in V$, on a $\psi(v) = 1 - 2^{-n}$ pour tout $v \in V_{m,n}$, ce qui montre que les $(V_{m,n})_n$ sont deux-à-deux disjoints pour m fixé.

En raison de l'invariance par translation de μ, on aurait pour tout n : $\mu(V_{m,n}) = \mu(V_{m,0})$. Et puisque $\bigcup_{n\in\mathbb{N}} V_{m,n} \subset [-m, m+1]$, on devrait avoir $\sum_n \mu(V_{m,n}) \leqslant 2m+1$ pour tout m, donc $\mu(V \cap [-m,m]) = \mu(V_{m,0}) = 0$, et $\mu(V) = \sup_m \mu(V \cap [-m,m]) = 0$.
Enfin, pour $q \in \mathbb{Q}$, on a $\psi^{-1}(q) = V + q$, donc $\mathbb{R} = \bigcup_{q\in\mathbb{Q}} \psi^{-1}(q) = \bigcup_{q\in\mathbb{Q}} V + q$ serait réunion dénombrable d'ensembles μ-négligeables, d'où la contradiction. ■

Corollaire 10.4.3. (AC) *Sur tout espace normé* E *de dimension infinie, il existe une forme linéaire discontinue.*

DÉMONSTRATION : Soit à nouveau B une base algébrique de E sur $\mathbb{K}$ ($=\mathbb{R}$ ou $\mathbb{C}$). Puisque E est de dimension infinie, B est une partie infinie de E, et on peut trouver une application injective ψ de $\mathbb{N}$ dans B. On peut alors définir sur E une forme linéaire f telle que $f(x) = 0$ si $x \in B \setminus \psi(\mathbb{N})$ et $f(\psi(n)) = 2^n \|\psi(n)\|$. Et puisque $\sup_{y\neq 0} \dfrac{|f(y)|}{\|y\|} \geqslant \sup\limits_n \dfrac{f(\psi(n))}{\|\psi(n)\|} = +\infty$, la forme linéaire f ne peut être continue. ■

10.5 Le théorème de Zermelo

L'axiome du choix est encore équivalent à l'axiome que tout ensemble peut être muni d'un *bon ordre*, c'est-à-dire d'un ordre total pour lequel toute partie non vide possède un plus petit élément (pas seulement une borne inférieure : $\mathbb{R}^+$ n'est pas bien ordonné !).

Lemme 10.5.1. *Si tout ensemble peut être bien ordonné, l'axiome du choix est satisfait.*

Soient $(E_i)_{i\in I}$ une famille d'ensembles non vides, et $E = \bigcup_{i\in I} E_i$. Notons $\leqslant$ un bon ordre sur E. Alors, si on pose $x_i = \min(E_i)$, qui est bien défini puisque la partie non vide E_i de E possède un plus petit élément, alors $\xi = (x_i)_{i\in I}$ est bien un point de $\prod_{i\in I} E_i$: la formule

$$G = \{(i,y) \in I \times E : y \in E_i \text{ et } \forall z \ (z \in E_i \Longrightarrow y \leqslant z)\}$$

définit bien un ensemble, qui est le graphe de la fonction ξ. ■

Inversement, si l'axiome du choix est vérifié, on montre que tout ensemble peut être bien ordonné.

Théorème 10.5.2. (Zermelo) *Supposons que l'axiome du choix soit satisfait. Alors tout ensemble E peut être muni d'un bon ordre.*

Notons $\mathscr{O}$ l'ensemble des couples (X, R) où X est une partie de E et R une relation de bon ordre sur X. L'ensemble $\mathscr{O}$ n'est pas vide puisque l'on peut prendre $X = \emptyset$ et $R = \emptyset$. Définissons ensuite la relation d'ordre $\leqslant$ sur $\mathscr{O}$ par $(X, R) \leqslant (Y, S)$ si $X \subset Y$ et si X est une section commençante de Y, c'est-à-dire : pour tout x de X et tout y de Y on a

$$y\ S\ x \iff (y \in X) \text{ et } (y\ R\ x)$$

On a alors en particulier, pour x et x' dans X : $x\ R\ x' \iff x\ S\ x'$.

L'ordre $\leqslant$ est inductif sur $\mathscr{O}$: si (X_i, R_i) est une chaîne dans $\mathscr{O}$ pour $\leqslant$, on prend $X = \bigcup_i X_i$ et on définit R sur X par $x\ R\ x'$ s'il existe $i \in I$ tel que $x\ R_i\ x'$. On vérifie sans peine que R est bien une relation d'ordre total sur X. De plus, si M est une partie non vide de X, il existe un $i \in I$ tel que $M \cap X_i \neq \emptyset$, donc un $m \in M$ qui est le plus petit élément de $M \cap X_i$. Alors, si $m' \in M$ il existe $j \geqslant i$ tel que $m' \in X_j$. Et si $m'\ R_j\ m$, on a nécessairement $m' \in X_i$ et $m'\ R_i\ m$, donc $m' = m$, ce qui montre que $m = \min M$.

Il reste à montrer qu'un élément maximal (X, R) vérifie nécessairement $X = E$. Supposons par l'absurde qu'existe un $a \in E \setminus X$. On posera alors $X' = X \cup \{a\}$ et on définira R' par "$x\ R'\ x' \iff (x \in X$ et $x' \in X$ et $x\ R\ x')$ ou $(x' = a)$". On vérifie alors sans peine que $(X', R') \in \mathscr{O}$, que $(X, R) \neq (X', R')$ et que $(X, R) \leqslant (X', R')$, ce qui contredit la maximalité de (X, R). ■

Voici, pour conclure, un exemple d'application du théorème de Zermelo.

Théorème 10.5.3. *Soient X un espace métrique et $(O_i)_{i \in I}$ un recouvrement ouvert de X. Alors il existe une partition de l'unité localement finie $(\varphi_i)_{i \in I}$ subordonnée au recouvrement $(O_i)_{i \in I}$.*

Démonstration : Quitte à remplacer la distance d sur X par $\inf(1, d)$, on peut supposer que $d \leqslant 1$. Munissons l'ensemble I d'un bon ordre $\leqslant$. Pour tout $i \in I$, la fonction $f_i : x \mapsto d(x, O_i^c)$ est 1-lipschitzienne, majorée par 1 et nulle hors de O_i.

On pose alors $\psi_i(x) = \sup_{j \leqslant i} f_j(x) - \sup_{j < i} f_j(x)$. Les fonctions $\sup_{j<i} f_j$ et $\sup_{j \leqslant i} f_j$ sont 1-lipschitziennes, et on en déduit que ψ_i est 2-lipschitzienne. Pour tout $a \in X$, il existe un plus petit $i \in I$ tel que $x \in O_i$. Alors $\sup_{j<i} f_j(a) = 0$ et $\sup_{j \leqslant i} f_j(a) = f_i(a) > 0$, d'où on tire $\psi_i(a) > 0$. La fonction $\psi = \sup_{i \in I} \psi_i$ est aussi 2-lipschitzienne et partout non nulle.

On voit aisément que $\psi_i = 0$ hors de O_i et que $0 < \sum_{i \in I} \psi_i = \sup_{i \in I} f_i \leqslant 1$, donc que, pour tout a, l'ensemble $J(a) = \{j \in I : \psi_j(a) > \psi(a)/3\}$ est non vide et fini (avec au plus $\dfrac{3}{\psi(a)}$ éléments). Alors, si $r = \psi(a)/18$ et si $U_j = \{x : \psi_j(x) > \psi(x)/2\}$ rencontre $B(a, r)$ en x, on a

$$\psi_j(a) - \frac{\psi(a)}{3} = \frac{\psi(a)}{6} + \psi_j(a) - \frac{\psi(a)}{2} \geqslant \frac{\psi(a)}{6} + \psi_j(x) - \frac{\psi(x)}{2} - 3d(x, a) > \frac{\psi(a)}{6} - 3r > 0\ ,$$

donc $j \in J(a)$. On en déduit que la famille $\psi_i' = (\psi_i - \frac{1}{2}\psi)_+ := \sup(0, \psi_i - \frac{1}{2}\psi)$ est localement finie. Il suffit alors, comme en 1.7.4, de poser $\varphi_i = \dfrac{\psi_i'}{\sum_{j \in I} \psi_j'}$, pour obtenir la partition de l'unité localement finie cherchée. ■

Index des définitions

absorbant 52, 79
adhérence 4
adjoint 64
algèbre de parties 91

base
  de voisinages 4
  hilbertienne 64
bidual 84
bon ordre 209
borné 83
borélien 91
boule
  fermée 1
  ouverte 1
  unité 41

$\mathscr{C}(K, F)$ 21
chaîne 205
compact 15
compactifié d'Alexandroff 23
complet 31
conjugués 125
continue 8
continuité uniforme 13
contraction 34
convergence
  normale 42
  uniforme 12
convolution 137
  de deux distributions 172
  distribution-fonction 168
croissance lente 181

$\mathscr{D}(\Omega)$ 151
décroissance rapide 179
dénombrable à l'infini 23
dense 4
dérivée d'une distribution 162
diamètre 1
discrète 3
distance 1
  à un fermé 3
distances équivalentes 1
distribution 155
  d'ordre m 156
  égale à une fonction 157
  homogène 167
  tempérée 183
dual
  d'un espace localement convexe 75
  d'un espace normé 49

$\mathscr{E}'(\Omega)$ 159
$\mathscr{E}(\Omega)$ 159
élément maximal 205
ensemble mesurable 92
équicontinu 36
équilibré 74
espace
  de Banach 41
  de Fréchet 77
  de Hilbert 58
  mesurable 92
  mesuré 96
  métrique 1
  métrisable 12
  préhilbertien 56
  vectoriel topologique 79

$f * g$ 137
fermé 2
filtre 25
fonction
  de test 151
  élémentaire 94
  intégrable 100, 102
  localement intégrable 142
  maximale 142
  mesurable 92

homéomorphe 12
homéomorphisme 12

inductif 205
intégrale
  d'une fonction 103
  sur un ensemble 100

intérieur 4
isométrie 13
isomorphisme 51

jauge 79

$\mathscr{L}(E,F)$ 48
ℓ^1 46
$L^1(X,\mu)$ 104
L^1_{loc} 142
$L^1_S(\mathbb{R}^d)$ 181
ℓ^2 58
limite
  d'une suite 6
  inductive d'espaces de Fréchet 88
lipschitzienne 11
localement
  compact 23
  convexe 80
  finie 29
$L^p(X,\mu)$ 127

maigre 35
mesure 95
  de Dirac 157
  de Lebesgue 116
  localement finie 112
  σ-additive 95
  σ-finie 95
$\mu \otimes \nu$ 109

négligeable 96
norme 41
  d'une application linéaire 48
  de la convergence uniforme 46
normes équivalentes 41

$\mathscr{O}_M(\mathbb{R}^d)$ 181
opérateur 51
  auto-adjoint 66
  compact 53
  normal 66
orthogonal 56
  d'un sous-espace 60
orthonormale 63
ouvert 2

partition
  de l'unité 29
  lisse de l'unité 154

partout dense 4
P-boule 73
point
  adhérent 4
  d'accumulation 6
  isolé 3
précompact 19
presque partout 96
produit
  d'espaces topologiques 5, 6
  distribution-fonction 158
  scalaire 55
  tensoriel de distributions 171
projecteur 61
projection orthogonale 58
propre 24
P-topologie 73

rare 35
recouvrement 14
  ouvert 14
réflexif 84
relativement compact 15
restriction d'une distribution 156

$\mathscr{S}'(\mathbb{R}^d)$ 183
$\mathscr{S}(\mathbb{R}^d)$ 179
semi-linéaire 55
semi-norme 73
séparable 13
séparé 3
séparer les points 22
série normalement convergente 42
sequilinéaire 62
$\sigma(E,F)$ 78
σ-algèbre 91
solution fondamentale 175
sous-espace topologique 5
sous-linéaire 79
sous-recouvrement 14
sous-suite 7
spectre d'un opérateur 51
suite
  convergente 6
  de Cauchy 30
  exhaustive de compacts 23
  extraite 7
  faiblement convergente 66

support
  d'une distribution 159
  d'une fonction 135
  singulier 161
$\mathfrak{S} \otimes \mathfrak{T}$ 108

topologie 3
  faible 78, 84
  préfaible 84
total 49
transformée de Fourier 145, 185
tribu 91
  borélienne 91

ultrafiltre 27
uniformément
  continue 13
  convexe 86
  équicontinue 36

valeur
  d'adhérence 6
  principale 157
voisinage 4
volume d'un pavé 116

$W^{s,p}(\Omega)$ 195

Table des Matières

Chapitre 1 *Topologie*
1.1 Espaces métriques 1
1.2 Ouverts 2
1.3 Espaces topologiques 3
1.4 Applications continues 8
1.5 Espaces compacts 14
1.6 Filtres et ultrafiltres. Théorème de Tychonoff 25
1.7 Partitions de l'unité 29
1.8 Espaces métriques complets 30
1.9 Le théorème de Baire 35
1.10 Espaces fonctionnels 36

Chapitre 2 *Espaces de Banach*
2.1 Espaces vectoriels normés 41
2.2 Applications linéaires continues 47
2.3 Spectre d'un opérateur 51
2.4 Le théorème de Banach-Steinhaus 52
2.5 Opérateurs compacts 53

Chapitre 3 *Espaces de Hilbert*
3.1 Orthogonalité 55
3.2 Séries orthogonales 62
3.3 Adjoint d'un opérateur 64
3.4 Compacité faible 66
3.5 Le théorème spectral pour les opérateurs normaux compacts 68

Chapitre 4 *Espaces localement convexes*
4.1 Topologie définie par une famille de semi-normes. 73
4.2 Topologies faibles 78
4.3 Jauge d'un ensemble convexe 79
4.4 Le théorème de Hahn-Banach 80
4.5 Ensembles bornés 83
4.6 Dualité des espaces normés 84
4.7 Le théorème de l'application ouverte pour les espaces de Fréchet 87
4.8 Limites inductives d'espaces de Fréchet 88

Chapitre 5 *Intégration*
5.1 Fonctions mesurables 91
5.2 Mesures positives sur un espace mesurable 95
5.3 Intégrale des fonctions élémentaires positives 97
5.4 Intégrale des fonctions positives 98
5.5 Fonctions intégrables 102
5.6 Fonctions définies par des intégrales 107
5.7 Produits d'espaces mesurables 108
5.8 Construction d'une mesure borélienne 112

5.9 La mesure de Lebesgue sur l'espace euclidien 116
5.10 Primitives 120
5.11 La formule du changement de variables 122
5.12 Inégalités de convexité 125
5.13 Inégalités de Clarkson 129
5.14 Le théorème de représentation de Riesz 131

Chapitre 6 *Convolution et transformation de Fourier*
6.1 Densité des fonctions continues à support compact 135
6.2 Convolution 137
6.3 Régularisation 139
6.4 Compacité des opérateurs de convolution 140
6.5 Fonctions maximales 142
6.6 Transformation de Fourier des fonctions 145

Chapitre 7 *Distributions*
7.1 L'espace des fonctions de test 151
7.2 Partitions différentiables de l'unité 154
7.3 Distributions sur un ouvert 155
7.4 Exemples de distributions 157
7.5 Produit d'une distribution par une fonction lisse 158
7.6 Support d'une distribution 159
7.7 Dérivation des distributions 161
7.8 Composition par un difféomorphisme 166
7.9 Convolution d'une distribution et d'une fonction de test 168
7.10 Produit tensoriel de distributions 171
7.11 Convolution des distributions 172
7.12 Quelques solutions fondamentales 175

Chapitre 8 *Transformation de Fourier des distributions*
8.1 Fonctions à décroissance rapide 179
8.2 Distributions tempérées 183
8.3 Transformation de Fourier des distributions tempérées 185
8.4 Exemples 189
8.5 Transformation de Fourier partielle 190

Chapitre 9 *Espaces de Sobolev*
9.1 Définition 195
9.2 Injections de Sobolev 196
9.3 Le cas s=1 pour le demi-espace 197
9.4 Le cas général 199
9.5 Le cas limite 200

Appendice *Le théorème de Zorn et l'axiome du choix*
10.1 L'axiome du choix 205
10.2 Le théorème de Zorn 205
10.3 Une version ordonnée du théorème de Hahn-Banach 207
10.4 Bases algébriques d'un espace vectoriel 208
10.5 Le théorème de Zermelo 209